Part 9: Particles, Fields & Interactions

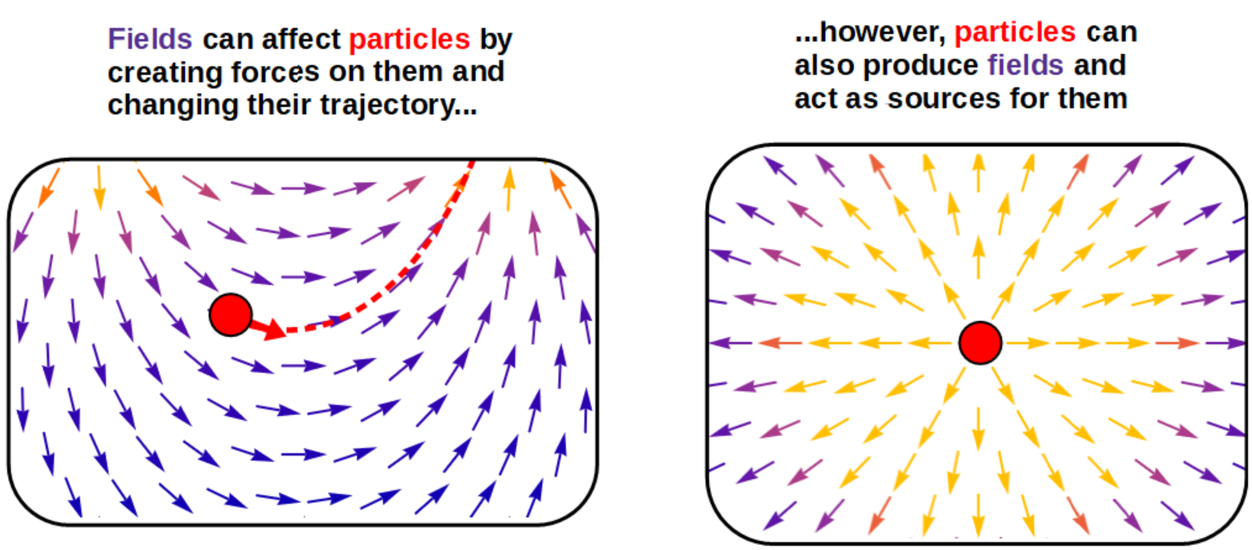
This will be the last main part of the course. If you've made it this far, I would like to congratulate you - you should now have a very good understanding of Lagrangian mechanics and field theory, as well as a solid foundation to study any topic in modern physics you choose!

Part 9 is all about *interactions* - interactions between particles, fields as well as other fields. We will essentially take everything we've learnt so far about various field theories and combine them to describe interacting field theories.

This part will also make use of the ordinary particle Lagrangian mechanics we discussed throughout Parts 1-6. This will be mainly for considering how particles and fields interact and affect each other in field theory. In a sense, this part will tie everything together - we'll take our newly learned concepts in field theory and combine them with what we learned about ordinary Lagrangian mechanics at the beginning of the course.

1. Particles & Fields - Tying Everything Together!

We will begin this part by discussing how particles and fields interact - in other words, we'll combine ordinary particle Lagrangian mechanics with Lagrangian field theory. The overarching theme we will discover here is that particle-field interactions should be thought of as a two-way street; fields affect particles by creating a force on them and particles affect fields essentially by acting as sources of the field.



In practice, we can study these particle-field interactions by constructing Lagrangians - as we've done all throughout this course - that combine both the dynamics of the particle and also the dynamics of the field. This is done by introducing **coupling terms** in the Lagrangian, and these terms are what allow us to study the interaction between the particle and the field.

For example, perhaps you've seen the Lagrangian for a charged particle in an electromagnetic field before:

$$L = \frac{1}{2}mv^2 + q\vec{v} \cdot \vec{A} - q\phi$$

This Lagrangian is just one example of a Lagrangian that describes a field-particle interaction - we have a field (well, two fields, $\vec{A}$ and ϕ) that interact with the particle due to a non-zero electric charge q (which acts as a "coupling constant" here).

At the end of this section, you'll discover where this Lagrangian fundamentally comes from. But first, we'll develop a general framework for describing these field-particle interactions. Let's get started!

1.1. Combining Particle & Field Lagrangians

As with everything in this course, the way we describe physics is by constructing a Lagrangian. In this case, we need a Lagrangian that allows us to describe a combined field-particle interaction.

The general way to do this is actually very simple; we simply take some ordinary Lagrangian for a particle and add to it an **interaction Lagrangian**. The interaction Lagrangian is a term that fully describes how the particle and the field interact, and how they affect each other. This gives us a total Lagrangian that describes the **dynamics of the particle** in the presence of this interaction:

$$L = L_{particle} + L_{field}$$

Essentially, the interaction term here "couples" the field to the particle, describing *how the field affects the particle*.

Now, we could do the same for the field as well - we take some Lagrangian density for some field and add to it an interaction Lagrangian density term. This gives us a total Lagrangian that describes the **dynamics of the field** in the presence of this particle-field interaction:

$$\mathcal{L} = \mathcal{L}_{field} + \mathcal{L}_{int}$$

This Lagrangian density describes the full dynamics of the field, including *how the particle affects the field*.

Perhaps you can already see the two-way street I mentioned earlier - we have Lagrangians for both the particle and the field, with the **particle Lagrangian** containing a term that describes how the field affects the particle and the **field Lagrangian** containing a term that describes how the particle affects the field.

The point here is that these Lagrangians essentially allow us to consider the same physical situation, but from two different prespectives; from the perspective of the particle or from the perspective of the field. We do this by constructing Lagrangians for both the particle and the field individually (which describe either the dynamics of the particle or the dynamics of the field by itself), with the two being related by the interaction Lagrangian.

As an example, a possible interaction Lagrangian for a particle could be (in non-index notation here):

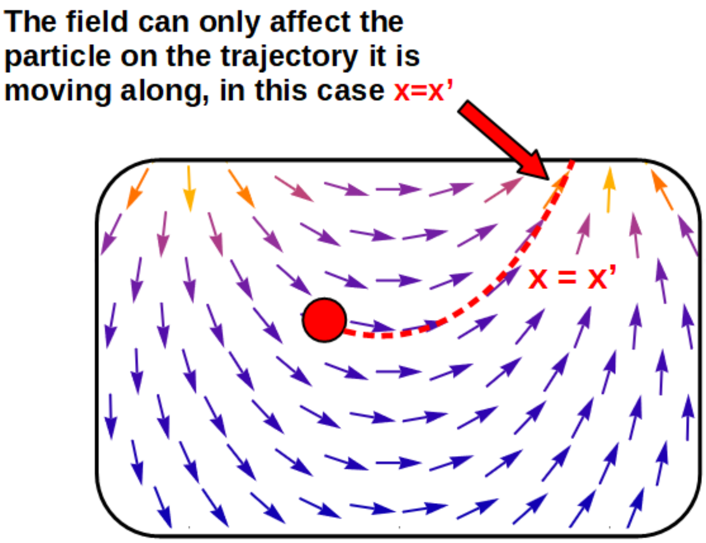
$$L_{int} = q\vec{v} \cdot \vec{A} - q\phi$$

This describes how the fields ϕ and $\vec{A}$ (the electric and magnetic potentials) affect the dynamics of the particle - it is an interaction Lagrangian that describes how a charged particle couples to the electromagnetic field.

The associated interaction Lagrangian density that describes the same interaction but from the perspective of the field, on the other hand, would be:

$$\mathcal{L}_{int} = (q\vec{v} \cdot \vec{A} - q\phi)\delta^3(x - x')$$

The object $\delta^3(x - x')$ here is called the **Dirac delta function**, with $x' = x'(t)$ being the spatial trajectory of the particle. You'll find more discussion about the Dirac delta funtion below, but the special property of it is that it's zero unless $x = x'$ - in other words, we are at the position where the particle is located at in space.



In simple terms, the delta function -factor here describes the fact that the particle can only directly affect the field at the position it is located at, and not anywhere else - this is needed here to ensure **locality**.

Math Interlude: The Dirac Delta Function

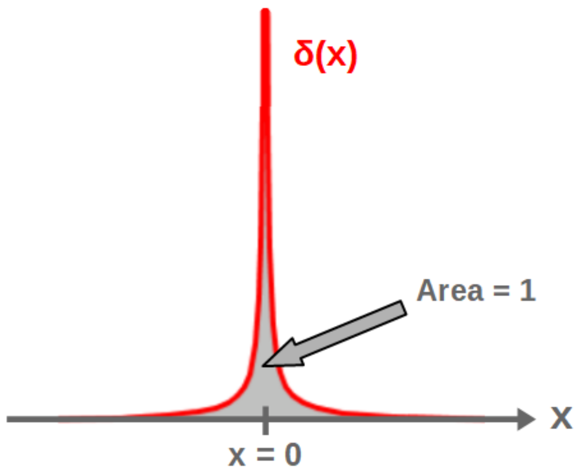
The Dirac Delta function (a distribution, to be more precise) is one of the most useful mathematical tools in physics. Essentially, it allows us to represent "point-like" things (like point particles) as well as all kinds of impulse-like signals and effects. The Dirac delta function is denoted as $\delta(x)$ and its defining properties are that it is zero at all points where its argument is NOT zero (so, $x \neq 0$) and goes to infinity when $x = 0$. In other words:

$$\delta(x) = \begin{cases} 0 & , x \neq 0 \\ \infty & , x = 0 \end{cases}$$

Now, this isn't the most mathematically precise way to define the delta function, but will work perfectly well for us. In addition to these, the defining property of the Dirac delta is that its integral over all of space is always one. In other words:

$$\int_{-\infty}^{\infty} \delta(x) dx = 1$$

Essentially, we can visualize the Dirac delta function as an infinite "spike" located at $x = 0$ that's also infinitely narrow (in such a way that the area under its curve remains exactly 1):



We can also shift the "spike" or zero location of the Dirac delta to any other point by writing it as $\delta(x-a)$. This is the same delta function as before, but its "spike" now occurs at $x = a$ instead of $x = 0$. One of the useful consequences of this is that if we now integrate any function with the delta function, we get back simply the value of the function itself, at the zero location of the delta function. Mathematically:

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)dx = f(a)$$

In other words, the delta function "picks out" the value of the function at a particular location. This turns out to be very useful in making integrals involving delta functions extremely simple to calculate.

We can also straightforwardly generalize the Dirac delta function to higher dimensions by writing it as a product of multiple delta functions. For example, in three-dimensional Cartesian coordinates, we'd have $\delta(x-x_0)\delta(y-y_0)\delta(z-z_0)$, which we usually write in terms of a vector-valued argument as $\delta^3(\vec{r}-\vec{r}_0)$. However, we will often suppress these vector arrows inside the arguments and just write $\delta^3(r-r_0)$.

One important thing to mention here is that there are multiple different ways of *representing* the Dirac delta function. For example, it could be represented as the limit of a Gaussian exponential or in terms of some other function with the same limit properties. However, the most useful representation - called the **integral representation of the delta function** - for us is the following:

$\delta(x-a) = \dfrac{1}{2\pi}\displaystyle\int_{-\infty}^{\infty} e^{i(x-a)x'}dx'$, where x' here is an arbitrary integration variable.

Intuitively, this definition makes sense if you think about it since it has the required properties of the delta function. Particularly, if $x \neq a$, this integral results in zero, but if $x = a$, we'd have $\int_{-\infty}^{\infty} dx' \rightarrow \infty$. In higher dimensions, for example, in 3D, this would be written as:

$\delta^3(x-a) = \frac{1}{(2\pi)^3} \int e^{i(\vec{x}-\vec{a})\cdot\vec{x}'} d^3x'$, where the integral is taken over all of 3D space.

One more particularly useful thing about the delta function for us is that the Fourier transform of a plane wave of the form $e^{-i(k_\mu - k'_\mu)x^\mu}$ is a four-dimensional delta function of the form $(2\pi)^4\delta^4(k'-k)$. We will later come to interpret that this can be used to describe a particle with momentum k in the case that we interpret plane wave excitations of fields as particles.

Now, perhaps the most important question now is - why would we actually *want* to define a function like the Dirac delta in the first place? Well, intuitively, we can use it to describe anything that's "point-like". We could think of the delta function as describing the "density" of a point particle, for example - zero everywhere except at the location of the point source where it is essentially infinite (since a point source is "infinitely dense").

Delta functions are used very commonly especially in electrodynamics. For example, we can describe the electric charge distribution of a point charge q at some location $\vec{r}' = \vec{r}'(t)$ (the charge can be moving, hence the function of time here) in terms of a delta function as:

$$\rho(\vec{r}, t) = q\delta^3(\vec{r} - \vec{r}'(t))$$

This could then be inserted into Maxwell's equations in order to solve them for the fields of a point charge.

Of course, there might not *really* exist any perfect delta functions in nature. It's just a model like anything else in physics. Perhaps a better way to view a "physical" delta function would be a very tall and narrow spike, but not infinite - in other words, an approximation.

Now, notice how the interaction Lagrangian for the particle and the interaction Lagrangian density for the field seem to be very closely related (at least in the example above). Indeed, these are actually the same - since they should describe the same interaction - but one is an *ordinary Lagrangian* and the other is a *Lagrangian density* (which is what we need for fields, generally speaking).

So, how are these two related? Well, it's actually simple - **the Lagrangian is just the integral of the Lagrangian density over all of space**:

$$L_{int} = \int \mathcal{L}_{int} d^3x$$

For example, for the electromagnetic interaction Lagrangian from above, we'd find (since the integral of something with the Dirac delta function gives you just the integrand, as explained above):

$$L_{int} = \int \mathcal{L}_{int} d^3x = \int (q\vec{v} \cdot \vec{A} - q\phi)\delta^3(x - x')d^3x = q\vec{v} \cdot \vec{A} - q\phi$$

Note that this should be evaluated at the position $x = x'$, even though we've suppressed the arguments of the functions here.

In more general terms, the interaction Lagrangian *density* will always contain a Dirac delta function to ensure locality. This means that we can write any interaction Lagrangian density generally in the form:

$$\mathcal{L}_{int} = \rho\delta^3(x - x')$$

Here, ρ is some function (which is to be specified for any given interaction) that describes the particular interaction we're considering. It is generally a function of both the particle's position, velocity, time and also whatever field we're considering in our theory, so $\rho = \rho(x, t, \dot{x}, \phi)$.

The interaction Lagrangian for the particle is then simply:

$$L_{int} = \int \mathcal{L}_{int} d^3x = \int \rho\delta^3(x - x')d^3x = \rho$$

Due to the integral with the delta function, the function ρ here is now to be evaluated at $x = x'$ (along the trajectory of the particle), so $\rho = \rho(x', t, \dot{x}', \phi)$ also with $\phi = \phi(x', t)$.

So, that's pretty simple - to obtain an interaction Lagrangian, we just have to come up with one for either the field or the particle, and the two are then simply related by just a factor of a Dirac delta function along the trajectory of the particle.

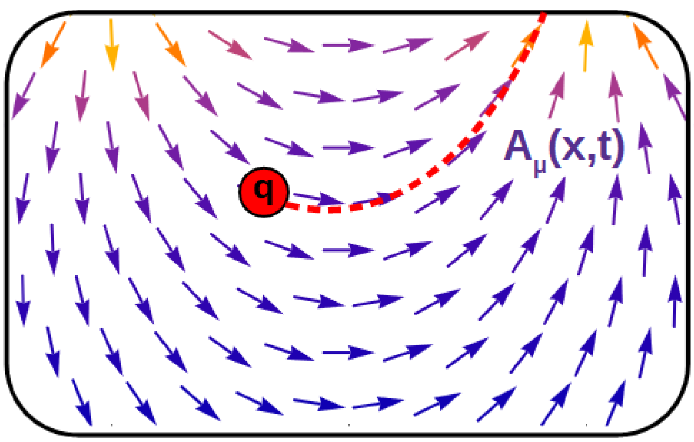
Generally when studying particle-field interactions, we want the interaction Lagrangian to be *linear* in the field. This gives the simplest possible non-trivial coupling between the field and the particle, which usually also gives us something that can be solved analytically. This property is just something to keep in mind when constructing these field-particle interaction Lagrangians.

In practice, this just means that the function $\rho(x, t, \dot{x}, \phi)$ can be separated into a product of the field and some function that describes only the particle, so of the form:

$$\rho(x, t, \dot{x}, \phi) = \rho(x, t, \dot{x})\phi(x, t)$$

1.1.1. Example: Lagrangian For Charged Particle In an Electromagnetic Field

Before move forward with this topic, let's do an example of constructing one of these interaction Lagrangians. We'll do this for a charged particle with electric charge q moving in an electromagnetic field - so this will describe the interaction between the particle and the field A^μ.



Note; the field A_μ is really a four-vector field, but due to four-vectors being quite difficult to draw, it is shown as an ordinary vector field here. So, this picture is not entirely accurate!

For our particle Lagrangian, we'll take the relativistic Lagrangian for a free particle (this was derived at the end of Part 6):

$$L = -mc^2\sqrt{1 - \frac{\dot{x}^2}{c^2}}$$

However, now that we're familiar with index notation, let's also write this one using it. The way to do this is pretty simple - we conside the velocity vector of the particle, which has components:

$$\dot{x}^\mu = \begin{pmatrix} c \\ \vec{v} \end{pmatrix}$$

With this, we can write the Lagrangian in the form (we can write here $\dot{x}^2 = -\dot{x}_i \dot{x}^i$):

$$L = -mc^2\sqrt{1 - \frac{\dot{x}^2}{c^2}} = -mc\sqrt{c^2 - \dot{x}^2} = -mc\sqrt{\dot{x}_0 \dot{x}^0 + \dot{x}_i \dot{x}^i} = -mc\sqrt{\dot{x}_\mu \dot{x}^\mu}$$

This is our relativistic Lagrangian for the free particle. But it doesn't contain an interaction term with the electromagnetic field yet - so how do we construct one? Well, perhaps the best approach to this is to look at the action directly, instead of the Lagrangian. After all, the action is what describes the actual physics. For the particle, this action is:

$$S = \int L dt = -mc\int \sqrt{\dot{x}_\mu \dot{x}^\mu} dt$$

Now, usually when we start constructing Lagrangians more or less out of thin air, we begin with the simplest possible valid Lagrangian. To construct the simplest possible interaction term in this action, we need to think about what the relevant four-vectors we have here - the action, of course, needs to be Lorentz invariant, so it needs to be built out of four-vectors.

In this case, probably the simplest four-vector we have would be the particle's coordinate displacement four-vector dx^μ (note; the coordinates x^μ themselves do not form a four-vector, but the displacements dx^μ do). It also makes sense to include an infinitesimal like this in the action, since we are doing an integral here. In fact, the action shown above is actually of the form $\sqrt{dx_\mu dx^\mu}$ if you bring the dt inside the square root.

So, a reasonable choice, perhaps the most sensible one in fact, to include in the interaction term is dx^μ. Because we also want to couple the particle to the field A^μ, we should also include A^μ in this term - this is the simplest four-vector related to the field itself and specifically, including only A^μ (and not some higher powers of it) gives us a term that is *linear* in the field.

To construct a Lorentz invariant out of these two, well, the simplest one is just a direct contraction, $A_\mu dx^\mu$. We will also include a constant prefactor q here (and a minus sign for conventional reasons), which acts as a "coupling constant" describing how strongly the particle interacts with the field (yes, this is the electric charge of the particle). So, our action with this interaction term is then:

$$S = -mc\int \sqrt{\dot{x}_\mu \dot{x}^\mu}dt + S_{int} = -mc\int \sqrt{\dot{x}_\mu \dot{x}^\mu}dt - q\int A_\mu dx^\mu$$

To obtain a Lagrangian from this, we can write it as follows (here, $\dot{x}^\mu = dx^\mu / dt$):

$$S = -mc\int \sqrt{\dot{x}_\mu \dot{x}^\mu}dt - q\int A_\mu \frac{dx^\mu}{dt}dt = \int \left(-mc\sqrt{\dot{x}_\mu \dot{x}^\mu} - qA_\mu \dot{x}^\mu\right)dt$$

The Lagrangian is then the expression inside this integral over time, so we have the full particle Lagrangian with an interaction term as:

$$\boxed{L = -mc\sqrt{\dot{x}_\mu \dot{x}^\mu} - qA_\mu \dot{x}^\mu}$$

For pedagogical reasons, it's interesting to write out this interaction term fully. We can do this using the components of the velocity four-vector $\dot{x}^\mu = (c, \vec{v})$ and the electromagnetic four-potential $A_\mu = (\phi / c, -\vec{A})$ - note the minus sign due to the lowered index μ:

$$L_{int} = -qA_\mu \dot{x}^\mu = -qA_0 \dot{x}^0 - qA_i \dot{x}^i = -q\phi + q\vec{A} \cdot \vec{v}$$

This, of course, matches with the usual Lagrangian for a charged particle in an electromagnetic field you'll typically see in the literature (without using index notation).

For the Lagrangian for our field A^μ, on the other hand, we'll take the Proca Lagrangian that also allows us to consider the case of a massive field. This turns out to be extremely interesting (for theoretical purposes, at least), because it allows us to see how electromagnetism would behave if the photon had mass, which we already touched on a bit in Part 8. The Lagrangian density for the field will then be of the form:

$$\mathcal{L} = \mathcal{L}_{field} + \mathcal{L}_{int} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}\mu^2 A_\mu A^\mu + \mathcal{L}_{int}$$

Now, this interaction Lagrangian density $\mathcal{L}_{int}$ can be obtained simply from the $L_{int} = -qA_\mu \dot{x}^\mu$ -term by multiplying with a delta function - so we already did all the hard work in figuring out the interaction Lagrangian for the particle, and the corresponding field interaction term just falls out directly!

We then have the full Lagrangian density describing the dynamics of the field as:

$$\boxed{\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}\mu^2 A_\mu A^\mu - qA_\mu \dot{x}^\mu \delta^3(x - x')}$$

We now have these two Lagrangians that fully describe the electromagnetic interaction betweem a charged particle and the electromagnetic field - how the field affects the particle and, on the other hand, how the particle affects the field. Keep these at the back of your mind, as we will return to them in the next example after discussing some theory first!

1.2. Inhomogeneous Field Equations

Let's get back to the general theory for a bit, before continuing with our example. In particular, let's talk about what the interaction terms in the Lagrangian actually do in terms of the dynamics of both the particle and the field - in other words, how these terms affect the equations of motion of the particle and the field equations of the field.

Here, we'll only consider interaction terms that are linear in the field, which means we can write them for both the particle and the field as follows:

$$\begin{cases} L_{int} = \rho\phi \\ \mathcal{L}_{int} = \rho\phi\delta^3(x - x') \end{cases}$$

Note that ϕ here represents a general field even though it looks like this is only valid for a scalar field. If we had a vector field A^μ, we'd have $j_\mu A^\mu$ instead of $\rho\phi$ here (to get something Lorentz invariant). However, the *form* of the term is exactly the same - it's a linear combination or product of the field with some function of x, t and $\dot{x}$.

The most direct way to see how these terms in the Lagrangian affect the dynamics of both the field and the particle is to, well, simply plug the full Lagrangians into the Euler-Lagrange equations. The full Lagrangians are, in this case:

$$\begin{cases} L = L_{particle} + \rho\phi \\ \mathcal{L} = \mathcal{L}_{field} + \rho\phi\delta^3(x - x') \end{cases}$$

The equations of motion for the particle are obtained from the Euler-Lagrange equations as:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}^\mu} - \frac{\partial L}{\partial x^\mu} = 0$$

$$\Rightarrow \quad \frac{d}{dt}\frac{\partial}{\partial \dot{x}^{\mu}}(L_{particle} + \rho\phi) - \frac{\partial}{\partial x^{\mu}}(L_{particle} + \rho\phi) = 0$$

$$\Rightarrow \quad \frac{d}{dt}\frac{\partial L_{particle}}{\partial \dot{x}^{\mu}} - \frac{\partial L_{particle}}{\partial x^{\mu}} = \frac{\partial(\rho\phi)}{\partial x^{\mu}} - \frac{d}{dt}\frac{\partial(\rho\phi)}{\partial \dot{x}^{\mu}}$$

The right-hand side here has exactly the form of a generalized (velocity-dependent) force, which we discussed way back in Part 5. In particular, this is a generalized force obtained from a velocity-dependent potential $U = -\rho\phi$:

$$Q_{\mu} = -\frac{\partial U}{\partial x^{\mu}} + \frac{d}{dt}\frac{\partial U}{\partial \dot{x}^{\mu}}$$

So, the field affects the particle essentially by creating a potential (which may or may not be velocity-dependent) that then results in a **force** acting on the particle - a particle and a field interact, from the particle's perspective, through a *potential field*. Of course, this isn't very surprising - it's exactly how would expect a gravitational field, an electromagnetic field or any other field to affect a particle, even in ordinary Newtonian mechanics.

The Euler-Lagrange equations for the particle then simply give us ordinary differential equations with additional force terms on the right-hand side - nothing we couldn't already solve based on everything discussed in, for example, Part 4.

But what about how the particle affects the field? How do the field equations change? In "ordinary physics", we would study this by calculating the fields from the field equations of the relevant theory - for example, Maxwell's equations of electromagnetism allow us to calculate how a charged particle affects the electric and magnetic fields. However, since we're using the Lagrangian formalism, we can consider this a but more generally since we don't necessarily have to specify a particular Lagrangian or field.

Okay, let's just get into it! The field equations for any given field theory can be obtained from the general form of the Lagrangian density $\mathcal{L}$ from above simply from the Euler-Lagrange equations:

$$\partial_{\mu}\frac{\partial \mathcal{L}}{\partial(\partial_{\mu}\phi)} - \frac{\partial \mathcal{L}}{\partial \phi} = 0$$

$$\Rightarrow \quad \partial_{\mu}\frac{\partial}{\partial(\partial_{\mu}\phi)}(\mathcal{L}_{field} + \rho\phi\delta(x - x')) - \frac{\partial}{\partial \phi}\left(\mathcal{L}_{field} + \rho\phi\delta^3(x - x')\right) = 0$$

$$\Rightarrow \quad \partial_{\mu}\frac{\partial \mathcal{L}_{field}}{\partial(\partial_{\mu}\phi)} - \frac{\partial \mathcal{L}_{field}}{\partial \phi} = \rho\delta^3(x - x')$$

On the left-hand side, we have just the ordinary field equations - for $\mathcal{L}_{field}$ being the Klein-Gordon Lagrangian, we'd get the Klein-Gordon equation and for the Maxwell Lagrangian, we'd get the Maxwell equation (as examples).

What's more interesting is the right-hand side - this is a source term with a function $\rho = \rho(x, t, \dot{x})$ and a delta function. The delta function here ensures that this source term is zero at any other point that where the particle is located at. However, at $x = x'$, this source term becomes non-zero and results in some very interesting physics, as we will see soon.

On a more general note, field equations (or partial differential equations) with a non-zero source term like this are called **inhomogeneous field equations**. If the right-hand side above were zero, it would be a *homogeneous* field equation.

The interaction term in the Lagrangian therefore has the effect of making the field equations **inhomogeneous** - in other words, it introduces additional source terms to the field equations. Physically, this means that the particle essentially behaves as a source for the field - we can think of the particle as interacting with the field by creating a field of its own, exactly like a charged particle would create an electromagnetic field of its own.

So, we while we found that fields interact with particles by creating forces on them, particles, on the other hand, interact with fields by acting as sources for the field. Any particle that interacts with a field will create that particular type of field by its own. This is the essence of any field-particle interaction and shows exactly how it is always a two-way street - we cannot have a field-particle interaction without both of them affecting each other in some way.

Now, we already know how to solve the equations of motion for the particle that result from these field-particle interactions. These are just ordinary differential equations just like all the other equations of motion we've solved throughout this course.

The inhomogeneous field equations, on the other hand, are much more tricky since they are partial differential equations. Previously, we solved the *homogeneous* field equations essentially just by guessing a form for the solution and plugging it in. However, this approach really won't work anymore for the *inhomogeneous* field equations due to the source term with a delta function.

Luckily, there is a very powerful and general tool for solving these inhomogeneous field equations - called **Green's functions** - that we can take advantage of. This will be discussed next and after that, we'll apply it to our example of electromagnetism from earlier.

For the last thing in this section, I want to present the general framework for describing field-particle interactions:

1. Specify an interaction Lagrangian $L_{int} = \rho(x, t, \dot{x}, \phi)$ or an interaction Lagrangian

density $\mathcal{L}_{int} = \rho(x, t, \dot{x}, \phi)\delta^3(x - x')$.

2. Construct Lagrangians for the particle and the field with this interaction term:

$$L = L_{particle} + L_{int}$$
$$\mathcal{L} = \mathcal{L}_{field} + \mathcal{L}_{int}$$

3. The equations of motion for the particle are then obtained from the (relativistic) Euler-Lagrange equations:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}^\mu} - \frac{\partial L}{\partial x^\mu} = 0$$

4. The field equations describing the dynamics of the field, on the other hand, are obtained from the Euler-Lagrange equations for the field:

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} - \frac{\partial \mathcal{L}}{\partial \phi} = 0$$

5. Solve the resulting equations of motion for the particle and the field equations for the field! This can be done generally using Green's functions.

1.2.1. Math Interlude: Green's Functions

Let's take a brief detour to discuss a powerful tool you'll encounter many many times in both classical and quantum field theory, called **Green's functions**. Some of the stuff discussed in this section is quite advanced and might require some prerequisites (such as knowledge about Fourier transformations), but I'll try my best to explain it in a way that is understandable evn if you're not familiar with these concepts beforehand.

In field theory, Green's functions are used to solve inhomogeneous field equations - which are exactly the field equations we are interested here in our case of particles acting as sources of a field.

Here, we will only discuss Green's functions at the most basic level, as a tool for solving these field equations. You'll find a much more comprehensive discussion of Green's functions in the Green's Functions & Contour Integration -bonus tutorial that's included with this course.

The idea behind Green's functions is that they are the "basic building block" solutions for a field created by an arbitrary source function. In particular, say we have an inhomogeneous field equation with a general source function $\rho(x,t)$:

$$L\phi(x,t) = \rho(x,t)$$

Here, L is an arbitrary differential operator. For example, if we had $L = \partial_\mu \partial^\mu + \mu^2$, we would get the inhomogeneous Klein-Gordon equation. However, here we do not necessarily have to specify what this differential operator is - the general idea behind Green's functions will apply to *any* differential operator.

Now, given that we have an inhomogeneous field equation of this form, the *general solution* for the field ϕ is then given by a superposition of the source function ρ with the Green's function G for that particular differential operator L:

$$\boxed{\phi(x,t) = \int \rho(y) G(x,y) d^4y}$$

, where y is a 4D integration variable, generally.

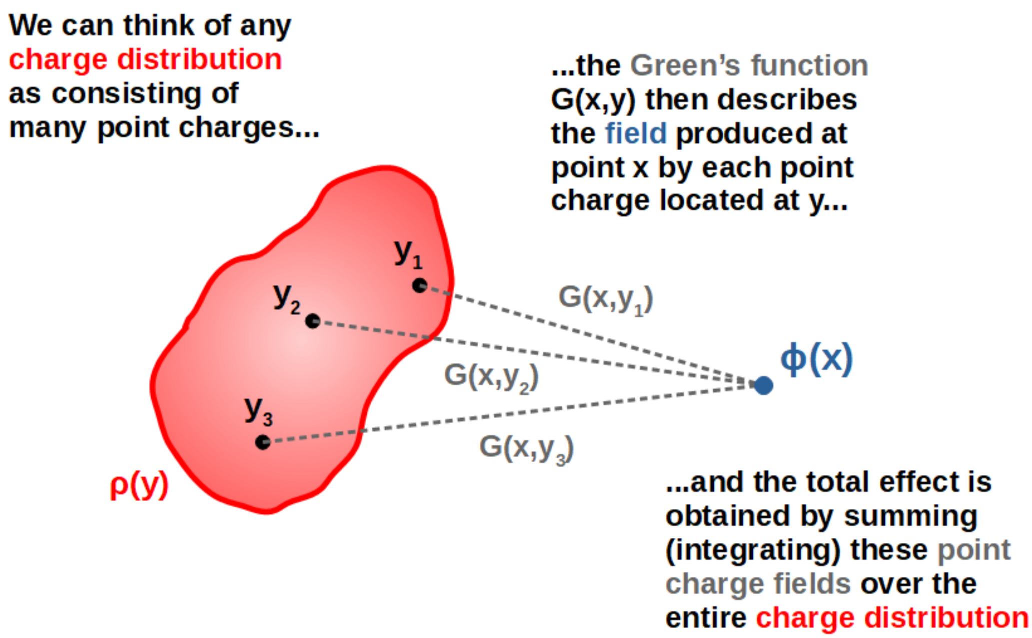
But what are these Green's functions? Well, they are the basic building block solutions for the field ϕ *with sources*. Analogously, we found plane waves as the basic building blocks for field equations *without sources* (for example, for the Klein-Gordon equation in Part 8) - Green's functions play exactly this role, but for inhomogeneous field equations.

These basic building block solutions, the Green's functions, are obtained when we set the source function to be a *point source* - in other words, a **Dirac delta function**. Thus, the defining equation for the Green's function *for a particular differential operator* L (this is important - a Green's function is always defined for a specific differential operator) is given by:

$$\boxed{LG(x,y) = \delta(x-y)}$$

Essentially, the trick here is that we're taking an inhomogeneous differential equation with some general source function and reducing it to the simplest possible choice of the source function - a point source. Then, we obtain the more complicated solution by summing up these "simpler", point-source solutions called Green's functions.

The intuitive picture here is that this works, because we can think of any arbitrary source function as a superposition of many many point sources. The Green's function then describes how each one of these point charges contributes to the total field $\phi(x,t)$ at some other location created by the full charge distribution - essentially, the Green's function represents the charge of a point source as well as "propagates" the effect to other points:



As a quick example - which you may actually have seen before in some form, at least - a very common use of Green's functions is found in electrostatics. In electrostatics, the fundamental differential equation we want to solve for an electrostatic potential $\phi(x)$ is Poisson's equation with some charge distribution function $\rho(x)$:

$$\nabla^2\phi = -\frac{\rho}{\varepsilon_0}$$

This can be solved using Green's functions. The Green's function $G(x, y)$ for this operator $L = \nabla^2$ satisfies the following equation:

$$\nabla^2 G = \delta^3(x - y)$$

The Green's function, in this case, can be found to be (the arguments x and y above are really vectors, we've just suppressed the vector arrows):

$$G(x, y) = \frac{1}{4\pi}\frac{1}{|\vec{x} - \vec{y}|}$$

Then, the general solution to Poisson's equation *for any* charge distribution $\rho(x)$ (with some appropriate boundary conditions) is:

$$\phi(x) = \int \rho(y)G(x, y)d^3y = -\frac{1}{4\pi\varepsilon_0}\int \frac{\rho(y)}{|\vec{x} - \vec{y}|}d^3y$$

Perhaps you've seen this general solution before when studying electrostatics. In fact, this is the origin of where the standard form of the electric potential comes from ($\sim 1/|\vec{x}-\vec{y}|$), even though Green's functions aren't often mentioned in elementary electromagnetism.

An interesting special case to consider is if we had a charged point particle of charge q (located at, for example, the origin $x = 0$) as our source function $\rho(x)$. The source function would then have the form:

$$\rho(x) = q\delta^3(x)$$

The potential would then have the form:

$$\phi(x) = -\frac{1}{4\pi\varepsilon_0}\int \frac{\rho(y)}{|\vec{x}-\vec{y}|}d^3y = -\frac{1}{4\pi\varepsilon_0}\int \frac{q\delta^3(y)}{|\vec{x}-\vec{y}|}d^3y$$

This can be simplified a lot more by carrying out this integral here, which is actually very simple - this is an integral over the variable y and the delta function here simply gives us the integrand function with $y = 0$. The answer is then:

$$\phi(x) = -\frac{q}{4\pi\varepsilon_0}\int \frac{\delta^3(y)}{|\vec{x}-\vec{y}|}d^3y = -\frac{q}{4\pi\varepsilon_0}\frac{1}{|\vec{x}|}$$

We can then see from this that a point particle produces an electrostatic potential that scales with the distance to the origin as $\sim 1/r$ - exactly as we would expect!

So, once we find a Green's function for a particular operator, we can then superpose lots of these Green's functions with some particular source function to obtain any solution we want for the inhomogeneous field equations.

Essentially, once we've found a Green's function for a particular operator L, we've solved a huge class of inhomogeneous differential equations - in fact, we've solved ALL differential equations of the form $L\phi = \rho$ for the operator L, regardless of what ρ is. That is powerful!

It's worth noting, however, that it isn't quite as simple as I've made it out to be. In particular, there are also boundary conditions that you have to take into account when solving for Green's functions, which we won't be discussing further here.

Now, for the important question - how do you actually find these Green's functions?

Generally, finding Green's functions is a very difficult task. However, in certain cases - which we have quite often in field theory - Green's functions can be found using **Fourier transformations**.

Essentially, Fourier transformations give us a way to turn differential equations into simple algebraic equations - a Fourier transform has the effect of turning the operation of differentiation into simple multiplication. The trick that enables this is that a Fourier transform transforms an equation to "frequency space" in which taking derivatives can be thought of as just multiplicative operations.

Here, I won't be covering Fourier transformations in much detail, I'll just tell you the rules for them that we need.

The outline for the process of finding Green's functions goes more or less as follows - first, we take a Fourier transform of the defining equation for the Green's function. This turns the Green's function $G(x, y)$ into another function $\widetilde{G}(k)$ and the differential operator L acting on G into a *multiplicative* operator $\widetilde{L}$ now acting on $\widetilde{G}(k)$:

$$\begin{aligned} & LG(x, y) = \delta(x - y) \\ & \Rightarrow \quad \mathcal{F}(LG(x, y)) = \mathcal{F}(\delta(x - y)) \\ & \Rightarrow \quad \widetilde{L}\widetilde{G}(k) = \mathcal{F}(\delta(x - y)) \end{aligned}$$

The Fourier transformation of the delta function, in this case, can be found to be $\mathcal{F}(\delta(x - y)) = e^{-ik_\mu y^\mu}$. Moreover, since $\widetilde{L}\widetilde{G}(k)$ is now just a multiplication, we can divide by $\widetilde{L}$ to find:

$$\widetilde{G}(k) = \frac{\mathcal{F}(\delta(x - y))}{\widetilde{L}} = \frac{e^{-ik_\mu y^\mu}}{\widetilde{L}}$$

With this approach, we've now solved for the Fourier transformed Green's function $\widetilde{G}(k)$. To find the original Green's function, we take an *inverse* Fourier transform, which by definition will give us:

$$G(x, y) = \frac{1}{(2\pi)^4} \int e^{ik_\mu x^\mu} \widetilde{G}(k) d^4k = \frac{1}{(2\pi)^4} \int \frac{e^{ik_\mu (x^\mu - y^\mu)}}{\widetilde{L}} d^4k$$

We've now solved for the Green's function for any operator L (although there are restrictions on where this solution applies, however, in many cases in field theory, this solution works perfectly fine).

Now, if you didn't quite understand what we did here, don't worry - the important result here is that as long as we can find the Fourier transformation $\widetilde{L}$ (which is a simple algebraic expression) of the differential operator L, then we can calculate the Green's function using the formula:

$$G(x,y) = \frac{1}{(2\pi)^4}\int \frac{e^{ik_\mu\left(x^\mu - y^\mu\right)}}{\widetilde{L}} d^4k$$

Without going into more detail, I've put the common Fourier transformations of some of the most common differential operators in the table below.

Differential operator (L):	**Fourier transformed operator ($\widetilde{L}$):**
∇^2	$-k^2$
$\nabla^2 - \mu^2$	$-\left(k^2 + \mu^2\right)$
∂_μ	ik_μ
$\partial_\mu \partial^\mu$	$-k_\mu k^\mu$
$\partial_\mu \partial^\mu + \mu^2$	$\mu^2 - k_\mu k^\mu$

With these Fourier transforms, we can find the Green's function for a particular operator L and with that, the general solution to the field equation $L\phi = \rho$. The most difficult part in all of this is actually doing the integral in the above formula for the Green's function analytically. We'll look at some simplified examples soon, but first, here's the general framework we will be using to solve inhomogeneous field equations:

1. Specify a differential operator L based on the inhomogeneous field equation you have. This will generally be of the form $L\phi = \rho$.

2. Find the Fourier transformation $\widetilde{L}$ of this operator. Some of the particularly useful ones for field theory are given in the above table.

3. Calculate the Green's function for that operator using the general formula:

$$G(x,y) = \frac{1}{(2\pi)^4}\int \frac{e^{ik_\mu\left(x^\mu - y^\mu\right)}}{\widetilde{L}} d^4k$$

4. Take a superposition of the Green's function with the source function of your inhomogeneous field equation to obtain the solution for the field ϕ (note; x and y here generally denote *spacetime* points):

$$\phi(x) = \int \rho(y)G(x,y)d^4y$$

Let's look at some examples next. Note that the following examples will be somewhat simplified, since finding the most general solutions is a very difficult task mathematically and doing all the rigogous math isn't really the purpose of this course either.

Examples of Solving Inhomogeneous Field Equations

Let's do a very simple example at first - Poisson's equation for electrostatics. We already (partially) looked at this example earlier, but let's do it using our above framework now. Now, Poisson's equation is a differential equation for the electrostatic potential $\phi(x)$ created by some charge distribution $\rho(x)$:

$$\nabla^2\phi = -\frac{\rho}{\varepsilon_0}$$

The differential operator here is $L = \nabla^2$, which according to our table above, has a Fourier transform $\widetilde{L} = -k^2$. Also, since we are dealing with electrostatics here, in which everything is time-independent, our formula for the Green's function should only contain a spatial integral over d^3k and the sum in the exponent reduces should have $e^{ik_\mu\left(x^\mu - y^\mu\right)} = e^{i\vec{k}\cdot(\vec{x}-\vec{y})}$.

Therefore, the Green's function for Poisson's equation is given by:

$$G(x,y) = \frac{1}{(2\pi)^4}\int \frac{e^{ik_\mu\left(x^\mu - y^\mu\right)}}{\widetilde{L}} d^4k \quad \Rightarrow \quad G(x,y) = -\frac{1}{(2\pi)^3}\int \frac{e^{i\vec{k}\cdot(\vec{x}-\vec{y})}}{k^2} d^3k$$

The difficult part is now actually calculating this integral analytically. Luckily, however, it

The difficult part is now actually calculating this integral analytically. Luckily, however, it can be done and the easiest way to do so is by switching to spherical coordinates. Note that the integral is done over the k-variable (and not a typical spatial coordinate variable), so this becomes an integral using spherical coordinates in "k-space".

In any case, the volume element is given in spherical coordinates by $d^3k = k^2 \sin\varphi dk d\theta d\varphi$ and the exponential term with a dot product is given by $e^{i\vec{k}\cdot(\vec{x}-\vec{y})} = e^{ik|\vec{x}-\vec{y}|\cos\varphi}$. We then have:

$$G(x,y) = -\frac{1}{(2\pi)^3}\int_0^{\pi}\int_0^{2\pi}\int_0^{\infty} e^{ik|\vec{x}-\vec{y}|\cos\varphi} \sin\varphi dk d\theta d\varphi$$

This integral can still be evaluated analytically using some tools from complex analysis (contour integration, in particular). However, I'll just tell you that the answer is:

$$G(x,y) = \frac{1}{4\pi|\vec{x}-\vec{y}|}$$

This is essentially the Coulomb potential and we can now see where it fundamentally comes from - from the Green's function for Poisson's equation! Notice that this has the form of a $\sim 1/r$ -type potential, which is in line with our earlier interpretation of the Green's functions describing fields of point sources.

With this, the general solution to Poisson's equation is then:

$$\phi(x) = \int \rho(y)G(x,y)d^4y \quad \Rightarrow \quad \phi(x) = -\frac{1}{4\pi\varepsilon_0}\int \frac{\rho(y)}{|\vec{x}-\vec{y}|}d^3y$$

Let's look at another, slightly more complicated example - the inhomogeneous Klein-Gordon equation with a general source function $\rho(x)$, where x denotes a *spacetime* point in general:

$$\left(\partial_\mu\partial^\mu + \mu^2\right)\phi = \rho$$

The differential operator here is $L = \partial_\mu\partial^\mu + \mu^2$, which has Fourier transform $\widetilde{L} = \mu^2 - k_\mu k^\mu$. Thus, the Green's function for the inhomogeneous Klein-Gordon equation is:

$$G(x,y) = \frac{1}{\cdots}\int \frac{e^{ik_\mu\left(x^\mu-y^\mu\right)}}{\cdots}d^4k = \frac{1}{\cdots}\int \frac{e^{ik_\mu\left(x^\mu-y^\mu\right)}}{\cdots}d^4k$$

$$(2\pi)^4 \int \quad \tilde{L} \quad (2\pi)^4 \int \quad \mu^2 - k_\mu k^\mu$$

Now, while it is possible to actually do this integral, the answer we'd get is quite complicated and wouldn't really tell us much right away. Instead, we can consider a simplified case - the Green's function for the *time-independent* Klein-Gordon equation. In that case, the integral above would be a spatial integral of the form:

$$G(x,y) = -\frac{1}{(2\pi)^3}\int \frac{e^{i\vec{k}\cdot(\vec{x}-\vec{y})}}{\mu^2 + k^2} d^3k$$

This looks quite similar to the Green's function we had above for Poisson's equation. In fact, the integral can be evaluated analytically in the same way by switching to spherical coordinates and using some techniques from complex analysis. The result we would find in the end is:

$$G(x,y) = \frac{1}{4\pi}\frac{e^{-\mu|\vec{x}-\vec{y}|}}{|\vec{x}-\vec{y}|}$$

This is called the **Yukawa potential**. The interesting thing about it is the exponential factor caused by a non-zero mass parameter μ. This has the effect of the potential going to zero much faster than the Coulomb potential would as the distance $|\vec{x}-\vec{y}|$ increases. What this essentially means is that *massive fields have a much shorter effective range*.

With this Green's function, the general solution to the time-independent inhomogeneous Klein-Gordon equation would then be:

$$\phi(x) = \int \rho(y) G(x,y) d^4y \;\; \Rightarrow \;\; \phi(x) = \frac{1}{4\pi}\int \frac{\rho(y) e^{-\mu|\vec{x}-\vec{y}|}}{|\vec{x}-\vec{y}|} d^3y$$

For last, I want to briefly mention the important role of Green's functions in quantum field theory. In quantum field theory, Green's functions play the role of *propagators*. The intuitive idea is that the Green's function tells us "how to propagate the effect of a source from one point to another" - hence why it is called a propagator. We can see this from the general solution for a field in terms of Green's functions:

$$\phi(x,t) = \int \rho(y) G(x,y) d^4y$$

Inside the integral here, we're looking at the source function at the point y. The Green's function then being a function of both x and y, it tells us how the source function at that point y produces the field at the point x - essentially, it "propagates" the effect of the source from one point to another. In quantum field theory, we can also think of these Green's functions as describing *force carrier particles* that "carry" the effects of a source from one point to another, producing a force.

For example, in quantum electrodynamics, these would be (virtual) photons that "carry" the effect of an electric charge distribution to produce an electric field or force at some other point. The Green's function, viewed as a propagator in quantum field theory, would then describe the probability amplitude that a force carrier photon goes from point y to x. All of this is to just say that understanding Green's functions is crucial if you want to study quantum field theory further.

1.2.2. Example: Massive Electromagnetism & The Lorentz Force Law

With the tool of Green's functions under our belt, let's now continue our example of the charged particle in an electromagnetic field from earlier. For reference, here are the Lagrangians we have for the particle and the field:

$$\begin{cases} L = -mc\sqrt{\dot{x}_\mu \dot{x}^\mu} - qA_\mu \dot{x}^\mu \\ \mathfrak{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}\mu^2 A_\mu A^\mu - qA_\mu \dot{x}^\mu \delta(x - x') \end{cases}$$

We'll begin with the equations of motion for the particle. These are obtained from the Euler-Lagrange equation as (note that the field A_μ here depends on x^μ):

$$\begin{aligned} &\frac{d}{dt}\frac{\partial L}{\partial \dot{x}^\nu} - \frac{\partial L}{\partial x^\nu} = 0 \\ &\Rightarrow \quad \frac{d}{dt}\frac{\partial}{\partial \dot{x}^\nu}\left(-mc\sqrt{\dot{x}_\mu \dot{x}^\mu} - qA_\mu \dot{x}^\mu\right) - \frac{\partial}{\partial x^\nu}\left(-mc\sqrt{\dot{x}_\mu \dot{x}^\mu} - qA_\mu \dot{x}^\mu\right) = 0 \\ &\Rightarrow \quad -mc\frac{d}{dt}\left(\frac{1}{2\sqrt{\dot{x}_\mu \dot{x}^\mu}}\frac{\partial}{\partial \dot{x}^\nu}\dot{x}_\mu \dot{x}^\mu\right) - q\frac{dA_\nu}{dt} + \frac{\partial A_\mu}{\partial x^\nu}\dot{x}^\mu = 0 \end{aligned}$$

This $\partial \dot{x}_\mu \dot{x}^\mu / \partial \dot{x}^\nu$ -derivative is just the same "derivative of a square" we've seen previously, which in this case gives us $2\dot{x}_\nu$. For the dA_ν / dt -derivative, we can use the chain rule since A_ν is a function of x^ν:

$$\frac{dA_\nu}{dt} = \frac{\partial A_\nu}{\partial x^\mu}\frac{dx^\mu}{dt} = \frac{\partial A_\nu}{\partial x^\mu}\dot{x}^\mu$$

Inserting these, we get:

$$-mc\frac{d}{dt}\left(\frac{1}{2\sqrt{\dot{x}_\mu\dot{x}^\mu}}\frac{\partial}{\partial \dot{x}^\nu}\dot{x}_\mu\dot{x}^\mu\right) - q\frac{dA_\mu}{dt} + q\frac{\partial A_\mu}{\partial x^\nu}\dot{x}^\mu = 0$$

$$\Rightarrow \quad -mc\frac{d}{dt}\left(\frac{1}{2\sqrt{\dot{x}_\mu\dot{x}^\mu}}2\dot{x}_\nu\right) - q\frac{\partial A_\nu}{\partial x^\mu}\dot{x}^\mu + q\frac{\partial A_\mu}{\partial x^\nu}\dot{x}^\mu = 0$$

$$\Rightarrow \quad \frac{d}{dt}\left(\frac{mc\dot{x}_\nu}{\sqrt{\dot{x}_\mu\dot{x}^\mu}}\right) = q\left(\frac{\partial A_\mu}{\partial x^\nu} - \frac{\partial A_\nu}{\partial x^\mu}\right)\dot{x}^\mu$$

The expression inside the parentheses on the left-hand side is actually just the relativistic *four-momentum* of the particle - if you calculate each component of the expression explicitly, you'll find them being exactly the four-momentum components, $p_\nu = (\gamma mc, -\gamma m\vec{v})$ with $\gamma = 1/\sqrt{1-v^2/c^2}$. The right-hand side can also be written in terms of the electromagnetic energy-momentum tensor $F_{\nu\mu} = \frac{\partial A_\mu}{\partial x^\nu} - \frac{\partial A_\nu}{\partial x^\mu}$, so that we get:

$$\boxed{\frac{dp_\nu}{dt} = qF_{\nu\mu}\dot{x}^\mu}$$

That's the equation of motion for the particle! This describes the dynamics of a charged particle in an electromagnetic field and in fact, this equation has a name - it's called the **Lorentz force law**, written in a Lorentz covariant form. The Lorentz force law describes how electric and magnetic fields affect a moving charged particles. You'll find some more discussion of this down below.

A Deeper Look At The Relativistic Lorentz Force Law

In ordinary electromagnetism, the Lorentz force law is a vector equation for the force vector created by an electromagnetic field:

$$\vec{F} = q\vec{E} + q\vec{v} \times \vec{B}$$

$F = qE + qv \times B$

Let's see how we can relate this to the Lorentz force law written above in index notation - they are, in fact, the same equation! First, recall from the end of Part 8 (in the discussion of the electromagnetic energy-momentum tensor) the following expressions relating the electric and magnetic fields to the electromagnetic field tensor:

$$F_{0i} = \frac{1}{c}E_i$$
$$F_{ij} = -\epsilon_{ijk}B^k$$

We can now write out the relativistic Lorentz force law equation in terms of these. Let's begin with the spatial component with $\nu = i$. These are given by:

$$\frac{dp_i}{dt} = qF_{i\mu}\dot{x}^\mu = qF_{i0}\dot{x}^0 + qF_{ij}\dot{x}^j$$

On the left-hand side, we have the time derivative of momentum components - these are just the components of a force, F_i! Also, the things on the right-hand side are $F_{i0} = -E_i/c$, $F_{ij} = -\epsilon_{ijk}B^k$, $\dot{x}^0 = c$ and $\dot{x}^j = v^j$. So, we can write this equation as:

$$F_i = -qE_i - q\epsilon_{ijk}B^k v^j$$

Now, there is a very useful mathematical formula that we can apply here. It tells us that the components of any cross product can be calculated in terms of the Levi-Civita symbol ϵ_{ijk} as:

$$(\vec{v} \times \vec{B})_i = \epsilon_{ijk}B^k v^j$$

We will also write the equation above in terms of upper indices (in which case, the only difference is that the force vector components get a negative sign, since $F_i = -F^i$ from lowering/raising a spatial index), Therefore, we find:

$$F_i = -qE_i - q\epsilon_{ijk}B^k v^j \quad \Rightarrow \quad F^i = qE^i + q(\vec{v} \times \vec{B})^i$$

This is, of course, nothing but the Lorentz force law equation, $\vec{F} = q\vec{E} + q\vec{v} \times \vec{B}$. So, in the end, we find something that is perhaps not all that surprising if you're familiar with electromagnetism - the equation of motion for a charged particle in an electromagnetic field is the Lorentz force law. What we did, however, was to derive this result starting from the very basic principles of field theory and Lagrangian mechanics.

Now, there's still one more thing to consider and that is the 0-component of our relativistic Lorentz force law:

$$\frac{dp_0}{dt} = qF_{0\mu}\dot{x}^{\mu} = qF_{00}\dot{x}^0 + qF_{0i}\dot{x}^i = qF_{0i}\dot{x}^i$$

To see what this tells us, let's consider the 0-component of the four-momentum. This is given by (here, $\dot{x}_0 = c$):

$$p_0 = \frac{mc\dot{x}_0}{\sqrt{\dot{x}_{\mu}\dot{x}^{\mu}}} = \frac{mc^2}{\sqrt{c^2 + v_i v^i}} = \frac{mc^2}{\sqrt{c^2 - v^2}} = \frac{mc}{\sqrt{1 - v^2 / c^2}}$$

What is this? Well, recall the formula for relativistic total energy that we derived at the end of Part 6, $E = \gamma mc^2$ with $\gamma = 1 / \sqrt{1 - v^2 / c^2}$. This is, in fact, nothing but the 0-component of the four-momentum we have above (with an additional $1 / c$ -factor). So, we have:

$$p_0 = \frac{E}{c}$$

With this and also $F_{0i} = E_i / c$ (the E_i -component not to be confused with the energy E), we have:

$$\frac{dp_0}{dt} = qF_{0i}\dot{x}^i$$
$$\Rightarrow \quad \frac{1}{c}\frac{dE}{dt} = q\frac{E_i}{c}\dot{x}^i$$
$$\Rightarrow \quad \frac{dE}{dt} = q\vec{E} \cdot \vec{v}$$

This tells us that the rate of change of the particle's energy is directly proportional to the electric field and most importantly, *only the electric field*. Therefore, we find that only electric fields, and not magnetic fields, can change the total energy of a particle.This is exactly what we would expect even in ordinary electromagnetism - **magnetic fields do no work**.

With the dynamics of the particle now pretty much covered, let's look at the dynamics of the electromagnetic field now. These are contained in the Lagrangian density:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}\mu^2 A_\mu A^\mu - qA_\mu \dot{x}^\mu \delta(x - x')$$

The dynamics of the field can then be obtained from the Euler-Lagrange equations as:

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} - \frac{\partial \mathcal{L}}{\partial A_\nu} = 0$$

$$\Rightarrow \partial_\mu \frac{\partial}{\partial(\partial_\mu A_\nu)}\left(-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}\mu^2 A_\mu A^\mu - qA_\mu \dot{x}^\mu \delta(x - x')\right)$$

$$-\frac{\partial}{\partial A_\nu}\left(-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}\mu^2 A_\mu A^\mu - qA_\mu \dot{x}^\mu \delta(x - x')\right) = 0$$

$$\Rightarrow \partial_\mu \left(-\frac{1}{4}\frac{\partial\left(F_{\mu\nu}F^{\mu\nu}\right)}{\partial(\partial_\mu A_\nu)}\right) - \frac{1}{2}\mu^2 \frac{\partial}{\partial A_\nu}\left(A_\mu A^\mu\right) + q\frac{\partial A_\mu}{\partial A_\nu}\dot{x}^\mu \delta(x - x') = 0$$

We've already calculated the derivative $\partial\left(F_{\mu\nu}F^{\mu\nu}\right) / \partial(\partial_\mu A_\nu)$ in the "Longer Calculations" - section associated with Part 8, so no point in repeating the calculations here. This derivative simply gives:

$$\frac{\partial\left(F_{\mu\nu}F^{\mu\nu}\right)}{\partial(\partial_\mu A_\nu)} = 4F^{\mu\nu}$$

For the other derivatives in our above equation, we get:

$$\frac{\partial}{\partial A_\nu}\left(A_\mu A^\mu\right) = 2A^\nu$$

$$\frac{\partial A_\mu}{\partial A_\nu} = \delta_\mu^\nu$$

With all of these, we have the field equation as:

$$\partial_\mu \left(-\frac{1}{4}\frac{\partial\left(F_{\mu\nu}F^{\mu\nu}\right)}{\partial(\partial_\mu A_\nu)}\right) - \frac{1}{2}\mu^2 \frac{\partial}{\partial A_\nu}\left(A_\mu A^\mu\right) + q\frac{\partial A_\mu}{\partial A_\nu}\dot{x}^\mu \delta(x - x') = 0$$

$$\Rightarrow \quad \partial_\mu \left(-\frac{1}{4} \cdot 4F^{\mu\nu} \right) - \frac{1}{2}\mu^2 \cdot 2A^\nu + q\delta^\nu_\mu \dot{x}^\mu \delta(x - x') = 0$$

$$\Rightarrow \quad \boxed{\partial_\mu F^{\mu\nu} + \mu^2 A^\nu = q\dot{x}^\nu \delta(x - x')}$$

This is the field equation - called the **inhomogeneous Proca equation** - that describes the dynamics of our (massive) vector field A^ν and also how the charged particle affects the field. In particular, we see that acts as a **velocity-dependent source for the field** due to the term on the right-hand side. This means that the faster the particle is moving, the stronger its interaction with the field will be - this is exactly what we'd expect with charged particles interacting with magnetic fields.

If we were to set $\mu = 0$ in the above equation, we would recover Maxwell's equations with a delta function as a source term. This would then describe how any moving charged particle produces an electromagnetic field. However, we'll leave the μ-parameter here to be non-zero for now to discover some interesting features later on.

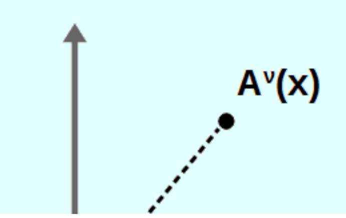
To make our calculations simpler (without losing any generality), we can still impose the Lorenz gauge ($\partial_\mu A^\mu = 0$). This is a perfectly valid gauge choice to make, as it applies for both the massive and massless cases (though it is mandated for the massive case). In the Lorenz gauge, the field equation reduces to:

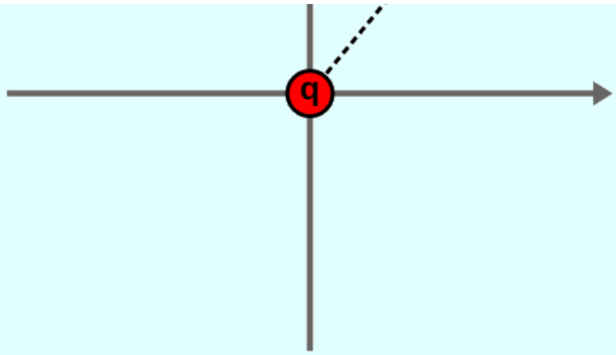
$$\partial_\mu F^{\mu\nu} + \mu^2 A^\nu = q\dot{x}^\nu \delta(x - x') \quad \Rightarrow \quad \partial_\mu \partial^\mu A^\nu + \mu^2 A^\nu = q\dot{x}^\nu \delta(x - x')$$

Now, in this form, this equation might not tell you much right away. The best way to see what it actually describes is to find particular solutions to it that are easy to intepret. You'll find an interesting example of this below, which also shows you how to use the Green's functions we discussed earlier in practice.

Example: Solving The Inhomogeneous Proca Equation

For this example, we'll consider a particle sitting stationary at the origin and see what type of field it produces:





For the stationary particle, its four-velocity has components:

$$\dot{x}^{\nu} = \begin{pmatrix} c \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

The field created by this particle will be time-independent, so $A^{\nu}(x, t) = A^{\nu}(x)$, and will only consist of the A^0-component, the electric potential. So, the field configuration in this case will be of the form:

$$A^{\nu} = \begin{pmatrix} \phi(x)/c \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

The only non-trivial component of the field equations is therefore the one with $\nu = 0$ (since the particle is at the origin, we have $x' = 0$ here):

$$\partial_{\mu}\partial^{\mu}A^0 + \mu^2 A^0 = q\dot{x}^0\delta(x - x')$$
$$\Rightarrow \left(\frac{1}{c^2}\partial_0\partial^0 + \partial_i\partial^i\right)\frac{\phi(x)}{c} + \mu^2\frac{\phi(x)}{c} = qc\delta(x)$$

Since the field $\phi(x)$ is time-independent, the $\partial_0 = \partial/\partial t$ -derivatives go to zero. Moreover, the second derivative operator $\partial_i\partial^i$ gives us the Laplacian, $-\nabla^2$, so the field equation can be written as:

$$\left(\frac{1}{c^2}\partial_0\partial^0 + \partial_i\partial^i\right)\frac{\phi(x)}{c} + \mu^2\frac{\phi(x)}{c} = qc\delta(x) \quad \Rightarrow \quad \left(\nabla^2 - \mu^2\right)\phi(x) = -qc^2\delta(x)$$

Now, what is this equation? It's exactly the defining equation for the Green's function of the time-independent Klein-Gordon equation (the Yukawa potential) we looked at earlier in the "Green's functions" -section. So, the solution for the potential $\phi(x)$ will then be:

$$\phi(x) = \frac{1}{4\pi}\int \frac{\rho(y)e^{-\mu|\vec{x}-\vec{y}|}}{|\vec{x}-\vec{y}|}d^3y$$

However, in this case, we have the source function $\rho(x)$ itself as a delta function, $\rho(x) = -qc^2\delta(x)$. Therefore, this potential is actually:

$$\phi(x) = \frac{1}{4\pi}\int \frac{\rho(y)e^{-\mu|\vec{x}-\vec{y}|}}{|\vec{x}-\vec{y}|}d^3y = \frac{1}{4\pi}\int \frac{-qc^2\delta(y)e^{-\mu|\vec{x}-\vec{y}|}}{|\vec{x}-\vec{y}|}d^3y$$

This delta function here results in this entire integral giving us just the integrand, $e^{-\mu|\vec{x}-\vec{y}|} / |\vec{x}-\vec{y}|$, but with $y = 0$ (i.e. the position of the source particle). Thus, we find:

$$\phi(x) = -\frac{qc^2}{4\pi}\int \frac{\delta(y)e^{-\mu|\vec{x}-\vec{y}|}}{|\vec{x}-\vec{y}|}d^3y = -\frac{qc^2}{4\pi}\frac{e^{-\mu|\vec{x}|}}{|\vec{x}|}$$

Here, $|\vec{x}|$ is just the radial distance from the origin, where the charged particle is sitting at, to the point where we are evaluating the potential at. Therefore, we can write it in spherical coordinates as:

$$\phi(r) = -\frac{qc^2}{4\pi}\frac{e^{-\mu r}}{r}$$

Here, we're using different units than those standardly used in electrodynamics. To use those units, this should have a factor of μ_0 as well, so it would have the form (using $c^2 = 1/\varepsilon_0\mu_0$):

$$\phi(r) = -\mu_0\frac{qc^2}{4\pi}\frac{e^{-\mu r}}{r} = -\frac{q}{4\pi\varepsilon_0}\frac{e^{-\mu r}}{r}$$

So, we find that for a *massive* electrostatic field, the potential as the form of the **Yukawa potential**. Now, for ordinary electromagnetism, we would have $\mu = 0$, such that this potential reduces to:

$$\phi(r) = -\frac{q}{4\pi\varepsilon_0 r}$$

This is the usual Coulomb potential created by a point charge, which we also found as a solution to the Coulomb equation. This is also the origin of where the typical $1/r^2$-dependence of electric fields come from, since electric fields can be obtained from the potentials as $\vec{E}(r) = -\nabla\phi(r)$.

Now, the unusual features arise if we consider the case with $\mu \neq 0$. In that case, the electric potential has the prefactor $e^{-\mu r}$, which comes from the fact that the field we're considering is *massive* - in other words, *if photons were massive*.

What this means is that if the photon had a mass, we would find that the field created by a charged particle would have the form of the Yukawa potential instead of a typical $1/r$-potential. The electric field, in that case, would have the form:

$$\vec{E}(r) = -\nabla\phi(r) = -\frac{\partial\phi(r)}{\partial r}\hat{r} = -\frac{q}{4\pi\varepsilon_0}\left(\frac{1}{r^2} + \frac{\mu}{r}\right)e^{-\mu r}\hat{r}$$

So, instead of electric fields behaving as $\sim 1/r^2$, they would have a much different form with, in fact, a much shorter range. This is because the $e^{-\mu r}$ drops off very quickly with increasing r, with the rate of drop-off essentially being determined by the mass parameter μ.

This then means that electric forces between charged particles would have a much shorter range and be much weaker than what we find in our world of massless photons. Now, since electric forces do NOT behave like this in the real world, this actually works as experimental evidence towards the fact that photons are massless - the world would be a much different place if they weren't.

To wrap things up with this example, the main point in all of this was to showcase the way in which particles and fields interact - particles behave as sources for a field, and fields produce forces on a particle. Specifically, in this example, we saw how an electrically charged particle acts as a source for the electromagnetic field. The electromagnetic field, on the other hand, influences the particle's equations of motion through the Lorentz force law.

Another important thing to note is that everything we discussed here was purely classical field theory, where we completely separate the idea of particles and fields. In quantum field theory, we cannot really talk about "field-particle interactions" in the same way anymore.

Instead, however, we can talk about fields interacting with other fields (in the case of quantum field theory, these would be quantum fields). We will discuss exactly this next.

2. Building Interacting Field Theories

The rest of Part 9 will be focused on looking at how fields interact with other fields, and how we can construct interacting field theories in various different ways. As usual, we'll also dive into lots of examples and applications - my personal favourite topic here is going to be spontanous symmetry breaking, which we can understand in the context of an interacting field theory.

2.1. How Do We Combine Field Theories?

The first question to tackle here is naturally; how do we even build interacting field theories in the first place? How do we combine different field theories?

Well, the idea is actually pretty similar to what we did when discussing field-particle interactions, though slightly more general. The central concept here is again an **interaction Lagrangian** $\mathcal{L}_{int}$ - this is now purely a Lagrangian density, since we're talking about fields only.

In field-particle interactions, we took two independent Lagrangians - one for a field and one for a particle - and added an interaction term to both, which gave us a theory of both the particle's and the field's dynamics:

$$\begin{cases} L = L_{particle} + L_{int} \\ \mathcal{L} = \mathcal{L}_{field} + \mathcal{L}_{int} \end{cases}$$

When combining field theories, we do the exact same - only now we add the Lagrangians for two (or more) fields together, since they are both Lagrangian densities. In the field-particle interactions, we had one being an ordinary Lagrangian, so we couldn't just add them together, but here we can.

So, given two Lagrangian densities, $\mathcal{L}_1$ and $\mathcal{L}_2$, that describe two different fields, we combine them into an interacting theory by constructing a new Lagrangian of the form:

$$\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_{int}$$

The idea here is that $\mathcal{L}_1$ and $\mathcal{L}_2$ describe the dynamics of both fields individually and have no effect on one another. The $\mathcal{L}_{int}$ -term then ties these fields together, since this term will be a function of both fields, say ϕ and ψ here:

$$\mathcal{L} = \mathcal{L}_1(\phi) + \mathcal{L}_2(\psi) + \mathcal{L}_{int}(\phi, \psi)$$

We can perhaps see this more clearly by considering the Euler-Lagrange equations. The dynamics of each of the fields ϕ and ψ *under the effects of this interaction*, are obtained from the "total" Lagrangian from both field's own Euler-Lagrange equations as:

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} - \frac{\partial \mathcal{L}}{\partial \phi} = 0 \quad \Rightarrow \quad \partial_\mu \frac{\partial \mathcal{L}_1}{\partial(\partial_\mu \phi)} - \frac{\partial \mathcal{L}_1}{\partial \phi} = \frac{\partial \mathcal{L}_{int}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}_{int}}{\partial(\partial_\mu \phi)}$$
$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} - \frac{\partial \mathcal{L}}{\partial \psi} = 0 \quad \Rightarrow \quad \partial_\mu \frac{\partial \mathcal{L}_2}{\partial(\partial_\mu \psi)} - \frac{\partial \mathcal{L}_2}{\partial \psi} = \frac{\partial \mathcal{L}_{int}}{\partial \psi} - \partial_\mu \frac{\partial \mathcal{L}_{int}}{\partial(\partial_\mu \psi)}$$

We see that the interaction Lagrangian introduces a non-zero right-hand side to both field's field equations. We can interpret these as some type of "generalized forces" - so, the fields affect one another by somehow again producing "forces" on one another (although speaking about forces in this context, is not very precise and shouldn't be taken too seriously - the field equations are what matter).

If we had more than two fields, say N fields interacting with one another with each described by a Lagrangian $\mathcal{L}_i$, we would combine their Lagrangians in exactly the same way (with the interaction Lagrangian now being a function of all the fields that take part in the interaction):

$$\mathcal{L} = \sum_{i=1}^{N} \mathcal{L}_i + \mathcal{L}_{int}$$

That's basically all there is to it! Very simple, isn't it? We just add together various Lagrangians and get all kinds of new theories out of it. This also shows us once again why Lagrangians are such powerful tools for describing physics.

Now, the difficult part in all of this is to actually come up with the interaction Lagrangian, $\mathcal{L}_{int}$, everything else is pretty much trivial - we just take existing Lagrangians for already existing field theories and add them up. However, for the interaction terms, there are a few types of them that are the most commonly used ones.

2.2. Types of Interactions

In general, constructing interaction Lagrangians for different fields is not much different than constructing the Lagrangians for the fields themselves - we have a set of rules the

Lagrangians have to obey and any Lagrangian within that set of rules is, in principle, a valid one.

The rules for interaction Lagrangians are pretty much the same ones as we discussed in Part 7 - the Lagrangian should be Lorentz invariant, real-valued and so on. Everything discussed there applies here as well.

All we really have to do when constructing these interaction Lagrangians is follow all the rules for building valid Lagrangians, but with the additional complication of considering Lagrangians that *couple* different fields together in some way. In general, these will pretty much always consist of *products of the fields*. For example, a valid interaction term between two *real* scalar fields, ϕ and ψ, would be of the form $\mathcal{L}_{int} \sim \phi\psi$.

Now, the most common types of interactions we have in field theory are the following:

- **Self-interactions** - these describe a field that interacts *with itself* (but not with other fields), resulting in non-linear dynamics. In general, self-interactions are represented by interaction Lagrangians containing powers of the field **higher than 2**, so $\mathcal{L}_{int} = \sum_{n=3} \lambda_n \phi^n$ (where λ_n are called *coupling constants*). The most common example of a self-interacting theory is the ϕ^4-theory with $\mathcal{L}_{int} \propto \phi^4$.

- **Yukawa interactions** - these describe different scalar fields interacting with one another or alternatively, scalar fields interacting with spinor fields. For example, we will later consider a Yukawa interaction between a complex scalar field ψ and a real scalar field ϕ, described by the interaction Lagrangian $\mathcal{L}_{int} = g\psi^\dagger\psi\phi$.

- **Gauge interactions** - these describe vector fields interacting with other fields, such as scalar or spinor fields, with the additional complication of the theory still retaining *gauge invariance*. The interaction term will generally be determined by applying *minimal coupling* and is of the form $\mathcal{L}_{int} = j^\mu A_\mu$, where j^μ is a conserved current.

We will discuss each of these interactions in detail next, beginning with self-interactions. Gauge interactions will be reserved for last, as they are the most complicated and perhaps the most interesting type of interaction as well.

A unique aspect of interacting field theories is that the resulting field equations we find are nearly always *non-linear*. Previously, we've only been dealing with linear field equations with solutions being some superposition of plane waves. For interacting fields, this will not be true anymore and the field equations are often not even solvable exactly.

Therefore, when solving interacting field theories, we will make use of some techniques in **perturbation theory** as well as apply the Green's functions discussed earlier. The key idea here is that interacting fields at can affect one another at different points and the Green's functions describe how this effect is "propagated" from one point to another.

3. Self-Interactions

Our first example of building and describing interacting field theories will be **self-interactions**. A self-interacting field theory is a theory in which we only have a single field, but this field can interact with itself in various ways. Now, what do we even mean by "interacting with itself"?

Well, to put it simply, self-interactions just mean non-linear dynamics for a field (as we'll discover shortly). Remember from earlier that due to a linear field equation, we were able to obtain the general solution for a field as a superposition of various plane wave solutions:

$$\phi \sim \int e^{ikx} d^4k$$

However, for a non-linear field equation, such a superposition of plane waves is not a valid solution anymore. Let's think about what a superposition of plane waves even means; essentially, it means that we're adding together a bunch of different plane waves that are not related to one another in any way - they do not interact.

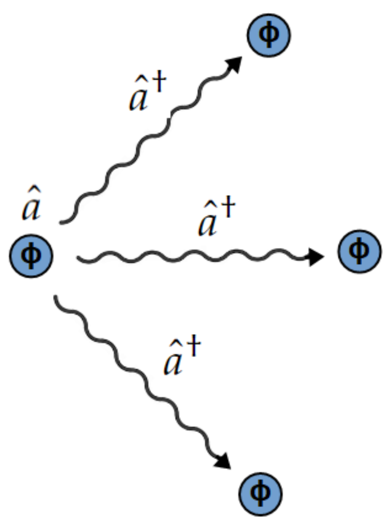
In particular, a superposition of plane waves NOT being a solution to a non-linear field equation then means that the plane waves must somehow be interacting with each other. Instead, we find all kinds of interesting and much more complicated solutions.

Another way to view these self-interactions would be from the quantum field theoretical perspective. In particular, let's say we have a self-interacting theory with the interaction Lagrangian $\mathcal{L}_{int} \sim \phi^4$. We could then imagine the solution for the field being something of the form $\phi \sim a + a^\dagger$, where the a's are the k-dependent coefficients we also had in the standard Klein-Gordon theory. Then, writing out the interaction term, we'll find various products of these coefficients, for example of the form:

$$\mathcal{L}_{int} \sim a^\dagger a^\dagger a^\dagger a$$

When quantizing this field theory, these coefficients turn into what are called **creation and annihilation operators**, so $a \implies \hat{a}$ and $a^\dagger \implies \hat{a}^\dagger$. The creation operator $\hat{a}^\dagger$ describes how a particle (quantum excitation of the field ϕ) is produced and the annihilation operator describes how a particle gets destroyed.

A product of the form $\hat{a}^\dagger\hat{a}^\dagger\hat{a}^\dagger\hat{a}$ in the interaction term would then describe how *one particle is destroyed* (corresponding to the single factor of $\hat{a}$) and *three particles are created* (corresponding to the product $\hat{a}^\dagger\hat{a}^\dagger\hat{a}^\dagger$). In other words, **a single particle turns into three of the same type of particle** (but with different momenta). Such a process would not be possible in the free Klein-Gordon theory.



In fact, this is what it means for a field to self-interact from the quantum field theoretical perspective - it means that particles are constantly being created and destroyed with the number of particles NOT remaining constant. In the free theory, however, a single particle would not be able to "split" into multiple particles or vice versa since we only find terms like $\hat{a}\hat{a}^\dagger$ - products of up to only two operators.

3.1. Non-Linear Field Equations

Now, let's begin developing the general theory for describing self-interactions. The most general way to construct a self-interacting field theory is by taking the interaction Lagrangian to be a *potential* $\mathcal{L}_{int} = -V(\phi)$ that's a function of the field - and importantly, a function of a power of the field greater than two.

This isn't much different from potentials in classical mechanics, which also describe interactions in some sense (essentially, interactions between a particle and its environment). For example, we could consider the Klein-Gordon Lagrangian with a general self-interaction potential:

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}\mu^2\phi^2 - V(\phi)$$

With this, the field equation would be:

$$\partial_\mu \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} - \frac{\partial\mathcal{L}}{\partial\phi} = 0 \quad \Rightarrow \quad \partial_\mu\partial^\mu\phi + \mu^2\phi = -V'(\phi)$$

This is now an inhomogeneous field equation with a source term $-V'(\phi) = -dV(\phi)/d\phi$. So, the field essentially behaves as a source *for itself* through the potential $V(\phi)$ - this is exactly what it means for the field to self-interact.

Now, how do we solve such a field equation? We've actually already discussed that - remember, inhomogeneous field equations can generally be solved using **Green's functions**. For example, we found the Green's function for the Klein-Gordon equation earlier to be:

$$G(x, y) = \frac{1}{(2\pi)^4}\int \frac{e^{ik_\mu\left(x^\mu - y^\mu\right)}}{\mu^2 - k_\mu k^\mu} d^4k$$

Then, the general solution to the above inhomogeneous Klein-Gordon equation with source term $\rho = -V'(\phi)$ would be:

$$\phi(x, t) = -\int V'(\phi)G(x, y)d^4y, \text{ where } V'(\phi) = V'(\phi(y)).$$

See any issue with this? Well, there is a pretty big issue here and it is that the field - which is our solution - appears also on the right-hand side of the solution through $V'(\phi)$. For example, if we had $V'(\phi) \propto \phi^3$, we would find the solution $\phi \sim \int \phi^3 G d^4y$.

This certainly isn't good - we have a general solution for ϕ that requires us to already know ϕ and integrate it. That sounds a bit pointless, doesn't it? However, there is a simple explanation for this. Due to the non-linearity of the field equations, we cannot simply take a superposition of the Green's functions. With the Green's functions approach, we were imagining any field configuration as being constructed out of many "point-source building blocks" and then summing over all of them. However, the same won't work anymore, because the field acts as its own source due to self-interactions.

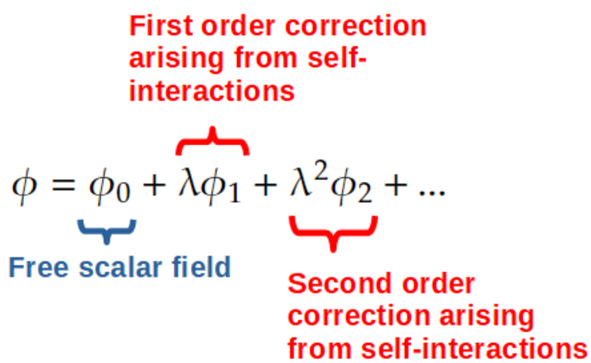
In fact, what self-interactions do is make the field equations not exactly solvable. The fix to this is perturbation theory, which allows us to solve this self-interacting theory but only in the case of weak self-interactions.

3.1.1. Finding Solutions Using Perturbation Theory

A common technique for solving the non-linear field equations of a self-interacting theory that we'll also be using a lot is by assuming the solution is a power series in the interaction parameter λ:

$$\phi = \sum_{n=0}^{\infty} \lambda^n \phi_n = \phi_0 + \lambda\phi_1 + \lambda^2\phi_2 + ...$$

Why would such a solution make sense? First of all, if we set $\lambda = 0$ (turn self-interactions off), this reduces to $\phi = \phi_0$ - so, ϕ_0 is the free Klein-Gordon scalar field. Therefore, each term in this sum represents a higher order correction to the free field due to self-interactions. What we are effectively doing in assuming a solution of this form is taking the free scalar field and adding to it as many correction terms as is needed to accurately describe the non-linear self-interactions of the field:



This is the general idea behind **perturbation theory**. We start with a known solution - in this case, the free Klein-Gordon field ϕ_0 that satisfies the standard homogeneous Klein-Gordon equation - and add these higher-order "perturbation" terms to obtain a more and more accurate solution to the full problem.

Okay, we know how to find the solution ϕ_0 of the perturbation series, but how do we determine these higher-order corrections? In other words, how do we find ϕ_1, ϕ_2 and so on?

In general, we will find have to find each higher-order term individually by iteration of the previous terms. It turns out that if we know ϕ_0 (the free field, which is a plane wave type solution), we can determine ϕ_1. Once we know ϕ_1, we can then determine ϕ_2 and so on.

We'll see examples soon, which should make this more clear.

The issue with this iterative approach is that there is an infinite number of higher-order terms in the series! More over, since the exponent of λ is getting bigger and bigger with each term, it seems that each higher-order term contributes more and more to the full field itself - so this approach looks totally hopeless as well, at least for finding exact solutions.

However, if we were to assume that the self-interactions of the field are somewhat weak, then a finite number of these higher-order terms is enough to accurately describe the full field. In particular, if $\lambda < 1$, then each term in the series would get successively smaller so that each higher-order term becomes less and less significant. This would then allow us to truncate the series after a finite number of terms.

So, this method works for weakly self-interaction theories. For strong self-interactions with $\lambda > 1$, it won't work since we would need to find an infinite number of higher order terms, so the whole approach is somewhat pointless.

We will therefore assume that $\lambda < 1$ and move forward with this assumption. Moreover, this is also a very reasonable assumption to make - as we will see soon, this leads to some very interesting and also useful results.

In fact, for many practical applications that we'll look at, we will take the self-interaction parameter λ to be *very small*, so $\lambda \ll 1$. This means that only the first-order perturbative correction is important and we can practically take $\lambda^2 \approx 0$. This describes a field that deviates only slightly from the free solution due to self-interactions. The solution for the field would then be of the form:

$$\boxed{\phi = \sum_{n=0}^{\infty} \lambda^n \phi_n \approx \phi_0 + \lambda \phi_1}$$

This approximation is actually going to be very useful when discussing *spontaneous symmetry breaking* and the **Higgs mechanism** - in particular, we take ϕ_0 to be a constant (representing a *vacuum solution*) and the extra term $\lambda\phi_1$ then describes small perturbations around this constant. This kind of scalar field configuration actually turns out to give a mass contribution to other fields it interacts with. But more on that later!

3.2. Example: The ϕ^4 Theory

Let's now look at a practical example of solving a self-interacting theory. In particular, we'll look at a scalar field theory with an interaction Lagrangian describing self-interactions that's of the form:

$$\mathcal{L}_{int} = -V(\phi) = -\frac{1}{4}\lambda\phi^4$$

The full Lagrangian for the field ϕ is then:

$$\boxed{\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}\mu^2\phi^2 - \frac{1}{4}\lambda\phi^4}$$

A theory with a Lagrangian of this form is called $\boldsymbol{\phi^4}$ **theory**, which the name for comes from the self-interacting term being $V(\phi) \propto \phi^4$. Now, it turns out that a self-interaction term quartic in the field represents by far the most useful type of self-interaction potential.

We could, of course, have taken $V(\phi) \propto \phi^3$, $V(\phi) \propto \phi^5$ or any number of higher powers. However, interaction potentials with odd powers of the field (so ϕ^3, ϕ^5 and so on) would violate U(1) symmetry in the case of the field ϕ being complex. This generally isn't what we want, since U(1) symmetry is such a common symmetry in so many real-world theories. Moreover, when going to quantum field theory, interaction terms with powers higher than 4 turn out to produce issues with the renormalizibility of the theory.

So, the bottom line is that the ϕ^4 theory is the most useful self-interacting scalar field theory, and in fact, also pretty much the only physically relevant self-interacting theory - physically, it can be used to, for example, describe self-interactions of the Higgs field!

Let's now take a look at the field equations for this theory. These are obtained from the Lagrangian and the Euler-Lagrange equations, as usual:

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi)} - \frac{\partial\mathcal{L}}{\partial\phi} = 0 \quad \Rightarrow \quad \boxed{\partial_\mu\partial^\mu\phi + \mu^2\phi = -\lambda\phi^3}$$

We can see that the field equations for this theory have become *non-linear* due to self-interactions, described by a source term $-\lambda\phi^3$. Essentially, the field will act as a source for itself and its own field equations - this is what self-interactions fundamentally mean.

In practice, the main effect of this non-linear term in field equation is that superpositions of plane waves are not solutions anymore. In fact, ordinary plane waves are not solutions at all, which you can easily see by trying a solution like $\phi = e^{ik_\mu x^\mu}$. Instead, we will find all kinds of interesting effects arising from this.

Okay, let's try to solve this field equation. First of all, the general solution can be expressed in terms of the Green's function for the Klein-Gordon equation as (we discussed this earlier):

$\phi = -\int \lambda\phi^3 G d^4y$, where $G = G(x,y)$ and $\phi = \phi(y)$

The issue, of course, is once again that the field ϕ itself appears inside the integral, making this solution sort of useless in this form. However, we can now try to apply our perturbation theory approach to this. Let's assume our solution for the field is of the form (note that we aren't assuming anything about λ - yet):

$$\phi = \phi_0 + \lambda\phi_1 + \lambda^2\phi_2 + ...$$

Plugging this into the integral expression above (on both the left- and right-hand sides) and writing out the cubic term, we find:

$$\begin{aligned} \phi &= -\int \lambda\phi^3 G d^4y \\ \Rightarrow \quad \phi_0 + \lambda\phi_1 + \lambda^2\phi_2 + ... &= -\int \lambda\left(\phi_0 + \lambda\phi_1 + \lambda^2\phi_2 + ...\right)^3 G d^4y \\ \Rightarrow \quad \phi_0 + \lambda\phi_1 + \lambda^2\phi_2 + ... &= -\int \left(\lambda\phi_0^3 + 3\lambda^2\phi_0^2\phi_1 + ...\right) G d^4y \end{aligned}$$

Now, what do we do with this? Well, each term on both sides include some power of λ. If this equation is to hold at all times, then the different powers of λ must match on both sides, so we can split this into multiple equations:

$$\begin{cases} \phi_1 = -\int \phi_0^3 G d^4y \\ \phi_2 = -3\int \phi_0^2\phi_1 G d^4y \\ \phi_3 = ... \end{cases}$$

This is nice - we've obtained expressions for each of the higher-order correction fields ϕ_1, ϕ_2 and so on. Moreover, notice that the expression for ϕ_1 involves the free field ϕ_0 and the Green's function for the Klein-Gordon equation, *both of which we know already*. Therefore, by doing this integral, we can find ϕ_1 - and, once we know ϕ_1, we can do another integral to find ϕ_2 and so on for any number of higher-order terms. So, we've effectively just solved this field theory!

Of course, to find the exact solution for the full field ϕ, we would have to do an infinite number of these iterations - unless, we take the self-interactions to be weak ($\lambda < 1$) like we discussed previously.

In particular, we'll study an example solution next where $\lambda \ll 1$. This means that the field is of the form $\phi \approx \phi_0 + \lambda\phi_1$, meaning that only the free solution and the first-order correction are important. Now, since we know ϕ_1 from the above integral, the solution for this weakly self-interaction field can be expressed as:

$$\boxed{\phi = \phi_0 - \lambda \int \phi_0^3 G d^4 y}$$

Again, since we know both ϕ_0 (a plane wave type solution, $\phi_0 \sim e^{-ik_\mu x^\mu}$) and G (the Green's function for the Klein-Gordon equation), we could in principle obtain an analytical solution for the field ϕ by carrying out this integral.

Let's try to get some intuition for what this solution represents. In particular, we'll take the simplest non-trivial free solution, a single plane wave $\phi_0 = e^{-ik_\mu x^\mu}$, along with the Green's function in its most general form (note; we have to use a different integration variable here to distinguish it from the wave number k' for our plane wave ϕ_0 - hence the label k' here):

$$G(x,y) = \frac{1}{(2\pi)^4} \int \frac{e^{ik'_\mu (x^\mu - y^\mu)}}{\mu^2 - k'_\mu k'^\mu} d^4 k'$$

The solution above then gives us:

$$\phi = \phi_0 - \lambda \int \phi_0^3(y) G(x,y) d^4 y = e^{-ik_\mu x^\mu} - \lambda \int e^{-3ik_\mu y^\mu} \frac{1}{(2\pi)^4} \int \frac{e^{ik'_\mu (x^\mu - y^\mu)}}{\mu^2 - k'_\mu k'^\mu} d^4 k' d^4 y$$

We can actually carry out these integrations here. To do so, we'll first write this in the following form:

$$\phi = e^{-ik_\mu x^\mu} - \frac{\lambda}{(2\pi)^4} \int \frac{e^{ik'_\mu x^\mu}}{\mu^2 - k'_\mu k'^\mu} \int e^{-i(3k_\mu + k'_\mu) y^\mu} d^4 y d^4 k'$$

Notice that this integral over $d^4 y$ here is just the integral representation of the delta function, in particular of $-(2\pi)^4 \delta^4(3k + k')$. This then gives us:

$$\phi = e^{-ik_\mu x^\mu} + \lambda \int \frac{e^{ik'_\mu x^\mu}}{\mu^2 - k'_\mu k'^\mu} \delta^4(k' + 3k) d^4k'$$

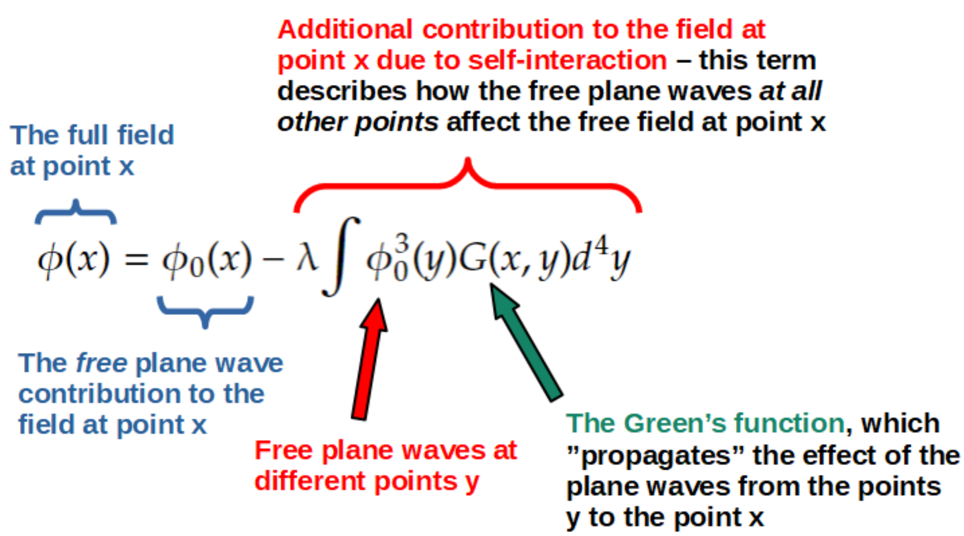
This integral is now very simple, since the delta function here just gives us back the integrand with $k' = -3k$, so we are left with:

$$\phi = e^{-ik_\mu x^\mu} + \lambda \frac{e^{-3ik_\mu x^\mu}}{\mu^2 - 9k_\mu k^\mu}$$

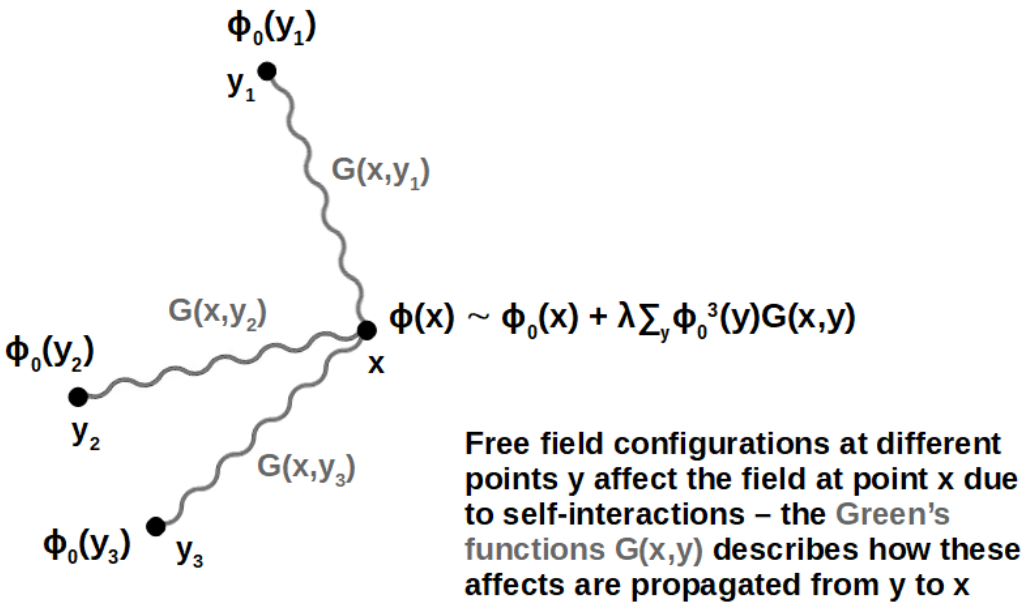
Now, how do we interpret such a solution? First of all, remember what kind of field configuration we started off with - a simple free plane wave with momentum k at the point x. However, due to self-interactions, we find that the plane waves at other points y also contribute to our field at point x. This is described by the first order self-interaction term:

$$\lambda\phi_1 = -\lambda \int \phi_0^3(y) G(x, y) d^4y$$

In particular, this term describes a superposition over all possible free plane waves at other points y (represented by the factor $\phi_0^3(y)$). The Green's function $G(x, y)$ describes how these plane waves at y affect the field at the point x we are interested in - it essentially "propagates" the effect of the field configuration at y over to the point x, hence why it's often called a *propagator*. The full field solution ϕ at the point x we are looking at it then consists of the free plane wave at point x, $\phi_0(x)$ as well as the contribution of these plane waves *at all other points* y, described by the superposition:



So, at a high level, the field ϕ in a self-interacting theory, at any given point, will consist of both the field configurations at that particular point, as well as of the field configurations that have propagated over to that point from all other points. The latter contribution to the field only arises from self-interactions of the field and this is indeed what it means for a field to self-interact - the field at one point acts as a source for the field at some other point.



For our particular example solution, the field configuration consists of a plane wave with momentum k originally at the point x as well as of a plane wave contribution with momentum $3k$, which comes from the $\phi_0^3(y)$ -factor. The Green's function $1/\left(\mu^2 - 9k_\mu k^\mu\right)$ then describes the "propagation" of these perturbative terms coming from point y over to the point x:

$$\phi = e^{-ik_\mu x^\mu} + \lambda \frac{e^{-3ik_\mu x^\mu}}{\mu^2 - 9k_\mu k^\mu}$$

Now, there is also a perhaps bit more intuitive picture hidden in all of this. We can understand this by thinking of these plane waves as describing particles with different momenta like we would do in quantum field theory. This also allows for a very natural way of representing these self-interactions in a "pictorial" form through simple schematics called **Feynman diagrams**. This is discussed below.

Self-Interactions In Terms of Particles & Feynman Diagrams

To see what, for example, the above field configuration ϕ says in terms of particles, we

To see what, for example, the above field configuration φ says in terms of particles, we can switch to "momentum space". All this really means is that instead of looking at the field as a function of space, $\phi(x)$, we look at it as a function of all the different momenta (wave numbers) that are present in the particular field configuration. This is described by the Fourier transform of $\phi(x)$, which we label here as $\widetilde{\phi}(k)$.

If you're not familiar with Fourier transforms, don't worry - all you really need to know is that the Fourier transform we are interested in is obtained by "integrating away" all the position-dependence of the field, giving us a new function of only the variable k. this is done using the following formula, which is just the definition of the Fourier transform:

$$\widetilde{\phi}(k) = \int \phi(x) e^{-ik_\mu x^\mu} d^4x$$

If we now plug in our field $\phi(x) = e^{-ik'_\mu x^\mu} + \lambda \frac{e^{-3ik'_\mu x^\mu}}{\mu^2 - 9k'_\mu k'^\mu}$ from earlier (note that we are using a different variable k' here, as it should be distinguished from the "Fourier variable" k we have above), we find:

$$\begin{aligned}
\widetilde{\phi}(k) &= \int \phi(x) e^{-ik_\mu x^\mu} d^4x \\
&= \int \left(e^{-ik'_\mu x^\mu} + \lambda \frac{e^{-3ik'_\mu x^\mu}}{\mu^2 - 9k'_\mu k'^\mu} \right) e^{-ik_\mu x^\mu} d^4x \\
&= \int e^{-i(k'_\mu + k_\mu)x^\mu} d^4x + \frac{\lambda}{\mu^2 - 9k'_\mu k'^\mu} \int e^{-i(3k'_\mu + k_\mu)x^\mu} d^4x
\end{aligned}$$

Notice that both the integrals here are just the integral representations of the delta functions again - in particular, of the delta functions $-(2\pi)^4 \delta^4(k' + k)$ and $-(2\pi)^4 \delta^4(3k' + k)$. We are then left with just:

$$\begin{aligned}
\widetilde{\phi}(k) &= \int e^{-i(k'_\mu + k_\mu)x^\mu} d^4x + \frac{\lambda}{\mu^2 - 9k'_\mu k'^\mu} \int e^{-i(3k'_\mu + k_\mu)x^\mu} d^4x \\
&= -(2\pi)^4 \delta^4(k' + k) + \frac{\lambda}{9k'_\mu k'^\mu - \mu^2} (2\pi)^4 \delta^4(3k' + k)
\end{aligned}$$

For simplicity, we can rescale our field as $\widetilde{\phi}(k) \rightarrow -(2\pi)^4 \widetilde{\phi}(k)$, our k-variable as $k \rightarrow -k$ and write $3k' = k$ in the denominator of our propagator (for now particular reason other than convenience). We can do this because the delta function in the second term only gives us something non-zero when $3k' = k$ (in terms of our rescaled k-variable). With these, we finally get:

$$\boxed{\tilde{\phi}(k) = \delta^4(k' - k) - \frac{\lambda}{k_\mu k^\mu - \mu^2}\delta^4(3k' - k)}$$

Here comes our particle interpretation - we can interpret fields essentially as particles through de Broglie's relations. This means that a plane wave with wave number k_μ describes a particle with four-momentum $p_\mu = \hbar k_\mu$. Here, we'll just set $\hbar = 1$ so that $p_\mu = k_\mu$ - in other words, each k directly corresponds to a given momenta p.

In momentum space, however, particles with a given momentum are not described by plane waves but instead by delta functions. Now, notice in our expression for $\tilde{\phi}(k)$ above that we only have something non-zero if either of the following conditions are satisfied (due to the delta functions):

$$\begin{cases} k' = k \\ 3k' = k \end{cases}$$

If we interpret k and k' here as the momenta of some *interacting particles*, after a bit of staring at the above conditions, we can see that they describe **momentum conservation**. More precisely, the delta functions can be thought of as being of the form $\delta^4(p_f - p_i)$, where p_f is the final (total) momentum in the field (particle) configuration and p_i the initial momentum.

This then leads to a very natural interpretation of this self-interacting theory in terms of particles - the field $\tilde{\phi}$ describes *processes between interacting particles, which conserve momentum*. Each delta function $\delta^4(p_f - p_i)$ describes one process with initial momentum p_i and final momentum p_f with momentum conservation being ensured by the condition $p_f = p_i$, which happens because the delta function can only be non-zero when this is true.

In particular, our field above describes two possible different processes since there are two delta functions. Generally, each delta function represents one possible process where momentum must be conserved. The full field configuration consists of *all* these possible processes:

A particle interaction process with initial total momentum k and final total momentum 3k' –

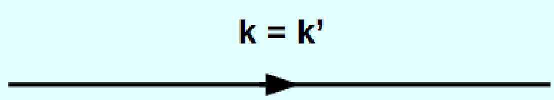
this process can only happen when 3k'=k, in other words, momentum is conserved

$$\widetilde{\phi}(k) = \delta^4(k' - k) - \frac{\lambda}{k_\mu k^\mu - \mu^2}\delta^4(3k' - k)$$

A particle interaction process with initial total momentum k and final total momentum k' – this process can only happen when k'=k, in other words, momentum is conserved

The first process here happens when $k' = k$, in which case the entire second term is zero (so no self-interaction happens, since the λ-term is not there) and the field configuration consists of just a single delta function, $\delta^4(k' - k)$. This is a process where the initial momentum in the system is k and the final momentum is k' and momentum is conserved.

We interpret this as a process with just a single particle moving from one point to another, with its momentum staying the same. We can draw this as a schematic with a single line that describes a free particle with momentum $k = k'$ (we'll talk more about what these diagrams actually represents soon!):



The second possible process we can have here is described by the term:

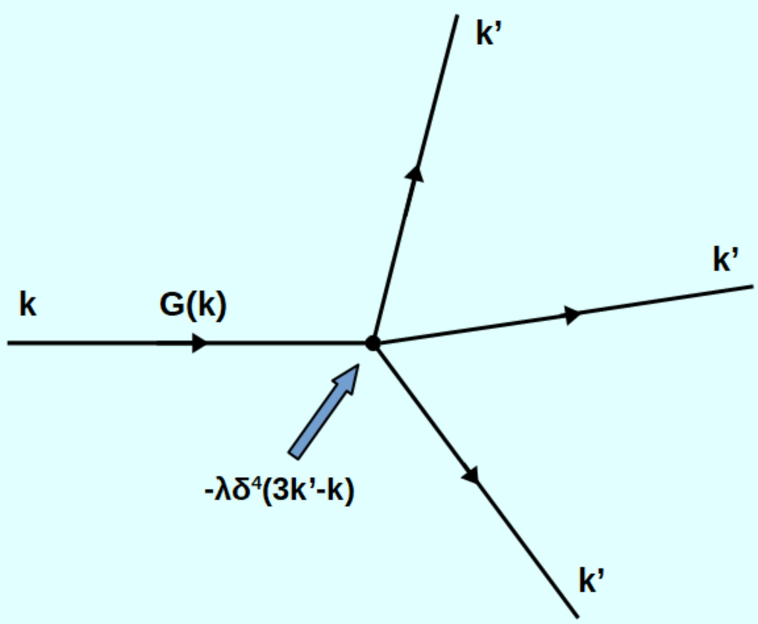
$$-\frac{\lambda}{k_\mu k^\mu - \mu^2}\delta^4(3k' - k)$$

This describes a process with initial total momentum k in the system and final total momentum $3k'$. This interaction can only happen when $3k' = k$, which again describes momentum conservation encoded into the delta function.

Now, we can interpret this term as describing a **particle scattering process** where we initially have just one particle with momentum k, which then "splits" into *three* new particles, each with momentum k'.

As far as the other factors present in this term go, the factor $-\lambda$ describes the interaction or scattering process itself and the factor $\dfrac{1}{k_\mu k^\mu - \mu^2}$ - the propagator - describes essentially how the single particle with momentum k "moves" from its initial point to the point where this interaction takes place.

We could draw this scattering process as a schematic with an inital line that represents the particle with momentum k that is "propagated" by the factor $G(k) = 1/\left(k_\mu k^\mu - \mu^2\right)$ to a *vertex* where the interaction takes place. The interaction at this vertex is described by the factor $-\lambda$ and a delta function $\delta^4(3k' - k)$ to ensure momentum conservation. Then, after the vertex, we draw three separate lines that each represent one of the particles with momentum k'. The diagram would look as follows:



This is actually what interactions mean generally in the context of quantum field theory more generally - scattering processes between particles in which the particle number may change. Particles can get created or destroyed.

In the context of self-interactions, these would be scattering processes between particles of the same type (since we are only describing a single field). These types of interactions with the total particle number changing would not happen in the free field theory (with $\lambda = 0$).

Also, in interacting theories with multiple fields, interactions can result in particles of one type "turning into" particles of a different type. An example of this would be electron-positron annihilation in which an electron and a positron interact and "turn into" photons. This is a possible type of scattering process in the theory of *quantum electrodynamics*.

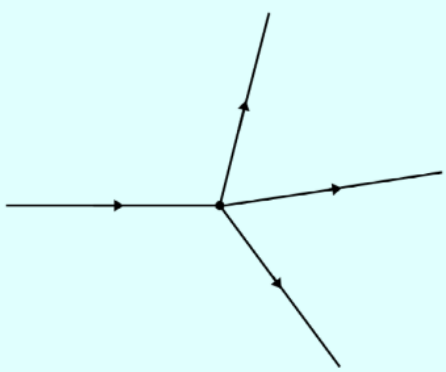
This is a possible type of scattering process in the theory of *quantum electrodynamics*.

Lastly, let's talk a bit about the diagrams with lines we draw above. These are actually really common in quantum field theories and they are called **Feynman diagrams**. A Feynman diagram is essentially just a pictorial way to represent an interaction or scattering process between particles in quantum field theory. They are not in any way *mandatory* for describing interaction processes, since all they really do is represent mathematical expressions in a pictorial way - they are just convenient for representing these processes in an intuitive and visual way.

Each Feynman diagram schematically represents a particular interaction process, which fundamentally is described by a term we have in some mathematical expression that describes a given interaction process. For example, we had above the possible first order self-interaction processes described by the field $\widetilde{\phi}(k)$:

$$\widetilde{\phi}(k) = \delta^4(k' - k) - \lambda G(k)\delta^4(3k' - k)$$

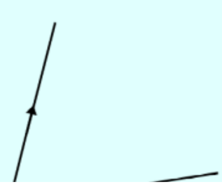
Both of these terms here can be represented as Feynman diagrams if we wanted to. In particular:



$$= -\lambda G(k)\delta^4(3k' - k)$$

$$= \delta^4(k' - k)$$

The real use of these Feynman diagrams is that we can actually do calculations using them in a much simpler way than with just standard equations. In quantum field theory, we are typically interested in calculating *scattering amplitudes* and this can be done simply by adding up all the Feynman diagrams of a given order of interaction. For example, schematically speaking, the scattering amplitude for a first-order self-interaction process in this particular example would be:



Okay, but how is "adding pictures" useful in any way for actual calculations? Well for that, we have a set of rules called the **Feynman rules**, which describe how to translate these Feynman diagrams to actual mathematical expressions. The Feynman rules are specific to each field theory and are used for doing calculations for that given theory.

4. Yukawa Interactions

Previously, we only talked about a single field that interacts with itself - a process called self-interaction. It is now time to talk about how different fields can interact with one another. Perhaps the simplest model for two fields that interact with one another is called the Yukawa interaction or Yukawa coupling.

Yukawa interactions can be used to describe how two scalar fields interact with one another (which we'll be studying in more detail soon), although more commonly, they are used to describe the interaction between a scalar field and a spinor field.

Yukawa interactions are also one very important for applications in particle physics. They are used to, for example, describe how the Higgs mechanism gives masses to otherwise massless fermions as well as for modeling the nuclear force between nucleons (although the detailed description of this is much more complicated).

At the most basic level, the Yukawa interaction term we include in our Lagrangian is just a product of the fields are interested in. For example, for two real scalar fields, their interaction term would be of the form $\mathcal{L}_{int} \propto \phi\psi$. However, the Yukawa interaction term is usually taken between a real scalar field and a complex scalar field (spinor, actually). Generally, it is of the form:

$$\boxed{\mathcal{L}_{Yukawa} = -g\phi\psi^\dagger\psi}$$

This describes the interaction between a real scalar field ϕ and a complex scalar field ψ. Note that we need $\psi^\dagger\psi$ (i.e. the norm of ψ, $|\psi|^2 = \psi^\dagger\psi$) here, because the Lagrangian should be real-valued and as such, we could not include just ψ if it is complex. The factor g here is called the coupling constant that describes the strength of the interaction between these fields. Its role in this theory is essentially the same as the role of the coupling constant λ in our self-interacting theory earlier.

There isn't much else to say about the Yukawa Lagrangian - it is essentially just the simplest interaction term that couples the real and complex scalar field. In most actual quantum field theory applications of Yukawa interactions, however, we would take ψ as a Dirac spinor instead of a scalar, so we are considering somewhat of a "toy model" here. Below, we will study this theory further.

4.1. Example: Interacting Scalar Fields

The full interacting theory of the real scalar field ϕ and the complex scalar field ψ is obtained by combining the free Lagrangians for these fields with the above Yukawa Lagrangian:

$$\mathcal{L} = \mathcal{L}_\phi + \mathcal{L}_\psi + \mathcal{L}_{Yukawa}$$

$$\Rightarrow \boxed{\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi + \partial_\mu\psi^\dagger\partial^\mu\psi - \frac{1}{2}\mu_\phi^2\phi^2 - \mu_\psi^2\psi^\dagger\psi - g\phi\psi^\dagger\psi}$$

The field equations for each field are then given by (note that the ψ-Lagrangian actually describes two different fields, ψ and $\psi^\dagger$, so they would each have their own field equations. However, for instructional purposes, it is enough for us to only consider the equation for $\psi^\dagger$ here, which gives us the dynamics of ψ):

$$\begin{cases} \partial_\mu \dfrac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} - \dfrac{\partial\mathcal{L}}{\partial\phi} = 0 \ \Rightarrow \ \partial_\mu\partial^\mu\phi + \mu_\phi^2\phi = -g\psi^\dagger\psi \\ \partial_\mu \dfrac{\partial\mathcal{L}}{\partial\left(\partial_\mu\psi^\dagger\right)} - \dfrac{\partial\mathcal{L}}{\partial\psi^\dagger} = 0 \ \Rightarrow \ \partial_\mu\partial^\mu\psi + \mu_\phi^2\psi = -g\phi\psi \end{cases}$$

These are two coupled field equations that we have to solve simultaneously for both fields. But how? Well, both equations have the form of a Klein-Gordon equation with potentials $V_1' = -g\psi^\dagger\psi$ and $V_2' = -g\phi\psi$ on the right-hand side. Therefore, the general solution to both equations are given in terms of their respective Green's functions (which are the same for both, actually, so $G_\phi = G_\psi = G$) as:

$$\begin{cases} \phi(x) = \displaystyle\int V_1(y)' G_\phi(x,y) d^4y = -g\int \psi^\dagger\psi G d^4y \\ \psi(x) = \displaystyle\int V_2(y)' G_\psi(x,y) d^4y = -g\int \phi\psi G d^4y \end{cases}$$

In this form, these solutions are not very useful as the fields themselves appear inside the integrals. However, just like before, we can try and see if the perturbation theory approach would fix this. Let's assume that both fields can be written as a power series, now in terms of the coupling constant g (instead of λ):

$$\begin{cases} \phi = \sum_{n=0}^{\infty} g^n \phi_n = \phi_0 + g\phi_1 + g^2\phi_2 + ... \\ \psi = \sum_{n=0}^{\infty} g^n \psi_n = \psi_0 + g\psi_1 + g^2\psi_2 + ... \end{cases}$$

Then, the above integral expressions become:

$$\begin{cases} \phi(x) = -g\int \psi^\dagger \psi G d^4y \\ \psi(x) = -g\int \phi\psi G d^4y \end{cases}$$

$$\Rightarrow \begin{cases} \phi_0 + g\phi_1 + g^2\phi_2 + ... \\ \quad = -g\int \left(\psi_0^\dagger + g\psi_1^\dagger + g^2\psi_2^\dagger + ...\right)\left(\psi_0 + g\psi_1 + g^2\psi_2 + ...\right) G d^4y \\ \psi_0 + g\psi_1 + g^2\psi_2 + ... \\ \quad = -g\int \left(\phi_0 + g\phi_1 + g^2\phi_2 + ...\right)\left(\psi_0 + g\psi_1 + g^2\psi_2 + ...\right) G d^4y \end{cases}$$

$$\Rightarrow \begin{cases} \phi_0 + g\phi_1 + g^2\phi_2 + ... = -\int \left(g\psi_0^\dagger\psi_0 + g^2\left(\psi_1^\dagger\psi_0 + \psi_0^\dagger\psi_1\right) + ...\right) G d^4y \\ \psi_0 + g\psi_1 + g^2\psi_2 + ... = -\int \left(g\phi_0\psi_0 + g^2(\phi_1\psi_0 + \phi_0\psi_1) + ...\right) G d^4y \end{cases}$$

Next, we simply match the powers of g on both sides, which gives us the following iterative solutions:

$$\begin{cases} \phi_1 = -\int \psi_0^\dagger\psi_0 G d^4y \\ \phi_2 = -\int \left(\psi_1^\dagger\psi_0 + \psi_0^\dagger\psi_1\right) G d^4y \\ \phi_3 = ... \end{cases} \quad \text{and} \quad \begin{cases} \psi_1 = -\int \phi_0\psi_0 G d^4y \\ \psi_2 = -\int (\phi_1\psi_0 + \phi_0\psi_1) G d^4y \\ \psi_3 = ... \end{cases}$$

Here, the fields ϕ_0 and ψ_0 are just the free real and complex scalar fields, which we know already. Therefore, we can find solutions iteratively using these relations - by knowing ϕ_0

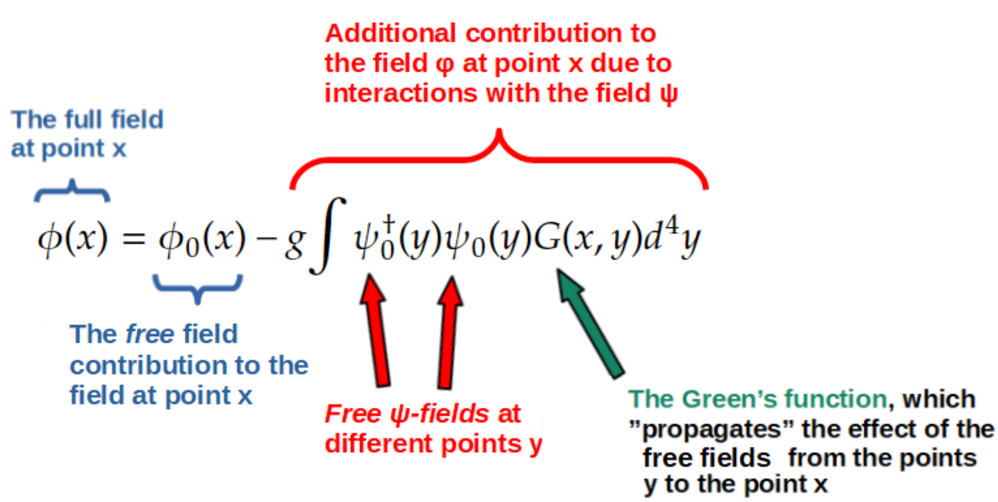
and ψ_0, we can calculate ϕ_1 and ψ_1, then by knowing ϕ_1 and ψ_1, we can calculate ϕ_2 and ψ_2 and so on, until infinity.

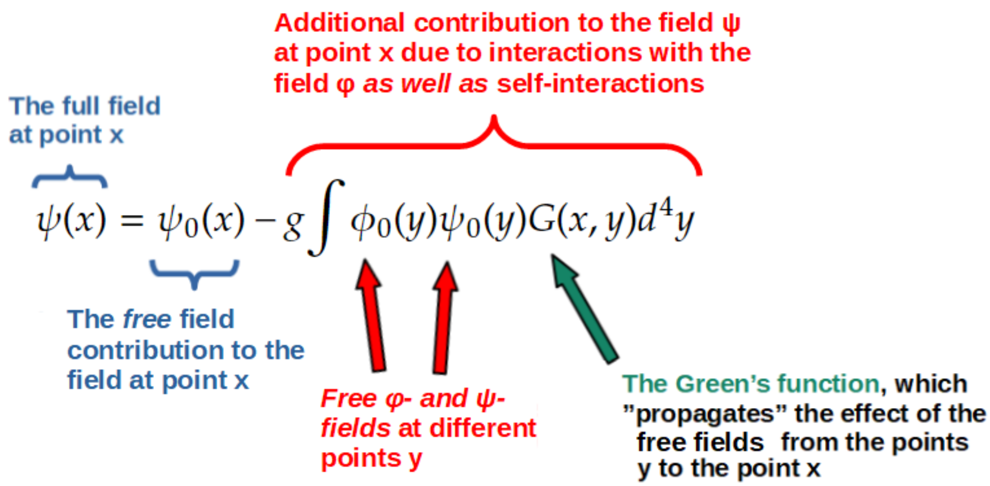
Of course, to actually obtain useful solutions, we have to cut the power series off somewhere - we cannot just calculate an infinite number of interative terms as that would be somewhat unpractical. The simplest interactive theory we could consider is one with only first-order interactions, so $\phi \approx \phi_0 + g\phi_1$ and $\psi \approx \psi_0 + g\psi_1$. Using our above iterative relations, the solutions to this theory are therefore given by:

$$\begin{cases} \phi = \phi_0 - g\int \psi_0^\dagger \psi_0 G d^4 y \\ \psi = \psi_0 - g\int \phi_0 \psi_0 G d^4 y \end{cases}$$

The interpretation of these is quite similar to the self-interacting theory - we have two fields that consist of free plane wave solutions ϕ_0 and ψ_0. However, due to the fields interacting through Yukawa interactions, we also have additional correction terms that arise from the fields at different points influencing each other. The Green's functions, once again, propagate these effects from one point to another.

For instance, the $\int \psi_0^\dagger \psi_0 G d^4 y$ describes how plane waves of the field ψ at different points affects the field ϕ. Notice that there are no self-interactions for the field ϕ in this case - only the other field, ψ, interacts with it. On the other hand, we notice from the other expression, $\int \phi_0 \psi_0 G d^4 y$, that the field ψ also self-interacts while interacting with ϕ as well.





$$\psi(x) = \psi_0(x) - g \int \phi_0(y)\psi_0(y)G(x,y)d^4y$$

To get a better intuitive understanding of what the Yukawa interactions actually describe, we will take a look at an example solution. We'll also do the same "Fourier transform trick" as before, which allows us to interpret the interactions in terms of particle processes.

Example: Yukawa Interactions In Terms of Particles

Let's begin with the simplest possible example solutions we could have for our free fields - single plane waves. Here, we'll assume the ϕ_0 plane waves and the ψ_0 plane waves have different momenta (k and q):

$$\begin{cases} \phi_0 = e^{-ik_\mu x^\mu} \\ \psi_0 = e^{-iq_\mu x^\mu} \\ \psi_0^\dagger = e^{iq_\mu x^\mu} \end{cases}$$

We can interpret these as describing free particles with momenta k, q and $-q$. The last one is curious - the field $\psi_0^\dagger$ describes a particle with the same momentum as ψ_0, but moving in the opposite direction.

We could, in principle, interpret this as the same kind of ψ-particle, but one that is *moving backwards through spacetime*. Well, we would actually interpret it as an antiparticle instead without needing to necessarily think about anything moving backwards in time, as that would be somewhat strange. However, the popular idea that "antiparticles are just particles moving backwards in time" comes exactly from this type of reasoning - particles and their antiparticles are related by complex conjugation of their respective fields.

Anyway, if we now write our (first-order) field solutions in terms of these plane waves, we find:

$$\begin{cases} \phi = \phi_0 - g\int \psi_0^\dagger \psi_0 G d^4y \\ \psi = \psi_0 - g\int \phi_0 \psi_0 G d^4y \end{cases} \Rightarrow \begin{cases} \phi = e^{-ik_\mu x^\mu} - g\int e^{iq_\mu y^\mu} e^{-iq_\mu y^\mu} G d^4y \\ \psi = e^{-iq_\mu x^\mu} - g\int e^{-ik_\mu y^\mu} e^{-iq_\mu y^\mu} G d^4y \end{cases}$$

Next, we want to insert our Green's function into this. It's the same Green's function for both cases (though with different masses for the fields, μ_ϕ and μ_ψ), given by:

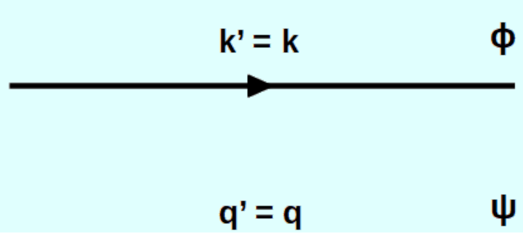
$$G = \frac{1}{(2\pi)^4}\int \frac{e^{ik'_\mu(x^\mu - y^\mu)}}{\mu^2 - k'_\mu k'^\mu} d^4k'$$

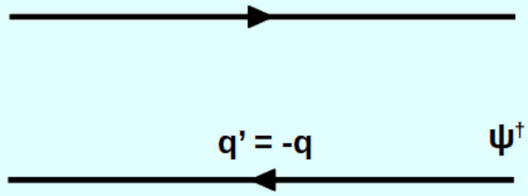
After this, we'll carry out the integration in full and then take the Fourier transform in the same way as we did previously for the self-interacting theory. You'll find the full calculation of this in the "Longer Calculations" -section, however, the answer we get is:

$$\begin{cases} \widetilde{\phi}(k) = \delta^4(k' - k) - \frac{g}{\mu_\phi^2}\delta^4(k) \\ \widetilde{\psi}(q) = \delta^4(q' - q) - \frac{g}{\mu_\psi^2 - q_\mu q^\mu}\delta^4(k' + q' - q) \\ \widetilde{\psi}^\dagger(q) = \delta^4(q' + q) - \frac{g}{\mu_\psi^2 - q_\mu q^\mu}\delta^4(k' + q' + q) \end{cases}$$

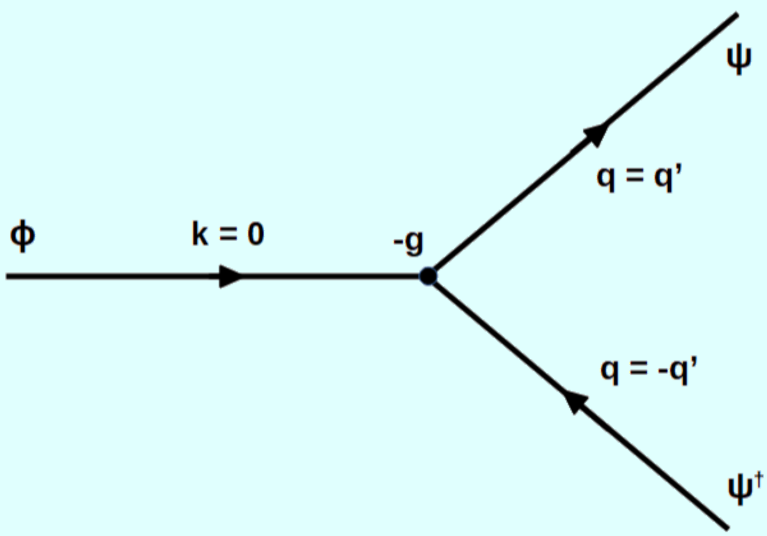
Notice that we find these delta functions again, which represent momentum conservation. This is a general feature in field theory - momentum conservation is built in through delta functions. But what kind of particle processes do these expressions represent?

First of all, since we have an interacting theory with two fields, each process must involve both of these fields simultaneously. The first possible process we can see here takes place when $k' = k$, $q' = q$ and $q' = -q$ (for the complex conjugate field). This simply represents all three fields describing free particles with no interaction taking place:

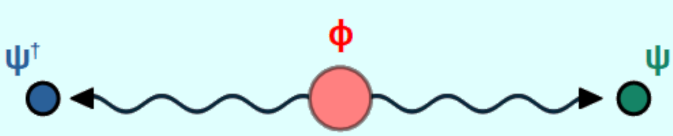




Another possible process here is an interaction process where $k = 0$, $q = k' + q'$ and for the complex conjugate field, $q = -k' - q'$. We can interpret this as describing a process in we first have a stationary ϕ-particle, because it has $k = 0$ (which also means that $k' = 0$). Then, an interaction (described by a factor $-g$) takes place and we end up with one ψ-particle with momentum q' and one $\psi^\dagger$-particle with momentum $-q'$:



This is an example of a *decay process* - a process in which one particle decays or "splits" into a set of different particles, while conserving momentum in the process. In this case, we have a stationary ϕ-particle decaying into a $\psi\psi^\dagger$-particle pair, which both must have equal and opposite momenta in order for the total momentum to be conserved. Essentially, we interpret this as describing the decay of a ϕ-particle into a ψ-particle-antiparticle pair. In Feynman diagrams, antiparticles are generally drawn with the arrows going *backwards* (like you see above).



An interesting thing to note is that since the Lagrangian has U(1) symmetry, we'll find the following conserved quantity:

$Q_\psi = N_a - N_b$

Here, N_a is the number of ψ-particles and N_b the number of associated antiparticles ($\psi^\dagger$-particles). This quantity is always conserved in any interaction process - for example, in the case we discussed above, we first have both $N_a = N_b = 0$, so $Q_\psi = 0$. After the decay of the ϕ-particle, we have $N_a = N_b = 1$, which still means that $Q_\psi = 0$. We can interpret this as the conservation of electric charge (since $\psi^\dagger$-particles have opposite electric charge to ψ-particles) - the possible processes that can take place must always conserve electric charge.

This is everything we're going to cover about Yukawa interactions for now. What we've discussed and the example we looked at is pretty much the simplest interacting field theory with more than one field. At a high level, the most important new idea that interacting field theories with more than a single field add are decay processes and particles of one type being able to turn into particles of a different type.

Now, we will stilll come back to Yukawa interactions for our discussion of spontaneous symmetry breaking and the Higgs mechanism later. So, keep what we discussed here at the back of your mind!

5. Gauge Interactions

Next, we will talk about **gauge interactions**. Gauge interactions are by far the most complicated of the interactions we've discussed so far, but also perhaps the most important. This is because gauge interactions describe how all other fields couple to *gauge fields* - which, in quantum field theory, describe forces between elementary particles. Therefore, gauge interactions essentially describe how all fundamental forces - the electromagnetic force, the weak force, the strong force and perhaps gravity as well - come about.

Here, we will talk about gauge interactions in the context of the electromagnetic field A_μ. For other gauge fields, the key ideas behind gauge interactions are exactly the same, just written in a bit more of a complicated way mathematically. So, what we discuss next serves as a great introduction and also directly explains how the theory of *quantum electrodynamics* comes about.

At the fundamental level, gauge interactions describe how a gauge field (i.e. a vector field that transforms in the form of a gauge transformation) interacts with other fields, such as scalar (or spinor) fields.

Since A_μ is a vector field, its interaction Lagrangian must contain a contraction over the index μ with some other four-vector. So, at the simplest possible level, the interaction Lagrangian describing how a vector field couples to other fields is of the form:

$$\boxed{\mathcal{L}_{int} = j^\mu A_\mu}$$

Here, j^μ is some four-vector that is a function of the other fields in our theory. To describe interactions, the four-vector j^μ thus has to contain the relevant fields we'd like to consider in our interactions and at the same time, must be a valid four-vector.

However, there is a glaring issue with this kind of interaction term - **it violates gauge invariance**! This is because any term involving A_μ directly is not invariant under a gauge transformation $A_\mu \ \rightarrow \ A_\mu + \partial_\mu \varphi$.

Now, we talked about why gauge invariance is important in the previous part already - it leads to directly observable physical consequences (such as longitudinal polarization being forbidden for photons), which we can easily verify in the real world. So, gauge invariance is certainly something we *need* to retain in the interacting theory as well if we are to describe a physically consistent theory. But how?

Luckily, there is actually a straightforward recipe for deriving the correct interaction term while still retaining gauge invariance. This is called the **minimal coupling procedure**, which we'll take a look at next.

5.1. Minimal Coupling

The notion of *minimal coupling* is essentially based on the idea that the electromagnetic field A_μ (and other gauge fields) couple to another field as a *connection*. A connection can be thought of as a geometric object that allows us to precisely compare local quantities at different points - in this case, it would be two locally gauge transformed fields. In minimally coupled theories, massless gauge fields are exactly these connections. We'll talk more about the geometry behind all of this a bit later.

The key idea is that the connection appears only through replacing ordinary derivatives by **gauge covariant derivatives** (denoted here by D_μ).

With this in mind, the **minimal coupling prescription** goes as follows:

1. Take the free Maxwell Lagrangian and the free Lagrangian of some other field ϕ and

combine them into a non-interacting Lagrangian:

$$\mathfrak{L}_{non-interacting} = \mathfrak{L}_{field} - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}$$

2. Replace all partial derivatives in $\mathfrak{L}_{field}$ by **gauge covariant derivatives**, defined as $D_\mu\phi = \partial_\mu\phi - iqA_\mu\phi$.

3. The result gives you the full Lagrangian with the correct interaction term:

$$\mathfrak{L}_{non-interacting} \xrightarrow{\partial_\mu\phi \rightarrow D_\mu\phi} \mathfrak{L} = \mathfrak{L}_{field} - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + j^\mu A_\mu$$

Again, the high-level idea here is that we first combine two Lagrangians into a non-interacting theory. Then, to retain gauge invariance in the full theory of both fields, we must introduce a *connection* through the gauge covariant derivative. This connection turns out to be exactly the electromagnetic field A_μ. The key is that the Lagrangian of the full theory that results from this approach is still gauge invariant - which is what we needed in the first place!

This is indeed how other fields couple to the electromagnetic field (to any gauge field, actually). All of this complicated stuff about connections and covariant derivatives comes from wanting to retain gauge invariance in our full, interacting theory.

Before looking at any practical examples, we'll dive into the intuition behind this entire approach of minimal coupling and specifically why it works. We'll also briefly discuss some of the geometry behind it and its relation to a few concepts from general relativity and differential geometry - there is actually a very close connection between these!

5.1.1. Why The Minimal Coupling Prescription Works

Perhaps the most intuitive way of understanding why the electromagnetic field couples to other fields through minimal coupling comes from looking at how electromagnetic fields couple to particles in the first place.

One way to view this is by thinking of the electromagnetic field as coupling to charged particles by introducing additional contributions to their energy and momenta. Specifically, a charged particle in an electromagnetic field has total energy $E = T - q\phi$, where T is its kinetic energy and ϕ the electric potential. So, the "electric potential -part" of an electromagnetic field adds to a particle's energy.

Similarly, the "magnetic vector potential -part" of the electromagnetic field adds to the particle's momentum. This can be seen by calculating the generalized momentum of a charged particle from its Lagrangian (the same Lagrangian we considered earlier), which turns out to be $\vec{p} = m\vec{v} + q\vec{A}$.

So, an electromagnetic field can be viewed as interacting with charged particles by contributing directly to their energy and momentum. We can actually combine both of these into a "more relativistic" statement in terms of the particle's **four-momentum** - an electromagnetic field interacts with a charged particle by changing its four-momentum as:

$$p_\mu \ \rightarrow \ p_\mu + qA_\mu$$

Okay, but where are we going with all of this? Well, it turns out that the minimal coupling prescription is *nothing but the field theoretical version* of this "electromagnetic contribution to a particle's four-momentum". Specifically, the p_μ here is the momentum of a particle described by another field ϕ - the field to which A_μ here couples to.

The key that relates these two together is the notion of viewing fields and particles as just two sides of the same coin. The thing that allows us to go from one description to another is the relation between a particle's four-momentum p_μ and the wave vector of the corresponding field k_μ - this relation is given by the **de Broglie relations**, $p_\mu = \hbar k_\mu$**,** like we discussed previously. For the rest of this section, we will set $\hbar = 1$ for convenience, so $p_\mu = k_\mu$.

Now, if this is the case, we would expect that the corresponding field describing the particle with momentum p_μ also couples to the electromagnetic field through its wave vector being modified as $k_\mu \ \rightarrow \ k_\mu + q\widetilde{A}_\mu$. Note that we are denoting the field here as $\widetilde{A}_\mu = \widetilde{A}_\mu(k)$ to distinguish it from the "actual" field that depends on spacetime, $A_\mu = A_\mu(x)$.

But how is this related to the gauge covariant derivative? Well, the answer comes from looking at how this coupling works in "physical space" through an inverse Fourier transform. The expression above for how k_μ changes takes place in "momentum space", but doesn't tell us anything about what happens in spacetime. In "physical space", this relation would be of the form (based on the Fourier transform rules from earlier):

$$k_\mu + q\widetilde{A}_\mu \ \ \Rightarrow \ \ i\partial_\mu + qA_\mu$$

Of course, this is nothing but the gauge covariant derivative multiplied by a factor of i, $iD_\mu = i(\partial_\mu - iqA_\mu)$. So, what is this telling us? Well, it tells us that the minimal coupling prescription is just the field theoretical "extension" of the fact that an electromagnetic field couples to charged particles by changing their energy and momenta. Specifically, it does this by the electromagnetic four-potential A_μ *appearing directly* in the particle's energy and momentum. This is the very reason why A_μ then couples to other fields specifically as a *gauge connection* through the gauge covariant derivative. From the particle side of things, electromagnetic coupling happens through a contribution to the particle's momentum, while from the perspective of fields, the coupling happens through a gauge connection - more on what this means soon!

Now, I want to just make clear the fact that this approach to deriving the gauge covariant derivative is sort of backwards. This is because the entire idea of the gauge covariant derivative is really based on the requirement of gauge invariance and fundamentally, that we should start with. This will be discussed next.

5.1.2. Geometry of The Gauge Covariant Derivative

If we take a brief step back here, all we are really trying to do is construct an interacting theory between a massless vector field and some other field. The idea with minimal coupling is that this can always be done by taking the non-interacting free Lagrangians of the respective fields and promoting them to an interacting theory by replacing partial derivatives to gauge covariant derivatives.

The gauge covariant derivative therefore describes the necessary interaction terms that arise when we couple a free field to a massless gauge field. In this sense, it is what "turns on" the interactions.

Now, the gauge covariant derivative is actually nothing but a consequence of **gauge invariance**. If we want to have a physically consistent theory, we *must* require that the full Lagrangian of the interacting theory also has gauge invariance. This is because we *know* that the massless vector field has gauge invariance, so any theory we build out of it should also have gauge invariance.

It's not that gauge invariance is just something we want to have in this theory for fun, it's that we require it for the interacting theory as well, since the massless vector field has gauge invariance already and we cannot change that fact. So, if we want a theory involving a massless vector field, **the theory must also come with gauge invariance** - there is simply no way around it.

It turns out that if gauge invariance is required, the minimal coupling prescription always results in the correct interaction terms that ensure gauge invariance. In fact, this is why the gauge covariant derivative is called what it's called. The reasoning behind this and a derivation of the gauge covariant derivative is explored in the section below.

Minimal Coupling From Gauge Invariance

Here, we will look at the reasoning behind minimal coupling through an example, which can be generalized to any other case in a very straightforward way - it's just easier to understand everything if we can see it concretely!

In particular, we'll by taking the non-interacting Lagrangians for a complex Klein-Gordon scalar field ϕ and for the gauge field A_μ (the Maxwell Lagrangian). It's worth noting that only complex scalar fields can actually couple to electromagnetism, real scalar fields cannot:

$$\mathcal{L} = \partial_\mu \phi^\dagger \partial^\mu \phi - \mu^2 \phi^\dagger \phi - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}$$

There are no interactions here yet. We will now see that interactions naturally fall in place if we require this Lagrangian to have gauge invariance. But doesn't this Lagrangian already have gauge invariance, since A_μ only appears through the gauge invariant field tensor $F_{\mu\nu}$?

Well, it actually doesn't. The key here is that we now have two different fields in our theory. A gauge transformation, fundamentally, is a transformation in the abstract "field space" where both of the fields ϕ and A_μ live. You could think of a gauge transformation (at least qualitatively) as a "coordinate transformation" in this abstract field space.

Because of this, both of the fields ϕ and A_μ will generally change under gauge transformations. We already know that A_μ changes as $A_\mu \ \rightarrow \ A_\mu + \partial_\mu \varphi$, but what about ϕ? Generally speaking, the gauge transformation acts on the field by the "local version" of the **global symmetry of the corresponding free field**. So, for a complex scalar field that has a global U(1) symmetry, the gauge transformation would act as a *local* U(1) transformation. This is simply a phase with the phase parameter being now the gauge function $\varphi = \varphi(x)$ with a constant factor q (a coupling constant) as well.

So, a gauge transformation changes the fields in our theory as:

$$\begin{cases} \phi \ \rightarrow \ e^{iq\varphi(x)} \phi \\ A_\mu \ \rightarrow \ A_\mu + \partial_\mu \varphi(x) \end{cases}$$

If you now look at the Lagrangian from above, it's pretty clear that it doesn't remain invariant under a gauge transformation of the form above due to the phase now begin spacetime-dependent. The issue arises from the partial derivatives in the term $\partial_\mu \phi^\dagger \partial^\mu \phi$.

Intuitively, the reason behind this comes from, well, the definition of what a gauge transformation is - a transformation that is different at every point in spacetime. The issue then is that a derivative, at a high level, always considers the difference between some quantity at two different points (though two very close-by points) - and because the field ϕ picks up a different phase due to gauge transformations at each point, comparing the field at two different points does not make sense due to this phase being different - we cannot meaningfully compare them.

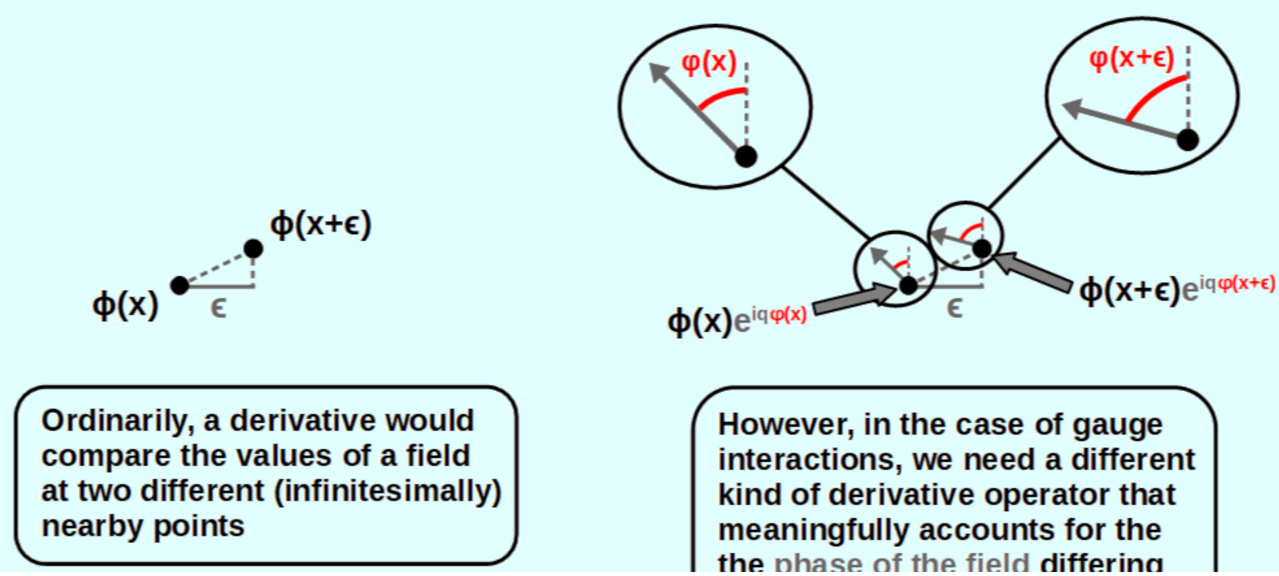
Intuitively, you can see this if you think about the derivative of the field as a "difference quotient" of the field at two different points, say x^μ and $x^\mu + \epsilon^\mu$:

$$\partial_\mu \phi \approx \frac{\phi(x+\epsilon) - \phi(x)}{x^\mu + \epsilon^\mu - x^\mu}$$

Now, since the field picks up a phase $e^{iq\varphi(x)}$ during the gauge transformation, the derivative would now be comparing the field at two different points, but also with different phases:

$$\partial_\mu \phi \approx \frac{e^{iq\varphi(x+\epsilon)}\phi(x+\epsilon) - e^{iq\varphi(x)}\phi(x)}{x^\mu + \epsilon^\mu - x^\mu}$$

This is why ordinary partial derivatives do not make sense anymore (at least in the context of gauge invariance) in a theory with gauge interactions - we would essentially be comparing apples to oranges.



from point to point!

Here, the arrows are used to represent complex phases $e^{iq\varphi}$, which are commonly known as phasors.

The interesting thing is that if we considered only the theory of the gauge vector field A_μ without any other field, then partial derivatives work just fine. This is because the field and its partial derivatives are only appearing through the field tensor $F_{\mu\nu}$, which is always gauge invariant. That's why we didn't have to care about any of this when considering the dynamics of only the gauge field. However, with interactions with a different field, things change a lot.

Now, to fix everything and make our interacting theory also have gauge invariance, we've identified that the issue lies with the partial derivatives - we therefore need something that allows us to compare the field at two different points regardless of the phase difference caused by the difference in the values of the gauge. But what would that something be?

Well, the approach we'll take is define a new derivative operator operator, D_μ, that essentially adds a correction term (call it Γ_μ) to the ordinary partial derivative:

$$D_\mu = \partial_\mu + \Gamma_\mu$$

We call this operator the gauge covariant derivative and the object Γ_μ a *connection* (if you're familiar with general relativity, there we also have connections, which are called Christoffel symbols). What the connection - which is really just a correction term - does here is that it "connects" the values of the field at two different points with different phases. The connection here allows us to compare the values of the field ϕ at different points regardless of the phase difference. So, in a sense, the gauge covariant derivative *ignores* the phase difference caused by gauge transformations and allows us to compute derivatives just like before - the only difference is that we need to include the additional gauge connection Γ_μ in our description.

Okay, but we do not know what this gauge connection is. How do we find this out? Well, we can look at how the gauge covariant derivative acts on the field ϕ under a gauge transformation $\phi \rightarrow e^{iq\varphi(x)}\phi$:

$$D_\mu \phi \rightarrow D_\mu\left(e^{iq\varphi(x)}\phi\right) = \partial_\mu\left(e^{iq\varphi(x)}\phi\right) + \Gamma_\mu e^{iq\varphi(x)}\phi$$

Let's write out this partial derivative using the product rule. This gives us:

$$D_\mu\left(e^{iq\varphi(x)}\phi\right) = e^{iq\varphi(x)}\partial_\mu\phi + iq\partial_\mu\varphi(x)e^{iq\varphi(x)}\phi + \Gamma_\mu e^{iq\varphi(x)}\phi$$
$$= (\partial_\mu + \Gamma_\mu + iq\partial_\mu\varphi(x))\phi e^{iq\varphi(x)}$$

Inside the parentheses here, we effectivaly have the *original* gauge covariant derivative $D_\mu = \partial_\mu + \Gamma_\mu$ plus an additional term, $iq\partial_\mu\varphi(x)$, which came from the gauge transformation. This means that generally, the gauge covariant derivative transforms under gauge transformations as $D_\mu \to D_\mu + iq\partial_\mu\varphi(x)$.

But what we need here is an object that is gauge invariant, so effectively $D_\mu \to D_\mu$. The way to achieve this is to demand that the connection Γ_μ transforms as $\Gamma_\mu \to \Gamma_\mu - iq\partial_\mu\varphi(x)$ as this would cancel out the extra gauge term in the covariant derivative. So, for gauge invariance, we require that the change in the connection due to a gauge transformation is $\delta\Gamma_\mu = -iq\partial_\mu\varphi(x)$. But what is an object that transforms this way?

Well, a massless gauge field transforms by the gradient of some scalar function under gauge transformations. So, we've found that the connection should transform under gauge transformations pretty much exactly the same way as the massless gauge field A_μ, which changes under gauge transformations by $\delta A_\mu = \partial_\mu\varphi(x)$. What we have now is a really interesting relation between the following transformation properties of these two objects under gauge transformations:

$$\delta\Gamma_\mu = -iq\partial_\mu\varphi(x) \quad \Leftrightarrow \quad \delta A_\mu = \partial_\mu\varphi(x)$$

Therefore, it's not really a far stretch to *define* the connection as $\Gamma_\mu = -iqA_\mu$. But is this enough for actually equating the two? Well, the connection is clearly some object that transforms like a massless gauge field and the only relevant massless gauge field we have in our theory is the electromagnetic potential A_μ. Since the electromagnetic potential fits so naturally here, this definition is certainly reasonable.

I guess the reasoning behind all of this would be best described by the phrase "if it looks like a duck, swims like a duck, and quacks like a duck, then it probably is a duck" - in other words, since we found something that behaves exactly like the gauge field A_μ, it's reasonable enough to conclude that *it is* the gauge field A_μ. Fundamentally, this is an assumption we make based on our findings here - you are free to accept it or not!

Again, the defining property of why we want the gauge covariant derivative is its property of effectively ignoring the phase difference in the field ϕ at two different points when computing derivatives. In fact, with our definition of the gauge connection as $\Gamma_\mu = -iqA_\mu$, the following property holds for the gauge covariant derivative:

$$D_\mu\left(e^{iq\varphi(x)}\phi\right) = e^{iq\varphi(x)}D_\mu\phi$$

In other words, the gauge covariant derivative allows us to compute derivatives *as if* the local U(1) gauge parameter $\varphi(x)$ was a constant, like we have in global U(1) transformations. Therefore, it's pretty clear that the Lagrangian from above is also gauge invariant if making the replacement $\partial_\mu\phi \rightarrow D_\mu\phi$:

$$\mathcal{L} = (D_\mu\phi)^\dagger D^\mu\phi - \mu^2\phi^\dagger\phi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}$$

$$\rightarrow \mathcal{L} = \left(D_\mu e^{iq\varphi(x)}\phi\right)^\dagger D^\mu\left(e^{iq\varphi(x)}\phi\right) - \mu^2\left(e^{iq\varphi(x)}\phi\right)^\dagger e^{iq\varphi(x)}\phi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}$$

$$= e^{-iq\varphi(x)}e^{iq\varphi(x)}(D_\mu\phi)^\dagger D^\mu(\phi) - \mu^2 e^{-iq\varphi(x)}e^{iq\varphi(x)}\phi^\dagger\phi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}$$

$$= (D_\mu\phi)^\dagger D^\mu(\phi) - \mu^2\phi^\dagger\phi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}$$

The bottom line here is that the above definition of the gauge covariant derivative as $D_\mu = \partial_\mu - iqA_\mu$ keeps everything gauge invariant like we required from the start.

So, with everything we found above, we define the gauge covariant derivative as:

$$\boxed{D_\mu = \partial_\mu - iqA_\mu}$$

Notice that if we now repalce partial derivatives by this according to the minimal coupling prescription, terms like $\partial_\mu\phi$ will become $D_\mu\phi = \partial_\mu\phi - iqA_\mu\phi$ - in other words, we find *interaction terms* that have the form of a product of the fields, $\sim A_\mu\phi$. This is exactly the sense in which minimal coupling actually *couples* the gauge field A_μ with other fields - by applying the minimal coupling prescription $\partial_\mu \rightarrow D_\mu$, we end up with interaction terms in the Lagrangian!

Now, one question you might have is why does the electric charge q appear here? Well, we don't necessarily know that it's the electric charge beforehand - it's just some coupling constant that describes how strongly the other field in our theory interacts with the gauge field (which, in this case, is the electromagnetic field).

For example, if $q = 0$, electromagnetic interactions would be "turned off" completely and the field ϕ also would not pick up any phase during the gauge transformation - telling us that it isn't interacting with the electromagnetic field in any way. However, as we'll see later, the coupling constant q here can really be interpreted as the electric charge.

Another quite interesting and also a much more general point to not here is specifically the way in which the electromagnetic field A_μ couples to the other field - as a *gauge connection*. In fact, this is not unique to the electromagnetic potential in any way, but is a general feature of all gauge theories - all gauge fields behave as connections for the gauge covariant derivative of the given interacting theory, and couple to other fields by appearing in the form of a gauge connection.

One last thing to mention - if you're familiar with general relativity, all this discussion about *covariant* derivatives may ring a bell for you. In general relativity, we also have a covariant derivative as well as a connection (called the Christoffel symbol). The idea behind these is pretty much the same as we discussed here with the gauge covariant derivative. This is explored more in the box below.

Advanced: Gauge Covariant Derivatives vs Covariant Derivatives In General Relativity

The main area of math used in general relativity is differential geometry. For our purposes, the important idea is that when taking derivatives of vectors, we have to account for how the basis vectors are changing due to the curvature of a space (or just in curvilinear coordinates in general). The connection - in this case called the Levi-Civita connection - is then what allows us to compare bases at different points in order to take derivatives correctly. Here, we will explore the analogy between this and the gauge covariant derivative discussed previously.

In differential geometry, when dealing with curvilinear coordinates or curved spaces, the derivative of a vector $\vec{v} = v^\alpha \vec{e}_\alpha$ is (this is a linear combination of the vector components v^α with the basis vectors $\vec{e}_\alpha$, which importantly, are not constant):

$$\partial_\mu \vec{v} = \partial_\mu v^\alpha \vec{e}_\alpha + v^\alpha \partial_\mu \vec{e}_\alpha$$

Here, $\partial_\mu \vec{e}_\alpha$ itself forms a set of new vectors for each value of α and μ - the components of these vectors can be obtained through dot products with the basis vectors, so $\vec{e}^\beta \cdot \partial_\mu \vec{e}_\alpha$. We call this set of vector components the **Christoffel symbols**. It is a three-index object labeled as $\Gamma^\beta_{\mu\alpha} = \vec{e}^\beta \cdot \partial_\mu \vec{e}_\alpha$ and it represents the *components of the vector formed by taking derivatives of basis vectors* - in this case, the β-component of the μ-derivative of the α-basis vector.

Now, since the Christoffel symbols are the components of a vector, we can also represent them as a linear combination in terms of basis vectors in the form:

$$\partial_\mu \vec{e}_\alpha = \Gamma^\beta_{\mu\alpha} \vec{e}_\beta$$

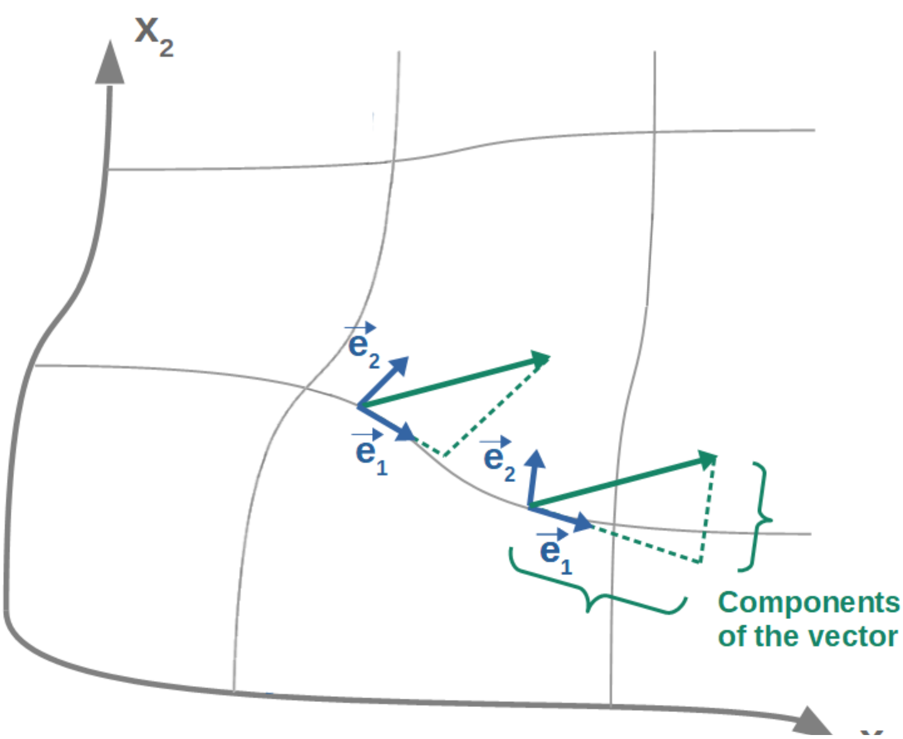
With this definition, we find that the derivative of the vector $\vec{v}$ is:

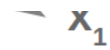
$$\partial_\mu \vec{v} = \partial_\mu v^\alpha \vec{e}_\alpha + v^\alpha \partial_\mu \vec{e}_\alpha = \partial_\mu v^\alpha \vec{e}_\alpha + v^\alpha \Gamma^\beta_{\mu\alpha} \vec{e}_\beta = \left(\partial_\mu v^\beta + \Gamma^\beta_{\mu\alpha} v^\alpha\right) \vec{e}_\beta$$

Here, we've factored out the basis vector $\vec{e}_\beta$ to also write this in the form of a linear combination of the components and basis vectors. Here, we see that the components formed by the vector $\partial_\mu \vec{v}$ are given by the expression inside the parentheses. We call this the covariant derivative, labeled as $\nabla_\mu v^\beta$:

$$\partial_\mu \vec{v} = \left(\partial_\mu v^\beta + \Gamma^\beta_{\mu\alpha} v^\alpha\right) \vec{e}_\beta = \left(\nabla_\mu v^\beta\right) \vec{e}_\beta$$

Taking a brief step back, the point in all of this is that the covariant derivative is needed when considering differentials in curvilinear coordinates, since the basis itself changes from point to point - the covariant derivative accounts for both how a vector is changing due to its components changing, but also how the vector is changing because the basis itself is changing.





Here, we have a vector in a curvilinear coordinate system. The schematic shows the components of the vector being different at different points, because the basis vectors $\vec{e}_1$ and $\vec{e}_2$ are different. The covariant derivative accounts for this when taking derivatives, as well as of course, for how the components may explicitly change in case they are functions of spacetime (not shown here).

In an exactly analogous way, the gauge covariant derivative accounts for both how a field is changing because the values of the field themselves are changing, but also how the field changes because the gauge parameter or phase is changing from point to point.

Because of this, both of them are called covariant derivatives - they allow us to construct *covariant* equations (discussed in Part 7) from derivatives, just in a slightly different way. The gauge covariant derivative D_μ ensures covariance under gauge transformations, while the differential geometry covariant derivative ∇_μ ensures covariance under general coordinate transformations.

In fact, the analogy is nearly exactly one-to-one if we were to think of the phase factor $e^{iq\varphi(x)}$ in the expression for the field, $e^{iq\varphi(x)}\phi$, as playing the same role as the basis vectors are in $\vec{v} = v^\alpha \vec{e}_\alpha$ - so, in a sense, the gauge factor $e^{iq\varphi(x)}$ is kind of like an abstract "curvilinear basis" for the field configurations ϕ that are related by a gauge transformation.

The interesting thing in what we discussed above are the connections. In the differential geometry case, the connection is called the **Levi-Civita connection** and it is represented by the Christoffel symbols $\Gamma^\beta_{\mu\alpha}$ (also called *connection coefficients*). It accounts for how the basis vectors vary from point to point.

In gauge theory, on the other hand, we have the **gauge connection**. The gauge connection is represented by the gauge field A_μ, which accounts for how the gauge parameter or phase varies from point to point. Both of these connections allow us to then construct a covariant derivative operator that plays the role of all differentiation in the given theory:

$$D_\mu = \partial_\mu - iqA_\mu \quad \Leftrightarrow \quad \nabla_\mu = \partial_\mu + \Gamma^\beta_{\mu\alpha}$$

The biggest distinction here, however, is that the gauge field or connection A_μ itself is a dynamical field in gauge theories - it has dynamics of its own and forms its completely own field theory. The Christoffel symbols, on the other hand, do not directly represent a dynamical field, but instead are built out of the *geometry of spacetime* itself (the metric tensor). This all, of course, makes sense only if you're familiar with general relativity.

There is also another, quite interesting analogy between the gauge covariant derivative and the covariant derivative from general relativity or differential geometry. This comes from considering the commutator of covariant derivatives, defined as $[\nabla_\mu, \nabla_\nu] = \nabla_\mu \nabla_\nu - \nabla_\nu \nabla_\mu$. Now, this is an operator that should act on something, namely a vector field. It turns out that for *any* vector field v^α, the following result is true:

$$[\nabla_\mu, \nabla_\nu] v^\alpha = R^\alpha_{\beta\mu\nu} v^\beta$$

Here, the four-index object $R^\alpha_{\beta\mu\nu}$ is called the **Riemann tensor**. The Riemann tensor fully describes how a geometric space (described by a given metric) is curved. In other words, it gives a full description of curvature in differential geometry - if all of its components are zero, the space is not curved (i.e. it is flat), but if any of its components are non-zero, the space is indeed curved. So, somehow, the non-commutativity of covariant derivatives describes the curvature of a geometric space.

Now, can we say something similar about the gauge covariant derivative? We indeed can! Let's consider the commutator of gauge covariant derivatives now, acting on some arbitrary field ϕ, and see what we find:

$$[D_\mu, D_\nu]\phi = (D_\mu D_\nu - D_\nu D_\mu)\phi$$

We can insert the definition $D_\mu = \partial_\mu - iqA_\mu$ into this and expand out everything:

$$\begin{aligned} [D_\mu, D_\nu]\phi &= (\partial_\mu - iqA_\mu)(\partial_\nu \phi - iqA_\nu \phi) - (\partial_\nu - iqA_\nu)(\partial_\mu \phi - iqA_\mu \phi) \\ &= \partial_\mu \partial_\nu \phi - iqA_\mu \partial_\nu \phi - iq\partial_\mu (A_\nu \phi) - q^2 A_\mu A_\nu \phi - \partial_\nu \partial_\mu \phi + iqA_\nu \partial_\mu \phi \\ &\quad + iq\partial_\nu (A_\mu \phi) + q^2 A_\nu A_\mu \phi \end{aligned}$$

Here, the double partial derivatives, $\partial_\mu \partial_\nu \phi$ and $\partial_\nu \partial_\mu \phi$, cancel because partial derivatives always commute. The same happens for the $q^2 A_\nu A_\mu \phi$-terms. We will also expand out the derivatives of the products of the form $\partial_\mu (A_\nu \phi)$ using the product rule, so that we're left with (the colored terms below also cancel each other):

$$\begin{aligned} [D_\mu, D_\nu]\phi &= -iqA_\mu \partial_\nu \phi - iqA_\nu \partial_\mu \phi - iq\phi \partial_\mu A_\nu + iqA_\nu \partial_\mu \phi + iqA_\mu \partial_\nu \phi \\ &\quad + iq\phi \partial_\nu A_\mu \\ &= -iq\phi \partial_\mu A_\nu + iq\phi \partial_\nu A_\mu \\ &= -iq(\partial_\mu A_\nu - \partial_\nu A_\mu)\phi \\ &= -iqF_{\mu\nu}\phi \end{aligned}$$

This is quite interesting - we've just found that the **electromagnetic field tensor** $F_{\mu\nu}$

represents the commutator of gauge covariant derivatives. In other words, only if all of its components are zero, the gauge covariant derivatives commute and otherwise, they do not.

This very closely resembles the differential geometry case, where the Riemann tensor represents the non-commutativity of covariant derivatives and hence, gives a full description of curvature. We could actually make a similar interpretation here - the non-commutativity of gauge covariant derivatives describes the "curvature of the field space" in which these electromagnetic interactions take place. Therefore, the electromagnetic field tensor somehow gives a measure of curvature, just like the Riemann tensor gives a measure of the curvature of a metric space. All in all, quite a far-reaching analogy, isn't it?

Next, we will look at an example of a theory with gauge interactions called **scalar electrodynamics**. The point of this is to show how everything discussed here actually works in practice. The nice thing is that scalar electrodynamics also works as a simplified model of a more complicated theory called **quantum electrodynamics** (QED), which is often said to be one of the most accurate theories in all of modern physics. So, the example we'll look at next also works as a fantastic introduction to the theory of QED!

5.2. Example: Scalar Electrodynamics

We will now briefly take a look at an example of constructing an interacting theory via gauge interactions. This example combines a complex scalar field ϕ with the electromagnetic gauge field A_μ to give the so-called theory of **scalar electrodynamics**.

This example works as a really good "toy model" for the full theory of *quantum electrodynamics*, which we would obtain simply by replacing the scalar field ϕ with a *spinor field* (and the Klein-Gordon Lagrangian with the *Dirac Lagrangian*). Thus, many of the (at least qualitative) aspects of quantum electrodynamics can be seen in perhaps in a more simple way by considering scalar electrodynamics.

Now, we already actually talked about scalar electrodynamics a bit when discussing the gauge covariant derivative. The simple idea is that we begin with the non-interacting complex
Klein-Gordon Lagrangian and the Maxwell Lagrangians, as the minimal coupling prescription tells us:

$$\mathcal{L} = \mathcal{L}_{KG} + \mathcal{L}_{Maxwell} = (\partial_\mu \phi)^\dagger \partial^\mu \phi - \mu^2 \phi^\dagger \phi - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}$$

Then, we "turn on" electromagnetic interactions by applying minimal coupling - partial derivatives $\partial_\mu \phi$ are replaced by gauge covariant derivatives $D_\mu \phi$. This gives us the

Lagrangian for scalar electrodynamics as:

$$\mathcal{L} = (D_\mu \phi)^\dagger D^\mu \phi - \mu^2 \phi^\dagger \phi - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}$$

Let's write out the gauge covariant derivatived here by using their definition, $D_\mu = \partial_\mu - iqA_\mu$, to see what this Lagrangian really represents. Doing this, we find that the Lagrangian can be written as:

$$\mathcal{L} = (\partial_\mu \phi - iqA_\mu \phi)^\dagger \left(\partial^\mu \phi - iqA^\mu \phi\right) - \mu^2 \phi^\dagger \phi - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}$$
$$= \partial_\mu \phi^\dagger \partial^\mu \phi + iqA_\mu \phi^\dagger \partial^\mu \phi - iqA^\mu \phi \partial_\mu \phi^\dagger + q^2 A_\mu A^\mu \phi^\dagger \phi - \mu^2 \phi^\dagger \phi - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}$$
$$= \partial_\mu \phi^\dagger \partial^\mu \phi - \mu^2 \phi^\dagger \phi - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \left(iq\left(\phi^\dagger \partial^\mu \phi - \phi \partial^\mu \phi^\dagger\right) + q^2 A^\mu \phi^\dagger \phi\right) A_\mu$$

Notice the form of this Lagrangian - we have the original, free Klein-Gordon and Maxwell Lagrangians as well as an additional term of the form $(\dots)^\mu A_\mu$. This additional term contains all the interactions and is exactly of the form $\mathcal{L}_{int} = j^\mu A_\mu$, which we postulated the form of the interaction Lagrangian to be at the very beginning of this section. Therefore, the minimal coupling presecription has automatically given us the correct interaction Lagrangian with the correct j^μ, in this case as:

Free Maxwell Lagrangian

$$\mathcal{L} = \underbrace{\partial_\mu \phi^\dagger \partial^\mu \phi - \mu^2 \phi^\dagger \phi}_{\text{Free Klein-Gordon Lagrangian}} - \overbrace{\frac{1}{4} F_{\mu\nu} F^{\mu\nu}}^{\text{Free Maxwell Lagrangian}} + \underbrace{j^\mu A_\mu}_{\text{Interaction Lagrangian}}$$

Free Klein-Gordon Lagrangian

Interaction Lagrangian, where:

$$j^\mu = iq\left(\phi^\dagger \partial^\mu \phi - \phi \partial^\mu \phi^\dagger\right) + q^2 A^\mu \phi^\dagger \phi$$

So, through the minimal coupling prescription, we've obtained the full Lagrangian for scalar electrodynamics that, among other things is **gauge invariant**, and fully describes how a complex scalar field ϕ interacts with the electromagnetic field A_μ.

Let's now look at the field equations for both ϕ and A_μ that we obtain from this Lagrangian.

First, we'll look at the Euler-Lagrange equation for $\phi^\dagger$, which gives us the field equation for ϕ (since the fields ϕ and $\phi^\dagger$ are related through complex conjugation, it's enough for us to look at just one of them):

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \phi^\dagger\right)} - \frac{\partial \mathcal{L}}{\partial \phi^\dagger} = 0$$

$$\Rightarrow \quad \partial_\mu\left(\partial^\mu \phi - iqA^\mu \phi\right) - \left(iqA_\mu \partial^\mu \phi + q^2 A_\mu A^\mu \phi - \mu^2 \phi\right) = 0$$

$$\Rightarrow \quad \partial_\mu \partial^\mu \phi - iq\partial_\mu\left(A^\mu \phi\right) - iqA_\mu \partial^\mu \phi - q^2 A_\mu A^\mu \phi + \mu^2 \phi = 0$$

$$\Rightarrow \quad \partial_\mu \partial^\mu \phi - iq\left(\phi \partial_\mu A^\mu + A^\mu \partial_\mu \phi\right) - iqA_\mu \partial^\mu \phi - q^2 A_\mu A^\mu \phi + \mu^2 \phi = 0$$

$$\Rightarrow \quad \partial_\mu \partial^\mu \phi + \mu^2 \phi = iq\phi \partial_\mu A^\mu + 2iqA_\mu \partial^\mu \phi + q^2 A_\mu A^\mu \phi$$

As you can see, the field equation for ϕ here is quite complicated. On the left-hand side, we have the usual Klein-Gordon equation. On the right-hand side, however, we have a complicated source term that involves both A_μ and ϕ as well as the derivatives $\partial_\mu A^\mu$ and $\partial_\mu \phi$.

Now, we can actually clean this expression up a bit by writing it in terms of the gauge covariant derivatives. This is easiest done by taking the field from above in the following form:

$$\partial_\mu\left(\partial^\mu \phi - iqA^\mu \phi\right) - iqA_\mu \partial^\mu \phi - q^2 A_\mu A^\mu \phi + \mu^2 \phi = 0$$

$$\Rightarrow \quad \partial_\mu\left(\partial^\mu \phi - iqA^\mu \phi\right) - iqA_\mu\left(\partial^\mu \phi - iqA^\mu \phi\right) + \mu^2 \phi = 0$$

$$\Rightarrow \quad \left(\partial_\mu - iqA_\mu\right)\left(\partial^\mu \phi - iqA^\mu \phi\right) + \mu^2 \phi = 0$$

We can now write this in terms of gauge covariant derivatives $D_\mu = \partial_\mu - iqA_\mu$ as:

$$\boxed{D_\mu D^\mu \phi + \mu^2 \phi = 0}$$

This is quite a nice way of writing the field equation as it really makes explicit the minimal coupling between the fields ϕ and A_μ - the coupling always happens just by partial derivatives getting replaced by gauge covariant derivatives. We see here that this is exactly the free Klein-Gordon equation, but with the minimal coupling rule $\partial_\mu \rightarrow D_\mu$ taking place.

Next, let's compute the field equation for A_μ:

$$\partial_\nu \frac{\partial \mathcal{L}}{\partial(\partial_\nu A_\mu)} - \frac{\partial \mathcal{L}}{\partial A_\mu} = 0$$

$$\Rightarrow \quad \partial_\nu F^{\mu\nu} - \left(iq\phi^\dagger\partial^\mu\phi - iq\phi\partial^\mu\phi^\dagger + 2q^2A^\mu\phi^\dagger\phi\right) = 0$$

$$\Rightarrow \quad \partial_\nu F^{\mu\nu} = iq\left(\phi^\dagger\partial^\mu\phi - \phi\partial^\mu\phi^\dagger\right) + 2q^2A^\mu\phi^\dagger\phi$$

We can once again write the right-hand side a bit more cleanly in terms of the covariant derivatives:

$$\partial_\nu F^{\mu\nu} = iq\left(\phi^\dagger\partial^\mu\phi - \phi\partial^\mu\phi^\dagger\right) + 2q^2A^\mu\phi^\dagger\phi$$

$$\Rightarrow \quad \partial_\nu F^{\mu\nu} = q\left(i\phi^\dagger\left(\partial^\mu\phi - iqA^\mu\phi\right) - i\phi\left(\partial^\mu\phi^\dagger + iqA^\mu\phi^\dagger\right)\right)$$

$$\Rightarrow \quad \boxed{\partial_\nu F^{\mu\nu} = iq\left(\phi^\dagger D^\mu\phi - \phi\left(D^\mu\phi\right)^\dagger\right)}$$

This is the field equation describing the dynamics of the electromagnetic field A_μ. It's quite a complicated one, since the source term $iq\left(\phi^\dagger D^\mu\phi - \phi\left(D^\mu\phi\right)^\dagger\right)$ depends on both ϕ, but also on A_μ itself.

There is actually an interesting result hiding in this field equation. We can find this out by taking the derivative on both sides and noting that $\partial_\mu\partial_\nu F^{\mu\nu} = 0$ always due to the antisymmetricity of $F^{\mu\nu}$:

$$\partial_\mu\partial_\nu F^{\mu\nu} = \partial_\mu\left(iq\left(\phi^\dagger D^\mu\phi - \phi\left(D^\mu\phi\right)^\dagger\right)\right) \quad \Rightarrow \quad \partial_\mu\left(iq\left(\phi^\dagger D^\mu\phi - \phi\left(D^\mu\phi\right)^\dagger\right)\right) = 0$$

But what is this? Well, it's a conservation law! In particular, it is the conservation of a four-current $J^\mu = iq\left(\phi^\dagger D^\mu\phi - \phi\left(D^\mu\phi\right)^\dagger\right)$ in the form $\partial_\mu J^\mu = 0$. In fact, this J^μ is exactly the **conserved Noether current in scalar electrodynamics** resulting from U(1) symmetry. The conservation of this current represents the conservation of total electric charge in this theory. This is explored in the example box below.

Interpretation of The Conserved Current In Scalar Electrodynamics

Okay, we have our conserved current in this theory as:

$$J^\mu = iq\left(\phi^\dagger D^\mu\phi - \phi\left(D^\mu\phi\right)^\dagger\right) = iq\left(\phi^\dagger\partial^\mu\phi - \phi\partial^\mu\phi^\dagger\right) + 2q^2A^\mu\phi^\dagger\phi$$

To see what this has to do with electric charge, we can look at an example solution. In particular, let's take the simplest possible configuration for the field ϕ - a simple plane wave, so that:

$$\phi = ae^{-ik_\mu x^\mu}$$
$$\phi^\dagger = a^\dagger e^{ik_\mu x^\mu}$$

In the particle world, these would represent ϕ-particles with momentum k_μ. Now, after some straightforward calculations, we find the conserved current to be for this case:

$$J^\mu = 2q\left(k^\mu + qA^\mu\right)a^\dagger a$$

In quantum field theory, the quantity $a^\dagger a$ is defined as the number operator N_a, which represents the number of ϕ-particles (we discussed this in Part 8 already). Also, the quantity $k^\mu + qA^\mu$ is the *total generalized momentum* of the particles we are describing here (i.e. the momentum $p^\mu \propto k^\mu$ plus the contribution coming from the coupling with the electromagnetic field, qA^μ). The current therefore has the form:

$$J^\mu \sim 2qN_a p^\mu$$

The conserved charge related to the 0-component of this current would then be proportional to:

$$Q = \int J^0 d^3x \propto \int qN_a d^3x$$

Now, since Q here is conserved according to Noether's theorem (here, conserved meaning $dQ/dt = 0$), it means that N_a - the number of particles does not change with time. This is nothing but the conservation of electric charge - if we have some number N_a of particles initially, then the same number of particles remains at all times. This also means that we can interpret q as the electric charge of a single particle, so that the total charge qN_a is conserved. Note, however, that in this example, we have N_b - the number of *antiparticles* - as zero. In the more general case, the conserved quantity would be related to $q(N_a - N_b)$, which is the total charge of all particles plus the total charge of all antiparticles (just like we had in Part 8 as well).

Here, we won't be discussing the solutions to scalar electrodynamics any further. The point here was to mainly illustrate the minimal coupling prescription and how it is used to obtain interacting theories for gauge fields. To conclude our discussion, below are presented the key results we found by applying minimal coupling to obtain the theory of scalar eletrodynamics.

Scalar electrodynamics

= The theory of a complex scalar field *minimally coupled* to the electromagnetic field

- Lagrangian:

$$\mathcal{L} = (D_\mu \phi)^\dagger D^\mu \phi - \mu^2 \phi^\dagger \phi - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}$$

- Field equations for ϕ:

$$D_\mu D^\mu \phi + \mu^2 \phi = 0$$

- Field equations for A_μ:

$$\partial_\nu F^{\mu\nu} = J^\mu$$

- Conserved U(1) current (appears as a source in the electromagnetic field equations):

$$J^\mu = iq\left(\phi^\dagger D^\mu \phi - \phi\left(D^\mu \phi\right)^\dagger\right)$$

6. Spontaneous Symmetry Breaking

In this section, which is also the last one in this Part, we discuss the concept of spontaneous symmetry breaking. Spontaneous symmetry breaking is an important topic in many different areas of physics, including quantum field theory, the theory of superconductors and so many more.

At a very high level, spontaneous symmetry breaking is the phenomenon of a particular system exhibiting *less* symmetry than the physical laws describing that system. In more practical terms, spontaneous symmetry breaking occurs when a particular solution to a set of equations of motion does not have the same symmetry as the equations of motion itself - if this is the case, we say that that solution has spontaneously broken some symmetry of the equations of motion.

In this section, we will talk about spontaneous symmetry breaking mainly in the context of field theory and particle physics. The solutions that spontaneously break a given symmetry of
the full equations of motion (or Lagrangian) we are interested in looking at here are going to be **vacuum solutions**, which are states of a field that minimize its energy. But more on this soon!

The plan for this section goes more or less as follows:

- We begin by discussing spontaneous symmetry breaking in general terms and why it is important.

- We then begin studying it in the context of field theory by looking at the complex scalar field theory - which is also the only type of field in the Standard Model that exhibits spontaneous symmetry breaking!

- After this, we dive into the Higgs mechanism - the description of how fields and particles can gain mass through spontaneous symmetry breaking. A few words about the Higgs mechanism in the Standard Model of particle physics are also given.

The key idea this section revolves around is that the Higgs field, a particular complex scalar field, naturally "likes" to sit at the minimum of its potential energy (as does pretty much everything in the universe). The crucial thing is that this minimum value of the field is not zero and these non-zero vacuum solutions exhibit **spontaneous symmetry breaking**.

Around this minimum of its potential, the Higgs field behaves like two non-interacting real scalar fields; a massive and a massless field. The particles describes by these fields are the **Higgs boson** and the *massless* **Goldstone boson**.

Other fields interacting with the Higgs field can then gain masses in two ways:

- Fields interacting with the Higgs field through Yukawa interactions gain their masses directly by coupling to the non-zero vacuum of the Higgs field.

- Gauge fields interacting with the Higgs field through gauge interactions gain their masses by "eating" the Goldstone boson of the Higgs field - intuitively, the Goldstone boson "turns into" a non-zero mass for the gauge field.

Let's dive into how all of this comes about!

6.1. What Is Spontaneous Symmetry Breaking?

The general idea behind spontaneous symmetry breaking has to do with discrepancies between the symmetries of a set of equations of motion (or equivalently, the Lagrangian) and the symmetries of the solutions to those equations of motion. In fact, we can spontaneous symmetry breaking as follows:

> *If a given solution to a set of equations of motion does not have the same symmetry as its corresponding equations of motion, the symmetry has been spontaneously broken.*

An example case of spontaneous symmetry breaking could be the Klein-Gordon equation (an equation of motion) having U(1) symmetry, as we very well know already, but a given plane wave solution to the Klein-Gordon equation does not.

Mathematically, U(1) of the Klein-Gordon equation means that we can multiply the field ϕ by an arbitrary (constant) phase, $\phi \rightarrow \phi e^{i\alpha}$. The Klein-Gordon equation still remains invariant:

$$\partial_\mu \partial^\mu \phi + \mu^2 \phi = 0$$
$$\rightarrow \ \partial_\mu \partial^\mu \left(\phi e^{i\alpha}\right) + \mu^2 \phi e^{i\alpha} = 0$$
$$\Rightarrow \ \left(\partial_\mu \partial^\mu \phi + \mu^2 \phi\right) e^{i\alpha} = 0$$
$$\Rightarrow \ \partial_\mu \partial^\mu \phi + \mu^2 \phi = 0$$

However, let's say we've found a plane wave solution of the form $\phi = e^{ik_\mu x^\mu}$ to the Klein-Gordon equation. Then, multiplication by a phase does not leave the solution strictly invariant
anymore:

$$\phi = e^{ik_\mu x^\mu} \ \rightarrow \ \phi = e^{i\left(k_\mu x^\mu + \alpha\right)}$$

Sure, the solution still has the form of a plane wave, but it is not *exactly* the same solution. In other words, this specific solution does not have U(1) symmetry anymore, even though the corresponding Klein-Gordon equationwe obtained it from does.

But where has the symmetry gone? Well, it has been spontaneously broken - this particular plane wave solution spontaneously broke U(1) symmetry, however, it's important to note that the underlying full theory (the Klein-Gordon theory) still has U(1) symmetry, so it hasn't "truly" been broken - hence the name *spontaneous* symmetry breaking and not *explicit* symmetry breaking. Spontaneous symmetry breaking is a feature of specific solutions to equations of motion, not of the equations of motion themselves.

This is fundamentally what it means for symmetries to be spontaneously broken. If there exists a symmetry in a given set of equations of motion, then in principle, there can exist solutions that spontaneously break that symmetry.

Now, when talking about spontaneous symmetry breaking here (and typically in field theory), we are mainly interested in the **spontaneous symmetry breaking of so-called vacuum solutions to field equations**. These are field configurations that correspond to the minima of a given potential, so they minimize the energy of the field. These vacuum solutions can also exhibit spontaneous symmetry breaking like we saw with the plane wave above.

Importantly, it turns out that a given vacuum solution exhibiting spontaneous symmetry breaking can act like a mass contribution for other fields it interacts with. This is the general idea behind the **Higgs mechanism** - the spontaneous symmetry breaking of the Higgs field's vacuum generates masses for other fields it interacts with (or particles).

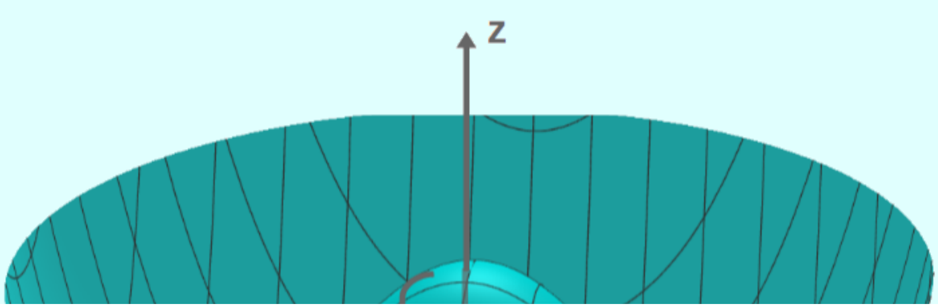
But before getting to any of that, we will look at an example of spontaneous symmetry breaking in ordinary Lagrangian mechanics. This example will look nearly exactly analogous as the spontaneous symmetry breaking of the Higgs field that we'll get to soon, just in a classical mechanics context. So, make sure you understand what is going on here!

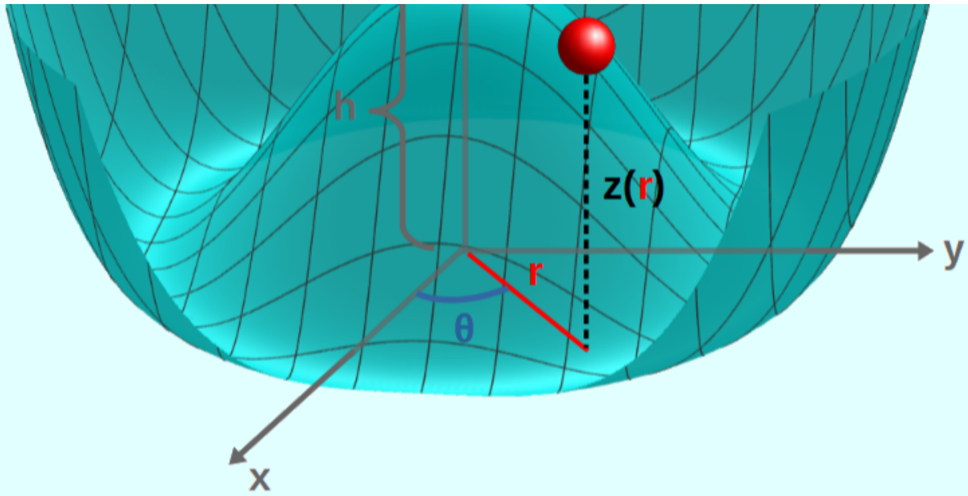
Example of Spontaneous Symmetry Breaking In Classical Mechanics

To introduce the idea of spontaneous symmetry breaking, let's consider a simple example from ordinary Lagrangian mechanics, which also turns out to be quite useful later on. The example we will look at is a ball rolling along a hill in three dimensions that has the shape of a "Mexican hat". We'll describe this hill mathematically in cylindrical coordinates (r, θ, z), such that the height z of the ball, as a function of the coordinate r, is given by:

$z(r) = h - \frac{1}{2}kr^2 + \frac{1}{4}\lambda r^4$, where k, λ and h are constants.

So, the height of the hill is given by this function $z(r)$ and since the ball is moving along the hill at all times, this will also be the height of the ball. The h here is the height of the hill located at $r = 0$, so $z(0) = h$. We can visualize this as follows:





The potential energy of the ball, as measured from the xy-plane, is given by:

$$V = mgz(r) = mgh - \frac{1}{2}mgkr^2 + \frac{1}{4}mg\lambda r^4$$

Therefore, the Lagrangian for the ball is (the kinetic energy here being expressed in cylindrical coordinates as well):

$$L = T - V = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2 + \dot{z}^2\right) - mgh + \frac{1}{2}mgkr^2 - \frac{1}{4}mg\lambda r^4$$

Note that we also have to insert $z(r)$ in the $\dot{z}^2$ term here. Doing so gives us:

$$L = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2 + \dot{z}^2\right) - mgh + \frac{1}{2}mgkr^2 - \frac{1}{4}mg\lambda r^4$$
$$= \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2 + \left(mgkr - mg\lambda r^3\right)^2\dot{r}^2\right) - mgh + \frac{1}{2}mgkr^2 - \frac{1}{4}mg\lambda r^4$$

With this, we've effectively eliminated the z-coordinate here. This is because the $z(r)$ specified earlier has the form of a constraint that allows us to express all z's in terms of r.

Now, here comes the important part. We can see that this Lagrangian has axisymmetry - that is, rotations about the z-axis, so shifts in the θ-coodinate of the form $\theta \rightarrow \theta + \delta\theta$ are symmetries. This is pretty clear from the fact that the height of the ball is not affected at all by the value of the θ-coordinate.

Let's consider now the equations of motion for the particle:

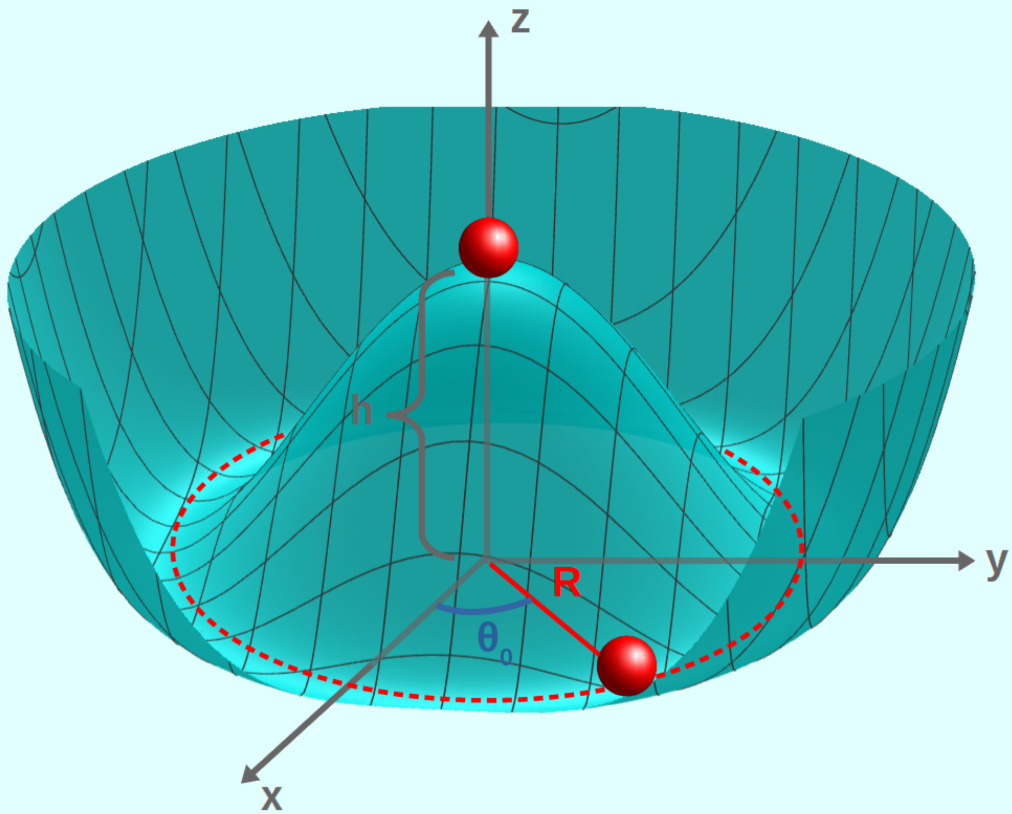
$$\frac{d}{dt}\frac{\partial L}{\partial \dot{r}} - \frac{\partial L}{\partial r} = 0 \quad \Rightarrow \quad \ddot{r} = \frac{\left(1 - mr^{2}\left(mgk - 3mg\lambda r^{2}\right)\right)\left(gkr - g\lambda r^{3}\right) + r\dot{\theta}\ldots}{1 + \left(mgkr - mg\lambda r^{3}\right)^{2}}$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}} - \frac{\partial L}{\partial \theta} = 0 \quad \Rightarrow \quad \ddot{\theta} = -\frac{2}{r}\dot{r}\dot{\theta}$$

The r-equation looks quite complicated, but let's not let that distract us. Importantly, we see again that both of these equations of motion are invariant under shifts of the coordinate θ.

Let's now consider solutions to these equations of motion. The ones we are interested here are constant solutions - ones where the ball can stay at a single location without moving.

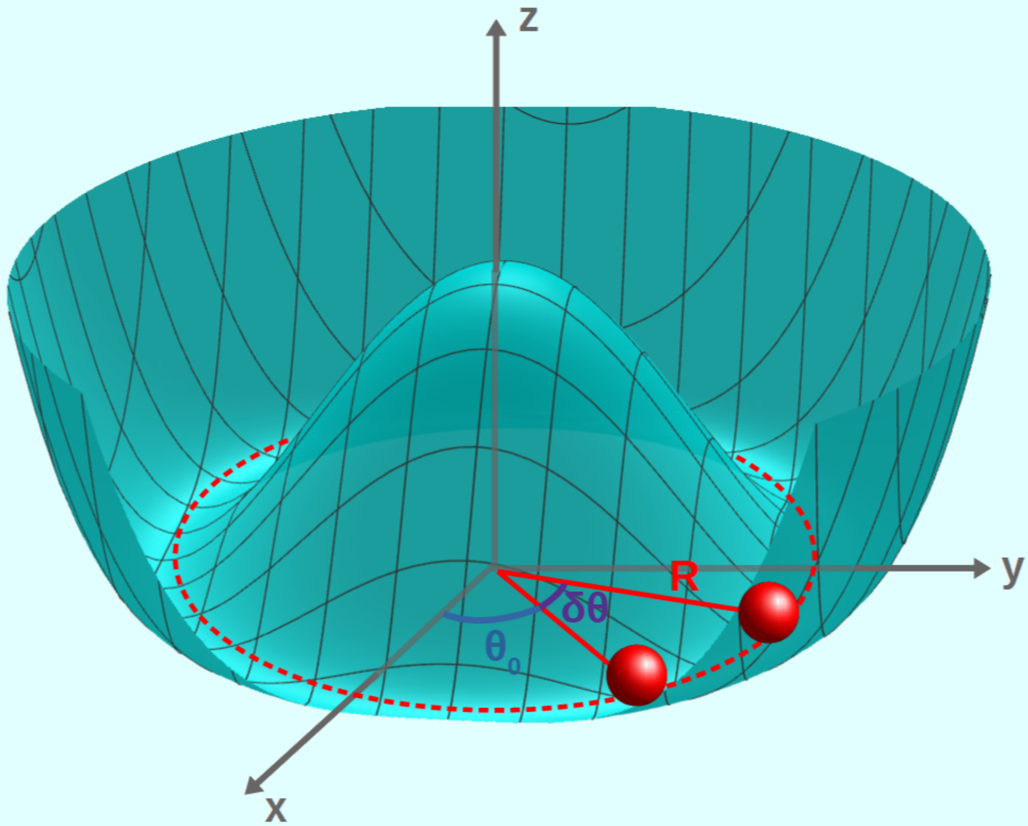
If we look at the "hill" picture from above, we might expect that there are two possible locations where the ball can stay without moving - the top of the hill at height h and $r = 0$ and the bottom of the "valley" at some constants $r = R$ and $\theta = \theta_0$:



If we plug in $r = 0, \theta = \theta_0$ as well as $r = R, \theta = \theta_0$, we indeed find that they are solutions of the equations of motion. We also find the value of R to be $R = \sqrt{k/\lambda}$ - this is the radius at which the particle can stay at the bottom of the valley.

Now, for the important thing for us - what are the symmetries of these solutions, *as compared to the symmetries of the equations of motion*? Clearly, for the $r = 0, \theta = \theta_0$ solution, rotations are still a symmetry. This is because shifting θ_0 by $\theta_0 \rightarrow \theta_0 + \delta\theta$ doesn't change the position of the ball at all (it is still at the top of the hill above), so it corresponds to exactly the same solution. So, the solution describing the ball sitting on top of the hill has the **same symmetry as the equations of motion** - nothing interesting to be found here.

However, for the $r = R, \theta = \theta_0$ solution, things get much more interesting. Clearly, if we shift θ_0 as $\theta_0 \rightarrow \theta_0 + \delta\theta$, then the ball is NOT at the same location anymore - it is still at the bottom of the valley at the same radius R but at a different point:



Okay, but what does this mean? Well, since the ball is now at a different location after a shift in θ - which corresponds to a different solution to the equations of motion - it means that **rotations are NOT symmetries of this solution anymore**. To reiterate; solutions where the ball is sitting at the minimum, lowest point on the hill, do not have the same symmetries as the equations of motion.

So somehow, we've gone from equations of motion with a particular symmetry to solutions to those equations of motion that do not have the same symmetry - it's as if the symmetry has been "lost" or perhaps, *broken*. Indeed, this is exactly the phenomenon of **spontaneous symmetry breaking**; if a given solution does not have the same symmetry as the corresponding equations of motion, then that symmetry has been spontaneously broken.

In our example here, we say that the axisymmetry or rotational symmetry has been spontaneously broken by this minimum of the potential - even though the equations of motion have axisymmetry, this particular solution corresponding to the minimum of the potential does not.

It's worth noting that spontaneous symmetry breaking is just a feature of a particular configuration or solution to the equations of motion - the system as a whole still has all the same symmetries (as described by the full equations of motion of Lagrangian), but a particular solution may not. In comparison, *explicit* symmetry breaking would mean that a symmetry has been broken also in the equations of motion themselves (which is not the case here!).

The reason we considered the particular example we did above is that pretty much the exact same kind of thing happens in the spontaneous symmetry breaking for the **Higgs field** that we will get to soon.

For the Higgs field (which is a complex scalar field), its potential looks exactly like the one above and the field naturally likes to "fall into" the minimum of the potential. Then, this particular field configuration that is at the minimum of the potential spontaneously breaks the U(1) symmetry exactly like the solution in the example above spontaneously broke the rotational symmetry.

One more additional thing to note here is that spontaneous symmetry breaking can only happen to "actual" symmetries, which are always **global symmetries** (i.e. transformation by a *constant* parameter). Local symmetries - gauge "symmetries" - CANNOT be spontaneously broken as they are not even real symmetries, strictly speaking. There is some talk I've seen out there about the Higgs mechanism being related to spontaneous symmetry breaking of gauge symmetry (gauge invariance).

So, it's important to mention that if you hear something like that, just know it is not correct! In fact, there is a theorem that proves gauge symmetries cannot be spontaneously broken, which goes by the name of *Elitzur's theorem*.

6.2. Spontaneous Symmetry Breaking For The Complex Scalar Field

We will begin our discussion of spontaneous symmetry breaking and the Higgs mechanism by looking at the simplest theory that exhibits spontaneous symmetry breaking of vacuum solutions - a self-interacting scalar field, specifically, a complex scalar field. The nice thing is that such a field is exactly the Higgs field in the Standard Model, so our discussion here works as an introduction to that.

Now, the self-interacting complex scalar field ϕ (essentially, the Higgs field) is describes by the complex version of the ϕ^4-Lagrangian we looked at earlier:

$$\boxed{\mathcal{L} = \partial_\mu \phi^\dagger \partial^\mu \phi + \mu^2 \phi^\dagger \phi - \lambda \left(\phi^\dagger \phi\right)^2}$$

Here, we essentially have $\phi^\dagger \phi \sim \phi^2$ and $\left(\phi^\dagger \phi\right)^2 \sim \phi^4$, so this has the same form as the ϕ^4-Lagrangian, which describes self-interactions. Therefore, this is a self-interacting theory. However, there is one crucial difference here as compared to the ordinary ϕ^4-Lagrangian - if you look closely, the μ^2-term has a different sign. Generally, the mass term in a Lagrangian should have a minus sign, but this has a plus sign. The only reason for the different sign here is that it is needed to have non-zero minima of the potential (i.e. non-zero vacuum solutions).

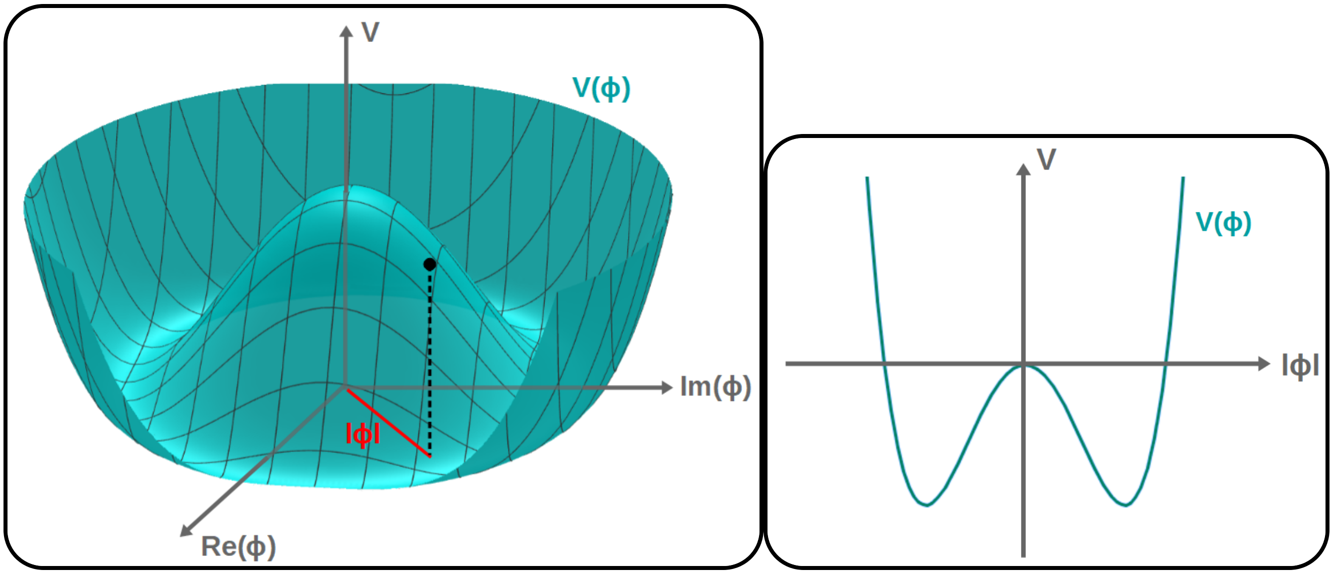
Speaking of the potential, we say that the potential of the theory of the theory is generally the part of the Lagrangian that depends only on ϕ, but not its derivatives. This is exactly analogous to potentials in classical mechanics usually only depending on position, but not velocity (derivatives of position). In this case, the potential is:

$$V(\phi) = -\mu^2 \phi^\dagger \phi + \lambda \left(\phi^\dagger \phi\right)^2$$

Another way to write it would be in terms of the "norm" of the field, $\phi^\dagger \phi = |\phi|^2$, as:

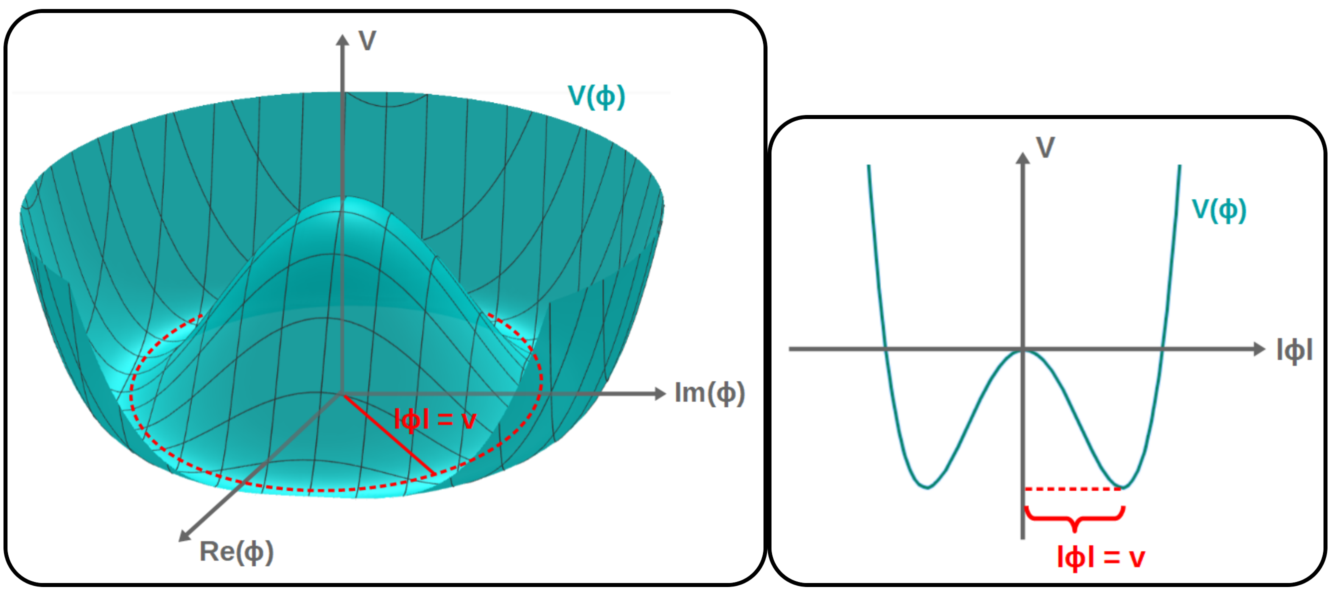
$$\boxed{V(\phi) = -\mu^2 |\phi|^2 + \lambda |\phi|^4}$$

If you look at this for a bit, you may notice that this has the same functional form as the function $z(r)$ describing the height of the ball in our earlier example - it looks like a "Mexican hat" type of hill. Hence, this kind of potential is actually often called a **Mexican hat potential**. We can plot it both as a function of only the "distance from the origin", $|\phi|$, and also in the full complex plane (the axes being the real and imaginary parts of the complex field ϕ):



Note; in the 3D picture, the potential has been shifted upwards to make the picture look nicer.

Now, the interesting thing for us here are these minimum points of the potential that lie on a circle of constant $|\phi|$. We'll call this constant v:



The reason these minima are interesting to us is that they minimize the energy of the field ϕ - at these minima, determined by $|\phi| = v$, the energy of the field is also minimized. As you may know by now, most things in the universe (which includes fields as well) like to minimize their energy. It is why objects fall to the ground under gravity, it is why atoms like to be in their ground state and so on.

Therefore, if a field *can* minimize its potential while still being in a stable field configuration, it naturally will. In this case, the minimum of the potential are stable points, so the field ϕ will indeed naturally "fall" into one of these minima and stay there is left by itself. This is exactly like the ball falling into the valley in our earlier example - only now, we have a more abstract field configuration rather than a simple ball.

So, given a self-interacting complex scalar field ϕ with a potential $V(\phi)$ of the form shown above, the field naturally likes to go into a configuration that minimizes its potential energy. We call such field configurations **vacuum solutions**. But what are these vacuum solutions of minimum energy? Well, they are minima of the potential function $V(\phi)$, which we obtain by differentiating and setting the derivative to zero:

$$\frac{dV}{d|\phi|} = \frac{d}{d|\phi|}\left(-\mu^2|\phi|^2 + \lambda|\phi|^4\right) = -2\mu^2|\phi| + 4\lambda|\phi|^3$$

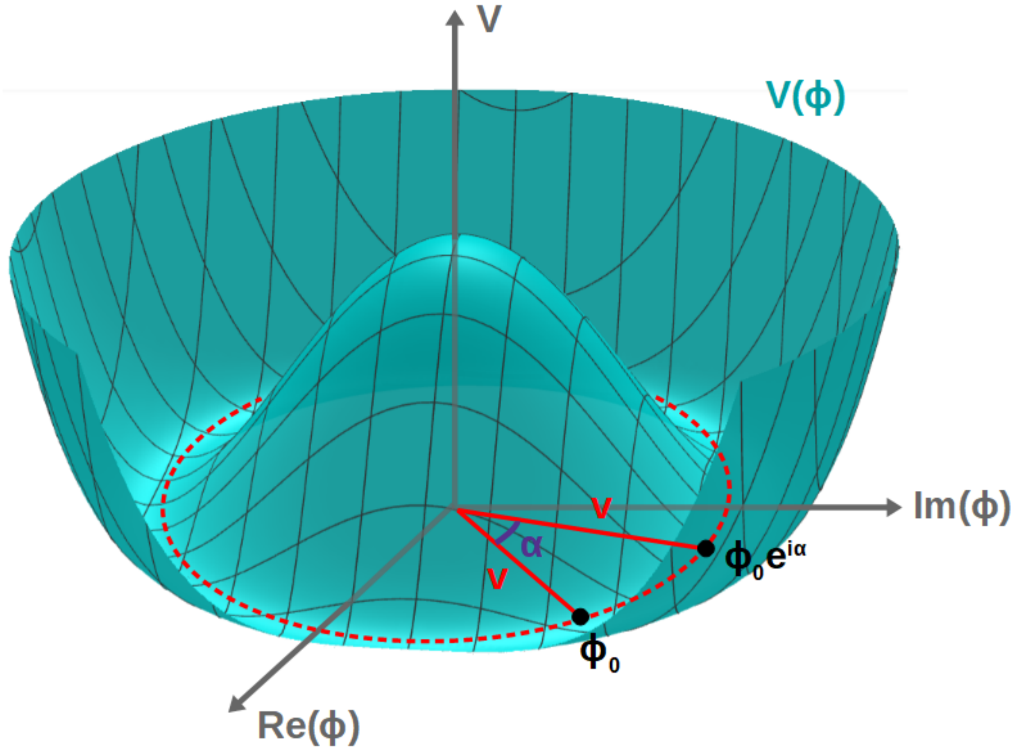
Setting this to zero then gives us the minimum of the potential, $|\phi_0| = v$:

$$-2\mu^2|\phi_0| + 4\lambda|\phi_0|^3 = 0 \quad \Rightarrow \quad \boxed{|\phi_0| = v = \frac{\mu}{\sqrt{2\lambda}}}$$

This is the *magnitude* of the vacuum field configuration ϕ_0. The vacuum field solution itself is a complex number ϕ_0, which just has a constant magnitude $|\phi_0| = v$. As seen from the graph above, the vacuum field configurations themselves lie on the circle of constant $|\phi| = v$, so there is an infinite number of different vacuum solutions.

The interesting thing for us, however, comes from analyzing the symmetries of these vacuum solutions. As we know, the full complex scalar field theory we are considering here has U(1) symmetry, which can be seen directly from the Lagrangian. But do the vacuum solutions also have the same symmetry?

The answer is no! Clearly, if we multiply the vacuum solution ϕ_0 by a phase, in other words, apply a U(1) transformation $\phi_0 \rightarrow \phi_0 e^{i\alpha}$, the solution changes. It's still a minimum of the potential, but now lying at a different point on the "minimum circle":



Since the solution is different after a U(1) transformation, the solution does not have U(1) symmetry anymore - it has been **spontaneously broken**, in exactly the same way as the rotational of the "ball system" was spontaneously broken in our previous example.

Note, however, that this spontaneous symmetry breaking only occurs when the magnitudeof the minimum of the field, $|\phi_0|$ is *non-zero*. If it were zero, the vacuum solutions would trivially be of the form $\phi_0 = 0$, which do indeed have U(1) symmetry. This is because $\phi_0 = 0$ and a U(1) transformed $\phi_0 e^{i\alpha} = 0$ are still exactly the same solution.

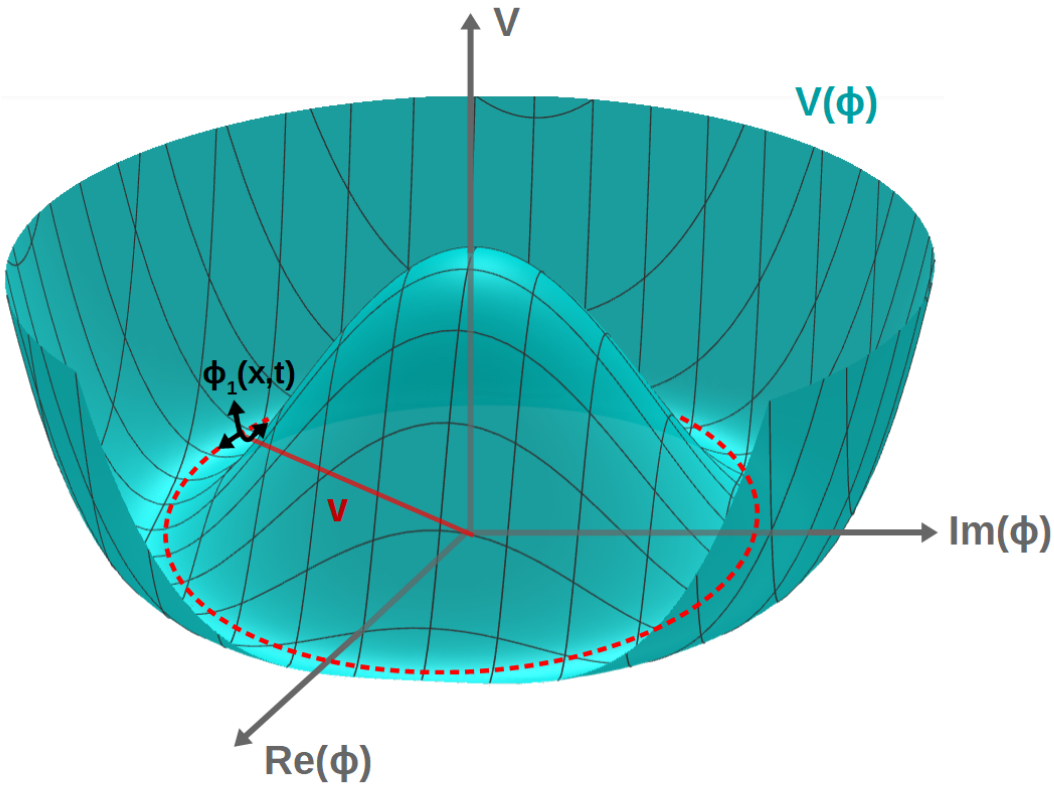
So, the crucial point to realize here is that spontaneous symmetry breaking only occurs in the case of a **non-zero vacuum minimum value**. In this case, we say that the non-zero vacuum minimum spontaneously breaks U(1) symmetry.

Okay, if the field naturally likes to "sit" at the minimum of the potential, what are its dynamics then? Well, the dynamics of the field arising from its self-interactions are then *oscillations* about these vacuum minima.

In this case, it makes sense for us to assume these oscillations are small, which corresponds to the self-interaction parameter λ being small. If this is the case, the field can be expanded around the minimum, as usual, as:

$$\phi \approx v + \lambda\phi_1$$

The field $\phi_1 = \phi_1(x, t)$ describes these small oscillations around the vacuum, which can occur in both the "radial" and "angular" directions:



In the context of particles and quantum field theory, the angular excitations turn out to correspond to massless particles called Goldstone bosons and the radial excitations to massive particles called Higgs bosons. This will be discussed more later.

Now, it turns out that we can get a much better understanding of everything moving forward if we reparameterize our field a bit first. In particular, the field ϕ_1 above is a full complex field. It is a complex-valued field, so it has two independent parameters - its real and imaginary parts.

Before, we've usually taken these two independent fields to be ϕ and $\phi^\dagger$ in most our discussions. However, here it is actually worth taking the two independent parameters to actually be the real and complex components of ϕ (or ϕ_1 in this case). This can be done by writing the field ϕ_1 as a complex number in terms of two *real* fields, its real and imaginary components, in the form of a typical complex number:

$$\boxed{\phi_1 = h + i\theta}$$

All the quantites here are fields, so $\phi_1(x, t) = h(x, t) + i\theta(x, t)$, although we will suppress the spacetime arguments. The h and θ here are our two new independent fields, which are both real. They correspond to the real and imaginary components of ϕ_1. Later, we will interpret them as describing particles called Goldstone bosons (for θ) and Higgs bosons (suggestively named h).

With these, we can now write our full field and its complex conjugate in terms of h and θ:

$$\begin{cases} \phi = v + \lambda\phi_1 = v + \lambda(h + i\theta) \\ \phi^\dagger = v + \lambda\phi_1^\dagger = v + \lambda(h - i\theta) \end{cases}$$

We've now essentially reparameterized our theory. We started with a theory of two independent scalar fields ϕ and $\phi^\dagger$, but we have reparameterized the theory to now being described by two other scalar fields h and θ. No information has been lost, so you can simply think of this as a "variable change" - we still have two independent parameters in our theory, which describe exactly the same physics.

The next step now would be to write our original Lagrangian in terms of these new fields h and θ to obtain a description of their dynamics. After a bit of work, the new Lagrangian we find turns out to be quite simple (the full calculation can found in the "Longer Calculations"):

$$\mathcal{L} = \partial_\mu\phi^\dagger\partial^\mu\phi + \mu^2\phi^\dagger\phi - \lambda\left(\phi^\dagger\phi\right)^2 \quad \Rightarrow \quad \boxed{\mathcal{L} = \partial_\mu h\partial^\mu h - 4\lambda v^2h^2 + \partial_\mu\theta\partial^\mu\theta}$$

Perhaps now you can see the point of this reparameterization - this new Lagrangian is much simpler, in particular, because it doesn't contain any interaction terms (unlike the old one, which has terms like $\phi^\dagger\phi$). You'll find some more discussion of this Lagrangian and the corresponding field equations below.

Interpretation of The Reparameterized Lagrangian & Field Equations

As mentioned earlier, the Lagrangian above has the form of describing two non-interacting fields h and θ. This is because it doesn't contain any products like $h\theta$ or $h\partial_\mu\theta$ or anything of the sort.

So, we've found that our theory - namely, the theory of our scalar field near its vacuum minima - actually describes two non-interacting independent fields. Now, because the reparameterization we did above did not change any of the physics of the theory (it was just a mathematical change of variables!), this was all hidden in our old parameterization in terms of ϕ and $\phi^\dagger$ - we just couldn't directly see it because of the form of our

Lagrangian.

Okay, let's have a look at the Lagrangian $\mathcal{L} = \partial_\mu h \partial^\mu h - 4\lambda v^2 h^2 + \partial_\mu \theta \partial^\mu \theta$ once more. There are two kinetic terms involving derivatives of the fields, one for h and one for θ as we would certainly expect in a dynamics theory - nothing too surprising with these!

We can also see that there is one term proportional to h^2 (with a minus sign). This has the form of a mass term $\sim \mu^2 h^2$, since it is quadratic in h. So, the field h here is **massive**, with mass parameter $\mu = \sqrt{4\lambda v^2}$.

What is also interesting is that there is no such mass term for θ - or equivalently, there is a mass term for θ, but with its mass parameter being exactly zero. So, we interpret the field θ as being **massless**.

Okay, we've found the description of two non-interacting fields, with one being massive and one being massless. The particles described by these fields will then also be massive (for h) and massless (for θ). We call the massive particles Higgs bosons and the massless particles Goldstone bosons. We'll discuss both of these in a bit more detail soon.

Lastly, let's look at the field equations for both of the fields:

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu h)} - \frac{\partial \mathcal{L}}{\partial h} = 0 \quad \Rightarrow \quad \partial_\mu \partial^\mu h + 4\lambda v^2 h = 0$$
$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \theta)} - \frac{\partial \mathcal{L}}{\partial \theta} = 0 \quad \Rightarrow \quad \partial_\mu \partial^\mu \theta = 0$$

As expected, of course, these have the form of a massive and massless non-interacting Klein-Gordon equation. As we know from our previous discussions, these field equations will have solutions of the form of plane waves, $h(x,t),\ \theta(x,t) \sim e^{-ik_\mu x^\mu}$. We also find the following dispersion relations for the fields:

$$\begin{cases} k_h k^\mu = 4\lambda v^2 \\ k_\mu k^\mu = 0 \end{cases}$$

Let's take a step back for a minute and summarize what we've learned here:

- We began by looking at a particular self-interacting complex scalar field theory and found that it has non-zero vacuum solutions (solutions that minimize its energy). The full theory itself has global U(1) symmetry, however, the *non-zero* vacuum solutions spontaneously break this symmetry.

- Since the field "likes" to stay near this state of minimum energy, its dynamics will typically be small excitations around this vacuum configuration. Around the vacuum, we then found out that the theory behaves exactly like a non-interacting theory of two independent real scalar fields - one being massive and the other massless.

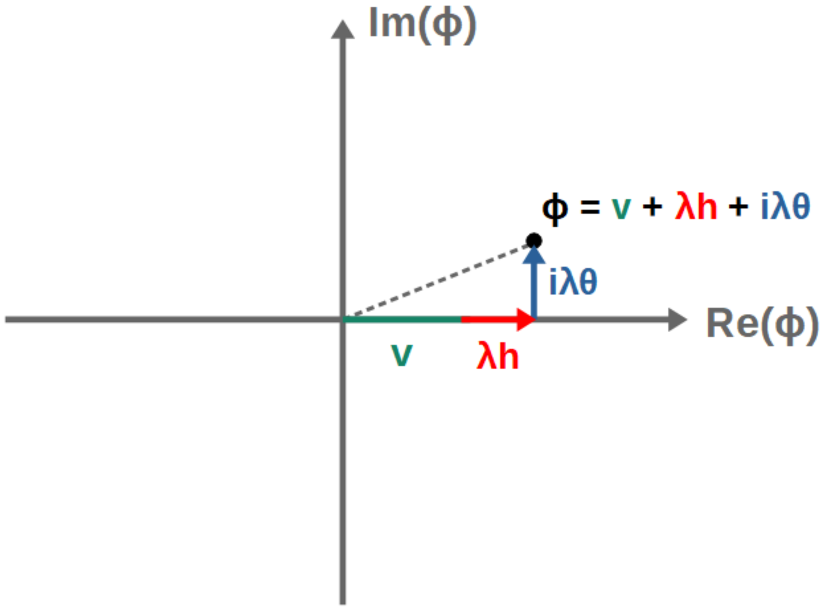
6.2.1. Higgs Bosons & Goldstone Bosons

Let's talk about the fields h and θ a bit more. These fields describe the dynamics of the full self-interacting complex scalar field around its non-zero vacuum configuration. As we discovered, the field h is massive and θ is massless.

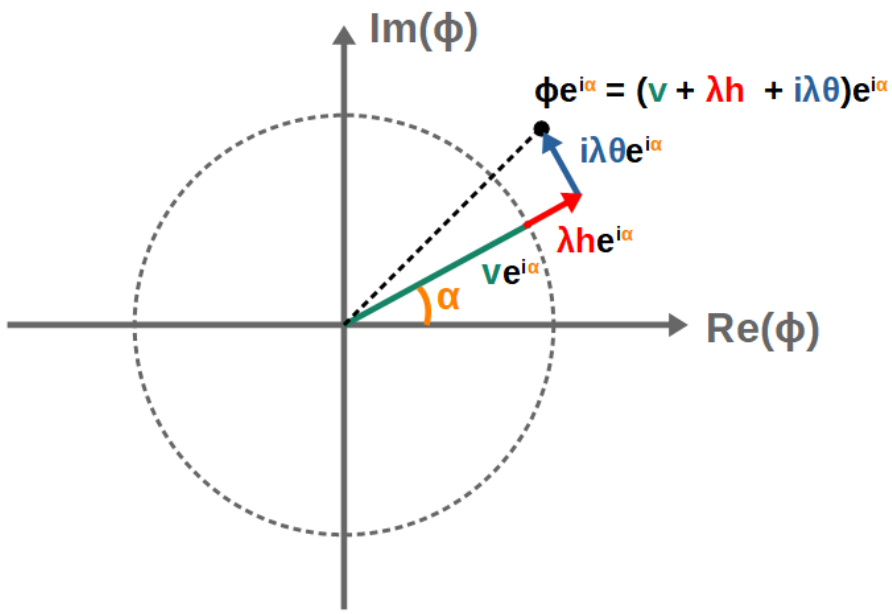
First of all, we can visualize these fields quite nicely. Remember, our full complex scalar field has the form:

$$\phi = v + \lambda h + i\lambda\theta$$

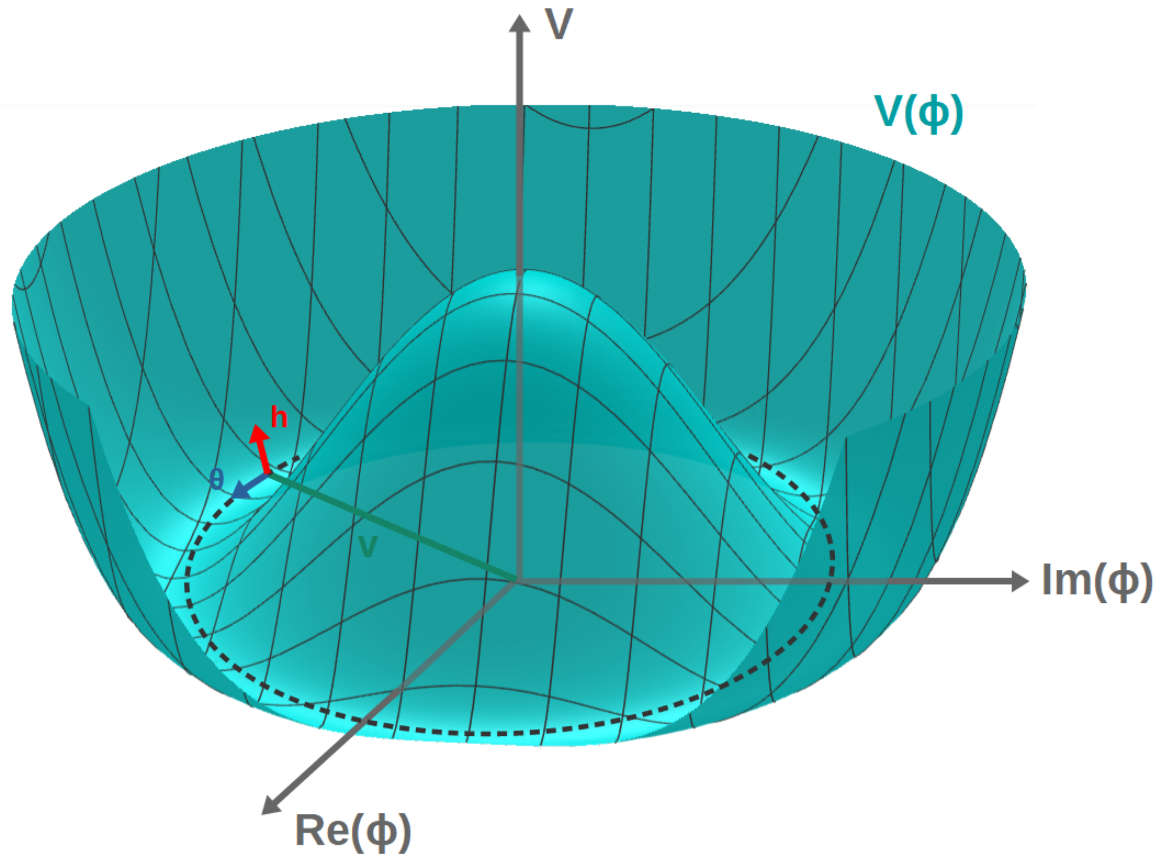
This has the form of a complex number ($z = a + ib$) with its real part being $v + \lambda h$ and its complex part being $\lambda\theta$. Now, how can we visualize complex numbers? Well, as points in the complex plane! In particular, in the two-dimensional complex plane (with the axes describing the real and imaginary parts of ϕ), this looks as follows:



Note, however, that the vacuum configuration v can be at any point on the "minimum circle" on the complex plane. To describe these excitations at any other point on this circle, we can simply rotate the above configuration by multiplying with $e^{i\alpha}$. In this case, these field excitations about the vacuum would look as follows:



What we see from this is that no matter which vacuum field configuration we are describing on this minimum circle, the field h always corresponds to a *radial* excitation in the complex plane, while the field θ corresponds to an *angular* excitation. In the 3D picture, this means that the h-excitation occurs upwards on the "potential hill" (in other words, *changing the value of the potential*), while the θ-excitation just takes the vacuum to another vacuum on the minimum circle:



What this means is that the h-excitation changes the energy of the field configuration both through its derivatives ($\partial_\mu h$), but also through changing the value of the potential $V(\phi)$. This means that even if this excitation was *stationary* (its derivatives are zero), such an excitation would still contribute to the energy of the field.

If this excitation is interpreted as a particle that is created, it would mean that creating a stationary particle (with no kinetic energy) would still add an energy contribution. But what is such a particle? Well, it is a **massive particle**, since massive particles have a non-zero rest energy given by their mass even with no kinetic energy. So, this is the intuitive reason why the field h is massive and its excitations always occur along the "radial" direction in the complex plane.

On the other hand, the θ-excitation occurs only in the angular direction, which does not change the value of the potential $V(\phi)$ at all. This would mean that the θ-excitation can only contribute to the energy of a field through its derivatives. In other words, if a *stationary* θ-particle (excitation in the θ-field) was to be created, it would have zero energy - this is exactly a **massless particle** with zero rest energy. So, intuitively, this is why the field θ is massless and its excitations are purely in the angular direction, from one vacuum configuration to another.

Now, in terms of particles, we call the massless θ-particles **Goldstone bosons**. On the other hand, we call the massive h-particles, suggestively labeled, **Higgs bosons** - yes, these are the same, famous Higgs bosons experimentally found in 2012. Although here we've considered just a general complex scalar field, the reason we call its radial excitations Higgs bosons is because the Higgs field in the Standard Model is describes exactly by such a complex field. Its massive, radial excitations are then interpreted as particles called Higgs bosons.

For the Goldstone bosons, there is an interesting and general result called **Goldstone's theorem**. It states that for every spontaneously broken symmetry, there exists a massless Goldstone boson. Here, we have exactly one symmetry (the U(1) symmetry) that is spontaneously broken and due to this, we find one massless Goldstone boson - an excitation in the θ-field.

One more important aspect to note as well - which is essentially the *Higgs mechanism* - is the fact that the mass parameter of the field h is directly proportional to the non-zero vacuum minimum value v of our scalar field, $\mu \propto v$, as we found previously. If, on the other hand, the vacuum configuration occured at $v = 0$, the field h would also be massless. So, it's as if the non-zero vacuum "gives" a mass to the h-field.

This is essentially the "mass generation" mechanism of the Higgs field - by going into a non-zero vacuum configuration that spontaneously breaks its U(1) symmetry, the Higgs field gives mass to **other fields it interacts with**. Here, we haven't yet looked at any other fields, which we'll do next in the context of various interacting theories. What we will discover doing so is exactly the famous Higgs mechanism that is responsible for the masses of many elementary particles.

6.3. The Higgs Mechanism

We are now ready to begin our discussion of the **Higgs mechanism** - the mechanism by which most elementary particles like electrons gain their masses. Fundamentally, of course, the Higgs mechanism is a part of quantum field theory, however, the nice thing is that we can still understand it perfectly well in terms of everything we've discussed so far in this course.

The Higgs mechanism is directly related to what we just discussed - the dynamics of a scalar field around its spontaneous symmetry breaking vacuum. Previously, however, we only looked at how the scalar field behaved due to s*elf-interactions*. The Higgs mechanism comes about when we ask the question of "what happens when this scalar field *interacts with other fields*?".

Specifically, the two ways the Higgs mechanism can take effect and give mass to other fields that we'll discuss are the following:

- Through **Yukawa interactions** - in this case, the non-zero vacuum value of the scalar field directly gives rise to a mass term for other fields interacting with the scalar field.
- Through **gauge interactions** - in this case, the massless Goldstone boson of the scalar field we discovered earlier becomes a simple gauge function and can be gauged away, while simultaneously, the gauge field interacting with our scalar field obtains a mass term. This is sometimes describes as the gauge field "eating" the Goldston boson.

6.3.1. Mass Generation Through Yukawa Coupling

We'll begin our discussion by considering Yukawa interactions - specifically, the Yukawa interactions between two scalar fields, ϕ and ψ.

The scalar field ϕ here is essentially the one with non-zero vacuum minima that we just considered before, only now we will take it to be **real** for the Yukawa interactions to work properly. This real scalar field still exhibits spontaneous symmetry breaking, however, its vacuum minima will now lie on the real axis and not on a circle in the complex plane.

The other field ψ is a complex scalar field, but it is a *free field* (with no quartic term like $\sim \left(\psi^\dagger \psi\right)^2$). We will also take this field to be *massless*, in order for this to resemble the mass generation of fermions as closely as possible. So, the individual Lagrangians we want to consider here are:

$$\begin{cases} \mathcal{L}_\phi = \frac{1}{2}\partial_\mu \phi \partial^\mu \phi + \frac{1}{2}\mu_\phi^2 \phi^2 - \frac{1}{4}\lambda \phi^4 \\ \mathcal{L}_\psi = \partial_\mu \psi^\dagger \partial^\mu \psi \end{cases}$$

The Yukawa interaction term here will be of the form $-g\phi\psi^\dagger\psi$, so the full Lagrangian for these two interacting fields is:

$$\boxed{\mathcal{L} = \frac{1}{2}\partial_\mu \phi \partial^\mu \phi + \frac{1}{2}\mu_\phi^2 \phi^2 - \frac{1}{4}\lambda \phi^4 + \partial_\mu \psi^\dagger \partial^\mu \psi - g\phi\psi^\dagger\psi}$$

This describes our real scalar field ϕ (which has a non-zero vacuum minimum) interacting with a massless complex scalar field ψ. Now, we can compute the vacuum minima of the field ϕ (similarly as before) to be:

$$\phi_0 = v = \frac{\mu_\phi}{\sqrt{\lambda}}$$

In this case, the spontaneously broken symmetry is the $\phi \rightarrow -\phi$ symmetry, which the Lagrangian has (but this vacuum solution does not). This is essentially just the same U(1) as before, but with only rotations of π allowed, since ϕ must be on the real axis (the phase factor would be $e^{i\pi} = -1$).

Anyway, we again have a vacuum solution that spontaneously breaks the symmetries of the original Lagrangian. We can now expand the field ϕ around this vacuum, assuming that λ is small:

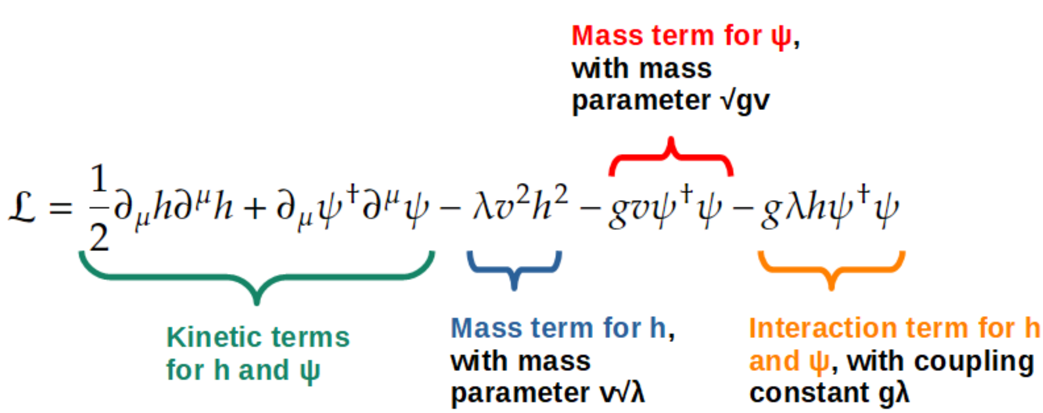
$$\phi \approx v + \lambda h$$

Here, we are calling the excitation about the vacuum $h = h(x, t)$. Note that since ϕ is real, the Goldstone boson field θ doesn't play any role here, so we do not need it. Let's now rewrite the Lagrangian in terms of this new field $\phi = v + \lambda h$. You'll again find the full calculation in the "Longer Calculations", but the result we get is:

$$\mathcal{L} = \frac{1}{2}\partial_\mu \phi \partial^\mu \phi + \frac{1}{2}\mu_\phi^2 \phi^2 - \frac{1}{4}\lambda \phi^4 + \partial_\mu \psi^\dagger \partial^\mu \psi - g\phi\psi^\dagger\psi$$

$$\Rightarrow \boxed{\mathcal{L} = \frac{1}{2}\partial_\mu h \partial^\mu h + \partial_\mu \psi^\dagger \partial^\mu \psi - \lambda v^2 h^2 - gv\psi^\dagger\psi - g\lambda h\psi^\dagger\psi}$$

Let's dissect what each of these terms represent here:



Perhaps the most interesting part of this new Lagrangian is the mass term $-gv\psi^\dagger\psi$. This is a mass term because it's quadratic in ψ (or $|\psi|$, to be exact) as all mass terms are. What makes this interesting is the fact that *the field ψ was originally massless* - in other words, it has somehow gained a non-zero mass due to interacting with the field ϕ. This is exactly the **Higgs mechanism** at play.

Its mass is now given by $\mu_\psi = \sqrt{gv}$, which depends on both the coupling strength with the ϕ-field (g) as well as the non-zero vacuum minimum value of the ϕ-field (v). So, we find that by interacting with a scalar field that has a non-zero vacuum minimum, the originally massless scalar field ψ has now gained a non-zero mass, with the value of the mass depending on both how strongly these two scalar fields interact (described by g) as well as how large the vacuum minimum v is.

Summary of the key results:

- We began by looking at a self-interacting scalar field ϕ and an originally *massless* scalar field ψ, which interact through a Yukawa coupling. The scalar field ϕ has a spontaneous symmetry breaking non-zero vacuum minimum v where it "likes to stay".
- Due to its non-zero vacuum value and its self-interactions around the vacuum, the scalar field ϕ gives rise to a non-zero mass for the other scalar field ψ.
- The result is a Yukawa interacting theory between a massive field h (corresponding to the small excitations of ϕ around the vacuum) and now a *massive* field ψ with mass parameter $\sqrt{gv}$. In other words, the scalar field ϕ has "generated" a mass for the other scalar field ψ it interacts with!

The main reason this "mass generation through Yukawa coupling" is important is that a very similar thing is what happens in the Standard Model to give fermions, such as electrons, their mass.

The actual Higgs mechanism for fermions is a bit more complicated since fermions are described by spinor fields, however, the conceptual idea is basically the same as we've discussed here.

It turns out that in the Standard Model, originally, a mass term for fermions (of the form $-\mu\psi^\dagger\psi$) would not be allowed in the Lagrangian as it would violate SU(2). However, what happens instead is that the fermion spinor field ψ couples to the Higgs field ϕ (which is really an *SU(2) doublet*) through a Yukawa coupling of the form $-g\phi\psi^\dagger\psi$ and by exactly the same mechanism as we discussed above, the Higgs field gives rise to a mass term for ψ.

6.3.2. Mass Generation of Gauge Fields

So far, we looked at how the Higgs mechanism works in the context of Yukawa couplings. Now we will look at how it works in the context of gauge couplings - in other words, how the Higgs field couples to gauge fields (i.e. vector fields) and how the non-zero vacuum affects the gauge fields it couples to.

Although slightly more complicated, in my opinion, the Higgs mechanis for gauge fields is incredibly fascinating. It is essentially a "loophole" for originally massless gauge fields to become massive, while still retaining their gauge invariance.

As a reminder, adding a quadratic mass term to the Lagrangian of a gauge field A_μ violates the gauge invariance of the theory (as we discussed in Part 8):

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}\mu^2 A_\mu A^\mu \quad \Rightarrow \quad \text{The term } \frac{1}{2}\mu^2 A_\mu A^\mu \text{ violates gauge invariance!}$$

In other words, it really appears as if there is no way to have a gauge invariant description of massive vector fields - but is there? Well, it turns out that there is and the key to this is the Higgs mechanism. It turns out that by coupling the *massless* gauge field A_μ to the Higgs field ϕ and expanding it around its non-zero vacuum (like we've done previously), the theory will generate a non-zero mass for A_μ - while *still* being gauge invariant!

Let's look at how this happens in detail. First of all, we need to construct a Lagrangian that couples the gauge field A_μ and the Higgs field ϕ (a complex scalar field). How do we do that? Well, through the **minimal coupling prescription**, as discussed previously.

First, we construct a non-interacting Lagrangian for these two fields. Importantly, note that A_μ is now taken as **massless**:

$$\mathcal{L} = \partial_\mu\phi^\dagger\partial^\mu\phi + \mu^2\phi^\dagger\phi - \lambda\left(\phi^\dagger\phi\right)^2 - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}$$

We now apply minimal coupling by promoting partial derivatives ∂_μ to gauge covariant derivatives $D_\mu = \partial_\mu - iqA_\mu$. This has the effect of changing the kinetic term for ϕ as:

$$\mathcal{L} = (D_\mu \phi)^\dagger D^\mu \phi + \mu^2 \phi^\dagger \phi - \lambda\left(\phi^\dagger \phi\right)^2 - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}$$

Here, we will expand out these gauge covariant derivatives. This gives us:

$$\mathcal{L} = (\partial_\mu \phi - iqA_\mu \phi)^\dagger \left(\partial^\mu \phi - iqA^\mu \phi\right) + \mu^2 \phi^\dagger \phi - \lambda\left(\phi^\dagger \phi\right)^2 - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}$$

$$\Rightarrow \begin{array}{l} \mathcal{L} = \partial_\mu \phi^\dagger \partial^\mu \phi + iqA_\mu \left(\phi^\dagger \partial^\mu \phi - \phi \partial^\mu \phi^\dagger\right) + \mu^2 \phi^\dagger \phi + q^2 \phi^\dagger \phi A_\mu A^\mu \\ \qquad -\lambda\left(\phi^\dagger \phi\right)^2 - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \end{array}$$

It's a bit of a lengthy Lagrangian, but let's not let that scare us. What we have here is a Lagrangian that describes the interactions between the Higgs field ϕ and the massless vector field A_μ - and most importantly, it is fully gauge invariant.

The next step is expanding the Higgs field around its non-zero vacuum again. In this case, it is actually advantageous to parameterize the field ϕ in complex *polar form* in terms of the fields h and θ:

$$\phi = (v + \lambda h) e^{i\theta}$$

This describes exactly the same physical content as the "Cartesian" complex number form we had earlier, $\phi = v + \lambda h + i\lambda\theta$, with h being the radial excitation and θ the angular excitation. The advantage of this form is that if we do a gauge transformation, the field ϕ will change as $\phi \rightarrow \phi e^{iq\varphi} = (v + \lambda h) e^{i(\theta + q\varphi)}$, where φ is the gauge parameter, $\varphi = \varphi(x)$.

In other words, a gauge transformation to ϕ can be expressed in terms of our new fields simply as $\theta \rightarrow \theta + q\varphi$, while nothing happens to h. This will turn out quite useful soon.

With this reparameterization, we can show that the Lagrangian can be written as (the full calculation is again found in the "Longer Calculations"):

$$\boxed{\mathcal{L} = \partial_\mu h \partial^\mu h - 4\lambda v^2 h^2 - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} + q^2 (v + \lambda h)^2 \left(\frac{1}{q} \partial_\mu \theta - A_\mu\right)\left(\frac{1}{q} \partial^\mu \theta - A^\mu\right)}$$

Before interpreting this any further, the most important question we need to ask is "is this gauge invariant?". Let's find out - this, as well as a particularly useful choice of gauge called the **unitary gauge**, is explored in the box below.

Gauge Invariance & The Unitary Gauge

Remember, in our particular parameterization for ϕ in terms of h and θ, gauge transformations are expressedgenerally in the form (here φ is the gauge parameter, as usual):

$$\begin{cases} h \ \to \ h \\ \theta \ \to \ \theta + q\varphi \\ A_\mu \ \to \ A_\mu + \partial_\mu \varphi \end{cases}$$

In our above Lagrangian, the entire $\partial_\mu h \partial^\mu h - 4\lambda v^2 h^2 - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}$ -part remains exactly the same under this transformation, since h doesn't change and $F_{\mu\nu}$ itself is gauge invariant. Let's concentrate on how the remaining term transforms:

$$q^2 (v + \lambda h)^2 \left(\frac{1}{q} \partial_\mu \theta - A_\mu \right) \left(\frac{1}{q} \partial^\mu \theta - A^\mu \right)$$

$$\to \ \ q^2 (v + \lambda h)^2 \left(\frac{1}{q} \partial_\mu (\theta + q\varphi) - (A_\mu + \partial_\mu \varphi) \right) \left(\frac{1}{q} \partial^\mu (\theta + q\varphi) - \left(A^\mu + \partial^\mu \varphi \right) \right)$$

$$\Rightarrow \ \ q^2 (v + \lambda h)^2 \left(\frac{1}{q} \partial_\mu \theta - A_\mu \right) \left(\frac{1}{q} \partial^\mu \theta - A^\mu \right)$$

So, nothing changes in this term either! This means that the entire Lagrangian is unchanged by the gauge transformation shown above, so **it is still gauge invariant**. This is, of course, what we would expect to have in a reasonable gauge invariant field theory.

Now, there is another particularly interesting aspect of the above Lagrangian. It turns out that the θ-field, the Goldstone boson, is not an actual physical field in this theory. This is because we can actually get rid of it completely by a gauge transformation. In other words, the Goldstone boson acts equivalently to a gauge function here, which we know should not be physically observable.

Let's see how this happens. If you stare at the above Lagrangian for a while, in particular the term with θ, you might notice that the form in which the θ-field appears looks a bit like a gauge, of the form $\partial_\mu \varphi$.

In fact, if we choose the gauge parameter as $\varphi = -\theta / q$, the θ-field would then transform as:

$$\theta \rightarrow \theta + q\left(-\frac{1}{q}\theta\right) = 0$$

In other words, by choosing a gauge of the form $\varphi = -\theta / q$, we can get rid of the Goldstone boson completely in this interacting theory! This is called the **unitary gauge** and the main advantage of this gauge is that it explicitly shows us which degrees of freedom are actually physical - in particular, the Goldstone boson degree of freedom is NOT physical.

Okay, we've managed to get rid of the Goldstone boson in this particular gauge - but the Goldstone boson is still there and has its own dynamics if we had chosen some other gauge like the Lorenz gauge, right?

Well, actually no. Remember, one of the key principles of gauge theory is that **all gauges are equivalent** in terms of their physical content. In other words, if our theory predicts some physical result in one gauge, it must be true in *all gauges*. Here, we've found that in the unitary gauge, the Goldstone boson degree of freedom θ is unphysical, which means that the same must be true in all other gauges as well. How amazing is that!

Now, it's true that in other gauges, the Goldstone boson will still appear in the Lagrangian and it *seems* as if it has its own dynamics. However, based on what we know now, these degrees of freedom of the Goldstone boson are completely unphysical. The Goldstone boson appears as a kind of "fake dynamical variable" in the Lagrangian. In the unitary gauge, however, we can completely get rid of it.

So, we've just concluded that the Goldstone boson θ is unphysical. For the rest of this section, we'll therefore work in the unitary gauge where $\theta = 0$. The Lagrangian then describes the coupled dynamics of the fields h and A_μ and is of the form shown below.

The Higgs mechanism Lagrangian in the unitary gauge:

$$\mathcal{L} = \partial_\mu h \partial^\mu h - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - 4\lambda v^2 h^2 + q^2 (v + \lambda h)^2 A_\mu A^\mu$$

To interpret this, let's expand out the term with $(v + \lambda h)^2$ and dissect the resulting terms (note that we get a term with λ^2 from this, which we can neglect):

Mass term for A_μ, with mass parameter qv – this is the crucial part regarding the Higgs mechanism!

$$\mathcal{L} = \partial_\mu h \partial^\mu h - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - 4\lambda v^2 h^2 + q^2 v^2 A_\mu A^\mu + 2q^2 \lambda v h A_\mu A^\mu$$

Kinetic terms for h and A_μ

Mass term for h, with mass parameter 2v√λ

Interaction term for h and A_μ, with coupling constant $2q^2\lambda v$

The terms highlighted with green, blue and orange here just describe the dynamics of the individual fields h and A_μ as well as their dynamics - nothing all that surprising here.

The interesting thing is the term highlighted in red, the mass term for A_μ. What we find here is that A_μ now has a non-zero mass, $\mu = qv$. The same thing has happened as saw previously - the non-zero vacuum value v of the Higgs field has produced a non-zero mass for what originally was a massless field. This is the essence in which the Higgs mechanism gives mass to vector fields - the result is an interacting theory of the Higgs boson and a massive vector field.

This is similar to what we found in the case of mass generation through Yukawa interactions as well, but the whole ordeal is just slightly more complicated because of gauge invariance. However, the key idea is the same - due to the non-zero vacuum of the Higgs field, the gauge field it interacts with gains a mass.

However, the REALLY interesting thing here is the fact that **this is a *gauge field* we're talking about**. To understand why this is so interesting, let's go back to the fully and explicitly gauge invariant version of this Lagrangian we had previously:

$$\mathcal{L} = \partial_\mu h \partial^\mu h - 4\lambda v^2 h^2 - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} + q^2 (v + \lambda h)^2 \left(\frac{1}{q} \partial_\mu \theta - A_\mu \right) \left(\frac{1}{q} \partial^\mu \theta - A^\mu \right)$$

Expanding out the $(v + \lambda h)^2$ term, we also find a mass term of the form $q^2 v^2 A_\mu A^\mu$. However, as we've seen previously, a gauge field cannot have mass due to gauge invariance (adding a mass term for it would violate gauge invariance). But what we have here is a massive gauge field - and the best part is that this Lagrangian is fully gauge invariant!

So, we've somehow managed to produce a gauge invariant theory of a massive vector field, which should not be possible based on our original thinking. However, as it turns out, the Higgs mechanism gives us a way to construct a fully gauge invariant theory of massive vector fields - a way to "go around" the fact that we cannot directly add a mass term for gauge fields.

Now, if we wanted to, we could also set $h = 0$ in the above Lagrangian, corresponding to the case where the Higgs field is sitting perfectly at its vacuum minimum, with no radial fluctuations (so, no Higgs bosons). This would then give us the **Lagrangian describing the dynamics of a massive gauge invariant vector field**:

$$\boxed{\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + q^2 v^2 \left(\frac{1}{q}\partial_\mu \theta - A_\mu\right)\left(\frac{1}{q}\partial^\mu \theta - A^\mu\right)}$$

The only role the Goldstone boson θ plays here is to "cancel out" the terms arising from gauge transformations, which makes this Lagrangian gauge invariant. However, as concluded earlier, this Goldstone mode is not physically meaningful by itself and all the physical degrees of freedom are contained in the vector field A_μ (which there are three of, corresponding to its possible polarizations).

In fact, let's count the degrees of freedom in our theory. Before, we had one complex scalar field ϕ (with two degrees of freedom) and a *massless* vector field A_μ (with two degrees of freedom), giving a total of four degrees of freedom.

After the Higgs mechanism has taken place, we now have one real scalar field h (with one degree of freedom) and a *massive* vector field A_μ (now with three degrees of freedom, as is the case with massive vector fields). So, we still have a total of four degrees of freedom, as we should - all good here!

However, the interesting thing here is **how these degrees of freedom are distributed among the fields in our theory**. Before the Higgs mechanism, the vector field had only two degrees of freedom, but afterwards, it has gained one more. However, the scalarn Higgs field instead has lost one degree of freedom, the Goldstone boson. So, in the Higgs mechanism, one degree of freedom (the Goldstone boson or "angular" degree of freedom) is essentially removed from the Higgs field and instead becomes an additional degree of freedom of the vector field that is now also massive. We say that the originally massless vector field has ***eaten* the Goldstone boson** and in doing so, it has gained a non-zero mass as well as an additional polarization degree of freedom.

With all of this, we can then state the Higgs mechanism (for gauge fields) in all of its glory:

A massless gauge field interacts with the Higgs field that has a spontaneously symmetry breaking non-zero vacuum and in doing so, eats the Goldstone boson of the Higgs field and becomes massive.

That is certainly a lot - no wonder the Higgs mechanism is so widely misunderstood. However, with everything we've covered in this part and in this course in general, it should now be possible for you to understand each piece of the above statement.

This concludes our main discussion of the Higgs mechanism. A few words about the Higgs mechanism in the Standard Model are still given below, but the key ideas have now been covered - congratulations if you made it this far!

6.3.3. Higgs Mechanism In The Standard Model

The Higgs mechanism plays an essential role in the Standard Model of particle physics, which is currently the most accurate and widely established theory of the fundamental forces of nature. The Higgs mechanism, in fact, gives masses to most of the other fields (and thus, also particles) in the theory - without it, the theory would predict these fields to be predicted. In particular:

- The W^+, W^- and Z^0 bosons are all originally massless gauge fields, as dictated by gauge invariance, but obtain their mass through the Higgs mechanism exactly like we discussed earlier.
- All elementary fermions like electrons and quarks - except neutrinos - which are described by spinor fields, obtain their mass also by the Higgs mechanism, but through a Yukawa coupling with the Higgs field (also in a similar fashion as to what we discussed earlier).

Now, when physicists speak of 'the Higgs mechanism', they are often referring only to the gauge field part. So, the mass generation of fermions through Yukawa interactions is, strictly speaking, usually not called the Higgs mechanism even though the idea is basically the same and also involves the Higgs field. But that's just arbitrary terminology anyway.

There are also some fields or particles in the Standard Model that do not participate in the Higgs mechanism. In particular, the Higgs mechanism doesn't give mass to photons or gluons - roughly speaking, the reason for this is that the component of the Higgs field that takes on a non-zero vacuum and thus gives masses to particles does NOT have **electric charge** or **color charge** (which are the charges needed in order to interact with the electromagnetic field and gluon fields). This means that the mass generating component of the Higgs field does not interact with the electromagnetic field or gluon field and therefore cannot give masses to their respective particles either. So, this is why the **photon and the eight different gluons remain massless** even after the Higgs mechanism.

Also, according to current research, neutrinos have a small mass but they don't seem to obtain it from the Higgs mechanism - in fact, it is not currently known for sure where neutrinos get their mass from. It is an extremely active area of research and may very well open the door for some new physics to be discovered in the future.

Sidenote: It's worth mentioning that for the mass generation of particles through the Higgs mechanism, the important part is specifically the non-zero vacuum value of the Higgs field - not the Higgs boson. I've often heard of misconceptions like "the Higgs boson gives masses to other particles", which just isn't true - the Higgs boson is described by the field h from earlier, which itself did not play any role in producing mass terms for other fields, the vacuum minimum v did. The Higgs boson itself doesn't do anything, it's just a particle - excitation of the Higgs field - like any other and it certainly doesn't "give mass" to anything.