ball, of the order $e^2/a$, was identified as the electron mass. Interpreting $1/a$ as $\Lambda$, we could say that in classical physics the electron mass is proportional to $\Lambda$ and diverges linearly. Thus, one way of stating the Weisskopf phenomenon is that "bosons behave worse than a classical charge, but fermions behave better."

As Weisskopf explained in 1939, the difference in the degree of the divergence can be understood heuristically in terms of quantum statistics. The "bad" behavior of bosons has to do with their gregariousness. A fermion would push away the virtual fermions fluctuating in the vacuum, thus creating a cavity in the vacuum charge distribution surrounding it. Hence its self-energy is less singular than would be the case were quantum statistics not taken into account. A boson does the opposite.

The "bad" behavior of bosons will come back to haunt us later.

## Exercises

III.3.1. Show that in $(1+1)$-dimensional spacetime the Dirac field $\psi$ has mass dimension $\frac{1}{2}$, and hence the Fermi coupling is dimensionless.

III.3.2. Derive (11) and (13).

III.3.3. Show that $B(p^2)$ in (14) vanishes when we set $m = 0$. Show that the same behavior holds in quantum electrodynamics.

III.3.4. We showed that the specific contribution (14) to $\delta m$ is logarithmically divergent. Convince yourself that this is actually true to any finite order in perturbation theory.

# Chapter III.4

# Gauge Invariance: A Photon Can Find No Rest

## When the central identity blows up

I explained in Chapter I.7 that the path integral for a generic field theory can be formally evaluated in what deserves to be called the Central Identity of Quantum Field Theory:

$$\int D\varphi e^{-\frac{1}{2}\varphi\cdot K\cdot\varphi - V(\varphi) + J\cdot\varphi} = e^{-V(\delta/\delta J)} e^{\frac{1}{2}J\cdot K^{-1}\cdot J} \qquad (1)$$

For any field theory we can always gather up all the fields, put them into one giant column vector, and call the vector $\varphi$. We then single out the term quadratic in $\varphi$ write it as $\frac{1}{2}\varphi\cdot K\cdot\varphi$, and call the rest $V(\varphi)$. I am using a compact notation in which spacetime coordinates and any indices on the field, including Lorentz indices, are included in the indices of the formal matrix $K$. We will often use (1) with $V = 0$:

$$\int D\varphi e^{-\frac{1}{2}\varphi\cdot K\cdot\varphi + J\cdot\varphi} = e^{\frac{1}{2}J\cdot K^{-1}\cdot J} \qquad (2)$$

But what if $K$ does not have an inverse?

This is not an esoteric phenomenon that occurs in some pathological field theory, but in one of the most basic actions of physics, the Maxwell action

$$S(A) = \int d^4x\mathcal{L} = \int d^4x\left[\tfrac{1}{2}A_\mu(\partial^2 g^{\mu\nu} - \partial^\mu\partial^\nu)A_\nu + A_\mu J^\mu\right]. \qquad (3)$$

The formal matrix $K$ in (2) is proportional to the differential operator $(\partial^2 g^{\mu\nu} - \partial^\mu\partial^\nu) \equiv Q^{\mu\nu}$. A matrix does not have an inverse if some of its eigenvalues are zero, that is, if when acting on some vector, the matrix annihilates that vector. Well, observe that $Q^{\mu\nu}$ annihilates vectors of the form $\partial_\nu\Lambda(x)$ : $Q^{\mu\nu}\partial_\nu\Lambda(x) = 0$. Thus $Q^{\mu\nu}$ has no inverse.

There is absolutely nothing mysterious about this phenomenon; we have already encountered it in classical physics. Indeed, when we first learned the concepts of electricity, we were told that only the "voltage drop" between two points has physical meaning. At a more sophisticated level, we learned that we can always add any constant (or indeed any function of time) to the electrostatic potential (which is of course just "voltage") since by definition its gradient is the electric field. At an even more sophisticated level, we see that solving Maxwell's equation (which of course comes from just extremizing the action) amounts to finding the inverse $Q^{-1}$. [In the notation I am using here Maxwell's equation $\partial_\mu F^{\mu\nu} = J^\nu$ is written as $Q_{\mu\nu}A^\nu = J^\nu$, and the solution is $A^\nu = (Q^{-1})^{\nu\mu} J_\mu$.]

Well, $Q^{-1}$ does not exist! What do we do? We learned that we must impose an additional constraint on the gauge potential $A^\mu$, known as "fixing a gauge."

## A mundane nonmystery

To emphasize the rather mundane nature of this gauge fixing problem (which some older texts tend to make into something rather mysterious and almost hopelessly difficult to understand), consider just an ordinary integral $\int_{-\infty}^{+\infty} dA e^{-A\cdot K\cdot A}$, with $A = (a, b)$ a 2-component vector and $K = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, a matrix without an inverse. Of course you realize what the problem is: We have $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} da\, db\, e^{-a^2}$ and the integral over $b$ does not exist. To define the integral we insert into it a delta function $\delta(b - \xi)$. The integral becomes defined and actually does not depend on the arbitrary number $\xi$. More generally, we can insert $\delta[f(b)]$ with $f$ some function of our choice. In the context of an ordinary integral, this procedure is of course ludicrous overkill, but we will use the analog of this procedure in what follows.

The necessity for imposing by hand a (gauge fixing) constraint in gauge theories can be seen from physical reasoning as well. In Chapter I.5 we sidestep this whole issue of fixing the gauge by treating the massive vector meson instead of the photon. In effect, we change $Q^{\mu\nu}$ to $(\partial^2 + m^2)g^{\mu\nu} - \partial^\mu\partial^\nu$, which does have an inverse (in fact we even found the inverse explicitly).

Consider a massive vector meson moving along. By a Lorentz boost, we can always bring it to its rest frame, where we can apply what we learned about the rotation group, namely that a spin 1 meson has three spin states or, classically, polarization states. But if the vector meson is massless, we can no longer find a Lorentz boost that would bring us to its rest frame. A photon can find no rest!

A massless spin 1 field is intrinsically different from a massive spin 1 field—that's the crux of the problem. Now we have only rotations around the direction of motion of the photon, that is, $O(2)$. The photon has only two polarization degrees of freedom. (You already learned in classical electrodynamics that an electromagnetic wave has two transverse degrees of freedom.) This is the true physical origin of gauge invariance.

In this sense, gauge invariance is, strictly speaking, not a "real" symmetry but merely a reflection of the fact that we used a redundant description: a Lorentz vector field to describe two physical degrees of freedom.

## Restricting the functional integral

I will now discuss the method for dealing with this redundancy invented by Faddeev and Popov. As you will see presently, it is the analog of the method we used in our baby problem above. Even in the context of electromagnetism this method is a bit of overkill, but it will prove to be essential for nonabelian gauge theories (as we will see in Chapter VII.1) and for gravity. I will describe the method using a completely general and somewhat abstract language. In the next section, I will then apply the discussion here to a specific example. If you have some trouble with this section, you might find it helpful to go back and forth between the two sections.

Suppose we have to do the integral $I \equiv \int DAe^{iS(A)}$; this can be an ordinary integral or a path integral. Suppose that under the transformation $A \to A_g$ the integrand and the measure do not change, that is, $S(A) = S(A_g)$ and $DA = DA_g$. The transformations obviously form a group, since if we transform again with $g'$, the integrand and the measure do not change under the combined effect of $g$ and $g'$ and $A_g \to (A_g)_{g'} = A_{gg'}$. We would like to write the integral $I$ in the form $I = (\int Dg)J$, with $J$ independent of $g$. In other words, we want to factor out the redundant integration over $g$. Note that $Dg$ is the invariant measure over the group of transformations and $\int Dg$ is the volume of the group. Be aware of the compactness of the notation in the case of a path integral: $A$ and $g$ are both functions of the spacetime coordinates $x$.

I want to emphasize that this hardly represents anything profound or mysterious. If you have to do the integral $I = \int dx\, dy\, e^{iS(x,y)}$ with $S(x, y)$ some function of $x^2 + y^2$, you know perfectly well to go to polar coordinates $I = (\int d\theta)J = (2\pi)J$, where $J = \int dr\, re^{iS(r)}$ is an integral over the radial coordinate $r$ only. The factor $2\pi$ is precisely the volume of the group of rotations in 2 dimensions.

Faddeev and Popov showed how to do this "going over to polar coordinates" in a unified and elegant way. Following them, we first write the numeral "one" as $1 = \Delta(A) \int Dg\delta[f(A_g)]$, an equality that merely defines $\Delta(A)$. Here $f$ is some function of our choice and $\Delta(A)$, known as the Faddeev-Popov determinant, of course depends on $f$. Next, note that $[\Delta(A_{g'})]^{-1} = \int Dg\delta[f(A_{g'g})] = \int Dg''\delta[f(A_{g''})] = [\Delta(A)]^{-1}$, where the second equality follows upon defining $g'' = g'g$ and noting that $Dg'' = Dg$. In other words, we showed that $\Delta(A) = \Delta(A_g)$ : the Faddeev-Popov determinant is gauge invariant. We now insert 1 into the integral $I$ we have to do:

$$\begin{aligned} I &= \int DAe^{iS(A)} \\ &= \int DAe^{iS(A)}\Delta(A)\int Dg\delta[f(A_g)] \\ &= \int Dg \int DAe^{iS(A)}\Delta(A)\delta[f(A_g)] \end{aligned} \tag{4}$$

integration.

At the physicist's level of rigor, we are always allowed to change integration variables until proven guilty. So let us change $A$ to $A_{g^{-1}}$; then

$$I = \left(\int Dg\right) \int DA e^{iS(A)} \Delta(A) \delta[f(A)] \tag{5}$$

where we have used the fact that $DA$, $S(A)$, and $\Delta(A)$ are all invariant under $A \to A_{g^{-1}}$.

That's it. We've done it. The group integration $(\int Dg)$ has been factored out.

The volume of a compact group is finite, but in gauge theories there is a separate group at every point in spacetime, and hence $(\int Dg)$ is an infinite factor. (This also explains why there is no gauge fixing problem in theories with global symmetries introduced in Chapter I.9.) Fortunately, in the path integral $Z$ for field theory we do not care about overall factors in $Z$, as was explained in Chapter I.3, and thus the factor $(\int Dg)$ can simply be thrown away.

## Fixing the electromagnetic gauge

Let us now apply the Faddeev-Popov method to electromagnetism. The transformation leaving the action invariant is of course $A_\mu \to A_\mu - \partial_\mu \Lambda$, so $g$ in the present context is denoted by $\Lambda$ and $A_g \equiv A_\mu - \partial_\mu \Lambda$. Note also that since the integral $I$ we started with is independent of $f$ it is still independent of $f$ in spite of its appearance in (5). Choose $f(A) = \partial A - \sigma$, where $\sigma$ is a function of $x$. In particular, $I$ is independent of $\sigma$ and so we can integrate $I$ with an arbitrary functional of $\sigma$, in particular, the functional $e^{-(i/2\xi)\int d^4x \sigma(x)^2}$.

We now turn the crank. First, we calculate

$$[\Delta(A)]^{-1} \equiv \int Dg \delta[f(A_g)] = \int D\Lambda \delta(\partial A - \partial^2 \Lambda - \sigma) \tag{6}$$

Next we note that in (5) $\Delta(A)$ appears multiplied by $\delta[f(A)]$ and so in evaluating $[\Delta(A)]^{-1}$ in (6) we can effectively set $f(A) = \partial A - \sigma$ to zero. Thus from (6) we have $\Delta(A)$ "=" $[\int D\Lambda \delta(\partial^2 \Lambda)]^{-1}$. But this object does not even depend on $A$, so we can throw it away. Thus, up to irrelevant overall factors that could be thrown away $I$ is just $\int DA e^{iS(A)} \delta(\partial A - \sigma)$.

Integrating over $\sigma(x)$ as we said we were going to do, we finally obtain

$$Z = \int D\sigma e^{-(i/2\xi)\int d^4x \sigma(x)^2} \int DA e^{iS(A)} \delta(\partial A - \sigma)$$
$$= \int DA e^{iS(A) - (i/2\xi)\int d^4x (\partial A)^2} \tag{7}$$

Nifty trick by Faddeev and Popov, eh?

$$S_{eff}(A) = S(A) - \frac{1}{2\xi} \int d^4x (\partial A)^2$$
$$= \int d^4x \left\{ \frac{1}{2} A_\mu \left[ \partial^2 g^{\mu\nu} - \left(1 - \frac{1}{\xi}\right) \partial^\mu \partial^\nu \right] A_\nu + A_\mu J^\mu \right\} \tag{8}$$

and $Q^{\mu\nu}$ by $Q^{\mu\nu}_{\text{eff}} = \partial^2 g^{\mu\nu} - (1 - 1/\xi)\partial^\mu \partial^\nu$ or in momentum space $Q^{\mu\nu}_{\text{eff}} = -k^2 g^{\mu\nu} + (1 - 1/\xi) k^\mu k^\nu$, which does have an inverse. Indeed, you can check that

$$Q^{\mu\nu}_{\text{eff}} \left[ -g_{\nu\lambda} + (1 - \xi) \frac{k_\nu k_\lambda}{k^2} \right] \frac{1}{k^2} = \delta^\mu_\lambda$$

Thus, the photon propagator can be chosen to be

$$\frac{(-i)}{k^2} \left[ g_{\nu\lambda} - (1 - \xi) \frac{k_\nu k_\lambda}{k^2} \right] \tag{9}$$

in agreement with the conclusion in Chapter II.7.

While the Faddeev-Popov argument is a lot slicker, many physicists still prefer the explicit Feynman argument given in Chapter II.7. I do. When we deal with the Yang-Mills theory and the Einstein theory, however, the Faddeev-Popov method is indispensable, as I have already noted.

## A reflection on gauge symmetry

As we will see later and as you might have heard, much of the world beyond electromagnetism is also described by gauge theories. But as we saw here, gauge theories are also deeply disturbing and unsatisfying in some sense: They are built on a redundancy of description. The electromagnetic gauge transformation $A_\mu \to A_\mu - \partial_\mu \Lambda$ is not truly a symmetry stating that two physical states have the same properties. Rather, it tells us that the two gauge potentials $A_\mu$ and $A_\mu - \partial_\mu \Lambda$ describe the same physical state. In your orderly study of physics, the first place where $A_\mu$ becomes indispensable is the Schrödinger equation, as I will explain in Chapter IV.4. Within classical physics, you got along perfectly well with just $\vec{E}$ and $\vec{B}$. Some physicists are looking for a formulation of quantum electrodynamics without using $A_\mu$, but so far have failed to turn up an attractive alternative to what we have. It is conceivable that a truly deep advance in theoretical physics would involve writing down quantum electrodynamics without writing $A_\mu$.

# Chapter III.5

# Field Theory without Relativity

## Slower in its maturity

Quantum field theory at its birth was relativistic. Later in its maturity, it found applications in condensed matter physics. We will have a lot more to say about the role of quantum field theory in condensed matter, but for now, we have the more modest goal of learning how to take the nonrelativistic limit of a quantum field theory.

The Lorentz invariant scalar field theory

$$\mathcal{L} = (\partial\Phi^\dagger)(\partial\Phi) - m^2\Phi^\dagger\Phi - \lambda(\Phi^\dagger\Phi)^2 \tag{1}$$

(with $\lambda > 0$ as always) describes a bunch of interacting bosons. It should certainly contain the physics of slowly moving bosons. For clarity consider first the relativistic Klein-Gordon equation

$$(\partial^2 + m^2)\Phi = 0 \tag{2}$$

for a free scalar field. A mode with energy $E = m + \varepsilon$ would oscillate in time as $\Phi \propto e^{-iEt}$. In the nonrelativistic limit, the kinetic energy $\varepsilon$ is much smaller than the rest mass $m$. It makes sense to write $\Phi(\vec{x}, t) = e^{-imt}\varphi(\vec{x}, t)$, with the field $\varphi$ oscillating much more slowly than $e^{-imt}$ in time. Plugging into (2) and using the identity $(\partial/\partial t)e^{-imt}(\cdots) = e^{-imt}(-im + \partial/\partial t)(\cdots)$ twice, we obtain $(-im + \partial/\partial t)^2\varphi - \vec{\nabla}^2\varphi + m^2\varphi = 0$. Dropping the term $(\partial^2/\partial t^2)\varphi$ as small compared to $-2im(\partial/\partial t)\varphi$, we find Schrödinger's equation, as we had better:

$$i\frac{\partial}{\partial t}\varphi = -\frac{\vec{\nabla}^2}{2m}\varphi \tag{3}$$

By the way, the Klein-Gordon equation was actually discovered before Schrödinger's equation.

Having absorbed this, you can now easily take the nonrelativistic limit of a quantum field theory. Simply plug

$$\Phi(\vec{x}, t) = \frac{1}{\sqrt{2m}}e^{-imt}\varphi(\vec{x}, t) \tag{4}$$



into (1). (The factor $1/\sqrt{2m}$ is for later convenience.) For example,

$$\frac{\partial\Phi^\dagger}{\partial t}\frac{\partial\Phi}{\partial t} - m^2\Phi^\dagger\Phi \to \frac{1}{2m}\left\{\left[\left(im + \frac{\partial}{\partial t}\right)\varphi^\dagger\right]\left[\left(-im + \frac{\partial}{\partial t}\right)\varphi\right] - m^2\varphi^\dagger\varphi\right\}$$

$$\simeq \frac{1}{2}i\left(\varphi^\dagger\frac{\partial\varphi}{\partial t} - \frac{\partial\varphi^\dagger}{\partial t}\varphi\right) \tag{5}$$

After an integration by parts we arrive at

$$\mathcal{L} = i\varphi^\dagger\partial_0\varphi - \frac{1}{2m}\partial_i\varphi^\dagger\partial_i\varphi - g^2(\varphi^\dagger\varphi)^2 \tag{6}$$

where $g^2 = \lambda/4m^2$.

As we saw in Chapter I.9 the theory (1) enjoys a conserved Noether current $J_\mu = i(\Phi^\dagger\partial_\mu\Phi - \partial_\mu\Phi^\dagger\Phi)$. The density $J_0$ reduces to $\varphi^\dagger\varphi$, precisely as you would expect, while $J_i$ reduces to $(i/2m)(\varphi^\dagger\partial_i\varphi - \partial_i\varphi^\dagger\varphi)$. When you first took a course in quantum mechanics, didn't you wonder why the density $\rho \equiv \varphi^\dagger\varphi$ and the current $J_i = (i/2m)(\varphi^\dagger\partial_i\varphi - \partial_i\varphi^\dagger\varphi)$ look so different? As to be expected, various expressions inevitably become uglier when reduced from a more symmetric to a less symmetric theory.

## Number is conjugate to phase angle

Let me point out some differences between the relativistic and nonrelativistic case.

The most striking is that the relativistic theory is quadratic in time derivative, while the nonrelativistic theory is linear in time derivative. Thus, in the nonrelativistic theory the momentum density conjugate to the field $\varphi$, namely $\delta\mathcal{L}/\delta\partial_0\varphi$, is just $i\varphi^\dagger$, so that $[\varphi^\dagger(\vec{x}, t), \varphi(\vec{x}', t)] = -\delta^{(D)}(\vec{x} - \vec{x}')$. In condensed matter physics it is often illuminating to write $\varphi = \sqrt{\rho}e^{i\theta}$ so that

$$\mathcal{L} = \frac{i}{2}\partial_0\rho - \rho\partial_0\theta - \frac{1}{2m}\left[\rho(\partial_i\theta)^2 + \frac{1}{4\rho}(\partial_i\rho)^2\right] - g^2\rho^2 \tag{7}$$

The first term is a total divergence. The second term tells us something of great importance[1] in condensed matter physics: in the canonical formalism (Chapter I.8), the momentum density conjugate to the phase field $\theta(x)$ is $\delta\mathcal{L}/\delta\partial_0\theta = -\rho$ and thus Heisenberg tells us that

$$[\rho(\vec{x}, t), \theta(\vec{x}', t)] = i\delta^{(D)}(\vec{x} - \vec{x}') \tag{8}$$

[1] See P. Anderson, *Basic Notions of Condensed Matter Physics*, p. 235.

Integrating and defining $N \equiv \int d^D x \rho(\vec{x}, t)$ = the total number of bosons, we find one of the most important relations in condensed matter physics

$$[N, \theta] = i \tag{9}$$

Number is conjugate to phase angle, just as momentum is conjugate to position. Marvel at the elegance of this! You would learn in a condensed matter course that this fundamental relation underlies the physics of the Josephson junction.

You may know that a system of bosons with a "hard core" repulsion between them is a superfluid at zero temperature. In particular, Bogoliubov showed that the system contains an elementary excitation obeying a linear dispersion relation.[2] I will discuss superfluidity in Chapter V.1.

## The sign of repulsion

In the nonrelativistic theory (7) it is clear that the bosons repel each other: Piling particles into a high density region would require an energy $g^2\rho^2$. But it is less clear in the relativistic theory that $\lambda(\Phi^\dagger\Phi)^2$ with $\lambda$ positive corresponds to repulsion. I outline one method in Exercise III.5.3, but here let's just take a flying heuristic guess. The Hamiltonian (density) involves the negative of the Lagrangian and hence goes as $\lambda(\Phi^\dagger\Phi)^2$ for large $\Phi$ and would thus be unbounded below for $\lambda < 0$. We know physically that a free Bose gas tend to condense and clump, and with an attractive interaction it surely might want to collapse. We naturally guess that $\lambda > 0$ corresponds to repulsion.

I next give you a more foolproof method. Using the central identity of quantum field theory we can rewrite the path integral for the theory in (1) as

$$Z = \int D\Phi D\sigma e^{i\int d^4x[(\partial\Phi^\dagger)(\partial\Phi) - m^2\Phi^\dagger\Phi + 2\sigma\Phi^\dagger\Phi + (1/\lambda)\sigma^2]} \tag{10}$$

Condensed matter physicists call the transformation from (1) to the Lagrangian $\mathcal{L} = (\partial\Phi^\dagger)(\partial\Phi) - m^2\Phi^\dagger\Phi + 2\sigma\Phi^\dagger\Phi + (1/\lambda)\sigma^2$ the Hubbard-Stratonovich transformation. In field theory, a field that does not have kinetic energy, such as $\sigma$, is known as an auxiliary field and can be integrated out in the path integral. When we come to the superfield formalism in Chapter VIII.4, auxiliary fields will play an important role.

Indeed, you might recall from Chapter III.2 how a theory with an intermediate vector boson could generate Fermi's theory of the weak interaction. The same physics is involved here: The theory (10) in which the $\Phi$ field is coupled to an "intermediate $\sigma$ boson" can generate the theory (1).

If $\sigma$ were a "normal scalar field" of the type we have studied, that is, if the terms quadratic in $\sigma$ in the Lagrangian had the form $\frac{1}{2}(\partial\sigma)^2 - \frac{1}{2}M^2\sigma^2$, then its propagator would be $i/(k^2 - M^2 + i\varepsilon)$. The scattering amplitude between two $\Phi$ bosons would be proportional to this propagator. We learned in Chapter I.4 that the exchange of a scalar field leads to an attractive force.

But $\sigma$ is not a normal field as evidenced by the fact that the Lagrangian contains only the quadratic term $+(1/\lambda)\sigma^2$. Thus its propagator is simply $i/(1/\lambda) = i\lambda$, which (for $\lambda > 0$) has a sign opposite to the normal propagator evaluated at low-momentum transfer $i/(k^2 - M^2 + i\varepsilon) \simeq -i/M^2$. We conclude that $\sigma$ exchange leads to a repulsive force.

Incidentally, this argument also shows that the repulsion is infinitely short ranged, like a delta function interaction. Normally, as we learned in Chapter I.4 the range is determined by the interplay between the $k^2$ and the $M^2$ terms. Here the situation is as if the $M^2$ term is infinitely large. We can also argue that the interaction $\lambda(\Phi^\dagger\Phi)^2$ involves creating two bosons and then annihilating them all at the same spacetime point.

## Finite density

One final point of physics that people trained as particle physicists do not always remember: Condensed matter physicists are not interested in empty space, but want to have a finite density $\bar{\rho}$ of bosons around. We learned in statistical mechanics to add a chemical potential term $\mu\varphi^\dagger\varphi$ to the Lagrangian (6). Up to an irrelevant (in this context!) additive constant, we can rewrite the resulting Lagrangian as

$$\mathcal{L} = i\varphi^\dagger\partial_0\varphi - \frac{1}{2m}\partial_i\varphi^\dagger\partial_i\varphi - g^2(\varphi^\dagger\varphi - \bar{\rho})^2 \tag{11}$$

Amusingly, mass appears in different places in relativistic and nonrelativistic field theories. To proceed further, I have to develop the concept of spontaneous symmetry breaking. Thus, adios for now. We will come back to superfluidity in due time.

## Exercises

III.5.1. Obtain the Klein-Gordon equation for a particle in an electrostatic potential (such as that of the nucleus) by the gauge principle of replacing $(\partial/\partial t)$ in (2) by $\partial/\partial t - ieA_0$. Show that in the nonrelativistic limit this reduces to the Schrödinger's equation for a particle in an external potential.

III.5.2. Take the nonrelativistic limit of the Dirac Lagrangian.

III.5.3. Given a field theory we can compute the scattering amplitude of two particles in the nonrelativistic limit. We then postulate an interaction potential $U(\vec{x})$ between the two particles and use nonrelativistic quantum mechanics to calculate the scattering amplitude, for example in Born approximation. Comparing the

---

[2] For example, L.D. Landau and E. M. Lifschitz, *Statisical Physics*, p. 238.

two scattering amplitudes we can determine $U(\vec{x})$. Derive the Yukawa and the Coulomb potentials this way. The application of this method to the $\lambda(\Phi^\dagger\Phi)^2$ interaction is slightly problematic since the delta function interaction is a bit singular, but it should be all right for determining whether the force is repulsive or attractive.

Chapter III.6

# The Magnetic Moment of the Electron

## Dirac's triumph

I said in the preface that the emphasis in this book is not on computation, but how can I not tell you about the greatest triumph of quantum field theory?

After Dirac wrote down his equation, the next step was to study how the electron interacts with the electromagnetic field. According to the gauge principle already used to write the Schrödinger's equation in an electromagnetic field, to obtain the Dirac equation for an electron in an external electromagnetic field we merely have to replace the ordinary derivative $\partial_\mu$ by the covariant derivative $D_\mu = \partial_\mu - ieA_\mu$:

$$(i\gamma^\mu D_\mu - m)\psi = 0 \qquad (1)$$

Recall (II.1.27).

Acting on this equation with $(i\gamma^\mu D_\mu + m)$, we obtain $-(\gamma^\mu\gamma^\nu D_\mu D_\nu + m^2)\psi = 0$. We have $\gamma^\mu\gamma^\nu D_\mu D_\nu = \frac{1}{2}(\{\gamma^\mu, \gamma^\nu\} + [\gamma^\mu, \gamma^\nu])D_\mu D_\nu = D_\mu D^\mu - i\sigma^{\mu\nu}D_\mu D_\nu$ and $i\sigma^{\mu\nu}D_\mu D_\nu = (i/2)\sigma^{\mu\nu}[D_\mu, D_\nu] = (e/2)\sigma^{\mu\nu}F_{\mu\nu}$. Thus,

$$\left(D_\mu D^\mu - \frac{e}{2}\sigma^{\mu\nu}F_{\mu\nu} + m^2\right)\psi = 0 \qquad (2)$$

Now consider a weak constant magnetic field pointing in the 3rd direction for definiteness, weak so that we can ignore the $(A_i)^2$ term in $(D_i)^2$. By gauge invariance, we can choose $A_0 = 0$, $A_1 = -\frac{1}{2}Bx^2$, and $A_2 = \frac{1}{2}Bx^1$ (so that $F_{12} = \partial_1 A_2 - \partial_2 A_1 = B$). As we will see, this is one calculation in which we really have to keep track of factors of 2. Then

$$\begin{aligned}(D_i)^2 &= (\partial_i)^2 - ie(\partial_i A_i + A_i\partial_i) + O(A_i^2)\\ &= (\partial_i)^2 - 2\frac{ie}{2}B(x^1\partial_2 - x^2\partial_1) + O(A_i^2)\\ &= \vec{\nabla}^2 - e\vec{B}\cdot\vec{x}\times\vec{p} + O(A_i^2)\end{aligned} \qquad (3)$$

Note that we used $\partial_i A_i + A_i\partial_i = (\partial_i A_i) + 2A_i\partial_i = 2A_i\partial_i$, where in $(\partial_i A_i)$ the partial derivative acts only on $A_i$. You may have recognized $\vec{L} \equiv \vec{x}\times\vec{p}$ as the orbital angular momentum operator. Thus, the orbital angular momentum generates an orbital magnetic moment that interacts with the magnetic field.

This calculation makes good physical sense. If we were studying the interaction of a charged scalar field $\Phi$ with an external electromagnetic field we would start with

$$(D_\mu D^\mu + m^2)\Phi = 0 \quad (4)$$

obtained by replacing the ordinary derivative in the Klein-Gordon equation by covariant derivatives. We would then go through the same calculation as in (3). Comparing (4) with (2) we see that the spin of the electron contributes the additional term $(e/2)\sigma^{\mu\nu}F_{\mu\nu}$.

As in Chapter II.1 we write $\psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix}$ in the Dirac basis and focus on $\phi$ since in the nonrelativistic limit it dominates $\chi$. Recall that in that basis $\sigma^{ij} = \varepsilon^{ijk}\begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix}$. Thus $(e/2)\sigma^{\mu\nu}F_{\mu\nu}$ acting on $\phi$ is effectively equal to $(e/2)\sigma^3(F_{12} - F_{21}) = (e/2)2\sigma^3 B = 2e\vec{B}\cdot\vec{S}$ since $\vec{S} = (\vec{\sigma}/2)$. Make sure you understand all the factors of 2! Meanwhile, according to what I told you in chapter II.1, we should write $\phi = e^{-imt}\Psi$, where $\Psi$ oscillates much more slowly than $e^{-imt}$ so that $(\partial_0^2 + m^2)e^{-imt}\Psi \simeq e^{-imt}[-2im(\partial/\partial t)\Psi]$. Putting it all together, we have

$$\left[-2im\frac{\partial}{\partial t} - \vec{\nabla}^2 - e\vec{B}\cdot(\vec{L} + 2\vec{S})\right]\Psi = 0 \quad (5)$$

There you have it! As if by magic, Dirac's equation tells us that a unit of spin angular momentum interacts with a magnetic field twice as much as a unit of orbital angular momentum, an observational fact that had puzzled physicists deeply at the time. The calculation leading to (5) is justly celebrated as one of the greatest in the history of physics.

The story is that Dirac did not do this calculation until a day after he discovered his equation, so sure was he that the equation had to be right. Another version is that he dreaded the possibility that the magnetic moment would come out wrong and that Nature would not take advantage of his beautiful equation.

Another way of seeing that the Dirac equation contains a magnetic moment is by the Gordon decomposition, the proof of which is given in an exercise:

$$\bar{u}(p')\gamma^\mu u(p) = \bar{u}(p')\left[\frac{(p'+p)^\mu}{2m} + \frac{i\sigma^{\mu\nu}(p'-p)_\nu}{2m}\right]u(p) \quad (6)$$

Looking at the interaction with an electromagnetic field $\bar{u}(p')\gamma^\mu u(p)A_\mu(p'-p)$, we see that the first term in (6) only depends on the momentum $(p'+p)^\mu$ and would have been there even if we were treating the interaction of a charged scalar particle with the electromagnetic field to first order. The second term involves spin and gives the magnetic moment. One way of saying this is that $\bar{u}(p')\gamma^\mu u(p)$ contains a magnetic moment component.



## The anomalous magnetic moment

With improvements in experimental techniques, it became clear by the late 1940's that the magnetic moment of the electron was larger than the value calculated by Dirac by a factor of $1.00118 \pm 0.00003$. The challenge to any theory of quantum electrodynamics was to calculate this so-called anomalous magnetic moment. As you probably know, Schwinger's spectacular success in meeting this challenge established the correctness of relativistic quantum field theory, at least in dealing with electromagnetic phenomena, beyond any doubt.

Before we plunge into the calculation, note that Lorentz invariance and current conservation tell us (see Exercise III.6.3) that the matrix element of the electromagnetic current must have the form (here $|p, s\rangle$ denotes a state with an electron of momentum $p$ and polarization $s$)

$$\langle p', s'|J^\mu(0)|p, s\rangle = \bar{u}(p', s')\left[\gamma^\mu F_1(q^2) + \frac{i\sigma^{\mu\nu}q_\nu}{2m}F_2(q^2)\right]u(p, s) \quad (7)$$

where $q \equiv (p' - p)$. The functions $F_1(q^2)$ and $F_2(q^2)$, about which Lorentz invariance can tell us nothing, are known as form factors. To leading order in momentum transfer $q$, (7) becomes

$$\bar{u}(p', s')\left\{\frac{(p'+p)^\mu}{2m}F_1(0) + \frac{i\sigma^{\mu\nu}q_\nu}{2m}[F_1(0) + F_2(0)]\right\}u(p, s)$$

by the Gordon decomposition. The coefficient of the first term is the electric charge observed by experimentalists and is by definition equal to 1. Thus $F_1(0) = 1$. The magnetic moment of the electron is shifted from the Dirac value by a factor $1 + F_2(0)$.

## Schwinger's triumph

Let us now calculate $F_2(0)$ to order $\alpha = e^2/4\pi$. First draw all the relevant Feynman diagrams to this order (Fig. III.6.1). Except for Figure 1b, all the Feynman diagrams are clearly proportional to $\bar{u}(p', s')\gamma^\mu u(p, s)$ and thus contribute to $F_1(q^2)$, which we don't care about. Happy are we! We only have to calculate one Feynman diagram.

It is convenient to normalize the contribution of Figure 1b by comparing it to the lowest order contribution of Figure 1a and write the sum of the two contributions as $\bar{u}(\gamma^\mu + \Gamma^\mu)u$. Applying the Feynman rules, we find

$$\Gamma^\mu = \int \frac{d^4k}{(2\pi)^4}\frac{-i}{k^2}\left(ie\gamma^\nu \frac{i}{\not{p}' + \not{k} - m}\gamma^\mu \frac{i}{\not{p} + \not{k} - m}ie\gamma_\nu\right) \quad (8)$$

I will now go through the calculation in some detail not only because it is important, but also because we will be using a variety of neat tricks springing

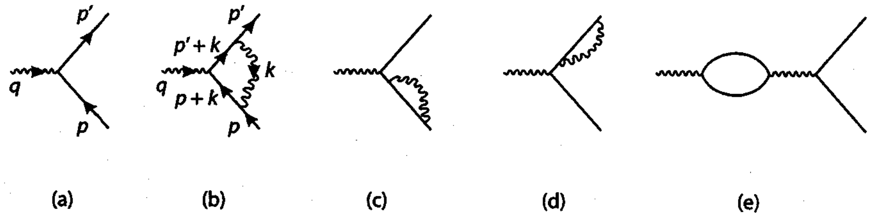


**Figure III.6.1**

from the brilliant minds of Schwinger and Feynman. You should verify all the steps of course.

Simplifying somewhat we obtain $\Gamma^\mu = -ie^2 \int [d^4k/(2\pi)^4](N^\mu/D)$, where

$$N^\mu = \gamma^\nu(\not{p}' + \not{k} + m)\gamma^\mu(\not{p} + \not{k} - m)\gamma_\nu \tag{9}$$

and

$$\frac{1}{D} = \frac{1}{(p'+k)^2 - m^2}\frac{1}{(p+k)^2 - m^2}\frac{1}{k^2} = 2\int d\alpha\, d\beta \frac{1}{\mathcal{D}}. \tag{10}$$

We have used the identity (D.16). The integral is evaluated over the triangle in the ($\alpha$-$\beta$) plane bounded by $\alpha = 0$, $\beta = 0$, and $\alpha + \beta = 1$, and

$$\mathcal{D} = [k^2 + 2k(\alpha p' + \beta p)]^3 = [l^2 - (\alpha+\beta)^2 m^2]^3 + O(q^2) \tag{11}$$

where we completed a square by defining $k = l - (\alpha p' + \beta p)$.

Our strategy is to massage $N^\mu$ into a form consisting of a linear combination of $\gamma^\mu$, $p^\mu$, and $p'^\mu$. Invoking the Gordon decomposition (6) we can write (7) as

$$\bar{u}\left\{\gamma^\mu[F_1(q^2) + F_2(q^2)] - \frac{1}{2m}(p'+p)^\mu F_2(q^2)\right\} u$$

Thus, to extract $F_2(0)$ we can throw away without ceremony any term proportional to $\gamma^\mu$ that we encounter while massaging $N^\mu$. So, let's proceed.

Eliminating $k$ in favor of $l$ in (9) we obtain

$$N^\mu = \gamma^\nu[\not{l} + \not{P}' + m]\gamma^\mu[\not{l} + \not{P} + m]\gamma_\nu \tag{12}$$

where $P'^\mu \equiv (1-\alpha)p'^\mu - \beta p^\mu$ and $P^\mu \equiv (1-\beta)p^\mu - \alpha p'^\mu$. I will use the identities in Appendix D repeatedly, without alerting you every time I use one. It is convenient to organize the terms in $N^\mu$ by powers of $m$. (Here I give up writing in complete grammatical sentences.)

1. The $m^2$ term: a $\gamma^\mu$ term, throw away.
2. The $m$ terms: organize by powers of $l$. The term linear in $l$ integrates to 0 by symmetry. Thus, we are left with the term independent of $l$:



$$m(\gamma^\nu \not{P}'\gamma^\mu\gamma_\nu + \gamma^\nu\gamma^\mu \not{P}\gamma_\nu) = 4m[(1-2\alpha)p'^\mu + (1-2\beta)p^\mu]$$
$$\rightarrow 4m(1-\alpha-\beta)(p'+p)^\mu \tag{13}$$

In the last step I used a handy trick; since $\mathcal{D}$ is symmetric under $\alpha \longleftrightarrow \beta$, we can symmetrize the terms we get in $N^\mu$.

3. Finally, the most complicated $m^0$ term. The term quadratic in $l$: note that we can effectively replace $l^\sigma l^\tau$ inside $\int d^4l/(2\pi)^4$ by $\frac{1}{4}\eta^{\sigma\tau}l^2$ by Lorentz invariance (this step is possible because we have shifted the integration variable so that $\mathcal{D}$ is a Lorentz invariant function of $l^2$.) Thus, the term quadratic in $l$ gives rise to a $\gamma^\mu$ term. Throw it away. Again we throw away the term linear in $l$, leaving [use (D.6) here!]

$$\gamma^\nu \not{P}'\gamma^\mu \not{P}\gamma_\nu = -2\not{P}\gamma^\mu \not{P}'$$
$$\rightarrow -2[(1-\beta)\not{p} - \alpha m]\gamma^\mu[(1-\alpha)\not{p}' - \beta m] \tag{14}$$

where in the last step we remembered that $\Gamma^\mu$ is to be sandwiched between $\bar{u}(p')$ and $u(p)$. Again, it is convenient to organize the terms in (14) by powers of $m$. With the various tricks we have already used, we find that the $m^2$ term can be thrown away, the $m$ term gives $2m(p'+p)^\mu[\alpha(1-\alpha) + \beta(1-\beta)]$, and the $m^0$ term gives $2m(p'+p)^\mu[-2(1-\alpha)(1-\beta)]$. Putting it altogether, we find that $N^\mu \rightarrow 2m(p'+p)^\mu(\alpha+\beta)(1-\alpha-\beta)$

We can now do the integral $\int[d^4l/(2\pi)^4](1/\mathcal{D})$ using (D.11). Finally, we obtain

$$\Gamma^\mu = -2ie^2\int d\alpha\, d\beta\left(\frac{-i}{32\pi^2}\right)\frac{1}{(\alpha+\beta)^2m^2}N^\mu$$
$$= -\frac{e^2}{8\pi^2}\frac{1}{2m}(p'+p)^\mu \tag{15}$$

and thus, trumpets please:

$$F_2(0) = \frac{e^2}{8\pi^2} = \frac{\alpha}{2\pi} \tag{16}$$

Schwinger's announcement of this result in 1948 had an electrifying impact on the theoretical physics community.

I gave you in this chapter not one, but two, of the great triumphs of twentieth century physics, although admittedly the first is not a result of field theory per se.

## Exercises

III.6.1. Evaluate $\bar{u}(p')(\not{p}'\gamma^\mu + \gamma^\mu \not{p})u(p)$ in two different ways and thus prove Gordon decomposition.