FIELD THEORY AND THE STANDARD MODEL

I.J.R.Aitchison, Department of Physics, University of Oxford, UK

Abstract
This is a course of six lectures given at the 2003 European School of High-Energy Physics, Tsakhkadzor, Armenia, 24th August - 6th September 2003. They aim to provide a compact introduction to quantum field theory (in the 'canonical' formalism) and the standard model, focusing on: field quantisation and the canonical route to the Feynman rules; Abelian symmetries and QED; one-loop renormalisation of QED; non-Abelian symmetries; spontaneously broken symmetries; and the electroweak theory.

1. OUTLINE OF THE COURSE

§2 (Lecture 1) Canonical quantisation of free spin-0 (scalar) field. Interacting scalar fields. The Dyson-Wick expansion of the S-matrix. Propagators. Tree graphs. The Yukawa potential.
§3 (Lecture 2) Complex scalar field. Global U(1) phase invariance. Number conservation laws. Fermions. Local U(1) phase invariance and the electromagnetic interaction. The Maxwell field. Elements of QED.
§4 (Lecture 3) One-loop graphs in QED: renormalisation, and running coupling constant.
§5 (Lecture 4) Non-Abelian symmetries, global and local. Local SU(2) symmetry. Gauge field self-interactions. Local SU(3) symmetry. QCD.
§(Lecture 5) Spontaneous symmetry breaking, global and local. Chiral symmetry breaking. The Abelian Higgs model. Spontaneously broken SU(2) x U(1).
§7 (Lecture 6) Introduction to the electroweak theory. The Higgs sector. One loop effects.

2. SCALAR FIELDS : TO TREE GRAPHS

A more leisurely treatment of the material in this section is given in chapters 5 and 6 of volume 1 of the new (third) edition of Aitchison and Hey [1].

2.1 The classical field as an assembly of non-interacting oscillators

Consider a familiar problem, that of a string stretched between points $x = 0$ and $x = L$. The transverse displacement y of the string at position x and time t, $y(x, t)$, satisfies the wave equation

$$\frac{\partial^2 y(x,t)}{c^2 \partial t^2} = \frac{\partial^2 y(x,t)}{\partial x^2} \tag{1}$$

for small displacements. Here $y(x, t)$ is a scalar field: 'scalar' because it has only one component, and 'field' because it varies continuously with x and t. The fundamental method of solving equations like (1) is first to find particular solutions called *modes*, and then to use the fact that (1) is *linear* to write the *general solution as a linear superposition of modes*. Here, the modes must satisfy the boundary conditions $y(0, t) = y(L, t) = 0$, so we try

$$y(x,t) = X_r(t) \sin(\frac{r\pi x}{L}) \tag{2}$$

for $r = 1, 2...$, which expresses the fact that any number of half-wavelengths must fit into the interval $(0, L)$. Substituting (2) into (1) we find

$$X_r(t) = -\omega_r^2 X_r(t) \tag{3}$$

where

$$\omega_r^2 = c^2 r^2 \pi^2 / L^2. \tag{4}$$

Thus each *mode amplitude* $X_r(t)$ executes simple harmonic motion with frequency $\omega_r = (cr\pi/L)$: it acts like the 'coordinate' of an oscillator! The general solution of (1) is then

$$y(x,t) = \sum_{r=1}^{\infty} X_r(t) \sin(\frac{r\pi\, x}{L}); \tag{5}$$

in short, a Fourier series.

Now let's consider the total energy of the vibrating string, which is given by the integral

$$E = \int_0^L [\frac{1}{2}\rho \left(\frac{\partial\, y}{\partial t}\right)^2 + \frac{1}{2}\rho c^2 \left(\frac{\partial\, y}{\partial\, x}\right)^2]dx, \tag{6}$$

where the first term is the kinetic energy 'T' (ρ is the mass per unit length) and the second is the potential energy 'V'. When (5) is placed in (6) and the integral over x done, a remarkable result is obtained (problem P1.1):

$$E = \frac{L}{2} \sum_{r=1}^{\infty} [\frac{1}{2}\, \rho \dot{X}_r^2 + \frac{1}{2}\, \rho\, \omega_r^2\, X_r^2]. \tag{7}$$

Equation (7) has a strikingly simple physical interpretation: the energy of the string is equal to the sum of the energies of individual 'mode oscillators' (recall the energy of one SHO is $\frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\omega^2 x^2$ so here $X_r \leftrightarrow x, \frac{L\rho}{2} \leftrightarrow m, \omega_r \leftrightarrow \omega$). For a general motion of the strings, all the oscillators (modes) will be present. Because the total energy is the *sum* of the individual mode energies, the *modes do not interact with each other*. So, from the point of view of the energy, at least, *the field is equivalent to an assembly of non-interacting oscillators*.

2.2 Quantisation

Let's write $M = \frac{L\rho}{2}$ so that (7) becomes

$$E = \sum_r \frac{1}{2} M \dot{X}_r^2 + \frac{1}{2} M w_r^2 X_r^2. \tag{8}$$

The essential idea is to treat the mode amplitudes 'X_r' as 'quantum coordinate-like variables'. The associated 'momentum-like variables' will be $P_r = M\dot{X}_r$. The energy (8) (which of course in classical physics is a number) becomes now an *operator*, namely the Hamiltonian operator

$$\hat{H} = \sum_{r=1}^{\infty} \frac{\hat{P}_r^2}{2\, M} + \frac{1}{2} M \omega_r^2\, \hat{X}_r^2. \tag{9}$$

We know all about the energy levels and states of a *single* quantum oscillator; the fact that we have here arbitrarily many oscillators doesn't worry us as they are not interacting with each other, so they can be treated quite independently. For a single oscillator of frequency ω, the energy levels are $E_n = (n + \frac{1}{2})\, \hbar\omega$, and the wavefunctions are well-known, in all q.m. textbooks. For our purposes, we prefer the 'operator approach' in terms of $\hat{a}$'s and $\hat{a}^\dagger$'s to the wavefunction one. The essentials are gone through in problem P1.2.

For our vibrating string, then, we simply have

$$\hat{H} = \frac{1}{2} \sum_{r=1}^{\infty} (\hat{a}_r^\dagger\, \hat{a}_r\ + \hat{a}_r\, \hat{a}_r^\dagger)\, \hbar\omega_r, \quad \text{with} \quad [\hat{a}_r, \hat{a}_s^\dagger] = \delta_{rs}. \tag{10}$$

The eigenstates of $\hat{H}$ are *products* of the single oscillator states $|n_1\rangle|n_2\rangle|n_3\rangle$ .... where $|n_1\rangle$ is the state of the oscillator with frequency ω_1, which has energy $(n_1 + \frac{1}{2})\hbar\omega_1$, etc. We can write this more briefly as $|n_1, n_2, ...\rangle$, which has energy $\sum_r (n_r + \frac{1}{2})\hbar\omega_r$. The ground state $|0\rangle$ has *all* n_r's $= 0$, and hence an energy (the 'zero point energy') equal to $\sum_r \frac{1}{2}\hbar\omega_r$.

Thus the energy eigenstates of the quantised field $\hat{y}(x,t)$ are characterised by saying *how many quanta of each frequency* are present; in the ground state there are *no* quanta of excitation present. Such *vibrational quanta* are called 'phonons' in condensed matter physics. *Our 'particles' are similar quanta of excitation of fields.* The state with no excitation quanta is a (too simple!) model of the *vacuum*.

2.3 Free massive real scalar field

We will from now on put $\hbar = c = 1$. The 'classical' field satisfies the Klein-Gordon equation

$$(\Box + m^2)\phi = (\frac{\partial^2\phi}{\partial\, t^2} - \boldsymbol{\nabla}^2\phi + m^2)\phi = 0, \quad \text{where } \Box = \partial_\mu\partial^\mu = \partial_t^2 - \boldsymbol{\nabla}^2 \tag{11}$$

which is the wave equation for a free massive spin-0 (scalar) field. We now consider the field to be in 'infinite space' so Fourier series $\to$ Fourier integrals and our *modes* have the form

$$\phi(\boldsymbol{x}, t) = X_{\boldsymbol{k}}(t)\exp[i\boldsymbol{k}\cdot\boldsymbol{x}]. \tag{12}$$

Plugging this into the K-G equation gives

$$\ddot{X}_{\boldsymbol{k}} = -(m^2 + \boldsymbol{k}^2)X_{\boldsymbol{k}} \tag{13}$$

which again shows that our mode amplitude acts like an SHO, this time with frequency

$$\omega_k = \pm(m^2 + \boldsymbol{k}^2)^{\frac{1}{2}}. \tag{14}$$

The total energy in the field is the obvious generalisation of the energy of the string:

$$E = \frac{1}{2}\int d^3\boldsymbol{x}[\dot{\phi}^2 + (\boldsymbol{\nabla}\phi)^2 + m^2\phi^2]. \tag{15}$$

Once again, this can be written as a 'sum' (in this case, an integral over the Fourier variable $\boldsymbol{k}$) of independent energies for each mode oscillator. So, when *quantised*, we get the Hamiltonian (compare (10))

$$\hat{H}_{KG} = \frac{1}{2}\int \frac{d^3\boldsymbol{k}}{(2\pi)^3}k_0\{\hat{a}^\dagger(\boldsymbol{k})\hat{a}(\boldsymbol{k}) + \hat{a}(\boldsymbol{k})\hat{a}^\dagger(\boldsymbol{k})\} \tag{16}$$

where $k_0 = +(m^2 + \boldsymbol{k}^2)^{\frac{1}{2}}$, and where the mode creation and annihilation operators satisfy

$$[\hat{a}(\boldsymbol{k}), \hat{a}^\dagger(\boldsymbol{k}')] = (2\pi)^3\delta(\boldsymbol{k} - \boldsymbol{k}'), \tag{17}$$

all other commutators vanishing: $[\hat{a}^\dagger(\boldsymbol{k}), \hat{a}^\dagger(\boldsymbol{k}')] = [\hat{a}(\boldsymbol{k}), \hat{a}(\boldsymbol{k}')] = 0$.

Problem P1.3 shows that the state $|p\rangle \propto \hat{a}^\dagger(\boldsymbol{p})|0\rangle$ is an eigenstate of $\hat{H}$ with eigenvalue $\sqrt{m^2 + \boldsymbol{p}^2}$, the expected energy for a particle of mass m and momentum $\boldsymbol{p}$ (note $\hbar = c = 1$). We actually choose the *particular normalisation*

$$|p\rangle = \sqrt{2p_0}\hat{a}^\dagger(\boldsymbol{p})|0\rangle. \tag{18}$$

The general (quantised) solution to the K-G field equation is then

$$\begin{aligned}\hat{\phi}(x) &= \int \frac{d^3\boldsymbol{k}}{(2\pi)^3(2k_0)^{\frac{1}{2}}}\{\hat{a}(\boldsymbol{k})\exp[-ik\cdot x] + \hat{a}^\dagger(\boldsymbol{k})\exp[ik\cdot x]\} \\ &= \hat{\phi}^\dagger(x)\end{aligned} \tag{19}$$

for a 'real' field, and where $k \cdot x = k_0 x_0 - \boldsymbol{k} \cdot \boldsymbol{x}$ and the $(2k_0)^{-\frac{1}{2}}$ is a conventional normalisation factor.

Problem P1.4 shows that

$$\langle 0|\hat{\phi}(x)|p\rangle = e^{-ip\cdot x}. \tag{20}$$

In ordinary quantum mechanics the RHS of this equation would be written as $\langle x|p\rangle$, the x-space wave-function for a state $|p\rangle$ of definite 4-momentum p (which is of course a 4-D plane wave). We can then reasonably assert that the operator

$$\int \frac{d^3\boldsymbol{k}}{(2\pi)^3(2k_0)^{\frac{1}{2}}}\hat{a}^\dagger(\boldsymbol{k})e^{ik\cdot x} \equiv \hat{\phi}^{(-)}(x) \tag{21}$$

creates a quantum at x: $\hat{\phi}^{(-)}(x)|0\rangle = |x\rangle$. (Note that the *other* part of $\hat{\phi}(x)$ gives 0 when acting on $|0\rangle$).

The commutation relations (17) imply that

$$|p_1, p_2\rangle = |p_2, p_1\rangle \tag{22}$$

so our particles are bosons!

2.4 Interactions

In the case of the freely vibrating string, or the free scalar field, the energy is the *sum* of individual mode energies - the modes do not interact. But our particles are precisely mode quanta, and we want them to interact, of course. So we must complicate our simple expressions for field energies in some way. The crucial feature of (6) and (15) which leads to the '$\sum$ mode energies' result is that they are *quadratic* in the fields and their derivatives. *Interactions* will generally be represented by expressions which are cubic or quartic in the fields. Correspondingly, quadratic or cubic expressions will appear in the equations of motion. (Compare the SHO: the 'free' SHO energy is $\frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\omega^2x^2$ with equation of motion $m\ddot{x} = -m\omega^2 x$; if it is perturbed by adding a cubic potential energy λx^3, this produces a force $-\frac{dV}{dx} = -3\lambda x^2$). In the case of lattice vibrations, such 'anharmonic terms' cause the *phonons to interact* - it is the same with our particles. We will introduce an interaction term $\hat{H}^{'}$ in the Hamiltonian:

$$\hat{H}^{'} = \int d^3\boldsymbol{x}\hat{\mathcal{H}}'(\hat{\phi}), \tag{23}$$

for example

$$\hat{\mathcal{H}}' = \lambda(\hat{\phi}(x))^3. \tag{24}$$

We treat $\hat{H}'$ as a perturbation on $\hat{H}_{KG}$.

2.5 Covariant perturbation theory: the Dyson-Wick expansion of the $\hat{S}$ operator, Feynman rules

There is a very compact and powerful formalism for doing relativistic perturbation theory, which we are not going to go through the details of here - just quote the essential results. Transitions are described by means of matrix elements (between free-particle states $|i\rangle$ and $|f\rangle$) of the $\hat{S}$ operator, $\langle f|\hat{S}|i\rangle$, where $\hat{S}$ has the expansion in powers of $\hat{\mathcal{H}}'$:

$$\hat{S} = 1 - i\int d^4x\hat{\mathcal{H}}'(\hat{\phi}(x')) + \frac{1}{2}\int\int d^4x_1 d^4x_2 T(-i\hat{\mathcal{H}}'(x_1). - i\hat{\mathcal{H}}'(x_2)) + \dots \tag{25}$$

where 'T' is the time-ordering operation

$$\begin{aligned} T(\hat{\phi}(x_1)\hat{\phi}(x_2)) &= \hat{\phi}(x_1)\hat{\phi}(x_2) \text{ for } t_1 > t_2 \\ &= \hat{\phi}(x_2)\hat{\phi}(x_1) \text{ for } t_1 < t_2 \end{aligned} \tag{26}$$

i.e. 'earlier on the right'.

Discussion Point: This is supposed to be covariant (relativistically invariant) perturbation theory. But the 'T' symbol seems to be singling out 'time' in some way, and doesn't look '4-D symmetric'. Should we be worried?

Example: 'ABC' theory

To have a little more variety than the single $\hat{\phi}$ field, let's imagine a world with three real scalar fields $\hat{\phi}_A(\text{mass } m_A), \hat{\phi}_B(m_B)$ and $\hat{\phi}_C(m_C)$ with an interaction $g\hat{\phi}_A(x)\hat{\phi}_B(x)\hat{\phi}_C(x)$. This interaction creates or annihilates one each of an A,B or C particle - for example C $\to$ A + B. Suppose $m_C > m_A + m_B$. Then C will be able to decay to A + B. The matrix element for this will be, to lowest order,

$$\int d^4x \langle p_A, p_B| - ig\hat{\phi}_A(x)\hat{\phi}_B(x)\hat{\phi}_C(x)|p_C\rangle. \tag{27}$$

Problem P1.5 shows that this matrix element is equal to $-ig(2\pi)^4\delta^4(p_C - p_A - p_B)$. (Note: creation and annihilation operators for the *different* fields commute with each other). So we have our first 'Feynman rule'!

(i) $-ig$ for an 'A-B-C' vertex

together with an overall factor of $(2\pi)^4\delta(p_{\text{initial}} - p_{\text{final}})$.

Now consider A B $\to$ A B scattering. The lowest order in perturbation theory at which this process can proceed is *second*, via the matrix element

$$\frac{1}{2}\int\int d^4x_1 d^4x_2 \langle p'_A, p'_B|T\{(-ig\hat{\phi}_A(x_1)\hat{\phi}_B(x_1)\hat{\phi}_C(x_1))(-ig\hat{\phi}_A(x_2)\hat{\phi}_B(x_2)\hat{\phi}_C(x_2))\}|p_A, p_B\rangle. \tag{28}$$

Suddenly we have a complicated expression on our hands! Remembering (18), we see that (28) is essentially

$$(16E_AE_BE'_AE'_B)^{\frac{1}{2}}\langle 0|\hat{a}_A(p'_A)\hat{a}_B(p'_B)T\{\hat{\phi}_A(x_1)\hat{\phi}_B(x_1)\hat{\phi}_C(x_1)\hat{\phi}_A(x_2)\hat{\phi}_B(x_2)\hat{\phi}_C(x_2)\}\hat{a}^\dagger_A(p_A)\hat{a}^\dagger_B(p_B)|0\rangle \tag{29}$$

which is the vacuum expectation value (vev) of 10 operators. Remarkably, it can be shown (*Wick's theorem*) that such vev's can be written as a sum of products of all possible choices of *pairwise vev's* (time-ordered vev's, in general). *One* such term is

$$\begin{aligned}\int\int d^4x_1 d^4x_2 &\langle 0|\hat{a}_A(p'_A)\hat{\phi}_A(x_1)|0\rangle\langle 0|\hat{\phi}_A(x_2)\hat{a}^\dagger_A(p_A)|0\rangle\langle 0|\hat{a}_B(p'_B)\hat{\phi}_B(x_2)|0\rangle\times \\ &\times\langle 0|\hat{\phi}_B(x_1)\hat{a}^\dagger_B(p_B)|0\rangle\langle 0|T(\hat{\phi}_C(x_1)\hat{\phi}_C(x_2))|0\rangle \times (16E_AE_BE'_AE'_B)^{\frac{1}{2}}.\end{aligned} \tag{30}$$

Problem 1.4 shows us that the terms with one field and one $\hat{a}$ or $\hat{a}^\dagger$ give just plane waves: two ingoing ones and two outgoing ones, yielding $\exp\, i\{p'_A\cdot x_1 - p_A\cdot x_2 + p'_B\cdot x_2 - p_B\cdot x_1\}$. The *interesting* bit is the remaining *vev of the time-ordered product of two $\hat{\phi}$ fields*, which is the *Feynman propagator* in coordinate space. The physical interpretation of the *two* terms in $\langle 0|T(\hat{\phi}_C(x_1)\hat{\phi}_C(x_2))|0\rangle$, one for $t_1 > t_2$ and one for $t_1 < t_2$ is as follows: A C-quantum is being produced at x_1 and destroyed by x_2, *or* the other way round (*Exercise:* explain why, with the aid of the mode expansion for $\hat{\phi}_C(x)$). So including the incoming and outgoing plane waves we have the physical processes shown in figure 1, and we have to integrate the whole expression in (30) *over all x_1 and x_2*. The result is the Feynman rule in *momentum space* for the scalar propagator (see textbooks):

(ii) a factor $i/[(4-\text{momentum carried by the propagating particle})^2 - (\text{its mass})^2]$

So for the C-exchange process we have the diagram of figure 2, corresponding to the Feynman amplitude $-ig/(q^2 - m_C^2)$ where $q = p_A - p'_B$. In addition, there is the overall factor $(2\pi)^4\delta^4(p_A + p_B - p'_A - p'_B)$.

Points to note:

1 A and B are interacting by 'exchanging a C'.

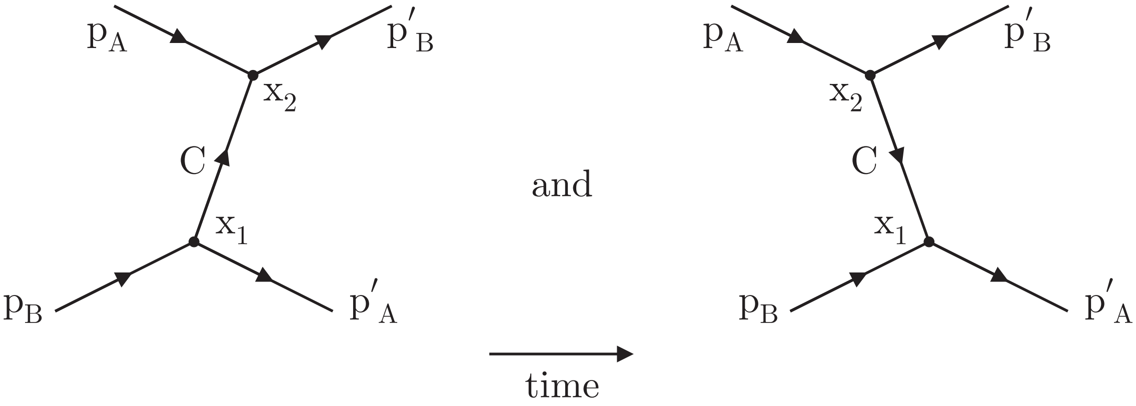

Fig. 1: The two physical processes included in the single Feynman C propagator.

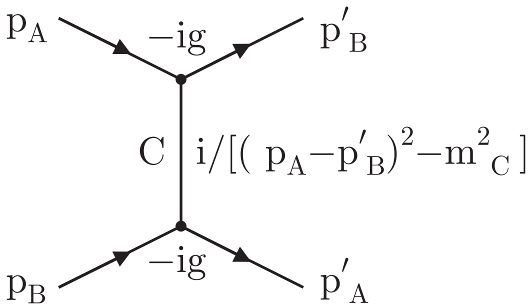

Fig. 2: One-C exchange process in A + B $\rightarrow$ A + B.

2 But the (4-momentum)^2 carried by the exchanged C is *not* equal to m_C^2 - it is 'off mass shell'.

3 *Both* time orderings are included in this *one* momentum space amplitude.

4 Suppose we evaluate the amplitude in the c.m. frame: $p_A = (E_A, \boldsymbol{p}), p_B = (E_B, -\boldsymbol{p}), p'_A = (E_A, -\boldsymbol{p}'), p'_B = (E_B, +\boldsymbol{p}'), |\boldsymbol{p}| = |\boldsymbol{p}'|$. Then $(p_A - p'_B)^2 = (E_A - E_B)^2 - (\boldsymbol{p} - \boldsymbol{p}')^2$. Now consider the static or non-relativistic limit $(E_A - E_B)^2 \ll (\boldsymbol{p} - \boldsymbol{p}')^2$. Our amplitude is now essentially

$$\sim \frac{1}{(\boldsymbol{p} - \boldsymbol{p}')^2 + m_C^2}. \tag{31}$$

We can interpret this in terms of a *potential* associated with the A-B interaction. According to the Born approximation in scattering theory, the amplitude to go from $\boldsymbol{p}$ to $\boldsymbol{p}'$ in the potential $V(\boldsymbol{r})$ is

$$\sim \int \exp\{-i\boldsymbol{p}' \cdot \boldsymbol{r}\} V(\boldsymbol{r}) \exp\{i\boldsymbol{p} \cdot \boldsymbol{r}\} d^3\boldsymbol{r} = \int \exp\{i(\boldsymbol{p} - \boldsymbol{p}') \cdot \boldsymbol{r}\} V(\boldsymbol{r}) d^3\boldsymbol{r} \tag{32}$$

which is some function of $(\boldsymbol{p} - \boldsymbol{p}')^2$. Question: what is $V(\boldsymbol{r})$ such that this function is equal to $[(\boldsymbol{p} - \boldsymbol{p}')^2 + m_C^2]^{-1}$? Answer: $V(\boldsymbol{r}) \propto \exp\{-m_C|\boldsymbol{r}|\}/|\boldsymbol{r}|$, the *Yukawa potential*, of range $1/m_C$; see problem P1.6.

5 Good exercise: think about some of the *other* terms in the Wick expansion of (28)!

Problems for Lecture 1

P1.1 A string is stretched between two points $x = 0$ and $x = L$. The transverse displacement of the string at the point x at time t is $y(x, t)$ where

$$\frac{\partial^2 y}{c^2 \partial t^2} = \frac{\partial^2 y}{\partial x^2}.$$

The general solution can be written as a superposition

$$y(x, t) = \sum_{r=1}^{\infty} X_r(t) \sin \frac{r\pi x}{L}.$$

The total energy of the vibrating string is

$$E = \int_0^L [\frac{1}{2}\rho \left(\frac{\partial y}{\partial t}\right)^2 + \frac{1}{2}\rho c^2 \left(\frac{\partial y}{\partial x}\right)^2] dx$$

where ρ is the mass per unit length. Show that

$$E = \frac{L}{2} \sum_r [\frac{1}{2}\rho \dot{X}_r^2 + \frac{1}{2}\rho \omega_r^2 X_r^2]$$

where $\omega_r = cr\pi/L$. [Hint: write the term $(\frac{\partial y}{\partial t})^2$, for example, as a product of *two independent summations* $(\sum_r \ldots)(\sum_s \ldots)$ and explain why there are no 'cross terms' of the form $X_r X_s \;\; r \neq s$ in the answer.)

P1.2 A one-dimensional harmonic oscillator has the Hamiltonian (energy operator) $\hat{H} = \hat{p}^2/2m + \frac{1}{2}\omega^2 \hat{x}^2$ where $[\hat{x}, \hat{p}] = i$ (units $\hbar = 1$). Define the operators $\hat{a}, \hat{a}^\dagger$ by

$$\hat{a} = \sqrt{\frac{m\omega}{2}} \left(\hat{x} + \frac{i\hat{p}}{m\omega}\right), \quad \hat{a}^\dagger = \sqrt{\frac{m\omega}{2}} \left(\hat{x} - \frac{i\hat{p}}{m\omega}\right).$$

(i) Show that $[\hat{a}, \hat{a}^\dagger] = 1$. (ii) Show that $\hat{H}$ can be written as $\frac{1}{2}\hbar\omega(\hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a})$ or as $\hbar\omega(\hat{a}^\dagger\hat{a} + \frac{1}{2})$. Deduce that $[\hat{a}^\dagger, \hat{H}] = -\omega\hat{a}^\dagger$ and hence show that if $\hat{H}|n\rangle = E_n|n\rangle$ then $\hat{H}\hat{a}^\dagger|n\rangle = (E_n + \omega)|n\rangle$, so

that $\hat{a}^\dagger|n\rangle \propto |n+1\rangle$. State and prove a similar result for $\hat{a}|\, n\rangle$. (iii) Explain why there must be a state $|0\rangle$ such that $\hat{a}|0\rangle = 0$. What is the energy eigenvalue of $|0\rangle$? Deduce the energy spectrum of the oscillator.
P1.3 $\hat{H}_{KG}$ is defined by

$$\hat{H}_{KG} = \frac{1}{2}\int \frac{d^3\boldsymbol{k}}{(2\pi)^3} k_0\{\hat{a}^\dagger(\boldsymbol{k})\hat{a}(\boldsymbol{k}) + \hat{a}(\boldsymbol{k})\hat{a}^\dagger(\boldsymbol{k})\}$$

where $k_0 = +\sqrt{m^2 + \boldsymbol{k}^2}$, with

$$[\hat{a}(\boldsymbol{k}), \hat{a}^\dagger(\boldsymbol{k}')] = (2\pi)^3\delta(\boldsymbol{k} - \boldsymbol{k}')$$

all other commutators vanishing. Show that

$$\hat{H}_{KG}\hat{a}^\dagger(\boldsymbol{p})|0\rangle = p_0\hat{a}^\dagger(\boldsymbol{p})|0\rangle$$

where $p_0 = +\sqrt{m^2 + \boldsymbol{p}^2}$ and $\hat{a}(\boldsymbol{k})|0\rangle = 0$ for all $\boldsymbol{k}$.
P1.4 The field $\hat{\phi}(x)$ has the mode expansion

$$\hat{\phi}(x) = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3(2k_0)^{\frac{1}{2}}}\{\hat{a}(\boldsymbol{k})\mathrm{exp}[-ik\cdot x] + \hat{a}^\dagger(\boldsymbol{k})\mathrm{exp}[ik\cdot x]\}$$

where $k\cdot x = k_0 x_0 - \boldsymbol{k}\cdot\boldsymbol{x}$. Show that

$$\langle 0|\hat{\phi}(x)|p\rangle = e^{-ip\cdot x}$$

where

$$|p\rangle = \sqrt{2p_0}\hat{a}^\dagger(\boldsymbol{p})|0\rangle.$$

P1.5 $\hat{\phi}_A, \hat{\phi}_B, \text{and } \hat{\phi}_C$ are three distinct scalar fields. Evaluate

$$\int d^4x\langle p_A, p_B| - ig\hat{\phi}_A(x)\hat{\phi}_B(x)\hat{\phi}_C(x)|p_C\rangle.$$

P1.6 Evaluate the Fourier transform

$$\int d^3\boldsymbol{r}\mathrm{exp}\{i\boldsymbol{q}\cdot\boldsymbol{r}\}\frac{\mathrm{exp}\{-r/a\}}{r}$$

of a Yukawa potential by following these steps: change $d^3\boldsymbol{r}$ to polar coordinates '$r^2 dr \sin\theta d\theta d\phi$' with the polar axis chosen along the direction of $\boldsymbol{q}$. So $\exp\{i\boldsymbol{q}\cdot\boldsymbol{r}\} = \exp\{i|\boldsymbol{q}|r\cos\theta\}$. Do the integral over θ. Then do the integral over r (the ϕ integral just gives 2π). [Answer: $4\pi/(\boldsymbol{q}^2 + a^{-2})$.]

3. LAGRANGIANS, COMPLEX SCALAR FIELDS, DIRAC AND MAXWELL FIELDS

See chapter 7 of [1].

We have managed to get this far without mentioning the word 'Lagrangian', but now we are going to have to start using this language, which is particularly well suited to the discussion of *symmetries*, and these are of fundamental importance in the Standard Model (SM).

3.1 Lagrangians

This is essentially a formulation of dynamics which is different from (but in the classical case equivalent to) Newton's. The basic quantity here is the Lagrangian function, which in most cases has the form '$L = T - V$' (instead of the energy which is '$E = T + V$'). For a classical particle with coordinate

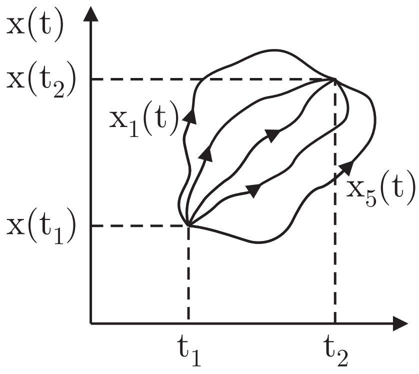

Fig. 3: Possible space-time trajectories between the fixed points $x(t_1)$ and $x(t_2)$.

$x(t)$, L is just $L[x(t), \dot{x}(t)] = \frac{1}{2}m\dot{x}(t)^2 - V(x(t))$. The 'path' $x(t)$ the particle takes is determined by the *principle* that the *action integral* S given by

$$S = \int_{t_1}^{t_2} L[x(t), \dot{x}(t)] = \int_{t_1}^{t_2} [\frac{1}{2}m\dot{x}(t)^2 - V(x(t))]dt \tag{33}$$

is a minimum as all paths $x(t)$ *are searched over*, subject to $x(t_1)$ and $x(t_2)$ being fixed (see figure 3). Problem P2.1 provides a simple example.

Although the action principle seems very different from the *differential equations* of Newton's laws, we can connect them by using a bit of calculus. The actual path must be determined from the condition that small changes away from it make no change in S, to first order (i.e. S is at a minimum). So consider an arbitrary change $x(t) \to x(t) + \delta x(t)$, which also implies $\dot{x}(t) \to \dot{x}(t) + \frac{d}{dt}\delta x(t)$. So then $\dot{x}^2 \to \dot{x}^2 + 2\dot{x}\frac{d}{dt}\delta x$ to first order, and $V(x) \to V(x) + \frac{dV}{dx}\delta x$, giving

$$\delta S = \int_{t_1}^{t_2} [m\dot{x}\frac{d}{dt}(\delta x) - \frac{dV}{dx}\delta x]dt. \tag{34}$$

Now do a partial integration in the first term to get

$$\delta S = \int_{t_1}^{t_2} -[\frac{d}{dt}(m\dot{x}) + \frac{dV}{dx}]\delta x(t)dt, \tag{35}$$

assuming that δx vanishes at the end points (all paths start and finish at the same points). Now it is important to realise that '$\delta x(t)$' here is an *arbitrary* (if 'small') *function of t*. But this change in S, δS, must be zero, by our principle. The only way the integral in (35) can be zero for *arbitrary* $\delta x(t)$ is if the quantity inside the square brackets vanishes, i.e.

$$\frac{d}{dt}(m\dot{x}) = -\frac{dV}{dx} \tag{36}$$

which is exactly Newton's law of motion!

In quantum mechanics, the action approach can also be used, as stated by Dirac and developed by Feynman. There, the amplitude to go from $x(t_1)$ to $x(t_2)$ is proportional to

$$\sum_{\text{all paths } x(t)} \exp\left(\frac{i}{\hbar}\int_{t_1}^{t_2} L(x(t), \dot{x}(t))dt\right) = \sum_{\text{paths}} \exp iS/\hbar. \tag{37}$$

The qualitative idea here is that if the integral is an essentially classical quantity, then its value will be a very large number of $\hbar$'s, so the phase factor will oscillate wildly as the x's change, and everything will cancel out *except* for trajectories such that the action is stationary to small variations around them, since for these ones the phases will 'add up' coherently; hence we get back to the classical action principle in that case.

The action approach can also be used for fields, both classical and quantum; for the latter, see Peter Hasenfratz's lectures. In this course we will not use it for dynamics (i.e. for deriving the Feynman rules), but we will use the Lagrangian language, because it is a very powerful one for discussing symmetries, and because it is quite simply the *lingua franca* of particle physics (at least insofar as the Standard Model is concerned). Before moving to that, we note that the general formulation of (36) is (problem P2.2)

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x} = 0. \tag{38}$$

For fields, we will have to introduce a Lagrangian *density* $\mathcal{L}$ such that (in one space dimension)

$$S = \int\int dt dx \mathcal{L}[\phi(x), \dot{\phi} = \frac{\partial \phi(x)}{\partial t}, \frac{\partial \phi(x)}{\partial x}]; \tag{39}$$

$\dot{\phi} = \frac{\partial \phi}{\partial t}$ is like $\dot{x}$ in (33), and $\frac{\partial \phi}{\partial x}$ is new, but analogous. Again, the field equation for $\phi(x)$ will be determined from the condition that $\delta S = 0$ under $\phi \to \phi + \delta\phi, \quad \dot{\phi} \to \dot{\phi} + \delta\dot{\phi}, \frac{\partial \phi}{\partial x} \to \frac{\partial \phi}{\partial x} + \delta\left(\frac{\partial \phi}{\partial x}\right)$:

$$0 = \delta S = \int dt \int \left[\frac{\partial \mathcal{L}}{\partial \phi}\delta\phi + \frac{\partial \mathcal{L}}{\partial(\partial\phi/\partial x)}\delta\left(\frac{\partial \phi}{\partial x}\right) + \frac{\partial \mathcal{L}}{\partial \dot{\phi}}\delta\dot{\phi}\right] dx. \tag{40}$$

(compare (34)). But $\delta\left(\frac{\partial \phi}{\partial x}\right) = \frac{\partial}{\partial x}\delta\phi$, and similarly for the $\dot{\phi}$ term, so that the second and third terms in (40) can both be integrated by parts, as in (35). As in that case, the variations vanish at the end-points, and since $\delta\phi$ is arbitrary, we deduce the *Euler-Lagrange equation of motion*

$$\frac{\partial \mathcal{L}}{\partial \phi} - \frac{\partial}{\partial x}\left(\frac{\partial \mathcal{L}}{\partial(\partial\phi/\partial x)}\right) - \frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}}\right) = 0. \tag{41}$$

Example. $\mathcal{L} = \frac{1}{2}\dot{\phi}^2 - \frac{1}{2}(\frac{\partial \phi}{\partial x})^2 - \frac{1}{2}m^2\phi^2$. The E-L equation is $\frac{\partial^2 \phi}{\partial t^2} - \frac{\partial^2 \phi}{\partial x^2} + m^2\phi = 0$, the KG equation.

This all generalises to 4-D via

$$\mathcal{L}_{\mathrm{KG}} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2. \tag{42}$$

Here $\partial^\mu = \frac{\partial}{\partial x_\mu}, \partial_\mu = \frac{\partial}{\partial x^\mu}, x^\mu = (x^0, \boldsymbol{x}), x_\mu = (x^0, -\boldsymbol{x}), \partial^\mu\partial_\mu = \frac{\partial}{\partial x^\mu}\frac{\partial}{\partial x_\mu} = \frac{\partial}{\partial x^0}\frac{\partial}{\partial x^0} + \frac{\partial}{\partial \boldsymbol{x}}\cdot(-\frac{\partial}{\partial \boldsymbol{x}}) = \frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2$.

And this generalises to *quantum fields* by putting hats on!

3.2 The complex scalar field

In section 2 we considered a 'real' scalar field for which $\hat{\phi}^\dagger = \hat{\phi}$. The next most complicated thing is a complex scalar field for which $\hat{\phi}^\dagger$ is different from $\hat{\phi}$. So here our mode expansion will have the form

$$\hat{\phi} = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3\sqrt{2\omega}}[\hat{a}(k)e^{-ik\cdot x} + \hat{b}^\dagger(k)e^{ik\cdot x}] \tag{43}$$

The physical interpretation of this is that '$\hat{a}$' will destroy a *particle* (quantum) of the field, while '$\hat{b}^\dagger$' will create an *antiparticle*. This is because states $\hat{a}^\dagger|0\rangle$ and $\hat{b}^\dagger|0\rangle$ are distinguished by having opposite signs of a certain *conserved quantum number*. Now conservation laws have to do with symmetries: what symmetry is at work here? The answer is that it is a symmetry under

$$\hat{\phi} \to e^{-i\alpha}\hat{\phi} \tag{44}$$

i.e a simple phase transformation. Any $\hat{\mathcal{L}}(\hat{\phi})$ which is a function of $\hat{\phi}^\dagger\hat{\phi}$ and $\partial_\mu\hat{\phi}^\dagger\partial^\mu\hat{\phi}$ only will be invariant (symmetric) under (44); for instance the Lagrangian for the free complex KG field

$$\hat{\mathcal{L}} = \partial_\mu\hat{\phi}^\dagger\partial^\mu\hat{\phi} - m^2\hat{\phi}^\dagger\hat{\phi} \tag{45}$$

is invariant under (44).

The symmetry (44) is called a *continuous* symmetry because the phase angle α can be anything (compare 'parity', where the transformation is $\boldsymbol{x} \to -\boldsymbol{x}$ and there's no such thing as a 'small change of parity'). It is also a *global* symmetry, meaning that the parameter α does not depend on the space-time point x; if it did, so that we had $\alpha \to \alpha(x)$ in (44), the symmetry would be called a *local* one. In the case of (45), the Lagrangian can't be invariant under such a local phase change because of the $\partial_\mu\phi$ terms, which will produce $\partial_\mu\alpha$ pieces which won't cancel. But, if we include the electromagnetic field, then we can get a Lagrangian which is invariant under local phase transformations (see section 2.4).

Another piece of jargon we need to introduce is the statement that (44) is a 'U(1)' transformation. The 'U' stands for 'unitary' as in 'unitary matrix'. We can write (44) as $\hat{\phi} \to \mathrm{U}(\alpha)\hat{\phi}$, where the 'matrix' $\mathrm{U}(\alpha)$ has only a single element - i.e. it is a ' 1×1' matrix. A genuine unitary matrix $\mathbf{U}$ satisfies $\mathbf{U}^\dagger\mathbf{U} = \mathbf{1}$, where $\mathbf{1}$ is the identity matrix and the dagger denotes the Hermitian conjugate. A one-dimensional matrix is of course a single number - in this case a complex number. The 'unitary' condition then reduces to $\mathrm{U}^*\mathrm{U} = 1$, which is to say that U is just a phase factor, as in (44). Such phase factors $e^{i\alpha}$ form a *group*: the product $e^{i\alpha}e^{i\beta}$ of any two of them is also a phase factor, and there is an obvious identity (when $\alpha = 0$) and an inverse (replace α by $-\alpha$). Furthermore, this group is *Abelian*, meaning that it doesn't matter in which order we multiply any two U's together: $\mathrm{U}(\alpha)\mathrm{U}(\beta) = \mathrm{U}(\beta)\mathrm{U}(\alpha)$. (As we shall see in Section 4, the symmetries of QCD and of the electroweak theory are precisely *non-abelian generalisations* of (44)). So finally, we say that (44) is a global U(1) transformation, and (45) has a global U(1) symmetry.

The basic theory of such continuous symmetries is supplied by *Noether's theorem*. Because the transformation is continuous, it is good enough to consider an infinitesimal transformation - finite ones can be built up by having lots of little ones. So let's consider an arbitrary $\hat{\mathcal{L}}$ which is invariant under

$$\begin{aligned} \hat{\phi} &\to \hat{\phi}' = \hat{\phi} - i\epsilon\hat{\phi} \\ \hat{\phi}^\dagger &\to \hat{\phi}'^\dagger = \hat{\phi}^\dagger + i\epsilon\hat{\phi}^\dagger. \end{aligned} \tag{46}$$

The change in $\hat{\mathcal{L}}(\hat{\phi}, \hat{\phi}^\dagger, \partial_\mu\hat{\phi}, \partial_\mu\hat{\phi}^\dagger)$ will then be *zero* (because it's invariant), and this change is

$$\begin{aligned} 0 = \delta\hat{\mathcal{L}} &= \frac{\partial\hat{\mathcal{L}}}{\partial(\partial_\mu\hat{\phi})}\delta(\partial_\mu\hat{\phi}) + \frac{\partial\hat{\mathcal{L}}}{\partial(\partial_\mu\hat{\phi}^\dagger)}\delta(\partial_\mu\hat{\phi}^\dagger) \\ &\quad + \frac{\partial\hat{\mathcal{L}}}{\partial\hat{\phi}}\delta\hat{\phi} + \frac{\partial\hat{\mathcal{L}}}{\partial\hat{\phi}^\dagger}\delta\hat{\phi}^\dagger. \end{aligned} \tag{47}$$

This is a bit like the manipulations leading up to the derivation of the Euler-Lagrange equation in Section 3.1, but now the changes $\delta\hat{\phi}$ and $\delta\hat{\phi}^\dagger$ have nothing to do with space-time trajectories - they mix up the two fields via (46). However, we can *use* the equations of motion for $\hat{\phi}$ and $\hat{\phi}^\dagger$ to rewrite $\delta\hat{\mathcal{L}}$ as

$$\begin{aligned} 0 &= \frac{\partial\hat{\mathcal{L}}}{\partial(\partial_\mu\hat{\phi})}\delta(\partial_\mu\hat{\phi}) + \frac{\partial\hat{\mathcal{L}}}{\partial(\partial_\mu\hat{\phi}^\dagger)}\delta(\partial_\mu\hat{\phi}^\dagger) \\ &\quad + \left[\partial_\mu\left(\frac{\partial\hat{\mathcal{L}}}{\partial(\partial_\mu\hat{\phi})}\right)\right]\delta\hat{\phi} + \left[\partial_\mu\left(\frac{\partial\hat{\mathcal{L}}}{\partial(\partial_\mu\hat{\phi}^\dagger)}\right)\right]\delta\hat{\phi}^\dagger. \end{aligned} \tag{48}$$

Since (see similar steps after (40)) $\delta(\partial_\mu \hat{\phi}_i) = \partial_\mu(\delta\hat{\phi}_i)$, the right hand side of (48) is just a total divergence, and (48) becomes

$$0 = \partial_\mu \left[\frac{\partial \hat{\mathcal{L}}}{\partial(\partial_\mu \hat{\phi})} \delta\hat{\phi} + \frac{\partial \hat{\mathcal{L}}}{\partial(\partial_\mu \hat{\phi}^\dagger)} \delta\hat{\phi}^\dagger \right]. \tag{49}$$

This means that the quantity inside the [.....] is a 'current' $\hat{j}^\mu$ which is *conserved* in the sense that $\partial_\mu \hat{j}^\mu = 0$.

This is a general result for *any* $\mathcal{L}$ invariant under (46), and it is an example of Noether's theorem (which states that continuous symmetries imply the existence of conserved currents). For our particular case, with the small changes (46), the quantity in the [...] brackets is

$$\begin{aligned} [\ldots] &= (\partial^\mu \hat{\phi}^\dagger). - i\epsilon\hat{\phi} + (\partial^\mu \hat{\phi}).i\epsilon\hat{\phi}^\dagger \\ &= i\epsilon \left((\partial^\mu \hat{\phi})\hat{\phi}^\dagger - (\partial^\mu \hat{\phi}^\dagger)\hat{\phi} \right) \equiv \epsilon \hat{j}^\mu_\phi. \end{aligned} \tag{50}$$

We drop the irrelevant constant parameter ϵ and arrive at the expression for the conserved current following from the symmetry under (46):

$$\hat{j}^\mu_\phi = i \left(\partial^\dagger \hat{\phi})\hat{\phi}^\dagger - (\partial^\mu \hat{\phi}^\dagger)\hat{\phi} \right). \tag{51}$$

What does all this have to do with conserved quantities? Written out in full, the conservation equation $\partial_\mu \hat{j}^\mu_\phi = 0$ is

$$\partial \hat{j}^0_\phi / \partial t + \boldsymbol{\nabla} \cdot \hat{\boldsymbol{j}}_\phi = 0. \tag{52}$$

Integrating this equation over all space, we obtain

$$\frac{d}{dt} \int_{V \to \infty} \hat{j}^0_\phi d^3\boldsymbol{x} + \int_{S \to \infty} \hat{\boldsymbol{j}}_\phi \cdot d\boldsymbol{S} = 0 \tag{53}$$

where we have used the divergence theorem in the second term. Normally the fields die off sufficiently fast at infinity that the surface integral vanishes, and we can therefore deduce that the quantity $\hat{N}_\phi$ is constant in time, where

$$\hat{N}_\phi = \int \hat{j}^0_\phi d^3\boldsymbol{x}; \tag{54}$$

that is, *the volume integral of the* $\mu = 0$ *component of a symmetry current is independent of time, so its eigenvalues are constants of the motion - i.e. conserved quantum numbers.*

We can calculate $\hat{N}_\phi$ given the field expansion (43). Here we must of course pay attention to the fact that the $\hat{a}$'s and $\hat{b}$'s are mode operators with the commutation relations

$$[\hat{a}(k), \hat{a}^\dagger(k')] = (2\pi)^3 \delta(\boldsymbol{k} - \boldsymbol{k}'), \;\; [\hat{b}(k), \hat{b}^\dagger(k')] = (2\pi)^3 \delta(\boldsymbol{k} - \boldsymbol{k}'), \tag{55}$$

all other commutators vanishing. Also we are defining the vacuum state $|0\rangle$ as being such that $\hat{a}(k)|0\rangle = \hat{b}(k)|0\rangle = 0$. Now, if we go ahead and calculate $\hat{N}_\phi$ in terms of the $\hat{a}$'s and $\hat{b}$'s from (54), we will get some terms in which the rightmost operator is a creation operator; such terms will not give zero when acting on $|0\rangle$. We want the vacuum to be a state with zero eigenvalue of this conserved quantity, and so we *re-order* the expression for $\hat{N}_\phi$, using the commutation relations, so as to arrive at a form in which all $\hat{a}$'s and $\hat{b}$'s appear to the right of all $\hat{a}^\dagger$'s and $\hat{b}^\dagger$'s (this is called 'normal ordered form'-note that we need to do this with the Hamiltonian also!). We discard (infinite) constant contributions arising from the δ-functions on the RHS of (55). Having done this, we find

$$\hat{N}_\phi = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} [\hat{a}^\dagger(k)\hat{a}(k) - \hat{b}^\dagger(k)\hat{b}(k)]. \tag{56}$$

while the Hamiltonian in normally ordered form is

$$\hat{H}_\phi = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} k_0[\hat{a}^\dagger(k)\hat{a}(k) + \hat{b}^\dagger(k)\hat{b}(k)]. \tag{57}$$

So $\hat{N}_\phi$ counts 1 for every 'a' and -1 for every 'b' particle in a state (remember that things like '$\hat{a}^\dagger\hat{a}$' are just number operators), while $\hat{H}_\phi$ counts $+k_0$ for every 'a' and also $+k_0$ for every 'b'. The interpretation then is that free a's and b's of momentum $\boldsymbol{k}$ have the same energy $\sqrt{m^2 + \boldsymbol{k}^2}$, but carry opposite values of the conserved quantum number N_ϕ, which is the eigenvalue of $\hat{N}_\phi$. This is why we interpret $\hat{b}^\dagger$ as the creation operator of an anti-a.

3.3 Fermions

The first step towards getting nearer to the SM is to introduce the quantised Dirac field, which is needed for spin-1/2 particles such as quarks and leptons. The free Dirac equation is

$$i\frac{\partial \psi}{\partial t}(x) = (-i\boldsymbol{\alpha} \cdot \boldsymbol{\nabla} + \beta m)\psi(x) \equiv H_D\psi(x) \tag{58}$$

where the Hamiltonian is thus $H_D = -i\boldsymbol{\alpha} \cdot \boldsymbol{\nabla} + \beta m$, and $\boldsymbol{\alpha}$ and β are the 4×4 Dirac matrices. As in the scalar case, we will promote the 'wave function field $\psi(x)$' into a quantum field operator $\hat{\psi}(x)$ with a mode expansion

$$\hat{\psi} = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3\sqrt{2k_0}} \sum_{s=1,2} [\hat{c}_s(k)u(k,s)e^{-ik\cdot x} + \hat{d}_s^\dagger(k)v(k,s)e^{ik\cdot x}], \tag{59}$$

where $k_0 = (m^2 + \boldsymbol{k}^2)^{1/2}$. Note: (i) $\hat{\psi} \neq \hat{\psi}^\dagger$ - it is a complex Dirac (spinor) field: as with the complex scalar field, this has to do with the fact that its quanta carry a conserved number which distinguishes particle quanta from antiparticle quanta; (ii) u and v are 4-component spinors of positive and negative 4-momentum respectively, such that

$$(\not{k} - m)u(k,s) = 0, \quad (\not{k} + m)v(k,s) = 0 \tag{60}$$

where $\not{k} = \gamma^0 k^0 - \boldsymbol{\gamma} \cdot \boldsymbol{k}$ and $\gamma^0 = \beta, \boldsymbol{\gamma} = \beta\boldsymbol{\alpha}$; (iii) there are two independent spinors u (and two independent v's) for given k, corresponding to the two possible spin states for a spin-1/2 particle, labelled by 's'.

We have written (59) in a form which mimics the complex spin-0 case, suggesting that the $\hat{c}$'s are mode annihilation operators and the $\hat{d}^\dagger$'s are mode creation operators. That is, we expect the vacuum to be such that $\hat{c}_s(k)|0\rangle = 0 = \hat{d}_s(k)|0\rangle$, and that particle states will be formed by applying $\hat{c}_s^\dagger$'s and $\hat{d}_s^\dagger$'s to $|0\rangle$. However, while this seems fine for single particle states, we know very well that a state such as

$$|k_1, s_1; k_2, s_2\rangle \propto \hat{c}_{s_1}^\dagger(k_1)\hat{c}_{s_2}^\dagger(k_2)|0\rangle \tag{61}$$

has to be *antisymmetric* under interchange of the labels $(k_1, s_1) \leftrightarrow (k_2, s_2)$: in particular, the state must be zero (fail to exist) if $k_1 = k_2$ and $s_1 = s_2$ (the Pauli exclusion principle). So these mode operators can't be just like the spin-0 ones.

The solution to this dilemma is simple but radical: for fermions, commutators are replaced by anticommutators! If two different $\hat{c}$'s anticommute, then

$$\hat{c}_{s_1}^\dagger(k_1)\hat{c}_{s_2}^\dagger(k_2) + \hat{c}_{s_2}^\dagger(k_2)\hat{c}_{s_1}^\dagger(k_1) = 0 \tag{62}$$

so that we have the desired antisymmetry

$$|k_1, s_1; k_2, s_2\rangle = -|k_2, s_2; k_1, s_1\rangle. \tag{63}$$

In general we postulate

$$\begin{aligned}\{\hat{c}_{s_1}(k_1), \hat{c}^\dagger_{s_2}(k_2)\} &= (2\pi)^3\delta^3(\boldsymbol{k}_1 - \boldsymbol{k}_2)\delta_{s_1 s_2} \\ \{\hat{c}_{s_1}(k_1), \hat{c}_{s_2}(k_2)\} &= \{\hat{c}^\dagger_{s_1}(k_1), \hat{c}^\dagger_{s_2}(k_2)\} = 0\end{aligned} \tag{64}$$

and similarly for the $\hat{d}$'s and $\hat{d}^\dagger$'s. The factor in front of the δ function depends on the convention for normalising Dirac wavefunctions.

Why does it have to be this way? This is a deep question and has a (rather technical) answer in the famous 'spin-statistics theorem' of quantum field theory. One can get some idea of what goes wrong if we use commutators for fermion modes, by considering the Hamiltonian operator which is

$$\hat{H}_D = \int \hat{\psi}^\dagger(x)(-i\boldsymbol{\alpha}\cdot\boldsymbol{\nabla} + \beta m)\hat{\psi}(x) d^3\boldsymbol{x}. \tag{65}$$

If we place the expansion (59) into (65) we find (after quite a lot of algebra)

$$\hat{H}_D = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} k_0 \sum_{s=1,2} [\hat{c}^\dagger_s(k)\hat{c}_s(k) - \hat{d}_s(k)\hat{d}^\dagger_s(k)]. \tag{66}$$

As with $\hat{H}_\phi$ and $\hat{N}_\phi$ for the scalar field, we would want to re-order the last term in (66) so as to ensure $\hat{H}_D|0\rangle = 0$. But if we do this assuming ordinary commutation relations for the $\hat{d}$'s, we get

$$\hat{H}_D = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} k_0 \sum_{s=1,2} [\hat{c}^\dagger_s(k)\hat{c}_s(k) - \hat{d}^\dagger_s(k)\hat{d}_s(k)]. \tag{67}$$

The problem with (67) is that, although indeed $\hat{H}_D|0\rangle = 0$, there are states with *negative* energy! - namely states with any number of d-quanta (because of the minus sign in front of the number operator $\hat{d}^\dagger\hat{d}$). On the other hand, if we re-order the $\hat{d}\hat{d}^\dagger$ term using *anticommutation* relations, we convert the - sign in (67) into a + sign, and all is well.

We can also see the same mechanism at work if we enquire about a conserved *fermion number*. The Dirac Lagrangian is

$$\hat{\mathcal{L}}_D = \hat{\bar{\psi}}(x)(i\gamma\cdot\partial - m)\hat{\psi}(x) \tag{68}$$

where $\hat{\psi}$ and $\hat{\psi}^\dagger$ are independent degrees of freedom (the E-L equation for $\hat{\bar{\psi}}$ is just the Dirac equation $(i\gamma\cdot\partial - m)\hat{\psi} = 0)$. The Lagrangian (68) is plainly invariant under the global U(1) transformation

$$\hat{\psi}(x) \to \hat{\psi}'(x) = e^{-i\alpha}\hat{\psi}(x). \tag{69}$$

The corresponding (Noether) symmetry current can be found by following the standard steps in Noether's theorem of §3.2, and is

$$\hat{N}^\mu_\psi = \hat{\bar{\psi}}(x)\gamma^\mu\hat{\psi}(x). \tag{70}$$

The associated symmetry operator is

$$\hat{N}_\psi = \int \hat{N}^0_\psi(x) d^3\boldsymbol{x} = \int \hat{\psi}^\dagger(x)\hat{\psi}(x) d^3\boldsymbol{x}, \tag{71}$$

which is just the usual Dirac number density, integrated over $\boldsymbol{x}$. If we now calculate $\hat{N}_\psi$ from (71), we find

$$\hat{N}_\psi = \int \frac{d^3\boldsymbol{k}}{(2\pi)^3} \sum_{s=1,2} [\hat{c}^\dagger_s(k)\hat{c}_s(k) + \hat{d}_s(k)\hat{d}^\dagger_s(k)]. \tag{72}$$

The first term is fine, but if we re-order the second to '$\hat{d}^\dagger \hat{d}$' so that $\hat{N}_\psi | \, 0\rangle = 0$, we will be counting +1 for both c's and d's. We clearly need, again, to use *anticommutators*, so that $\hat{N}_\psi \sim \hat{c}^\dagger \hat{c} - \hat{d}^\dagger \hat{d}$, which counts +1 for each c (particles), and -1 for each d (antiparticles).

We also need the Dirac propagator $\langle 0|T(\hat{\psi}(x_1)\hat{\bar{\psi}}(x_2))|0\rangle$. This may be compared with the analogous propagator for the complex scalar field, namely $\langle 0|T(\hat{\phi}(x_1)\hat{\phi}^\dagger(x_2))|0\rangle$ - see problem P2.3. But note that in the Dirac case, each of $\hat{\psi}$ and $\hat{\psi}^\dagger$ carries an independent spinor index (telling which of the four components it is), so the Dirac propagator is a 4×4 *matrix* in this spinor space. For the Feynman rule appropriate to a propagating fermion we need the momentum space version, as usual. In the scalar case, the propagator is proportional to $1/(q^2 - m^2)$ where q is the momentum carried by the internal particle and m is its mass. The 'poor man's' way of getting this is to take the equation of motion for a free scalar particle (the KG equation)

$$(\partial_t^2 - \boldsymbol{\nabla}^2 + m^2)\phi(x) = 0 \tag{73}$$

and consider a plane wave solution (4-momentum eigenfunction) of the form

$$\phi = A\exp(-iq^0 t + i\boldsymbol{q}\cdot\boldsymbol{x}) = A\exp(-iq\cdot x) \tag{74}$$

giving

$$(-(q^0)^2 + \boldsymbol{q}^2 + m^2)A = (-q^2 + m^2)A = 0 \tag{75}$$

and the propagator is basically the *inverse* of the expression (....) multiplying A in (75), namely $(-q^2 + m^2)^{-1}$. In the Dirac case, an analogous plane wave solution has the form

$$\psi = \exp(-iq\cdot x)u, \tag{76}$$

where u is a 4-component spinor. Inserting (76) into (58) we find

$$(\not{q} - m)u = 0 \tag{77}$$

as in (60), and the inverse of the LHS of (77) is $(\not{q} - m)^{-1}$ (remember that $\not{q}$ is a matrix!). The actual answer is

(iii) a factor $i/(\not{q} - m)$ for an internal fermion line carrying 4-momentum q.

3.4 Local U(1) phase invariance (U(1) gauge theory): QED

Consider the Dirac Lagrangian

$$\hat{\mathcal{L}}_{\mathrm{D}} = \hat{\bar{\psi}}(i\gamma\cdot\partial - m)\hat{\psi}. \tag{78}$$

It is certainly invariant under $\hat{\psi} \to e^{-i\alpha}\psi$ with *constant* α, which is a global U(1) symmetry associated with conservation of the number of ψ-fermions, as we have seen. Let's explore the possibility of invariance under the *local* phase transformation

$$\hat{\psi}(x) \to e^{-i\hat{\alpha}(x)}\hat{\psi}(x) \tag{79}$$

where $\hat{\alpha}(x)$ is a scalar quantum field. Clearly $\hat{\mathcal{L}}_{\mathrm{D}}$ is not invariant under (79): it changes by

$$\delta\hat{\mathcal{L}}_{\mathrm{D}} = \hat{\bar{\psi}}(x)\gamma^\mu\hat{\psi}(x)\partial_\mu\hat{\alpha}(x). \tag{80}$$

Now, in classical electrodynamics, the way in which electromagnetic interactions are introduced in the Hamiltonian formulation of dynamics is via the replacement of the momentum variable p^μ by $p^\mu - eA^\mu$, where $e(>0)$ is the particle's charge and $A^\mu = (V, \boldsymbol{A})$ is the 4-vector of electromagnetic potentials V

and $\boldsymbol{A}$ such that $\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}$ and $\boldsymbol{E} = -\boldsymbol{\nabla} V - \partial \boldsymbol{A}/\partial t$. In quantum mechanics, we follow the same prescription, but now $p^\mu \to \hat{p}^\mu = i\partial^\mu$ and electromagnetism is introduced via $i\partial^\mu \to i\partial^\mu - eA^\mu$, or

$$\partial^\mu \to \partial^\mu + ie\hat{A}^\mu \equiv \hat{D}^\mu. \tag{81}$$

Applying this prescription to $\hat{\mathcal{L}}_{\mathrm{D}}$, we generate an *interaction*

$$\hat{\mathcal{L}}_{\mathrm{int}} = -e\bar{\hat{\psi}}\gamma^\mu\hat{\psi}\hat{A}_\mu. \tag{82}$$

Now, if $\hat{A}_\mu$ were also to change by exactly the rule

$$\hat{A}_\mu \to \hat{A}_\mu + \frac{1}{e}\partial_\mu\hat{\alpha} \tag{83}$$

when $\hat{\psi}$ changes by (79), the term (80) will be cancelled and the complete Lagrangian $\hat{\mathcal{L}}_{\mathrm{D}} + \hat{\mathcal{L}}_{\mathrm{int}}$ would be *locally* U(1) invariant.

Of course, this is indeed the case. The electromagnetic potentials *are* arbitrary up to 'gauge transformations' of the form (83) (consider for example just the 3-vector part: $\hat{\boldsymbol{A}} \to \hat{\boldsymbol{A}} + \frac{1}{e}\boldsymbol{\nabla}\hat{\alpha}$, and $\hat{\boldsymbol{B}} = \boldsymbol{\nabla} \times \hat{\boldsymbol{A}}$ remains the same because curl grad =0). So the combined transformations

$$\begin{aligned} \hat{\psi}(x) &\to e^{-i\hat{\alpha}(x)}\hat{\psi}(x) \\ \hat{A}_\mu(x) &\to \hat{A}_\mu + \frac{1}{e}\partial_\mu\hat{\alpha}(x) \end{aligned} \tag{84}$$

are what we mean by a U(1) gauge transformation. *Note* that the interaction is the 4-dimensional dot product of the gauge field $\hat{A}_\mu(x)$ and the 'global U(1) symmetry current' $\bar{\hat{\psi}}\gamma^\mu\hat{\psi}$.

Like our other quantum fields, $\hat{A}^\mu(x)$ has a mode expansion:

$$\hat{A}^\mu(x) = \sum_{\lambda=0}^{3}\int \frac{d^3\boldsymbol{k}}{(2\pi)^3\sqrt{2\omega}}[\epsilon^\mu(k,\lambda)\hat{\alpha}_\lambda(k)e^{-ik\cdot x} + \epsilon^{*\mu}(k,\lambda)\hat{\alpha}_\lambda^\dagger(k)e^{ik\cdot x}] \tag{85}$$

where $\epsilon^\mu(k,\lambda)$ is the 'polarisation vector' of the plane wave solution ($\lambda = 0, 1, 2, 3$). $\hat{A}^\mu$ is real (because the photon is its own antiparticle), and $\epsilon^\mu(k,\lambda)$ is a 'spin-1 analogue' of the spinor $u(p,s)$ for the Dirac field.

But this '$\hat{A}^\mu$' is itself a dynamical field, of course. What is *its* Lagrangian? To answer this, we need to find an $\hat{\mathcal{L}}_A$ such that, if that was all we had, the E-L equations of motion would give us the free-space (source-free) Maxwell equations. Now Maxwell's equations are for the field strengths $\hat{\boldsymbol{E}}$ and $\hat{\boldsymbol{B}}$, not the potentials, so they are automatically unchanged under the transformation (83) - that is, they are gauge invariant. This suggests that we need to use the gauge invariant object

$$\hat{F}_{\mu\nu} = \partial_\mu\hat{A}_\nu - \partial_\nu\hat{A}_\mu \tag{86}$$

to build our $\hat{\mathcal{L}}_A$ (it is easy to check that $\hat{F}_{\mu\nu}$ is invariant under (83)). Indeed, the Maxwell Lagrangian is

$$\hat{\mathcal{L}}_A = -\frac{1}{4}\hat{F}_{\mu\nu}\,\hat{F}^{\mu\nu}. \tag{87}$$

How do we know? By verifying that indeed the E-L equations for $\hat{A}_\mu$ following from $\hat{\mathcal{L}}_A$ *are* the free-space Maxwell equations (warning: this needs some patience to do correctly, first time!).

So actually we are now in possession of the QED Lagrangian

$$\hat{\mathcal{L}}_{QED} = \bar{\hat{\psi}}(i\,\partial\!\!\!/ - m)\hat{\psi} - e\bar{\hat{\psi}}\gamma^\mu\hat{\psi}\hat{A}_\mu - \frac{1}{4}\hat{F}_{\mu\nu}\,\hat{F}^{\mu\nu} \equiv \bar{\hat{\psi}}(i\,D\!\!\!\!/ - m)\hat{\psi} - \frac{1}{4}\hat{F}_{\mu\nu}\,\hat{F}^{\mu\nu} \tag{88}$$

for one fermion of charge e and mass m. *It is invariant under local U(1) transformations - i.e. it is gauge invariant*. What are the Feynman rules? We have the fermion propagator: we need the interaction vertex, and the $\hat{A}^\mu$ (photon) propagator. First, the vertex. Remember that '$\mathcal{L} = T - V$', so the interaction Hamiltonian is

$$\hat{H}' = \int e\bar{\hat{\psi}}\gamma^\mu\hat{\psi}\hat{A}_\mu d^4x. \tag{89}$$

In perturbation theory we always get '$-i\hat{H}'$'. So a lowest order matrix element will be

$$\langle f| - ie\int \bar{\hat{\psi}}\gamma^\mu\hat{\psi}\hat{A}_\mu d^4x|i\rangle. \tag{90}$$

Just as in the 'ABC' case, the amplitude for the elementary building block '$e^- \to e^- + \gamma$' will be just

(iv) $ie\gamma^\mu$

with appropriate factors for an incoming fermion (a u spinor), an outgoing fermion (a $\bar{u}$ spinor), and the γ (ϵ_μ for an ingoing γ, ϵ^*_μ for an outgoing one).

The only other thing we need is the photon propagator, and here we hit an unpleasant snag, which should not be concealed. Let's try to follow the 'poor man's' way of getting propagators in this case. We start with the E-L equation of motion for the A^μ field, which turns out to be

$$\Box A^\nu - \partial^\nu(\partial_\mu A^\mu) = 0 \tag{91}$$

(see problem P2.4). Now try plugging in a free particle plane wave solution $A^\nu \sim \exp(iq\cdot x)\epsilon^\nu$. We get

$$(-q^2\delta^\nu_\mu + q^\nu q_\mu)\epsilon^\mu = 0. \tag{92}$$

The propagator should be basically $(q^2\delta^\nu_\mu + q^\nu q_\mu)^{-1}$. But this inverse doesn't exist! It's obvious that

$$(-q^2\delta^\nu_\mu + q^\nu q_\mu)q^\mu) = 0 \tag{93}$$

so that treated as a matrix it has a zero eigenvalue; hence its determinant must vanish, and its inverse therefore will not exist.

The propagator *should* be something like $\langle 0|T(\hat{A}^\mu(x_1)\hat{A}^\nu(x_2))|0\rangle$, but as we have seen the $\hat{A}^\mu$'s are not unique, and can be altered by a gauge transformation (83). So *the propagator is in fact gauge dependent*, not a unique quantity, and that's why the naive poor man's approach failed. In classical electromagnetic theory, one 'fixes the gauge', for example by imposing the condition $\partial_\mu A^\mu = 0$, which reduces (91) to $\Box A^\nu = 0$, and then the plane wave solution gives $-q^2\epsilon^\mu = 0$ and the propagator $\sim 1/q^2$ (as expected!). But in general we must acknowledge the gauge dependence. A standard form for the γ propagator is

(v) a factor $i[-g^{\mu\nu} + (1-\xi)q^\mu q^\nu/q^2]/q^2$ for an internal photon line carrying 4-momentum q, where ξ is a 'gauge parameter' ($\xi = 1$ gives the simple $1/q^2$ form).

Results for physical quantities will always be independent of ξ (i.e. will be gauge invariant) , but it is not so simple to give a general proof of this.

Problems for Lecture 2

P2.1 The 'action' in classical mechanics is defined by

$$S = \int_{t_1}^{t_2}[\frac{1}{2}m(\dot{x}(t))^2 - V(x(t))]dt.$$

Consider one-dimensional motion under gravity with $V = -mgx(t)$. Evaluate S for $t_1 = 0, t_2 = T$, for three alternative trajectories: (a) $x(t) = at$; (b) $x(t) = \frac{1}{2}gt^2$ (the Newtonian one); and (c) $x(t) = bt^3$. [Take care to choose a and b so that all trajectories end at the same point.] P2.2 The classical action is

$$S = \int_{t_1}^{t_2} L[x(t), \dot{x}(t)]dt$$

where L is the Lagrangian. Under an infinitesimal change of trajectory $x(t) \rightarrow x(t) + \delta x(t), \dot{x}(t) \rightarrow \dot{x}(t) + \frac{d}{dt}\delta x(t)$ the action changes by

$$\delta S = \int_{t_1}^{t_2} [\frac{\partial L}{\partial x}\delta x + \frac{\partial L}{\partial \dot{x}}\frac{d}{dt}\delta x]dt.$$

The classical path is determined from the condition $\delta S = 0$. Show that this implies

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x} = 0.$$

P2.3 Discuss the interpretation of $\langle 0|T(\hat{\phi}(x_1)\hat{\phi}^\dagger(x_2))|0\rangle$ for both time-orderings. P2.4 Maxwell's equations are

$$\boldsymbol{\nabla} \cdot \boldsymbol{E} = 0,\ \boldsymbol{\nabla} \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t},\ \boldsymbol{\nabla} \cdot \boldsymbol{B} = 0,\ \boldsymbol{\nabla} \times \boldsymbol{B} = \frac{\partial \boldsymbol{E}}{\partial t}.$$

In quantum mechanics, electromagnetic interactions are introduced via the *potentials* V and $\boldsymbol{A}$ defined by

$$\boldsymbol{E} = -\boldsymbol{\nabla} V = \frac{\partial \boldsymbol{A}}{\partial t},\ \boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}.$$

Then $\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0$ and $\boldsymbol{\nabla} \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}$ are satisfied automatically, while the other two Maxwell equations become

$$(\partial_t^2 - \boldsymbol{\nabla}^2)\boldsymbol{A} + \boldsymbol{\nabla}(\frac{\partial V}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{A}) = 0$$

and

$$(\partial_t^2 - \boldsymbol{\nabla}^2)V - \frac{\partial}{\partial t}(\frac{\partial V}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{A}) = 0.$$

(i) Verify these last two equations.
(ii) Verify that they can be put into a neat *covariant* form by introducing the 4-vector $A^\mu = (V, \boldsymbol{A})$, namely

$$\Box A^\nu - \partial^\nu(\partial_\mu A^\mu) = 0$$

where $\Box \equiv \partial_t^2 - \boldsymbol{\nabla}^2,\ \partial^\nu \equiv \frac{\partial}{\partial x_\nu},\ \partial_\mu A^\mu = \frac{\partial V}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{A}$. [Note that $x_i = -x^i$ for $i = 1, 2, 3$; so $\partial^i = \frac{\partial}{\partial x_i} = -\frac{\partial}{\partial x^i} = (-\boldsymbol{\nabla})_{i \text{ component}}$, and $\partial_i = \frac{\partial}{\partial x^i} = (+\boldsymbol{\nabla})_{i \text{ component}}$. So $\partial_\mu A^\mu = \partial_0 A^0 + \partial_i A^i = \frac{\partial V}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{A}$].
P2.5 Show that

$$[\hat{D}_\mu, \hat{D}_\nu] = ie\hat{F}_{\mu\nu}$$

(see (81)). Hint: in working with such commutators of differential operators, it is best to put in an arbitrary function for the operators to act on, on both sides.

P2.6 A photon mass term in the Lagrangian would give a term proportional to $\hat{A}^\mu \hat{A}_\mu$. Show that this is *not* gauge invariant.

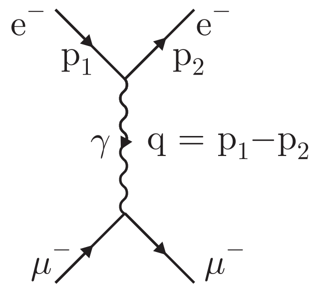

Fig. 4: γ-exchange amplitude in $e^-\mu^- \to e^-\mu^-$.

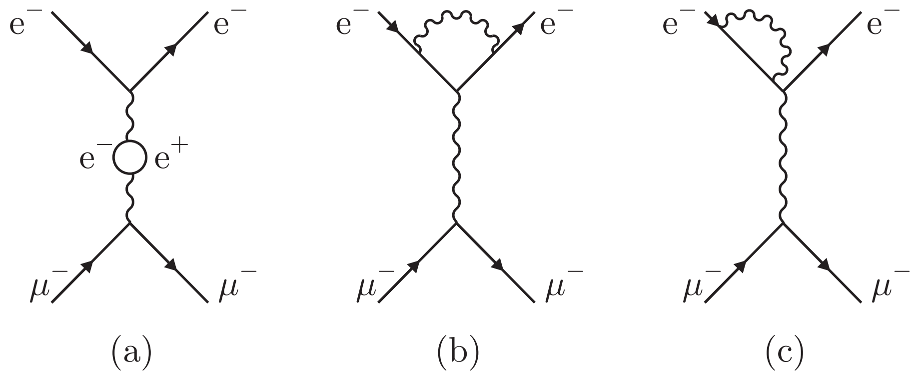

Fig. 5: One-loop corrections to figure 4.

4. ONE-LOOP GRAPHS IN QED: RENORMALISATION, AND RUNNING COUPLING CONSTANT

See chapters 10 and 11 of [1].

Feynman diagrams represent terms in a perturbation theory expansion of physical amplitudes, where the expansion parameter is the relevant 'charge' of the theory - 'e' for QED, or more precisely the fine structure constant $\alpha = e^2/4\pi$. The lowest order graphs for any process are always the ones with the fewest vertices, and this means, in fact, that for given external 'legs', each vertex must be joined to only one other vertex by a single internal line (propagator); for example, the γ-exchange amplitude in $e^-\mu^- \to e^-\mu^-$ shown in figure 4. Such graphs are called 'tree' graphs.

But tree graphs will only give us the lowest order contribution to the amplitudes. As soon as we go to the next order in perturbation theory, we meet *loops* - for example, those shown in figures 5 (a), (b) and (c), which are O(α^2) (four powers of e) diagrams in $e^-\mu^- \to e^-\mu^-$. Admittedly, since $\alpha \sim 1/137$ is quite small, such corrections would seem to be relatively insignificant, perhaps. But, as you all know very well, there are certain quantities (such as the anomalous magnetic moments of the e and the μ) which are known with truly remarkable precision (typically 0.1%), well beyond that represented by the simplest lowest order calculation. More to the point for this school, LEP and SLAC experiments had an accuracy sensitive to one-loop corrections; hence an understanding of this physics is now essential for phenomenolgy.

As soon as one tries to calculate a loop, in nearly all quantum field theories, one finds that it is infinite! This is pretty disastrous, particularly as loops are supposed to be a small correction to the tree graphs (if the expansion parameter is small, as α is). Thus at once we are faced with the whole business of *renormalisation*, which is a systematic procedure for 'taming' these infinities. All three gauge theories of the standard model are 'renormalisable', meaning that higher order corrections can in fact be reliably

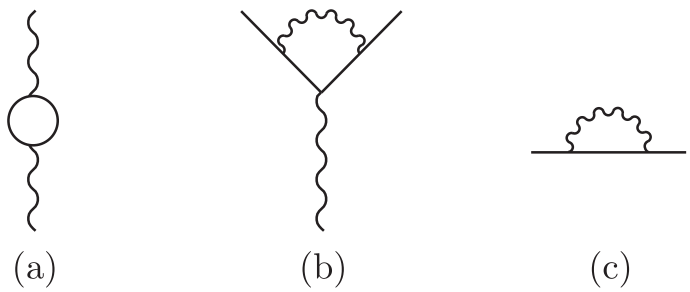

Fig. 6: The loop parts of figure 5.

calculated. The remarkable agreement between theory and experiment is impressive confirmation that the rather elaborate theoretical structure of these theories is actually a good model of nature at this scale. However, the renormalisation of non-abelian gauge theories is too technical for this course, and here I shall sketch how it works for QED only.

The loop bits in figure 5 are, in fact, the only divergent one-loop graphs in QED; we redraw them separately in figures 6 (a)-(c). Figure 6(a) is clearly a correction to the photon propagator, and is called generically a 'vacuum polarisation' graph (see section 5.3), (b) is a 'vertex correction' and (c) is a correction to a fermion propagator. We are going to concentrate on (a).

4.1 Vacuum polarisation and the photon self-energy

We shall use the gauge $\xi = 1$ in which the unmodified photon propagator is $-ig^{\mu\nu}/q^2$. The amplitude for figure 6(a) is (omitting Dirac spinor factors for the fermion lines)

$$\frac{-ig^{\nu\rho}}{q^2}\left(i\Pi^{[2]}_{\rho\sigma}(q^2)\right)\frac{-ig^{\sigma\mu}}{q^2} \tag{94}$$

where

$$i\Pi^{[2]}_{\rho\sigma}(q^2) = (-1)(-ie)^2\mathrm{Tr}\int\frac{d^4k}{(2\pi)^4}\frac{i}{\not{q}+\not{k}-m}\gamma_\rho\frac{i}{\not{k}-m}\gamma_\sigma. \tag{95}$$

Note: (i) When we attach external legs to figure 6(a), as in figure 5(a), 'q' will be determined in terms of the 4-momenta of the external particles, but this q is *shared* by the e^+ and e^- in the loop in all possible ways: the e^+ has 4-momentum k, say, in the direction indicated, and the e^- has $q + k$, but nothing determines k - it has to be *integrated over*. (ii) The (-1) factor has to be included for all closed fermion loops, as does the Tr (which means 'take the trace - i.e. sum the diagonal elements - of the Dirac matrix product').

The $\int d^4x$ in (95) extends over the (presumably) infinite 4-D 'volume'; in particular, all components of k can go to infinity. So a crude 'counting of powers' seems to show that (95) will diverge as

$$\int d^4k/k^2 \sim \int k^3dk/k^2 \sim \int kdk \sim \Lambda^2 \tag{96}$$

if we 'cut-off' the integral at an upper limit Λ. This would be a (divergent) *constant* contribution, multiplying $g_{\rho\sigma}$ to get the indices right. What would such a constant loop correction mean , in this case? Suppose we consider a whole series of such 'insertions', as shown in figure 7 - which is, in fact, a geometric series of the form '$\frac{1}{a}+\frac{1}{a}b\frac{1}{a}+\frac{1}{a}b\frac{1}{a}b\frac{1}{a}\ldots$', summing to $\frac{1}{a(1-b/a)} = \frac{1}{a-b}$. In the present case, then, this would mean that a constant part of $\Pi^{[2]}_{\rho\sigma}$ will correct the propagator (after summing) to something of the form $(q^2 - \mathrm{const})^{-1}$ - in other words, the photon will apparently acquire a mass!

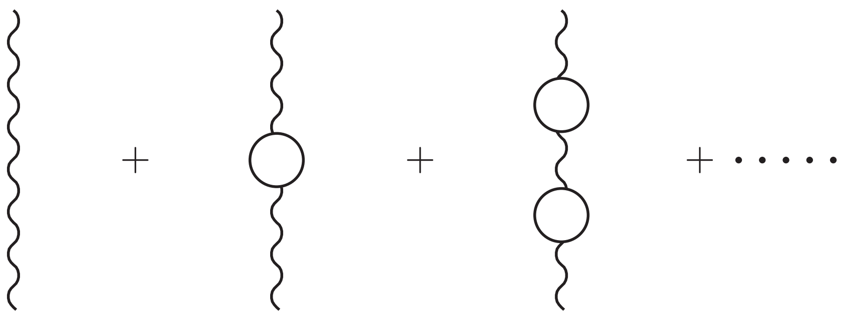

Fig. 7: Sum of vacuum polarisation 'bubble' insertions.

Actually, such insertions into propagators usually do have the effect of shifting the mass of the particle in question, and they are generically called 'self-energies' (e.g. figure 6(c) is a fermion self-energy, which will indeed modify the original fermion mass). But the real photon is massless! We know this to a very high accuracy experimentally. Theoretically, this is fundamentally related to *gauge invariance* - see problem P2.6! So, provided we introduce the cut-off in a *gauge-invariant* way, it turns out that this apparent Λ^2 divergence of (95) is not there after all. Instead, what one finds is that

$$i\Pi^{[2]}_{\rho\sigma}(q^2) = i(q^2 g_{\rho\sigma} - q_\rho q_\sigma)\Pi^{[2]}_\gamma(q^2) \tag{97}$$

where $\Pi^{[2]}_\gamma(q^2)$ is a Lorentz scalar, and is given by an integral which diverges more 'weakly', namely as $\ln\Lambda$. Note that the dimensions of $\Pi^{[2]}_{\rho\sigma}(q^2)$ are M^2: in the 'naive' cut-off approach this was visible in the Λ^2, whereas in (97) quadratic factors of q appear, and this is why the divergence can only be logarithmic. These factors ensure that

$$q^\rho \Pi^{[2]}_{\rho\sigma} = q^\sigma \Pi^{[2]}_{\rho\sigma} = 0 \tag{98}$$

(assuming $\Pi^{[2]}_\gamma$ is finite!); this guarantees that the ξ-dependent part of the propagator (rule(v)) disappears - i.e. the result is gauge invariant, as required.

When all the bubbles are added up, and bits proportional to q are omitted because of (98) (gauge invariance), one finds the net result that the photon propagator is modified according to

$$\frac{-ig_{\mu\nu}}{q^2} \to \frac{-ig_{\mu\nu}}{q^2\left(1 - \Pi^{[2]}_\gamma(q^2)\right)}. \tag{99}$$

What is the physics of this? When (99) appears inside a scattering graph such as figure 5(a), we would still be able to say that the (corrected) exchanged photon had zero mass, since near the 'mass shell' point $q^2 = 0$ (99) does indeed behave like (a constant times) the massless propagator $1/q^2$, provided that $\Pi^{[2]}_\gamma(q^2 = 0)$ is finite.

Discussion point: what happens if $\Pi^{[2]}_\gamma(q^2 = 0)$ itself has a term like A/q^2? and how might this happen?

On the other hand, the propagator will have a peculiar *normalisation*: it will be

$$\sim \left(\frac{1}{1 - \Pi^{[2]}_\gamma(0)}\right) \cdot \frac{-ig_{\mu\nu}}{q^2} \tag{100}$$

for $q^2 \approx 0$ instead of the familiar $-ig_{\mu\nu}/q^2$. Why is this? The propagator in the *free* case was the Fourier transform of $\langle 0|T(A_\mu(x_1)A_\nu(x_2))|0\rangle$. Take one time-ordering, say $\langle 0|A_\mu(x_1)A_\nu(x_2)|0\rangle$, and insert a complete set of free states via '$\sum_n |n\rangle\langle n| = 1$':

$$\sum_n \langle 0|A_\mu(x_1)|n\rangle\langle n|A_\nu(x_2)|0\rangle. \tag{101}$$

The *only* state 'n' which can contribute is the state of one free photon - and indeed we know that matrix elements of the form $\langle 0|\text{field operator}|\text{particle state}\rangle$ are always just the corresponding *wavefunction*. But now consider the interacting case. Here the full propagator is $\langle\Omega|T(A_\mu(x_1)A_\nu(x_2))|\Omega\rangle$ where $|\Omega\rangle$ is the exact 'interacting' vacuum. Insert a complete set of interacting states $\sum_n |\bar{n}\rangle\langle\bar{n}| = 1$: then the analogue of (101) is

$$\sum_n \langle\Omega|A_\mu(x_1)|\bar{n}\rangle\langle\bar{n}|A_\nu(x_2)|\Omega\rangle \tag{102}$$

and now the crucial point is that in addition to the one-photon state in $|\bar{n}\rangle$ there will also be a whole lot of *other* states to which the photon can couple - for instance, precisely the e^+e^- state in our vacuum polarisation graph! This must mean that the $|1\gamma\rangle$ state cannot any longer, by itself, produce all of the '1' in the completeness sum. So the 'strength' of the matrix element $\langle 1\gamma|A_\mu(x)|\Omega\rangle$ cannot be unity (in the normalisation we are adopting, like problem P1.4).

To take account of this 'diminished single particle strength', we write

$$\langle\gamma, k, \lambda|A_\mu(x)|\Omega\rangle = \sqrt{Z_3}\epsilon^*_\mu(\lambda)e^{ik.x} \tag{103}$$

where Z_3 is called the wavefunction renormalisation constant. This will mean that the interacting propagator has the form

$$\begin{aligned}&\text{F.T. of}\langle\Omega|T(A_\mu(x_1)A_\nu(x_2))|\Omega\rangle\\ &= \tfrac{-iZ_3 g_{\mu\nu}}{q^2} + \text{contributions from non single particle states,}\end{aligned} \tag{104}$$

for $q^2 \approx 0$. So we can identify

$$Z_3 = \frac{1}{1-\Pi_\gamma^{[2]}(0)}. \tag{105}$$

This is the interpretation of the change in normalisation of the photon propagator.

This is all innocent-sounding enough ... but of course $\Pi_\gamma^{[2]}(0)$ depends on Λ and is *divergent* as the cut-off $\Lambda \to \infty$. To bury this divergence, which after all is occurring as a multiplicative factor in the wavefunction, we introduce the 'physical' (*renormalised*) photon field operator $A_{\mu,\text{ph}}$ defined by

$$A_{\mu,\text{ph}}(x) = \frac{1}{\sqrt{Z_3}}A_\mu(x) \tag{106}$$

for which the propagator will be of the expected form

$$\text{F.T. of}\langle\Omega|T(A_{\mu,\text{ph}}(x_1)A_{\nu,\text{ph}}(x_2))|\Omega\rangle \approx \frac{-ig_{\mu\nu}}{q^2} + \text{multiparticle bits} \tag{107}$$

for $q^2 \approx 0$. Formally this will work even if $\Lambda \to \infty$; the *physical* matrix elements are OK. Note that $Z_3 = Z_3(\Lambda)$, from (105), since $\Pi_\gamma^{[2]}$ depends on Λ.
Discussion point: Do we actually envisage $\Lambda \to \infty$, really?

Now let's tidy up. Our results so far tell us that the *renormalised* γ-propagator is $Z_3^{-1}\times$ the one we have been calculating to $O(\alpha)$, that is

$$\frac{1}{Z_3}\cdot\frac{-ig_{\mu\nu}}{q^2\left(1-\Pi_\gamma^{[2]}(q^2,\Lambda)\right)} \tag{108}$$

where we now indicate the Λ dependence explicitly. Now

$$Z_3(\Lambda) = [1-\Pi_\gamma^{[2]}(0,\Lambda)]^{-1} \approx [1+\Pi_\gamma^{[2]}(0,\Lambda)] \tag{109}$$

since $\Pi^{[2]} \sim \alpha$ and we are doing a systematic order-by-order perturbative approach. So (108) becomes

$$\approx \frac{-ig_{\mu\nu}}{q^2\left(1 - \Pi_\gamma^{[2]}(q^2,\Lambda) + \Pi_\gamma^{[2]}(0,\Lambda)\right)} \tag{110}$$

again dropping the $O(\alpha^2)$ term $\Pi_\gamma^{[2]}(q^2)\Pi_\gamma^{[2]}(0)$. So finally our renormalised propagator is

$$\frac{-ig_{\mu\nu}}{q^2\left(1 - \bar{\Pi}_\gamma^{[2]}(q^2)\right)} \tag{111}$$

where

$$\bar{\Pi}_\gamma^{[2]}(q^2) = \lim_{\Lambda\to\infty} [\Pi_\gamma^{[2]}(q^2,\Lambda) - \Pi_\gamma^{[2]}(0,\Lambda)] \tag{112}$$

is called the 'once-subtracted self-energy', and is *finite* and independent of Λ as $\Lambda \to \infty$. We will come back to (111) in section 4.3.

4.2 The fermion self-energy and the vertex correction

Let's now briefly examine the other two one-loop divergent graphs, figures 6(b) and 6(c), beginning with the latter, the fermion self-energy. In analogy with $\Pi_\gamma^{[2]}$, we call the amplitude for figure 6(c) $-i\Sigma^{[2]}(p)$ where

$$-i\Sigma^{[2]}(p) = (-ie)^2 \int \gamma^\nu \frac{-ig_{\mu\nu}}{k^2} \frac{i}{\not p - \not k - m}\gamma^\mu \frac{d^4k}{(2\pi)^4}. \tag{113}$$

As in the γ case, when the string of self-energy insertions is summed up, the result is a modified fermion propagator equal to

$$\frac{i}{\not p - m - \Sigma^{[2]}(p)}. \tag{114}$$

As expected, $\Sigma^{[2]}$ as given by (113) diverges: there are four powers of k in the numerator and three in the denominator, so we might expect a divergent term proportional to Λ (note that $\Sigma^{[2]}$ has dimensions of mass, as is also evident from (114)). Actually the leading p-independent divergence is, instead, proportional to $m\ln(\Lambda/m)$. The reason for this is important, and it has interesting generalisations. Suppose that m in the Dirac Lagrangian $\bar\psi(i\not\partial - m)\psi$ were were set equal to zero. Then (see problem P3.1) the two 'left' and 'right' helicity components $\psi_L = \left(\frac{1-\gamma_5}{2}\right)\psi$ and $\psi_R = \left(\frac{1+\gamma_5}{2}\right)\psi$ of the electron field will not be coupled by the QED interaction. It follows that no terms of the form $\bar\psi_L\psi_R$ or $\bar\psi_R\psi_L$ can be generated - and these are just of the 'Dirac mass' type (problem P4.2). Hence no perturbatively-induced fermion mass term can be generated by higher-order e-m interactions, and the $\Sigma^{[2]}$ correction must vanish as $m \to 0$. So it must behave as $\sim m\ln(\Lambda/m)$ on dimensional grounds, which gives a logarithmically divergent correction to m in (114), call it $\delta m^{[2]}(\Lambda)$.

We can agree to call the resulting 'on shell point $\not p = m + \delta m^{[2]}(\Lambda)$' the *physical* mass $m_{\rm ph}$, such that

$$m_{\rm ph} = m(\Lambda) + \delta m^{[2]}(\Lambda) \tag{115}$$

is independent of Λ as $\Lambda \to \infty$ - which of course means that the *original* parameter m has in fact to be Λ-dependent, and in just such a way as to compensate for that of $\delta m^{[2]}$.

There is also a p-dependent logarithmic divergence of the form $\not p \ln \Lambda/m$. This can be soaked up in a fermion wavefunction renormalisation constant Z_2, analogous to Z_3, and having the same interpretation:

$$\psi_{\rm ph} = \frac{1}{(Z_2)^{\frac{1}{2}}}\psi. \tag{116}$$

In this way the physical fermion propagator is indeed

$$i/(\not{p} - m_{\mathrm{ph}}). \tag{117}$$

Finally there is the vertex part shown in figure 6(b). In this case, power counting indicates a new logarithmic divergence. We have one more card to play, in order to sweep it up. Consider the QED interaction term

$$-e\bar{\psi}(x)\,\not{A}(x)\psi(x) = -e\bar{\psi}_{\mathrm{ph}}\,\not{A}_{\mathrm{ph}}\psi_{\mathrm{ph}} \cdot Z_2 Z_3^{\frac{1}{2}}. \tag{118}$$

This generates a 'lowest order' vertex (in terms of the physical renormalised fields) equal to $-ie\gamma^{\mu} Z_2 Z_3^{\frac{1}{2}}$ to which figure 6(b) must be added. Now the physical charge e_{ph} is going to be determined experimentally from the Coulomb scattering contribution as $q^2 \to 0$ (the classical limit). Figure 6(b) contributes a logarithmically divergent correction to the charge in this limit, call it $\delta e(\Lambda)$. So, once again, we are going to assume that the 'original' e had a Λ-dependence just right to cancel out the Λ-dependence of the total contribution, leaving a finite Λ-independent physical charge as $\Lambda \to \infty$. We express this formally by introducing the vertex renormalisation constant Z_1 such that the physical charge is defined by

$$e_{\mathrm{ph}} = Z_2 Z_3^{\frac{1}{2}}(e/Z_1) \tag{119}$$

The interaction (118) then becomes

$$-Z_1 e_{\mathrm{ph}}\bar{\psi}_{\mathrm{ph}}\,\not{A}_{\mathrm{ph}}\psi_{\mathrm{ph}}. \tag{120}$$

Now some alarm bells should be ringing! The free Dirac part of the QED Lagrangian is now

$$\bar{\psi}(i\,\not{\partial} - m)\psi = Z_2\bar{\psi}_{\mathrm{ph}}(i\,\not{\partial} - m)\psi_{\mathrm{ph}} \tag{121}$$

to which we must add (120) (as well as the Maxwell term). But then the result is *not* gauge invariant! - since $\not{\partial}$ doesn't appear in the gauge invariant combination '$\not{\partial} + ie\,\not{A}$' (see section 3). For this to work we need a kind of small miracle - the equality

$$Z_1 = Z_2 \tag{122}$$

between two quite different wavefunction renormalisation constants. Of course, (122) *is* true; it is a *Ward identity*, and can be proved to follow from the gauge invariance of the original QED Lagrangian.

Relation (122) has a remarkable consequence: the 'rescaling' relation (119) now becomes

$$e_{\mathrm{ph}} = \sqrt{Z_3}e \tag{123}$$

showing that the corrections to 'e' associated with the fermion propagator and the vertex cancel out, leaving only the γ-propagator correction. Now this correction is the same whatever the external particles are, in a Feynman graph. So (123) is a statement of 'universality' of radiative corrections: they do not spoil the gauge invariance of the original Lagrangian, and the ration of e to e_{ph} is independent of the types of external particles. If a set of unrenormalised charges are all equal (or 'universal'), the renormalised ones will be too. Universality survives renormalisation - and this is a very big clue as to why the weak interactions have to be described by a gauge theory too, since quarks and leptons do seem to couple in some 'universal' way to W's and Z's: the strong interactions, experienced only by the quarks, do not seem to spoil that, just as - in the e-m case - the charge on a proton is the same as that on a positron.

4.3 The physics of $\bar{\Pi}_{\gamma}^{[2]}(q^2)$

We will only be able to offer brief notes:

(i) How does the renormalised γ-propagator affect physical processes? Let's imagine using it in $e^-\mu^- \to$

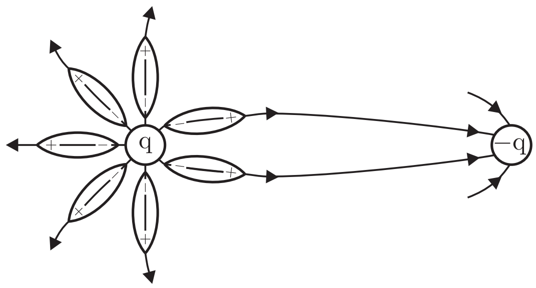

Fig. 8: Screeing of a charge in a dipolar medium.

$e^-\mu^-$ scattering via figure 4 with the corrected propagator (111), for instance. Then, the amplitude will be (omitting the spinor factors)

$$(-ie)^2 \frac{-ig_{\mu\nu}}{q^2\left(1 - \bar{\Pi}_\gamma^{[2]}(q^2)\right)} \tag{124}$$

where now we have changed the notation so that 'e' means the *physical* charge (which we previously called e_{ph}), and m is the physical mass (previously m_{ph}). In the static limit $q_0 = 0$, the photon propagator $\sim 1/\boldsymbol{q}^2$ has a simple interpretation - it is the Fourier transform of the $1/r$ Coulomb potential (see 'Point 4' at the end of section 2). So the form (124) must, in the static limit, represent *corrections* to Coulomb's law. Indeed, with $q_0 = 0$ and evaluating $\bar{\Pi}_\gamma^{[2]}$ for $\boldsymbol{q}^2 \ll m^2$, one finds that (124) becomes, approximately,

$$(-ie^2)\frac{ig_{\mu\nu}}{-\boldsymbol{q}^2}\left(1 + \frac{\alpha}{15\pi}\,\boldsymbol{q}^2/m^2\right) \tag{125}$$

$$\sim \frac{e^2}{\boldsymbol{q}^2} + \text{constant.} \tag{126}$$

The $e^2/\boldsymbol{q}^2$ in (126) gives us back the Coulomb $1/r$ in x-space: the Fourier transform of the 'constant' is a δ function. This very short distance correction, affecting only s-states in atomic physics, is responsible for a small (but entirely detectable) contribution to the famous *Lamb shift* between hydrogenic $2^2S_{\frac{1}{2}}$ and $2^2P_{\frac{1}{2}}$ levels. See problem P3.3.

(ii) Without making the low-$\boldsymbol{q}^2$ approximation, the form $\sim \frac{e^2}{\boldsymbol{q}^2}\left(1 + \bar{\Pi}_\gamma^{[2]}(q^2)\right)$ indicates that the charged leptons have effectively developed a 'form factor' (or spatial extension, when Fourier transformed) due to radiative corrections. Sharing it equally between the two e's in 'e^2', we can say that the radiatively induced charge form factor is $\mathcal{F}_1(q^2) \approx 1 + \frac{1}{2}\bar{\Pi}_\gamma^{[2]}(q^2)$. Examination of the Fourier transform of this shows that the spatial extension is of order $\sim m^{-1}$, the fermion Compton wavelength.

(iii) An alternative interpretation is in terms of a 'q^2-dependent charge', or 'q^2-dependent α', given by

$$\alpha(q^2) = \alpha\left[1 + \bar{\Pi}_\gamma^{[2]}(q^2)\right]. \tag{127}$$

The idea that a charge is q^2-dependent may seem strange at first, but it is analogous to the way in which a charge placed in a polarisable medium can give rise to a *space*-dependent effective charge, due to screening (see figure 8). The screening length here is just m^{-1}, the distance over which the e^+e^- pairs can be 'fluctuated' out of the cacuum, and which measures the extension of the radiatively induced form factor. This is why the photon self-energy e^+e^- bubble is called a vacuum polarisation graph!

For $|q^2| \gg m^2$, (127) becomes

$$\alpha(q^2) \approx \alpha\left[1 + \frac{\alpha}{3\pi}\ln\left(\frac{-q^2}{m^2}\right)\right] \tag{128}$$

showing that $\alpha(q^2)$ increases at large $-q^2$ (which is short distances, when Fourier transformed), just as indicated in figure 8.
(iv) However, a better approximation at large $-q^2$ is to return to the form (124) and write

$$\alpha(Q^2) = \frac{\alpha}{[1 - (\alpha/3\pi)\ln(Q^2/m^2)]} \quad \text{for} \quad Q^2 \gg m^2 \tag{129}$$

where $Q^2 = -q^2$. Equation (129) is the standard 'leading log' expression for the *running coupling constant* in QED. This shows a slow logarithmic increase as Q^2 increases. For example, $\alpha(M_Z^2) \sim 1/128.8$, as compared with $\alpha(= \alpha(0)) \sim 1/137$. In QCD, the effect of gluon self-interactions is to make α_s (the QCD analogue of α) *decrease* as Q^2 increases ('asymptotic freedom'). There, the analogous formula is

$$\alpha_{\rm s}(Q^2) = \frac{\alpha_{\rm s}}{[1 + \frac{\alpha_{\rm s}}{12\pi}(33 - 2f)\ln(Q^2\mu^2)]} \tag{130}$$

where f is the number of fermion-antifermion pairs (in the loops) considered, and μ is a 'renormalisation scale'. If $f < 16$, α_s will decrease as Q^2 increases, leaving the quarks weakly interacting at very short distances.

4.4 Renormalisability

We have tried to give some idea of how we can make sense of a theory with divergences. At the one-loop level, some of the steps seemed quite trivial. More generally, however, we can ask: how do we know that we can *go on* soaking up these divergences into redefinitions of 'physical' quantities, as we proceed on to higher order loops? The answer is really rather remarkable: there *are* classes of theory ('renormalisable theories') which are such that *all* divergences, encountered at each successive order in systematic perturbation theory, can be tamed by this procedure of redefining finite physical quantities (and doing wavefunction rescalings), and then re-expressing all amplitudes in terms of these physical quantities. Furthermore, there is a surprisingly simple criterion for telling (almost) which theory is renormalisable and which isn't. This criterion has to do with the *dimensionality of the coupling constant* (in units $\hbar = c = 1$) - see problem P3.4.

The result is simply stated: if the dimensionality of the coupling is M^a where $a > 0$, then the theory is 'super-renormalisable' (like the ABC theory - there are fewer divergences than we could in fact deal with, for instance Z_C and the vertex correction are finite); if $a = 0$ (dimensionless) then the theory *may* be renormalisable, and often is (e.g. QED, where the coupling is α); and if $a < 0$, the theory is not renormalisable.

Consider a hypothetical theory, similar to the original four-fermion theory of β-decay, describing interactions between the ν_e and a neutron (assumed pointlike for this purpose). The interaction density is

$$G_F \bar{\psi}_n(x)\psi_n(x)\bar{\psi}_{\nu_e}(x)\psi_{\nu_e}(x). \tag{131}$$

To find the dimensionality of G_F, we need to remember that the mass term in the Dirac Hamiltonian is $m\bar{\psi}\psi$, so that the dimension of a ψ field is $M^{\frac{3}{2}}$. This implies that the dimension of G_F is M^{-2} so that this theory is non-renormalisable. Is this in fact so bad? Consider what happens when we calculate $n + \nu_e \to n + \nu_e$ in perturbation theory. The lowest order ('tree') graph is figure 9(a); next is figure 9(b); and then at third order figure 9(c). Let's count powers in the loop of figure 9(b). Since each fermion propagator $\sim k^{-1}$, we expect the graph to diverge as Λ^2. Fine ... what about figure 9(c)? Here we have two loops, with therefore 8 momentum integrals, and four fermion propagators each contributing only one power of k in the denominator, so it diverges as Λ^4! The first point to note, then, is clearly that as we go up in order of perturbation theory, the divergence gets *worse*. To control the Λ^4 divergence, we would have to 'subtract' the amplitude for figure 9(c) *three* times. Each subtraction means that we have to take one parameter from experiment (the amplitude at a certain point, its derivative at that point, its second derivative, etc). Very soon we need more parameters than are appearing in the original

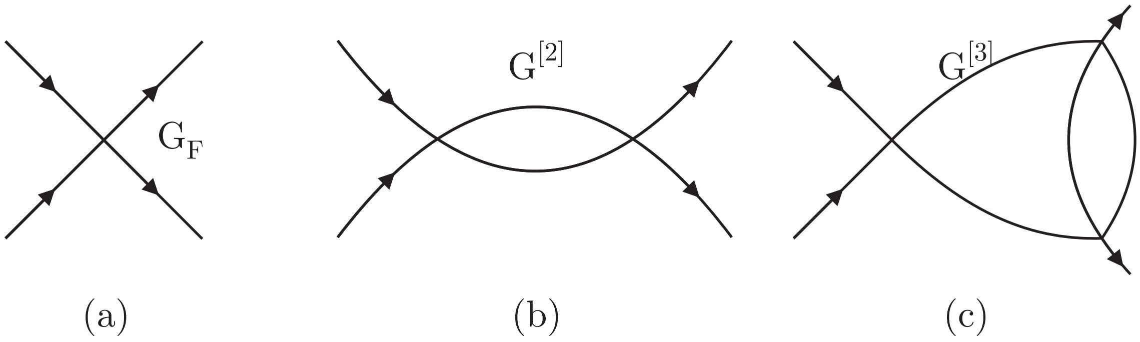

Fig. 9: Contributions to $n + \nu_e \to n + \nu_e$ in perturbation theory, using (131).

Lagrangian (masses, couplings). So simply defining a 'physical' set of Lagrangian parameters won't get us off the hook in this case. A renormalisable theory is one whose infinities can all be tamed by redefinitions of the parameters in the original Lagrangian (plus wavefunction rescalings); if infinities arise which need new parameters (not in the original Lagrangian) to be taken from experiment, then the theory is non-renormalisable.

The reason for this worsening divergence in higher orders in G_F is, of course, related to the dimensionality of G_F. All the amplitudes of figure 9 have to have the same dimension, obviously. But since each G_F brings in two powers of a mass 'M' in the denominator, these must be compensated by two powers of momentum in the numerator, making the divergence successively worse.

Is the situation really hopeless? Actually no. We know quite well that people lived with the Fermi theory reasonably happily for years, *until* the advent of high energy experiments probing weak interactions. The reason can again be found in dimensional analysis. Consider the amplitude for figure 9(b), call it $G^{[2]}(s)$, where $s = (p_1 + p_2)^2$. This needs two subtractions to tame it into a finite quantity $\bar{G}_F(s) = G^{[2]} - G^{[2]}(s_0) - (s - s_0)\frac{dG^{[2]}}{ds}\Big|_{s=s_0}$, where s_0 is the point we choose to define our amplitudes at. This means that, expanding $\bar{G}^{[2]}(s)$ about $s = s_0$, we can *calculate* terms of order $(s - s_0)^2$ and higher (the two lowest terms in the expansion have to be taken from experiment). But the worse divergence of figure 9(c) (amplitude $G^{[3]}$) would require us to do *three* subtractions before arriving at a finite part we could calculate: in this case, the first calculable bit would be $\sim G_F^3(s - s_0)^3 \frac{d^3G^{[3]}}{ds^3}\Big|_{s=s_0}$ - and the process has to be repeated each time we go up an order. Assuming that all the derivatives are about the same order of magnitude, we see that we can get away with using only low order corrections provided $G_F s \ll 1$, i.e.

$$\sqrt{s} \ll \frac{1}{\sqrt{G_F}}. \tag{132}$$

This is an important idea - and in the case of the real Fermi constant ($G_F \sim 1.17 \times 10^{-5}\mathrm{GeV}^{-2}$), $\frac{1}{\sqrt{G_F}} \sim 300\mathrm{GeV}$. So a non-renormalisable theory can be useful at energies well below its 'natural energy scale', as set by the inverse coupling constant; but the nearer we approach this scale, the less predictive the theory will become. And we are, after all, always striving to *reduce* the number of parameters in our theories that have to be taken from experiment.

From this perspective, it may be less of a mystery why renormalisable theories are generally the *relevant* ones at present energies. We may imagine that a 'true' theory exists at some enormously high scale Λ (the Planck scale?) which, though not itself a local quantum field theory, can be written out in terms of all possible fields and their couplings, as allowed by the operative symmetry principles. Our particular renormalisable *subset* of these theories then emerges as a low energy effective theory, due to the strong suppression of the non-renormalisable terms (which are damped like (s/Λ^2) to some power).

Nonrenormalisable theories may be physically detectable at low energies if they involve processes

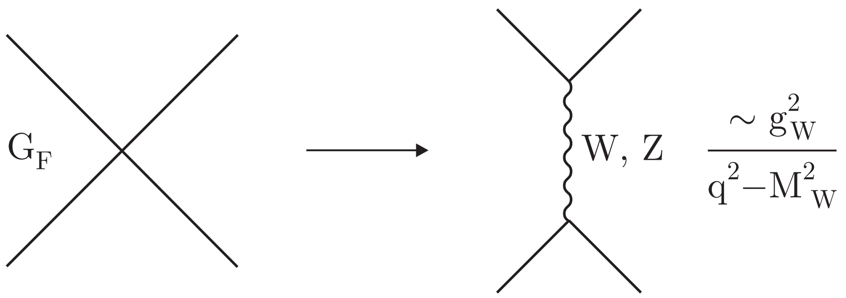

Fig. 10: Relation between four-fermi coupling and Yukawa-like coupling.

that would be otherwise forbidden. For example, the fact that (as far as we know) neutrinos have neither electromagnetic nor strong interactions, but only weak ones, allowed the four-fermi interaction to be detected - but amplitudes were suppressed by powers of s/M_W^2 relative to e-m ones, and this is precisely why it was 'weak'! As we'll discuss later, the four-fermi model is superseded in the Standard Model by a Yukawa-type theory involving exchanges of $W^{\pm}, Z^0$ (see figure 10). For $q^2 \ll M_W^2, G_F \sim g_W^2/M_W^2$, explaining the origin of the M^{-2} dimensionality of G_F, *and* telling us the actual scale, in this case. Thus this theory changes from being an effective non-renormalisable four-fermion theory at very low energies, to being an effective renormalisable one at $q^2 \sim M_W^2$.

Problems for Lecture 3

P3.1 For a Dirac field $\psi(x)$, define $\psi_R = \left(\frac{1+\gamma_5}{2}\right)\psi$, $\psi_L = \left(\frac{1-\gamma_5}{2}\right)\psi$. Show that

$$\bar{\psi}_L \gamma^\mu \psi_R = 0,$$

where $\bar{\psi}_L = \psi_L^\dagger \gamma_0$.

P3.2 Rewrite $m\bar{\psi}\psi$ in terms of the ψ_R and ψ_L fields, and deduce that e-m interactions cannot generate such a 'Dirac' mass in perturbation theory.

P3.3 Coulomb's law is corrected by the vacuum polarisation (e^+e^-) to

$$-\left\{\frac{\alpha}{r} + \frac{4\alpha^2}{15m^2}\delta^3(\boldsymbol{r})\right\}$$

where m is the electron mass. Treating the δ function piece as a perturbation on the Coulomb term, calculate the shift in energy (to first order) of an $l = 0$ hydrogenic state with principal quantum number n, given that the Coulomb wave function at $r = 0$ is

$$\phi_n(0) = \frac{1}{\sqrt{\pi}}\left(\frac{\alpha m}{n}\right)^{\frac{3}{2}}.$$

Give the answer in eV for the $n = 2$ shift.

P3.4 What is the (mass) dimension of a scalar field ϕ in four space-time dimensions? What is the dimension of the coupling constant λ in a '$\lambda\phi^3$' interaction? And of g in a '$g\phi^4$' interaction? What is the dimension of G in a '$G(\bar{\psi}\psi)^3$' interaction?

P3.5 Consider a $\lambda\phi^4$ theory. Given that it is renormalisable, explain why any graph contributing to the process $\phi + \phi \to \phi + \phi + \phi + \phi$ must be finite.

5. GLOBAL AND LOCAL NON-ABELIAN SYMMETRIES

For a much fuller treatment of the material in this section see chapters 12 and 13 of volume 2 of the new (third) edition of Aitchison and Hey [2].

Having introduced QED as an example of a *gauge theory* with a *local phase invariance*, we now consider the generalisations of QED which describe the weak and strong interactions between quarks and leptons. These involve a more complicated kind of local phase symmetry, in which the phase factors are (x-dependent) *matrices*, which in general don't commute - that's what 'non-abelian' means in this context. We shall limit the treatment to the particular ingredients needed for the Standard Model. Note: from now on we shall **omit the hats** on quantum field operators!

5.1 Global non-Abelian symmetry

Consider the Lagrangian for two free fermions of the same mass $m_1 = m_2 = m$

$$\mathcal{L}_2 = \bar{\psi}_1(i\partial\!\!\!/ - m)\psi_1 + \bar{\psi}_2(i\partial\!\!\!/ - m)\psi_2\,; \tag{133}$$

in terms of the 'doublet' field

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \tag{134}$$

it can easily be rewritten as

$$\mathcal{L}_2 = \bar{\psi}(i\partial\!\!\!/ - m)\psi\,. \tag{135}$$

Note that although (135) looks formally like the single-field $\mathcal{L}_{\mathrm{D}}$ of (78), it is of course quite different physically, representing two different sorts of particle (e.g. up and down quarks, and their antiparticles). Nevertheless, (135) is invariant under a symmetry rather like (79), namely the 2×2 unitary transformation

$$\psi \to \psi' = U\psi, \quad UU^\dagger = U^\dagger U = 1\,. \tag{136}$$

The U in (136) is a 2×2 matrix of numbers (not field operators) acting on the 2 components of ψ in (134), and they commute with the Dirac γ's. Such unitary 2×2 U's form a group, U(2). Since U in (136) does not involve x, we call (136) a global symmetry. In general, two U's do not commute with each other, and it is called a non-Abelian symmetry.

From elementary properties of determinants we have

$$\mathrm{det}UU^\dagger = \mathrm{det}U.\mathrm{det}U^\dagger = \mathrm{det}U.\mathrm{det}U^* = |\mathrm{det}U|^2 = 1 \tag{137}$$

so that $\mathrm{det}U = e^{-2i\alpha}$, say. We can therefore write

$$U = e^{-i\alpha}\tilde{U} \tag{138}$$

where $\tilde{U}$ has determinant $+1$. Matrices of the form $\tilde{U}$ form the SU(2) group, where the S just means they have unit determinant. The phase factor in (138) corresponds to a simultaneous U(1) transformation of ψ_1 and ψ_2 (with the same phase angle) and leads, as in Section 3.3, to a conservation law of the total number of '1' particles and '2' particles. (For quarks this would be part of baryon number conservation). The new physics is contained in the $\tilde{U}$ part.

Groups such as SU(2) (and, later, SU(3)) have the important feature that their physically important properties can be found by studying infinitesimal transformations, of the form (cf (46))

$$\tilde{U} = 1 - i\xi \tag{139}$$

where ξ is a 2×2 matrix with small entries. The condition $\mathrm{det}U = 1$ gives Tr$\xi = 0$ (neglecting terms of order ξ^2 - see problem P4.1), while $\tilde{U}\tilde{U}^\dagger = 1$ reduces (problem P4.1) to $\xi = \xi^\dagger$. So ξ is a Hermitian traceless matrix. Such a thing depends on only three real parameters (problem P4.1) and can be written as

$$\xi = \boldsymbol{\epsilon}\cdot\boldsymbol{\tau}/2 \tag{140}$$

where $\boldsymbol{\epsilon} = (\epsilon_1, \epsilon_2, \epsilon_3)$ are the three parameters, and $\boldsymbol{\tau} = (\tau_1, \tau_2, \tau_3)$ are the Pauli matrices (problem P4.2). Thus an infinitesimal SU(2) transformation on the doublet ψ is

$$\psi \to \psi' = (1 - i\boldsymbol{\epsilon}\cdot\boldsymbol{\tau}/2)\psi\,. \tag{141}$$

This should be compared with the infinitesimal version of (69), namely $\psi \to \psi' = (1 - i\epsilon)\psi$, from which it is clear that the 'ϵ' in that case becomes a *matrix* in (141). The form for a finite SU(2) transformation is

$$\psi \to \psi' = e^{-i\boldsymbol{\alpha}\cdot\boldsymbol{\tau}/2}\psi \tag{142}$$

which generalises (69) (note that for a matrix A, exp $A = 1 + A + A^2/2! + \cdots$).

Since (141) or (142) are invariances of $\mathcal{L}_2$ we expect an associated conservation law. Indeed, since we have *three* independent transformations (using each of ϵ_i in turn) we expect *three* conservation laws. Following the same steps used in deriving the Noether current for the complex scalar field in §3.2, but this time for the doublet Dirac field ψ, one finds that the three quantities $T_1^\mu(x), T_2^\mu(x), T_3^\mu(x)$ defined by (cf (70))

$$T_i^\mu(x) = \bar{\psi}(x)(\tau_i/2)\gamma^\mu\psi(x) \tag{143}$$

satisfy

$$\partial_\mu T_i^\mu(x) = 0 \tag{144}$$

and are therefore symmetry currents. The corresponding 'charges'

$$T_i = \int \psi^\dagger(x)\frac{\tau_i}{2}\psi(x)\,d^3x \tag{145}$$

are conserved. These are the (field theoretic) 'isospin' operators, which have the very interesting property

$$[T_i, T_j] = i\epsilon_{ijk}T_k \tag{146}$$

as can be explicitly checked from (145) (using the proper commutation relations for the ψ fields). A simple example is provided in problem P4.2. The relations (146) are of course exactly the commutation relations of the familiar angular momentum operators, which is why the name iso*spin* was coined; (146) is called the 'SU(2) algebra'. Not coincidentally, the τ's satisfy $[\tau_i/2, \tau_j/2] = i\epsilon_{ijk}\tau_k/2$, the same algebra.

In thinking about more complicated SU(2) multiplets than doublets (which we shan't need to do much) this angular momentum analogy is very helpful. The essential step is to find larger matrices than the 2×2 $\frac{\tau\text{'s}}{2}$, which satisfy commutation relations of the form (146). For example, the three 3×3 matrices t_1, t_2 and t_3, defined by

$$(t_i)_{jk} = -i\epsilon_{ijk} \tag{147}$$

satisfy $[t_i, t_j] = i\epsilon_{ijk}t_k$ (see problem P4.3). Then if we consider a triplet of three real degenerate fields (bosonic, say)

$$\boldsymbol{\phi} = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{pmatrix} \tag{148}$$

with Lagrangian

$$\mathcal{L}_B = \frac{1}{2}\partial_\mu\boldsymbol{\phi}\cdot\partial^\mu\boldsymbol{\phi} - \frac{1}{2}m^2\boldsymbol{\phi}\cdot\boldsymbol{\phi}\,, \tag{149}$$

$\mathcal{L}_B$ is invariant under

$$\boldsymbol{\phi} \to \boldsymbol{\phi}' = (1 - i\boldsymbol{\epsilon}\cdot\boldsymbol{t})\boldsymbol{\phi}\,. \tag{150}$$

Using (147), (150) is equivalent to (problem P4.4)

$$\boldsymbol{\phi}' = \boldsymbol{\phi} + \boldsymbol{\epsilon} \times \boldsymbol{\phi} \tag{151}$$

which should be familiar as the 'infinitesimal rotation' of an ordinary vector.

The SU(2) transformation of (142) can be generalised to the case of *three* degenerate fermion fields. If $\mathcal{L}_3$ is (133) with the addition of $\bar{\psi}_3(i\not\partial - m)\psi_3$, it too can can be written as in (135) where now

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \end{pmatrix} . \tag{152}$$

Note particularly that unlike the ϕ's in (148), the ψ's in (152) are complex: each ψ_i contains c_i and $d_i^\dagger$ operators as in (59). $\mathcal{L}_3$ is invariant under $\psi \to \psi' = U\psi$ where U is now an x-independent 3×3 unitary matrix. Extracting the overall phase again, we are left with a global SU(3) transformation. An infinitesimal SU(3) matrix has the form

$$\tilde{U} = 1 - i\chi \tag{153}$$

where χ is a Hermitean traceless 3×3 matrix. Such a χ involves *eight* parameters and can be written as

$$\chi = \boldsymbol{\eta}\cdot\boldsymbol{\lambda}/2 \tag{154}$$

where $\boldsymbol{\eta} = (\eta_1, \ldots \eta_8)$ are the arbitrary parameters and the eight $\boldsymbol{\lambda}$'s are 3×3 Hermitean traceless matrices generalising the three τ's. They obey the commutation relations

$$\left[\frac{\lambda_a}{2}, \frac{\lambda_b}{2}\right] = i f_{abc}\frac{\lambda_c}{2} \tag{155}$$

where the f_{abc} are numbers characteristic of SU(3) (a, b, c all run from 1 to 8). If ψ_1, ψ_2, ψ_3 are taken to be the u, d, s quarks, this global SU(3) symmetry would be the SU(3) of strong interaction flavour symmetry (which however is not exact as m_u, m_d and m_s are not equal). Similarly, if we take 1, 2, 3 to be colour indices we have the exact $\mathrm{SU(3)_c}$ colour symmetry of QCD, which we shall shortly see is a local symmetry. The currents corresponding to the SU(3) symmetry of $\mathcal{L}_3$ are (cf (143))

$$G_a^\mu(x) = \bar{\psi}(x)(\lambda_a/2)\gamma^\mu\psi(x) \tag{156}$$

and the associated eight 'charges'

$$G_a = \int \psi^\dagger(x)(\lambda_a/2)\psi(x)\, d^3x \tag{157}$$

generalise the three isospin operators, and obey the commutation relations

$$[G_a, G_b] = i f_{abc} G_c\,. \tag{158}$$

which is called the 'SU(3) algebra'. Note the similarity between (146) and (158).

As in the case of SU(2), larger multiplets are possible too. The key requirement is to find matrices which satisfy (158), since these commutation relations effectively define the group. For SU(3), the only larger multiplet in which we shall be interested is the octet, **8**, which is analogous to the triplet of SU(2). The matrices for the **8** are defined analogously to the t's of (147), namely $(F_a)_{bc} = -if_{abc}$ where the f's are as in (158). Notice that since there are eight 'charges' G_a, and all the indices a, b, c in (158) run from 1 to 8, the eight matrices F_a are each 8×8. In the same way, the three matrices t_i of (147) are each 3×3, since there are three SU(2) charges. This kind of pattern can be extended to arbitrary SU(N); the 'representation' in which the matrices are equal (with a factor of $-i$) to the 'structure constants' (the ϵ's and f's in (147) and (158)) is generally called the adjoint or regular representation.

5.2 Local non-Abelian SU(2) symmetry

Global symmetries and their associated (possibly approximate) conservation laws are certainly interesting, but they do not have the *dynamical* significance of local symmetries. We saw in section 3.4 how the 'requirement' of local U(1) symmetry seemed to lead almost automatically to QED, with the symmetry current of the ψ matter fields now playing the role of the dynamical current which, when dotted into the A-field, gives the interaction term in $\mathcal{L}_{QED}$. A similar link between symmetry and dynamics follows if we generalise the preceding non-Abelian global symmetries to local ones. In this section we carry through the analysis for SU(2).

We begin by considering again a fermion doublet as in (135), without yet specifying exactly what the physical application will be. We want to extend the global SU(2) symmetry transformation (142) to the local one

$$\psi(x) \rightarrow \psi'(x) = e^{-ig\boldsymbol{\alpha}(x)\cdot\boldsymbol{\tau}/2}\psi(x) \tag{159}$$

by analogy with (79); note that we have slipped in a constant g in the exponent - it will be analogous to the e-m charge e. Clearly, although the $\bar{\psi}m\psi$ part of (135) is still invariant under (159), the $\bar{\psi}i\partial\!\!\!/\psi$ part is not - just as in the U(1) case (80), since the $\partial\!\!\!/$ will pull down a $\partial\!\!\!/\boldsymbol{\alpha}(x)$ factor. As in the U(1) case, we try to compensate this factor by introducing some vector field whose change under an appropriate transformation (accompanying (159)), exactly cancels this $\partial\!\!\!/\boldsymbol{\alpha}(x)$ part. This time, since there are three $\boldsymbol{\alpha}(x)$'s ($\alpha_1(x)$, $\alpha_2(x)$, $\alpha_3(x)$) we immediately see that we need three vector (gauge) fields, called $W_1^\mu(x)$, $W_2^\mu(x)$, $W_3^\mu(x)$, or $\boldsymbol{W}^\mu(x)$ for short.

The key step in constructing the locally U(1) invariant Lagrangian of QED was the replacement of '∂^μ' by '$D^\mu = \partial^\mu + ieA^\mu$' (cf (81)), together with the transformation '$A^\mu \rightarrow A^\mu + \frac{1}{e}\partial^\mu\alpha(x)$' (cf (83)) for the A-field. Let's have another look at the combination $D^\mu\psi$ in the QED Lagrangian (88). Under the gauge transformation (84) ,

$$\begin{aligned} D^\mu &= (\partial^\mu + ieA^\mu)\psi \rightarrow (\partial^\mu + ieA'^\mu)\psi' \\ &= (\partial^\mu + ieA^\mu + i(\partial^\mu\alpha(x)))e^{-i\alpha(x)}\psi \\ &= [-i(\partial^\mu\alpha(x))e^{-i\alpha(x)}\psi] + e^{-i\alpha(x)}\partial^\mu\psi + ieA^\mu e^{-i\alpha(x)}\psi + [i(\partial^\mu)e^{-i\alpha(x)}\psi] \\ &= e^{-i\alpha(x)}D^\mu\psi \end{aligned} \tag{160}$$

since the bracketed terms cancel. So we have

$$D'^\mu\psi' = e^{-i\alpha(x)}D^\mu\psi. \tag{161}$$

In words, this says that the quantity '$D^\mu\psi$' transforms under a *local* U(1) phase transformation just like ψ would under a global one (i.e. it just gets multiplied by a phase factor). So to construct a locally U(1) invariant Lagrangian all we needed to do was multiply $D^\mu\psi$ by $\bar{\psi}$ from the left, since then under the local transformation

$$\bar{\psi}D^\mu\psi \rightarrow \bar{\psi}'D'^\mu\psi' = \bar{\psi}e^{i\alpha(x)}\, e^{-i\alpha(x)}D^\mu\psi = \bar{\psi}D^\mu\psi, \tag{162}$$

showing that $\bar{\psi}D^\mu\psi$ is indeed locally U(1) invariant. Of course, we also need the γ_μ to get rid of the loose Lorentz index μ, and make $\mathcal{L}$ a Lorentz invariant.

So the key to constructing a locally SU(2) phase-invariant theory is to generalise '$D^\mu\psi$'. The required generalisation is

$$D^\mu\psi = (\partial^\mu + ig\boldsymbol{\tau}\cdot\boldsymbol{W}^\mu/2)\psi \tag{163}$$

when acting on an SU(2) doublet field such as ψ. The property required of (163) is that $D^\mu\psi$ should transform under the local symmetry (159) exactly as $\partial^\mu\psi$ does under the global one (142), as we have seen happening in the U(1) case. Then, a term like $\bar{\psi}D\!\!\!\!/^\mu\psi$ is automatically invariant under local SU(2).

This requirement on $D^\mu\psi$ determines the transformation law of the fields $\boldsymbol{W}^\mu$. The algebra is easier if we consider an infinitesimal transformation

$$\delta\psi = (-ig\boldsymbol{\epsilon}(x)\cdot\boldsymbol{\tau}/2)\psi(x)\,; \tag{164}$$

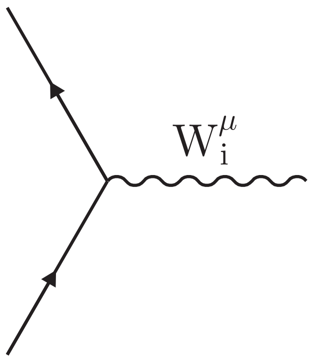

Fig. 11: ψ-ψ-W vertex.

we then require

$$\delta(D^\mu\psi) = (-ig\boldsymbol{\epsilon}(x)\cdot\boldsymbol{\tau}/2)D^\mu\psi\,. \tag{165}$$

It is a good exercise (problem P4.5) to verify that (165) implies that

$$\delta\boldsymbol{W}^\mu(x) = \partial^\mu\boldsymbol{\epsilon}(x) + g\boldsymbol{\epsilon}(x)\times\boldsymbol{W}^\mu(x)\,, \tag{166}$$

which tells us how the $\boldsymbol{W}^\mu$'s must transform. The first term in (166) is the straightforward analogue of the infinitesimal version of (84), with $\alpha(x)\to e\epsilon(x)$. Comparing the second term of (166) with (151), we see that it implies that the three W-fields form the components of an SU(2) triplet. Thus *the W's carry SU(2) 'charge'*.

We now know the generalisation of (135) which makes it locally SU(2) invariant:

$$\mathcal{L}_{2W} = \bar\psi(i\not{D}-m)\psi = \bar\psi(i\not{\partial}-m)\psi - g\bar\psi\gamma_\mu\boldsymbol{\tau}/2\psi\cdot\boldsymbol{W}^\mu, \tag{167}$$

the last term being the generalisation of $\mathcal{L}_{\text{int}}$ in QED (equation (82)). We can immediately read off the ψ-ψ-W vertex factor as (figure 11)

$$-ig\frac{\tau_i}{2}\gamma^\mu\,. \tag{168}$$

In (168) the index 'i' refers to the SU(2) component of the W field quantum, and 'μ' to the Lorentz component of its polarisation vector. Each W-field will have the same kind of mode expansion as the A-field did (equation (85)).

We can easily generalise (163) to other SU(2) multiplets than doublets, by using appropriately larger matrices instead of the $\boldsymbol{\tau}/2$. For example, for an SU(2) triplet of fields $\boldsymbol{\phi}=(\phi_1,\phi_2,\phi_3)$, (163) becomes

$$D^\mu\phi_i = (\partial^\mu + ig\boldsymbol{t}\cdot\boldsymbol{W}^\mu)\phi_i \tag{169}$$

where the three 3×3 matrices $\boldsymbol{t}$ are defined in (147). Under infinitesimal transformations, this changes by

$$\begin{aligned}\delta(D^\mu\phi_i) &= (-ig\boldsymbol{\epsilon}(x)\cdot\boldsymbol{t})(D^\mu\phi_i) &(170)\\ &= (g\boldsymbol{\epsilon}(x)\times D^\mu\boldsymbol{\phi})_i &(171)\end{aligned}$$

(cf (150), (151), and (164)).

However, there is still an important part of the non-Abelian analogue of $\mathcal{L}_{\text{QED}}$ unaccounted for - namely the bit corresponding to the Maxwell-term $-\frac{1}{4}F.F$ for the gauge fields $\boldsymbol{W}^\mu$. Note that, as in the QED case (problem P2.6), a simple mass term involving $\boldsymbol{W}^\mu\cdot\boldsymbol{W}_\mu$ will violate invariance under (166), so these quanta are massless. Clearly we have a problem here in applying this local SU(2) - as we eventually will - to weak interactions, which are very short ranged, and whose quanta are therefore massive. This is where we will need the Higgs mechanism - see Section 6.

To get the non-Abelian '$F \cdot F$' term, the obvious thing might be to consider

$$\partial^\mu A^\nu - \partial^\nu A^\mu \rightarrow D^\mu \boldsymbol{W}^\nu - D^\nu \boldsymbol{W}^\mu \tag{172}$$

with D^μ given by (169), since the W's are an SU(2) triplet. The hope would be that by using the D's, $D^\mu \boldsymbol{W}^\nu - D^\nu \boldsymbol{W}^\mu$ would transform under local SU(2) transformations exactly as $\partial^\mu \boldsymbol{W}^\nu - \partial^\nu \boldsymbol{W}^\mu$ does under global ones - i.e. like (171). Then the 'dot product' $(D^\mu \boldsymbol{W}^\nu - D^\nu \boldsymbol{W}^\mu) \cdot (D_\mu \boldsymbol{W}_\nu - D_\nu \boldsymbol{W}_\mu)$ would be a locally invariant '$F \cdot F$' term. Unfortunately it is not quite that simple. The problem is that the W's are a rather special triplet: whereas an ordinary triplet $\boldsymbol{\phi}$ would transform via only the second term in (166), the W's *also* have the first ('non-homogeneous') term as well. You can verify that in fact

$$\delta(D^\mu \boldsymbol{W}^\nu - D^\nu \boldsymbol{W}^\mu) \neq g\boldsymbol{\epsilon}(x) \times (D^\mu \boldsymbol{W}^\nu - D^\nu \boldsymbol{W}^\mu) \tag{173}$$

so that the proposed '$F \cdot F$' term will not work.

With the aid of some hindsight, we can be led to the right answer as follows. Consider, in the U(1) case, the quantity

$$(D^\mu D^\nu - D^\nu D^\mu)\phi \tag{174}$$

where ϕ is any field of charge e and $D^\mu = \partial^\mu + ieA^\mu$. Evaluating (174) one finds (problem P2.5)

$$(D^\mu D^\nu - D^\nu D^\mu)\phi = ieF^{\mu\nu}\phi \tag{175}$$

where $F^{\mu\nu} = \partial^\mu A^\mu - \partial^\nu A^\mu$. This suggests that we should look at the commutator of two covariant derivatives $[D^\mu, D^\nu]$. It does not matter whether we use the D from (163) or (169) - the result is essentially the same for all cases. Using the D^μ from (163) one finds (problem P4.6)

$$[D^\mu, D^\nu] = ig\boldsymbol{\tau}/2 \cdot \boldsymbol{F}^{\mu\nu} \tag{176}$$

where

$$\boldsymbol{F}^{\mu\nu} = \partial^\mu \boldsymbol{W}^\nu - \partial^\nu \boldsymbol{W}^\mu - g\boldsymbol{W}^\mu \times \boldsymbol{W}^\nu . \tag{177}$$

(Had we used (169) we would have got (176) with $\boldsymbol{\tau}/2 \rightarrow \boldsymbol{t}$.) When we now investigate the effect of the local SU(2) transformation (166) on $\boldsymbol{F}^{\mu\nu}$ we find (problem P4.7)

$$\delta \boldsymbol{F}^{\mu\nu}(x) = g\boldsymbol{\epsilon}(x) \times \boldsymbol{F}^{\mu\nu}(x) \tag{178}$$

precisely as desired (but not accomplished) in (173) - i.e. the inhomogeneous part in (166) has been got rid of. Thus $\boldsymbol{F}^{\mu\nu}$ does transform under local SU(2) transformations exactly as if it were an ordinary triplet under global SU(2) transformations and so the quantity

$$\mathcal{L}_W = -\frac{1}{4}\boldsymbol{F}_{\mu\nu} \cdot \boldsymbol{F}^{\mu\nu} \tag{179}$$

is indeed locally SU(2) invariant. This is the famous *Yang-Mills Lagrangian*, the non-Abelian generalisation of the Maxwell Lagrangian. $\boldsymbol{F}^{\mu\nu}$ is the non-Abelian field strength tensor.

The argument leading to (179) has been given in some detail since the result is of fundamental importance. Looking at (177) and (179) it is clear that, unlike the Maxwell term $\mathcal{L}_A$ of (87), the Yang-Mills term $\mathcal{L}_W$ of (179) *includes interactions between the gauge fields* - in addition, of course, to the expected 'free' part

$$-\frac{1}{4}(\partial_\mu \boldsymbol{W}_\nu - \partial_\nu \boldsymbol{W}_\mu) \cdot (\partial^\mu \boldsymbol{W}^\nu - \partial^\nu \boldsymbol{W}^\mu) . \tag{180}$$

The free part leads to a W-propagator which is the same as that in rule (v) of section 3.4, with a δ_{ij} factor to 'dot' the W's together. The interactions included in (179) are of two types: W-W-W (trilinear) and

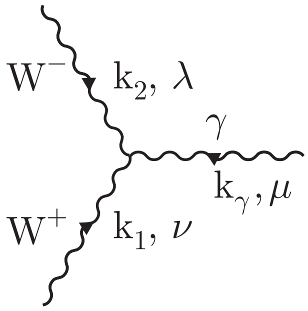

Fig. 12: W-W-γ vertex.

W-W-W-W (quadrilinear). This is quite unlike QED, where no fundamental γ-γ vertices are present. It arises here because the W's both 'transmit' the gauge field force and feel it themselves since they are not SU(2) neutral (as the γ was U(1) neutral). Another important point to note is that these self-interactions among the W's come in with a coupling constant which is the same one as appears in the ψ-ψ-W vertex (168)—the W's 'couple universally'.

The physics application of all this is to the SU(2) of the weak interactions (see section 7). There, the W_1^μ and W_2^μ fields correspond to the charged gauge bosons $W^{\pm\mu}$ (the combination $\frac{1}{\sqrt{2}}(W_1 - iW_2)$ destroys W^+ or creates W^-). As we shall see, the field W_3^μ is a linear combination of the photon γ and Z^0 fields:

$$W_3^\mu = \sin\theta_W A^\mu + \cos\theta_W Z^\mu \tag{181}$$

where θ_W is the 'weak angle', and the SU(2) gauge coupling constant g is related to e by

$$g\sin\theta_W = e\,. \tag{182}$$

We can then pick out the W-W-γ vertex from (179), and find that it is given by

$$ie\,[g_{\nu\lambda}(k_1 - k_2)_\mu + g_{\lambda\mu}(k_2 - k_\gamma)_\nu + g_{\mu\nu}(k_\gamma - k_1)_\lambda] \tag{183}$$

where the momenta and indices are as in figure 12. This *unique* e-m coupling of the $W^\pm$ is of precisely the kind needed to make a *renormalisable* (see section 4) theory of the 'electromagnetic interactions of charged vector bosons'.

5.3 Local SU(3) Symmetry: the QCD Lagrangian

Using what has been said about global SU(3) in section (5.1), and about how to make a global SU(2) symmetry into a local one in section 5.2, it is straightforward to discuss local SU(3). This is the gauge group of QCD (see the course on QCD), the labels 1, 2, 3 in (152) standing for colour, the ψ's being one flavour of quark. Under a local SU(3)$_c$ transformation, the triplet (152) transforms by

$$\delta\psi = (-ig_s\boldsymbol{\eta}(x)\cdot\boldsymbol{\lambda}/2)\psi \tag{184}$$

(cf (154) and (164)), where now there are eight field parameters $\eta_1(x), \eta_2(x)\ldots\eta_8(x)$ going with the eight λ's. To cancel off the unwanted $\partial\!\!\!/\boldsymbol{\eta}$ parts which occur when we try to make $\bar\psi\partial\!\!\!/\psi$ invariant under (184), we now need eight vector gauge fields $A_a^\mu(x)$, $a = 1, 2, \ldots 8$. These A's transform according to

$$\delta A_a^\mu(x) = \partial^\mu\eta_a(x) + g_s f_{abc}\eta_b(x)A_c^\mu(x) \tag{185}$$

(cf (166) and (155)). The SU(3)$_c$ covariant derivative acting on a triplet is

$$D^\mu\psi = (\partial^\mu + ig_s\boldsymbol{\lambda}/2\cdot\boldsymbol{A}^\mu)\psi \tag{186}$$

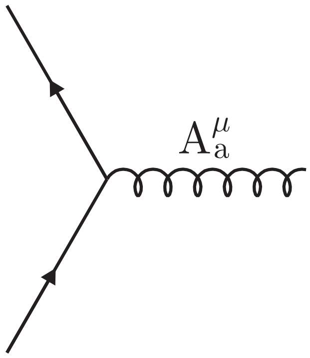

Fig. 13: A-ψ-ψ vertex.

giving the A-ψ-ψ vertex (cf (168)) of figure 13:

$$-ig_s \frac{\lambda_a}{2}\gamma^\mu\,. \tag{187}$$

The quanta of the A_a^μ field are the (eight different) gluons. As in local SU(2), there is an $\mathrm{SU(3)_c}$ field strength tensor which is (cf (177))

$$F_a^{\mu\nu} = \partial^\mu A_a^\nu - \partial^\nu A_a^\mu - g_s f_{abc} A_b^\mu A_c^\nu\,. \tag{188}$$

The $\mathrm{SU(3)_c}$ Yang-Mills term is then

$$-\frac{1}{4}F_a^{\mu\nu}F_{a\mu\nu} \tag{189}$$

and it contains triple and quadruple gluon couplings, all involving the same 'strong' coupling g_s, and the constants f_{abc} determined from (155). Once again, there is no mass term allowed by invariance under (185), and the gluons are massless. Their propagator is the same as the photon one in rule (v), with a colour factor δ_{ab}.

For one $\mathrm{SU(3)_c}$ triplet ψ, then, our Lagrangian so far is

$$\mathcal{L} = \bar\psi(i\not{D} - m)\psi - \frac{1}{4}F_a^{\mu\nu}F_{a\mu\nu} \tag{190}$$

with $D^\mu\psi$ given by (186). For many different quark flavours f, the Dirac term is repeated for each, giving

$$\mathcal{L}_{\mathrm{QCD}} = \sum_f \bar\psi_f(i\not{D} - m_f)\psi_f - \frac{1}{4}F_a^{\mu\nu}F_{a\mu\nu}\,. \tag{191}$$

Actually, however, matters are not quite that simple. As in QED, we need a gauge-fixing term to produce the gauge field propagator; in the non-Abelian case this turns out to be a more complicated affair, necessitating additional pieces in $\mathcal{L}_{\mathrm{QCD}}$ called 'ghost terms'. We shall not give their form here: they are needed only for loop calculations, the details of which we shall not need. The Lagrangian of (191) is adequate at the tree level.

Problems for Lecture 4

P4.1 An 'infinitesimal' SU(2) transformation means one very close to the identity, $\tilde U = 1 - i\xi$ where ξ is a matrix whose entries are infinitesimally small. So $\tilde U = \begin{pmatrix} 1 - i\xi_{11} & -i\xi_{12} \\ -i\xi_{21} & 1 - i\xi_{22} \end{pmatrix}$. Show that to first order in the ξ's, $\tilde U \tilde U^\dagger = I$ implies that $\xi = \xi^\dagger$ (i.e. ξ is Hermitean). Also, show (again to first order in the ξ's) that $\det U = 1$ implies $\xi_{11} + \xi_{22} = 0$ (i.e. ξ is traceless). So ξ is a traceless Hermitean matrix, 2×2. Explain why ξ is specified by three real parameters. How many parameters are needed for an infinitesimal SU(N) matrix?

P4.2 The τ-matrices are

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

(a) Verify that $[\tau_1/2, \tau_2/2] = i\tau_3/2$. (b) A simple model of the isospin raising operator $\hat{T}_+$ is

$$\hat{T}_+ = (\hat{a}_u^\dagger \ \hat{a}_d^\dagger)(\tau_1/2 + i\tau_2/2) \begin{pmatrix} \hat{a}_u \\ \hat{a}_d \end{pmatrix}$$

where the $\hat{a}^\dagger$'s create u's and d's. Check that $\hat{T}_+ = \hat{a}_u^\dagger \hat{a}_d$ and interpret this. Define also

$$\hat{T}_- = (\hat{a}_u^\dagger \ \hat{a}_d^\dagger)(\tau_1/2 - i\tau_2/2) \begin{pmatrix} \hat{a}_u \\ \hat{a}_d \end{pmatrix}.$$

Show that $\hat{T}_- = \hat{a}_d^\dagger \hat{a}_u$. (c) Evaluate $[\hat{T}_+, \hat{T}_-]$, and check that it is compatible with $[\hat{T}_i, \hat{T}_j] = i\epsilon_{ijk}\hat{T}_k$, where

$$\hat{T}_i = (\hat{a}_u^\dagger \ \hat{a}_d^\dagger)(\tau_i/2) \begin{pmatrix} \hat{a}_u \\ \hat{a}_d \end{pmatrix}.$$

P4.3 The 3×3 matrices $t_1, \ t_2, \ t_3$ are defined by $(t_i)_{jk} = -i\epsilon_{ijk}$ for $i, j, k = 1, 2, 3$ where the index i stands for which t it is, and the j, k indices specify the row and column, respectively, of that ith t matrix. Here ϵ_{ijk} is defined to be 0 if any of i, j, k are equal, +1 if they are a cyclic permutation of '123', and -1 if they are a cyclic permutation of '213'. Write down the 3×3 matrices $t_1, \ t_2, \ t_3$, and verify that $[t_1, t_2] = t_3$.

P4.4 The infinitesimal transformation law of an SU(2) triplet $\boldsymbol{\phi}$ is

$$\begin{pmatrix} \phi_1' \\ \phi_2' \\ \phi_3' \end{pmatrix} = (1 - i\epsilon_1 t_1 - i\epsilon_2 t_2 - i\epsilon_3 t_3) \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{pmatrix}.$$

Calculate the 3×3 transformation matrix explicitly, and show that the transformation can also be written in 'cross product' form $\boldsymbol{\phi}' = \boldsymbol{\phi} + \boldsymbol{\epsilon} \times \boldsymbol{\phi}$.

P4.5 The 'SU(2) covariant derivative' acting on an SU(2) doublet is $D^\mu \psi = (\partial^\mu + ig\boldsymbol{\tau} \cdot \boldsymbol{W}^\mu(x)/2)\psi$. Under an infinitesimal local SU(2) transformation, ψ transforms by

$$\delta\psi = -ig\boldsymbol{\tau} \cdot \boldsymbol{\epsilon}(x)/2 \ \psi.$$

The transformation law of $\boldsymbol{W}^\mu$ is determined from the requirement that

$$\delta(D^\mu \psi) = -ig\boldsymbol{\tau} \cdot \boldsymbol{\epsilon}(x)/2 \ (D^\mu \psi).$$

Now the LHS of this equation is

$$\begin{aligned} &\delta[(\partial^\mu + ig\boldsymbol{\tau} \cdot \boldsymbol{W}^\mu(x)/2)\psi] \\ &\qquad = \ ig\boldsymbol{\tau} \cdot (\delta\boldsymbol{W}(x)^\mu/2) \ \psi + (\partial^\mu + ig\boldsymbol{\tau} \cdot \boldsymbol{W}^\mu(x)/2)\delta\psi \\ &\qquad = \ ig\boldsymbol{\tau} \cdot (\delta\boldsymbol{W}(x)^\mu/2) \ \psi + (\partial^\mu + ig\boldsymbol{\tau} \cdot \boldsymbol{W}^\mu(x)/2)(-ig\boldsymbol{\tau} \cdot \boldsymbol{\epsilon}(x)/2)\psi \end{aligned}$$

while the RHS is

$$-ig\boldsymbol{\tau} \cdot \boldsymbol{\epsilon}(x)/2 \ (\partial^\mu + ig\boldsymbol{\tau} \cdot \boldsymbol{W}^\mu(x)/2)\psi.$$

Verify that this implies

$$\delta\boldsymbol{W}^\mu(x) = \partial^\mu \boldsymbol{\epsilon}(x) + g\boldsymbol{\epsilon}(x) \times \boldsymbol{W}^\mu(x).$$

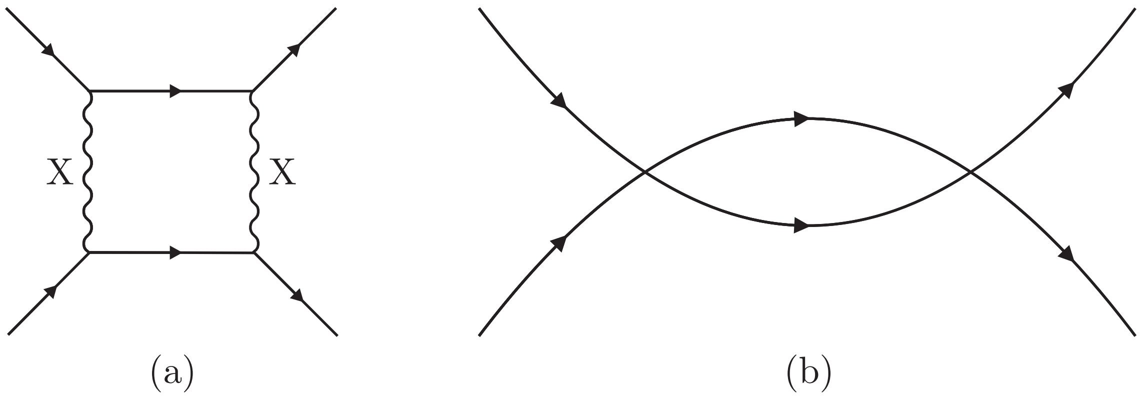

Fig. 14: Two-X exchange in fermion-fermion scattering, and effective four-fermion structure.

P4.6 Check that

$$[\partial^\mu + ig\boldsymbol{\tau}\cdot\boldsymbol{W}^\mu(x)/2, \partial^\nu + ig\boldsymbol{\tau}\cdot\boldsymbol{W}^\nu(x)/2] = ig\boldsymbol{\tau}\cdot\boldsymbol{F}^{\mu\nu}/2$$

where

$$\boldsymbol{F}^{\mu\nu} = \partial^\mu\boldsymbol{W}^\nu(x) - \partial^\nu\boldsymbol{W}^\mu(x) - g\boldsymbol{W}^\mu(x)\times\boldsymbol{W}^\nu(x).$$

P4.7 Verify that, under an infinitesimal local SU(2) transformation, $\delta\boldsymbol{F}^{\mu\nu} = g\boldsymbol{\epsilon}(x)\times\boldsymbol{F}^{\mu\nu}$.

6. SPONTANEOUS SYMMETRY BREAKING

See chapter 21 of [2].

6.1 Some motivation

In the previous section, an indication was given as to why the relevant theories at current energy scales should be renormalisable theories (a small subclass, incidentally, out of all possible quantum field theories!). We also pointed out how 'universality' phenomena in weak interactions suggested that they are described by a gauge theory, which presumably should be a renormalisable one. On the other hand, we also know that weak interactions are very short-ranged, so their mediating quanta must be massive - and this at once seems to present a barrier to the 'gauge' idea, because (see problem P2.6) a simple gauge boson mass term violates gauge invariance. Perhaps, then, we can have a theory involving massive charged $W^\pm$ bosons, for instance, without it being a gauge theory? Yes, we can, but *it will not be renormalisable*. In fact, the renormalisability of QED has a great deal to do with the gauge symmetry it possesses. Let's try and explain what's wrong with a 'non-gauge theory of massive $W^\pm$'s'.

Consider figure 14, which shows some kind of fermion-fermion scattering proceeding, in fourth order of perturbation theory (one loop), via the exchange of two massive vector bosons that we'll call X^μ. To calculate this diagram, we need to know the propagator for X^μ.

For this we need the wave equation for X^μ, which is quite simple to write down. We just replace $\Box$ in the wave equation (91) for A^μ by $\Box + M^2$ where M is the mass of the X^μ:

$$(\Box + M^2)X^\mu - \partial^\mu\partial_\nu X^\nu = 0. \tag{192}$$

To find the propagator, we follow the poor-man's route, putting in a plane wave solution for X^μ, which yields

$$\left[(-q^2 + M^2)\delta^\mu_\nu + q^\mu q_\nu\right]\epsilon^\nu e^{-iq\cdot x} = 0. \tag{193}$$

The propagator should now be proportional to the inverse of the [...] bracket in (193), and (*unlike* the corresponding inverse in (92)!) this does exist and is given by (problem P5.1)

$$\frac{-g^{\mu\nu} + q^\mu q^\nu/M^2}{q^2 - M^2}. \tag{194}$$

Note (i) that trouble ensues (the numerator blows up) when $M \to 0$, so already we see that a massless vector particle seems to be a very different kind of thing from a massive one (you can't just simply take the massless limit); (ii) that if we 'dot' (193) with q_μ we easily deduce $q \cdot \epsilon = 0$ (see below, after (198)).

Now consider the loop integral in figure 14. At each vertex we will have a coupling constant factor 'g', which is in fact dimensionless (the interaction will be something like $g\bar{\psi}\gamma_\mu\psi X^\mu$). But, as we warned in section 4.4, this may not guarantee renormalisabilty, and this is a case where it does not. To get an idea of why not, consider the leading divergent behaviour of figure 14. This will be associated with the '$q^\mu q^\nu$' terms in the numerator of (194), so that the leading divergence is effectively

$$\sim \int d^4q \left(\frac{q^\mu q^\nu}{q^2}\right)\left(\frac{q^\mu q^\nu}{q^2}\right)\frac{1}{q}\frac{1}{q} \tag{195}$$

for high q (we are of course not troubling to get all the indices etc right). But the first two (...)'s in (195) behave like a constant, at large q, so that the asymptotic behaviour is effectively

$$\sim \int d^4q \frac{1}{q}\frac{1}{q} \tag{196}$$

which is *exactly what we would get in a four-fermion theory* ! - see figure 14, and we know that such a theory is non-renormalisable.

Where have these dangerous powers of q come from? The answer is simple and important. They come from the *longitudinal* polarisation state of the massive X particle. We can see this as follows. Consider a free X particle with 4-momentum $q = (q^0, 0, 0, |\boldsymbol{q}|)$, so that the x and y directions are transverse, and the z direction is longitudinal. In the rest frame of the X, the three polarisation states can be taken to be

$$\boldsymbol{\epsilon}(\lambda = \pm 1) = \mp 2^{-\frac{1}{2}}(1, \pm i, 0), \quad \boldsymbol{\epsilon}(\lambda = 0) = (0, 0, 1). \tag{197}$$

Boosting to the frame with 4-momentum q, the transverse polarisation vectors remain the same, but the longitudinal one becomes

$$\epsilon^\mu(q, \lambda = 0) = M^{-1}(|\boldsymbol{q}|, 0, 0, q^0). \tag{198}$$

Note that $q \cdot \epsilon(q, \lambda = 0) = 0$ is satisfied. At large values of q, $\epsilon^\mu(q, \lambda)$ is therefore proportional to q^μ/M, and this is the origin of such factors in the propagator.

Consider now the photon propagator given by rule (v): there are apparently quite similar factors there too, but they are gauge dependent, and in fact *can be 'gauged away' entirely by choice of* ξ! But, as we have seen, such 'gauging' seems to be possible only in a massless vector theory. A closely related point is that, as we all know, electromagnetic waves are purely transverse: equivalently, free photons exist in only two independent polarisation states, instead of the three we might have expected (from the three orientations of their unit spin). The longitudinal state is missing, and it turns out (see Aitchison and Hey [1] page 188) that this is precisely related to the masslessness of the photon. In the massive X case, all three polarisation states are present - and this gives another way of seeing why a massless vector particle is really different from even a very light massive one: there is no smooth naive $M \to 0$ limit.

This above considerations therefore suggest the following line of thought:

• can we somehow create a gauge theory involving massive vector quanta, such that the offending $q^\mu q^\nu$ bits could be gauged away, making the theory renormalisable?

The answer is yes, via the idea of *spontaneous breaking* of the gauge symmetry.

This terminology is contrasted with 'explicit symmetry breaking', in which the observed symmetry breaking is associated with a term in the Lagrangian, in the absence of which the theory would possess some exact symmetry. For example, to the extent that the up and down quark masses are equal, we have approximate SU(2) flavour symmetry of the QCD Lagrangian. But it is also possible to have a symmetrical Lagrangian, while the particle states and other physical observables seem to show no obvious (even approximate) sign of the symmetry. This is the 'spontaneously broken' case. This language is borrowed from condensed matter physics, where the ferromagnet is the frequently quoted example. The (Heisenberg) Hamiltonian is certainly rotationally invariant, yet below the transition temperature the spins are thought of as lining up in some particular direction, breaking the rotational symmetry 'spontaneously'.

In the case of a field theory, there are striking differences in the physical consequences depending on whether the symmetry that is spontaneously broken is a global or a local one. In the global case, a general result due to Goldstone [3] and others states that spontaneous breaking of a continuous symmetry is always associated with the appearance of a massless particle, or particles, called 'Goldstone bosons'. In the local case, these Goldstone bosons become the longitudinal components of the gauge field(s) - which, before symmetry breaking, always had only the two transverse components. The total of three 'spin' components in all is exactly what is required for a *massive* vector field. This is the essence of the theoretical loophole which allows gauge bosons to be massive even though the Lagrangian is locally (gauge-) invariant (cf problem P2.6), and which is invoked to give masses to the W and the Z bosons in the Standard Model.

We begin with the simpler case of spontaneously broken global symmetry, which is of physical importance in its own right in the non-Abelian case (section 6.3).

6.2 Spontaneously broken global U(1) symmetry

See chapter 17 of [2].

We consider a simple classical field theory which shows the effect we want to study. Let ϕ be a complex scalar field, described by the Lagrangian

$$\mathcal{L}_\phi = \partial_\mu \phi^* \partial^\mu \phi - V(\phi) \tag{199}$$

where the potential is taken to have the form ($\lambda > 0$)

$$V(\phi) = -\mu^2 \phi^* \phi + \frac{\lambda}{4} (\phi^* \phi)^2 . \tag{200}$$

Clearly $\mathcal{L}_\phi$ is invariant under the global U(1) symmetry

$$\phi \to \phi' = e^{-i\alpha} \phi . \tag{201}$$

(Note that a term like $(\phi^* \phi)^3$ would also be invariant under (201), but this would be a non-renormalisable interaction in the quantum theory of $\mathcal{L}_\phi$, so we exclude it.)

Application of the Euler-Lagrange equation yields the equation of motion

$$(\Box - \mu^2)\phi = -\frac{\lambda}{2} |\phi|^2 \phi . \tag{202}$$

This is nearly the standard Klein-Gordon equation for ϕ (with an interaction term on the right-hand side) - except for the fact that '$-\mu^2$' has the wrong sign for a mass term! This prevents us from making any quantum interpretation of (199) as yet; we therefore concentrate on $V(\phi)$ regarded simply as the potential energy of the classical field.

As a first step to understanding (199), we try to identify the configuration(s) of minimum energy, about which the system might be expected to oscillate. Generally, the energy will be a minimum when

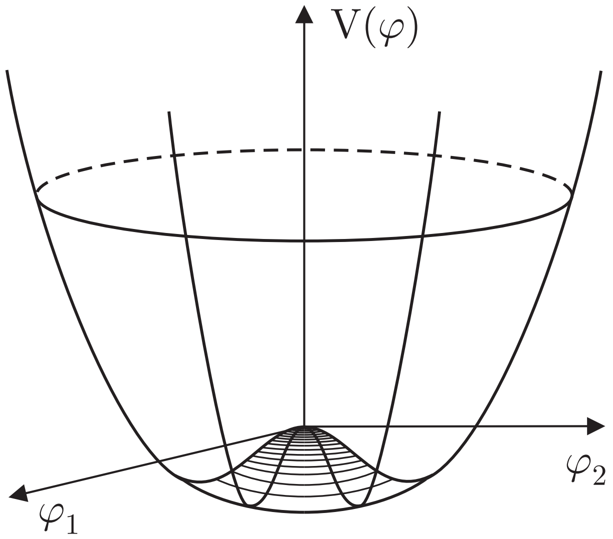

Fig. 15: The ‘wine-bottle’ potential of (200).

ϕ is a constant, which reduces the kinetic terms to zero. The minimum energy is then reached at the minimum of $V(\phi)$. This occurs at

$$|\phi| = v/\sqrt{2}, \quad v = 2\mu/\lambda^{1/2}, \tag{203}$$

where v is referred to as the ‘symmetry breaking parameter’. To have a clearer picture, it is helpful to introduce two real fields ϕ_1 and ϕ_2 by

$$\phi = (\phi_1 - i\phi_2)/\sqrt{2} \tag{204}$$

and also the ‘polar’ variables

$$\phi = (\rho/\sqrt{2})e^{i\theta/v}, \tag{205}$$

where the v is inserted so that θ has the same dimensions as ρ. Figure 15 shows $V(\phi)$ versus ϕ_1 and ϕ_2, from which it is obvious that the minimum of V is not at $\phi_1 = \phi_2 = 0$. In fact, there is *no* unique minimum point - rather, any value on the circle $\phi_1^2 + \phi_2^2 = v^2$ or equivalently $\rho = v$ will do. Before proceeding further, we briefly outline the condensed matter analogue of (199) and (200) which we mentioned earlier - namely the ferromagnet. In this case, one considers the free energy as a function of the magnetisation $\boldsymbol{M}$ at a given temperature T, and makes an expansion of the form

$$F \approx F_0(T) + \frac{1}{2}\mu^2(T)\boldsymbol{M}^2 + \frac{1}{4}\lambda(T)(\boldsymbol{M}^2)^2 + \cdots, \tag{206}$$

valid for small magnetisation. If the parameter μ^2 is positive, it is easy to see that F has a simple ‘bowl’ shape as a function of $|\boldsymbol{M}|$, with a minimum at $|\boldsymbol{M}| = 0$. This is the case for T greater than the ferromagnetic transition temperature T_{C}. However, if one assumes that $\mu^2(T)$ becomes negative for $T < T_{\mathrm{C}}$ (so that $\mu^2(T_{\mathrm{C}}) = 0$), then F will now look like figure 15 and the minimum free energy will occur for $|\boldsymbol{M}| \neq 0$. The interpretation is that in this case the ground state will be magnetised. Any direction of $\boldsymbol{M}$ is possible (only $|\boldsymbol{M}|$ is specified); but when the system does settle into one actual configuration with $\boldsymbol{M} \neq \mathbf{0}$ the original full rotational invariance of (206) is lost - the magnetisation, and the breaking of the symmetry, has occurred ‘spontaneously’.

In the same way, any particular minimum on the circle $\rho = v$ will select out a particular θ in (205), breaking ‘spontaneously’ the invariance (201).

In quantum field theory, particles are thought of as excitations from a ground state, which we call 'the vacuum'. Figure 15 strongly suggests that if we want a decent quantum interpretation of (199), we should consider expanding the fields about a point on the circle of minima, about which stable oscillations are likely. Any such point represents a possible vacuum state in which

$$\langle 0|\,\phi_1^2+\phi_2^2\,|0\rangle = v^2\,, \qquad \text{or } \langle 0|\,\rho\,|0\rangle = v\,. \tag{207}$$

Bearing in mind (cf (200)) that for a field with a conventional (positive) mass^2 parameter the potential would be U-shaped, we might guess that 'radial' oscillations in figure 15 would correspond to a conventional massive field, while 'angle' oscillations - which pass through all the degenerate minima (vacua) - have no 'restoring force' and are massless. Accordingly, we set (cf (205))

$$\phi(x) = \frac{1}{\sqrt{2}}(v+h(x))e^{-i\theta(x)/v} \tag{208}$$

and find that $\mathcal{L}_\phi$ becomes (problem P5.2)

$$\mathcal{L}_\phi = \frac{1}{2}\partial_\mu h\partial^\mu h - \mu^2 h^2 + \frac{1}{2}\partial_\mu\theta\partial^\mu\theta + \frac{\mu^4}{\lambda} + \text{terms cubic and quartic in } \theta, h\,. \tag{209}$$

Equation (209) exhibits the desired form of a conventional scalar field h with mass $\sqrt{2}\mu$ and a massless field θ, together with interaction terms. In particular, the quantum version of (209) will have $\langle 0|h(x)|0\rangle = \langle 0|\theta(x)|0\rangle = 0$, consistent with (207), so that h and θ will have the usual mode expansions (of the form (19) for example), allowing the usual particle interpretation. (The constant term in (209), which does not affect equations of motion, reflects the fact that $V(\mathrm{min}) = -\mu^4/\lambda$). Note that the symmetry (201), which is evident in (199), is well and truly *hidden* in (209)!

This model (due originally to Goldstone [3]) contains the essence of spontaneous symmetry breaking in field theory: a non-zero value of a field in the ground state (vacuum), a zero mass mode or modes (the Goldstone bosons), and a massive excitation or excitations in the directions 'perpendicular' to the degenerate ground states.

It is interesting to find out what happens to the symmetry current corresponding to the invariance (201). Following the usual procedure, this current is

$$j^\mu_\phi \;=\; i\left\{\phi^\dagger\partial^\mu\phi - (\partial^\mu\phi)^\dagger\phi\right\} \;=\; v\partial^\mu\theta + 2h\partial^\mu\theta + h^2\partial^\mu\theta/v\,. \tag{210}$$

The presence of the term involving just the *single* field θ is very remarkable: it tells us that (in the quantum theory) there is a non-zero matrix element of the form

$$\langle 0|\,j^\mu_\phi(0)\,|\theta\rangle = -ip^\mu v\,, \tag{211}$$

where $|\theta\rangle$ stands for a state with one Goldstone boson θ, with momentum p^μ. That is, the symmetry current connects the Goldstone boson to the vacuum, with an amplitude proportional to the symmetry breaking parameter. In the case of spontaneously broken chiral $\mathrm{SU(2)_{f5}}$ symmetry (section 6.3 below), the analogue of j^μ_ϕ is the current of the global axial SU(2) symmetry A^μ_i, and there are three θ modes which are identified with the physical pions. The parameter v in the corresponding equation (211) is then f_π ($\sim$ 94MeV), the constant which enters into the pion decay $\pi \to \ell\nu$.

Although by the ansatz (208) we seem to have arrived at a viable particle interpretation of (199), we might well ask: how would such a negative (mass)^2 term arise in quantum field theory? One possible answer is that, as with the ferromagnetic analogy, the coefficient μ^2 in (200) could be temperature dependent: perhaps at extremely high temperatures, such as prevailed in the early universe, μ^2 had the opposite sign, corresponding to a conventional mass term. In that case the potential would have a simple minimum at the origin, and the symmetry would not be spontaneously broken until T dropped below some $T_{\rm C}$, where $\mu^2(T_{\rm C}) = 0$. This simple picture is indeed popular in models of the early universe, where such phase transitions are proposed. On the other hand, it may be that some theory might predict the coefficient μ^2 in (200) to be negative, in a particular case. Or, one might simply postulate a $V(\phi)$ of the form (200), so as to 'trigger' the desired breakdown. The last alternative is essentially what is done in the Higgs sector of the Standard Model - as we will discuss in section 6.5 and section 7.

6.3 Spontaneously broken global chiral symmetry

See section 12.3.2, and chapter 17, of [2].

The Dirac Lagrangian for a single massless fermion,

$$\bar{\psi} i \not{\partial} \psi \tag{212}$$

is invariant not only under the ordinary global U(1) symmetry of (69), but also under the 'γ_5-version' of it, namely

$$\psi \to \psi' = e^{-i\eta\gamma_5}\psi \,. \tag{213}$$

This can be easily verified directly, using

$$\gamma^0\gamma^5 = -\gamma^5\gamma^0, \quad \gamma^i\gamma^5 = -\gamma^5\gamma^i \,, \tag{214}$$

but it will be useful later to expand the discussion now to cover this type of symmetry, not considered previously. We may write

$$\psi = \frac{(1-\gamma_5)}{2}\psi + \frac{(1+\gamma_5)}{2}\psi \;\equiv\; \psi_{\rm L} + \psi_{\rm R}. \tag{215}$$

The ordinary (infinitesimal) U(1) symmetry (69) is then

$$\delta\psi_{\rm R} = -i\epsilon\psi_{\rm R}, \quad \delta\psi_{\rm L} = -i\epsilon\psi_{\rm L} \tag{216}$$

while the infinitesimal version of (213) is

$$\delta\psi_{\rm R} = -i\eta\psi_{\rm R}\,, \quad \delta\psi_{\rm L} = +i\eta\psi_{\rm L}\,. \tag{217}$$

Transformations such as (217), which act differently on the L and R components are called 'chiral'. Using (214), (215) can be written as

$$\bar{\psi} i\not{\partial}\psi = \bar{\psi}_{\rm L} i\not{\partial}\psi_{\rm L} + \bar{\psi}_{\rm R} i\not{\partial}\psi_{\rm R}\,, \tag{218}$$

which clearly exhibits both the symmetries (216) and (217). It is also manifestly L $\leftrightarrow$ R symmetric, which means it conserves parity. On the other hand, a mass term $m\bar{\psi}\psi$ becomes

$$m\bar{\psi}\psi = m(\bar{\psi}_{\rm L}\psi_{\rm R} + \bar{\psi}_{\rm R}\psi_{\rm L}) \tag{219}$$

which is invariant under (216) but not under (217), while still preserving parity.

Consider then $\mathcal{L}_{\rm QCD}$ of (191), in the limit in which some quark masses - in particular the lightest, m_u and m_d - are regarded as negligible. The fact that $\not{\partial}$ in (212) is replaced by $\not{D}$ clearly makes no difference to the preceding discussion, which depended only on (214). Thus in this limit $\mathcal{L}_{\rm QCD}$ will be invariant under the γ_5-version of (141), namely

$$\delta\psi = -i\boldsymbol{\eta}\cdot\boldsymbol{\tau}/2\,\gamma_5\,\psi\,, \tag{220}$$

which is a chiral '$\mathrm{SU(2)}_{\mathrm{f}5}$' transformation. Now this cannot be realised as an exact symmetry in nature, or else for every non-strange baryon made of u and d quarks there would have to exist another one, degenerate in mass, but with the opposite parity. The reason is worth pausing over.

Associated with the invariance (220) will be three conserved charges, just as in (141)-(145), namely

$$T_i^5 = \int \psi^\dagger(x) \frac{\tau^i}{2} \gamma_5 \psi(x) d^3x. \tag{221}$$

In this case, however, these objects are 'pseudoscalars' (because of the γ_5) - meaning that they will change the parity of any state they act on. Thus whereas the ordinary isospin raising operator $T_+ = T_1 + iT_2$ has the action $T_+|d\rangle = |u\rangle$, where u and d are degenerate in mass because $[T_+, H] = 0$ (T_+ is a constant of the motion), in the case of T_+^5 we must have

$$T_+^5|d\rangle = |\tilde{u}\rangle \tag{222}$$

where $\tilde{u}$ is an 'up' state, degenerate in mass with $|u\rangle$ (because $[T_+^5, H] = 0$ also), but with opposite parity.

Such negative parity analogues of all non-strange baryons are not seen experimentally. One might of course blame this on the finite mass of the u and d quarks, but this is implausible. Instead, we try the idea that this chiral symmetry is spontaneously broken. In that case, we expect three massless Goldstone bosons (corresponding to the three independent SU(2) chiral transformations), and we can interpret $\tilde{u}\rangle$ of (222) as being really $|u + \text{massless pseudoscalar boson}\rangle$, thus producing a state degenerate with u in mass, but of opposite parity! These three massless Goldstone bosons are identified with the *pions* - thereby explaining their anomalously low mass (by comparison with that of the ρ-meson, for example). The mass of the physical pion is not, of course, strictly zero, and this is attributed to small non-zero quark masses in the original QCD Lagrangian. Still useful, though more 'explicitly' broken than this chiral SU(2), is the chiral flavour SU(3) analogue, in which we suppose $m_s \approx 0$ - the Goldstone bosons are then the kaons.

Remarkably enough, these ideas are also relevant to the weak interactions. In this case, as we shall see, the interaction is most definitely not left-right symmetric (it violates parity) - indeed the 'V$-$A' structure means that the weak gauge fields couple only to the $\psi_{\rm L}$ components of the fermions, and not to the $\psi_{\rm R}$ components at all. This means that the corresponding local gauge symmetry is of the form

$$\begin{aligned} \delta\psi_{\rm L} &= -i\boldsymbol{\epsilon}\cdot\boldsymbol{\tau}(x)/2\,\psi_{\rm L} && (223)\\ \delta\psi_{\rm R} &= 0\,, && (224) \end{aligned}$$

for a 'weak doublet' such as

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}. \tag{225}$$

But this implies that any mass term of the form (219), which treats $\psi_{\rm L}$ and $\psi_{\rm R}$ the same, will break this 'left-handed' gauge symmetry. Although the neutrinos were usually taken to be massless, the other leptons are definitely not, nor are the quarks. Thus, curiously enough, there is another 'mass problem' with the weak interactions: they would like not only the W and Z bosons but also the fermions to be massless. Once again, we shall have to suppose that the fermion masses arise 'spontaneously', if we want to save the (weak) gauge symmetry. In the Standard Model, one appeals to the same mechanism (the Higgs field) to give mass to the gauge bosons and to the fermions, which is an economical but not necessary step; see section 7.

It is now time to turn to spontaneously broken local symmetries, concentrating on those relevant to the Standard Model.

6.4 Spontaneously broken local U(1) symmetry: the Abelian Higgs model

See section 19.3 of [2].

The U(1) Higgs model is just $\mathcal{L}_\phi$ of (199) extended so as to be locally U(1) invariant; it provides a beautifully simple model for investigating what happens when a *gauge* symmetry is spontaneously broken. To make (199) locally U(1) invariant, we need only replace ∂'s by D's as in (81), and add the Maxwell piece, giving

$$\mathcal{L}_h = [(\partial_\mu + ieA_\mu)\phi]^\dagger [(\partial^\mu + ieA^\mu)\phi] - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - V(\phi) \tag{226}$$

where V is still (200), and of course $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$. (226) is invariant under the local version of (201), namely

$$\phi \to \phi'(x) = e^{-i\alpha(x)}\phi(x) \tag{227}$$

when accompanied by a gauge transformation on A^μ

$$A^\mu \to A^{\mu\prime} = A^\mu + \frac{1}{e}\partial^\mu \alpha \tag{228}$$

as in section 3.4. Before proceeding further, we note at this stage that we have four field degrees of freedom - two in ϕ and two in the massless A^μ $(F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu)$.

Now we have learned that the form of V in (200) does not lend itself to a natural particle interpretation, which only appears after making the 'shift to the minimum', as in (208). But there is a remarkable difference between the local and global cases. In the local case, the phase of ϕ is completely arbitrary, since any change in $\theta(x)$ in (208) can be compensated by an appropriate transformation (228) on A^μ, leaving $\mathcal{L}_h$ the same as before. Thus in fact the 'θ' field in (208) can be 'gauged away' altogether, if we like! This must mean that the massless Goldstone boson, described precisely by θ in the quantum theory, somehow no longer appears. This is the first unexpected result in the local case (and it reminds us of our desire to 'gauge away' those longitudinal polarisation states . . .).

However, we cannot simply 'lose' degrees of freedom. Somehow the system must keep track of the fact that we started with four. To see what has happened, we substitute (208) into (226) with $\theta = 0$; i.e. set

$$\phi = \frac{1}{\sqrt{2}}(v + h(x)) \tag{229}$$

in $\mathcal{L}_h$. We find then (problem P5.3)

$$\mathcal{L}_h = \frac{1}{2}\partial_\mu h\, \partial^\mu h - \mu^2 h^2 + \frac{\mu^4}{\lambda} - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}e^2 v^2 A_\mu A^\mu + \text{ interaction terms,} \tag{230}$$

where A^μ has to be understood as the gauge field after the transformation needed to reduce ϕ to (229). Equation (230) shows the second 'Higgs miracle': we see that the A^μ field now has a *mass*, equal to ev where v is the symmetry breaking parameter. The missing degree of freedom has reappeared as the third (longitudinal) polarisation state of the massive field A^μ. The fourth degree of freedom is still there, the massive h field as in (209).

Can such miracles ever occur? The answer is undoubtedly yes, at least in the non-relativistic case. The low-energy version of $\mathcal{L}_h$ is just the Ginzburg-Landau (GL) approximation for (again) the free energy in a superconductor. In this case (see section 19.2 of Aitchison and Hey [2] for example) ϕ represents a composite (rather than elementary) field, such that $|\phi|^2$ is the density of bound Cooper pairs (of e^-e^-). Also, the mass for the A-field implies that the field is exponentially attenuated inside the superconductor, with a penetration length of order $1/ev$; this is the Meissner effect. It is worth noting that the GL free energy is not to be regarded as a fundamental theory, which must of course be derived from the physical electron-electron and electron-lattice interactions; this is what the BCS theory is all about, and the GL free energy is a phenomenological expression embodying much of the important physics of the BCS theory. In particle physics the question of whether the ϕ field in the Standard Model (see section 7) is elementary or composite is completely unknown. However, whatever the truth of that may be, it seems pretty well inevitable that some such field, or effective field, is required to give mass to the

W and Z (see section 6.5 , and section 7)—and in that case it should have its own excitation quantum, the *Higgs boson*: hence the intense interest in hunting for it!

Before proceeding further we can at this stage read off from (230) the propagator for the massive vector A-field. As in the discussion following (193), we need to invert the quantity $P_{\mu\nu}(M_A) = [(-k^2 + M_A^2)g_{\mu\nu} + k_\mu k_\nu]$, where $M_A = ev$ here. As we saw, this does have a straightforward inverse, leading to the propagator

$$i\frac{(-g_{\mu\nu} + \frac{k_\mu k_\nu}{M_A^2})}{k^2 - M_A^2} . \tag{231}$$

We see that (231) makes no sense as $M_A \to 0$, reflecting the difficulty with the massless limit of the massive theory. A more technical point concerns the fact that (231) obtains only when the special choice of gauge, $\theta = 0$, is made as in (229). In general, the vector propagator will contain a gauge parameter ξ like the massless propagator of rule (v): this is after all a gauge theory! Rule (v) becomes

- rule (v)$'$ a factor $i[-g^{\mu\nu} + \frac{(1-\xi)q^\mu q^\nu}{q^2-\xi M^2}]/(q^2 - M^2)$ for an internal massive gauge boson carrying 4-momentum q, where ξ is a gauge parameter ($\xi \to \infty$ gives the 'naive' vector boson propagator).

Note that for finite ξ, this propagator has a large q behaviour $\sim 1/q^2$, which is good enough to make figure 14 convergent! This, then, is the essential clue as to how we can have a renormalisable theory with massive gauge bosons. The gauge $\xi \to \infty$ is called 'unitary gauge': in this gauge there is no visible sign of the scalar ϕ-field. But note that in gauges with ξ finite, the scalar field will also be present with a ξ-dependent propagator (associated with the degree of freedom suppressed in (229)); the complete theory is nevertheless always ξ-independent. Further discussion of this is contained in section 19.5 of Aitchison and Hey [2] for example.

Returning to (226), we can again look at the electromagnetic current in this 'spontaneously broken local U(1)' model. The gauge invariant form of (210) is

$$\begin{aligned} j^\mu_{\text{e.m.}} &= ie\left[\phi^\dagger(\partial^\mu + ieA^\mu)\phi - \text{ complex conjugate}\right] \\ &= ie(\phi^\dagger\partial^\mu\phi - (\partial^\mu\phi^\dagger)\phi) - 2e^2A^\mu\phi^*\phi . \end{aligned} \tag{232}$$

Inserting (208) into (232) (this time in a gauge such that $\theta \neq 0$) we find (cf (210))

$$j^\mu_{\text{e.m.}} = -e^2v^2A^\mu + ev\partial^\mu\theta + \text{interaction terms} . \tag{233}$$

(233) tells us that there is a 'screening current' (the first term on the RHS) which leads to a mass ev of the A-field, once again; the second term shows that - as in (211) - the vacuum couples to the 'would-be Goldstone boson' (which has become the longitudinal part of the A-field) via the electromagnetic current.

This is an important observation as it leads to a somewhat different way of understanding the 'mechanism' whereby a gauge particle can become massive. In section 5.1 we introduced the photon self-energy $\Pi_{\rho\sigma}$ which had the general form

$$\Pi_{\rho\sigma} = (g_{\rho\sigma}q^2 - q_\rho q_\sigma)\Pi^\gamma(q^2) . \tag{234}$$

When all the self-energy insertions are summed up, and after renormalisation, the photon propagator has the form (cf (111))

$$-ig^{\mu\nu}/q^2\left(1 - \bar{\Pi}_\gamma(q^2)\right) , \tag{235}$$

in the Feynman gauge. The existence of the matrix element

$$\langle 0|j^\mu_{\text{e.m.}}(0)|\theta\rangle = -iq^\mu ev \tag{236}$$

means that $\Pi_{\rho\sigma}$ will now receive a contribution from the diagram of figure 16, where the dotted line represents the massless θ quantum. This is now a tree diagram, not a loop as in the e^+e^- contribution

Fig. 16: Massless Goldstone boson coupling to photon.

of figure 9(a), and so the contribution to $\Pi_{\rho\sigma}$ will involve simply the (massless) θ-propagator, with no momentum integration. The γ-θ vertex is given by (236), with the result that the contribution to $\bar{\Pi}_\gamma(q^2)$ in (235) is

$$\bar{\Pi}_\gamma^\theta(q^2) = e^2 v^2 / q^2 \,, \tag{237}$$

so that the pole in the photon propagator (235) is now at $q^2 = e^2 v^2$, and the photon has a mass ev, as before. We have been casual about questions of gauge choice in this argument, but the essential point is valid: a gauge quantum can acquire mass if (for some reason) its vacuum polarisation function has a zero mass pole (see the *Discussion point* after (99)). This pole can be associated with the 'elementary' massless quantum in a Higgs potential of the form (200), but it does not have to be. The massless quantum could equally well be a bound state in some strongly-interacting fermion-antifermion channel - in particular, a Goldstone boson arising from the spontaneous breaking of some global symmetry in a purely fermionic theory, for instance. All that is necessary is that it has a coupling of the form (236). The point of this latter interpretation is that only the product 'ev' has significance - there is no sign of figure 15, or of 'v' alone as the vacuum value of a scalar field. Theories of this latter type do seem to produce a natural 'dynamical' mechanism for gauge boson mass generation. Both the '$t\bar{t}$' models (Nambu [4]; Miransky et al [5], [6]; Bardeen et al [7]), and technicolour (Farhi and Susskind [8]), are of this type, but neither seem to be favoured by experiment. In the electroweak theory it is of course the W and Z particles that we want to be massive (while still being gauge bosons), not the photon. We therefore need to extend the above to the (non-Abelian) SU(2) case.

6.5 Spontaneously broken SU(2)$\times$U(1) symmetry: the gauge and Higgs field sectors of the electroweak theory

See section 19.6 of [2].

We shall confine ourselves to the particular case which we need for the electroweak theory. We consider a complex scalar (spin-0) SU(2) doublet

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{238}$$

where the complex ϕ^+ field destroys positively charged particles and creates negatively charged ones, and the complex ϕ^0 field creates neutral particles and antiparticles (a hadronic analogy would be the K^+ and K^0 fields under hadronic SU(2)$_\mathrm{f}$). The Lagrangian

$$\mathcal{L}_\Phi = (\partial_\mu \phi)^\dagger (\partial^\mu \phi) + \mu^2 \phi^\dagger \phi - \frac{\lambda}{4} (\phi^\dagger \phi)^2 \tag{239}$$

then exhibits a global SU(2) invariance of the form (cf (159))

$$\phi \to \phi' = \exp(-i\boldsymbol{\alpha}\cdot\boldsymbol{\tau}/2)\phi \,, \tag{240}$$

but this is spontaneously broken, the minimum of the potential in (239) occurring at (cf (207))

$$(\phi^\dagger \phi)_{\min} = 2\mu^2/\lambda \equiv v^2/2 \,. \tag{241}$$

As in the U(1) case, we interpret (241) in the quantum theory as (cf (207))

$$\langle 0 | \, \phi^\dagger \phi \, | 0 \rangle = v^2/2 \,, \tag{242}$$

so that the ϕ-field has a non-zero value in the vacuum. Once again, we exclude higher powers of $\phi^\dagger\phi$ in (237) on grounds of renormalisability.

As before, in order to get a sensible particle spectrum we must 'shift' the fields so as to deal with stable oscillations about the minimum (vacuum) given by (242). So we need to define '$\langle 0|\phi|0\rangle$' and expand about it, as in (207) and (208). In the present case, however, the situation is more complicated than (208), since the complex doublet (238) contains four real fields, parametrised for example as

$$\phi^+ = \frac{1}{\sqrt{2}}(\phi_1 - i\phi_2)\,, \quad \phi^0 = \frac{1}{\sqrt{2}}(\phi_3 - i\phi_4)\,; \tag{243}$$

(242) then becomes

$$\langle 0|\,\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2\,|0\rangle = v^2\,. \tag{244}$$

It is evident that we have a lot of freedom in choosing the $\langle 0|\phi_i|0\rangle$ so that (244) holds, and it is not at first obvious what an appropriate generalisation of (207) and (208) might be.

Furthermore, in this more complicated (non-Abelian) situation a qualitatively new feature can arise: it may happen that the chosen condition $\langle 0|\phi_i|0\rangle \neq 0$ is *invariant* under some subset of the allowed symmetry transformations. This would effectively mean that this particular choice of the vacuum state respected that subset of symmetries, which would therefore not be 'spontaneously broken' after all. Since each broken symmetry is associated with a massless Goldstone boson, we would then get fewer of these bosons than expected.

Just this happens (by design!) in the present case. To understand how it works, we must first recognize that, in addition to the global SU(2) symmetry of (4.41), $\mathcal{L}_\Phi$ of (240) is also invariant under a completely independent global U(1) symmetry of the form

$$\phi \to \phi' = e^{-i\beta}\phi \tag{245}$$

which just means that the phases of the upper and lower components of ϕ in (238) change simultaneously by the same amount. Thus the full symmetry of (239) is global SU(2)$\times$U(1) (which will be made local in a moment, as is required in the Standard Model).

Suppose then that we could choose the $\langle 0|\phi_i|0\rangle$ so as to break this SU(2)$\times$U(1) symmetry completely: we would then expect four massless fields. Actually, however, it is not possible to make such a choice. An analogy may make this point clearer. Suppose we were considering just SU(2), and the field ϕ was an SU(2)-triplet. Then we could always write $\langle 0|\,\boldsymbol{\phi}\,|0\rangle = v\boldsymbol{n}$ where $\boldsymbol{n}$ is a unit vector; but this form is invariant under rotations about the $\boldsymbol{n}$-axis, irrespective of where that points. In the present case, by using the freedom of global SU(2)$\times$U(1) phase changes, an arbitrary $\langle 0|\phi|0\rangle$ can be brought to the form

$$\langle 0|\phi|0\rangle = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix}. \tag{246}$$

In considering what symmetries are respected or broken by (246), it is easiest to look at infinitesimal transformations. It is then clear that the particular transformation

$$\delta\phi = -i\epsilon(1+\tau_3)\phi \tag{247}$$

(which is a combination of (245) and the 'third component' of (240)) is still a symmetry of (246) since

$$(1+\tau_3)\begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{248}$$

so that

$$\langle 0|\phi|0\rangle = \langle 0|\phi + \delta\phi|0\rangle\;; \tag{249}$$

we say that 'the vacuum is invariant under (247)', and when we look at the spectrum of oscillations about that vacuum we expect to find only three massless bosons, not four.

Oscillations about (246) are conveniently parametrised by

$$\phi = \exp(-i(\boldsymbol{\theta}(x)\cdot\boldsymbol{\tau}/2)v)\begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}(v+H(x)) \end{pmatrix}, \tag{250}$$

which is to be compared with (208). Inserting (250) into (239) (see problem P5.4), we easily find that no mass term is generated for the $\boldsymbol{\theta}$ fields, while the H field piece is

$$\mathcal{L}_H = \frac{1}{2}\partial_\mu H \partial^\mu H - \mu^2 H^2 + \text{ interactions} \tag{251}$$

just as in (209), showing that $m_H = \sqrt{2}\mu$.

As noted in section 6.3, there is an interesting physical example of a spontaneously broken global SU(2) symmetry, the SU(2)$_{\mathrm{f5}}$ symmetry of $\mathcal{L}_{\mathrm{QCD}}$, in which the three massless modes are identified with the pions. We cannot consider this in any more detail here, however, being concerned rather to proceed to the local version of the SU(2)$\times$U(1) model of (239). Such an extension is easily written down, just by using the SU(2) covariant form (3.28) and the U(1) covariant derivative of the form (163). In the notation we shall use in the next section, this means replacing (239) by

$$\mathcal{L}_{\mathrm{G}\Phi} = (D_\mu\phi)^\dagger(D^\mu\phi) + \mu^2\phi^\dagger\phi - \frac{\lambda}{4}(\phi^\dagger\phi)^2 - \frac{1}{4}\boldsymbol{F}_{\mu\nu}\cdot\boldsymbol{F}^{\mu\nu} - \frac{1}{4}G_{\mu\nu}G^{\mu\nu} \tag{252}$$

where

$$D_\mu\phi = (\partial_\mu + ig\boldsymbol{\tau}\cdot\boldsymbol{W}_\mu/2 + ig' B_\mu/2)\phi\,, \tag{253}$$

$\boldsymbol{F}_{\mu\nu}$ is as in (177), and $G_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$. Thus the $\boldsymbol{W}$'s are the SU(2) gauge fields, and the B is the U(1) gauge field. (252) is, in fact, the gauge and Higgs field sector of the Standard Model. As in the local U(1) case, the particle spectrum is most easily found by exploiting the local gauge freedom to choose the $\boldsymbol{\theta}$ fields in (250) to vanish, as in the ansatz (229): that is, we set

$$\phi = \begin{pmatrix} 0 \\ (v+H(x))/\sqrt{2} \end{pmatrix}. \tag{254}$$

Substituting (254) into (252) and retaining only terms which are of second order in the fields (i.e. kinetic energies or mass terms) we find

$$\begin{aligned}
\mathcal{L}_{G\Phi} \;=\;& \frac{1}{2}\partial_\mu H\partial^\mu H - \mu^2 H^2 \\
& -\frac{1}{4}F_{1\mu\nu}F_1^{\mu\nu} + \frac{1}{8}g^2v^2W_{1\mu}W_1^\mu \\
& -\frac{1}{4}F_{2\mu\nu}F_2^{\mu\nu} + \frac{1}{8}g^2v^2W_{2\mu}W_2^\mu \\
& -\frac{1}{4}F_{3\mu\nu}F_3^{\mu\nu} - \frac{1}{4}G_{\mu\nu}G^{\mu\nu} + \frac{1}{8}v^2(gW_{3\mu} - g'B_\mu)(gW_3^\mu - g'B^\mu)\,.
\end{aligned} \tag{255}$$

The first line of (255) tells us that we have a scalar field of mass $\sqrt{2}\mu$ (the Higgs boson, again). The next two lines tell us that the components W_1 and W_2 of the triplet (W_1, W_2, W_3) acquire a mass

$$M_1 = M_2 = gv/2 \equiv M_W\,. \tag{256}$$

The last line shows us that the fields W_3 and B are mixed. But they can easily be unmixed by noting that the last term in (255) involves only the combination $gW_3 - g'B$, which evidently acquires a mass. This suggests introducing the linear combinations

$$Z^\mu = \cos\theta_W W_3^\mu - \sin\theta_W B^\mu \tag{257}$$

$$A^\mu = \sin\theta_W W_3^\mu + \cos\theta_W B^\mu \tag{258}$$

where

$$\cos\theta_W = g/(g^2+g'^2)^{1/2}, \quad \sin\theta_W = g'/(g^2+g'^2)^{1/2}\,. \tag{259}$$

We then find that the last line of (255) becomes

$$-\frac{1}{4}F_{Z\mu\nu}F_Z^{\mu\nu} + \frac{1}{8}v^2(g^2+g'^2)Z_\mu Z^\mu - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{260}$$

where

$$F_{Z\mu\nu} = \partial_\mu Z_\nu - \partial_\nu Z_\mu \quad \text{and} \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu\,. \tag{261}$$

Thus

$$M_Z = \frac{1}{2}v(g^2+g'^2)^{1/2} = M_W/\cos\theta_W \tag{262}$$

and

$$M_A = 0\,. \tag{263}$$

Counting degrees of freedom as in the local U(1) case, we originally had 12 in (252) - three massless W's and one massless B, which is 8 in all, together with 4 ϕ-fields. After symmetry breaking, we have three massive vector fields W_1, W_2 and Z making 9 degrees of freedom, one massless vector field A with 2, and one massive scalar H. Of course, the physical application will be to identify the W and Z fields with those physical particles, and the A field with the massless photon. In the gauge (254), the W and Z particles have propagators of the form (231).

The identification of A^μ with the photon field is made clearer if we look at the form of $D_\mu\phi$ written in terms of A_μ and Z_μ, discarding the W_1, W_2 pieces:-

$$D_\mu\phi = \left\{\partial_\mu + ig\sin\theta_W\left(\frac{1+\tau_3}{2}\right)A_\mu + \frac{ig}{\cos\theta_W}\left[\frac{\tau_3}{2} - \sin^2\theta_W\left(\frac{1+\tau_3}{2}\right)\right]Z_\mu\right\}\phi\,. \tag{264}$$

Now the operator $(1+\tau_3)$ acting on $\langle 0|\phi|0\rangle$ gives zero, as observed in (248), and this is why A_μ does not acquire a mass when $\langle 0|\phi|0\rangle \neq 0$ (gauge fields coupled to *unbroken* symmetries of $\langle 0|\phi|0\rangle$ do *not* become massive). Although certainly not unique, this choice of ϕ and $\langle 0|\phi|0\rangle$ (due to Weinberg (1967)) is undoubtedly very economical and natural. The zero eigenvalue of $(1+\tau_3)$ can be interpreted as the electromagnetic charge of the vacuum, which we would not wish to be non-zero. We would then tentatively expect the identification

$$e = g\sin\theta_W \tag{265}$$

in order to get the right 'electromagnetic D_μ' in (264).

We have at last assembled all the conceptual ingredients we need for the electroweak theory, to which we now turn.

Problems for Lecture 5

P5.1 Verify that the inverse of the bracket [...] in (193) is as given in (194).
P5.2 Let

$$\mathcal{L}_\phi^{(1)} = \partial_\mu\phi^\dagger\partial^\mu\phi + \mu^2\phi^\dagger\phi - \frac{\lambda}{4}(\phi^\dagger\phi)^2.$$

Set

$$\phi(x) = \frac{1}{\sqrt{2}}(v+h(x))e^{-i\theta(x)/v}.$$

Show that

$$\mathcal{L}_\phi^{(1)} = \frac{1}{2}\partial_\mu h\partial^\mu h - \mu^2h^2 + \frac{1}{2}\partial_\mu\theta\partial^\mu\theta + \frac{\mu^4}{\lambda} + \text{ non-quadratic terms}$$

(μ^4/λ is an irrelevant constant).
P5.3 Let

$$\mathcal{L}_h = [(\partial_\mu + ieA_\mu)\phi]^\dagger[(\partial^\mu + ieA^\mu)\phi] - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \mu^2\phi^\dagger\phi - \frac{\lambda}{4}(\phi^\dagger\phi)^2.$$

Set $\phi = 1/\sqrt{2}(v + h(x))$. Show that

$$\mathcal{L}_h = \frac{1}{2}\partial_\mu h \partial^\mu h - \mu^2 h^2 + \frac{\mu^4}{\lambda} - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}e^2 v^2 A_\mu A^\mu + \text{non-quadratic terms.}$$

So $m_A = ev$.
P5.4 Let

$$\mathcal{L}_\phi^{(2)} = (\partial\phi)^\dagger(\partial^\mu\phi) + \mu^2(\phi^\dagger\phi) - \frac{\lambda}{4}(\phi^\dagger\phi)^2$$

where $\phi = \binom{\phi^+}{\phi^0}$ and $(\phi^+)^\dagger = \phi^-, \ (\phi^0)^\dagger = \bar{\phi}^0$. Set

$$\phi = \exp(-i\boldsymbol{\theta}(x)\cdot\boldsymbol{\tau}/v)\begin{pmatrix} 0 \\ \frac{v+\sigma(x)}{\sqrt{2}} \end{pmatrix}.$$

Show that

$$\mathcal{L}_\phi^{(2)} = \frac{1}{2}\partial_\mu\sigma\partial^\mu\sigma - \mu^2\sigma^2 + \frac{1}{2}\partial_\mu\boldsymbol{\theta}\cdot\partial^\mu\boldsymbol{\theta} + \text{non-quadratic terms.}$$

7. THE ELECTROWEAK THEORY

See chapter 22 of [2].

We have seen that the original 4-fermion theory of weak interactions is non-renormalisable, and useful only at energies well below 100 GeV. Replacing the 4-fermion coupling by a Yukawa-like coupling to massive W's and Z's gave us a theory with a dimensionless coupling constant, but it was not renormalisable either. In fact, the only known way of getting a renormalisable theory of massive charged vector bosons is to regard them as gauge quanta of a spontaneously broken gauge theory. This necessitates the existence of a scalar field, the Higgs field, three of whose components correspond to the longitudinal components of the $W^\pm$ and Z^0, and the fourth of which survives as a scalar particle in the physical spectrum, but of unknown mass. In a sense, the mass of the Higgs boson m_H acts like a cut-off; but we shall see that there are quite persuasive reasons to think that at least the simplest Higgs sector model of section 6.5 does not make sense for m_H much beyond 500-1000 GeV.

7.1 The electroweak theory for one fermion family

So far, in section 6.5, we have only introduced the gauge and Higgs field sectors of the electroweak theory; we now need to include the quarks and leptons. Here the crucial new phenomenological input is that the weak interactions violate parity (while the electromagnetic ones of course do not). This means that the weak interaction is different for the left-handed components of fermion fields and for right-handed components. Electroweak interactions are described by a gauge theory based on a spontaneously broken local $\mathrm{SU(2)_L}\times\mathrm{U(1)}$ invariance. The 'L' means that the SU(2) part (with the gauge fields $\boldsymbol{W}^\mu$ of section 6.5) acts only on the left-handed parts ψ_L of fermion fields (see problem P4.1); it is therefore 'maximally' parity violating. The U(1) part (with the gauge field B^μ) acts on both right-(if any) and left-handed components, in such a way that the particular combination (258) conserves parity, as is required for the electromagnetic interaction; the other combination (257), which mediates neutral weak interactions, will turn out not to couple in the 'pure V$-$A' form, as is indeed observed. The simplest structure allowing connection between the parity violating weak force and the parity conserving e-m one is the $\mathrm{SU(2)_L}\times\mathrm{U(1)}$ one, originally proposed by Glashow [9], with brave disregard for the non-renormalisability problem. The $\mathrm{SU(2)_L}$ part is often called 'weak isospin' and the U(1) 'weak hypercharge'.

In this theory, the basic fields are fermions (leptons and quarks), gauge bosons, and Higgs fields. The left-handed parts of the fermion fields form (weak isospin) doublets under $\mathrm{SU(2)_L}$

$$\text{`}\psi_{\mathrm{L}}\text{'} = \begin{pmatrix}\nu_e \\ e^-\end{pmatrix}_{\mathrm{L}}, \quad \begin{pmatrix}\nu_\mu \\ \mu^-\end{pmatrix}_{\mathrm{L}}, \quad \begin{pmatrix}\nu_\tau \\ \tau^-\end{pmatrix}_{\mathrm{L}}, \quad \begin{pmatrix}u \\ \tilde{d}\end{pmatrix}_{\mathrm{L}}, \quad \begin{pmatrix}c \\ \tilde{s}\end{pmatrix}_{\mathrm{L}}, \quad \begin{pmatrix}t \\ \tilde{b}\end{pmatrix}_{\mathrm{L}}, \tag{266}$$

where the ˜ denotes states which are mixed with respect to the strong interaction states d, s and b (see the following section, and note that the colour labels will be suppressed throughout), while the right-handed components are $\mathrm{SU(2)_L}$ singlets

$$`\psi_{\mathrm{R}}\text{'} = e^-_{\mathrm{R}},\ \ \mu^-_{\mathrm{R}}, \ldots, \tag{267}$$

where for simplicity we shall generally assume in this section that the neutrinos are massless (see also section 7.2). We shall confine the discussion in the present section to just one 'family', comprising ν_e, e^-, u and d (which should really be $\tilde{d}$ but we are ignoring mixing for the moment).

The Lagrangian can be looked at in many ways, but we shall write it as

$$\mathcal{L} = \mathcal{L}_{\mathrm{S}} + \mathcal{L}_{\mathrm{SB}} \tag{268}$$

where S stands for 'symmetrical' under SU(2)×U(1) and SB stands for 'symmetry breaking'. In $\mathcal{L}_{\mathrm{S}}$ we have a gauge invariant Lagrangian $\mathcal{L}_{\mathrm{f}}$ describing the interactions of the fermions with the $\boldsymbol{W}$ and B fields, together with the SU(2) Yang-Mills Lagrangian $\mathcal{L}_W$ (179) for the $\boldsymbol{W}$ fields and the U(1) Lagrangian $\mathcal{L}_B$ for the B field as in (252); in $\mathcal{L}_{\mathrm{SB}}$ we will have everything involving the Higgs fields. In section 4.2 we learned how to construct a locally SU(2) invariant gauge theory with a fermion doublet (see (163)). The difference now is that we want the $\mathrm{SU(2)_L}$ to act only on the L-component of the doublet. However, there is no problem with this for *massless* fields: (218) shows us that the 'kinetic' operator $\not\partial$ does not mix L and R components, and hence there is no objection to 'gauging' each of them differently (i.e. using a different $\not D$ on ψ_{L} and on ψ_{R}). On the other hand, (219) shows that this is *not* true for the mass terms - a difficulty we will deal with shortly by getting the mass terms from $\mathcal{L}_{\mathrm{SB}}$. First, we simply state that the appropriate D's are in fact

$$D_\mu = \partial_\mu + ig\boldsymbol{\tau}\cdot\boldsymbol{W}_\mu/2 + ig'yB_\mu/2 \quad \text{on } \psi_{\mathrm{L}}\text{'s} \tag{269}$$

and

$$D_\mu = \partial_\mu + ig'yB_\mu/2 \quad \text{on } \psi_{\mathrm{R}}\text{'s}, \tag{270}$$

where the condition

$$Q = \tau_3/2 + y/2 \tag{271}$$

is imposed, Q being the electric charge in units of e (the positron charge). The factor of $\frac{1}{2}$ in the B-term of (269) is conventional, but (271) fixes the normalisation of the coupling g'. The eigenvalues of the $\tau_3/2$ operator in (269) are as indicated by the placings in (266): namely $+\frac{1}{2}$ for $(\nu_e, \nu_\mu, \nu_\tau, u, c, s)_{\mathrm{L}}$ and $-\frac{1}{2}$ for e^-_{L}, etc. For the $(\text{lepton})_{\mathrm{L}}$ fields the y eigenvalue is -1, while for the $(\text{quark})_{\mathrm{L}}$ fields it is $+\frac{1}{3}$; for the R-fields y is just $2Q$ since the $\tau_3/2$ eigenvalue is zero.

The gauge invariant Lagrangian $\mathcal{L}_{\mathrm{f}}$ (for massless fermions) is therefore

$$\mathcal{L}_{\mathrm{f}} = \bar{\ell}_{e\mathrm{L}} i\not D \ell_{e\mathrm{L}} + \bar{q}_{\mathrm{L}} i\not D q_{\mathrm{L}} + \bar{e}_{\mathrm{R}} i\not D e_{\mathrm{R}} + \bar{u}_{\mathrm{R}} i\not D u_{\mathrm{R}} + \bar{d}_{\mathrm{R}} i\not D d_{\mathrm{R}} \tag{272}$$

where

$$\ell_{e\mathrm{L}} = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_{\mathrm{L}}, \quad q_{\mathrm{L}} = \begin{pmatrix} u \\ d \end{pmatrix}_{\mathrm{L}} \tag{273}$$

and a $\nu_{e\mathrm{R}}$ term can be added to (272) if desired. From (272) we can already read off the couplings of the charged W's to the fermions (the W_3 and B will mix, as we saw in section 6.5). The correct normalisation for charged fields is that $W^\mu = (W_1 - iW_2)/\sqrt{2}$ destroys the W^+ or creates W^-, so that the $\boldsymbol{\tau}\cdot\boldsymbol{W}/2$ terms are

$$\frac{1}{\sqrt{2}}\left\{\tau_+ \frac{(W_1 - iW_2)}{\sqrt{2}} + \tau_- \frac{(W_1 + iW_2)}{\sqrt{2}}\right\} + \tau_3 \frac{W_3}{2} \tag{274}$$

where $\tau_\pm = (\tau_1 \pm i\tau_2)/2$ are the raising and lowering operators for the doublet. Thus the first term in (274) picks out the process $e^- \to \nu_e W^-$ for example, with the result that the corresponding vertex is

$$-\frac{ig}{\sqrt{2}}\gamma_\mu \frac{(1-\gamma_5)}{2}\,, \tag{275}$$

and similarly for the quarks (if unmixed), and other families. Hence we can immediately make a connection with the original V-A Fermi theory of these charged current processes, namely

$$G_{\mathrm{F}}/\sqrt{2} = g^2/8M_W^2\,. \tag{276}$$

Although the quark couplings can also be read off from (272), they are unphysical at this stage since mixing has not yet been introduced.

There are also couplings of the Z^0 to fermions. To find these, we need to rewrite the neutral part of the D's in (269) and (270) in terms of the Z and A fields defined in (257) and (258) (cf (264)). We find

$$D_\mu(\mathrm{neutral}) = \partial_\mu + ieQA_\mu + \frac{igZ_\mu}{2\cos\theta_W}(v_f - a_f\gamma_5) \tag{277}$$

where

$$v_f = \frac{\tau_3}{2} - 2Q\sin^2\theta_W \tag{278}$$

and

$$a_f = \frac{\tau_3}{2}\,. \tag{279}$$

We see that, as remarked earlier, the Z (or 'neutral-current') coupling is not pure V$-$A. The Z-couplings analogous to (275) are therefore

$$\frac{-ig}{2\cos\theta_W}\gamma_\mu(v_f - a_f\gamma_5)\,. \tag{280}$$

(280) is the coupling observed around the Z^0 peak.

Finally, we may write effective four-fermion interactions (valid for energies much less than M_W, M_Z) as

$$\frac{G_{\mathrm{F}}}{\sqrt{2}} j^C_{\mu+}\, j^{C\mu}_- \tag{281}$$

for the charged current processes, with

$$j^C_{\mu\pm} = (\overline{\psi}_2\gamma_\mu(1-\gamma_5)\tau_\pm\psi_1)\,, \tag{282}$$

and as

$$\sqrt{2}G_{\mathrm{F}}\rho\, j^N_\mu\, j^{N\mu} \tag{283}$$

for the neutral current processes, where

$$j^N_\mu = \overline{\psi}_f\gamma_\mu(v_f - a_f\gamma_5)\psi_f \tag{284}$$

and the quantity

$$\rho = M_W^2/M_Z^2\cos^2\theta_W \tag{285}$$

has the value 1 in the Standard Model, at tree level.

The vector boson masses arise through symmetry breakdown via the Higgs sector, in the standard model, as discussed in section 6.5. After spontaneous symmetry breaking, we have

$$\begin{aligned} M_W &= gv/2 = \cos\theta_W M_Z && (286)\\ \cos\theta_W &= g/(g^2 + g'^{\,2})^{1/2} && (287)\\ e &= g\sin\theta_W && (288)\\ m_H &= \sqrt{2}\mu && (289) \end{aligned}$$

in terms of the fundamental coupling parameters g, g' of the SU(2)$\times$U(1) gauge group, and the parameters v and μ of the Higgs potential. There is also the low-energy connection (276), which we can write as

$$\frac{v}{\sqrt{2}} = 2^{-3/4} G_{\rm F}^{-1/2} = 174.1 \text{ GeV} \,, \tag{290}$$

using $G_{\rm F} \approx 1.17 \times 10^{-5} \text{GeV}^{-2}$. This gives us the scale of $\langle 0|\phi|0\rangle$, for which as yet there is no theoretical explanation. We may also write (276) as

$$M_W = (\pi\alpha/\sqrt{2}\, G_{\rm F})^{1/2} / \sin\theta_W \tag{291}$$

$$= 37.2802 \text{ GeV} / \sin\theta_W \tag{292}$$

using the conventional low-energy value of α. Note that all the above relations are between parameters in the Lagrangian, and hold at the tree level only; they can be changed by loop corrections (see section 7.4).

We must now consider how to bring fermion masses into this theory. We begin by noting, again, that a typical Dirac mass term has the form (219), which is clearly not invariant under transformations which treat $\psi_{\rm L}$ and $\psi_{\rm R}$ differently. Would it matter if we just added in such a mass term? The answer is that if we did this the theory would, once again, not be renormalisable. And, once again, we can arrange for the fermions to 'acquire mass spontaneously', this time via couplings of the generic 'Yukawa' type $g_f \bar{\psi}\psi\phi$. This can be made SU(2)$_{\rm L}\times$U(1) invariant, and then if the scalar field acquires a vacuum value v we have a mass term (in such a vacuum) equal to $g_f v$. Some such treatment of fermion masses is necessary for the theory to make sense much beyond the $W - Z$ mass range.

It is obviously most economical if we can 'blame' fermion masses on the same Higgs field that generates the W and Z masses, but it must be recognised that the Yukawa coupling 'mechanism' is on a very different footing from the symmetry-inspired gauge couplings - at least in the absence of any further symmetry that might relate these two types of coupling. At any rate, consider the case of the ν_e, e^- doublet, in the simple case that the ν_e is massless, with a Yukawa coupling between these fields and the standard doublet Higgs, of the type

$$-g_e(\overline{\ell}_{e\rm L}\phi e_{\rm R} + \overline{e}_{\rm R}\phi^\dagger \ell_{e\rm L}) \,. \tag{293}$$

Remembering that $e_{\rm R}$ is an SU(2) scalar, we see that (293) is Lorentz invariant, and invariant under global SU(2) transformations (because $\bar{\ell}\phi$ and $\phi^\dagger\ell$ are invariant); it is also invariant under U(1)$_y$ transformations, with the y assignments made after (271), if $y(\phi) = 1$ (which is what we actually assumed in (253)). In fact, since no derivatives are involved in (293), it is also invariant under local SU(2)$\times$U(1) transformations. But the Higgs sector contains the potential $V(\phi)$ of (239), which 'triggers' spontaneous symmetry breaking. The vacuum value (246) for ϕ when inserted into (293), yields

$$-(g_e v/\sqrt{2})(\overline{e}_{\rm L} e_{\rm R} + \overline{e}_{\rm R} e_{\rm L}) \tag{294}$$

which is precisely a mass term for the electron if we identify

$$g_e = m_e\sqrt{2}/v \,. \tag{295}$$

When oscillations about this vacuum are considered, in the simple gauge of (254), one easily finds that the H-field couples to the electron with a vertex

$$-im_e/v \,. \tag{296}$$

Sure enough, the coupling is proportional to the electron mass - and on dimensional grounds to v^{-1}.

It might seem from the foregoing that only a mass for the $t_3 = -\frac{1}{2}$ component of the fermion doublets could be generated this way, because of the form of $\langle 0|\phi|0\rangle$. Remarkably enough, however, the same Higgs field can also provide a mass for the $t_3 = +\frac{1}{2}$ component (and this is of course necessary for the quarks, if not for the neutrinos). It can be shown that the field ϕ_c defined by

$$\phi_c = i\tau_2\phi^* = \begin{pmatrix} (\phi_3 + i\phi_4)/\sqrt{2} \\ -(\phi_1 + i\phi_2)/\sqrt{2} \end{pmatrix} = \begin{pmatrix} \overline{\phi}^0 \\ -\phi^- \end{pmatrix}, \tag{297}$$

where (243) has been used, is also an isodoublet. (The notation in (297) is reminiscent of the K-meson doublet $(\bar{K}^0, K^-)$; alternatively, we may think of a quark isospin doublet like $\begin{pmatrix} u \\ d \end{pmatrix}$ and its conjugate doublet $\begin{pmatrix} \bar{d} \\ -\bar{u} \end{pmatrix}$, with the $I = 0$ combination being $(\bar{d}d - \bar{u}u)$. With the help of ϕ_c we can write down another gauge invariant coupling in the ν_e-e sector, namely

$$-g_{\nu_e}\left(\overline{\ell}_{e\mathrm{L}}\phi_c\nu_{e\mathrm{R}} + \overline{\nu}_{e\mathrm{R}}\phi_c^\dagger\ell_{e\mathrm{L}}\right) \tag{298}$$

which produces

$$-\left(g_{\nu_e}v/\sqrt{2}\right)\left(\overline{\nu}_{e\mathrm{L}}\nu_{e\mathrm{R}} + \overline{\nu}_{e\mathrm{R}}\nu_{e\mathrm{L}}\right) \tag{299}$$

in the Higgs vacuum (246), which is a neutrino mass term (if required) provided $g_{\nu_e} = \sqrt{2}m_{\nu_e}/v$. Once again, the H-field will couple with an amplitude of the form (296), with $m_e \to m_{\nu_e}$. The procedure can obviously be repeated for the u and d quarks.

It is clearly possible to go on like this, and arrange for as many fermion families to have a mass as is required - and we will look at this a little more closely in the next section. However, one must note that the theory does no more than accommodate itself to the mass difficulty: in no sense do the fermion masses ‘come out’ of the theory, since each has simply to be inserted by hand via a new Yukawa coupling. In essence, these Yukawa couplings are *not* gauge interactions, and hence not universal.

The Higgs coupling to fermions can now be written generally as

$$-iem_f/2\sin\theta_W M_W\,. \tag{300}$$

There are also trilinear and quadrilinear Higgs self-couplings arising from the $\lambda(\phi^\dagger\phi)^2$ term in (252). Recalling that $\lambda = 4\mu^2/v^2$ and that $m_H = \sqrt{2}\mu$, we can write the trilinear coupling as

$$-i3m_H^2 e/8M_W\sin\theta_W \tag{301}$$

and the quadrilinear as

$$-i3m_H^2 e^2/16M_W^2\sin^2\theta_W\,. \tag{302}$$

There are also the trilinear H-W^+-W^-

$$ieM_W g_{\lambda\mu}/\sin\theta_W \tag{303}$$

and H-Z-Z

$$i2eM_Z g_{\lambda\mu}/\sin 2\theta_W \tag{304}$$

couplings, together with quadrilinear ϕ^2W^2, ϕ^2Z^2 couplings which we shall not give here. Note that all these couplings are determined by the existing set of parameters—and, in particular, that the Higgs couples most strongly to the heaviest particles, so that decays to heavy channels offer the largest rates.

7.2 The three-family model

We now extend the preceding discussion to the three family case, which will involve the important subjects of quark flavour mixing in charged current processes (and of no mixing - the GIM mechanism (Glashow et al [10]) - in neutral current processes), and CP violation. We shall here assume that there are just three families. We introduce three doublets of left handed fields

$$q_{\mathrm{L}1} = \begin{pmatrix} u_{\mathrm{L}1} \\ d_{\mathrm{L}1} \end{pmatrix}, \quad q_{\mathrm{L}2} = \begin{pmatrix} u_{\mathrm{L}2} \\ d_{\mathrm{L}2} \end{pmatrix}, \quad q_{\mathrm{L}3} = \begin{pmatrix} u_{\mathrm{L}3} \\ d_{\mathrm{L}3} \end{pmatrix} \tag{305}$$

and the corresponding six singlets

$$u_{\mathrm{R}1}, \quad d_{\mathrm{R}1}, \quad u_{\mathrm{R}2}, \quad d_{\mathrm{R}2}, \quad u_{\mathrm{R}3}, \quad d_{\mathrm{R}3}\,, \tag{306}$$

which transform in the now familiar way under SU(2)$_{\mathrm{L}}\times$U(1). The u-fields correspond to the $t_3 = +\frac{1}{2}$ components of SU(2)$_{\mathrm{L}}$, the d ones to the $t_3 = -\frac{1}{2}$ components, and to their 'R' partners. The labels 1, 2, and 3 refer to the family number; for example, with no mixing at all, $u_{\mathrm{L}1} = u_{\mathrm{L}}$, $d_{\mathrm{L}1} = d_{\mathrm{L}}$, etc. (We are thinking of (305) and (306) as quark fields, but the discussion will be quite general and could just as well apply to leptons if they should need mixing too - we return to leptons later). We have to consider what is the most general SU(2)$_{\mathrm{L}}\times$U(1)-invariant interaction between the Higgs field (assuming we can still get by with only one) and these various fields. Apart from the symmetry, the only other theoretical requirement is renormalisability - for, after all, if we drop this we might as well abandon the whole motivation for the 'gauge' concept. This implies (as in the discussion of the Higgs potential V) that we cannot have terms like $(\bar{\psi}\psi\phi)^2$ appearing - which would have a coupling with dimensions (mass)$^{-4}$ and would be non-renormalisable. In fact the only renormalisable Yukawa coupling is of the form '$\bar{\psi}\psi\phi$', which has a dimensionless coupling (as in the g_e and g_{ν_e} of (293) and (298)). However, there is no *a priori* requirement for it to be 'diagonal' in the weak interaction family index i. The allowed generalisation of (293) and (298) is therefore an interaction of the form (summing on repeated indices)

$$\mathcal{L}_{\psi\phi} = a_{ij}\bar{q}_{\mathrm{L}i}\phi^c u_{\mathrm{R}j} + b_{ij}\bar{q}_{\mathrm{L}i}\phi d_{\mathrm{R}j} + \mathrm{h.c.} \tag{307}$$

where

$$q_{\mathrm{L}i} = \begin{pmatrix} u_{\mathrm{L}i} \\ d_{\mathrm{L}i} \end{pmatrix} \tag{308}$$

and a sum on the family indices i and j (from 1 to 3) in (307) is assumed. After symmetry breaking, using the gauge (254), we find

$$\mathcal{L}_{f\phi} = -\left(1 + \frac{H}{v}\right)\left[\bar{u}_{\mathrm{L}i} m^u_{ij} u_{\mathrm{R}j} + \bar{d}_{\mathrm{L}i} m^d_{ij} d_{\mathrm{R}j} + \mathrm{h.c.}\right] \tag{309}$$

where the 'mass matrices' are

$$m^u_{ij} = -\frac{v}{\sqrt{2}}\, a_{ij}, \quad m^d_{ij} = -\frac{v}{\sqrt{2}}\, b_{ij}\,. \tag{310}$$

Although we have not indicated it, the m^u and m^d matrices could involve a 'γ_5' part as well as a '1' part in Dirac space. It can be shown (Weinberg [11], Feinberg et al [12]) that m^u and m^d can both be made Hermitean, γ_5-free, and diagonal by making four separate unitary transformations on the 'family triplets'

$$u_{\mathrm{L}} = \begin{pmatrix} u_{\mathrm{L}1} \\ u_{\mathrm{L}2} \\ u_{\mathrm{L}3} \end{pmatrix}, \quad d_{\mathrm{L}} = \begin{pmatrix} d_{\mathrm{L}1} \\ d_{\mathrm{L}2} \\ d_{\mathrm{L}3} \end{pmatrix}, \; etc. \tag{311}$$

via

$$u_{\mathrm{L}\alpha} = \left(U^{(u)}_{\mathrm{L}}\right)_{\alpha i} u_{\mathrm{L}i}, \quad u_{\mathrm{R}\alpha} = \left(U^{(u)}_{\mathrm{R}}\right)_{\alpha i} u_{\mathrm{R}i}, \tag{312}$$

$$d_{\mathrm{L}\alpha} = \left(U_{\mathrm{L}}^{(d)}\right)_{\alpha i} d_{\mathrm{L}i}, \quad d_{\mathrm{R}\alpha} = \left(U_{\mathrm{R}}^{(d)}\right)_{\alpha i} d_{\mathrm{R}i} \tag{313}$$

In this notation, 'α' is the index of the 'mass diagonal' basis, and 'i' is the 'weak interaction' basis. Then (309) becomes

$$\mathcal{L}_{q\psi} = -\left(1 + \frac{H}{v}\right)\left[m_u \bar{u}u + \cdots + m_b \bar{b}b\right] \ . \tag{314}$$

Rather remarkably, we can still manage with only the one Higgs field. It couples to each fermion with a strength proportional to the mass of that fermion, divided by M_W.

Now consider the $\mathrm{SU(2)_L} \times \mathrm{U(1)}$ gauge invariant interaction part of the Lagrangian. Written out in terms of the 'weak interaction' fields $u_{\mathrm{L,R}\ i}$ and $d_{\mathrm{L,R}\ i}$ (cf (269) and (270)), it is

$$\begin{aligned} \mathcal{L}_{fW,B} &= i\,(\bar{u}_{\mathrm{L}j}, \bar{d}_{\mathrm{L}j})\,\gamma^\mu(\partial_\mu + ig\boldsymbol{\tau}\cdot\boldsymbol{W}_\mu/2 + ig'yB_\mu/2)\begin{pmatrix} u_{\mathrm{L}j} \\ d_{\mathrm{L}j} \end{pmatrix} \\ &\quad + i\bar{u}_{\mathrm{R}j}\gamma^\mu(\partial_\mu + ig'yB_\mu/2)u_{\mathrm{R}j} + i\bar{d}_{\mathrm{R}j}\gamma^\mu(\partial_\mu + ig'yB_\mu/2)d_{\mathrm{R}j} \end{aligned} \tag{315}$$

where a sum on j is understood. This now has to be rewritten in terms of the mass-eigenstates $u_{\mathrm{L,R}\ \alpha}$ and $d_{\mathrm{L,R}\ \alpha}$.

Problem P6.1 shows that the neutral current part of (315) is diagonal in the mass basis - that is, the neutral current interactions do not change the flavour of the physical (mass eigenstates) quarks. The charged current processes, however, involve the *non*-diagonal matrices τ_1 and τ_2 in (315), and this spoils the argument used in problem P6.1. Indeed, using (274) we find that the charged current piece is

$$\begin{aligned} \mathcal{L}_{\mathrm{cc}} &= -\frac{g}{\sqrt{2}}(\bar{u}_{\mathrm{L}j}, d_{\mathrm{L}j})\gamma_\mu\tau_+W_\mu\begin{pmatrix} u_{\mathrm{L}j} \\ d_{\mathrm{L}j} \end{pmatrix} + \mathrm{h.c.} \\ &= -\frac{g}{\sqrt{2}}\bar{u}_{\mathrm{L}j}\gamma^\mu d_{\mathrm{L}j}W_\mu + \mathrm{h.c.} \\ &= -\frac{g}{\sqrt{2}}\bar{u}_{\mathrm{L}\alpha}\left[\left(U_{\mathrm{L}}^{(u)\dagger}\right)_{\alpha i}\left(U_{\mathrm{L}}^{(d)}\right)_{i\beta}\right]\gamma^\mu d_{\mathrm{L}\beta}W_\mu + \mathrm{h.c.} \end{aligned} \tag{316}$$

where the matrix

$$V_{\alpha\beta} \equiv \left[U_{\mathrm{L}}^{(u)\dagger}U_{\mathrm{L}}^{(d)}\right]_{\alpha\beta} \tag{317}$$

is not diagonal, though it is unitary. V therefore has 9 real parameters, which can be reduced to 4 - three 'rotational angles' and one phase - by redefinitions of the quark fields (Jarlskog [13]). This is the famous CKM matrix, (Cabibbo [14], Kobayashi and Maskawa [15]) the interaction (316) having the form

$$-\frac{g}{\sqrt{2}}W_\mu\,(\bar{u}_{\mathrm{L}}\ \bar{c}_{\mathrm{L}}\ \bar{t}_{\mathrm{L}})\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}\begin{pmatrix} d_{\mathrm{L}} \\ s_{\mathrm{L}} \\ b_{\mathrm{L}} \end{pmatrix} + \mathrm{h.c.} \tag{318}$$

The entries in the V-matrix modify the vertex (275) in an obvious way. The single phase δ in the V-matrix accommodates CP-violation. In the case of only two flavours, V has only 1 real parameter, which is the Cabibbo angle, and there is no freedom to have a CP violation phase in the family mixing matrix. It is an important challenge to experiment to find out whether all CP-violating phenomena can be described with just this one parameter δ in the CKM matrix (see the lectures on CP violation).

Returning finally to the leptons, all of the above will apply (with three more mixing angles and one more phase) if the neutrinos do in fact have a mass. We would then have leptonic flavour mixing in c.c. processes, involving a term of the form $[\bar{\ell}_{\mathrm{L}}V_\ell\gamma^\mu\ell'_{\mathrm{L}}W_\mu + \mathrm{h.c.}]$ (cf (316)), and lepton mass terms $[\bar{\ell}_{\mathrm{L}}m_\ell\ell_{\mathrm{R}} + \mathrm{h.c.}]$ and $[\bar{\ell}'_{\mathrm{L}}m'_\ell\ell'_{\mathrm{R}} + \mathrm{h.c.}]$, where V_ℓ is the leptonic analogue of (317), and m_ℓ, m'_ℓ are the analogues of the quark masses. There is nothing in the standard model that requires the neutrinos to be massless, and indeed the experimental data now imply that more than one is not; in GUTs they generally do have (small) masses - see the lectures on neutrinos.

7.3 One remark about the Higgs sector

The Higgs sector is the one big unknown still hanging over the Standard Model, starting with the question: what is the Higgs mass? There is an interesting theoretical argument here which is worth a mention.

We first note that, for a given vacuum value v as in (290), the Higgs mass is (cf (241) and (251))

$$m_H = v\lambda^{\frac{1}{2}}/\sqrt{2} \sim \lambda^{\frac{1}{2}} \times 174 \text{ GeV}. \tag{319}$$

Now λ is a dimensionless constant: if it is $O(\alpha)$ we would say that the theory is perturbative, while if it is $O(1)$ we would say it was strongly coupled. It is clear from (319), and the present experimental lower bounds on m_H, that we are already not far from the strongly coupled region. But we can ask: can λ (the renormalised coupling) take *any* value at all? That is, can m_H (for fixed v) be arbitrarily large?

To answer this we must recall that, in a renormalisable theory, 'the' value of λ has to be defined at a certain scale, and the value at another scale is different (i.e. λ 'runs'). For the interaction (239), calculation shows that the analogue of (129) is

$$\lambda(E) = \lambda \Big/ \left[1 - \frac{3}{8\pi^2}\lambda \ln\left(\frac{E}{v}\right)\right]. \tag{320}$$

taking the 'physical' λ to be defined at the scale v. Note that this theory, like QED, is *not* asymptotically free. It follows from (320) that the theory breaks down (or, more conservatively, $\lambda(E)$ becomes so large that all perturbative expectations are useless) at an energy E^* such that $E^* \sim v\exp(\frac{8\pi^2}{3\lambda})$. But, for given v, we also have from (319) that λ is related to m_H. So the theory breaks down at

$$E^* \sim v\exp(\frac{4\pi^2 v^2}{3m_H^2}). \tag{321}$$

This is a very remarkable formula, because it is exponentially sensitive to the unknown m_H - and it is particularly interesting that the Higgs mass is in the denominator of the exponent. For 'small' m_H the breakdown scale is high - e.g. for $m_H \sim 150$ Gev, $E^* \sim 6 \times 10^{17}$ GeV. But for $m_H \sim 700$ GeV, E^* is already as low as 1 TeV. Clearly, at such a value of m_H, the Higgs mass is essentially equal to the 'breakdown scale' itself, and m_H cannot get any higher without new physics intervening in one form or another: maybe non-perturbative phenomena, or maybe supersymmetry.

7.4 Two remarks on one-loop corrections in the Standard Model

The precision of LEP and other data (of order 0.1%) was such that the measurements were sensitive to one-loop effects - and the very high quality of the fits to all the data confirm the presence of these corrections very convincingly. What is particularly interesting is that the loop corrections could be used to make *predictions* about as yet unseen particles: for example, the top quark mass was predicted to be something like 175 ± 10 GeV via its virtual effects in loops, *before* it was discovered as a real particle! (and the errors on the experimental mass determination were similar!). A typical fit to all data (Grünewald [16]) has a $\chi^2/\text{d.o.f}$ of 14.9/15, corresponding to a probability of 46%. This extremely strong numerical consistency lends impressive support to the belief that we are indeed dealing with a renormalisable spontaneously broken gauge theory, because *no extra parameters, not in the original Lagrangian, have had to be introduced*. In fact, one can turn this around. It is widely believed that, remarkably successful as it is, the Standard Model is not the end of physics, and that consequently further parameters will be required at some stage. The close agreement between the data and the existing Standard Model means that the new physics is proving very hard to see, at present energies.

As we have seen, we obtain cut-off independent results from loop corrections in a renormalisable theory by taking certain parameters (those appearing in the original Lagrangian) from experiment. In the electroweak case, it is usual to take the set

$$\alpha,\ G_F,\ m_Z,\ m_H,\ m_f, \text{parameters of mixing matrices;} \tag{322}$$

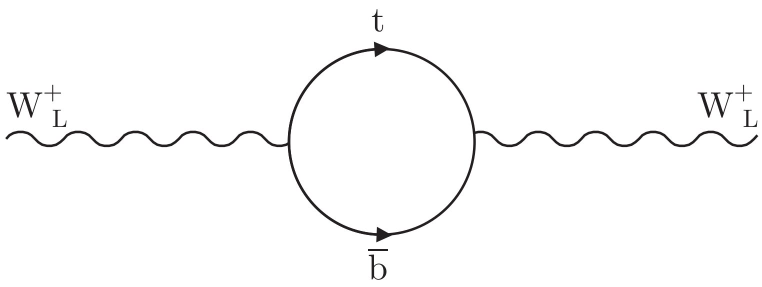

Fig. 17: $t - \bar{b}$ vacuum polarisation loop.

(α_s of QCD and the QCD θ-parameter need to be added for the full Standard Model). After renormalisation, one can derive radiatively-corrected values for physical quantities in terms of the set (322). For example, the tree-level relation (291) takes the following form at one loop:

$$M_W^2 = \left[(\pi\alpha/\sqrt{2}\,G_{\rm F})/\sin^2\theta_W\right]/(1-\Delta r) \tag{323}$$

where $\sin\theta_W$ has been *defined* as $\sin^2\theta_W \equiv 1 - M_W^2/M_Z^2$. Δr is the one-loop correction.

We cannot go into all the details of Δr, but we do want to focus on two important features of the result (which are typical of other radiatively-corrected formulae). The leading terms in Δr have the form

$$\Delta r = \Delta\alpha - \cot^2\theta_W\,\Delta\rho + (\Delta r)_{\rm rem}\,. \tag{324}$$

In (324), $\Delta\alpha$ is precisely the quantity $\bar{\Pi}_\gamma^2(M_Z^2)$ which entered into the running QED constant α discussed in Section 5.3 (see (127) and after (129)). $\Delta\rho$ is given by

$$\Delta\rho = \frac{3G_{\rm F}(m_t^2 - m_b^2)}{8\pi^2\sqrt{2}}\,, \tag{325}$$

while the 'remainder' $(\Delta r)_{\rm rem}$ contains a non-negligible term proportional to $\ln(m_t/M_Z)$, and a contribution from the Higgs boson which is (for $m_H \gg M_W$)

$$(\Delta r)_{{\rm rem},H} \approx \frac{\sqrt{2}\,G_F M_W^2}{16\pi^2}\frac{11}{3}\left[\ln\left(\frac{m_H^2}{M_W^2}\right) - \frac{5}{6}\right]. \tag{326}$$

The running of α is no surprise, but (325) and (326) contain unexpected features.

As regards (325), it is associated with top-bottom quark loops in vacuum polarisation amplitudes, of the kind discussed for $\bar{\Pi}_\gamma^{[2]}$, but in weak boson propagators. In the QED case, referring to (125) we see that the contribution of very heavy fermions (e.g. the top) in a vacuum polarisation loop should be suppressed, appearing as '$O(q^2/m_t^2)$'. This seems plausible enough: after all, the mass appears in the fermion propagator and hence in the denominator of the loop integral expression. Yet in fact m_f^2 appears in the *numerator* of (325)! the usual case ($\sim q^2/m^2$) is termed 'decoupling' of heavy matter, and it is certainly what we'd expect intuitively; in (325) we have 'non-decoupling'.

We can understand the appearance of the fermion masses (squared) in the numerator as follows. The shift $\Delta\rho$ is associated with vector boson vacuum polarisation contributions, for example the one shown in figure 16. Consider in particular the contribution from the longitudinal polarisation components of the W's. As we have seen, these components are nothing but three of the four Higgs components which the $W^\pm$ and Z^0 'swallowed' to become massive. But the couplings of these 'swallowed' Higgs fields to fermions are determined by just the same Higgs-fermion Yukawa couplings as we introduced to generate the fermion masses via spontaneous symmetry breaking. Hence we expect the fermion loops to

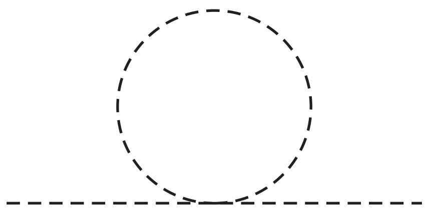

Fig. 18: One-loop self-energy graph in ϕ^4 theory.

contribute (to these longitudinal W states) something of order $g_f^2/4\pi$ where g_f is the Yukawa coupling. Since $g_f \sim m_f/v$ (see (295)) we arrive at an estimate $\sim m_f^2/4\pi v^2 \sim G_{\rm F} m_f^2/4\pi$ as in (325). An important message is that particles whose mass is proportional to their coupling to some field (ie in this case the Higgs field) do not 'decouple'.

But we still have to explain why $\Delta\rho$ vanishes if $m_t = m_b$. This has to do with a further symmetry of the assumed Higgs sector. As the notation suggests, $\Delta\rho$ is a leading order correction to the ρ parameter introduced in (283) and (285). At tree level, ρ has the value 1, which is a reflection of the fact that the (mass)^2 matrix, in terms of the original SU(2)$_{\rm L}\times$U(1) fields $\boldsymbol{W}^\mu$ and B^μ was (cf (255))

$$\frac{v^2}{4}\begin{pmatrix} g^2 & 0 & 0 & 0 \\ 0 & g^2 & 0 & 0 \\ 0 & 0 & g^2 & -gg' \\ 0 & 0 & -gg' & g'^2 \end{pmatrix} \tag{327}$$

acting in the $(W_1^\mu\ W_2^\mu\ W_3^\mu\ B^\mu)$ space. Notice now that the leading 3×3 block of this matrix, acting on the $\boldsymbol{W}$'s alone, is proportional to the unit matrix. This would be the natural consequence of an unbroken SU(2) symmetry in which the $\boldsymbol{W}$'s form an SU(2) triplet. Now, with the doublet Higgs of the form (243), it is a striking fact that the Higgs potential only involves the (globally) SO(4)-symmetric combination

$$\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^3. \tag{328}$$

The vacuum expectation value (246) singles out one of the four components, and breaks the SO(4) symmetry of the Higgs sector down to an SO(3), which is equivalent to the SU(2) of the $\boldsymbol{W}$'s, above. This (global) symmetry is called the 'custodial symmetry' of the (assumed) Higgs sector. It is this symmetry, in fact, that guarantees $\rho = 1$ *to all orders*.

However, examination of the behaviour of the quark mass terms under such global SU(2) transformations shows that the symmetry is explicitly broken by a difference in the masses of two quarks in the same doublet. This explains the '$m_t^2 - m_b^2$' dependence of the non-decoupled $t - \bar{b}$ loop correction. Phenomenologically this m_t^2 dependence was of great importance, because of course it meant that (paradoxically!) the heavier the top was, the more visible its effect in such loops would be. Its 'virtual' discovery was a wonderful cooperative achievement between theory and experiment.

The case is unfortunately 'reversed', in a sense, for the Higgs - and this is our second remark about loops. Without the Higgs particle, the Standard Model is non-renormalisable, and hence one might expect to see some radiative correction becoming large $O(m_H^2)$ as one tried to 'banish' the Higgs from the theory by sending $m_H \to \infty$ (m_H would be acting like a cut-off Λ). The reason is that in such a 'ϕ^4' theory, the simplest loop we meet is that shown in figure 17, and it is easy to see by counting powers as usual that it diverges as the square of the cut-off.

However, even without a Higgs contribution it turns out that the theory is renormalisable at the one-loop level for zero fermion masses (Veltman [17], [18]). Thus one suspects that the large m_H^2 effects will not be so dramatic after all. In fact, calculation shows (Veltman [19]; Chanowitz et al [20], [21]) that one-loop radiative corrections grow at most like $\ln m_H^2$ for large m_H. While there are finite corrections which are approximately $O(m_H^2)$ for $m_H^2 \ll M_{W,Z}^2$, for $m_H^2 \gg M_{W,Z}^2$ the $O(m_H^2)$ pieces cancel out from all observable quantities, leaving only $\ln m_H^2$ terms. This is just what we have in (326), and it means, unfortunately, that the sensitivity of the data to the last remaining parameter of the Standard Model (not counting the neutrino parameters!) is only logarithmic. Fits to data typically give m_H in the region of 100 GeV at the minimum of the χ^2 curve, but the error (which is not simple to interpret) is of the order of 50 GeV. Direct searches now rule out a Higgs mass less than about 110 GeV, while the $\sim$ 2.5 s.d. effect seen just before LEP closed down gave $m_H \sim 114$ GeV.

At the two-loop level, the expected $O(m_H^4)$ behaviour becomes $O(m_H^2)$ instead (van der Bij and Veltman [22], van der Bij [23]) - and of course appears (relative to the one-loop contributions) with an additional factor of $O(\alpha)$. This relative insensitivity of the radiative corrections to m_H, in the limit of large m_H, was discovered by Veltman [19] and called a 'screening' phenomenon by him: for large m_H (which also means, as we have seen, large λ) we have an effectively strongly interacting theory whose principal effects are screened off from observables at lower energy. It was shown by Einhorn and Wudka [24] that this screening is also a consequence of the (approximate) isospin-SU(2) symmetry we have just discussed in connection with (325). Phenomenologically, the upshot is that it is unfortunately very difficult to get a good handle on the value of m_H from fits to the precision data.

Problems for Lecture 6

P6.1 Show that the neutral current couplings are diagonal in the 'mass' basis.

P6.2 Suppose that we took the Higgs field to be a triplet of SU(2)$_{\mathrm{L}}$ instead of a doublet; and suppose $\langle 0|\phi|0\rangle = \begin{pmatrix} 0 \\ 0 \\ f \end{pmatrix}$ in the gauge in which it is real. The non-vanishing component has $t_3 = -1$, using

$$t_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

in the familiar 'spherical' basis. Since we want the charge of the vacuum to be zero ($Q|0\rangle = 0$) and $Q = t_3 + y/2$, we need to pick $y(\phi) = 2$. So the covariant derivative on ϕ is

$$(\partial_\mu + ig\boldsymbol{t} \cdot \boldsymbol{W}^\mu - ig' B^\mu)\phi$$

where

$$t_1 = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 0 \end{pmatrix}, \quad t_2 = \begin{pmatrix} 0 & \frac{-i}{\sqrt{2}} & 0 \\ \frac{i}{\sqrt{2}} & 0 & \frac{-i}{\sqrt{2}} \\ 0 & \frac{i}{\sqrt{2}} & 0 \end{pmatrix}$$

and t_3 is as above (this is the more familiar set of three matrices satisfying $[t_1, t_2] = it_3$, a change of basis from the set $(t_i)_{jk} = -i\epsilon_{ijk}$. Show that the photon and Z fields are still (257) and (258), with the same $\sin\theta_W$ as in (259), but that now

$$M_Z = \sqrt{2} M_W / \cos\theta_W.$$

What would be the parameter ρ, at tree level, for this model?

ACKNOWLEGEMENTS

It is a very great pleasure to thank all those responsible for organising this School, and for making it so successful and memorable. In particular, I should like to thank George Pogosyan for his personal kindness and hospitality, and for being such a wonderful Director. Special thanks also to Tatyana Donskova, Danielle Métral and Egil Lillestol for all their hard work before, during and after the School, which ensured that everything went smoothly and that everyone gained the maximum profit from the physics, and the location.

References

[1] I.J.R. Aitchison and A.J.G. Hey, *Gauge Theories in Particle Physics*, third edition, volume 1, Institute of Physics (2003).

[2] I.J.R. Aitchison and A.J.G. Hey, *Gauge Theories in Particle Physics*, third edition, volume 2, Institute of Physics (2004).

[3] J. Goldstone, Nuov. Cim. **19** (1961) 154.

[4] Y. Nambu, in *New Theories in Physics*, Proc. XI Int. Symp. on Elem. Part. Phys., Kazimierz, Poland, eds Z. Ajduk, S. Pokorski and A. Trauman, World Scientific, Singapore (1989).

[5] V.A. Miransky, M. Tanabashi and K. Yamawaki, Mod. Phys. Lett. **A4** (1989) 1043.

[6] V.A. Miransky, M. Tanabashi and K. Yamawaki, Phys. Lett. **B 221** (1989) 177.

[7] W.A. Bardeen, C. Hill and M. Linder, Phys. Rev. **D 41** (1990) 1647.

[8] E. Farhi and L. Susskind, Phys. Rep. **74 C** (1981) 277.

[9] S.L. Glashow, Nucl. Phys. **22** (1961) 579.

[10] S.L. Glashow, J. Iliopoulos and L. Maiani, Phys. Rev. **D 2** (1970) 1285.

[11] S. Weinberg, Phys. Rev. **D 8** (1973) 605 (especially footnote 8).

[12] G. Feinberg, S. Weinberg and P. Kabir, Phys. Rev. Lett. **3** (1959) 527 (especially footnote 9).

[13] C. Jarlskog, in *Fundamental Forces*, Proc. XXVII Scott. Univ. Sum. Sch. in Phys. 1 (1985).

[14] N. Cabibbo, Phys. Rev. Lett. **10** (1963) 531.

[15] M. Kobayashi and K. Maskawa, Prog. Theor. Phys. **49** (1973) 652.

[16] M.W. Grünewald, Phys. Rep. **322** (1999) 125.

[17] M. Veltman, Nucl. Phys. **B 7** (1968) 637.

[18] M. Velyman, Nucl. Phys. **B 21** (1970) 288.

[19] M. Veltman, Nucl. Phys. **B 123** (1977) 89.

[20] M. Chanowitz, E. Furman and I. Hinchliffe, Phys. Lett. **B 78** (1978) 285.

[21] M. Chanowitz, E. Furman and I. Hinchliffe, Nucl. Phys. **B 153** (1979) 402.

[22] J.J. van der Bij and M. Veltman, Nucl. Phys. **B 231** (1984) 205.

[23] J.J. van der Bij, Nucl. Phys. **B 248** (1984) 141.

[24] M.B. Einhorn and J. Wudka, Phys. Rev. **D 39** (1989) 2758.