1. The Standard Model

1.1. **All the phenomena of nature, except gravity, are described by the standard model of elementary particles.** The Lagrangian defines a particular quantum field theory. The precise definition of a quantum field theory involves the mysterious procedure of renormalization; we only understand it fully in perturbation theory.

1.2. **It is a Yang-Mills Theory with gauge group $SU(3) \times SU(2) \times U(1)$, a scalar field and a set of spin $\frac{1}{2}$ fields.** The theory is specified by the representations of the scalar and fermion fields; and the couplings among them. The three states $SU(3)$ symmetry is labelled by ' color'; $SU(2)$is called weak isospin and the $U(1)$is called weak hypercharge. We denote the corresponding gauge fields by a set of vector fields: $A^{\alpha}_{\mu\beta}$, each component of which is a traceless hermitean 3×3 matrix; $W^{a}_{\mu b}$, whose components are traceless hermitean 2×2 matrices; and Y_μ each component of which is just a real function.

1.3. **The scalar field is invariant under $SU(3)$; it is in the fundamental (doublet) representation of $SU(2)$; has hypercharge one.** That is, ϕ transforms as $(1,2,1)$: the trivial representation of color, the defining representation of $SU(2)$and of hypercharge one.

This means that if $g \in U(2) = SU(2) \times U(1)/Z_2$, the scalar field transforms as the doublet $\phi \to g\phi$. The covariant derivative on scalars is

$$\nabla_\mu \phi_b = \partial_\mu \phi_b + i[W^{a}_{\mu b} + Y_\mu \delta^{a}_{b}]\phi_a$$

1.4. **There are two kinds of fermions: quarks and leptons whose left and right-handed components couple differently to gauge bosons, as listed below.**

(1) The left-handed quarks $q_{\alpha b A}$ where $\alpha = 1,2,3$ labels the fundamental representation of color SU(3); b labels the fundamental representation of weak isospin $SU(2)$; the hypercharge $\frac{1}{3}$; and $A = 1,2,3$ labels the generations.

$$\gamma_5 q = q, \quad \nabla_\mu q_{\alpha b A} = \partial_\mu q_{\alpha b A} + i\left[\left\{W^{a}_{\mu b} + \frac{1}{3}Y_\mu \delta^{a}_{b}\right\}\delta^{\beta}_{\alpha} + A^{\beta}_{\mu\alpha}\right]q_{\beta a A}$$

(2) The left-handed leptons l_{bA} are trivial under color; in the fundamental of weak isospin; and has hypercharge -1.

$$\gamma_5 l = l, \quad \nabla_\mu l_{bA} = \partial_\mu l_{bA} + i\left\{W^{a}_{\mu b} - Y_\mu \delta^{a}_{b}\right\} l_{aA}$$

(3) The right-handed up-quarks $u_{\alpha A}$ are in the fundamental of color; are trivial under weak isospin; and has hypercharge $\frac{4}{3}$

$$\gamma_5 u = -u, \quad \nabla_\mu u_{\alpha A} = \partial_\mu u_{\alpha A} + i\left[A^\beta_{\mu\alpha} + \frac{4}{3} Y_\mu \delta^\beta_\alpha\right] u_{\beta A}$$

(4) The right-handed down-quarks$d_{\alpha A}$ are also fundamental in color; trivial under weak isospin; and hypercharge

$$\gamma_5 d = -d, \quad \nabla_\mu d_{\alpha A} = \partial_\mu d_{\alpha A} + i\left[A^\beta_{\mu\alpha} - \frac{2}{3} Y_\mu \delta^\beta_\alpha\right] d_{\beta A}$$

(5) The right-handed charged leptons $e_A$are trivial under both color and weak isospin; and have hypercharge -2.

$$\gamma_5 e = -e, \quad \nabla_\mu e_A = \partial_\mu e_A + i\left[-\frac{2}{3} Y_\mu\right] e_A$$

(6) The right-handed neutrinos ν_A are trivial under color and weak isospin, as well as of zero hypercharge. They do not couple to any gauge bosons at all.

$$\gamma_5 \nu = -\nu, \quad \nabla_\mu \nu_A = \partial_\mu \nu_A$$

1.5. **The Lagrangian can then be written down.**

$$L = \frac{1}{4\alpha_1} Y^{\mu\nu} Y_{\mu\nu} + \frac{1}{4\alpha_2} W^{\mu\nu} W_{\mu\nu} + \frac{1}{4\alpha_3} F^{\mu\nu} F_{\mu\nu} + |\nabla\phi|^2 - \frac{\lambda}{4}\left[\phi^\dagger\phi - v^2\right]^2 +$$

$$\bar{q}^{\alpha b A} i\gamma^\mu \nabla_\mu q_{\alpha b A} + \bar{u}^{\alpha A} i\gamma^\mu \nabla_\mu u_{\alpha A} + \bar{d}^{\alpha b A} i\sigma\gamma^\mu \nabla_\mu d_{\alpha b A} +$$

$$\bar{l}^{\alpha b A} i\gamma^\mu \nabla_\mu l_{\alpha b A} + \bar{e}^A i\sigma^\mu \nabla_\mu e_A + \bar{\nu}^A i\gamma^\mu \partial_\mu \nu_A +$$

$$+ \frac{1}{v}\sum_A M_A \bar{u}^{\alpha A} \phi_a \varepsilon^{ab} q_{\alpha b A} + \frac{1}{v}\sum_{AB} \tilde{M}_B V^B_A \bar{d}^{\alpha A} \phi^{\dagger b} q_{\alpha b B} +$$

$$\frac{1}{v}\sum_A m_A \bar{e}^A \phi_a \varepsilon^{ab} l_{bA} + \frac{1}{v}\sum_A \tilde{m}_A U^B_A \bar{\nu}^A \phi^{\dagger b} l_{bB}$$

Note that the right-handed neutrinos do not couple to any gauge bosons. The basis is chosen so that mass matrix of the up quarks is diagonal, with entries M_A. Then those of the down quarks will be the diagonal matrix $\tilde{M}$ times the CKM matrix V which is a unitary matrix; because of the freedom to choose phases for the basis of quarks, V is defined only modulo a left and a right action by diagonal unitary matrices: $V \in U(1)^N \backslash U(N)/U(1)^N$. Similarly, the mass matrix of the charged leptons is chosen to be diagonal, leading to a neutrino mixing matrix $U \in U(1)^N \backslash U(N)/U(1)^N$. The dimension of the double coset space $U(1)^N \backslash U(N)/U(1)^N$ is $(N-1)^2$.

1.6. **The parameters are $\alpha_1, \alpha_2, \alpha_3, \lambda, v, M_A, \tilde{M}_{A,} V_A^B, m_A, \tilde{m}_A, U_A^B$.** That is $5+4N+2(N-1)^2 = 2N^2+7$ parameters where N is the number of generations. For $N=3$ these are 25 parameters. Most of them are known to reasonable accuracy: the unknown parameters are the Higgs mass (λ)and all except two combinations of $\widetilde{m_A}, U_B^A$ in the neutrino sector: only the two combinations of the neutrino masses and mixing angles that give neutrino oscillation lengths are known.

1.7. **All of the particles predicted by the standard model except the right handed neutrinos and the Higgs boson have been observed.** Most of them were not known at the time the model was proposed. They were found over the last three decades, as the energy of the accelerators and the size of neutrino detectors increased. The aim of the next generation of high energy physics experiments is to detect the Higgs boson and the right handed neutrinos; or, to prove that the standard model is wrong.

1.8. **In the following sections we will develop the mathematical and physical apparatus to explain the meaning of the above Lagrangian.**

2. Quantum Mechanics

A good reference is R. Shankar, *Principles of Quantum Mechanics* . The classic text remains Dirac's book of the same title. You will be a walking encyclopedia of quantum mechanics if you master the third volume of the *Course in Theoretical Physics* by Landau and Lifshitz. The summary below is not a substitute for a proper course in quantum mechanics: it can take a year to learn the material summarized in this section.

2.1. **The Postulates.** Once you understand the basic structure of a physical or mathematical theory, it is useful to summarize the basic laws as axioms: independent facts from which all others can be derived. This was first achieved for plane geometry by Euclid. For mechanics by Newton. There is always a period of experimentation and discovery before a subject become mature enough to be axiomatized.

In the case of non-relativistic quantum mechanics the works of Planck, Einstein, Bohr, Heisenberg, Pauli, Born,Jordan, Schrodinger.... formed this period of discovery. It was Dirac, in his book *Principles of Quantum Mechanics,* who put it all together in more or less the form we think of it now.

Quantum theory is still not completely developed. Questions about measurement and interpretation are still being worked out (e.g.,"weak measurement"). Also, combining relativity with quantum mechanics (quantum field theory) leads to infinities that have to be removed by a mysterious procedure known as renormalization. No one is satisfied with this situation. Even worse, we have not yet been able to reconcile quantum mechanics with general relativity to get a quantum theory of gravity. Also, the discovery of dark energy suggests that something is seriously wrong in the way we think of the energy of the vaccuum.

Nevertheless we can say, after almost a century of experimental tests, a few things for certain about how quantum theory works. It is not too early to summarize them as a set of postulates.

2.1.1. *The states of a physical system are represented by vectors in a complex Hilbert space.* This means the we can take linear combinations

$$\alpha|\psi> + \beta|\phi>$$

of two states $|\psi>$ and $\phi>$ to get another state. The quantities α, β are complex numbers. There is a way to take the inner product (scalar product) of two states to get a complex number

$$<\psi|\phi> .$$

This inner product is linear in the second argument

$$<\psi|\alpha\phi + \beta\chi> = \alpha <\psi|\phi> + \beta <\psi|\chi>$$

and anti-linear in the first entry

$$<\alpha\psi + \beta\chi|\phi> = \alpha^* <\psi|\phi> + \beta^* <\chi|\phi> .$$

Remark 1. Be aware that mathematicians use the opposite convention: for them it is the second entry in an inner product that is anti-linear. Mathematics and physics are two neighboring cultures divided by a common language.

Moreover, the inner product of any vector with itself is positive; it is only zero for the zero vector. Thus

$$|\psi|^2 = <\psi|\psi>$$

can be thought of as the square of the length of a vector.

Remark 2. Strictly speaking states are represented by rays (directions) in Hilbert space. It is a fine point though.

A typical situation is that the state is a complex valued function of some real variable (e.g., position), The inner product is then

$$<\psi|\psi> = \int \psi^*(x)\phi(x)dx.$$

Exercise 3. Verify that this integral has the properties of an inner product.

Or, the states may be represented by a column vector with complex components $\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$ etc.

Exercise 4. Prove that

$$\frac{|<\psi|\phi>|^2}{|\psi|^2|\phi|^2} \leq 1$$

for all non-zero states $|\phi>, |\psi>$. This is called the Cauchy-Schwarz inequality.

2.1.2. *If a system is in state $|\phi>$, the probability of finding it in another state $|\psi>$is $\frac{|<\psi|\phi>|^2}{|\psi|^2|\phi|^2}$.*. This is one of the confusing things about quantum mechanics, until you get used to it. A classical analogue is the polarization of light. About half of circularly polarized light will pass through a filter that allows only linearly polarized light.

2.1.3. *The observables of a physical system are hermitean linear operators on the states.* A linear operator (or matrix) acting on a state produces another state, such that

$$L(\alpha|\psi> + \beta|\chi>) = \alpha L|\psi> + \beta L|\chi> .$$

Hermitean operators satisfy in addition

$$<\psi|L|\chi> = <\chi|L|\psi>^* .$$

That is, the conjugate-transpose of a matrix elements is itself. If

$$L|\psi_\lambda> = \lambda|\psi_\lambda>$$

for some complex number λ and non-zero vector $|\psi_\lambda>$, we say that $\psi_\lambda>$is an eigenvector of $L$with eigenvalue λ. The most important property of a hermitean operator is that it has real eigenvalues. That comes in handy because

2.1.4. *The possible outcomes of measuring an observable are its eigenvalues.* Even if if we know the state of a system, we may not be able to predict the outcome of measuring an observable. The best we can do is to give probabilities. With $\lambda, |\psi_\lambda>$ defined as above, 2.1.5. *If the system is in some state $|\phi>$,the probability of getting the outcome λ upon measuring L is.*

$$\frac{|<\psi_\lambda|\phi>|^2}{|\psi_\lambda|^2|\phi|^2}.$$

Exercise 5. Recall that this is always real and less than one. Why do the probabilities add up to one?

2.1.6. *There is a hermitean operator called the hamiltonian which represents energy; the time dependence of a state is given by.*

$$i\hbar\frac{\partial|\psi>}{\partial t} = H|\psi> .$$

Thus if you know the state at some time, you can in principle predict what it will be at some later time. If you know the exact hamiltonian and if it is simple enough to make the equation solvable.

2.2. **Electron in a Magnetic Field.** As an example, think of an electron in a magnetic field. It is bound to an atom (e.g., Sodium) and we ignore the change in its position: only the rotation of its spin. The wave function has two components. The energy of an eletron in a magnetic field is proportional to the dot product of the spin and the magnetic field

$$H = \mu\sigma\cdot B \tag{2.1}$$

where

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

are the Pauli matrices.

Exercise 6. Find the eigenvalues and eigenfunctions of the hamiltonian (2.1). If the initial state at time $t = 0$ is $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and the magnetic field is in along the $x-$axis $B = (B,0,0)$ what is the probability that a measurement of σ_3 at a time t will yield the value -1? This illustrates the phenomenon of oscillationof quantum states, also important for neutrinos.

2.3. **Symmetries and Conservation Laws.**

2.3.1. *Symmetries are represented by unitary operators that commute with the hamiltonian.* Recall that the probability of finding a particle in state ψ in another state ϕ is $|<\phi|\psi>|^2$ (assuming that the state vectors are of length one.) If the symmetry is represented by a linear transformation L it must satisfy

$$|<L\phi|L\psi>| = 1.$$

Also recall the definition of the hermitean conjugate (adjoint)

$$< L\phi|\psi >=< \phi|L^\dagger\psi >$$

Thus we see that one way to satisfy the condition is to have

$$L^\dagger L = 1$$

That is, **unitary transformations**. Most symmetries are of this type. (See below for an exception.)

Recall that a state of energy E is an eigenstate of the hamiltonian.

$$H\psi = E\psi$$

A symmetry must take it to another state of the same energy:

$$H(L\psi) = E(L\psi).$$

This is satisfied if

$$HL = LH.$$

That is, if the hamiltonian commutes with the symmetry operator. Thus a symmetry is represented by a unitary operator that commutes with the hamiltonian:

$$L^\dagger L = 1, \quad HL - LH = 0.$$

2.3.2. *An exceptional case is time reversal, which is an anti-linear operator.*

$$\Theta(a\psi + b\phi) = a^*\Theta\psi + b^*\Theta\phi.$$

We won't have much more to say about this case for now; we will only consider the case of linear operators for now.

2.3.3. *An example is Parity.* It reverses the sign of the co-ordinates of a particle

$$P\psi(x) = \psi(-x).$$

Clearly $P^2 = 1$.

The Schrödinger equation for a free particle is invariant under this transformation

$$-\frac{\hbar^2}{2m}\nabla^2\psi = -i\hbar\frac{\partial\psi}{\partial t}.$$

Another way of seeing that this is a symmetry is that the operator P commutes with the hamiltonian

$$H = -\frac{\hbar^2}{2m}\nabla^2, \quad PH = HP.$$

Thus if ψ is a state with energy E

$$H\psi = E\psi$$

so will be $P\psi$. Even with a potential parity continues to be a symmetry if

$$V(-x) = V(x).$$

For example consider a particle in one dimension with a potential

$$H = -\frac{\hbar^2}{2m}\nabla^2 + V, \quad V(x) = \lambda(x^2 - a^2)^2, \quad \lambda > 0.$$

There are two minima at $x = \pm a$. The eigenstates of energy can also be simultaneously eigenstates of parity because $[H,P] = 0$. It turns out that the ground state is of even parity

$$\psi(-x) = \psi(x)$$

while the first excited state is of odd parity

$$\psi(-x) = -\psi(x).$$

2.3.4. *Translation invariance leads to conservation of momentum.* The translation by a is represented by the operator

$$T(a)\psi(x) = \psi(x+a).$$

A free particle on a line has hamiltonian

$$H = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V$$

with a constant potential. Thus whether we apply the hamiltonian before or after a translation we get the same effect on a wavefunction:

$$HT(a) = T(a)H.$$

For a particle moving in one dimension, an infinitesimal translation is represented by the derivative operator:

$$\psi(x+a) \approx \psi(x) + a\frac{\partial\psi}{\partial x} + \cdots$$

Thus if a system is invariant under translation, its hamiltonian must satisfy

$$[H, \frac{\partial}{\partial x}] = 0.$$

The operator $\frac{\partial}{\partial x}$is anti-hermitean. The corresponding hermitean operator is $-i\frac{\partial}{\partial x}$. If we multiply by $\hbar$ we get the momentum operator

$$p = -i\hbar\frac{\partial}{\partial x}.$$

Thus translation invariance implies the conservation of the momentum:

$$[H, p] = 0.$$

Similar arguments apply to each component of momentum of a free particle moving in R^3.

2.3.5. *Rotation invariance implies conservation of angular momentum.* The infinitesimal generators of rotation are

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}, \quad \mathbf{p} = -i\hbar \frac{\partial}{\partial \mathbf{r}}.$$

They satisfy the relations

$$[L_1, L_2] = i\hbar L_3, \quad [L_2, L_3] = i\hbar L_1, \quad [L_3, L_1] = i\hbar L_2.$$

2.3.6. *In quantum mechanics, a particle can have angular momentum even when its momentum is zero.* Total angular momentum is the sum of the orbital angular momentum and an intrinsic angular momentum

$$\mathbf{J} = \mathbf{L} + \mathbf{S}.$$

The components of S are a set of three matrices satisfying

$$[S_1, S_2] = i\hbar S_3, \quad [S_2, S_3] = i\hbar S_1, \quad [S_3, S_1] = i\hbar S_2.$$

The simplest choice is $S = 0$. There are several such spin zero particles; e.g., the alpha particle. The next simplest choice is

$$S_1 = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad S_2 = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad S_3 = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

These are 'spin half' particles, since the maximum eigenvalue of a component of spin is half of $\hbar$. An electron, a proton, a neutron are all examples of such particles.

The photon has spin one. But it cannot be described by the above theory because it moves at the speed of light. We need relativistic quantum mechanics for that.

There are a set of particles called Δ that have spin $\frac{3}{2}$. Their spin is represented by 4x4 matrices. There are particles with even higher spin but they tend to be unstable.

2.4. **Rotations.**

2.4.1. *The distance between two points* $x = (x_1, x_2 x_3)$ *and* $y = (y_1, y_2, y_3)$ *is given by* $(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2$. If we translate both vectors by the same amount , the distance is unchanged. Similarly if we rotate both the same way, the distance is unchanged.

2.4.2. *The square of the length of a vector, thought of as a* 3×1 *matrix is* $x^T x$*.* A rotation is described by a 3×3 matrix R

$$x \to Rx.$$

The length of x is unchanged if

$$R^T R = 1$$

Such matrices are called *orthogonal*. Not all orthogonal matrices describe rotations though. The matrix

$$P = -1$$

reverses the sign of all three co-ordinates, which cannot be achieved by any rotation. Yet it is orthogonal. Every orthogonal matrix has determinant ± 1. Every rotation can be built up by rotations through very small angles. Since the sign of the determinant cannot suddenly jump from 1 to -1, the determinant of a rotation matrix must be the same as for the identity. That is,

2.4.3. *A rotation is a matrix that is both orthogonal and of determinant one.* The set of all 3×3 orthogonal matrices is called $O(3)$. Matrices of determinant one used to be called special matrices. Thus the set of rotation matrices is called $SO(3)$.

2.4.4. *If* $g, h \in SO(3)$*, then* $gh \in SO(3)$ *and* $g^{-1}, h^{-1} \in SO(3)$ *as well.* In mathematical terminology, this means that $SO(3)$is a *group*: it is closed under multiplication of these matrices as well as taking their inverses. This idea of a group is essential in particle physics: all the important symmetries are described by groups.

2.4.5. *An infinitesimal transformation* $R = 1 + A$ *is orthogonal* $A^T + A = 0$*; i.e., infinitesimal rotations are described by anti-symmetric matrices.* An arbitrary anti-symmetric matrix can be written as a linear combination of the basic ones

$$S_{12} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad S_{23} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad S_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$$

which describe rotations in each co-ordinate plane. They satisfy the commutation relations (verify this)

$$[S_{12}, S_{23}] = S_{13}, \quad [S_{23}, S_{13}] = S_{12}, \quad [S_{13}, S_{23}] = -S_{12}$$

This is an example of what is called a Lie algebra. It has turned out that elementary particles are classified by algebras $SU(3), SU(2)$ such as these, only more complicated. Note that the Pauli matrices satisfy very similar commutation relations.

Exercise 7. What are the commutation relations of rotations in four dimensions?

3. Relativity

3.1. **It is an astonishing physical fact that the speed of light is the same for all observers.** This is not true of other waves. For example, the speed of sound measured by someone on standing on Earth is different from that measured by someone in an airplane. This is because sound is the vibration of molecules of air while light is the oscillation of electric and magnetic fields: there is no medium (ether) needed for its propagation. If an observer is moving with velocity v_1 relative to air, the speed of sound measured by him/her would be

$$v = v_1 + c_s$$

where c_s is the speed of sound measured by a static observer.

3.2. **The law of addition of velocities has to be modified to take account of this fact.** The usual law (Galeleo) for addition of velocities (considering only one component for simplicty)

$$u + v$$

would lead to velocities greater than or less than c , the speed of light. The correct law is

$$\frac{u+v}{1+\frac{uv}{c^2}}.$$

3.3. **Rapidity is a more convenient variable than velocity in relativistic mechanics.** If either u or v is of magntitude c, the sum is also of magnitude c. If we make the change of variables (considering only one component for simplicity)

$$v = c\tanh\theta$$

this formula becomes simple addition:

$$\tanh[\theta_1+\theta_2] = \frac{\tanh\theta_1+\tanh\theta_2}{1+\tanh\theta_1\tanh\theta_2}.$$

The variable θ is called *rapidity.* Although velocities cannot exceed c,rapidity can be as big as you want. As $\theta \to \infty$, $v \to c$. This also suggests that a boost (change of velocity) is some kind of rotation, through an imaginary angle. Remember that $\tan i\theta = i\tanh\theta$.

3.4. **The wavefront of light has the same shape for all observers.** Imagine that you turn on and off quickly a light bulb at $t = 0$,at the position $\mathbf{x} = 0$. Light propagates along cone

$$c^2t^2 - x_1^2 - x_2^2 - x_3^2 = 0,$$

3.5. **The laws of physics must be the same for all observers.** This is the principle of relativity. It is not a discovery of Einstein: what is new is really how he reconciled this principle with the fact that the velocity of light is the same for all observers. Einstein realized that Newton's laws of mechanics need to be modified to fit with the new law for the addition of velocities. Minkowski realized that the theory of relativity can be understood geometrically: it says that the square of the distance between two events (points in space-time) is

$$c^2(t-t')^2 - (x_1 - x_1')^2 - (x_2 - x_2')^2 - (x_3 - x_3')^2$$

Lorentz transformations (changes of velocities) are like rotations in the $x - t$ plane. Because of the relative sign difference between the time and space components, these rotations are through an imaginary angle; this angle is rapidity. More generally

3.6. **The Minkowski inner product of four-vectors is.**

$$u \cdot v = u_0v_0 - u_1v_1 - u_2v_2 - u_3v_3$$

It is useful to write this as

$$u \cdot v = u^T \eta v$$

where

$$\eta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

is called the *Minkowski metric.*

Remark 8. Be warned that all physicists except particle physicists, use the opposite convention, with three positive and one negative sign. Even some particle physicists use the opposite convention (e.g., S. Weinberg).

The points of space time are four vectors with components ct, x_1, x_2, x_3.

3.7. **Rapidity can be thought of as an angle in the $x\ t$ plane.** Again consider just one direction for space. Supose $x\ t$ and $x'\ t'$ are space and time as measured by two observers. Since they must agree that the velocity of light is the same, the shape of the wave front must be the same as well:

$$c^2t^2 - x^2 = c^2t'^2 - x'^2$$

This means that

$$x' = \cosh\theta x - \sinh\theta ct$$

$$ct' = \cosh\theta ct - \sinh\theta x$$

for some real number θ. Similar to the way a rotation in the plane leaves the square of the distance $x^2 + y^2$ unchanged. We can identify this angle with rapidity. Along $x' = 0$ (the position of one of the observers) the relation between $x$and t is

$$\frac{x}{t} = c\tanh\theta.$$

In other words $v = c\tanh\theta$. The above is an example of a Lorentz transformation. More generally,

3.8. **A Lorentz transformation is a 4×4 matrix that leaves the Minkowski distance unchanged.** Thus a Lorentz transformations is much like a rotation, except that the matrices must satisfy the condition

$$\Lambda^T\eta\Lambda = \eta.$$

3.9. **The products and inverses of Lorentz transformations are also Lorentz transformations.** This means that the set of Lorentz transformations forms a *group*. It is denoted by $O(1,3)$: orthogonal matrices with respect to a metric $\eta$with 1 positive sign and three negative signs.

3.9.1. *O(3) is contained as a special case of O(1,3).* A matrix that does not mix space and time

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & R & \\ 0 & & & \end{pmatrix}$$

is a Lorentz transformation if R is orthogonal.

3.9.2. *We can split Lorentz transformations into four types using the signs of* $\det\Lambda$ *and* Λ_{00}.. If $\Lambda_{00} < 0$ the transformation includes a time reversal. If $\Lambda_{00} > 0$ and $\det\Lambda < 0$, it includes a space reversal (Parity). The subgroup with

$$\det L > 0, \quad \Lambda_{00} > 0$$

is called the set of *proper Lorentz transformations* $SO(1,3)$. Although at first all four types of Lorentz transformations appeared to be at he same footing, it turned out neither Parity nor time reversal is an exact symmetry of nature.

3.9.3. *Only the proper Lorentz transformtions are exact symmeties of nature.* Parity is violated by the weak interactions: essentially the fact that neutrinos are left handed. Time reversal is violated by the phase in the mass matrix (Kobayashi-Mazkawa matrix) of quarks.

3.9.4. *Infinitesimal Lorentz transformations form a six dimensional Lie algebra.* If $\Lambda = 1 + \lambda$ for some small matrix λ, the Lorentz condition becomes

$$\lambda^T\eta + \eta\lambda = 0$$

This means that $\eta\lambda$ is anti-symmetric. A four by four antis-symmetric matrix has six independent components. Three of them represent infinitesimal rotations. The remaining three are boosts (changes of velocity) in the three co-ordinate directions. Any infinitesimal Lorentz transformation can be written as the sum of the six independent matrices:

$$L_{12} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad L_{23} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad L_{13} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

(3.1)

$$L_{01} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad L_{02} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad L_{03} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

They satisfy the commutation relations

$$[L_{12}, L_{23}] = -L_{13}$$

$$[L_{12}, L_{01}] = L_{02}$$

$$[L_{01}, L_{02}] = -L_{12}$$

and their cyclic permutations. These commutation relations define the Lorentz Lie algebra. They can also be written as

$$[L_{\mu\nu}, L_{\rho\sigma}] = \eta_{\nu\rho}L_{\mu\sigma} - \eta_{\mu\rho}L_{\nu\sigma} - \eta_{\nu\sigma}L_{\mu\rho} + \eta_{\mu\sigma}L_{\nu\rho}$$

3.9.5. *In addition to Lorentz transformations, translations are also symmetries. Together they form a ten dimensional algebra of symmetries, the Poincare' algebra.*

$$[L_{\mu\nu}, L_{\rho\sigma}] = \eta_{\nu\rho}L_{\mu\sigma} - \eta_{\mu\rho}L_{\nu\sigma} - \eta_{\nu\sigma}L_{\mu\rho} + \eta_{\mu\sigma}L_{\nu\rho}$$

$$[L_{\mu\nu}, P_\rho] = \eta_{\nu\rho}P_\mu - \eta_{\mu\rho}P_\nu$$

$$[P_\mu, P_\nu] = 0. \tag{3.2}$$

3.10. **Invariance under translations leads to the conservation of energy and momentum.**

3.11. **Energy and momentum transform together as a four-vector under Lorentz transformations.**

$$p = (E, cp_1, cp_2, cp_3).$$

3.12. **The relation between energy and momentum is.**

$$p \cdot p = m^2c^4, \quad E > 0$$

Geometically, this one sheet of a hyperboloid in four-dimensional space, called the *mass shell*.

$$E = \sqrt{m^2c^4 + c^2\mathbf{p}^2}$$

In particular, even a particle at rest has energy

$$E = mc^2.$$

For velocities small compared to c,

$$E \approx mc^2 + \frac{\mathrm{p}^2}{2m}$$

The second term is the Newtonian formula for kinetic energy. Since the mass of particles usually do not change,in most situations we do can ignore the first term. But in nuclear reactions, this energy can be released with spectacular results.

3.13. **For massless particles the momentum vector is null.** The inner product of momentum with itself is zero

$$p \cdot p = 0.$$

Geometrically, the set of null momenta is a cone in four-dimensional space.

The relation of energy to momentum is

$$E = c|\mathbf{p}|$$

3.14. **Free particles move along straight lines in Minkowski space.** Massive particles move along time-like straight lines: the tangent vector has positive inner product with itself. Massless particles move along null lines.

3.15. **Conservation of energy-momentum places important restrictions on decays and scattering of elementary particles.**

4. The Schrodinger equation has to be changed to take account of relativity

Recall that in quantum mechanics

$$\mathbf{p} = -i\hbar\frac{\partial}{\partial \mathbf{x}}, \quad E = i\hbar\frac{\partial}{\partial t}.$$

The relation

$$E = \frac{\mathbf{p}^2}{2m}$$

of non-relativistic mechanics gives the Schrodinger equation for a free particle:

$$i\hbar\frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2 \psi}{\partial \mathbf{x}^2}.$$

For a relativistic particle instead

$$E^2 = c^2\mathbf{p}^2 + m^2c^4$$

leading to

4.1. **The Klein Gordon Equation.**

$$\frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} = \frac{\partial^2\psi}{\partial \mathbf{x}^2} - \left(\frac{mc}{\hbar}\right)^2 \psi.$$

The quantity $\frac{\hbar}{mc}$ has the dimension of length; it is a fundamental property of a particle determined by its mass, called its Compton wavelength. It is called that because this combination first appeared in Compton's explanation of the scattering of gamma rays by elecrons. Now we know that this equation only describes spin zero particles. Dirac discovered the correct equation for spin one half particles like the electron.

4.2. **For a massless particle this becomes the wave equation.**

$$\frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} = \frac{\partial^2\psi}{\partial \mathbf{x}^2}$$

4.3. **Mass and spin define the transformation properties of an elementary particle.** Mass is defined by the norm of the momentum four-vector.

$$p \cdot p = m^2$$

Spin s is defined in a similar way by

$$W \cdot W = m^2 s(s+1)$$

where

$$W^\mu = \varepsilon^{\mu\nu\rho\sigma} p_\nu L_{\rho\sigma}$$

is the Pauli-Lubanski vector. Its time component is the dot-product of angular momentum and momentum. Thus it picks out the intrisic or spin component of angular momentum: orbital angular momentum has zero dot product with momentum.

4.4. **In the quantum theory, infinitesimal translations and Lorentz transformations are represented by hermitean operators on the Hilbert space.**

4.5. **An elementary particle is such an irreducible representation of the Poincare Lie algebra.** Irreducible means that every state in the Hilbert space can be turned into any other state by some Poincare transformation. If the representation is not irreducible, there would be some subset of states that only mix with each other (form an invariant subspace) and then the system can be broken up into two pieces (the invariant subspace and its complement). So it would not be elementary or indivisible.This mathematical realization of the physical concept of an elementary particle is due to Wigner.

5. Maxwell's Equations

Read the book by Jackson on *Classical Electrodynamics.* Or the second volume of the series by Landau and Lifshitz *Classical Theory of Fields.*

5.1. **All magnetic fields must have zero divergence.**

$$\nabla \cdot \mathbf{B} = 0$$

This means in particular that there is no analogue to an isolated electric charge in magnetism: a permanent magnet has to be a dipole. If you cut a dipole into two we will not get an isolated North pole and South pole. Instead we will get two dipoles again. Some theories that go beyond the standard model do allow for magnetic monopoles; but none have yet been observed.

5.2. **This equation can be solved by postulating that the magnetic field is a curl of a vector potential.**

$$\mathbf{B} = \nabla \times \mathbf{A}$$

5.3. **Two vector potentials that differ only by the gradient of a scalar give the same magnetic field.** This is called a gauge transformation

$$\mathbf{A}' = \mathbf{A} + \nabla\Lambda, \quad \mathbf{B}' = \mathbf{B}$$

$$\nabla \times \nabla\Lambda = 0.$$

It turns out that invariance under this transfomation is a fundamental symmetry of nature. We will see that gauge transformations that generalize this are the fundametnal symmetries of the standard model.

5.4. **Another equation of Maxwell relates the time derivative of the magnetic field to the eletric field.**

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times \mathbf{E}$$

5.5. **We can solve this by postulating in addition a scalar potential V.**

$$\mathbf{E} = \frac{\partial \mathbf{A}}{\partial t} - \nabla V$$

Remark 9. Recall that we are using units such that $c = 1$. Otherwise there will be some factors of $c$all over the place.

The gauge transformations must now change the scalar potential as well

$$V' = V + \frac{\partial \Lambda}{\partial t}$$

so that the electric field is unchanged.

$$\frac{\partial \nabla \Lambda}{\partial t} = \nabla \frac{\partial \Lambda}{\partial t}.$$

5.6. **Under Lorentz transformations the scalar and vector potentials combine into a four-vector** $A = (V, \mathbf{A})$**.** We will introduce an index $\mu = 0, 1, 2, 3$ such that

$$A_0 = V, \quad A = (A_0, A_1, A_2, A_3)$$

Then the gauge transformation can be written as

$$A'_\mu = A_\mu + \partial_\mu \Lambda$$

where ∂_μ denotes differentiation along the μ th direction. Gauge invariance is based on the identity

$$\partial_\mu \partial_\nu \Lambda = \partial_\nu \partial_\mu \Lambda.$$

The electric and magnetic fields are then

$$E_i = \partial_0 A_i - \partial_i A_0, \quad i = 1, 2, 3.$$

$$B_1 = \partial_2 A_3 - \partial_3 A_2, \quad B_2 = \partial_3 A_1 - \partial_1 A_3, \quad B_3 = \partial_1 A_2 - \partial_2 A_1$$

This suggests that we combine them into a single matrix $F_{\mu\nu}$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

It is an anti-symmetric matrix:

$$F = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & B_3 & -B_2 \\ -E_2 & -B_3 & 0 & B_1 \\ -E_3 & B_2 & -B_1 & 0 \end{pmatrix}$$

5.7. **Scalar Products of vectors are Lorentz invariant.** Recall that there is also a symmetric matrix $\eta^{\mu\nu}$ that allows us to take products of vectors. Its indices are written above as a way of keeping track of them:

$$\eta^{\mu\nu} p_\mu q_\nu = p_0 q_0 - p_1 q_1 - p_2 q_2 - p_3 q_3, \quad \eta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

It is also useful to think of the combination

$$\eta^{\mu\nu} p_\mu = p^\nu, \quad p^\nu = (p_0, -p_1, -p_2, -p_3)$$

as a vector with indices above and the scalar product as

$$p^\nu q_\nu = p_0 q_0 - p_1 q_1 - p_2 q_2 - p_3 q_3$$

A pair of indices that are repeated are summed over.

$$p^\nu q_\nu = \sum_\nu p^\nu q_\nu.$$

This convention due to Einstein simplifies the appearance of equations. It means that you must be careful not to use the same index more than twice.

5.8. **The remaining Maxwell's equations can be written in Lorentz invariant form as.**

$$\partial^\mu F_{\mu\nu} = j_\nu$$

Expanded in terms of three dimensional quantities

$$\frac{\partial \mathbf{E}}{\partial t} = -\nabla \times \mathbf{B} + \mathbf{j}$$

$$\nabla \cdot \mathbf{E} = j_0$$

The scalar j_0 is proportional to charge density and the vector $\mathbf{j}$ to current density.

5.9. **The potential A satisfies a wave equation.**

5.10. **The electromagnetic field describes a particle of mass zero and spin one.** Mass zero because it travels at the velocity of light. (Duh. it is light.) Spin one because in three dimensional language it includes a vector field, which has spin one.

6. The Dirac Equation

We saw earlier that a free spin zero massive particle is described by the Klein-Gordon equation

$$\partial^\mu \partial_\mu \phi + m^2 \phi = 0.$$

It was Dirac who discovered the correct relativistic wave equation for spin $\frac{1}{2}$ particles. Recall that such a particle has angular momentum even when it is at rest, given by the Pauli matrices $\frac{1}{2}\sigma$. Also, their wave function is not a single complex number but a pair of complex numbers. Since the spin matrices act on such a pair, it is called a *spinor.* Since both the spin and the momentum are vectors a combination such as

$$\sigma\cdot\mathbf{p}=\begin{pmatrix} p_3 & p_1-ip_2\\ p_1+ip_2 & -p_3\end{pmatrix}$$

is a scalar under rotations. But it changes sign under parity (reflection of all three spatial co-ordinates). Note that

$$(\sigma\cdot\mathbf{p})^2=|\mathbf{p}|^{\mathbf{2}}.$$

Moreover,

$$\det\sigma\cdot\mathbf{p}=-|\mathbf{p}|^2.$$

Is there a way to generalize this to get something invariant under Lorentz transformations?

6.1. **The set of four matrices** $(1,\sigma_1,\sigma_2,\sigma_3)=\sigma^\mu$ **transform as a vector under Lorentz transformations.**

6.1.1. $\sigma\cdot p==\begin{pmatrix} p_0+p_3 & p_1-ip_2\\ p_1+ip_2 & p_0-p_3\end{pmatrix}$ *is Lorentz invariant.* Note that

$$\det\sigma\cdot p=p_0^2-|\mathbf{p}|^2=p\cdot p$$

is the Minkowski scalar product.Using the usual quantum mechanical rule momentum can be thought of as differentiation

$$p_\mu=-i\hbar\partial_\mu$$

leading to a wave equation origially discovered (but not published) by Pauli.

Remark 10. We will use units such that $\hbar=1$so that it will not usually appear explicitly. I put in in here just for clarity.

6.2. **The Pauli Wave Equation.**

$$\sigma\cdot\partial\chi=0$$

Or expanded out,

$$\frac{\partial\chi_1}{\partial t}+\frac{\partial\chi_1}{\partial z}+\frac{\partial\chi_2}{\partial x}-i\frac{\partial\chi_2}{\partial y}=0$$

$$\frac{\partial\chi_2}{\partial t}-\frac{\partial\chi_2}{\partial z}+\frac{\partial\phi\chi_1}{\partial x}+i\frac{\partial\chi_1}{\partial y}=0$$

6.2.1. *This describes a massless spin $\frac{1}{2}$ particle: each component of the spinor satisfies the wave equation.* The equation is Lorentz invariant, but is not invariant under parity. This because the combination $E - \sigma \cdot \mathbf{p}$ is invariant under rotation and Lorentz boosts. Under a reflection momentum changes sign but not angular momentum. (Remember that spin $\frac{1}{2}\sigma$ and orbital angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ must transform the same way). Therefore $\sigma \cdot \mathbf{p}$ changes sign. But not E.

Pauli thought this equation could describe the neutrino but rejected it because it violated parity. But in 1957 it was discovered that Parity symmetry is broken, and precisely in beta decays: that involve the neutrino!

6.2.2. *We can get a Parity invariant equation by putting together two Pauli spinors: under parity we just exchange them.*

$$(\partial_0 + \sigma \cdot \nabla)\,\chi = 0$$

$$(\partial_0 - \sigma \cdot \nabla)\,\tilde{\chi} = 0$$

These still decsribe a pair of massless spin one half particles. We can get a parity invariant equation for a massive spin one half particle by making mass mix the two components:

6.3. **The Dirac equation is.**

$$(\partial_0 + \sigma \cdot \nabla)\,\chi = im\tilde{\chi}$$

$$(\partial_0 - \sigma \cdot \nabla)\,\tilde{\chi} = im\chi$$

The factor of i is chosen such that

6.3.1. *Each component satisfies the wave equation for massive particles.*

$$\left(\partial_0^2 - \nabla^2\right)\chi = -m^2\tilde{\chi}$$

$$\left(\partial_0^2 - \nabla^2\right)\tilde{\chi} = -m^2\chi$$

Thus a Dirac spinor has four components, which can be broken up into two Pauli spinors.

$$\psi = \begin{pmatrix} \chi \\ \tilde{\chi} \end{pmatrix}$$

$$\begin{pmatrix} 0 & \partial_0 - \sigma \cdot \nabla \\ \partial_0 + \sigma \cdot \nabla & 0 \end{pmatrix} \psi = im\psi$$

6.3.2. *The Dirac equation can also be written as.*

$$\gamma^{\mu}\partial_{\mu}\psi = im\psi$$

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\gamma = \begin{pmatrix} 0 & -\sigma \\ \sigma & 0 \end{pmatrix}$$

Recall that the Pauli matrices satisfy (indeed are defined by) the conditions

$$\sigma^i\sigma^j + \sigma^j\sigma^i = 2\delta^{ij}.$$

That is, each Pauli matrix has square one; and they ani-commute with each other. In the same spirit

6.3.3. *The Dirac matrices satisfy the conditions.*

$$\gamma^{\mu}\gamma^{\nu} + \gamma^{\nu}\gamma^{\mu} = 2\eta^{\mu\nu}.$$

Exercise. Use this identity to derive the massive wave equation for each component of the Dirac spinor, without using the explicit form in terms of Pauli matrices above.

Up to a choice of basis, all the properties of Dirac matrices follow from these conditions. In fact people use use other representations than the one above (which is called the chiral representation because it makes the parity or handedness explicit) when it suits some other purpose.

6.4. **The left and right handed components are eigentates of $\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3$.** Note that $\gamma_5^{12} = 1$.

6.5. **The wave function of the electron satisfies the Dirac equation.** It predicts the correct value for the magnetic moment; the fine structure of the hydrogen atom. The problems with negative energy solutions are resolved by the hole theory of Dirac.

6.5.1. *The muon, the tau and the six quarks all satisfy the Dirac equation, but with vastly different masses.* The masses vary from that of the electron, 0.5MeV, up to 180 GeV for the top quark. More precisely the mass is a matix whose eigenvalues have these magnitudes. It turns out that in fact the quark mass matrix has complex eigenvalues: the phase represents a violation of *CP*. This clever way of explaning *CP*violation won Kobayashi and Mazkawa a Nobel prize last year. More on all this later.

7. Quantum Electrodynamics

So far we know the equations for the wave functions of spin $0,\frac{1}{2}$ and 1 particles. To understand the interactions of these particles with each other we must introduce non-linearities. The key is gauge invariance. A complete study of the resulting theory, quantum electrodynamics is well outside the scope of this course. Itzykson and Zuber *Introduction to Quantum Field Theory* is still a good reference. At a level closer to this course is the book by Kerson Huang, *Quarks and Leptons*.

Exercise. The Dirac equation implies the conservation of a current

$$j^{\mu} = \bar{\psi}\gamma^{\mu}\psi, \quad \bar{\psi} = \begin{pmatrix} \chi^* & \phi^* \end{pmatrix}$$

That is,

$$\partial_{\mu} j^{\mu} = 0.$$

This implies that

$$\frac{\partial}{\partial t}\int j^0 d^3x = 0.$$

Thus we can think of $Q = e\int j^0 d^3x$ as the electric charge and $j^0,\mathbf{j}$ as the charge and current densities respectively. The constant e is the electric charge of the electron (or whatever other particle to which we will apply this equation).Thus

7.1. The Maxwell's equations in the presence of electrons is.

$$\partial^{\mu} F_{\mu\nu} = e\bar{\psi}\gamma_{\mu}\psi. \tag{7.1}$$

Just as electrons create electric and magnetic fields,these fields must affect their motion. The change in the Dirac equation due to the presence of electric and magnetic fields is more subtle. Gauge invariance is the key to understanding this. Recall that under gauge transformation

$$A'_{\mu} = A_{\mu} + \partial_{\mu}\Lambda$$

where Λ is an arbitrary function. We want to preserve this symmetry when we introuduce A_{μ} into the Dirac equation. We must transform ψas well so that the changes in ψ and A_{μ}compensate for each other. Notice that if

$$\psi' = e^{ie\Lambda}\psi$$

$$\partial_{\mu}\psi' = e^{ie\Lambda}\left[\partial_{\mu}\psi + \left(ie\partial_{\mu}\Lambda\right)\psi\right]$$

Remark 11. Sensible people can handle the double use of the symbol e here. The e in the exponent is the electric charge and that below is the base of natural logarithms. Their values of course, have nothing to do with each other.

Thus in the combination below the derivatives of Λ cancel out:

$$\left[\partial_\mu - ieA'_\mu\right]\psi' = e^{ie\Lambda}\left[\partial_\mu - ieA_\mu\right]\psi$$

7.2. **The Dirac equation in the presence of an electromagnetic field is.**

$$\gamma^\mu\left[\partial_\mu - ieA_\mu\right]\psi = im\psi \tag{7.2}$$

Under a gauge transformation both sides are mutiplied by the same factor, so it cancels out. The pair of equations (7.1,7.2) describe Quantum ElectroDynamics (QED) of charged spin one half particles and photons.

7.3. **The equation of a charged massive spin zero particle is.**

$$\eta^{\mu\nu}\left[\partial_\mu - ieA_\mu\right]\left[\partial_\nu - ieA_\nu\right]\phi = -m^2\phi$$

This also follows using gauge invariance. Of course, here $\phi$is a scalar not a spinor.

7.4. **The proper interpretation of the equations of Quantum Electrodynamics involves renormalization.** The trouble is that the equations as describes above lead to infinities when quantum effects are fully included. They have to removed by a strange set of rules called "renormalization". These rules work remarkably well and agree with experiments to high precision: fifteen decimal point accuracy is the best science has ever achieved. Yet the correct mathematical formulation is still not clear. Dirac himself was very unsatisfied by this situation. New ideas in analysis are needed. But that is another story.

8. Atomic Nuclei

8.1. **Atomic Nuclei are made of neutrons and protons, of roughly equal masses. The protons carry an electric charge but the neutrons do not.** This explains why many chemical elements have isotopes: the chemical properties are determined by the number of electrons which is equal to the number of protons (the atomic number Z). The atomic mass number (A) is determined by the sum of neutrons and protons. The proportion of neutrons in a nucleus increases with the atomic number. The stable isotopes of hydrogen have either zero or one neutron; for Uranium, even the longest lived isotope has atomic mass number 238 compared with 92 protons. If there are too few neutrons, a nucleus is in danger of decaying by fission: breaking up into two nuclei each with smaller atomic number which can be stable with a smaller number of neutrons. If a nucleus has too many neutrons, it will emit an electron converting a neutron into a proton (beta decay). For example, Tritium decays into He_3 with a half-life of about seven years. This is why the warhead in hydrogen bombs have to be replenished every few years.

8.2. **The neutron is slightly heavier than a proton. A free neutron decays into a proton emitting an electron and an anti-neutrino with a half life of about ten minutes.** The neutrons inside nuclei do not decay as fast or are stable because their binding energy is greater than the $n-p$ mass difference. The decay is so slow because it is mediated by weak interactions and because there is very little room in phase space for it:

$$\frac{m_n-(m_p+m_e)}{m_n}\approx 0.1\%$$

Typical weak lifetimes (e.g. muon) is about a microsecond, already very long by the standards of particle physics. Typical strong decays has lifetimes (10^{-23} secs) (e.g., $\rho\to\pi\pi$).

8.3. **The proton is stable.** The lifetime is 6.6×10^{33} years: the longest confirmed lifetime of any particle. (Except perhaps the electron which no one seems to imagine can decay.)

8.4. **Both the neutron and the proton are spin half particles.**

8.5. **Although the neutron has no electric charge, it has a magnetic moment: suggesting that it is made of charged particles (quarks) whose electric charges cancel out.**

8.6. **The neutron has no electric dipole moment. This is a test of *CP* invariance of strong interactions.** The standard model predicts, through CP violation of weak interactions, a small electric dipole moment for the neutron. But it is too small to be observed.

8.7. **It became clear in the 1930s that the neutrons and protons have a strong attractive force.** Otherwise a nucleus would fall apart under Coulomb repulsion. The range of this force has to be about the size of a nucleus: so that all the neutrons and protons of the world don't fall into one gigantic nucleus. This can be explained if the force were to decrease exponentially with a range $a\sim 1$ fm (fm is a femtometer,10^{-15}m which conveniently is also called a Fermi).

8.8. **Yukawa suggested that the strong force is due to exchange of a massive particle, of mass $\mu\sim\frac{\hbar}{ac}\sim 100$ MeV.**

8.8.1. *It is useful for conversions to note that* $\hbar = 197.3269631(49) MeV\, fm.$. For simplicity we will for now ignore the fact that there are two kinds of particles (n and p) inside the nucleus. In the next section we will return to this doubling.

8.8.2. *The Klein-Gordon equation with a point source has an exponential decreasing static solution* $\phi = g\frac{e^{-\mu r}}{4\pi r}$. Here g is a constant (Yukawa coupling constant) that measures the strength of the field, analogous to electric charge for the Coulomb field.

8.8.3. *Similar to the photon which mediates the electromagnetic interactions, except the photon is massless and the Coulomb force has infinite range.*

8.8.4. *The exchange of photons can lead to repulsive as well as attractive interactions. Because the nuclear force is always attractive, the spin of the particle must be even. Yukawa suggested it must be spin zero.* Gravity is also always attractive: it is mediated by a hypothetical spin two particle.

8.9. **This particle has since been discovered and is called the π meson.** It has a mass of about 140 MeV. There was some confusion about its discovery. In fact another particle with a very close mass was discovered first in cosmic rays, called the muon. But the muon did not get absorbed by nuclei. It was Marshak (former Chair of our Department) who resolved the confusion: the muon is a lepton, a copy of the electron only with a higher mass. It has no strong interactions with the nuclei. But pions which are caused by cosmic ray collisions in the upper atmosphere decay into the muons, which are detected at lower altitudes.

8.10. **We can modify the Dirac equation to include a coupling to a spin zero particle.** One way to do this is to add the scalar field to the mass of the fermion:

$$\gamma^\mu \partial_\mu \psi = i(m + \phi)\psi$$

The scalar field of this type would be invariant under parity. This does not quite work because

8.10.1. *The pi meson is a pseudo-scalar.* This means that it changes sign under parity. Measuring the polarizations of the photons in the decay $\pi^0 \to 2\gamma$ tells us this fact. So the pion couples with opposite signs to the left and right handed components of the proton.

8.11. **The Dirac equation coupled to a pseudo-scalar is.**

$$\gamma^{\mu}\partial_{\mu}\psi = im\psi + g\gamma_5\phi\psi$$

$$\gamma_5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

At infinity $\phi \to 0$ and we get back the free Dirac equation. g is a constant that measures the strength of the interaction. Conversely the fermion serves as a source for the scalar field

8.12. **The wave equation of the scalar field is modified by the fermion as well.**

$$\left(\partial_0^2 - \nabla^2\right)\phi + m^2\phi = -g\bar{\psi}\gamma_5\psi$$

9. ISOSPIN

9.1. **Heisenberg suggested that the neutron and the proton should be thought of as two states of the same particle.** If we were to ignore the electromagnetic interactions (which are only about 1% as strong as the strong interactions) and weak interactions (which are even smaller) and also ignore the mass difference (less than 1%) the neutron and proton appear to be two distinct states of the same particle. Much like the spin up and spin down states of an electron. Heisenberg proposed an approximate symmetry similar to spin, *isospin,*which is broken by electromagnetic and weak interactions but is respected by strong interactions. There is a single particle called the nucleon whose isospin up state is the proton and the isospin down state is the neutron.(In addition, N carries spin $\frac{1}{2}$, so that each isospin state can be spin up or spin down.) Isospin is just like spin as far as mathematical properties go. But it is not related to angular momentum in any way. It is just another conserved quantity like electric charge. Unlike the electric charge, isospin conservation is not an exact law of nature: electromagnetic and weak interactions will violate it.

$$\psi = \begin{pmatrix} p \\ n \end{pmatrix}$$

The electric charge is equal to the component of isospin plus a constant shift:

$$Q = I_3 + \frac{1}{2}.$$

The free Dirac equation becomes

$$\gamma^{\mu}\partial_{\mu}\psi_a = im\psi_a, \quad a = 1,2$$

with $a = 1$ for the proton and $a = 2$ for the neutron. We ignore the fact that they have slightly different masses.

9.2. **The pi meson has isospin one.** Thus there are three possible isospin states: there are actually three pi mesons, with almost equal masses and electric charges $\pm 1, 0$.

$$\phi = \begin{pmatrix} \pi^+ \\ \pi^0 \\ \pi^- \end{pmatrix}$$

For them the formula for electric charge is

$$Q = I_3.$$

There is no shift, unlike for the fermions. In reality the mass of the charged pions are a few percent different from that of the neutral pion but we ignore that for now. The strong interactions are caused by exchanges of pions:

$$n \to p + \pi^-, \quad p \to n + \pi^+.$$

Because there may not be enough energy to create a free pion in a nucleus, the pions are often virtual: they exist only for a time of order $\frac{1}{\mu}$. But that is enough to produce the attractive interactions of range $\frac{1}{\mu}$.

9.3. **The Dirac equation of the nucleon including the coupling to the pion triplet is.**

$$\gamma^\mu \partial_\mu \psi = i(m + g\tau \cdot \phi \gamma_5)\psi$$

Here, $\tau$are Pauli matrices of isospin. For example $\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ interchanges n with p of the same spin.

To complete the story we must say how the nucleon is a source for the pion field:

9.4. **The equation for the pion field is.**

$$\left(\partial_0^2 - \nabla^2\right)\phi + \mu^2\phi = -g\bar{\psi}\tau\gamma_5\psi$$

9.5. **This Yukawa theory is an effective theory for strong interactions.** The proton and the pion are not elementary particles: both are made of quarks. The correct modern theory is Quantum ChromoDynamics. A "effective theory" is high energy physics jargon for an approximate theory, valid in some cases but not to be thought of as the final word. Thus, hydrodynamics is an effective theory for the collective theory of large numbers of molecules; chemists have theories of chemical bonds which approximate quantum mechanics; a planet can be treated as a point particle in celestial mechanics. In each of these cases, the complexities at short distances being ignored do not affect the behavior of the system at large enough scales. Actually every theory of physics except for the standard model and general relativity is an effective theory. Even today's exact theory of particle physics (the standard model) might be the effective theory of tomorrow when a more accurate theory is discovered. We can only speculate now about what that next level theory might be.

10. Breaking of Isospin

10.1. **We can combine the two formulas for electric charge, of the nucleon and the pion into a single formula.**

$$Q = I_3 + \frac{B}{2}$$

where B is the baryon number. Baryons are half integer spin particles which participate in strong interactions: the nucleon is the lighest baryon. The mesons are not baryons so they have $B = 0$. We will see later in terms of the quark model why there are so many baryons and mesons and why the mesons have baryon number zero.

10.2. **The masses of the nucleon and pion have a small dependence on isospin.**

$$m = m_0 + \varepsilon_1 I_3$$

The masses of the pion are more subtle: we will postpone that story till we talk about the quark model.

10.3. **Isospin breaking is best understood in terms of the quark model: the up and down quarks have different masses.**

11. The Sigma Model

An excellent reference for this lecture is the book "Chiral Dynamics" by Benjamin Lee.

11.1. **The pi mesons are much lighter than the nucleons.** 140MeV $<<$ 940Mev. To a first approximation we can treat the pions as massless. Nambu suggested that they are so light because they are the remnants of a spontaneously broken symmetry. This is one of the deepest ideas of modern physics and is going to play a role also in understanding the weak interactions (the Higgs et. al. mechanism) So we are going to digress to explain it. Let us go back to basics, classical mechanics.

11.2. **The ground state is the minimum of the potential.** For a classical mechanical system the state of least energy (ground state) has zero kinetic energy and is also at the minimum of the potential.

$$\frac{\partial V}{\partial \phi_a} = 0. \tag{11.1}$$

The second derivative of the potential at this point V'' is a symmetric matrix all of whose eigenvalues are ≥ 0 : this is the condition for a minimum. These eigenvalues also have another physical meaning: their square roots are proportional to the frequencies of oscillations around the ground state. Of special interest is the situation when there is a "zero mode": one or more of the eigenvalues of ,V'' is zero. What happens in that direction depends on higher derivatives. Suppose there is a symmetry of V that is broken by the ground state. For example, suppose

$$V(\phi) = \lambda(\phi^2 - a^2)^2$$

Any point on the sphere $|\phi| = a$ is a minimum: the choice of any one direction breaks the symmery of rotation by choosing a direction. In this case, the second derivative V'' at any minimum will have two zero modes: there are two eigenvectors with zero eigenvalue, corresponding to two infinitesimal rotations that change the direction of ϕ.

Exercise 12. Calculate the eigenvalues and eigenvectors when $\phi = (0,0,a)$

11.3. **There are zero modes whenever a continuous symmetry is spontaneously broken.** When we pick one of many solutions of (11.1) we can always move to a nearby solution at no cost of energy. Thus there are always "zero modes", which are excitations of arbitrarily low energy.

11.4. **In quantum mechanics the ground state is unique.** In quantum mechanics, the ground state is a linear superposition of all the points of equal energy: there is an equal probability of finding the system at any one point. Tunneling will prevent the ground state from being concentrated at any one point. Thus in quantum systems with a finite number of degrees of freedom we cannot have spontaneous symmetry breaking.

11.5. **Quantum systems with an infinite number of degrees of freedom can break symmetry spontaneously.** In these cases tunneling is suppressed by $e^{-\Omega}$ where $\Omega$is proportional to the volume of the system. This is why a magnet can develop spontaneous magnetization: all directions have the same energy. The quantum wave function of the magnet is concentrated along one direction because tunnelling to the other states is suppressed. In this case a continuous symmetry-rotation- is spontaneously broken.

11.6. **When a continuous symmetry is spontaneously broken there are waves of arbirarily small energy.** The frequencies are still given by the eigenvalues of V''. More precisely, taking into account of translation invariance they are given by

$$\sqrt{k^2 + V''}$$

where k is the wave number. To understand this, think of changing the magnetization by a smal amount. This change will propagate along the magnet as a wave: a "spin wave". If the change is only one of direction of the magnetization, we can make it as slowly varying in time as we want: the restoring force is zero as the second derivative vanishes.Such low frequency (and long wavelength) waves are a feature of all quantum systems with spontaneously broken symmetry.

Small oscillations in translation invariant many body systems correspond to bosonic particles: phonons for sound waves and "magnons" for magnetic waves. They are called Nambu-Goldstone bosons. Nambu's Nobel Prize was in part for this work.In a relativistic quantum theory, the Namu-Goldstone bosons are massless particles: mass being the smallest possible energy for a particle.

11.7. **Pions are the Nambu-Goldstone bosons of a spontaneously broken $O(4)$ symmetry.** In a magnet there are two zero eigenvectors for V'' ,corresponding to the two independent directions of a sphere. To get three massless bosons, we must imagine a sphere of fixed length vectors in four dimensions. Imagine a scalar field with four components $(\sigma, \phi_1, \phi_2, \phi_3)$ and a potential

$$V(\phi) = \frac{\lambda}{4}\left(\sigma^2 + \phi^2 - F_\pi^2\right)^2.$$

It is invariant under rotations in four dimensions $O(4)$. All the points on the three dimensional sphere S^3

$$\sigma^2 + \phi^2 = F_\pi^2$$

have the same energy. The constant F_π is the radius of this sphere: it has the value of about 150 MeV. It is called the pion decay constant because it also determines the probability of weak decays of the pion. (Weak decays are a quite different phenomenon which we will discuss later. It just so happens that this constant also shows up there and got its name that way.) We can define the direction chosen by the ground state (i.e., the vacuum) to be the fourth: the vacuum expectation value of σ is

$$<\sigma>=F_\pi.$$

The second derivative of the potential at this point is

$$V''=\mathrm{diag}(3\lambda F_\pi^2,0,0,0)$$

This leads to three massless particles (the pions) and on massive particle (called the sigma). The constant λ determines the mass of the sigma particle.

$$m_\sigma=\sqrt{3\lambda}F_\pi$$

11.8. **The sigma particle is very unstable.** It decays into pairs of pions with a very high probability: only after many years of work have we been able to identify it unambigously.

11.9. **The sigma is a scalar while the pions are pseudo-scalars.** Under parity

$$(\sigma,\phi)\mapsto(\sigma,-\phi)$$

We know the pions are pseudoscalars by measuring the polarization of the photon in the decay $\pi^0\to 2\gamma$. (More on this important decay later). The sigma decays into two pions so it must be a scalar. The parity properties are an important clue on how to complete the Yukawa theory to include the sigma.

11.10. **The Dirac equation of the nucleon including the coupling to the pion triplet is.**

$$\gamma^\mu\partial_\mu\psi=g(i\sigma+\tau\cdot\phi\gamma_5)\psi \tag{11.2}$$

In the ground state of the bosons, $\sigma=F_\pi,\phi=0$. Thus the mass of the nucleon is related to the Yukawa coupling constant and the pion decay constant

$$m=gF_\pi.$$

This is an important prediction of the sigma model, which can be verified experimentally: each of the numbers are independently measurable.

11.11. **The equation for the pion field and sigma fields are.**

$$\left(\partial_0^2 - \nabla^2\right)\phi = \frac{\partial V}{\partial \phi} - g\bar{\psi}\tau\gamma_5\psi \tag{11.3}$$

$$\left(\partial_0^2 - \nabla^2\right)\sigma = \frac{\partial V}{\partial \sigma} - ig\bar{\psi}\psi \tag{11.4}$$

The masses and interactions of the scalars are encoded in V. The field equations (11.2,11.3,11.4) descrbe the “linear sigma model”, also called the Gell-Mann-Levy model. The word linear here does not mean that the equations are linear: only that the scalar fields take values in the linear space R^4. This model was first invented to explain strong interactions. But with some tweaking-replace nucleons by quarks and scalar fields by Higgs bosons- it describes the weak interactions. Old ideas of physics don’t just die: they just get reincarnated.

11.12. **We can add a correction to the potential to include a mass for the pion.** This will break the $O(4)$ symmetry of the potential by picking out some direction. The main effect is to change the eigenvalues of V'' at the minimum, although the value of sigma field at the minimum is also changed slightly, as a higher order effect:

$$V(\sigma,\phi) = \frac{\lambda}{4}(\sigma^2 + \phi^2 - F_\pi^2)^2 + \varepsilon\sigma. \tag{11.5}$$

Since ε is very small, we should think of this as a constained minimization of $\varepsilon\sigma$ subject to the condition that $\sigma^2 + \phi^2 = F_\pi^2$: the condition for the dominant term to be a minimum.

11.12.1. *We can think of a pendulum as a mechanical analogue.* Imagine a pendulum suspended from a rod of length R. Actually the pendulum bob has some elasticity. So what we mean by this is that the elastic potential energy is minimized when the length is R. We can model this by an elastic potential energy

$$V_1 = \frac{\lambda}{4}(y^2 + x^2 - R^2)^2$$

where (x,y) is the position of the pendulum bob; the origin is at the point of suspension. All points of the circle $x^2 + y^2 = R^2$ have the same elastic energy. Adding the gravitational potential energy breaks this symmetry:

$$V(x,y) = \frac{\lambda}{4}(y^2 + x^2 - R^2)^2 + mgy$$

The first and second derivatives are

$$\frac{\partial V}{\partial x} = \lambda (y^2 + x^2 - R^2)x$$

$$\frac{\partial V}{\partial y} = \lambda (y^2 + x^2 - R^2)y + mg$$

$$V'' = \begin{pmatrix} \lambda (y^2 + x^2 - R^2) + 2\lambda x^2 & 2\lambda xy \\ 2\lambda xy & \lambda (y^2 + x^2 - R^2) + 2\lambda y^2 \end{pmatrix}$$

The condition for the minimum

$$\lambda (y^2 + x^2 - R^2)x = 0$$

implies that either $x = 0$ or $(y^2 + x^2 - R^2) = 0$. We must choose $x = 0$ because if $x^2 + y^2 = R^2$ exactly, there is no way to satisfy the second equation! Thus y is determined by

$$\lambda (y^2 - R^2)y + mg = 0$$

There is a root $y_- \approx -R$ which is the minimum; the other root is a maximum. By rewriting the equation as

$$y = \frac{R^2}{y} - \frac{mg}{\lambda y^2}$$

we can find an approximation to this solution as

$$y_- \approx -R - \frac{mg}{\lambda R^2}$$

The second term represents the small extension of the length of the pendulum due to the weight of the bob.

The calculation of the eigenvalues of V'' is simplified if we put

$$\lambda (y^2 - R^2) = -\frac{mg}{y}$$

into the formula for the second derivative *before* making the approximations:

$$V''(0, y_-) = \begin{pmatrix} -\frac{mg}{y_-} & 0 \\ 0 & -\frac{mg}{y_-} + 2\lambda y_-^2 \end{pmatrix} \approx \begin{pmatrix} -\frac{mg}{y_-} & 0 \\ 0 & 2\lambda y_-^2 \end{pmatrix} \approx \begin{pmatrix} \frac{mg}{R} & 0 \\ 0 & 2\lambda R^2 \end{pmatrix}$$

Thus one of the eigenvalues is large and the other small. The smaller eigenvalue leads to the usual formula $\sqrt{\frac{g}{R}}$ for the angular frequency of the gravitational oscillations of the pendulum. The other one cooresponds to the elastic oscillation of the pendulum and is usually ignored. In the case of the sigma model, the angular frequency of the usual oscillations of the pendulum are analogous to the pion mass; that of the elastic oscillations is like the σmass. The approximation of treating the length of the rod as fixed (rigid rod) is analogous to the nonlinear sigma model. The rigid pendulum has another kind of motion, in addition to the small oscillations. There are rotations that go all the way around: past the maximum of the potential. These are like the baryon (soliton) solutions of the nonlinear sigma model. There are two kinds of rotations, clockwise and anti-clockwise, which are roughly like the baryon and anti-baryon.

Exercise 13. Find the masses of the pi mesons predicted by (11.5), assuming that ε is small.

12. The Chiral Model

The sigma particle is so unstable that it is hardly there: it is a broad resonance whose width is not smaller than its mass. If we were to eliminate the σ field by constraining the potential to be at the minimum

$$\sigma^2 + \phi^2 = F_\pi^2$$

we would get the "nonlinear sigma model", also called the "chiral model". This is a much more elegant model with much interesting geometry and topology. Althought it is only an "effective theory"-an approximation to QCD- of strong interactions, it has much life left in it still. In recent years, the same ideas have found applications to new areas of physics: not only magnetism but also atomic condensates show broken symmetries. Since accelerator based experimental particle physics has not discovered no new phenomena in forty years, theorists are inspired more by the exciting new tools of atomic condensates. Ideas discovered validated there in turn should help us in resolving the puzzles posed by particle physics. It is a bad idea to look only at one set of tools when confronting the unknown: physics is a unified subject. Thus there has been a revival of interest in these old ideas of strong interaction physics.

12.1. **The self-interactions of the pions are described by making pion field takes values in a three dimensional sphere.** The radius of this sphere is a measure of the strength of this interaction: the larger the radius, the higher the pion momentum has to be for the interactions to kick in. This number is called F_π , the pion decay constant. (It is called that because it also controls the probability of weak decay of the pion: a quite different phenomenon we are ignoring for now.) Its value is about 150 MeV. A variational principle tells us how to change the wave equation (in stereographic co-ordinates)

$$\partial_\mu \left[\frac{\partial^\mu \xi}{\left(1+\frac{\xi^2}{F_\pi^2}\right)^2} \right] = 0$$

The actual derivation of these equations require some geometry of the sphere. We skip details here. The essential point is that the distance ds between two neighboring points on the sphere is given by

$$ds^2 = \frac{d\xi^2}{\left(1+\frac{\xi^2}{F_\pi^2}\right)^2}$$

in the co-ordinate system we are using. Such geometric modifications of the wave equations are a fashionable topic among mathematicians right now: they call this the "wave map" equation.

12.2. **At infinity the field much approach a constant.** Otherwise the total energy would be infinite. The limiting value may be chosen to the "North Pole" of the sphere which is the origin of our co-ordinate system

$$\lim_{|x|\to\infty} \xi = (0,0,0).$$

12.3. **Any continuous map $\phi : R^3 \to S^3$ which approaches a constant at infinity has an integer associated with it.**

$$Q = \frac{2}{\pi} \int \det \partial \xi d^3x$$

This is the integral of the Jacobian of the map divided by the volume of the sphere: it counts how many times R^3 is wrapped around the sphere, a kind of winding number. Under continuous evolution, this number would be conserved.

12.3.1. *A one dimensional analogue is the winding number of a map ξ : $R \to S^1$.*

$$w(\xi) = \frac{\xi(\infty) - \xi(-\infty)}{2\pi} = \frac{\int_{-\infty}^{\infty} \frac{d\xi}{dx} dx}{2\pi}$$

12.4. **Skyrme suggested that nucleons are solitons of pions.** So far we have thought of the nucleons and pions as independent particles. In QCD they are both made of quarks. Here is a bizarre third possibility: the nucleons are simply pion configurations with non-zero winding number. They are solutions to the pion wave equations that are very different from the ground state. Nonlinear equations often have such solutions that are stable and behave as particles on there on right, in addition to the small oscillations around the ground state. An important example are the solitary waves of the ocean: tsunamis. Pions are like the more familiar ripples on the surface of the oceans. These cost very little energy. But given a lot of energy (in the case of ocean wave it is provided by an earth quake) we can also create a huge new wave that looks like an entirely new particle. Skyrme suggested that nucleons are just such solitons of the pion field. At the time this idea was suggested (1960s) it was ignored because it was just too bizarre. But about mid eighties a group of us revived it and showed how it is in fact another way to look at QCD. Although QCD itself is too difficult to solve, in simpler cases where it is possible to solve gauge theories, we have shown that nucleons indeed are solitons.

12.5. **Solitons of bosonic fields can have half integer spin and be fermions.** This is because of very subtle topological facts. When an infinite number of particles act in unison, the usual rules of quantum mechanics can be overcome.

12.6. **The pion field equations have to be modified to include a short range repulsion.** Otherwise the soliton would shrink to zero and disappear. This modification does not affect the other parts of the theory. Mathematical proof that this avoids the singularity in the soliton solution is difficult but there has been recent progress.

12.7. **Properties of nuclei are well described by the soliton model.** There have been numerical solutions of the equations of the Skyrme model that reproduce quite well the messy phenemenology of nuclear physics.

12.8. **Physics is full of surprises even in areas that we think we understand.**

13. Hadrons

In the 1930s it looked as though we were on the verge of a simple description of the fundamental constituents of matter: the proton, the neutron,the pion, the electron (and possibly the neutrino) along with the photon would make up all matter. In the 1940s the muon was identified. (I. I. Rabi famously asked "Who ordered that?") Throughout 1940s, 1950s and 1960s experimentalists discovered a whole zoo of strongly interacting particles.

13.1. **Strongly interacting particles are collectively known as hadrons.** Electrons, neutrinos etc. are leptons, not hadrons.

13.2. **Hadrons of half integer spin are called baryons; those of integer spin are called mesons.**

13.2.1. *The baryon number B is defined to be equal to 1 for the half integer spin hadrons (baryons) and equal to zero for mesons.* Anti-baryons have baryon number minus one. The baryon number is just the atomic mas number of nuclear physics: the Deuteron has $B = 2$, the α particle has $B = 4$ and so on:

13.3. **The nucleon** $N = \begin{pmatrix} p \\ n \end{pmatrix}$ **is the lightest baryon with a mass of about** 940**MeV..**

13.4. **The pion** $\pi = \begin{pmatrix} \pi^+ \\ \pi^0 \\ \pi^- \end{pmatrix}$ **is the lightest meson at about** 140**MeV..**

In addition,

13.5. **There is a set of four baryons** $\Delta = \begin{pmatrix} \Delta^{++} \\ \Delta^+ \\ \Delta^0 \\ \Delta^- \end{pmatrix}$ **of spin and isospin both equal to** $\frac{3}{2}$**..** These decay into a nucleon and a pion. Since the nucleon has $I = \frac{1}{2}, J = \frac{1}{2}$ and the pion $I = 1, J = 0$ this strong decay respects both spin and isospin conservation.

13.6. **There is a set of three spin one mesons** $\rho = \begin{pmatrix} \rho^+ \\ \rho^0 \\ \rho^- \end{pmatrix}$ **of isospin 1.**

Their mass is ≈ 770MeV : about two thirds of the mass of a nucleon. They decay strongly into pions.

13.7. **The charges are related to isospin by the relation.**

$$Q = I_3 + \frac{B}{2}.$$

The shift of charge by a constant for the case of baryons (but not mesons) is explained by the quark model.

13.8. **There are hadrons of spins** $J = 0, \frac{1}{2}, 1, \frac{3}{2}2, \cdots$**.**

13.8.1. *If the spin is a half integer, so is the isospin.* Again explained by the quark model.

13.8.2. *As the spin grows the masses grow approximately proportionately.* The high mass hadrons are more and more unstable to decay to lower mass ones. Such very unstable particles are called resonances. As the numbers of hadrons grew into the hundreds, physicists accepted that there must be in principle an infinite number of them.

13.9. **String theory arose as an explanation for the infinitely rising spectrum of hadrons.**

13.9.1. *A string is a surface in space time whose action is proportonal to its area.*

13.9.2. *Nambu and Goto showed that this implies that the masses of its excited states are proportional to the angular momentum.* The Nambu-Goto model only allowed integer spins. Theirs was a "bosonic string theory".

13.9.3. *Supersymmetry was invented by Ramond to include fermions.* The idea did not quite work: no one has found a string theory that works in four space time dimensions. Finding the correct string theory of hadrons remains an important theoretical challenge.

13.9.4. *The 10 dimensional version of superstring theory is logically consistent and is a candidate for a quantum theory of gravity.*

14. Quarks

14.1. **All of the hadrons are bound states of more elementary particles known as quarks.** This gives a very simple explanation for the proliferation of hadrons.

14.1.1. *Quarks have spin $\frac{1}{2}$.*. This is reasonable for an elementary particle: it is the smallest non-zero value of spin allowed by quantum mechanics.

14.1.2. *Mesons are bound states of quarks and anti-quarks.* Which explains why they are bosons.

14.1.3. *Baryons contain three quarks.* It must be an odd number since baryons are fermions. Since there are baryons of spin $\frac{3}{2}$ and even parity (so that the orbital angular momentum is even) we need thre to be three quarks in a baryon. It follows that

14.1.4. *The baryon number of a quark is $\frac{1}{3}$.*. Anti-quarks have $B = -\frac{1}{3}$.

14.2. **There are a pair of quarks $\begin{pmatrix} u \\ d \end{pmatrix}$ forming an isospin $\frac{1}{2}$ system.** This explains why spin and isospin are equal for the lowest lying baryons.

14.2.1. *Their charges are given by $Q = I_3 + \frac{B}{2}$.*

$$Q_u = \frac{2}{3}, \quad Q_d = -\frac{1}{3}.$$

This then explains why this formula holds for all hadrons. Note that the piece proportional to B cancels out in mesons since anti-quarks have $B = -\frac{1}{3}$.

14.2.2. *Some group theory will allow us to get the spins and isospins of the hadrons out of those of the quarks.*

14.3. **But there is a surprise: color.**

15. THE STATIC QUARK MODEL

15.1. **A first approximation is to treat the hadrons as non-relativistic bound states of quarks.** So the different spin states have the same energy (spin-orbit coupling is a relativistic correction). This gives an $SU(4)$ symmetry: the four states of the quarks (spin up and down, isospin up and down) all have the same energy in this approximation.

15.2. **Quarks are fermions.** Spin half particles should satisfy the Dirac equation and satisfy the exclusion principle.

15.3. **But then how do we explain the Δ^{++}?** It should be impossible to put three up quarks into a state of spin $\frac{3}{2}$: all the spin are the same, violating the exclusion principle. One idea was that quarks obeyed some exotic statistics that violates the Pauli exclusion principle. That turned out to be wrong. Another possibility is that there is a extra degree of freedom.

15.4. **Each quark comes in three colors.** Thus there are three states for the up quark (not counting the spin states) and three for the down quark.The word color is used in a figurative way here: this quantum number has nothing at all to do with light: nothing to do with electromagnetism.

15.5. **There is an SU(3) symmetry corresponding to rotations among the color states.** Since quarks of different colors have the same masses, isospin, charges etc.

15.6. **Hadrons are color neutral.** Nucleons and mesons do not have this extra degree of freedom: we would have seen this in nuclear physics. Hadron states are invariant under the color SU(3) symmetry. This means that color cannot be directly measured:it can be inferred indirectly from properties of hadrons.

15.7. **The grond state of a three quark system must be a symmetric combination of three fundamental representations of $SU(4)$.** The wave function of quarks in a baryon is completely antisymmetric in color: that is the way to make it invariant under $SU(3)$ symmetry of color. The wave function of fermions is anti-symmetric overall. Thus in spin and isospin it must be symmetric. (We are assuming that the orbital angular momentum is zero for the lowest lying states). Thus we can calculate the number of independent states of the baryon by ignoring color and pretending that quarks are bosons:there are $\frac{4(4+1)(4+2)}{3!} = 20$ such states. These can be split into

$I = \frac{3}{2}, J = \frac{3}{2}$ and $I = \frac{1}{2}, J = \frac{1}{2}$ states. The first are the Δ and the second set the nucleons. There are $4 \times 4 = 16$ states for the Δ and $2 \times 2 = 4$ states for the nucleon which add up to twenty.

15.7.1. *The number of independent states of a system of N bosons, each with M states, is* $\frac{M(M+1)\cdots(M+N-1)}{N!}$. There are many ways to establish this formula.

To begin with, it holds for $M = 1$. There is exactly one indepedent way of occupying a single state with N bosons : we put them all into that one available state. More generally, suppose we put N_1 bosons in state 1, N_2 in state 2, and so on. The number of independent states $b_N(M)$ we seek is just the number of solutions to the equation

$$N = N_1 + N_2 \cdots + N_M$$

where each $N_i = 0, 1, \cdots$. Thus the generating function is

$$\sum_{N=0}^{\infty} b_N(M) x^N = \sum_{N_i=0}^{\infty} x^{N_1+N_2\cdots+N_n}$$

But

$$\sum_{N_i=0}^{\infty} x^{N_1+N_2\cdots+rN_n} = \sum_{N_1=0}^{\infty} x^{N_1} \sum_{N_2=0}^{\infty} x^{N_2} \cdots \sum_{N_M=0}^{\infty} x^{N_M} = \frac{1}{(1-x)^M}.$$

since each factor is a geometric series. Now expanding the r.h.s. as a binomial series

$$(1-x)^{-M} = \sum_{N=0}^{\infty} \frac{M(M+1)\cdots(M+N-1)}{N!} x^N$$

which proves the result.

15.7.2. *The states of a bosonic system where can be represented as polynomials in complex variables.* The degree of the polynomial is the number of bosons; the nmber of variables is the number of states available to each boson. Thus 1 is the empty state (vacuum) z_i $i = 1, \cdots M$ are the one particles, and so on. The inner product of the states is determined by declaring that the following is an orthonormal basis

$$|N_1, N_2, \cdots N_M\rangle = \frac{z^{N_1}}{\sqrt{N_1!}} \frac{z^{N_2}}{\sqrt{N_2!}} \cdots \frac{z^{N_M}}{\sqrt{N_M!}}$$

This is the coherent state description of a harmonic oscillator; the connection to bosonic states is important also in quantum optics (Glauber and Sudarshan). It is part of a general technique known as second quantization.

15.7.3. *If the number of colors where some odd number N , there would have been baryons with $I = J$ for $I = \frac{1}{2}, \frac{3}{2}, \cdots \frac{N}{2}$.* Show that

$$\sum_{I=\frac{1}{2},\frac{3}{2},\cdots\frac{N}{2}} (2I+1)^2 = \frac{(N+1)(N+2)(N+3)}{3!} = \frac{4(4+1)(4+2)\cdots(4+N-1)}{N!}$$

15.8. **The proton and the neutron have anamolous magnetic moments.**

15.8.1. *The magnetic moment of the proton is roughly* 3 *times the nuclear magneton.* $\mu_p = 1.410\ 607\ 61\ (41) \times 10^{-26} JT^{-1}$ .The nuclear magneton $\frac{e\hbar}{2m} = 5.050\ 783\ 24(13) \times 10^{-27} JT^{-1}$. (These values are from the NIST data base online.) For a fundamental spin half particle this ratio is predicted by the Dirac equation to be 2: this works for the electron and muon but not the proton .

15.8.2. *The ratio of magnetic moments of the neutron to the proton is* $\frac{\mu_n}{\mu_p} \approx -\frac{2}{3}$. NIST lists the values is a strange way: take the ratio of the two gyromagnetic ratios, $-\frac{1.832\ 471\ 85(43)\times 10^8}{2.675\ 222\ 099(70)\times 10^8}$ to get this number. The negative sign means that the neutron magnetic moment is pointed opposite to its spin, while the proton's is parallel. The Dirac theory would predict the neutron not to have a magnetic moment at all, since it is electrically neutral.

15.9. **The quarks model predicts the correct magnetic moments for the nucleons.** The accuracy is a few percent.

15.9.1. *Assume that the quarks have the magnetic moments of fundamental particles.* The Dirac equation predicts their magnetic moments given the mass and charges. We know the charges already.

15.9.2. *In the static quark model the mass of the quark is approximately a third of the nucleon mass.* Neglects the binding energy and kinetic energy as well as spin/isospin dependence. It is convenient to suppose that there are N quarks in a baryon. Of course $N = 3$, but it is useful to the calculation for a general value of N. It teaches us how it depends on N: also we are less likely to make arithmetic mistakes.

The charge and magnetic moment of a quark are (using τ for the Pauli matrices of isospin)

$$Q = e\left[\frac{B}{2} + I_3\right] = e\left[\frac{1}{2N} + \frac{\tau_3}{2}\right]$$

For each quark (a runs over $1,2\cdots N$ since there are N quarks in the baryon) the component of magnetic moment along some direction (say third) is

$$\mu_a = 2\frac{eQ_a\hbar}{m_a}\frac{\sigma_{3a}}{2} = 2\frac{e\hbar}{\frac{m}{N}}\left[\frac{1}{2N}+\frac{\tau_{3a}}{2}\right]\frac{\sigma_{3a}}{2} = \frac{e\hbar}{m}\left\{\frac{\sigma_{3a}}{2}+\frac{N}{2}\tau_{3a}\sigma_{3a}\right\}.$$

We must sum over all the quarks in a baryon to get its magnetic moment

$$\mu = \frac{e\hbar}{m}\sum_{a=1}^{N}\left\{\frac{\sigma_{3a}}{2}+\frac{N}{2}\tau_{3a}\sigma_{3a}\right\}$$

The first term is independent of isospin and the second depends on isospin. This matrix on the space of states of the baryon will describe the magnetic moments of states such as the nucleon and the Delta, as well as magnetic transitions such as $\Delta \to N\gamma$.The first term is easy:

$$\sum_{a=1}^{N}\frac{\sigma_{3a}}{2} = J_3$$

J_3 being just the baryon angular momentum along the third direction.

$$\mu = \mu_0+\mu_1, \quad \mu_0 = \frac{e\hbar}{m}J_3, \quad \mu_1 = \frac{e\hbar}{m}\frac{N}{2}\sum_{a=1}^{N}\tau_{3a}\sigma_{3a}$$

We need the matrix elements for the neutron and the proton.

15.9.3. *The states of the baryon are in one-one correspondence with Nth degree polynomials in four complex variables.* This is just a way of thinking of states of a system of N bosons, each with four states. We can conveniently think of the four variables as the elements of a 2×2 matrix. Under isospin and spin this matrix transforms as

$$z \mapsto gzh^{\dagger} \; g,h \in SU(2)$$

The orthonormal basis is given by

$$|n_{11},n_{12},n_{21},n_{22}> = \frac{z_{11}^{n_{11}}}{\sqrt{n_{11}!}}\frac{z_{12}^{n_{12}}}{\sqrt{n_{12}!}}\frac{z_{11}^{n_{21}}}{\sqrt{n_{21}!}}\frac{z_{22}^{n_{22}}}{\sqrt{n_{22}!}}$$

where the occupation numbers can take values $0,2\cdots$ subject to

$$n_{11}+n_{12}+n_{21}+n_{22} = 3$$

The left action is isospin and the right action is spin (say). Then $\det z$ is invariant under both. The polynomials that describe the neutron and proton states are $z(\det z)^k$ where $N = 2k+1$. You can check that these sates transform with spin and isospin both equal to $\frac{1}{2}$. In this point of view

$$D \equiv \sum_{a=1}^{N} \tau_{3a}\sigma_{3a} = \sum_{ij} A_{ij} z_{ij} \frac{\partial}{\partial z_{ij}}, \quad A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

(Verify that this is correct on the one-particle states) The differential operator will mix the nucleon states with the other states (the magnetic moment can induce the transitions such as $\Delta \to N$.) For $N = 3$,(states are not normalized to have length one) the spin up states of the proton and the neutron are

$$|p\rangle = z_{11}(z_{11}z_{22} - z_{12}z_{21}) = z_{11}^2 z_{22} - z_{12}z_{21}z_{11}$$

$$|n\rangle = z_{21}(z_{11}z_{22} - z_{12}z_{21}) = z_{11}z_{22}z_{21} - z_{12}z_{21}^2$$

$$< p|p >= 2+1 = 3 =< n|n >$$

$$D|p\rangle = 3z_{11}^2 z_{22} + z_{12}z_{21}z_{11}$$

$$D|n\rangle = z_{11}z_{22}z_{21} + 3z_{12}z_{21}^2$$

$$< p|D|p >= 3(2!) - 1 = 5$$

$$< n|D|n >= 1 - 3(2!) = -5$$

$$\frac{< p|D|p >}{< p|p >} = \frac{5}{3}$$

$$\frac{< n|D|n >}{< n|n >} = -\frac{5}{3}$$

$$< \mu >_p = \frac{e\hbar}{m}\left[\frac{1}{2} + \frac{3}{2}\frac{5}{3}\right] = 3\frac{e\hbar}{m}$$

$$< \mu >_n = \frac{e\hbar}{m}\left[\frac{1}{2} - \frac{3}{2}\frac{5}{3}\right] = -2\frac{e\hbar}{m}$$

16. K MESONS

16.1. $K^{\pm}$ **are pseudo-scalar , isospin** $\frac{1}{2}$ **particles of mass** 494 **Mev that only decay by weak interactions.** They were called "strange particles" when they were discovered. What was strange about them is that they were unusually long lived ($10^{-8}s$): suggesting that they carry a quantum number that is approximately conserved. This number was called "strangeness" (Gell-Mann). They are each other's anti-particles. K^{+} was assigned strangenesss $S=+1$and therefore K^{-} would have $S=-1$. Unlike $\pi^{\pm}$ the $K^{\pm}$ have form an isospin $\frac{1}{2}$ doublet.

16.2. $K^{0},\bar{K^{0}}$ **is another pair of pseudo-scalar , isospin** $\frac{1}{2}$ **particles of mass** 498 **MeV that are also stable under strong interactions.** K^{0} has $I_3=-\frac{1}{2}, S=1$ and $\bar{K^{0}}$has $I_3=\frac{1}{2}, S=-1$ .The charges of all the K-mesons can be fit by changing the formula for electric charge (Gell-Mann-Nihijima)

$$Q=I_3+\frac{B+S}{2}$$

16.3. **The new quantum number is counts the net number of a new kind quark, the strange quark.** By a twist of fate, the strange quark has $S=-1$ and the strange anti-quark hs $S=+1$. It has baryon number $\frac{1}{3}$ like the u and d quarks. From the above formula we see that its electric charge is $-\frac{1}{3}$. That is the same charge as the d quark. Thus we have the constituents of the Kaons:

$$K^{+}=\bar{s}u,\ K^{-}=\bar{u}s,\ K^{0}=\bar{s}d\ \bar{K^{0}}=\bar{d}s$$

16.4. **There is also a neutral pseudoscalar meson that has strangeness zero and isospin zero.**

$$\eta^{0}=\bar{s}s$$

with a mass $\approx$ 548MeV. It decays mostly into 2γ which can be thought of as the strange quark and anti-quark annihilating each other. A more accurate description of the η^{0} includes mixing with $\bar{u}u$ and $\bar{d}d$. More on mixing later.

16.5. **The** s **quark is heavier than the** u**and** d **quarks.** which explains why particles that contain it as a few hundred MeV heavier than corresponding particles made from u and $d$quarks alone.For example, $m_{K^{+}}-m_{\pi^{+}}\approx$ 350 MeV.Recall that the d quark is slightly heavier (by a few MeV) than the u quark to explain the neutron-proton mass difference. For strong interactions, the three quarks behave the same way. If we also ignore their mass differences, the isospin symmetry is enlarged to a symmetry that rotates three quarks into each other. Since these transformations can involve complex matrices, the symmetry must involve 3×3 complex matrices. One natural choice is to generalize the $SU(2)$ of isospin to $SU(3)$. This is not the only possibility: there several rank two Lie groups (with two commuting quantum numbers such as I_3 and S) to choose from. But $SU(3)$ is what worked.

16.6. **The 8 pseudo-scalar mesons form a representation of $SU(3)$.** Analogous to the way the pions form a three dimensional representation of $SU(2)$.

17. $SU(n)$

17.1. **$U(n)$ is the set of $n \times n$ unitary matrices.** Unitary means that

$$gg^{\dagger} = 1$$

A set of matrices whose products and inverses are also contained within it is called a group. Thus $U(n)$ is a group.

Unitarity already implies that the determinant of g is a complex number of magnitude one. (Prove it.)

17.2. **$SU(n)$ is the set of $n \times n$ unitary matrices which are of determinant one.**

17.2.1. *The familiar example from spin and isospin is the group $SU(2)$.*

17.2.2. *The case of $SU(3)$ is of much interest as well in particle physics.*

17.3. **A unitary matrix that is infinitesimally close to the identity is of the form $g = 1 + iA$ where A is hermitean.** For

$$gg^{\dagger} = (1 + iA)(1 - iA) = 1 + \mathrm{O}(A^2)$$

17.3.1. *Matrices of determinant one which are inifinitesimally close to the identity are of the form $1 + iA$ with tr A=0.* Prove that

$$\det[1 + A] = 1 + \mathrm{tr\ A} + \mathrm{O}(A^2)$$

17.4. **The set of traceless hermitean matrices is called $su(n)$.** We use lower case letters to denote infinitesimal matrices.

17.5. **A basis for $su(2)$ is provided by the Pauli matrices.**

17.5.1. *More precisely any traceless hermitean* 2×2 *matrix can be written as.*

$$A = a_1\frac{\sigma_1}{2} + a_2\frac{\sigma_2}{2} + a_3\frac{\sigma_3}{2}$$

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

The coefficients a_i are real. The factor of i is needed to make the term anti-hermitean. the factor of $\frac{1}{2}$ is a convention, which assures that

$$\mathrm{tr}\, A^\dagger A = a_1^2 + a_2^2 + a_3^2.$$

17.5.2. *We can think of the up and down quarks as eigenstates of* σ_3*with eigenvalues* ± 1.

17.5.3. *As more quarks were added, the approximate symmetry of isospin came to be enlarged to larger groups. The up, down and strange quarks correspond to* $su(3)$.

17.6. **The number of linearly independent elements of $u(n)$ is n^2.** A hermitean mtrix has n^2 independent components: there are n real entries along the diagonal and $\frac{n(n-1)}{2}$ complex numbers above the diagonal. The entries below the diagonal are not independent because they are just complex conjugates of the ones above, so the total is $n + 2\frac{n(n-1)}{2} = n^2$. Since an anti-hermitean matrix is simply i times a hermitean one, its number of independent components is also n^2.This is called the dimension of $u(n)$.

17.6.1. *The dimension of* $su(n)$ *is* $n^2 - 1$. The condition of being traceless imposes one condition among the diagonal entries, so the number of independent components of $su(n)$ is $n^2 - 1$.

17.6.2. *The dimension of* $su(3)$ *is* 8.

17.7. **The Gell-Mann matrices provide a basis for $su(3)$.**

$$A = a_1\frac{\lambda_1}{2} + a_2\frac{\lambda_2}{2} + a_3\frac{\lambda_3}{2} + a_4\frac{\lambda_4}{2} + a_5\frac{\lambda_5}{2} + a_6\frac{\lambda_6}{2} + a_7\frac{\lambda_7}{2} + a_8\frac{\lambda_8}{2}$$

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}$$

$$\lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}$$

$$\lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

These are normalized such that

$$\mathrm{tr}\, \lambda_\alpha \lambda_\beta = 2\delta_{\alpha\beta}.$$

17.7.1. *λ_3 and λ_8 are diagonal. They are related to isospin and strangeness of quarks.*

Exercise 14. Derive the commutation relations of $su(3)$in the Gell-Mann basis. That is, write the commutators $\left[\frac{\lambda_\alpha}{2}, \frac{\lambda_\beta}{2}\right] = i f_{\alpha\beta\gamma} \frac{\lambda_\gamma}{2}$ as linear combinations ofthe Gell-Mann matrices. Identify subsets of generators that are transformed among each other by the Pauli matrices.

18. Gell-Mann-Okubo Formula

18.1. **The most obvious consequence of the strange quark is that there is a spin $\frac{3}{2}$ baryon sss.** This is analogous to $ddd = \Delta^{--}$ and so should be negatively charged. It is called the Ω^-. The spin is $\frac{3}{2}$ because the quark is a fermion and its wavefunction is anti-symmetric in color. Same argument as for uuu or ddd. Obviously, Ω^- should not carry isospin and has strangeness -3.

18.2. **Next there should be an isospin $\frac{1}{2}$ pair ssu and ssd of charges zero and -1 respectively.** These are called Ξ^{*0}, Ξ^{*-}. They have strangeness $S = -2$. The star is to distinguish it from a similar particle of spin $\frac{1}{2}$.

18.3. **There is an isospin 1 triplet suu, sud, sdd of charges $1, 0, -1$.** These are called $\Sigma^{*+}, \Sigma^{*0}, \Sigma^{*-}$.

18.4. **Along with the original quartet $\begin{pmatrix} \Delta^{++} \\ \Delta^{+} \\ \Delta^{0} \\ \Delta^{-} \end{pmatrix} = \begin{pmatrix} uuu \\ uud \\ udd \\ ddd \end{pmatrix}$ we get a set of ten spin $\frac{3}{2}$ baryons.** To a good approximation we can think of the $u$and $d$quarks as having the same mass, but the s quark is heavier so that $SU(3)$is broken down to its isospin subgroup.We should expect that the $\Delta$are all of the mass. The three Σ^*, would of somewhat larger mass, then $\Xi^*$and

Ω^-. The static quarks model would say that each time we add replace a u or d quark by an s quark we are increasing the mass of a particle by some fixed amount: the $s-u$ mass difference. Thus we get a particular case of the Gell-Mann-Okubo mass relations

$$m_{\Sigma^*} - m_\Delta = m_{\Xi^*} - m_\Sigma = m_{\Omega^-} - m_{\Xi^*}$$

At the time that this relation was discovered (by deeper group theoretic arguments rather than the static quark model) it was known that (all in MeV)

$$m_\Delta = 1230,\ m_{\Sigma^*} = 1385,\ m_{\Xi^*} = 1530$$

but the Ω^- had not even been seen yet. Thus the first equality is a post-diction that can be verified

$$m_{\Sigma^*} - m_\Delta = 155 \approx 145 = m_{\Xi^*} - m_\Sigma$$

And there is a prediction of the mass of the Ω^-.

$$m_{\Omega^-} \approx 1675$$

The discovery of the Ω^- with a mass of 1672MeV at Brookhaven was spectacular confirmation of the Gell-Mann-Okubo relations.

18.5. **Similar mass relations can be derived for the spin $\frac{1}{2}$ baryons.** There are eight of them, in one-one correspondence with the meson octet, with names like N, Σ, Ξ. The mass relations are also verified here. The mass relations for the pseudo-scalar mesons are not so easy to derive due to spontaneous breaking of chiral symmetry which makes them lighter than expected.

18.6. **To go deeper into this subject we will need the representation theory of $SU(3)$.** This is analogous to the theory of angular momentum operators. See the course “Symmetries in Physics” .

19. Quarkonium

19.1. **The discovery of the spin one particle J/ψ at a mass of 3.096 GeV led to the immediate acceptance of the quark model.** Theorists had already predicted the existence of a fourth quark (charm) to be the upper counterpart to the strange quark: it was needed to cancel certain unobserved phenomena (flavor changing neutral currents). Identifying $J/\psi = \bar{c}c$ gave an immediate explanation for its excited states. Indeed non-relativistic quantum mechanics suffices to explain the spectroscopy of the charmonium states which were soon found.

19.2. **Subsequently the Υ state was found at spin one and mass** 9.46 **GeV.** This was identified with the $b$quark. Again the spectroscopy is explained by quantum mechanics with linear potential between b and $\bar{b}$.

19.3. **The discovery of the top quark with a mass** 175**GeV completes the quark periodic table as far as we know.** The weak decay of the top quark occurs so fast that it does not form $\bar{t}t$ bound states through strong interactions. So there is no quarkonium here.

20. Potential Model For Quarks

20.1. **For the c and b quarks, a linearly rising potential gives a good fit to spectra.**

$$-\frac{1}{2m}\nabla^2\psi + V(r)\psi = E\psi, \quad V(r) = a|r|$$

with

$$a \approx 100\mathrm{MeV}$$

fits the data. Recall that this equation can be solved using Airy's equation.

20.2. **A linear potential fits with string theory.**

20.3. **The decay of J/ψ and Υ well explained as the production of three gluons which then becomes hadrons.** The spin one state decays into an odd number of gluons, due to parity conservation.

21. Quantum ChromoDynamics

21.1. **QCD is the fundamental theory of strong interactions.** It is named Quantum Chromodynamics by analogy to Quantum ElectroDynamics (QED), the theory of electromagnetism. Like QED, it describes a spin $\frac{1}{2}$ particle (quark) interacting with a massless spin 1particle (the gluon). Instead of the electric charge, the source of the interaction is color.

21.2. **Each quark (u,d etc.) comes in three colors, leading to an exact $SU(3)$ symmetry.** Recall that this is needed to reconcile the $\Delta^{++} = uuu$ bound state with the fermionic nature of quarks. The symmetry is exact because these three states have the same masses: unlike the approximate $SU(3)$ symmetry of 'flavor' that combines u,d,s.

21.3. **Yang-Mills Theory describes the self-interaction of spin one particles of a non-abelian symmetry.** The photon does not couple with itself: two photons pass right though each other. (There is a small chance of scattering through pair craetion, but that is not a fundamental interaction). With a non-abelian group this is not the case any more.

21.4. **The Self-Interactions arise through commutators.**

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + i[A_\mu, A_\nu]$$

The last term would be zero if the components were just real numbers as in Maxwell's theory.

21.5. **The field equations are.**

$$\partial^\mu F_{\mu\nu} + i[A^\nu, F_{\mu\nu}] = e\bar{\psi}\gamma_\nu\psi, \quad \gamma^\mu \left[i\partial_\mu - eA_\mu\right]\psi + m\psi = 0$$

A detailed study of Yang-Mills theories such as QCD is outside the scope of this course. Some of the deepest theoretical physics ever developed arise here. Despite that even deeper new ideas are still needed to fully understand QCD. We have not even scratched the surface. The most progress has been though computer simulations. See the book by K. Huang "Quarks and Leptons" for a more.

21.6. **Divergences arise when we study this theory at the quantum level.**

21.7. **'t Hooft and Veltmann showed that these divergences can be removed by a procedure known as renormalization.** This generalizes the previous renormalization of QED due to Tomonaga, Schwinger and Feynman.

21.8. **Surprisingly, QCD has the property that interactions become weaker at shirt distances.** Only a handful of theories are asymptotically free. Other than QCD the examples are in condensed matter physics (e.g., the Kondo problem).

22. BETA DECAY

22.1. **Heavy nuclei decay by emitting beta rays.** Nuclei were found to emit three kinds of radiation: α, β and γ. Now we know that the α particle is the nucleus of He_4. The γ particle is just a photon. And the β particle is an electron.

An apparent lack of energy conservation in beta decay was explained as due to the simultaneous emission of a massless chargeless particle (the antineutrino in moern terminology). We know that a nucleus is a composite of protons and neutrons. In terms of this a beta decay occurs when

$$n \to p + e^- + \bar{\nu}_e$$

22.2. **A neutron converts to a proton while emitting an electron and an anti-electron-neutrino.** But the nucleon is made of quarks too: $\begin{pmatrix} p \\ n \end{pmatrix} = \begin{pmatrix} uud \\ udd \end{pmatrix}$. Thus

22.3. **An down quark converts to an up quark emitting an electron and an anti-electron-neutrino.** An even deeper explanation is provided by the Glashow-Salam-Weinberg theory: the $d \to u + W^-, \quad W^- \to e^- + \bar{\nu}_e$. The W^- is a spin one particle of mass (80GeV) much higher than any of the particles involved. Yet it still can be produced because quantum mechanics allows energy conservation to be temporarily violated in intermediate transitions of a reaction. The probability of the reaction is supressed by a factor of $\frac{1}{M_W^4}$ which explains why the beta decay is a “weak interaction”: much rarer than strong decays such as $\Delta^+ \to p\pi^0$.

22.4. **The muon decays to the electron through the weak interaction.** In this case

$$\mu \to e + \bar{\nu}_e + \nu_\mu$$

Again we understand this today as a two step process

$$\mu \to \nu_\mu + W^-, \quad W^- \to e^- + \bar{\nu}_e$$

the second step is the same that in beta decay.

22.5. **There are three kinds of neutrinos: the electron neutrino, the muon neutrino and the tau neutrino.** There is some mixing among them, but to a good approximation, the three kinds of fermions numbers are separately conserved. We postpone the questions of quark mixing and neutrino oscillations.

22.6. **To a good approximation the weak interactions can be understood as the interaction of $W^\pm$with fermion doublets.**

$$\begin{pmatrix} u \\ d \end{pmatrix}, \begin{pmatrix} c \\ s \end{pmatrix}, \begin{pmatrix} t \\ b \end{pmatrix}$$

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}$$

Only the top quark has a high enough mass to produce a real W in its decay $t \to bW^+$. In all other cases the W is virtual: its creation violates enery conservation and is suppressed by the large mass. Because the charge of the particles changes (u to $d$etc.) in such decays these are called the “Charged Current” interactions. This suggests that there is an $SU(2)$ symmetry in the weak sector. Unlike the isospin of the strong interactions, this would apply also to leptons.

23. V-A THEORY

23.1. **Marshak and Sudarshan showed that only the left handed part of fermions take part in charged current interactions.** Once parity violation was discovered in weak interactions, it became possible that the left and right handed components could couple with different strengths to weak interactions. This is to be contrasted with the electromagnetic interactions which, because of parity conservation, couple equally to the left and right handed components. It was a revolutionary suggession of Marshak and Sudarshan that the beta decay only involves the left handed component: parity violation is maximal. This would allow the neutrino to be purely left handed: have half as many degrees of freedom as the electron. At that time this was called the $V-A$ theory: the left handed component of a current is the vector minus axial component.

In the middle seventies neutral current interactions were discovered as well $\nu_e + p \to \nu_e + p$. (A neutrino beam is a sent into a bubble chamber full of liquid hydrogen. Occasionally you see the track of a proton recoiling: the neutrinos don't leave tracks.)

23.2. **Neutral currents are weak interactions caused by the Z boson.** The Z boson is a spin one particle of mass 91GeV. Again, the Z boson produced is virtual and so the strength of the neutral weak interactions is suppressed by $\frac{1}{M_z^4}$. The Glashow-Salam-Weinberg theory predicted its mass, once the strength of neutral currents was determined. Neutral currents couple to a linear combination of left and right handed components. More precisely, there are two neutral spin bosons, one of which couples only to the left handed component and the other only to the rigt handed component of the fermions. The photon is the linear combination of these two which respects parity and the Z boson is the orthogonal linear combination. It is natural now to combine the neutral boson coupling to the left handed fermions along with the W bosons into a triplet

$$\begin{pmatrix} L^+ \\ L^0 \\ L^- \end{pmatrix}$$

24. GLASHOW-SALAM-WEINBERG THEORY

24.1. **Thus we can account for weak interactions by having three vector bosons that couple to the left handed components of fermions.** The left handed components of fermions fall naturally into weak isospin doublets

$$\begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L, \begin{pmatrix} t \\ b \end{pmatrix}_L$$

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L$$

We can couple them to the gauge bosons of an $SU(2)$Yang-Mills theory: the corresponding bosons are $L_{\mu a}$ above.

24.2. **There is another vector boson that couples to the right handed component.** The right handed components (the neutrino has no right handed component)

$$u_R, d_R, c_R, s_R, t_R, b_R$$

$$e_R, \mu_R, \tau_R$$

couple to R . This is the symmetry $U(1)$. Thus the complete theory has gauge invariance under $U(2) \approx SU(2) \times U(1)$

24.3. **The photon and the Z boson are linear combinations of L^0 and R.** The mixing angle θ_W of this linear combination (the Weinberg angle) determines the ratio of the masses of W and Z.

$$Z = \cos\theta_W R + \sin\theta_W L^0$$

$$V = -\sin\theta_W R + \cos\theta_W L^0$$

$$\cos\theta_W = \frac{M_W}{M_Z}, \quad \theta_W \approx 30^\circ.$$

24.4. **The Higgs et. al. mechanism allows the photon to be massless while the other three vector bosons are massive.** At the time Glashow proposed the above theory of weak interactions, it was known that gauge invariance implies massless vector bosons:that would not work for weak interactions. Breaking gauge invariance 'by hand' as Glashow did led to various inconsistencies: the divergences in the theory did not cancel (not renormalizable). This bothered Goldstone, Salam and Weinberg particularly: the leading experts on renormalization theory. A way of breaking gauge invariance without losing renormalizability was found by Higgs and independently by Englert-Brout and Guralnik-Hagen-Kibble. This was at the cost of adding a neutral spin zero particle: the 'Higgs boson'. Salam and Weinberg used this mechanism to find the unified theory of weak interactions. 't Hooft and Veltman later proved the renormalizability and invented methods for doing accurate ("higher loop") calculations with the theory. The discovery of neutral currents and then the $W^{\pm}, Z$ bosons exactly as predicted by GSW theory gave spectacular confirmation of the theory. The last piece of the puzzle is the discovery of the Higgs boson, which is a main goal of the LHC. Theory makes no prediction for its mass, except that it should be less than a TeV.

24.5. **Quarks and leptons get their masses through the Higgs mechanism as well.** The Yukawa coupling emerges again as the way quarks and leptons couple to the Higgs boson.

24.6. **It turns out that the mass matrices of quarks and neutrinos are not diagonal.** That is, the eigenstates of weak interaction (the coupling matrix to the gauge bosons) and of the mass matrix (the Yukawa coupling matrix) are not the same: these matrices do not commute. This leads to very important phenomena of quark mixing and of neutrino oscillations. As it is the focus of the current experimental research, we will study deeper the Higgs mechanism and fermion mixing in separate lectures.

25. Lagrangian Formalism

25.1. **Hamilton's Variational Principle gives a concise formulation of equations of motion.** Define the Lagrangian L to be some function of position and velocity; and action to be its integral:

$$S = \int L(q, \dot{q}) dt$$

The condition that the action be stationary w.r.t. to small changes in q leads to the condition

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0$$

With the choice

$$L = \frac{1}{2} m \dot{q}^2 - V(q)$$

this gives the Newtonian equations of motion

$$m\ddot{q} = -\frac{\partial V}{\partial q}.$$

25.2. **In a relativistic theory we the unknown quantities are fields: functions of space and time.**

25.3. **The lagrangian depends on the fields and their derivatives.** The Lagrangian is a Lorentz scalar.

25.4. **The action is the integral of the Lagrangian over space and time.**

$$S = \int L(\phi, \partial\phi) d^4x$$

$$\partial_\mu \left[\frac{\partial L}{\partial\left(\partial_\mu \phi\right)} \right] = \frac{\partial L}{\partial \phi}$$

The Lagrangian of a free massive scalar field is

$$L = \frac{1}{2}\eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - \frac{1}{2}m^2\phi^2$$

leading to the Klein-Gordon equation

$$\partial_\mu\partial^\mu\phi + m^2\phi = 0$$

More generally, an interacting scalar theory will have a lagrangian that has terms higher degree than two:

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

$$\partial_\mu\partial^\mu\phi + \frac{\partial V}{\partial\phi} = 0$$

Example 15. For the Higgs field of the standard model (a complex doublet)

$$L = \eta^{\mu\nu}\partial_\mu\phi^\dagger\partial_\nu\phi - V(\phi), \quad V(\phi) = \frac{\lambda}{2}\left[\phi^\dagger\phi - v^2\right]^2$$

We can see directly that the ground states are on the sphere

$$\phi^\dagger\phi = v^2$$

25.5. **The Lagrangian of Maxwell's theory is.**

$$L = \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + j^\mu A_\mu, \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

leading to the equation

$$\partial^\mu F_{\mu\nu} = j_\nu$$

25.6. **The Lagrangian of Dirac field is.**

$$L = \bar{\psi}\left[i\gamma^\mu\partial_\mu + m\right]\psi$$

25.7. **To get interacting theories we add the free lagrangians plus terms that depend on several fields.** For the Yukawa theory,

$$L = \bar{\psi}\left[i\gamma^\mu\partial_\mu + g\phi\right]\psi + \eta^{\mu\nu}\partial_\mu\phi^\dagger\partial_\nu\phi - V(\phi)$$

For QED

$$L = \bar{\psi}\left[i\gamma^\mu\partial_\mu + eA_\mu\right]\psi + m\bar{\psi}\psi + \frac{1}{4}F_{\mu\nu}F^{\mu\nu}$$

26. Charged Scalar Field

26.1. **A free charged spin zero particle is represented by a complex scalar field satisfying the Klein-Gordon equation.**

$$\partial\cdot\partial\phi + m^2\phi = 0$$

26.2. **The Lagrangian is.**

$$L = \partial\phi^*\cdot\partial\phi - m^2\phi^*\phi$$

26.3. **There is an invariance under the global transformation.**

$$\phi \mapsto e^{i\Lambda}\phi$$

It is a very general fact (Noether's theorem) that symmetries lead to conservation laws.

26.4. **And there is a conserved current.**

$$j_\mu = i\left[\phi^*\partial_\mu\phi - \phi\partial_\mu\phi^*\right]$$

26.5. **We can replace the mass term by a potential $V(\phi)$ that depends only on $|\phi|^2$ without losing this symmetry.**

$$L = \partial\phi^*\cdot\partial\phi - V(\phi)$$

An example is

$$V(\phi) = m^2|\phi|^2 + \frac{\lambda}{2}|\phi|^4$$

The last term represents a self-interaction of the scalar field: it leads to cubic non-linearities in the field equations

$$\partial\cdot\partial\phi - \frac{\partial V}{\partial\phi^*} = 0$$

$$\partial\cdot\partial\phi - m^2\phi - \lambda|\phi|^2\phi = 0$$

26.6. **By allowing the symmetry transformation to be position dependent we get a theory of electromagnetism coupled to the charged scalar field.**

$$L = \nabla\phi^* \cdot \nabla\phi - V(\phi) - \frac{1}{4}\mathrm{tr}\, F^2, \quad V(\phi) = m^2|\phi|^2 + \frac{\lambda}{2}|\phi|^4$$

$$\phi(x) \mapsto e^{ie\Lambda(x)}\phi(x)$$

is not a symmetry.

$$\partial_\mu\phi \mapsto e^{i\Lambda(x)}\left[\partial_\mu\phi(x) + ie\partial_\mu\Lambda\phi\right]$$

But if we bring in a gauge field and connect $\Lambda(x)$ into a gauge transformation

$$A_\mu \mapsto A_\mu + \partial_\mu\Lambda$$

$$\nabla_\mu\phi = \partial_\mu\phi - ieA_\mu\phi$$

$$\nabla_\mu\phi \mapsto e^{ie\Lambda}\nabla_\mu\phi$$

26.6.1. *The field equations are.*

$$\nabla\cdot\nabla\phi - m^2\phi - \lambda|\phi|^2\phi = 0$$

$$\partial^\mu F_{\mu\nu} = ie\left[\phi^*\partial_\nu\phi - \phi\partial_\nu\phi^*\right] + 2e^2|\phi|^2A_\nu$$

27. The Higgs Mechanism

27.1. **If $m^2 > 0$, small perturbations around the ground state describe a massive spin zero particle and a massless spin one particle.** The minimum of the potential is at

$$\phi = 0$$

Then all the fields can be expanded around this point; the nonlinear terms are small.

27.2. **But if $m^2 < 0$ we get spontaneous symmetry breaking.** The point is that the minimum is no longer at $\phi = 0$. Indeed

$$V(\phi) = m^2|\phi|^2 + \frac{\lambda}{2}|\phi|^4 = \frac{\lambda}{2}\left[\phi^*\phi + \frac{m^2}{\lambda}\right]^2$$

When $m^2 < 0$, it is convenient to use the parameter $v^2 = -\frac{m^2}{\lambda}$

$$V(\phi) = \frac{\lambda}{2}\left[\phi^*\phi - v^2\right]^2$$

The minmum is along the circle

$$|\phi|^2 = v^2$$

27.3. **The Abelian Higgs Model describes a massive spin one particle and a massive neutral scalar field.**

$$L = \frac{1}{4}F^{\mu\nu}F_{\mu\nu} + |\nabla\phi|^2 - \frac{\lambda}{4}\left[|\phi|^2 - v^2\right]^2$$

where $\nabla_\mu = \partial_\mu + ieA_\mu$ is the covariant derivative as before.

27.3.1. *We change to variables centered at a minimum of the potential in scalar potential.* By passing to the variables

$$\phi = \left[v + \frac{1}{\sqrt{2}}H\right]e^{i\theta}$$

$$\left[|\phi|^2 - v^2\right]^2 = \left\{[v + \frac{1}{\sqrt{2}}H]^2 - v^2\right\}^2 = \left\{\frac{H^2}{2} + \sqrt{2}Hv\right\}^2$$

$$= \frac{H^4}{4} + \sqrt{2}vH^3 + 2v^2H^2$$

$$\frac{\lambda}{4}\left[|\phi|^2 - v^2\right]^2 = \frac{\lambda v^2}{2}H^2 + \frac{\sqrt{2}}{4}\lambda vH^3 + \frac{\lambda}{16}H^4$$

Note that θdrops out: consequence of the symmetry of the potential under rotations.

27.3.2. *We can remove $\theta$from the term inlving derivatives of ϕ as well by a change of the vector field variable.*

$$A_\mu = Z_\mu + \frac{1}{e}\partial_\mu \theta$$

$$\partial_\mu \phi = \left\{ \frac{1}{\sqrt{2}}\partial_\mu H + \left[v + \frac{1}{\sqrt{2}}H \right] i\partial_\mu \theta \right\} e^{i\theta}$$

$$\nabla_\mu \phi = \left\{ \frac{1}{\sqrt{2}}\partial_\mu H + \left[v + \frac{1}{\sqrt{2}}H \right] i\partial_\mu \theta \right\} e^{i\theta} - ie\left[Z_\mu + \frac{1}{e}\partial_\mu \theta \right]\left[v + \frac{1}{\sqrt{2}}H \right] e^{i\theta}$$

$$= \left\{ \frac{1}{\sqrt{2}}\partial_\mu H - ieZ_\mu \left[v + \frac{1}{\sqrt{2}}H \right] \right\} e^{i\theta}$$

So that

$$|\nabla\phi|^2 = \frac{1}{2}\left[\partial H\right]^2 + e^2 Z^2 \left[v^2 + \sqrt{2} vH + \frac{H^2}{2} \right]$$

27.3.3. *The Maxwell Larangian also is independent of θ.*

$$\frac{1}{4}F^{\mu\nu}F_{\mu\nu} = \frac{1}{4}Z^{\mu\nu}Z_{\mu\nu}, \quad Z_{\mu\nu} = \partial_\mu Z_\nu - \partial_\nu Z_\mu$$

27.3.4. *Combining all this the Lagrangian is, in the new variables.* we get

$$L = \frac{1}{2}[\partial H]^2 - \frac{1}{2}\lambda v^2 H^2 + \frac{1}{4}\eta^{\mu\nu}[\partial_\mu Z_\nu - \partial_\nu Z_\mu][\partial_\mu Z_\nu - \partial_\nu Z_\mu] + e^2 v^2 Z^2$$

$$-\frac{\sqrt{2}}{4}\lambda v H^3 - \frac{\lambda}{16}H^4 + \sqrt{2}e^2 v H Z^2 + \frac{1}{2}e^2 H^2 Z^2$$

The term $e^2v^2Z^2$ becomes a mass for the vector boson. Also the term $\lambda v^2 H^2$ is a mass term for the scalar boson.

$$m_H = \sqrt{\lambda} v, \quad m_Z = ev.$$

The remaining terms describe interactions; such as the decay

$$H \to ZZ$$

and scatterings

$$HH \to HH$$

$$HZ \to HZ$$

27.3.5. *The scalar is electrically neutral.* There is no cubic term such as $H\partial HZ$ which would have described a charge interaction .

27.3.6. *Gauge invariance assures us that the variable θ will cancel out, as it did.* In fact we could have simplified the calculation by exploiting this. That is "choose the gauge" where $\phi = v + \frac{H}{\sqrt{2}}$ is real.

28. FERMION MASSES IN THE ABELIAN HIGGS MODEL

28.1. **The same Higgs field that gives mass to the vector bosons can also give a mass to fermions.** Recall that the Dirac equation for a massive fermion is

$$[i\gamma^\mu \partial_\mu + m]\psi = 0$$

This follows from the Lagrangian

$$L = \bar{\psi}[i\gamma^\mu \partial_\mu + m]\psi$$

Here

$$\bar{\psi} = \psi^\dagger \gamma_0$$

so that

$$\bar{\psi}\psi, \quad \bar{\psi}\gamma^\mu \psi$$

are a Lorentz scalar and Lorentz vectors respectively.

If we replace the constant m by a complex scalar field we get a version of Yukawa coupling

$$L = \bar{\psi}\left[i\gamma^\mu \partial_\mu + g\left(\phi_1 + i\gamma_5 \phi_2\right)\right]\psi$$

Note that the imaginary part of ϕcouples as a pseud-scalar while the real part is a scalar. This is similar to the$\sigma$and π mesons in the Yukawa theory.

When the value v of the scalar field in the vacuum ϕ_1 is not zero, we get a mass for the fermion

$$m = gv.$$

28.2. **To get the equations of the scalar fields as well, we must add its Lagrangian.**

$$L = \bar{\psi}\left[i\gamma^\mu \partial_\mu + g\left(\phi_1 + i\gamma_5 \phi_2\right)\right]\psi + |\partial\phi|^2 - \frac{\lambda}{4}\left[|\phi|^2 - v^2\right]^2$$

Note the $U(1)$ symmetry

$$\phi \to e^{i\Lambda}\phi, \quad \psi \to e^{-\frac{i}{2}\gamma_5 \Lambda}\psi, \quad \bar{\psi} \to \bar{\psi}e^{-\frac{i}{2}\gamma_5 \Lambda}$$

for constant Λ. Reval that

$$\gamma_5\gamma_\mu = -\gamma_\mu\gamma_5.$$

28.3. **We can combine this with the Abelian Higgs model to get a theory of massive spin zero, half and one fields.**

$$L = \frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \bar{\psi}[i\gamma^\mu\nabla_\mu + g(\phi_1 + i\gamma_5\phi_2)]\psi + |\nabla\phi|^2 - \frac{\lambda}{4}\left[|\phi|^2 - v^2\right]^2$$

where

$$\nabla_\mu\psi = \left[\partial_\mu + \frac{i}{2}\gamma_5 eA_\mu\right]\psi$$

is the covariant derivative of the fermion field: the factor $\frac{i}{2}\gamma_5 e$ is determined by the transformation of the fermion field under the $U(1)$ symmetry above.. By passing to the variables Z, H, χ

$$\phi = \left[v + \frac{1}{\sqrt{2}}H\right]e^{i\theta}, \quad \psi = e^{-\frac{i}{2}\gamma_5\theta}\chi$$

$$A_\mu = Z_\mu + \frac{1}{e}\partial_\mu\theta$$

$$\left[\partial_\mu + \frac{i}{2}\gamma_5 eA_\mu\right]\psi = \left[\partial_\mu\chi - \frac{i}{2}\gamma_5\partial_\mu\theta\chi\right]e^{i\theta} + \frac{i}{2}\gamma_5 e[Z_\mu + \frac{1}{e}\partial_\mu\theta]\chi e^{i\theta}$$

$$= \left[\partial_\mu\chi + \frac{i}{2}\gamma_5 eZ_\mu\chi\right]e^{i\theta}$$

we get

$$\bar{\psi}\left[i\gamma^\mu\nabla_\mu + g(\phi_1 + i\gamma_5\phi_2)\right]\psi = \bar{\chi}\left[i\gamma^\mu\partial_\mu + \frac{i}{2}\gamma^\mu\gamma_5 eZ_\mu + g\left(v + \frac{1}{\sqrt{2}}H\right)\right]\chi$$

Combining with the bosonic part,

$$L = \frac{1}{2}[\partial H]^2 - \frac{1}{2}\lambda v^2H^2 + \frac{1}{4}\eta^{\mu\nu}[\partial_\mu Z_\nu - \partial_\nu Z_\mu][\partial_\mu Z_\nu - \partial_\nu Z_\mu] + e^2v^2Z^2 + \bar{\chi}\left[i\gamma^\mu\partial_\mu + gv\right]\chi$$

$$-\frac{\sqrt{2}}{4}\lambda vH^3 - \frac{\lambda}{16}H^4 + \sqrt{2}e^2vHZ^2 + \frac{1}{2}e^2H^2Z^2$$

$$+\bar{\chi}\left[\frac{i}{2}\gamma^\mu\gamma_5 eZ_\mu + \frac{g}{\sqrt{2}}H\right]\chi$$

The first line describes particle masses :

spin	mass
0	$\sqrt{\lambda}v$
$\frac{1}{2}$	gv
1	ev

The rest describes the interactions.

Exercise 16. Derives the masses of the spin zero, spin $\frac{1}{2}$ and spin one particles in this Lagrangian.

28.3.1. *The larger the fermion mass the stronger its coupling to the Higgs boson.* Both are proportional to g.

28.4. **The Abelian Higgs model is only a toy; it is a good place to learn properties of the much more complex realistic theory: the standard model.** The standard model must include many fermions and vector bosons, but it has only one Higgs particle: it is the model with the absolute minimum number of fundamental spin zero particles we need to explain masses of the known particles. So v is the same for all particles in the above table for masses.An important consequence is that the more massive a particle is, the stronger is its coupling to the Higgs field. Thus, you should expect that the top quark couples 10^5 times as strongly as the electron to the Higgs boson. Since λ only appears in the self-interaction of the scalar and in its mass, the currently known parameters cannot be used to predict it: we have no clue about the mass of the Higgs boson. Using higher order quantum corrections some indirect bounds can be obrained but they are only a general guideline for the Higgs hunter: theorists can always wiggle out of any such indirect argument. Any Higgs mass under a TeV is fine theoretically.

29. YANG-MILLS THEORY

29.1. **Yang-Mills Theory is the foundation of the theory of elementary particles.** It describes the self-interaction of spin 1 particles: the photon, $Z, W^{\pm}$ and the gluons. The principle of gauge invariance also determines the interactions of these spin one particles with those of spin zero and spin 1: the quarks and leptons.There is also a theory of interactions of spin zero particles (Higgs fields) and spin two particles (General Relativity).

29.2. **Maxwell's theory of electromagnetism is invariant under an abelian gauge group.** Let $\Lambda : R^4 \to R$ be a real valued function. Recall that under the gauge transformation

$$A_\mu \to A_\mu + \partial_\mu \Lambda$$

the field strength

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

is unchanged. Thus the Lagrangian

$$L = \frac{1}{4} F^{\mu\nu} F_{\mu\nu}$$

is invariant under both gauge and Lorentz transformations. Two successive gauge transformations is equivalent to one under the sum

$$\Lambda_1 + \Lambda_2$$

This is a commutative (abelian) group. Suppose a scalar field transforms as

$$\phi \to e^{i\Lambda} \phi$$

Then the covariant derivative

$$\nabla_\mu \phi = \partial_\mu \phi + i A_\mu \phi$$

transforms as

$$\nabla_\mu \phi \to e^{i\Lambda} \nabla_\mu \phi.$$

The Lagrangian

$$L = \frac{1}{4e^2} F^{\mu\nu} F_{\mu\nu} + \frac{1}{2} |\nabla \phi|^2 - V(|\phi|)$$

is gauge invariant. We saw a version of this in the discussion of the Higgs mechanism.

29.2.1. *The value of e determines the strength of the interaction.* We have chosen to define the gauge potential such that the coupling constant appears as a constant factor in the Lagrangian. For any e the gauge invariance holds. and is determined experimentally to be about a third. More precisely

$$\frac{e^2}{4\pi} \approx \frac{1}{137}.$$

29.2.2. *The commutator of covariant derivatives is just a multiplication by the field strength:*

$$\nabla_\mu \nabla_\nu \phi - \nabla_\nu \nabla_\mu \phi = i F_{\mu\nu} \phi$$

This is similar to the definition of curvature in Riemannian geometry.

29.3. **In Yang-Mills theory, the gauge fields are matrix-valued.** Let g be a function on space-time whose value is a unitary matrix of determinant one. That is $g(x) \in SU(n)$. Suppose we have a scalar field which is a vector with $n$complex vector components. It transforms as

$$\phi \to g\phi$$

We can define a covariant derivative by analogy

$$\nabla_\mu \phi = \partial_\mu \phi + iA_\mu \phi$$

where A_μ is a traceless hermitean matrix. (We have absorbed the constant e into the definition of A .) How should A_μ transform in order that this covariant derivative transform as before?

$$\nabla_\mu \phi \to g\nabla_\mu \phi$$

A short calculation gives the answer

$$A_\mu \to gA_\mu g^{-1} + g\partial_\mu (g^{-1})$$

If $g = e^{i\Lambda}$ this reduces to the transformation of Maxwell's theory. What then is the analogue of the field strength? We can calculate

$$\nabla_\mu \nabla_\nu \phi - \nabla_\nu \nabla_\mu \phi = iF_{\mu\nu}\phi$$

where

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + i[A_\mu, A_\nu]$$

The commutator term on the r.h.s. makes all the difference: it implies interactions among spin one particles that have no analogue in Maxwell's theory.

Under gauge transformations,

$$F_{\mu\nu} \to gF_{\mu\nu}g^{-1}.$$

29.4. **The Lagrangian of Yang-Mills theory is.**

$$L_{YM} = \frac{1}{4\alpha} \mathrm{tr} F^{\mu\nu} F_{\mu\nu}$$

It is invaraiant under gauge transformations. The constant αcontrols the strength of the field: how likely it is to deviate from the value $F = 0$. It is called the coupling constant.

29.4.1. *The Yang-Mills field equations are.*

$$\partial^\mu F_{\mu\nu} + [A^\mu, F_{\mu\nu}] = 0.$$

This is the analogue of Maxwell's equatins without sources. It is already nonlinear.All the nonlinearities arise from commutators.

29.5. **Using covariant derivatives we can bring spin zero and spin one fields as sources.**

$$L = \frac{1}{4\alpha}\mathrm{tr}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}|\nabla\phi|^2 - V(|\phi|)$$

$$L = \frac{1}{4\alpha}\mathrm{tr}F^{\mu\nu}F_{\mu\nu} + \bar{\psi}[i\gamma^\mu\nabla_\mu + m]\psi$$

etc. We will study some special cases in detail as they are part of the standard model.

30. Quantum ChromoDynamics

30.1. **Yang-Mills Theory with gauge group $SU(3)$ is Quantum Chromodynamics, the theory of strong interactions.**

$$L = \frac{1}{8\alpha}\mathrm{tr}\ F^{\mu\nu}F_{\mu\nu} + \sum_{a=1}^{N_f} \bar{\psi}[i\gamma^\mu\nabla_\mu + m_a]\psi$$

Each quark field ψ_a is a three component vector under $SU(3)$in addition to being a Dirac spinor. There are six kinds of such quarks $a = 1, \cdots 6$ corresponding to u, d, c, s, t, b with widely varying masses:

$$m_a \sim 5, 10, 1500, 250, 175000, 5000$$

in MeV. In most cases of interest in Nuclear Physics, only the lightest two or three quarks needs to be considered.

30.2. **Understanding the dynamics of non-abelian Yang-Mills theories is one of the deepest unsolved problems of theoretical physics.** It is one of the seven millenium problems of the Clay Math Institute: one of the seven hardest and most important problems in all of mathematics.The only other physics problem in this list is fluid mechanics.

31. An $SU(2)$ Gauge Theory

31.1. **A gauge theory with $SU(2)$ invariance is.**

$$L = \frac{1}{4\alpha}\mathrm{tr}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}|\nabla\phi|^2 - V(|\phi|)$$

where $\phi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}$ is a vector with two complex components.

31.2. **The vacuum value of the Higgs field is non-zero.** With the potential

$$V(\phi) = \frac{\lambda}{4}\left[\phi^\dagger\phi - v^2\right]^2$$

any scalar satisfying

$$|\phi_1|^2 + |\phi_2|^2 = v^2$$

is a minimum. This is a vector of length v in four dimensions (the real and imaginary parts of the two components). In the field the field will point in some constant direction, which we choose to be the second direction. (This will pove to be convenient later.)

31.2.1. *The isotropy group of the vacuum is trivial.*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ v \end{pmatrix} \Rightarrow d = 1, \quad b = 0$$

The only matrix of the form

$$\begin{pmatrix} a & 0 \\ c & 1 \end{pmatrix}$$

which is unitary has

$$c = 0, \quad |a| = 1.$$

Thus the unbroken subgroup is of the form $\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}$ with $a \in U(1)$. To have determinant one we need $a = 1$ as well.

31.3. **It is convenient to change variables adapted to the vacuum.**

$$\phi = U\begin{pmatrix} 0 \\ v + \frac{H}{\sqrt{2}} \end{pmatrix}, \quad U \in SU(2)$$

U has three degrees of freedom and H has one that add up to the four in ϕ.

$$V(\phi) = \frac{\lambda v^2}{2}H^2 + \frac{\sqrt{2}}{4}\lambda v H^3 + \frac{\lambda}{16}H^4$$

Again, U drops out because of the $SU(2)$ symmetry of the potential

31.3.1. *We can remove U from the term inlving derivatives of ϕ as well by a change of the vector field variable.*

$$A_\mu = eUZ_\mu U^{-1} - i[\partial_\mu U]U^{-1}$$

$$\partial_\mu \phi = U\left\{\frac{1}{\sqrt{2}}\partial_\mu H + \left[v + \frac{1}{\sqrt{2}}H\right] iU^{-1}\partial_\mu U\right\}\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$A_\mu \phi = U\left\{eZ_\mu - iU^{-1}\partial_\mu U\right\}\left[v + \frac{1}{\sqrt{2}}H\right]\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\nabla_\mu \phi = U\left\{\frac{1}{\sqrt{2}}\partial_\mu H + eZ_\mu\left[v + \frac{1}{\sqrt{2}}H\right]\right\}\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

So that

$$|\nabla\phi|^2 = \frac{1}{2}[\partial H]^2 + e^2\left[v^2 + \sqrt{2}vH + \frac{H^2}{2}\right](0,1)Z^\mu Z_\mu \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

and

$$L = \frac{e^2}{4\alpha}\mathrm{tr}Z^{\mu\nu}Z_{\mu\nu} + e^2v^2(0,1)Z^\mu Z_\mu \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \frac{1}{2}[\partial H]^2 - \frac{\lambda v^2}{2}H^2$$

$$+e^2\left[\sqrt{2}vH + \frac{H^2}{2}\right](0,1)Z^\mu Z_\mu \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \frac{\sqrt{2}}{4}\lambda vH^3 - \frac{\lambda}{16}H^4$$

31.3.2. *Choose e to normalize the kinetic energy of the YM field conventionally.*

$$Z_\mu = Z_{\mu a}\frac{\sigma_a}{2}$$

$$\frac{e^2}{4\alpha}\mathrm{tr}Z^{\mu\nu}Z_{\mu\nu} = \frac{1}{4}Z_a^{\mu\nu}Z_{\mu\nu a}$$

That is,

$$e^2 = 2\alpha$$

31.3.3. *The easily checked identity.*

$$(1,0)Z^\mu Z_\mu \begin{pmatrix} 1 \\ 0 \end{pmatrix} = (Z_1^\mu)^2 + (Z_2^\mu)^2 + (Z_3^\mu)^2$$

31.4. **Thus we have the Lagrangian.**

$$L = \frac{1}{4} Z_a^{\mu\nu} Z_{\mu\nu a} + e^2 v^2 Z_{\mu a} Z_a^{\mu} + \frac{1}{2}[\partial H]^2 - \frac{\lambda v^2}{2} H^2$$

$$+ e^2 \left[\sqrt{2} v H + \frac{H^2}{2}\right] Z_{\mu a} Z_a^{\mu} - \frac{\sqrt{2}}{4} \lambda v H^3 - \frac{\lambda}{16} H^4$$

Again the first line describes a spin zero particle of mass $\sqrt{\lambda} v$ and a set of three spin one particles of mass ev. The rest are interactions.

Exercise 17. Show that the masses of all the gauge bosons are non-zero. Are they equal?**Hint** Expand $Z_\mu = Z_{\mu a} \frac{\sigma_a}{2}$ in terms of Pauli matrices and express $(1,0) Z^\mu Z_\mu \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ in terms of these components.

32. $SU(2) \times U(1)$ Gauge Theory

32.1. **Weak interactions are mediated by a pair of charged massive particles ($W^\pm$) and a spin zero particle Z^0.** Beta decay is due to $W^\pm$; the Z^0 bosons mediate neutral current interactions which were discovered later. The masses are not equal.

32.2. **Electromagnetism is already described as a $U(1)$ gauge theory.**

32.3. **The only consistent theory of massive spin one charged particles is a broken gauge theory.** Veltmann did heroic calculations of quantum corrections to magnetic moments of spin one particles to show that the only sensible value is that determined by such a Yang-Mills theory. In modern language, no other theory is renormalizable.

32.4. **There must be a gauge theory based on a Lie algebra of dimension 4 which is broken to a $u(1)$ subalgebra by a Higgs field.** This left over symmetry is responsible for the photon being massless. In particular the $SU(2)$ model of the last section cannot be the right theory of weak interactions because it does not have a massless gauge boson.

32.5. **The simplest choice is $U(2) \to U(1)$.** The gauge fields are 2×2 hermitean matrices, but they are not traceless. We can use exactly the same scalar Lagrangian as above. Without the condition that the determinant be one the unbroken symmetry group is

$$\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}, \quad a \in U(1).$$

Thus the photon is the gauge boson corresponding to this direction in the symmery group.

32.6. **There are two independent gauge invariant actions for the Yang-Mills field.** This is because the gauge group is not simple. Both

$$\mathrm{tr} F^{\mu\nu} F_{\mu\nu}$$

and

$$\mathrm{tr} F^{\mu\nu} \mathrm{tr} F_{\mu\nu}$$

are gauge invariant. The latter would have been zero for the $SU(2)$ gauge theory, because of the tracelessness. This is why the $W^{\pm}$ and Z^0 can have different masses. Or equivalently, why the charged current and neutral current can have different, but universal, coupling constants. It is useful to split the gauge field into a traceless part and a piece that is the multiple of the identity

$$A_\mu = g\sigma_a W_{\mu a} + g' Y_\mu$$

with two independent coupling constans g, g'.

32.6.1. *The $SU(2)$ symmetry arising here is called weak isospin; the $U(1)$ is called weak hypercharge.* We will see that electric is a linear combination of the third generator of $SU(2)$ and the weak hypercharge

$$Q = I_{3W} + \frac{Y}{2}.$$

The general Yang-Mills Lagrangian for $U(2)$ becomes

$$L_{YM} = \frac{1}{4} W_a^{\mu\nu} W_{\mu\nu a} + \frac{1}{4} Y^{\mu\nu} Y_{\mu\nu}$$

32.7. **The Lagrangian including the Higgs field is.**

$$L = \frac{1}{4} W_a^{\mu\nu} W_{\mu\nu a} + \frac{1}{4} Y^{\mu\nu} Y_{\mu\nu} + |\nabla\phi|^2 - \frac{\lambda}{4}\left[\phi^\dagger\phi - v^2\right]^2$$

Again, the minimum of the potential is on the sphere $\phi^\dagger\phi = v^2$.

Exercise 18. Go to the gauge $\phi = \begin{pmatrix} 0 \\ v + \frac{H}{\sqrt{2}} \end{pmatrix}$ and diagonalize the quadratic part of the Lagrangian to get

$$m_W = gv, \quad m_Z = \sqrt{g^2 + g'^2}v, \quad m_V = 0$$

for appropriate linear combinations Z, V of $W_{\mu 3}$ and Y_μ.

$$\nabla_\mu \phi = \partial_\mu \phi + i\left[g\sigma_a L_{\mu a} + g' Y_\mu\right]\phi$$

$$= \begin{pmatrix} 0 \\ \frac{\partial_\mu H}{\sqrt{2}} \end{pmatrix} + \begin{pmatrix} ig'Y_\mu + igW_{\mu 3} & ig\left[W_{\mu 1} - iW_{\mu 2}\right] \\ ig\left[W_{\mu 1} + iW_{\mu 2}\right] & igY_\mu - ig'W_{\mu 3} \end{pmatrix} \begin{pmatrix} 0 \\ v + \frac{H}{\sqrt{2}} \end{pmatrix}$$

$$= \begin{pmatrix} \left[v + \frac{H}{\sqrt{2}}\right] ig(W_\mu - iW_{\mu 2}) \\ \frac{\partial_\mu H}{\sqrt{2}} + \left[v + \frac{H}{\sqrt{2}}\right] (ig'Y_\mu - igW_{\mu 3}) \end{pmatrix}$$

Writing out in detail only the quadratic terms,

$$L = \frac{1}{4}\left[\partial_\mu W_{\mu 1} - \partial_\nu W_{\mu 1}\right]^2 + \frac{1}{4}\left[\partial_\mu W_{\mu 2} - \partial_\nu W_{\mu 2}\right]^2 + g^2 v^2 (W_{\mu 1}^2 + W_{\mu 2}^2) + \frac{1}{4}\left[\partial_\mu W_{\nu 3} - \partial_\nu W_{\mu 3}\right]^2 + \frac{1}{4} Y^\mu$$

We see that $W_{\mu 1}, W_{\mu 2}$ correspond to spin one bosons of mass gv. But $W_{\mu 3}$ and Y_μ are mixed. The combination

$$Z_\mu = \frac{g'Y_\mu - gW_{\mu 3}}{\sqrt{g^2 + g'^2}}$$

is massive while the orthogonal combination

$$V_\mu = \frac{gY_\mu + g'W_{\mu 3}}{\sqrt{g^2 + g'^2}}$$

is massless. The mixing angle

$$\tan\theta_W = \frac{g'}{g}$$

measures the relative strengths of the charged and neutral weak interactions.

$$L = \frac{1}{4}\left[\partial_\mu W_{\mu 1} - \partial_\nu W_{\mu 1}\right]^2 + \frac{1}{4}\left[\partial_\mu W_{\mu 2} - \partial_\nu W_{\mu 2}\right]^2 + g^2 v^2 (W_{\mu 1}^2 + W_{\mu 2}^2) + \frac{1}{4}\left[\partial_\mu V_\nu - \partial_\nu V_\mu\right]^2 + \frac{1}{4}\left[\partial_\mu Z\right.$$

32.7.1. *Strictly speaking the symmetry group of the electroweak theory is $U(2)$ and not $SU(2) \times U(1)$ as it is commonly called.* The difference is a discrete symmetry $Z_2 = \{1, -1\}$:

$$U(2) = SU(2) \times U(1)/Z_2$$

It so happens that all the standard model particles happen to be unchanged by this symmetry so that the true gauge group is $U(2)$. This does not affect any perturbative calculation, but can slightly change some facts about large excitations such as cosmic strings.

33. Fermions In the ElectroWeak Theory

33.1. **The quarks and leptons are Dirac fermions, but their left and right handed parts transform differently under $SU(2)\times U(1)$.** Consider for simplicity just the fermions of the first generation.We will include others later

33.2. **The left handed components are doublets under weak isospin.** The left handed components of the leptons form a doublet under $SU(2)$. So do the left handed parts of the up and down quarks, except there are three copies of them, one for each color.

$$\left(\begin{array}{c} l_1 \\ \nu_1 \end{array}\right)_L, \quad \left(\begin{array}{c} u_\alpha \\ d_\alpha \end{array}\right)_L$$

l_1 here stands for the ani-electron (positron) and $\nu_1$for the anti-electron-neutrino That is, these particles have lepton number -1. This choice makes it easier to write down the Lagrangian.

33.2.1. *Recall* $Q=I_{3W}+\frac{Y}{2}$. To get the correct electric charges, the weak hypercharges are 1 and $\frac{1}{3}$ for the left-handed leptons and quarks respectively.

33.2.2. *The right handed components are invariant under $SU(2)$: they do not take part in the charged current interactions.*

$$\nu_{1R}, l_{1R}, u_{R\alpha}$$

Because $I_{3W}=0$, the right electric charges the hypercharge assignments are $0,2,\frac{4}{3},-\frac{2}{3}$ respectively.

33.3. **Fermions get their masses from coupling to the Higgs boson as well.**

33.3.1. *The cross-product of spinors is invariant under $SU(2)$.* If we have two spinors

$$\left(\begin{array}{c} \phi_1 \\ \phi_2 \end{array}\right) \quad \left(\begin{array}{c} \chi_1 \\ \chi_2 \end{array}\right)$$

the product

$$\phi_1\chi_2-\phi_2\chi_1$$

is invariant under the transformation

$$\phi\to g\phi, \quad \chi\to g\chi.$$

For then,

$$\phi_1\chi_2-\phi_2\chi_1\to \det g[\phi_1\chi_2-\phi_2\chi_1].$$

We can write

$$\phi_1\chi_2 - \phi_2\chi_1 = \phi^T i\sigma_2\chi$$

This will come in useful in forming Yukawa couplings.

33.3.2. *The electron mass.* Let us consider the simplest case before we look at the general situation. The left-handed fermions as well as the scalar field transform as doublets. By the way this determines the weak hypercharge of the scalar field to be 1. So the Yukawa coupling

$$-g_e\bar{l}_{1R}\phi^T i\sigma_2 \begin{pmatrix} l_1 \\ \nu_1 \end{pmatrix}_L + h.c.$$

is invariant under the gauge symmetry as well as Lorentz transformations. In the gauge where the scalar is

$$\phi = \begin{pmatrix} 0 \\ v + \frac{H}{\sqrt{2}} \end{pmatrix}$$

this becomes

$$g_e v\bar{l}_{1R}l_{1L} + g_e\frac{H}{\sqrt{2}}\bar{e_R}e_L + h.c$$

Thus the electron mass is

$$m_e = g_e v$$

and its coupling to the Higgs is also proportional to the Yukawa constant.

33.3.3. *The Up Quark masses.* The natural analogue

$$-g_u\bar{u}_R^\alpha\phi^T i\sigma_2 \begin{pmatrix} u_\alpha \\ d_\alpha \end{pmatrix}_L + h.c.$$

will give a mass to the up quark. Of course, the different colors have the same coupling to the scalar. Again this leads to $m_u = g_u v$

33.3.4. *The down Quark Masses.* But the up quark is massive too.

$$g_d\bar{d}_R^\alpha\phi^\dagger \begin{pmatrix} u_\alpha \\ d_\alpha \end{pmatrix}_L + h.c.$$

is gauge invariant well. In the unitary gauge this is

$$g_d v\bar{d}_R u_L + g_d\frac{H}{\sqrt{2}}\bar{d}_R d_L + h.c$$

leading to a mass for the up quark.

33.3.5. *Neutrino masses.* In the original version of the standard model, there was no provision for neutrino masses. Now that we know at least two of the neutrinos are massive, we must ask what is the minimal modification that will include them? Of course the truth may be more complicated, or elegant, but we must always start with the minimal model; that is the lesson of the success of the standard model itself. Attempts to grand unify it have so far not wored out well.

The simplest idea is to just imitate what we did for the down quarks.

$$g_{1\nu}\bar{\nu}_{1R}\phi^{\dagger}\begin{pmatrix} l_1 \\ \nu_1 \end{pmatrix}_L + h.c$$

leading to

$$m_{\nu_e} = g_{1\nu}v.$$

This has got to be one of the smallest fundamental constants in all of physics: $g_{1\nu} < \frac{2}{2.5\times 10^{11}}$.

34. Quark Mixing

34.1. **The situation is complicated by the fact that the mass matrices are not diagonal.** Since the quarks of the different generations transform the same way under $SU(2)\times U(1)$ a Yukawa coupling can mix them.

Thus we should think of the Yukawa couplings of the standard model as 3×3 matrices. There is one such matrix for the up quarks and another for the down quarks.

$$g^a_b\bar{d_{Rb}}\phi^{\dagger}\begin{pmatrix} u_{La} \\ d_{La} \end{pmatrix} - i\tilde{g}^a_b\bar{u_{Rb}}\phi^T\sigma_2\begin{pmatrix} u_{La} \\ d_{Lb} \end{pmatrix} + h.c.$$

They need not be diagonal in the same basis: they do not need to commute. The unitary transformation that realates the bases where they are diagonal is called the CKM matrix. (Cabibbo-Kobayashi-Mazkawa).

34.1.1. *Number of degrees of freedom.* It is useful to generalize to the case where there are N generations and only put $N = 3$at the end. We can diagonalize each mass matrix by a unitary transformation

$$g = U\mathrm{diag}(m_1,\cdots m_N)U^{\dagger}$$

$$\tilde{g} = \tilde{U}\mathrm{diag}(\tilde{m}_1,\cdots \tilde{m}_N)\tilde{U}$$

Each of the eigenvalues are distinct: no two quark masses are equal. The ratio of these two transformations

$$V = \tilde{U}^{-1}U$$

measures the mixing. It is the CKM matrix. But recall that$U$and $\tilde{U}$ are not unique. If we multiply them of the right by a diagonal unitary matrix, g remains unaffected. That is, by a matrix which is diagonal and each entry along the diagonal is a complex number of unit magnitude. Thus there is an equivalence among CKM matrices.

$$V_{ab} \sim V_{ab} e^{i[\theta_a - \tilde{\theta}_b]}$$

Of the N^2independent parameters in V, we can remove $2N-1$ this way. (The minus one is because an overall constant added to the $\theta, \tilde{\theta}$ do not matter.) This leaves

$$N^2 - (2N-1) = (N-1)^2$$

independent parameters. For the original case of $N = 2$, this only leaves one parameter, the Cabibbo angle. In today's realistic case $N = 3$ we have 4 independent parameters.

34.2. **CP Violation.** Of great importance in physics is the question: **can V be chosen to be real using the transformation above?** The reason this is so important is that a complex phase in V would lead to CP violation. Kobayashi and Mazkawa realized that for $N = 3$ this matrix cannot be chosen to be real: there are only $\frac{N(N-1)}{2}$ parameters in a real unitray (i.e., orthogonal) matrix. Thus for $N = 3$,the CKM matrix has 4 independent parameters, of which only three can be accounted for by a real orthogonal matrix. The remaining complex parameter now provides a natural explanation to the observed CP violation in physics.

34.3. **Measuring the CKM mixing angles.** The magnitudes of matrix elements $|V_{ab}|$ can be measured using decay widths of the mesons containing these quarks. The oldest example (Cabibbo mixing) is the decay of the charged Kmesons. Many years of careful work has now pinned down all these mixing angles. Along with the CP violation (observed in $K\bar{K}$ and$B\bar{B}$ oscillations) we now understand reasonably well the value of the CKM matrix.

34.3.1. *Double Coset spaces.* What we are talking about here mathematically is the double coset space $[U(1)]^N \backslash U(N) / [U(1)]^N$. The parameter count above is just the dimension of this "orbifold". There is a standard theory of these, but Kobayashi-Mazkawa reinvented the part they needed on ther own.

35. The Lagrangian Of the Standard Model

35.1. **All the phenomena of nature, except gravity, are described the the Lagrangian of the standard model.** The Lagrangian defines a particular quantum field theory. The precise definition of a quantum field theory involves the mysterious procedure of renormalization; we only understand it fully in perturbation theory.

35.2. **It is a Yang-Mills Theory with gauge group $SU(3) \times SU(2) \times U(1)$, a scalar field and a set of spin $\frac{1}{2}$ fields.** The theory is specified by the representations of the scalar and fermion fields; and the couplings among them. The three states $SU(3)$ symmetry is labelled by ‘ color’; $SU(2)$is called weak isospin and the $U(1)$is called weak hypercharge. We denote the corresponding gauge fields by a set of vector fields: $A^{\alpha}_{\mu\beta}$, each component of which is a traceless hermitean 3×3 matrix; $W^{a}_{\mu b}$, whose components are traceless hermitean 2×2 matrices; and Y_μ each component of which is just a real function.

35.3. **The scalar field is invariant under $SU(3)$; it is in the fundamental (doublet) representation of $SU(2)$; has hypercharge one.** That is, ϕ transforms as $(1,2,1)$: the trivial representation of color, the defining representation of $SU(2)$and of hypercharge one.

This means that if $g \in U(2) = SU(2) \times U(1)/Z_2$, the scalar field transforms as the doublet $\phi \to g\phi$. The covariant derivative on scalars is

$$\nabla_\mu \phi_b = \partial_\mu \phi_b + i[W^{a}_{\mu b} + Y_\mu \delta^{a}_{b}]\phi_a$$

35.4. **There are two kinds of fermions: quarks and leptons whose left and right-handed components couple differently to gauge bosons, as listed below.**

(1) The left-handed quarks $q_{\alpha bA}$ where $\alpha = 1,2,3$ labels the fundamental representation of color SU(3); b labels the fundamental representation of weak isospin $SU(2)$; the hypercharge $\frac{1}{3}$; and $A = 1,2,3$ labels the generations.

$$\gamma_5 q = q, \quad \nabla_\mu q_{\alpha bA} = \partial_\mu q_{\alpha bA} + i\left[\left\{W^{a}_{\mu b} + \frac{1}{3}Y_\mu \delta^{a}_{b}\right\}\delta^{\beta}_{\alpha} + A^{\beta}_{\mu\alpha}\right] q_{\beta aA}$$

(2) The left-handed leptons l_{bA} are trivial under color; in the fundamental of weak isospin; and has hypercharge -1.

$$\gamma_5 l = l, \quad \nabla_\mu l_{bA} = \partial_\mu l_{bA} + i\left\{W^{a}_{\mu b} - Y_\mu \delta^{a}_{b}\right\} l_{aA}$$

(3) The right-handed up-quarks $u_{\alpha A}$ are in the fundamental of color; are trivial under weak isospin; and has hypercharge $\frac{4}{3}$

$$\gamma_5 u = -u, \quad \nabla_\mu u_{\alpha A} = \partial_\mu u_{\alpha A} + i\left[A^{\beta}_{\mu\alpha} + \frac{4}{3} Y_\mu \delta^{\beta}_{\alpha}\right] u_{\beta A}$$

(4) The right-handed down-quarks$d_{\alpha A}$ are also fundamental in color; trivial under weak isospin; and hypercharge

$$\gamma_5 d = -d, \quad \nabla_\mu d_{\alpha A} = \partial_\mu d_{\alpha A} + i\left[A^{\beta}_{\mu\alpha} - \frac{2}{3} Y_\mu \delta^{\beta}_{\alpha}\right] d_{\beta A}$$

(5) The right-handed charged leptons $e_A$are trivial under both color and weak isospin; and have hypercharge -2.

$$\gamma_5 e = -e, \quad \nabla_\mu e_A = \partial_\mu e_A + i\left[-\frac{2}{3} Y_\mu\right] e_A$$

(6) The right-handed neutrinos ν_A are trivial under color and weak isospin, as well as of zero hypercharge. They do not couple to any gauge bosons at all.

$$\gamma_5 \nu = -\nu, \quad \nabla_\mu \nu_A = \partial_\mu \nu_A$$

35.5. **The Lagrangian can then be written down.**

$$L = \frac{1}{4\alpha_1} Y^{\mu\nu} Y_{\mu\nu} + \frac{1}{4\alpha_2} W^{\mu\nu} W_{\mu\nu} + \frac{1}{4\alpha_3} F^{\mu\nu} F_{\mu\nu} + |\nabla\phi|^2 - \frac{\lambda}{4}\left[\phi^\dagger\phi - v^2\right]^2 +$$

$$\bar{q}^{\alpha bA} i\gamma^\mu \nabla_\mu q_{\alpha bA} + \bar{u}^{\alpha A} i\gamma^\mu \nabla_\mu u_{\alpha A} + \bar{d}^{\alpha bA} i\sigma\gamma^\mu \nabla_\mu d_{\alpha bA} +$$

$$\bar{l}^{\alpha bA} i\gamma^\mu \nabla_\mu l_{\alpha bA} + \bar{e}^A i\sigma^\mu \nabla_\mu e_A + \bar{\nu}^A i\gamma^\mu \partial_\mu \nu_A +$$

$$+\frac{1}{v}\sum_A M_A \bar{u}^{\alpha A} \phi_a \varepsilon^{ab} q_{\alpha bA} + \frac{1}{v}\sum_{AB} \tilde{M}_B V^B_A \bar{d}^{\alpha A} \phi^{\dagger b} q_{\alpha bB} +$$

$$\frac{1}{v}\sum_A m_A \bar{e}^A \phi_a \varepsilon^{ab} l_{bA} + \frac{1}{v}\sum_A \tilde{m}_A U^B_A \bar{\nu}^A \phi^{\dagger b} l_{bB}$$

Note that the right-handed neutrinos do not couple to any gauge bosons. The basis is chosen so that mass matrix of the up quarks is diagonal, with entries M_A. Then those of the down quarks will be the diagonal matrix $\tilde{M}$ times the CKM matrix V which is a unitary matrix; because of the freedom to choose phases for the basis of quarks, V is defined only modulo a left and a right action by diagonal unitary matrices: $V \in U(1)^N \backslash U(N)/U(1)^N$. Similarly, the mass matrix of the charged leptons is chosen to be diagonal, leading to a neutrino mixing matrix $U \in U(1)^N \backslash U(N)/U(1)^N$. The dimension of the double coset space $U(1)^N \backslash U(N)/U(1)^N$ is $(N-1)^2$.

35.6. **The parameters are** $\alpha_1, \alpha_2, \alpha_3, \lambda, v, M_A, \tilde{M}_{A,} V_A^B, m_A, \tilde{m}_A, U_A^B$**.** That is $5 + 4N + 2(N-1)^2 = 2N^2 + 7$ parameters where N is the number of generations. For $N = 3$ these are 25 parameters. Most of them are known to reasonable accuracy: the unknown parameters are the Higgs mass (λ)and all except two combinations of $\widetilde{m_A}, U_B^A$ in the neutrino sector: only the two combinations of the neutrino masses and mixing angles that give neutrino oscillation lengths are known.