8. Forces, Connections and Gauge Fields

8.0 Preliminary

In general relativity, gravitational forces arise naturally from the geometrical structure of spacetime. As described in chapter 4, the logical steps that lead to this conclusion are

1. Physical quantities (tensors) at different points in spacetime are related by an affine connection, which defines parallel transport.
2. Connection coefficients that cannot be set equal to zero everywhere by a suitable coordinate transformation indicate the presence of gravitational forces.
3. Such effects can be described by a principle of least action.

In essence, *gravitational forces arises from communication between different points in spacetime*. It was found that all known fundamental forces can be described in a similar manner in terms of gauge theories.

8.1. Electromagnetism

8.1.1. Internal Space

Consider a particle described by a complex wavefunction

$$\phi(x) = \phi_1(x) + i\,\phi_2(x) = |\phi(x)|\,e^{i\theta(x)} \qquad (8.1)$$

where ϕ_1, ϕ_2 and $\theta(x)$ are real, and $x = (ct, \mathbf{x})$. By definition, a constant overall phase is not observable so that ϕ and $\phi\,e^{i\theta_0}$ represents the same state if $\theta_0 = const$. However, a varying phase angle $\theta(x)$ is observable through its finite contribution to the momentum $\langle p^\mu \rangle = i\hbar\eta^{\mu\nu}\int d^3\mathbf{x}\,\phi^*(x)\partial_\nu\phi(x)$ of the particle.

Alternatively, we can think of $\phi(x)$ as a vector in the 2-D **internal space** (ϕ_1, ϕ_2) of the spacetime point x. We are then dealing with a fibre bundle with spacetime as the base manifold and the internal space as the typical fibre [see §3.7.3]. The wave function $\phi(x)$ as a whole is a vector field, or cross section of the bundle. Obviously, the direction of the vector $\phi(x)$ at x is given by its phase $\theta(x)$ there. To express the fact that an overall constant phase has no physical meaning, we need to define a parallel transport so that physically significant changes in $\phi(x)$ between points x_1 and x_2 are given by

$$\delta\phi = \phi(x_2) - \phi(x_1 \to x_2) \qquad (8.2)$$

where $\phi(x_1 \to x_2)$ is the vector at x_2 that is "parallel" to $\phi(x_1)$. Analogous to eq(2.23), we define

$$\phi_i(x \to x + \Delta x) = \phi_i(x) - \Gamma_{i\,j\mu}(x)\,\phi_j(x)\,\Delta x^\mu \qquad (8.3)$$

where $\Gamma_{i\,j\mu}$ are the **connection coefficients**. As in the theory of general relativity, we shall interpret this as a geometrical description of the dynamics of the particle subject to some external potential field. To this end, consider 1st the case of a "flat" space so that the directions of $\phi(x)$ at all x can be referred to a single global coordinate system, i.e., two vectors $\phi(x_1)$ and $\phi(x_2)$ are parallel if $\theta(x_1) = \theta(x_2) + 2n\pi$, where n is any integer. The internal space is therefore the same for all x. Obviously, this describes a free particle. A “curved” space then denotes the presence of forces, which we shall associate with electromagnetism.

8.1.2. Connection Coefficients

Unlike the phase θ, the magnitude

$$|\phi| = \sqrt{\phi^* \phi} = \sqrt{\phi_1^2 + \phi_2^2}$$

denotes the measurable probability amplitude. Since $\phi(x_1 \to x_2)$ should be physically equivalent to $\phi(x_1)$, we must have $|\phi(x_1 \to x_2)|^2 = |\phi(x_1)|^2$. Using (8.3), we have

$$|\phi(x_1 \to x_2)|^2 = [\phi_1(x_1 \to x_2)]^2 + [\phi_2(x_1 \to x_2)]^2$$

$$= [\phi_1(x_1)]^2 - 2\phi_1(x_1)\Gamma_{1j\mu}(x_1)\phi_j(x_1)\Delta x^\mu$$

$$+[\phi_2(x_1)]^2 - 2\phi_2(x_1)\Gamma_{2j\mu}(x_1)\phi_j(x_1)\Delta x^\mu + O[\Delta x]^2$$

$$= |\phi(x_1)|^2 - 2[\phi_1(x_1)\Gamma_{1j\mu}(x_1) + \phi_2(x_1)\Gamma_{2j\mu}(x_1)]\phi_j(x_1)\Delta x^\mu + O[\Delta x]^2$$

Thus, the requirement $|\phi(x_1 \to x_2)|^2 = |\phi(x_1)|^2$ can be satisfied only if

$$0 = (\phi_1\Gamma_{1j\mu} + \phi_2\Gamma_{2j\mu})\phi_j = \Gamma_{ij\mu}\phi_i\phi_j$$

$$\Rightarrow \qquad \Gamma_{ij\mu} = -\Gamma_{ji\mu}$$

Thus, we can write

$$\Gamma_{ij\mu}(x) = -\lambda\varepsilon_{ij}A_\mu(x) \qquad (8.4)$$

where ε_{ij} is the 2-D Levi-Civita symbol. A_μ is a 4-vector which will be identified later with the electromagnetic vector potential. λ is a constant proportional the charge of the particle [See §8.1.6 for details].

8.1.3. Group Manifold

Since the parallel transport preserves the magnitude $|\phi|$, it affects only the phase θ.

Thus, we need only study a fibre bundle whose typical fibre is the unit circle $|\phi| = 1$ in the complex ϕ plane. Now, to each phase function $\theta(x)$ is associated a (phase) transformation $\phi(x) \to e^{i\theta(x)}\phi(x)$ such that each $\phi(x)$ is rotated by an angle $\theta(x)$.

Since, for a constant θ, $e^{i\theta}\phi(x)$ is equivalent to $\phi(x)$, the operator $e^{i\theta}$ represents a symmetry transformation of the system. The set $\{e^{i\theta}\}$ of all such transformations forms a Lie group called $U(1)$ [unitary group of 1 dimension]. Thus, the elements of the typical fibre $\theta \in [0, 2\pi]$ also serves as the parameters of the symmetry group. In other words, the typical fibre is also the **group manifold**.

In anticipation of subsequent generalization, we shall consider transformations of the form $\phi(x) \to e^{i\lambda\theta(x)}\phi(x)$ and refer to them as **local gauge transformations**. [A **global** gauge transformation is one with $\theta(x) = const$]. Taken as coordinate transformations, these gauge transformations can be used to define (gauge) tensors in our fibre bundle. Since the fibre is 1-D, the only independent gauge vector we have is ϕ. Treating the number $\phi^* \phi$ as a contraction, we identify the gauge 1-form as ϕ^*. A gauge tensor field of rank $\begin{pmatrix} n \\ m \end{pmatrix}$ is therefore $\Phi_{mn}(x) = \phi^{*m}(x)\phi^n(x)$, which transforms according to

$$\Phi'_{mn}(x) = e^{i(n-m)\lambda\theta(x)}\Phi_{mn}(x) \qquad (8.5)$$

Note that the subscripts of Φ denote its rank, not its components. Reminder: in a 1-D space, every tensor has exactly 1 component.

8.1.4. Covariant Derivative

The parallel transport (8.3) leads to a covariant derivative

$$D_\mu \phi_i(x) = \lim_{\Delta x^\mu \to 0} \frac{\phi_i(x+\Delta x) - \phi_i(x \to x+\Delta x)}{\Delta x^\mu} \qquad (8.6)$$

$$= \partial_\mu \phi_i(x) + \Gamma_{i\,j\mu}(x)\phi_j(x) \qquad (8.7)$$

$$= \partial_\mu \phi_i(x) - \varepsilon_{i\,j}\lambda A_\mu(x)\phi_j(x) \qquad \text{[eq(8.4) used]} \qquad (8.7a)$$

Thus,

$$D_\mu \phi_1(x) = \partial_\mu \phi_1(x) - \lambda A_\mu(x)\phi_2(x)$$

$$D_\mu \phi_2(x) = \partial_\mu \phi_2(x) + \lambda A_\mu(x)\phi_1(x)$$

$$\Rightarrow \qquad D_\mu \phi(x) = D_\mu\left[\phi_1(x) + i\phi_2(x)\right] = \partial_\mu \phi(x) + i\lambda A_\mu(x)\phi(x) \qquad (8.8)$$

Consider the gauge transformation

$$\phi(x) \to \phi'(x) = e^{i\lambda\theta(x)}\phi(x) \qquad \phi^*(x) \to \phi'^*(x) = e^{-i\lambda\theta(x)}\phi^*(x)$$

In the new gauge coordinates, we have

$$D'_\mu \phi' \equiv \partial_\mu \phi' + i\lambda A'_\mu \phi' = \left(\partial_\mu + i\lambda A'_\mu\right)\left(e^{i\lambda\theta}\phi\right) \qquad (8.9a)$$

where

$$\Gamma'_{i\,j\mu}(x) = -\varepsilon_{i\,j}\lambda A'_\mu(x)$$

As in the case of the affine connection, the transformation of A_μ is determined by the condition that the covariant derivative of a tensor is another tensor. It is obvious from its definition that D_μ does not change the rank of the gauge tensor on which it operates. Hence, $D_\mu \phi$ is a gauge vector so that

$$D'_\mu \phi' = e^{i\lambda\theta}\left(D_\mu \phi\right) = e^{i\lambda\theta}\left(\partial_\mu + i\lambda A_\mu\right)\phi$$

Comparing this with (8.9a) gives

$$\partial_\mu e^{i\lambda\theta} + i\lambda A'_\mu e^{i\lambda\theta} = i\lambda A_\mu e^{i\lambda\theta}$$

$$\Rightarrow \qquad A'_\mu = A_\mu + \frac{i}{\lambda} e^{-i\lambda\theta}\partial_\mu e^{i\lambda\theta} = A_\mu - \partial_\mu \theta \qquad (8.10)$$

Similarly, the covariant derivative of a rank $\begin{pmatrix} n \\ m \end{pmatrix}$ gauge tensor can be easily found to be

$$D_\mu \Phi_{mn} = \partial_\mu \Phi_{mn} + i\lambda (n-m) A_\mu \Phi_{mn} \tag{8.11}$$

Note that eq(8.10) has the same form as the gauge transformation (3.60) in electromagnetism. This is the reason why $A_\mu(x)$ is called the **gauge field**. At this point, we have shown that the phases of a complex wavefunction constitue a U(1) fibre bundle, whose geometry is determined by the gauge fields.

8.1.5. Spin 1/2 Particles

One advantage of a geometric point of view of interactions is that it allows for easy generalizations. In particular, it provides a classification of tensors that is indispensable in the construction of a scalar action for a new theory. As an example, consider the generalization of the results in the previous sections to the case of a spin 1/2 particle. To include the effects of gauge fields, we need only convert the Dirac action

$$S = \int d^4x\, \bar{\psi}\left(i\partial\!\!\!/ - m\right)\psi \qquad (7.61)$$

into a gauge scalar. Since only $\partial\!\!\!/$ is not a gauge tensor, the simplest choice is to replace it with the gauge scalar $D\!\!\!\!/ = \partial\!\!\!/ + i\lambda A\!\!\!/$. Hence, the desired action is

$$S = \int d^4x\, \bar{\psi}\left(i\partial\!\!\!/ - \lambda A\!\!\!/ - m\right)\psi \qquad (8.12)$$

The Euler-Lagrange equation for the $\bar{\psi}$ degree of freedom is simply

$$\left(i\partial\!\!\!/ - \lambda A\!\!\!/ - m\right)\psi = 0 \qquad (8.13)$$

which will be identified as the Dirac equation for a spin 1/2 particle of charge λ moving in an electromagnetic field of vector potential A_μ. The recipe

$$\partial \to D = \partial + i\lambda A \qquad \text{(a)}$$

is called the **minimal coupling prescription**.

Although no firm evidence has yet been presented, it is convenient here to accept the identification of the gauge fields with the electromagnetic vector potentials in order to facilitate the following discussions. The transformation (a) promotes the *global* gauge symmetry of the original Dirac equation to a *local* gauge symmetry of eq(8.13). This is analogous to the promotion of a global metric $\eta_{\mu\nu}$ to the local one $g_{\mu\nu}(x)$ in the relativistic theory. Now, in the absence of gravitational fields, there is a Cartesian coordinate system such that $\Gamma^{\mu}_{\alpha\beta}(x) = 0$ for all x. Similarly, in the absence of electromagnetic fields, there is a gauge such that $\Gamma_{i\,j\mu}(x) = 0$ for all x. Combining eqs(8.4,10), we have

$$\Gamma'_{i\,j\mu}(x) = -\varepsilon_{i\,j}\lambda A'_\mu(x) = -\varepsilon_{i\,j}\lambda\left[A_\mu(x) - \partial_\mu\theta(x)\right]$$

Setting this to zero gives $A_\mu(x) = \partial_\mu\theta(x)$. Thus, a vector potential of the form

$A_\mu = \partial_\mu \theta$ should denote the case of zero electromagnetic fields. Indeed, we see that

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad = \partial_\mu \partial_\nu \theta - \partial_\nu \partial_\mu \theta = 0$$

Furthermore, given $A_\mu = \partial_\mu \theta$, a gauge transformation $A_\mu \to A'_\mu = A_\mu - \partial_\mu \theta$ gives $A'_\mu = 0$ so that the original Dirac equation is recovered. This is tantamount to choosing the inertial Cartesian coordinates in special relativity.

8.1.6. Field Equations

We now attempt to construct the field equations for the gauge fields. We begin with the 1-D Riemann tensor $R_{\mu\nu}$ defined by [c.f. eq(2.34)],

$$\left[D_\mu, D_\nu\right]\phi(x) = R_{\mu\nu}(x)\phi(x)$$

$$\Rightarrow \qquad L.H.S. = \left[\left(\partial_\mu + i\lambda A_\mu\right)\left(\partial_\nu + i\lambda A_\nu\right) - \left(\partial_\nu + i\lambda A_\nu\right)\left(\partial_\mu + i\lambda A_\mu\right)\right]\phi$$

$$= i\lambda\left[\partial_\mu\left(A_\nu\phi\right) + A_\mu\partial_\nu\phi - \partial_\nu\left(A_\mu\phi\right) - A_\nu\partial_\mu\phi\right]$$

$$= i\lambda\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right)\phi$$

$$\therefore \qquad R_{\mu\nu} = i\lambda\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right) \qquad (8.14)$$

Since A_μ transforms according to (8.10), we see that $R_{\mu\nu}$ is gauge invariant. The simplest Lagrangian density that is a scalar under both Lorentz and gauge transformations is therefore $\alpha R_{\mu\nu}R^{\mu\nu}$, where α is a constant. Thus, if we identify A_μ as the electromagnetic vector potential, then

$$F_{\mu\nu} = \frac{1}{i\lambda}R_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

will be the Maxwell field tensor. The total action can therefore be written as

$$S = \int d^4x\left[-\frac{1}{4e^2}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}\left(i\not{\partial} - \lambda\not{A} - m\right)\psi\right] \qquad (8.15a)$$

where the coupling constant e^2 has the dimension of $(\text{charge})^2$. Note that $F_{\mu\nu}$ scales with A_μ, i.e.,

$$A_\mu \to \lambda A_\mu \qquad \Rightarrow \qquad F_{\mu\nu} \to \lambda F_{\mu\nu}$$

Thus, a scale change $A_\mu \to \lambda A_\mu$ in (8.15a) may be taken as a change of the coupling strength between the particle and the electromagnetic fields. Therefore, for the case where n types of spin 1/2 particles are present, we can write

$$S = \int d^4x\left[-\frac{1}{4e^2}F_{\mu\nu}F^{\mu\nu} + \sum_{j=1}^{n}\bar{\psi}_j\left(i\not{\partial} - \lambda_j\not{A} - m_j\right)\psi_j\right] \qquad (8.15)$$

where λ_j indicate the different coupling strengths of the particles to the *same* field $F_{\mu\nu}$. Finally, to make contact with the conventional notations, we make the rescaling

$$A_\mu \to eA_\mu \qquad \Rightarrow \qquad F_{\mu\nu} \to eF_{\mu\nu} \tag{8.16}$$

so that eq(8.15) becomes

$$S = \int d^4x \left[-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \sum_{j=1}^{n} \bar{\psi}_j \left(i\partial\!\!\!/ - e\lambda_j A\!\!\!/ - m_j \right) \psi_j \right] \tag{8.17}$$

It is left as an exercise to show that the Euler-Lagrange equations for the gauge fields are just the Maxwell equations with current density of the *i*-type particles given by [cf. (7.62)]

$$j_i^\mu(x) = e\lambda_i \bar{\psi}(x) \gamma^\mu \psi(x)$$

Thus, we can interpret e as some elementary charge unit and $\lambda_j e$ the charge of particle j. However, since there is no restrictions on λ_j, our theory does not provide an explanation of the observation that all charges are integral multiples of the electron charge. An explanation of this "charge quantization" is offered by the grand unified theory [see Chapter 12].

8.2. Non-Abelian Gauge Theories

8.2.1. Isospin

Nuclei are made up of protons and neutrons, known collectively as nucleons. Except for their charges, proton and neutron are quite similar. Their masses differ by only a few percent and, more significantly, they are interchangeable as far as the strong interaction is concerned. This leads to the idea that they are just different states of the same particle, called conveniently the nucleon. Let $\psi_p(x)$ and $\psi_n(x)$ be spin 1/2 Dirac spinors denoting the proton and neutron states, respectively. The nucleon wavefunction is defined as

$$\psi_N(x) = \begin{pmatrix} \psi_p(x) \\ \psi_n(x) \end{pmatrix} \qquad (8.18)$$

with the understanding that $\psi_N(x) = \begin{pmatrix} \psi_p(x) \\ 0 \end{pmatrix}$ denotes the pure proton state, and $\psi_N(x) = \begin{pmatrix} 0 \\ \psi_n(x) \end{pmatrix}$ the neutron state. Now, ψ_N can be taken as a 2-component matrix analogous to that describing the non-relativistic spin 1/2 states. In this context, ψ_N is called an **isotopic spin** (**isospin**) state. As discussed in appendix B, the complete set of independent operators in the corresponding **isospin space** can be chosen as the unit matrix I and the Pauli matrices $\boldsymbol{\tau}$. The isospin operator is then $\mathbf{T} = \frac{1}{2}$. Since $\boldsymbol{\tau}$ are hermitian, any unitary operator that leaves the magnitude $\sqrt{\psi_N^+ \psi_N}$ unchanged can be written as

$$U(\theta, \boldsymbol{\alpha}) = \exp\left[i \left(\theta I + \frac{1}{2} \boldsymbol{\alpha} \cdot \boldsymbol{\tau} \right) \right] \qquad (8.19)$$

where θ and $\boldsymbol{\alpha}$ are real parameters. In particular, θ specifies a phase transformation and $\boldsymbol{\alpha}$ a rotation in the 3-D isospin space. With respect to the quantization (z-) axis of the isospin space, the proton and neutron states are the isospin up and down states, respectively.

8.2.2. Isospin Connection

Consider now the fibre bundle with spacetime as the base manifold and the isospin space as the typical fibre. In contrast with the complex phase, directions in the isospin space have observable physical meanings, i.e., a proton, a neutron, or some mixed state. As in (8.3), we introduce a parallel transport

$$\psi_i(x \to x+\Delta x) = \psi_i(x) - \Gamma_{ij\mu}(x)\psi_j(x)\Delta x^\mu \qquad (8.21a)$$

where $i, j = p$ or n. Now, the only meaningful change in isospin space is a rotation. Thus, keeping only terms linear in $\boldsymbol{\alpha}$ in eq(8.20), we write,

$$\psi_N(x \to x+\Delta x) \simeq \left(I + i\frac{1}{2}\boldsymbol{\alpha}\cdot\boldsymbol{\tau}\right)\psi_N(x) = \left(I + i\frac{1}{2}\alpha^a\tau^a\right)\psi_N(x)$$

where $\boldsymbol{\alpha}$ is an infinitesimal rotation vector and $a = 1, 2, 3$. Writing $\alpha^a = -A_\mu^a \Delta x^\mu$ and comparing with (8.21a), we have

$$\Gamma_{ij\mu}(x) = i\frac{1}{2}A_\mu^a(x)\left(\tau^a\right)_{ij} \qquad (8.21')$$

where $A_\mu^a(x)$ are gauge fields. Note that there is no scale factor λ in (8.21') because it is found that the field tensor does not scale with the gauge fields [see §8.2.3].

The typical fibre can be generated by the set of all rotations $\left\{U(\boldsymbol{\alpha}) \,\middle|\, \sqrt{\boldsymbol{\alpha}\cdot\boldsymbol{\alpha}} < 2\pi\right\}$, which forms the group SU(2). This is a non-Abelian group, i.e., some of its elements do not commute. The theory is easily generalized by replacing $\frac{1}{2}\boldsymbol{\tau}$ with a general isospin $\mathbf{T}$ of (2T+1) components. In which case, the gauge covariant derivative can be written as

$$D_\mu\psi_i(x) = \partial_\mu\psi_i(x) + \Gamma_{ij\mu}(x)\psi_j(x) \qquad i, j = -T, -T+1, \cdots, T-1, T$$

$$= \partial_\mu\psi_i(x) + iA_\mu^a(x)\left(T^a\right)_{ij}\psi_j(x) \qquad a = 1, 2, 3$$

$$= \partial_\mu\psi_i(x) + i\left[A_\mu(x)\right]_{ij}\psi_j(x) \qquad (8.22a)$$

where

$$A_\mu(x) = A_\mu^a(x)T^a \qquad (8.23)$$

so that (8.22a) can be written as

$$D_\mu \psi(x) = \partial_\mu \psi(x) + iA_\mu(x)\psi(x)$$

$$\Rightarrow \qquad D_\mu = \partial_\mu + iA_\mu(x) \qquad (8.22)$$

Under a gauge transformation, we have

$$\psi(x) \to \psi'(x) = U(\boldsymbol{\alpha})\psi(x) \quad = \exp\left[i\boldsymbol{\alpha}(x)\cdot\mathbf{T}\right]\psi(x) \qquad (8.24)$$

Thus,

$$D'_\mu \psi' \equiv \left(\partial_\mu + iA'_\mu\right)\psi' = \left(\partial_\mu + iA'_\mu\right)U\psi \quad = \left(U\partial_\mu + \partial_\mu U + iA'_\mu U\right)\psi$$

Since D_μ is a gauge scalar, we have

$$D'_\mu \psi' = UD_\mu \psi \quad = \left(U\partial_\mu + iUA_\mu\right)\psi$$

$$\Rightarrow \qquad iUA_\mu = \partial_\mu U + iA'_\mu U$$

$$A'_\mu = UA_\mu U^{-1} + i\left(\partial_\mu U\right)U^{-1} \qquad (8.26)$$

which is the gauge transformation rule for the gauge fields. Note that by setting $U = e^{i\theta(x)}$, eq(8.10) is recovered.

8.2.3. Field Tensor

As in §8.1.6, we now attempt to construct the field equations for the gauge fields. The Riemann tensor $R_{\mu\nu}$ is defined by [c.f. eq(2.34)],

$$\left[D_\mu, D_\nu\right]\psi(x) = R_{\mu\nu}(x)\psi(x)$$

$$\Rightarrow \qquad L.H.S. = \left[\left(\partial_\mu + iA_\mu\right)\left(\partial_\nu + iA_\nu\right) - \left(\partial_\nu + iA_\nu\right)\left(\partial_\mu + iA_\mu\right)\right]\psi$$

$$= i\left[\partial_\mu\left(A_\nu\psi\right) + A_\mu\partial_\nu\psi - \partial_\nu\left(A_\mu\psi\right) - A_\nu\partial_\mu\psi\right] - \left(A_\mu A_\nu - A_\nu A_\mu\right)\psi$$

$$= i\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right)\psi - \left(A_\mu A_\nu - A_\nu A_\mu\right)\psi$$

$$\therefore \qquad R_{\mu\nu} = i\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right) - \left[A_\mu, A_\nu\right]$$

The field tensor is defined by

$$F_{\mu\nu} = -iR_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + i\left[A_\mu, A_\nu\right] \qquad (8.27)$$

Note that the presence of the commutator makes $F_{\mu\nu}$ nonlinear in A_μ and destroys its gauge invariance. Thus, $F_{\mu\nu}$ no longer scales with A_μ. A rescaling $A_\mu^a \to \lambda A_\mu^a$ would indicate coupling to a *different* field so that there is no point in including a scaling factor λ in (8.21). Consequently, different particles of the same isospin must have the same isospin connection, i.e., only particles of different isospins can have different connections. Another way to arrive at these conclusions is via the observation that directions in the isospin space have definite physical meanings. Thus, the angles of rotations, and hence the gauge fields, cannot be re-scaled while leaving the system unchanged.

The exact form of eq(8.27) depends on the particular representation of the gauge group we use. Since the generators of the gauge (Lie) group are **T**, the corresponding Lie algebra is defined by

$$\left[T^a, T^b\right] = iC^{abc}T^c \quad = i\varepsilon^{abc}T^c \qquad (8.28)$$

where we have used the fact that the **structure constants** C^{abc} for SU(2) are simply ε^{abc} [see Appendix B]. Using (8.23), we can write eq(8.27) as

$$F_{\mu\nu} = \partial_\mu\left(A_\nu^a T^a\right) - \partial_\nu\left(A_\mu^a T^a\right) + iA_\mu^a A_\nu^b\left[T^a, T^b\right]$$

$$= \left(\partial_\mu A_\nu^a\right)T^a - \left(\partial_\nu A_\mu^a\right)T^a - A_\mu^a A_\nu^b C^{abc}T^c$$

$$= \left(\partial_\mu A_\nu^a - \partial_\nu A_\mu^a - A_\mu^b A_\nu^c C^{bca}\right) T^a$$

$$= F_{\mu\nu}^a T^a \qquad (8.29)$$

where, with the help the antisymmetry of C^{abc}, we have set

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a - C^{abc} A_\mu^b A_\nu^c \qquad (8.30)$$

8.2.4. Gauge Transformation

By definition, a gauge transformation is a rotation on ψ given by

$$\psi \to \psi' = U(\boldsymbol{\alpha})\psi = \exp(i\boldsymbol{\alpha}\cdot\mathbf{T})\psi$$

Thus, ψ is a gauge vector. Since T^a is a generator of the transformation, it is a gauge tensor of rank 2 so that

$$T^{a\prime} = UT^aU^{-1} \qquad \text{and} \qquad U(T^a\psi) = UT^aU^{-1}U\psi = T^{a\prime}\psi'$$

As show in §8.2.2,

$$A'_\mu = UA_\mu U^{-1} + i(\partial_\mu U)U^{-1} \qquad (8.26)$$

so that A_μ is not a gauge tensor. According to (8.27),

$$F_{\mu\nu} \to F'_{\mu\nu} = \partial_\mu A'_\nu - \partial_\nu A'_\mu + i\left[A'_\mu, A'_\nu\right]$$

We shall leave it as an exercise [see Ex.8.3] to show that $F_{\mu\nu}$ is a gauge tensor of rank 2, i.e.,

$$F'_{\mu\nu} = UF_{\mu\nu}U^{-1} \qquad (8.31)$$

Hence,

$$F^{a}_{\mu\nu}{}' T^{a\prime} = UF^a_{\mu\nu}T^aU^{-1} = UF^a_{\mu\nu}U^{-1}UT^aU^{-1} = UF^a_{\mu\nu}U^{-1}T^{a\prime}$$

$$\Rightarrow \qquad F^{a}_{\mu\nu}{}' = UF^a_{\mu\nu}U^{-1}$$

On the other hand, we can use the three T^a as basis for vector operators on the isospin space. A gauge transformation is then a rotation operator $\mathcal{U}$ defined by its action on the basis:

$$T^a \to T^{a\prime} = \mathcal{U}T^a = T^b\mathcal{U}^{ba}(\boldsymbol{\alpha}) \qquad (8.33)$$

where $\mathcal{U}^{ba}(\boldsymbol{\alpha})$ are the coefficients of the transformation. They can be found by comparing (8.33) with $T^{a\prime} = U(\boldsymbol{\alpha})T^aU^{-1}(\boldsymbol{\alpha})$. In this sense, the relation $F_{\mu\nu} = F^a_{\mu\nu}T^a$ is to be taken as expressing the vector $F_{\mu\nu}$ in terms of its components $F^a_{\mu\nu}$ with respect to the basis T^a. Under the gauge transformation, we have,

$$F_{\mu\nu} \to F'_{\mu\nu} = \mathcal{U} F_{\mu\nu} = F^a_{\mu\nu} \mathcal{U} T^a = F^a_{\mu\nu} T^b \mathcal{U}^{ba}(\boldsymbol{\alpha}) \equiv F^a_{\mu\nu}{}' T^a$$

so that the rotated vector has components

$$F^a_{\mu\nu}{}' = \mathcal{U}^{ab}(\boldsymbol{\alpha}) F^b_{\mu\nu} \qquad (8.32)$$

or $$F_{\mu\nu}{}' = \mathcal{U}(\boldsymbol{\alpha}) F_{\mu\nu}$$

where $\mathcal{U}(\boldsymbol{\alpha})$ is the matrix with elements $\mathcal{U}^{ab}(\boldsymbol{\alpha})$. Clearly, there is an isomorphism between U and $\mathcal{U}$. Thus, analogous to (8.20), we can write

$$\mathcal{U}(\boldsymbol{\alpha}) = \exp(i\boldsymbol{\alpha} \cdot \mathcal{T}) \qquad (8.34)$$

The SU(2) representation formed by the three 3×3 matrices $\mathcal{T}^a$ is just the **adjoint representation**, so called because their elements can be constructed from the structural constants as

$$(\mathcal{T}^a)_{bc} = -iC^{abc} \qquad (8.35)$$

Proof of this is left as exercise [see Ex.8.4].

8.2.5. Intermediate Vector Bosons

Our task now is to construct a gauge invariant action for the gauge fields. To begin, consider the obvious generalization of (8.15):

$$S = -\frac{1}{4g^2}\int d^4x\, Tr\left(F_{\mu\nu}F^{\mu\nu}\right)$$

where g is a coupling constant. Using (8.29), we have

$$Tr\left(F_{\mu\nu}F^{\mu\nu}\right) = F^a_{\mu\nu}F^{b\mu\nu}Tr\left(T^aT^b\right)$$

Therefore, the simplest way to ensure that $Tr\left(F_{\mu\nu}F^{\mu\nu}\right)$ is a gauge scalar is to demand

$$Tr\left(T^aT^b\right) = \lambda\delta^{ab} \qquad (8.40)$$

so that

$$S = -\frac{1}{4g^2}\int d^4x\, F^a_{\mu\nu}F^{a\mu\nu} \qquad (8.36)$$

It is straightforward to show that the Pauli matrices satifsy (8.40). For an arbitary set of generators **T′**, we can always make a linear transformation to another set **T** that satisfies (8.40). Henceforth, we shall always assume this is done.

Under the scaling $A^a_\mu \to gA^a_\mu$, we have

$$F^a_{\mu\nu} \to g\left(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu - gC^{abc}A^b_\mu A^c_\nu\right) = g\tilde{F}^a_{\mu\nu}$$

$$S \to -\frac{1}{4}\int d^4x\, \tilde{F}^a_{\mu\nu}\tilde{F}^{a\mu\nu}$$

Dropping the ~ for the sake of clarity, we have

$$S = -\frac{1}{4}\int d^4x\, F^a_{\mu\nu}F^{a\mu\nu} \qquad (8.38)$$

where

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - gC^{abc}A^b_\mu A^c_\nu \qquad (8.37)$$

The quantized gauge fields correspond to **intermediate vector bosons**, which, like photons in electromagnetism, mediate the nuclear weak interaction. Putting (8.37) into (8.38), we see that the action contains terms of the form

$$g\left(\partial A\right)AA \qquad \text{and} \qquad g^2AAAA$$

which represent interactions between the vector bosons themselves [see Chapter 9]. Thus, the carriers of the interaction are themselves “charged”. In contrast, the carriers (photons) of electromagnetic fields are “neutral” and there are no interaction terms between photons in the abelian gauge theory. The nonlinearity of the non-abelian gauge theory is more akin to that of the gravitational theory. Indeed, we expect interaction terms between the gravitons to be added to (7.119) when the assumption of small $h_{\mu\nu}$ is lifted. Thus, gravitons are expected to be “charged”. This is inevitable since the gravitational charge is just the mass / energy of the particle. Unfortunately, quantum theory of gravity developed along this direction is unviable mathematically.

8.2.6. Action

Including n spin 1/2 Dirac particles into the action (8.36) gives

$$S = \int d^4x \left[-\frac{1}{4g^2} F^a_{\mu\nu} F^{a\mu\nu} + \sum_{j=1}^{n} \bar{\psi}_j \left(i\partial\!\!\!/ - A\!\!\!/ - m_j \right) \psi_j \right]$$

Rescaling by $A \to gA$ gives

$$S = \int d^4x \left[-\frac{1}{4} F^a_{\mu\nu} F^{a\mu\nu} + \sum_{j=1}^{n} \bar{\psi}_j \left(i\partial\!\!\!/ - gA\!\!\!/^{(j)} - m_j \right) \psi_j \right] \tag{8.41}$$

where $F^a_{\mu\nu}$ is given by (8.37) and each ψ_j is a $2T^{(j)}+1$ multiplet of 4-component Dirac spinors. Schematically, we have

$$\psi_j = \left. \begin{pmatrix} (\;\;) \\ (\;\;) \\ \vdots \\ (\;\;) \end{pmatrix} \begin{matrix} \}4 \\ \}4 \\ \\ \}4 \end{matrix} \right\} \left[2T^{(j)} + 1 \right] \times 4$$

The matrix $A\!\!\!/^{(j)}$ is the shorthand for

$$A\!\!\!/^{(j)} = \gamma^\mu A_\mu = A^a_\mu \gamma^\mu T^{(j)a} \tag{8.42}$$

where $a = -T^{(j)}, -T^{(j)}+1, \cdots, T^{(j)}-1, T^{(j)}$. The Euler-Lagrange equations for the field degrees of freedom can be found by inspection to be

$$D_\mu F^{\mu\nu} = J^\nu \qquad \text{or} \qquad \partial_\mu F^{a\mu\nu} - gC^{abc} A^b_\mu F^{c\mu\nu} = J^{a\nu} \tag{8.43}$$

where the current density is given by

$$J^\nu = g \sum_{j=1}^{n} \bar{\psi}_j \gamma^\nu \psi_j \qquad \text{or} \qquad J^{a\nu} = g \sum_{j=1}^{n} \bar{\psi}_j \gamma^\nu T^{(j)a} \psi_j \tag{8.44}$$

For the nucleon doublet, we have

$$j^{3\nu} = g \begin{pmatrix} \bar{\psi}_p & \bar{\psi}_n \end{pmatrix} \gamma^\nu \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \psi_p \\ \psi_n \end{pmatrix}$$

$$= g \frac{1}{2} \left(\bar{\psi}_p \gamma^\nu \psi_p - \bar{\psi}_n \gamma^\nu \psi_n \right)$$

$$= g \sum_{p,n} T^3 \times \left(\textit{probability current density} \right) \tag{8.45}$$

Finally, the Euler-Lagrange eqautions for the spinor degrees of freedom are simply the Dirac equations

$$\left(i\not{\partial} - g\not{A}^{(j)} - m_j\right)\psi_j = 0$$

8.2.7. Conserved Currents

In classical electrodynamics, gauge invariance implies the conservation of charges, which can be expressed by a conserved current as $\partial_\mu j^\mu = 0$. For gauge fields, the conservation law that can be derived from the equation of motion is $D_\mu J^\mu = 0$. In which case, J^μ is said to be **covariantly conserved**. In contrast to $\partial_\mu j^\mu = 0$, the equation $D_\mu J^\mu = 0$ does not imply the conservation of any physical scalar quantity.

For a Dirac particle governed by eq(8.13), we see that $D_\mu J^\mu = 0$ with $j^\mu = \lambda e \bar{\psi} \gamma^\mu \psi$. However, we also have $D_\mu j^\mu = \partial_\mu j^\mu = 0$ so that we again recover the conservation of charges.

For the non-abelian SU(2) gauge group, the current (8.44) is covariantly conserved and satsifies

$$D_\mu J^\mu = 0 \qquad \text{or} \qquad \partial_\mu J^{a\mu} - gC^{abc} A_\mu^b J^{c\mu} = 0 \qquad (8.46)$$

On the other hand, if we take the partials of the non-Abelian Maxwell equations (8.43) , we get, with the help of $\partial_\nu \partial_\mu F^{a\mu\nu} = 0$, that

$$\partial_\nu \left(J^{a\nu} + gC^{abc} A_\mu^b F^{c\mu\nu} \right) = 0$$

Hence, the modified current

$$\tilde{J}^{a\nu} = J^{a\nu} + gC^{abc} A_\mu^b F^{c\mu\nu} \qquad (8.47)$$

is conserved in the meaning sense. In fact, $\tilde{J}^{a\nu}$ is the **Noether current** associated with the non-Abelian symmetry. As flows of isospin, the 1st and 2nd term on the right side of (8.47) represent the contributions from the Fermion and the vector bosons, respectively.

Analogous to the abelian theory, the components of the field tensor (8.37) can be thought of as ‘electric’ and ‘magnetic’ fields $\mathbf{E}^a$ and $\mathbf{B}^a$. Thus,

$$E^{ai} = F^{ai0} = -F^{a0i} \qquad\qquad B^{ai} = -\frac{1}{2}\varepsilon^{ijk}F^{ajk}$$

As show in chapter 3, the equation $\nabla \cdot \mathbf{B} = 0$ forbids the existence of monopoles. In the non-Abelian theory, the corresponding equation is easily shown to be

$$\partial_i B^{ai} = gC^{abc}A_i^b B^{ci} \qquad (8.49)$$

Hence, 'magnetic monopoles' are allowed in non-abelian gauge fields. Obviously, B^{ai} in (8.49) are not what one would call magnetic fields. However, in unified theories, the electromagnetic fields can be combined with the weak interaction in a non-abelian gauge theory. In that case, genuine magnetic monopoles are allowed [see chapter 13].

8.3. Non-Abelian Theories and Electromagnetism

Replacing $\frac{1}{2}\boldsymbol{\tau}$ with the more general isospin **T**, the general unitary transformation (8.19) becomes

$$U(\theta,\boldsymbol{\alpha}) = \exp\left[i\left(\theta I + \boldsymbol{\alpha}\cdot\mathbf{T}\right)\right] \qquad (8.19a)$$

which contains a phase transformation $e^{i\theta I} = e^{i\theta} I$ that we have so far ignored. Now, $\{e^{i\theta}\}$ forms the group U(1) and $\{e^{i\boldsymbol{\alpha}\cdot\mathbf{T}}\}$ the group SU(2). Since $e^{i\theta} I$ commutes with $e^{i\boldsymbol{\alpha}\cdot\mathbf{T}}$, the group $\{U(\theta,\boldsymbol{\alpha})\}$ can be written as a direct product

$$\{U(\theta,\boldsymbol{\alpha})\} = SU(2)\times U(1) \qquad (8.50a)$$

which suggests a unification of electromagnetism with the interaction represented by non-Abelian gauge fields.

However, there is a technical detail that must be taken care of first. This can be illustrated by the case of isospin 1/2. From

$$e^{i\theta} I\,\psi = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{i\theta} \end{pmatrix}\begin{pmatrix} \psi_p \\ \psi_n \end{pmatrix} = \begin{pmatrix} e^{i\theta}\psi_p \\ e^{i\theta}\psi_n \end{pmatrix}$$

we see that $e^{i\theta} I$ does not distinguish between a charged proton and a neutral neutrons. Hence, it cannot be the gauge transformation *G* that represents electromagnetism, which affects only charged particles. Indeed, we expect

$$G(\omega) = \begin{pmatrix} e^{i\omega} & 0 \\ 0 & 1 \end{pmatrix} \qquad (8.52)$$

Comparing with (8.19a), we see that $G(\omega) = U(\theta,\boldsymbol{\alpha})$ with

$$\theta I + \frac{1}{2}\boldsymbol{\alpha}\cdot\boldsymbol{\tau} = \omega\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad (8.51)$$

Using the standard representations (7.28) of $\boldsymbol{\tau}$, we have

$$\theta I + \frac{1}{2}\boldsymbol{\alpha}\cdot\boldsymbol{\tau} = \begin{pmatrix} \theta + \frac{1}{2}\alpha_3 & \frac{1}{2}(\alpha_1 - i\alpha_2) \\ \frac{1}{2}(\alpha_1 + i\alpha_2) & \theta - \frac{1}{2}\alpha_3 \end{pmatrix}$$

so that (8.51) can be satisfied only if

$$\alpha_1 = \alpha_2 = 0 \qquad \theta = \frac{1}{2}\omega \qquad \alpha_3 = \omega \qquad (8.50a)$$

Obviously, $\{G(\omega)\}$ forms a group U(1). Indeed, it is the U(1) subgroup of

SU(2)×U(1) that represents electromagnetism. The unification scheme is thus verified.

For a general isospin *T*, eq(8.52) is generalized to

$$G(\omega) = \begin{pmatrix} e^{iQ_1\omega} & 0 & \cdots \\ 0 & e^{iQ_2\omega} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \qquad (8.53)$$

where Q_j, with $j = -T, -T+1, \cdots, T-1, T+1$, is the charge of the *j*-th isospin multiplet, measured as multiples of the fundamental charge. In a representation where T^3 is diagonal, the diagonal part of eq(8.50a) is easily generalized to

$$\theta = \frac{1}{2}Y\omega \qquad Q_j = T_j^3 + \frac{1}{2}Y \qquad (8.54a)$$

where we have introduced the **hypercharge** *Y* to coup with more general situations [see (8.54)]. Thus, the largest charge of the multiplets is

$$Q = T^3 + \frac{1}{2}Y \qquad (8.54)$$

which is just the **Gell-Mann- Nishijima relations** proposed originally for a phenomenological classification of the observed particles.

8.3.a. Gell-Mann- Nishijima Law

The Gell-Mann- Nishijima law

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

was proposed in 1953 to explain the “8-fold way” grouping of “stable” hadrons. Here, “stable” means no decay if electroweak interactions were absent. One such grouping is

$$\begin{array}{ccccc} & n^0 & & p^+ & \\ & & \Sigma^0 & & \\ \Sigma^- & & & & \Sigma^+ \\ & & \Lambda^0 & & \\ & \Xi^- & & \Xi^0 & \end{array}$$

The corresponding values of (Q, I_3, Y) are

$$\begin{array}{ccccc} & \left(0,-\frac{1}{2},1\right) & & \left(1,\frac{1}{2},1\right) & \\ & & (0,0,0) & & \\ (-1,-1,0) & & & & (1,1,0) \\ & & (0,0,0) & & \\ & \left(-1,-\frac{1}{2},-1\right) & & \left(0,\frac{1}{2},-1\right) & \end{array}$$

where the directions of increasing values are $Q\nearrow$, $I_3\rightarrow$, and $Y\uparrow$. Note that the hypercharge Y and strangeness S are related by $Y = S$ for mesons and $Y = S + 1$ for baryons.

8.4. Relevance of Non-Abelian Theories to Physics

In §8.1, we have shown that pure geometrical consideration of the wavefunction leads to the abelian gauge fields and hence the existence of electromagnetic forces. Furthermore, applying the same idea to the isospin also results in non-abelian gauge fields that seemingly described the nuclear weak interaction. These non-abelian gauge theories are also known as **Yang-Mills theories**, in honor of their creators.

When the Yang-Mills theory was 1st proposed in 1954, nucleons were considered as fundamental particles that cannot be further divided. Nowadays, the truly fundamental particles are thought to be quarks, leptons and the quanta of fundamental interactions such as photons. However, the idea of gauge fields remains valid so that we need only shift our attention to these more fundamental building blocks. The rest of this book will be devoted to this endeavor. [See §8.4 of Lawrie for an introductory discussion].

8.5. The Theory of Kaluza and Klein

The classical (non-quantum mechanical) theory of Kaluza and Klein unifies gravity and electromagnetism by means of a 5-D spacetime.

To begin, consider the metric tensor $\tilde{g}_{AB}$ of the spacetime. In general, an overhead ~ is used to indicate a quantity that is, or specific to, 5-D. Furthermore, to showcase the extra dimension, we set A, B = 0, 1, 2, 3, 5. Next, we set

$$\tilde{g}_{5\mu} = \tilde{g}_{\mu 5} = \tilde{g}_{55} A_\mu \qquad \text{and} \qquad \tilde{g}_{\mu\nu} = g_{\mu\nu} + \tilde{g}_{55} A_\mu A_\nu \qquad (8.55)$$

where $\mu = 0,1,2,3$ and $g_{\mu\nu}$ is the metric tensor of the Einstein's 4-D spacetime. In analogy with eqs(4.14, 16), the action for "gravity" is assumed to be

$$S = -\frac{1}{16\pi\tilde{G}} \int d^5x \sqrt{|\tilde{g}|}\, \tilde{R} \qquad (8.56)$$

Next, we make the following assumptions:

1. The 5th dimension is space-like, i.e., $g_{55} < 0$ so that $\tilde{g} > 0$.

2. $g_{\mu\nu}$ and A_μ are independent of x^5 and $\tilde{g}_{55} = const$.

3. The 5th dimension rolls into a circle of radius r_5 [see Fig.8.1].

To conform with our usual perception of a 4-D spacetime, the extend ($2\pi r_5$) of x^5 must be much smaller than the smallest length measurable at present. With the help of these assumptions and (8.55), eq(8.56) can be written as

$$S = -\int d^4x \sqrt{|g|} \left(\frac{1}{16\pi G} R + \frac{1}{4e^2} F_{\mu\nu} F^{\mu\nu} \right) \qquad (8.57)$$

where

$$G = \frac{\tilde{G}}{2\pi r_5 \sqrt{|\tilde{g}_{55}|}} \qquad e^2 = \frac{8\tilde{G}}{r_5 |\tilde{g}_{55}|^{3/2}} \qquad (8.58)$$

Now, (8.57) is simply the action for an Einsteinian spacetime with both gravity and electromagnetism. Hence, the claimed unification. Derivation of (8.57) is much too complicated to be described here. [Lawrie called (8.57) a miracle]. Interested readers are referred to the original paper: O. Klein, Z. Phys. 37, 985 (1926).

One deeply unsatisfying feature of the theory is that there is no physical justification to the required assumptions listed above. Furthermore, it offers no new observable effects. Thus, for a long time, its status is no more than that of a beautiful toy. However, interest in it was revived after the advent of the **supergravity** and

superstring theories, both of which made use of spacetimes of more than 4 dimensions [see chapter 15].