Connections, Gauges and Field Theories

Lionel Brits

August 3, 2005

Abstract

The theory of gauges and connections in the principal bundle formalism is reviewed. The geometrical aspects of gauge potential, such as curvature are explored. Finally, gauge field theories such as the Yang-Mills and General Relativistic theories are reviewed in terms of connections.

1 Introduction

Physical systems exist independently of the mathematical objects used to represent them. It often happens that these objects expose additional degrees of freedom that the physical system is insensitive to. These parameters, which do not enter the predictions of the theory, are known as *gauge parameters*, and theories that contain such parameters are known as *gauge theories*. The existence of these degrees of freedom may nevertheless have important physical consequences to the theory.

Gauge parameters arise in classical mechanics, although their treatment here is more subtle. For example, we often use an absolute spatial coordinate x^i to describe the position of a particle. However, if the Lagrangian of the system does not depend explicitly on x^i, then we say that the system is invariant under a transformation $x^i \to x^i + \Delta x^i$. It follows from the Euler-Lagrange equations of motion that canonical momentum p_i conjugate to x^i is conserved:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right) = \frac{\partial L}{\partial x^i} = 0 \tag{1}$$

Similarly, if the Lagrangian does not depend explicitly on the absolute polar angle ϕ, then the system is invariant under coordinate rotations, and the angular momentum p_ϕ conjugate to ϕ is conserved. In general, the existence of a conserved quantity arising from the invariance or symmetry of a system under a group of transformations is known as Noether's theorem. It finds broad applicability outside of classical mechanics, especially in quantum mechanics and particle physics. For example, the invariance of the wavefunction under a global U(1) phase transformation gives rise to conservation of charge [Ryder, 1985]. As will also be seen, turning the invariance into a local one also gives rise to the existence of electromagnetic theory.

2 Mathematical Preliminaries

2.1 Motivation

The gauge formalism is usually encountered in the study of fields. More formally, given a fibre bundle over a manifold $\mathcal{M}$, a *field* is a cross-section of the bundle, i.e., a continuous map from an open subset of $\mathcal{M}$ to points in the fibre bundle. To each point on the manifold, an element of the fibre at that point is picked out. In addition to scalar, vector and tensor fields, quantum field theories also make use of spinor fields, where the fibre elements are spinors. In general, fields will transform in a specific manner upon transformations of the manifold, and the components of the field may also transform into each other independently of the manifold. The latter type of transformation gives rise to internal symmetries of the field, and is the type of gauge transformation normally considered.

While we can at any point in the manifold transform the internal components of the field in any way that is consistent with the symmetries of the field, there is no direct way to carry such a transformation over to a neighbouring part of the field, since the values of the field at different points reside in different spaces altogether. What is needed, then, is a means by which to relate transformations at different points on the manifold, and, in doing so, relate elements of the fibres at these points. This is accomplished by a connection, and is most clearly introduced in terms of the principal bundle formalism.

2.2 The Principal Bundle

Given a fibre bundle $E \xrightarrow{\pi} \mathcal{M}$ with a fibre F and a base space $\mathcal{M}$, let G be its structure group, i.e., a Lie group which acts on F on the left. For instance, if E is the tangent bundle $T\mathcal{M}$ then G is the general linear group $Gl(n, \mathbb{R})$ of $n \times n$ real, invertible matrices that transform tangent vectors in $T_p\mathcal{M}$ into each other. We can then define the *principal bundle* $P \xrightarrow{\pi} \mathcal{M}$ over $\mathcal{M}$ as a fibre bundle with the structure group G as its fibre [Nakahara, 2003].

Now if we pick a particular element $g \in G_p$ at a point $p \in \mathcal{M}$ as our gauge transformation[1], then we can ask how g should vary as we move away from p. Since group elements at different points in $\mathcal{M}$ reside in different fibres of P, there is no *a priori* way to relate them, so it is necessary to introduce the concept of a *connection*: given a curve $\gamma(t) : [0, 1] \to \mathcal{M}$ and a point $u_0 \in P$, a connection allows us to define the *horizontal lift* $\tilde{\gamma}(t) : [0, 1] \to P$ of $\gamma(t)$ where $\tilde{\gamma}(0) = u_0$, i.e., a way to lift curves on $\mathcal{M}$ to curves in P. We can then speak of u_0 as being parallel transported along the curve γ, tracing out points on $\tilde{\gamma}$ as it is moved.

In order to make the notion of a horizontal lift more concrete, we need to decompose the tangent space T_uP of P at a point $u \in P$ into a subspace V_uP which is tangent

[1] Technically, a gauge transformation of the first type.

to the fibre G_u at u, and its complement, H_uP. These are known as the vertical and horizontal subspaces respectively. The parallel transport of the fibre at $p = \pi(u)$ to the fibre at $p' = \pi(u')$ can be defined uniquely by requiring that the curve $\tilde{\gamma}(t)$ connecting u and u' remain tangent to H_vP at any point v along the path [Nash and Sen, 1983].

One way to visualize the decomposition of the principal bundle into horizontal and vertical subspaces is to think of a ball rolling over a curved two-dimensional surface, where the group of rotations of the ball at a point forms the group G that acts on the rigid tetrad attached to the ball. The orientation of the ball at any point is given by $(\hat{x}, \hat{y}, \hat{z}) = (g\hat{x}_0, g\hat{y}_0, g\hat{z}_0)$ where $(\hat{x}_0, \hat{y}_0, \hat{z}_0)$ is some reference orientation, and g is some element of G. We can therefore think about g as being the orientation of the ball at a particular point, so that the complete configuration at a point with coordinates x^α is simply $(x^\alpha, g(x^\alpha))$. A curve $\tilde{\gamma}(t)$ through the configuration space has a tangent vector at any point p that lies in T_uP where $u = \gamma(t) \in \pi^{-1}(p)$. The vertical subspace V_uP at u is formed by all tangent vectors of the form $\frac{dg}{dt}\frac{\partial}{\partial g}$. If we apply a constraint upon $g(x^\alpha)$, for instance by requiring that the ball roll without slipping, then the horizontal subspace H_uP forms the space of all allowable "velocities" $(\dot{x}^\alpha, \dot{g})$ [Yamashita, 2002].

While V_uP can be found uniquely, it is the function of the connection to determine H_uP. To that end, we define the Lie algebra valued 1-form $\boldsymbol{\omega}$, called the connection 1-form, with the property that $\boldsymbol{\omega}(\mathbf{X}) = 0$ if $\mathbf{X} \in H_uP$. Let $g \in G$ be an element of the fibre at u. Then,

$$\boldsymbol{\omega} = g^{-1}\mathbf{d}g + g^{-1}\mathbf{A}g, \tag{2}$$

where $\mathbf{A} = A^a_\mu \lambda_a \mathbf{d}x^\mu$ is a Lie-algebra valued 1-form known as the *gauge potential* 1-form, and λ_a are the generators of G at u. Now, the basis of V_uP is $\frac{\partial}{\partial g_{ij}}$ by definition. Let the basis of H_uP be $\frac{\partial}{\partial x^\mu} + C_{\mu ij}\frac{\partial}{\partial g_{ij}}$ for some unknown $C_{\mu ij}$. Then any vector in T_uP can be written as

$$\mathbf{X} = X_{ij}\frac{\partial}{\partial g_{ij}} + X^\mu\left(\frac{\partial}{\partial x^\mu} + C_{\mu ij}\frac{\partial}{\partial g_{ij}}\right). \tag{3}$$

But if $X \in H_uP$, then $\boldsymbol{\omega}(\mathbf{X}) = 0$, hence $X_{ij} = 0$ and $C_{\mu ij} = -A^a_\mu \lambda_{aik} g_{kj}$. With H_uP constructed, we find that the vectors

$$D_\mu = \frac{\partial}{\partial x^\mu} - A^a_\mu \lambda_{aik} g_{kj} \frac{\partial}{\partial g_{ij}} \tag{4}$$

form a basis for H_uP, and we can write the tangent to any curve $\tilde{\gamma}(t)$ at u as $X^\mu D_\mu$ where the tangent to the curve $\gamma(t)$ at $\pi(u)$ in $\mathcal{M}$ is $X^\mu \frac{\partial}{\partial x^\mu}$ [Nash and Sen, 1983].

2.3 The Covariant Derivative on the Tangent Bundle

With the connection defined on the principal bundle P, we can obtain the connection on the tangent bundle $T\mathcal{M}$, since this is an associated vector bundle of P, i.e., the structure group of TM is the fibre of P (nash). We can obtain a horizontal lift $\gamma_{T\mathcal{M}}$ to $T\mathcal{M}$ from the horizontal lift γ_P to P. Given a tangent vector $\mathbf{X}$ from one of the fibres in $T\mathcal{M}$, and an operator $g \in Gl(n, \mathbb{R})$, $g\mathbf{X}$ is another tangent vector in the fibre. We can therefore define a horizontal lift $\gamma_{T\mathcal{M}}$ to $T\mathcal{M}$ by

$$\gamma_{T\mathcal{M}}(t) = \gamma_P(t)\mathbf{X}. \tag{5}$$

To see how $\gamma_{T\mathcal{M}}(t)$ is induced by $\gamma_P(t)$, consider

$$\frac{d\gamma_{T\mathcal{M}}(t)}{dt} = \frac{d\gamma_P(t)\mathbf{X}}{dt} = \frac{dx^\mu}{dt} D_\mu \gamma_P(t)\mathbf{X}. \tag{6}$$

Denoting $\mathbf{Y} = \frac{dx^\mu}{dt}\frac{\partial}{\partial x^\mu} = Y^\mu \frac{\partial}{\partial x^\mu}$ and $\mathbf{Z} = \gamma_P(t)\mathbf{X} = Z^\mu \frac{\partial}{\partial x^\mu}$, this can be abbreviated as $\nabla_{\mathbf{Y}}\mathbf{Z}$. Now,

$$D_\mu = \frac{\partial}{\partial x^\mu} - A^a_\mu \lambda_a g \frac{\partial}{\partial g}, \tag{7}$$

which, in local coordinates where $g = g^\mu_\nu$ and $A^a_\mu \lambda_a = A^\nu_{\mu\kappa}$, becomes

$$D_\mu = \frac{\partial}{\partial x^\mu} - A^\nu_{\mu\kappa} g^\kappa_\lambda \frac{\partial}{\partial g^\lambda_\nu}. \tag{8}$$

Hence,

$$\begin{aligned}
\nabla_{\mathbf{Y}}\mathbf{Z} &= Y^{\mu}\left(\frac{\partial}{\partial x^{\mu}} - A^{\nu}_{\mu\kappa}g^{\kappa}_{\lambda}\frac{\partial}{\partial g^{\nu}_{\lambda}}\right)g^{\sigma}_{\rho}X^{\rho}\frac{\partial}{\partial x^{\sigma}} \\
&= Y^{\mu}\left(\frac{\partial Z^{\sigma}}{\partial x^{\mu}} - A^{\nu}_{\mu\kappa}g^{\kappa}_{\lambda}\delta^{\sigma}_{\nu}\delta^{\lambda}_{\rho}X^{\rho}\right)\frac{\partial}{\partial x^{\sigma}} \\
&= Y^{\mu}\left(\frac{\partial Z^{\sigma}}{\partial x^{\mu}} - A^{\sigma}_{\mu\kappa}g^{\kappa}_{\lambda}X^{\lambda}\right)\frac{\partial}{\partial x^{\sigma}} \\
&= Y^{\mu}\left(\frac{\partial Z^{\sigma}}{\partial x^{\mu}} - A^{\sigma}_{\mu\kappa}Z^{\kappa}\right)\frac{\partial}{\partial x^{\sigma}}. \qquad (9)
\end{aligned}$$

Making the identification $A^{\sigma}_{\mu\kappa} \to -\Gamma^{\sigma}_{\mu\kappa}$, the convential covariant derivative on the tangent bundle is obtained [Nash and Sen, 1983].

2.4 Curvature

Since $\frac{d}{dt}\tilde{\gamma}(t) = X^{\mu}D_{\mu}\tilde{\gamma}(t)$, we can Taylor expand $\tilde{\gamma}(t)$ around some point in order to see how it varies in the neighbourhood around that point. Let the curve lying in $\mathcal{M}$ be parameterized by t so that $X^{\mu}\frac{\partial}{\partial x^{\mu}} = \frac{d}{dt}$. Then

$$\tilde{\gamma}(t) = exp(tX^{\mu}D_{\mu})\tilde{\gamma}(0), \qquad (10)$$

defines the parallel transport of $\tilde{\gamma}(0)$ along the curve in $\mathcal{M}$. Consider a closed loop $ABCD$ in $\mathcal{M}$ formed by the intersection of two congruences with tangents $\mathbf{U} = \frac{d}{d\lambda}$ and $\mathbf{V} = \frac{d}{d\alpha}$. Let $\mathbf{U}$ and $\mathbf{V}$ be part of a coordinate basis, so that $[\mathbf{U}, \mathbf{V}] = 0$. Then the value of $\tilde{\gamma}$ after parallel transport along the curve ABC is given by $\tilde{\gamma}_{ABC} = e^{\Delta\lambda U^{\mu}D_{\mu}}e^{\Delta\alpha V^{\nu}D_{\nu}}\tilde{\gamma}(0)$. Similarly, the value of $\tilde{\gamma}$ after parallel transport along the curve ADC is given by $\tilde{\gamma}_{ADC} = e^{\Delta\alpha V^{\mu}D_{\mu}}e^{\Delta\lambda U^{\nu}D_{\nu}}\tilde{\gamma}(0)$. If we take the limit as $\Delta\alpha, \Delta\lambda \to 0$, then to first order in $\Delta\alpha$ and $\Delta\lambda$,

$$\begin{aligned}\tilde{\gamma}_{ABC} - \tilde{\gamma}_{ADC} &\approx (1 + \Delta\lambda U^\mu D_\mu)\,(1 + \Delta\alpha V^\nu D_\nu)\,\tilde{\gamma}(0) \\ &\quad - (1 + \Delta\alpha V^\nu D_\nu)\,(1 + \Delta\lambda U^\mu D_\mu)\,\tilde{\gamma}(0), \\ &= \Delta\lambda\Delta\alpha\,[U^\mu D_\mu, V^\nu D_\nu]\,\tilde{\gamma}(0). \end{aligned} \tag{11}$$

If we use λ, α as our coordinates, and write $U^\mu D_\mu = D_\lambda$ etc, then

$$\tilde{\gamma}_{ABC} - \tilde{\gamma}_{ADC} \approx \Delta\lambda\Delta\alpha\,[D_\lambda, D_\alpha]\,\tilde{\gamma}(0), \tag{12}$$

so that the term $[D_\lambda, D_\alpha]$ measures the failure of $\tilde{\gamma}(0)$ to be parallel transported along the loop, and is therefore a measure of the curvature of the connection [Schutz, 1980, Nash and Sen, 1983].

Denoting the quantity $\lambda_a g \frac{\partial}{\partial g}$ by R_a, we can express the curvature as

$$[D_\mu, D_\nu] = -F^a_{\mu\nu} R_a, \tag{13}$$

where

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + f_{abc} A^a_\mu A^b_\nu \tag{14}$$

where f_{abc} are the structure constants of G [2]. We can define the curvature 2-form $\mathbf{F}$ as

$$\mathbf{F} = F^a_{\mu\nu} \lambda_a \mathbf{d}x^\mu \wedge \mathbf{d}x^\nu. \tag{15}$$

Since $\mathbf{A}\wedge\mathbf{A} = A^a_\mu \lambda_a \mathbf{d}x^\mu \wedge A^b_\nu \lambda_b \mathbf{d}x^\nu = A^a_\mu A^b_\nu [\lambda_a, \lambda_b] \mathbf{d}x^\mu \wedge \mathbf{d}x^\nu = A^a_\mu A^b_\nu f_{abc} \lambda_c \mathbf{d}x^\mu \wedge \mathbf{d}x^\nu$, we can write $\mathbf{F}$ in the compact form

$$\mathbf{F} = \mathbf{dA} + \mathbf{A}\wedge\mathbf{A}, \tag{16}$$

[2] That is, $[\lambda_a, \lambda_b] = f_{abc}\lambda_c$

or in terms of ω,

$$\begin{aligned}\mathbf{F} &= g(\mathbf{d}\boldsymbol{\omega} + \boldsymbol{\omega} \wedge \boldsymbol{\omega})g^{-1}, \\ &= g\boldsymbol{\Omega}g^{-1}(\text{definition}),\end{aligned} \tag{17}$$

(see appendix).

2.5 Noether's Theorem

To see how symmetries in geometry can give rise to conserved quantities, consider the Lagrangian density $\mathcal{L}$ of some scalar field ψ that depends on ψ and $\partial_\mu\psi$, as well as x^μ. The action is given by the integral of $\mathcal{L}$ over the manifold $\mathcal{M}$:

$$S[\psi] = \int d^n x \mathcal{L}. \tag{18}$$

We consider such fields ψ for which $\delta S[\psi] = 0$. Consider an infinitesimal transformation of ψ generated by a functional derivation Q such that $Q[\mathcal{L}] = \partial_\mu f^\mu$. Then, on any submanifold $\mathcal{N}$,

$$\begin{aligned}Q\left[\int_{\mathcal{N}} d^n x \mathcal{L}\right] &= \int d^n x \left(\frac{\partial \mathcal{L}}{\partial \psi}\right) Q[\psi] + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} Q[\partial_\mu \psi], \\ &= \int_{\mathcal{N}} d^n x \left(\frac{\partial \mathcal{L}}{\partial \psi}\right) Q[\psi] + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} \partial_\mu Q[\psi], \\ &= \int_{\mathcal{N}} d^n x \left(\frac{\partial \mathcal{L}}{\partial \psi} - \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)}\right]\right) Q[\psi] + \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} Q[\psi]\right], \\ &= \int_{\mathcal{N}} d^n x \left(\frac{\partial \mathcal{L}}{\partial \psi} - \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)}\right]\right) Q[\psi] + \int_{\partial\mathcal{N}} d\sigma_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} Q[\psi],\end{aligned} \tag{19}$$

where in the last step Stokes' Theorem has been applied. The first integrand vanishes since ψ is a solution to the Euler-Lagrange equations. Therefore,

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} Q[\psi] - f^\mu \right) = 0, \tag{20}$$

and thus the continuity equation for the current density $J^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} Q[\psi] - f^\mu$ is obtained. If the infinitesimal transformations form a Lie group, then we may write $Q[\psi] = iT_a\psi$ where T_a is a generator of the group. If the action is invariant under the group, then to each generator T_a there is a corresponding current,

$$J^\mu_a = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} iT_a\psi. \tag{21}$$

3 Examples of Gauge Theories

3.1 The Klein Gordon Scalar Field

To see how connections arise in field theories, consider the Klein-Gordon equation which describes a scalar field ψ with mass m [Ryder, 1985]:

$$\left(\nabla_\mu \nabla^\mu - m^2\right) \psi = 0. \tag{22}$$

The Lagrangian that gives rise to this equation is

$$\mathcal{L} = (\nabla^\mu \psi)^\dagger (\nabla_\mu \psi) - m^2 \psi^\dagger \psi. \tag{23}$$

Varying $\mathcal{L}$ with respect to $\psi^\dagger$ yields the Klein-Gordon equation, whereas varying $\mathcal{L}$ with respect to ψ yields it's complex conjugate. Any global $U(1)$ unitary transformation of the form $\psi \to e^{i\phi}\psi$ leaves the form of Eq. 22 unchanged, and furthermore preserves the normalization of ψ. Such a transformation is merely a change of internal coordinates on the complex unit circle, and should not affect the measurable properties of the field. However, there is no *a priori* reason to require that if one transformation is applied at one point in space, then the same transformation must be applied at another, and a "rigid" transformation of the form $\psi \to e^{i\phi}\psi$ is certainly not the most general. To that end, we allow ϕ to vary over spacetime and become $\phi(x)$. Upon transformation, the quantity $\nabla_\mu \psi = \mathbf{d}\psi$ becomes

$$\mathbf{d}(\psi e^{i\phi}) = (\mathbf{d}\psi)e^{i\phi} + i\psi(\mathbf{d}\phi)e^{i\phi}. \tag{24}$$

[Kaku, 1993, Ryder, 1985]. In order for Eq. 22 to remain unchanged, we need to replace $\mathbf{d}$ with $D_\mu = \mathbf{d} - iA_\mu$, where, upon the transformation $\psi \to e^{i\phi}\psi$, we require that $A_\mu \to A_\mu + \mathbf{d}\phi$. Note that here D_μ refers to the induced covariant derivative on the fibre bundle rather than the covariant derivative on the principal bundle, because ∇_μ has already been used. Then

$$\begin{aligned}
(\mathbf{d} - iA_\mu)\,\psi \to (\mathbf{d} - iA_\mu - i\mathbf{d}\phi)\,\psi e^{i\phi} \\
&= (\mathbf{d}\psi)e^{i\phi} + i\psi(\mathbf{d}\phi)e^{i\phi} - iA_\mu\psi e^{i\phi} - i(\mathbf{d}\phi)\psi e^{i\phi} \\
&= (\mathbf{d}\psi)e^{i\phi} + i\psi(\mathbf{d}\phi)e^{i\phi} - iA_\mu\psi e^{i\phi} - i(\mathbf{d}\phi)\psi e^{i\phi} \\
&= e^{i\phi}\,(\mathbf{d} - iA_\mu)\,\psi,
\end{aligned} \tag{25}$$

i.e., $D_\mu \to e^{i\phi}D_\mu$, and we obtain a form of the Klein-Gordon equation that is gauge-invariant:

$$\left(D_\mu D^\mu - m^2\right)\psi = 0. \tag{26}$$

The term A_μ is known as a guage connection for the gauge transformation. In analogy with differential geometry on the tangent bundle, we can also speak of the curvature of the gauge connection, i.e., the failure of ψ to be parallel transported along a closed loop:

$$(D_\mu D_\nu - D_\nu D_\mu)\psi = [D_\mu, D_\nu]\psi = F_{\mu\nu}\psi. \tag{27}$$

Since $\mathbf{d}^2 = 0$, the curvature is gauge invariant, and is also a tensor. We began this discussion with the assumption that Eq. 22 held true, so there exists one global transformation in which A_μ vanishes everywhere. Therefore $F_{\mu\nu}$ is zero also, and the Klein-Gordon equation is gauge-flat. However, if we regard Eq. 26 as more fundamental, then $F_{\mu\nu}$ need not vanish. Indeed, Eq. 26 is the general form of the Klein-Gordon equation of a charged particle in an electromagnetic field where the canonical momentum is $\vec{p}_{EM} = \vec{p} + q\vec{A}$. We can therefore regard the electromagnetic potential 4-vector to be the connection on the $U(1)$ bundle, and the Faraday tensor to be its curvature.

The Lagrangian density that gives rise to the free (gauge invariant) Klein-Gordon equation is

$$\mathcal{L}_{local} = (D^\mu\psi)^\dagger(D_\mu\psi) - m^2\psi^\dagger\psi, \tag{28}$$

which can also conveniently be written as

$$\mathcal{L}_{local} = \eta^{\mu\nu}(D_\mu\psi)^\dagger(D_\nu\psi) - m^2\psi^\dagger\psi. \tag{29}$$

Note that $\mathcal{L}_{local}$ has introduced additional terms to $\mathcal{L}$ that involve both A^μ and ψ. If we view ∂_μ as the correct derivative of ψ, then these extra terms may be interpreted as an interaction between ψ and the gauge field, in the same way that a particle in freefall in curved spacetime appears to be under the influence of a force due to the curvature of the $\Gamma^\beta_{\alpha\gamma}$ terms. However, as is the case in GR, it is more fruitful to consider $\mathcal{L}_{local}$ as a free Lagrangian.

To see what quantities are conserved due to the symmetry properties of $\mathcal{L}_{local}$, we can make use of Noether's theorem. The generator of the $U(1)$ transformation $\psi \to e^{iq\phi}\psi$ is simply q, the electric charge operator, and $Q[\psi] = iq\psi$ while $Q[\psi^\dagger] = -iq\psi^\dagger$.

Therefore, the invariance of $\mathcal{L}_{local}$ under $U(1)$ yields the following current, from Eqn ?:

$$\begin{aligned} J^\mu &= \frac{\partial \mathcal{L}_{local}}{\partial(D_\mu\psi)}(iq)\psi + \frac{\partial \mathcal{L}_{local}}{\partial(D_\mu\psi)^\dagger}(-iq)\psi^\dagger \\ &= iq\left[\psi(D^\mu\psi)^\dagger - \psi^\dagger(D^\mu\psi)\right], \end{aligned} \tag{30}$$

from which we can immediately see that a real scalar field carries no charge. In order to make the connection between the gauge and electrodynamics complete, we need to consider gauge field as a physical object in itself, in the same way that the curvature tensor arises in the transition from Special to General Relativity. While we require that the contribution of the gauge field to the Lagrangian be both gauge and coordinate invariant, this is not sufficient to single out a candidate. On the other hand, if we require that the corresponding quantum theory be renormalizable, and that the Lagrangian be invariant under parity, then the only possibility is of the form $F^{\mu\nu}F_{\mu\nu}$. Note that the gauge field has to be massless, since a term such as $m_g A^\mu A_\mu$ is not gauge invariant[3]. Guided by the electrodynamics Lagrangian *in vacuo*, we see that the complete Lagrangian, hereafter just called $\mathcal{L}$, is given by

$$\mathcal{L} = (D^\mu\psi)^\dagger(D_\mu\psi) - m^2\psi^\dagger\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}, \tag{31}$$

or equivalently,

$$\begin{aligned} \mathcal{L} &= (\partial^\mu\psi^\dagger)(\partial_\mu\psi) - iqA_\mu\psi(D^\mu\psi)^\dagger + iqA^\mu\psi^\dagger(D_\mu\psi) - m^2\psi^\dagger\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}, \\ &= (\partial^\mu\psi^\dagger)(\partial_\mu\psi) - A_\mu J^\mu - m^2\psi^\dagger\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \end{aligned} \tag{32}$$

[3]However, the field can become massive by means of, e.g., the Higgs process.

With the Lagrangian in this form we can vary with respect to A_ν:

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} \right) - \frac{\partial \mathcal{L}}{\partial A_\nu} = 0,$$
$$\partial_\mu F^{\mu\nu} - J^\nu = 0. \tag{33}$$

This result yields the two inhomogeneous Maxwell equations. The other two equations result from the fact that $\mathbf{d}F_{\mu\nu} = 0$.

3.2 $SU(2)$ Isospin

We can consider more complicated fields where the associated symmetry groups are non-Abelian, i.e., non-commutative. Such gauge theories are known as Yang-Mills theories. The earliest such theory made use of the approximate symmetry of the proton and neutron in the strong interaction under the exchange of particle species, i.e., $p \leftrightarrow n$. Such a relabling amounts to an $SU(2)$ symmetry group, and the theory is known as isospin, in analogy with the $SU(2)$ electronic spin. Since a particle can be found in one of two states $|p\rangle$ or $|n\rangle$, and may exist in a superposition of these two states, we may assign to the field, in addition to its amplitude, a position in this internal isospin space at each point in space.

In order for the theory to be locally invariant under a transformation of the form $\psi \to \exp(i\phi^a \sigma_a)\psi$, where ϕ^a is a set of parameters describing the transformation, and σ_a are the Pauli matrices, the covariant derivative $D_\mu = \partial_\mu - igA^a_\mu \sigma_a$ must replace the ordinary derivative of the field in the Lagrangian. Here g denotes the coupling strength of the three gauge potentials A^a_μ. Since the $SU(2)$ generators obey $[\sigma_i, \sigma_j] = \varepsilon_{ijk}\sigma_k$, we can write the 3 gauge fields as a single field strength tensor $F^a_{\mu\nu}$, given by

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g\varepsilon_{abc} A^b_\mu A^c_\nu \tag{34}$$

3.3 SU(3) Chromodynamics

Unlike isospin, which is an approximate symmetry and is broken by the electromagnetic interaction, the strong interaction is exactly symmetric with regards to interchange of colour. The quarks may exist in a superposition of the states $|r\rangle$, $|g\rangle$, and $|b\rangle$, and the Lagrangian is invariant under a relabling of the form $r \to b$, $b \to g$, $g \to r$. The quark field therefore possesses an $SU(3)$ colour symmetry. There is also an approximate $SU(3)$ flavour symmetry, but this symmetry is broken by the unequal masses of the u, d, and s quarks. The group $SU(3)$ has 8 generators, which can be represented by the Gell-Mann matrices λ_a. Corresponding to these generators are 8 gauge potentials, which, when quantized, become a variety of 8 gauge bosons known as the gluons [Ryder, 1985].

3.4 General Relativity

Although the concept of a connection arises naturally when considering the parallel transport of a vector in the tangent bundle, the treatment becomes more interesting when we consider non-coordinate bases. Specifically, given a metric $g_{\mu\nu}$ and coordinate basis vectors $\mathbf{e}_\mu = \partial_\mu$, we can construct an orthonormal basis $\mathbf{e}_{\hat{a}}$ at every point such that $\mathbf{e}_{\hat{a}} \cdot \mathbf{e}_{\hat{b}} = \eta_{ab}$. Similarly, we can define a dual basis $\mathbf{e}^{\hat{a}}$ such that $\mathbf{e}^{\hat{a}}(\mathbf{e}_{\hat{b}}) = \delta^a_b$. The set of orthonormal vectors $\mathbf{e}_{\hat{a}}$ is known as a *tetrad* or *vierbein* [Carroll, 1997]. Following the convention of Carrol, we use greek letters for coordinate basis indices, and latin letters for tetrad indices. We can then express our coordinate basis in terms of the tetrad as follows:

$$\mathbf{e}_\mu = e^a_\mu \mathbf{e}_{\hat{a}}, \tag{35}$$

where e^a_μ is an $n \times n$ invertible matrix. Since the basis forms a set of vector fields, these fields, together with the base manifold $\mathcal{M}$, form a *frame bundle* over $\mathcal{M}$. Now, in addition to general coordinate transformations we can also allow local changes of basis. If we require that our basis remains orthonormal, then the appropriate set of allowed transformations is the Lorentz group $O(1,3)$, and we may write a change of basis as

$$\mathbf{e}_{\hat{a}'} = \Lambda^{b}_{a'} \mathbf{e}_{\hat{b}}, \tag{36}$$

where $\Lambda^{b}_{a'}$ corresponds to an inverse Lorentz transformation. We can then define the covariant derivative of a vector in the tetrad basis, in analogy with Eqn. 9, as

$$\nabla_\mu X^a = \partial_\mu X^a + \omega_{\mu}{}^{a}_{c} X^c, \tag{37}$$

where $\omega_{\mu}{}^{a}_{c}$ is known as a *spin connection*, the reason for which will be explained below. We can write $\omega_{\mu}{}^{a}_{c}$ as a Lie-algebra valued oneform $\omega_\mu{}^{ij}\ell_{ij}$ where ℓ_{ij} are the generators of the Lorentz group.

It is natural to attempt to formulate a general Yang-Mills theory using the curved spacetime General Relativity in place of flat Minkowski space as the base manifold. However, while the full Lorentz group naturally admits a representation in terms of spinors, there are no finite-dimensional spinorial representations of $Gl(4,\mathbb{R})$, the group that expresses the general covariance of General Relativity. Nevertheless, the tangent space at any point in spacetime is flat and Lorentzian (the key concept behind the vierbein formalism) and the Lorentz group *does* allow finite dimensional spinorial representations.

As a first step towards finding a generally covariant form of the Dirac equation, we elevate the Dirac matrices from their vierbein components into spacetime components, thus:

$$\gamma^\mu(x^\nu) = e^{\mu}_{a}(x^\nu)\gamma^a. \tag{38}$$

[Kaku, 1993]. Requiring that a spinor transform as a scalar under coordinate transformations (ones which do not alter the vierbein), and as a spinor under Lorentz transformations, we obtain the action of the covariant derivative on a general spinor,

$$\nabla_\mu \Psi = \left(\partial_\mu - \frac{i}{4} \omega_\mu^{ab} \sigma_{ab} \right) \Psi, \tag{39}$$

which results in the generally covariant Dirac equation:

$$\left(i \gamma^\mu \nabla_\mu - m \right) \Psi = 0. \tag{40}$$

The Riemann curvature tensor can then be defined in terms of the curvature of the spin connection:

$$\left[\nabla_\mu, \nabla_\nu \right] \Psi = -\frac{i}{4} R_{\mu\nu}^{ab} \sigma_{ab} \Psi, \tag{41}$$

where

$$R_{\mu\nu}^{ab} = \partial_\mu \omega_\nu^{ab} - \partial_\nu \omega_\mu^{ab} + \omega_\mu^{ac} \omega_{\nu c}^{\ \ b} - \omega_\nu^{ac} \omega_{\mu c}^{\ \ b}. \tag{42}$$

If we drop spatial indices we obtain the following compact form:

$$\mathbf{R}_b^a = \mathbf{d}\boldsymbol{\omega}_b^a + \boldsymbol{\omega}_c^a \wedge \boldsymbol{\omega}_b^c. \tag{43}$$

To make the connection compatible with the metric, we can further require that $\nabla_\mu e_\nu^a = 0$.

Although Equation 43 casts General Relativity squarely in terms of the principal bundle formalism, it should be noted that the dynamics of spacetime differ from other Yang-Mills theories in the sense that neither the connection nor the curvature itself appear as a source for curvature in the Einstein equations. However tempting it may be to consider spacetime as being background for the physics, the Einstein equations are nevertheless non-linear so that the connection can interact with itself.

4 Conclusion

Symmetry has been a powerful guide throughout modern physics, and by combining the natural symmetries of a physical system with the notion of general covariance, we are naturally led to consider a larger group of symmetries, and with it, a new set of conservation laws. Under this general framework, many seemingly disparate systems have been studied in a unified manner, giving new insight into the physics of these systems.

5 Appendix

5.1 Proof that $\mathbf{dA} + \mathbf{A} \wedge \mathbf{A} = g(\mathbf{d}\boldsymbol{\omega} + \boldsymbol{\omega} \wedge \boldsymbol{\omega})g^{-1}$

$$\begin{aligned}\mathbf{d}\boldsymbol{\omega} &= \mathbf{d}(g^{-1}) \wedge \mathbf{d}g + \mathbf{d}(g^{-1}) \wedge \mathbf{A}g + g^{-1} \wedge (\mathbf{dA} \wedge g - \mathbf{A} \wedge \mathbf{d}g), \\ &= -g^{-1}(\mathbf{d}g)g^{-1} \wedge \mathbf{d}g - g^{-1}(\mathbf{d}g)g^{-1} \wedge \mathbf{A}g + g^{-1} \wedge (\mathbf{dA} \wedge g - \mathbf{A} \wedge \mathbf{d}g),\end{aligned} \tag{44}$$

while

$$\begin{aligned}\boldsymbol{\omega} \wedge \boldsymbol{\omega} &= (g^{-1}\mathbf{d}g + g^{-1}\mathbf{A}g) \wedge (g^{-1}\mathbf{d}g + g^{-1}\mathbf{A}g), \\ &= g^{-1}\mathbf{d}g \wedge g^{-1}\mathbf{d}g + g^{-1}\mathbf{d}g \wedge g^{-1}\mathbf{A}g + g^{-1}\mathbf{A}g \wedge g^{-1}\mathbf{d}g + g^{-1}\mathbf{A}g \wedge g^{-1}\mathbf{A}g,\end{aligned} \tag{45}$$

where upon adding we get,

$$\mathbf{d}\boldsymbol{\omega} + \boldsymbol{\omega} \wedge \boldsymbol{\omega} = g^{-1} \wedge \mathbf{dA} \wedge g + g^{-1}\mathbf{A}g \wedge g^{-1}\mathbf{A}g,$$
$$= g^{-1}\mathbf{dA}g + g^{-1}\mathbf{A} \wedge \mathbf{A}g, \tag{46}$$

which proves the result.

References

L. Ryder. *Quantum Field Theory.* Cambridge University Press, Camridge, 1985.

M. Nakahara. *Geometry, Topology and Physics, Second Edition.* Institute of Physics Publishing, London, 2003.

C. Nash and S. Sen. *Topology and Geometry for Physicists.* Academic Press, London, 1983.

H. Yamashita. A (amazingly) brief introduction to principal bundles, connections, and gauge theory. Retrieved from http://www.sun-inet.or.jp/~leibniz/math/geometry/pbundle.pdf&e=10384, Aichi-Gakuin University Japan, 2002.

B. Schutz. *Geometrical Methods of Mathematical Physics.* Cambridge University Press, Cambridge, 1980.

M. Kaku. *Quantum Field Theory - A modern Introduction.* Oxford University Press, Oxford, 1993.

Sean. M. Carroll. Lecture notes on general relativity. Retrieved from http://pancake.uchicago.edu/~carroll/notes/ University of Chicago, 1997.