Spring 4-27-2017

Maxwell's Equations, Gauge Fields, and Yang-Mills Theory

Nicholas Alexander Gabriel

MAXWELL'S EQUATIONS, GAUGE FIELDS, AND YANG-MILLS THEORY

An honors paper submitted to the Department of Mathematics
of the University of Mary Washington
in partial fulfillment of the requirements for Departmental Honors

Nicholas Alexander Gabriel

April 2017

By signing your name below, you affirm that this work is the complete and final version of your paper submitted in partial fulfillment of a degree from the University of Mary Washington. You affirm the University of Mary Washington honor pledge: "I hereby declare upon my word of honor that I have neither given nor received unauthorized help on this work."

Nicholas A. Gabriel
(digital signature)

04/27/17

Maxwell's Equations, Gauge Fields, and Yang-Mills Theory

Nicholas A. Gabriel

submitted in partial fulfillment of the requirements for Honors in Mathematics at the University of Mary Washington

Fredericksburg, Virginia

April 2017

This thesis by **Nicholas A. Gabriel** is accepted in its present form as satisfying the thesis requirement for Honors in Mathematics.

DATE	APPROVED
____________	________________________________ Yuan-Jen Chiang, Ph.D. thesis advisor
____________	________________________________ Randall Helmstutler, Ph.D. committee member
____________	________________________________ Suzanne Sumner, Ph.D. committee member

Contents

Abstract

Starting from Maxwell's theory of electromagnetism in a Minkowski spacetime, we generalize to arbitrary spacetimes and gauge groups. The gauge groups $U(1)$ and $SU(3)$ and their associated Yang-Mills theories are discussed in detail.

1 Introduction

Arguably the most fundamental pursuit of theoretical physics is that of unification of the laws of nature. The development of special relativity by Einstein in the early 20th century and the later development of Yang-Mills theory both have origins in Maxwell's theory of electromagnetism. Naturally then, we begin by developing some of the formalism of special relativity and Maxwell's theory of electromagnetism. Then, we develop some tools from differential geometry to generalize the vector calculus used in the description of Maxwell's theory. This development will in turn allow us to understand what is considered the cornerstone of modern theoretical physics, Yang-Mills theory, which more or less gives us a prescription for developing theories of the behavior of matter all around us (besides behavior that is due to gravity). Particularly we will discuss the theories of Quantum Electrodynamics (QED), which describes how electrically charged particles and photons interact, and Quantum Chromodynamics, which describes how the nuclei of atoms are formed and behave. As we will see, these particular Yang-Mills theories are $U(1)$ and $SU(3)$ gauge invariant, respectively, and in some sense the only difference between them are described by the properties of these groups!

We will find many pieces of notation useful, but not by any means universal, so we shall clarify maybe the most ubiquitous of these now, since the following we will begin using it immediately. We will work in what some call "god-given" units, where the speed of light, $c \approx 3 \times 10^8$ m/s and the reduced Planck's constant, $\hbar \approx 4 \times 10^{-15}$ eV·s are set to unity. This will make Maxwell's equations, Lorentz transforms, and the wave equations of quantum mechanics much less cluttered. For example, Einstein's famous mass-energy equivalence, $E = mc^2$ now reduces to

$$E = m. \tag{1.1}$$

In practice this notation makes calculations less tedious, and in order to get back the desired units of some quantity, one simply multiplies by the correct (unique) factor of c's and $\hbar$'s. We will not really discuss the nuances of units after this point, but it is worth mentioning since some readers may be unfamiliar with this somewhat confusing (albeit convenient) practice. When it makes sense, we will keep these units around to make concepts more transparent.

2 Maxwell's Equations and Special Relativity

2.1 Special Relativity

Special relativity is concerned with how measures of space and time differ from one inertial reference frame to another (that is, two frames moving at constant velocities). The fundamental quantities in relativity are four component vectors (or 4-vectors), like space-time, denoted using a Greek index that runs from 0 to 3, (e.g., $\mu = 0,1,2,3$). The space-time vector is defined as $x^\mu = (t,x,y,z)^T = (t,\vec{r})^T$, in terms of indices,

$$x^0 = t, \quad x^1 = x, \quad x^2 = y, \quad x^3 = z. \tag{2.1}$$

Specific points in spacetime $x^\mu \in \mathbb{R}^4$, are called *events*. Particles follow continuous trajectories of events called *world lines*. We will be concerned with events and world lines as they are seen by two different inertial reference frames, O and O'. As we will see shortly, Minkowski spacetime is a manifold, specifically, a manifold with a Lorentz metric

$$\eta_{\mu\nu} = \eta^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{2.2}$$

Using this metric we can lower the index of our space-time vector in the following way

$$x_\mu = \sum_{\mu=0}^{3} \eta_{\mu\nu} x^\nu = (t, -\vec{r}). \tag{2.3}$$

Following Einstein's summation convention, we will suppress the summation notation whenever we see a repeated index in an expression, so the above becomes

$$x_\mu = \eta_{\mu\nu} x^\nu = (t, -\vec{r}). \tag{2.4}$$

2.1.1 Lorentz Transformations

We distinguish the vector x^μ (index up) vector as being *contravariant* while x_μ (index down) as being *covariant*. We can take the *Minkowski inner product* of a contravariant quantity and its covariant counterpart to obtain an *invariant* scalar quantity,

$$\begin{aligned} x_\mu x^\mu &= t^2 - \vec{r}\cdot\vec{r} \\ &= t^2 - x^2 - y^2 - z^2. \end{aligned} \tag{2.5}$$

This quantity is called a *space-time interval* and is said to be invariant because it remains constant under a Lorentz transformation, Λ, (which we will define shortly) from one frame to another

$$x^{\mu'} = \Lambda x^\mu \tag{2.6}$$

just as length $r^2 = x^2 + y^2 + z^2$ is invariant under rotations[1]. We can classify space-time intervals according to their sign, i.e.

$$\begin{aligned} x_\mu x^\mu &> 0, \qquad x^\mu \text{ is timelike} \\ x_\mu x^\mu &< 0, \qquad x^\mu \text{ is spacelike} \\ x_\mu x^\mu &= 0, \qquad x^\mu \text{ is lightlike.} \end{aligned} \tag{2.7}$$

Definition 2.1. Let $x^\mu \in \mathbb{R}^4$. A *Lorentz transformation* is a linear transformation that maps a space-time interval in some frame O to a space-time interval in another frame O' by a transformation $\Lambda\colon \mathbb{R}^4 \to \mathbb{R}^4$ that leaves a space-time interval constant, so that

$$\begin{gathered} x_\mu x^\mu = \Lambda x_\mu \Lambda x^\mu = x_{\mu'} x^{\mu'} \\ (t)^2 - (x)^2 - (y)^2 - (z)^2 = (t')^2 - (x')^2 - (y')^2 - (z')^2 \end{gathered} \tag{2.8}$$

[1] We will better define the idea of rotations in $\mathbb{R}^3$ in the discussion of the rotation group $SO(3)$ in sections 4.1 and 4.2.

where the frame O' moves with velocity $v \in [0, c]$ relative to O. We can see then that a rotation in the xy plane by an angle $-\theta$

$$R = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.9}$$

is a Lorentz transform since

$$x^{\mu'} = Rx^{\mu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} \tag{2.10}$$

implies that

$$\begin{aligned} t' &= t \\ x' &= x\cos\theta - y\sin\theta \\ y' &= x\sin\theta + y\cos\theta \\ z' &= z. \end{aligned} \tag{2.11}$$

Moreover, we have

$$\begin{aligned} x_{\mu'}x^{\mu'} &= (t')^2 - (x')^2 - (y')^2 - (z')^2 \\ &= (t)^2 - (x\ \cos\theta - y\ \sin\theta)^2 - (y\ \cos\theta + x\ \sin\theta)^2 - (z)^2 \\ &= t^2 - x^2\cos^2\theta - y^2\sin^2\theta + 2xy\sin\theta\cos\theta \\ &\quad - x^2\sin^2\theta - y^2\cos^2\theta - 2xy\sin\theta\cos\theta - (z)^2 \\ &= (t)^2 - (x^2 + y^2)(\cos^2\theta + \sin^2\theta) - (z)^2 \\ &= (t)^2 - (x)^2 - (y)^2 - (z)^2. \end{aligned} \tag{2.12}$$

This transformation corresponds to the frame O' having a coordinate system that differs from O by an angle of $-\theta$ in the xy plane and zero relative velocity. Typically we are not worried about two frames that have different coordinate systems, but two frames with the same coordinate system and a non-zero relative velocity between one another. These transformations are called *boosts*, and we will outline a derivation here.

Suppose that we have a stationary frame O and a frame O' with a velocity v in the x direction and that their world lines intersect at $x^{\mu} = 0^{\mu}$. Suppose further that a spherical wavefront of light is emitted at $x^{\mu} = 0^{\mu}$. By Einstein's second postulate, this wavefront has a constant velocity c in all reference frames, and so from the perspective of O and O' the wavefront is at a distance

$$r = ct \tag{2.13}$$

$$r' = ct' \tag{2.14}$$

in each frame, respectively. The equations for the spherical wavefront in O is

$$r^2 = x^2 + y^2 + z^2 \tag{2.15}$$

and in O',

$$r'^2 = x'^2 + y'^2 + z'^2. \tag{2.16}$$

Now

$$\begin{aligned} y' &= y, \\ z' &= z \end{aligned} \tag{2.17}$$

since relative motion is in the x direction. We then have

$$x'^2 - (ct')^2 = x^2 - (ct)^2. \tag{2.18}$$

The most general, linear[2] solution for x' and ct' in terms of x and ct is

$$\begin{aligned} x' &= x\cosh\phi - ct\sinh\phi \\ ct' &= -x\sinh\phi + ct\cosh\phi \end{aligned} \tag{2.19}$$

which can be verified using the identity

$$\cosh^2\phi - \sinh^2\phi \equiv 1 \tag{2.20}$$

and (2.18). To obtain the relation of ϕ to v, we can use the fact that in the frame of O, $x = 0$ and $x' = vt'$, so

$$x' = -ct\sinh\phi, \tag{2.21}$$

$$ct' = ct\cosh\phi. \tag{2.22}$$

Dividing (2.21) by (2.22) we have

$$\frac{x'}{ct'} = \frac{-\sinh\phi}{\cosh\phi} = -\tanh\phi = -\frac{v}{c} \tag{2.23}$$

We can then make use of two more hyperbolic identities

$$\sinh\phi \equiv \frac{\tanh\phi}{\sqrt{1-\tanh^2\phi}}, \quad \cosh\phi \equiv \frac{1}{\sqrt{1-\tanh^2\phi}} \tag{2.24}$$

to obtain

$$\sinh\phi = \frac{v/c}{\sqrt{1-(v/c)^2}}, \quad \cosh\phi = \frac{1}{\sqrt{1-(v/c)^2}}. \tag{2.25}$$

Often the ratio v/c is denoted β and $(1-(v/c)^2)^{-1/2}$ as γ so that

$$\sinh\phi = \beta\gamma, \quad \cosh\phi = \gamma. \tag{2.26}$$

But arguably the hyperbolic parameterization is more transparent. We can why by writing the boost as a matrix

$$B = \begin{pmatrix} \cosh\phi & \sinh\phi & 0 & 0 \\ \sinh\phi & \cosh\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.27}$$

from which we can immediately read off that $\det B = 1$ by (2.20). Our regular Euclidean rotation above also has a determinant equal to unity, and in fact all Lorentz transformations have

[2] We require linearity, otherwise we do not have translational symmetry between frames. If we did not require translational symmetry, we would basically be saying that physical laws depend on where we define the origin of our coordinate system to be.

this property, since each transformation Λ can all be expressed as a product of rotations and boosts.

We could have just as well chosen the metric tensor

$$\eta^{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.28}$$

because in multiplying both sides of equation (2.8) by a negative sign we still have equality. For this metric tensor we have the spacetime interval

$$x_\mu x^\mu = -(t)^2 + x^2 + y^2 + z^2. \tag{2.29}$$

We say (2.2) has a metric signature $(1, -1, -1, -1)$ and (2.28) has a metric signature $(-1, 1, 1, 1)$. It is sometimes justified to regard time as existing in $\mathbb{C}$, with our"usual" time existing on the real axis. We can then perform a so-called Wick rotation of $-\pi/2$:

$$t \to -i\tau \ , \tag{2.30}$$

in which case the interval (2.29) becomes

$$x_\mu x^\mu = \tau^2 + x^2 + y^2 + z^2. \tag{2.31}$$

which takes our Minkowski metric to a Euclidean metric in the variables (τ, x, y, z).

2.2 Maxwell's Equations in $\mathbb{R}^4$

Maxwell's equations are the following:

$$\nabla \cdot \vec{B} = 0, \quad \nabla \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = \vec{0} \tag{2.32}$$

$$\nabla \cdot \vec{E} = \rho, \quad \nabla \times \vec{B} - \frac{\partial \vec{E}}{\partial t} = \vec{\jmath} \tag{2.33}$$

where $\vec{\jmath}, \vec{E}, \vec{B} \colon \mathbb{R}^4 \to \mathbb{R}^3$ and $\rho \colon \mathbb{R}^4 \to \mathbb{R}$. The vectors $\vec{E}$ and $\vec{B}$ are the *electric field* and *magnetic field*, respectively, ρ is the *electric charge density*, and $\vec{\jmath}$ the *current density*. We have grouped the homogeneous and inhomogeneous equations together for reasons that will be come apparent shortly.

When ρ and $\vec{\jmath}$ are zero everywhere, there is an apparent symmetry in the equations:

$$\nabla \cdot \vec{B} = 0, \quad \nabla \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = \vec{0} \tag{2.34}$$

$$\nabla \cdot \vec{E} = 0, \quad \nabla \times \vec{B} - \frac{\partial \vec{E}}{\partial t} = \vec{0} \tag{2.35}$$

characterized by the transformation

$$\vec{B} \mapsto \vec{E}, \quad \vec{E} \mapsto -\vec{B} \tag{2.36}$$

which takes the first pair of equations into the second and vice versa. By defining a complex valued vector field in $\mathbb{C}^3$

$$\vec{\mathcal{E}} = \vec{E} + i\vec{B} \tag{2.37}$$

duality now amounts to the transformation

$$\vec{\mathcal{E}} \mapsto i\vec{\mathcal{E}} \tag{2.38}$$

and our four equations reduce to two:

$$\nabla \cdot \vec{\mathcal{E}} = 0, \quad \nabla \times \vec{\mathcal{E}} = i\frac{\partial \mathcal{E}}{\partial t} \tag{2.39}$$

This trick[3] can be extended to the inhomogeneous case by defining an electromagnetic charge and current density,

$$\mathcal{P} = \rho + i\mu, \tag{2.40}$$

$$\vec{\mathcal{J}} = \vec{\jmath} + i\vec{k} \tag{2.41}$$

where ρ and $\vec{\jmath}$ are the electric charge density, and μ and $\vec{k}$ are analogous quantities caused by the presence and motion of *magnetic charges*. The resulting Maxwell's equations are then

$$\nabla \cdot \vec{\mathcal{E}} = \mathcal{P}, \quad \nabla \times \vec{\mathcal{E}} = i(\frac{\partial \mathcal{E}}{\partial t} + \vec{\mathcal{J}}) \tag{2.42}$$

or, separating into real and imaginary parts to compare with our original four equations,

$$\nabla \cdot \vec{B} = \mu, \quad \nabla \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = -\vec{k} \tag{2.43}$$

$$\nabla \cdot \vec{E} = \rho, \quad \nabla \times \vec{B} - \frac{\partial \vec{E}}{\partial t} = \vec{\jmath}. \tag{2.44}$$

We see that we now have 4 inhomogeneous equations, and the transformations

$$\mathcal{P} \mapsto i\mathcal{P}, \quad \vec{\mathcal{J}} \mapsto i\vec{\mathcal{J}} \tag{2.45}$$

together with (2.38) defines the duality in the the presence of electric and magnetic charges. These magnetic charges, so called magnetic monopoles, have many satisfactory properties in addition to allowing a duality between the inhomogeneous Maxwell's equations. Unfortunately, these magnetic monopoles have never been observed!

2.2.1 Potentials

The form of Maxwell's equations can be used to develop a mathematical tool, called the *potential*, that will often allow us to solve Maxwell's equations with much more ease, and also introduce the idea of a gauge transformation, which will be the topic of section 4.3.1. To start, we will begin with the so-called static case of Maxwell's equations, where there is no time dependence of our fields and sources (an example is charge moving around a fixed circle in space):

$$\begin{aligned} \nabla \cdot \vec{B} &= 0, & \nabla \times \vec{E} &= 0, \\ \nabla \cdot \vec{E} &= \rho, & \nabla \times \vec{B} &= \vec{\jmath}. \end{aligned} \tag{2.46}$$

[3] This trick may not seem so clever once one realizes it essentially amounts to using the fact that $\mathbb{R}^2 \cong \mathbb{C}$ (or more generally that $\mathbb{R}^{2k} \cong \mathbb{C}^k$) and that linear transformations can be represented by multiplication in $\mathbb{C}$. However, as is often the case, moving from a real vector space to a complex one often makes solving problems much easier.

Let ϕ be a C^2 scalar function $\phi : \mathbb{R}^3 \to \mathbb{R}$. We then have the following identity from vector calculus:

$$\nabla \times (\nabla \phi) \equiv 0. \tag{2.47}$$

It then follows from our second homogeneous equation, $\nabla \times \vec{E} = 0$, that the curl of $\vec{E}$ can be expressed as the gradient of some scalar function $V \colon \mathbb{R}^3 \to \mathbb{R}$ called the electric potential. By the charge convention we pick up a negative sign and we have

$$\vec{E} = -\nabla V. \tag{2.48}$$

Let Ψ be a smooth vector function $\Psi \colon \mathbb{R}^3 \to \mathbb{R}^3$. Similar to our last definition, we can use an identity

$$\nabla \cdot (\nabla \times \Psi) \equiv 0. \tag{2.49}$$

together with our first homogeneous equation $\nabla \cdot \vec{B} = 0$ to conclude that there exists a vector function $\vec{A} : \mathbb{R}^3 \to \mathbb{R}^3$ such that

$$\vec{B} = \nabla \times \vec{A}. \tag{2.50}$$

We should note that these potentials are not uniquely defined, we need to specify the value of V at the boundary $\partial \mathcal{V}$ of some volume $\mathcal{V}$ to obtain a unique solution within the volume. Returning to the more general dynamical Maxwell's equations

$$\begin{aligned} \nabla \cdot \vec{B} &= 0, &\qquad \nabla \times \vec{E} + \frac{\partial \vec{B}}{\partial t} &= 0, \\ \nabla \cdot \vec{E} &= \rho, &\qquad \nabla \times \vec{B} - \frac{\partial \vec{E}}{\partial t} &= \vec{\jmath} \end{aligned} \tag{2.51}$$

we can see now that our result (2.50) still holds, since the divergence of $\vec{B}$ is still zero, but (2.48) no longer holds, since the curl of $\vec{E}$ is non-zero. We can still find $\vec{E}$ in terms of a scalar potential V by substituting (2.50) into the second homogeneous equation to obtain

$$\begin{aligned} \nabla \times \vec{E} + \frac{\partial}{\partial t}(\nabla \times \vec{A}) &= \\ \nabla \times \left(\vec{E} + \frac{\partial \vec{A}}{\partial t} \right) &= 0 \end{aligned} \tag{2.52}$$

and now we may apply the identity (2.47) to express $\vec{E} + \frac{\partial \vec{A}}{\partial t}$ as the gradient of a scalar potential V:

$$\vec{E} + \frac{\partial \vec{A}}{\partial t} = -\nabla V. \tag{2.53}$$

Now we have $\vec{E}$ in terms of our potentials:

$$\vec{E} = -\nabla V - \frac{\partial \vec{A}}{\partial t}. \tag{2.54}$$

We can tidy up our notation a bit by rewriting the fields, potentials, and sources in relativistic notation. First the electric charge and current densities will constitute a four-vector:

$$J^\mu = (\rho, \vec{\jmath}) \tag{2.55}$$

Similarly, the four-potential is written

$$A^\mu = (V, \vec{A}), \tag{2.56}$$

and the four-derivative

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \Big(\frac{\partial}{\partial t}, \nabla\Big). \tag{2.57}$$

The four-potential can undergo certain transformations of the form

$$A^\mu \to \tilde{A}^\mu = A^\mu + \partial^\mu \lambda \tag{2.58}$$

without changing the underlying physical fields $\vec{E}$ and $\vec{B}$. This scalar function $\lambda : \mathbb{R}^4 \to \mathbb{R}$ is called a gauge function and is a function of space and time. This transformation is our first example of a *gauge transformation*. In the next section we will show that λ can be *any* function of space and time.

2.2.2 The Field Strength Tensor

The *field strength tensor* is an antisymmetric rank (2,0) tensor which allows us to compactly express our fields:

$$\begin{aligned}
F^{\mu\nu} &= \partial^\mu A^\nu - \partial^\nu A^\mu \\
&= \begin{pmatrix} F^{00} & F^{01} & F^{02} & F^{03} \\ -F^{01} & F^{11} & F^{12} & F^{13} \\ -F^{02} & -F^{12} & F^{22} & F^{23} \\ -F^{03} & -F^{13} & -F^{23} & F^{33} \end{pmatrix} \\
&= \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}
\end{aligned} \tag{2.59}$$

where we have used equations (2.50) and (2.54) to determine the coefficients of the tensor. Now when we make the transformation $A^\mu \to \tilde{A}^\mu = A^\mu + \partial^\mu \lambda$ we have

$$\begin{aligned}
F^{\mu\nu} &\to \partial^\mu \tilde{A}^\nu - \partial^\nu \tilde{A}^\mu \\
&= \partial^\mu (A^\nu + \partial^\nu \lambda) - \partial^\nu (A^\mu + \partial^\mu \lambda) \\
&= \partial^\mu A^\nu - \partial^\nu A^\mu + \partial^\mu \partial^\nu \lambda - \partial^\mu \partial^\nu \lambda
\end{aligned} \tag{2.60}$$

but we can interchange the order of partial derivatives, so $\partial^\mu \partial^\nu = \partial^\nu \partial^\mu$, and hence, the field strength tensor $F^{\mu\nu}$ is invariant under the gauge transformation $A^\mu \to \tilde{A}^\mu$.

In order to write Maxwell's equations in terms of the field strength tensor, we also need the *dual field strength tensor*, defined as

$$G^{\mu\nu} = \frac{1}{2} \epsilon^{\mu\nu\alpha\beta} F_{\alpha\beta} \quad , \tag{2.61}$$

where $\epsilon^{\mu\nu\alpha\beta}$ is the (rank-4) *Levi-Civita symbol*:

$$\epsilon^{\mu\nu\alpha\beta} = \begin{cases} +1, & \mu\nu\alpha\beta \text{ is an even permutation of } 0123 \\ -1, & \mu\nu\alpha\beta \text{ is an odd permutation of } 0123 \\ 0, & \text{any index is repeated.} \end{cases} \tag{2.62}$$

An odd permutation is obtained by interchanging any two adjacent indices an odd number of times, e.g. interchanging 0 and 1 in 0123 yield 1023. An even permutation then comes from interchanging any two adjacent indices an even number of times, e.g. interchanging 0 and 1 in 0123 and then 2 and 3 yield 1032. Some examples of repeated indices are 0023, 0111, 3333, etc. Hence there are $4^4 = 128$ total elements, $4! = 24$ non-zero elements, and $128 - 24 = 104$ zero elements. There is also a rank-3 symbol, which involves permutations of 123, and rank-n symbols involving permutations of $123\ldots n$ all sharing the same basic rules as (2.62).

Once the smoke clears, equation (2.61) becomes

$$G^{\mu\nu} = \begin{pmatrix} 0 & -E_z & E_y & -B_z \\ E_z & 0 & -E_x & -B_y \\ -E_y & E_x & 0 & -B_x \\ B_z & B_y & B_x & 0 \end{pmatrix}. \tag{2.63}$$

Now Maxwell's equations can be written as

$$\begin{aligned} \partial_\mu G^{\mu\nu} &= 0^\mu, \\ \partial_\mu F^{\mu\nu} &= J^\nu. \end{aligned} \tag{2.64}$$

The benefit of this form, besides being more aesthetically pleasing than (2.32) and (2.33), is that we can potentially tell which symmetries are obeyed by Maxwell's equations upon inspection, e.g. how the components transform under a Lorentz transformation.

3 Rewriting Maxwell's Equations

So far we have talked about Maxwell's equations as existing on Minkowski spacetime, which is essentially just $\mathbb{R}^4$ with a Minkowski metric. As it turns out, Minkowksi spacetime is a *flat* spacetime. In Einstein's general theory of relativity spacetime is only approximately flat in the near absence of energy and momentum, and always locally flat. This first statement is that our metric tensor in curved spacetime can be written as

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{3.1}$$

where $h_{\mu\nu}$ is a small perturbation to the Minkowski metric, i.e. $|\det(h_{\mu\nu})| \ll 1$. The second statement that spacetime is locally a Minkowski space is the Einstein Equivalence principle. Stated in the language of differential geometry, the *tangent space* at some point on a manifold such as curved spacetime will be a pseudo-Euclidean vector space. For example, the sphere given by

$$x^2 + y^2 + z^2 = 1 \tag{3.2}$$

has tangent planes that look like $\mathbb{R}^2$. This sphere is called S^2, or the 2-sphere, since it is a 2-manifold. If we wish to write Maxwell's equations on a curved spacetime (or any theory, for that matter), we will need to develop some tools from differential geometry.

3.1 Manifolds

In this section we will define the general idea of a smooth, n-dimensional manifold, or n-manifold. The basic idea is that an n-manifold looks locally like $\mathbb{R}^n$. This does not mean that it has the same *metric* as $\mathbb{R}^n$, but instead that there exist certain functions from the manifold to $\mathbb{R}^n$

that have certain properties.

To develop a well defined notion of a smooth manifold, we start with the definition of a topological space.

Definition 3.1. A *topological space* is a set X together with a specified collection of open sets $U_i \subseteq X$, satisfying the following conditions [1]:

$$\begin{aligned} &1) \text{ The empty set and } X \text{ itself are open,} \\ &2) \text{ If } U_j, U_k \subseteq X \text{ with fixed } j \text{ and } k \text{ are open, so is } U_j \cap U_k, \\ &3) \text{ If the subsets } U_i \subseteq X \text{ with } i = 1, 2, ..., n \text{ are open, so is the union} \cup U_i. \end{aligned} \tag{3.3}$$

The open sets collectively are called the *topology* of X. An open set containing a point $x \in X$ is called a neighborhood of x. For example, in $\mathbb{R}^n$, a set U is open if there exists some $\epsilon > 0$ such that all points in U within an ϵ of the center of U are also in U. A collection of open subsets $U_i \subseteq X$ *covers* X if

$$\bigcup U_i = X.$$

The reason we need a topology to define a smooth manifold is so that we can define continuous functions $\varphi\colon U \to \mathbb{R}^n$. The function φ is called a *chart*, and we can effectively work in $\mathbb{R}^n$ as long as we work in $\varphi(U)$, called a *chart-set*. We are now ready to formally define an n-manifold.

Definition 3.2. An n-manifold is a topological space M covered by open subsets U_i, each having charts $\varphi_i\colon U_i \to \mathbb{R}^n$, which are related in the following way:

$$\begin{aligned} &\text{For some } j \text{ and } k, \text{ the so-called transition function, } \varphi_j \circ \varphi_k^{-1}\colon \varphi_k(U_j \cap U_k) \to \mathbb{R}^n \\ &\text{is smooth where it is defined. Separate collections of charts are called an } \textit{atlas}. \end{aligned} \tag{3.4}$$

This definition allows us to "patch together" a manifold M with open subsets that look like $\mathbb{R}^n$. This also guarantees that if a scalar function $f\colon U_j \to \mathbb{R}$ is smooth, then it will smooth on the intersection of U_j with some other open subset of M. The necessity of smoothness arises because we want the two principal kinds of fields that live on manifolds, *vector fields* and *differential forms*, to be defined everywhere on a manifold.

3.2 Vector Fields

As far as we are interested here, a vector field will be a *differential operator* defined at each point on a manifold whose sole ambition in life is to differentiate smooth functions. We will start by naming a few things. The set of smooth real valued functions on a manifold M is denoted $C^\infty(M)$, and is a *commutative algebra* over the real numbers. Formally, this property means the following:

let $f, g, h \in C^\infty(M)$ and $\alpha, \beta \in \mathbb{R}$. Then at each point in M we have

$$\begin{aligned}
f + g &= g + f \\
f + (g + h) &= (f + g) + h \\
f(gh) &= (fg)h \\
f(g + h) &= fg + fh \\
(f + g)h &= fh + gh \\
1f &= f \\
\alpha(\beta f) &= (\alpha\beta)f \\
\alpha(f + g) &= \alpha f + \alpha g \\
(\alpha + \beta)f &= \alpha f + \beta f \\
fg &= gf
\end{aligned} \tag{3.5}$$

All a vector field does, in some sense, is take one function on a manifold to another function of the same manifold. Abstractly, we say $v\colon C^\infty(M) \to C^\infty(M)$, where v is a vector field. But before we define vector fields on a manifold, we will consider the familiar case in $\mathbb{R}^n$. The directional derivative of a function $f\colon \mathbb{R}^n \to \mathbb{R}$ in the direction of the vector field v is written

$$vf = v^\mu \partial_\mu f \tag{3.6}$$

where $\mu = 1, 2, ..., n$ (in general, but we will have $n = 4$ for our purposes). Actually, this formula will hold for all $f \in C^\infty(\mathbb{R}^n)$, so we can just write.

$$v = v^\mu \partial_\mu \tag{3.7}$$

A *vector field* on M will have the same basic properties as a differential operator in $\mathbb{R}^n$. Namely, linearity and the Leibniz law define a vector field on M.

Definition 3.3. Let v be a vector field on M, $\alpha \in \mathbb{R}$, and $f, g \in C^\infty(M)$. Then we have:

$$\begin{aligned}
v(f + g) &= vf + vg \\
v(\alpha f) &= \alpha v(f) \\
v(fg) &= (vf)g + f(vg),
\end{aligned} \tag{3.8}$$

where the first two are simply linearity, and the last is the Leibniz law or product rule. We should also note that the components of the vector field, v^μ, can themselves be functions of M. We now let $\text{Vec}(M)$ be the set of all vector fields on M. Then $\text{Vec}(M)$ is a module over $C^\infty(M)$, that is, for $f, g \in C^\infty(M)$ and $v, w \in \text{Vec}(M)$ we have

$$\begin{aligned}
f(v + w) &= fv + fw \\
(f + g)v &= fv + gv \\
(fg)v &= f(gv) \\
1v &= v,
\end{aligned} \tag{3.9}$$

where "1" is the constant function equal to $1 \in \mathbb{R}$ on all of M.

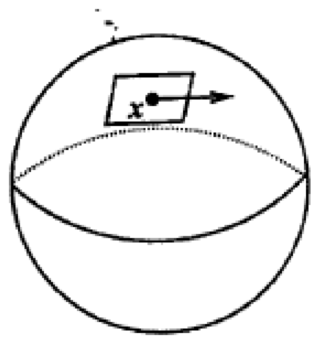

Figure 1: S^2 with a tangent vector at the point $x \in S^2$ [1]

3.2.1 Tangent Vectors

A convenient way to imagine a vector field on M is simply as an arrow assigned to each point $p \in M$. This arrow is called a *tangent vector* at the point p, and is denoted v_p. Where do these tangent vectors live? They belong to the set of all tangent vectors at p, called the *tangent space* at p, denoted $T_p(M)$. We should note, however, that unlike the vector space $\mathbb{R}^n$, it only makes sense to add and subtract vectors that are at the same point on a manifold. The reason we can get away with adding and subtracting vectors at different points in $\mathbb{R}^n$ is because each tangent space has the same basis.

3.2.2 Contravariance and Covariance

We mentioned earlier that an index up quantity like x^μ is contravariant, while an index down quantity like ∂_μ is covariant in a very cursory manner, and we are now almost ready to understand what this terminology really means. First we need one more definition, which is that of a curve on a manifold.

Definition 3.4. A *curve* γ on a manifold is a smooth function $\gamma\colon \mathbb{R} \to M$ such that for any $t \in \mathbb{R}$, $\gamma'(t) \in T_{\gamma(t)}(M)$.

Now suppose we have a function $\phi\colon M \to N$ from one manifold to another and a function $f \in C^\infty(N)$. Then by composing ϕ with f, we obtain a $C^\infty(M)$ function. This process is called *pulling back* f from N to M by ϕ, defined by

$$\phi^* f = f \circ \phi \tag{3.10}$$

so that $\phi^*\colon C^\infty(N) \to C^\infty(M)$, whereas $\phi\colon M \to N$ (see figure 2). For this reason, real-valued functions on manifolds are said to be *contravariant*. We can actually use this fact to check whether or not a function ϕ from one manifold to another is smooth, which is necessary for practically any purpose, especially when doing physics on manifolds. If the pullback of some $f \in C^\infty(N)$ is *not* in $C^\infty(M)$, then ϕ is not smooth. Such functions from one manifold to another are called *maps*. On the other hand, we can use a map ϕ to take curves and tangent vectors from one manifold to another. This operation is called a pushforward.

More generally, we can pull back any $(n, 0)$ tensor, and push forward any $(0, m)$ tensor, where $n, m \in \mathbb{Z}^+$. For example, we can pullback $F^{\mu\nu}$ and x^μ, and pushforward $G_{\mu\nu}$ and ∂_μ. But what

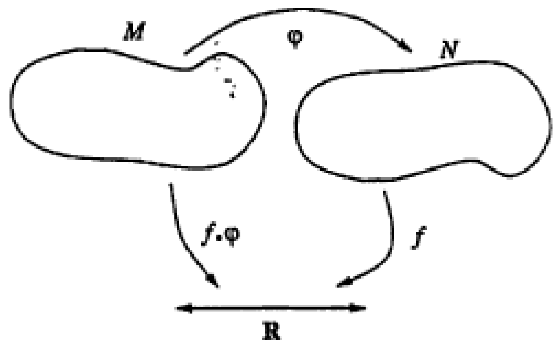

Figure 2: A map ϕ from M to N, a function f, and the pullback of f. [1]

about a mixed tensor quantity, like $F^{\alpha}{}_{\beta}$ or $\Gamma^{\rho}{}_{\alpha\beta}$? We can neither pullback or pushforward these quantities, but we can obtain contravariant and covariant quantities by raising or lowering indices.

3.2.3 The Lie Bracket

Given two vector fields $v, w \in \mathrm{Vec}(M)$, how can we construct a new vector field? Suppose we have two vector fields $v, w \in \mathrm{Vec}(M)$ and $f \in C^{\infty}(M)$. Then the *Lie bracket* of v and w is defined by

$$[v, w](f) = vw(f) - wv(f), \tag{3.11}$$

or, since $[v, w]$ itself is a vector field we can just write

$$[v, w] = vw - wv. \tag{3.12}$$

which we can verify using (3.9). If v and w commute, as is the case when $v = \partial_{\mu}$ and $w = \partial_{\nu}$, then $[v, w] = 0$. The Lie bracket is sometimes called the commutator for this reason. But what is a Lie bracket really? In a sense, it is just a measure of the failure of mixed directional derivatives to commute. The relationship between three (or more) vector fields is summed up by the *Jacobi identity*,

$$[u, [v, w]] + [w, [u, v]] + [v, [w, u]] \equiv 0, \tag{3.13}$$

or in other words, the sum over all even permutations of operators in the above way is zero.

3.3 Differential Forms

3.3.1 1-forms

In elementary calculus we encounter expressions involving differentials, which formally look like a function with a "d" in front of it, as in dy, df, and dr. Our first encounter was probably differentiation

$$\frac{dy}{dx}, \tag{3.14}$$

then integration

$$\int f(x)\ dx, \tag{3.15}$$

and possibly more perplexing expressions like

$$d(\mathrm{sin}x) = \mathrm{cos}x\ dx. \tag{3.16}$$

What, then, is "d" really doing here? We start first by recalling the gradient in $\mathbb{R}^n$, ∇f, which we can think of as an instruction to take $v \in \mathrm{Vec}(\mathbb{R}^n)$ to the directional derivative vf as

$$\nabla f(v) = vf. \tag{3.17}$$

The essential properties of the map are, once again, linearity and the Leibniz law. That is, if $f, g \in C^\infty(\mathbb{R}^n)$ and $v, w \in \mathrm{Vec}(\mathbb{R}^n)$, we have

$$\begin{aligned} \nabla f(v+w) &= \nabla f(v) + \nabla f(w) \\ \nabla f(gv) &= g\nabla f(v) \\ \nabla(fg)(v) &= g\nabla f(v) + f\nabla g(v) \end{aligned} \tag{3.18}$$

or in other words, the gradient $\nabla f \colon \mathrm{Vec}(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n)$ is linear over $C^\infty(\mathbb{R}^n)$.

We now define the differential forms on M, using only the above three properties to define our generalized gradient. Here it is:

$$df(v) = vf. \tag{3.19}$$

The 1-form df is called the *differential* of f, or the *exterior derivative* of f. So what is a 1-form on $\mathbb{R}^n$? It is what you obtain when attempting to make (3.17) and (3.19), the gradient multiplied by the basis of differentials in $\mathbb{R}^n$, $\{dx^\mu\}$:

$$df = \partial_\mu f dx^\mu. \tag{3.20}$$

Using just (3.19), we can show that this truly is consistent with our old definition, (3.17). Let $v = v^\nu \partial_\nu \in \mathrm{Vec}(\mathbb{R})$. Then we have

$$df(v) = vf = v^\nu \partial_\nu f \tag{3.21}$$

by our previous definition. Now on the right hand side we need to do a bit more work:

$$\begin{aligned} \partial_\mu f dx^\mu(v) &= \partial_\mu f v^\nu \partial_\nu x^\mu \\ &= \partial_\mu f v^\nu \delta^\mu_\nu \\ &= v^\nu \partial_\nu f. \end{aligned} \tag{3.22}$$

where the Kronecker delta is defined as

$$\delta^\mu_\nu = \begin{cases} 1, & \mu = \nu \\ 0, & \mu \neq \nu. \end{cases} \tag{3.23}$$

To give a concrete example we show that (3.16) follows from this definition. Let $v \in \mathrm{Vec}(\mathbb{R})$. Then there exists some $f \in C^\infty(\mathbb{R})$ such that $v = f(x)\partial_x$, by our previous definition. So on the left hand side of (3.16) we have

$$d(\sin x)(v) = v \sin x = f(x)\partial_x \sin x = f(x)\cos x \tag{3.24}$$

and on the right hand side we have

$$(\cos x)\ dx(v) = \cos x\, v(x) = \cos x f(x)\partial_x x = \cos x f(x) = f(x)\ \cos x. \tag{3.25}$$

The set of all one-forms is denoted $\Omega^1(\mathbb{R}^n)$, so we have

$$d \colon C^\infty(\mathbb{R}^n) \to \Omega^1(\mathbb{R}^n). \tag{3.26}$$

We have noted that functions are a contravariant quantity on manifolds, and that vector fields are covariant, but there is actually a bit more to it than that. We also say that they are *dual* to one another, which we will now define. Just as a tangent vector at $p \in M$ is in the tangent space T_pM, a one form ω of v_p is in the cotangent space $T_p^*(M)$, which is a special case of what we shall now describe. If V and W are vector spaces over $\mathbb{R}$, and $\nu \in V^*$ and $\omega \in W^*$, then $\nu\colon V \to \mathbb{R}$ and $\omega\colon W \to \mathbb{R}$. Further, if g is a linear map

$$g\colon V \to W, \tag{3.27}$$

then the dual of g is

$$g^*\colon W^* \to V^* \tag{3.28}$$

is defined by

$$(g^*\omega)(v) = \omega(g(v)). \tag{3.29}$$

Notice how g goes from V to W, g^* goes "backwards" from W^* to V^*, so g^* is contravariant. Additionally, if e_i is the basis of V and e^j the basis of V^*, then we have

$$e_i e^j = \delta_i^j. \tag{3.30}$$

The language of differential forms may seem unenlightening at this point, but as we shall see in the next sections, it is a powerful and unifying principle. It will allow us to generalize the gradient, curl, and divergence and their associated identities to arbitrary dimensions and curved geometries. This is a hint that we will also be able to generalize Maxwell's equations in this way, since they are in some sense just the divergences and curls of the electric and magnetic fields.

3.3.2 p-forms

Why do cross products require a "right-hand" rule? How does one take cross products in higher dimensions? Both of these questions can be answered using differential forms. More precisely, by defining a generalized cross product for differential forms (rather than vector fields), we will be able to answer these questions.

Let V be a vector space and $v, w \in V$. Then the *exterior algebra* over V is the algebra generated by V with the relation

$$v \wedge w = -w \wedge v, \tag{3.31}$$

where $\wedge$ is called the *wedge product* or *exterior product*. The exterior algebra is denoted ΛV and can be thought of as V together with all non-zero wedge products of vectors in V. Equation (3.31) immediately implies that

$$v \wedge v = 0 \tag{3.32}$$

for any vector space. Now suppose that V is a 3-dimensional vector space with a basis being the one forms dx, dy, dz. Then any vector in ΛV can be written as a linear combination of wedge products in V. So we have (all of the linear combinations of)

$$\begin{aligned} 1 \in& \Lambda V, \\ dx, dy, &dz \in \Lambda V, \end{aligned} \tag{3.33}$$

and

$$v \wedge w \in \Lambda V \tag{3.34}$$

for all $v, w \in V$. Writing v and w out explicitly,

$$\begin{aligned} v &= v_x dx + v_y dy + v_z dz, \\ w &= w_x dx + w_y dy + w_z dz, \end{aligned} \tag{3.35}$$

the wedge product is then

$$\begin{aligned} v \wedge w &= (v_x dx + v_y dy + v_z dz) \wedge (w_x dx + w_y dy + w_z dz) \\ &= (v_x w_y - v_y w_x) dx \wedge dy + (v_z w_x - v_x w_z) dz \wedge dx + (v_y w_z - v_z w_y) dy \wedge dz \end{aligned} \tag{3.36}$$

where we have used (3.32), (3.33), and the definition of an algebra. We can then obtain another element in ΛV by taking the wedge product of $u \in V$ and (3.36) to obtain

$$u \wedge v \wedge w = \det \begin{pmatrix} u_x & u_y & u_z \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{pmatrix} dx \wedge dy \wedge dz \tag{3.37}$$

or what is the same thing

$$u \wedge v \wedge w = \epsilon^{ijk} \, u_i v_j w_k \, dx \wedge dy \wedge dz, \tag{3.38}$$

where ϵ^{ijk} is the rank-3 Levi-Civita symbol. The rank-3 symbol has the same properties as the rank-4 symbol, but the range of the indices $i, j, k = 1, 2, 3$. If we take the wedge of four or more vectors in V, the result is equal to zero, since

$$dx \wedge dy \wedge dz \wedge dx^i \equiv 0 \tag{3.39}$$

by (3.32) and (3.33).

Now what we have really described here is a linear combination of subspaces $\Lambda^p V \subset \Lambda V$ where $\Lambda^p V$ is the set of p-fold products of vectors in V. Explicitly,

$$\begin{aligned} \Lambda^0 V &= \mathbb{R} \\ \Lambda^1 V &= V. \end{aligned} \tag{3.40}$$

and if $\{e_i\}$ is the set of basis vectors of V, then

$$e_i \wedge ... \wedge e_p \in \Lambda^p V, \tag{3.41}$$

and $\Lambda^p V$ itself is the set of all linear combinations of such wedge products. Interpreted literally, we are saying $\mathbb{R}$ is the wedge product of no vectors, $\Lambda^1 V$ is just V itself, and $\Lambda^p V$ is the space of all wedge product of p vectors.

For two different vector spaces V and V' with bases e_i and e'_j, we define the the tensor product $V \otimes V'$ to be all expressions of the form $e_i \otimes e'_j$ [1]. Additionally, given any bilinear function

$$f \colon V \times V' \to W \tag{3.42}$$

to some vector space W, there is a unique Linear function

$$F \colon V \otimes V' \to W \tag{3.43}$$

such that

$$f(v, v') = F(v \otimes v'). \tag{3.44}$$

Now if V is n-dimensional, how many elements are in the basis of $\Lambda^p V$? As we saw earlier, if $p > n$, then $\Lambda^p V$ is empty. Now since there are n basis vectors in V, we cannot repeat any vectors. Additionally, since any permutation of a wedge product is linearly dependent on the original wedge product, we know the order of the wedge product does not matter. So for $0 \le p \le n$ we have

$$\dim \Lambda^p V = \binom{n}{p} = \frac{n!}{p!(n-p)!} \tag{3.45}$$

which we see is consistent with our definitions (3.40) and (3.41). The dimension of ΛV is then

$$\dim \Lambda V = \sum_{p=0}^{n} \binom{n}{p} = 2^n. \tag{3.46}$$

Now what we really want to do is follow the same procedure as above, but using differential forms.

Definition 3.5. The set of *differential forms* on M, denoted $\Omega(M)$, is the algebra generated by $\Omega^1(M)$ with the relationship

$$\omega \wedge \mu = -\mu \wedge \omega \tag{3.47}$$

for all $\omega, \mu \in \Omega^1(M)$. Recall that the one-forms on M are linear over the functions $C^\infty(M)$. So for the algebra $\Omega(M)$ our coefficients will be functions rather than scalars. Then in analogy to the exterior algebra over a vector space we have

$$\begin{aligned}
\Omega^0(M) &= C^\infty(M) \\
\Omega^1(M) &= d(C^\infty(M)) \\
\Omega^2(M) &= \Omega^1(M) \wedge \Omega^1(M) \\
\Omega^3(M) &= \Omega^1(M) \wedge \Omega^1(M) \wedge \Omega^1(M) \\
&= \Omega^1(M) \wedge \Omega^2(M) \\
&\dots \\
\Omega^p(M) &= \Omega^1(M) \wedge \Omega^{p-1}(M).
\end{aligned} \tag{3.48}$$

Hence, we can take the wedge product of any p form with any q form. Additionally, the anti-commutation of 1-forms implies that $\Omega(M)$ is a *supercommutative* algebra, that is, if $\omega \in \Omega^p(M)$ and $\mu \in \Omega^q(M)$, then

$$\omega \wedge \mu = (-1)^{pq} \mu \wedge \omega. \tag{3.49}$$

We can demonstrate this result by starting with $\omega^i, \omega^j, \omega^k \in \Omega^1(M)$, and noting that $\omega^i \wedge \omega^j \in \Omega^2(M)$. Then,

$$\begin{aligned}
(\omega^i \wedge \omega^j) \wedge \omega^k &= \omega^i \wedge \omega^j \wedge \omega^k \\
&= -\omega^i \wedge \omega^k \wedge \omega^j \\
&= \omega^k \wedge \omega^i \wedge \omega^j \\
&= \omega^k \wedge (\omega^i \wedge \omega^j)
\end{aligned} \tag{3.50}$$

so 2-forms commute with one-forms. A useful way to think about this property is that we have simply transposed two one-forms with a single one-form, each time picking up a minus sign. Hence, the overall sign change is $(-1)^2$ =1. So if we imagine permuting p one-forms with q one-forms, we can see that we pick up an overall sign change of $(-1)^{pq}$.

3.3.3 The Exterior Derivative

The previous sections might leave one asking: what is the exterior derivative of a p-form? The answer is a $(p+1)$-form, and moreover we define the exterior derivative of a p-form to be the unique map

$$d\colon \Omega^p(M) \to \Omega^{p+1}(M) \tag{3.51}$$

with the following properties for all $\omega, \mu \in \Omega(M)$ and $\nu \in \Omega^p(M)$

$$\begin{aligned}
&1)\ d\colon \Omega^0(M) \to \Omega^1(M) \text{ to agree with our previous definition}\\
&2)\ d(\omega+\mu) = d\omega + d\mu \text{ and } d(c\omega) = c\,d\omega,\ (d \text{ is linear})\\
&3)\ d(\nu \wedge \mu) = d\nu \wedge \mu + (-1)^p \nu \wedge d\mu\\
&4)\ d(d\omega) = 0.
\end{aligned} \tag{3.52}$$

Now, using these properties we should be able to show that

$$d(\omega \wedge \mu) = d(-\mu \wedge \omega). \tag{3.53}$$

Let ω and μ be 1-forms. Then on the right hand side of (3.52) we have

$$\begin{aligned}
d(-\mu \wedge \omega) &= -d(\mu \wedge \omega)\\
&= -d\mu \wedge \omega + \mu \wedge d\omega\\
&= -\omega \wedge d\mu + d\omega \wedge \mu\\
&= d\omega \wedge \mu - \omega \wedge d\mu\\
&= d(\omega \wedge \mu),
\end{aligned} \tag{3.54}$$

where we have used 2), 3), and the fact that 2-forms commute with 1-forms. What we have done in this section may not seem very exciting, but this next statement may be more compelling. The gradient, curl, and divergence from vector calculus are related to the exterior derivative in the following way:

$$\begin{aligned}
&\text{Gradient} - && d\colon \Omega^0(M) \to \Omega^1(M)\\
&\text{Curl} - && d\colon \Omega^1(M) \to \Omega^2(M)\\
&\text{Divergence} - && d\colon \Omega^2(M) \to \Omega^3(M).
\end{aligned} \tag{3.55}$$

If you are still not impressed, then consider the fact that all of the information of the previously used vector calculus identity having to do with the composition of gradients, curls, and divergences, is contained in a single operator d.

3.4 The First Equation

We have now developed enough differential geometry to rewrite the first pair of Maxwell's equations,

$$\nabla \cdot \vec{B} = 0, \qquad \nabla \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0, \tag{3.56}$$

on any manifold. From the formulation of the electric and magnetic fields in terms of the scalar and vector potentials and (3.55), we know that $\vec{B}$ must be a 2-form, and $\vec{E}$ a 1-form. Specifically we will write them as

$$\begin{aligned}
B &= B_x\,dy \wedge dz + B_y\,dz \wedge dx + B_z\,dx \wedge dy = \epsilon^{ijk} B_i\,dx^j \wedge dx^k,\\
E &= E_x\,dx + E_y\,dy + E_z\,dz = E_i\,dx^i.
\end{aligned} \tag{3.57}$$

Then we can write the field strength as the 2-form on $\mathbb{R}^4$

$$F = B + E \wedge dt. \tag{3.58}$$

In terms of the rank two tensor

$$F_{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{pmatrix} \tag{3.59}$$

the 2-form is

$$F = \frac{1}{2} F_{\mu\nu} dx^\mu \wedge dx^\nu. \tag{3.60}$$

Taking the exterior derivative of this two-form yields

$$\begin{aligned} dF &= d(B + E \wedge dt) \\ &= dB + d(E \wedge dt) \\ &= dB + dE \wedge dt - d(dt) \wedge E \\ &= dB + dE \wedge dt. \end{aligned} \tag{3.61}$$

Now evaluating each term (in index notation) we have

$$\begin{aligned} dB &= \epsilon^{ijk} B^i \, d(dx^j \wedge dx^k) \\ &= \epsilon^{ijk} B^i \left(d(dx^j) \wedge dx^k - d(dx^k) \wedge dx^j \right) \\ &= 0, \end{aligned} \tag{3.62}$$

since $d(d\omega) = 0$, and similarly

$$\begin{aligned} dE &= E^i \, d(dE \wedge dt) \\ &= E^i \left(d(dE) \wedge dt - d(dt) \wedge dE \right) \\ &= 0. \end{aligned} \tag{3.63}$$

Our first pair of equations can now be written as

$$dF = 0. \tag{3.64}$$

3.5 The Inner Product and the Hodge Star Operator

The wedge product of forms is essentially a generalization of the cross product in $\mathbb{R}^3$, so is there a generalization of the dot product? There is, and we can compute it anywhere on a manifold using the metric tensor $g_{\mu\nu}$ (which is formally defined below). For any 1-forms μ and ω the *inner product* is

$$\langle \omega, \mu \rangle = g^{\alpha\beta} \omega_\alpha \mu_\beta. \tag{3.65}$$

Now we define the wedge product for any p-forms to be

$$\langle \omega^1 \wedge \ldots \wedge \omega^p, \mu^1 \wedge \ldots \wedge \mu^p \rangle = \det[\langle \omega^i, \mu^j \rangle] \tag{3.66}$$

where $\{\omega^i\}$ and $\{\mu^j\}$ are sets of p one-forms and $\langle \omega^i, \mu^j \rangle = g_{ij} \omega^i \mu^j$ is a $p \times p$ matrix. Let E and B be the forms on $\mathbb{R}^3$ as in (3.57). Then the inner products of E and B with themselves are

$$\langle E, E \rangle = E_x^2 + E_y^2 + E_z^2 \tag{3.67}$$

and

$$\langle B,B\rangle = B_x^2 + B_y^2 + B_z^2, \tag{3.68}$$

since the metric on $\mathbb{R}^n$ is just the $n\times n$ identity matrix. The *Lagrangian* for the vacuum equations is

$$L = \frac{1}{2}(\langle E,E\rangle - \langle B,B\rangle) = -\frac{1}{2}F_{\mu\nu}F^{\mu\nu} \tag{3.69}$$

which we will discuss in greater detail in chapter 5.

If we are working in Minkowski spacetime, the metric is just $\eta_{\mu\nu}$, but in general the metric on a chart-set of a manifold is defined in the following way. For some chart set U_α on a n-manifold M, with the set $\{\partial_\mu\}$ forming the basis of all vector fields on that chart, the metric is given by

$$g_{\mu\nu} = g^{\mu\nu} = g(\partial_\mu, \partial_\nu) \tag{3.70}$$

where $g : \mathrm{Vec}(M)\times\mathrm{Vec}(M)\to\mathbb{R}$ is a function that is bilinear, symmetric, and non-degenerate. That is, for some $v, w, v', w' \in \mathrm{Vec}(M)$ and $f\in C^\infty(M)$

$$\begin{aligned} &g(fv+v',w) = fg(v,w)+g(v',w),\\ &g(v,fw+w') = fg(v,w)+g(v,w'),\\ &g(v,w) = g(w,v) \end{aligned} \tag{3.71}$$

define bilinearity and symmetry. Non-degeneracy means that for all $w\in\mathrm{Vec}(M)-\{0\}$, if $g(v,w) = 0$, then

$$v = 0. \tag{3.72}$$

We then define the volume form on the chart set to be

$$\Delta = \sqrt{|\mathrm{det} g_{\mu\nu}|}\; dx^1\wedge dx^2\wedge ...\wedge dx^n. \tag{3.73}$$

For an n-dimensional manifold M, the *Hodge star operator* is the unique map

$$\star : \Omega^p(M)\to\Omega^{n-p}(M) \tag{3.74}$$

such that for all $\omega, \mu\in\Omega^p(M)$

$$\omega\wedge\star\mu = \langle\omega,\mu\rangle\Delta \tag{3.75}$$

Now this definition does not ensure us that $\star$ actually exists, and even less obvious is how to compute the star product of a p-form. We shall explain how to do this calculation for an n-manifold, and then use the example of Minkowski spacetime as a demonstration.

Let M be an n-manifold, $\{dx^0, dx^1, ..., dx^n\}$ be an orthonormal basis of 1-forms on a chart $U_\alpha\subset M$, and $1\le p\le n$. Then for any distinct set of p integers $1\le i_1,\ldots,i_p\le n$, the star product is defined by

$$\star(dx^{i_1}\wedge\ldots\wedge dx^{i_p}) = \det[\langle dx^i, dx^j\rangle]\epsilon^{(i_1,\ldots,i_n)}(dx^{i_{p+1}}\wedge\ldots\wedge dx^{i_n}) \tag{3.76}$$

where on the left hand side the wedge product is over the set $\{i_1,\ldots,i_p\}$ and the right hand side is a wedge product over the complement of those integers

$$\{i_{p+1},\ldots,i_n\} = \{1,\ldots,n\}-\{i_1,\ldots,i_p\}. \tag{3.77}$$

Once a set of p integers is chosen we can determine the sign of the right hand side by multiplying the rank-n Levi-Civita symbol and $\det[\langle dx^i, dx^j\rangle]$ where i and j run over $\{i_1, \ldots, i_p\}$.

If that seems confusing, consider the example of $\mathbb{R}^4$ with $\{dt, dx, dy, dz\}$ forming the basis of 1-forms and the Minkowski metric $\eta^{\mu\nu} = (-1, 1, 1, 1)$. Now letting ω be the wedge product of all basis 1-forms, as in the first column of Table 1

$$\langle \omega, \mu\rangle = \eta^{\mu\nu}\omega_\mu\mu_\nu = -\omega_0\mu_0 + \omega_i\mu_i \tag{3.78}$$

and the volume form

$$\Delta = dt \wedge dx \wedge dy \wedge dz. \tag{3.79}$$

Now we can use (3.75) to calculate the *dual* of ω, $\star\omega$. In this case we can write (3.75) as

$$\omega \wedge \star\omega = \langle \omega, \omega\rangle \, dt \wedge dx \wedge dy \wedge dz \tag{3.80}$$

and determine $\star\omega$ by inspection. Then the star products of all wedge products are as they appear in Table 1. We can then calculate $\star^2\omega$ in a similar way via

$$\star\omega \wedge \star^2\omega = \langle \star\omega, \star\omega\rangle \, dt \wedge dx \wedge dy \wedge dz. \tag{3.81}$$

ω	$\langle\omega,\omega\rangle$	$\star\omega$	$\star^2\omega$
dt	-1	$-dx \wedge dy \wedge dz$	dt
dx	1	$-dt \wedge dy \wedge dz$	dx
dy	1	$dt \wedge dx \wedge dz$	dy
dz	1	$-dt \wedge dx \wedge dy$	dz
$dt \wedge dx$	-1	$-dy \wedge dz$	$-dt \wedge dx$
$dt \wedge dy$	-1	$dx \wedge dz$	$-dt \wedge dy$
$dt \wedge dz$	-1	$-dx \wedge dy$	$-dt \wedge dz$
$dx \wedge dy$	1	$dx \wedge dy$	$-dx \wedge dy$
$dx \wedge dz$	1	$dx \wedge dy$	$-dx \wedge dz$
$dy \wedge dz$	1	$dt \wedge dx$	$-dy \wedge dz$
$dt \wedge dx \wedge dy$	-1	$-dz$	$dt \wedge dx \wedge dy$
$dt \wedge dx \wedge dz$	-1	dy	$dt \wedge dx \wedge dz$
$dt \wedge dy \wedge dz$	-1	$-dx$	$dt \wedge dy \wedge dz$
$dx \wedge dy \wedge dz$	1	$-dt$	$dx \wedge dy \wedge dz$

Table 1: Wedge products of the basis and their star products

From the table we can show that

$$\star^2\omega = (-1)^{p(4-p)+1}\omega, \tag{3.82}$$

or since this equation is valid for all p-forms,

$$\star^2 = (-1)^{p(4-p)+1}. \tag{3.83}$$

The reason for the form of (3.82) is that it is a special case of $n = 4$. There are also 0-forms and and 4-forms, but there is really only one of each, and their star products are summarized by

$$\star\, 1 = -\Delta, \quad \star\Delta = -1 \tag{3.84}$$

which implies that

$$\star^2\, 1 = -1, \quad \star^2\Delta = -\Delta. \tag{3.85}$$

3.6 The Second Equation

Earlier we discussed the dual field strength tensor, $G^{\mu\nu}$, which we calculated without much insight into where it really came from. Now we actually have a set of instructions for taking the dual of the two-form F, or any form for that matter. Returning to our second pair of Maxwell's equations

$$\nabla \cdot \vec{E} = \rho, \quad \nabla \times \vec{B} - \frac{\partial \vec{E}}{\partial t} = \vec{\jmath}, \tag{3.86}$$

we might guess that we should have the one-form analog of $J^\mu = (\rho, \vec{\jmath})$ on the right hand side of our second equation by noting that we can write any one form as $\omega = \omega_i dx^i$. Then possibly the only sensible way to define the current density one form is the following:

$$\begin{aligned} J &= J^\mu \eta_{\mu\nu} dx^\nu \\ &= -\rho\, dt + j^i dx^i. \end{aligned} \tag{3.87}$$

Now the dual of F

$$\star F = \star B + \star(E \wedge dt) \tag{3.88}$$

is still a 2-form since B and $E \wedge dt$ are two forms. If we try by brute force to write a differential operator to apply to $\star F$ in order to obtain a 1-form, we might first try to apply the exterior derivative, but $d\colon \Omega^p(M) \to \Omega^{p+1}(M)$, and so $d \star F \in \Omega^3(M)$. Now if we recall that $\star\colon \Omega^p(M) \to \Omega^{n-p}(M)$, then we will see that $\star d \star F \in \Omega^1(M)$, which is where we want to be, and so we conclude that

$$\star d \star F = J. \tag{3.89}$$

One might regard this "derivation" as a joke, but this result is the correct answer, and we can actually see why in the following way. Consider the divergence of the second inhomogeneous Maxwell equation

$$\nabla \cdot \left(\nabla \times \vec{B} - \frac{\partial \vec{E}}{\partial t} \right) = \nabla \cdot \vec{\jmath} \tag{3.90}$$

to obtain

$$-\nabla \cdot \frac{\partial \vec{E}}{\partial t} = \nabla \cdot \vec{\jmath}. \tag{3.91}$$

Interchanging the order of the derivatives on the left hand side and using $\nabla \cdot \vec{E} = \rho$ we then have

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \vec{\jmath} = 0, \tag{3.92}$$

or in index notation

$$\partial_\mu J^\mu = 0. \tag{3.93}$$

This result is called the continuity equation, and was actually Maxwell's breakthrough in unifying the four Maxwell equations, which before had been regarded as four descriptions of different phenomena. Maxwell noticed that if the term

$$-\frac{\partial \vec{E}}{\partial t} \tag{3.94}$$

was added to the left hand side of the equation

$$\nabla \times \vec{B} = \vec{\jmath} \tag{3.95}$$

the continuity equation becomes automatic in the theory. In addition, it allows for the description of electromagnetic radiation, or electric and magnetic fields that propagate without the presence of electric charge.

After scratching our heads for a while, we realize that we can generalize this identity by applying $d\star$ to both sides of (3.86) to obtain

$$d \star J = 0 \tag{3.96}$$

since on the left hand side of (3.86) we have

$$d \star^2 d \star F = -d^2 \star F = 0, \tag{3.97}$$

which tells us that (3.86) really is the generalization of (3.83).

4 Gauge Fields and Yang-Mills Theory

4.1 Lie Groups

Certain groups are of central importance in modern physics. As it happens, many of these groups are *Lie groups*, or groups of continuous transformations that are also manifolds. Further, the Lie groups are all subgroups of the $n \times n$ real and complex general linear groups, denoted $\mathrm{GL}(n, \mathbb{R})$ and $\mathrm{GL}(n, \mathbb{C})$. These are just $n \times n$ invertible matrices with entries from $\mathbb{R}$ and $\mathbb{C}$. In fact, the study of $\mathrm{GL}(2, \mathbb{R})$ and $\mathrm{GL}(3, \mathbb{R})$ (and $\mathrm{GL}(n, \mathbb{R})$, $n \geq 4$, if your professor was not concerned with your weekend plans) comprises much of elementary linear algebra, we are just assigning a convenient name to these groups of matrices. The subgroup of the general linear groups having determinant 1 are called the *special* linear groups, and are denoted as in $\mathrm{SL}(n, \mathbb{R})$, for example.

Before we specify which of these groups we are particularly interested in, we will state the group axioms. We will use multiplicative notation to specify the operation, with the group operation denoted by "$\cdot$". In general the group operation is just a set of instructions telling us how to combine two elements in the group, and can be multiplication, addition, or any abstract operation on two elements.

Definition 4.1. A *group* is a set G together with an operation $\cdot$ with the following stipulations: For all $g, h, k \in G$

$$\begin{aligned}
&(G0) \quad \text{Closure: } g \cdot h \in G \\
&(G1) \quad \text{Associativity: } g \cdot (h \cdot k) = (g \cdot h) \cdot k \\
&(G2) \quad \text{Identity: There exists an element } 1 \in G \text{ such that } g \cdot 1 = 1 \cdot g = g \\
&(G3) \quad \text{Inverse: There exists an element } g^{-1} \in G \text{ such that } g \cdot g^{-1} = g^{-1} \cdot g = 1
\end{aligned} \tag{4.1}$$

where in (G3) and (G4) we have specified the identity element as "1" and the inverse of an element g as g^{-1}. Now without further ado, let us define some of the most elementary Lie groups that we are interested in, which are groups of $n \times n$ matrices:

The relationships between the various groups we have mentioned so far are summarized in Figure 1.

Group	definition	coefficients from
$O(n)$	$n \times n$ unitary	$\mathbb{R}$
$SO(n)$	$n \times n$ unitary, special	$\mathbb{R}$
$U(n)$	$n \times n$ unitary	$\mathbb{C}$
$SU(n)$	$n \times n$ unitary, special	$\mathbb{C}$

Table 2: The most important groups in gauge theory.

$$\begin{array}{ccc} GL(n,\mathbb{C}) & \supset & SL(n,\mathbb{C}) \\ \cup & & \cup \\ U(n) & \supset & SU(n) \\ \cup & & \cup \\ O(n) & \supset & SO(n) \end{array}$$

Figure 1: The relationships between different Lie groups

To clarify what we mean by the descriptions in Table 2, by a *unitary* matrix we mean that the conjugate transpose (denoted by a superscript †) of a matrix is the inverse of the original matrix, e.g.

$$A = \frac{1}{\sqrt{3}} \begin{pmatrix} i & -1+i \\ 1+i & -i \end{pmatrix}, \quad A^{\dagger} = \frac{1}{\sqrt{3}} \begin{pmatrix} -i & 1-i \\ -1-i & i \end{pmatrix}, \quad A^{\dagger}A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{4.2}$$

Of course if a matrix has real coefficients, as the elements of $O(n)$ do, then the conjugate transpose is just the transpose, since there are no imaginary components to conjugate.

Now we will start with the 1×1 matrices, also known as scalars. We will skip over $SO(1) = 1$, since it is not particularly interesting as a gauge group. Complex numbers of magnitude 1 constitute a Lie group that is an interesting gauge group[4], and in general we write elements of this group as

$$U(1) = \{e^{i\theta} = \cos\theta + i\sin\theta : \theta \in \mathbb{R}\} \tag{4.3}$$

Note that this group is special "automatically", since it is unitary. This group is also abelian, since complex numbers commute. The 2×2 special unitary matrices,

$$SU(2) = \left\{ \begin{pmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{pmatrix} : \alpha, \beta \in \mathbb{C}, |\alpha|^2 + |\beta|^2 = 1 \right\} \tag{4.4}$$

is our first example of a non-abelian Lie group. We can demonstrate that $SU(2)$ is a manifold, or show that it is diffeomorphic to the 3-sphere by setting $\alpha = x + iy$ and $\beta = z + iw$. Then the condition that $|\alpha|^2 + |\beta|^2 = 1$ is equivalent to $x^2 + y^2 + z^2 + w^2 = 1$, and hence $(x, y, z, w) \in S^3$. This establishes a smooth invertible map between $SU(2)$ and S^3, which is the definition of a diffeomorphism.

[4]The gauge group for Electromagnetism.

Drawing a comparison between $U(1)$ and $SU(2)$, one might find the definition of the first more pleasing, since it is very clear how to obtain all elements of the group simply by the definition. We can establish how to obtain all elements of a Lie group in terms of the *Lie algebra*, which we will show in the next section.

As it happens, the group $SO(3)$, which we first encountered as "rotations in $\mathbb{R}^3$", has a special relationship with the group $SU(2)$. We could write a generic $SO(3)$ element as the product of 3 matrices describing rotation in three orthogonal planes in $\mathbb{R}^3$, but this descriptor would be ugly. Instead we will wait until we have some understanding of using Lie algebras to express Lie groups.

4.2 Lie Algebra

The Pauli σ-matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.5}$$

make frequent appearances in physics for describing the spin of particles in quantum mechanics, and, together with the 2×2 identity matrix, form the basis for 2×2 Hermitian matrices. The Pauli matrices satisfy the following relation:

$$\sigma_i \sigma_j = \delta_{ij} I_{2\times 2} - i\epsilon^{ijk}\sigma_k \tag{4.6}$$

where $I_{2\times 2}$ is the 2×2 identity matrix. These imply all three of the Lie algebra axioms, which are as follows

$$\begin{aligned} &1)\ [\sigma_i, \sigma_j] = -[\sigma_i, \sigma_j] \\ &2)\ [\sigma_i,\ \alpha\sigma_j + \beta\sigma_k] = \alpha[\sigma_i, \sigma_j] + \beta[\sigma_i, \sigma_k] \\ &3)\ [\sigma_i, [\sigma_j, \sigma_k]] + [\sigma_j, [\sigma_k, \sigma_i]] + [\sigma_k, [\sigma_i, \sigma_j]] = 0, \end{aligned} \tag{4.7}$$

where $\alpha, \beta \in \mathbb{C}$ and we have used the fact that scalars and matrices commute. Stated precisely, a *Lie algebra* is any vector space $\mathfrak{g}$ over a field equipped with a map $[\cdot,\cdot]\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ satisfying the three axioms above. The map that takes us from a Lie algebra to a Lie group (at least for the Lie groups in Table 2) is exponentiation, just as it was for $U(1)$. For $n \times n$ special matrices, exponentiation is shorthand for the power series

$$\exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!}. \tag{4.8}$$

which we can show converges for the groups of Table 2 by grouping together odd and even terms of the power series and rewriting them as sines and cosines. Now we can write the group $SU(2)$ in a way that can be interpret as rotations. First we will let $\tau_i = -\frac{i}{2}\sigma_i$, then

$$SU(2) = \Big\{ \exp(\theta^i \tau_i) : \theta^i \in \mathbb{R} \Big\} \tag{4.9}$$

Geometrically, the Lie algebra is the tangent space of a Lie group at the identity and we can obtain the Lie algebra from a parameterized element of the Lie group by differentiation. For example, a rotation by an angle t about the z axis in $\mathbb{R}^3$ can be written as

$$\gamma(t) = \begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{4.10}$$

Then, differentiating and setting $t = 0$ we obtain

$$\gamma'(0) = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{4.11}$$

Repeating this procedure for the x and y axes, we obtain the basis of the Lie algebra $\mathfrak{su}(3)$:

$$J_x = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad J_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad J_z = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}. \tag{4.12}$$

Now we can write

$$SO(3) = \Big\{ \exp(\theta^i J_i) : \theta^i \in \mathbb{R} \Big\}. \tag{4.13}$$

We can say that $SO(3)$ "acts" on the real vector space $\mathbb{R}^3$, or that any element of the group defines a linear transformation of $\mathbb{R}^3$. In general, we say that a group G acts on a vector space V if there exists a homomorphism

$$\rho \colon G \to GL(V), \tag{4.14}$$

i.e., $g, h \in G$ and $v \in V$ implies that

$$\rho(gh)v = \rho(g)\rho(h)v. \tag{4.15}$$

Such a homomorphism is called a *representation* of G on V.

4.3 Bundles and Connections

Non-abelian gauge theories have a conceptually identical structure to the theory of fiber bundles. In this way we can understand Yang-Mills theory by studying fiber bundles (so long as we keep the physical interpretation in mind).

Consider two manifolds M and E and an onto map $\pi \colon E \to M$. We call this structure (the manifolds and the map, together) a *bundle*. If $p \in M$ we say that

$$E_p = \{q \in E \colon \pi(q) = p\} \tag{4.16}$$

is the *fiber over* p. Then since π is an onto map, the union of all fibers is the entirety of E itself:

$$E = \bigcup_{p \in M} E_p. \tag{4.17}$$

For the bundles we are interested in, M is a manifold and each fiber is a vector space. One such example of a fiber that is a vector space is the tangent space at p, $T_p(M)$. It is then quite natural to define the *tangent bundle* as the union of each tangent space:

$$TM = \bigcup_{p \in M} T_p M. \tag{4.18}$$

If we have two manifolds M and F such that E can be expressed as their cartesian product

$$E = M \times F, \tag{4.19}$$

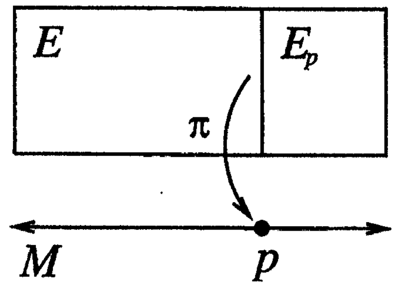

Figure 3: A bundle. [1]

then we say that the bundle is a *trivial bundle* over M with a *standard fiber* F. A vector bundle is a trivial bundle with a standard fiber that is a vector space, for example, Figure 3 is an example of a trivial bundle. Another example is a cylinder, which is a trivial bundle over S^1 with a standard fiber $\mathbb{R}$.
The bundles that we will be concerned with will all be locally trivial with a standard fiber that is a vector space. That is, for each point $p \in M$ there is a neighborhood U of p and a map ϕ such that

$$\phi \colon E|_U \to U \times \mathbb{R}^n \tag{4.20}$$

is a homeomorphism, where for some $U \subset M$ we define

$$E|_U = \{q \in E \colon \pi(q) \in U\}. \tag{4.21}$$

Note that the fiber can also be $\mathbb{C}^n$. Now if E_p is a vector space, then it has a dual E_p^*, and we let the union be defined as

$$E^* = \bigcup E_p^*. \tag{4.22}$$

The physical fields that we described earlier can be seen as particular aspects of a bundle called *sections*. A section, s, of a bundle $\pi \colon E \to M$ is a function $s \colon M \to E$ such that

$$s(p) \in E_p \tag{4.23}$$

for all $p \in M$. This definition implies that the composition $\pi \circ s$ is the identity map on M. For a

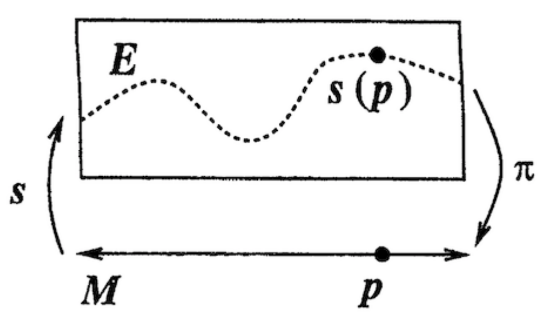

Figure 4: A set of sections of p (dotted line) [1].

vector bundle with a standard fiber $\mathbb{R}^n$, we have $s\colon M \to \mathbb{R}^n$.

As it happens, the maps of some set of sections s_i could have the property that $\pi \circ s_i(p) = p$ so a section is not unique in that sense. The set of all sections is denoted $\Gamma(E)$.

4.3.1 The Connection and Exterior Covariant Derivative

A connection, D, gives us a way to differentiate sections. Stated formally, for a given vector field $v \in \mathrm{Vec}(M)$, a *connection* D on M is a function $D_v\colon \Gamma(E) \to \Gamma(E)$ with the following properties:

$$\begin{aligned} D_v(\alpha s) &= \alpha D_v(s) \\ D_v(s+t) &= D_v(s) + D_v(t) \\ D_v(fs) &= v(f)s + fD_v(s) \\ D_{v+w}(s) &= D_v(s) + D_w(s) \\ D_{fv}(s) &= fD_v(s) \end{aligned} \tag{4.24}$$

for all $v, w \in \mathrm{Vec}(M)$, $s, t \in \Gamma(E)$, $f \in C^\infty(M)$, and $\alpha \in \mathbb{C}$ or $\alpha \in \mathbb{R}$ if s maps to a complex or real valued fiber, respectively. We call $D_v s$ the covariant derivative of s in the direction of v. We will use an abbreviation for the flat or local coordinate case:

$$D_\mu = D_{\partial_\mu}. \tag{4.25}$$

The exterior covariant derivative is defined in terms of the connection in the following way:

Definition 4.2. The *exterior covariant derivative* d_D of a section s is an E-valued 1-form such that

$$d_D s(v) = D_v s \tag{4.26}$$

for any vector field $v \in \mathrm{Vec}(M)$. In local coordinates x^μ on an open subset $U \subset M$

$$d_D s(\partial_\mu) = D_\mu s \otimes dx^\mu. \tag{4.27}$$

In local coordinates, if the basis of sections is $\{e_i\}$, then the covariant derivative is defined in terms of components

$$D_\mu e_i = A^j_{\mu i} e_j \tag{4.28}$$

where $A^j_{\mu i}$ are called the *connection coefficients* or *Christoffel symbols*. The covariant derivative of a section s is

$$\begin{aligned} D_\mu s &= D_\mu(s^i e_i) \\ &= ((\partial_\mu s^i)e_i + A^j_{\mu i} s^i e_i)) \\ &= (\partial_\mu s^i + A^i_{\mu j} s^j)e_i \end{aligned} \tag{4.29}$$

where we have used the Leibniz law and interchanged the indices i and j of the Christoffel symbol in the last line so that we could factor out e_i. We can define

$$D_\mu s = (D_\mu s)^i e_i \tag{4.30}$$

so that

$$(D_\mu s)^i = \partial_\mu s^i + A^i_{\mu j} s^j. \tag{4.31}$$

4.4 Origins of Yang-Mills Theory

In 1954 Chen Ning Yang and Robert Mills published a paper that would eventually lead to the extension of the $U(1)$ gauge symmetry of electromagnetism to other physical theories, most notably those with $SU(n)$ gauge symmetry. In this section we will begin by writing down the Yang-Mills equations and systematically comparing them to Maxwell's equations.

4.5 The Yang-Mills equations

Formally the Yang-Mills equations look almost identical to Maxwell's equations,

$$\begin{aligned} d_D F &= 0 \\ \star d_D \star F &= J. \end{aligned} \tag{4.32}$$

Actually, the only real difference between the two sets of equations is the gauge group, and the way we have written them does not restrict us to any particular manifold.

4.5.1 Curvature

The field strength tensor for electromagnetism is just a special case of the *curvature*, which is an operator acting on the sections of E

$$F(v,w)s = D_v D_w s - D_w D_v s - D_{[v,w]} s \tag{4.33}$$

which as usual we can just write

$$F(v,w) = D_v D_w - D_w D_v - D_{[v,w]}. \tag{4.34}$$

Now when $v = \partial_\mu$ and $v = \partial_\nu$, we have

$$\begin{aligned} F_{\mu\nu} = F(\partial_\mu, \partial_\nu) &= D_\mu D_\nu - D_\nu D_\mu - D_{[\partial_\mu, \partial_\nu]} \\ &= D_\mu D_\nu - D_\nu D_\mu \\ &= [D_\mu, D_\nu] \end{aligned} \tag{4.35}$$

where we have used the fact that $[\partial_\mu, \partial_\nu] = 0$. In attempting to find a field strength tensor that would generalize to non-abelian gauge groups, Yang and Mills guessed that they should add an additional "quadratic term", $A \wedge A$, making the curvature

$$F = dA + A \wedge A \tag{4.36}$$

where A is the $\mathrm{End}(E)$ valued 1-form $A = A_\mu \otimes dx^\mu$. It was a fortuitous guess, since this result is the correct generalization of Maxwell's equations to non-abelian gauge theories. In a local coordinate system x^μ,

$$F = \frac{1}{2} F_{\mu\nu}\, dx^\mu \wedge dx^\nu \tag{4.37}$$

as before, but the field strength is now

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu] \tag{4.38}$$

and we can see that the quadratic term is the commutator of the vector potential, which is zero for an abelian gauge theory like electromagnetism.

4.5.2 The Bianchi Identity

For Maxwell's first pair of equations, we could demonstrate that $dF = 0$ by using the identity $d^2 = 0$ and the definition of the field strength tensor as the derivative of the vector potential. This same procedure does not work in general for the exterior covariant derivative d_D. The Bianchi identity

$$[D_u, [D_v, D_w]] + [D_v, [D_w, D_u]] + [D_w, [D_u, D_v]] = 0 \tag{4.39}$$

is true for any linear operator like D_u. Now taking $u = \partial_\mu$, $v = \partial_\nu$, $w = \partial_\lambda$, and using the fact that $F_{\mu\nu} = [D_\mu, D_\nu]$ we have

$$[D_\mu, F_{\nu\lambda}] + [D_\nu, F_{\lambda\mu}] + [D_\lambda, F_{\mu\nu}] = 0. \tag{4.40}$$

Now let us work out $d_D F$ in local coordinates:

$$\begin{aligned} d_D F &= D_\mu F \otimes dx^\mu \\ &= \frac{1}{2} D_\mu F_{\lambda\nu}\, dx^\lambda \wedge dx^\nu \otimes dx^\mu \\ &= \frac{1}{3!}(D_\mu F_{\lambda\nu} + D_\lambda F_{\nu\mu} + D_\nu F_{\mu\lambda}) \otimes dx^\lambda \wedge dx^\nu \wedge dx^\mu. \end{aligned} \tag{4.41}$$

It can be shown that

$$D_\mu F_{\lambda\nu} + D_\lambda F_{\nu\mu} + D_\nu F_{\mu\lambda} = 0 \tag{4.42}$$

follows from the Bianchi identity, and so we have $d_D F = 0$.

5 Quantized Yang-Mills Theories

Strictly speaking, what we have developed in the last chapter is a way to write physical theories on arbitrary spacetimes. The applications of physical theories on curved spacetimes are generally either astrophysical in nature, or Beyond the Standard Model (BSM) of particle physics. An example may be helpful. One object of study that requires Maxwell's equations in curved space are magnetars [11], which are extremely dense remnants of supernovae that have a very strong magnetic field (about 10^{10} times stronger than any magnet on Earth).

What we have developed so far is considered to be a classical Yang-Mills theory. In order to develop physical theories, we need to quantize the Yang-Mills field. Proper quantization of a field is a very involved effort, so rather than actually quantizing a field, we will sketch how to do this and interpret the results.

5.1 Lagrangian Mechanics

In classical physics, there is a coordinate-free generalization of Newtonian mechanics called *Lagrangian Mechanics* which relies on knowing the quantity

$$L = T - V, \tag{5.1}$$

called the *Lagrangian* of a system of particles, where T is the kinetic energy and V is the potential energy. The Lagrangian of a system of N particles in $\mathbb{R}^3$ will then be described by $3N$ coordinates $\{q_i(t) : i = 1, ..., 3N\}$ and $3N$ velocities $\{\dot{q}_i(t) = \frac{dq_i}{dt} : i = 1, ..., 3N\}$, where q_i and $\dot{q}_i$ are treated as independent variables. Given the Lagrangian of a physical system, we can determine the paths that the particles take by minimizing the *action* [8]

$$S = \int_{t_1}^{t_2} L(q_i, \dot{q}_i) dt. \tag{5.2}$$

We say an action is minimized when the variation $\delta S = 0$ to linear order in $\delta q_i(t)$ under the transformation

$$\delta q_i(t) \rightarrow q_i(t) + \delta q_i(t) \tag{5.3}$$

where $\delta q_i(t)$ is a smooth function that is zero at the limits of integration, i.e.

$$\delta q_i(t_1) = \delta q_i(t_2) = 0. \tag{5.4}$$

The variation of the action is

$$\begin{aligned} \delta S &= \int_{t_1}^{t_2} \delta q_i \frac{\partial L}{\partial q_i} + \delta \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} dt \\ &= \int_{t_1}^{t_2} \delta q_i \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) dt + \left[\delta q_i \frac{\partial L}{\partial \dot{q}_i} \right]_{t_1}^{t_2} \end{aligned} \tag{5.5}$$

where we have performed integration by parts on the second term of the integrand. The term outside the integrand is 0 by (5.4), and since the function $\delta q_i(t)$ is arbitrary we have

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = 0. \tag{5.6}$$

Hence, the action is minimized when (5.6), known as the *Euler-Lagrange equation*, is satisfied.

For relativistic fields[5], we cannot treat time as an entity distinct from spacial dimensions. In other words a field $\phi(x^\mu)$ is a function of spacetime. Instead of working with a Lagrangian, we will now work with a Lagrangian density

$$L = \int_V \mathcal{L} \ d^3x, \tag{5.7}$$

and make the following substitutions in the Euler-Lagrange equation

$$\begin{aligned} L &\rightarrow \mathcal{L}, \\ q_i &\rightarrow \phi_i, \\ \frac{d}{dt} &\rightarrow \partial_\mu, \end{aligned} \tag{5.8}$$

we can obtain the relativistic Euler-Lagrange equation

$$\frac{\partial \mathcal{L}}{\partial \phi_i} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)} \right) = 0 \tag{5.9}$$

for N fields. Though this equation is relativistically invariant, like the electromagnetic field, it is still a classical field. In order to obtain a quantum field, it needs to be "quantized", which formally amounts to promoting the dynamical variables (fields) to operators and imposing a canonical commutation relation [4].

[5] That is, we want to work with fields instead of particles.

The first field that we will look at is the Dirac field, which is a fermionic field whose excitations are spin 1/2 particles like electrons (e), as well as their heavier cousins the μ and τ, but also quarks, which make up particles like protons and neutrons. Fermionic fields are perhaps the most fundamental in all of quantum field theory; as a matter of fact, all physical matter (at least the kind we are familiar with on Earth) is composed of fermions.

5.2 The Dirac Equation

In the early 20th century, physicists sought a quantum mechanical wave equation that was compatible with special relativity. To obtain the non-relativistic Schrödinger equation, we start with the classical energy momentum relation for a single particle

$$E = T + V = \frac{p^2}{2m} + V \tag{5.10}$$

and apply the "quantum prescription", which amounts to the substitutions

$$\begin{aligned} p &\to i\nabla \\ E &\to i\frac{\partial}{\partial t} \end{aligned} \tag{5.11}$$

to obtain the operator

$$i\frac{\partial}{\partial t} = -\frac{1}{2m}\nabla^2 + V. \tag{5.12}$$

The equation obtained by applying this operator to a function

$$-\frac{1}{2m}\nabla^2\Psi + V\Psi = i\frac{\partial}{\partial t}\Psi. \tag{5.13}$$

yields the Schrödinger equation, where Ψ is called the wave function, and is a function of position and time (which are treated as independent). The integral of the square of the wave function is set to unity

$$\int_{\text{All space}} \Psi^*\Psi \; d^3x = 1 \tag{5.14}$$

and we can find the probability that a particle will be in some region of space by simply integrating over that region. As the wave function evolves with time the integral remains unitary. Following the same procedure with the relativistic energy momentum relation

$$p^\mu p_\mu - m^2c^2 = 0 \tag{5.15}$$

yields an equation that is second order in time. This turns out to be a problem, because a wave equation constructed using (5.15) will not remain unitary as time goes on. Schrödinger actually tried the second procedure before the first, and gave up when it failed to predict the hydrogen emission spectrum. Dirac later realized that an equation that was first order in time could be obtained by factoring the energy momentum relation,

$$p^\mu p_\mu - m^2c^2 = (\gamma^\alpha p_\alpha + mc)(\gamma^\kappa p_\kappa - mc). \tag{5.16}$$

and determining the coefficients γ^α such that

$$p^\mu p_\mu = \gamma^\alpha\gamma^\kappa p_\alpha p_\kappa. \tag{5.17}$$

By expanding both sides and comparing coefficients, we realize that the coefficients must satisfy an anticommutation relation

$$\{\gamma^\mu, \gamma^\nu\} = \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu} \tag{5.18}$$

where $\eta^{\mu\nu}$ is the Lorentz metric. As it turns out, the coefficients have to be matrices, and the smallest ones that work are the following 4×4 matrices

$$\begin{aligned} \gamma^0 &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad \gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \\ \gamma^2 &= \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}, \quad \gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}. \end{aligned} \tag{5.19}$$

Notice that this implies the right hand side of (5.18) is multiplied by the 4×4 unit matrix. We then obtain the Dirac equation by picking one of the factors of (5.16), substituting $p_\mu \to i\partial_\mu$, and applying the resulting operator to a wave function ψ to obtain

$$(i\gamma^\mu\partial_\mu - mc)\psi = 0, \tag{5.20}$$

known as the Dirac equation. The fact that γ^μ are 4×4 matrices also implies that the wave function is a 4 component column vector:

$$\psi = \begin{pmatrix} \psi_0 \\ \psi_1 \\ \psi_2 \\ \psi_3 \end{pmatrix}. \tag{5.21}$$

To write down the Dirac Lagrangian, we also need the Hermitian conjugate of ψ,

$$\bar{\psi} \equiv \psi^\dagger\gamma^0 = (\psi_3^*\, \psi_2^*\, \psi_1^*\, \psi_0^*). \tag{5.22}$$

5.3 Quantum Electrodynamics

Writing down the Dirac Lagrangian[6] only takes 3 more strokes of a pencil than the Dirac equation,

$$\mathcal{L}_{\text{Dirac}} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi \tag{5.23}$$

but we can get considerably more mileage from a Lagrangian than a wave equation, primarily for two reasons. The first is that we can tell the symmetries of a theory from the Lagrangian as a consequence of Noether's theorem, which says that every conservation law (of energy, momentum, charge, etc.) correspond to a continuous symmetry of a Lagrangian. The second is that we can read the Feynman rules for calculating different processes directly off of the Lagrangian.

Together with the Maxwell Lagrangian,

$$\mathcal{L}_{\text{Maxwell}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{5.24}$$

[6] We can verify that a Lagrangian is correct by taking ψ and $\partial_\mu\psi$ as the dynamical variables in the Euler-Lagrange equation, and seeing that it spits the wave equation back out.

we almost have all of the ingredients of the QED Lagrangian,

$$\mathcal{L}_{\text{QED}} = \mathcal{L}_{\text{Dirac}} + \mathcal{L}_{\text{Maxwell}} + \mathcal{L}_{\text{int}} \tag{5.25}$$

where the last term, $\mathcal{L}_{\text{int}}$ is a yet unspecified interaction term. Without it we have a Lagrangian for a theory of two fields minding their own business. The interaction term can be determined by replacing ∂_μ with the gauge covariant derivative [4]

$$D_\mu = \partial_\mu + ieA_\mu \tag{5.26}$$

and seeing what "extra" term we pick up, where e is the coupling constant (the electrical charge) between the Dirac and electromagnetic field. Notice that if $e = 0$, the fields "decouple", or go back to minding their own business. Actually, now we can just write the QED Lagrangian as

$$\mathcal{L}_{\text{QED}} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{5.27}$$

and see what extra term the gauge covariant produces when compared to (5.23) and (5.24) alone. When we work this out, the interaction term is

$$\mathcal{L}_{\text{int}} = -e\bar{\psi}\gamma^\mu\psi A_\mu. \tag{5.28}$$

The theory of Quantum Electrodynamics is in some sense "solved". This claim is due to the fact that a quantitative description of all electromagnetic processes can be calculated to an incredible degree of precision by only evaluating a handful of integrals that can be read off from the Lagrangian. More accurately, these integrals[7] are terms in a series expansion of the scattering matrix, S: [4]

$$S = \sum_{n=0}^{\infty} S^{(n)} \tag{5.29}$$

with each term

$$S^{(n)} \propto \left(\frac{e^2}{4\pi}\right)^n \approx \left(\frac{1}{137}\right)^n. \tag{5.30}$$

Since the coupling constant e is relatively small, we really only have to evaluate the first terms in this series to get an accurate approximation of S. For the theory of strong nuclear force (Quantum Chromodynamics, or QCD) this kind of method is only possible at energies on the order of several GeV, which are generally only achievable by particle accelerators. Thus, there has been a historic struggle to understand QCD at low energy scales.

5.4 Quantum Chromodynamics

We will proceed to the theory of quarks and gluons (QCD) that describe the strong interaction of atomic nuclei by analogy with QED. First, while there is only one charge in QED, the electrical charge, there are three "color" charges in QCD, denoted r, b and g, and their negatives, $\bar{r}, \bar{b}$ and $\bar{g}$. Quarks carry a single color charge. Another difference arises from the fact that the gauge group is $SU(3)$, rather than $U(1)$. Whereas $U(1)$ has a single element in the basis of it's Lie algebra, $SU(3)$ has 8, which correspond to the number of "force carriers", or *gauge bosons* in the theory.

[7] We will not explicitly write out these integrals since we have not developed the formalism of scattering theory here.

The Lagrangian for the theory of QCD does look almost identical to that of QED, save for three new sets of indices that we are summing over

$$\mathcal{L}_{\text{QCD}} = \sum_q \bar{\psi}_{q,i}((i\gamma^\mu D_\mu)^{ij} - m_q\delta^{ij})\psi_{q,j} - \frac{1}{4}F^a_{\mu\nu}F^{\mu\nu}_a \tag{5.31}$$

where the flavor index $q = u, d, s, c, b, t$ runs over all flavors of quarks, $i, j = r, b, g$ run over all quark color charges and $a = r\bar{b}, g\bar{r}...$ runs over 8 color-anticolor pairs of gluons. The reason we sum over flavors in QCD but not QED is because quarks undergo flavor mixing [4], or they spontaneously change from one flavor to another, whereas we do not see mixing between the charged leptons' flavors e, μ, and τ[8]. The gluon field strength tensor is

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + gf^{abc}A^b_\mu A^c_\nu \tag{5.32}$$

where f^{abc} is the structure constants of $SU(3)$. By comparison with $SU(2)$, the "Pauli matrices" for $SU(3)$ are called the Gell-Mann matrices:

$$\begin{aligned}
\lambda_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
\lambda_4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \\
\lambda_6 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}
\end{aligned} \tag{5.33}$$

and the structure constant is

$$f^{abc} = -\frac{i}{4}\,\text{Tr}([\lambda_a, \lambda_b]\lambda_c). \tag{5.34}$$

The gauge covariant derivative for QCD is

$$D_\mu = \partial_\mu - gA^a_\mu\frac{\lambda_a}{2}. \tag{5.35}$$

Now that we have enumerated what every term in the Lagrangian is, we still do not know what the difference between QED and QCD, besides the fact that the latter has more indices.

First, quarks and gluons are never seen in isolation. They come primarily in pairs or triplets of quarks called pions and nucleons (like protons and neutrons) that exchange gluons with one another. Additionally, gluons interact with one another, whereas photons do not. This means we can see "glue-balls", or bound states of multiple gluons [10]. In any case, we never see net color charged particles of any kind. This phenomenon is called color confinement, and though we do not have a rigorous mathematical proof that QCD should have this property, it has become abundantly clear from experimental evidence as well as numerical evidence that this is the case [9].

A rigorous mathematical framework for much of quantum Yang-Mills theory is still very much lacking. Attempts to do so have, however, led to new and interesting insights into pure mathematics,

[8] Or at least this is rare enough not to be observed. Their neutral counterparts, the neutrinos, undergo very strong flavor mixing.

particularly the study of three and four- manifolds [9]. Even on the practical side, the low energy dynamics in QCD are only accessible by numerical Monte Carlo calculations which require millions of hours of compute time[9]. Indeed, many more advances in the underlying theory of physical simulation need to be better developed in order to make a broader range of phenomena in nuclear physics (among others) accessible by this method [12].

[9] Luckily these calculations can be done in parallel, so this translates to hundreds of hours on a massively parallel supercomputer.

References

[1] J. Baez and J. P. Muniain, *Gauge Fields, Knots And Gravity,* World Scientific, 1994.

[2] G. 't Hooft, editor. *50 Years of Yang-Mills theory,* World Scientific, 2005.

[3] D. J. Griffiths, *Introduction to Elementary Particles*, Wiley-VCH, 2010.

[4] M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory,* Addison-Wesley, 1995

[5] R. Millman and G. D. Parker, *Elements of Differential Geometry,* Prentice-Hall, 1979

[6] J. Oprea, *Differential Geometry and Its Applications*, Prentice-Hall, 2000

[7] J. M. Lee, *An introduction to Topological Manifolds,* Springer-Verlag, New York, 2000.

[8] J. M. Rabin, *Introduction to Quantum Field Theory for Mathematicians,* American Mathematical Society, Rhode Island, 1995.

[9] A. Jaffe and E. Witten, *Quantum Yang Mills Theory,* Clay Mathematics Institute, United Kingdom, 2000.

[10] Y. Chen et. al., *Glueball Spectrum and Matrix Elements on Anisotropic Lattices,* Physical Review D73, 2006.

[11] M. Hoven et. al., *Magnetar Oscillations II: spectral method,* Royal Astronomical Society, 2012

[12] A. Alexandru et. al., *Monte Carlo calculations of the finite density Thirring model,* Physical Review D95, 2017.