

Notes on Gauge Field Theory

Matt Noonan

October 8, 2004

1 Group Actions and Bundles on Manifolds

1.1 The Group Action Picture

Suppose P is a smooth manifold and G a (usually compact, finite-dimensional) Lie group acting smoothly on P by diffeomorphisms. We also ask that the action of G on P is free, so that (except for the identity) no element of G has a fixpoint in P . In this case, you can see that the orbit of any $p \in P$ is diffeomorphic to G . For reasons to be discussed later, We call P a *principal G -bundle*, and G is called the *structure group* of the bundle.

Since we asked G to act smoothly on P , around any $p \in P$ we can find a local cross-section ψ to the action. More precisely, suppose that P is n -dimensional, G is k -dimensional, $p \in P$ is arbitrary and U^m is the open unit disk in \mathbb{R}^n . Then there is a smooth map $\psi : U^{n-k} \rightarrow P$ which picks one element (smoothly!) from each orbit near p . This is known as *local triviality*: around every point, P is locally diffeomorphic to $U^{n-k} \times G$ in a way which preserves the group action. Note that if $g : U^{n-k} \rightarrow G$ is any smooth map then $g \cdot \psi$ is another local cross-section and, in fact, every local cross-section near p is of this form.

Local triviality is essentially a tameness condition: it tells us that the P has no interesting local structure, so complications can only arise through larger-scale topological considerations. Much more about this will be said later.

1.2 The Bundle Picture

Because of the existence of local cross-sections, we can “factor out” by the action of G to get a new smooth manifold M of dimension $n - k$. The points of M are G -orbits in P , and the open sets in M are generated by sets of the form $G \cdot \psi(U^{n-k}) \subset P$. This process of factoring out gives us a smooth surjective map $P \xrightarrow{\pi} M$. In this context, P is called the *total space* and M the *base space*.

By considering the map π in more detail, we can see where the “bundle” terminology comes from. First, consider the *fiber* $\pi^{-1}(p)$ for any $p \in M$. Of course this is just the G -orbit of p in P , and by our assumptions on the G -action, $\pi^{-1}(p) \cong G$ for every $p \in M$. Thus, the fibers of π are all alike. Furthermore, the local triviality tells us that for any sufficiently small open set U around p , $\pi^{-1}(U) \cong U \times G$.

These observations will actually let us reverse the process: starting with an arbitrary Lie group G and smooth manifold M , we can build a principal bundle $P \xrightarrow{\pi} M$. Begin by covering M with open sets U_i , $i = 1, 2, 3, \dots$. To each of these sets, associate a smooth function $\psi_i : U_i \rightarrow G$.

We wish to declare these ψ_i to be local cross-sections in our bundle, but there is a subtlety. Suppose that U_i and U_j intersect, and call this intersection U_{ij} . For $p \in U_{ij}$, $\psi_i(U_{ij})$ and $\psi_j(U_{ij})$ each represent local cross-sections to p , and thus should be related by a *gluing map* $g_{ij} : U_{ij} \rightarrow G$:

$$\psi_j(p) = \psi_i(p) \cdot g_{ij}(p)$$

Since the gluing maps act by right multiplication, the bundle action will be on the left (though this point sounds innocent enough, it is worth pondering!)

By our definition, the gluing maps must satisfy the *cocycle condition*

$$g_{ik} = g_{ij} \cdot g_{jk}$$

from which it also (generically) follows that $g_{ij}^{-1} = g_{ji}$ and $g_{ii} = \text{id}$.

Now we can define our bundle. Take the sets $U_i \times G$ and for every $p \in U_{ij}$ define an equivalence

$$(p, \alpha) \in U_i \times G \sim (p, \beta) \in U_j \times G \iff \beta = \alpha \cdot g_{ij}(p)$$

This equivalence respects the left group action and the manifold structure of the local products, and therefore gives us a smooth manifold P with a free left G -action such that the factor space is M .

Further consideration of the construction above shows that we can do away with the choice of ψ_i , so the data needed to build a principal bundle is the following:

1. The base space: a smooth manifold M
2. A trivializing cover: an open covering $\{U_i\}$ of M
3. The structure group: a Lie group G
4. Gluing maps: smooth functions $g_{ij} : U_i \cap U_j \rightarrow G$ satisfying the three cocycle conditions

Given these four things, we can produce a unique principle G -bundle P over M .

1.3 An Example: The Hopf Bundle

Let us put all this abstraction to work on an example which will resurface many times in both geometry and electromagnetism (though we cannot explain why until later!).

There are only three spheres which are also Lie groups:

$$S^0 = \{\pm 1\} = \{r \in \mathbb{R} : |r| = 1\}$$

$$S^1 = U(1) = \{z \in \mathbb{C} : |z| = 1\}$$

$$S^3 = Spin(3) = SU(2) = Sp(1) = \{q \in \mathbb{H} : |q| = 1\}$$

where \mathbb{H} is the quaternions. Recall that the quaternions are the real 4-dimensional division algebra with basis $\{1, i, j, k\}$ such that $i^2 = j^2 = k^2 = ijk = -1$. The quaternions admit a conjugation $\overline{a + bi + cj + dk} = a - bi - cj - dk$ and the norm is defined by $|q| = \sqrt{q\bar{q}}$.

Every element of S^3 is of the form $z + wj$ for unique $z, w \in \mathbb{C}$ with $|z|^2 + |w|^2 = 1$. Now let S^1 act on S^3 by $z + wj \mapsto (e^{i\theta}z) + (e^{i\theta}w)j$. Clearly this action is smooth, but it is also free: if there were a fixpoint q , it would mean $e^{i\theta}q = q$. But the quaternions form a division algebra, so we may right-cancel the two q s to get $e^{i\theta} = 1$. We have therefore presented S^3 as the total space of a principal S^1 bundle.

At this point it is natural to ask “what is the base space of this bundle?” We have identified S^3 with the (real) unit sphere $\mathbb{C}^2/\mathbb{R}^{>0}$ by means of $z + wj \mapsto (z, w)$. The S^1 action then takes $(z, w) \mapsto (e^{i\theta}z, e^{i\theta}w)$. Thus, if we factor out by the S^1 action we have shown that the base space is

$$M = (\mathbb{C}^2/\mathbb{R}^{>0})/e^{i\theta} = \mathbb{C}^2/\mathbb{C}^\times = \mathbb{CP}^1$$

\mathbb{CP}^1 is simply the Riemann sphere $S^2 = \mathbb{C} \cup \{\infty\}$, so we have shown that there is a principal S^1 bundle over S^2 such that the total space is S^3 ! This is the *Hopf bundle* (also known as the Hopf fibration).

★

Let us next work out the transition functions of the Hopf bundle for a simple trivializing cover of S^2 . This will allow us to easily identify the Hopf bundle later on. But first we need a good working model of S^2 .

We may coordinatize S^2 as a complex manifold using two charts. Let $[z : w]$ denote the equivalence class $\{(\lambda z, \lambda w) \in \mathbb{C}^2 : \lambda \in \mathbb{C}^\times\}$. Then if we set $N = [0 : 1]$ (the North pole) and $S = [1 : 0]$ (the South pole), we can cover S^2 with the two open sets $X_N = S^2 - S$ and $X_S = S^2 - N$. We get coordinate charts $\phi_{N,S} : X_{N,S} \rightarrow \mathbb{C}$ by

$$\phi_N[z : w] = z/w$$

$$\phi_S[z : w] = w/z$$

so the coordinate transition from north to south is $\phi_{NS}(z) = 1/z$ which is holomorphic on its domain of definition (that is, \mathbb{C}^\times).

Now, let us take X_N and X_S , cross them with S^1 and see if we can find a gluing map g_{NS} which gives us back the Hopf bundle. First, consider a point z in \mathbb{C} , thought of as a point in X_N by use of the ϕ_N chart. If $S^3 \xrightarrow{\pi} S^2$ is the Hopf bundle, then

$$(\phi_N \circ \pi)^{-1}(z) = \frac{(e^{i\theta} z) + e^{i\theta} j}{\sqrt{1 + |z|^2}}$$

We can use this observation to describe a local trivialization of the Hopf bundle over X_N . If we define $\Phi_N : \mathbb{C} \times S^1 \rightarrow S^3$ by

$$\Phi_N(z, e^{i\theta}) = e^{i\theta} \frac{z + j}{\sqrt{1 + |z|^2}}$$

then Φ_N is clearly a smooth diffeomorphism which only omits the points of the form $e^{i\theta} j \in S^3$. Similarly, define

$$\Phi_S(w, e^{i\theta}) = e^{i\theta} \frac{1 + wj}{\sqrt{1 + |w|^2}}$$

Φ_S is again a smooth diffeomorphism from $\mathbb{C} \times S^1$ (identified with $X_S \times S^1$) to S^3 , which only omits the orbit $e^{i\theta} j$. Thus, we have found a trivializing cover for the Hopf bundle $S^3 \xrightarrow{\pi} S^2$.

Finally, we can compute the gluing function g_{NS} for the bundle. This is the map which makes $\Phi_S^{-1} \circ \Phi_N = (\phi_{NS}, g_{NS})$:

$$\Phi_S^{-1}(\Phi_N(z, e^{i\theta})) = (1/z, e^{i\theta} \cdot g_{NS}(z))$$

Proceed by direct computation.

$$\begin{aligned} \Phi_S^{-1}(\Phi_N(z, e^{i\theta})) &= \Phi_S^{-1} \left(e^{i\theta} \frac{z + j}{\sqrt{1 + |z|^2}} \right) \\ &= \Phi_S^{-1} \left(\frac{e^{i\theta} z}{|z|} \cdot \frac{|z| + (1/z)|z|j}{\sqrt{1 + |z|^2}} \right) \\ &= \Phi_S^{-1} \left(\frac{e^{i\theta} z}{|z|} \cdot \frac{1 + (1/z)j}{\sqrt{1 + |1/z|^2}} \right) \\ &= \left(\frac{1}{z}, e^{i\theta} \frac{z}{|z|} \right) \end{aligned}$$

So the gluing map for the Hopf bundle is $g_{NS}(z) = z/|z|$.

The meaning of all of this can be clarified by considering local sections $\psi_{N,S} : X_{N,S} \rightarrow S^1$. These glue together into a global section (that is, they agree on the overlap $X_N \cap X_S$) if and only if $\psi_S = \psi_N \cdot g_{NS}$. Suppose we defined $\psi_N \equiv 1$. Then on the overlap, $\psi_S(re^{it}) = e^{it}$. From the perspective of X_S , a section which was untwisted in X_N will appear to have a twist of $+1$!

Can we find ψ_N and ψ_S which glue together to give a global section ψ ? The previous paragraph proves we cannot: local sections are maps of the form $\psi_i : U \rightarrow S^1$ where U is the open unit disk in the plane. Since U is contractible, ψ_i cannot be twisted along the boundary. That is to say, the map $\psi_i|_{\partial U} : S^1 \rightarrow S^1$ must have winding number zero. But if $\psi_N|_{\partial U}$ has winding number k then $\psi_S|_{\partial U}$ will have winding number $k + 1$, and so if $k = 0$, $\psi_S|_{\partial U}$ will not extend to a local section ψ_S .

Actually, we would be distressed to learn that there was a global section! In general, if $P \xrightarrow{\pi} M$ is a G -bundle with global section ψ , then P is isomorphic to the trivial bundle $M \times G$ by

$$g \cdot \psi(p) \in P \longleftrightarrow (p, g) \in M \times G$$

The Hopf bundle is (clearly?) nontrivial, so it cannot support a global section.

It is very worthwhile to spend some time playing with the group structure of S^3 , the Hopf bundle, different trivializations and so forth. The Hopf bundle will be an important starting in the study of topological charge, so it is beneficial to make the computations your own.

1.4 Associated Bundles

Actually, the principal bundles are more often implicit in another bundle rather than explicitly defined. Suppose that M and X are manifolds, and we wish to make a bundle of X s over M . We may proceed as we did for principal bundles: choose a trivializing cover U_i and a cocycle of gluing maps $g_{ij} : U_i \cap U_j \rightarrow G$ where G is some subgroup of the diffeomorphisms of X . This defines a bundle $B \xrightarrow{\pi} M$ with fiber X . Furthermore, any properties of X which are invariant under G may be extended to the bundle B .

The most common example of this is the vectorbundle. A vectorbundle is a bundle with fiber some fixed vector space V and with gluing maps $g_{ij} \in GL(V)$. Of course, $GL(V)$ preserves the linear structure of V , so B picks up this linear structure fiber-wise. Sections of B can therefore be added pointwise or scalar-multiplied by functions on M . If we supposed additionally that V carried an inner product and $g_{ij} \in O(V) \subset GL(V)$, we would find that the inner product extends to sections of B : the inner product of two such sections would be a function on M .

Another example you already know is the tangent bundle TM to a manifold M . The gluing maps for TM are just the differentials of the coordinate transition maps on

M , and therefore define a unique principal $GL(V)$ -bundle over M , the *frame bundle* \mathcal{FM} . Choosing a local section of \mathcal{FM} is the same as choosing a local ordered basis of vectorfields in TM .

Note that there is something strange about our definition of X -bundles: it never directly uses X itself! Rather, what we have defined is a G -bundle along with a particular representation/action $\rho : G \longrightarrow \text{Diffeo}(X)$. Any such action will define an X -bundle over M , and any X -bundle over M yields a principal bundle. These X -bundles are therefore called *associated bundles*.

1.5 Things to Ponder

1. (*Factor Spaces*) Let G be a Lie group, H a compact Lie subgroup. Then the quotient map $G \xrightarrow{\pi} G/H$ presents G as a principal H -bundle over the coset space G/H . Take G to be the Möbius group of fractional linear transformations acting on $\mathbb{C} \cup \{\infty\}$, and let H the stabilizer of zero in G . Identify H as a group, identify the factor space G/H as a manifold, and find a trivializing cover and gluing maps for G as an H -bundle over G/H . (This example is closely related to the Erlangen program, Klein's answer to "what is geometry?")
2. (*Retractabile Base Spaces*) Suppose that M is a manifold, N a submanifold, and M retracts smoothly onto N . What can you say about principal G -bundles over M and N ? What are the principal G -bundles over the open n -ball U^n ?
3. (*Homotopy of Gluing Maps*) Two principal G -bundles are bundle-isomorphic if there is a diffeomorphism between them which respects the action of G . Suppose that P_1 and P_2 are principal G -bundles over M . Fix an open cover U_i on M which simultaneously trivializes P_1 and P_2 , and let g_{ij}, h_{ij} be the respective gluing maps. Prove that if the cover is fine enough that every U_i retracts to a point, then P_1 and P_2 are bundle-isomorphic if and only if we can find a map $\Phi : M \longrightarrow G$ such that $g_{ij} = \Phi \cdot h_{ij}$.
4. (*G -bundles over n -Spheres*)

2 Calculus with Bundles

2.1 Vanilla Differential Forms

Suppose that M is a smooth manifold, TM the tangent bundle of M , and $\Gamma(TM)$ the space of vectorfields on M . We can also define a *cotangent* bundle TM^* , whose

sections are $C^\infty(M)$ -linear maps

$$\omega : \Gamma(TM) \longrightarrow C^\infty(M)$$

The sections of TM^* are called (*differential*) *1-forms*, and we will usually write $\Omega^1(M)$ for $\Gamma(TM^*)$.

Now suppose that f is a smooth function from M to \mathbb{R} . We can then cook up a 1-form df by

$$df \left(\frac{\partial}{\partial x} \right) = \frac{\partial f}{\partial x}$$

If we write $\Omega^0(M)$ for $C^\infty(M)$, we now have defined a map

$$d : \Omega^0(M) \longrightarrow \Omega^1(M)$$

called the *exterior differential*. d is \mathbb{R} -linear and satisfies the Liebniz equality $d(fg) = df \cdot g + f \cdot dg$ on function f, g .

If we have chosen a local basis of vectorfields $\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}$ then we can define the dual basis dx^1, \dots, dx^n by $dx^i \left(\frac{\partial}{\partial x^j} \right) = \delta_i^j$. In this local basis we can then write

$$df = \sum_i \frac{\partial f}{\partial x^i} dx^i$$

Given a metric, it is clear that df is the same as the operator

$$v \mapsto (\text{grad} f, v)$$

so why don't we just concern ourselves with $\text{grad} f$? The reason is that the definition of $\text{grad} f$ depends on a metric, while the definition of df does not. Therefore, we can always talk about the differential of a function, even in contexts where the gradient does not make sense.

2.2 Higher Forms

Before defining higher-order differential forms, we consider a related problem. Let $T : \mathbb{R} \longrightarrow \text{Hom}(V, W)$ be a smooth path, V and W vectorspaces, and consider the problem of defining T' . T' should itself be a map from \mathbb{R} to $\text{Hom}(V, W)$, and we would expect that if $v : \mathbb{R} \longrightarrow V$ is a path in V then $(Tv)' = T'(v) + T(v')$. This encourages us to define T' by the formula

$$T'(v) = (Tv)' - T(v')$$

or somewhat more abstractly,

$$d_{\text{Hom}(V,W)}T = d_W \circ T - T \circ d_V$$

where d_* is the appropriate derivative on the space $*$. In the special case of $W = \mathbb{R}$, $\text{Hom}(V, W) = V^*$ and our formula reads

$$d_{V^*}\omega = d_{\mathbb{R}} \circ \omega - \omega \circ d_V$$

In practice we will drop the decoration on the d whenever the meaning can be made clear from context.

Given two vectorfields v, w on M , we can define a 2-vector field $v \wedge w$ to be the section

$$v \wedge w = \frac{1}{2}(v \otimes w - w \otimes v) \in \Gamma(TM \otimes TM)$$

Actually, since $v \wedge w = -w \wedge v$, $v \wedge w \in \Gamma(\bigwedge^2 TM)$, the space of antisymmetric $(2, 0)$ tensors on M . $v \wedge w$ represents an infinitesimal bit of plane (a 2-plane field!) on TM spanned by v and w , just as v represents an infinitesimal bit of line in TM . Also note that if $w = \lambda v$ then

$$v \wedge \lambda v = -\lambda v \wedge v = -v \wedge \lambda v$$

so $v \wedge w$ is either 0 or represents a nondegenerate bit of plane.

Similarly, we can define k -vector fields on TM as sections of $\bigwedge^k TM$. If v_i form a local basis of vectorfields in TM then $\bigwedge^k TM$ is the vectorbundle generated locally by expressions of the form

$$\sum_{\sigma \in S_k} (-1)^\sigma v_{\sigma(1)} \otimes \dots \otimes v_{\sigma(k)}$$

where S_k is the group of permutations of k items and $(-1)^\sigma$ is the sign of the permutation σ .

Note that if v_1, \dots, v_k are linearly dependant then $v_1 \wedge \dots \wedge v_k = 0$.

We are now ready to define higher forms, and the definition is clear: a k -form on M is a map $\omega : \bigwedge^k(M) \rightarrow C^\infty(M)$. The vectorbundle of k -forms on M is denoted $\Omega^k(M)$ and is isomorphic to $\bigwedge^k TM^*$. Thus, if dx^i is a local basis of 1-forms then every k -form looks something like a sum of terms like $f \cdot dx^1 \wedge \dots \wedge dx^k$. This form would act on the k -vector $V = v_{I_1} \wedge \dots \wedge v_{I_k}$ by

$$\omega(V) = \frac{1}{k!} \sum_{\sigma \in S_k} (-1)^\sigma \cdot f \cdot dx^{\sigma(1)}(v_{I_1}) \dots dx^{\sigma(k)}(v_{I_k})$$

There is a question of normalization still to be resolved. A convention which I will try to follow is that the wedge product of vectors is

$$v_1 \wedge \dots \wedge v_k = \frac{1}{k!} \sum \dots$$

while a wedge product of forms is

$$\omega^1 \wedge \dots \wedge \omega^k = \sum \dots$$

This means that if x_i is a basis of vectorfields with dual basis dx^i then

$$(dx^1 \wedge \dots \wedge dx^k)(x_1 \wedge \dots \wedge x_k) = 1$$

2.3 The Exterior Algebra

We now set about extending $d : \Omega^0(M) \longrightarrow \Omega^1(M)$ to higher forms. We ask for an operator satisfying the axioms:

1. (*Linearity*): $d : \Omega^k(M) \longrightarrow \Omega^{k+1}(M)$ should be linear.
2. (*Graded Leibniz*): If ω is a k -form then

$$d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta$$

3. (*Generalizes "old" d*): If f is a function (0-form) then

$$df \left(\frac{\partial}{\partial x} \right) = \frac{\partial f}{\partial x}$$

4. (*d is a Coboundary Map*): $d \circ d = d^2 : \Omega^k(M) \longrightarrow \Omega^{k+2}(M)$ is the zero map.

In fact, such an operator does exist and is even unique. All this goes to show we have obtained a complex

$$0 \longrightarrow \Omega^0(M) \xrightarrow{d} \Omega^1(M) \xrightarrow{d} \dots \xrightarrow{d} \Omega^n(M) \longrightarrow 0$$

where $n = \dim M$. The cohomology of this complex is called the *deRham cohomology* of M and is denoted $H_{dR}^\bullet(M)$, The example below shows how it contains important information about the topology of M .

2.4 An Example of deRham Cohomology

Let us do some computations to clarify the use of differential forms. Suppose $M = \mathbb{R}^2$ and we have chosen basis vectors ∂_x, ∂_y and a dual basis dx, dy . $\Omega^0(\mathbb{R}^2) = C^\infty(\mathbb{R}^2)$, $\Omega^1(\mathbb{R}^2) = C^\infty(\mathbb{R}^2) \cdot dx \oplus C^\infty(\mathbb{R}^2) \cdot dy$ and $\Omega^2(\mathbb{R}^2) = C^\infty(\mathbb{R}^2) \cdot dx \wedge dy$. All other Ω^k are 0, so we have the complex

$$0 \longrightarrow C^\infty \longrightarrow C^\infty dx \oplus C^\infty dy \longrightarrow C^\infty dx \wedge dy \longrightarrow 0$$

Let us do some computations with the deRham cohomology. The first deRham group $H_{dR}^0(\mathbb{R}^2)$ is just the kernel of d as it acts on $\Omega^0(M)$. In this case, $df = 0$ if and only if f is constant, so $H_{dR}^0(\mathbb{R}^2) = \mathbb{R}$.

Next, suppose ω is a closed 1-form, so $d\omega = 0$. This means that if $\omega = fdx + gdy$,

$$d\omega = \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} \right) dx \wedge dy = 0$$

Given f and g with $f_y = g_x$ on the plane, we can always find a function h such that $h_x = f$, $h_y = g$. The condition that $f_y = g_x$ simply ensures that the mixed partial derivatives commute with each other, and it can be proven that this is the only condition needed (Frobenius theorem or Poincare lemma, depending on your perspective). In this case, $dh = fdx + gdy = \omega$, so $d\omega = 0$ implies that $\omega = dh$. Therefore, $H_{dR}^1(\mathbb{R}^2) = 0$.

Finally, we consider $d(\Omega^1(\mathbb{R}^2)) \subset \Omega^2(\mathbb{R}^2)$. If $gdx \wedge dy$ is any 2-form and $\partial f/\partial x = g$ then

$$d(fdy) = \frac{\partial f}{\partial x} dx \wedge dy = gdx \wedge dy$$

so $d : \Omega^1 \longrightarrow \Omega^2$ is surjective, and thus $H_{dR}^2(\mathbb{R}^2) = 0$.

Now let us try a different space for comparison. Let $M = \mathbb{R}^2 - 0$ be the punctured plane. Again we find that $H_{dR}^0(M) = \mathbb{R}$ and $H_{dR}^2(M) = 0$, but the computations for $H_{dR}^1(M)$ do not go through so simply. In particular, consider the inappropriately named 1-form

$$d\theta = \frac{xdy - ydx}{x^2 + y^2}$$

is closed, as a direct calculation will show. In fact, on any open set which does not enclose 0 $d\theta$ really is the derivative of the function $\theta(x, y) = \tan^{-1}(y/x)$, but as we go around 0 the function θ picks up a period of 2π until, when we are back at the starting point, we cannot close θ up into a continuous function on the entire plane. Some more computation could convince us that $d\theta$ generates the space of closed 1-forms on M which are not exact, and we can then conclude that $H_{dR}^1(M) = \mathbb{R}$.

For \mathbb{R}^2 , the dimensions were $h_{dR}^\bullet(\mathbb{R}^2) = (1, 0, 0)$, while for $\mathbb{R}^2 - 0$ the dimensions were $h_{dR}^\bullet(\mathbb{R}^2 - 0) = (1, 1, 0)$. How should these dimensions be interpreted?

In general, we can think of the dimension of $H_{dR}^k(M)$ as a count of the number of k -dimensional holes in M — areas of M around which surfaces of dimension k can get stuck. Both \mathbb{R}^2 and $\mathbb{R}^2 - 0$ had $h_{dR}^0 = 1$, meaning that there is one way a 0-dimensional object could get stuck on them (i.e., there is one connected component of each space). Again, both \mathbb{R}^2 and $\mathbb{R}^2 - 0$ had $h_{dR}^2 = 0$, meaning that there is no way for a closed 2-dimensional surface to get stuck in either the plane or the punctured plane. Finally, $h_{dR}^1(\mathbb{R}^2) = 0$ while $h_{dR}^1(\mathbb{R}^2 - 0) = 1$, meaning that there is one fundamental way that a loop can get stuck on $\mathbb{R}^2 - 0$ (namely, wrapping around 0 some number of times) but there is no way such a loop could get stuck on \mathbb{R}^2 .

The deRham cohomology is just one particularly nice way that the algebra of differential forms is intertwined with the topology of the base space. We can expect similarly deep connections in mechanics between the topology of the configuration space and the resulting behavior of particles moving in that space.

2.5 Bundle Valued Forms

For most of our calculations, vanilla forms will not be sufficiently powerful or flexible. We extend their power by introducing vectorbundle valued differential forms:

If V and W are vectorbundles over M then we can form the bundle $V \otimes W$ over M . A local section of $V \otimes W$ is a sum of terms of the form $v \otimes w$ where $v \in \Gamma(V)$, $w \in \Gamma(W)$. Unless otherwise noted, we are tensoring over \mathbb{R} pointwise.

Now if V is a vectorbundle over M we can define the V -valued k -forms $\Omega^k(M; V)$ by

$$\Omega^k(M; V) = \Gamma(V \otimes TM^*)$$

For example, if $f : M \rightarrow N$ then $df : TM \rightarrow TN$ is a section of f^*TN over M , defined by

$$df = \sum_i \frac{\partial f}{\partial x^i} \otimes dx^i$$

where of course $\partial f / \partial x^i$ is a tangent vector on N .

The notion of bundle-valued differential forms lets us define first-order linear differential operators on vectorbundles. Suppose that V is a vectorbundle and we define an operator $\nabla : \Gamma(V) \rightarrow \Omega^1(M; V)$ which satisfies the following properties:

1. (*Linearity*): If c is constant and v, w are sections of V , then $\nabla(cv + w) = c\nabla v + \nabla w$.
2. (*Leibniz Rule*): If f is a function and v a section, $\nabla(fv) = f\nabla v + v \otimes df$

Such a ∇ is called an *affine connection* on V .

Given an affine connection ∇ , we can combine it with d to get a *covariant exterior derivative* $d^\nabla : \Omega^k(M; V) \longrightarrow \Omega^{k+1}(M; V)$ by letting d^∇ treat \otimes like a product for the Leibniz rule:

$$d^\nabla(\psi \otimes \omega) = (\nabla\psi) \wedge \omega + \psi \otimes d\omega$$

We now have tools powerful enough to revisit vector calculus in a unified framework.

2.6 Vector Calculus in Two Symbols

Let us now focus on \mathbb{R}^3 -valued differential forms on \mathbb{R}^3 . Both \mathbb{R}^3 have a differential d , so we know how to differentiate anything in $\Omega^k(\mathbb{R}^3; \mathbb{R}^3)$ (we will now drop the $(\mathbb{R}^3; \mathbb{R}^3)$). We also introduce the useful *Hodge star* operator $\star : \Omega^k \longrightarrow \Omega^{3-k}$, defined by $\omega \wedge \star\omega = dx \wedge dy \wedge dz$.

In order to compare our computations with nineteenth-century vector calculus, let us define the operator \flat which takes the vectorfield (v_1, v_2, v_3) and gives us the 1-form $v_1 dx + v_2 dy + v_3 dz$. Of course, \flat has inverse \sharp . Our previous discussion on d and grad can then be rephrased as saying that if f is a function from \mathbb{R}^3 to \mathbb{R}^3 then $df = (\text{grad}f)^\flat$.

Next, consider a 1-form $\omega = a dx + b dy + c dz$. Taking the exterior derivative gives us

$$d\omega = (b_x - a_y) dx \wedge dy + (c_y - b_z) dy \wedge dz + (a_z - c_x) dz \wedge dx$$

Applying the Hodge star yields

$$\star d\omega = (c_y - b_z) dx + (a_z - c_x) dy + (b_x - a_y) dz$$

and so we have shown that

$$\star d(v^\flat) = (\text{curl } v)^\flat$$

so in fact d also generalizes the curl!

Here we can see where the equation $d^2 = 0$ becomes useful:

$$(\text{curl grad}f)^\flat = \star d((\text{grad}f)^\flat) = \star ddf = 0$$

so $d^2 = 0$ has encoded the fact that the curl of a gradient is zero.

How about applying d and \star in the opposite order? Then we have

$$\star\omega = a dy \wedge dz + b dz \wedge dx + c dx \wedge dy$$

and

$$d\star\omega = (a_x + b_y + c_z) dx \wedge dy \wedge dz$$

This proves that $\star d \star (v^\flat) = (\operatorname{div} v)^\flat$, and we have recovered the three “derivative” operations from vector calculus!

Can we also recover the div curl relationship?

$$(\operatorname{div} \operatorname{curl} v)^\flat = \star d \star \star d (v^\flat) = 0$$

where we used the fact that $\star\star = \pm 1$, depending only on the degree of the forms on which \star is acting.

The operator $\star d \star$ which we have used above is sometimes called the *codifferential* and denoted δ or d^\dagger . It essentially computes the source of a form, as opposed to d which calculates the derivative. $\delta : \Omega^k \rightarrow \Omega^{k-1}$ and also satisfies $\delta^2 = 0$. Combining the two lets us define the *Laplace-Beltrami operator* $\Delta : \Omega^k \rightarrow \Omega^k$ by $\Delta = d\delta + \delta d$. In the special case $\Delta : \Omega^0 \rightarrow \Omega^0$ we find that

$$\Delta f = (d\delta + \delta d)f = \delta df = \pm \sum_i \frac{\partial^2 f}{\partial x_i^2}$$

so Δ is an extension of the regular Laplacian to differential forms. The differential and codifferential on a compact manifold work together to again give a relation between the analytic structure of the manifold and its topology via Hodge theory. In particular, you can use Hodge theory to show that each deRham cohomology class has a unique harmonic representative, a trick which we may have use for later.

2.7 The Forms Picture of Electromagnetism

As a final grand application of our machinery of differential forms, we set out to see how Maxwell’s equations look in terms of differential forms. We define the star operator on Minkowski spacetime in the expected way: $\omega \wedge \star\omega = dt \wedge dx \wedge dy \wedge dz$. Let us also define the electromagnetic field 2-form

$$F = \vec{E}^\flat \wedge dt + \star_3(\vec{B}^\flat)$$

Here \vec{E} and \vec{B} are the electric and magnetic fields, and \star_3 is the Hodge star acting only on space (the \star_3 operator would of course change if we changed what we meant by “straight ahead through time”). Thus, $\star_3 dx = dy \wedge dz$ while $\star dx = -dt \wedge dy \wedge dz$. We define $E = \vec{E}^\flat$ and $B = \star_3(\vec{B}^\flat)$ to simplify notation, so $F = E \wedge dt + B$ and more explicitly,

$$\begin{aligned} E &= E^1 dx + E^2 dy + E^3 dz \\ B &= B^1 dy \wedge dz + B^2 dz \wedge dx + B^3 dx \wedge dy \end{aligned}$$

Let us proceed by randomly applying operators which we know about to see what we find, and later on try to make sense of things. First, we can take the derivative of F to find a 3-form which is obviously of interest:

$$\begin{aligned}
dF &= d(E^1 dx \wedge dt + E^2 dy \wedge dt + E^3 dz \wedge dt + B^1 dy \wedge dz + B^2 dz \wedge dx + B^3 dx \wedge dy) \\
&= (B_x^1 + B_y^2 + B_z^3) dx \wedge dy \wedge dz \\
&\quad + (B_t^1 - E_z^2 + E_y^3) dt \wedge dy \wedge dz \\
&\quad + (B_t^2 - E_x^3 + E_z^1) dt \wedge dz \wedge dx \\
&\quad + (B_t^3 - E_y^1 + E_x^2) dt \wedge dx \wedge dy
\end{aligned}$$

Note that the field 2-form F is closed ($dF = 0$) if and only if the two vector equations

$$\begin{aligned}
\nabla \cdot \vec{B} &= 0 \\
\nabla \times \vec{E} + \frac{\partial \vec{B}}{\partial t} &= 0
\end{aligned}$$

hold, and these are exactly the first two Maxwell equations!

Emboldened, we might next try computing the source of F , $\star d \star F$:

$$\star F = E^1 dz \wedge dy + E^2 dx \wedge dz + E^3 dy \wedge dx + B^1 dt \wedge dx + B^2 dt \wedge dy + B^3 dt \wedge dz$$

so we find that

$$\begin{aligned}
d \star F &= d(E^1 dz \wedge dy + E^2 dx \wedge dz + E^3 dy \wedge dx + B^1 dt \wedge dx + B^2 dt \wedge dy + B^3 dt \wedge dz) \\
&= (-E_x^1 - E_y^2 - E_z^3) dx \wedge dy \wedge dz \\
&\quad + (-E_t^1 + B_z^2 - B_y^3) dt \wedge dy \wedge dz \\
&\quad + (-E_t^2 + B_x^3 - B_z^1) dt \wedge dz \wedge dx \\
&\quad + (-E_t^3 + B_y^1 - B_x^2) dt \wedge dx \wedge dy
\end{aligned}$$

This is of course implies that

$$\star d \star F = (-\nabla \cdot \vec{E}) dt + (\nabla \times \vec{B} - \frac{\partial \vec{E}}{\partial t}) \lrcorner$$

If \vec{j} is the electric current density and ρ is the electric charge density we can define the current $J = \vec{j} \lrcorner - \rho dt$ and the second pair of Maxwell's equations

$$\nabla \cdot \vec{E} = \rho$$

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

boil down to the one equation

$$\delta F = J$$

Note that we have not said anything special about J , but the identity $d^2 = 0$ (in the form $\delta^2 = 0$) gives us more information about the behavior of this current: since $J = \delta F$, we know that $\delta J = 0$. A direct computation shows that

$$-\delta J = \frac{d\rho}{dt} + \nabla \cdot \vec{j}$$

so J must satisfy the classical *equation of continuity*

$$\frac{d\rho}{dt} = -\nabla \cdot \vec{j}$$

which states that electric charge can only change by having current flow into or out of a region. In other words, electric charge is conserved (and locally conserved, as required by Einstein's equivalence principle).

We have therefore shown that the electromagnetic field can be modelled as a 2-form F on spacetime such that

$$dF = 0$$

$$\delta F = J$$

for some 1-form J . When we revisit these equations, we will try to answer the question “why aren't Maxwell's equations symmetric?” or the equivalent question “why doesn't a magnetic current appear in the equations?”

2.8 The Vector Potential, Semi-Classically

Suppose that F is an electromagnetic field with $\delta F = J$. We know that $dF = 0$ so, if we assume F is defined on a contractible region of spacetime, $F = dA$ for some 1-form A , the *vector potential*.

3 Connections, Curvature and Gauge Fields

3.1 The Maurer-Cartan Form

Suppose that G is a Lie group with Lie algebra \mathfrak{g} . \mathfrak{g} can be identified with the set of left G -invariant vectorfields on G , and is a vectorspace of the same dimension as

G . Furthermore, \mathfrak{g} is equipped with a bilinear, anticommutative and nonassociative product

$$[,] : \mathfrak{g} \otimes \mathfrak{g} \longrightarrow \mathfrak{g}$$

called the *Lie bracket*, defined as the commutator of two vectorfields in \mathfrak{g} .

There is also a unique \mathfrak{g} -valued 1-form $\Theta \in \Omega^1(\mathfrak{g})$ which acts as the identity on the space of left-invariant vectorfields. In the case of a matrix Lie group with coordinate g , this form is given by the equation

$$\Theta = g^{-1}dg$$

Θ is called the *Maurer-Cartan* form of the Lie group.

The exterior derivative of the Maurer-Cartan form may be calculated easily: since

$$d(g^{-1}) = -g^{-1}dgg^{-1}$$

we find that

$$d\Theta = d(g^{-1}dg) = -g^{-1}dgg^{-1} \wedge dg = -\Theta \wedge \Theta$$

so the form Θ satisfies the so-called *Maurer-Cartan equation*

$$d\Theta + \Theta \wedge \Theta = 0$$

Note that the wedge product here is achieved by wedging over the matrix product. In the case of an abstract Lie group, this formula reads

$$d\Theta + \frac{1}{2}[\Theta, \Theta]$$

where

$$[\alpha, \beta](X, Y) = [\alpha(X), \beta(Y)] - [\alpha(Y), \beta(X)]$$

Since the exterior derivative and the wedge product are natural with respect to pullbacks, for any map $f : M \longrightarrow G$ from a manifold into G we can take the “Darboux derivative” $f^*\Theta \in \Omega^1(M; f^*\mathfrak{g})$ and see that it satisfies the Maurer-Cartan equation. In fact, if M is contractible this is the full integrability condition: If $\omega \in \Omega^1(\mathfrak{g})$ is a Lie-algebra valued 1-form on M which satisfies the Maurer-Cartan equation, then locally $\omega = f^*\Theta$ for some $f : M \longrightarrow G$.

The Maurer-Cartan equation is essentially just the integrability condition for a first-order exponential PDE as opposed to a first-order “independent” PDE: $\omega = \omega_1 dx^1 + \dots + \omega_n dx^n$ gives a PDE

$$\begin{aligned} \frac{\partial f}{\partial x^1} &= \omega_1 \\ &\vdots \\ \frac{\partial f}{\partial x^n} &= \omega_n \end{aligned}$$

and this PDE has local solutions if and only if $d\omega = 0$. Similarly, ω could represent the “exponential” PDE

$$\begin{aligned} \frac{\partial f}{\partial x^1} &= f \cdot \omega_1 \\ &\vdots \\ \frac{\partial f}{\partial x^n} &= f \cdot \omega_n \end{aligned}$$

and *this* PDE has local solutions if and only if $d\omega + \omega \wedge \omega = 0$. This exponential PDE picture will give us critical insight into the seemingly strange form of Ehresmann connections and curvature of principal bundles.

3.2 Path-Lifting Procedures on Principal Bundles

Suppose that $\gamma : \mathbb{R} \rightarrow M$ is a path on M , P a principal G -bundle over M , and we wish to find a lift of γ into P . Suppose we also have chosen a local section ψ . We want the path-lifting procedure to be entirely local: the lift of γ near p should depend only on the germ of γ around p . Thus, the path-lifting procedure will be given by a system of PDE on M . By using ψ as a local coordinate, we actually get a system of PDE on M mapping to G : there is some \mathfrak{g} -valued 1-form ω such that if $\Gamma = g \cdot \psi$ is the lift of γ ,

$$g'(t) = g \cdot -\omega_{\gamma(t)}(\gamma'(t))$$

Why does this differential equation appear in exponential form? Because we also want it to be invariant under global “phase changes”: if we replace the initial condition g_0 with $h \cdot g_0$, then the new solution \hat{g} should be related to the old solution g by $\hat{g} = h \cdot g$.

We have therefore constructed a “connection” operator $\nabla = d + \omega$ such that $g \cdot \Psi$ is a lift of γ if and only if

$$dg + g \cdot \omega(\gamma') = 0$$

This is the differential equation for our path-lifting procedure in the local frame (“local gauge”) ψ . But what if we change gauge, replacing ψ with $\phi = h^{-1}\psi$?

In this case, g is replaced by gh and the differential equation becomes

$$d(gh) + gh \cdot \omega_h = 0$$

We proceed by direct computation to solve for ω_h :

$$\begin{aligned} 0 = d(gh) + gh \cdot \omega_h &= dg \cdot h + g \cdot dh + g \cdot h\omega_h \\ &= g \cdot -\omega \cdot h + g \cdot dh h^{-1} \cdot h + gh\omega_h h^{-1} \cdot h \end{aligned}$$

Or after rearranging,

$$\omega_h = h^{-1}\omega h - h^{-1}dh = Ad(h^{-1})(\omega) - h^*\Theta$$

This is the condition which tells us when two connections written in different gauges are “the same”.

In particular, this tells us how to glue together locally defined connections into a global connection. Suppose that ψ_1, ψ_2 are local gauges on U_1, U_2 and $U_1 \cap U_2$ is nonempty. Furthermore, let ω_1, ω_2 be connection forms on U_1, U_2 (of course, in the respective gauges) and let g_{12} be the transition function

$$g_{12}\psi_1 = \psi_2$$

on $U_1 \cap U_2$. Then ω_1, ω_2 agree on the overlap if and only if

$$\omega_2 = Ad(g_{12}^{-1})(\omega_1) - g_{12}^*\Theta$$

There is a fundamental result due to Ehresmann that if ω_1, ω_2 agree on the overlap then there is a connection form ω defined on all of $U_1 \cup U_2$ such that ω_i are the pullbacks of ω by ψ_i . Thus, if we have a covering of M which trivializes P and a bunch of ω_i which are all compatible then we can find a globally defined \mathfrak{g} -valued 1-form ω which defines a global connection on P .

As an example, we consider the principal S^1 -bundle over S^2 from the previous section: the Hopf bundle. We already saw that this bundle trivializes over the north and south subsets of S^2 , and on the overlap the transition function is given by $g_N S(re^{i\theta}) = e^{i\theta}$. A connection form for this bundle locally looks like an $Im\mathbb{C}$ -valued 1-form, and if ω_N, ω_S are two such connection forms then they agree if and only if on the overlap, $\omega_S = \omega_N - id\theta$.

Let ϕ, ψ be the standard coordinates on the sphere, so

$$\phi, \psi \longleftrightarrow (\cos \phi \cdot \cos \psi, \sin \phi \cdot \cos \psi, \sin \psi)$$

We would like to define something like $\omega_S = -\frac{i}{2}d\theta$ and $\omega_N = \frac{i}{2}d\theta$, but of course $d\theta$ is not continuous near the poles. So we might add a correction term:

$$\omega_N = \frac{i}{2}(1 - \sin \psi)d\theta$$

$$\omega_S = \frac{i}{2}(-1 - \sin \psi)d\theta$$

These are each continuous on all but one pole, and the union of their domains covers S^2 . Furthermore, on the equator they satisfy the compatibility condition and therefore glue together into a connection on the Hopf bundle.

3.3 Curvature on Abelian Principal Bundles

If P is a principal G -bundle and G is abelian, then we have already seen that the equations for a connection simplify somewhat. In particular, Ad acts as the identity so $\omega_2 = \omega_1 - g_{12}^* \Theta$ is the abelian compatibility condition. What happens if we take the exterior derivative of a connection form to get a \mathfrak{g} -valued 2-form?

In this abelian case,

$$d\omega_2 = d\omega_1 - d(g_{12}^* \Theta) = d\omega_1 + g_{12}^*(\Theta \wedge \Theta) = d\omega_1$$

since when α and β commute in the Lie algebra, $\alpha \wedge \beta = 0$. So given a connection on an abelian principal bundle, we have obtained a 2-form $R = d\omega$ which takes values in the Lie algebra \mathfrak{g} . Note that $dR = dd\omega = 0$, so R is always a closed form! R measures the lack of integrability in ω , which is to say R is a measure of how non-commuting ω -flows are. Because of this, R is called the *curvature form* of ω .

Let us compute the curvature for our connection on the Hopf bundle. First, note that the area 2-form on S^2 is given by

$$A_{S^2} = -\cos \psi d\psi \wedge d\phi$$

So we compute:

$$\begin{aligned} d\omega_N &= -\frac{i}{2} \cos \psi d\psi \wedge d\phi = \frac{i}{2} A_{S^2} \\ d\omega_S &= -\frac{i}{2} \cos \psi d\psi \wedge d\phi = \frac{i}{2} A_{S^2} \end{aligned}$$

We can then compute the *total curvature* of our connection on the Hopf bundle:

$$\int_{S^2} R = \frac{i}{2} \int_{S^2} A_{S^2} = 2\pi i$$

What happens if we compute the total curvature with a different connection? In the abelian case where we have been working, the answer is fairly straightforward. Suppose that ω and η are both connection forms on the same principal bundle P with abelian structure group, and let R^ω, R^η be the corresponding curvatures. Then $R^\omega = d\omega, R^\eta = d\eta$ so

$$\frac{1}{2\pi i} \int_X (R^\omega - R^\eta) = \frac{1}{2\pi i} \int_X d(\omega - \eta) = 0$$

where X is a closed submanifold (actually, an element of $H_2(M; \mathbb{Z})$). Chern was the first to realize that the normalizing factor of $\frac{1}{2\pi i}$ causes the total curvature to actually be an integer(!), so we have shown that there is some cohomology class

$$c_1(P) = \left[\frac{1}{2\pi i} R^\omega \right] \in H^2(M; \mathbb{Z})$$

called the *first Chern class* of P . We can calculate $c_1(P)$ with any connection we choose, but our previous calculation shows that the choice doesn't matter. In particular, the integrality of the first Chern class gives us a bijection between principal S^1 -bundles over S^2 and elements of \mathbb{Z} .

3.4 Dirac and Magnetic Monopoles

Suppose that F is an electromagnetic field 2-form with source $\delta F = J$. Quantum mechanics gives us a somewhat strange formula involving F when we look at the phase of a particle after dragging it around a loop (of course, we can't look at the *actual* phase of the particle, but by interfering it with reference particles we can determine a relative phase). If Ψ is a particle responding to the electromagnetic field with charge q and we drag Ψ around a spacetime loop γ then a basic formula of quantum mechanics tells us that the phase of Ψ is multiplied by a factor of

$$\exp \left\{ \frac{iq}{\hbar} \int_D F \right\}$$

where D is, somewhat mysteriously, any disk in spacetime bounded by γ . The Maxwell equation $dF = 0$ guarantees that this formula is well-defined on homotopy classes since if D is homotopic to D' then $D - D' = \partial X$ for some region X and

$$\int_D F - \int_{D'} F = \int_{\partial X} F = \int_X dF = 0$$

But what if we suppose that $dF = m dx \wedge dy \wedge dz$ — that is, what if we propose that there is a magnetic monopole of strength m ? Then if we send Ψ along a loop around the equator, we can compute the phase change in two different ways and get an answer of

$$e^{\frac{iqm}{2\hbar}}$$

if we use the “north” hemispherical disk, versus

$$e^{-\frac{iqm}{2\hbar}}$$

if we use the south hemisphere (the orientation change is responsible for the sign change). Therefore if quantum mechanics has the proper formula for phase changes *and* if we propose that a single magnetic monopole of strength m exists, we see that

$$e^{\frac{iqm}{2\hbar}} = e^{-\frac{iqm}{2\hbar}}$$

so for some integer N ,

$$iqm = 2\pi i N \hbar$$

or, writing $h = 2\pi\hbar$,

$$qm = Nh$$

In other words, if a single magnetic monopole exists *anywhere in the universe*, then electric charge is quantized! This is Dirac’s quantization argument, and it remains a very compelling reason to leave the possibility of magnetic monopoles existing open to debate. Unfortunately, this monopole which we tried to model doesn’t satisfy Maxwell’s equation $dF = 0$, so it certainly isn’t an object which fits in to the classical framework of electromagnetism.

3.5 Parallel Transport in Quantum Mechanics

Before we consider another method for building monopoles, let us revisit the strange phase-shift equation. Why should a particle traversing a path care about the electromagnetic field inside of the path? In fact, if there is high spacetime curvature in some region (for example, a particle in orbit around a black hole), it might be the case that the particle travels around the loop in less time than it would take a photon to travel from the loop to the center of the disk. In other words, we are proposing that a particle is using information about the electromagnetic field *nonlocally*. Since locality is one of the only large-scale unifying principles of physics, we can suspect that a different mechanism is at work here.

Actually, we already have the mathematical tools onhand to understand this “paradox”. Suppose we wanted to model a particle responding to the electromagnetic field classically: this would give us some state space M . For a particle in Minkowski space $\mathbb{R}^{3,1}$, this is just $M = T\mathbb{R}^{3,1} \times \mathbb{R}$: position, momentum and charge. When dealing with phase shifts, we then presume that sitting “over” each point of M is a copy of the unit circle $S^1 = U(1)$, and the classical dynamics simply ignore the S^1 . These S^1 may or may not have a globally nontrivial topology: in other words, the electromagnetic configuration space is some principal S^1 -bundle $P \rightarrow M$. The early assumption in quantum mechanics was that $P = M \times S^1$. But one would be hard-pressed to motivate this assumption physically!

Now, as a particle moves from point to point in the configuration space M , its phase must also change in P . What does this mean? It exactly means that we are searching for a path-lifting procedure (“connection”) on $P \rightarrow M$ which tells us how a particle’s phase will respond to being dragged along some path. If we write $\omega \in \Omega^1(M; Im\mathbb{C})$ for this connection form, then we see that a particle, upon going around a loop γ in M , undergoes a phase shift equal to

$$\exp \left\{ \int_{\gamma} \omega \right\}$$

(of course, we could make this phase shift proportional to the charge by choosing an appropriate connection).

Where has the old formula for phase shift gone? Well, if D is a disk which bounds γ then the phase shift is given by

$$\exp \left\{ \int_{\gamma} \omega \right\} = \exp \left\{ \int_D d\omega \right\} = \exp \left\{ \int_D R \right\}$$

Where have the i , q and \hbar gone? We already noted that the q can be subsumed into ω , and we can choose units with $\hbar = 1$. And the i is part of ω now, which is a pure imaginary 1-form.

Furthermore, since R is the curvature of ω , it automatically satisfies the equation $dR = 0$. In other words, if we cleverly choose a connection so that $\delta d\omega = J$, then ω gives us back all of Maxwell's equations, where the curvature of ω has replaced the electromagnetic field! Note the added benefit: the previously mysterious phase-shift formula is now a direct consequence of our bundle model of the electromagnetic field.

3.6 The Hopf Bundle as a Monopole

Recall that, once we have chosen a reference frame, the $dx \wedge dy \wedge dz$ part of dF is given by $\nabla \cdot \vec{B}$: in other words, if S is a spacelike sphere then $\int_S F = \text{Area}(S) \cdot m$, where m is the total amount of "magnetic charge" inside the ball bounded by S . Of course, if F is a classical electromagnetic field, $dF = 0$ implies that this magnetic charge is zero (if B is the ball inside of S , then $\int_S F = \int_B dF = 0$).

Now, suppose that our electromagnetic configuration space $P \rightarrow M$ is a non-trivial principal S^1 -bundle. In particular, let us look at the case where P is the Hopf bundle. We already saw that the curvature R is given by $\frac{i}{2}A_{S^2}$. We can compute the source of this electromagnetic field easily:

$$\delta R = \star d \star \frac{i}{2}A_{S^2} = \star d \frac{i}{2} = 0$$

so R is a free electromagnetic field. Of course, this R still satisfies $dR = 0$, so is a perfectly well-defined Maxwell electromagnetic field. But what happens if we integrate R around S^2 ? As we already saw, this gives the total curvature:

$$\frac{1}{\text{Area}(S^2)} \int_{S^2} R = \frac{i}{8\pi} \int_{S^2} A_{S^2} = \frac{4\pi i}{8\pi} = i \cdot \frac{1}{2}$$

In other words, the Hopf bundle supports an otherwise classical, free electromagnetic field which nonetheless behaves like a monopole of strength $m = \frac{1}{2}$!

3.7 The Yang-Mills Functional

The previous sections hopefully provide strong evidence that electromagnetism can be modeled by a S^1 gauge field. But if we want to fully reconstruct electromagnetism, we also should provide a dynamic principle to govern the evolution of fields and particles over time. The easiest way to proceed is by searching for a Lagrangian which generalizes the electromagnetism Lagrangian to this connection-on-a- S^1 -bundle context.

The only constraint on our connection form ω was that it satisfied the equality $\delta d\omega = J$ where J is the electromagnetic current. Let us first deal with the free field case, $\delta d\omega = 0$. If you have seen some Hodge theory before, you already know a variational principle which gives these solutions: ω satisfies $\delta d\omega = 0$ exactly at the critical points of the functional

$$\int_{\gamma} \langle d\omega \wedge \star d\omega \rangle$$

where $\langle \alpha \wedge \star \beta \rangle = \langle \alpha, \beta \rangle \text{Vol}$. This is exactly like the Dirichlet energy,

$$DE(f) = \int \langle df \wedge \star df \rangle$$

and the Euler-Lagrange equation is fairly easy to calculate: if $\epsilon\eta$ is a variation of ω and we only look at terms on the order of ϵ we get

$$\begin{aligned} \int_{\gamma} \langle (d\omega + \epsilon d\eta) \wedge (\star d\omega + \epsilon \star d\eta) \rangle &= \epsilon \int \langle d\omega \wedge \star d\eta \rangle + \langle d\eta \wedge \star d\omega \rangle \\ &= 2\epsilon \int \langle d\eta \wedge \star d\omega \rangle \\ &= 2\epsilon \left(\int d(\eta \wedge \star d\omega) - \eta \wedge d\star d\omega \right) \\ &= -2\epsilon \int \langle \eta \wedge d\star d\omega \rangle \\ &= 2\epsilon \int \langle \eta \wedge \star(\star d\star d\omega) \rangle \\ &= \epsilon \int \langle \eta \wedge \star(2\delta d\omega) \rangle \end{aligned}$$

By the standard variational argument, this proves that ω is a critical point of the action if and only if it satisfies the equation $2\delta d\omega = 0$. Note that the path γ is not on M or P , but is actually a path in the space of connections on P ! Actually, the

situation is even more subtle: γ is a path in the space of connections on P , modulo gauge equivalence. This space is of fundamental importance in the related fields of Donaldson theory and Seiberg-Witten theory, and we will probably return to it later.

Now, suppose that ψ is a path in M and $g \cdot \Psi$ a lift of ψ into P . We think of Ψ as describing the evolution of a particle of unit charge which is responding to the electromagnetic field. The current J is then given by $J = \nabla \Psi = (d + \omega)g = d^\omega g$, the charge-moment of $g \cdot \Psi$ through P . J and $\delta d\omega$ are both \mathfrak{g} -valued 1-forms, so it is reasonable to write down a variational principle of the form

$$\frac{1}{2} \int |R^\omega|^2 - |d^\omega g|^2$$

Let us compute the dynamics of the connection form ω by replacing it with $\omega + \epsilon\eta$. First, note that $R^{\omega+\epsilon\eta} = R^\omega + \epsilon d\eta$ and $d^{\omega+\epsilon\eta}g = d^\omega g + \epsilon g \cdot \eta$. So the order- ϵ terms of the action are

$$\int \langle \eta, \delta d\omega \rangle - \langle \eta, d^\omega g \rangle$$

so ω is critical for this action if and only if $\delta d\omega = d^\omega g$. In other words, writing $J = d^\omega g$ as above, ω is critical exactly when

$$\delta R^\omega = J !$$

Since $dR^\omega = dd\omega = 0$, this variational principle gives us back all of Maxwell's equations with very little choice. In this framework of connections on principal bundles, electromagnetism appears to fall out almost naturally!

This action which we have been discussing is a quasi-classical variant of the *Yang-Mills action* for a principal S^1 -bundle. The real Yang-Mills action deals with particles as wavefunctions instead of points moving along a path, but the essential form is exactly the same.