

Geometric Concepts in Physics

Paul Duffell

Based on lectures by Brian Greene

Illustrations by Andrew Macfarlane

Table of Contents

	Introduction	3
1.	Review of Manifold Concepts	6
1.1	The Exponential Map	6
1.2	Pushforward and Pullback	10
1.3	The Lie Derivative	16
2.	Group Theory	22
2.1	Groups	22
2.2	Extracting Groups out of Groups	27
2.3	Morphisms	28
2.4	Quotient Groups	32
2.5	Group Action on a Set	33
2.6	Fields and Algebras	34
3.	Lie Groups	36
3.1	Matrix Groups	37
3.2	The Tangent Space of a Lie Group	38
3.3	From the Group to the Algebra and Back Again	41
3.4	The Rest of the Lie Algebra Story	47
3.5	Realizations and Representations	49
3.6	Summary and Application	53
4.	Fiber Bundles	55
4.1	Product Manifolds: A Visual Picture	55
4.2	Fiber Bundles: The Informal Description	56
4.3	Fiber Bundles: The Bloated Description	59
4.4	Special Types of Fiber Bundles	60

4.5	Fiber Bundles: The More Elegant Description	64
4.6	Curvature: Generalizing Past the Tangent Bundle	65
4.7	Parallel Transport on General Fiber Bundles	69
4.8	Connections on Associated Vector Bundles	73

Introduction

This text is intended for a student who has already studied some geometry. It will be assumed that the reader is familiar with topology, manifolds, tangent spaces, tensors, differential forms, metrics, connections, and curvature. If it is useful to the interested reader, I may eventually supplement this work with background material, but this will most likely be unnecessary, as there are a great many sources which treat these basic foundational concepts.

Einstein's general theory of relativity was an incredibly beautiful discovery. The notion that gravity can be described without forces, purely geometrically, is as compelling a concept today as it was in the 1920's. Even more impressive is the fact that this simple observation of a beautiful mathematical description of a physical phenomenon lead to a theory of gravity which was more precise and accurate than any theory ever considered, and remains so to this day. The construction of this theory didn't require any physical observation; Einstein was just trying to look at gravity in a more mathematically appealing way. In doing so, he improved the theory in ways that went far beyond this elegant aesthetic.

It is in this spirit that we seek to study Einstein's final accomplishment: a classical unified theory. We shall see how it is possible to characterize every force (not just gravity) in an entirely geometric language. The mathematical machinery we'll need will not be trivial, but the payoff is certainly worth the effort. We'll begin by reviewing a few useful concepts which may or may not be second-nature to you already. As stated before, you should already know all about manifolds, but you may not have a very strong background in the theory of

groups, so I also include a decent introduction to this subject (or perhaps just a refresher). We will then plunge deep into the subject of Lie groups, followed by fiber bundles, and finally reviewing the ideas of connections and curvature (in which you should be fluent), generalizing these concepts as far as they will go. When we are done, it will be possible to write down a geometric physical theory which describes all of the forces of the universe. At some point, a section will be included which gives some treatment of the physical applications, but currently this part will have to be found in supplementary material. This physical theory is the big payoff, and I don't wish to rob you of it, so I tentatively label this work a "first edition", with improvements to come.

The problems provided are generally not very challenging, but are meant to keep the reader involved in the material. I personally find that a page full of equations is as aesthetically displeasing as it is unenlightening, and whenever the opportunity arises to trade a block of equations in exchange for a guided exercise for the reader, I feel it's a trade worth making. In later editions, I may add more challenging exercises to the end of each section, but as of now I don't feel the necessity.

As stated before, this is a work in progress. I take pride in my attention to detail, but this does not mean I'm always right. Contact me, if you find any problem with my clarity, consistency of notation, or even grammar. Good writing requires good feedback, especially in the case of technical writing. In any case, I'd like to avoid too long an introduction, so let us get right to the mathematics; I hope your experience with these concepts is as enjoyable as mine has been.

Review of Manifold Concepts

There are some preliminary odds and ends that we need to cover before leaping into lie groups. Specifically, we'll need to clearly understand the exponential map, the pushforward and pullback maps, and finally the lie derivative. These concepts will be used again and again, and so a fairly thorough treatment is given, in the hopes that your understanding becomes as complete as possible.

1.1 The Exponential Map

This is a map which provides a concrete relationship between the tangent space of a manifold, and the manifold itself. There are many ways to conceptually approach the exponential map, so a couple of different approaches are given.

From the Infinitesimal to the Macroscopic

Often in mathematics (and very often in physics), one deals with the effect on a function $f(\lambda)$ when we change λ by a small parameter, ϵ . Formally, we can expand $f(\lambda + \epsilon)$ in a Taylor series about λ :

$$f(\lambda + \epsilon) = f(\lambda) + \epsilon \frac{df}{d\lambda} + \frac{1}{2} \epsilon^2 \frac{d^2 f}{d\lambda^2} + \dots \quad (1.1)$$

If ϵ is infinitesimally small, we can ignore terms of order greater than ϵ^2 and the result is simply:

$$f(\lambda + \epsilon) = \left(1 + \epsilon \frac{d}{d\lambda}\right) f(\lambda) \quad (1.2)$$

It is possible to regard this as an operator, $T(\epsilon)$, acting on $f(\lambda)$, whose operation is to translate

λ by an infinitesimal amount, ϵ .

$$f(\lambda + \epsilon) = T(\epsilon)f(\lambda)$$

What if we want a more general operator, $T(\Delta\lambda)$, whose operation translates λ by a finite amount, $\Delta\lambda$? We could easily read off this operator from the fully-expanded power series above, but it is more instructive to think of this finite-translation operator as a product of a large amount of successive infinitesimal-translation operators:

$$T(\Delta\lambda) = [T(\epsilon)]^N, \text{ where } \Delta\lambda = N \cdot \epsilon \quad (1.3)$$

We then take the limit as $N \rightarrow \infty$. In other words,

$$T(\Delta\lambda) = \lim_{N \rightarrow \infty} \left[1 + \frac{\Delta\lambda}{N} \frac{d}{d\lambda} \right]^N \quad (1.4)$$

Does this formula look familiar yet? Let us pretend for the time being that the operator $(\Delta\lambda \frac{d}{d\lambda})$ is just a number, k . Then the formula looks like:

$$T(\Delta\lambda) = \lim_{N \rightarrow \infty} \left[1 + \frac{k}{N} \right]^N \quad (1.5)$$

This you should recognize as a definition of the exponential function, e^k . We carry the notation over to describe the translation operator:

$$T(\Delta\lambda) = e^{\Delta\lambda \frac{d}{d\lambda}} \quad (1.6)$$

This is typically evaluated by expanding the exponential in a power series. It is straightforward to check that its action on a function gives the full Taylor series expansion of that function.

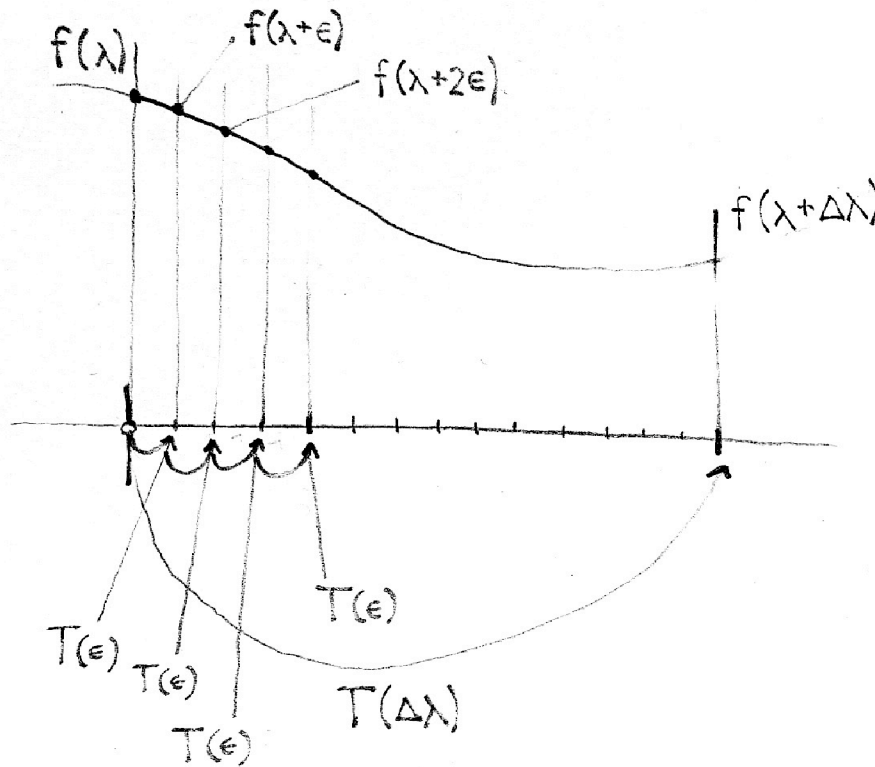


Figure 1.1 The translation of a function, found by a large number of infinitesimal translations.

Problem 1.1 Check that the exponential map's action on a function gives the Taylor series expansion of the function; that is, check that it appropriately translates $f(\lambda)$ by a distance $\Delta\lambda$.

Now, we make a small conceptual transition. Instead of thinking of this as an operator-valued function with respect to the interval, $\Delta\lambda$, it is more natural to think of it as an operator-valued function on derivatives, $d/dt = \Delta\lambda d/d\lambda$. In other words, we can vary the translation distance by varying our reparameterization, $t = t(\lambda)$. So, we symbolically write this as

$$T\left(\frac{d}{dt}\right) = e^{d/dt} \quad (1.7)$$

This is a more natural form of our translation operator, the exponential map. You can think of this as generating a translation, T , given a derivative, d/dt .

From Vector Fields to Integral Curves

On an arbitrary differential manifold M , imagine a smoothly varying family of curves $\Phi(p, \lambda)$, covering the manifold (or at least filling some open set in the manifold) without intersecting. Such a family of curves is known as a congruence of curves. For example, imagine the lines of latitude or longitude on the sphere (minus the poles). At each point p in this open set in the manifold, there is exactly one curve passing through p . As we know, such a curve is always associated with a particular vector V_p in the tangent space, corresponding to the velocity of the curve $V_p \in T_p M$. Since we can do this at every point p in the region in question, this determines a smoothly varying vector field $V(p)$.

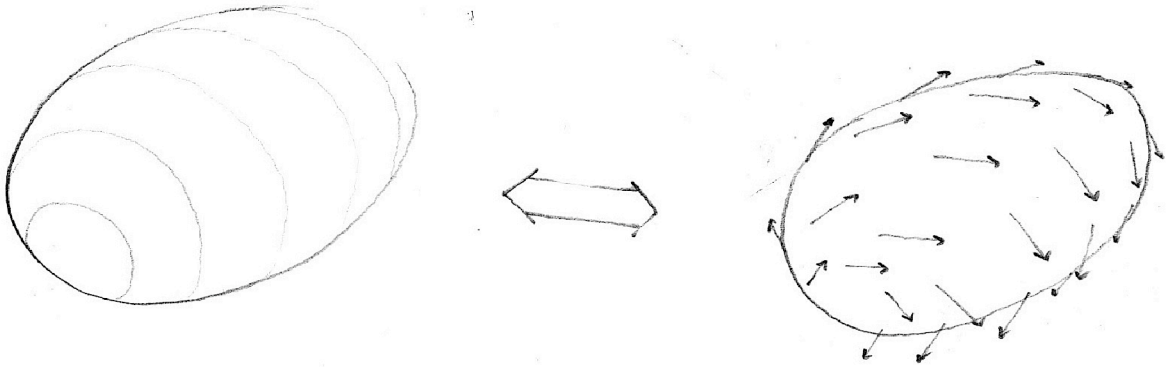


Figure 1.2 A congruence of curves defines a vector field, and vice versa.

We can go the other direction, too. On a differential manifold M , a smooth vector field $V(p)$ determines a smoothly varying family of curves $\Phi_p: \mathbb{R} \rightarrow M$, called the integral curves of V . To help guide your physical intuition, you might regard this set of curves as describing the “flow” of the vector field. In a given coordinate system $\{x^i\}$,

$$\Phi_p(\lambda) = (x_p^1(\lambda), x_p^2(\lambda), \dots, x_p^n(\lambda)) \quad (1.8)$$

This family of curves provides a map from the tangent space to the manifold, which we call the exponential map. The curves Φ_p are determined by demanding that the velocity of each curve $d\Phi_p/d\lambda$ is equal to the vector field evaluated at that point, V_p . This demand can be represented in a coordinate-dependent manner:

$$\frac{dx^i}{d\lambda} = V^i(x_p^1(\lambda), x_p^2(\lambda), \dots, x_p^n(\lambda)) \quad (1.9)$$

This is a set of coupled first-order ordinary differential equations for $x^i(\lambda)$. There always exists a unique solution about a sufficiently small neighborhood of p . Note that this requirement implies that the directional derivative $d/d\lambda = V^i \partial/\partial x^i$, i.e. that the curve parameter λ appears in the directional derivative associated with the vector field V . We can find the solution to these differential equations by expanding in a Taylor series in $\Delta\lambda$:

$$x^i(\lambda_0 + \Delta\lambda) = x^i(\lambda_0) + \Delta\lambda \frac{dx^i}{d\lambda} + \dots = \left[1 + \Delta\lambda \frac{d}{d\lambda} + \frac{1}{2} \Delta\lambda^2 \frac{d^2}{d\lambda^2} + \dots \right]_{\lambda_0} x^i \quad (1.10)$$

$$x^i(\lambda_0 + \Delta\lambda) = e^{\Delta\lambda d/d\lambda} x^i$$

As before, we notice that $\Delta\lambda d/d\lambda = \Delta\lambda V$ is a vector by itself. That is, instead of thinking of this as a map which inputs a vector V and gives us a curve, then inputs a distance $\Delta\lambda$ and moves us this distance along the curve to produce a point in M , we can just think of this as a map which inputs vectors $W = (\Delta\lambda V)$ and outputs the point on our manifold by moving a unit distance along its integral curve. This is nothing more than repackaging, but we can cut through all the unnecessary notation by just setting $\Delta\lambda$ equal to unity:

$$x^i(\lambda_0 + 1) = e^{d/d\lambda} x^i \quad (1.11)$$

This is the exponential map of $d/d\lambda$ acting on x^i . We could be more explicit by expressing $d/d\lambda$ as $V^i \partial/\partial x^i$:

$$x^i(\lambda_0 + 1) = \exp\left(V^k \frac{\partial}{\partial x^k}\right) x^i \quad (1.12)$$

This expression may seem strange-looking, as we are taking partial derivatives with respect to x^k of x^i , which we expect to just give us a Kronecker delta, but don't forget that V^k is also dependent on the $\{x^i\}$. Thus, the expansion of this formula should look like:

$$x^i(\lambda_0 + 1) = \left[x^i(\lambda_0) + V^i + \frac{1}{2} V^k \frac{\partial V^i}{\partial x^k} + \frac{1}{6} V^k \frac{\partial}{\partial x^k} \left(V^l \frac{\partial V^i}{\partial x^l} \right) + \dots \right] \quad (1.13)$$

This formula was only guaranteed to work in a small neighborhood of p (meaning we cannot justify setting $\Delta\lambda$ equal to unity the way we did), but we can get around this by restricting the domain, i.e. requiring that our vector fields be small enough to keep within some neighborhood of p in M . Moreover, we can often find solutions which cover a large portion of the manifold. For example, if we choose a coordinate vector field $\partial/\partial x^i$, then the integral curves produced are simply the coordinate curves,

$$x^{i \neq k} = \text{constant}, \quad x^k = \text{constant} + \lambda \quad (1.14)$$

This exponential map will be well-defined as far as the coordinate chart reaches, which may be nearly the entire manifold (For example, S^2 can be covered minus one point, by stereographic projection). For this reason, the exponential map is often viewed as a map from the local structure of $T_p M$ to the more global structure of M itself.

1.2 Pushforward and Pullback

The pushforward and pullback maps are natural ways of mapping objects like vectors and forms from one differential manifold to another, given a map $\Phi: M_1 \rightarrow M_2$ between points on the manifolds. The concepts here are somewhat abstract, but they can be concretely represented in computational form when we look at specific coordinate systems on the manifolds.

The Pullback of a Function

We start with two manifolds, M_1 and M_2 , of dimension n_1 and n_2 , respectively. Additionally, we have a smooth map $\Phi: M_1 \rightarrow M_2$. This map need only be well-defined and smooth; it does not have to be a one-to-one map, nor does it have to map onto the entire manifold M_2 . Now, let's look at the space of smooth functions $f(q)$ on points $\{q\}$ in M_2 . Is there a natural way to map these to the set of functions $g(p)$ on points $\{p\}$ in M_1 ? That is, given a function $f(q)$, can we use the map Φ to naturally produce a function $g(p)$ defined on M_1 ?

The answer is much simpler than the question. Simply note that the function $f(\Phi(p))$ is a function well-defined on M_1 . In other words, we can use the original map Φ from the points in M_1 to the points in M_2 to produce a new map which we call Φ^* , the pullback map from functions on M_2 to functions on M_1 , by the formula:

$$\Phi^* f(p) = f(\Phi(p)) \quad (1.15)$$

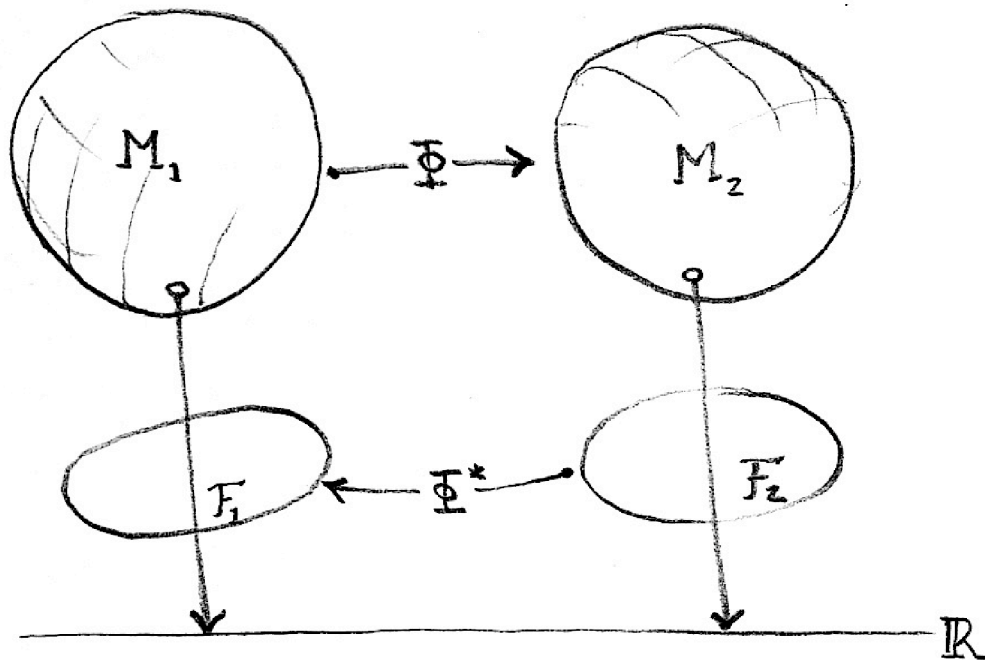


Figure 1.3 The pullback map on functions.

This is clearly well-defined; given any smooth function $f(q)$ on M_2 , we can always use this formula to produce a function $g(p) = f(\Phi(p))$ on M_1 . The pullback map is not generally one-to-one, nor does it always map onto the entire space of smooth functions on M_1 . In other words, we cannot generally “push forward” functions. In the special case that Φ is invertible, so is Φ^* ; we can push functions forward simply by using the pullback of the inverse map.

Problem 1.2 Show that when Φ is injective, Φ^* is surjective, and when Φ is surjective, Φ^* is injective. Thus, when Φ is bijective, so is Φ^* .

To summarize, the pullback of a function defined on M_2 is simply the function's representation on points in M_1 which get mapped to points in M_2 by the map Φ .

The Pushforward of a Vector

We now abstract the concept further by considering the space of tangent vectors at a given point p in the manifold M_1 , i.e. the tangent space $T_p M_1$. We know that the space tangent vectors at this point can be considered the space of directional derivatives on smooth functions, evaluated at p . In other words, vectors are smooth maps acting on functions on M_1 . The space of smooth functions on M_1 is itself a manifold, as is the space of smooth functions on M_2 . Let's call these manifolds F_1 and F_2 . We have already constructed a map between these spaces; this is just the pullback map:

$$\Phi^* : F_2 \rightarrow F_1 \quad (1.16)$$

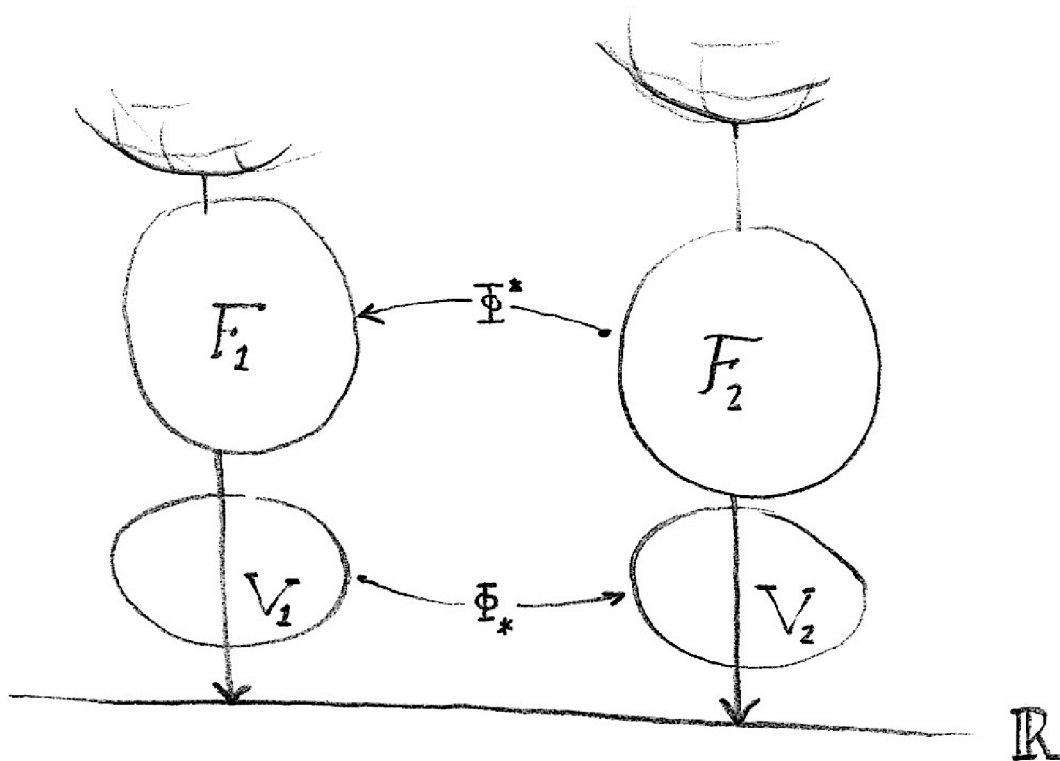


Figure 1.4 The pushforward map on vectors.

Since vectors in $T_p M_1$ are themselves linear maps on F_1 , we should be able to find a natural map to vectors on F_2 , by pulling back again. We have to check that this map indeed produces a directional derivative, and not just some general map, but for now accept that it will.

What does this map look like? Given a vector V in $T_p M_1$, this is a directional derivative map on functions on M_1 , given by the formula:

$$V[g(p)] = V^i \frac{\partial g}{\partial x^i} \Big|_p \quad (1.17)$$

where $\{x^i\}$ is a local coordinate system in the vicinity of p on M_1 . Now use the pullback map on functions to get a vector acting on functions in M_2 . This is what we will call the pushforward map, Φ_* :

$$\Phi_* V[f(q)] = V[\Phi^* f(p)] = V[f(\Phi(p))] \quad (1.18)$$

Let's get our head straight about things. V is a map on functions on M_1 . $f(q)$ is a function on M_2 . Thus, V can't act directly on f . $\Phi_* V$ is a vector field defined on M_2 , which can act on functions $f(q)$. $\Phi^* f$ is a function on M_1 , which can be acted on by V .

Computation

We can just use the chain rule to evaluate this:

$$\Phi_* V[f(q)] = V[f(\Phi(p))] = V^i \frac{\partial f}{\partial x^i} = V^i \left(\frac{\partial \Phi^k}{\partial x^i} \right) \frac{\partial f}{\partial \Phi^k} \quad (1.19)$$

$$= V^i \left(\frac{\partial y^k}{\partial x^i} \right) \frac{\partial f}{\partial y^k} \quad (1.20)$$

Here, we are also using a local coordinate chart on M_2 , $\{y^k\}$. We have simplified the notation by expressing $\Phi(p)$ in both of the coordinate representations as $y^k(x^i)$. The last expression is clearly a directional derivative, thus this genuinely pushes vectors from $T_p M_1$ to $T_{\Phi(p)} M_2$ (not just to the space of general maps acting on functions in M_2).

We now have a natural map from $T_p M_1$ to $T_{\Phi(p)} M_2$. We can realize this map as a matrix $A_{ik} = \partial y^k / \partial x^i$ acting on vectors in $T_p M_1$. A_{ik} is an $n_2 \times n_1$ matrix, mapping the n_1 components of V to the n_2 components of $\Phi_* V$. Note that A_{ik} is coordinate-dependent. Note also that we can push vectors forward, but we cannot pull them back, unless Φ has an inverse. In other words, A_{ik} is not generally an invertible matrix (it might not even have the same number of rows as columns).

Another way to see how the pushforward map acts on vectors is to look at another natural manifestation of the tangent vector space $T_p M_1$: velocities of curves passing through p . Since Φ maps points in M_1 to points in M_2 , it naturally maps curves in M_1 to curves in M_2 , hence it can map velocities from one manifold to the other. It is easy to show that this leads to the same pushforward map that we defined on directional derivatives. The velocity of the mapped curve is given by the same chain rule as above. This is a nice way of picturing

the pushforward map on vectors, as it requires no computation to properly visualize.

Problem 1.3 Show that the pushforward map defined as the naturally induced map from velocities of curves in M_1 to velocities of curves in M_2 gives the same chain rule as above.

The Pullback of a One-Form

Just when you thought we couldn't take the definitions any deeper, we are about to define yet another pullback map. We know that one-forms are linear maps on vector fields, and therefore we can define the pullback of a one-form to be the one-form's action on pushed-forward vector fields:

$$\Phi^* \omega(V) = \omega(\Phi_* V) \quad (1.21)$$

Again, let's make sure we know what we're looking at. ω is a one-form in M_2 ; that is, it is a linear map on vectors in M_2 . V is a vector in M_1 . ω cannot act directly on V , because they live on different manifolds. $\Phi^* \omega$ is a one-form in M_1 , which acts on vectors V in M_1 . The action of $\Phi^* \omega$ on V in M_1 is dictated by ω 's action on the vector pushed forward to M_2 . That is exactly what we've written down. We can now evaluate this in some choice of coordinates:

$$\Phi^* \omega(V) = \omega(A_{ik} V^i \frac{\partial}{\partial x^k}) \quad (1.22)$$

$$= V^i A_{ik} \omega_k \quad (1.23)$$

$$= (A^T \omega)_i V^i \quad (1.24)$$

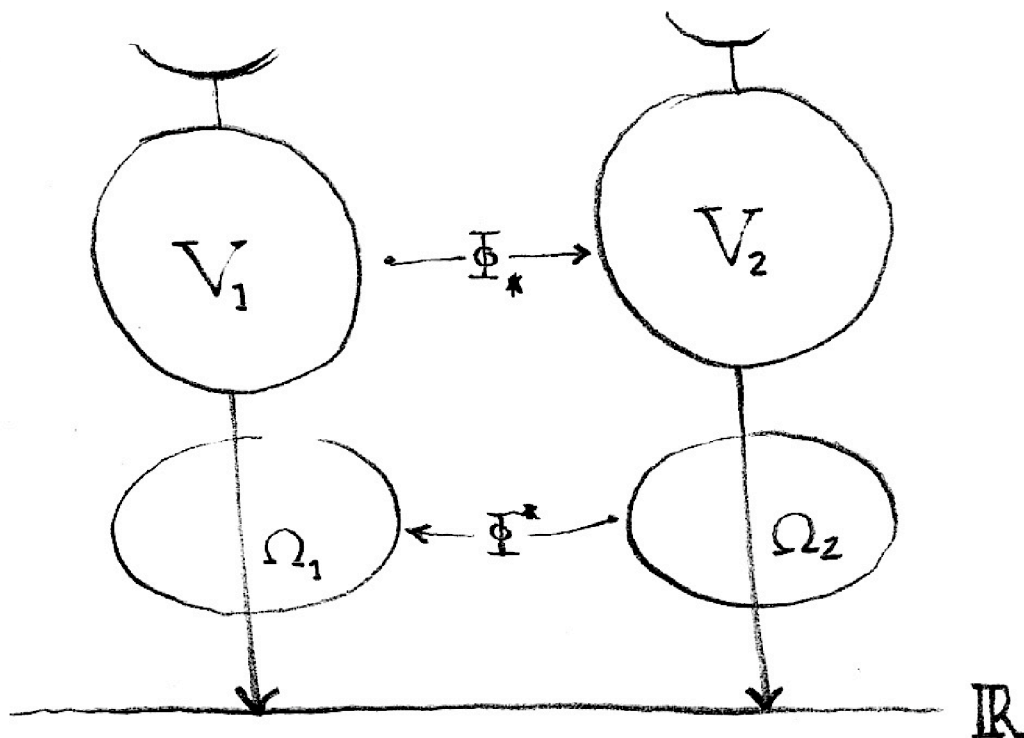


Figure 1.5 The pullback map on one-forms.

The pullback of a one-form is given by the action of the transposed matrix A^T acting

on its components. We might have expected this, especially if we note that A is an $n_2 \times n_1$ matrix, while A^T is an $n_1 \times n_2$ matrix, which is exactly what we'd need to send the n_2 components of ω to an n_1 -component one-form defined on M_1 .

In general, we can pull back any $(0,k)$ tensor, and push forward any $(k,0)$ tensor; the generalization being to simply multiply by additional copies of the matrix A or A^T . However, we cannot do the reverse, nor can we push or pull any (l,m) tensor, for nonzero l and m , unless Φ is invertible. In the case where Φ is invertible, it is possible to push or pull tensors of any rank, essentially by multiplying by inverse matrices, where applicable. Obviously there are some details to be filled in, but this is the basic idea.

Problem 1.4 Derive the pushforward and pullback formulas for tensors of arbitrary rank, given an invertible map Φ between the manifolds, with a corresponding pushforward matrix A .

A good example shows up in the coordinate charts on a differential manifold, M . Notice that any coordinate chart (usually also labeled Φ , which might normally confuse us, but in this case they are referring to the same function) is an invertible map between an open set in M and an open set in $\mathbb{R}^n$. Therefore, we can pull functions on open sets in M back to functions on open sets in $\mathbb{R}^n$. This is how we define things like continuity, differentiability, and in the case of complex manifolds, analyticity. We pull functions back to $\mathbb{R}^n$, and evaluate these properties in a more concrete setting.

We can also push vectors forward from $\mathbb{R}^n$ to T_pM , using a coordinate chart in a neighborhood of p . This is essentially what we are doing when we determine the components of a vector V in a given coordinate system, given by Φ . Different coordinate charts will generally give us different pushforward maps, which in turn give us different components for V .

The Special Case $M_1 = M_2$

Set $M_1 = M_2 = M$. In other words, look at bijective maps $\Phi: M \rightarrow M$ from a manifold onto itself. In particular, look at a smoothly varying family of such maps, $\{\Phi_{pq}\}$, where Φ_{pq} is a smooth, bijective map from M onto itself, which maps p to q . An example of this is the family of rotations on the manifold of S^1 , the circle. For any two points p and q on the circle, there is a unique rotation Φ_{pq} which sends p to q .

We can define Φ -invariant vector fields on M to be vector fields which are invariant under the pushforward map $\Phi_{pq*}: T_pM \rightarrow T_qM$. Thus, $V(q) = \Phi_{pq*}V(p)$. The set of Φ -invariant vector fields forms a vector space, since any two invariant vector fields can be added together to find another invariant vector field, due to Φ 's linearity. It is easily seen that this vector space is equivalent to T_pM at any point p in the manifold, since we can push

any tangent vector V_p forward from p to every point q in the manifold, using Φ_{pq} , producing a manifestly invariant vector field: $V(q) = \Phi_{pq}^* V_p$. Since we can do this for any vector V at p , and since Φ_{pq} is bijective (by assumption here), we can output a unique invariant vector field $V(q)$ given any vector V_p at p . Going back the other direction is even easier. Given an invariant vector field, we can get a unique vector V_p in $T_p M$, simply by evaluating the vector field at p : $V_p = V(p)$. Notice that both of these identifications are linear.

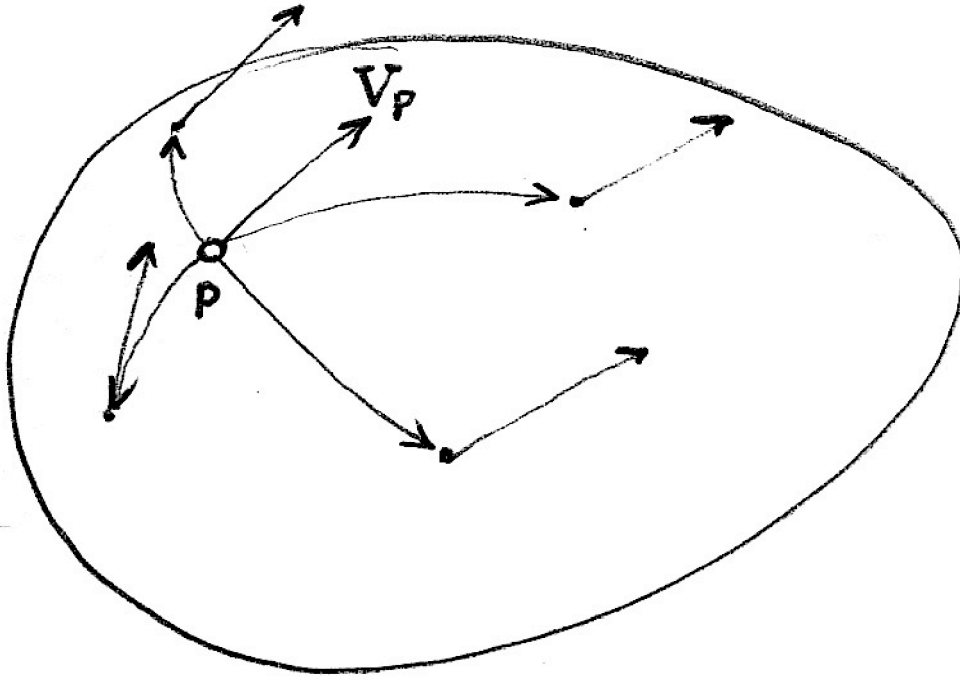


Figure 1.6 Φ -invariant vector fields can be found by pushing vectors forward from a single given point to everywhere else on the manifold.

Let's make this more concrete with the example of rotations on the circle. Given a point p on the circle, and a vector V_p at p , we should be able to produce a Φ -invariant vector field for all $\{q\}$ on the circle. In this simple case, the vector space is one-dimensional, so

$$V_p = V \frac{\partial}{\partial \theta} \Big|_p \quad (1.25)$$

and the vector field is just constant:

$$V(q) = V \frac{\partial}{\partial \theta} \Big|_q \quad (1.26)$$

In higher-dimensional examples, the Φ -invariant vector fields are less trivial.

We have shown that there is a natural linear correspondence between $T_p M$ and the set of Φ -invariant fields on the manifold M . This is useful because it provides yet another manifestation of the tangent space, whenever the manifold possesses such a smooth family of maps Φ_{pq} . This can be a powerful manifestation, as it provides a connection between the tangent space at a single point and the entire global structure of the manifold.

Problem 1.5 For the manifold $\mathbb{R}^n$, the set of n -dimensional translations provides such a smooth family of maps. Show that Φ -invariant vector fields in this case are given simply by constant vector fields.

Problem 1.6 Consider the manifold given by $\mathbb{R}^2 \setminus \{0\}$, the plane minus the origin. Consider mapping this space to itself, mapping a point p to another point q by the following algorithm: rotate the plane by the angle p and q make with the origin. Then scale the plane by the ratio of the distances from q and p (respectively) to the origin. First, show that this always gives a smooth map from the plane to itself which sends p to q . Then, pick two linearly independent vectors at any point, and find their corresponding Φ -invariant vector fields.

1.3 The Lie Derivative

A lie derivative provides a way of taking derivatives of vector fields on a manifold. Computationally, it can be quite simple, though conceptually it's actually very subtle. On differential manifolds, it will be useful to be able to take derivatives of vector fields, but until we provide additional information, the concept of a “derivative” of a vector field is not well-defined.

Comparing Tangent Vectors

We want to think of the derivative as “the rate at which a vector field changes as we move in some direction” in the manifold. This implies we are comparing tangent vectors defined at different points p and q in M . There is no canonical way to do this, since tangent vectors at p live in the tangent space T_pM , and vectors at q live in a different space, T_qM . There are many ways of defining maps between these two spaces, but there is no special or natural map. Choosing a particular map between tangent spaces imposes additional structure, and this is usually fleshed out in the form of a connection, which gives us the covariant derivative. The lie derivative provides an alternative method for differentiating vector fields, which does not require a connection. Instead, the additional information specified to compare tangent vectors will be a congruence of curves.

Congruence of Curves

We mentioned this concept earlier, but it's worth restating. A congruence of curves defined on some neighborhood in a manifold M is a smoothly varying family of curves which fill this neighborhood, without intersecting. Each point in this neighborhood lies on exactly one such curve.

The key concept which allows the lie derivative to function conceptually is the relationship between vector fields and congruences of curves. Given a congruence of curves, we can find a vector field by taking at each point the velocity of the integral curve passing through that point. We can also get a congruence of curves from a vector field,

using the exponential map defined at the beginning of this chapter.

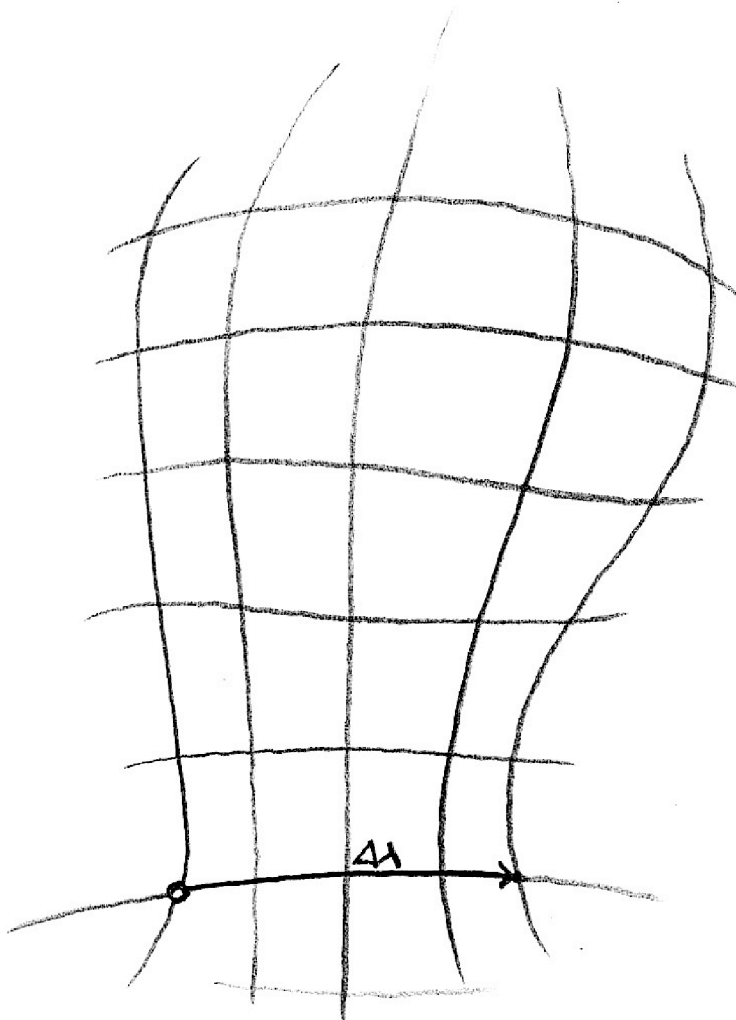


Figure 1.7 The lie derivative measures how the integral curves of one vector field change as we move along the integral curves of another vector field.

Let us take our vector field W that we want to differentiate, and transform it into a congruence of curves using the exponential map. It doesn't really matter the size of the neighborhood on which we define this congruence of curves, since the derivative only cares what's happening locally. As we know, W is also associated with a directional derivative operator at this point, which we will call $\partial/\partial\mu$. The integral curve of W can then be parameterized by the parameter μ . We can write an integral curve of W passing through p as $\alpha_p(\mu)$. Now, as stated previously, we need to provide an additional congruence of curves in order to differentiate W . Equivalently, we could provide a vector field V , since we know we can go back and forth between these two objects. We shall write V as $\partial/\partial\lambda$, and the integral curves of V will be parameterized by λ .

We now have two congruences of curves. How do we take the derivative of one with

respect to the other? Conceptually, we want to look at how the curves of W change when we move a small distance $\Delta\lambda$ along curves of V . We still have to deal with the issue of comparing vectors at different points in the manifold; we've just transformed the problem into comparing *curves* at different points in the manifold. Fortunately, the congruence of curves given by V gives us a natural way of transporting a curve of W to different points on the manifold. We can define a *new* congruence of curves about the point p in the following manner:

Lie Dragging

At p , look at the integral curve of W passing through p . Call this curve $\alpha_\lambda(\mu)$. Move a distance $\Delta\lambda$ along the curve of V passing through p . Call this new point q . To produce a transported curve $\tilde{\alpha}_{\lambda+\Delta\lambda}(\mu)$, simply transport each point in the μ -curve $\alpha_\lambda(\mu)$ this same distance $\Delta\lambda$ along the integral curve of V passing through $\alpha_\lambda(\mu)$. This produces a new curve which we can compare with $\alpha_{\lambda+\Delta\lambda}(\mu)$ simply by taking the difference between their velocities:

$$\Delta_v[W] = \frac{\partial \alpha(\mu)}{\partial \mu} \Big|_{\lambda+\Delta\lambda} - \frac{\partial (\tilde{\alpha}_{\lambda+\Delta\lambda}(\mu))}{\partial \mu} \quad (1.27)$$

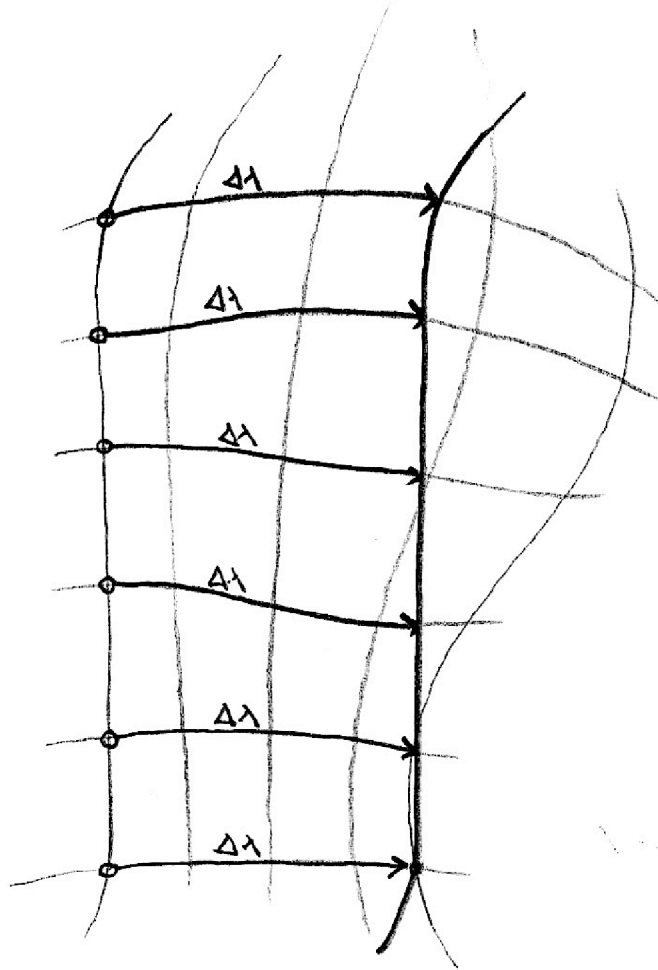


Figure 1.8 Lie Dragging one of the vector fields.

A note about notation: we've introduced several concepts at once, and it's good to keep our head on straight about why things are written down the way they are. $\alpha_\lambda(\mu)$ is an integral curve of W passing through p with parameter μ . We write the subscript “ λ ” instead of “ p ” to accentuate the fact that p is given as a point on an integral curve of V , parameterized by λ . $\alpha_{\lambda+\Delta\lambda}(\mu)$ is simply another integral curve of W , this one instead passing through q , the point found by moving a distance $\Delta\lambda$ along an integral curve of V . $\tilde{\alpha}_{\lambda+\Delta\lambda}(\mu)$ is *not* an integral curve. It is the curve found by transporting $\alpha_\lambda(\mu)$ a distance $\Delta\lambda$ along integral curves of V passing through each μ of $\alpha_\lambda(\mu)$. The family of curves produced in this manner is said to be lie dragged. $\alpha_{\lambda+\Delta\lambda}(\mu)$ and $\tilde{\alpha}_{\lambda+\Delta\lambda}(\mu)$ intersect each other at q , which is why we can compare their velocities.

It is important to understand why the velocities of α and $\tilde{\alpha}$ are written the way they are. Since $\alpha_{\lambda+\Delta\lambda}(\mu)$ is an integral curve of W , the velocity of this curve is exactly what we mean by W_q , the vector field evaluated at q . This is found by taking the derivative with respect to μ of integral curves at arbitrary λ , then evaluating it specifically at the point q , which corresponds to $\lambda + \Delta\lambda$. For $\tilde{\alpha}$, we must first transport the curve *before* computing its velocity. This is noted symbolically by putting the subscript $\lambda + \Delta\lambda$ inside the parentheses. This will become important shortly.

Computation of the Lie Derivative

We can turn this difference into a derivative by dividing by $\Delta\lambda$ and taking the limit as $\Delta\lambda$ vanishes. This specifies the lie derivative:

$$\mathcal{L}_V[W] = \lim_{\Delta\lambda \rightarrow 0} \frac{1}{\Delta\lambda} \left[\frac{\partial \alpha(\mu)}{\partial \mu} \Big|_{\lambda+\Delta\lambda} - \frac{\partial \tilde{\alpha}_{\lambda+\Delta\lambda}}{\partial \mu} \right] \quad (1.28)$$

The simplest way to compute this is to expand these terms to first order in a taylor series in $\Delta\lambda$, since higher-order terms will vanish in the limit $\Delta\lambda \rightarrow 0$. This is an exercise left to the reader. Eventually, we arrive at the following:

$$\mathcal{L}_V[W] = \left(\frac{\partial}{\partial \lambda} \frac{\partial}{\partial \mu} - \frac{\partial}{\partial \mu} \frac{\partial}{\partial \lambda} \right) \alpha_\lambda(\mu) \quad (1.29)$$

Problem 1.7 Expand equation (1.28) in a first-order taylor series, deriving equation (1.29).

If we express this result in a particular coordinate system, we find that the second derivatives cancel:

$$\mathcal{L}_V[W]^i = \left[V^j \frac{\partial W^k}{\partial x^j} - W^j \frac{\partial V^k}{\partial x^j} \right] \frac{\partial \alpha^i_\lambda(\mu)}{\partial x^k} \quad (1.30)$$

We've been interpreting this as the velocity of a curve, α , but we can now think of this as a directional derivative operator acting on the coordinate function $\alpha^i_\lambda(\mu) = x^i_\lambda(\mu)$. The derivative will just give us a kronecker delta, giving the following result:

$$\mathcal{L}_V[W]^i = V^j \frac{\partial W^i}{\partial x^j} - W^j \frac{\partial V^i}{\partial x^j} \quad (1.31)$$

We usually express this as the commutator $[V,W]$, the result of commuting directional derivative operators V and W . Written this way, it is often simply called the lie bracket of V with W .

Problem 1.8 Given the following vector fields,

$$V = \frac{\partial}{\partial x}, W = \frac{\partial}{\partial \phi} \quad (1.32)$$

where ϕ is the standard axial coordinate for the x - y plane, find the lie derivative of W with respect to V .

Lie Derivatives of Other Tensors

We don't have to stop here; we can now take the lie derivative of a one-form. We first must fix our lie derivative with two reasonable requirements. First, the lie derivative of a scalar is simply a directional derivative:

$$\mathcal{L}_V[f] = \frac{\partial f}{\partial \lambda} \quad (1.33)$$

Then we note that a scalar function can be formed by operating with a one-form on a vector field:

$$\omega(W) = \omega_i W^i \quad (1.34)$$

Then we finally require that our derivative satisfies a leibnitz rule,

$$\mathcal{L}_V[\omega_i W^i] = \mathcal{L}_V[\omega_i] W^i + \omega_i \mathcal{L}_V[W^i] \quad (1.35)$$

We can then use these equations to find the lie derivative of a one-form:

$$\mathcal{L}_V[\omega_i] = V^j \frac{\partial \omega_i}{\partial x^j} + \omega_j \frac{\partial V^j}{\partial x^i} \quad (1.36)$$

Problem 1.9 Derive equation (1.36) from equations (1.33) and (1.35), by choosing W^i to be a coordinate basis vector field $W^i = \delta^i_j$.

Problem 1.10 Find equations for the lie derivative of a (2,0) tensor, a (1,1) tensor, and a (0,2) tensor.

In a similar fashion, we can compute the lie derivative of tensors of arbitrary rank. Generally, the lie derivative is most useful in its rank-(1,0) interpretation, the change in the congruence of curves as described above. In this case, it is also simpler computationally, as it is just given by the lie bracket $[V,W]$.

Coordinate Bases

Now that we have a new way of comparing tangent spaces, how can we make use of it? A very common use appears when we look at basis vectors that we want to use in different tangent spaces. Say we choose a set of basis vectors at each tangent space in a neighborhood of a point p , and this set of basis vectors varies smoothly in the manifold. Under what conditions can we find a coordinate chart in a neighborhood of p whose

coordinate basis vectors correspond to our choice of basis vectors? In other words, given a set of basis vectors $\{e_i\}$, can we find a coordinate system $\{x^i\}$ whose partial derivatives $\{\partial/\partial x^i\}$ are the associated directional derivative operators of $\{e_i\}$?

The necessary and sufficient condition for this to be possible is that the lie brackets of all the basis vectors vanish:

$$[e_\mu(p), e_\nu(p)] = 0, \text{ for all } (\mu, \nu) \quad (1.37)$$

Computationally, it's easy to see why this is a necessary condition. If it is possible to write $\{e_i\}$ as a set of partial derivatives $\{\partial_i\}$, then:

$$[e_\mu(p), e_\nu(p)] = [\partial_\mu, \partial_\nu] = 0 \quad (1.38)$$

because partial derivatives commute (when acting on smooth functions). So, if the lie bracket does not vanish, clearly we cannot write the vectors in terms of partial derivatives of a given coordinate system. However, if the lie bracket *does* vanish, how do we know we can always find an appropriate coordinate system?

It should be clear that such a coordinate system can always be found via the integral curves of $\{e_i\}$. The reason this works (and fails when the lie bracket does not vanish) is that the integral curves agree when we lie-drag them in the way we did before, when calculating $\mathcal{L}_V[W]$. Since the lie derivative is zero, this means that

$$\alpha(\mu)|_{\lambda+\Delta\lambda} = \tilde{\alpha}_{\lambda+\Delta\lambda}(\mu) \quad (1.39)$$

This ensures that our coordinate system is not ambiguous (when we move a parameter distance $\Delta\mu$ along one coordinate, then a distance $\Delta\lambda$ along another, we get the same result as if we reverse the order).

How the Lie Derivative Differs from the Covariant Derivative

If we are given a vector field V , we specify the lie derivative, $\mathcal{L}_V$. If we are given a connection, Γ , we specify the covariant derivative, ∇ . You might now be tempted to ask, is there a relationship between V and Γ ? That is, given a vector field, V , can we produce a connection, Γ , such that $\mathcal{L}_V = \nabla$?

The simple answer is no. The lie derivative and the covariant derivative are simply two different beasts. One way of understanding this is to note that the lie derivative is a map from (p,q) tensors to (p,q) tensors, and the covariant derivative is a map from (p,q) tensors to $(p,q+1)$ tensors. The “equation” $\mathcal{L}_V = \nabla$ simply makes no sense. It is possible to write down some relationships between the two, but it is really best to think of them as different objects which live in different spaces.