Diffeomorphisms and Lie derivative

Maps between manifolds

Definition. Let M, N be differentiable manifolds of dimension m, n respectively. A function $\phi : M \to N$ is *smooth* if, and only if, $\psi_A \circ \phi \circ \psi_\alpha^{-1}$ is smooth for all charts ψ_α of M and all charts ψ_A of N (note that this is a map from a subset of $\mathbb{R}^m$ to a subset of $\mathbb{R}^n$).

If we have such a map then we can "pull-back" a function on N to define a function on M:

Definition. Let $\phi : M \to N$ and $f : N \to \mathbb{R}$ be smooth functions. The *pull-back* of f by ϕ is the function $\phi^*(f) : M \to \mathbb{R}$ defined by $\phi^*(f) = f \circ \phi$, i.e., $\phi^*(f)(p) = f(\phi(p))$.

Furthermore, ϕ allows us to "push-forward" a curve λ in M to a curve $\phi \circ \lambda$ in N. Hence we can push-forward vectors from M to N (Figs. 7.1, 7.2)

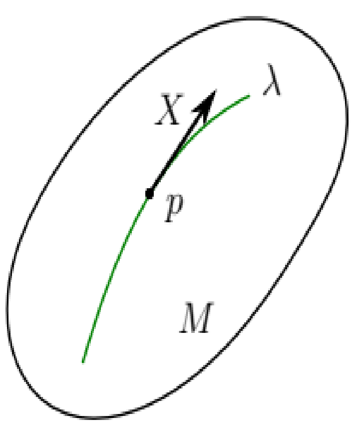

Figure 7.1: A curve in M

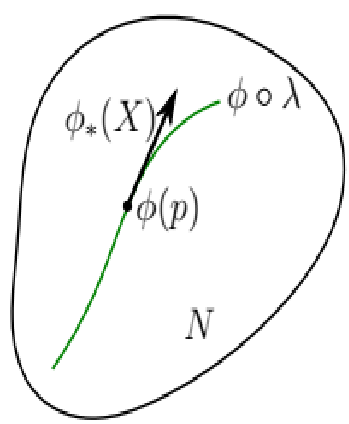

Figure 7.2: The curve in N

Definition. Let $\phi : M \to N$ be smooth. Let $p \in M$ and $X \in T_p(M)$. The *push-forward* of X with respect to ϕ is the vector $\phi_*(X) \in T_{\phi(p)}(N)$ defined as follows. Let λ be a smooth curve in M passing through p with tangent X at p. Then $\phi_*(X)$ is the tangent vector to the curve $\phi \circ \lambda$ in N at the point $\phi(p)$.

Lemma. Let $f : N \to \mathbb{R}$. Then $(\phi_*(X))(f) = X(\phi^*(f))$.

Proof. Wlog $\lambda(0) = p$.

$$\begin{aligned} (\phi_*(X))(f) &= \left[\frac{d}{dt}(f \circ (\phi \circ \lambda))(t)\right]_{t=0} \\ &= \left[\frac{d}{dt}((f \circ \phi) \circ \lambda)(t)\right]_{t=0} \\ &= X(\phi^*(f)) \end{aligned} \tag{7.1}$$

Exercise. Let x^μ be coordinates on M and y^α be coordinates on N (we use different indices α, β etc for N because N is a different manifold which might not have the same dimension as M). Then we can regard ϕ as defining a map $y^\alpha(x^\mu)$. Show that the components of $\phi_*(X)$ are related to those of X by

$$(\phi_*(X))^\alpha = \left(\frac{\partial y^\alpha}{\partial x^\mu}\right)_p X^\mu \tag{7.2}$$

The map on covectors works in the opposite direction:

Definition. Let $\phi : M \to N$ be smooth. Let $p \in M$ and $\eta \in T^*_{\phi(p)}(N)$. The *pull-back* of η with respect to ϕ is $\phi^*(\eta) \in T^*_p(M)$ defined by $(\phi^*(\eta))(X) = \eta(\phi_*(X))$ for any $X \in T_p(M)$.

Lemma. Let $f : N \to \mathbb{R}$. Then $\phi^*(df) = d(\phi^*(f))$.

Proof. Let $X \in T_p(M)$. Then

$$(\phi^*(df))(X) = (df)(\phi_*(X)) = (\phi_*(X))(f) = X(\phi^*(f)) = (d(\phi^*(f)))(X) \tag{7.3}$$

The first equality is the definition of ϕ^*, the second is the definition of df, the third is the previous Lemma and the fourth is the definition of $d(\phi^*(f))$. Since X is arbitrary, the result follows.

Exercise. Use coordinates x^μ and y^α as before. Show that the components of $\phi^*(\eta)$ are related to the components of η by

$$(\phi^*(\eta))_\mu = \left(\frac{\partial y^\alpha}{\partial x^\mu}\right)_p \eta_\alpha \tag{7.4}$$

Remarks.

1. In all of the above, the point p was arbitrary so push-forward and pull-back can be applied to vector and covector *fields*, respectively.

2. The pull-back can be extended to a tensor S of type $(0, s)$ by defining $(\phi^*(S))(X_1, \dots X_s) = S(\phi_*(X_1), \dots \phi_*(X_s))$ where $X_1, \dots, X_s \in T_p(M)$. Similarly, one can push-forward a tensor of type $(r, 0)$ by defining $\phi_*(T)(\eta_1, \dots, \eta_r) = T(\phi^*(\eta_1), \dots, \phi^*(\eta_r))$ where $\eta_1, \dots, \eta_r \in T_p^*(N)$. The components of these tensors in a coordinate basis are given by

$$(\phi^*(S))_{\mu_1 \dots \mu_s} = \left(\frac{\partial y^{\alpha_1}}{\partial x^{\mu_1}}\right)_p \dots \left(\frac{\partial y^{\alpha_s}}{\partial x^{\mu_s}}\right)_p S_{\alpha_1 \dots \alpha_s} \tag{7.5}$$

$$(\phi_*(T))^{\alpha_1 \dots \alpha_r} = \left(\frac{\partial y^{\alpha_1}}{\partial x^{\mu_1}}\right)_p \dots \left(\frac{\partial y^{\alpha_s}}{\partial x^{\mu_s}}\right)_p T^{\mu_1 \dots \mu_r} \tag{7.6}$$

Example. The embedding of S^2 into Euclidean space. Let $M = S^2$ and $N = \mathbb{R}^3$. Define $\phi : M \to N$ as the map which sends the point on S^2 with spherical polar coordinates $x^\mu = (\theta, \phi)$ to the point $y^\alpha = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) \in \mathbb{R}^3$. Consider the Euclidean metric g on $\mathbb{R}^3$, whose components are the identity matrix $\delta_{\alpha\beta}$. Pulling this back to S^2 using (7.5) gives $(\phi^* g)_{\mu\nu} = \mathrm{diag}(1, \sin^2\theta)$ (check!), the unit round metric on S^2.

Diffeomorphisms, Lie Derivative

Definition. A map $\phi : M \to N$ is a *diffeomorphism* iff it 1-1 and onto, smooth, and has a smooth inverse.

Remark. This implies that M and N have the same dimension. In fact, M and N have identical manifold structure.

With a diffeomorphism, we can extend our definitions of push-forward and pull-back so that they apply for any type of tensor:

Definition. Let $\phi : M \to N$ be a diffeomorphism and T a tensor of type (r, s) on M. Then the *push-forward* of T is a tensor $\phi_*(T)$ of type (r, s) on N defined by (for arbitrary $\eta_i \in T^*_{\phi(p)}(N)$, $X_i \in T_{\phi(p)}(N)$)

$$\phi_*(T)(\eta_1, \dots, \eta_r, X_1, \dots, X_s) = T(\phi^*(\eta_1), \dots, \phi^*(\eta_r), (\phi^{-1})_*(X_1), \dots, (\phi^{-1})_*(X_s)) \tag{7.7}$$

Exercises.

1. Convince yourself that push-forward commutes with the contraction and outer product operations.

2. Show that the analogue of equation (7.6) for a $(1,1)$ tensor field is

$$[(\phi_*(T))^\mu{}_\nu]_{\phi(p)} = \left(\frac{\partial y^\mu}{\partial x^\rho}\right)_p \left(\frac{\partial x^\sigma}{\partial y^\nu}\right)_p (T^\rho{}_\sigma)_p \tag{7.8}$$

(We don't need to use indices α, β etc because now M and N have the same dimension.) Generalize this result to a (r,s) tensor.

Remarks.

1. Pull-back can be defined in a similar way, with the result $\phi^* = (\phi^{-1})_*$.

2. We've taken an "active" point of view, regarding a diffeomorphism as a map taking a point p to a new point $\phi(p)$. However, there is an alternative "passive" point of view in which we consider a diffeomorphism simply as a change of chart at p. Consider a coordinate chart x^μ defined near p and another chart y^μ defined near $\phi(p)$ (Fig. 7.3). Regarding the coordinates y^μ as functions on N, we can pull them back to define corresponding coordinates, which we also call y^μ, on M. So now we have two coordinate systems defined near p. The components of tensors at p in the new coordinate basis are given by the tensor transformation law, which is exactly the RHS of (7.8).

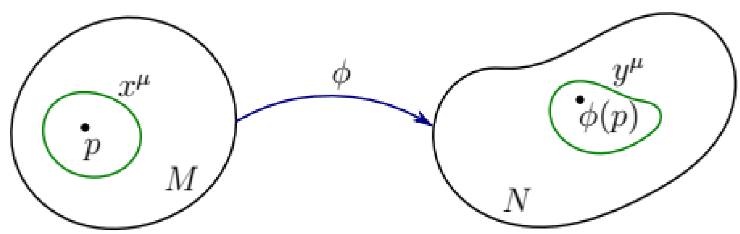

Figure 7.3: Active versus passive diffeomorphism.

Definition. Let $\phi : M \to N$ be a diffeomorphism. Let ∇ be a covariant derivative on M. The push-forward of ∇ is a covariant derivative $\tilde{\nabla}$ on N defined by

$$\tilde{\nabla}_X T = \phi_* \left(\nabla_{\phi^*(X)} \left(\phi^*(T)\right)\right) \tag{7.9}$$

where X is a vector field and T a tensor field on N. (In words: pull-back X and T to M, evaluate the covariant derivative there and then push-forward the result to N.)

Exercises (examples sheet 3).

1. Check that this satisfies the properties of a covariant derivative.

2. Show that the Riemann tensor of $\tilde{\nabla}$ is the push-forward of the Riemann tensor of ∇.

3. Let ∇ be the Levi-Civita connection defined by a metric g on M. Show that $\tilde{\nabla}$ is the Levi-Civita connection defined by the metric $\phi_*(g)$ on N.

Remark In GR we describe physics with a manifold M on which certain tensor fields e.g. the metric g, Maxwell field F etc. are defined. If $\phi : M \to N$ is a diffeomorphism then there is no way of distinguishing $(M, g, F, \ldots)$ from $(N, \phi_*(g), \phi_*(F), \ldots)$; they give equivalent descriptions of physics. If we set $N = M$ this reveals that the set of tensor fields $(\phi_*(g), \phi_*(F), \ldots)$ is physically indistinguishable from $(g, F, \ldots)$. If two sets of tensor fields are not related by a diffeomorphism then they *are* physically distinguishable. It follows that diffeomorphisms are the gauge symmetry (redundancy of description) in GR.

Example. Consider three particles following geodesics of the metric g. Assume that the worldlines of particles 1 and 2 intersect at p and that the worldlines of particles 2 and 3 intersect at q. Applying a diffeomorphism $\phi : M \to M$ maps the worldlines to geodesics of $\phi_*(g)$ which intersect at the points $\phi(p)$ and $\phi(q)$. Note that $\phi(p) \neq p$ so saying "particles 1 and 2 coincide at p" is not a gauge-invariant statement. An example of a quantity that *is* gauge invariant is the proper time between the two intersections along the worldline of particle 2.

Remark. This gauge freedom raises the following puzzle. The metric tensor is symmetric and hence has 10 independent components. Consider the vacuum Einstein equation - this appears to give 10 independent equations, which looks good. But the Einstein equation should not determine the components of the metric tensor uniquely, but only up to diffeomorphisms. The resolution is that not all components of the Einstein equations are truly independent because they are related by the contracted Bianchi identity.

Note that diffeomorphisms allow us to compare tensors defined at different points via push-forward or pull-back. This leads to a notion of a tensor field possessing symmetry:

Definition. A diffeomorphism $\phi : M \to M$ is a *symmetry transformation* of a tensor field T iff $\phi_*(T) = T$ everywhere. A symmetry transformation of the metric tensor is called an *isometry*.

Definition. Let X be a vector field on a manifold M. Let ϕ_t be the map which sends a point $p \in M$ to the point parameter distance t along the integral curve of X through p (this might be defined only for small enough t). It can be shown that ϕ_t is a diffeomorphism.

Remarks.

1. Note that ϕ_0 is the identity map and $\phi_s \circ \phi_t = \phi_{s+t}$. Hence $\phi_{-t} = (\phi_t)^{-1}$. If ϕ_t is defined for all $t \in \mathbb{R}$ (in which case we say the integral curves of X are *complete*) then these diffeomorphisms form a 1-parameter abelian group.

2. Given X we've defined ϕ_t. Conversely, if one has a 1-parameter abelian group of diffeomorphisms ϕ_t (i.e. one satisfying the rules just mentioned) then through any point p one can consider the curve with parameter t given by $\phi_t(p)$. Define X to be the tangent to this curve at p. Doing this for all p defines a vector field X. The integral curves of X generate ϕ_t in the sense defined above.

3. If we use $(\phi_t)_*$ to compare tensors at different points then the parameter t controls how near the points are. In particular, in the limit $t \to 0$, we are comparing tensors at infinitesimally nearby points. This leads to the notion of a new type of derivative:

Definition. The *Lie derivative* of a tensor field T with respect to a vector field X at p is

$$(\mathcal{L}_X T)_p = \lim_{t \to 0} \frac{((\phi_{-t})_* T)_p - T_p}{t} \tag{7.10}$$

Remark. The Lie derivative wrt X is a map from (r, s) tensor fields to (r, s) tensor fields. It obeys $\mathcal{L}_X(\alpha S + \beta T) = \alpha \mathcal{L}_X S + \beta \mathcal{L}_X T$ where α and β are constants.

The easiest way to demonstrate other properties of the Lie derivative is to introduce coordinates in which the components of X are simple. Let Σ be a hypersurface that has the property that X is nowhere tangent to Σ (in particular $X \neq 0$ on Σ). Let x^i, $i = 1, 2, \ldots, n-1$ be coordinates on Σ. Now assign coordinates (t, x^i) to the point parameter distance t along the integral curve of X that starts at the point with coordinates x^i on Σ (Fig. 7.4).

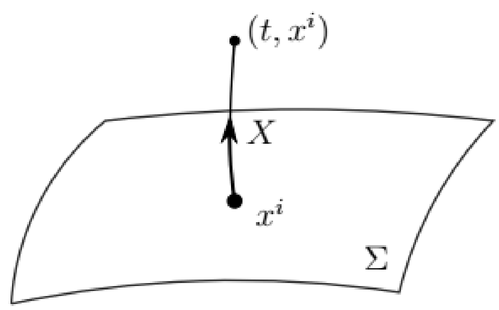

Figure 7.4: Coordinates adapted to a vector field

This defines a coordinate chart (t, x^i) at least for small t, i.e., in a neighbourhood of Σ. Furthermore, the integral curves of X are the curves (t, x^i) with fixed

x^i and parameter t. The tangent to these curves is $\partial/\partial t$ so we have constructed coordinates such that $X = \partial/\partial t$. The diffeomorphism ϕ_t is very simple: it just sends the point p with coordinates $x^\mu = (t_p, x^i_p)$ to the point $\phi_t(p)$ with coordinates $y^\mu = (t_p + t, x^i_p)$ hence $\partial y^\mu / \partial x^\nu = \delta^\mu_\nu$. The generalization of (7.8) to a (r, s) tensor then gives

$$[((\phi_t)_*(T))^{\mu_1,\dots,\mu_r}{}_{\nu_1,\dots\nu_s}]_{\phi_t(p)} = [T^{\mu_1,\dots,\mu_r}{}_{\nu_1,\dots\nu_s}]_p \tag{7.11}$$

and hence

$$[((\phi_t)_*(T))^{\mu_1,\dots,\mu_r}{}_{\nu_1,\dots\nu_s}]_p = [T^{\mu_1,\dots,\mu_r}{}_{\nu_1,\dots\nu_s}]_{\phi_{-t}(p)} \tag{7.12}$$

It follows that, if p has coordinates (s, x^i) in this chart,

$$\begin{aligned}(\mathcal{L}_X T)^{\mu_1,\dots,\mu_r}{}_{\nu_1,\dots\nu_s} &= \lim_{t\to 0} \frac{1}{t}\left(T^{\mu_1,\dots,\mu_r}{}_{\nu_1,\dots\nu_s}(s+t, x^i) - T^{\mu_1,\dots,\mu_r}{}_{\nu_1,\dots\nu_s}(s, x^i)\right) \\ &= \left[\frac{\partial}{\partial t} T^{\mu_1,\dots,\mu_r}{}_{\nu_1,\dots\nu_s}(t, x^i)\right]_{(s,x^i)}\end{aligned} \tag{7.13}$$

So *in this chart*, the Lie derivative is simply the partial derivative with respect to the coordinate t. It follows that the Lie derivative has the following properties:

1. It obeys the Leibniz rule: $\mathcal{L}_X(S \otimes T) = (\mathcal{L}_X S) \otimes T + S \otimes \mathcal{L}_X T$.

2. It commutes with contraction.

Now let's derive a basis-independent formula for the Lie derivative. First consider a function f. In the above chart, we have $\mathcal{L}_X f = (\partial/\partial t)(f)$. However, in this chart we also have $X(f) = (\partial/\partial t)(f)$. Hence

$$\mathcal{L}_X f = X(f) \tag{7.14}$$

Both sides of this expression are scalars and hence this equation must be valid in any basis. Next consider a vector field Y. In our coordinates above we have

$$(\mathcal{L}_X Y)^\mu = \frac{\partial Y^\mu}{\partial t} \tag{7.15}$$

but we also have

$$[X, Y]^\mu = \frac{\partial Y^\mu}{\partial t} \tag{7.16}$$

If two vectors have the same components in one basis then they are equal in all bases. Hence we have the basis-independent result

$$\mathcal{L}_X Y = [X, Y] \tag{7.17}$$

Remark. Let's compare the Lie derivative and the covariant derivative. The former is defined on any manifold whereas the latter requires extra structure (a connection). Equation (7.17) reveals that the Lie derivative wrt X at p depends on X_p *and the first derivatives of* X at p. The covariant derivative wrt X at p depends only on X_p, which enables us to remove X to define the tensor ∇T, a covariant generalization of partial differentiation. It is not possible to define a corresponding tensor $\mathcal{L}T$ using the Lie derivative. Only $\mathcal{L}_X T$ makes sense.

Exercises (examples sheet 3).

1. Derive the formula for the Lie derivative of a covector ω in a coordinate basis:
$$(\mathcal{L}_X \omega)_\mu = X^\nu \partial_\nu \omega_\mu + \omega_\nu \partial_\mu X^\nu \tag{7.18}$$
Show that this can be written in the basis-independent form (where ∇ is the Levi-Civita connection)
$$(\mathcal{L}_X \omega)_a = X^b \nabla_b \omega_a + \omega_b \nabla_a X^b \tag{7.19}$$

2. Show that the Lie derivative of the metric in a coordinate basis is
$$(\mathcal{L}_X g)_{\mu\nu} = X^\rho \partial_\rho g_{\mu\nu} + g_{\mu\rho}\partial_\nu X^\rho + g_{\rho\nu}\partial_\mu X^\rho \tag{7.20}$$
and that this can be written in the basis-independent form
$$(\mathcal{L}_X g)_{ab} = \nabla_a X_b + \nabla_b X_a \tag{7.21}$$

Remark. If ϕ_t is a symmetry transformation of T (for all t) then $\mathcal{L}_X T = 0$. If ϕ_t are a 1-parameter group of isometries then $\mathcal{L}_X g = 0$, i.e.,
$$\nabla_a X_b + \nabla_b X_a = 0 \tag{7.22}$$
This is *Killing's equation* and solutions are called *Killing vector fields*. Consider the case in which there exists a chart for which the metric tensor does not depend on some coordinate z. Then equation (7.20) reveals that $\partial/\partial z$ is a Killing vector field. Conversely, if the metric admits a Killing vector field then equation (7.13) demonstrates that one can introduce coordinates such that the metric tensor components are independent of one of the coordinates.

Lemma. Let X^a be a Killing vector field and let V^a be tangent to an affinely parameterized geodesic. Then $X_a V^a$ is constant along the geodesic.

Proof. The derivative of $X_a V^a$ along the geodesic is
$$\begin{aligned}\frac{d}{d\tau}(X_a V^a) = V(X_a V^a) &= \nabla_V (X_a V^a) = V^b \nabla_b (X_a V^a) \\ &= V^a V^b \nabla_b X_a + X_a V^b \nabla_b V^a \end{aligned} \tag{7.23}$$

The first term vanishes because Killing's equation implies that $\nabla_b X_a$ is antisymmetric. The second term vanishes by the geodesic equation.

Exercise. Let $J^a = T^a{}_b X^b$ where T_{ab} is the energy-momentum tensor and X^b is a Killing vector field. Show that $\nabla_a J^a = 0$, i.e., J^a is a *conserved current*.