Vector fields and Differential Forms on $\mathbb{R}^n$

Robert Hladky

September 29, 2007

1 Introduction

These notes are intended to supplement Chapter 1 of Barret O'Neill. The definitions and notations given here are often at a higher level of abstraction than those given in the text. While on a practical level, all definitions will prove to be equivalent, those given here are easiest to generalize.

This is a first draft. Please bring needed corrections, typos and suggestions to the attention of the author.

2 Vector Fields and 1-forms

A derivation at $p \in \mathbb{R}^n$ is a linear map

$$D \colon C^\infty(\mathbb{R}^n) \to \mathbb{R}$$

such that

$$D(fg) = f(p)Dg + g(p)Df$$

The principle examples of a derivation at p are the directional derivatives at p. For $v \in \mathbb{R}^n$ we define

$$D_{p,v} f = \frac{d}{dt}\Big|_{t=0} f(p + tv).$$

The chain and product rules of differentiation imply that each $D_{p,v}$ is a derivation at p.

Lemma 2.1. *Every derivation at $p \in \mathbb{R}^n$ can be expressed as $D_{p,v}$ for some $v \in \mathbb{R}^n$.*

Proof: For a derivation D at p, define $v_i = D(x_i)$ for $i = 1, \ldots, n$ and $v = (v_1, \ldots, v_n)$. Let f be a smooth function. Then we can consider the Taylor expansion of f near p

$$f(x) = f(p) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(p)\,(x_i - p_i) + O(|x|)^2$$

where the notation $O(|x|^2)$ denotes a term that vanishes quadratically as $x \to 0$. i.e. has magnitude $\leq C|x|^2$ for some constant C when x is small.

Since derivations all annihilate constants (do you see why?) we have that

$$D(f) = \sum v_i \frac{\partial f}{\partial x_i}(p) = D_{p,v} f.$$

■

Definition 2.2. *The tangent space to $\mathbb{R}^n$ at $p \in \mathbb{R}^n$, denoted by $T_p\mathbb{R}^n$, is the vector space of derivations at $p \in \mathbb{R}^n$.*

From Lemma 2.1, it is clear that we can associate to any derivation at p, a vector v. Thus there is a linear isomorphism which we can view as an identification

$$T_p\mathbb{R}^n \cong \{v_p = (p, v) \colon v \in \mathbb{R}^n\}$$

This identification is made explicit by the observation that for $v = (v_1, \ldots, v_n)$,

$$D_{v,p} f = \sum v_i \frac{\partial f}{\partial x_i}(p)$$

Thus we can regard the partial derivatives $\left(\frac{\partial}{\partial x_i}\right)_p$ at the point p, as a basis for $T_p\mathbb{R}^n$. This identification means that we shall often want to regard tangent vectors, simply as vectors rather than as derivations. To emphasize when we are employing the derivation definition we shall denote the the application of a tangent vector to a function using the notation $v_p[f]$. Thus $v_p[f] = D_{p,v} f$.

Another and very useful way of viewing and computing tangent vectors is as tangent vectors to arbitrary curves. If $\gamma \colon \mathbb{R} \to \mathbb{R}^n$ is a curve with $\gamma(0) = p$ and $\gamma'(0) = v$, then the tangent vector v_p is the derivation that acts by

$$v_p[f] = \frac{d}{dt}\Big|_{t=0} f(\gamma(t)).$$

That this is equivalent to the directional derivative $D_{p,v}$ is an exercise in basic calculus. The usefulness of this version kicks in when you have to work with tangent vectors on spaces other than $\mathbb{R}^n$ where lines are hard to compute explicitly.

Definition 2.3. *The tangent bundle*

$$T\mathbb{R}^n = \{(p, v_p) \colon p \in \mathbb{R}^n, v_p \in T_p\mathbb{R}^n\}$$

is the disjoint union of all tangent spaces to $\mathbb{R}^n$.

Definition 2.4. *A vector field X on $\mathbb{R}^n$ is a map $X \colon \mathbb{R}^n \to T\mathbb{R}^n$ such that*

$$X(p) \in T_p\mathbb{R}^n$$

So a vector field associates to each point $p \in \mathbb{R}^n$, a tangent vector $X(p)$ at p.

Often we shall consider the action of vector fields on smooth functions $X[f]$, which returns the function $p \mapsto X(p)[f]$.

We introduce canonical vector fields $U_i = \frac{\partial}{\partial x_i}$ with the property that

$$U_i[f] = \frac{\partial f}{\partial x_i}.$$

Every vector field can then be expressed as

$$X = \sum X_i U_i$$

with X_i, $i = 1, \dots, n$ being a smooth function.

Example 2.5. On $\mathbb{R}^3$, let $X = y^2 \frac{\partial}{\partial x} + z \frac{\partial}{\partial y}$ then

$$X[xz + y] = y^2(z) + z(1) = zy^2 + z.$$

□

Recall from linear algebra that the dual space V^* of a vector space V is the vector space of all linear transformations $\phi \colon V \to \mathbb{R}$.

Definition 2.6. *The cotangent space of* $\mathbb{R}^n$ *at* $p \in \mathbb{R}$, $T_p^*\mathbb{R}^n$, *is the dual space of the tangent space at* p, *i.e.*

$$T_p^*\mathbb{R}^n = (T_p\mathbb{R}^n)^*.$$

The cotangent bundle of $\mathbb{R}^n$

$$T^*\mathbb{R}^n = \left\{(p, \phi) \colon p \in \mathbb{R}^n, \phi \in T_p^*\mathbb{R}^n\right\}$$

is the disjoint union of all the cotangent spaces. A 1*-form on* $\mathbb{R}^n$ *is a map* $\varphi \colon \mathbb{R}^n \to T^*\mathbb{R}^n$ *such that*

$$\varphi(p) \in T_p^*\mathbb{R}^n$$

Thus a cotangent vector eats a vector and spits out a real number while a 1-form eats a vector field and spits out a real-valued function.

We define the coordinate 1-forms on $\mathbb{R}$ by

$$dx_j(U_i) = \delta_{ij}$$

where $\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$

Since any linear map on a vector space is exactly determined by what it does to a basis, we see that any 1-form on $\mathbb{R}^n$ can be written

$$\varphi = \sum \varphi_i dx_i$$

where the φ_i are smooth functions.

3 Alternating multilinear Maps, the Wedge Product and Differential Forms

A k-multilinear (or k-linear) map on a vector space V is a map

$$\varphi\colon \prod_1^k V = V \times V \times \cdots \times V \to \mathbb{R}$$

such that for every j and $v_1, \dots v_k$ the map

$$v \mapsto (a_1, \dots, a_{j-1}, v, a_{j+1}, \dots a_k)$$

is a linear.

A k-linear map is called alternating (or asymmetric) if for any permutation $\pi \in S(k)$ we have

$$\varphi \circ \pi = \operatorname{sgn}(\pi)\varphi$$

Remember that the sign of the permutation is $+1$ if it built from an even number of pair swaps and is -1 if built from an odd number.

So for 2-linear maps this is equivalent to $\varphi(v,w) = -\varphi(w,v)$. For 3-linear maps we get

$$\begin{aligned}\varphi(u,v,w) &= \varphi(v,w,u) = \varphi(w,u,v)\\ &= -\varphi(v,u,w) = -\varphi(u,w,v) = -\varphi(w,v,u).\end{aligned}$$

Alternating multilinear maps can be built out of linear maps using the wedge product. If $\varphi_1, \dots \varphi_k \in V^*$, i.e. are linear maps $V \to \mathbb{R}$, then we define the k-fold wedge product $\varphi_1 \wedge \cdots \wedge \varphi_k$ by

$$\begin{aligned}\varphi_1 \wedge \cdots \wedge \varphi_k(v_1, \dots v_k) &= \sum_{\pi \in S(k)} \left(\operatorname{sgn}(\pi) \prod_{i=1}^k \varphi_i(v_{\pi(i)}) \right)\\ &= \det \begin{pmatrix} \varphi_1(v_1) & \varphi_1(v_2) & \dots & \varphi_1(v_k) \\ \varphi_2(v_1) & \varphi_2(v_2) & \dots & \varphi_2(v_k) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_k(v_1) & \varphi_k(v_2) & \dots & \varphi_k(v_k) \end{pmatrix}\end{aligned}$$

When $k = 2$, this implies that

$$\phi \wedge \psi(X,Y) = \det \begin{pmatrix} \phi(X) & \phi(Y) \\ \psi(X) & \psi(Y) \end{pmatrix} = \phi(X)\psi(Y) - \phi(Y)\psi(X). \tag{1}$$

It is a useful fact that all alternating k-linear can be built out of linear combinations of k-fold wedge products of linear maps.

The permutation property also applies to these wedge products: if $\pi \in S(k)$ then

$$\varphi_{\pi(1)} \wedge \varphi_{\pi(2)} \wedge \cdots \wedge \varphi_{\pi(k)} = \operatorname{sgn}(\pi)\varphi_1 \wedge \dots \varphi_k$$

when $k = 2$ for instance, this implies

$$\phi \wedge \psi = -\psi \wedge \phi.$$

Lemma 3.1. *If φ is an alternating k-linear map on V and $v_1, \dots v_k \in V$ are linearly dependent then*

$$\varphi(v_1, \dots, v_k) = 0$$

Proof: Since φ is alternating any single swap of order of two of $v_1, \dots, v_n$, switches the sign of the output. This means that if any $v_i = v_j$ for $i \neq j$ then the output is zero. But if the vectors are linearly dependent then one of them can be written as a linear combination of the others. Thus by $\varphi(v_1, \dots, v_k)$ can be broken into pieces each having a repeated input. The whole thing must therefore vanish.

■

When working with a tangent space $T_p\mathbb{R}$, this means that all alternating k-linear maps can be expressed as linear combinations of $dx_1(p), \dots dx_n(p)$.

We declare a k-form to be a map that assigns to each $p \in \mathbb{R}^n$ an alternating k-linear map on $T_p\mathbb{R}^n$. A k-form can then be thought of as something that eats k vector fields and spits out a function.

To compute on $\mathbb{R}^3$, it is usually easiest to use the following methods:

If $k = 2$ use (1). For example

$$(dx + zdy) \wedge (ydz) \left(y\frac{\partial}{\partial x} + \frac{\partial}{\partial y}, 2\frac{\partial z}{\partial +}3\frac{\partial}{\partial y} \right) = \det \begin{pmatrix} y+z & 3y \\ 0 & 2 \end{pmatrix} = 2y + 2z$$

If $k = 3$, note that every 3-form on $\mathbb{R}^3$ can be expressed as $f(x, y, z)dx \wedge dy \wedge dz$, then use the determinant formula. For example

$$\begin{aligned}
(xdx + dy) \wedge (dy - ydz) \wedge (zdx + dz) &\left(3\frac{\partial}{\partial x} + z\frac{\partial}{\partial y}, \frac{\partial}{\partial z}, \frac{\partial}{\partial y} + 2\frac{\partial}{\partial z} \right) \\
&= (x - yz)dx \wedge dy \wedge dz \left(3\frac{\partial}{\partial x} + z\frac{\partial}{\partial y}, \frac{\partial}{\partial z}, \frac{\partial}{\partial y} + 2\frac{\partial}{\partial z} \right) \\
&= (x - yz)\det \begin{pmatrix} 3 & z & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix} \\
&= (x - yz)(-3) = 3yz - 3x
\end{aligned}$$

4 Exterior Derivative

The exterior differential operator, d, is really a family of maps that differentiates k-forms into $k+1$-forms for all k. It is defined by stating it has the following properties:

- d is linear.
- For a smooth function $df(v_p) = v_p[f]$.
- For any coordinate function x_i, $d(dx_i) = 0$.
- $d(\phi \wedge \psi) = d\phi \wedge \psi + (-1)^{deg(\phi)}\phi \wedge d\psi$.

We can use these to work out how to compute more explicitly.

Lemma 4.1. *For a smooth function f,*

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i$$

Proof: We compute what happens to both sides when we apply it to the basis tangent vectors $\left(\frac{\partial}{\partial x_k}\right)_p$

$$df\left(\left(\frac{\partial}{\partial x_k}\right)_p\right) = \left(\frac{\partial}{\partial x_k}\right)_p [f] = \frac{\partial f}{\partial x_k}(p)$$

$$\left(\sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i\right)\left(\frac{\partial}{\partial x_k}\right)_p = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(p)\delta_{ik} = \frac{\partial f}{\partial x_k}(p)$$

Since these are the same for all k and p and 1-forms are linear, we can deduce that the 1-forms act the same way on all tangent vectors so must be the same. ∎

From the remaining defining properties of the exterior derivative we can immediately see how to differentiate all other forms. Note that a function can be regarded as a 0-form (or 0 degree form). Rather than try to produce a general formula, this is best illustrated by examples. Take $n = 3$

$$\begin{aligned}
d(f dx \wedge dy) &= df \wedge dx \wedge dy + f d(dx) \wedge dy - f dx \wedge d(dy) \\
&= \left(\frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz\right) \wedge dx \wedge dy + 0 + 0 \\
&= \frac{\partial f}{\partial z} dz \wedge dx \wedge dy \\
&= \frac{\partial f}{\partial z} dx \wedge dy \wedge dz.
\end{aligned}$$

$$\begin{aligned}
d(f dy) &= df \wedge dy + f d(dy) \\
&= \left(\frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz\right) \wedge dy + 0 \\
&= \frac{\partial f}{\partial x} dx \wedge dy - \frac{\partial f}{\partial z} dy \wedge dz.
\end{aligned}$$

In general, if you want to compute the action of an exterior derivative on vectors, e.g. $d\phi(X, Y)$, it is best to write out the differential and vectors in coordinates and compute using the techniques of the wedge product.

Lemma 4.2. *For all degrees of form $d \circ d = 0$.*

Proof: First prove this for functions (see solutions to HW1). But then note that every p-form is of the type

$$\eta = \sum_k f_k dx_{k1} \wedge dx_{k2} \wedge \cdots \wedge dx_{kp}$$

for some choices of coordinate functions x_{kp}. Then

$$d\eta = \sum_k df_k \wedge dx_{k1} \wedge dx_{k2} \wedge \cdots \wedge dx_{kp}.$$

Now since d is linear, from the fourth defining property, when you compute $d(d\eta)$ the exterior derivative must always hit either a df_k or a dx_{kp}. But since $d \circ d = 0$ on functions, this means $d(d\eta) = 0$ also.

∎

Remark 4.3. *The adjective exterior is used as there are other notions of derivatives of forms. We'll see the "interior derivative" later on when we talk about Stokes' theorem, but for the most part the adjective "exterior" can be regarded as redundant and d should really be thought of as just taking the derivative of a differential form.*

5 Pushforward and Pullback

Let $F = (F_1, \ldots, F_m)\colon \mathbb{R}^n \to \mathbb{R}^m$ be a smooth map, then the pushforward of F at the point p is the linear map

$$F_* : T_p\mathbb{R}^n \to T_{F(p)}\mathbb{R}^m$$

defined as follows

$$F_*(v_p) = \frac{d}{dt}\Big|_{t=0} F(p + tv)$$

Notice that $t \mapsto F(p+tv)$ is a curve into the image space $\mathbb{R}^m$, so the derivative on the right hand side can be viewed as a tangent vector in $T_{F(p)}\mathbb{R}^m$.

From the chain-rule, we note that

$$\frac{d}{dt} F(p + t(cv + w)) = c\frac{d}{dt} F(p + tv) + \frac{d}{dt} F(p + tw)$$

which means that F_* is indeed a linear map.

Remark 5.1. *In general, the curve $p+tv$ can be replaced by any curve $\gamma\colon \mathbb{R} \to \mathbb{R}^n$ such that $\gamma(0) = p$ and $\gamma'(0) = v$.*

For the case $n = 2$ and $m = 3$, it is easy to visualize what this means. Stretch a piece of paper (or better a stretchable cloth) out flat on a table. This is our $\mathbb{R}^2$. Mark on the paper a point p and draw a vector v based at p. This is our tangent vector v_p in $T_p\mathbb{R}^2$. (If you like you can now draw a straight line through p in the direction of v.) Now pick up the paper and do some combination of bending, twisting, folding and stretching to your paper. This is applying a function $F\colon \mathbb{R}^2 \to \mathbb{R}^3$. Now hold the paper still. The marked point p is now at location $F(p)$. The marked vector coming out of the point is now $F_*(v_p)$.

Lemma 5.2. *If $\mathbb{R}^n$ has coordinate functions $x_1, \ldots, x_n$ and $\mathbb{R}^m$ coordinate functions $y_1, \ldots, y_m$, then*

$$F_* \left(\frac{\partial}{\partial x_j} \right)_p = \sum_{k=1}^{m} \frac{\partial F_k}{\partial x_j}(p) \left(\frac{\partial}{\partial y_k} \right)_{F(p)}$$

Proof: Go back to the definition and the chain-rule

$$\frac{d}{dt} F_k(p + tU_j(p)) = \delta_{jm} \frac{F_k}{x_m}(p) = \frac{F_k}{x_j}(p).$$

But this implies the kth coefficient of $F_* \left(\frac{\partial}{\partial x_j} \right)_p$ is exactly $\frac{\partial F_k}{\partial x_j}(p)$.

∎

This means that we can regard F_* as a matrix

$$F_* = \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \cdots & \frac{\partial F_1}{\partial x_m} \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} & \cdots & \frac{\partial F_2}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1} & \frac{\partial F_m}{\partial x_2} & \cdots & \frac{\partial F_m}{\partial x_n} \end{pmatrix}$$

If $X = \sum_1^n a_i \left(\frac{\partial}{\partial x_i} \right)_p$ then $F_* X = \sum_1^m b_j \left(\frac{\partial}{\partial y_j} \right)_{F(p)}$ where

$$\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} = F_*(p) \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

Example 5.3. Consider spherical polar coordinates which we can regard as a map

$$F(r, \theta, \phi) = (r \cos\theta \sin\phi, r \sin\theta \sin\phi, r \cos\phi)$$

from $\mathbb{R}^3$ to $\mathbb{R}^3$. Then the pushforward matrix is

$$F_* = \begin{pmatrix} \cos\theta \sin\phi & -r \sin\theta \sin\phi & r \cos\theta \cos\phi \\ \sin\theta \sin\phi & r \cos\theta \sin\phi & r \sin\theta \cos\phi \\ \cos\phi & 0 & -r \sin\phi \end{pmatrix}$$

so for instance to compute $F_* \left(2\frac{\partial}{\partial r} + 3\frac{\partial}{\partial \phi} \right)_{(2,0,\pi/2)}$ we compute

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ -6 \end{pmatrix}$$

so

$$F_*\left(2\frac{\partial}{\partial r}+3\frac{\partial}{\partial \phi}\right)_{(2,0,\pi/2)} = 2\left(\frac{\partial}{\partial x}\right)_{(2,0,0)} - 6\left(\frac{\partial}{\partial z}\right)_{(2,0,0)}.$$

□

There's another way to think about vectors, namely as directional derivatives. The pushforward can be viewed in this way too.

Lemma 5.4. *If f is a smooth function on $\mathbb{R}^m$ and v_p is a tangent vector on $\mathbb{R}^n$ then*

$$(F_* v_p)[f] = v_p[f \circ F]$$

Proof: We go back to the definitions, the vector $F_* v_p$ can be viewed as the derivative at $t = 0$ of the curve $\gamma(t) = F(p + tv)$. The directional derivative $F_* v_p$ applied to f is then the derivative at $t = 0$ of the function $f \circ \gamma(t)$, thus

$$(F_* v_p)[f] = \frac{d}{dt}\Big|_{t=0}(f \circ \gamma)(t) = \frac{d}{dt}\Big|_{t=0}(f \circ F)(p + tv) = v_p[f \circ F].$$

■

Definition 5.5. *A smooth map is regular at p if F_* is injective at p. A map is said to be regular if it is regular at every point.*

If F_* is not injective at p, then there are two distinct vectors at p that are pushed forward to the same vector at $F(p)$. On the piece of paper model this would correspond to a sharp fold in the paper or a cone point/crumpling. Non-regular points are where the image does not look smooth.

Definition 5.6. *A smooth map $F\colon U \to V$ with $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ both open is a diffeomorphism if there is a smooth inverse map $F^{-1}\colon V \to U$.*

Lemma 5.7. *A smooth 1-1 and onto map $F\colon U \to V$ that is regular at every point $p \in U$ is a diffeomorphism.*

The proof of this is a bit of fiddly analysis and is omitted.

We note that regularity at p requires that $n \leq m$ and is equivalent to either being able to find n linearly independent rows or columns in the matrix of F_* or being able to find a $n \times n$ submatrix with non-zero determinant.

The problem with pushforwards is that although they are well defined for pushing vectors to vectors, they behave badly on vector fields. Given a vector field X on $\mathbb{R}^n$, we would like to define a vector field $F_* X$ by $(F_* X)_{F(p)} = F_*(X_p)$. However there are several problems with doing this. If F is not surjective, then this doesn't even define $F_* X$ at points outside the range of F. If F is not injective, then there be many possible candidates for $F_* X$ at certain points. Even if F is bijective, if F is not regular then there are problems with the coefficients of $F_* X$ not being smooth.

Because of this problem, we move away from vector fields and pushforwards with their intuitive descriptions. Where we move is into the more abstract realm of differential forms. Computationally this turns out to be a great boon, but it does come at the cost of obscuring the pictures.

If $F\colon \mathbb{R}^n \to \mathbb{R}^m$ is a smooth map then the pullback is the linear map

$$F^*\colon T^*_{F(p)}\mathbb{R}^m \to T^*_p\mathbb{R}^n$$

defined by

$$F^*\varphi(v_p) = \varphi(F_* v_p).$$

Although this definition seems very abstract and requires the pushforward, the computational formulas turn out to be very simple. Furthermore since every point $p \in \mathbb{R}^n$ is mapped to exactly one $F(p) \in \mathbb{R}^m$, if we pullback a 1-form on $\mathbb{R}^m$ point by point, we do indeed get a 1-form on $\mathbb{R}^n$.

Lemma 5.8. *If f is a smooth function on $\mathbb{R}^m$ then*

$$F^* df = d(f \circ F)$$

Proof: We unwind all the formal definitions using Lemma 5.4

$$F^* df(v_p) = df(F_* v_p) = (F_* v_p)[f] = v_p[f \circ F] = d(f \circ F)(v_p).$$

■

We extend the pullback to act on k-forms by declaring the following

- F^* is linear
- $F^* f = f \circ F$
- $F^*(\phi \wedge \psi) = F^*\phi \wedge F^*\psi$

This means we can commute pullbacks by simple differentiation.

Example 5.9. For the spherical polar coordinate map

$$F(r, \theta, \phi) = (r\cos\theta\sin\phi, r\sin\theta\sin\phi, r\cos\phi)$$

we have

$$\begin{aligned} F^* dx &= \cos\theta\sin\phi\, dr - r\sin\theta\sin\phi\, d\theta + r\cos\theta\cos\phi\, d\phi \\ F^*(ydz) &= r\sin\theta\sin\phi\,(\cos\phi\, dr - r\sin\phi\, d\phi) \end{aligned}$$

$$\begin{aligned} F^*(dx \wedge dy \wedge dz) &= (\cos\theta\sin\phi\, dr - r\sin\theta\sin\phi\, d\theta + r\cos\theta\cos\phi\, d\phi) \\ &\quad \wedge (\sin\theta\sin\phi\, dr + r\cos\theta\sin\phi\, d\theta + r\sin\theta\cos\phi\, d\phi) \\ &\quad \wedge (\cos\phi\, dr - r\sin\phi\, d\phi) \\ &= \det\begin{pmatrix} \cos\theta\sin\phi & -r\sin\theta\sin\phi & r\cos\theta\cos\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & \sin\theta\cos\phi \\ \cos\phi & 0 & -r\sin\phi \end{pmatrix} dr \wedge d\theta \wedge d\phi \\ &= -r^2\sin\phi\, dr \wedge d\theta \wedge d\phi \\ &= r^2\sin\phi\, dr \wedge d\phi \wedge d\theta \end{aligned}$$

□

Another computational consequence of all of this is the following

Lemma 5.10. *The pullback commutes with the exterior derivative, i.e.*

$$F^*(d\eta) = d(F^*\eta)$$

Proof: We've already seen that this is true for 1-forms. The general case now follows from the defining property of F^* and the fact that d is linear and anti-commutes with $\wedge$.

■

6 Inner Products

Definition 6.1. *An inner product on a vector space V is a strictly positive, symmetric bilinear function, g on V. i.e. bilinear map*

$$g \colon V \times V \to \mathbb{R}$$

such that

$$g(v,w) = g(v,w), \qquad g(v,v) > 0 \text{ for } v \neq 0$$

Example 6.2. The dot product is the inner product on $\mathbb{R}^n$, $\bar{g}(a,b) = a \cdot b$ where

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 & \dots & a_n \end{pmatrix} \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \sum_{i=1}^{n} a_i b_i$$

Other inner products can be created replacing the identity matrix with a strictly positive, symmetric $n \times n$ matrix G. So G must satisfy $GG^\top = Id$ and the $j \times j$ square matrix in the top left corner of G must have positive determinant for all $1 \leq j \leq n$.

$$g_G(a,b) = a^\top G b.$$

□

We can view any tangent space $T_p\mathbb{R}^n$ as just a copy of $\mathbb{R}^n$ with the vectors $U_j(p)$ as the canonical basis. Thus we can view the dot product as being an inner product on every $T_p\mathbb{R}^n$.

Definition 6.3. *If $v_1, \dots, v_n$ is a basis for a vector space V, the dual basis is the basis $v_1^*, \dots v_n^*$ for V^* such that*

$$v_j^*(v_k) = \delta_{jk}$$

Example 6.4. Recall that the standard basis for each $T_p\mathbb{R}^n$ is given by $U_i(p) = \left(\frac{\partial}{\partial x_i}\right)_p$, $i = 1, \dots, n$. The dual basis for $T_p^*\mathbb{R}^n$ is just $dx_1, \dots, dx_n$. □

Lemma 6.5. *Suppose*

$$\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = M \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$$

is a linear change of basis for $\mathbb{R}^n$, *i.e.* $u_1, \dots u_n$ *is a basis of vectors of* $\mathbb{R}^n$ *and the* $n \times n$ *matrix* M *has* $\det(M) \neq 0$. *Then the dual basis changes are given by*

$$\vec{u}^* = M^\top \vec{v}^*, \qquad \vec{v}^* = (M^\top)^{-1} \vec{u}^*$$

or equivalently

$$u_j^* = \sum_{k=1}^n M_{kj} v_k^*, \qquad v_j^* = \sum_{k=1}^n (M^{-1})_{kj} u_k^*$$

Proof: Note that $u_j^*(v_k) = u_j^*\left(\sum M_{km} u_m\right) = M_{kj}$ so $u_j^* = \sum_k M_{kj} v_k^*$.

■

There is another way to build bilinear forms on the tangent spaces. We can make them out of covectors.

Recall a covector ϕ at p is a linear map $\phi \colon T_p\mathbb{R}^n \to \mathbb{R}$. We define the tensor product $\otimes$ of two covectors as follows

$$(\phi \otimes \psi)(v_p, w_p) = \phi(v_p)\psi(w_p).$$

We've seen something like this before with the wedge product. In fact

$$\phi \wedge \psi = \phi \otimes \psi - \psi \otimes \phi.$$

These tensor products are bilinear maps on $T_p\mathbb{R}^n$ and all bilinear maps can be built out of these. For example, the bilinear map

$$g\left(\sum a_i U_i(p), \sum b_i U_i(p)\right) = a^\top G(p) b$$

with the $n \times n$ matrix of functions $G(p) = (G_{jk}(p))$ can be re-expressed as

$$g = \sum_{j,k=1}^n G_{jk}(p)\, (dx_j(p) \otimes dx_k(p))$$

Why is this true? Well

$$\begin{aligned}
g\left(\sum a_i U_i, \sum b_i U_i\right) &= \sum_{j,k} G_{jk}\, (dx_j(p) \otimes dx_k(p)) \left(\sum_i a_i U_i, \sum_m b_m U_m\right) \\
&= \sum_{j,k,i,m} G_{jk} a_i b_m dx_j \otimes dx_k(U_i, U_m) \\
&= \sum_{j,k,i,m} G_{jk} a_i b_m \delta_{ji} \delta_{km} \\
&= \sum_{j,k} G_{jk} a_j b_k \\
&= a^\top G b
\end{aligned}$$

Remark 6.6. *For notational ease, we denote $\eta \otimes \eta$ as η^2. We insist on the convention that df^2 always denotes $(df)^2 = df \otimes df$ and not $d(f^2)$. If we want the second then we must put the brackets in.*

Example 6.7. The dot product on $\mathbb{R}^3$ can be viewed as

$$\bar{g} = dx^2 + dy^2 + dz^2$$

Noting that $dx^2 = dx \otimes dx$, etc. □

One huge advantage of viewing an inner product this way is so that we can compute what happens when we change coordinates. How do we do this? Well our inner products are built from 1-forms and 1-forms can be pulled back by maps between spaces.

Example 6.8. Consider $F(r, \theta) = (r\cos\theta, r\sin\theta)$. Let's compute the Euclidean inner product in polar coordinates, i.e. $g_{PC} = F^*\bar{g}$

$$\begin{aligned} F^*\bar{g} = F^*(dx^2 + dy^2) &= (d(x \circ F))^2 + (d(y \circ F))^2 \\ &= (\cos\theta dr - r\sin\theta d\theta) \otimes (\cos\theta dr - r\sin\theta d\theta) \\ &+ (\sin\theta dr + r\cos\theta d\theta) \otimes (\sin\theta dr + r\cos\theta d\theta) \\ &= (\cos^2\theta + \sin^2\theta) dr \otimes dr + (r^2\sin^2\theta + r^2\cos^2\theta) d\theta \otimes d\theta \\ &\quad + ((\cos\theta)(-r\sin\theta) + (\sin\theta)(r\cos\theta))\, dr \otimes d\theta \\ &\quad + ((-r\sin\theta)(\cos\theta) + (r\cos\theta)(\sin\theta))\, d\theta \otimes dr \\ &= dr^2 + r^2 d\theta^2 \end{aligned}$$

Now lets think about what this means. If $X = X_r\frac{\partial}{\partial r} + X_\theta\frac{\partial}{\partial\theta}$ and $Y = Y_r\frac{\partial}{\partial r} + Y_\theta\frac{\partial}{\partial\theta}$ then

$$g_{PC}(X, Y) = (X_r)(Y_r) + r^2(X_\theta)(Y_\theta) = \begin{pmatrix} X_r & X_\theta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix} \begin{pmatrix} Y_r \\ Y_\theta \end{pmatrix}.$$

We note in passing that g_{PC} fails to be a genuine inner product when $r = 0$. This is because $g_{PC}(\frac{\partial}{\partial\theta}, \frac{\partial}{\partial\theta}) = 0$ at any point with $r = 0$. But notice that the polar coordinate function F fails to be regular at $r = 0$. □

We can also think of these in terms of matrices. Let

$$F\colon \mathbb{R}^n_{x_1,\ldots x_n} \to \mathbb{R}^m_{y_1,\ldots,y_m}$$

We can regard the pushforward at p as being the matrix F_* and any inner product g on $\mathbb{R}^m$ is associated to a symmetric, positive matrix $G = (G_{jk})$ by $g = \sum_{j,k=1}^{m} G_{jk} dy_j \otimes dy_k$. Then F^*g is an inner product on $\mathbb{R}^n$ associated to the matrix $F_*^\top G F_*$ by

$$F^*g = \sum_{j,k=1}^{n} (F_*^\top G F_*)_{jk} dx_j \otimes dx_k$$

This is a messy piece of linear algebra. If you want to check that you really understand all of this try to prove it!

Example 6.9. In spherical polars, we can compute $g = F^*\bar{g}$, either by computing with pullbacks

$$\begin{aligned} g &= (\cos\theta\sin\phi\, dr - r\sin\theta\sin\phi\, d\theta + r\cos\theta\cos\phi\, d\phi)^2 \\ &\quad + (\sin\theta\sin\phi\, dr + r\cos\theta\sin\phi\, d\theta + r\sin\theta\cos\phi\, d\phi)^2 \\ &\quad + (\cos\phi\, dr - r\sin\phi\, d\phi)^2 \\ &= dr^2 + (r^2\sin^2\phi)d\theta^2 + r^2 d\phi^2 \end{aligned}$$

(The computation showing that all the cross terms do indeed cancel out is tedious but essentially the same as Example 6.8 and left to the reader.) Or alternatively, we can use the matrix method. The matrix associated to $\bar{g}$ with respect to the coordinates x, y and z is just the identity, so we must compute $F_*^\top(Id)F_*$ or

$$\begin{pmatrix} \cos\theta\sin\phi & \sin\theta\sin\phi & \cos\phi \\ -r\sin\theta\sin\phi & r\cos\theta\sin\phi & 0 \\ r\cos\theta\cos\phi & r\sin\theta\cos\phi & -r\sin\phi \end{pmatrix} \begin{pmatrix} \cos\theta\sin\phi & -r\sin\theta\sin\phi & r\cos\theta\cos\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\phi & 0 & -r\sin\phi \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2\sin^2\phi & 0 \\ 0 & 0 & r^2 \end{pmatrix}.$$

So again we deduce

$$g = F^*\bar{g} = dr^2 + (r^2\sin^2\phi)d\theta^2 + r^2 d\phi^2$$

□

7 $\mathbb{R}^3$: div, grad and curl

Differential forms and the exterior derivative provide a robust mathematical framework to generalize older notation in the calculus of several variables. To describe these we shall need a few more operators. We shall mostly use them when $n = 3$, but they can be defined for all n.

Mathematicians prefer to work in the world of differential forms and the exterior derivative as these provide clean and unifying statements for many old theorems. It is generally preferable to state theorems and results in this new language, but as a useful exercise in manipulating our new tools, we shall work out in detail how to move back and forwards between the new and the old.

Given an inner product g, at every point p, we can define a mapping $\flat$ from $T_p\mathbb{R}^n \to T_p^*\mathbb{R}^n$ by the equation

$$X_p^\flat(v_p) = g(X_p, v_p). \tag{2}$$

As any inner product g is a bilinear map, this does indeed define $X_p^\flat$ as a covector and also implies that $\flat$ itself is a linear map. Since this works at every point p, we can also think of $\flat$ as a map from vector fields to 1-forms.

Example 7.1. Suppose we are working with the inner product $\bar{g}(v,w) = v \cdot w$. Since

$$dx_i\left(\frac{\partial}{\partial x_j}\right) = \delta_{ij} = \frac{\partial}{\partial x_i} \cdot \frac{\partial}{\partial x_j}$$

we can immediately deduce that

$$\left(\frac{\partial}{\partial x_i}\right)^\flat = dx_i \tag{3}$$

From linearity, we then get the useful formula that for $\mathbb{R}^n$ with the inner product $\bar{g}$ we have

$$\left(\sum f_i \frac{\partial}{\partial x_i}\right)^\flat = \sum f_i dx_i. \tag{4}$$

□

Example 7.2. We can work out a formula for $\flat$ in spherical polars (with the dot product) as follows. First recall that

$$g = F^*\bar{g} = dr^2 + (r^2 \sin^2\phi)d\theta^2 + r^2 d\phi^2$$

then

$$g(\frac{\partial}{\partial r}, \frac{\partial}{\partial r}) = 1, \quad g(\frac{\partial}{\partial \theta}, \frac{\partial}{\partial \theta}) = r^2 \sin^2\phi, \quad g(\frac{\partial}{\partial \phi}, \frac{\partial}{\partial \phi}) = r^2$$

and all mixed terms vanish, we have

$$\left(\frac{\partial}{\partial r}\right)^\flat = dr, \quad \left(\frac{\partial}{\partial \theta}\right)^\flat = r^2 \sin^2\theta\, d\theta, \quad \left(\frac{\partial}{\partial \phi}\right)^\flat = r^2\, d\phi$$

Because it is often useful to use an orthonormal frame, the following vector fields are introduced (especially in physics)

$$e_r = \frac{\partial}{\partial r}, \quad e_\theta = \frac{1}{r\sin\theta}\frac{\partial}{\partial \theta}, \quad e_\phi = \frac{1}{r}\frac{\partial}{\partial \phi}$$

If $X = a_1 e_r + a_2 e_\theta + a_3 e_\phi$ and $Y = b_1 e_r + b_2 e_\theta + b_3 e_\phi$ then the inner product $g(X,Y) = \sum a_i b_i$ can be computed just like the dot product. □

There is an exceptionally useful result in linear algebra

Lemma 7.3 (Riesz representation theorem)**.** *If g is an inner product on a finite dimensional vector space V, then for every $\varphi \in V^*$ there exists a unique $x \in V$ such that*

$$\varphi(v) = g(x,v).$$

An immediate consequence of this is that every 1-form can be written as the "flat" of some vector field. Since "flat" is also 1-1. This follows from it being linear and if $X_p^\flat = 0$ then $g(X_p, Y_p) = 0$ for all Y_p, in particular $Y_p = X_p$. But by properties of inner products this means that $X_p = 0$. Thus "flat" is a 1-1 and onto linear map. This means it has an inverse which we call "sharp" or $\#$.

Formulas for "sharp" can often be derived in the same way as for "flat". For the dot product $\bar{g}$, we can immediately deduce

$$\left(\sum f_i dx_i\right)^{\#} = \sum f_i \frac{\partial}{\partial x_i}.$$

Or for spherical polar coordinates

$$dr^{\#} = \frac{\partial}{\partial r} = e_r, \quad d\theta^{\#} = \frac{1}{r^2 \sin^2\theta}\frac{\partial}{\partial\theta} = \frac{1}{r\sin\theta}e_\theta, \quad d\phi^{\#} = \frac{1}{r^2}\frac{\partial}{\partial\phi} = \frac{1}{r}e_\phi$$

There is yet another useful map between vector spaces and differential forms. This time between vector spaces and $(n-1)$-forms. There is a special n-form on Euclidean space called the volume form

$$dV = dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n.$$

For a vector field X, we define an $(n-1)$-form $X \lrcorner dV$ by the formula

$$X \lrcorner dV(Y_1, \ldots Y_{n-1}) = dV(X, Y_1, \ldots Y_{n-1}).$$

Remark 7.4. *This formula can be extended to define $X \lrcorner \eta$ as a $k-1$-form whenever η is a k-form. The new form $X \lrcorner \eta$ is often called the interior derivative of η in the direction of X. This is why we use the seeming redundant adjective "exterior" when talking about the operator d.*

The choice of the special n-form dV actually depends in a subtle fashion on the choice of the inner product.

Example 7.5. When $n = 3$, this new equivalence takes the form

$$\frac{\partial}{\partial x}\lrcorner dV = dy \wedge dz, \quad \frac{\partial}{\partial y}\lrcorner dV = -dx \wedge dz = dz \wedge dx, \quad \frac{\partial}{\partial z}\lrcorner dV = dx \wedge dy$$

so if $X = a\frac{\partial}{\partial x} + b\frac{\partial}{\partial y} + c\frac{\partial}{\partial z}$ then

$$X \lrcorner dV = cdx \wedge dy - bdx \wedge dz + ady \wedge dz.$$

□

Example 7.6. If we want to compute in spherical polar coordinates instead, we first have to pullback the special 3-form dV. But as computed earlier we have

$$F^*dV = r^2 \sin\phi \, dr \wedge d\phi \wedge d\theta.$$

Thus

$$\begin{aligned} &\frac{\partial}{\partial r}\lrcorner F^*dV = r^2\sin\phi\, d\phi \wedge d\theta, \quad \frac{\partial}{\partial\theta}\lrcorner F^*dV = r^2 \sin\phi\, dr \wedge \phi \\ &\frac{\partial}{\partial\phi}\lrcorner F^*dV = -r^2\sin\phi\, dr \wedge d\theta \end{aligned}$$

and equivalently

$$\begin{aligned} &e_r \lrcorner F^*dV = r^2 \sin\phi\, d\phi \wedge d\theta, \quad e_\theta \lrcorner F^*dV = r\, dr \wedge d\phi \\ &e_\phi \lrcorner F^*dV = -r\sin\phi\, dr \wedge d\theta \end{aligned} \tag{5}$$

□

Theorem 7.7. *On $\mathbb{R}^3$, the exterior derivative is related to the gradient, curl and divergence as follows: computing using the standard Euclidean inner product $\bar{g}$*

(a) *grad* $f = df^\#$

(b) *curl* $X \lrcorner dV = d(X^\flat)$

(c) $(\textit{div}\ X)dV = d(X \lrcorner dV)$

Proof: For (a), we compute

$$(df)^\# = \left(\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz\right)^\# = \frac{\partial f}{\partial x}\frac{\partial}{\partial x} + \frac{\partial f}{\partial y}\frac{\partial}{\partial y} + \frac{\partial f}{\partial z}\frac{\partial}{\partial z} = \text{grad } f.$$

For (b), let $X = a\frac{\partial}{\partial x} + b\frac{\partial}{\partial y} + c\frac{\partial}{\partial z}$ then

$$\text{curl } X = \left(\frac{\partial c}{\partial y} - \frac{\partial b}{\partial z}\right)\frac{\partial}{\partial x} + \left(\frac{\partial b}{\partial x} - \frac{\partial a}{\partial z}\right)\frac{\partial}{\partial y} + \left(\frac{\partial a}{\partial y} - \frac{\partial b}{\partial x}\right)\frac{\partial}{\partial z}$$

whereas

$$\begin{aligned} d(X^\flat) &= d\left(adx + bdy + cz\right) \\ &= \left(\frac{\partial c}{\partial y} - \frac{\partial b}{\partial z}\right) dy \wedge dz + \left(\frac{\partial b}{\partial x} - \frac{\partial a}{\partial z}\right) dz \wedge dx + \left(\frac{\partial a}{\partial y} - \frac{\partial b}{\partial x}\right) dx \wedge dy. \end{aligned}$$

The result then follows from the computation of Example 7.5.

For (c), again let $X = a\frac{\partial}{\partial x} + b\frac{\partial}{\partial y} + c\frac{\partial}{\partial z}$ so

$$\begin{aligned} d(X \lrcorner dV) &= d\left(ady \wedge dz - bdx \wedge dz + cdx \wedge dy\right) \\ &= \frac{\partial a}{\partial x}dx \wedge dy \wedge dz - \frac{\partial b}{\partial y}dy \wedge dx \wedge dz + \frac{\partial c}{\partial z}dz \wedge dz \wedge dy \\ &= \left(\frac{\partial a}{\partial x} + \frac{\partial b}{\partial y} + \frac{\partial c}{\partial z}\right) dx \wedge dy \wedge dz \\ &= (\text{div } X)\, dV. \end{aligned}$$

■

Example 7.8. As out final exercise of this section, we shall work out how to compute the gradient, curl and divergence in spherical polar coordinates.

First the gradient: if f is a smooth function in polar coordinates then

$$df = \frac{\partial f}{\partial r}dr + \frac{\partial f}{\partial \theta}d\theta + \frac{\partial f}{\partial \phi}d\phi$$

so using Example 7.2

$$\begin{aligned} \text{grad f} = df^\# &= \frac{\partial f}{\partial r}\frac{\partial}{\partial r} + \frac{1}{r^2 \sin^2\theta}\frac{\partial}{\partial \theta} + \frac{1}{r^2}\frac{\partial}{\partial \phi} \\ &= \frac{\partial f}{\partial r}e_r + \frac{1}{r\sin\theta}e_\theta + \frac{1}{r}e_\phi \end{aligned} \tag{6}$$

Now the curl: let $X = X_r e_r + X_\theta e_\theta + X_\phi e_\phi$. Then

$$\begin{aligned}
d(X^\flat) &= d(X_r\,dr + r\sin\phi X_\theta\,d\theta + rX_\phi\,d\phi) \\
&= \left(-\frac{\partial X_r}{\partial\theta} + \sin\phi\frac{\partial}{\partial r}(rX_\theta)\right) dr\wedge d\theta \\
&\quad + \left(-\frac{\partial X_r}{\partial\phi} + \frac{\partial}{\partial r}(rX_\phi)\right) dr\wedge d\phi \\
&\quad + \left(-r\frac{\partial}{\partial\phi}(\sin\phi X_\theta) + r\frac{\partial X_\phi}{\partial\theta}\right) d\theta\wedge d\phi
\end{aligned}$$

Now we use (5) to see that

$$\begin{aligned}
\text{curl } X &= \frac{1}{r\sin\phi}\left(\frac{\partial}{\partial\phi}(\sin\phi X_\theta) - \frac{\partial X_\phi}{\partial\theta}\right) e_r \\
&\quad + \frac{1}{r}\left(-\frac{\partial X_r}{\partial\phi} + \frac{\partial}{\partial r}(rX_\phi)\right) e_\theta \\
&\quad + \left(\frac{1}{r\sin\phi}\frac{\partial X_r}{\partial\theta} - \frac{1}{r}\frac{\partial}{\partial r}(rX_\theta)\right) e_\phi
\end{aligned} \tag{7}$$

After a break to let my hands recover from typing that, we then tend to the divergence: first recall $F^*dV = r^2\sin\phi\,dr\wedge d\phi\wedge d\theta$.

Now with X as before

$$\begin{aligned}
d(X \lrcorner F^*dV) &= d\left(r^2\sin\phi X_r\,d\phi\wedge d\theta + rX_\theta\,dr\wedge d\phi - r\sin\phi X_\phi\,dr\wedge d\theta\right) \\
&= \left(\frac{\partial}{\partial r}\left(r^2\sin\phi X_r\right) + r\frac{\partial X_\theta}{\partial\theta} + r\frac{\partial}{\partial\phi}(\sin\phi X_\phi)\right) dr\wedge d\phi\wedge d\theta
\end{aligned}$$

Again using $F^*dV = r^2\sin\phi\,dr\wedge d\phi\wedge d\theta$, we deduce

$$\text{div } X = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 X_r\right) + \frac{1}{r\sin\phi}\left(\frac{\partial X_\theta}{\partial\theta} + \frac{\partial}{\partial\phi}(\sin\phi X_\phi)\right) \tag{8}$$

Compare remembering the formulas of (6), (7) and (8) to remembering how to compute the exterior derivative. If that doesn't convince you that differential forms are the way to go then nothing ever will. □