Pushforwards and Pullbacks Tutorial

Boris Ragula

February 2025

Abstract

These notes were designed as supporting material for AMath 475: Introduction to General Relativity at the University of Waterloo. In these notes, we review the concepts of the pushforward of vectors, and the pullback of forms between smooth manifolds. Additionally, we do explicit calculations of the pushforward and pullback given smooth maps between $\mathbb{R}^2$ and $\mathbb{R}^3$, as well as show an example of the commutation between the exterior derivative and the pullback of forms.

1 Pushforwards: The Idea and It's Properties

We'll first start by examining the pushforward of vectors. If we have a smooth map $\varphi : \mathcal{M} \longrightarrow \mathcal{N}$ and v is a vector on $\mathcal{M}$, then we can define a vector $[\varphi_*]_p(\mathsf{v}_p) \in T_{\varphi(p)}N$. This is what is known as the pushforward of v at p by φ. If φ is not injective, as you'll show on the assignment, this map is not guaranteed to define a vector field on $\mathcal{N}$. If φ is a diffeomorphism, then, by definition, we have that $\forall q \in \mathcal{N}$, $q = \varphi(p)$ for some $p \in \mathcal{M}$. This now allows us to define a vector field y on $\mathcal{N}$ using by defining that $\mathsf{y}_q = [\varphi_*]_{\varphi^{-1}(q)}(\mathsf{v}_{\varphi^{-1}(q)})$.

Notice that we require the fact that φ is a diffeomorphism, because in general φ^{-1} may not exist. As a result, when we consider the pushforward by an arbitrary smooth map between manifolds, it will be necessary to consider the pushforward of a vector field at a point.

Finally, we have that at each point $p \in \mathcal{M}$, $[\varphi_*]_p$ acts on $\mathsf{v} \in T_p\mathcal{M}$ as $[\varphi_*]_p(\mathsf{v})(f) = \mathsf{v}_p(f \circ \varphi)$. Notice that if we pick local coordinates $(x^1, \ldots, x^m)$ for $\mathcal{M}$ and $(y^1, \ldots, y^n)$ for $\mathcal{N}$, then the pushforward can be understood as

$$\begin{aligned} ([\varphi_*]_p\mathsf{v})f &= v_p(f \circ \varphi) \\ &= v_p^i \frac{\partial}{\partial x^i}\bigg|_p (f \circ \varphi) \\ &= v_p^i \frac{\partial}{\partial x^i}\bigg|_p (f(y^1(x), \ldots, y^m(x))) \\ &= v_p^i \frac{\partial f}{\partial y^j}\bigg|_{\varphi(p)} \frac{\partial y^j}{\partial x^i}\bigg|_p \end{aligned} \tag{1}$$

So we have

$$[\varphi_*]_p\mathsf{v} = v_p^i \frac{\partial y^j}{\partial x^i}\bigg|_p \frac{\partial}{\partial y^j}\bigg|_{\varphi(p)} \tag{2}$$

Now that we've gotten all these definitions out of the way, let's look at some of the properties of the pushfoward which you may find useful for later calculations.

Property 1: Linearity As a nice warm-up for proving statements about the pushforward, we will show that this is a linear map between tangent spaces of manifolds.

Proof:
Let $a, b \in \mathbb{R}$, $\mathsf{v}_p, \mathsf{y}_p \in T_p\mathcal{M}$, and $h \in C^\infty(\mathcal{N})$. Then,

$$\begin{aligned}\big[[\varphi_*]_p(a\mathsf{v}_p + b\mathsf{y}_p))\big]h &= (a\mathsf{v}_p + b\mathsf{y}_p)(h \circ \varphi) = a\mathsf{v}_p(h \circ \varphi) + b\mathsf{y}_p(h \circ \varphi) \\ &= \big(a[\varphi_*]_p\mathsf{v}_p + b[\varphi_*]_p\mathsf{y}_p\big)h \end{aligned} \tag{3}$$

Since this is true $\forall h \in C^\infty(\mathcal{N})$, we have

$$[\varphi_*]_p(a\mathsf{v}_p + b\mathsf{y}_p)) = a[\varphi_*]_p\mathsf{v}_p + b[\varphi_*]_p\mathsf{y}_p \tag{4}$$

Property 2: Composition Rule If you have a smooth map $\varphi : \mathcal{M} \longrightarrow \mathcal{N}$ and a smooth map $\psi : \mathcal{N} \longrightarrow \mathcal{L}$, then we know that $\psi \circ \varphi : \mathcal{M} \longrightarrow \mathcal{L}$ is a smooth map. Now what if we wanted to pushforward a vector from $\mathcal{M}$ to $\mathcal{L}$?

Proof:
Let $\mathsf{v}_p \in T_p\mathcal{M}$ and $h \in C^\infty(\mathcal{L})$. Then,

$$\big[[(\psi \circ \varphi)_*]_p\mathsf{v}_p h\big] = \mathsf{v}_p(h \circ \psi \circ \varphi) = [\varphi_*]_p\mathsf{v}_p(h \circ \psi) = [\psi_*]_{\varphi(p)}[\varphi_*]_p\mathsf{v}_p h \tag{5}$$

Since this is true $\forall h \in C^\infty(\mathcal{L})$, and $\forall \mathsf{v}_p \in T_p\mathcal{M}$ we obtain

$$\big[(\psi \circ \varphi)_*\big]_p = [\psi_*]_{\varphi(p)}[\varphi_*]_p \tag{6}$$

Now it is again possible to view the pushforward in terms of a coordinate transformation, by choosing local coordinates on the manifolds $\mathcal{M}, \mathcal{N}, \mathcal{L}$. Let $(x^1, \ldots, x^m)$, $(y^1, \ldots, y^n)$, and $(z^1, \ldots, z^l)$ be local coordinates on the manifolds respectively. Then what we can see the pushforward acts as follows:

$$\begin{aligned}\big[[(\psi \circ \varphi)_*]_p v_p h\big] &= v_p^i \frac{\partial}{\partial x^i}(h \circ \psi \circ \varphi(x^1, \ldots, x^m)) \\ &= v_p^i \frac{\partial}{\partial x^i}(h \circ \psi(y^1(\boldsymbol{x}), \ldots, y^n(\boldsymbol{x})) \\ &= v_p^i \frac{\partial}{\partial x^i}(h(z^1(\boldsymbol{y}(\boldsymbol{x})), \ldots, z^l(\boldsymbol{y}(\boldsymbol{x})))) \\ &= v_p^i \frac{\partial y^j}{\partial x^i}\bigg|_p \frac{\partial z^k}{\partial y^j}\bigg|_{\varphi(p)} \frac{\partial h}{\partial z^k}\bigg|_{\psi(\varphi(p))} \end{aligned} \tag{7}$$

Property 3: Pushforward by the Identity
The pushforward by the identity defines an identity map on the tangent space of $\mathcal{M}$ at p. If we have the identity map $\mathrm{Id}_\mathcal{M} : \mathcal{M} \longrightarrow \mathcal{M}$, we would like to know how the pushforward by the Identity works.
Proof:

$$\big[(\mathrm{Id}_\mathcal{M})_*\big]_p\mathsf{v}_p h = \mathsf{v}_p(h \circ \mathrm{Id}_\mathcal{M}) = \mathsf{v}_p h \tag{8}$$

That is,

$$\left[(\mathrm{Id}_{\mathcal{M}})_*\right]_p \mathsf{v}_p = \mathsf{v}_p, \forall \mathsf{v}_p \in T_p\mathcal{M} \tag{9}$$

Thus,

$$\left[(\mathrm{Id}_{\mathcal{M}})_*\right]_p = \mathrm{Id}_{T_p\mathcal{M}} \tag{10}$$

Property 4: Pushforward by a Diffeomorphism
The pushforward by a diffeomorphism defines a linear isomorphism between the tangent space of $\mathcal{M}$ at p and the tangent space of $\mathcal{N}$ at $\varphi(p)$. Moreover, the pushforward by a diffeomorphism defines a smooth vector field on $\mathcal{N}$. So far, we have only considered smooth maps between manifolds of arbitrary dimension. Now let us look at what happens when we pushforward by a diffeomorphism.

Proof:
Since φ is a diffeomorphism, we know that $\exists \varphi^{-1}$ s.t. $\varphi \circ \varphi^{-1} = \mathrm{Id}_{\mathcal{N}}$ and $\varphi^{-1} \circ \varphi = \mathrm{Id}_{\mathcal{M}}$. So we have that

$$\left[(\mathrm{Id}_{\mathcal{M}})_*\right]_p = \left[(\varphi^{-1} \circ \varphi)\right]_p = [\varphi_*^{-1}]_{\varphi(p)}[\varphi_*]_p = \mathrm{Id}_{T_p\mathcal{M}}, \tag{11}$$

$$\left[(\mathrm{Id}_{\mathcal{N}})_*\right]_{\varphi(p)} = \left[(\varphi \circ \varphi^{-1})\right]_{\varphi(p)} = [\varphi_*]_p[\varphi_*^{-1}]_{\varphi(p)} = \mathrm{Id}_{T_{\varphi(p)}\mathcal{N}}, \tag{12}$$

Therefore, we obtain

$$[\varphi_*]_p^{-1} = [\varphi_*^{-1}]_{\varphi(p)} \tag{13}$$

Then since φ is a diffeomorphism $\dim(\mathcal{M}) = \dim(\mathcal{N})$ which implies that $\dim(T_p\mathcal{M}) = \dim(T_{\varphi(p)}\mathcal{N})$. Therefore, the pushforward by a diffeomorphism defines a linear isomorphism between the tangent space at p of $\mathcal{M}$ and the tangent space at $\varphi(p)$ of $\mathcal{N}$. This is true $\forall p \in \mathcal{M}$ and hence we have that the pushfoward by a diffeomorphism defines a smooth vector field on $\mathcal{N}$.

Now that we have seen the important properties of the pushforward, let's look at how we might actually compute the pushforward of a vector at a point. Let $\varphi : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ be the map defined by

$$\varphi(x, y) = (u = x^2 + y^2, v = x^2 - y^2, w = xy). \tag{14}$$

Now, let $p = (1, 2)$, and $\mathsf{v}_p = (3\partial_x - 2\partial_y)_p$. We know the pushforward at a point defines a vector $\mathsf{y}_{\varphi(p)} = a\partial_u + b\partial_v + c\partial_w$ for some $u, v, w \in \mathbb{R}$. We can then compute the pushforward using the definitions that have been given above as follows. First we compute the pushforward of the elements of the basis $\{\partial_x, \partial_y\}$

$$\begin{aligned}
\varphi_*\partial_x &= \frac{\partial u}{\partial x}\partial_u + \frac{\partial v}{\partial x}\partial_v + \frac{\partial w}{\partial x}\partial_w \\
&= 2\partial_u + 2\partial_v + 2\partial_w && (15)\\
\varphi_*\partial_y &= \frac{\partial u}{\partial y}\partial_u + \frac{\partial v}{\partial y}\partial_v + \frac{\partial w}{\partial y}\partial_w \\
&= 4\partial_u - 4\partial_v + \partial_w && (16)
\end{aligned}$$

Now we simply substitute in the results for the basis vectors into our vector v_p to obtain the final result:

$$\begin{aligned}
(3\partial_x - 2\partial_y)_p &= 3\left(2\partial_u + 2\partial_v + 2\partial_w\right)_{\varphi(p)} - 2\left(4\partial_u - 4\partial_v + \partial_w\right)_{\varphi(p)} \\
&= \left(-2\partial_u + 14\partial_v + 4\partial_w\right)_{(5,-3,2)} && (17)
\end{aligned}$$

For another example, let v be the vector $x\partial_y - y\partial_x$, and $\varphi(x, y) = (xy, \sin(x), \cos(y)) = (u, v, w)$.

$$\begin{aligned}
\varphi_*\partial_x &= \frac{\partial u}{\partial x}\frac{\partial}{\partial u} + \frac{\partial v}{\partial x}\frac{\partial}{\partial v} + \frac{\partial w}{\partial x}\frac{\partial}{\partial w} \\
&= y\partial_u + \cos(x)\partial_v \qquad (18) \\
\varphi_*\partial_y &= \frac{\partial u}{\partial y}\frac{\partial}{\partial u} + \frac{\partial v}{\partial y}\frac{\partial}{\partial v} + \frac{\partial w}{\partial y}\frac{\partial}{\partial w} \\
&= x\partial_u - \sin(y)\partial_w \qquad (19)
\end{aligned}$$

So we end up with

$$\begin{aligned}
\varphi_*\mathsf{v} &= x\left(x\partial_u - \sin(y)\partial_w\right) - y\left(y\partial_u + \cos(x)\partial_v\right) \\
&= (x^2 - y^2)\partial_u - x\sin(y)\partial_w - y\cos(x)\partial_v \qquad (20)
\end{aligned}$$

Now, the question is, why do we have an expression that is still contains the terms (x, y, z) when we have mapped the vector $\mathsf{v}_p \in T_p\mathcal{M}$ to a vector $\mathsf{y}_{\varphi(p)} \in T_{\varphi(p)}\mathcal{N}$? Well this has to do with the fact that the map φ in this problem is not a diffeomorphism. As mentioned earlier, this means that, in general, we cannot define a smooth vector field $\mathsf{y} \in T\mathcal{N}$ since it is possible to have more than one vector at each point in $T_p\mathcal{N}$. If, however, our map was a diffeomorphism, then you can actually rewrite the vector in terms of its local coordinates in the codomain of φ, as

$$\varphi_*\mathsf{v} = f(x(u, v, w), y(u, v, w))\partial_u + g(x(u, v, w), y(u, v, w))\partial_v + h(x(u, v, w), y(u, v, w))\partial_w \qquad (21)$$

where $f, g, h \in C^\infty(\mathcal{N})$.

2 Pullbacks: The Idea and It's Properties

Now we'll move on to understanding the pullback of forms. By definition, we have that given a map $\varphi : \mathcal{M} \longrightarrow \mathcal{N}$,a form $\omega \in T^*\mathcal{N}$, local coordinates on $\mathcal{N}$ so that $\omega = \omega_j \mathrm{d}x^j$, then the pullback is given by $\varphi^*\omega = (\omega_j \circ \varphi)\mathrm{d}(x^j \circ \varphi)$. And so we can see that the pullback is actually nothing more than a change of local coordinates from the coordinates of $\mathcal{N}$ to the coordinates of $\mathcal{M}$.

2.1 Neat example with the pullback of an n-form

Consider the n-form $\omega = \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n$, and a smooth map $\varphi = (\varphi^1, \ldots, \varphi^n)$. If we take the pullback of ω by φ we obtain the following:

$$\begin{aligned}
\varphi^*\omega &= \varphi^*(\mathrm{d}x^1) \wedge \cdots \wedge \varphi^*(\mathrm{d}x^n) \\
&= \mathrm{d}(x^1 \circ \varphi) \wedge \cdots \wedge \mathrm{d}(x^n \circ \varphi) \\
&= \mathrm{d}\varphi^1 \wedge \cdots \wedge \mathrm{d}\varphi^n \qquad (22)
\end{aligned}$$

While this appears simple enough, we can actuall simplify this into a much nicer results by actually taking the differentials of each of the φ^i's.

$$
\begin{aligned}
\mathrm{d}\varphi^1 \wedge \cdots \wedge \mathrm{d}\varphi^n &= \frac{\partial \varphi^1}{\partial x^{j_1}} \mathrm{d}x^{j_1} \wedge \cdots \wedge \frac{\partial \varphi^n}{\partial x^{j_n}} \mathrm{d}x^{j_n} \\
&= \frac{\partial \varphi^1}{\partial x^{j_1}} \cdots \frac{\partial \varphi^n}{\partial x^{j_n}} \mathrm{d}x^{j_1} \wedge \cdots \wedge \mathrm{d}x^{j_n} \\
&= \frac{\partial \varphi^1}{\partial x^{j_1}} \cdots \frac{\partial \varphi^n}{\partial x^{j_n}} \epsilon^{j_1 \dots j_n} \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n \\
&= \det\left(\frac{\partial \varphi^i}{\partial x^j}\right) \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n
\end{aligned} \tag{23}
$$

2.2 An example of computing the pullback

Given the map $\varphi : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ defined by $\varphi(u,v) = (u^2, v^2, uv) = (x,y,z)$, and the smooth 1-form $\omega = x\mathrm{d}x + \mathrm{d}y$, the pullback of ω by φ is given by:

$$
\begin{aligned}
\varphi^*\omega &= \varphi^* x \mathrm{d}x + \varphi^* \mathrm{d}y \\
&= (x \circ \varphi)\mathrm{d}(x \circ \varphi) + \mathrm{d}(y \circ \varphi) \\
&= u^2 \mathrm{d}(u^2) + \mathrm{d}(v^2) \\
&= 2u^3 \mathrm{d}u + 2v\mathrm{d}v
\end{aligned} \tag{24}
$$

Now, because the pullback has the property that it distributes over the wedge product, we can take the pullback of higher order forms. We simply apply the rules of the pullback to each term in the wedge product. For example, if we have the 2-form $\omega = \mathrm{d}y \wedge \mathrm{d}z$, and we take the pullback by the same map φ we obtain

$$
\begin{aligned}
\varphi^*\omega &= \varphi^*(\mathrm{d}y \wedge \mathrm{d}z) \\
&= \varphi^*(\mathrm{d}y) \wedge \varphi^*(\mathrm{d}z) \\
&= \mathrm{d}(y \circ \varphi) \wedge \mathrm{d}(z \circ \varphi) \\
&= \mathrm{d}(v^2) \wedge \mathrm{d}(uv) \\
&= 2v\mathrm{d}v \wedge [v\mathrm{d}u + u\mathrm{d}v] \\
&= 2v^2 \mathrm{d}v \wedge \mathrm{d}u
\end{aligned} \tag{25}
$$

2.3 Example of the Pullback Commuting with the Exterior Derivative

Consider the map $\varphi : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ given by $\varphi(x,y) = (x+y, x^2-y^2, xy) = (u,v,w)$. Now consider the one-form on $\mathbb{R}^3$ given by $\omega = uv\mathrm{d}u + w\mathrm{d}v + \mathrm{d}w$. We will show that $\varphi^*(\mathrm{d}\omega) = \mathrm{d}(\varphi^*\omega)$.

$$
\begin{aligned}
\varphi^*(\mathrm{d}\omega) &= \varphi^*(u\mathrm{d}u \wedge \mathrm{d}v) + \varphi^*(\mathrm{d}w \wedge \mathrm{d}v) \\
&= (u \circ \varphi)\mathrm{d}(v \circ \varphi) \wedge \mathrm{d}(u \circ \varphi) + \mathrm{d}(w \circ \varphi) \wedge \mathrm{d}(v \circ \varphi) \\
&= (x+y)\mathrm{d}(x^2 - y^2) \wedge \mathrm{d}(x+y) + \mathrm{d}(xy) \wedge \mathrm{d}(x^2 - y^2) \\
&= 4xy\mathrm{d}x \wedge \mathrm{d}y
\end{aligned} \tag{26}
$$

Now, on the other hand we have

$$\begin{aligned}
\mathrm{d}(\varphi^*\omega) &= \mathrm{d}\left[(uv \circ \varphi)\mathrm{d}(u \circ \varphi) + (w \circ \varphi)\mathrm{d}(v \circ \varphi) + \mathrm{d}(w \circ \varphi)\right] \\
&= \mathrm{d}\left[(x+y)(x^2 - y^2)\mathrm{d}(x+y) + xy\mathrm{d}(x^2 - y^2) + d(xy)\right] \\
&= 4xy\mathrm{d}x \wedge \mathrm{d}y
\end{aligned} \tag{27}$$

Hence we have shown in this example that $\mathrm{d}(\varphi^*\omega) = \varphi^*(\mathrm{d}\omega)$.