This equation shows how contravariant tensor components transform under a change of basis. Again, often contravariant tensors are simply defined as objects that transform according to this equation. Had you been introduced to contravariant tensors this way would it have been obvious to you that a contravariant tensor was really just a vector field?

Finally, we would like to see how basis elements of T^*M transform with this change of basis. Using exterior differentiation on the coordinate functions u^i we have that

$$\begin{aligned}
du^1 &= \frac{\partial u^1}{\partial x^1}dx^1 + \frac{\partial u^1}{\partial x^2}dx^2 + \cdots + \frac{\partial u^1}{\partial x^n}dx^n \\
du^2 &= \frac{\partial u^2}{\partial x^1}dx^1 + \frac{\partial u^2}{\partial x^2}dx^2 + \cdots + \frac{\partial u^2}{\partial x^n}dx^n \\
&\;\vdots \\
du^n &= \frac{\partial u^n}{\partial x^1}dx^1 + \frac{\partial u^n}{\partial x^2}dx^2 + \cdots + \frac{\partial u^n}{\partial x^n}dx^n,
\end{aligned}$$

which of course give the matrix equation

$$\begin{bmatrix} du^1 \\ du^2 \\ \vdots \\ du^n \end{bmatrix} = \begin{bmatrix} \frac{\partial u^1}{\partial x^1} & \frac{\partial u^1}{\partial x^2} & \cdots & \frac{\partial u^1}{\partial x^n} \\ \frac{\partial u^2}{\partial x^1} & \frac{\partial u^2}{\partial x^2} & \cdots & \frac{\partial u^2}{\partial x^n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial u^n}{\partial x^1} & \frac{\partial u^n}{\partial x^2} & \cdots & \frac{\partial u^n}{\partial x^n} \end{bmatrix} \begin{bmatrix} dx^1 \\ dx^2 \\ \vdots \\ dx^n \end{bmatrix}.$$

In Einstein summation notation this becomes

$$\boxed{\text{One-Form Basis Transformation Rule:} \qquad du^i = \frac{\partial u^i}{\partial x^j}dx^j.}$$

So this is the way that basis elements of T^*M transform. Compare this with how components of contravariant tensors transform

$$\widetilde{T}^i = \frac{\partial u^i}{\partial x^j}T^j.$$

Contravariant tensors eat one-forms and contravariant tensor components transform like one-form basis elements.

Let us compare this transformation rule with that of covariant tensors.

Covariant Tensors	Contravariant Tensors
$\widetilde{T}_i = \frac{\partial x^j}{\partial u^i}T_j$	$\widetilde{T}^i = \frac{\partial u^i}{\partial x^j}T^j$
Vector Basis Elements	One-Form Basis Elements
$\frac{\partial}{\partial u^i} = \frac{\partial x^j}{\partial u^i}\frac{\partial}{\partial x^j}$	$du^i = \frac{\partial u^i}{\partial x^j}dx^j$

A.3 Rank-Two Tensors

Before getting into general tensors we will take a closer look at rank-two tensors. There are three possibilities for rank-two tensors:

1. $(0, 2)$-Tensors (Rank-Two Covariant Tensor),
2. $(2, 0)$-Tensors (Rank-Two Contravariant Tensor),
3. $(1, 1)$-Tensors (Mixed-Rank Covariant-Contravariant Tensor).

Rank (0, 2)-Tensors (Rank-Two Covariant Tensors)
Once we get a firm grasp on how these three tensors work moving to the general case is little more than adding indices. First we will consider the $(0, 2)$-tensor, which is a rank two covariant tensor. It is a multilinear map

$$T : TM \times TM \longrightarrow \mathbb{R}.$$

So a $(0, 2)$-tensor takes as input two vectors and gives as output a number. We already know one thing that does exactly this, a two-form. After our success with $(0, 1)$-tensors being one-forms we might be tempted to guess that $(0, 2)$-tensors are two-forms. But if we did that we would be wrong. Two-forms are a subset of $(0, 2)$-tensors but not the whole set. There are many $(0, 2)$-tensors that are not two-forms. In fact, the most important of these is called the metric tensor, which we will talk about in a later section. Similarly, three-forms are a subset of $(0, 3)$-tensors, four-forms are a subset of $(0, 4)$-tensors, etc. Had you learned about tensors first instead of differential forms then differential forms would have been defined to be skew-symmetric covariant tensors.

Here we are getting very close to the abstract mathematical approach to defining tensors. Since that is beyond the scope of this book and we want to avoid that, we will not be mathematically rigorous, which will thereby necessitate some degree of vagueness in the following discussion. However, we hope that you will at least get a general idea of the issues involved.

We begin by defining the Cartesian product of two vector spaces V and W to be

$$V \times W = \Big\{(v, w)\Big| v \in V \text{ and } w \in W\Big\}.$$

So $V \times W$ is the set of pairs of elements, the first of which is from V and the second of which is from W. If $(v_1, \ldots, v_n)$ is a basis of V and $(w_1, \ldots, w_m)$ is a basis of W then we can write

$$V \times W = \text{span}\Big\{(v_i, 0), (0, w_j)\Big| 1 \le i \le n, 1 \le j \le m\Big\}.$$

Let us consider a $(0, 2)$-tensor T which is a mapping $T : TM \times TM \longrightarrow \mathbb{R}$. We might be tempted to say that T is an element of the vector space

$$T^*M \times T^*M = \text{span}\Big\{(dx^i, 0), (0, dx^j)\Big| 1 \le i, j \le n\Big\}.$$

Unfortunately, this actually does not make much sense. Now is the time to remember that tensors are multilinear mappings. This means that the mapping T is linear in each "slot,"

$$T\Big(1^{\text{st}} \text{ slot} , 2^{\text{st}} \text{ slot}\Big).$$

In other words, suppose that $v, v_1, v_2, w, w_1, w_2 \in TM$ are vectors and $a, b \in \mathbb{R}$ are scalars. Then we have that

1. $T\big(av_1 + bv_2, w\big) = aT\big(v_1, w\big) + bT\big(v_2, w\big)$,
2. $T\big(v, aw_1 + bw_2\big) = aT\big(v, w_1\big) + bT\big(v, w_2\big)$.

We want to somehow get these multilinear mappings from the set $T^*M \times T^*M = \text{span}\big\{(dx^i, 0), (0, dx^j)\big\}$. We will completely hand-wave through this bit, but what is done is that something called a free vector space is constructed from this Cartesian product space and then the elements of this space are put into equivalence classes in a procedure similar to what was done when we discussed equivalence classes of curves on manifolds. A great deal of mathematical machinery is employed in this precise but abstract mathematical approach to defining tensors. If you did not follow the last couple paragraphs, don't worry, none of it is necessary for what follows.

The space made up of these equivalence classes is the tensor space $T^*M \otimes T^*M$, which can be thought of as the set of multilinear elements of the form $dx^i \otimes dx^j$, called the tensor product of dx^i and dx^j. Thus we have

$$T \in T^*M \otimes T^*M = \text{span}\left\{ dx^i \otimes dx^j \;\middle|\; 1 \le i, j \le n \right\}.$$

How do the elements of the form $dx^i \otimes dx^j$ work? If v and w are two vector fields on M then

$$dx^i \otimes dx^j(v, w) \equiv dx^i(v) dx^i(w).$$

So, what would $T \in T^*M \otimes T^*M$ actually look like?

For the moment let us suppose that M is a two dimensional manifold. Then we would have

$$T \in T^*M \otimes T^*M = \text{span}\left\{dx^1 \otimes dx^1,\ dx^1 \otimes dx^2,\ dx^2 \otimes dx^1,\ dx^2 \otimes dx^2\right\}$$

so that

$$T = T_{11} dx^1 \otimes dx^1 + T_{12} dx^1 \otimes dx^2 + T_{21} dx^2 \otimes dx^1 + T_{22} dx^2 \otimes dx^2.$$

Pay close attention to the way that the subscript indices work. In general we would write a $(0, 2)$-tensor, using Einstein summation notation, as

$$T = T_{ij} dx^i \otimes dx^j.$$

We could, if we wanted, write the tensor T as a matrix.

Like before suppose we had the change of basis mappings $M_{(x_1, x_2, \ldots, x_n)} \longrightarrow M_{(u_1, u_2, \ldots, u_n)}$ given by the n invertible functions $u_1(x_1, \ldots, x_n), \ldots, u_n(x_1, \ldots, x_n)$. We know how to write

$$T = T_{ij} dx^i \otimes dx^j \in T^*M_{(x_1, x_2, \ldots, x_n)} \otimes T^*M_{(x_1, x_2, \ldots, x_n)}$$

and we want to write T as

$$T = \widetilde{T_{ij}} du^k \otimes du^l \in T^*M_{(u_1, u_2, \ldots, u_n)} \otimes T^*M_{(u_1, u_2, \ldots, u_n)}.$$

That is, we want to see how the components of T transform. Since T is a rank two covariant tensor, each index transforms covariantly and we have

$$\boxed{\text{Covariant Rank 2 Transformation Rule:} \qquad \widetilde{T_{kl}} = \frac{\partial x^i}{\partial u^k} \frac{\partial x^j}{\partial u^l} T_{ij}}$$

Finally, we point out again that the two-forms are the skew-symmetric elements of $T^*M \otimes T^*M$ they are a subset of the $(0, 2)$-tensors,

$$\underbrace{\bigwedge^2(M)}_{\text{two-forms}} \subsetneq \underbrace{T^*M \otimes T^*M}_{(0,2)-\text{tensors}}.$$

What do we mean by that? Well, as we know, for a two-form $dx^i \wedge dx^i$ and vectors v and w, we have that

$$dx^i \wedge dx^i(v, w) = -dx^i \wedge dx^i(w, v).$$

The sign of our answer changes if we switch the order of our input vectors. This need not happen for a general tensor. For a general tensor there is no reason why we should expect the value $dx^i \otimes dx^j(v, w)$ to be related to the value $dx^i \otimes dx^j(w, v)$. They may not be related at all, much less related by a simple sign change. The tensors for which this happens, that is, the tensors for which

$$dx^i \otimes dx^j(v, w) = -dx^i \otimes dx^j(w, v)$$

are said to be skew-symmetric, or anti-symmetric. Tensors that have this property are a very special and unusual subset of the set of all tensors. The covariant tensors that have this property are so special that they have their own name, differential forms.

Rank (2, 0)-Tensors (Rank-Two Contravariant Tensors)
Now we consider the (2, 0)-tensor, which is a rank-two contravariant tensor. This is a multilinear map

$$T : T^*M \times T^*M \longrightarrow \mathbb{R}.$$

So a (2, 0)-tensor takes as input two one-forms and gives as output a number. Without repeating the whole discussion from above, something analogous happens. The (2, 0)-tensor is an element of the vector space

$$TM \otimes TM = \text{span}\left\{ \frac{\partial}{\partial x^i} \otimes \frac{\partial}{\partial x^j} \;\Big|\; 1 \le i, j \le n \right\},$$

which consists of the multilinear elements of the vector product space $TM \times TM$. Consider a two-dimensional manifold M. Then we have

$$T \in TM \otimes TM = \text{span}\left\{ \frac{\partial}{\partial x^1} \otimes \frac{\partial}{\partial x^1}, \; \frac{\partial}{\partial x^1} \otimes \frac{\partial}{\partial x^2}, \; \frac{\partial}{\partial x^2} \otimes \frac{\partial}{\partial x^1}, \; \frac{\partial}{\partial x^2} \otimes \frac{\partial}{\partial x^2} \right\},$$

so that

$$T = T^{11} \frac{\partial}{\partial x^1} \otimes \frac{\partial}{\partial x^1} + T^{12} \frac{\partial}{\partial x^1} \otimes \frac{\partial}{\partial x^2} + T^{21} \frac{\partial}{\partial x^2} \otimes \frac{\partial}{\partial x^1} + T^{22} \frac{\partial}{\partial x^2} \otimes \frac{\partial}{\partial x^2}.$$

In general we would write a (2, 0)-tensor, in Einstein summation notation, as

$$T = T^{ij} \frac{\partial}{\partial x^i} \otimes \frac{\partial}{\partial x^j}.$$

Like before suppose we have an invertible change of basis $M_{(x_1,x_2,\ldots,x_n)} \longrightarrow M_{(u_1,u_2,\ldots,u_n)}$ and we know how to write

$$T = T^{ij} \frac{\partial}{\partial x^i} \otimes \frac{\partial}{\partial x^j} \in TM_{(x_1,x_2,\ldots,x_n)} \otimes TM_{(x_1,x_2,\ldots,x_n)}$$

and we want to write T as

$$T = \widetilde{T^{kl}} \frac{\partial}{\partial u^k} \otimes \frac{\partial}{\partial u^l} \in TM_{(u_1,u_2,\ldots,u_n)} \otimes TM_{(u_1,u_2,\ldots,u_n)}.$$

That is, we want to see how the components of T transform. Since T is a rank two contravariant tensor, each index transforms contravariantly and we have

$$\boxed{\text{Contravariant Rank 2 Transformation Rule:} \qquad \widetilde{T^{kl}} = \frac{\partial u^k}{\partial x^i} \frac{\partial u^l}{\partial x^j} T^{ij}.}$$

Rank (1, 1)-Tensors (Mixed Rank Tensors)
Now we will look at a (1, 1)-tensor, or a mixed-rank covariant-contravariant tensor. This tensor is a multilinear map

$$T : T^*M \times TM \longrightarrow \mathbb{R}.$$

A $(1, 1)$-tensor takes as input one vector and one one-form and gives as output a number. Again, the $(1, 1)$-tensor is an element of the vector space

$$TM \otimes T^*M = \text{span}\left\{ \frac{\partial}{\partial x^i} \otimes dx^j \;\Big|\; 1 \leq i, j \leq n \right\},$$

which consists of the multilinear elements of the vector product space $TM \times T^*M$. Considering a two dimensional manifold M we have that

$$T \in TM \otimes T^*M = \text{span}\left\{ \frac{\partial}{\partial x^1} \otimes dx^1,\ \frac{\partial}{\partial x^1} \otimes dx^2,\ \frac{\partial}{\partial x^2} \otimes dx^1,\ \frac{\partial}{\partial x^2} \otimes dx^2 \right\},$$

so that

$$T = T_1^1 \frac{\partial}{\partial x^1} \otimes dx^1 + T_2^1 \frac{\partial}{\partial x^1} \otimes dx^2 + T_1^2 \frac{\partial}{\partial x^2} \otimes dx^1 + T_2^2 \frac{\partial}{\partial x^2} \otimes dx^2.$$

In general we would write T, using Einstein summation notation, as

$$T = T_j^i \frac{\partial}{\partial x^i} \otimes dx^j.$$

Like before suppose we have an invertible change of basis $M_{(x_1, x_2, \ldots, x_n)} \longrightarrow M_{(u_1, u_2, \ldots, u_n)}$ and we know how to write

$$T = T_j^i \frac{\partial}{\partial x^i} \otimes dx^j \in TM_{(x_1, x_2, \ldots, x_n)} \otimes T^*M_{(x_1, x_2, \ldots, x_n)}$$

and we want to write T as

$$T = \widetilde{T_l^k} \frac{\partial}{\partial u^k} \otimes du^l \in TM_{(u_1, u_2, \ldots, u_n)} \otimes T^*M_{(u_1, u_2, \ldots, u_n)}.$$

That is, we want to see how the components of T transform. The upper index transforms contravariantly and the lower index transforms covariantly. Thus we have

$$\boxed{\text{Rank } (1, 1)\text{-Tensor Transformation Rule:} \qquad \widetilde{T_l^k} = \frac{\partial u^k}{\partial x^i} \frac{\partial x^j}{\partial u^l} T_j^i.}$$

A.4 General Tensors

Now we will finally take a look at a general (r, s)-tensor; it is a multilinear map

$$\mathscr{T} : \underbrace{T^*M \times \cdots \times T^*M}_{\substack{r \\ \text{contravariant} \\ \text{degree}}} \times \underbrace{TM \times \cdots \times TM}_{\substack{s \\ \text{covariant} \\ \text{degree}}} \longrightarrow \mathbb{R}.$$

This means that the general (r, s)-tensor is an element of the space

$$\begin{aligned} \mathscr{T} &\in \underbrace{TM \otimes \cdots \otimes TM}_{r} \otimes \underbrace{T^*M \otimes \cdots \otimes T^*M}_{s} \\ &= \text{span}\left\{ \frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_s} \Big| 1 \leq i_1, \ldots, i_r, j_1, \ldots, j_s \leq n \right\}. \end{aligned}$$

A general (r, s)-tensor is written

$$\begin{aligned}\mathscr{T} &= \mathscr{T}^{i_1\cdots i_r}_{j_1\cdots j_s} \frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_s} \\ &= \mathscr{T}^{i_1\cdots i_r}_{j_1\cdots j_s} \partial x^{i_1} \otimes \cdots \otimes \partial x^{i_r} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_s}.\end{aligned}$$

Given the same invertible change of basis $M_{(x_1,x_2,\ldots,x_n)} \longrightarrow M_{(u_1,u_2,\ldots,u_n)}$ we know how to write

$$\mathscr{T} = \widetilde{\mathscr{T}}^{k_1\cdots k_r}_{l_1\cdots l_s} \frac{\partial}{\partial u^{k_1}} \otimes \cdots \otimes \frac{\partial}{\partial u^{k_r}} \otimes du^{l_1} \otimes \cdots \otimes du^{l_s}.$$

Tensor $\mathscr{T}$'s components transform according to

$$\boxed{\text{Rank } (r, w)\text{-Tensor Transformation Rule:} \qquad \widetilde{\mathscr{T}}^{k_1\cdots k_r}_{l_1\cdots l_s} = \frac{\partial u^{k_1}}{\partial x^{i_1}} \cdots \frac{\partial u^{k_r}}{\partial x^{i_r}} \frac{\partial x^{j_1}}{\partial u^{l_1}} \cdots \frac{\partial x^{j_s}}{\partial u^{l_s}} \mathscr{T}^{i_1\cdots i_r}_{j_1\cdots j_s}.}$$

Suppose that we have a mapping $\phi : M \to M$. The pullback of a rank $(0, t)$-tensor $\mathscr{T}$ at the point p is defined exactly as we defined the pullback of differential forms, by

$$\big(\phi^* \mathscr{T}_{\phi(p)}\big)_p (v_{1_p}, \ldots, v_{t_p}) = \mathscr{T}_{\phi(p)}(\phi_* v_{1_p}, \ldots, \phi_* v_{t_p}).$$

Now suppose we had two covariant tensors, a rank $(0, t)$-tensor $\mathscr{T}$ and a rank $(0, s)$-tensor $\mathscr{S}$. Then we have

$$\begin{aligned}&\Big(\phi^*\big(\mathscr{T} \otimes \mathscr{S}\big)_{\phi(p)}\Big)_p (v_{1_p}, \ldots, v_{(t+s)_p}) \\ &\quad = \big(\mathscr{T} \otimes \mathscr{S}\big)_{\phi(p)} (\phi_* v_{1_p}, \ldots, \phi_* v_{(t+s)_p}) \\ &\quad = \mathscr{T}_{\phi(p)}(\phi_* v_{1_p}, \ldots, \phi_* v_{t_p}) \mathscr{S}_{\phi(p)}(\phi_* v_{(t+1)_p}, \ldots, \phi_* v_{(t+s)_p}) \\ &\quad = \phi^* \mathscr{T}_{\phi(p)}(v_{1_p}, \ldots, v_{t_p}) \phi^* \mathscr{S}_{\phi(p)}(v_{(t+1)_p}, \ldots, v_{(t+s)_p}) \\ &\quad = \big(\phi^* \mathscr{T}_{\phi(p)} \otimes \phi^* \mathscr{S}_{\phi(p)}\big)(v_{1_p}, \ldots, v_{(t+s)_p}).\end{aligned}$$

Thus we have shown for covariant tensors that

$$\phi^*\big(\mathscr{T} \otimes \mathscr{S}\big) = \phi^* \mathscr{T} \otimes \phi^* \mathscr{S}.$$

Let us now consider contravariant tensors. Even though the building blocks of contravariant tensors are vector fields, and we generally think of pushing-forward vector fields, we will still think in terms of pulling back contravariant tensors. The reason we do this is because we will want to pull-back mixed rank tensor fields as well, and in order to define the pullback of a mixed rank tensor field we have to know how to pullback a contravariant tensor field. It would be entirely possible to alter the definitions so that we always think of pushing forward tensor fields, but the general convention is to think of pulling back tensor fields.

We define the pullback of a rank $(t, 0)$-tensor by

$$\big(\phi^* \mathscr{T}_{\phi(p)}\big)_p (\alpha_{1_p}, \ldots, \alpha_{t_p}) = \mathscr{T}_{\phi(p)}\big((\phi^{-1})^* \alpha_{1_p}, \ldots, (\phi^{-1})^* \alpha_{t_p}\big).$$

Now suppose we had two contravariant tensors, a rank $(t, 0)$-tensor $\mathscr{T}$ and a rank $(s, 0)$-tensor $\mathscr{S}$. Then, using an identical line of reasoning, we have

$$\begin{aligned}&\Big(\phi^*\big(\mathscr{T} \otimes \mathscr{S}\big)_{\phi(p)}\Big)_p (\alpha_{1_p}, \ldots, \alpha_{(t+s)_p}) \\ &\quad = \big(\mathscr{T} \otimes \mathscr{S}\big)_{\phi(p)} \big((\phi^{-1})^* \alpha_{1_p}, \ldots, (\phi^{-1})^* \alpha_{(t+s)_p}\big) \\ &\quad = \mathscr{T}_{\phi(p)}\big((\phi^{-1})^* \alpha_{1_p}, \ldots, (\phi^{-1})^* \alpha_{t_p}\big) \mathscr{S}_{\phi(p)}\big((\phi^{-1})^* \alpha_{(t+1)_p}, \ldots, (\phi^{-1})^* \alpha_{(t+s)_p}\big)\end{aligned}$$

$$= \phi^* \mathscr{T}_{\phi(p)}(\alpha_{1_p}, \ldots, \alpha_{t_p}) \phi^* \mathscr{S}_{\phi(p)}(\alpha_{(t+1)_p}, \ldots, \alpha_{(t+s)_p})$$
$$= \big(\phi^* \mathscr{T}_{\phi(p)} \otimes \phi^* \mathscr{S}_{\phi(p)}\big)(\alpha_{1_p}, \ldots, \alpha_{(t+s)_p}).$$

Thus we have also shown for contravariant tensors that

$$\phi^*\big(\mathscr{T} \otimes \mathscr{S}\big) = \phi^* \mathscr{T} \otimes \phi^* \mathscr{S}.$$

Now we consider a mixed rank (s, t)-tensor $\mathscr{T}$. We define the pullback of a (s, t)-tensor as follows

$$\big(\phi^* \mathscr{T}_{\phi(p)}\big)_p (\alpha_{1_p}, \ldots, \alpha_{s_p}, v_{1_p}, \ldots, v_{t_p})$$
$$= \mathscr{T}_{\phi(p)}\big((\phi^{-1})^* \alpha_{1_p}, \ldots, (\phi^{-1})^* \alpha_{t_p}, \phi_* v_{1_p}, \ldots, \phi_* v_{t_p}\big).$$

Question A.1 Given a rank (q, r)-tensor $\mathscr{T}$ and a rank (s, t)-tensor $\mathscr{S}$, show that

$$\phi^*\big(\mathscr{T} \otimes \mathscr{S}\big) = \phi^* \mathscr{T} \otimes \phi^* \mathscr{S}.$$

Putting this all together, for any tensors $\mathscr{T}$ and $\mathscr{S}$ we have

$$\boxed{\phi^*\big(\mathscr{T} \otimes \mathscr{S}\big) = \phi^* \mathscr{T} \otimes \phi^* \mathscr{S}.}$$

A.5 Differential Forms as Skew-Symmetric Tensors

Quite often you will find books introduce tensors first and then introduce and define differential forms in terms of tensors. This is of course not the approach we have taken in this book, but we would like to devote a section to the definition of k-forms in terms of tensors for completeness sake. Also, some of the formulas look quite different from this perspective and it is important that you at least be exposed to these so that when you do see them in the future you will have some idea of where they came from.

Suppose we have a totally covariant tensor $\mathscr{T}$ of rank k, that is, a $(0, k)$-tensor. Thus, $\mathscr{T} \in \underbrace{T^*M \otimes \cdots \otimes T^*M}_{k}$. Since $\mathscr{T}$ is a tensor it is multilinear. But differential forms are skew-symmetric in addition to being multilinear. The tensor $\mathscr{T}$ is called anti-symmetric or skew-symmetric if it changes sign whenever any pair of its arguments are switched. That is,

$$\mathscr{T}\big(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k\big) = -\mathscr{T}\big(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k\big)$$

when v_i and v_j have switched their places and all other inputs have stated the same. Sometimes skew-symmetry is defined in terms of permutations π of $(1, 2, 3, \ldots, k)$. Thus, T is skew-symmetric if

$$\mathscr{T}\big(v_{\pi(1)}, v_{\pi(2)}, \ldots, v_{\pi(k)}\big) = \text{sgn}(\pi) \mathscr{T}\big(v_1, v_2, \ldots, v_k\big)$$

$$\boxed{\text{A } k\text{-form is a skew-symmetric rank } k \text{ covariant tensor.}}$$

Thus we have that the set of k-forms is a subset of the set of $(0, k)$-tensors,

$$\underbrace{\bigwedge^{s}(M)}_{k-\text{forms}} \subsetneq \underbrace{T^*M \otimes \cdots \otimes T^*M}_{(0,k)-\text{tensors}}.$$

Question A.2 Show that for a $(0,2)$-tensor this implies that $\mathscr{T}_{ij} = -\mathscr{T}_{ji}$.

Question A.3 For $\mathscr{T}$ a $(0,3)$-tensor show that the two definitions of skew-symmetry are equivalent. Then show the two definitions are equivalent for a $(0,4)$-tensor.

In Sect. 3.3.3 on the general formulas for the wedgeproduct, we have the following formula for the wedgeproduct of a k-form α and an ℓ-form β,

$$(\alpha \wedge \beta)\big(v_1, \ldots, v_{k+\ell}\big) = \frac{1}{k!\ell!} \sum_{\sigma \in S_{k+\ell}} \operatorname{sgn}(\sigma)\alpha\big(v_{\sigma(1)}, \ldots, v_{\sigma(k)}\big)\beta\big(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+\ell)}\big).$$

In that section we then proceeded to show that this definition was equivalent to the determinant based formula that finds volumes of a parallelepiped. We also said that an alternative definition of the wedgeproduct, given in terms of tensors, was

$$\alpha \wedge \beta = \frac{(k+\ell)!}{k!\ell!}\mathscr{A}(\alpha \otimes \beta),$$

where $\mathscr{A}$ was the skew-symmetrization operator. We said that we would explain that formula in the appendix on tensors.

We begin by defining the skew-symmetrization operator. The skew-symmetrization operator takes a p-tensor $\mathscr{T}$ and turns it into a skew-symmetric tensor $\mathscr{A}(\mathscr{T})$ according to the following formula,

$$\mathscr{A}(\mathscr{T})(v_1, \ldots, v_p) = \frac{1}{p!} \sum_{\pi \in S_p} \operatorname{sgn}(\pi)\ \mathscr{T}(v_{\pi(1)}, \ldots, v_{\pi(p)}),$$

where the sum is over all the permutations π of p elements. We divide by $p!$ because there are $p!$ different permutations of p elements, that is, $|S_p| = p!$, and hence there are $p!$ terms in the sum.

Question A.4 Show that $\mathscr{A}(\mathscr{T})$ is skew-symmetric when $\mathscr{T} = dx^1 \otimes dx^2$; when $\mathscr{T} = dx^1 \otimes dx^3 + dx^2 \otimes dx^3$; when $\mathscr{T} = dx^1 \otimes dx^2 \otimes dx^3$.

We now show that the definition of the wedgeproduct in terms of the skew-symmetrization operator is equivalent to the other definition from the previous chapter,

$$\begin{aligned}
\alpha \wedge \beta(v_1, \ldots, v_{k+\ell}) &= \frac{(k+\ell)!}{k!\ell!}\mathscr{A}(\alpha \otimes \beta)(v_1, \ldots, v_{k+\ell}) \\
&= \frac{(k+\ell)!}{k!\ell!} \frac{1}{(k+\ell)!} \sum_{\pi \in S_{k+\ell}} \operatorname{sgn}(\pi)\ (\alpha \otimes \beta)(v_{\pi(1)}, \ldots, v_{\pi(k+l)}) \\
&= \frac{1}{k!\ell!} \sum_{\pi \in S_{k+\ell}} \operatorname{sgn}(\pi)\ \alpha(v_{\pi(1)}, \ldots, v_{\pi(k)})\beta(v_{\pi(k+1)}, \ldots, v_{\pi(k+l)}).
\end{aligned}$$

Thus we see that this definition of the wedgeproduct is exactly equivalent to the other definitions we have used throughout this book. It should be noted that the various properties of the wedgeproduct can also be proved using this definition, though we will not attempt to do so here. In books that introduce tensors first, the above formula is often given as the definition of the wedgeproduct.

We finally turn to showing that the pullback of the wedgeproduct of two forms is the wedgeproduct of the pullbacks of the two forms. Another way of saying this is that pullbacks distribute across wedgeproducts. The proof of this fact proceeds quite nicely and cleanly when we use the tensor definition of the wedgeproduct. Dropping the base point from the notation, pullback of the sum of two covariant $(0,t)$-tensors is written as

$$\begin{aligned}
\phi^*\big(\mathscr{T} + \mathscr{S}\big)(v_1, \ldots, v_t) &= \big(\mathscr{T} + \mathscr{S}\big)(\phi_* v_1, \ldots, \phi_* v_t) \\
&= \mathscr{T}(\phi_* v_1, \ldots, \phi_* v_t) + \mathscr{S}(\phi_* v_1, \ldots, \phi_* v_t) \\
&= \phi^*\mathscr{T}(v_1, \ldots, v_t) + \phi^*\mathscr{S}(v_1, \ldots, v_t).
\end{aligned}$$

Hence we have $\phi^*(\mathscr{T} + \mathscr{S}) = \phi^*\mathscr{T} + \phi^*\mathscr{S}$, which means that pullbacks distribute over addition.

$$
\begin{aligned}
\mathscr{A}(\phi^*\mathscr{T})(v_1, \dots, v_p) &= \frac{1}{p!} \sum_{\pi \in S_p} \operatorname{sgn}(\pi)\ \phi^*\mathscr{T}(v_{\pi(1)}, \dots, v_{\pi(p)}) \\
&= \frac{1}{p!}\phi^* \left(\sum_{\pi \in S_p} \operatorname{sgn}(\pi)\ \mathscr{T}(v_{\pi(1)}, \dots, v_{\pi(p)}) \right) \\
&= \phi^* \left(\frac{1}{p!} \sum_{\pi \in S_p} \operatorname{sgn}(\pi)\ \phi^*\mathscr{T}(v_{\pi(1)}, \dots, v_{\pi(p)}) \right) \\
&= \phi^*\mathscr{A}(\mathscr{T})(v_1, \dots, v_p)
\end{aligned}
$$

and so we have $\mathscr{A}\phi^* = \phi^*\mathscr{A}$.

Now suppose we have two skew-symmetric covariant rank $(0, k)$-tensors, in other words, two k-forms, α and β. Then using the identity $\phi^*(\mathscr{T} \otimes \mathscr{S}) = \phi^*\mathscr{T} \otimes \phi^*\mathscr{S}$ we derived at the end of the last section we have

$$
\begin{aligned}
\phi^*(\alpha \wedge \beta) &= \phi^* \left(\frac{(k+\ell)!}{k!\ell!} \mathscr{A}(\alpha \otimes \beta) \right) \\
&= \frac{(k+\ell)!}{k!\ell!} \phi^*\mathscr{A}(\alpha \otimes \beta) \\
&= \frac{(k+\ell)!}{k!\ell!} \mathscr{A}\phi^*(\alpha \otimes \beta) \\
&= \frac{(k+\ell)!}{k!\ell!} \mathscr{A}(\phi^*\alpha \otimes \phi^*\beta) \\
&= \phi^*\alpha \wedge \phi^*\beta.
\end{aligned}
$$

Hence we have the important identity

$$
\boxed{\phi^*(\alpha \wedge \beta) = \phi^*\alpha \wedge \phi^*\beta.}
$$

So we have shown that pullbacks distribute across wedgeproducts.

A.6 The Metric Tensor

In this section we finally get around to defining and introducing metrics on manifolds, something that we promised to do in the introduction to the chapter on manifolds. We have also alluded to metrics several times as well, particularly in the chapter on electromagnetism when we discussed special relativity. There we gave a very informal introduction to both the Euclidian metric and the Minkowski metric as matrices that were used in finding the inner product of two vectors. The concept of a metric is extremely important and very fundamental. It underlies many of the operations and concepts that you have worked with throughout both this book and your calculus classes. But there is a good possibility that before this book you have not even seen the word.

There are a couple good and practical reasons why you probably have never been formally introduced to metrics before. The first is that most of high school math and geometry, as well as calculus, is built on our comfortable familiar Euclidian spaces, and Euclidian space, particularly when it comes equipped with our nice comfortable familiar Cartesian coordinates, has the Euclidian metric seamlessly built in. The Euclidian metric is built into the formulas that you have always used for the distance between two points, or the length or norm of a vector, or the inner product between two vectors, or even into our ideas of what the shortest distance between two points is or our idea of what a straight line is. All of these things rely, implicitly, on the Euclidian metric.

Another reason is that as there are only so many new ideas and concepts that a human brain can assimilate at once, and when you are trying to learn calculus in $\mathbb{R}^n$ it is best that you concentrate on learning calculus. Side tracks into the theoretical underpinnings as to why things work would only muddy up the calculus concepts that you are trying to learn. And of course, unless you are going to go on to become a mathematician or a theoretical physicists, there is actually little need to go off and deal with other sorts of spaces or manifolds.

And finally, as you probably know, for a long time people thought Euclidian geometry was the only kind of geometry there was. Even after they were discovered, it took many decades for people to become comfortable that there were other kinds of "non-Euclidian" geometry out there, in particular hyperbolic geometry and elliptic geometry. And then later on differential geometry, and in particular Riemannian geometry were introduced. It actually took some time for mathematician to get all the underlying theoretical concepts, like metrics, sorted out. Thus the idea of metrics is several centuries more recent than the major ideas of calculus.

Now onto the definition of a metric. A ***metric*** *on the manifold M is a smooth, symmetric, non-degenerate, rank-two covariant tensor g*. Metric tensors are generally denoted with a lower-case g. We explain each of these terms in turn. Suppose we had two smooth vector fields v and w. Recall that in each coordinate neighborhood of M we have a coordinate system $(x_1, \ldots, x_n)$, which allows us to write the vectors as $v = v^i(x_1, \ldots, x_n)\partial_{x^i}$ and $w = w^i(x_1, \ldots, x_n)\partial_{x^j}$, where Einstein summation notation is used. The vector field v is called smooth if each of the functions $v^i(x_1, \ldots, x_n)$ is differentiable an infinite number of times with respect to the arguments $x_1, \ldots, x_n$.

The $(0, 2)$-tensor g is called **smooth** if for the smooth vector fields v and w then the real-valued function $g(v, w)$ is also differentiable an infinite number of times. Since g is a $(0, 2)$-tensor we can write $g = g_{ij}(x_1, \ldots, x_n)dx^i \otimes dx^i$, so another way to say that g is smooth is to say that the component functions $g_{ij}(x_1, \ldots, x_n)$ are infinitely differentiable in the arguments. We have already discussed skew-symmetry, so symmetry should be clear. The tensor g is symmetric if $g(v, w) = g(w, v)$ for all smooth vector fields v and w. Non-degeneracy is a little more complicated, and is discussed in a little more detail when the symplectic form is introduced in Sect. B.3, but we will state its definition here. Suppose at some point $p \in M$ that $g_p(v_p, w_p) = 0$ for any possible choice of vector v_p. If g is non-degenerate at p then that means the only way that this could happen is if the vector $w_p = 0$. g is called non-degenerate if it is non-degenerate at every point $p \in M$.

A manifold that has such a tensor g on it is called a **pseudo-Riemannian manifold** and the tensor g is called the metric or sometimes the **pseudo-Riemannian metric**. If the metric g also has one additional property, that $g(v, w) \geq 0$ for all vector fields v and w then it is called a **Riemannian metric** and the manifold is called a **Riemannian manifold**.

Question A.5 Is the Minkowski metric from Sect. 12.3 a Riemannian metric? Why or why not?

The metric tensor g gives an inner product on every vector space T_pM in the tangent bundle of M. The inner product of $v_p, w_p \in T_pM$ is given by $g(v_p, w_p)$. Most often the inner product of two vectors is denoted with $\langle \cdot, \cdot \rangle$ where

$$\langle v_p, w_p \rangle \equiv g(v_p, w_p).$$

For basis vectors we clearly have

$$\begin{aligned}
\left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle &\equiv g\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) \\
&= g_{k\ell}dx^k \otimes dx^\ell \left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) \\
&= g_{ij}.
\end{aligned}$$

When $g_{ij} = \delta_{ij}$ then g is called the Euclidian metric, the associated inner product is called the Euclidian inner product, which is none other than the dot product that you are familiar with.

Since g is a rank-two tensor we can write g as a matrix. We start by noticing

$$\begin{aligned}
g(v, w) &= g(v^i\partial_{x^i}, w^j\partial_{x^j}) \\
&= v^i w^j g(\partial_{x^i}, \partial_{x^j}) \\
&= v^i w^j g_{ij}.
\end{aligned}$$

Another way of writing this would be as

$$\begin{aligned} g(v, w) &= [v^1, v^2, \cdots, v^n] \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{bmatrix} \begin{bmatrix} w^1 \\ w^2 \\ \vdots \\ w^n \end{bmatrix} \\ &= v^T [g_{ij}] w. \end{aligned}$$

Often you see metrics introduced or given simply as a matrix, as we did earlier when discussing the Euclidian and Minkowski metrics in the chapter on electromagnetism. This particularly happens when one wishes to avoid any discussion of tensors. In the case of the Euclidian metric, if $g_{ij} = \delta_{ij}$ then the matrix $[g_{ij}]$ is none other than the identity matrix, which is exactly the matrix which we had given for the Euclidian metric in the chapter on electromagnetism in Sect. 12.3. But this kind of representation of the metric implicitly depends on a particular coordinate system being used. In the case where the Euclidian metric was represented as the identity matrix the Cartesian coordinate system was being used. To find the matrix representation of a metric in another coordinate system one would have to perform a coordinate transformation of the metric tensor.

We can also use the metric to define what the length of a vector is. The length of a vector is also called the norm of the vector, and is most often denoted by $\|\cdot\|$. The length of a vector v_p at a point p is defined to be

$$\|v_p\| \equiv \sqrt{|g(v_p, v_p)|},$$

where the $|\cdot|$ in the square root is simply the absolute value, which is necessary if g is a pseudo-Riemannian metric instead of a Riemannian metric.

Question A.6 Show that the Euclidian metric gives the usual dot product of two vectors, $g(v, w) = v \cdot w$. In particular, show that the matrix which defined the dot product in Chap. 5 is exactly the Euclidian metric on $\mathbb{R}^3$ defined above by having $g_{ij} = \delta_{ij}$.

Question A.7 Recall that the Minkowski metric is given by the matrix

$$\begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Show that the absolute value in the definition of the norm of a vector is necessary for this metric.

When a manifold has a metric defined on it we can also use this metric to find the distance between two points on the manifold, at least in theory. Suppose we have two points p and q that are both in the same coordinate patch of M and they are connected by a curve $\gamma : [a, b] \subset \mathbb{R} \to M$ where $\gamma(a) = p$ and $\gamma(b) = q$. The curve $\gamma(t)$ has tangent velocity vectors $\dot{\gamma}(t)$ along the curve. To ensure the tangent velocity vectors actually exist at the endpoints the curve needs to be extended a tiny amount ϵ to $(a - \epsilon, b + \epsilon) \subset \mathbb{R}$. The length of the curve γ from p to q is defined to be

$$L(\gamma) = \int_a^b \sqrt{\left|g\big(\dot{\gamma}(t), \dot{\gamma}(t)\big)\right|}\, dt.$$

We will not actually try to solve any equations of this nature, though in all probability you did solve a few similar to this in your multi-variable calculus course. It can be shown that the length of the curve, $L(\gamma)$, is independent of the parametrization we use to define the curve, though we will not show that here. We also point out that this is an example of an integral where the integrand is not a differential form. As you are of course aware of by this point, differential forms are very natural things to integrate because they handle changes in variable so well, but they are not the only objects that can be integrated - this example is a case in point.

We then define the distance between the two points p and q as

$$d(p,q) = \inf_{\gamma} L(\gamma),$$

where γ is any piecewise continuous curve that connects p and q. If you don't know what infimum is, it is basically the lower limit of lengths of all possible curves that connect p and q. This is of course a nice theoretical definition for the distance between two points, but in general one certainly wouldn't want to compute the distance between any two points this way. This procedure turns out to give exactly the distance formula that you are used to from Euclidian geometry.

However, the main point to understand is that distances between points on a manifold rely on the fact that there is a metric defined on the manifold. If a manifold does not have a metric on it the very idea distances between points simply does not make sense.

A.7 Lie Derivatives of Tensor Fields

We now have all the pieces in place to define a new kind of derivative called the Lie derivative, named after the mathematician Sophus Lie. The Lie derivative does exactly what we think of derivatives doing, measuring how something changes in a particular direction. In a sense, the Lie derivative conforms a little more naturally to our calculus ideas of what derivatives do than exterior derivatives, even though exterior derivatives are such natural generalizations of directional derivatives. The idea that there can be different ways measuring how objects change, and hence different ways to define differentiation, seems very odd at first. But it should come as no surprise. After all, how the derivative of a function $f : \mathbb{R}^n \to \mathbb{R}^m$ is defined is somewhat different from how the derivative of a function $f : \mathbb{R} \to \mathbb{R}$ is defined. And the more complicated a mathematical object is then it stands to reason that the more ways there are of measuring how that object changes.

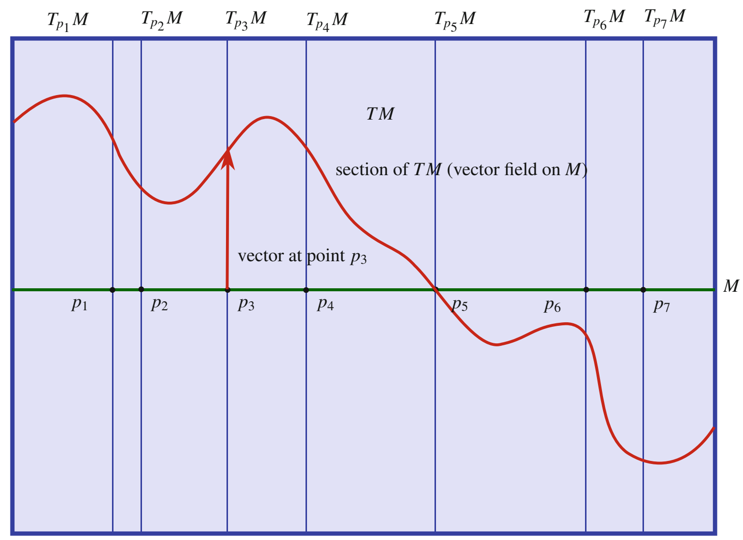
The most obvious difference is that Lie derivatives apply to any tensor field at all, whereas the exterior derivative only applies to differential forms. Of course, since differential forms are skew-symmetric covariant tensors then one can also take the Lie derivative of a differential form as well. Hence, after this section we will know two different concepts of differentiation for differential forms. As a book on differential forms we believe it is important that you at least be introduced to the idea of Lie derivatives since it is one of the ways differential forms can be differentiated.

Our introduction to Lie derivatives will be follow a standard path and be very similar to introduction that you will see in other books. Because tensors are mappings

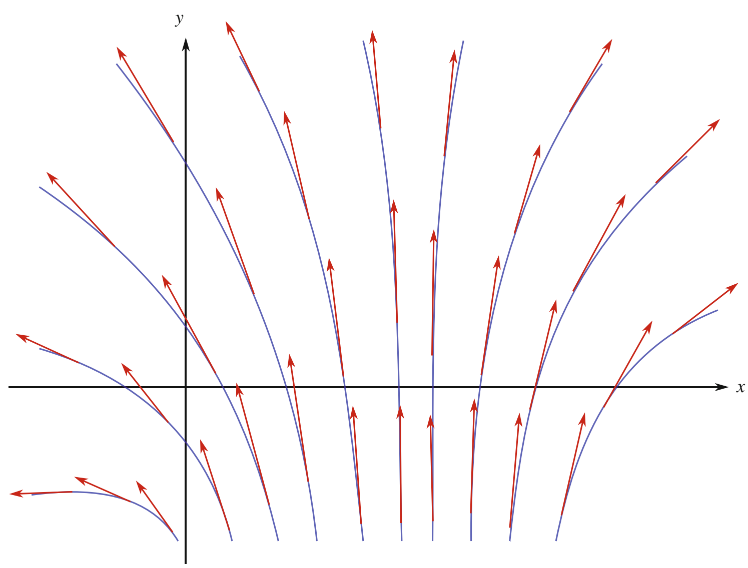
$$T : \underbrace{T^*M \times \cdots \times T^*M}_{\substack{r \\ \text{contravariant} \\ \text{degree}}} \times \underbrace{TM \times \cdots \times TM}_{\substack{s \\ \text{covariant} \\ \text{degree}}} \longrightarrow \mathbb{R}$$

and thus are made up of building blocks of the form $\frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_r}} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_s}$ we will first consider the Lie derivative of vectors, and in particular vectors of the form ∂_x. We will then consider lie derivatives of one-forms, and in particular elements of the form dx. It is then conceptually easy, though notationally cumbersome, to extend the concept to general tensors. We then consider the Lie derivative of a function and then will derive some identities that involve the Lie derivative. Finally we will compare the Lie derivative to the exterior derivative. For the remainder of this section we will always assume we are in a single coordinate chart on a manifold. It is not difficult, though again somewhat cumbersome, to patch together coordinate charts, though we will not do that here. Furthermore, due to the amount of material we will cover in this section we will be substantially more terse than we have been throughout this book so far, but hopefully it will serve as a first introduction.

We begin by considering a smooth integrable vector field v. What we mean by this is that the vector field has nice integral curves. But what are integral curves? Of course the vector field is simply a section of the tangent bundle TM, see Fig. A.1, but smooth integrable vector fields are a special kind of vector field. Suppose that the manifold coordinates of the chart we are in are given by $(x^1, \ldots, x^n)$. Thus we can write the vector field as $v = \sum v^i(x^1, \ldots, x^n)\frac{\partial}{\partial x^i} = v^i \partial_{x^i}$ where the v^i are real-valued functions on M. The smooth integrable vector field v is a vector field for which there exists functions $\gamma : \mathbb{R} \to M$ of a parameter t, usually thought of as time, such that the velocity vectors $\dot{\gamma}$ of the curves $\gamma(t) = \big(\gamma^1(t), \ldots, \gamma^n(t)\big)$ are



A smooth section v in the tangent bundle TM.



The vector field v shown along with its integral curves.

Fig. A.1 An integrable section v of the tangent bundle TM (top) is simply a vector field v on the manifold M that has integral curves γ on M (bottom)

given by the vector field v. This means that at each point $p \in M$ we have

$$\left. \frac{d\,\gamma(t)}{dt} \right|_p = v_p.$$

Since $\gamma(t) = \big(\gamma^1(t), \ldots, \gamma^n(t)\big)$ we can write this equation as the system of n differential equations,

$$\frac{d}{dt}\gamma^i = v^i.$$

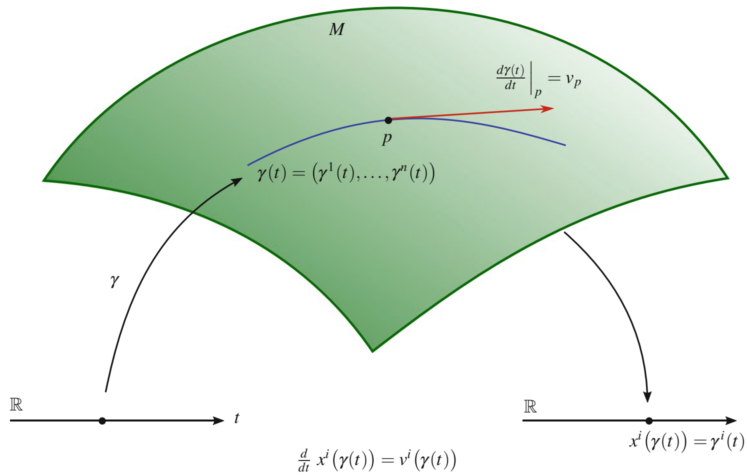


Fig. A.2 The manifold M shown with one of the integral curves γ which is obtained by solving the differential equation $\frac{d}{dt}x^i\Big(\gamma(t)\Big) = v^i\Big(\gamma(t)\Big)$. In essence, the integral curves γ are obtained by integrating the vector field v. Even though only one curve is shown here, the whole manifold M is filled with these curves

Thinking of γ as a function onto the manifold, $\gamma : \mathbb{R} \to M$, and x^i as coordinate functions on the manifold $x^i : M \to \mathbb{R}$ then we could also write this system of n differential equations as

$$\frac{d}{dt}x^i\Big(\gamma(t)\Big) = v^i\Big(\gamma(t)\Big),$$

where the $v^i : M \to \mathbb{R}$ are the real-valued component functions of the vector v. Solving this set of n differential equations gives the n integral functions γ^i which make up the integral curve γ. See Fig. A.2. Thus this particular idea of integration of a vector field on a manifold boils down to differential equations.

Integral curves γ are just that, curves on the manifold M, but we can also think of them as a family of mappings from the manifold M to itself, or to another copy of M. We begin by fixing some time t_0 as our zero time. Then for each time t we have a mapping γ_t that sends $\gamma(t_0)$ to $\gamma(t_0 + t)$,

$$M \xrightarrow{\gamma_t} M$$
$$p = \gamma(t_0) \longmapsto \gamma_t(p) = \gamma(t_0 + t).$$

Each different time t gives a different mapping γ_t, thereby giving us a whole family of mappings, as shown in Fig. A.3. As soon as we have a mapping $\gamma_t : M \to M$ between manifolds we have the push-forward mapping

$$T_p\gamma_t \equiv \gamma_{t*} : T_pM \to T_{\gamma_t(p)}M,$$

which allows us to push-forward vectors, and the pullback mapping

$$T_p^*\gamma_t \equiv \gamma_t^* : T_{\gamma_t(p)}^*M \to T_p^*M,$$

which allows us to pull-back one-forms. We generally drop the point p from the notation and infer it from context. Doing so we would simply write $T\gamma_t \equiv \gamma_{t*}$ and $T^*\gamma_t \equiv \gamma_t^*$. See Fig. A.4.

Of course the mapping γ_t is invertible. Suppose we have $\gamma_t(p) = q$. Then we can define $\gamma_t^{-1}(q) = p$. Let us look closely at the definitions, $p = \gamma(t_0)$ and $q = \gamma(t_0 + t)$ for some t_0. We can run time backwards just as well as forward, so if we defined $\tilde{t}_0 = t_0 + t$ then we would have $\gamma(\tilde{t}_0) = q$ and $\gamma(\tilde{t}_0 - t) = p$. But this is exactly the same as saying $\gamma_{-t}(q) = p$, which is exactly what we want γ_t^{-1} to do. Hence, we have $\gamma_t^{-1} = \gamma_{-t}$. Therefore we also can define the push-forward and pullback of $\gamma_t^{-1} = \gamma_{-t}$. We only need the push-forward,

$$T_q\gamma_t^{-1} \equiv T_q\gamma_{-t} \equiv \gamma_{-t*} : T_qM \to T_{\gamma_{-t}(q)}M,$$

where of course $q = \gamma_t(p)$. Dropping the point q from the notation we write $T\gamma_t^{-1} \equiv T\gamma_{-t}$ and are left to infer the base point from context.

Lie Derivatives of Vector Fields

Now we are ready for the definition of the Lie derivative. We begin by giving the definition of the Lie derivative of a vector field w in the direction of v, though perhaps it is more accurate to call it the Lie derivative of w along the flow of the vector field v, at the point p,

$$\boxed{\text{Lie Derivative of Vector Field:} \qquad (\mathscr{L}_v w)_p = \lim_{t\to 0} \frac{T\gamma_{-t} \cdot w_{\gamma_t(p)} - w_p}{t}.}$$

You will of course come across a fairly wide variety of notations for the Lie derivative of a vector field, such as

$$\begin{aligned}(\mathscr{L}_v w)_p &= \lim_{t\to 0} \frac{\left(\gamma_t^{-1}\right)_* w_{\gamma_t(p)} - w_p}{t} \\ &= \lim_{t\to 0} \frac{\left(\gamma_{-t}\right)_* w_{\gamma_t(p)} - w_p}{t}\end{aligned}$$

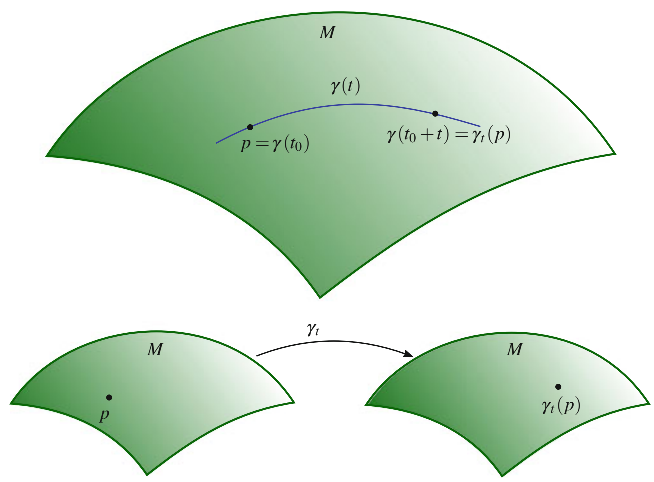


Fig. A.3 The integral curves γ obtained by integrating the vector field v can be used to define a whole family of mappings $\gamma_t : M \to M$, one for each value of t. The mapping γ_t sends each point $p = \gamma(t_0)$ of the manifold to $\gamma_t(p) = \gamma(t_0 + t)$

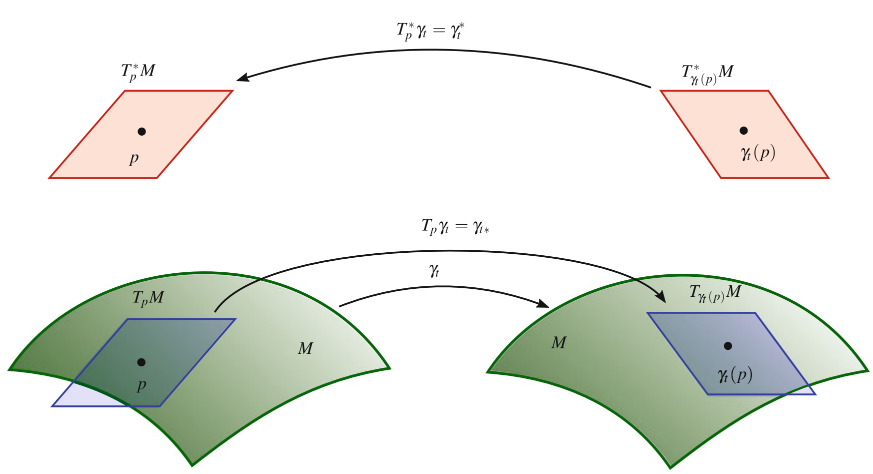


Fig. A.4 The mapping γ_t can be used to push-forward vectors with $T_p\gamma_t \equiv \gamma_{t*}$ and pullback one-forms with $T_p^*\gamma_t \equiv \gamma_t^*$

$$
\begin{aligned}
&= \lim_{t\to 0} \frac{\left(\gamma_t^{-1}\right)_* w_{\gamma_t(p)} - \left(\gamma_0^{-1}\right)_* w_{\gamma_0(p)}}{t} \\
&\equiv \frac{d}{dt}\left(\left(\gamma_t^{-1}\right)_* w_{\gamma_t(p)}\right)\Big|_0 \\
&= \frac{d}{dt}\left(\left(\gamma_{-t}\right)_* w_{\gamma_t(p)}\right)\Big|_0 .
\end{aligned}
$$

Question A.8 Show that these different notations all mean the same thing.

What is happening here? We want to understand how the vector field w is changing at the point p. In order to do this we need to first choose a direction v_p. Actually, we need just a little more than that, we need a little piece of an integral curve of a vector field v defined close to the point p. See Fig. A.5. With this we can measure how the vector field w is changing along the integral curve of v. We do this by pulling back the vector $w_{\gamma_t(p)}$ to the point p by $T\gamma_{-t}$. Thus both w_p and $T\gamma_{-t} \cdot w_{\gamma_t(p)}$ are both vectors at the point p, that is, both vectors in T_pM, so we can subtract one from the other, as shown graphically in Fig. A.6. This becomes the numerator of the difference quotient, with t being the denominator. Dividing a vector by a number t simply means dividing each component of the vector by t. Taking the limit as $t \to 0$ gives us the Lie derivative of w along v at the point p. A picture of this is shown in Fig. A.5. Thus we have that the Lie derivative is in fact another vector field, one that we have encountered already. More will be said about this in a few pages.

Lie Derivatives of One-Forms

The Lie derivative of a differential form α in the direction of v at the point p is defined in exactly the same way, only we have to use the pullback instead of push-forward,

$$
\boxed{\text{Lie Derivative of One-Form:} \qquad (\mathscr{L}_v\alpha)_p = \lim_{t\to 0} \frac{T^*\gamma_t \cdot \alpha_{\gamma_t(p)} - \alpha_p}{t},}
$$

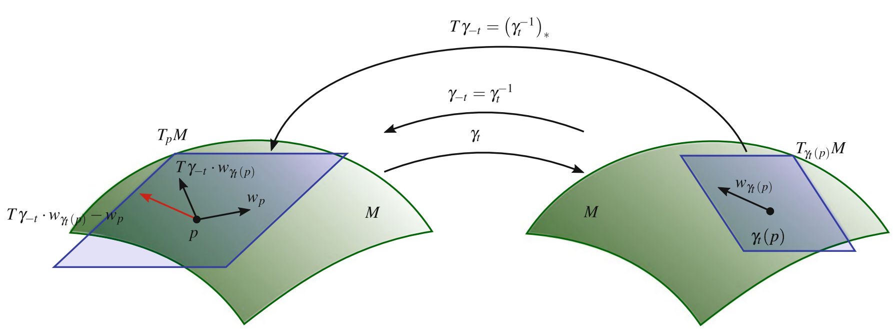


Fig. A.5 To find the Lie derivative of a vector field w in the direction v we first find the integral curves γ of v. With γ we have a whole family of maps γ_t that simply move points p of M along the flow γ by time t as shown in Fig. A.4. We can then push-forward a vector $w_{\gamma_t(p)}$ by $T\gamma_{-t}$ to the point p and subtract w_p from it. This becomes the numerator of the difference quotient for the Lie derivative

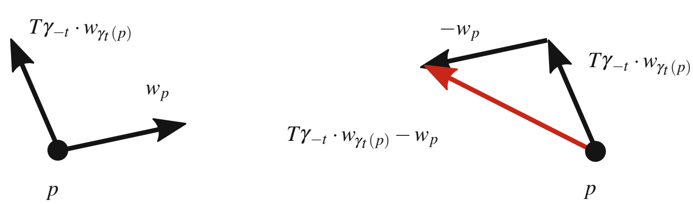


Fig. A.6 Here we graphically show how the vector w_p is subtracted from the vector $T\gamma_{-t} \cdot w_{\gamma_t(p)}$ to give the vector $T\gamma_{-t} \cdot w_{\gamma_t(p)} - w_p$

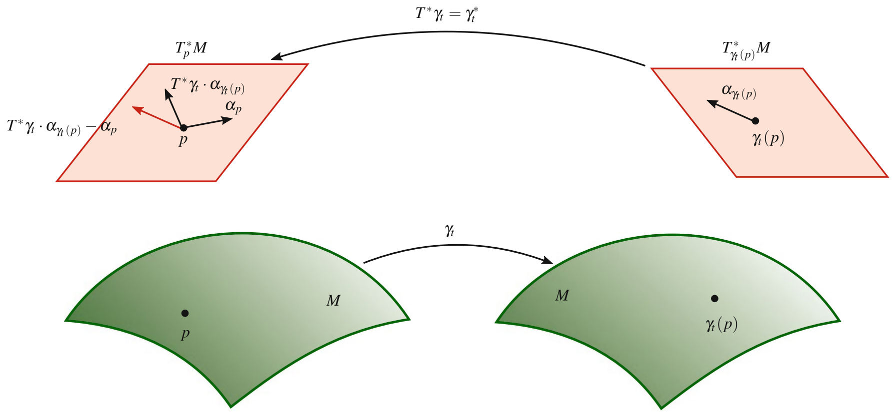


Fig. A.7 To find the Lie derivative of a one-form $w\alpha$ in the direction v we first find the integral curves γ of v. With γ we have a whole family of maps γ_t that simply move points p of M along the flow γ by time t as shown in Fig. A.4. We can then pullback a one-form $\alpha_{\gamma_t(p)}$ by $T^*\gamma_t$ to the point p and subtract α_p from it. This becomes the numerator of the difference quotient for the Lie derivative

which of course can also be written as

$$\begin{aligned}(\mathscr{L}_v\alpha)_p &= \lim_{t\to 0} \frac{\gamma_t^*\alpha_{\gamma_t(p)} - \alpha_p}{t} \\ &= \lim_{t\to 0} \frac{\gamma_t^*\alpha_{\gamma_t(p)} - \gamma_0^*\alpha_{\gamma_0(p)}}{t} \\ &\equiv \frac{d}{dt}\Big((\gamma_t)^*\alpha_{\gamma_t(p)}\Big)\Big|_0 .\end{aligned}$$

Notice that the Lie derivative of the one-form α is defined in an completely analogous way to the Lie derivative of the vector field w, see Fig. A.7.

Question A.9 Show that the above notations are all equivalent.

Lie Derivatives of Functions

What would the Lie derivative of a function be? Since we have previously viewed functions as zero-forms we could also view them as rank-zero covariant tensors and use the formula for the Lie derivative of a one-form to get

$$(\mathscr{L}_v f)_p = \lim_{t\to 0} \frac{T^*\gamma_t \cdot f_{\gamma_t(p)} - f_p}{t}.$$

But $f_{\gamma_t(p)}$ is simply $f\big(\gamma_t(p)\big) = f(q)$, or the numerical value of f evaluated at $\gamma_t(p) = q$. The pull-back of a number by γ_t^* is simply that number, so we have $\gamma_t^* f_{\gamma_t(p)} = f\big(\gamma_t(p)\big) = f(q)$. Similarly, f_p is simply the function f evaluated at point p, or $f(p)$. Since $p = \gamma(t_0)$ and $q = \gamma(t_0 + t)$, putting everything together we can write

$$\begin{aligned}(\mathscr{L}_v f)_p &= \lim_{t\to 0} \frac{f\big(\gamma(t_0+t)\big) - f\big(\gamma(t)\big)}{t} \\ &= \frac{\partial f\big(\gamma(t)\big)}{\partial t}\bigg|_{t_0} \\ &= \frac{\partial f}{\partial \gamma^i}\bigg|_{\gamma(t_0)} \frac{\partial \gamma^i}{\partial t}\bigg|_{t_0} .\end{aligned}$$

The last equality is simply the chain rule. The integral curve γ can be written as $\gamma(t) = \big(\gamma^1(t), \ldots, \gamma^n(t)\big)$, but what is γ^i? It is simply the x^i coordinate of the curve, so we could relabel and write the curve in more standard notation as $\gamma(t) = \big(x^1(t), \ldots, x^n(t)\big)$. Rewriting this curve in more standard notation simply allows us to recognize the end result easier, which can be written as

$$(\mathscr{L}_v f)_p = \left.\frac{\partial f}{\partial x^i}\right|_{\gamma(t_0)} \left.\frac{\partial x^i}{\partial t}\right|_{t_0}.$$

Next recall how we got the x^i coordinate γ^i in the first place, we got it from the vector v by solving the set of differential equations

$$\frac{d}{dt}\gamma^i = v^i,$$

thereby giving us

$$\begin{aligned}(\mathscr{L}_v f)_p &= \frac{\partial f}{\partial x^i}\, v^i \\ &= v^i\, \frac{\partial f}{\partial x^i} \\ &= \left(v^i \frac{\partial}{\partial x^i}\right) f \\ &= v_p[f],\end{aligned}$$

which is simply our directional derivative of f in the direction v. So for functions there is no difference between Lie derivatives and our familiar directional derivative. Thus we have obtained

$$\boxed{\text{Lie Derivative of Function:} \qquad (\mathscr{L}_v f)_p = v_p[f].}$$

Lie Derivatives of (r, s)-Tensors
Now we turn to the Lie derivative of a (r, s)-tensor $\mathscr{T}$. We will define the pull-back of the tensor $\mathscr{T}$ by γ_t with the following

$$\left(\gamma_t^* \mathscr{T}\right)_p\left(\alpha_1, \ldots, \alpha_r, v_1, \ldots, v_s\right) = \mathscr{T}_{\gamma_t(p)}\left(\gamma_{-t}^*\alpha_1, \ldots, \gamma_{-t}^*\alpha_r, \gamma_{t*}v_1, \ldots, \gamma_{t*}v_s\right).$$

Compare this to our original definition of the pullback of a one-form. Defining the pull-pack of a tensor this way we define the Lie derivative of the tensor $\mathscr{T}$ with the same formula that we used to define the Lie derivative of a one-form,

$$\boxed{\text{Lie Derivative of Tensor:} \qquad (\mathscr{L}_v \mathscr{T})_p = \lim_{t\to 0} \frac{\gamma_t^* \mathscr{T}_{\gamma_t(p)} - \mathscr{T}_p}{t}.}$$

Of course we can also write the Lie derivative of a general tensor field as

$$\begin{aligned}(\mathscr{L}_v \mathscr{T})_p &= \lim_{t\to 0} \frac{\gamma_t^* \mathscr{T}_{\gamma_t(p)} - \gamma_0^* \mathscr{T}_{\gamma_0(p)}}{t} \\ &= \frac{d}{dt}\left.\left(\gamma_t^* \mathscr{T}_{\gamma_t(p)}\right)\right|_0.\end{aligned}$$

Some Lie Derivative Identities
Now we turn to showing a number of important identities that involve the Lie derivative. We move thorough these identities rather quickly, leaving some as exercises. The first identity is that regardless of the kind of tensors $\mathscr{S}$ and $\mathscr{T}$ are, the Lie derivative is linear. That is, for $a, b \in \mathbb{R}$, we have

$$\boxed{\mathscr{L}_v\left(a\mathscr{S} + b\mathscr{T}\right) = a\mathscr{L}_v\mathscr{S} + b\mathscr{L}_v\mathscr{T}.}$$

Question A.10 Using the definition of the Lie derivative of a tensor, show that $\mathscr{L}_v\left(a\mathscr{S} + b\mathscr{T}\right) = a\mathscr{L}_v\mathscr{S} + b\mathscr{L}_v\mathscr{T}$.

Question A.11 Using the definition of the Lie derivative of a tensor, show that $\phi^*\big(\mathscr{T} \otimes \mathscr{S}\big) = \phi^*\mathscr{T} \otimes \phi^*\mathscr{S}$

Turning to the Lie derivative of the tensor product of two tensors, and dropping the base points from our notation and relying on the identity $\phi^*\big(\mathscr{T} \otimes \mathscr{S}\big) = \phi^*\mathscr{T} \otimes \phi^*\mathscr{S}$, we find that

$$
\begin{aligned}
\mathscr{L}_v(\mathscr{S} \otimes \mathscr{T}) &= \lim_{t\to 0} \frac{\gamma_t^*(\mathscr{S} \otimes \mathscr{T}) - \mathscr{S} \otimes \mathscr{T}}{t} \\
&= \lim_{t\to 0} \frac{\gamma_t^*\mathscr{S} \otimes \gamma_t^*\mathscr{T} - \mathscr{S} \otimes \mathscr{T}}{t} \\
&= \lim_{t\to 0} \frac{\big(\mathscr{S} + (\gamma_t^*\mathscr{S} - \mathscr{S})\big) \otimes \big(\mathscr{T} + (\gamma_t^*\mathscr{T} - \mathscr{T})\big) - \mathscr{S} \otimes \mathscr{T}}{t} \\
&= \lim_{t\to 0} \frac{(\gamma_t^*\mathscr{S} - \mathscr{S}) \otimes \mathscr{T} + \mathscr{S} \otimes (\gamma_t^*\mathscr{T} - \mathscr{T}) + (\gamma_t^*\mathscr{S} - \mathscr{S}) \otimes (\gamma_t^*\mathscr{T} - \mathscr{T})}{t} \\
&= \lim_{t\to 0} \left(\frac{\gamma_t^*\mathscr{S} - \mathscr{S}}{t} \otimes \mathscr{T}\right) + \lim_{t\to 0} \left(\mathscr{S} \otimes \frac{\gamma_t^*\mathscr{T} - \mathscr{T}}{t}\right) + \lim_{t\to 0} \left(\frac{\gamma_t^*\mathscr{S} - \mathscr{S}}{t} \otimes (\gamma_t^*\mathscr{T} - \mathscr{T})\right) \\
&= \mathscr{L}_v\mathscr{S} \otimes \mathscr{T} + \mathscr{S} \otimes \mathscr{L}_v\mathscr{T} + \mathscr{L}_v\mathscr{S} \otimes 0 \\
&= \mathscr{L}_v\mathscr{S} \otimes \mathscr{T} + \mathscr{S} \otimes \mathscr{L}_v\mathscr{T}.
\end{aligned}
$$

It is clear that as $t \to 0$ we have $(\gamma_t^*\mathscr{T} - \mathscr{T}) \to (\mathscr{T} - \mathscr{T}) = 0$ in the next to last line.

Question A.12 In the fifth line of the above string of equalities we used that

$$
\lim_{t\to 0} \frac{(\gamma_t^*\mathscr{S} - \mathscr{S}) \otimes (\gamma_t^*\mathscr{T} - \mathscr{T})}{t} = \lim_{t\to 0} \left(\frac{\gamma_t^*\mathscr{S} - \mathscr{S}}{t} \otimes (\gamma_t^*\mathscr{T} - \mathscr{T})\right).
$$

Show that this is true. Also show that

$$
\lim_{t\to 0} \frac{(\gamma_t^*\mathscr{S} - \mathscr{S}) \otimes (\gamma_t^*\mathscr{T} - \mathscr{T})}{t} = \lim_{t\to 0} \left((\gamma_t^*\mathscr{S} - \mathscr{S}) \otimes \frac{\gamma_t^*\mathscr{T} - \mathscr{T}}{t}\right).
$$

Explain that the final result in the proof of the above identity does not depend on which of the equalities in this question is used.

Putting everything together we have shown the identity

$$
\boxed{\mathscr{L}_v(\mathscr{S} \otimes \mathscr{T}) = \mathscr{L}_v\mathscr{S} \otimes \mathscr{T} + \mathscr{S} \otimes \mathscr{L}_v\mathscr{T}.}
$$

Similarly, relying on the identity $\phi^*(\alpha \wedge \beta) = \phi^*\alpha \wedge \phi^*\beta$ from Sect. 6.7 we have

$$
\begin{aligned}
\mathscr{L}_v(\alpha \wedge \beta) &= \lim_{t\to 0} \frac{\gamma_t^*(\alpha \wedge \beta) - \alpha \wedge \beta}{t} \\
&= \lim_{t\to 0} \frac{\gamma_t^*\alpha \wedge \gamma_t^*\beta - \alpha \wedge \beta}{t} \\
&= \lim_{t\to 0} \frac{\big(\alpha + (\gamma_t^*\alpha - \alpha)\big) \wedge \big(\beta + (\gamma_t^*\beta - \beta)\big) - \alpha \wedge \beta}{t} \\
&= \lim_{t\to 0} \frac{(\gamma_t^*\alpha - \alpha) \wedge \beta + \alpha \wedge (\gamma_t^*\beta - \beta) + (\gamma_t^*\alpha - \alpha) \wedge (\gamma_t^*\beta - \beta)}{t} \\
&= \lim_{t\to 0} \left(\frac{\gamma_t^*\alpha - \alpha}{t} \wedge \beta\right) + \lim_{t\to 0} \left(\alpha \wedge \frac{\gamma_t^*\beta - \beta}{t}\right) + \lim_{t\to 0} \left(\frac{\gamma_t^*\alpha - \alpha}{t} \wedge (\gamma_t^*\beta - \beta)\right) \\
&= \mathscr{L}_v\alpha \wedge \beta + \alpha \wedge \mathscr{L}_v\beta + \mathscr{L}_v\alpha \wedge 0 \\
&= \mathscr{L}_v\alpha \wedge \beta + \alpha \wedge \mathscr{L}_v\beta.
\end{aligned}
$$

Again it is clear that as $t \to 0$ we have $(\gamma_t^*\beta - \beta) \to (\beta - \beta) = 0$ in the next to last line.

Question A.13 The above proof used that

$$\lim_{t\to 0} \frac{(\gamma_t^*\alpha - \alpha)\wedge(\gamma_t^*\beta - \beta)}{t} = \lim_{t\to 0}\left(\frac{\gamma_t^*\alpha - \alpha}{t}\wedge(\gamma_t^*\beta - \beta)\right).$$

Show this is true. Also, show that

$$\begin{aligned}\lim_{t\to 0} \frac{(\gamma_t^*\alpha - \alpha)\wedge(\gamma_t^*\beta - \beta)}{t} &= \lim_{t\to 0}\left(\frac{\gamma_t^*\alpha - \alpha}{t}\wedge(\gamma_t^*\beta - \beta)\right)\\ &= \lim_{t\to 0}\left((\gamma_t^*\alpha - \alpha)\wedge\frac{\gamma_t^*\beta - \beta}{t}\right).\end{aligned}$$

Explain that the final result in the proof of the above identity does not depend on which of the equalities in this question are used.

Putting everything together we have shown the identity

$$\boxed{\mathscr{L}_v(\alpha\wedge\beta) = \mathscr{L}_v\alpha\wedge\beta + \alpha\wedge\mathscr{L}_v\beta.}$$

The next identity involves functions f and vector fields v and w. It is

$$\boxed{\mathscr{L}_v(fw) = f\mathscr{L}_v w + v[f]w.}$$

Question A.14 Using the definitions and identities already given and proved, show that $\mathscr{L}_v(fw) = f\mathscr{L}_v w + v[f]w$.

Another important identity is

$$\boxed{d\big(\mathscr{L}_v\alpha\big) = \mathscr{L}_v\big(d\alpha\big).}$$

This follows from the commutativity of pull-backs and exterior differentiation, $\phi^*d = d\phi^*$.

Question A.15 Using the definition of the Lie derivative for a covariant tensor along with the fact that $\phi^*d = d\phi^*$, show that $d\big(\mathscr{L}_v\alpha\big) = \mathscr{L}_v\big(d\alpha\big)$.

Cartan's Homotopy Formula

We now prove a very important identity called either **Cartan's homotopy formula** or sometimes **Cartan's magic formula**. Given a k-form α and a vector field v, and recalling the definition of the interior product ι from Sect. 3.4, Cartan's magic formula is

$$\mathscr{L}_v\alpha = \iota_v(d\alpha) + d(\iota_v\alpha).$$

While there are several ways to show this identity we will use induction. First we consider our base case for a zero-form $\alpha = f$. The interior product $\iota_v f$ is not defined for a zero-form hence the identity is simply

$$\mathscr{L}_v f = \iota_v(df).$$

But this is obviously true, as long as you remember all the different ways of writing the same thing. Recall, that we found the Lie derivative of a function is simply the directional derivative of the function, that is, $\mathscr{L}_v f = v[f]$. Also recall from early in the book that $v[f] \equiv df(v)$ and of course $df(v) = \iota_v df$ by definition of the interior product. Hence we get the following string of equalities that are little more than different ways of writing the same thing,

$$\mathscr{L}_v f = v[f] = df(v) = \iota_v df.$$

This classic example of "proof by notation" establishes our base case. Also notice that if the function f was a coordinate function x^i then we would have $\mathscr{L}_v x^i = \iota_v dx^i$. We will need this fact in a moment.

Now assume that the identity $\mathscr{L}_v\alpha = \iota_v(d\alpha) + d(\iota_v\alpha)$ holds for all $(k-1)$-forms α. Suppose we have a k-form $\omega = f dx^1\wedge\cdots\wedge dx^k$, which is clearly a base term for a general k-form. We can write $\omega = dx^1\wedge\alpha$ where $\alpha = f dx^2\wedge\cdots\wedge dx^k$, a $(k-1)$-form. Recalling one of the identities we just proved, the left hand side of Cartan's magic formula becomes

$$\begin{aligned}\mathscr{L}_v\omega &= \mathscr{L}_v(dx^1\wedge\alpha)\\ &= \mathscr{L}_v dx^1\wedge\alpha + dx^1\wedge\mathscr{L}_v\alpha.\end{aligned}$$

Before turning to the right hand side of Cartan's magic formula we recall a couple of identities we will need. The first is an identity of the interior product proved in Sect. 3.4. If α is a k-form and β is another differential form, then

$$\begin{aligned}\iota_v(\alpha\wedge\beta) &= (\iota_v\alpha)\wedge\beta + (-1)^k\alpha\wedge(\iota_v\beta)\\ &= \iota_v\alpha\wedge\beta + (-1)^k\alpha\wedge\iota_v\beta.\end{aligned}$$

The second is an identity of the exterior derivative. This identity was either derived as a consequence of how we defined exterior derivative, as in Sect. 4.2, or it was viewed as an axiom for exterior differentiation that allowed us to find the formula for the exterior derivative, as in Sect. 4.3. If α is a k-form and β is another differential form, then

$$\begin{aligned}d(\alpha\wedge\beta) &= (d\alpha)\wedge\beta + (-1)^k\alpha\wedge(d\beta)\\ &= d\alpha\wedge\beta + (-1)^k\alpha\wedge d\beta.\end{aligned}$$

Now we consider the right hand side of Cartan's magic formula,

$$\begin{aligned}\iota_v(d\omega) + d(\iota_v\omega) &= \iota_v d\big(dx^1\wedge\alpha\big) + d\iota_v\big(dx^1\wedge\alpha\big)\\ &= \iota_v\Big(\underbrace{ddx^1}_{=0}\wedge\alpha + (-1)^1 dx^1\wedge d\alpha\Big) + d\Big(\iota_v dx^1\wedge\alpha + (-1)^1 dx^1\wedge i_v\alpha\Big)\\ &= -\iota_v\big(dx^1\wedge d\alpha\big) + d\big(\iota_v dx^1\wedge\alpha\big) - d\big(dx^1\wedge\iota_v\alpha\big)\\ &= -\cancel{\iota_v dx^1\wedge d\alpha} + dx^1\wedge\iota_v d\alpha\\ &\qquad + d\underbrace{\iota_v dx^1}_{\mathscr{L}_v x^1}\wedge\alpha + (-1)^0\cancel{\iota_v dx^1\wedge d\alpha}\\ &\qquad\qquad - \underbrace{ddx^1}_{=0}\wedge\iota_v\alpha - (-1)^1 dx^1\wedge d\iota_v\alpha\\ &= d\big(\mathscr{L}_v x^1\big)\wedge\alpha + dx^1\wedge\iota_v d\alpha + dx^1\wedge d\iota_v\alpha\\ &= \mathscr{L}_v dx^1\wedge\alpha + dx^1\wedge\Big(\underbrace{\iota_v d\alpha + d\iota_v\alpha}_{\substack{=\mathscr{L}_v\alpha\\ \text{induction hypothesis}}}\Big)\\ &= \mathscr{L}_v dx^1\wedge\alpha + dx^1\wedge\mathscr{L}_v\alpha.\end{aligned}$$

So both the left and the right hand side of Cartan's magic formula for $\omega = dx^1\wedge\alpha$ are equal to the same thing, namely $\mathscr{L}_v dx^1\wedge\alpha + dx^1\wedge\mathscr{L}_v\alpha$, and are therefore equal to each other.

Question A.16 Use linearity of exterior differentiation, the Lie derivative, and the interior product, to show that Cartan's magic formula applies to every differential form of the form $\omega = \sum f_{i_1\cdots i_k} dx^{i_1}\wedge\cdots\wedge dx^{i_k}$.

Thus we have shown for any differential form α that

$$\boxed{\text{Cartan's Homotopy Formula:}\qquad \mathscr{L}_v\alpha = \iota_v(d\alpha) + d(\iota_v\alpha).}$$

This identity is very often simply written as $\mathscr{L}_v = \iota_v d + d\iota_v$.

Lie Derivative Equals Commutator

We now turn our attention to showing that the Lie derivative of the vector field w in the direction v is equal to the commutator of v and w. In other words, we will show the identity

$$\boxed{\mathscr{L}_v w = [v, w] = vw - wv.}$$

We first encountered the commutator of v and w, also called the Lie-brackets of v and w, in Sect. 4.4.2 where we attempted to explain what $[v, w]$ was from a somewhat geometrical perspective. The proof of this identity is almost too slick, amounting to little more than some clever playing with definitions and notation.

First let us consider the situation where we have a map $\phi : M \to N$ and a vector w_p on M and a real-valued function $f : N \to \mathbb{R}$. We can use ϕ to push-forward the vector w_p on M to the vector $T_p\phi \cdot w_p$ on N. We can then find the directional derivative of f in the direction of $T_p\phi \cdot w_p$ at the point $\phi(p)$. This is a numerical value. By definition, the push-forward of a vector by a map ϕ acting on a function f gives the same numerical value as the vector acting on $f \circ \phi$ at the point p, that is,

$$w_p[f \circ \phi] = \big(T_p\phi \cdot w_p\big)_{\phi(p)}[f].$$

We reiterate, as numerical values these are equal. But as we know, the derivative of a function is another function, the numerical value is only obtained when we evaluate this derived function at a point. If w were actually a vector field on M then $w[f \circ \phi]$ is a function on M. Similarly, we have $\big(T\phi \cdot w\big)[f]$ is a function on N. But clearly, since these two functions are not even on the same manifold

$$w[f \circ \phi] \neq \big(T\phi \cdot w\big)[f].$$

How do we fix this? Since $\big(T\phi \cdot w\big)[f]$ is a function on N, which is nothing more than a zero-form, we can pull it back to M. Recalling how the pull-backs of zero-forms are defined we have

$$T^*\phi \cdot \Big(\big(T\phi \cdot w\big)[f]\Big) = \Big(\big(T\phi \cdot w\big)[f]\Big) \circ \phi,$$

which is now a function on M that gives the same numerical values as $w[f \circ \phi]$ does, which we can also rewrite as $w[T^*\phi \cdot f]$. See Fig. A.8. Thus we end up with the rather slick equality of functions on M,

$$w[T^*\phi \cdot f] = T^*\phi \cdot \Big(\big(T\phi \cdot w\big)[f]\Big).$$

This procedure illustrates the general idea that we need to use.

Now we consider the specific situation that now concerns us. Suppose we have an integrable vector field v on M with integral curves γ parameterized by time t. Then for each time t we get a mapping $\gamma(t) : M \to \widetilde{M}$, where we simply have the range manifold $\widetilde{M} = M$. We can run time backwards to get the inverse map $(\gamma_t)^{-1} = \gamma_{-t} : \widetilde{M} \to M$. Now suppose we have a vector field w on $M = \widetilde{M}$ and a real-valued mapping $f : M \to \mathbb{R}$, see Fig. A.9. We have the following numerical values being equal,

$$\big(T_{\gamma_t(p)}\gamma_{-1} \cdot w_{\gamma_t(p)}\big)_p[f] = w_{\gamma_t(p)}[f \circ \gamma_{-t}].$$

But if we take away the base point we no longer have numerical values but functions on each side of this equality; on left hand side $\big(T\gamma_{-1} \cdot w\big)[f]$ is a function on M and the right hand side $w[f \circ \gamma_{-t}]$ as a function on $\widetilde{M}$. To get an equality between functions we can pull-back the right hand side to M we get an equality of function on M,

$$\big(T\gamma_{-1} \cdot w\big)[f] = \big(w[f \circ \gamma_{-t}]\big) \circ \gamma_t,$$

which can also be rewritten as

$$\big(T\gamma_{-1} \cdot w\big)[f] = T^*\gamma_t \cdot \big(w[T^*\gamma_{-t} \cdot f]\big).$$

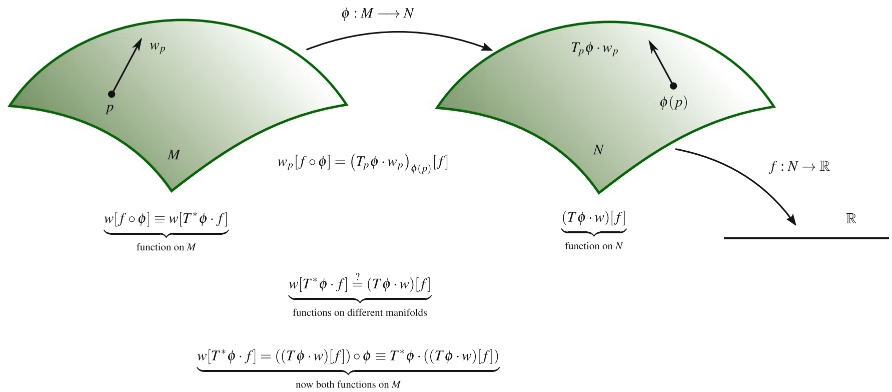


Fig. A.8 From one of the questions in Sect. 10.3 we know that $w_p[f \circ \phi] = \big(T_p\phi \cdot w_p\big)_{\phi(p)}[f]$. But this is an equality of values. By eliminating the base point the right hand side becomes $w[f \circ \phi]$, the directional derivative of $f \circ \phi$ in the w direction, which is itself a function on M. Similarly, the left hand side becomes $(T\phi \cdot w)[f]$, the directional derivative of f in the $T\phi \cdot w$ direction, which is itself a function on N. Thus $v[f \circ \phi]$ and $T\phi \cdot w[f]$ can not be equal as functions. But this can be fixed by pulling-back the zero-form (function) $(T\phi \cdot w)[f]$ by $T^*\phi$. The pull-back of a zero-form was also defined in Sect. 10.3

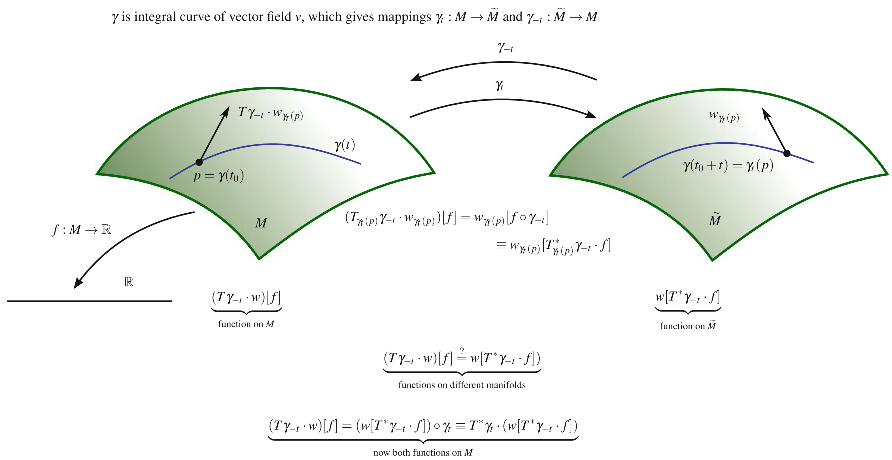


Fig. A.9 Here a procedure similar to that done in Fig. A.8 allows us to equate two functions on M, obtaining $T\gamma_{-t} \cdot w[f] = T^*\gamma_t \cdot (w[T^*\gamma_{-t} \cdot f])$. This is the particular case we need in the proof of $\mathscr{L}_v w = [v, w]$

This is the exact identity we need.

Now we are finally ready for the proof of $\mathscr{L}_v w = [v, w] = vw - wv$. We begin by writing the definition of the Lie derivative of a vector field w in the direction v,

$$(\mathscr{L}_v w)_p = \lim_{t \to 0} \frac{T\gamma_{-t} \cdot w_{\gamma_t(p)} - w_p}{t}.$$

Since we are interested in a general formula independent of base point p we will use this definition without the base point. If you wanted to put the base points back in you would need to think carefully about where the vectors w are located at. Also, since $\mathscr{L}_v w$ is itself a vector field we can use it to take the directional derivative of a function f, so we have

$$\begin{aligned}
\big(\mathscr{L}_v w\big)[f] &= \left(\lim_{t \to 0} \frac{T\gamma_{-t} \cdot w - w}{t}\right)[f] \\
&= \lim_{t \to 0} \frac{(T\gamma_{-t} \cdot w)[f] - w[f]}{t} \\
&= \lim_{t \to 0} \frac{T^*\gamma_t \cdot \big(w[T^*\gamma_{-t} \cdot f]\big) - w[f]}{t}
\end{aligned}$$

$$
\begin{aligned}
&= \lim_{t\to 0} \frac{T^*\gamma_t \cdot \big(w[T^*\gamma_{-t}\cdot f]\big) - T^*\gamma_t \cdot \big(w[f]\big) + T^*\gamma_t \cdot \big(w[f]\big) - w[f]}{t} \\
&= \lim_{t\to 0} \frac{T^*\gamma_t \cdot \big(w[T^*\gamma_{-t}\cdot f]\big) - T^*\gamma_t \cdot \big(w[f]\big)}{t} + \lim_{t\to 0} \frac{T^*\gamma_t \cdot \big(w[f]\big) - w[f]}{t} \\
&= \lim_{t\to 0} T^*\gamma_t \cdot \left(w \left[\frac{T^*\gamma_{-t}\cdot f - f}{t} \right] \right) + \lim_{t\to 0} \frac{T^*\gamma_t \cdot \big(w[f]\big) - w[f]}{t} \\
&= \lim_{t\to 0} T^*\gamma_t \cdot \left(w \left[\frac{f\circ\gamma_{-t} - f}{t} \right] \right) + \lim_{t\to 0} \frac{\big(w[f]\big)\circ\gamma_t - w[f]}{t} \\
&= T^*\gamma_0 \cdot (w\,[-v[f]]) + v\big[w[f]\big] \\
&= -w\big[v[f]\big] + v\big[w[f]\big] \\
&= [v, w]f.
\end{aligned}
$$

Writing without the function f gives us $\mathscr{L}_v w = [v, w] = vw - wv$.

Question A.17 Show that

$$
\frac{T^*\gamma_t \cdot \big(w[T^*\gamma_{-t}\cdot f]\big) - T^*\gamma_t \cdot \big(w[f]\big)}{t} = T^*\gamma_t \cdot \left(w \left[\frac{T^*\gamma_{-t}\cdot f - f}{t} \right] \right).
$$

Question A.18 Recalling that γ is the integral curve of the vector field v show that

$$
\lim_{t\to 0} \frac{\big(w[f]\big)\circ\gamma_t - w[f]}{t} = v\big[w[f]\big].
$$

Question A.19 Recalling that γ is the integral curve of the vector field v show that

$$
\lim_{t\to 0} \frac{f\circ\gamma_{-t} - f}{t} = -v[f].
$$

Here we will simply take a moment to write a formula for the commutator. Suppose $v = \sum v^i \frac{\partial}{\partial x^i} = v^i \frac{\partial}{\partial x^i}$ and $w = \sum w^j \frac{\partial}{\partial x^j} = w^j \frac{\partial}{\partial x^j}$. The Lie derivative $\mathscr{L}_v w$ is itself a vector field and as such it can act on a function f; that is, we can find $\mathscr{L}_v w[f]$, the directional derivative of f in the direction $\mathscr{L}_v w$. What the identity means is that

$$
\begin{aligned}
\mathscr{L}_v w[f] &= v[w[f]] - w[v[f]] \\
&= v\left[w^j \frac{\partial}{\partial x^j}[f] \right] - w\left[v^i \frac{\partial}{\partial x^i}[f] \right] \\
&= v\left[w^j \frac{\partial f}{\partial x^j} \right] - w\left[v^i \frac{\partial f}{\partial x^i} \right] \\
&= v^i \frac{\partial}{\partial x^i}\left[w^j \frac{\partial f}{\partial x^j} \right] - w^j \frac{\partial}{\partial x^j}\left[v^i \frac{\partial f}{\partial x^i} \right] \\
&= v^i \frac{\partial w^j}{\partial x^i}\frac{\partial f}{\partial x^j} + v^i w^j \frac{\partial^2 f}{\partial x^i \partial x^j} - w^j \frac{\partial v^i}{\partial x^j}\frac{\partial f}{\partial x^i} - w^j v^i \frac{\partial^2 f}{\partial x^j \partial x^i} \\
&= v^i \frac{\partial w^j}{\partial x^i}\frac{\partial f}{\partial x^j} - w^j \frac{\partial v^i}{\partial x^j}\frac{\partial f}{\partial x^i} \\
&= \left(v^i \frac{\partial w^j}{\partial x^i}\frac{\partial}{\partial x^j} - w^j \frac{\partial v^i}{\partial x^j}\frac{\partial}{\partial x^i} \right) f.
\end{aligned}
$$

Thus we could also write

$$\boxed{\mathscr{L}_v w = v^i \frac{\partial w^j}{\partial x^i}\frac{\partial}{\partial x^j} - w^j \frac{\partial v^i}{\partial x^j}\frac{\partial}{\partial x^i}.}$$

Global Formula for Lie Derivative of Differential Forms

Now we turn our attention to the global formula for Lie derivatives of differential forms. This formula will be used in the proof of the global formula for exterior differentiation. The global formula for Lie derivatives of differential forms is given by

$$\big(\mathscr{L}_v\alpha\big)(w_1,\ldots,w_k) = v\big[\alpha(w_1,\ldots,w_k)\big] - \sum_{i=1}^{k} \alpha(w_1,\ldots,[v,w_i],\ldots,w_k).$$

We will actually carry out the proof for the two-form case and leave the general case as an exercise. Other than being notationally more complex, the general case is similar to the two-form case. Letting γ be the integral curves of v, α a two-form, and w_1, w_2 vector fields, we note that $\alpha(w_1, w_2)$ is a function. Thus we have

$$\begin{aligned}
\big(\mathscr{L}_v\alpha(w_1,w_2)\big)_p &= \lim_{t\to 0} \frac{T^*\gamma_t\big(\alpha(w_1,w_2)\big)_{\gamma_t(p)} - \big(\alpha(w_1,w_2)\big)_p}{t} \\
&= \lim_{t\to 0} \frac{\alpha_{\gamma_t(p)}\left(w_{1_{\gamma_t(p)}}, w_{2_{\gamma_t(p)}}\right) - \alpha_p\left(w_{1_p}, w_{2_p}\right)}{t} \\
&= \lim_{t\to 0} \frac{\alpha_{\gamma_t(p)}\left(w_{1_{\gamma_t(p)}}, w_{2_{\gamma_t(p)}}\right) - \alpha_p\left(T\gamma_{-t}\cdot w_{1_{\gamma_t(p)}}, T\gamma_{-t}\cdot w_{2_{\gamma_t(p)}}\right)}{t} \\
&\quad + \lim_{t\to 0} \frac{\alpha_p\left(T\gamma_{-t}\cdot w_{1_{\gamma_t(p)}}, T\gamma_{-t}\cdot w_{2_{\gamma_t(p)}}\right) - \alpha_p\left(w_{1_p}, T\gamma_{-t}\cdot w_{2_{\gamma_t(p)}}\right)}{t} \\
&\quad + \lim_{t\to 0} \frac{\alpha_p\left(w_{1_p}, T\gamma_{-t}\cdot w_{2_{\gamma_t(p)}}\right) - \alpha_p\left(w_{1_p}, w_{2_p}\right)}{t}.
\end{aligned}$$

The first term following the last equality gives us

$$\begin{aligned}
&\lim_{t\to 0} \frac{\alpha_{\gamma_t(p)}\left(w_{1_{\gamma_t(p)}}, w_{2_{\gamma_t(p)}}\right) - \alpha_p\left(T\gamma_{-t}\cdot w_{1_{\gamma_t(p)}}, T\gamma_{-t}\cdot w_{2_{\gamma_t(p)}}\right)}{t} \\
&= \lim_{t\to 0} \frac{T^*\gamma_t\cdot\alpha_{\gamma_t(p)}\left(T\gamma_{-t}\cdot w_{1_{\gamma_t(p)}}, T\gamma_{-t}\cdot w_{2_{\gamma_t(p)}}\right) - \alpha_p\left(T\gamma_{-t}\cdot w_{1_{\gamma_t(p)}}, T\gamma_{-t}\cdot w_{2_{\gamma_t(p)}}\right)}{t} \\
&= \lim_{t\to 0} \left(\frac{T^*\gamma_t\cdot\alpha_{\gamma_t(p)} - \alpha_p}{t}\right)\left(T\gamma_{-t}\cdot w_{1_{\gamma_t(p)}}, T\gamma_{-t}\cdot w_{2_{\gamma_t(p)}}\right) \\
&= \big(\mathscr{L}_v\alpha\big)_p(w_{1_p}, w_{2_p}).
\end{aligned}$$

The second term following the last equality gives us

$$\begin{aligned}
&\lim_{t\to 0} \frac{\alpha_p\left(T\gamma_{-t}\cdot w_{1_{\gamma_t(p)}}, T\gamma_{-t}\cdot w_{2_{\gamma_t(p)}}\right) - \alpha_p\left(w_{1_p}, T\gamma_{-t}\cdot w_{2_{\gamma_t(p)}}\right)}{t} \\
&= \lim_{t\to 0} \alpha_p\left(\frac{T\gamma_{-t}\cdot w_{1_{\gamma_t(p)}} - w_{1_p}}{t}, T\gamma_{-t}\cdot w_{2_{\gamma_t(p)}}\right) \\
&= \alpha_p\left((\mathscr{L}_v w_1)_p, w_{2_p}\right).
\end{aligned}$$

The third term following the last equality is very similar, giving us

$$
\begin{aligned}
\lim_{t\to 0} \frac{\alpha_p\left(w_{1_p}, T\gamma_{-t}\cdot w_{2_{\gamma_t(p)}}\right) - \alpha_p\left(w_{1_p}, w_{2_p}\right)}{t}
&= \lim_{t\to 0} \alpha_p\left(w_{1_p}, \frac{T\gamma_{-t}\cdot w_{2_{\gamma_t(p)}} - w_{2_p}}{t}\right) \\
&= \alpha_p\left(w_{1_p}, (\mathscr{L}_v w_2)_p\right).
\end{aligned}
$$

Thus we have, leaving off the base point,

$$\mathscr{L}_v\alpha(w_1, w_2) = \big(\mathscr{L}_v\alpha\big)(w_1, w_2) + \alpha(\mathscr{L}_v w_1, w_2) + \alpha(w_1, \mathscr{L}_v w_2).$$

This can be rewritten as

$$\big(\mathscr{L}_v\alpha\big)(w_1, w_2) = \mathscr{L}_v \underbrace{\alpha(w_1, w_2)}_{\text{a function}} - \alpha(\mathscr{L}_v w_1, w_2) - \alpha(w_1, \mathscr{L}_v w_2).$$

Recalling the definition of the Lie derivative of a function and the Lie derivative commutator identity this can again be rewritten as

$$\big(\mathscr{L}_v\alpha\big)(w_1, w_2) = v\big[\alpha(w_1, w_2)\big] - \alpha([v, w_1], w_2) - \alpha(w_1, [v, w_2]).$$

This is the global formula for the Lie derivative of a two-form.

Question A.20 Prove the global formula for the Lie derivative of a general k-form.

Global Lie Derivative Formula for differential forms:

$$\big(\mathscr{L}_v\alpha\big)(w_1, \dots, w_k) = v\big[\alpha(w_1, \dots, w_k)\big] - \sum_{i=1}^{k} \alpha(w_1, \dots, [v, w_i], \dots, w_k).$$

The global Lie derivative formula for a general rank (r, s)-tensor $\mathscr{T}$ is given by

Global Lie Derivative Formula for (r, s)-tensors:

$$
\begin{aligned}
\big(\mathscr{L}_v\mathscr{T}\big)(\alpha_1, \dots, \alpha_r, w_1, \dots, w_s) &= v\big[\mathscr{T}(\alpha_1, \dots, \alpha_r, w_1, \dots, w_s)\big] \\
&\quad - \sum_{i=1}^{r} \mathscr{T}(\alpha_1, \dots, \mathscr{L}_v\alpha_i, \dots, \alpha_r, w_1, \dots, w_s) \\
&\quad - \sum_{i=1}^{s} \mathscr{T}(\alpha_1, \dots, \alpha_r, v_1, \dots, \mathscr{L}_v w_i, \dots, w_s).
\end{aligned}
$$

Question A.21 Prove the global Lie derivative formula for a general rank (r, s)-tensor $\mathscr{T}$.

Global Formula for Exterior Differentiation

There are actually quite a number of other important identities and formulas that involve Lie derivatives, and if this were a full-fledged course in differential geometry they would need to be presented and proved. However, we will forego them. We have now covered Lie derivatives in sufficient detail that you have a basic understanding of what they are. We now will close this appendix with exploring the relationship between Lie derivatives and exterior differentiation. Since exterior derivatives only apply to differential forms, or skew-symmetric covariant tensors, then it is obvious that Lie derivatives are, in a sense, more general then exterior derivatives since you can take the Lie derivatives of any kind of tensors.

Let's take another look at Cartan's magic formula,

$$\mathscr{L}_v\alpha = \iota_v(d\alpha) + d(\iota_v\alpha).$$

The Lie derivatives of the k-form α along the flow of v is equal to the interior product by v of the exterior derivative of α plus the exterior derivative of the interior product by v of α. In other words, Cartan's magic formula allows us to write the Lie derivative of α in terms of the exterior derivatives of α. Thus we start to see that there really is a relationship between the two kinds of derivatives.

The question is, can we go the other way and write the exterior derivative in terms of the Lie derivative? The answer is of course yes, but it is in terms of the Lie derivative of vector fields, not k-forms. In fact, this formula is also called the global formula for exterior differentiation and was given in Sect. 4.4.2. The proof of the global formula for exterior differentiation for a general k-form requires induction. We begin by showing the formula for a one-form α,

$$\begin{aligned} d\alpha(v,w) &= v[\alpha(w)] - w[\alpha(v)] - \alpha\big([v,w]\big) \\ &= v\alpha(w) - w\alpha(v) - \alpha\big([v,w]\big). \end{aligned}$$

Question A.22 Show that both sides of the above equation are linear. To do this you need to show that if $\alpha = \sum f_i dx^i$ then $d\big(\sum f_i dx^i\big)(v,w) = \sum d\big(f_i dx^i\big)(v,w)$, $v\big[\big(\sum f_i dx^i\big)(w)\big] = \sum v\big[f_i dx^i(w)\big]$, and $\big(\sum f_i dx^i\big)\big([v,w]\big) = \sum \big(f_i dx^i\big)\big([v,w]\big)$.

It is enough to assume we are in a single coordinate patch and so a general one-form is written as $\alpha = \sum f_i dx^i$. Since both sides of the above equation are linear we only need to prove the equation for a single term $f_i dx^i$. But remember that x^i is simply the coordinate function. For simplicity's sake we will prove the equation for the one-form $\alpha = f dg$, where both f and g are functions on the manifold. Thus we have

$$\begin{aligned} d\alpha(v,w) &= d(f dg)(v,w) \\ &= \big(df \wedge dg\big)(v,w) \\ &= df(v)dg(w) - df(w)dg(v) \\ &= v[f]\cdot w[g] - w[f]\cdot v[g] \\ &= (vf)(wg) - (wf)(vg). \end{aligned}$$

Next we look at the term $v\alpha(w)$,

$$\begin{aligned} v\alpha(w) &= v\big(f dg\big)(w) \\ &= v\big(f dg(w)\big) \\ &= v\big(f \cdot \underbrace{w[g]}_{\text{a function}}\ \big) \\ &= v[f]\cdot w[g] + f\cdot v\big[w[g]\big] && \text{Product Rule} \\ &= (vf)(wg) + f\cdot vwg. \end{aligned}$$

Similarly we have $w\alpha(v) = (wf)(vg) + f\cdot wvg$. Since $[v,w] = vw - wv$ we also have

$$\begin{aligned} \alpha\big([v,w]\big) &= (f dg)\big([v,w]\big) \\ &= f\cdot [v,w][g] \\ &= f\cdot \big(vw - wv)[g] \\ &= f\cdot vw[g] - f\cdot wv[g] \\ &= f\cdot vwg - f\cdot wvg. \end{aligned}$$

Putting everything together we have

$$\begin{aligned} v\alpha(w) - w\alpha(v) - \alpha\big([v,w]\big) &= (vf)(wg) + f\cdot vwg \\ &\qquad - (wf)(vg) - f\cdot wvg \\ &\qquad\quad - f\cdot vwg + f\cdot wvg \\ &= (vf)(wg) - (wf)(vg) \\ &= d\alpha(v,w), \end{aligned}$$

which is the identity we wanted to show. This identity is the base case of the global formula for exterior derivatives. Notice that we could of course also write this using the Lie derivative as

$$d\alpha(v,w) = v\alpha(w) - w\alpha(v) - \alpha\big(\mathscr{L}_v w\big).$$

Of course, if we let $g = x^i$, the coordinate function, and then use the linearity that you proved in the last question we can see that this formula applies for all one-forms $\alpha = \sum f_i dx^i$, and hence we have completed the proof of our base case. Notice that this formula for the exterior derivative of a one-form does not depend at all on which coordinate system we use, thus it is called a coordinate independent formula.

We now turn to proving the global formula for the exterior derivative of a k-form α. The global formula is

Global Exterior Derivative Formula:

$$\begin{aligned} (d\alpha)(v_0,\ldots,v_k) = \sum_{i=0}^{k} &(-1)^i\ v_i\big[\alpha(v_0,\ldots,\widehat{v_i},\ldots,v_k)\big] \\ + \sum_{0\le i<j\le k} &(-1)^{i+j}\ \alpha\big([v_i,v_j],v_0,\ldots,\widehat{v_i},\ldots,\widehat{v_j},\ldots,v_k\big). \end{aligned}$$

Question A.23 Show that for a one-form α the global formula above simplifies to our base case formula.

We have already proved the base case for a one-form. Our induction hypothesis is that this formula is true for all $(k-1)$-forms. We will use the induction hypothesis, along with the global Lie derivative formula of a k-form, to show that this formula is true for a k-form α. Using Cartan's magic formula we have

$$\begin{aligned} d\alpha(v_0,v_1,\ldots,v_k) &= \big(\iota_{v_0} d\alpha\big)(v_1,\ldots,v_k) \\ &= \big(\mathscr{L}_{v_0}\alpha - d\iota_{v_0}\alpha\big)(v_1,\ldots,v_k) \\ &= \big(\mathscr{L}_{v_0}\alpha\big)(v_1,\ldots,v_k) - \big(d\iota_{v_0}\alpha\big)(v_1,\ldots,v_k). \end{aligned}$$

The first term is given by the global Lie derivative formula,

$$\begin{aligned} \big(\mathscr{L}_{v_0}\alpha\big)(v_1,\ldots,v_k) &= v_0\big[\alpha(v_1,\ldots,v_k)\big] - \sum_{i=1}^{k} \alpha(v_1,\ldots,[v_0,v_i],\ldots,v_k) \\ &= v_0\big[\alpha(v_1,\ldots,v_k)\big] - \sum_{i=1}^{k} (-1)^{i-1}\alpha([v_0,v_i],v_1,\ldots,\widehat{v_i},\ldots,v_k) \end{aligned}$$

and the second term is given by the induction hypothesis. That means we assume the global exterior derivative formula is true for the $(k-1)$-form $\iota_{v_0}\alpha$,

$$\begin{aligned} \big(d\iota_{v_0}\alpha\big)(v_1,\ldots,v_k) = \sum_{i=1}^{k} &(-1)^{i-1}\ v_i\big[\iota_{v_0}\alpha(v_1,\ldots,\widehat{v_i},\ldots,v_k)\big] \\ + \sum_{1\le i<j\le k} &(-1)^{i+j-1}\ \iota_{v_0}\alpha\big([v_i,v_j],v_1,\ldots,\widehat{v_i},\ldots,\widehat{v_j},\ldots,v_k\big) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{k} (-1)^{i-1}\ v_i\big[\alpha(v_0, v_1, \ldots, \widehat{v_i}, \ldots, v_k)\big] \\
&\quad + \sum_{1\le i<j\le k} (-1)^{i+j-1}\ (-1)\alpha\big([v_i, v_j], v_0, v_1, \ldots, \widehat{v_i}, \ldots, \widehat{v_j}, \ldots, v_k\big).
\end{aligned}
$$

Combining we end up with the following mess,

$$
\begin{aligned}
d\alpha(v_0, v_1, \ldots, v_k) &= \big(\mathscr{L}_{v_0}\alpha\big)(v_1, \ldots, v_k) - \big(d\iota_{v_0}\alpha\big)(v_1, \ldots, v_k) \\
&= v_0\big[\alpha(v_1, \ldots, v_k)\big] - \sum_{i=1}^{k} (-1)^{i-1}\ v_i\big[\alpha(v_0, v_1, \ldots, \widehat{v_i}, \ldots, v_k)\big] \\
&\quad + \sum_{j=1}^{k} (-1)^{0+j}\alpha([v_0, v_i], v_1, \ldots, \widehat{v_i}, \ldots, v_k) \\
&\quad - \sum_{1\le i<j\le k} (-1)^{i+j}\ \alpha\big([v_i, v_j], v_0, \ldots, \widehat{v_i}, \ldots, \widehat{v_j}, \ldots, v_k\big).
\end{aligned}
$$

The first two terms sum to the first term of the global exterior derivative formula while the third and forth terms sum to the second term of the global exterior derivative formula. Pay attention to the signs and how the dummy indices are used and make sure you understand each step. Thus we have shown the global exterior derivative formula. Sometimes the definition of the exterior derivative is given in terms of this formula. The reason this is sometimes done is that this formula is totally independent of what basis you are using. Notice that the commutator $[v_i, v_j]$ of two vector fields v_i and v_j can be introduced and explained without referring to the Lie derivative.

A.8 Summary and References

A.8.1 Summary

A tensor on a manifold is a multilinear map

$$
\begin{aligned}
T : \underbrace{T^*M \times \cdots \times T^*M}_{\substack{r \\ \text{contravariant} \\ \text{degree}}} \times \underbrace{TM \times \cdots \times TM}_{\substack{s \\ \text{covariant} \\ \text{degree}}} &\longrightarrow \mathbb{R} \\
T\big(\alpha_1, \ldots, \alpha_r, v_1, \ldots, v_s\big) &\longrightarrow \mathbb{R}.
\end{aligned}
$$

That is, the map T eats r one-forms and s vectors and produces a real number.

Covariant tensors eat vectors and covariant tensor components transform like vector basis elements.

Contravariant tensors eat one-forms and contravariant tensor components transform like one-form basis elements.

Let us compare this transformation rule with that of covariant tensors.

Covariant Tensors	Contravariant Tensors
$\widetilde{T}_i = \frac{\partial x^j}{\partial u^i} T_j$	$\widetilde{T}^i = \frac{\partial u^i}{\partial x^j} T^j$
Vector Basis Elements	**One-Form Basis Elements**
$\frac{\partial}{\partial u^i} = \frac{\partial x^j}{\partial u^i} \frac{\partial}{\partial x^j}$	$du^i = \frac{\partial u^i}{\partial x^j} dx^j$

A general rank (r, w)-tensor $\mathscr{T}$'s components transform according to

Rank (r, w)-Tensor Transformation Rule: $$\widetilde{\mathscr{T}}^{k_1 \cdots k_r}_{l_1 \cdots l_s} = \frac{\partial u^{k_1}}{\partial x^{i_1}} \cdots \frac{\partial u^{k_r}}{\partial x^{i_r}} \frac{\partial x^{j_1}}{\partial u^{l_1}} \cdots \frac{\partial x^{j_s}}{\partial u^{l_s}} \mathscr{T}^{i_1 \cdots i_r}_{j_1 \cdots j_s}.$$

The tensor $\mathscr{T}$ is called anti-symmetric or skew-symmetric if it changes sign whenever any pair of its arguments are switched. That is,

$$\mathscr{T}\big(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k\big) = -\mathscr{T}\big(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k\big)$$

when v_i and v_j have switched their places and all other inputs have stayed the same.

A k-form is a skew-symmetric rank-k covariant tensor.

Thus we have that the set of k-forms is a subset of the set of $(0, k)$-tensors. Using the tensor definition of differential forms it is easy to prove an extremely important identity, that pullbacks distribute across wedgeproducts,

$$\boxed{\phi^*(\alpha \wedge \beta) = \phi^*\alpha \wedge \phi^*\beta.}$$

A metric on the manifold M is a smooth, symmetric, non-degenerate, rank-two covariant tensor g. Metric tensors are generally denoted with a lower-case g. A manifold that has such a metric on it is called a pseudo-Riemannian manifold. If the metric g also has one additional property, that $g(v, w) \geq 0$ for all vector fields v and w then it is called a Riemannian metric and the manifold is called a Riemannian manifold. Metrics are used to define an inner-product on the manifold, which in term is used to define the lengths of curves on a manifold. The length of the curve γ from p to q is defined to be

$$L(\gamma) = \int_a^b \sqrt{\big|g\big(\dot{\gamma}(t), \dot{\gamma}(t)\big)\big|}\, dt.$$

With this we can then define the distance between the two points p and q as

$$d(p, q) = \inf_{\gamma} \; L(\gamma),$$

where γ is any piecewise continuous curve that connects p and q. Thus a metric is essential for many of the concepts we take for granted on Euclidian manifolds $\mathbb{R}^n$.

Another concept of differentiation on a manifold called the Lie derivative was introduced. The lie derivative depends on there existing integral curves γ for a vector field v. The Lie derivative can be applied to vector fields, differential forms, functions, and tensors.

Lie Derivative of Vector Field: $$(\mathscr{L}_v w)_p = \lim_{t \to 0} \frac{T\gamma_{-t} \cdot w_{\gamma_t(p)} - w_p}{t}.$$

Lie Derivative of One-Form: $$(\mathscr{L}_v \alpha)_p = \lim_{t \to 0} \frac{T^*\gamma_t \cdot \alpha_{\gamma_t(p)} - \alpha_p}{t},$$

The Lie derivative of a function was found to be exactly equivalent to directional derivative,

$$\boxed{\text{Lie Derivative of Function:} \qquad (\mathscr{L}_v f)_p = v_p[f].}$$

We will define the pull-back of the tensor $\mathscr{T}$ by γ_t with the following

$$\big(\gamma_t^* \mathscr{T}\big)_p \big(\alpha_1, \ldots, \alpha_r, v_1, \ldots, v_s\big) = \mathscr{T}_{\gamma_t(p)}\big(\gamma_{-t}^* \alpha_1, \ldots, \gamma_{-t}^* \alpha_r, \gamma_{t*} v_1, \ldots, \gamma_{t*} v_s\big).$$

Defining the pull-pack of a tensor this way we define the Lie derivative of the tensor $\mathscr{T}$ with the same formula that we used to define the Lie derivative of a one-form,

$$\boxed{\text{Lie Derivative of Tensor:} \qquad (\mathscr{L}_v \mathscr{T})_p = \lim_{t\to 0} \frac{\gamma_t^* \mathscr{T}_{\gamma_t(p)} - \mathscr{T}_p}{t}.}$$

Notice that since k-forms are a special kind of tensor this definition covers differential k-forms as well.

The following identities describe how the Lie derivative act over addition of tensors, tensor products, and wedgeproducts,

$$\boxed{\mathscr{L}_v\big(a\mathscr{S} + b\mathscr{T}\big) = a\mathscr{L}_v \mathscr{S} + b\mathscr{L}_v \mathscr{T},}$$

$$\boxed{\mathscr{L}_v(\mathscr{S} \otimes \mathscr{T}) = \mathscr{L}_v \mathscr{S} \otimes \mathscr{T} + \mathscr{S} \otimes \mathscr{L}_v \mathscr{T},}$$

$$\boxed{\mathscr{L}_v(\alpha \wedge \beta) = \mathscr{L}_v \alpha \wedge \beta + \alpha \wedge \mathscr{L}_v \beta.}$$

The next identity involves the Lie derivative of the product of functions f and vector fields w,

$$\boxed{\mathscr{L}_v(fw) = f\mathscr{L}_v w + v[f]w.}$$

The exterior derivative of the lie derivative of a differential form α is given by

$$\boxed{d\big(\mathscr{L}_v \alpha\big) = \mathscr{L}_v\big(d\alpha\big).}$$

Next is one of the most important identities. It shows the relationship between the Lie derivative, the exterior derivative, and the interior product,

$$\boxed{\text{Cartan's Homotopy Formula:} \qquad \mathscr{L}_v \alpha = \iota_v(d\alpha) + d(\iota_v \alpha).}$$

This identity is very often simply written as $\mathscr{L}_v = \iota_v d + d\iota_v$. Finally the Lie derivative of a vector field can be written as the commutator of the two vector fields,

$$\boxed{\mathscr{L}_v w = [v, w] = vw - wv.}$$

Global formulas are formulas that do not use coordinates. They are sometimes called coordinate-free formulas. The global formula for the Lie derivative of differential forms is given by

$$\boxed{\begin{array}{l}\text{Global Lie Derivative Formula for differential forms:} \\ \qquad \big(\mathscr{L}_v \alpha\big)(w_1, \ldots, w_k) = v\big[\alpha(w_1, \ldots, w_k)\big] - \displaystyle\sum_{i=1}^{k} \alpha(w_1, \ldots, [v, w_i], \ldots, w_k).\end{array}}$$

The global Lie derivative formula for a general rank (r, s)-tensor $\mathscr{T}$ is given by

Global Lie Derivative Formula for differential forms:

$$\begin{aligned}\big(\mathscr{L}_v \mathscr{T}\big)(\alpha_1, \ldots, \alpha_r, w_1, \ldots, w_s) &= v\big[\mathscr{T}(\alpha_1, \ldots, \alpha_r, w_1, \ldots, w_s)\big] \\ &\quad - \sum_{i=1}^{r} \mathscr{T}(\alpha_1, \ldots, \mathscr{L}_v \alpha_i, \ldots, \alpha_r, w_1, \ldots, w_s) \\ &\quad - \sum_{i=1}^{s} \mathscr{T}(\alpha_1, \ldots, \alpha_r, v_1, \ldots, \mathscr{L}_v w_i, \ldots, w_s).\end{aligned}$$

The global formula for the exterior derivative of a k-form α is given by

Global Exterior Derivative Formula:

$$\begin{aligned}(d\alpha)(v_0, \ldots, v_k) &= \sum_{i=0}^{k} (-1)^i \; v_i\big[\alpha(v_0, \ldots, \widehat{v_i}, \ldots, v_k)\big] \\ &\quad + \sum_{0 \le i < j \le k} (-1)^{i+j} \; \alpha\big([v_i, v_j], v_0, \ldots, \widehat{v_i}, \ldots, \widehat{v_j}, \ldots, v_k\big).\end{aligned}$$

A.8.2 References and Further Reading

Many books on advanced calculus, manifolds, and physics begin with tensors and then introduce differential forms as a particular kind of tensor. For example, Martin [33] and Bushop and Goldberg [6] take this approach. Several additional very good references for tensors which we relied on are Tu [46], Kay [29], Renteln [37], and Abraham, Marsden, and Ratiu [1]. The very short book by Domingos [15] is also a nice introduction.

Appendix B
Some Applications of Differential Forms

In this appendix we will attempt to broadly show some of the fields of mathematics and physics where differential forms are used. After reading a book this size on differential forms it would be nice for the reader to have some idea of what some of the applications are, other than electromagnetism. None of these sections aim for anything approaching completeness. Instead, they are meant to broaden the reader's horizons, to give just a taste of each subject, and hopefully to whet the reader's appetite for more.

Section one introduces de Rham cohomology in a very general way. In a sense this is one of the "hard core" mathematical applications of differential forms. Basically, differential forms can be used to study the topology of manifolds. In section two we look at some examples of de Rham cohomology groups and see how they can be used to learn about the global topology of the manifold. In section three the idea of a symplectic manifold is introduced, as is the most common symplectic manifold around, the cotangent bundle of any manifold. Symplectic manifolds lead to symplectic geometry. Section four covers the Darbaux theorem, which is a fundamental result in symplectic geometry. And then section five discusses a physics application of symplectic geometry, namely geometric mechanics.

B.1 Introduction to de Rham Cohomology

In this next section we will give a very brief introduction of a field of mathematics called de Rham cohomology. In essence, de Rham cohomology creates a link between differential forms on a manifold M and the global topology of that manifold. This is an incredible and rich area of mathematics and we will not be able to do any more than give an introductory glimpse of it. We will be using differential forms, and in particular the relation between closed and exact forms, to determine global topological information about manifolds. This basically allows us to use what are in essence calculus-like computations to determine global topological properties. But now what do we mean by the words "global" and "topological"? We will not attempt to give rigorous definition but will try to give you a feel for what is meant.

You may roughly think of **topology** as the study of how manifolds, or subsets of $\mathbb{R}^n$, are connected and how many "holes" the manifold or subset has. And what do we mean by global, as opposed to local, properties of a manifold? Recall our intuitive idea of a manifold is something that is locally Euclidian. A little more precisely, an n-dimensional manifold is covered with overlapping patches U_i which are the image of mappings $\phi_i : \mathbb{R}^n \to U_i$. For example, in Fig. B.1 the manifold $M = \mathbb{R}^2 - \{(0, 0)\}$, which is the plane with the origin point missing, is shown with four coordinate patches. Every point $p \in M$ is contained in some open neighborhood which is essentially identical to an open neighborhood of $\mathbb{R}^2$.

So the local information of $M = \mathbb{R}^2 - \{(0, 0)\}$ is the same as the local information in $\mathbb{R}^2$. But clearly in M we have one point, the origin, missing. Since there is only one point missing you may be tempted to think, using the common usage of the word local, that the missing point is a local property, but it is not. The missing origin is, in fact, a global property; it says something about the whole manifold, namely that the manifold has a hole in it. The fact that M has a hole in it is a global topological property.

We can use certain differential forms defined on a manifold to detect global properties like holes. When it comes to detecting global properties certain forms are interesting and certain forms are not interesting at all. The interesting forms turn out to be those whose exterior derivative is zero, that is, the closed forms.

J. P. Fortney, *A Visual Introduction to Differential Forms and Calculus on Manifolds*,
https://doi.org/10.1007/978-3-319-96992-3

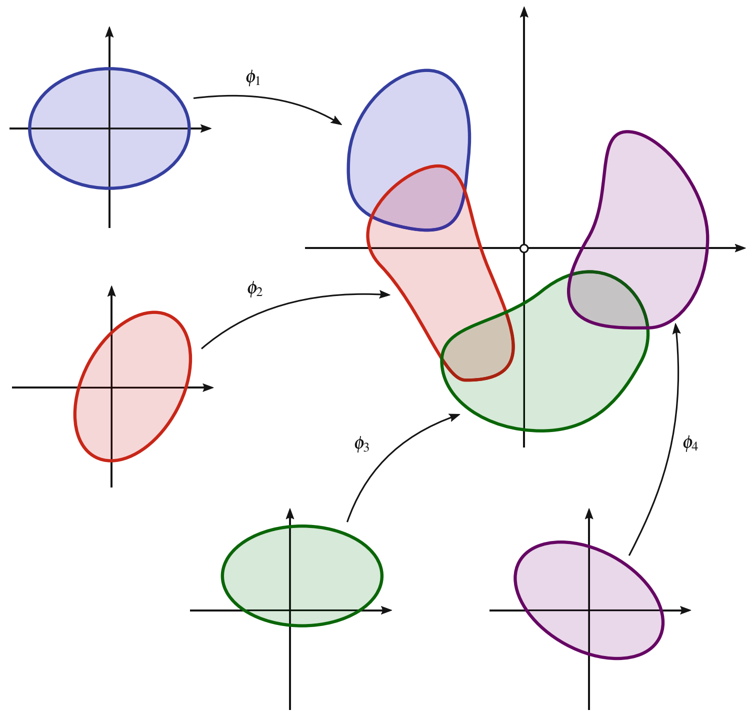


Fig. B.1 The manifold $M = \mathbb{R}^2 - \{(0, 0)\}$ shown with four coordinate patches

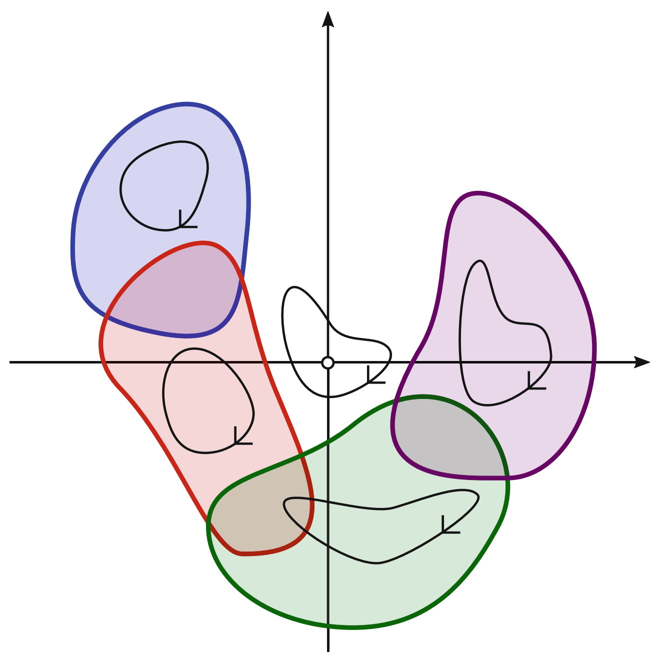


Fig. B.2 The manifold $M = \mathbb{R}^2 - \{(0, 0)\}$ is shown with five small closed curves. For of these curves are contained in a single coordinate patch, which the fifth curve encloses the hole at the origin. This curve can not be drawn entirely inside a single coordinate patch

For the moment we will set aside the comment that closed forms give global information and instead try to illustrate to you that closed forms do not give local information. Suppose α is a closed form. Looking locally at M, that is, inside some patch U_i, we integrate α over some small closed curve C that lies entirely inside the patch U_i; that is, over a "local" curve. See Fig. B.2.

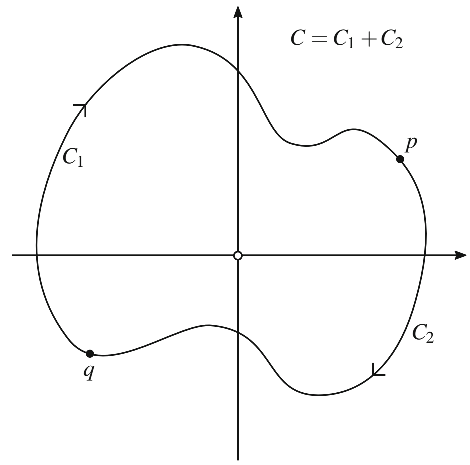


Fig. B.3 The closed curve C broken into two curves C_1 and C_2

Using Stokes' theorem on the local curve C, which is inside a single coordinate patch and thus the boundary of some domain D, we have

$$\int_{C=\partial D} \alpha = \int_D d\alpha = \int_D 0 = 0.$$

Since the integral of α over a local curve is zero this tells us nothing about the manifold locally. So closed forms do not give us any local information about the manifold. Of course by now you may already be guessing how the interesting "global" nature of closed forms arises. What if we integrate a closed α on a curve that goes around the missing point at the origin? First of all, this curve is no longer contained entirely in a single patch U_i. If we had $\int_C \alpha \neq 0$ then we would know that $C \neq \partial D$ for some domain D and that there must be something strange going on inside the curve C. The very fact that there exists a curve C such that $\int_C \alpha \neq 0$ for a closed α tells us some global topological information about M. If C is not the boundary for some region D then C must enclose a hole of some sort.

Now suppose that α were exact, that is, there exists a β such that $\alpha = d\beta$. We will see that exact forms are not interesting. For any closed curve C, whether or not C encloses a hole, by choosing two points p and q on the curve C we can split the curve into two parts, C_1, which goes from q to p and C_2, which goes from p to q, as is shown in Fig. B.3. Since C is now made up of the two curves C_1 and C_2 together. We often write $C = C_1 + C_2$. We then have

$$\begin{aligned}
\int_C \alpha &= \int_{C_1+C_2} \alpha \\
&= \int_{C_1} \alpha + \int_{C_2} \alpha \\
&= \int_{C_1} d\beta + \int_{C_2} d\beta \\
&= \int_{\partial C_1} \beta + \int_{\partial C_2} \beta \\
&= \int_{\{p\}-\{q\}} \beta + \int_{\{q\}-\{p\}} \beta \\
&= \int_{\{p\}-\{q\}} \beta - \int_{\{p\}-\{q\}} \beta \\
&= 0.
\end{aligned}$$

In other words, for an exact form α we have $\int_C \alpha = 0$ whether or not C encloses a hole. In other worlds, exact forms are not interesting, at least when it comes to telling us something about holes.

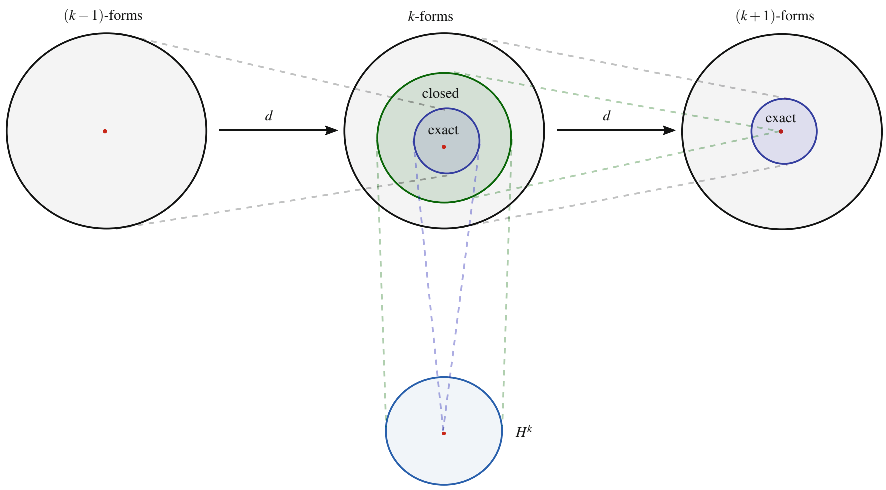


Fig. B.4 We want to measure how different the set of closed k-forms is from the set of exact k-forms. We do this by means of the equivalence class $H^k_{dR}(M)$ (Bachman [4], 2006, p.118)

If we were in the case where our manifold were $\mathbb{R}^n$ then by the Poincaré lemma every closed form is exact. That means that the set of closed forms is the same as the set of exact forms, so then there is no curve C such that for closed α we have $\int_C \alpha \neq 0$. But that should not surprise us too much since we know that there are no holes in $\mathbb{R}^n$.

We are basically interested in the case where the set of closed forms is different than the set of exact forms. We want some way to measure how different these two sets are. In Fig. B.4 we use our Venn-like picture to aid us in visualizing the situation.

What we are interested in is some way to measure how far away from exact a set of closed forms are. We do this by means of what is called an **equivalence relation**. An equivalence relation is simply a way of talking a set of objects and creating another set of objects by defining certain objects in the original set to be "the same." In fact, we have already seen this concept in action when we used equivalence classes of curves to define tangent vectors in Sect. 10.2. There we defined two curves as equivalent, or "the same", if certain criteria were met. Here, two closed k-forms, α and β on a manifold M are considered to be the same, or equivalent, if the difference between them is exact. In other words, we will consider α and β to be equivalent if $\alpha - \beta$ is an exact k-form on M. To be more precise, suppose that α and β are two closed k-forms on M. Then we say that α is equivalent to β, which we write as $\alpha \sim \beta$, if

$$\alpha - \beta = d\omega$$

for some $(k-1)$-form ω on M. We denote the equivalence class of α by

$$[\alpha] = \left\{ \beta \in \bigwedge^k(M) \;\middle|\; \alpha \sim \beta \right\}.$$

When we do this we end up with a set whose elements are the equivalence classes of closed forms that differ by an exact form. This set is denoted $H^k_{dR}(M)$

$$H^k_{dR}(M) = \frac{\left\{ \alpha \in \bigwedge^k(M) \;\middle|\; d\alpha = 0 \right\}}{\left\{ d\omega \in \bigwedge^k(M) \;\middle|\; \omega \in \bigwedge^{k-1}(M) \right\}}.$$

When read out loud the horizonal line in the middle is pronounced modulo, or mod for short. The right hand side would be read "the closed forms mod the exact forms." This $H^k_{dR}(M)$ is called the kth **de Rham cohomology group**. What this vector space $H^k_{dR}(M)$ is actually doing is telling us something about the connectedness and number of holes the manifold M has.

We will not, in this book, explain what the word group means, but if you have not had an abstract algebra class yet then be aware that the word group has a very precise meaning in mathematics and is not to be used lightly or imprecisely. Unfortunately, or perhaps fortunately, for us the Rham cohomology group is actually also a vector space, which was defined in Sect. 1.1. It turns out that a vector space is indeed an abelian group with a scalar multiplication defined on it. That is, a vector space is an abelian group with some additional structure defined on it; the additional structure is scalar multiplication. Scalar multiplication just means that you can multiply the elements of the space by a scalar, which is simply another word for a real number $c \in \mathbb{R}$. So, calling the de Rham cohomology group a group, while not inaccurate, is not as accurate as it could be either. The de Rham cohomology group is a vector space. But as happens in language, once a name gets attached to an object and everyone gets used to that name it is basically impossible to change it, so any attempt to change the terminology to the de Rham cohomology vector space would be pointless.

But why do we care about these de Rham cohomology groups? It turns out that the de Rham cohomology groups on a manifold tell us something about certain global topological properties of the manifold. In the next section we will look at a few examples to see how this works.

B.2 de Rham Cohomology: A Few Simple Examples

Now we will compute a few de Rham cohomology groups and try to explain what they mean. While most of the detailed computations in his section have been left as exercises for you to try to do if you are so inclined, don't let yourself get bogged down in details. Many of these exercises may be quite difficult. And the reality is that de Rham cohomology has a great deal of advanced mathematical machinery that is actually used in doing these sorts of computations that is far beyond the purview of this book. This is only an extremely short introduction to a vast subject. For the moment it is sufficient to simply walk away with a general feel for the fact that computations involving differential forms can actually tell you something interesting about a manifold's global topology.

Suppose that M is a manifold. For convenience' sake we will denote the vector space of closed k-forms on M by

$$Z^k(M) = \left\{ \alpha \in \bigwedge^k(M) \,\middle|\, d\alpha = 0 \right\}$$

and the set of exact k-forms on M by

$$B^k(M) = \left\{ d\omega \in \bigwedge^k(M) \,\middle|\, \omega \in \bigwedge^{k-1}(M) \right\}.$$

This makes it easier to discuss the vector spaces of closed and exact forms. Thus we have

$$H^k_{dR}(M) = Z^k(M)/B^k(M).$$

The Manifold $\mathbb{R}$

We start with the manifold $M = \mathbb{R}$. To find $H^0_{dR}(\mathbb{R})$ we need to find both $Z^0(\mathbb{R})$ and $B^0(\mathbb{R})$. We first turn our attention to $Z^0(\mathbb{R})$, the set of closed zero-forms on $\mathbb{R}$. What are the zero-forms on $\mathbb{R}$? They are just the functions f on $\mathbb{R}$. Thus we have

$$Z^0(\mathbb{R}) = \left\{ f : \mathbb{R} \to \mathbb{R} \,\middle|\, df = 0 \right\}.$$

To find these we take the exterior derivative of f, $df = \frac{\partial f}{\partial x^1} dx^1$ which of course implies that $\frac{\partial f}{\partial x^1} = 0$. For what real-valued functions is this true? Those functions that do not change at all, in other words, the constant functions $f(x^1) = c$ for some constant $c \in \mathbb{R}$. Thus $Z^0(\mathbb{R})$ is simply the set of all constant functions. Since any real number can be a constant we have that $Z^0(\mathbb{R})$ is isomorphic to $\mathbb{R}$, written as $Z^0(\mathbb{R}) \simeq \mathbb{R}$. Isomorphisms of vector spaces were discussed in Sect. 3.1. Now we consider $B^0(\mathbb{R})$, the set of exact 0-forms on $\mathbb{R}$. But since there are no such thing as (-1)-forms then no zero-form can be exact so we say that $B^0(\mathbb{R}) = \{0\}$. This gives

$$H^0_{dR}(\mathbb{R}) = Z^0(\mathbb{R})/B^0(\mathbb{R}) = \mathbb{R}/\{0\} = \mathbb{R}.$$

To find $H^1_{dR}(\mathbb{R})$ we turn our attention to $Z^1(\mathbb{R})$ and $B^1(\mathbb{R})$. But here things become easy because of the Poincaré lemma which says that every closed form on $\mathbb{R}^n$ is exact. This of course means that $Z^1(\mathbb{R}) = B^1(\mathbb{R})$ and hence

$$H^1_{dR}(\mathbb{R}) = \{0\} \equiv 0.$$

In fact, by the Poincaré lemma we have that $Z^k(\mathbb{R}) = B^k(\mathbb{R})$ for all $k > 0$ so we actually have

$$H^k_{dR}(\mathbb{R}) = \{0\} \equiv 0.$$

for every $k > 0$.

The Manifold $\mathbb{R}^n$

Question B.1 For the manifold $\mathbb{R}^n$ show that

$$H^k_{dR}(\mathbb{R}^n) = \begin{cases} \mathbb{R} & \text{if } k = 0, \\ 0 & \text{if } k \geq 1. \end{cases}$$

The Manifold $\mathbb{R} - \{(0)\}$

Now we turn our attention to the manifold $\mathbb{R} - \big\{(0)\big\} = (-\infty, 0) \cup (0, \infty)$, which is the real line $\mathbb{R}$ with the origin removed. We have not spent any time discussing manifolds of this nature, but it is perfectly possible to have a manifold with components that are not connected to each other. Consider $\mathbb{R} - \big\{(0)\big\}$; we say that this manifold has two connected components. Implicit in this is that these two connected components are not connected to each other. In topology there are several different ways to define what **connected** mean, but here we will stick with the simplest and most intuitive idea, that of **path-connectedness**. A space is called path-connected if, for any two points in that space, it is possible to find a path between those two points. Consider the left side of $\mathbb{R} - \big\{(0)\big\}$, which is simply the negative numbers $(-\infty, 0)$. It is possible to find a path in $(-\infty, 0)$ that connects any two negative numbers. Similarly, the right side of $\mathbb{R} - \big\{(0)\big\}$, which is the set of positive numbers $(0, \infty)$, is path-connected as well. However, any potential path that would connect a negative number to a positive number would have to go through the origin, which has been removed. Hence $(-\infty, 0)$ and $(0, \infty)$ are not connected to each other and so we say the manifold $\mathbb{R} - \big\{(0)\big\}$ has two connected components.

Of course it is possible to define both functions and differential forms on a manifold like $\mathbb{R} - \big\{(0)\big\}$ that has multiple connected components. At times it makes sense to write a function f on $\mathbb{R} - \big\{(0)\big\}$ as

$$f = \begin{cases} f(x) & x \in (-\infty, 0), \\ \tilde{f}(x) & x \in (0, \infty). \end{cases}$$

In order to find $Z^0\Big(\mathbb{R} - \big\{(0)\big\}\Big)$ we want to find all closed functions on $\mathbb{R} - \big\{(0)\big\}$. By definition of closed we need to have $df = 0$, which means that

$$0 = df = \begin{cases} \frac{\partial f(x)}{\partial x} dx & x \in (-\infty, 0) \\ \frac{\partial \tilde{f}(x)}{\partial x} dx & x \in (0, \infty), \end{cases}$$

which leads to

$$\frac{\partial f(x)}{\partial x} = 0 \qquad \text{and} \qquad \frac{\partial \tilde{f}(x)}{\partial x} = 0,$$

which only occurs if both $f(x)$ and $\tilde{f}(x)$ are constant functions. But there is no reason to expect that the function f takes on the same constant on each component.

Question B.2 Find df for the following functions,

(i) $f = \begin{cases} 7 & x \in (-\infty, 0) \\ -3 & x \in (0, \infty), \end{cases}$

(ii) $f = \begin{cases} -6 & x \in (-\infty, 0) \\ 5 & x \in (0, \infty), \end{cases}$

(iii) $f = \begin{cases} \frac{14}{3} & x \in (-\infty, 0) \\ -6\pi & x \in (0, \infty). \end{cases}$

Question B.3 Explain why each closed zero-form on $\mathbb{R} - \{(0)\}$ can be exactly specified by two real numbers, and thus that $Z^0\Big(\mathbb{R} - \{(0)\}\Big) \simeq \mathbb{R}^2$.

Since no zero-form can be exact we have $B^0(\mathbb{R} - \{(0)\}) = \{0\}$. This means that

$$\begin{aligned} H^0_{dR}\Big(\mathbb{R} - \{(0)\}\Big) &= \frac{Z^0\Big(\mathbb{R} - \{(0)\}\Big)}{B^0\Big(\mathbb{R} - \{(0)\}\Big)} \\ &= \mathbb{R}^2 \big/ \{0\} \\ &= \mathbb{R}^2. \end{aligned}$$

The Manifold $\mathbb{R} - \{p\} - \{q\}$

Suppose $p, q \in \mathbb{R}$ and $p < q$. Consider the manifold $\mathbb{R}$ with the points p and q removed. As before we can write a function f on $\mathbb{R} - \{p\} - \{q\}$ as

$$f = \begin{cases} f_1(x) & x \in (-\infty, p), \\ f_2(x) & x \in (p, q), \\ f_3(x) & x \in (q, \infty). \end{cases}$$

Finding the closed zero-forms on $\mathbb{R} - \{p\} - \{q\}$ reduces to finding functions f_1, f_2, f_3 such that $\frac{\partial f_1(x)}{\partial x} = 0$, $\frac{\partial f_2(x)}{\partial x} = 0$, and $\frac{\partial f_3(x)}{\partial x} = 0$, which are clearly the constant functions. And of course on each component of $\mathbb{R} - \{p\} - \{q\}$ the constant may be a different value.

Question B.4 Show that $Z^0\big(\mathbb{R} - \{p\} - \{q\}\big) \simeq \mathbb{R}^3$ and hence that $H^0_{dR}\big(\mathbb{R} - \{p\} - \{q\}\big) = \mathbb{R}^3$.

Manifolds with *m* Connected Components

Question B.5 Let $p_1, p_2, \ldots, p_m \in \mathbb{R}$ be distinct points and let $M = \mathbb{R} - \{p_1\} - \{p_2\} - \cdots - \{p_m\}$. Show that $H^0_{dR}(M) = \mathbb{R}^m$.

Question B.6 Argue that for a general manifold M with m connected components that $H^0_{dR}(M) = \mathbb{R}^m$.

Thus we can see that the zeroth de Rham cohomology group actually tells us how many connected components our manifold has.

Connected *n*-Dimensional Manifolds

If M is a connected n-dimensional manifold, then by arguments already familiar to us we have

$$H^0_{dR}(M) = \mathbb{R}.$$

Question B.7 Find $Z^k(M)$ and $B^k(M)$ for $k > n$. Show that $H^k_{dR}(M) = 0$.

The Manifold S^1

The manifold S^1 is simply a circle, which we can consider to be the unit circle, see Fig. 2.8. That is, we will consider

$$S^1 = \left\{ (x, y) \;\middle|\; x^2 + y^2 = 1 \right\} \subset \mathbb{R}^2.$$

This manifold is connected, one-dimensional, and has a "hole" in it. Since it is connected and one-dimensional we already know that

$$H^0_{dR}(M) = \mathbb{R},$$

and

$$H^k_{dR}(M) = 0 \text{ for } k \geq 2.$$

Now we actually want to find what $H^1_{dR}(M)$ is. This will take some work. Let us begin by considering the one-form

$$\omega = \frac{-y}{x^2 + y^2}\, dx + \frac{x}{x^2 + y^2}\, dz.$$

It is easy to show that $d\omega = 0$ and hence that ω is closed.

Question B.8 Show that $d\omega = 0$.

Our next goal is to show that ω is not exact. We shall do this by assuming that ω is exact and then reason our way to a contradiction, thereby showing that ω can not be exact. Since ω is a one-form defined on S^1 and we are assuming it is exact then there must be some zero-form f_0 such that $\omega = df_0$. Now suppose that there is some other zero-form f_1 such that $\omega = df_1$ as well. This of course means that $df_1 = df_0$ and hence $f_1 = f_0 + f_2$ for some closed zero-form f_2. But since S^1 is a connected manifold then we know that the closed zero-forms on S^1 are the constants, and hence $f_1 = f_0 + c$ for some constant $c \in \mathbb{R}$. So, if we can find even a single zero-form f_0 such that $df_0 = \omega$ then we automatically know every possible zero-form f_1 such that $df_1 = \omega$.

We begin by simply presenting a candidate function for f_0. Let f_0 be the polar coordinate θ. From trigonometry we have that $\tan\theta = \frac{y}{x}$ for $x \neq 0$ and $\cot\theta = \frac{x}{y}$ for $y \neq 0$.

Question B.9 Using $\tan\theta = \frac{y}{x}$ show that $d\theta = \omega$ when $x \neq 0$. You will need to recall how to take derivatives of trigonometric functions and use the identity $\sec^2\theta = 1 + \tan^2\theta$.

Question B.10 Using $\cot\theta = \frac{x}{y}$ show that $d\theta = \omega$ when $y \neq 0$. You will need to recall how to take derivatives of trigonometric functions and use the identity $\csc^2\theta = 1 + \cot^2\theta$.

Clearly θ indeed works since $d\theta = \omega$. Of course the real problem, and the contradiction, lies with the actual "function" θ.

Question B.11 Using trigonometry argue that θ takes on multiple values at the point $(x, y) = (1, 1)$.

The "function" θ takes on multiple values at any point on S^1 and so it actually is not a function at all since it is not well-defined. But by our reasoning we know that any function f_1 that satisfies $df_1 = \omega$ must be of the form $\theta + c$ hence there is no function that satisfies $df_1 = \omega$ which means that ω must not be exact.

Question B.12 Suppose we had the manifold $\mathbb{R}^2 - \big\{x, 0) \bigm| x \leq 0\big\}$, that is, the plane with the non-positive portion of the x-axis removed. Show that the one-form ω is actually exact on $\mathbb{R}^2 - \big\{(x, 0) \bigm| x \leq 0\big\}$. Similarly, if S^1 is the unit circle, show that on the manifold $S^1 - \big\{(-1, 0)\big\}$ the one-form ω is exact.

Question B.13 Suppose that α is a closed one-form on S^1. Define $r = \frac{1}{2\pi}\int_{S^1} \alpha$. While a rigorous proof would be difficult, give a plausibility argument that $\psi = \alpha - r\omega$ is an exact one-form on S^1. Use this to then argue that $H^1_{dR}(M)$ is a one-dimensional vector space with basis being the equivalence class $[\omega]$.

So for the manifold S^1 we have

$$\begin{aligned} H^0_{dR}(S^1) &= \mathbb{R}, \\ H^1_{dR}(S^1) &= \mathbb{R}, \\ H^k_{dR}(S^1) &= 0, \quad k > 1. \end{aligned}$$

Putting It All Together

Now we are ready to put together the ideas from the last section and the examples from this section and try to understand, at least in a very general and somewhat imprecise way, what topological properties of the manifold the de Rham cohomology groups tell us. It is clear that the dimension of the vector space $H^0_{dR}(M)$ tells us how many connected components the manifold M has.