Appendix A

Some background results

A.1 Linear algebra

A.1.1 Vectors and co-vectors

Suppose we have a vector $\boldsymbol{v} \in V$, belonging to an $n-$dimensional real[1] vector space and let $B := \{\boldsymbol{e}_i\}_{i=1,\ldots n}$ be a basis for V. We can write

$$\boldsymbol{v} = \sum_{i=1}^{n} v^i \boldsymbol{e}_i,$$

where $v^i \in \mathbb{R}$ are the uniquely determined components with respect to the basis B. The dual space V^* is the $n-$dimensional real vector space of linear maps $\boldsymbol{\omega} : V \to \mathbb{R}$. Such maps are sometimes called one-forms or covectors. We can define the dual basis $B^* := \{\boldsymbol{e}^i\}_{i=1,\ldots n}$ uniquely by the requirement

$$\boldsymbol{e}^i(\boldsymbol{e}_j) = \delta^i{}_j, \qquad i, j = 1, \ldots, n,$$

where

$$\delta^i{}_j = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

is the *Kronecker delta.* For any $\boldsymbol{\omega} \in V^*$ we can write

$$\boldsymbol{\omega} = \sum_{i=1}^{n} \omega_i \boldsymbol{e}^i,$$

[1] for simplicity. Similar constructions exist over other fields.

so that

$$\begin{aligned}\boldsymbol{\omega}(\boldsymbol{v}) &= \sum_{i=1}^{n} \omega_i \boldsymbol{e}^i \left(\sum_{j=1}^{n} v^j \boldsymbol{e}_j \right), \\ &= \sum_{i,j=1}^{n} \omega_i v^j \boldsymbol{e}^i(\boldsymbol{e}_j) = \sum_{i,j=1}^{n} \omega_i v^j \delta^i{}_j, \\ &= \sum_{i=1}^{n} \omega_i v^i.\end{aligned}$$

A.1.2 Tensors

The tensor product

Suppose V, W are real vector spaces of dimensions n, m respectively. From them, we can form a new vector space: the tensor product, denoted $V \otimes W$. The space $V \otimes W$ consists of formal sums of the form

$$\boldsymbol{v}_1 \otimes \boldsymbol{w}_1 + \ldots + \boldsymbol{v}_k \otimes \boldsymbol{w}_k,$$

with $\boldsymbol{v}_p \in V$, $\boldsymbol{w}_p \in W$ for $p = 1, \ldots, k$. The tensor product $\otimes$ is bilinear, obeying:

$$\begin{aligned}\boldsymbol{v} \otimes (\boldsymbol{w}_1 + \lambda \boldsymbol{w}_2) &= \boldsymbol{v} \otimes \boldsymbol{w}_1 + \lambda (\boldsymbol{v} \otimes \boldsymbol{w}_2), \\ (\boldsymbol{v}_1 + \lambda \boldsymbol{v}_2) \otimes \boldsymbol{w} &= \boldsymbol{v}_2 \otimes \boldsymbol{w} + \lambda (\boldsymbol{v}_2 \otimes \boldsymbol{w}_2).\end{aligned}$$

If $\{\boldsymbol{e}_i\}_{i=1,\ldots,n}$ and $\{\boldsymbol{f}_a\}_{a=1,\ldots,m}$ are bases for V, W respectively, then we have a natural basis:

$$\{\boldsymbol{e}_i \otimes \boldsymbol{f}_a\}_{i=1,\ldots,n;\ a=1,\ldots,m}$$

for $V \otimes W$.

The space $T^p{}_q(V)$

From a real $n-$dimensional vector space V, we can naturally form a $(p+q)n-$dimensional vector space $T^p{}_q(V)$ by taking the tensor product of p copies of V and q copies of V^*:

$$T^p{}_q(V) = \underbrace{V \otimes \cdots \otimes V}_{p \text{ copies}} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{q \text{ copies}}.$$

An element of $T^p{}_q(V)$ is called a $(p, q)-$tensor, or a tensor of rank (p, q). A basis $B := \{\boldsymbol{e}_n\}_{i=1,\ldots n}$ induces a basis on $T^p{}_q(V)$:

$$B^p{}_q = \left\{\boldsymbol{e}_{i_1} \otimes \cdots \otimes \boldsymbol{e}_{i_p} \otimes \boldsymbol{e}^{j_1} \otimes \cdots \otimes \boldsymbol{e}^{j_q}\right\}_{i_k, j_l = 1,\ldots,n}.$$

With respect to this basis, we can write any $(p, q)-$tensor $\boldsymbol{T} \in T^p{}_q(V)$ in terms of its components:

$$\boldsymbol{T} = \sum_{i_1,\ldots,i_p,j_1,\ldots,j_q=1}^{n} T^{i_1 \ldots i_p}{}_{j_1 \ldots j_q} \boldsymbol{e}_{i_1} \otimes \cdots \otimes \boldsymbol{e}_{i_p} \otimes \boldsymbol{e}^{j_1} \otimes \cdots \otimes \boldsymbol{e}^{j_q}.$$

Often the distinction between a tensor and its components is elided, so one will speak of ‘the tensor T_{ij}’.

A.1.3 Change of basis

Suppose we have a basis $B := \{\boldsymbol{e}_i\}_{i=1,\ldots,n}$ and we wish to instead work with a new basis, say $B' := \{\boldsymbol{e}'_i\}_{i=1,\ldots,n}$. By virtue of the fact that B' is a basis, we must be able to write

$$\boldsymbol{e}_i = \sum_{j=1}^{n} \boldsymbol{e}'_j \Lambda^j{}_i, \tag{A.1}$$

for some real numbers $\Lambda^j{}_i$ with $i, j = 1, \ldots n$. Equally, we can write

$$\boldsymbol{e}'_i = \sum_{k=1}^{n} \boldsymbol{e}_k \widetilde{\Lambda}^k{}_i, \tag{A.2}$$

for some other real numbers $\widetilde{\Lambda}^j{}_i$ with $i, j = 1, \ldots n$. Substituting (A.1) into (A.2) we find

$$\boldsymbol{e}_i = \sum_{j=1}^{n} \left(\sum_{k=1}^{n} \boldsymbol{e}_k \widetilde{\Lambda}^k{}_j \right) \Lambda^j{}_i = \sum_{k,j=1}^{n} \boldsymbol{e}_k \widetilde{\Lambda}^k{}_j \Lambda^j{}_i$$

Since $\boldsymbol{e}_i$ are linearly independent, we conclude that

$$\sum_{j=1}^{n} \widetilde{\Lambda}^k{}_j \Lambda^j{}_i = \delta^k{}_i.$$

A similar calculation inserting (A.2) into (A.1) shows that

$$\sum_{j=1}^{n} \Lambda^k{}_j \widetilde{\Lambda}^j{}_i = \delta^k{}_i,$$

so that thinking of $\Lambda^i{}_j$ and $\widetilde{\Lambda}^i{}_j$ as the components of a matrix, we have $\Lambda^{-1} = \widetilde{\Lambda}$. In this way, we can identify the set of basis transformations with the group $GL(n, \mathbb{R})$ of invertible linear transformations (or equivalently matrices) on $\mathbb{R}^n$.

Suppose that $\boldsymbol{v} \in V$ has components v^i with respect to the basis B, and components v'^i with respect to the basis B'. We can relate these two sets of components by

$$\boldsymbol{v} = \sum_i \boldsymbol{e}_i v^i = \sum_{i,j=1}^{n} \boldsymbol{e}'_j \Lambda^j{}_i v^i = \sum_{i=1}^{n} \boldsymbol{e}'_j v'^j.$$

Using the linear independence of B', we deduce

$$v'^j = \sum_{i=1}^{n} \Lambda^j{}_i v^i, \tag{A.3}$$

so that the components of v change by matrix multiplication on the left by Λ (thinking of v^i as a column vector).

Now let's look at how the dual bases transform. Associated to B and B' are the dual bases $B^* = \{\boldsymbol{e}^i\}_{i=1,\dots n}$ and $B'^* = \{\boldsymbol{e}'^i\}_{i=1,\dots n}$ defined by

$$\boldsymbol{e}^i(\boldsymbol{e}_j) = \delta^i{}_j, \qquad \boldsymbol{e}'^i(\boldsymbol{e}'_j) = \delta^i{}_j, \qquad i,j = 1,\dots,n.$$

I claim that the dual bases are related by the relation:

$$\boldsymbol{e}'^i = \sum_{j=1}^n \Lambda^i{}_j \boldsymbol{e}^j.$$

To show this, it is enough to check that $\boldsymbol{e}'^i(\boldsymbol{e}'_j) = \delta^i{}_j$. We have:

$$\begin{aligned} \boldsymbol{e}'^i(\boldsymbol{e}'_j) &= \sum_{j=1}^n \Lambda^i{}_k \boldsymbol{e}^k \left(\sum_{k=1}^n \boldsymbol{e}_l \widetilde{\Lambda}^l{}_j \right), \\ &= \sum_{k,l=1}^n \Lambda^i{}_k \widetilde{\Lambda}^l{}_j \boldsymbol{e}^k(\boldsymbol{e}_l), \\ &= \sum_{k=1}^n \Lambda^i{}_k \widetilde{\Lambda}^k{}_j = \delta^i{}_j. \end{aligned}$$

We can invert the relationship between the bases using the fact that $\widetilde{\Lambda}$ is the matrix inverse of Λ, to find:

$$\boldsymbol{e}^i = \sum_{j=1}^n \widetilde{\Lambda}^i{}_j \boldsymbol{e}'^j.$$

Suppose that $\boldsymbol{\omega} \in V^*$ has components ω_i with respect to the basis B^*, and components ω'_i with respect to the basis B'^*. We can relate these two sets of components by

$$\boldsymbol{\omega} = \sum_i \omega_i \boldsymbol{e}^i = \sum_{i,j=1}^n \omega_i \widetilde{\Lambda}^i{}_j \boldsymbol{e}'^j = \sum_{i=1}^n \omega'_j \boldsymbol{e}'^j$$

Using the linear independence of B'^*, we deduce

$$\omega'_j = \sum_{i=1}^n \omega_i \widetilde{\Lambda}^i{}_j, \tag{A.4}$$

so that the components of v change by matrix multiplication on the right by $\widetilde{\Lambda}$ (thinking of ω_i as a row vector).

Extending these arguments to the space $T^p{}_q(V)$, we deduce

Lemma A.1. *Under the change of basis (A.1), the components of a tensor $\boldsymbol{T} \in T^p{}_q(V)$ transform according to:*

$$T'^{i_1\dots i_p}{}_{j_1\dots j_q} = \sum_{k_1,\dots,k_p,l_1,\dots,l_q=1}^n T^{k_1\dots k_p}{}_{l_1\dots l_q} \Lambda^{i_1}{}_{k_1} \cdots \Lambda^{i_p}{}_{k_p} \widetilde{\Lambda}^{l_1}{}_{j_1} \cdots \widetilde{\Lambda}^{l_q}{}_{j_q}.$$

We say that an upstairs index is covariant and a downstairs index is contravariant, reflecting the different transformation laws for the two types of index under a change of basis B for V. For each space $T^p{}_q(V)$, the transformation law gives a *representation* of the group $GL(n, \mathbb{R})$.

Exercise A.1. a) Suppose that a tensor $\boldsymbol{T} \in T^1{}_1(V)$ has components $T^i{}_j = \lambda \delta^i{}_j$ with respect to a given basis B. Show that $\boldsymbol{T}$ has the same components with respect to any other basis. Deduce that the Kronecker delta is invariant under a change of coordinates.

b) Suppose that $\boldsymbol{v} \in V$, with components v^i and $\omega \in V^*$ with components ω_j. Show that the numbers

$$T^i{}_j := v^i \omega_j,$$

transform as the components of a $(1,1)-$tensor. In this way we can build tensors of higher rank from lower rank tensors.

c) Suppose that $\boldsymbol{T} \in T^0{}_2(V)$ has components T_{ij}. Show that the numbers

$$\overline{T}_{ij} := T_{ji},$$

transform as a $(0,2)-$tensor. Deduce that

$$T_{(ij)} := \frac{1}{2}\left(T_{ij} + T_{ji}\right), \quad \text{and} \quad T_{[ij]} := \frac{1}{2}\left(T_{ij} - T_{ji}\right),$$

transform as the components of $(0,2)-$tensors. We call the tensors with components $T_{(ij)}$ and $T_{[ij]}$ the symmetric and antisymmetric part of $\boldsymbol{T}$ respectively.

Contracting indices

Let's suppose we have a tensor $\boldsymbol{T} \in T^1{}_2(V)$, so that its components can be written $T^i{}_{jk}$. Let us consider the following set of numbers:

$$S_k := \sum_{j=1}^{n} T^j{}_{jk}.$$

Now, let us see how S_k transforms when we change our basis. We have

$$T'^i{}_{jk} = \sum_{p,q,r=1,\ldots,n} T^p{}_{qr} \Lambda^i{}_p \widetilde{\Lambda}^q{}_j \widetilde{\Lambda}^r{}_k,$$

so that

$$\begin{aligned} S'_k := \sum_{j=1}^{n} T'^j{}_{jk} &= \sum_{j,p,q,r=1}^{n} T^p{}_{qr} \Lambda^j{}_p \widetilde{\Lambda}^q{}_j \widetilde{\Lambda}^r{}_k, \\ &= \sum_{p,q,r=1}^{n} T^p{}_{qr} \delta^q{}_p \widetilde{\Lambda}^r{}_k = \sum_{pr=1}^{n} T^p{}_{pr} \widetilde{\Lambda}^r{}_k, \\ &= \sum_{r=1}^{n} S_r \widetilde{\Lambda}^r{}_k. \end{aligned}$$

In other words, the set of numbers S_k transforms in the same way as the components of a covector. This means that the map

$$\sum_{i,j,k=1}^{n} T^i{}_{jk} \boldsymbol{e}_i \otimes \boldsymbol{e}^j \otimes \boldsymbol{e}^k \mapsto \sum_{i,j=1}^{n} T^i{}_{ij} \boldsymbol{e}^j,$$

is defined *independently of the basis* with respect to which we express the tensors. We can calculate the sum over components in any basis that we like, and we will still have the same covector as a result. This map is known as a *contraction over indices.* More generally we have

Lemma A.2. *Suppose* $\boldsymbol{T} \in T^p{}_q(V)$ *has components* $T^{i_1 \cdots i_p}{}_{j_1 \cdots j_q}$ *with respect to some basis* B *for* V*. The numbers*[2]

$$T^{i_1 \ldots, \hat{i}_r, \ldots, i_p}{}_{j_1 \ldots, \hat{j}_s, \ldots, j_q} = \sum_{i_r, j_s = 1}^{n} T^{i_1 \ldots i_p}{}_{j_1 \ldots, j_q} \delta^{i_r}{}_{j_s},$$

transform in the same fashion as the components of a $(p-1, q-1)-$*tensor. As a result, this defines a natural map* $T^p{}_q(V) \to T^{p-1}{}_{q-1}(V)$ *which does not depend on the choice of basis. This map is called 'contraction of the* r*'th covariant and* s*'th contravariant indices'.*

Summation convention

If we examine the various formulae that appear in this section, we will notice certain patterns. In particular:

- Whenever an index appears exactly once on the left hand side of an equation, the same index appears exactly once on the right hand side, and this index is *not* summed over. These indices are called 'free indices'.

- Whenever an index appears exactly twice, it occurs once in the upstairs position and once in the downstairs position and is summed over from 1 to n. These indices are called 'dummy indices'.

- No index appears more than twice.

These features are the basis for a very powerful notation known as Einstein's summation convention. The basic idea is very simple: we follow the rules as stated above, and we simply leave out any summation signs. Since the summations only appear when we see a repeated index pair with one up and one down, we can always put them back in if we want to. The great power of this convention is that so long as we stick to the rules, the objects we form will always have predictable transformation rules under a change of

[2] $\hat{i}_r$ means omit this index

basis[3]. For example, if $\boldsymbol{T} \in T^3{}_1(V)$ and $\boldsymbol{S} \in T^0{}_2(V)$, then

$$P_l = T^{jk}{}_l S_{jk} = \sum_{j,k=1}^{n} T^{jk}{}_l S_{jk},$$

transform as the components of a covector.

A.1.4 Metric tensors

When looking at vectors on $\mathbb{R}^n$, a useful concept is that of the inner product (sometimes called the dot product). If we take B to be the canonical basis for $\mathbb{R}^n$, then the inner product of two vectors $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^n$ is

$$\boldsymbol{v} \cdot \boldsymbol{w} = \sum_{i=1}^{n} v^i w^i.$$

It would be nice to write this using summation convention, however we have the problem that both indices are upstairs, so we can't simply drop the sum. To fix this, we introduce a new $(0,2)-$tensor, $\boldsymbol{g}$, called the *metric tensor*. We define $\boldsymbol{g}$ to have components with respect to the canonical basis:

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

Notice that under a change of basis g_{ij} will change in general. This is in contrast to $\delta^i{}_j$ which is the same with respect to any basis. They are different tensors, even though in this basis it looks like they have the same components: the position of indices matters! Having defined $\boldsymbol{g}$, we can write the inner product of two vectors as

$$\boldsymbol{v} \cdot \boldsymbol{w} = g_{ij} v^i w^j.$$

More generally, we will say that a *metric tensor* is a symmetric, non-degenerate, $(0,2)-$tensor. A $(0,2)-$tensor is symmetric if $g_{ij} = g_{ji}$ and non-degenerate if

$$g_{ij} v^i w^j = 0 \text{ for all } w^j \implies v^i = 0.$$

This condition is equivalent to requiring the matrix with components g_{ij} to be invertible. We denote the components of the inverse of g_{ij} by g^{ij}. These satisfy

$$g^{ij} g_{jk} = g_{kj} g^{ji} = \delta^i{}_k. \tag{A.5}$$

Exercise A.2. Verify that if the numbers g^{ij} are the unique solution of (A.5) then they indeed transform as the components of a $(2,0)-$tensor. This tensor is called the cometric.

[3] In this regard, summation convention is similar to Newspeak:

> "In the end we shall make thoughtcrime literally impossible, because there will be no words in which to express it."
>
> *1984, George Orwell*

The metric tensor allows us to identify V and V^* as follows. An element $\boldsymbol{v} \in V$ with components v^i is identified with the element $\boldsymbol{v}^\flat$ with components

$$v_i := g_{ij} v^j.$$

Similarly, an element $\boldsymbol{\omega} \in V^*$ with components ω_i is identified with the element $\boldsymbol{\omega}^\sharp \in V$ with components

$$\omega^i = g^{ij} \omega_j.$$

The notation $\sharp$, $\flat$ is inspired by musical notation and is designed to recall 'raising' and 'lowering' an index. The identification between V and V^* induced by $\boldsymbol{g}$ is occasionally called, somewhat whimsically, the *musical isometry.* This obviously extends to higher rank tensors, and allows us to identify elements of ${T^p}_q(V)$ with elements of ${T^{p'}}_{q'}(V)$ so long as $p + q = p' + q'$. By convention, when writing the components of two tensors identified via $\boldsymbol{g}$, we use the same core letter, relying on the raised and lowered indices to indicate which space the tensor belongs to. When we have a metric tensor, we will often speak of a rank k–tensor, without specifying the index structure.

A.1.5 The orthogonal groups

Given a metric tensor, $\boldsymbol{g}$, we naturally have a bilinear form on V, defined by

$$\boldsymbol{g}(\boldsymbol{v}, \boldsymbol{w}) = g_{ij} v^i w^i = v^i w_i = v_i w^i.$$

By a Gram-Schmidt type of process, it is possible to construct a basis $B = \{\boldsymbol{e}_i\}_{i=1,\ldots,n}$ such that for some $r \in \{1, \ldots, n+1\}$ we have

$$\boldsymbol{g}(\boldsymbol{e}_i, \boldsymbol{e}_j) = g_{ij} = \begin{cases} -1 & i = j, \quad j < r, \\ 1 & i = j, \quad j \geq r, \\ 0 & i \neq j. \end{cases} \tag{A.6}$$

Exercise A.3. Adapt the Gram-Schmidt process to show that there always exists a basis such that (A.6) holds.

We call such a basis, B, an orthonormal basis. With respect to an orthonormal basis, the metric tensor $\boldsymbol{g}$ (thought of as a matrix) is diagonal, with entries ± 1. The number of positive and negative signs is fixed, independent of which orthonormal basis we choose, by Silvester's law of inertia. The number of positive and negative signs is known as the *signature* of the metric, and is usually written as $(+,+,+)$ or $(-,-,+,+)$, or alternatively as (r, s), where r is the number of negative signs and s the number of positive. There are two cases that are most often studied. If all entries are positive, we say the metric has Riemannian signature, and in this case it defines a positive-definite inner product on V. If we have one negative entry and the rest positive, or one positive and the rest negative, we say the metric has Lorentzian signature. This is the case of interest for special and general relativity.

Exercise A.4. Consider $\mathbb{R}^3$ with its canonical basis. Find an orthonormal basis for the metric whose components are given by

$$g_{12} = g_{21} = g_{33} = 1, \qquad \text{all other components } 0.$$

and write down the signature of the metric.

Having established that every metric admits at least one orthonormal basis, a natural question arises concerning other orthonormal bases. Suppose we have a basis B with respect to which $\boldsymbol{g}$ has components given by (A.6). Let B' be a new basis which is also orthonormal, and suppose $\Lambda^i{}_j$ is the matrix corresponding to this change of basis. Then by the transformation law for components of $T^0{}_2(V)$, we have that

$$g'_{ij} = \widetilde{\Lambda}^k{}_i \widetilde{\Lambda}^l{}_j g_{kl},$$

so that if $g'_{ij} = g_{ij}$, we deduce:

$$g_{ij} = \widetilde{\Lambda}^k{}_i \widetilde{\Lambda}^l{}_j g_{kl}. \tag{A.7}$$

Exercise(∗). Show that (A.7) holds if and only if:

$$g_{ij} = \Lambda^k{}_i \Lambda^l{}_j g_{kl}, \tag{A.8}$$

If $\Lambda^i{}_j$ satisfies either (A.7) or (A.8), we say that the corresponding transformation is *orthogonal*. Orthogonal transformations form a group, which can be identified with a subgroup of $GL(n, \mathbb{R})$. We denote this group by $O(r, s)$, where (r, s) is the signature of the metric tensor $\boldsymbol{g}$. In the case that $r = 0$, we have $O(0, n) \equiv O(n)$, the standard orthogonal group. To see this, we note that

$$\Lambda^k{}_i \Lambda^l{}_j \delta_{kl} = \left(\Lambda^T \Lambda\right)_{ij}.$$

The group at the heart of Special Relativity is $O(1, 3)$, which is sufficiently important that it has its own name, the Lorentz group.

A.2 Review of differentiation in $\mathbb{R}^n$

We will review some material about differentiation, and fix some notation that will (hopefully) be familiar if you attended the Manifolds course.

Suppose U is an open subset of $\mathbb{R}^n$ and suppose we have a function $f \in C^1(U)$. This implies that at each point in $\boldsymbol{x}$ there exists a linear map $df|_{\boldsymbol{x}} : \mathbb{R}^n \to \mathbb{R}$ such that if $\boldsymbol{V} \in \mathbb{R}^n$ is any vector, and s is sufficiently small, we have

$$f(\boldsymbol{x} + s\boldsymbol{V}) = f(\boldsymbol{x}) + s \; df|_{\boldsymbol{x}} \boldsymbol{V} + o(s).$$

The linear map $df|_{\boldsymbol{x}}$ is called the *differential* of f at $\boldsymbol{x}$. Geometrically, we should think of the vector $\boldsymbol{V}$ in this formula as having its base at $\boldsymbol{x}$. We call the space of such vectors $T_{\boldsymbol{x}}U$, which is isomorphic, as a vector space, with $\mathbb{R}^3$. Since $df|_{\boldsymbol{x}}$ is a linear map from $T_{\boldsymbol{x}}U$ to $\mathbb{R}$, it belongs to the dual space of $T_{\boldsymbol{x}}U$, which we denote $T^*_{\boldsymbol{x}}U$. Given any vector $\boldsymbol{V} \in T_{\boldsymbol{x}}U$, we can define the *directional derivative* of f along $\boldsymbol{V}$ at $\boldsymbol{x}$ to be:

$$\boldsymbol{V}[f](\boldsymbol{x}) := df|_{\boldsymbol{x}} \boldsymbol{V} = \lim_{s \to 0} \frac{f(\boldsymbol{x} + s\boldsymbol{V}) - f(\boldsymbol{x})}{s}.$$

A.2.1 Working in coordinates

Now, suppose $\{\boldsymbol{e}_i\}_{i=1,\ldots,n}$ is a basis for $\mathbb{R}^n$, and let $\{\boldsymbol{e}^j\}_{i=1,\ldots,n}$ be the dual basis for $(\mathbb{R}^n)^*$, defined by

$$\boldsymbol{e}^i(\boldsymbol{e}_j) = \delta^i{}_j.$$

We define the functions

$$\begin{array}{rcccl} x^i & : & \mathbb{R}^3 & \to & \mathbb{R} \\ & & \boldsymbol{x} & \mapsto & \boldsymbol{e}^i(\boldsymbol{x}) \end{array}$$

for $i = 1, \ldots, n$. It follows[4] that we can write:

$$\boldsymbol{x} = x^i \boldsymbol{e}_i. \tag{A.9}$$

For any vector $\boldsymbol{V} \in T^*_{\boldsymbol{x}}U$, we have

$$dx^i\big|_{\boldsymbol{x}}(\boldsymbol{V}) = \boldsymbol{V}[x^i](\boldsymbol{x}) = \lim_{s\to 0} \frac{\boldsymbol{e}^i(\boldsymbol{x}+s\boldsymbol{V}) - \boldsymbol{e}^i(\boldsymbol{x})}{s} = \boldsymbol{e}^i(\boldsymbol{V}).$$

In other words, we have $dx^i = \boldsymbol{e}^i$.

Now consider a function f which is continuously differentiable in a neighbourhood of $\boldsymbol{x} \in U$, and let us look at $\boldsymbol{V}[f](\boldsymbol{x})$, the directional derivative of f along $\boldsymbol{V}$. We can write $\boldsymbol{V} = V^i \boldsymbol{e}_i$, so that

$$\boldsymbol{V}[f](\boldsymbol{x}) = df|_{\boldsymbol{x}}(V^i \boldsymbol{e}_i) = V^i\, df|_{\boldsymbol{x}}(\boldsymbol{e}_i) = V^i \boldsymbol{e}_i[f](\boldsymbol{x}). \tag{A.10}$$

In other words, to calculate the directional derivative along an arbitrary direction, it is enough to know the three directional derivatives $\boldsymbol{e}_i[f](\boldsymbol{x})$. These directional derivatives are useful enough that we give them a special symbol. We write

$$\frac{\partial}{\partial x^i} f(\boldsymbol{x}) := \boldsymbol{e}_i[f](\boldsymbol{x}) = \frac{d}{ds} f(\boldsymbol{x} + s\boldsymbol{e}_i)\bigg|_{s=0}.$$

This of course agrees with our usual notion of partial derivative. We can go further though, and declare that the object $\frac{\partial}{\partial x^i}$ is itself a vector, which we can identify with $\boldsymbol{e}_i$ i.e. we have:

$$\frac{\partial}{\partial x^i} = \boldsymbol{e}_i \in T_{\boldsymbol{x}}U.$$

We will often find it useful to write as shorthand:

$$\frac{\partial}{\partial x^i} := \partial_i$$

Of course, we have:

$$dx^i(\partial_j) = \delta^i{}_j.$$

Returning to (A.10), we write

$$\boldsymbol{V}[f](\boldsymbol{x}) = V^i \boldsymbol{e}_i[f](\boldsymbol{x}) = \boldsymbol{e}^i(\boldsymbol{V})\boldsymbol{e}_i[f](\boldsymbol{x}) = \frac{\partial f}{\partial x^i}\, dx^i\bigg|_{\boldsymbol{x}}(\boldsymbol{V})$$

[4] Check this. You may need to refer to §A.1

so that we can write

$$df|_{\boldsymbol{x}} = \left.\frac{\partial f}{\partial x^i} dx^i \right|_{\boldsymbol{x}}.$$

Because we have respected the Einstein summation convention, the right hand side of this formula is independent of the choice of basis that we originally made (as, of course it must be).

Exercise A.5. a) Suppose that we consider a new basis $\{\boldsymbol{e}_i'\}$ for $\mathbb{R}^3$. Show that the set of numbers

$$\frac{\partial f}{\partial x^i},$$

transform in the same way as the components of a *co-vector*.

b) Fix $\boldsymbol{x} \in U$ and suppose that $\boldsymbol{V}, \boldsymbol{W} \in T_{\boldsymbol{x}}U$ are two vectors such that for any $f \in C^1(U)$ we have

$$\boldsymbol{V}[f](\boldsymbol{x}) = \boldsymbol{W}[f](\boldsymbol{x}).$$

Show that $\boldsymbol{V} = \boldsymbol{W}$.

A.3 Differential geometry

We will briefly review the basic definitions of a manifold, the tangent bundle and higher rank tensor bundles.

A.3.1 Manifolds

Definition 21. A C^k-*atlas* of a second countable, Hausdorff, topological space $\mathcal{M}$ is a collection of *charts* $\{(\mathcal{U}_\alpha, \phi_\alpha)\}$ which satisfy:

i) Each $\mathcal{U}_\alpha \subset \mathcal{M}$ is an open subset of $\mathcal{M}$, and the $\mathcal{U}_\alpha$ cover $\mathcal{M}$.

ii) φ_α is a homeomorphism from $\mathcal{U}_\alpha$ onto an open subset of $\mathbb{R}^n$.

iii) If $\mathcal{U}_\alpha \cap \mathcal{U}_\beta \neq \emptyset$, then

$$\varphi_{\alpha\beta} := \varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha (\mathcal{U}_\alpha \cap \mathcal{U}_\beta) \to \varphi_\beta (\mathcal{U}_\alpha \cap \mathcal{U}_\beta)$$

is a C^k-diffeomorphism between subsets of $\mathbb{R}^n$, that is to say that $\varphi_{\alpha\beta}^{-1}$ exists and both $\varphi_{\alpha\beta}$ and $\varphi_{\alpha\beta}^{-1}$ are C^k-functions on their respective domains. The functions $\varphi_{\alpha\beta}$ are called *transition functions.*

We say that two C^k-atlases $\{(\mathcal{U}_\alpha, \varphi_\alpha)\}$ and $\{(\mathcal{V}_\alpha, \psi_\alpha)\}$ for $\mathcal{M}$ are compatible if their union is again a C^k-atlas. Clearly compatibility defines an equivalence relation.

Definition 22. A C^k-manifold is a second countable, Hausdorff, topological space $\mathcal{M}$, equipped with an equivalence class of C^k-atlases. The dimension of the manifold is n, where we understand the charts as mapping into $\mathbb{R}^n$.

We can define in an exactly analogous way, a smooth (C^∞) or real analytic[5] (C^ω) atlas (and hence manifold) by requiring that the transition functions be smooth or real analytic diffeomorphisms between open subsets of $\mathbb{R}^n$. We will always assume that $k \geq 1$, which means that objects such as the tangent space are well defined.

Obviously if we have a C^k-manifold, we can always extract from it a $C^{k'}$-manifold for $k' < k$ by picking a representative C^k-atlas and considering it's equivalence class among $C^{k'}$-atlases. There is a pitfall for the unwary in that these two manifolds, although based on the same topological space, are *not* the same. It is, however, common practice to leave the regularity of the manifold unstated in many circumstances so one should be careful.

Example 14. Let $U \subset \mathbb{R}^n$ be an open set with the subset topology. This is certainly a second countable, Hausdorff, topological space. We can equip U with a trivial C^ω-atlas, given by $\{(U, id)\}$. Thus U, with the equivalence class of atlases defined by the trivial atlas, is a real analytic manifold.

Example 15. We define S^1 to be the second countable, Hausdorff[6], topological space $\mathbb{R}/_{\sim_\circ}$, where

$$x \sim_\circ y \quad \iff \exists\, n \in \mathbb{Z} \text{ s.t. } \quad x = y + 2\pi n.$$

We can define an atlas as follows. Let $0 \leq \alpha < 2\pi$. We set $\mathcal{U}_\alpha = S^1 \setminus [\alpha]_{\sim_\circ}$, and define

$$\begin{array}{rccl} \varphi_\alpha & : & \mathcal{U}_\alpha & \to & (0, 2\pi) \\ & & [x]_{\sim_\circ} & \mapsto & \theta - \alpha \end{array}$$

where θ is the unique real number satisfying $\theta \sim_\circ x$ and $\alpha < \theta < \alpha + 2\pi$. Suppose $0 \leq \alpha < \alpha' < 2\pi$. We have that

$$\begin{aligned} \varphi_\alpha \left(\mathcal{U}_\alpha \cap \mathcal{U}_\beta\right) &= (0, 2\pi) \setminus \{\alpha' - \alpha\} \\ \varphi_\beta \left(\mathcal{U}_\alpha \cap \mathcal{U}_\beta\right) &= (0, 2\pi) \setminus \{2\pi + \alpha - \alpha'\} \end{aligned}$$

and

$$\varphi_{\alpha\beta}(s) = \begin{cases} s + 2\pi + \alpha - \alpha' & 0 < s < \alpha' - \alpha, \\ s - \alpha' + \alpha & \alpha' - \alpha < s < 2\pi. \end{cases}$$

This can be easily verified to be a real analytic diffeomorphism, so that $\mathscr{A} = \{\mathcal{U}_\alpha, \varphi_\alpha\}_{\alpha \in [0, 2\pi)}$ is a real analytic atlas for S^1. Taking $\mathscr{A}$ to define an equivalence class of C^ω-atlases, we can make S^1 into a real analytic manifold.

A.3.2 Mappings between manifolds, and their derivatives

Smooth functions

Suppose that we have two manifolds $\mathcal{M}, \mathcal{N}$, which are both at least C^k-regular[7] and are equipped with representative atlases $\{(U_\alpha, \varphi_\alpha)\}, \{(\mathcal{V}_\beta, \psi_\beta)\}$. Consider a function:

$$f : \mathcal{M} \to \mathcal{N}.$$

[5] A function f is real analytic if there is a neighbourhood of every point on which f can be expressed as a convergent Taylor series.

[6] you should check that this is true

[7] From now on, we'll abuse notation and allow '$k = \infty$' and '$k = \omega$' in statements like this.

We say that f is C^k-smooth if for any α, β, the function

$$\psi_\beta \circ f \circ \varphi_\alpha^{-1} : \mathcal{U}_\alpha \to \mathcal{V}_\beta$$

is C^k, understood in the usual way for maps between subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$. We denote by $C^k(\mathcal{M};\mathcal{N})$ the set of all C^k-smooth maps from $\mathcal{M}$ to $\mathcal{N}$.

Exercise(∗). Show that this definition is independent of the choice of atlases. That is, if we choose a different atlas $\{(U'_\alpha, \varphi'_\alpha)\}$ for $\mathcal{M}$ which is compatible with $\{(U_\alpha, \varphi_\alpha)\}$, our definition of $C^k(\mathcal{M};\mathcal{N})$ agrees for both.

Exercise A.6. a) Show that the identity map on a C^k-manifold, $\mathcal{M}$, always belongs to $C^k(\mathcal{M};\mathcal{M})$.

b) Show that if $f, g \in C^k(\mathcal{M};\mathbb{R})$, we have $fg \in C^k(\mathcal{M};\mathbb{R})$, where we define the product pointwise $fg(p) = f(p)g(p)$.

c) Suppose $\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3$ are all at least C^k regular, and that $f \in C^k(\mathcal{M}_1;\mathcal{M}_2)$, $g \in C^k(\mathcal{M}_2;\mathcal{M}_3)$. Show that $g \circ f \in C^k(\mathcal{M}_1;\mathcal{M}_3)$.

d) Let $\mathcal{M}$ be a C^k-manifold of dimension n and let $\pi^i : \mathbb{R}^n \to \mathbb{R}$ be the projection onto the i^{th} coordinate. Suppose that $(\mathcal{U}_\alpha, \varphi_\alpha)$ is a chart. Show that

$$\varphi^i_\alpha = \pi^i \circ \varphi_\alpha : \mathcal{U}_\alpha \to \mathbb{R}$$

is C^k-smooth.

Example 16. Take S^1 and $\mathbb{R}^2$ with their real analytic structures as previously defined. We define a map

$$\begin{array}{rcccl} f & : & S^1 & \to & \mathbb{R}^2, \\ & & [x]_{\sim_\circ} & \mapsto & (\sin x, \cos x). \end{array}$$

First, note that this is well defined regardless of which representative x for $[x]_{\sim_\circ}$ we choose. Now consider the functions

$$\begin{array}{rcccl} id \circ f \circ \varphi_\alpha^{-1} & : & (0, 2\pi) & \to & \mathbb{R}^2, \\ & & \theta & \mapsto & (\sin(\theta + \alpha), \cos(\theta + \alpha)). \end{array}$$

which are clearly real analytic. Thus, $f \in C^\omega(S^1;\mathbb{R}^2)$.

The tangent and co-tangent space at a point

Now that we have defined our C^k-smooth functions, we can define the tangent vectors to the manifold $\mathcal{M}$ at a point $p \in M$. There are various ways of doing this, some more concrete than others. We'll give the definition here, and then show how this definition fits with our intuition from the case of $\mathbb{R}^n$.

We say that a C^k-smooth map $\gamma : (-\epsilon, \epsilon) \to \mathcal{M}$ is a curve in $\mathcal{M}$. Fix $p \in \mathcal{M}$. We define the tangent space at p, $T_p\mathcal{M}$ to be the space of curves in $\mathcal{M}$ with $\gamma(0) = p$, modulo the equivalence relation:

$$\gamma_1 \sim \gamma_2 \quad \text{if} \quad \left.\frac{d}{dt}\left[f \circ \gamma_1\right](t)\right|_{t=0} = \left.\frac{d}{dt}\left[f \circ \gamma_2\right](t)\right|_{t=0} \qquad \text{for all } f \in C^k(\mathcal{M};\mathbb{R}).$$

Each element of $T_p\mathcal{M}$ (i.e. each equivalence class of curves under $\sim$) is called a tangent vector at p, and we write $[\gamma]_\sim = \dot{\gamma}(0)$. For any tangent vector $V \in T_p\mathcal{M}$, and function $f \in C^k(\mathcal{M};\mathbb{R})$, we define the directional derivative of f along V at p to be:

$$V[f](p) := \frac{d}{dt}\left[f \circ \gamma\right](t)\Big|_{t=0}$$

for any γ such that $\dot{\gamma}(0) = V$. This is a useful way to think of a tangent vector at p. It is a direction in which we can differentiate a function.

We can endow $T_p\mathcal{M}$ with a vector space structure using the following result:

Lemma A.3. *Let $(\mathcal{U}_\alpha, \varphi_\alpha)$ be a chart with $p \in \mathcal{U}_\alpha$. The chart induces a canonical identification between $T_p\mathcal{M}$ and $\mathbb{R}^n$.*

Proof. Let us set $\varphi_\alpha(p) = \boldsymbol{x}$. Suppose we are given two curves γ_1, γ_2 with $\gamma_i(0) = p$. We can use the fact that $\varphi_\alpha^{-1} \circ \varphi_\alpha = id_{\mathcal{U}_\alpha}$ to see that

$$\frac{d}{dt}\left[f \circ \gamma_1\right](t)\Big|_{t=0} = \frac{d}{dt}\left[f \circ \gamma_2\right](t)\Big|_{t=0} \qquad \text{for all } f \in C^k(\mathcal{M};\mathbb{R})$$

holds if and only if

$$\frac{d}{dt}\left[f \circ \varphi_\alpha^{-1} \circ \varphi_\alpha \circ \gamma_1\right](t)\Big|_{t=0} = \frac{d}{dt}\left[f \circ \varphi_\alpha^{-1} \circ \varphi_\alpha \circ \gamma_2\right](t)\Big|_{t=0} \qquad \text{for all } f \in C^k(\mathcal{M};\mathbb{R})$$

which is true if and only if

$$\frac{d}{dt}\left[\widetilde{f} \circ \widetilde{\gamma}_1\right](t)\Big|_{t=0} = \frac{d}{dt}\left[\widetilde{f} \circ \widetilde{\gamma}_2\right](t)\Big|_{t=0} \qquad \text{for all } \widetilde{f} \in C^k\left(\varphi_\alpha(\mathcal{U}_\alpha);\mathbb{R}\right).$$

Where for $i = 1, 2$ we have defined

$$\begin{array}{rcrcl} \widetilde{\gamma}_i & : & (-\epsilon, \epsilon) & \to & \varphi_\alpha(\mathcal{U}_\alpha), \\ & & t & \mapsto & (\varphi_\alpha \circ \gamma_i)(t) \end{array}$$

which is a curve in a subset of $\mathbb{R}^n$. As a result, we can simply apply the chain rule to deduce that $\gamma_1 \sim \gamma_2$ if and only if we have equality of the directional derivatives:

$$\boldsymbol{V}_1[f](\boldsymbol{x}) = \boldsymbol{V}_2[f](\boldsymbol{x}), \qquad \text{for all } f \in C^k(\varphi_\alpha(\mathcal{U}_\alpha);\mathbb{R}),$$

where $\boldsymbol{V}_i = \dot{\widetilde{\gamma}}(0)$ are the tangent vectors to the curves $\widetilde{\gamma}_i$, in the usual sense of curves in $\mathbb{R}^n$. By Exercise A.5 this holds if and only if $\boldsymbol{V}_1 = \boldsymbol{V}_2$. Thus, with each equivalence class $[\gamma]_\sim$ we can associate a unique vector in $\mathbb{R}^n$ via the coordinate chart $(\mathcal{U}_\alpha, \varphi_\alpha)$. Conversely, from $\boldsymbol{V} \in \mathbb{R}^n$, we can construct the curve

$$\gamma_{\boldsymbol{V}}(t) = \varphi_\alpha^{-1}\left(\boldsymbol{x} + \boldsymbol{V}t\right)$$

Which satisfies $\dot{\widetilde{\gamma}}_{\boldsymbol{V}}(0) = \boldsymbol{V}$. Thus, the chart $(\mathcal{U}_\alpha, \varphi_\alpha)$ induces a canonical identification of $T_p\mathcal{M}$ with $\mathbb{R}^n$. □

We deduce from this Lemma the important result:

Theorem A.1. *Given $\gamma_i : (-\epsilon, \epsilon) \to \mathcal{M}$ with $\gamma_i(0) = p$ for $i = 1, 2$ and $\lambda \in \mathbb{R}$, there exists a curve γ such that*

$$\frac{d}{dt}\left[f \circ \gamma\right](t)\Big|_{t=0} = \frac{d}{dt}\left[f \circ \gamma_1\right](t)\Big|_{t=0} + \lambda \frac{d}{dt}\left[f \circ \gamma_2\right](t)\Big|_{t=0}$$

for all $f \in C^k(\mathcal{M}; \mathbb{R})$. Clearly, this defines $[\gamma]_\sim = \dot{\gamma}(0)$ uniquely. We then write

$$\dot{\gamma}_1(0) + \lambda\dot{\gamma}_2(0) := \dot{\gamma}(0)$$

This provides us with a definition of addition of tangent vectors and scalar multiplication. With these definitions, T_pM is a real vector space of dimension n.

Proof. We may simply take (with the notation of the Lemma)

$$\gamma(t) = \gamma_{\boldsymbol{V}}(t)$$

where

$$\boldsymbol{V} = \dot{\tilde{\gamma}}_1(0) + \lambda\dot{\tilde{\gamma}}_2(0).$$

This gives us a vector space isomorphism between $T_p\mathcal{M}$ and $\mathbb{R}^n$. □

We can use this canonical identification induced by $(\mathcal{U}_\alpha, \varphi_\alpha)$ to provide us with a basis for $T_p\mathcal{M}$ for each $p \in \mathcal{U}_\alpha$. We pick a basis $\{\boldsymbol{e}_i\}_{i=1,\dots,n}$ for $\mathbb{R}^n$ and define

$$\left(\frac{\partial}{\partial x^i}\right)_p = (\partial_i) := \dot{\gamma}_{\boldsymbol{e_i}}(0).$$

We say that $\left\{\left(\frac{\partial}{\partial x^i}\right)_p\right\}$ is a coordinate basis for $T_p\mathcal{M}$.

The reason for this notation is made clear by the following result

Lemma A.4. *Suppose $f \in C^k(\mathcal{M}; \mathbb{R})$ and let $(\mathcal{U}_\alpha, \varphi_\alpha)$ be a chart with $p \in \mathcal{U}_\alpha$ and $\varphi_\alpha(p) = \boldsymbol{x}$. We define $\widetilde{f} = f \circ \varphi_\alpha^{-1}$, which is a C^k function defined on some open set about $\boldsymbol{x} = x^i\boldsymbol{e}_i$. We then have:*

$$\left(\frac{\partial}{\partial x^i}\right)_p [f](p) = \left(\frac{\partial \widetilde{f}}{\partial x^i}\right)(\boldsymbol{x})$$

where on the right hand side, we simply have the partial derivative of $\widetilde{f}$ as a function on $\mathbb{R}^n$.

Proof. We simply calculate using the definitions:

$$\begin{aligned}\left(\frac{\partial}{\partial x^i}\right)_p [f](p) &= \frac{d}{dt}\left[f \circ \gamma_{\boldsymbol{e_i}}\right]\Big|_{t=0} \\ &= \frac{d}{dt}\left[(f \circ \varphi_\alpha^{-1})(\boldsymbol{x} + \boldsymbol{e_i}t)\right]\Big|_{t=0} \\ &= \left(\frac{\partial \widetilde{f}}{\partial x^i}\right)(\boldsymbol{x}).\end{aligned}$$

□

Now that we have the vector space $T_p\mathcal{M}$, we are free to construct its dual, and the higher rank tensor spaces associated with it. Particularly important is the dual space $T_p^*\mathcal{M}$, known as the co-tangent space at p. The cotangent space is the natural place in which the differential of a function lives:

Definition 23. Given $f \in C^k(\mathcal{M};\mathbb{R})$, we define the differential of f at p, $df|_p$ to be the linear map

$$\begin{array}{rcrcl} df|_p & : & T_p\mathcal{M} & \to & \mathbb{R}, \\ & & V & \mapsto & V[f](p). \end{array}$$

We can define a basis for $T_p^*\mathcal{M}$ from the coordinate basis for $T_p\mathcal{M}$ by duality. We define $dx^i\big|_p \in T_p^*\mathcal{M}$ by the requirement that

$$dx^i\big|_p \left[\left(\frac{\partial}{\partial x^j} \right)_p \right] = \delta^i{}_j.$$

With this definition we have

$$df|_p = \left(\frac{\partial}{\partial x^i} \right)_p [f](p) \;\; dx^i\big|_p$$

The tangent and co-tangent bundles

In the previous subsection, we restricted our attention to vectors and co-vectors defined at a single point. More often than not, rather than considering a vector at a single point we want to consider a *vector field*, where we have a vector defined at each point. Before discussing vector fields, we introduce the *tangent bundle*. This is the disjoint union over the tangent spaces at each point in the manifold:

$$T\mathcal{M} = \bigsqcup_{p\in\mathcal{M}} T_p\mathcal{M} = \bigcup_{p\in\mathcal{M}} \{p\} \times T_p\mathcal{M}.$$

We shall show that a C^k-atlas for $\mathcal{M}$ induces a C^{k-1}-atlas for $T\mathcal{M}$, so that the tangent bundle may be given a manifold structure in a natural fashion. Suppose $\{(\mathcal{U}_\alpha, \varphi_\alpha)\}$ is a C^k-atlas for $\mathcal{M}$. Pick a chart $(\mathcal{U}_\alpha, \varphi_\alpha)$. We know that at each point $p \in \mathcal{U}_\alpha$ the chart induces a canonical identification between $T_p\mathcal{M}$ and $\mathbb{R}^n$. A tangent vector $V \in T_p\mathcal{M}$ is identified uniquely with a vector $\boldsymbol{V} \in \mathbb{R}^n$. We now define:

$$\hat{\mathcal{U}}_\alpha = \bigsqcup_{p\in\mathcal{U}_\alpha} T_p\mathcal{M}$$

and

$$\begin{array}{rcrcl} \hat{\varphi}_\alpha & : & \hat{\mathcal{U}}_\alpha & \to & \varphi_\alpha(\mathcal{U}_\alpha) \times \mathbb{R}^n, \\ & & (p, V) & \mapsto & (\varphi_\alpha(p), \boldsymbol{V}). \end{array}$$

Clearly $\hat{\varphi}_\alpha$ maps $\hat{\mathcal{U}}_\alpha$ bijectively onto an open subset of $\mathbb{R}^{2n}$. We define a topology on $T\mathcal{M}$ by declaring that $\mathcal{A} \subset T\mathcal{M}$ is open if and only if $\hat{\varphi}_\alpha\left(\mathcal{A} \cap \hat{\mathcal{U}}_\alpha\right)$ is open in $\mathbb{R}^{2n}$ for any α. One can verify that this definition turns $T\mathcal{M}$ into a second countable, Hausdorff, topological space.

Lemma A.5. *The collection $\{(\hat{\mathcal{U}}_\alpha, \hat{\varphi}_\alpha)\}$ is a C^{k-1}-atlas for $T\mathcal{M}$.*

Proof. By construction, $\hat{\varphi}_\alpha$ maps $\hat{\mathcal{U}}_\alpha$ homeomorphically onto an open subset of $\mathbb{R}^{2n}$ and moreover the sets $\hat{\mathcal{U}}_\alpha$ cover $T\mathcal{M}$. It remains to show that the transition functions are C^{k-1}-diffeomorphisms between subsets of $\mathbb{R}^{2n}$. I claim that we have

$$\begin{array}{rcrcl} \hat{\varphi}_{\alpha\beta} & : & \varphi_\alpha(\mathcal{U}_\alpha \cap \mathcal{U}_\alpha) \times \mathbb{R}^n & \to & \varphi_\beta(\mathcal{U}_\alpha \cap \mathcal{U}_\alpha) \times \mathbb{R}^n, \\ & & (\boldsymbol{x}, \boldsymbol{V}_\alpha) & \mapsto & (\varphi_{\alpha\beta}(\boldsymbol{x}), D\varphi_{\alpha\beta}|_{\boldsymbol{x}}(\boldsymbol{V}_\alpha)). \end{array}$$

To see this, recall that if $\varphi_\alpha(p) = \boldsymbol{x}$, then $\boldsymbol{V}_\alpha \in \mathbb{R}^n$ corresponds to $V \in T_p\mathcal{M}$ under the identification induced by φ_α implies that $V = \dot{\gamma}(0)$ for the curve

$$\begin{array}{rcrcl} \gamma & : & (-\epsilon, \epsilon) & \to & \mathcal{M}, \\ & & t & \mapsto & \varphi_\alpha^{-1}(\boldsymbol{x} + t\boldsymbol{V}_\alpha). \end{array}$$

Recall also that, going in the other direction, if $V \in T_p\mathcal{M}$, then under the identification induced by φ_β it corresponds to the vector $\boldsymbol{V}_\beta \in \mathbb{R}^n$ defined by

$$\boldsymbol{V}_\beta = \frac{d}{dt}(\varphi_\beta \circ \gamma)(t)\bigg|_{t=0}$$

for any $\gamma : (-\epsilon, \epsilon) \to \mathcal{M}$ such that $\gamma(0) = p$ and $\dot{\gamma}(0) = V$. Combining these two facts, we have that

$$\begin{aligned} \boldsymbol{V}_\beta &= \frac{d}{dt}(\varphi_\beta \circ \varphi_\alpha^{-1})(\boldsymbol{x} + t\boldsymbol{V}_\alpha)\bigg|_{t=0} \\ &= \frac{d}{dt}(\varphi_{\alpha\beta})(\boldsymbol{x} + t\boldsymbol{V}_\alpha)\bigg|_{t=0} \\ &= D\varphi_{\alpha\beta}|_{\boldsymbol{x}}(\boldsymbol{V}_\alpha). \end{aligned}$$

Now, since $\varphi_{\alpha\beta}$ is a C^k-diffeomorphism, $D\varphi_{\alpha\beta}$ is a C^{k-1} map into the space of invertible matrices, thus the transition function $\hat{\varphi}_{\alpha\beta}$ is a C^{k-1}-diffeomorphism. □

As a consequence, we have:

Theorem A.2. *The tangent bundle naturally inherits the structure of a C^{k-1}-manifold of dimension $2n$.*

An entirely analogous construction can be performed in which we glue together the co-tangent spaces to define the co-tangent bundle, $T^*\mathcal{M}$. This again inherits the structure of a C^{k-1}-manifold of dimension $2n$ from $\mathcal{M}$. The only difference in our development above is that the relevant transition functions are given by[8]:

$$\begin{array}{rcrcl} \hat{\varphi}^*_{\alpha\beta} & : & \varphi_\alpha(\mathcal{U}_\alpha \cap \mathcal{U}_\alpha) \times (\mathbb{R}^n)^* & \to & \varphi_\beta(\mathcal{U}_\alpha \cap \mathcal{U}_\alpha) \times (\mathbb{R}^n)^*, \\ & & (\boldsymbol{x}, \boldsymbol{\omega}_\alpha) & \mapsto & (\varphi_{\alpha\beta}(\boldsymbol{x}), D\varphi_{\alpha\beta}^{-1}\big|_{\boldsymbol{x}}^*(\boldsymbol{\omega}_\alpha)). \end{array}$$

[8] recall that the adjoint of a linear map $\Lambda : V \to V$ is a linear map $\Lambda^* : V^* \to V^*$ whose action on $\omega \in V^*$ is given by:

$$(\Lambda^*\omega)[v] = \omega[\Lambda v], \qquad \text{for all } v \in V.$$

Glueing together tensor products of $T_p\mathcal{M}$ and $T_p^*\mathcal{M}$ in the same manner, we can construct tensor bundles of arbitrary rank.

Definition 24. Suppose $\mathcal{M}$ is at least $k+1$ regular.

i) A C^k-vector field is a C^k-map $X : \mathcal{M} \to T\mathcal{M}$ such that at every point $p \in \mathcal{M}$, we have $X(p) \in T_p(\mathcal{M})$. The set of all C^k-vector fields is denoted $\mathfrak{X}(\mathcal{M})$, or $\mathfrak{X}_k(\mathcal{M})$ if the regularity is not obvious from context.

ii) A C^k-one-form is a C^k-map $\omega : \mathcal{M} \to T^*\mathcal{M}$ such that at every point $p \in \mathcal{M}$, we have $\omega(p) \in T_p^*(\mathcal{M})$. The set of all C^k-one-form fields is denoted $\mathfrak{X}^*(\mathcal{M})$, or $\mathfrak{X}_k^*(\mathcal{M})$ if the regularity is not obvious from context.

Notice that there is a natural paring $\mathfrak{X}_k(\mathcal{M}) \times \mathfrak{X}_k^*(\mathcal{M}) \to C^k(\mathcal{M};\mathbb{R})$ defined by

$$p \in \mathcal{M} \mapsto \omega|_p \left(X|_p \right).$$

We can extend the definition to higher rank tensor fields in the obvious fashion.

Suppose that $\mathcal{M}$ is a C^k-manifold and that $\varphi : \mathcal{U} \subset \mathcal{M} \to U \subset \mathbb{R}^n$ is a coordinate chart, and pick a basis $\{\boldsymbol{e}_i\}$ for $\mathbb{R}^n$. At each point, we have a basis for $T_p\mathcal{M}$ given by

$$\left\{ \left(\frac{\partial}{\partial x^i} \right)_p \right\}.$$

For each i, the map

$$\begin{array}{rccl} \frac{\partial}{\partial x^i} & : \mathcal{U} & \to & T\mathcal{U} \subset T\mathcal{M} \\ & p & \mapsto & \left(\frac{\partial}{\partial x^i}\right)_p \end{array}$$

defines a $C^{k-1}-$vector field on $\mathcal{U}$, i.e.

$$\frac{\partial}{\partial x^i} \in \mathfrak{X}_{k-1}(\mathcal{U}).$$

This vector field is also sometimes written ∂_i. In a similar fashion, dual to these vector fields we have the one-forms

$$dx^i \in \mathfrak{X}_{k-1}^*(\mathcal{U})$$

satisfying

$$dx^i \left(\frac{\partial}{\partial x^j} \right) = \delta^i{}_j.$$

We know that at each point $p \in \mathcal{U}$ that $\{(\partial_i)_p\}_{i=1,\ldots,n}$, $\{dx^i\big|_p\}_{i=1,\ldots,n}$ span $T_p\mathcal{M}$, $T_p^*\mathcal{M}$ respectively. As a consequence, any C^r-smooth $(p,q)-$tensor field with $r \leq k-1$ can be written locally as

$$T = T^{\mu_1,\ldots\mu_p}{}_{\nu_1,\ldots\nu_q} \frac{\partial}{\partial x^{\mu_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{\mu_p}} \otimes dx^{\nu_1} \otimes \cdots \otimes dx^{\mu_q}$$

Where $T^{\mu_1,\ldots\mu_p}{}_{\nu_1,\ldots\nu_q} : \mathcal{U} \to \mathbb{R}$ are C^r-functions, called the components of T in the coordinate chart $(\varphi,\mathcal{U})$. Using the chain rule, it is straightforward to demonstrate that the condition that the components are C^r-smooth is independent of the coordinate chart, provided $r \leq k-1$.

A.3.3 Derivations and commutators

Definition 25. A C^r-*derivation* on the C^k-manifold $\mathcal{M}$ is an $\mathbb{R}-$linear map

$$D : C^{s+1}(\mathcal{M}; R) \to C^s(\mathcal{M}; R)$$

for any $s \leq r$, which satisfies

i) $D(f_1 + \lambda f_2) = Df_1 + \lambda Df_2$,

ii) $D(f_1 f_2) = f_1 Df_2 + f_2 Df_1$,

for all $f_i \in C^{s+1}(\mathcal{M}; R)$, $\lambda \in \mathbb{R}$.

A vector field $V \in \mathfrak{X}_r(\mathcal{M})$ naturally defines a C^r-derivation. We can check that

$$D_V f(p) := V_p[f] = \frac{d}{dt} f \circ \gamma(t)\bigg|_{t=0}$$

where $\gamma(0) = p, \dot{\gamma}(0) = V_p$ satisfies the conditions above. Conversely, we can identify a C^r-derivation with a C^r-vector field. To see this, we first note that any C^1-function in $\mathbb{R}^n$ may be written locally about a point y as

$$g(x) = g(y) + g_i(x)(x^i - y^i),$$

where $g_i \in C^0(\mathbb{R}^n)$ with $g_i(y) = \frac{\partial g}{\partial x^i}(y)$. Now consider a point $p \in \mathcal{M}$ and a coordinate chart $(\mathcal{U}, \varphi)$ with $p \in \mathcal{U}$ and $\varphi = (\varphi^1, \ldots, \varphi^n)$. We deduce that any function $f \in C^1(\mathcal{M}; \mathbb{R})$ may be written for q in a neighbourhood of p as:

$$f(q) = f(p) + f_i(q)(\varphi^i(q) - \varphi^i(p))$$

for some $f_i \in C^0(\mathcal{M}; \mathbb{R})$ with $f_i(p) = \frac{\partial}{\partial x^i}\big|_p [f]$. Applying an arbitrary derivation to this formula, we deduce that

$$Df = (\varphi^i - \varphi^i(p))Df_i + f_i D\varphi^i$$

Evaluating this formula at p we deduce:

$$Df(p) = f_i(p) D\varphi^i(p) = \frac{\partial}{\partial x^i}\bigg|_p [f] D\varphi^i(p) = V|_p [f]$$

where $V|_p \in T_p\mathcal{M}$ is the tangent vector given in local coordinates by:

$$V|_p = D\varphi^i(p) \left.\frac{\partial}{\partial x^i}\right|_p$$

Defining $V \in \mathfrak{X}_r(\mathcal{U})$ by

$$V := D\varphi^i \frac{\partial}{\partial x^i}$$

we have

$$D = D_V.$$

We have shown that in a neighbourhood of any point there exists a V such that $D = D_V$. Since a vector field is uniquely determined by its action on functions, V is uniquely fixed by this condition.

Lemma A.6. *Let $\mathcal{M}$ be a C^k manifold. Suppose $X, Y \in \mathfrak{X}_r(\mathcal{M})$ for $r < k$. There exists a unique vector field $[X, Y] \in \mathfrak{X}_{r-1}(\mathcal{M})$ defined by the condition:*

$$[X, Y]f = X(Yf) - Y(Xf), \qquad \forall f \in C^r(\mathcal{M}; \mathbb{R})$$

Proof. By the previous discussion, it suffices to check that the operation $[X, Y]f$ defines a $C^{r-1}-$derivation. We calculate

$$\begin{aligned}[X, Y](f_1 + \lambda f_2) &= X(Y(f_1 + \lambda f_2)) - Y(X(f_1 + \lambda f_2)) \\ &= X(Yf_1) - Y(Xf_1) + \lambda\left[X(Yf_2) - Y(Xf_2)\right] \\ &= [X, Y]f_1 + \lambda[X, Y]f_2,\end{aligned}$$

using the $\mathbb{R}-$linearity of the action of a vector field on a function. We also find

$$\begin{aligned}[X, Y](fg) &= X(Y(fg)) - Y(X(fg)) \\ &= X(fYg + gYf) - Y(fXg + gXf) \\ &= fX(Yg) + (Xf)(Yg) + (Xg)(Yf) + gX(Yf) \\ &\quad - [fY(Xg) + (Yf)(Xg) + (Yg)(Xf) + gY(Xf)] \\ &= f\left[X(Yg) - Y(Xg)\right] + g\left[X(Yf) - Y(Xf)\right] \\ &= f[X, Y]g + g[X, Y]f\end{aligned}$$

Hence the operation $f \mapsto [X, Y]f$ is a derivation, to which we can associate the unique vector field $[X, Y]$. □

Exercise(∗). Working in a coordinate patch, $\mathcal{U}$ so that we can write

$$X = X^i \frac{\partial}{\partial x^i}, \qquad X = Y^i \frac{\partial}{\partial x^i},$$

for $X^i, Y^i \in C^r(\mathcal{U}; \mathbb{R})$, show that

$$[X, Y] = \left(X^i \frac{\partial Y^j}{\partial x^i} - Y^i \frac{\partial X^j}{\partial x^i}\right) \frac{\partial}{\partial x^j}.$$

A.3.4 Immersions and embeddings

Push forward, pull back

For this section, we shall make the assumption that all manifolds and objects that we discuss are smooth. We return to the situation we considered earlier, where we have two smooth manifolds $\mathcal{N}, \mathcal{M}$, of dimension $n, n + d$ respectively. Suppose we have a function $\phi \in C^\infty(\mathcal{N}; \mathcal{M})$. The mapping ϕ naturally induces relations between various geometric objects defined on the two manifolds $\mathcal{N}$, $\mathcal{M}$. We start with the pull back of a function by ϕ. Suppose $f \in C^\infty(\mathcal{M}; \mathbb{R})$, we define the *pull back* of f by ϕ, written $\phi^* f \in C^\infty(\mathcal{N}; \mathbb{R})$ by

$$\phi^* f := f \circ \phi.$$

Now let us consider vectors. Suppose $X \in T_p\mathcal{N}$, and that $\gamma \in C^\infty((-\epsilon, \epsilon), \mathcal{N})$ satisfies $\gamma(0) = p$, $\dot{\gamma}(0) = X$. We can define a curve, $\widetilde{\gamma} \in C^\infty((-\epsilon, \epsilon), \mathcal{M})$ by:

$$\widetilde{\gamma} := \phi \circ \gamma$$

We define the *push-forward* of X by ϕ, written $\phi_* X \in T_{\phi(p)}\mathcal{M}$ by:

$$\phi_* X = \dot{\widetilde{\gamma}}(0).$$

Lemma A.7. *Suppose that $X \in \mathfrak{X}(\mathcal{N})$, $f \in C^\infty(\mathcal{M}; \mathbb{R})$ and $\phi \in C^\infty(\mathcal{N}, \mathcal{M})$. Fix $p \in \mathcal{N}$. Then:*

$$X(\phi^* f)|_p = [(\phi_* X)\, f]|_{\phi(p)}\,,$$

and

$$f \phi_* X|_{\phi(p)} = \phi_* \left[(\phi^* f) X\right]|_{\phi(p)}$$

Proof. Fix $p \in \mathcal{N}$. Suppose that γ satisfies $\gamma(0) = p$, $\dot{\gamma}(0) = X|_p$. From the definition of a vector acting on a function, we have:

$$\begin{aligned} X(\phi^* f)|_p &= \left.\frac{d}{ds}(\phi^* f) \circ \gamma(s)\right|_{s=0} \\ &= \left.\frac{d}{ds}(f \circ \phi) \circ \gamma(s)\right|_{s=0} \\ &= \left.\frac{d}{ds} f \circ (\phi \circ \gamma)(s)\right|_{s=0} \\ &= \left.\frac{d}{ds} f \circ \widetilde{\gamma}(s)\right|_{s=0} \\ &= (\phi_* X)\, f|_{\phi(p)} \end{aligned}$$

For the second part, suppose that $g \in C^\infty((M); \mathbb{R})$, and calculate using the previous result:

$$\begin{aligned} [f\,(\phi_* X)\, g]|_{\phi(p)} &= f(\phi(p)) X(\phi^* g)|_p \\ &= (\phi^* f)\,(p)\; X(\phi^* g)|_p \\ &= [(\phi^* f)\, X(\phi^* g)]|_p \\ &= \phi_* \left[(\phi^* f) X\right] g|_{\phi(p)} \end{aligned}$$

□

Now, suppose that $\omega \in T^*_{\phi(p)}\mathcal{M}$ is a one-form. We can define the *pull-back* of ω by ϕ, written $\phi^* \omega \in T_p\mathcal{N}$ by:

$$[\phi^* \omega]\,(X) = \omega\,(\phi_* X)\,, \qquad \forall\ X \in T_p\mathcal{N}.$$

These definitions readily extend to allow us to define the push forward of any $(p, 0)-$tensor, and the pull-back of any $(0, q)-$tensor. Notice that the push forward and pull-back at each point p are linear maps between vector spaces.

Exercise(∗). Suppose $f \in C^\infty(\mathcal{M};\mathbb{R})$ and $\phi \in C^\infty(\mathcal{N},\mathcal{M})$. Show that:

$$d\left(\phi^* f\right) = \phi^* df.$$

Notation: Where the choice of map ϕ is clear from context, we will write $f^* = \phi^* f$, $X_* = \phi_* X$, etc.

Definition 26. We say that a map $\imath \in C^\infty(\mathcal{N};\mathcal{M})$ is a smooth *immersion* if the push forward map:

$$\imath_* : T_p\mathcal{N} \to T_{\phi(p)}\mathcal{M}$$

is an injective linear map for each $p \in \mathcal{N}$. We say that $\imath$ is a smooth *embedding* if moreover $\imath$ is injective. In this case, we write $\imath : \mathcal{N} \hookrightarrow \mathcal{M}$.

If we have that $\imath \in C^\infty(\mathcal{N};\mathcal{M})$ is an immersion, then for each $p \in \mathcal{N}$, we can identify $T_p\mathcal{N} \simeq T_{\imath(p)}\imath(\mathcal{N})$, where $T_{\imath(p)}\imath(\mathcal{N})$ is a linear subspace of $T_{\phi(p)}\mathcal{M}$.

Canonical immersions and extensions

Lemma A.8 (Canonical Immersion Theorem)**.** *Suppose $\imath : \mathcal{N} \hookrightarrow \mathcal{M}$ is an immersion, and fix $p \in \mathcal{N}$. There exist coordinate charts $(\mathcal{U},\varphi)$ and $(\mathcal{V},\psi)$ for $\mathcal{N}$ and $\mathcal{M}$ respectively, with $\varphi(p) = 0$, such that*

$$\begin{array}{rcrcl} \psi\circ\imath\circ\varphi^{-1} & : & \varphi(\mathcal{U}) & \to & \psi(\mathcal{V}), \\ & & (x_1,\ldots,x_n) & \mapsto & (x_1,\ldots,x_n,0,\ldots,0). \end{array}$$

Proof. Let us fix $(\mathcal{U},\varphi)$ to be a coordinate chart for $\mathcal{N}$ centred at p, i.e., $\varphi(p) = 0$. Now consider any $(\mathcal{V},\psi)$, a coordinate chart for $\mathcal{M}$ centred at $\imath(p)$, and set $f = \psi\circ\imath\circ\varphi^{-1}$. We will denote by $(x_i)_{i=1,\ldots n}$ the coordinates on $\varphi(\mathcal{U}) \subset \mathbb{R}^n$, and by $(y_a)_{a=1,\ldots,n+d}$ the coordinates of $\psi(\mathcal{V}) \subset \mathbb{R}^{n+d}$.

The fact that $\imath$ is an immersion implies that $Df(0)$ is injective, so by a linear transformation on $\mathbb{R}^{n+d}$, we may assume

$$Df(0) = \begin{pmatrix} I_{n\times n} & 0_{d\times n} \end{pmatrix}.$$

Now set

$$h(y_1,\ldots,y_{n+d}) = f(y_1,\ldots y_n) + (0,\ldots,0,y_{n+1},\ldots,y_{n+d}),$$

defined on some neighbourhood of 0 in $\mathbb{R}^{d+n}$. Clearly $h(0) = 0$ and $Dh(0) = I_{(n+d)\times(n+d)}$, so by restricting the domain to a subset if necessary, we can assume that h is smoothly invertible. We can easily verify that on a sufficiently small neighbourhood of 0, the map $h^{-1}\circ f$ takes $(x_1,\ldots,x_n)$ to $(x_1,\ldots,x_n,0,\ldots,0)$, so by redefining ψ to be $\psi\circ h$ and shrinking $\mathcal{U},\mathcal{V}$ as necessary, we are done. □

From here, we can derive various extension Lemmas which allow us to locally extend objects defined on $\imath(\mathcal{N})$ to a neighbourhood in $\mathcal{M}$ in a smooth fashion. Obviously such extensions are highly non-unique.

Corollary A.3. *Suppose $\imath : \mathcal{N} \to \mathcal{M}$ is an immersion, and let $p \in \mathcal{N}$. There exists a neighbourhood $\mathcal{W}$ of p in M such that for any $f \in C^\infty(\mathcal{N};\mathbb{R})$ we can find a function $\widetilde{f} \in C^\infty(\mathcal{M};\mathbb{R})$ with $f = \imath^*\widetilde{f}$ in $\mathcal{W}$.*

Proof. For $p \in \mathcal{N}$ we have local coordinate charts $(\mathcal{U},\varphi)$, $(\mathcal{V},\psi)$ centred at $p, \imath(p)$ such that $\imath$ is given by the canonical immersion. Pick $W \Subset \varphi(\mathcal{U})$, and define $\mathcal{W} = \varphi^{-1}(W)$. Let $\chi_T : U \to [0,1]$ be a smooth cut-off function equal to unity on W and vanishing near the boundary of $\varphi(\mathcal{U})$. We define $\overline{f} = f \circ \varphi^{-1}$ and finally we set $\widetilde{f} = 0$ on $\mathcal{M} \setminus \mathcal{V}$, and $\widetilde{f} = \overline{\overline{f}} \circ \psi$, for

$$\overline{\overline{f}}(y_1, \dots y_{n+d}) = \chi_N(y_{n+1},\dots,y_{n+d})\chi_T(y_1,\dots,y_n)\overline{f}(y_1,\dots,y_n).$$

Here $\chi_N : \mathbb{R}^d \to [0,1]$ is a smooth cut-off function, equal to 1 at the origin and supported on a sufficiently small set that

$$\text{supp } (\chi_N(y_{n+1},\dots,y_{n+d})\chi_T(y_1,\dots,y_n)) \subset \psi(\mathcal{V}).$$

□

Let $\mathcal{U} \subset \mathcal{N}$ be an open set, and $X \in \mathfrak{X}(\mathcal{U})$. We say that $\widetilde{X} \in \mathfrak{X}(\mathcal{M})$ is an extension of X away from $\imath(\mathcal{U})$ if $X(f^*) = \imath^*\left[\widetilde{X}(f)\right]$ in $\mathcal{U}$ for any $\phi \in C^\infty(\mathcal{M};\mathbb{R})$. Equivalently, $X_* = \widetilde{X}$ on $\imath(\mathcal{U})$.

Corollary A.4. *Suppose $\imath : \mathcal{N} \to \mathcal{M}$ is an immersion and let $p \in \mathcal{N}$. There exists a neighbourhood $\mathcal{W}$ of p in $\mathcal{N}$ such that for any $X \in \mathfrak{X}(\mathcal{N})$ we can find $\widetilde{X} \in \mathfrak{X}(\mathcal{M})$ which is an extension of X away from $\imath(\mathcal{W})$.*

Proof. Take the same charts as in the previous proof, so that (x^i) are local coordinates on $\mathcal{N}$ and (y^a) are local coordinates on $\mathcal{M}$. We have that:

$$X = \sum_{i=1}^n X^i \frac{\partial}{\partial x^i},$$

for $X^i \in C^\infty(\mathcal{U},\mathbb{R})$. By the previous result, there exists $\mathcal{W} \subset \mathcal{N}$ containing p, and $\widetilde{X}^i \in C^\infty(\mathcal{N};\mathbb{R})$ such that $\imath^*\widetilde{X}^i = X^i$. Defining $\widetilde{X} \in \mathfrak{X}(\mathcal{M})$ by:

$$\widetilde{X} = \sum_{i=1}^n \widetilde{X}^i \frac{\partial}{\partial y^i}$$

we have the desired extension away from $\imath(\mathcal{W})$. □

This allows us to prove the following result

Lemma A.9. *Suppose $X, Y \in \mathfrak{X}(\mathcal{N})$, and pick $p \in \mathcal{N}$. Let $\widetilde{X}, \widetilde{Y} \in \mathfrak{X}(\mathcal{M})$ be extensions of X, Y away from $\imath(\mathcal{U})$ for some neighbourhood $\mathcal{U}$ of p. Then $\left[\widetilde{X}, \widetilde{Y}\right]$ is an extension of $[X, Y]$ away from $\imath(\mathcal{U})$.*

Proof. We calculate in $\mathcal{U}$:

$$\begin{aligned}\left(\left[\widetilde{X},\widetilde{Y}\right](\phi)\right)^* &= \left[\widetilde{X}\left(\widetilde{Y}(\phi)\right)\right]^* - \left[\widetilde{Y}\left(\widetilde{X}(\phi)\right)\right]^* \\ &= X\left(\left[\widetilde{Y}(\phi)\right]^*\right) - Y\left(\left[\widetilde{X}(\phi)\right]^*\right) \\ &= X\left(Y\left(\phi^*\right)\right) - Y\left(X\left(\phi^*\right)\right) \\ &= [X,Y]\left(\phi^*\right)\end{aligned}$$

whence we are done. □

Diffeomorphisms

We say that the map $\phi \in C^\infty(\mathcal{N};\mathcal{M})$ is a *diffeomorphism* if ϕ^{-1} exists and belongs to $C^\infty(\mathcal{M};\mathcal{N})$. A diffeomorphism allows us to identify $T_p\mathcal{N} \simeq T_{\phi(p)}\mathcal{M}$ for all $p \in \mathcal{N}$.

We can construct a family of diffeomorphisms taking $\mathcal{M}$ to itself from a vector field $X \in \mathfrak{X}(\mathcal{M})$. For this, we require the following result:

Lemma A.10. *Let $X \in \mathfrak{X}(\mathcal{M})$, and suppose $p \in \mathcal{M}$. Then there exists a parameterised curve $\gamma \in C^\infty((a,b);\mathcal{M})$, where $a < 0 < b$, satisfying:*

$$\gamma(0) = p, \qquad \dot{\gamma}(s) = X|_{\gamma(s)}\,.$$

This is called an integral curve *of X starting at p, and is unique up to extension.*

Proof. Take a coordinate chart $(\varphi,\mathcal{U})$ with $p \in \mathcal{U}$, and write $X = X^i\frac{\partial}{\partial x^i}$, for $X^i \in C^k(\mathcal{U};\mathbb{R})$, and set $\varphi \circ \gamma(s) = x(s) = (x^i(s))$. The condition on γ becomes:

$$\dot{x}^i(s) = X^i \circ \varphi^{-1}(x(s)), \qquad x(0) = \varphi(p),$$

which is an ordinary differential equation, with Lipshitz right hand side, whose solutions are unique up to extension. By this local uniqueness property, the curve is defined independent of the coordinate chart, and is unique up to extension. □

Lemma A.11. *Suppose that $X \in \mathfrak{X}(\mathcal{M})$ has the property that for each $p \in \mathcal{M}$, the integral curve of X through p, written γ_p can be extended so that it belongs to $C^k((-\epsilon,\epsilon);\mathcal{M})$ for some ϵ independent of p. Then the map*

$$^X\phi_s : p \mapsto \gamma_p(s)$$

is a diffeomorphism, referred to as the one parameter family of diffeomorphisms induced by X.

A.4 Matrix Lie Groups

This section is meant to give a very brief introduction to Lie groups. We bypass a lot of important theory and focus our attention on the matrix Lie groups. As a result, some of the definitions are not standard, but allow us to get to our goal more quickly. It is certainly no substitute for a proper study of the beautiful topic of Lie groups. Think of the approach here as 'good enough' for our purposes.

We'll start by extending the familiar definition of a group to encompass an additional requirement, that of differentiability.

Definition 27. A *group* is a set G, together with a binary operation $\cdot$ such that

i) $a \cdot b \in G$ for all $a, b \in G$ [Closure]

ii) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for all $a, b \in G$ [Associativity]

iii) There exits $e \in G$ such that $e \cdot a = a \cdot e = a$ for all $a \in G$ [Existence of an identity]

iv) For each $a \in G$ there exists $b \in G$ such that $a \cdot b = b \cdot a = e$ [Existence of an inverse]

The identity element is unique, as is the inverse, and we write $b = a^{-1}$ if $a \cdot b = b \cdot a = e$.

A *Lie group* is a group, where the set G has a smooth[9] structure making it into a manifold in such a way that the operations $\cdot$ and $()^{-1}$ are smooth.

The many of the most important Lie groups are groups of matrices:

Example 17. Consider the set $GL(n)$ of invertible $n \times n$ real matrices. This is naturally a group with the group operation of matrix multiplication. Thinking of $GL(n)$ as an open subset of $\mathbb{R}^{n^2}$ with the standard real analytic structure, $GL(n)$ is also a manifold in a natural way. Matrix multiplication and inversion are smooth with respect to this structure, as can be seen by writing out the operations in components.

In fact, this example (and its subgroups) will be sufficiently rich to cover the situations that we are interested in for this course. We define in the obvious way

Definition 28. A *Lie subgroup* of a Lie group G is a subgroup H of G endowed with a topology and smooth structure making it into a Lie group in such a way that the inclusion map is an immersion. A *matrix Lie group* is an embedded Lie subgroup of $GL(n)$.

An important (and deep result) tells us when a subgroup of a Lie group is in fact a Lie subgroup:

Proposition 2 (Closed subgroup theorem)**.** *If H is a subgroup of G which is closed (in the topology of G), then H is an embedded Lie subgroup of G.*

This gives us the following result

Lemma A.12. *Suppose that H is a subgroup of $GL(n)$ with the property that for any sequence $A_n \in H$ with $A_n \to A \in Mat(n \times n)$ component-wise, either $A \in H$ or $A \notin GL(n)$. Then H is a matrix Lie group.*

Example 18.

1. $GL^+(n) = \{A \in GL(n) : \det A > 0\}$ is a matrix Lie group
2. $SL(n) = \{A \in GL(n) : \det A = 1\}$ is a matrix Lie group

[9]It turns out that we can assume that the manifold is C^ω without any loss of generality.

3. $O(n) = \{A \in GL(n) : A^T A = I\}$ is a matrix Lie group
4. The set of matrices $A \in GL(2)$ of the form

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \qquad \theta \in \mathbb{Q},$$

is a subgroup of $GL(2)$, but *not* a matrix Lie group.

A.4.1 The matrix exponential

A very useful tool to understand the structure of matrix Lie groups is the matrix exponential. Before we define the matrix exponential, it's first useful to introduce a norm on $Mat(n \times n)$, called the operator norm. We define

$$||A||_{op.} = \sup_{x\in\mathbb{R}^n, x\neq 0} \frac{||Ax||}{||x||}$$

where $||\cdot||$ is the usual Euclidean norm. We clearly have that for any non-zero $x \in \mathbb{R}^n$:

$$||Ax|| \leq ||A||_{op.} ||x|| .$$

we deduce that

$$||(A+B)x|| \leq ||Ax|| + ||Bx|| \leq ||A||_{op.} ||x|| + ||B||_{op.} ||x||$$

so that

$$||A+B||_{op.} \leq ||A||_{op.} + ||B||_{op.} .$$

so the triangle inequality is satisfied by this norm. The other criteria for $||\cdot||_{op.}$ to define a norm are straightforward to verify.

We can also show by a simple induction that

$$||A^n x|| \leq ||A||^n_{op.} ||x||$$

holds for any x and thus

$$||A^n||_{op.} \leq ||A||^n_{op.} .$$

Definition 29. The matrix exponential of an element $A \in Mat(n \times n)$ is defined to be

$$\text{Exp } A = e^A := \sum_{k=0}^{\infty} \frac{A^n}{n!}. \tag{A.11}$$

Exercise($*$). i) Show that the sum on the right hand side of (A.11) converges in the operator norm for any A (and hence with respect to any other norm on $Mat(n\times n)$). Show also that

$$\left(e^A\right)^T = e^{A^T}$$

ii*) Show that the matrix exponential is a real analytic function $Mat(n \times n) \to GL(n)$, where both spaces inherit the canonical real analytic structure of $\mathbb{R}^{n^2}$, and show that

$$\frac{d}{dt}e^{tA} = Ae^{tA} = e^{tA}A$$

and

$$e^{tA}e^{sA} = e^{(t+s)A} = e^{sA}e^{tA}$$

for $t, s \in \mathbb{R}$. Deduce that

$$\left(e^{A}\right)^{-1} = e^{-A}.$$

iv) Show that

$$\lim_{n\to\infty}\left(1 + \frac{1}{n}A\right)^n = e^A,$$

v) Deduce that if $\Gamma : (-\epsilon, \epsilon) \to GL(n)$ is a C^1-curve with $\Gamma(0) = I$ and $\dot{\Gamma}(0) = A$, we have

$$\lim_{n\to\infty}\left[\Gamma\left(\frac{1}{n}\right)\right]^n = e^A$$

A.4.2 The Lie algebra

With the matrix exponential in our pocket, we are ready to define the Lie algebra of a matrix Lie group.

Definition 30. The Lie algebra $\mathfrak{h}$ of a matrix Lie group H is defined to be the set of all matrices $a \in Mat(n \times n)$ such that $e^{ta} \in H$ for all $t \in \mathbb{R}$.

Theorem A.5. *Let $\mathfrak{h}$ be the Lie algebra of a matrix Lie group H. Then*

i) $a \in \mathfrak{h} \Leftrightarrow a \in T_I H$, i.e. a is in the Lie algebra if and only if there exists a C^1-curve $\Gamma : (-\epsilon, \epsilon) \to H$ with $\Gamma(0) = I$ and $\dot{\Gamma}(0) = a$.

ii) $\mathfrak{h}$ is a vector subspace of $Mat(n \times n)$, i.e. if $a, b \in \mathfrak{h}$ then $a + \lambda b \in \mathfrak{h}$

iii) $\mathfrak{h}$ is closed under the matrix commutator, i.e. if $a, b \in \mathfrak{h}$ then $[a, b] = ab - ba \in \mathfrak{h}$

Proof. i) "⇒" is trivial, take $\Gamma(t) = e^{ta}$. For "⇐", we use the fact that $\Gamma\left(\frac{t}{n}\right) \in H$ for any $t \in \mathbb{R}$ and n a sufficiently large integer to deduce that

$$\left[\Gamma\left(\frac{t}{n}\right)\right]^n \in H$$

for n sufficiently large. We know

$$\lim_{n\to\infty}\left[\Gamma\left(\frac{t}{n}\right)\right]^n = e^{ta}$$

which is invertible, so by the closeness of H, we have $e^{ta} \in H$ and we're done.

ii) Suppose $a, b \in \mathfrak{h}$ and fix $\lambda \in \mathbb{R}$. Consider the curve $\Gamma : (-\epsilon, \epsilon) \to H$ given by

$$\Gamma(t) = e^{ta} e^{t\lambda b}$$

This is clearly C^1 and we readily calculate that

$$\Gamma(0) = I, \quad \dot{\Gamma}(0) = a + \lambda b,$$

so that $a + \lambda b \in \mathfrak{h}$.

iii) Now consider the curve $\Gamma : (-\epsilon, \epsilon) \to H$ given by

$$\Gamma(t) = \begin{cases} e^{\sqrt{t}a} e^{\sqrt{t}\lambda b} e^{-\sqrt{t}(a+b)} & t > 0 \\ I & t = 0 \\ e^{\sqrt{-t}b} e^{\sqrt{-t}\lambda a} e^{-\sqrt{-t}(a+b)} & t < 0 \end{cases}$$

This is smooth for $t \neq 0$. Expanding for t small and positive, we have

$$\begin{aligned} \Gamma(t) &= \left[I + \sqrt{t}a + \frac{1}{2}ta^2 + o(t)\right]\left[I + \sqrt{t}b + \frac{1}{2}tb^2 + o(t)\right]\left[I - \sqrt{t}(a+b) + \frac{1}{2}t(a+b)^2 + o(t)\right] \\ &= I + t\left\{\frac{1}{2}\left[a^2 + b^2 + (a+b)^2\right] + ab - a(a+b) - b(a+b)\right\} + o(t) \\ &= I + \frac{t}{2}\{ab - ba\} + o(t) \end{aligned}$$

Noting that we can obtain the expansion for t small and negative by interchanging $t \leftrightarrow -t$ and $a \leftrightarrow b$, we deduce that Γ is in fact C^1 with $\dot{\Gamma}(0) = \frac{1}{2}[a, b]$.

□

Thus the Lie algebra of a matrix Lie group is naturally a vector space endowed with an antisymmetric bilinear operation $[,]$, which moreover satisfies the Jacobi identity

$$[[a, b], c] + [[b, c], a] + [[c, a], b] = 0$$

One can abstractly define a Lie algebra to be a vector space endowed with such an operation, but the concrete realisation as a space of matrices endowed with the matrix commutator is the most useful for us.

A.4.3 The orthogonal group

As a brief example, we will give a brief treatment of the orthogonal group. Recall

$$O(n) = \{A \in GL(n) : A^T A = I\}.$$

Let us find the Lie algebra, $\mathfrak{o}(n)$. Suppose $\Gamma : (-\epsilon, \epsilon) \to O(n)$ with $\Gamma(0) = I$. We have

$$\Gamma(t)^T \Gamma(t) = I, \quad t \in (-\epsilon, \epsilon)$$

so we can differentiate this condition to find

$$\dot{\Gamma}(t)^T \Gamma(t) + \Gamma(t)^T \dot{\Gamma}(t) = 0, \quad t \in (-\epsilon, \epsilon)$$

so that, setting $t = 0$ we deduce that if $\dot{\Gamma}(0) = a \in \mathfrak{o}(n)$ we must have

$$a^T + a = 0$$

so that a is antisymmetric. Conversely, suppose that a is antisymmetric. We have

$$\left(e^{ta}\right)^T = e^{ta^T} = e^{-ta} = \left(e^{ta}\right)^{-1},$$

so that $e^{ta} \in O(n)$. Thus $\mathfrak{o}(n)$ is precisely the set of antisymmetric matrices.

We should be careful here. Although $e^{ta} \in O(n)$ whenever $a \in \mathfrak{o}(n)$, not every element of $O(n)$ can be written as the exponential of an antisymmetric matrix. We know that if $A \in O(n)$ a simple calculation shows that $\det A = \pm 1$. The determinant of a matrix is a continuous real valued function, so on any continuous curve $\Gamma : (a, b) \to O(n)$ the determinant must be constant. In particular for $a \in \mathfrak{o}(n)$ we have:

$$\det \left(e^{at}\right) = 1.$$

In fact, we can show that by exponentiating elements of $\mathfrak{o}(n)$ it is possible to construct any element of $SO(n)$.

Exercise(∗). Let

$$a_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad a_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad a_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Show that

$$[a_i, a_j] = \epsilon_{ijk} a_k,$$

where ϵ_{ijk} is the totally antisymmetric tensor with $\epsilon_{123} = 1$. Show also that:

$$e^{\theta a_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

and find similar expressions for $e^{\theta a_2}$, $e^{\theta a_3}$. Deduce that by exponentiating $\mathfrak{o}(3)$ we can produce a matrix representing an arbitrary rotation, i.e. an arbitrary element of $S0(3)$.

Appendix B

Bonus material

B.1 Integration in $\mathbb{E}^{1,3}$ and the divergence theorem

We will need to integrate over surfaces embedded in $\mathbb{E}^{1,3}$ and relate these surface integrals to volume integrals using the divergence theorem in the usual way. We would like the expressions we write down to be invariant under a change of frame. This is slightly complicated when the metric has Lorentzian signature. The main issue is that the notion of *outwards unit normal* becomes tricky when vectors can have zero norm.

First, suppose we have an open set $U \subset \mathbb{E}^{1,3}$ and that $f : U \mapsto \mathbb{R}$ is a function defined on U. Picking an inertial frame $\{\vec{e}_\mu\}$, we naturally have that $\widetilde{f} : (x^\mu) \mapsto f(x^\mu \vec{e}_\mu)$ is a map from some subset of $\widetilde{U} \subset \mathbb{R}^4$ to $\mathbb{R}$. We say that f is measurable if $\widetilde{f}$ is Lebesgue measurable. Consider now

$$I[f] := \int_{\widetilde{U}} \widetilde{f}(x) dx$$

where dx is the standard Lebesgue measure on $\mathbb{R}^4$. Now consider a different choice of inertial frame $\{\vec{e}_\mu{}'\}$. We define $\widetilde{f}' : (x'^\mu) \mapsto f(x'^\mu \vec{e}_\mu{}')$, which maps $\widetilde{U}' \subset \mathbb{R}^4$ to $\mathbb{R}$. We have

$$\widetilde{f}(x) = \widetilde{f}'(\Lambda x)$$

Where $\Lambda \in O(1,3)$ is the matrix representing the change of basis. Now, we have

$$I'[f] = \int_{\widetilde{U}'} \widetilde{f}'(x) dx = \int_{\widetilde{U}} \widetilde{f}(y) \left|\det \Lambda\right| dy = \int_{\widetilde{U}} \widetilde{f}(y) dy = I[f]$$

where we change variables $x^\mu = \Lambda^\mu{}_\nu y^\nu$ and use that $|\det \Lambda| = 1$. Thus, we can define

$$\int_U f dX := I[f],$$

and this definition is independent of the inertial frame.

Next, let us introduce the *alternating tensor*. This is a $(0,4)-$tensor, defined to be totally antisymmetric (i.e. antisymmetric under exchange of any pair of indices), and such that in a given inertial frame we have

$$\varepsilon_{0123} = 1.$$

Under a Lorentz transformation with metric $\Lambda^\mu{}_\nu$, the alternating tensor transforms as[1]

$$\varepsilon_{\mu\nu\sigma\tau} = (\det \Lambda)\varepsilon'_{\mu\nu\sigma\tau},$$

so that provided we restrict the transformations to belong to $SO(1,3)$, this tensor is invariant.

Now suppose that we have a surface Σ which may be written as $\Sigma = \varphi(U)$ for a smooth parameterisation φ:

$$\begin{array}{rccc} \vec{\varphi} : & U \subset \mathbb{R}^3 & \to & \Sigma \subset \mathbb{E}^{1,3} \\ & (y^1, y^2, y^3) & \mapsto & \varphi^\mu(y)\boldsymbol{e}_\mu \end{array}$$

The choice of parameterisation naturally gives an orientation to the surface. At any point, the ordered set of vectors

$$\left\{ \frac{\partial \vec{\varphi}}{\partial y^1}, \frac{\partial \vec{\varphi}}{\partial y^2}, \frac{\partial \vec{\varphi}}{\partial y^3} \right\}$$

defines an oriented basis for $T\Sigma$.

The vector surface measure of Σ with respect to this parameterisation is the vector measure (defined on a subset of $\mathbb{R}^3$):

$$dS^{\vec{\varphi}}_\mu = \varepsilon_{\mu\nu\sigma\tau} \frac{\partial \varphi^\nu}{\partial y^1} \frac{\partial \varphi^\sigma}{\partial y^2} \frac{\partial \varphi^\tau}{\partial y^3} dy$$

Where dy is the standard Lebesgue measure on $\mathbb{R}^3$.

For example, suppose we consider the plane $\Sigma = \{t = 0\}$, which we can parameterize by the map:

$$\begin{array}{rccc} \vec{\varphi} : & U \subset \mathbb{R}^3 & \to & \Sigma \subset \mathbb{E}^{1,3} \\ & (y^1, y^2, y^3) & \mapsto & y^i \boldsymbol{e}_i \end{array}$$

so that $\varphi^0(y) = 0, \varphi^i(y) = y^i$. Then we have that for this parameterisation

$$dS^\varphi_0 = dy, \qquad dS^\varphi_i = 0.$$

Exercise(∗). a) Show that if $F : W \to U$ is an orientation preserving diffeomorphism between subsets of $\mathbb{R}^3$ then

$$\int_{y \in U} V^\mu(\vec{\varphi}(y)) dS^{\vec{\varphi}}_\mu = \int_{y \in W} V^\mu(\vec{\varphi} \circ F(y)) dS^{(\vec{\varphi} \circ F)}_\mu.$$

b) Show that if P is the orientation reversing diffeomorphism $P : W \to U$ given by $y \to -y$, then we have

$$\int_{y \in U} V^\mu(\vec{\varphi}(y)) dS^{\vec{\varphi}}_\mu = -\int_{y \in W} V^\mu(\vec{\varphi} \circ P(y)) dS^{(\vec{\varphi} \circ P)}_\mu.$$

Deduce that

$$\int_{y \in U} V^\mu(\vec{\varphi}(y)) dS^{\vec{\varphi}}_\mu$$

depends only on the orientation of Σ and not on the choice of parameterisation.

[1] Check that you understand why this is!

With the vector surface measure, we can define the flux of a sufficiently smooth vector field V through Σ:

$$\int_\Sigma \vec{V} \cdot d\vec{S} := \int_{y\in U} V^\mu(\vec{\varphi}(y)) dS^{\vec{\varphi}}_\mu$$

The flux of $\vec{V}$ through Σ does not depend on the choice of parameterisation, but it does depend on the orientation for the surface Σ. By taking a partition of unity to localise $\vec{V}$ in coordinate patches[2], we can define the flux through any oriented surface[3] $\Sigma \subset \mathbb{E}^{1,3}$.

If Σ is an orientable surface which is everywhere spacelike or timelike, we can define a scalar measure on Σ, which we'll denote by $d\sigma$:

$$\int_\Sigma f d\sigma = \int_\Sigma f \vec{N} \cdot d\vec{S}$$

where $\vec{N}$ is a unit normal of Σ with respect to η (i.e. $\eta(\vec{N}, \vec{N}) = \pm 1$) with the direction chosen such that $N^\mu dS^\varphi_\mu$ is a positive measure on $U \subset \mathbb{R}^3$ for each parameterisation respecting the orientation. Notice that if Σ is not either everywhere spacelike or timelike then this definition does not make sense because a null surface does not have a well defined unit normal. We still have a perfectly reasonable vector measure, but no scalar measure.

Theorem B.1. *Suppose that $\vec{V}$ is a vector field on a bounded domain $U \subset \mathbb{E}^{1,3}$, whose boundary $\Sigma = \partial U$ is piecewise smooth. Suppose also that $V^\mu \in C^1(\overline{U})$. Then we have*

$$\int_U \nabla_\mu V^\mu dX = \int_\Sigma \vec{V} \cdot d\vec{S},$$

where the orientation of Σ is chosen such that if $\vec{K}$ is an outwards pointing vector transverse to Σ, then

$$\{\vec{K}, \vec{V}_1, \vec{V}_2, \vec{V}_3\}$$

is a positively oriented basis for $T_p\mathbb{E}^{1,3}$ whenever $\{\vec{V}_1, \vec{V}_2, \vec{V}_3\}$ is a positively oriented basis for $T_p\Sigma$.

We shall omit the proof here. It can be deduced from the usual divergence theorem on $\mathbb{R}^4$ keeping careful account of changes of sign.

Lemma B.1. *With the same set-up as in Theorem B.1, the choice of orientation for Σ is equivalent to the requirement that*

$$K^\mu dS^\varphi_\mu$$

is a positive measure on $U \subset \mathbb{R}^3$, whenever $\vec{\varphi} : U \to \Sigma$ is a local parameterisation respecting the orientation.

---

[2]This is a technical point, which can usually be avoided. If the surface Σ cannot be smoothly parameterised by a single coordinate patch then we have to write our vector field $\vec{V}$ as a sum of terms each of which is supported on a single patch. We can then define the integral of $\vec{V}$ over the surface by linearity.

[3]i.e. a surface with a consistent global choice of orientation

Corollary B.2. *Suppose that $\vec{V}$ is a vector field on a bounded domain $U \subset \mathbb{E}^{1,3}$, whose boundary $\Sigma = \partial U$ is piecewise smooth and can be written as $\Sigma = \Sigma_s \cup \Sigma_t$ where Σ_s is spacelike and Σ_t is timelike. Suppose also that $V^\mu \in C^1(\overline{U})$. Then we have*

$$\int_U \nabla_\mu V^\mu dX = \int_{\Sigma_s} \vec{V} \cdot \vec{N} d\sigma - \int_{\Sigma_t} \vec{V} \cdot \vec{N} d\sigma,$$

where $\vec{N}$ is the unit outwards *normal.*

Notice here that there is a sign change for the timelike surfaces relative to the spacelike surfaces.

Example 19. Fix an inertial frame for $\mathbb{E}^{1,3}$. Consider the cylinder $\mathcal{C} = \{x^\mu : 0 < x^0 < 1, |\boldsymbol{x}| < 1\}$. We have:

$$\int_{\mathcal{C}} \nabla_\mu V^\mu dX = \int_{x^0=1, \boldsymbol{x} \in B_1(0)} V^t dx - \int_{x^0=1, \boldsymbol{x} \in B_1(0)} V^t dx + \int_{[0,1] \times \partial B_1(0)} \boldsymbol{V} \cdot \boldsymbol{n} \, d\sigma dt$$

where $B_r(\boldsymbol{x})$ is the Euclidean ball of radius r centred at $\boldsymbol{x}$ and $\boldsymbol{n}, d\sigma$ are the usual outward directed normal and surface measure of the unit Euclidean sphere.

Index