Differential Geometry IV: General Relativity

G. Moschidis

1.1 Let V be an n-dimensional vector space and $m : V \times V \to \mathbb{R}$ a Lorentzian inner product on V. Recall that, for any timelike vector $v \in V$, we have defined

$$|v| \doteq \sqrt{-m(v,v)}$$

(a) Let $v \in V$ be a *timelike* vector in V. Show that the hyperplane

$$v^{\perp} \doteq \big\{ w \in V : m(v,w) = 0 \big\}$$

is a spacelike subspace of V.

(b) Show that that, for any two timelike vectors $v, w \in V$, the inverse Cauchy–Schwarz inequality

$$|m(v,w)| \geqslant |v||w|$$

and (in the case v, w belong to the same component of the timelike cone I) the inverse triangle inequality

$$|v+w| \geqslant |v| + |w|$$

hold, with equality only in the case when v and w are collinear.

Solution. (a) In order to show that $v^{\perp}$ is a spacelike subspace, we merely have to show that the restriction of the inner product m on $v^{\perp}$ is positive definite, namely that

$$m(w,w) > 0 \quad \text{for all } w \in v^{\perp} \setminus 0. \tag{1}$$

Assume that, in a given orthonormal basis $\{e_0, e_1, \dots, e_{n-1}\}$ of V (with the convention that $m(e_0, e_0) = -1$ and $m(e_i, e_i) = +1$ for $i \geqslant 1$), the components of the vector v are $(v^0, \dots, v^{n-1})$; then the fact that v is timelike translates into

$$m(v,v) < 0 \Leftrightarrow -(v^0)^2 + \sum_{i=1}^{n-1} (v^i)^2 < 0.$$

Thus,

$$\Big(\sum_{i=1}^{n-1} (v^i)^2 \Big)^{\frac{1}{2}} < |v^0| \tag{2}$$

which implies that

$$|v^0| > 0 \tag{3}$$

(since $v \neq 0$ by our convention for timelike vectors)

If $w = (w^0, w^1, \dots, w^{n-1})$ belongs to $v^{\perp} \setminus 0$ then

$$0 = m(v,w) = -v^0 w^0 + \sum_{i=1}^{n-1} v^i w^i.$$

Moving the term $v^0 w^0$ to the left hand side and using the Cauchy–Schwarz inequality, we can therefore bound

$$|v^0 w^0| = \Big| \sum_{i=1}^{n-1} v^i w^i \Big| \leqslant \Big(\sum_{i=1}^{n-1} (v^i)^2 \Big)^{\frac{1}{2}} \Big(\sum_{i=1}^{n-1} (w^i)^2 \Big)^{\frac{1}{2}}. \tag{4}$$

We can now distinguish two cases:

1. If $w^0 = 0$, then w is necessarily spacelike, since

$$m(w,w) = -(w^0)^2 + \sum_{i=1}^{n-1}(w^i)^2 = \sum_{i=1}^{n-1}(w^i)^2 > 0$$

(since we assumed that $w \neq 0$).

2. If $w^0 \neq 0$, the bound (2) implies that

$$\Big(\sum_{i=1}^{n-1}(v^i)^2\Big)^{\frac{1}{2}}|w^0| < |v^0 w^0|. \tag{5}$$

Moreover, the left hand side of (4) in this case cannot be equal to 0 (recall (3)), thus the right hand side of (4) cannot vanish; this implies that

$$\Big(\sum_{i=1}^{n-1}(v^i)^2\Big)^{\frac{1}{2}} > 0.$$

Combining (4) and (5) we therefore obtain

$$\Big(\sum_{i=1}^{n-1}(v^i)^2\Big)^{\frac{1}{2}}|w^0| < \Big(\sum_{i=1}^{n-1}(v^i)^2\Big)^{\frac{1}{2}}\Big(\sum_{i=1}^{n-1}(w^i)^2\Big)^{\frac{1}{2}},$$

from which we infer (after dividing with $\Big(\sum_{i=1}^{n-1}(v^i)^2\Big)^{\frac{1}{2}}$) that

$$|w^0| < \Big(\sum_{i=1}^{n-1}(w^i)^2\Big)^{\frac{1}{2}},$$

which is equivalent to $m(w,w) > 0$.

(b) Let $\lambda \in \mathbb{R}$ be the unique number such that $w - \lambda v \in v^\perp$; by solving the equation $m(w-\lambda v, v) = 0$, we readily calculate

$$\lambda = \frac{m(w,v)}{m(v,v)}.$$

In view of part (a) of this exercise, $\lambda \neq 0$ (since two timelike vectors cannot be perpdendicular to each other) and the vector $\tilde{w} = w - \lambda v$ is spacelike (since it belongs to $v^\perp$). Therefore,

$$\begin{aligned}
0 &\leqslant m(\tilde{w},\tilde{w}) \\
&= m(w-\lambda v, w-\lambda v) \\
&= m(w,w) - 2\lambda m(v,w) + \lambda^2 m(v,v) \\
&= m(w,w) - 2\frac{m(v,w)}{m(v,v)}m(v,w) + \Big(\frac{m(w,v)}{m(v,v)}\Big)m(v,v)
\end{aligned}$$

$$= m(w,w) - \frac{(m(v,w))^2}{m(v,v)}.$$

Thus (since $m(w,w), m(v,v) < 0$),

$$(m(v,w))^2 \geqslant (-m(v,v))(-m(w,w)),$$

with equality only if $\tilde{w} = 0 \Rightarrow w = \lambda v$.

If v, w lie on the same timecone, then $v + w$ is also a timelike vector and $m(v,w) < 0$. We can then compute:

$$\begin{aligned} |v+w|^2 &= -m(v+w, v+w) \\ &= -m(w,w) - 2m(v,w) - m(v,v) \\ &= |w|^2 - 2m(v,w) + |v|^2. \end{aligned}$$

Noting that $-m(v,w) = |m(v,w)|$ since v, w belong to the same timecone, we infer:

$$\begin{aligned} |v+w|^2 &= |w|^2 + 2|m(v,w)| + |v|^2 \\ &\geqslant |w|^2 + 2|v||w| + |v|^2, \end{aligned}$$

where, in the last line above, we made use of the inverse Cauchy–Schwarz inequality that we established earlier. Thus,

$$|v+w| \geqslant |v| + |w|,$$

with equality holding only in the case when the Cauchy–Schwarz inequality used for v, w becomes an equality.

1.2 Let V be an $(n+1)$-dimensional vector space equipped with a Lorentzian inner product m.

(a) Prove that any two *null* vectors v, w of V that are orthogonal are also *collinear*.

(b) Prove that if v and w are *causal* vectors that are orthogonal, then they have to be *null* and *collinear*.

(c) Prove that if v is a *null* vector, then its orthogonal complement

$$v^{\perp} = \big\{ w \in V : m(v,w) = 0 \big\}$$

is a null hyperplane containing v.

Solution. (a) First, we will pick an orthonormal basis for V in which the expression for one of the vectors (say v) becomes the simplest possible: Let e_0 be a *unit* timelike vector (i.e. $m(e_0, e_0) = -1$) in the same component of the timecone as v (i.e. $m(e_0, v) < 0$; recall that, as we proved earlier, any vector orthogonal to a timelike vector must be spacelike, therefore we cannot have $m(e_0, v) = 0$). Note that the vector

$$x = -e_0 - \frac{1}{m(e_0, v)} v \tag{6}$$

is orthogonal to e_0 and satisfies

$$\begin{aligned} m(x,x) &= m\big(e_0 + \frac{1}{m(e_0,v)}v, e_0 + \frac{1}{m(e_0,v)}v\big) \\ &= m(e_0,e_0) + \frac{2}{m(e_0,v)}m(e_0,v) + \frac{1}{(m(e_0,v))^2}m(v,v) \\ &= 1, \end{aligned}$$

i.e. the pair $\{e_0, x\}$ is orthonormal. Therefore, if we set

$$e_1 = x,$$

we can use the Gram–Schmidt process to extend $\{e_0, e_1\}$ to an orthonormal basis $\{e_\alpha\}_{\alpha=0}^n$ of V; since e_0 is timelike and m has signature $(1,n)$, the vectors $\{e_i\}_{i=1}^n$ are necessarily spacelike. Moreover, in view of (6)

$$v = -m(e_0,v)\big(e_0 + e_1\big),$$

i.e. in the $\{e_\alpha\}_{\alpha=0}^n$ basis v takes the form

$$v = (\lambda, \lambda, 0, \dots, 0),$$

where $\lambda = -m(e_0, v) > 0$.

Remark. In general, when confronted with calculations in some Lorentzian inner product space (i.e. the tangent space $T_p\mathcal{M}$ at a point p of a Lorentzian manifold $(\mathcal{M}, g)$), it is always useful to be able to choose an orthonormal basis adapted to the vectors in question; we can always choose an orthonormal basis where e_0 is parallel to a given timelike vector or, as shown here, $e_0 + e_1$ is parallel to a given null vector.

Let $w = (w^0, \dots, w^n) \in V \setminus 0$ be a vector such that $v \perp w$. We can then calculate (since $\{e_\alpha\}_{\alpha=0}^n$ is an orthonormal basis and hence, in this basis, $m = \mathrm{diag}(-1, +1, \dots, +1)$):

$$0 = m(v,w) = -v^0 w^0 + \sum_{i=1}^n v^i w^i = \lambda(-w^0 + w^1).$$

Thus, since $\lambda = -m(e_0, v) \neq 0$, we conclude

$$w^0 = w^1. \tag{7}$$

If the vector w is causal, i.e. $m(w,w) \leqslant 0$, then

$$\begin{aligned} 0 &\geqslant m(w,w) \\ &= -(w^0)^2 + \sum_{i=1}^n (w_i)^2 \\ &\overset{(7)}{=} \sum_{i=2}^n (w_i)^2 \end{aligned}$$

and, therefore,

$$w^2 = \ldots = w^n = 0.$$

Thus, w is of the form $w = (w^0, w^0, 0, \ldots, 0)$ and is therefore null and collinear with $v = (\lambda, \lambda, 0, \ldots, 0)$.

(b) As we have shown in class, any vector which is orthogonal to a timelike vextor is spacelike. Therefore, that none of the vectors v, w can be timelike (since then the other would have to be spacelike, i.e. non-causal). Therefore, v and w are null, so from part (a) of this exercise we infer that they are collinear.

(c) By our convention for a null vector, $v \neq 0$. Thus, the linear functional $v_\flat \doteq m(v, \cdot) : V \to 0$ cannot be identically zero (since m is non-degenerate), therefore its kernel (which is precisely $v^\perp$) is of codimension 1 (i.e. it is a hyperplane). Moreover, since v is null, $m(v, v) = 0$ and, thus, $v \in V^\perp$.

In order to show that $v^\perp$ is a *null* hyperplane, it remains to show that $m|_{v^\perp}$ is degenerate, i.e. that there exists a vector $L \neq 0$ in $v^\perp$ such that $m(L, x) = 0$ for all $x \in v^\perp$. It is clear from the definition of $v^\perp$ that $L = v$ has exactly this property.

Remark. In the basis $\{e_\alpha\}_{\alpha=0}^n$ constructed in part (a) of this exercise, where $v \parallel e_0 + e_1$, the space $v^\perp$ is spanned by the vectors $\{e_0 + e_1, e_2, \ldots, e_n\}$. Using those vectors as a basis for $v^\perp$, the associated matrix of the inner product $m|_{v^\perp}$ on $v^\perp$ takes the form $m = \text{diag}(0, +1, \ldots, +1)$.

1.3 Let $\mathcal{M}$ be a differentiable manifold of dimension n and $p \in \mathcal{M}$. Recall that the tangent space $T_p\mathcal{M}$ at p is defined as the set of all functionals $X : C^\infty(\mathcal{M}) \to \mathbb{R}$ satisfying the product rule

$$X(f \cdot g) = X(f) \cdot g(p) + f(p) \cdot X(g).$$

Prove that the set $T_p\mathcal{M}$ is a vector space of dimension n. (*Hint: Use the fact that, in any given local coordinate chart* $\phi : \mathcal{U} \to \phi(\mathcal{U}) \subset \mathbb{R}^n$ *on a neighborhood* $\mathcal{U}$ *around* p *with* $\phi(p) = 0$*, any smooth function* $f : \phi(\mathcal{U}) \to \mathbb{R}$ *can be expanded as* $f(x) = f(0) + A_a x^a + B_{ab}(x) x^a x^b$ *for constants* $\{A_a\}_{a=1}^n$ *and smooth functions* $\{B_{ab}(x)\}_{a,b=1}^n$*.*)

Remark. This exercise is in fact a standard theorem in the study of differentiable manifolds. In order to solve it properly, we will reprove a number of fundamental results from that field (such as the fact that $Z_p(f) = 0$ for any $Z_p \in T_p\mathcal{M}$ when the function $f \in C^\infty(\mathcal{M})$ is constant in an open neighborhood of the point $p \in \mathcal{M}$). The aim of the exercise is to remind you of those results; you should be able to use them without having to reprove them in the rest of the exercises of this course.

Solution. The fact that $T_p\mathcal{M}$ is a vector space follows easily by its definition: If $X, Y \in T_p\mathcal{M}$ and $\lambda \in \mathbb{R}$, then the linear functional $X + \lambda Y : C^\infty(\mathcal{M}) \to \mathbb{R}$, defined by

$$(X + \lambda Y)(f) \doteq X(f) + \lambda \cdot Y(f)$$

also belongs to $T_p\mathcal{M}$, i.e. it satisfies the product rule, since

$$\begin{aligned}
(X + \lambda Y)(f \cdot g) &= X(f \cdot g) + \lambda \cdot Y(f \cdot g) \\
&= X(f) \cdot g(p) + f(p) \cdot X(g) + \lambda Y(f) \cdot g(p) + \lambda \cdot f(p) \cdot Y(g) \\
&= (X + \lambda Y)(f) \cdot g(p) + f(p) \cdot (X + \lambda Y)(g).
\end{aligned}$$

Let $\phi : \mathcal{U} \to \phi(\mathcal{U}) \subset \mathbb{R}^n$ be a local coordinate chart around p, with associated coordinates $(x^1, \dots, x^n)$; recall that the coordinate functions $x^i : \mathcal{U} \to \mathbb{R}$ are defined as

$$x^i \doteq \bar{x}^i \circ \phi,$$

where $\bar{x}^i : \mathbb{R}^n \to \mathbb{R}$ are the Cartesian projections on the i-th coordinate. Let us also fix a smooth function $\chi : \mathcal{M} \to \mathbb{R}$ satisfying the following properties:

- $\chi(q) = 1$ in an open neighborhood of p,
- supp is compact and contained in $\mathcal{U}$.

Such a function can be readily constructed on $\phi(\mathcal{U}) \subset \mathbb{R}^n$, and then pulled-back to $\mathcal{U}$ via ϕ and extended to 0 on $\mathcal{M} \setminus \mathcal{U}$.

We will use the following fundamental results:

- If $f = c$ is a constant function on $\mathcal{M}$, then $Z_p(f) = 0$ for all $Z_p \in T_p\mathcal{M}$; this can be shown by arguing as follows: Using the fact that $Z_p : C^\infty(M) \to \mathbb{R}$ is a *linear* functional, it suffices to show that $Z_p(1) = 0$; using the product rule for Z_p, we can calculate

$$Z_p(1) = Z_p(1 \cdot 1) = Z_p(1) \cdot 1(p) + 1(p) \cdot Z_p(1) = 2Z_p(1) \quad \Leftrightarrow \quad Z_p(1) = 0.$$

- Let $\psi \in C^\infty(\mathcal{M})$ be such that $\psi(q) = 1$ for all q belonging to an open neighborhood $\mathcal{W}$ of p. Then, for all $Z_p \in T_p\mathcal{M}$:

$$Z_p(\psi) = 0.$$

  This can be shown by first introducing an auxiliary function $\chi' \in C^\infty(\mathcal{M})$ supported in $\mathcal{W} \cap \mathcal{U}$ and satisfying $\chi'(p) = 1$. We can then calculate (since supp$\chi' \subset \mathcal{W}$ and $\psi \equiv 1$ on $\mathcal{W}$)

$$\begin{aligned} Z_p(\chi') &= Z_p(\chi' \cdot \psi) \\ &= Z_p(\chi') \cdot \psi(p) + \chi'(p) \cdot Z_p(\psi) \\ &= Z_p(\chi') + Z_p(\psi) \\ \Rightarrow \quad & Z_p(\psi) = 0. \end{aligned}$$

- As a consequence of the previous result, if $f \in C^\infty(\mathcal{M})$ and $\chi : \mathcal{M} \to \mathbb{R}$ is the cut-off function introduced earlier, then we have for any $Z_p \in T_p\mathcal{M}$:

$$Z_p(f) = Z_p(\chi \cdot f + (1 - \chi) \cdot f) = Z_p(\chi \cdot f) \tag{8}$$

  since $1 - \chi$ vanishes in an open neighborhood of p and thus $1 - \chi(p) = 0 = Z_p(1 - \chi)$.

- A tangent vector $Z_p \in T_p\mathcal{M}$ can also be viewed as a linear function $Z_p : C^\infty(\mathcal{U}) \to \mathbb{R}$, namely as a functional on the space of smooth functions defined only on $\mathcal{U}$. This is because, for any $h \in C^\infty(\mathcal{U})$, we can define the function $\mathbb{E}_\chi h \in C^\infty(\mathcal{M})$ by the relation

$$\mathbb{E}_\chi h(q) = \begin{cases} \chi(q)h(q), & \text{when } q \in \mathcal{U}, \\ 0, & q \in \mathcal{M} \setminus \mathcal{U} \end{cases}$$

(Ex: show that supp$\chi \subset \mathcal{U}$ implies that $\mathbb{E}_\chi(h)$ is smooth on $\mathcal{M}$). Therefore, we can simply define for any $Z_p \in T_p\mathcal{M}$:

$$Z_p(h) \doteq Z_p(\mathbb{E}_\chi h).$$

Note that, as a consequence of the previous remarks, the value of $Z_p(h)$ for $h \in C^\infty(\mathcal{U})$ is *independent* of the choice of the cut-off function χ since, for any two functions χ_1, χ_2 which are both equal to 1 in an open neighborhood $\mathcal{W}$ of p, $\mathbb{E}_{\chi_1}h - \mathbb{E}_{\chi_2}h$ vanishes on $\mathcal{W}$ and, hence

$$Z_p(\mathbb{E}_{\chi_1}h - \mathbb{E}_{\chi_2}h) = 0.$$

Moreover, (8) implies that if $h = f|_{\mathcal{U}}$, then

$$Z_p(h) = Z_p(f).$$

Therefore,

$$Z_p(h) = 0 \text{ for all } h \in C^\infty(\mathcal{U}) \quad \Leftrightarrow \quad Z_p(f) = 0 \text{ for all } f \in C^\infty(\mathcal{M}).$$

We will now proceed to show that $\dim T_p\mathcal{M} = n$.

∘ We will first show that $\dim T_p\mathcal{M} \geqslant n$. To this end, it suffices to show that the coordinate tangent vectors $\left\{\frac{\partial}{\partial x^i}\right\}_{i=1}^n$ at p are linearly independent. Recall that $\frac{\partial}{\partial x^i}$ satisfies at p:

$$\frac{\partial}{\partial x^i}(x^j) = \delta_i^j.$$

Suppose, for the sake of contradiction, that $\left\{\frac{\partial}{\partial x^i}\right\}_{i=1}^n$ are linearly dependent at p, i.e. there exist constants $\lambda^1, \dots, \lambda^n \in \mathbb{R}$, not all identically zero, such that

$$\lambda^i \frac{\partial}{\partial x^i} = 0$$

(recall that repeated indices are assumed to be summed over their domain of definition, which here is $i \in \{1, \dots, n\}$). We then have, for any $j \in \{1, \dots, n\}$,

$$0 = \lambda^i \frac{\partial}{\partial x^i}(x^j) = \lambda^i \delta_i^j = \lambda^j.$$

Hence, all λ^j's have to vanish, which is a contradiction; the tangent vectors $\left\{\frac{\partial}{\partial x^i}\right\}_{i=1}^n$ are therefore linearly independent.

∘ We will now show that $\dim T_p\mathcal{M} \leqslant n$. To this end, it suffices to show that any $X \in T_p\mathcal{M}$ can be written as a linear combination of $\left\{\frac{\partial}{\partial x^i}\right\}_{i=1}^n$; in particular, we will show that

$$X = X(x^i) \cdot \frac{\partial}{\partial x^i},$$

or, equivalently, that

$$Y \doteq X - X(x^i)\cdot\frac{\partial}{\partial x^i} = 0. \tag{9}$$

Note that

$$Y(x^j) = X(x^j) - X(x^i)\cdot\frac{\partial}{\partial x^i}(x^j) = 0 \quad \text{for all } j\in\{1,\dots,n\}. \tag{10}$$

In order to establish (9), it suffices to show that

$$Y(f) = 0 \quad \text{for all } f\in C^\infty(\mathcal{U}). \tag{11}$$

Let $f\in C^\infty(\mathcal{U})$ and let us consider the function $f\circ\phi^{-1}:\phi(\mathcal{U})\subset\mathbb{R}^n\to\mathbb{R}$. Applying Taylor's expansion theorem for $f\circ\phi^{-1}$ around the point $\bar{p}=\phi(p)=(x^1(p),\dots,x^n(p))\in\mathbb{R}^n$, we can express express $f\circ\phi^{-1}$ as

$$f\circ\phi^{-1}(\bar{x}) = f\circ\phi^{-1}(p) + A_i(\bar{x}^i - x^i(p)) + B_{ij}(\bar{x})(\bar{x}^i - x^i(p))(\bar{x}^j - x^j(p)),$$

where $\bar{x}^i$ are the Cartesian coordinates on $\phi(\mathcal{U})\subset\mathbb{R}^n$, $\{A_i\}_{i=1}^n$ are constants and $\{B_{ij}(\cdot)\}_{i,j=1}^n$ are smooth functions on $\phi(\mathcal{U})$. Composing the above expression with ϕ, we obtain on $\mathcal{U}$:

$$f = f(p) + A_i(x^i - x^i(p)) + B_{ij}\circ\phi\cdot(x^i - x^i(p))(x^j - x^j(p)).$$

Applying the product rule and using the fact that $Y(c)=0$ for all constant functions c and $Y(x^i)=0$ for all the coordinate functions x^i (see (10)), we obtain

$$\begin{aligned}
Y(f) &= Y\big(f(p)\big) + Y\big(A_i(x^i - x^i(p))\big) + Y\big(B_{ij}\circ\phi\cdot(x^i - x^i(p))(x^j - x^j(p))\big)\\
&= Y\big(f(p)\big) + A_i\cdot\big(Y(x^i) - Y(x^i(p))\big) + Y(B_{ij}\circ\phi)\cdot(x^i(p) - x^i(p))(x^j(p) - x^j(p))\big)\\
&\qquad + B_{ij}\circ\phi(p)\cdot Y(x^i - x^i(p))(x^j(p) - x^j(p)) + B_{ij}\circ\phi(p)\cdot(x^i(p) - x^i(p))Y(x^j(p) - x^j(p))\\
&= 0.
\end{aligned}$$

Therefore, (11) holds.

1.4 Let $\mathcal{M}^n$ be a differentiable manifold and V be a smooth vector field on M. Assume that $V(p)\neq 0$ for some $p\in\mathcal{M}$. Show that there exists an open neighborhood $\mathcal{U}$ of p and a coordinate chart $(x^1,\dots,x^n)$ on $\mathcal{U}$ such that $V=\frac{\partial}{\partial x^1}$ in $\mathcal{U}$.

Solution. Let us start by fixing a coordinate chart $\phi':\mathcal{U}'\to\phi'(\mathcal{U}')\subset\mathbb{R}^n$ on an open neighborhood $\mathcal{U}'$ of p in $\mathcal{M}$. By composing ϕ' on the left with a translation $y\to y+y_0$ in $\mathbb{R}^n$, we can assume without loss of generality that $\phi'(p)=0$. Let $(y^1,\dots,y^n)$ be the local coordinate system on $\mathcal{U}'$ associated to ϕ (note that $y^i(p)=0$ for $i=1,\dots,n$). In this coordinate system, the vector field V can be expressed as

$$V = V^i\frac{\partial}{\partial y^i}.$$

Since $V(p)\neq 0$, at least one of the components $V^i(p)$ must ne non-zero; without loss of generality we can assume that $V^1(p)\neq 0$ (otherwise, we can simply relabel the coordinate functions). Since V is a smooth vector field, $V^1(p)\neq 0$ in an open neighborhood $\mathcal{W}$ of p.

We will construct the coordinate system $(x^1, \dots, x^n)$ by introducing an appopriate change of coordinates on a neighborhood of 0 in $\mathbb{R}^n$ and then pulling back these new coordinates to $\mathcal{M}$ via the chart ϕ'. More precisely, let $\Psi : \mathcal{V} \subset \mathbb{R}^n \to \mathcal{V}' \subset \phi'(\mathcal{U}')$ be a diffeomorphism between subsets of $\mathbb{R}^n$. Then, it is easy to verify that, in the local coordinate system $(x^1, \dots, x^n)$ on $(\phi')^{-1}(\mathcal{V}') \subset \mathcal{U}' \subset \mathcal{M}$ associated to the coordinate chart $\phi = \Psi^{-1} \circ (\phi')^{-1}$ on $(\phi')^{-1}(\mathcal{V})$,[1] the coordinate vector fields $\{\frac{\partial}{\partial x^i}\}_{i=1}^n$ can be expressed in terms of $\{\frac{\partial}{\partial y^i}\}_{i=1}^n$ by the relation

$$\frac{\partial}{\partial x^i} = \partial_i \Psi^j \cdot \frac{\partial}{\partial y^j}$$

(since the expression of the coordinates y^i as functions of x^i is $y^i = \Psi^i(x)$). Therefore, in order to construct a local coordinate system $(x^1, \dots, x^n)$ around p in which $V = \frac{\partial}{\partial x^1}$, it suffices to construct a smooth function $\Psi : \mathcal{W} \to \mathbb{R}^n$ for a domain $\mathcal{V} \subset \mathbb{R}^n$ containing 0 such that:

1. $\Psi(0) = 0$,
2. $D\Psi|_{x=0}$ is invertible,
3. $\partial_1 \Psi^i = V^i \circ (\phi')^{-1} \circ \Psi$ in an open neighborhood $\mathcal{V} \subset \mathcal{W}$ of 0.

In view of the inverse function theorem, Condition 2 above would imply that Ψ is a local diffeomorphism when restricted to a (possibly small) open neighborhood $\mathcal{V}$ of 0. Since $0 \in \phi'(\mathcal{U}')$ and $\Psi(0) = 0$ (according to Condition 1), by possibly choosing $\mathcal{V}$ even smaller, we can guarantee that $\Psi(\mathcal{V}) \subset \phi'(\mathcal{U}')$; hence $V^i \circ (\phi')^{-1} \circ \Psi$ (in the statement of Condition 3) would be a well defined function on $\mathcal{V}$.

In order to construct a local diffeomorphism Ψ as above, we will make use of the flow map associated to the vector field $\bar{V} = (V^1 \circ (\phi')^{-1}, \dots, V^n \circ (\phi')^{-1})$ on $\phi'(\mathcal{U}') \subset \mathbb{R}^n$ (note that this is simply the pushforward of the vector field V via the map ϕ'). For a smooth vector field $\bar{V}$ defined on an open domain Ω of $\mathbb{R}^n$, the classical theory of ODEs guarantees the existence of an open set $\overline{\Omega} \subset \mathbb{R} \times \Omega$ containing $\{0\} \times \Omega$ and a smooth map $\tilde{\Psi} : \overline{\Omega} \to \Omega$ such that

$$\begin{cases} \partial_t \tilde{\Psi}(t; \bar{x}) = \bar{V}(\tilde{\Psi}(t; \bar{x})), \\ \tilde{\Psi}(0; \bar{x}) = \bar{x}. \end{cases} \tag{12}$$

(this statetement can be equivalently stated in a more familiar language as follows: The initial value problem

$$\begin{cases} \partial_t x = \bar{V}(x), \\ x(0) = x_0 \in \Omega \end{cases}$$

admits a unique smooth solution $x[x_0, \cdot] : I_{x_0} \to \Omega$ on a maximal open interval $I_{x_0} \subseteq \mathbb{R}$ containing 0; moreover, $x[x_0, \cdot]$ and I_{x_0} depend smoothly on the initial value x_0.)

[1] Recall that, in this case, $x^i = (\Psi^{-1})^i \circ \phi'$; thus, $y^i = (\phi')^i = (\Psi \circ \Psi^{-1} \circ \phi')^i = (\Psi(x))^i$.

Let $\tilde{\Psi} : \overline{\Omega} \to \mathbb{R}^n$ be the map obtained by applying the above result with $\Omega = \phi'(\mathcal{U}')$. Let $\delta > 0$ be small enough so that $(-\delta, \delta) \times B_\delta[0] \subset \overline{\Omega}$ (where $B^{(n)}_\delta[0]$ is the Euclidean ball around $0 \in \mathbb{R}^n$ of radius δ). Let us consider the map $\Psi : (-\delta, \delta) \times B^{(n-1)}_\delta[0] \to \mathbb{R}^n$ defined by

$$\Psi(x^1, \ldots, x^n) = \tilde{\Psi}(x^1; 0, x^2, \ldots, x^n)$$

(this is simply the map that takes each point on the surface $\{\bar{x}^1 = 0\} \cap B_\delta[0]^{(n)}$ and maps it to its image under the flow of the vector field $\bar{V}$ for time $t = x^1$). In view of (12), we can readily compute:

1. $\Psi(0) = \tilde{\Psi}(0; 0) = 0$.
2. We can calculate at $(x^1, \ldots, x^n) = (0, \ldots, 0)$:

$$\partial_1 \Psi^j(0) = \partial_t \tilde{\Psi}^j(t; \bar{x}^1, \ldots, x^n)|_{(t;\bar{x}^1,\ldots,x^n)=(0;0,\ldots,0)} = \bar{V}^j(0) \quad \text{for } j = 1, \ldots, n$$

   and, for $i \geqslant 2$:

$$\begin{aligned} \partial_i \Psi^j(0) &= \partial_{\bar{x}^i} \tilde{\Psi}^j(t; \bar{x}^1, \ldots, x^n)|_{(t;\bar{x}^1,\ldots,x^n)=(0;0,\ldots,0)} \\ &= \delta^j_i. \end{aligned}$$

   Therefore, the matrix of the differential $D\Psi$ at 0 takes the form

$$[D\Psi]|_{x=0} = \begin{bmatrix} \bar{V}^1(0) & \bar{V}^2(0) & \ldots & \bar{V}^n(0) \\ 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \ldots & 1 \end{bmatrix},$$

   which is invertible since $\bar{V}^1(0) = V^1(p) \neq 0$.

3. We have everywhere on $(-\delta, \delta) \times B^{(n-1)}_\delta[0]$:

$$\begin{aligned} \partial_1 \Psi(x^1, \ldots, x^n) &= \partial_t \tilde{\Psi}(t; \bar{x}^1, \ldots, \bar{x}^n)|_{(t;\bar{x}^1,\bar{x}^2,\ldots,\bar{x}^n)=(x^1;0,x^2,\ldots,x^n)} \\ &= \bar{V}\big(\tilde{\Psi}(x^1; 0, x^2, \ldots, x^n)\big) \\ &= \bar{V}\big(\Psi(x^1, \ldots, x^n)\big) \end{aligned}$$

   and, hence

$$\partial_1 \Psi^i = \bar{V}^i \circ \Psi = V^i \circ (\phi')^{-1} \circ \Psi.$$

Therefore, setting $\mathcal{V} \doteq (-\delta, \delta) \times B^{(n-1)}_\delta[0]$, the map Ψ defined above satisfies Conditions 1–3; hence, as explained earlier, $\phi = \Psi^{-1} \circ \phi' : (\phi')^{-1}(\Psi(\mathcal{V})) \subset \mathcal{U}' \to \mathcal{V}$ is a coordinate chart around p in which

$$\frac{\partial}{\partial x^1} = V.$$

1.5 Let X, Y, Z be smooth vector fields on a differentiable manifold $\mathcal{M}$. We define the commutator (or *Lie bracket*) $[X, Y]$ of X and Y to be the vector field satisfying for any function $f \in C^\infty(\mathcal{M})$

$$[X, Y](f) = X(Y(f)) - Y(X(f)).$$

(a) Show that $[X, Y]$ satisfies the following identities:

1. $[X, Y] = -[Y, X]$ (*anticommutativity*).
2. $[X, aY + bZ] = a[X, Y] + b[X, Z]$ for any constans a, b (*linearity*).
3. $[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0$ (*Jacobi identity*).

(b) Let X^a and Y^b be the components of X and Y, respectively, in a local coordinate chart $(x^1, \ldots, x^n)$ on $\mathcal{M}$ (i.e. $X = X^a \frac{\partial}{\partial x^a}$ and $Y = Y^a \frac{\partial}{\partial x^a}$). Compute the components of $[X, Y]$ in the same coordinate chart.

Solution. (a) Let us first verify that indeed $[X, Y]$ is a vector field on $\mathcal{M}$ (recall that a vector field Z is simply an assignment $p \to Z_p$ for all $p \in \mathcal{M}$ such that $Z_p : C^\infty(\mathcal{M}) \to \mathbb{R}$ is a linear functional satisfying the product rule; note that, for any $f \in C^\infty(\mathcal{M})$, $Z(f)$ then defines a smooth function $p \to Z_p(f)$). To this end, we simply have to verify that, for any point $p \in \mathcal{M}$, the functional $[X, Y]_p : C^\infty(\mathcal{M}) \to \mathbb{R}$ defined by

$$[X, Y]_p(f) = X_p(Y(f)) - Y_p(X(f))$$

is linear (which is obvious) and satisfies the product rule. Indeed, for any $f, h \in C^\infty(\mathcal{M})$:

$$\begin{aligned}
[X, Y]_p(f \cdot h) &= X_p(Y(f \cdot h)) - Y_p(X(f \cdot h)) \\
&= X_p\big(Y(f) \cdot h + f \cdot Y(h)\big) - Y_p\big(X(f) \cdot h + f \cdot X(h)\big) \\
&= X_p(Y(f)) \cdot h(p) + Y_p(f) \cdot X_p(h) + X_p(f) Y_p(h) + f(p) X_p(Y(h)) \\
&\quad - Y_p(X(f)) \cdot h(p) - X_p(f) \cdot Y_p(h) - Y_p(f) X_p(h) - f(p) Y_p(X(h)) \\
&= \big(X_p(Y(f)) - Y_p(X(f))\big) \cdot h(p) + f(p) \cdot \big(X_p(Y(h)) - Y_p(X(h))\big),
\end{aligned}$$

where, above, we made use of the fact that the functionals $X, Y : C^\infty(\mathcal{M}) \to C^\infty(\mathcal{M})$ and $X_p, Y_p : C^\infty(\mathcal{M}) \to \mathbb{R}$ satisfy the product rule.

Identities 1–3 follow easily by using the definition of $[X, Y]$ and the fact that any vector field X on $\mathcal{M}$ defines a *linear* function $X : C^\infty(\mathcal{M}) \to C^\infty(\mathcal{M})$ satisfying the product rule $X(f \cdot g) = X(f) \cdot g + f \cdot X(g)$.

(b) We can readily calculate in the $(x^1, \ldots, x^n)$ coordinates:

$$\begin{aligned}
[X, Y]^i &= [X, Y](x^i) \\
&= X(Y(x^i)) - Y(X(x^i)) \\
&= X\Big(Y^j \frac{\partial x^i}{\partial x^j}\Big) - Y\Big(X^j \frac{\partial x^i}{\partial x^j}\Big) \\
&= X(Y^i) - Y(X^i) \\
&= X^j \frac{\partial Y^i}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j}.
\end{aligned}$$