

Chapter 9

Calculus on Euclidean Space

More than the algebraic operations on \mathbb{R}^n are needed for differential geometry. We need to know how to *differentiate* various geometric objects, and we need to know the relationship between differentiation and the algebraic operations. This chapter establishes the theoretical foundations of the theory of differentiation in \mathbb{R}^n that will be the basis for subsequent chapters.

In Section 9.1, we define the notion of *tangent vector* to \mathbb{R}^n . The interpretation of a tangent vector to \mathbb{R}^n as a directional derivative is given in Section 9.2. This allows us to define tangent maps as the derivatives of differentiable maps between Euclidean spaces \mathbb{R}^n and \mathbb{R}^m in Section 9.3.

It is important to understand how tangent vectors vary from point to point, so we begin the study of *vector fields* on \mathbb{R}^n in Section 9.4. Derivatives of vector fields on \mathbb{R}^n with respect to tangent vectors are considered in Section 9.5, in which we also prove some elementary facts about the linear map on vector fields induced by a diffeomorphism. The classical notions of gradient, divergence and Laplacian, are defined in Sections 9.4 and 9.5. Finally, in Section 9.6 we return to our study of curves, using tangent vectors to provide a different perspective and make more rigorous some previous definitions.

9.1 Tangent Vectors to \mathbb{R}^n

In Section 1.1, we defined the notion of a *vector* \mathbf{v} in \mathbb{R}^n . So far, we have generally considered vectors in \mathbb{R}^n to be *points* in \mathbb{R}^n ; however, a vector \mathbf{v} can be regarded as an arrow or *displacement* which starts at the origin and ends at the point \mathbf{v} . From this point of view, and with the theory of surfaces in mind, it makes sense to study displacements that start at points other than the origin, and to retain the starting point in the definition of the tangent vector.

Definition 9.1. A **tangent vector** $\mathbf{v}_{\mathbf{p}}$ to Euclidean space \mathbb{R}^n consists of a pair of elements \mathbf{v}, \mathbf{p} of \mathbb{R}^n ; \mathbf{v} is called the **vector part** and \mathbf{p} is called the **point of application** of $\mathbf{v}_{\mathbf{p}}$.

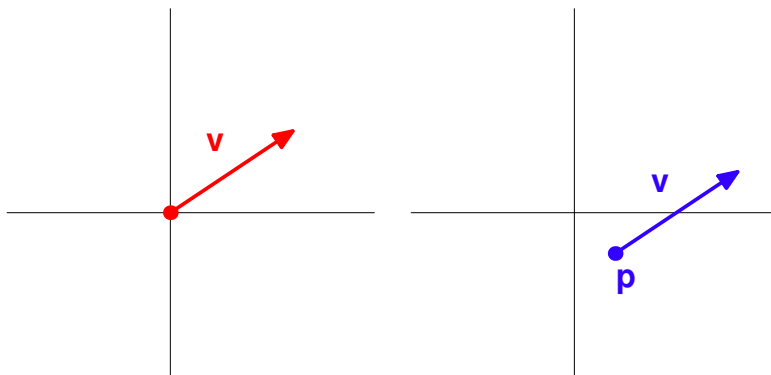


Figure 9.1: Vectors applied at different points

The tangent vector can be identified with the arrow from \mathbf{p} to $\mathbf{p} + \mathbf{v}$. Furthermore, two tangent vectors $\mathbf{v}_{\mathbf{p}}$ and $\mathbf{w}_{\mathbf{q}}$ are declared to be equal if and only if *both* their vector parts and points of application coincide: $\mathbf{v} = \mathbf{w}$ and $\mathbf{p} = \mathbf{q}$.

Definition 9.2. Let $\mathbf{p} \in \mathbb{R}^n$. The **tangent space** of \mathbb{R}^n at \mathbf{p} is the set

$$\mathbb{R}_{\mathbf{p}}^n = \{ \mathbf{v}_{\mathbf{p}} \mid \mathbf{v} \in \mathbb{R}^n \}.$$

The tangent space $\mathbb{R}_{\mathbf{p}}^n$ is a carbon copy of the vector space \mathbb{R}^n ; in fact, there is a canonical isomorphism between $\mathbb{R}_{\mathbf{p}}^n$ and \mathbb{R}^n given by $\mathbf{v}_{\mathbf{p}} \mapsto \mathbf{v}$. This isomorphism allows us to turn $\mathbb{R}_{\mathbf{p}}^n$ into a vector space; we can add elements in $\mathbb{R}_{\mathbf{p}}^n$ and multiply by scalars in the obvious way, retaining the common point of application:

$$\begin{aligned} \mathbf{v}_{\mathbf{p}} + \mathbf{w}_{\mathbf{p}} &= (\mathbf{v} + \mathbf{w})_{\mathbf{p}}, \\ \lambda \mathbf{v}_{\mathbf{p}} &= (\lambda \mathbf{v})_{\mathbf{p}}. \end{aligned}$$

To sum up, \mathbb{R}^n has a tangent space $\mathbb{R}_{\mathbf{p}}^n$ attached to each of its points, and the tangent space looks like \mathbb{R}^n itself. For this reason, it is possible to transfer the dot product and norm to each tangent space. We simply put

$$\mathbf{v}_{\mathbf{p}} \cdot \mathbf{w}_{\mathbf{p}} = \mathbf{v} \cdot \mathbf{w}$$

for $\mathbf{v}_{\mathbf{p}}, \mathbf{w}_{\mathbf{p}} \in \mathbb{R}_{\mathbf{p}}^n$. It follows that there is a Cauchy–Schwarz inequality for tangent vectors, namely,

$$|\mathbf{v}_{\mathbf{p}} \cdot \mathbf{w}_{\mathbf{p}}| \leq \|\mathbf{v}_{\mathbf{p}}\| \|\mathbf{w}_{\mathbf{p}}\|$$

for $\mathbf{v}_p, \mathbf{w}_p \in \mathbb{R}_p^n$. Angles between tangent vectors can now be defined in exactly the same way as angles between ordinary vectors (see page 3).

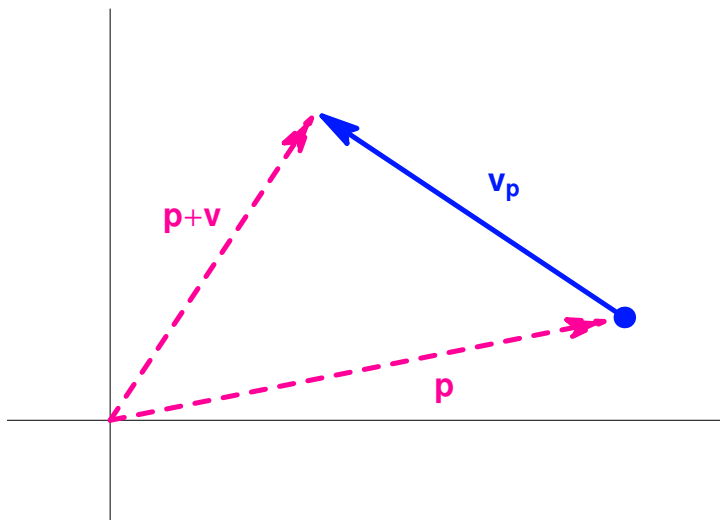


Figure 9.2: A tangent vector \mathbf{v}_p at \mathbf{p}

To finish this introductory section, consider tangent vectors to \mathbb{R}^2 and \mathbb{R}^3 . The complex structure of \mathbb{R}^2 that we defined on page 3 can be naturally extended to tangent vectors by putting

$$J\mathbf{v}_p = (J\mathbf{v})_p$$

for a tangent vector \mathbf{v}_p to \mathbb{R}^2 . It is easy to see that $J\mathbf{v}_p$ is also a tangent vector to \mathbb{R}^2 and that $-J^2$ is the identity map on tangent vectors. Furthermore, $J\mathbf{v}_p \cdot J\mathbf{w}_p = \mathbf{v}_p \cdot \mathbf{w}_p$ for $\mathbf{v}_p, \mathbf{w}_p \in \mathbb{R}_p^2$. Similarly, on \mathbb{R}^3 we extend the vector cross product to tangent vectors via

$$\mathbf{v}_p \times \mathbf{w}_p = (\mathbf{v} \times \mathbf{w})_p,$$

for $\mathbf{v}_p, \mathbf{w}_p \in \mathbb{R}_p^3$. The usual identities, such as Lagrange's

$$\|\mathbf{v}_p \times \mathbf{w}_p\|^2 = \|\mathbf{v}_p\|^2 \|\mathbf{w}_p\|^2 - (\mathbf{v}_p \cdot \mathbf{w}_p)^2$$

(see page 193) will now hold for tangent vectors.

9.2 Tangent Vectors as Directional Derivatives

Frequently, it is important to know how a real-valued function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ varies in different directions. For example, if we write the values of f as $f(u_1, \dots, u_n)$,

the partial derivative $\partial f / \partial u_i$ measures how much f varies in the u_i direction, for $i = 1, \dots, n$. However, it is possible to measure the variation of f in other directions as well. To measure the variation of f at $\mathbf{p} \in \mathbb{R}^n$ along the straight line $t \mapsto \mathbf{p} + t\mathbf{v}$, we need to make the tangent vector $\mathbf{v}_{\mathbf{p}}$ operate on functions in a suitable way.

Throughout this section, the words ‘ f is a differentiable function’ or ‘ f is differentiable’ will mean that f has partial derivatives of all orders on its domain. For details on why this is a reasonable definition, see the introduction to Chapter 24.

Definition 9.3. Let $f: \mathbb{R}^n \rightarrow \mathbb{R}$ be a differentiable function, and let $\mathbf{v}_{\mathbf{p}}$ be a tangent vector to \mathbb{R}^n at $\mathbf{p} \in \mathbb{R}^n$. We put

$$(9.1) \quad \mathbf{v}_{\mathbf{p}}[f] = \left. \frac{d}{dt} f(\mathbf{p} + t\mathbf{v}) \right|_{t=0}.$$

In elementary calculus, (9.1) goes under the name of *directional derivative in the $\mathbf{v}_{\mathbf{p}}$ direction* (at least if $\|\mathbf{v}_{\mathbf{p}}\| = 1$), but we prefer to consider (9.1) as defining how a tangent vector operates on functions.

A more explicit way to write (9.1) is

$$\mathbf{v}_{\mathbf{p}}[f] = \lim_{t \rightarrow 0} \frac{f(\mathbf{p} + t\mathbf{v}) - f(\mathbf{p})}{t}.$$

For example, let $n = 2$, $f(x, y) = \sin(xy)$, $\mathbf{p} = (1, 2)$ and $\mathbf{v} = (2, 3)$. Then

$$\begin{aligned} f(\mathbf{p} + t\mathbf{v}) &= f((1, 2) + t(2, 3)) &= f(2t + 1, 3t + 2) \\ &= \sin((2t + 1)(3t + 2)) \\ &= \sin(6t^2 + 7t + 2), \end{aligned}$$

so that

$$\frac{d}{dt} f(\mathbf{p} + t\mathbf{v}) = (12t + 7) \cos(6t^2 + 7t + 2).$$

Thus

$$\mathbf{v}_{\mathbf{p}}[f] = (12t + 7) \cos(6t^2 + 7t + 2) \Big|_{t=0} = 7 \cos 2 = -2.913.$$

Although it is possible to compute $\mathbf{v}_{\mathbf{p}}[f]$ directly from the definition, it is usually easier to use a general formula, which we now derive.

Lemma 9.4. Let $\mathbf{v}_{\mathbf{p}} = (v_1, \dots, v_n)_{\mathbf{p}}$ be a tangent vector to \mathbb{R}^n , and $f: \mathbb{R}^n \rightarrow \mathbb{R}$ a differentiable function. Then

$$(9.2) \quad \mathbf{v}_{\mathbf{p}}[f] = \sum_{j=1}^n v_j \frac{\partial f}{\partial u_j}(\mathbf{p}).$$

Proof. Let $\mathbf{p} = (p_1, \dots, p_n)$, so that $\mathbf{p} + t\mathbf{v} = (p_1 + tv_1, \dots, p_n + tv_n)$. The chain rule implies that

$$\begin{aligned} \frac{d}{dt}(f(\mathbf{p} + t\mathbf{v})) &= \frac{d}{dt}f(p_1 + tv_1, \dots, p_n + tv_n) \\ &= \sum_{j=1}^n \frac{\partial f}{\partial u_j}(p_1 + tv_1, \dots, p_n + tv_n) \frac{d(p_j + tv_j)}{dt} \\ &= \sum_{j=1}^n \frac{\partial f}{\partial u_j}(p_1 + tv_1, \dots, p_n + tv_n) v_j. \end{aligned}$$

When we put $t = 0$ in this equation, we get (9.2). ■

Next, we list some algebraic properties of the operation (9.1).

Lemma 9.5. *Let $f, g: \mathbb{R}^n \rightarrow \mathbb{R}$ be differentiable functions, $\mathbf{v}_\mathbf{p}, \mathbf{w}_\mathbf{p}$ tangent vectors in $\mathbb{R}_\mathbf{p}^n$, and a, b real numbers. Then*

$$(9.3) \quad (a\mathbf{v}_\mathbf{p} + b\mathbf{w}_\mathbf{p})[f] = a\mathbf{v}_\mathbf{p}[f] + b\mathbf{w}_\mathbf{p}[f],$$

$$(9.4) \quad \mathbf{v}_\mathbf{p}[af + bg] = a\mathbf{v}_\mathbf{p}[f] + b\mathbf{v}_\mathbf{p}[g],$$

$$(9.5) \quad \mathbf{v}_\mathbf{p}[fg] = f(\mathbf{p})\mathbf{v}_\mathbf{p}[g] + g(\mathbf{p})\mathbf{v}_\mathbf{p}[f].$$

Proof. For example, to prove (9.5):

$$\begin{aligned} \mathbf{v}_\mathbf{p}[fg] &= \left. \frac{d}{dt}(fg)(\mathbf{p} + t\mathbf{v}) \right|_{t=0} \\ &= \left(f(\mathbf{p} + t\mathbf{v}) \frac{d}{dt}g(\mathbf{p} + t\mathbf{v}) + g(\mathbf{p} + t\mathbf{v}) \frac{d}{dt}f(\mathbf{p} + t\mathbf{v}) \right) \Big|_{t=0} \\ &= f(\mathbf{p})\mathbf{v}_\mathbf{p}[g] + g(\mathbf{p})\mathbf{v}_\mathbf{p}[f]. \quad \blacksquare \end{aligned}$$

There is a chain rule for tangent vectors:

Lemma 9.6. *Let $g_1, \dots, g_k: \mathbb{R}^n \rightarrow \mathbb{R}$ and $h: \mathbb{R}^k \rightarrow \mathbb{R}$ be differentiable functions and write $f = h \circ (g_1, \dots, g_k)$ for the composed function $\mathbb{R}^n \rightarrow \mathbb{R}$. Let $\mathbf{v}_\mathbf{p} \in \mathbb{R}_\mathbf{p}^n$ where $\mathbf{p} \in \mathbb{R}^n$. Then*

$$\mathbf{v}_\mathbf{p}[f] = \sum_{j=1}^k \frac{\partial h}{\partial u_j}(g_1(\mathbf{p}), \dots, g_k(\mathbf{p})) \mathbf{v}_\mathbf{p}[g_j].$$

Proof. Let $\mathbf{v}_\mathbf{p} = (v_1, \dots, v_n)_\mathbf{p}$. We use the ordinary chain rule to compute

$$\begin{aligned} \mathbf{v}_\mathbf{p}[f] &= \sum_{i=1}^n v_i \frac{\partial f}{\partial u_i}(\mathbf{p}) = \sum_{i=1}^n \sum_{j=1}^k v_i \frac{\partial h}{\partial u_j}(g_1(\mathbf{p}), \dots, g_k(\mathbf{p})) \frac{\partial g_j}{\partial u_i}(\mathbf{p}) \\ &= \sum_{j=1}^k \frac{\partial h}{\partial u_j}(g_1(\mathbf{p}), \dots, g_k(\mathbf{p})) \mathbf{v}_\mathbf{p}[g_j]. \quad \blacksquare \end{aligned}$$

9.3 Tangent Maps or Differentials

We shall need to deal with functions $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$ other than linear maps. For any such function, we can write

$$(9.6) \quad F(\mathbf{p}) = (f_1(\mathbf{p}), \dots, f_m(\mathbf{p})),$$

where $f_j: \mathbb{R}^n \rightarrow \mathbb{R}$ for $j = 1, \dots, m$, and $\mathbf{p} \in \mathbb{R}^n$. It will be convenient to use the notation $F = (f_1, \dots, f_m)$ as a shorthand for (9.6).

Clearly, (9.6) also makes sense for a function $F: \mathcal{U} \rightarrow \mathbb{R}^m$ if \mathcal{U} is an open subset of \mathbb{R}^n . Recall that a subset \mathcal{U} of \mathbb{R}^n is said to be **open** if for any $\mathbf{p} \in \mathcal{U}$ it is possible to ‘squeeze’ a small ball with center \mathbf{p} into \mathcal{U} , meaning that there exists $\varepsilon > 0$ such that $\{\mathbf{q} \in \mathbb{R}^n \mid \|\mathbf{q} - \mathbf{p}\| < \varepsilon\} \subseteq \mathcal{U}$.

Definition 9.7. Let $\mathcal{U} \subseteq \mathbb{R}^n$ be open. A function $F: \mathcal{U} \rightarrow \mathbb{R}^m$ is said to be **differentiable** provided that each f_j is differentiable. More generally, if A is any subset of \mathbb{R}^n , we say that $F: A \rightarrow \mathbb{R}^m$ is differentiable if there exists an open set \mathcal{U} containing A and a differentiable map $\tilde{F}: \mathcal{U} \rightarrow \mathbb{R}^m$ such that the restriction of \tilde{F} to A is F . A **diffeomorphism** F between open subsets \mathcal{U} and \mathcal{V} of \mathbb{R}^n is a differentiable map $F: \mathcal{U} \rightarrow \mathcal{V}$ possessing an inverse $F^{-1}: \mathcal{V} \rightarrow \mathcal{U}$ that is also differentiable.

It is easy to prove that the composition of differentiable functions is differentiable:

Lemma 9.8. Suppose that there are differentiable maps

$$\begin{array}{ccccc} \mathcal{U} & \xrightarrow{F} & \mathcal{V} & \xrightarrow{G} & \mathbb{R}^n \\ \cap & & \cap & & \\ \mathbb{R}^\ell & & \mathbb{R}^m & & \end{array}$$

Then the composition $G \circ F: \mathcal{U} \rightarrow \mathbb{R}^n$ is differentiable.

Any differentiable map $F: \mathcal{U} \rightarrow \mathbb{R}^m$ gives rise in a natural way to a linear map between tangent spaces. To establish this, first note that if $F: \mathcal{U} \rightarrow \mathbb{R}^m$ is differentiable, then so is $t \mapsto F(\mathbf{p} + t\mathbf{v})$ for $\mathbf{p} + t\mathbf{v} \in \mathcal{U}$, because it is the composition of differentiable functions. Therefore, the following makes sense:

Definition 9.9. Let $F: \mathcal{U} \rightarrow \mathbb{R}^m$ be a differentiable map, where \mathcal{U} is an open subset of \mathbb{R}^n , and let $\mathbf{p} \in \mathcal{U}$. For each tangent vector $\mathbf{v}_{\mathbf{p}}$ to \mathbb{R}^n set

$$F_{*\mathbf{p}}(\mathbf{v}_{\mathbf{p}}) = F(\mathbf{p} + t\mathbf{v})'(0).$$

Then $F_{*\mathbf{p}}: \mathbb{R}_{\mathbf{p}}^n \rightarrow \mathbb{R}_{F(\mathbf{p})}^m$ is called the **tangent map** or **differential** of F at \mathbf{p} .

Note that $F_{*\mathbf{p}}(\mathbf{v}_{\mathbf{p}})$ is the initial velocity of the curve $t \mapsto F(\mathbf{p} + t\mathbf{v})$.

One often abbreviates $F_{*\mathbf{p}}$ to F_* if the identity of the point in question is clear. For example, let $\mathbf{p} = (1, 2, 3)$, $\mathbf{v} = (1, 0, -1)$ and $F(x, y, z) = (yz, zx, xy)$. Then

$$F(\mathbf{p} + t\mathbf{v}) = (6 - 2t, 3 + 2t - t^2, 2 + 2t),$$

so that

$$F_*(\mathbf{v}_{\mathbf{p}}) = \left(\frac{d}{dt}(6 - 2t, 3 + 2t - t^2, 2 + 2t) \Big|_{t=0} \right)_{F(\mathbf{p})} = (-2, 2, 2)_{(6, 3, 2)}.$$

The following lemma provides a useful way of computing tangent maps.

Lemma 9.10. *Let $F: \mathcal{U} \rightarrow \mathbb{R}^m$ be a differentiable map defined on an open subset $\mathcal{U} \subseteq \mathbb{R}^n$, and write $F = (f_1, \dots, f_m)$. If $\mathbf{p} \in \mathcal{U}$ and $\mathbf{v}_{\mathbf{p}}$ is a tangent vector to \mathbb{R}^n at \mathbf{p} , then*

$$(9.7) \quad F_*(\mathbf{v}_{\mathbf{p}}) = (\mathbf{v}_{\mathbf{p}}[f_1], \dots, \mathbf{v}_{\mathbf{p}}[f_m])_{F(\mathbf{p})}.$$

Proof. Replace \mathbf{p} by $\mathbf{p} + t\mathbf{v}$ in (9.6) and differentiate with respect to t :

$$(9.8) \quad F(\mathbf{p} + t\mathbf{v})' = (f_1(\mathbf{p} + t\mathbf{v})', \dots, f_m(\mathbf{p} + t\mathbf{v})').$$

When we evaluate both sides of (9.8) at $t = 0$ and use (9.1), we obtain (9.7). ■

The next result is an easy consequence of (9.7) and (9.3):

Corollary 9.11. *If $F: \mathcal{U} \rightarrow \mathbb{R}^m$ is differentiable, then at each point $\mathbf{p} \in \mathcal{U}$ the tangent map $F_{*\mathbf{p}}: \mathbb{R}_{\mathbf{p}}^n \rightarrow \mathbb{R}_{F(\mathbf{p})}^m$ is linear.*

The tangent map of a differentiable map F can be thought of as the *best linear approximation* to F at a given point \mathbf{p} . We have already seen examples of tangent maps; for example, the velocity of a curve is a tangent map in disguise. To explain this, we let $1_{\mathbf{p}}$ be the tangent vector to \mathbb{R} whose vector part is 1 and whose point of application is \mathbf{p} . This makes perfect logical sense, and it follows from the definitions that

$$(9.9) \quad 1_{\mathbf{p}}[f] = \frac{d}{dt}f(\mathbf{p} + t1) \Big|_{t=0} = f'(\mathbf{p})$$

for any differentiable function $f: \mathbb{R} \rightarrow \mathbb{R}$.

Lemma 9.12. *Let $\alpha: (a, b) \rightarrow \mathbb{R}^n$ be a curve. Then for $\mathbf{p} \in (a, b)$ we have*

$$\alpha_*(1_{\mathbf{p}}) = \alpha'(\mathbf{p}).$$

Proof. If we write $\alpha = (a_1, \dots, a_n)$, then from (9.9) and (9.7) we get

$$\alpha_*(1_{\mathbf{p}}) = (1_{\mathbf{p}}[a_1], \dots, 1_{\mathbf{p}}[a_n]) = (a'_1(\mathbf{p}), \dots, a'_n(\mathbf{p})) = \alpha'(\mathbf{p}). \blacksquare$$

If $\alpha: (a, b) \rightarrow \mathcal{U} \subseteq \mathbb{R}^n$ is a differentiable curve, and $F: \mathcal{U} \rightarrow \mathbb{R}^m$ a differentiable map, then $F \circ \alpha: (a, b) \rightarrow \mathbb{R}^m$ is a differentiable curve by Lemma 9.8. It is reasonable to suspect that F_* maps the velocity of α into the velocity of $F \circ \alpha$. We show that this is indeed the case.

Lemma 9.13. *Let $F: \mathcal{U} \rightarrow \mathbb{R}^m$ be a differentiable map, where \mathcal{U} is an open subset of \mathbb{R}^n , and let $\alpha: (a, b) \rightarrow \mathcal{U}$ be a curve. Then*

$$(F \circ \alpha)'(t) = F_*(\alpha'(t))$$

for $a < t < b$.

Proof. Write $F = (f_1, \dots, f_m)$ and $\alpha = (a_1, \dots, a_n)$. Lemmas 9.4, 9.10 and the chain rule imply that

$$\begin{aligned} F_*(\alpha'(t)) &= (\alpha'(t)[f_1], \dots, \alpha'(t)[f_m]) \\ &= \left(\sum_{j=1}^n a'_j(t) \frac{\partial f_1}{\partial u_j}, \dots, \sum_{j=1}^n a'_j(t) \frac{\partial f_m}{\partial u_j} \right)_{\alpha(t)} \\ &= ((f_1 \circ \alpha)'(t), \dots, (f_m \circ \alpha)'(t)) \\ &= (F \circ \alpha)'(t). \blacksquare \end{aligned}$$

Corollary 9.14. *Let $F: \mathcal{U} \rightarrow \mathbb{R}^m$ be a differentiable map, where \mathcal{U} is an open subset of \mathbb{R}^n . Let $\mathbf{p} \in \mathcal{U}$, $\mathbf{v}_{\mathbf{p}} \in \mathbb{R}_{\mathbf{p}}^n$, and let $\alpha: (a, b) \rightarrow \mathcal{U}$ be a curve such that $\alpha'(0) = \mathbf{v}_{\mathbf{p}}$. Then*

$$F_{*\mathbf{p}}(\mathbf{v}_{\mathbf{p}}) = (F \circ \alpha)'(0).$$

We originally defined $F_*(\mathbf{v}_{\mathbf{p}}) = F_{*\mathbf{p}}(\mathbf{v}_{\mathbf{p}})$ to be the velocity vector at 0 of the image under F of the straight line $t \rightarrow \mathbf{p} + t\mathbf{v}$. Corollary 9.14 says that $F_*(\mathbf{v}_{\mathbf{p}})$ equals the velocity vector at 0 of the image under F of *any* curve α such that $\alpha(0) = \mathbf{p}$ and $\alpha'(0) = \mathbf{v}_{\mathbf{p}}$.

Corollary 9.15. *Suppose that there are differentiable maps*

$$\begin{array}{ccccc} \mathcal{U} & \xrightarrow{F} & \mathcal{V} & \xrightarrow{G} & \mathbb{R}^n \\ \cap & & \cap & & \\ \mathbb{R}^\ell & & \mathbb{R}^m & & \end{array}$$

Then $(G \circ F)_* = G_* \circ F_*$.

Proof. Let $\mathbf{v}_{\mathbf{p}}$ be a tangent vector to \mathbb{R}^{ℓ} at $\mathbf{p} \in \mathbb{R}^{\ell}$. Then $t \mapsto F(\mathbf{p} + t\mathbf{v})$ is a curve in \mathbb{R}^m such that $F_*(\mathbf{v}_{\mathbf{p}}) = F(\mathbf{p} + t\mathbf{v})'(0)$. Hence

$$\begin{aligned}(G \circ F)_*(\mathbf{v}_{\mathbf{p}}) &= ((G \circ F)(\mathbf{p} + t\mathbf{v}))'(0) = G'(F(\mathbf{p} + t\mathbf{v}))'(0) \\ &= G_*(F(\mathbf{p} + t\mathbf{v})'(0)) \\ &= G_*(F_*(\mathbf{v}_{\mathbf{p}})). \blacksquare\end{aligned}$$

Recall from Chapter 1 that if $L: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a linear map, then the matrix associated with L is the $m \times n$ matrix $A = (a_{ij})$ such that

$$L(\mathbf{v}) = A\mathbf{v}$$

for all $\mathbf{v} \in \mathbb{R}^n$. (Here, for clarity, we use different symbols for the map and the matrix.) Let $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be the standard basis of \mathbb{R}^n , that is, \mathbf{e}_i is the n -tuple with 1 in the i^{th} spot and zeros elsewhere; similarly, let $\{\mathbf{e}'_1, \dots, \mathbf{e}'_m\}$ be the standard basis of \mathbb{R}^m . Then the formula we are adopting to relate A to L is

$$L(\mathbf{e}_j) = \sum_{i=1}^m a_{ij} \mathbf{e}'_i$$

for $j = 1, \dots, n$. Thus, it is the j^{th} column of A that represents the image of the j^{th} basis element.

If $F: \mathcal{U} \rightarrow \mathbb{R}^m$ is a differentiable map, where \mathcal{U} is an open subset of \mathbb{R}^n , then for each $\mathbf{p} \in \mathcal{U}$ the tangent map $L = F_{*\mathbf{p}}$ is a linear map between the vector spaces $\mathbb{R}_{\mathbf{p}}^n$ and $\mathbb{R}_{F(\mathbf{p})}^m$. Since $\mathbb{R}_{\mathbf{p}}^n$ is canonically isomorphic to \mathbb{R}^n , and $\mathbb{R}_{F(\mathbf{p})}^m$ is \mathbb{R}^m , we can associate to $F_{*\mathbf{p}}$ an $m \times n$ matrix A , which has the property that $F_{*\mathbf{p}}(\mathbf{v}_{\mathbf{p}}) = A\mathbf{v}_{\mathbf{p}}$ for all $\mathbf{v}_{\mathbf{p}} \in \mathbb{R}_{\mathbf{p}}^n$. More precisely,

Definition 9.16. Let $F: \mathcal{U} \rightarrow \mathbb{R}^m$ be a differentiable map, where \mathcal{U} is an open subset of \mathbb{R}^n , and write $F = (f_1, \dots, f_m)$. The **Jacobian matrix**¹ is the matrix-valued function $\mathcal{J}(F)$ of F given by

$$\mathcal{J}(F)(\mathbf{p}) = \begin{pmatrix} \frac{\partial f_1}{\partial u_1}(\mathbf{p}) & \cdots & \frac{\partial f_1}{\partial u_n}(\mathbf{p}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial u_1}(\mathbf{p}) & \cdots & \frac{\partial f_m}{\partial u_n}(\mathbf{p}) \end{pmatrix}$$

for $\mathbf{p} \in \mathcal{U}$.



Carl Gustav Jacob Jacobi (1804–1851). Professor of Mathematics at Königsberg and Berlin. He is remembered for his work on dynamics and elliptic functions. His *Fundamenta Nova Theoria Functionum Ellipticarum*, in which elliptic function theory is based on four theta functions, was published in 1829.

Recall that the **rank** of an $m \times n$ matrix A can be defined as the dimension ρ of the space spanned by the columns of A . A fundamental theorem of linear algebra asserts that it is also equal to the dimension of the space spanned by the rows of A , and is therefore no larger than the minimum of m and n . Accordingly, if ρ equals $\min(m, n)$, then A is said to have **maximum rank**. The situation described by the next result can only occur when $m \geq n$.

Lemma 9.17. *Let $F: \mathcal{U} \rightarrow \mathbb{R}^m$ be a differentiable map, where \mathcal{U} is an open subset of \mathbb{R}^n . Then each tangent map $F_{*\mathbf{p}}: \mathbb{R}_{\mathbf{p}}^n \rightarrow \mathbb{R}_{F(\mathbf{p})}^m$ is injective if and only if $\mathcal{J}(F)(\mathbf{p})$ has rank equal to n .*

Proof. Let $\mathbf{v}_{\mathbf{p}} \in \mathbb{R}_{\mathbf{p}}^n$ and write $\mathbf{v} = (v_1, \dots, v_n)$; also put $F = (f_1, \dots, f_m)$. Suppose that $F_{*}(\mathbf{v}_{\mathbf{p}}) = 0$ and $\mathbf{v} \neq 0$. Then Lemma 9.10 implies that

$$\mathbf{v}_{\mathbf{p}}[f_1] = \dots = \mathbf{v}_{\mathbf{p}}[f_m] = 0,$$

and by Lemma 9.4 we have

$$\sum_{i=1}^n v_i \frac{\partial f_j}{\partial u_i}(\mathbf{p}) = 0$$

for $j = 1, \dots, m$. It follows that

$$\sum_{i=1}^n v_i \left(\frac{\partial f_1}{\partial u_i}(\mathbf{p}), \dots, \frac{\partial f_m}{\partial u_i}(\mathbf{p}) \right) = 0.$$

Therefore, the n columns of the Jacobian matrix $\mathcal{J}(F)(\mathbf{p})$ are linearly dependent, and consequently $\mathcal{J}(F)(\mathbf{p})$ has rank less than n .

Conversely, if the rank of $\mathcal{J}(F)(\mathbf{p})$ is less than n , the steps in the above proof can be reversed to conclude that $F_{*}(\mathbf{v}_{\mathbf{p}}) = 0$ for some nonzero $\mathbf{v}_{\mathbf{p}} \in \mathbb{R}_{\mathbf{p}}^n$. ■

9.4 Vector Fields on \mathbb{R}^n

It will be necessary to consider tangent vectors as they vary from point to point.

Definition 9.18. *A **vector field** \mathbf{V} on an open subset \mathcal{U} of \mathbb{R}^n is a function that assigns to each $\mathbf{p} \in \mathcal{U}$ a tangent vector $\mathbf{V}(\mathbf{p}) \in \mathbb{R}_{\mathbf{p}}^n$. If $f: \mathcal{U} \rightarrow \mathbb{R}$ is differentiable, we let \mathbf{V} act on f via*

$$\mathbf{V}[f](\mathbf{p}) = \mathbf{V}(\mathbf{p})[f].$$

*The vector field \mathbf{V} is said to be **differentiable** provided that $\mathbf{V}[f]: \mathcal{U} \rightarrow \mathbb{R}$ is differentiable whenever f is.*

It is important to distinguish between the i^{th} coordinate of a point $\mathbf{p} \in \mathbb{R}^n$ and the *function* that assigns to \mathbf{p} its i^{th} coordinate:

Definition 9.19. The **natural coordinate functions** of \mathbb{R}^n are the functions u_i defined by

$$u_i(\mathbf{p}) = p_i, \quad i = 1, \dots, n,$$

for $\mathbf{p} = (p_1, \dots, p_n)$.

In the special cases \mathbb{R} or \mathbb{R}^2 or \mathbb{R}^3 , we shall often denote the natural coordinate functions by symbols such as t or u, v or x, y, z , respectively.

Next, we need some ‘standard’ vector fields on \mathbb{R}^n .

Definition 9.20. The vector field \mathbf{U}_j on \mathbb{R}^n is defined by

$$\mathbf{U}_j(\mathbf{p}) = (0, \dots, 0, 1, 0, \dots, 0)_{\mathbf{p}},$$

where 1 occurs in the j^{th} spot. We shall call $\{\mathbf{U}_1, \dots, \mathbf{U}_n\}$ the **natural frame field** of \mathbb{R}^n .

It is obvious from Lemma 9.4 that $\mathbf{U}_j[f] = \partial f / \partial u_j$ for any differentiable function f , and it is common practice to write $\partial / \partial u_j$ in place of \mathbf{U}_j . Note that the restrictions of $\mathbf{U}_1, \dots, \mathbf{U}_n$ to any open subset $\mathcal{U} \subseteq \mathbb{R}^n$ are vector fields on \mathcal{U} .

The proofs of the following algebraic properties of vector fields are easy.

Lemma 9.21. Let \mathbf{X} and \mathbf{Y} be vector fields on an open subset $\mathcal{U} \subseteq \mathbb{R}^n$, let $a \in \mathbb{R}$ and let $g: \mathbb{R}^n \rightarrow \mathbb{R}$. Define $\mathbf{X} + \mathbf{Y}$, $a\mathbf{X}$, $g\mathbf{X}$ and $\mathbf{X} \cdot \mathbf{Y}$ by

$$\begin{aligned} (\mathbf{X} + \mathbf{Y})[f] &= \mathbf{X}[f] + \mathbf{Y}[f], & (a\mathbf{X})[f] &= a\mathbf{X}[f], \\ (g\mathbf{X})[f] &= g\mathbf{X}[f], & (\mathbf{X} \cdot \mathbf{Y})(\mathbf{p}) &= \mathbf{X}(\mathbf{p}) \cdot \mathbf{Y}(\mathbf{p}), \end{aligned}$$

where $f: \mathcal{U} \rightarrow \mathbb{R}$ is differentiable and $\mathbf{p} \in \mathcal{U}$. Then $\mathbf{X} + \mathbf{Y}$, $a\mathbf{X}$, $g\mathbf{X}$ are vector fields on \mathcal{U} and $\mathbf{X} \cdot \mathbf{Y}: \mathcal{U} \rightarrow \mathbb{R}$ is a function. If \mathbf{X} , \mathbf{Y} and g are differentiable, so are the vector fields $\mathbf{X} + \mathbf{Y}$, $a\mathbf{X}$, $g\mathbf{X}$, and the function $\mathbf{X} \cdot \mathbf{Y}$.

It is not hard to express a vector field in terms of the \mathbf{U}_j ’s:

Lemma 9.22. If \mathbf{V} is any vector field on an open subset $\mathcal{U} \subseteq \mathbb{R}^n$, there exist functions $v_j: \mathcal{U} \rightarrow \mathbb{R}$ for $j = 1, \dots, n$ such that

$$(9.10) \quad \mathbf{V} = \sum_{j=1}^n v_j \mathbf{U}_j.$$

Proof. Since $\mathbf{V}(\mathbf{p}) \in \mathbb{R}_{\mathbf{p}}^n$, we can write $\mathbf{V}(\mathbf{p}) = (v_1(\mathbf{p}), \dots, v_n(\mathbf{p}))_{\mathbf{p}}$. In this way the functions $v_j: \mathcal{U} \rightarrow \mathbb{R}$ are defined. On the other hand,

$$\begin{aligned} (v_1(\mathbf{p}), \dots, v_n(\mathbf{p}))_{\mathbf{p}} &= v_1(\mathbf{p})(1, 0, \dots, 0)_{\mathbf{p}} + \dots + v_n(\mathbf{p})(0, \dots, 0, 1)_{\mathbf{p}} \\ &= \sum_{j=1}^n v_j(\mathbf{p})\mathbf{U}_j(\mathbf{p}). \end{aligned}$$

Since this equation holds for arbitrary $\mathbf{p} \in \mathcal{U}$, we get (9.10). ■

The next lemma yields an important practical criterion for the differentiability of a vector field:

Corollary 9.23. *A vector field \mathbf{V} on an open subset $\mathcal{U} \subseteq \mathbb{R}^n$ is differentiable if and only if the natural coordinate functions v_1, \dots, v_n of \mathbf{V} given in (9.10) are differentiable.*

There is a natural way to construct vector fields from functions on \mathbb{R}^n .

Definition 9.24. *Let $g: \mathbb{R}^n \rightarrow \mathbb{R}$ be a differentiable function. The **gradient** of g is the vector field $\text{grad } g$ given by*

$$\text{grad } g = \sum_{j=1}^n \frac{\partial g}{\partial u_j} \mathbf{U}_j.$$

This operator is characterized by

Lemma 9.25. *The gradient of a function g is the unique differentiable vector field $\text{grad } g$ on \mathbb{R}^n for which*

$$(\text{grad } g) \cdot \mathbf{V} = \mathbf{V}[g]$$

for any differentiable vector field \mathbf{V} on \mathbb{R}^n .

Proof. Write $\mathbf{V} = \sum_{j=1}^n v_j \mathbf{U}_j$. Then

$$\begin{aligned} \text{grad } g \cdot \mathbf{V} &= \left(\sum_{j=1}^n \mathbf{U}_j \frac{\partial g}{\partial u_j} \right) \cdot \mathbf{V} = \sum_{j=1}^n v_j \frac{\partial g}{\partial u_j} \\ &= \sum_{j=1}^n v_j \mathbf{U}_j[g] = \mathbf{V}[g]. \end{aligned}$$

Conversely, let \mathbf{Z} be a vector field for which $\mathbf{Z} \cdot \mathbf{V} = \mathbf{V}[g]$ for any vector field \mathbf{V} on \mathbb{R}^n . We can write

$$\mathbf{Z} = \sum_{i=1}^n f_i \mathbf{U}_i,$$

where $f_i: \mathbb{R}^n \rightarrow \mathbb{R}$ is differentiable for $i = 1, \dots, n$. Then for $j = 1, \dots, n$ we have

$$f_j = \mathbf{Z} \cdot \mathbf{U}_j = \mathbf{U}_j[g] = \frac{\partial g}{\partial u_j},$$

so that $\mathbf{Z} = \text{grad } g$. ■

Finally, we note some special operations on vector fields on \mathbb{R}^2 and \mathbb{R}^3 , which extend the pointwise ones defined at the end of Section 9.1.

Lemma 9.26. For a vector field \mathbf{Y} on an open subset $\mathcal{U} \subseteq \mathbb{R}^2$, define $J\mathbf{Y}$ by

$$(J\mathbf{Y})(\mathbf{p}) = J(\mathbf{Y}(\mathbf{p}))$$

for $\mathbf{p} \in \mathcal{U}$. Similarly, for vector fields \mathbf{Y}, \mathbf{Z} on \mathcal{U} , define their cross product in a pointwise fashion, namely

$$(\mathbf{Y} \times \mathbf{Z})(\mathbf{p}) = \mathbf{Y}(\mathbf{p}) \times \mathbf{Z}(\mathbf{p})$$

for $\mathbf{p} \in \mathcal{U}$.

It is immediate that $J\mathbf{Y}$ and $\mathbf{Y} \times \mathbf{Z}$ are differentiable vector fields on \mathcal{U} whenever \mathbf{Y} and \mathbf{Z} are differentiable.

9.5 Derivatives of Vector Fields

In Section 9.2, we learned how to differentiate *functions* by means of tangent vectors. A key concept in differential geometry is the notion of how to differentiate a *vector field*. For vector fields on open subsets of \mathbb{R}^n , the definition is natural.

If \mathbf{W} is a vector field on an open subset $\mathcal{U} \subseteq \mathbb{R}^n$, one frequently does not distinguish the element $\mathbf{W}(\mathbf{p})$ inside the tangent space $\mathbb{R}_{\mathbf{p}}^n$ from its vector part $\mathbf{v} \in \mathbb{R}^n$. Strictly speaking this is incorrect, since $\mathbf{W}(\mathbf{p}) = \mathbf{v}_{\mathbf{p}}$ and one loses knowledge of the point \mathbf{p} . It is however a convenient abuse of notation, which (as will be explained in Section 9.6) was carried out repeatedly in Chapter 1.

Restrict \mathbf{W} to that portion of the straight line $t \mapsto \mathbf{p} + t\mathbf{v}$ contained in \mathcal{U} so that, with the convention above, $t \mapsto \mathbf{W}(\mathbf{p} + t\mathbf{v})$ is now a curve in \mathbb{R}^n .

Definition 9.27. Let \mathbf{W} be a differentiable vector field on an open subset \mathcal{U} of \mathbb{R}^n , and let $\mathbf{v}_{\mathbf{p}} \in \mathbb{R}_{\mathbf{p}}^n$. The derivative of \mathbf{W} with respect to $\mathbf{v}_{\mathbf{p}}$ is the element of $\mathbb{R}_{\mathbf{p}}^n$ given by

$$\mathbf{W}(\mathbf{p} + t\mathbf{v})'(0)_{\mathbf{p}} = \left(\lim_{t \rightarrow 0} \frac{\mathbf{W}(\mathbf{p} + t\mathbf{v}) - \mathbf{W}(\mathbf{p})}{t} \right)_{\mathbf{p}}.$$

This is denoted by $\mathbf{D}_{\mathbf{v}_{\mathbf{p}}} \mathbf{W}$, or $\mathbf{D}_{\mathbf{v}} \mathbf{W}$ if the identity of \mathbf{p} is understood.

It is possible to compute the tangent vector $\mathbf{D}_{\mathbf{v}} \mathbf{W}$ directly from its definition, but there is an alternative formula (similar to equation (9.7)), which is usually more convenient.

Lemma 9.28. Let \mathbf{W} be a differentiable vector field on an open subset $\mathcal{U} \subseteq \mathbb{R}^n$, and write

$$\mathbf{W} = \sum_{i=1}^n w_i \mathbf{U}_i.$$

Then for $\mathbf{p} \in \mathcal{U}$ and $\mathbf{v}_{\mathbf{p}} \in \mathbb{R}_{\mathbf{p}}^n$,

$$(9.11) \quad \mathbf{D}_{\mathbf{v}} \mathbf{W} = \sum_{i=1}^n \mathbf{v}_{\mathbf{p}}[w_i] \mathbf{U}_i(\mathbf{p}) = (\mathbf{v}_{\mathbf{p}}[w_1], \dots, \mathbf{v}_{\mathbf{p}}[w_n])_{\mathbf{p}}.$$

Proof. We have

$$\begin{aligned} \mathbf{W}(\mathbf{p} + t\mathbf{v}) &= \sum_{i=1}^n w_i(\mathbf{p} + t\mathbf{v}) \mathbf{U}_i(\mathbf{p} + t\mathbf{v}) \\ &= (w_1(\mathbf{p} + t\mathbf{v}), \dots, w_n(\mathbf{p} + t\mathbf{v}))_{\mathbf{p} + t\mathbf{v}}. \end{aligned}$$

Hence $\mathbf{W}(\mathbf{p} + t\mathbf{v})' = (w_1(\mathbf{p} + t\mathbf{v})', \dots, w_n(\mathbf{p} + t\mathbf{v})')$. But by the definition of tangent vector, we have

$$w_j(\mathbf{p} + t\mathbf{v})'(0) = \mathbf{v}_{\mathbf{p}}[w_j]$$

for $j = 1, \dots, n$, so

$$\mathbf{W}(\mathbf{p} + t\mathbf{v})'(0) = (\mathbf{v}_{\mathbf{p}}[w_1], \dots, \mathbf{v}_{\mathbf{p}}[w_n]),$$

as required. ■

It is an easy matter to compute the derivatives of the standard vector fields $\mathbf{U}_1, \dots, \mathbf{U}_n$.

Corollary 9.29. *We have $\mathbf{D}_{\mathbf{p}} \mathbf{U}_j = 0$ for any tangent vector $\mathbf{v}_{\mathbf{p}}$ to \mathbb{R}^n , and $j = 1, \dots, n$.*

Proof. If we use (9.10) to express \mathbf{U}_j in terms of $\mathbf{U}_1, \dots, \mathbf{U}_n$, then each of the coefficients v_1, \dots, v_n in (9.10) is either 0 or 1, and is in any case constant. The result follows from (9.11). ■

Next, we show that the operator \mathbf{D} behaves naturally with respect to the operations on vector fields given in Lemma 9.21.

Lemma 9.30. *Let \mathbf{Y}, \mathbf{Z} be differentiable vector fields on an open subset \mathcal{U} of \mathbb{R}^n , and $\mathbf{v}_{\mathbf{p}}, \mathbf{w}_{\mathbf{p}}$ tangent vectors to \mathbb{R}^n at $\mathbf{p} \in \mathcal{U}$. Then*

$$\begin{aligned} \mathbf{D}_{a\mathbf{v}+b\mathbf{w}} \mathbf{Y} &= a\mathbf{D}_{\mathbf{v}} \mathbf{Y} + b\mathbf{D}_{\mathbf{w}} \mathbf{Y}, \\ \mathbf{D}_{\mathbf{v}}(a\mathbf{Y} + b\mathbf{Z}) &= a\mathbf{D}_{\mathbf{v}} \mathbf{Y} + b\mathbf{D}_{\mathbf{v}} \mathbf{Z}, \\ \mathbf{D}_{\mathbf{v}}(f\mathbf{Y}) &= \mathbf{v}_{\mathbf{p}}[f]\mathbf{Y}(\mathbf{p}) + f(\mathbf{p})\mathbf{D}_{\mathbf{v}} \mathbf{Y}, \\ (9.12) \quad \mathbf{v}_{\mathbf{p}}[\mathbf{Y} \cdot \mathbf{Z}] &= \mathbf{D}_{\mathbf{v}} \mathbf{Y} \cdot \mathbf{Z}(\mathbf{p}) + \mathbf{Y}(\mathbf{p}) \cdot \mathbf{D}_{\mathbf{v}} \mathbf{Z}, \end{aligned}$$

where $a, b \in \mathbb{R}$ and $f: \mathcal{U} \rightarrow \mathbb{R}$ is a differentiable function.

Proof. We prove (9.12); the proofs of the other formulas are similar. Write

$$\mathbf{Y} = \sum_{i=1}^n y_i \mathbf{U}_i \quad \text{and} \quad \mathbf{Z} = \sum_{i=1}^n z_i \mathbf{U}_i.$$

From (9.4) and (9.5) it follows that

$$\begin{aligned} \mathbf{v}_p[\mathbf{Y} \cdot \mathbf{Z}] &= \sum_{i=1}^n \mathbf{v}_p[y_i z_i] \\ &= \sum_{i=1}^n \left\{ \mathbf{v}_p[y_i] z_i(\mathbf{p}) + y_i(\mathbf{p}) \mathbf{v}_p[z_i] \right\} \\ &= \left\{ \sum_{i=1}^n \mathbf{v}_p[y_i] \mathbf{U}_i(\mathbf{p}) \right\} \cdot \left\{ \sum_{i=1}^n z_i(\mathbf{p}) \mathbf{U}_i(\mathbf{p}) \right\} \\ &\quad + \left\{ \sum_{i=1}^n y_i(\mathbf{p}) \mathbf{U}_i(\mathbf{p}) \right\} \cdot \left\{ \sum_{i=1}^n \mathbf{v}_p[z_i] \mathbf{U}_i(\mathbf{p}) \right\} \\ &= \mathbf{D}_v \mathbf{Y} \cdot \mathbf{Z}(\mathbf{p}) + \mathbf{Y}(\mathbf{p}) \cdot \mathbf{D}_v \mathbf{Z}. \quad \blacksquare \end{aligned}$$

The operator \mathbf{D} also behaves naturally with respect to the complex structure J of \mathbb{R}^2 and the vector cross product of \mathbb{R}^3 :

Lemma 9.31. *Let \mathbf{Y} and \mathbf{Z} be differentiable vector fields on an open subset $\mathcal{U} \subseteq \mathbb{R}^n$, where $n = 2$ or 3 , and let $\mathbf{v}_p, \mathbf{w}_p$ be tangent vectors to \mathbb{R}^n . Then*

$$(9.13) \quad \mathbf{D}_v(J\mathbf{Y}) = J(\mathbf{D}_v \mathbf{Y}) \quad (\text{for } n = 2),$$

$$(9.14) \quad \mathbf{D}_v(\mathbf{Y} \times \mathbf{Z}) = \mathbf{D}_v \mathbf{Y} \times \mathbf{Z}(\mathbf{p}) + \mathbf{Y}(\mathbf{p}) \times \mathbf{D}_v \mathbf{Z} \quad (\text{for } n = 3).$$

Proof. Given a vector field \mathbf{Y} on \mathbb{R}^2 , we can write $\mathbf{Y} = y_1 \mathbf{U}_1 + y_2 \mathbf{U}_2$; then $J\mathbf{Y} = -y_2 \mathbf{U}_1 + y_1 \mathbf{U}_2$. It follows from (9.11) that

$$\mathbf{D}_v(J\mathbf{Y}) = (-\mathbf{v}_p[y_2], \mathbf{v}_p[y_1])_p = J(\mathbf{v}_p[y_1], \mathbf{v}_p[y_2])_p = J(\mathbf{D}_v \mathbf{Y}),$$

giving (9.13). The proof of (9.14) is similar. \blacksquare

It is also possible to define the derivative of one vector field by another.

Lemma 9.32. *Let \mathbf{V} and \mathbf{W} be vector fields defined on an open subset $\mathcal{U} \subseteq \mathbb{R}^n$, and assume that \mathbf{W} is differentiable. Define $\mathbf{D}_v \mathbf{W}$ by*

$$(\mathbf{D}_v \mathbf{W})(\mathbf{p}) = \mathbf{D}_{\mathbf{v}(\mathbf{p})} \mathbf{W}$$

for $\mathbf{p} \in \mathcal{U}$. Then $\mathbf{D}_v \mathbf{W}$ is a vector field on \mathcal{U} which is differentiable if \mathbf{V} and \mathbf{W} are differentiable. Moreover,

$$(9.15) \quad \mathbf{D}_v \left(\sum_{i=1}^n w_i \mathbf{U}_i \right) = \sum_{i=1}^n \mathbf{V}[w_i] \mathbf{U}_i.$$

Proof. Equation (9.15) is an immediate consequence of (9.11). \blacksquare

Let $F: \mathcal{U} \rightarrow \mathcal{V}$ be a diffeomorphism between open subsets \mathcal{U} and \mathcal{V} of \mathbb{R}^n , and let \mathbf{V} be a differentiable vector field on \mathcal{U} . We want to define the image $F_*(\mathbf{V})$ of \mathbf{V} under F , in a manner consistent with (9.7). The following does the job:

Definition 9.33. Let $F: \mathcal{U} \rightarrow \mathcal{V}$ be a diffeomorphism between open subsets \mathcal{U} and \mathcal{V} of \mathbb{R}^n and let \mathbf{V} be a differentiable vector field on \mathcal{U} . Write $F = (f_1, \dots, f_n)$. Then $F_*(\mathbf{V})$ is the vector field on \mathcal{V} given by

$$F_*(\mathbf{V}) = \sum_{i=1}^n (\mathbf{V}[f_i] \circ F^{-1}) \mathbf{U}_i.$$

It is easy to prove

Lemma 9.34. Suppose that $F: \mathcal{U} \rightarrow \mathcal{V}$ is a diffeomorphism between open subsets \mathcal{U} and \mathcal{V} of \mathbb{R}^n . Let \mathbf{Y} and \mathbf{Z} be differentiable vector fields on an open subset \mathcal{U} . Then

$$(9.16) \quad F_*(a\mathbf{Y} + b\mathbf{Z}) = aF_*(\mathbf{Y}) + bF_*(\mathbf{Z}),$$

$$(9.17) \quad F_*(f\mathbf{Y}) = (f \circ F^{-1})F_*(\mathbf{Y}),$$

$$F_*(\mathbf{Y})[f \circ F^{-1}] = \mathbf{Y}[f] \circ F^{-1},$$

where a, b are real numbers and $f: \mathcal{U} \rightarrow \mathbb{R}$ is a differentiable function.

Next, we show that certain maps of \mathbb{R}^n into itself have a desirable effect on the derivatives of vector fields.

Lemma 9.35. Suppose that $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is an affine transformation. Then

$$(9.18) \quad \mathbf{D}_{F_*(\mathbf{V})}(F_*(\mathbf{W})) = F_*(\mathbf{D}_{\mathbf{V}}\mathbf{W}).$$

Proof. By assumption, $F(\mathbf{p}) = A\mathbf{p} + \mathbf{q}$, where $A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is linear. We can therefore write

$$(9.19) \quad F_*(\mathbf{U}_j) = \sum_{i=1}^n a_{ij} \mathbf{U}_i, \quad j = 1, \dots, n$$

where each a_{ij} is constant. From (9.16) and (9.17), we get

$$\begin{aligned} (9.20) \quad \mathbf{D}_{F_*(\mathbf{V})}F_*(\mathbf{W}) &= \mathbf{D}_{F_*(\mathbf{V})}\left(F_*\left(\sum_{j=1}^n w_j \mathbf{U}_j\right)\right) \\ &= \mathbf{D}_{F_*(\mathbf{V})}\left(\sum_{j=1}^n (w_j \circ F^{-1})F_*(\mathbf{U}_j)\right) \\ &= \sum_{j=1}^n \mathbf{D}_{F_*(\mathbf{V})}((w_j \circ F^{-1})F_*(\mathbf{U}_j)). \end{aligned}$$

Then (9.15) and (9.19) imply that

$$\begin{aligned}
 (9.21) \quad \mathbf{D}_{F_*(\mathbf{V})}((w_j \circ F^{-1})F_*(\mathbf{U}_j)) &= \mathbf{D}_{F_*(\mathbf{V})}\left((w_j \circ F^{-1})\sum_{i=1}^n a_{ij}\mathbf{U}_i\right) \\
 &= \sum_{i=1}^n F_*(\mathbf{V})[(w_j \circ F^{-1})a_{ij}]\mathbf{U}_i.
 \end{aligned}$$

Since each a_{ij} is constant,

$$(9.22) \quad F_*(\mathbf{V})[(w_j \circ F^{-1})a_{ij}] = F_*(\mathbf{V})[w_j \circ F^{-1}]a_{ij}.$$

From (9.20)–(9.22), we get

$$\begin{aligned}
 (9.23) \quad \mathbf{D}_{F_*(\mathbf{V})}F_*(\mathbf{W}) &= \sum_{i,j=1}^n F_*(\mathbf{V})[w_j \circ F^{-1}]a_{ij}\mathbf{U}_i \\
 &= \sum_{j=1}^n (\mathbf{V}[w_j] \circ F^{-1})F_*(\mathbf{U}_j) \\
 &= \sum_{j=1}^n F_*(\mathbf{V}[w_j]\mathbf{U}_j).
 \end{aligned}$$

Now (9.18) follows from (9.23) and (9.15). ▀

We conclude this section by defining some well-known operators from vector analysis.

Definition 9.36. Let \mathbf{V} be a differentiable vector field defined on an open subset $\mathcal{U} \subseteq \mathbb{R}^n$. The **divergence** of \mathbf{V} is the function

$$\operatorname{div} \mathbf{V} = \sum_{i=1}^n \left(\mathbf{D}_{\mathbf{U}_i} \mathbf{V} \cdot \mathbf{U}_i \right).$$

If $f: \mathcal{U} \rightarrow \mathbb{R}$ is differentiable, then the **Laplacian** of f is the function

$$\Delta f = \operatorname{div}(\operatorname{grad} f).$$

Lemma 9.37. If $f: \mathcal{U} \rightarrow \mathbb{R}$ is a differentiable function, where $\mathcal{U} \subseteq \mathbb{R}^n$ is open, then

$$\Delta f = \sum_{i=1}^n \mathbf{U}_i^2 f = \sum_{i=1}^n \frac{\partial^2 f}{\partial u_i^2}.$$

Proof. We have

$$\Delta f = \operatorname{div}(\operatorname{grad} f) = \sum_{i=1}^n \mathbf{D}_{\mathbf{U}_i}(\operatorname{grad} f) \cdot \mathbf{U}_i$$

$$\begin{aligned}
&= \sum_{i=1}^n \mathbf{U}_i [\operatorname{grad} f \cdot \mathbf{U}_i] \\
&= \sum_{i=1}^n \mathbf{U}_i [\mathbf{U}_i[f]]. \blacksquare
\end{aligned}$$

9.6 Curves Revisited

In this section, we assume for simplicity that all curves are differentiable in the manner explained on page 266, so that partial derivatives to all orders exist.

In the study of curves and surfaces, it is common practice to make no distinction between a tangent vector $\mathbf{v}_{\mathbf{p}}$ and its vector part. For example, let $\alpha: (a, b) \rightarrow \mathbb{R}^n$ be a curve. On page 6, we defined its velocity $\alpha'(t)$ to be a vector in \mathbb{R}^n . More properly, it should be a tangent vector at the point $\alpha(t)$:

Definition 9.38. (Revised) The **velocity** of a curve α is the function

$$t \mapsto \alpha'(t)_{\alpha(t)},$$

and the **acceleration** of α is the function

$$t \mapsto \alpha''(t)_{\alpha(t)}.$$

Since $\alpha'(t)_{\alpha(t)}$ is a tangent vector, it can be applied to a differentiable function f :

Lemma 9.39. Let $\alpha: (a, b) \rightarrow \mathbb{R}^n$ be a curve and $f: \mathbb{R}^n \rightarrow \mathbb{R}$ a differentiable function. Then

$$\alpha'(t)_{\alpha(t)}[f] = (f \circ \alpha)'(t).$$

Proof. We write $\alpha'(t) = (a'_1(t), \dots, a'_n(t))$. It follows from Lemma 9.4 that

$$(9.24) \quad \alpha'(t)_{\alpha(t)}[f] = \sum_{j=1}^n a'_j(t) \frac{\partial f}{\partial u_j}(\alpha(t)).$$

But the chain rule says that the right-hand side of (9.24) is $(f \circ \alpha)'(t)$. \blacksquare

In Section 1.2 we defined the notion of vector field along a curve α in \mathbb{R}^n . A vector field \mathbf{W} on \mathbb{R}^n gives rise to a vector field on α ; it is simply $\mathbf{W} \circ \alpha$. Note that for each t , $(\mathbf{W} \circ \alpha)(t)$ is a vector in \mathbb{R}^n . A closely related concept is the **restriction** of \mathbf{W} to α , which is the function

$$t \mapsto \mathbf{W}(\alpha(t))_{\alpha(t)}.$$

Frequently, it is useful to make no distinction between the vector $(\mathbf{W} \circ \alpha)(t)$ in \mathbb{R}^n , and the tangent vector $\mathbf{W}(\alpha(t)) = \mathbf{W}(\alpha(t))_{\alpha(t)}$ in $\mathbb{R}^n_{\alpha(t)}$, mainly because it is too much trouble to write the subscript. Similarly, unless complete clarity is required, we shall not distinguish between $\alpha'(t)$ and $\alpha'(t)_{\alpha(t)}$.

In Section 1.2, we also defined the derivative of a vector field along a curve. In particular, we know how to compute $(\mathbf{W} \circ \alpha)'$. Let us make precise the relation between $(\mathbf{W} \circ \alpha)'$ and the derivative of \mathbf{W} with respect to $\alpha'(t)$.

Lemma 9.40. *If \mathbf{W} is a differentiable vector field on \mathbb{R}^n and $\alpha: (a, b) \rightarrow \mathbb{R}^n$ is a curve, then*

$$\mathbf{D}_{\alpha'(t)} \mathbf{W} = (\mathbf{W} \circ \alpha)'(t)_{\alpha(t)}.$$

Proof. By Lemma 9.28 we have

$$\mathbf{D}_{\alpha'(t)} \mathbf{W} = \sum_{i=1}^n \alpha'(t)[w_i] \mathbf{U}_i.$$

On the other hand, Lemma 9.39 implies that $\alpha'(t)[w_i] = (w_i \circ \alpha)'(t)$. Thus,

$$\begin{aligned} \mathbf{D}_{\alpha'(t)} \mathbf{W} &= \sum_{i=1}^n \alpha'(t)[w_i] \mathbf{U}_i \\ &= \sum_{i=1}^n (w_i \circ \alpha)'(t) \mathbf{U}_i \\ &= (\mathbf{W} \circ \alpha)'(t)_{\alpha(t)}. \quad \blacksquare \end{aligned}$$

9.7 Exercises

1. Let $\mathbf{v} = (v_1, v_2, v_3)$ and $\mathbf{p} = (2, -1, 4)$. Define $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ by

$$f(x, y, z) = xy^2z^4.$$

Compute $\mathbf{v}_{\mathbf{p}}[f]$.

2. Fill in the details of the proofs of Lemma 9.5, Lemma 9.8, Corollary 9.11 and Corollary 9.14.
3. Define $F: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ by $F(x, y, z) = (xy, yz, zx)$. Determine the following sets:

$$A = \{ \mathbf{p} \in \mathbb{R}^3 \mid \|\mathbf{p}\| = 1, F(\mathbf{p}) = \mathbf{0} \},$$

$$B = \{ \mathbf{p} \in \mathbb{R}^3 \mid \|\mathbf{p}\| = 1, F_*(\mathbf{v}_{\mathbf{p}}) = \mathbf{0} \text{ for some } \mathbf{v}_{\mathbf{p}} \in \mathbb{R}_{\mathbf{p}}^3 \}.$$

4. Compute by hand the gradients of the following functions:

$$(a) \quad (x, y, z) \mapsto e^{az} \cos ax - \cos ay.$$

$$(b) \quad (x, y, z) \mapsto \sin z - \sinh x \sinh y.$$

- (c) $(x, y, z) \mapsto \frac{x^2}{a^2} - \frac{y^2}{b^2} - c z$.
- (d) $(x, y, z) \mapsto x^m + y^n + z^p$.

M 5. Compute the gradient of

- (a) $g_1(x, y, z) = x^2 + 2y^2 + 3z^2 - xz + yz - xy$.
- (b) $g_2(x, y, z) = x^3 + y^3 + z^3 - 3axyz$.
- (c) $g_3(x, y, z) = xyz e^{x+y+z}$.
- (d) $g_4(x, y, z) = \arctan \left(x + y + z - \frac{xyz}{1 - xy - yz - xz} \right)$.

6. Compute by hand the divergence of each of the following vector fields:

- (a) $(x, y, z) \mapsto (ayz \cos xy, byz \sin xy, cxy)$.
- (b) $(x, y, z) \mapsto (\sin(xy), \cos(yz), -3 \cos^2(yz) \sin(xy) + \sin^3(xy))$.
- (c) $(x, y, z) \mapsto ((xyz)^a, (xyz)^b, (xyz)^c)$.
- (d) $(x, y, z) \mapsto (\log(x + y - z), e^{x-y+z}, (-x + y + z)^5)$.

M 7. Compute the divergence of

- (a) $\mathbf{V}_1(x, y, z) = (xyz, 2x + 3y + z, x^2 + z^2)$.
- (b) $\mathbf{V}_2(x, y, z) = (6x^2y^2 - z^3 + yz - 5, 4x^3y + xz + 2, xy - 3xz^2 - 3)$.

8. Compute the Laplacian of the functions in Exercises 4 and 5.

9. Let $f: \mathbb{R}^n \rightarrow \mathbb{R}$ be a differentiable function. A point $\mathbf{p} \in \mathbb{R}^n$ is said to be a **critical point** of f if $f_*: \mathbb{R}_{\mathbf{p}}^n \rightarrow \mathbb{R}_{F(\mathbf{p})}$ maps some nonzero tangent vector to zero. If \mathbf{p} is a critical point of f , define the **Hessian** of f at \mathbf{p} by

$$\text{Hessian}[f](\mathbf{v}_{\mathbf{p}}, \mathbf{w}_{\mathbf{p}}) = \mathbf{v}_{\mathbf{p}}[\mathbf{W}[f]],$$

for $\mathbf{v}_{\mathbf{p}}, \mathbf{w}_{\mathbf{p}} \in \mathbb{R}_{\mathbf{p}}^n$. Here \mathbf{W} is any vector field on \mathbb{R}^n such that $\mathbf{W}(\mathbf{p}) = \mathbf{w}_{\mathbf{p}}$.

(a) Show that $\text{Hessian}[f](\mathbf{v}_{\mathbf{p}}, \mathbf{w}_{\mathbf{p}}) = \sum_{i,j=1}^n v_i w_j \frac{\partial^2 f}{\partial u_i \partial u_j}$.

(b) Conclude that the definition of the Hessian is independent of the choice of the vector field \mathbf{W} .

10. Fill in the details of the proofs of Lemmas 9.30, 9.31 and 9.34.