Chapter 1
Calculus on Euclidean Space

As mentioned in the Preface, the purpose of this initial chapter is to establish the mathematical language used throughout the book. Much of what we do is simply a review of that part of elementary calculus dealing with differentiation of functions of three variables and with curves in space. Our definitions have been formulated so that they will apply smoothly to the later study of surfaces.

1.1 Euclidean Space

Three-dimensional space is often used in mathematics without being formally defined. Looking at the corner of a room, one can picture the familiar process by which rectangular coordinate axes are introduced and three numbers are measured to describe the position of each point. A precise definition that realizes this intuitive picture may be obtained by this device: instead of saying that three numbers *describe the position* of a point, we define them to *be* a point.

1.1 Definition *Euclidean 3-space* $\mathbf{R}^3$ is the set of all ordered triples of real numbers. Such a triple $\mathbf{p} = (p_1, p_2, p_3)$ is called a *point* of $\mathbf{R}^3$.

In linear algebra, it is shown that $\mathbf{R}^3$ is, in a natural way, a vector space over the real numbers. In fact, if $\mathbf{p} = (p_1, p_2, p_3)$ and $\mathbf{q} = (q_1, q_2, q_3)$ are points of $\mathbf{R}^3$, their *sum* is the point

$$\mathbf{p} + \mathbf{q} = (p_1 + q_1, p_2 + q_2, p_3 + q_3).$$

The *scalar multiple* of a point $\mathbf{p} = (p_1, p_2, p_3)$ by a number a is the point

$$a\mathbf{p} = (ap_1, ap_2, ap_3).$$

It is easy to check that these two operations satisfy the axioms for a vector space. The point $\mathbf{0} = (0, 0, 0)$ is called the *origin* of $\mathbf{R}^3$.

Differential calculus deals with another aspect of $\mathbf{R}^3$ starting with the notion of differentiable real-valued functions on $\mathbf{R}^3$. We recall some fundamentals.

1.2 Definition Let x, y, and z be the real-valued functions on $\mathbf{R}^3$ such that for each point $\mathbf{p} = (p_1, p_2, p_3)$

$$x(\mathbf{p}) = p_1, \quad y(\mathbf{p}) = p_2, \quad z(\mathbf{p}) = p_3.$$

These functions x, y, z are called the *natural coordinate functions* of $\mathbf{R}^3$. We shall also use index notation for these functions, writing

$$x_1 = x, \quad x_2 = y, \quad x_3 = z.$$

Thus the value of the function x_i on a point $\mathbf{p}$ is the number p_i, and so we have the identity $\mathbf{p} = (p_1, p_2, p_3) = (x_1(\mathbf{p}), x_2(\mathbf{p}), x_3(\mathbf{p}))$ for each point $\mathbf{p}$ of $\mathbf{R}^3$. Elementary calculus does not always make a sharp distinction between the *numbers* p_1, p_2, p_3 and the *functions* x_1, x_2, x_3. Indeed the analogous distinction on the real line may seem pedantic, but for higher-dimensional spaces such as $\mathbf{R}^3$, its absence leads to serious ambiguities. (Essentially the same distinction is being made when we denote a function on $\mathbf{R}^3$ by a single letter f, reserving $f(\mathbf{p})$ for its value at the point $\mathbf{p}$.)

We assume that the reader is familiar with partial differentiation and its basic properties, in particular the chain rule for differentiation of a composite function. We shall work mostly with first-order partial derivatives $\partial f/\partial x$, $\partial f/\partial y$, $\partial f/\partial z$ and second-order partial derivatives $\partial^2 f/\partial x^2$, $\partial^2 f/\partial x \partial y$, . . . In a few situations, third- and even fourth-order derivatives may occur, but to avoid worrying about exactly how many derivatives we can take in any given context, we establish the following definition.

1.3 Definition A real-valued function f on $\mathbf{R}^3$ is *differentiable* (or *infinitely differentiable*, or *smooth*, or *of class* C^∞) provided all partial derivatives of f, of all orders, exist and are continuous.

Differentiable real-valued functions f and g may be added and multiplied in a familiar way to yield functions that are again differentiable and real-valued. We simply add and multiply their values at each point—the formulas read

$$(f+g)(\mathbf{p}) = f(\mathbf{p}) + g(\mathbf{p}), \quad (fg)(\mathbf{p}) = f(\mathbf{p})g(\mathbf{p}).$$

The phrase "differentiable real-valued function" is unpleasantly long. Hence we make the convention that *unless the context indicates otherwise*, "function" shall mean "real-valued function," and (unless the issue is explicitly raised) the functions we deal with will be assumed to be differentiable. We do not intend to overwork this convention; for the sake of emphasis the words "differentiable" and "real-valued" will still appear fairly frequently.

Differentiation is always a *local* operation: To compute the value of the function $\partial f/\partial x$ at a point $\mathbf{p}$ of $\mathbf{R}^3$, it is sufficient to know the values of f at all points $\mathbf{q}$ of $\mathbf{R}^3$ that are sufficiently near $\mathbf{p}$. Thus, Definition 1.3 is unduly restrictive; the domain of f need not be the whole of $\mathbf{R}^3$, but need only be an *open set* of $\mathbf{R}^3$. By an *open set* $\mathcal{O}$ of $\mathbf{R}^3$ we mean a subset of $\mathbf{R}^3$ such that if a point $\mathbf{p}$ is in $\mathcal{O}$, then so is every other point of $\mathbf{R}^3$ that is sufficiently near $\mathbf{p}$. (A more precise definition is given in Chapter 2.) For example, the set of all points $\mathbf{p} = (p_1, p_2, p_3)$ in $\mathbf{R}^3$ such that $p_1 > 0$ is an open set, and the function $yz \log x$ defined on this set is certainly differentiable, even though its domain is not the whole of $\mathbf{R}^3$. Generally speaking, the results in this chapter remain valid if $\mathbf{R}^3$ is replaced by an arbitrary open set $\mathcal{O}$ of $\mathbf{R}^3$.

We are dealing with *three-dimensional* Euclidean space only because this is the dimension we use most often in later work. It would be just as easy to work with *Euclidean n-space* $\mathbf{R}^n$, for which the points are n-tuples $\mathbf{p} = (p_1, \ldots, p_n)$ and which has n natural coordinate functions $x_1, \ldots, x_n$. All the results in this chapter are valid for Euclidean spaces of arbitrary dimensions, although we shall rarely take advantage of this except in the case of the *Euclidean plane* $\mathbf{R}^2$. In particular, the results are valid for the *real line* $\mathbf{R}^1 = \mathbf{R}$. Many of the concepts introduced are designed to deal with higher dimensions, however, and are thus apt to be overelaborate when reduced to dimension 1.

Exercises

1. Let $f = x^2y$ and $g = y \sin z$ be functions on $\mathbf{R}^3$. Express the following functions in terms of x, y, z:

(a) fg^2.

(b) $\dfrac{\partial f}{\partial x} g + \dfrac{\partial g}{\partial y} f$.

(c) $\dfrac{\partial^2 (fg)}{\partial y \partial z}$.

(d) $\dfrac{\partial}{\partial y}(\sin f)$.

2. Find the value of the function $f = x^2y - y^2z$ at each point:
(a) $(1, 1, 1)$. (b) $(3, -1, \frac{1}{2})$.
(c) $(a, 1, 1 - a)$. (d) (t, t^2, t^3).

3. Express $\partial f/\partial x$ in terms of x, y, and z if
(a) $f = x\sin(xy) + y\cos(xz)$.
(b) $f = \sin g$, $g = e^h$, $h = x^2 + y^2 + z^2$.

4. If g_1, g_2, g_3, and h are real-valued functions on $\mathbf{R}^3$, then

$$f = h(g_1, g_2, g_3)$$

is the function such that

$$f(\mathbf{p}) = h(g_1(\mathbf{p}), g_2(\mathbf{p}), g_3(\mathbf{p})) \quad \text{for all } \mathbf{p}.\dagger$$

Express $\partial f/\partial x$ in terms of x, y, and z, if $h = x^2 - yz$ and
(a) $f = h(x + y, y^2, x + z)$. (b) $f = h(e^z, e^{x+y}, e^x)$.
(c) $f = h(x, -x, x)$.

1.2 Tangent Vectors

Intuitively, a vector in $\mathbf{R}^3$ is an oriented line segment, or "arrow." Vectors are used widely in physics and engineering to describe forces, velocities, angular momenta, and many other concepts. To obtain a definition that is both practical and precise, we shall describe an "arrow" in $\mathbf{R}^3$ by giving its starting point $\mathbf{p}$ and the change, or vector $\mathbf{v}$, necessary to reach its end point $\mathbf{p} + \mathbf{v}$. Strictly speaking, $\mathbf{v}$ is just a point of $\mathbf{R}^3$.

2.1 Definition‡ A *tangent vector* $\mathbf{v}_p$ to $\mathbf{R}^3$ consists of two points of $\mathbf{R}^3$: its *vector part* $\mathbf{v}$ and its *point of application* $\mathbf{p}$.

We shall always picture $\mathbf{v}_p$ as *the arrow from the point* $\mathbf{p}$ to the point $\mathbf{p} + \mathbf{v}$. For example, if $\mathbf{p} = (1, 1, 3)$ and $\mathbf{v} = (2, 3, 2)$, then $\mathbf{v}_p$ runs from $(1, 1, 3)$ to $(3, 4, 5)$ as in Fig. 1.1.

We emphasize that tangent vectors are equal, $\mathbf{v}_p = \mathbf{w}_q$, if and only if they have the same vector part, $\mathbf{v} = \mathbf{w}$, and the same point of application, $\mathbf{p} = \mathbf{q}$.

† A consequence is the identity $f = f(x, y, z)$.

‡ The term "tangent" in this definition will acquire a more direct geometric meaning in Chapter 4.

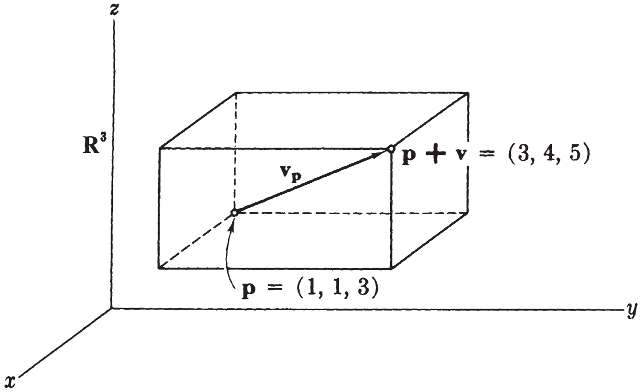

FIG. 1.1

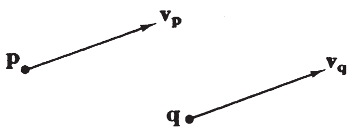

FIG. 1.2

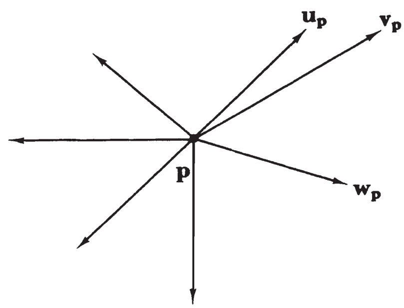

FIG. 1.3

Tangent vectors $\mathbf{v}_p$ and $\mathbf{v}_q$ with the same vector part, but different points of application, are said to be *parallel* (Fig. 1.2). It is essential to recognize that $\mathbf{v}_p$ and $\mathbf{v}_q$ are different tangent vectors if $\mathbf{p} \neq \mathbf{q}$. In physics the concept of moment of a force shows this clearly enough: The same force **v** applied at different points **p** and **q** of a rigid body can produce quite different rotational effects.

2.2 Definition Let **p** be a point of $\mathbf{R}^3$. The set $T_p(\mathbf{R}^3)$ consisting of all tangent vectors that have **p** as point of application is called the *tangent space* of $\mathbf{R}^3$ at **p** (Fig. 1.3).

We emphasize that $\mathbf{R}^3$ has a different tangent space at each and every one of its points.

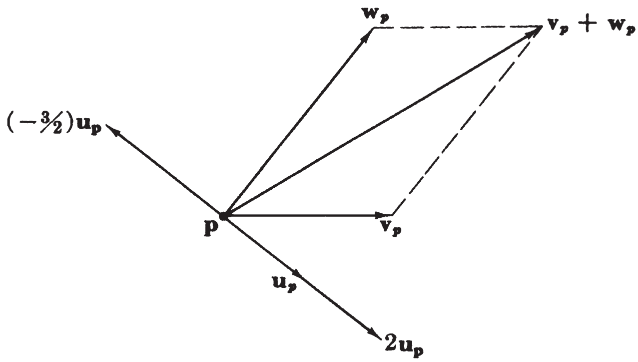

FIG. 1.4

Since all the tangent vectors in a given tangent space have the same point of application, we can borrow the vector addition and scalar multiplication of $\mathbf{R}^3$ to turn $T_p(\mathbf{R}^3)$ into a vector space. Explicitly, we define $\mathbf{v}_p + \mathbf{w}_p$ to be $(\mathbf{v} + \mathbf{w})_p$. and if c is a number we define $c(\mathbf{v}_p)$ to be $(c\mathbf{v})_p$. This is just the usual "parallelogram law" for addition of vectors, and scalar multiplication by c merely stretches a tangent vector by the factor c—reversing its direction if $c < 0$ (Fig. 1.4).

These operations on the tangent space $T_p(\mathbf{R}^3)$ make it a vector space isomorphic to $\mathbf{R}^3$ itself. Indeed, it follows immediately from the definitions above that for a fixed point $\mathbf{p}$, the function $\mathbf{v} \to \mathbf{v}_p$ is a linear isomorphism from $\mathbf{R}^3$ to $T_p(\mathbf{R}^3)$—that is, a linear transformation that is one-to-one and onto.

A standard concept in physics and engineering is that of a force field. The gravitational force field of the earth, for example, assigns to each point of space a force (vector) directed at the center of the earth.

2.3 Definition A *vector field* V on $\mathbf{R}^3$ is a function that assigns to each point $\mathbf{p}$ of $\mathbf{R}^3$ a tangent vector $V(\mathbf{p})$ to $\mathbf{R}^3$ at $\mathbf{p}$.

Roughly speaking, a vector field is just a big collection of arrows, one at each point of $\mathbf{R}^3$.

There is a natural algebra of vector fields. To describe it, we first reexamine the familiar notion of addition of real-valued functions f and g. It is possible to add f and g because it is possible to add their values at each point. The same is true of vector fields V and W. At each point $\mathbf{p}$, the values $V(\mathbf{p})$ and $W(\mathbf{p})$ are in the same vector space—the tangent space $T_p(\mathbf{R}^3)$—hence we can add $V(\mathbf{p})$ and $W(\mathbf{p})$. Consequently, we can add V and W by adding their values at each point. The formula for this addition is thus the same as for addition of functions,

$$(V+W)(\mathbf{p}) = V(\mathbf{p}) + W(\mathbf{p}).$$

This scheme occurs over and over again. We shall call it the *pointwise principle:* If a certain operation can be performed on the values of two functions at each point, then that operation can be extended to the functions themselves; simply apply it to their values at each point.

For example, we invoke the pointwise principle to extend the operation of *scalar multiplication* (on the tangent spaces of $\mathbf{R}^3$). If f is a real-valued function on $\mathbf{R}^3$ and V is a vector field on $\mathbf{R}^3$, then fV is defined to be the vector field on $\mathbf{R}^3$ such that

$$(fV)(\mathbf{p}) = f(\mathbf{p})V(\mathbf{p}) \quad \text{for all } \mathbf{p}.$$

Our aim now is to determine in a concrete way just what vector fields look like. For this purpose we introduce three special vector fields that will serve as a "basis" for all vector fields.

2.4 Definition Let U_1, U_2, and U_3 be the vector fields on $\mathbf{R}^3$ such that

$$U_1(\mathbf{p}) = (1, 0, 0)_p$$
$$U_2(\mathbf{p}) = (0, 1, 0)_p$$
$$U_3(\mathbf{p}) = (0, 0, 1)_p$$

for each point $\mathbf{p}$ of $\mathbf{R}^3$ (Fig. 1.5). We call U_1, U_2, U_3—collectively—the *natural frame field* on $\mathbf{R}^3$.

Thus, U_i ($i = 1, 2, 3$) is the unit vector field in the positive x_i direction.

2.5 Lemma If V is a vector field on $\mathbf{R}^3$, there are three uniquely determined real-valued functions, v_1, v_2, v_3 on $\mathbf{R}^3$ such that

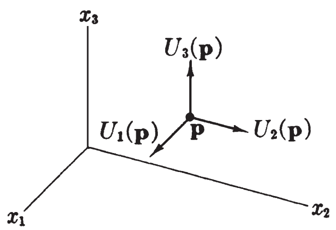

FIG. 1.5

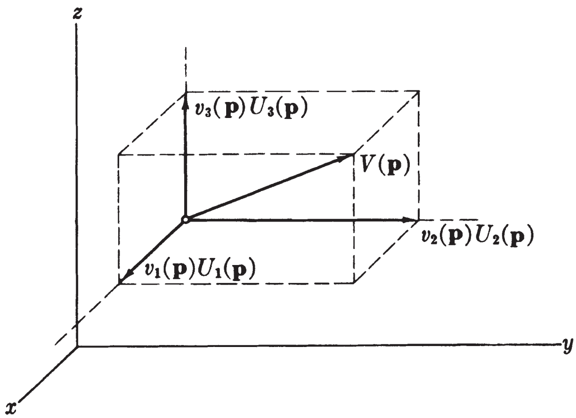

FIG. 1.6

$$V = v_1U_1 + v_2U_2 + v_3U_3.$$

The functions v_1, v_2, v_3 are called the *Euclidean coordinate functions of V.*

Proof. By definition, the vector field V assigns to each point $\mathbf{p}$ a tangent vector $V(\mathbf{p})$ at $\mathbf{p}$. Thus, the vector part of $V(\mathbf{p})$ depends on $\mathbf{p}$, so we write it $(v_1(\mathbf{p}), v_2(\mathbf{p}), v_3(\mathbf{p}))$. (This defines v_1, v_2, and v_3 as real-valued *functions* on $\mathbf{R}^3$.) Hence

$$\begin{aligned} V(\mathbf{p}) &= (v_1(\mathbf{p}), v_2(\mathbf{p}), v_3(\mathbf{p}))_p \\ &= v_1(\mathbf{p})(1,0,0)_p + v_2(\mathbf{p})(0,1,0)_p + v_3(\mathbf{p})(0,0,1)_p \\ &= v_1(\mathbf{p})U_1(\mathbf{p}) + v_2(\mathbf{p})U_2(\mathbf{p}) + v_3(\mathbf{p})U_3(\mathbf{p}) \end{aligned}$$

for each point $\mathbf{p}$ (Fig. 1.6). By our (pointwise principle) definitions, this means that the vector fields V and $\sum v_iU_i$ have the same (tangent vector) value at each point. Hence $V = \sum v_iU_i$. ◆

This last sentence uses two of our standard conventions: $\sum v_iU_i$ means sum over $i = 1, 2, 3$; the symbol (◆) indicates the end of a proof.

The tangent-vector identity $(a_1, a_2, a_3)_p = \sum a_iU_i(\mathbf{p})$ appearing in this proof will be used very often.

Computations involving vector fields may always be expressed in terms of their Euclidean coordinate functions. For example, addition and multiplication by a function, are expressed in terms of coordinates by

$$\sum v_iU_i + \sum w_iU_i = \sum (v_i + w_i)U_i,$$

$$f\left(\sum v_iU_i\right) = \sum (fv_i)U_i.$$

Since this is differential calculus, we shall naturally require that the various objects we deal with be differentiable. A vector field V is *differentiable* provided its *E*uclidean coordinate functions are differentiable (in the sense of Definition 1.3). From now on, we shall understand "vector field" to mean "differentiable vector field."

Exercises

1. Let $\mathbf{v} = (-2, 1, -1)$ and $\mathbf{w} = (0, 1, 3)$.
(a) At an arbitrary point $\mathbf{p}$, express the tangent vector $3\mathbf{v}_p - 2\mathbf{w}_p$ as a linear combination of $U_1(\mathbf{p})$, $U_2(\mathbf{p})$, $U_3(\mathbf{p})$.
(b) For $\mathbf{p} = (1, 1, 0)$, make an accurate sketch showing the four tangent vectors $\mathbf{v}_p$, $\mathbf{w}_p$, $-2\mathbf{v}_p$, and $\mathbf{v}_p + \mathbf{w}_p$.

2. Let $V = xU_1 + yU_2$ and $W = 2x^2U_2 - U_3$. Compute the vector field $W - xV$, and find its value at the point $\mathbf{p} = (-1, 0, 2)$.

3. In each case, express the given vector field V in the standard form $\sum v_i U_i$.
(a) $2z^2U_1 = 7V + xyU_3$.
(b) $V(\mathbf{p}) = (p_1, p_3 - p_1, 0)_p$ for all $\mathbf{p}$.
(c) $V = 2(xU_1 + yU_2) - x(U_1 - y^2U_3)$.
(d) At each point $\mathbf{p}$, $V(\mathbf{p})$ is the vector from the point (p_1, p_2, p_3) to the point $(1 + p_1, p_2p_3, p_2)$.
(e) At each point $\mathbf{p}$, $V(\mathbf{p})$ is the vector from $\mathbf{p}$ to the origin.

4. If $V = y^2U_1 - x^2U_3$ and $W = x^2U_1 - zU_2$, find functions f and g such that the vector field $fV + gW$ can be expressed in terms of U_2 and U_3 only.

5. Let $V_1 = U_1 - xU_3$, $V_2 = U_2$, and $V_3 = xU_1 + U_3$.
(a) Prove that the vectors $V_1(\mathbf{p})$, $V_2(\mathbf{p})$, $V_3(\mathbf{p})$ are linearly independent at each point of $\mathbf{R}^3$.
(b) Express the vector field $xU_1 + yU_2 + zU_3$ as a linear combination of V_1, V_2, V_3.

1.3 Directional Derivatives

Associated with each tangent vector $\mathbf{v}_p$ to $\mathbf{R}^3$ is the straight line $t \to \mathbf{p} + t\mathbf{v}$ (see Example 4.2). If f is a differentiable function on $\mathbf{R}^3$, then $t \to f(\mathbf{p} + t\mathbf{v})$ is an ordinary differentiable function on the real line. Evidently the derivative of this function at $t = 0$ tells the initial rate of change of f as $\mathbf{p}$ moves in the $\mathbf{v}$ direction

3.1 Definition Let f be a differentiable real-valued function on $\mathbf{R}^3$, and let $\mathbf{v}_p$ be a tangent vector to $\mathbf{R}^3$. Then the number

$$\mathbf{v}_p[f] = \frac{d}{dt}(f(\mathbf{p} + t\mathbf{v}))|_{t=0}$$

is called the *derivative of f with respect* to $\mathbf{v}_p$.

This definition appears in elementary calculus with the additional restriction that $\mathbf{v}_p$ be a unit vector. Even though we do not impose this restriction, we shall nevertheless refer to $\mathbf{v}_p[f]$ as a *directional derivative.*

For example, we compute $\mathbf{v}_p[f]$ for the function $f = x^2yz$, with $\mathbf{p} = (1, 1, 0)$ and $\mathbf{v} = (1, 0, -3)$. Then

$$\mathbf{p} + t\mathbf{v} = (1, 1, 0) + t(1, 0, -3) = (1 + t, 1, -3t)$$

describes the line through $\mathbf{p}$ in the $\mathbf{v}$ direction. Evaluating f along this line, we get

$$f(\mathbf{p} + t\mathbf{v}) = (1 + t)^2 \cdot 1 \cdot (-3t) = -3t - 6t^2 - 3t^3.$$

Now,

$$\frac{d}{dt}(f(\mathbf{p} + t\mathbf{v})) = -3 - 12t - 9t^2;$$

hence at $t = 0$, we find $\mathbf{v}_p[f] = -3$. Thus, in particular, the function f is initially decreasing as $\mathbf{p}$ moves in the $\mathbf{v}$ direction.

The following lemma shows how to compute $\mathbf{v}_p[f]$ in general, in terms of the partial derivatives of f at the point $\mathbf{p}$.

3.2 Lemma If $\mathbf{v}_p = (v_1, v_2, v_3)_p$ is a tangent vector to $\mathbf{R}^3$, then

$$\mathbf{v}_p[f] = \sum v_i \frac{\partial f}{\partial x_i}(\mathbf{p}).$$

Proof. Let $\mathbf{p} = (p_1, p_2, p_3)$; then

$$\mathbf{p} + t\mathbf{v} = (p_1 + tv_1, p_2 + tv_2, p_3 + tv_3).$$

We use the chain rule to compute the derivative at $t = 0$ of the function

$$f(\mathbf{p} + t\mathbf{v}) = f(p_1 + tv_1, p_2 + tv_2, p_3 + tv_3).$$

Since

$$\frac{d}{dt}(p_i + tv_i) = v_i,$$

we obtain

$$\mathbf{v}_p[f] = \frac{d}{dt}(f(\mathbf{p} + t\mathbf{v}))|_{t=0} = \sum \frac{\partial f}{\partial x_i}(\mathbf{p})v_i. \qquad \blacklozenge$$

Using this lemma, we recompute $\mathbf{v}_p[f]$ for the example above. Since $f = x^2yz$, we have

$$\frac{\partial f}{\partial x} = 2xyz, \quad \frac{\partial f}{\partial y} = x^2z, \quad \frac{\partial f}{\partial z} = x^2y.$$

Thus, at the point $\mathbf{p} = (1, 1, 0)$,

$$\frac{\partial f}{\partial x}(\mathbf{p}) = 0, \quad \frac{\partial f}{\partial y}(\mathbf{p}) = 0, \quad \text{and} \quad \frac{\partial f}{\partial z}(\mathbf{p}) = 1.$$

Then by the lemma,

$$\mathbf{v}_p[f] = 0 + 0 + (-3)1 = -3,$$

as before.

The main properties of this notion of derivative are as follows.

3.3 Theorem Let f and g be functions on $\mathbf{R}^3$, $\mathbf{v}_p$ and $\mathbf{w}_p$ tangent vectors, a and b numbers. Then

(1) $(a\mathbf{v}_p + b\mathbf{w}_p)[f] = a\mathbf{v}_p[f] + b\mathbf{w}_p[f]$.
(2) $\mathbf{v}_p[af + bg] = a\mathbf{v}_p[f] + b\mathbf{v}_p[g]$.
(3) $\mathbf{v}_p[fg] = \mathbf{v}_p[f]\cdot g(\mathbf{p}) + f(\mathbf{p})\cdot \mathbf{v}_p[g]$.

Proof. All three properties may be deduced easily from the preceding lemma. For example, we prove (3). By the lemma, if $\mathbf{v} = (v_1, v_2, v_3)$, then

$$\mathbf{v}_p[fg] = \sum v_i \frac{\partial (fg)}{\partial x_i}(\mathbf{p}).$$

But

$$\frac{\partial (fg)}{\partial x_i} = \frac{\partial f}{\partial x_i}\cdot g + f\cdot \frac{\partial g}{\partial x_i}.$$

Hence

$$\begin{aligned}\mathbf{v}_p[fg] &= \sum v_i\left(\frac{\partial f}{\partial x_i}(\mathbf{p})\cdot g(\mathbf{p}) + f(\mathbf{p})\cdot\frac{\partial g}{\partial x_i}(\mathbf{p})\right)\\ &= \left(\sum v_i \frac{\partial f}{\partial x_i}(\mathbf{p})\right)g(\mathbf{p}) + f(\mathbf{p})\left(\sum v_i \frac{\partial g}{\partial x_i}(\mathbf{p})\right)\\ &= \mathbf{v}_p[f]\cdot g(\mathbf{p}) + f(\mathbf{p})\cdot \mathbf{v}_p[g]. \qquad \blacklozenge\end{aligned}$$

The first two properties in the preceding theorem may be summarized by saying that $\mathbf{v}_p[f]$ is *linear* in $\mathbf{v}_p$ and in f. The third property, as its proof makes clear, is essentially just the usual Leibniz rule for differentiation of a product. *No matter what form differentiation may take, it will always have suitable* linear *and* Leibnizian *properties.*

We now use the pointwise principle to define the *operation of a vector field V on a function f.* The result is the real-valued function $V[f]$ whose value at each point $\mathbf{p}$ is the number $V(\mathbf{p})[f]$, that is, the derivative of f with respect to the tangent vector $V(\mathbf{p})$ at $\mathbf{p}$. This process should be no surprise, since for a function f on the real line, one begins by defining the derivative of f *at a point*—then the derivative *function df/dx* is the function whose value at each point is the derivative at that point. Evidently, the definition of $V[f]$ is strictly analogous. In particular, if U_1, U_2, U_3 is the natural frame field on $\mathbf{R}^3$, then $U_i[f] = \partial f/\partial x_i$. This is an immediate consequence of Lemma 3.2. For example, $U_1(\mathbf{p}) = (1, 0, 0)_p$; hence

$$U_1(\mathbf{p})[f] = \frac{d}{dt}(f(p_1 + t, p_2, p_3))|_{t=0},$$

which is precisely the definition of $(\partial f/\partial x_1)(\mathbf{p})$. This is true for all points $\mathbf{p} = (p_1, p_2, p_3)$; hence $U_1[f] = \partial f/\partial x_1$.

We shall use this notion of directional derivative more in the case of vector fields than for individual tangent vectors.

3.4 Corollary If V and W are vector fields on $\mathbf{R}^3$ and f, g, h are real-valued functions, then

(1) $(fV + gW)[h] = fV[h] + gW[h]$.
(2) $V[af + bg] = aV[f] + bV[g]$, for all real numbers a and b.
(3) $V[fg] = V[f]\cdot g + f\cdot V[g]$.

Proof. The pointwise principle guarantees that to derive these properties from Theorem 3.3 we need only be careful about the placement of parentheses. For example, we prove the third formula. By definition, the value of the function $V[fg]$ at $\mathbf{p}$ is $V(\mathbf{p})[fg]$. But by Theorem 3.3 this is

$$\begin{aligned} V(\mathbf{p})[f]\cdot g(\mathbf{p}) + f(\mathbf{p})\cdot V(\mathbf{p})[g] &= V[f](\mathbf{p})\cdot g(\mathbf{p}) + f(\mathbf{p})\cdot V[g](\mathbf{p}) \\ &= (V[f]\cdot g + f\cdot V[g])(\mathbf{p}). \end{aligned}$$ ◆

If the use of parentheses here seems extravagant, we remind the reader that a meticulous proof of Leibniz's formula

$$\frac{d}{dx}(fg) = \frac{df}{dx} \cdot g + f \cdot \frac{dg}{dx}$$

must involve the same shifting of parentheses.

Note that the linearity of $V[f]$ in V and f is for *functions* as "scalars" in the first formula in Corollary 3.4 but only for *numbers* as "scalars" in the second. This stems from the fact that fV signifies merely multiplication, but $V[f]$ is differentiation.

The identity $U_i[f] = \partial f/\partial x_i$ makes it a simple matter to carry out explicit computations. For example, if $V = xU_1 - y^2U_3$ and $f = x^2y + z^3$, then

$$\begin{aligned} V[f] &= xU_1[x^2y] + xU_1[z^3] - y^2U_3[x^2y] - y^2U_3[z^3] \\ &= x(2xy) + 0 - 0 - y^2(3z^2) = 2x^2y - 3y^2z^2. \end{aligned}$$

3.5 Remark Since the subscript notation $\mathbf{v}_p$ for a tangent vector is somewhat cumbersome, from now on we shall frequently omit the point of application $\mathbf{p}$ from the notation. This can cause no confusion, since $\mathbf{v}$ and $\mathbf{w}$ will always denote tangent vectors, and $\mathbf{p}$ and $\mathbf{q}$ points of $\mathbf{R}^3$. In many situations (for example, Definition 3.1) the point of application is crucial, and will be indicated by using either the old notation $\mathbf{v}_p$ or the phrase "a tangent vector $\mathbf{v}$ to $\mathbf{R}^3$ at $\mathbf{p}$."

Exercises

1. Let $\mathbf{v}_p$ be the tangent vector to $\mathbf{R}^3$ with $\mathbf{v} = (2, -1, 3)$ and $\mathbf{p} = (2, 0, -1)$. Working directly from the definition, compute the directional derivative $\mathbf{v}_p[f]$, where

(a) $f = y^2z$. (b) $f = x^7$.
(c) $f = e^x \cos y$.

2. Compute the derivatives in Exercise 1 using Lemma 3.2.

3. Let $V = y^2U_1 - xU_3$, and let $f = xy$, $g = z^3$. Compute the functions

(a) $V[f]$. (b) $V[g]$.
(c) $V[fg]$. (d) $fV[g] - gV[f]$.
(e) $V[f^2 + g^2]$. (f) $V[V[f]]$.

4. Prove the identity $V = \sum V[x_i]U_i$, where x_1, x_2, x_3 are the natural coordinate functions. (*Hint:* Evaluate $V = \sum v_iU_i$ on x_j.)

5. If $V[f] = W[f]$ for every function f on $\mathbf{R}^3$, prove that $V = W$.

1.4 Curves in $\mathbf{R}^3$

Let I be an open interval in the real line R. We shall interpret this liberally to include not only the usual finite open interval $a < t < b$ (a, b real numbers), but also the infinite types $a < t$ (a half-line to $+\infty$), $t < b$ (a half-line to $-\infty$), and also the whole real line.

One can picture a curve in $\mathbf{R}^3$ as a trip taken by a moving point α. At each "time" t in some open interval, α is located at the point

$$\alpha(t) = (\alpha_1(t), \alpha_2(t), \alpha_3(t))$$

in $\mathbf{R}^3$. In rigorous terms then, α is a function from I to $\mathbf{R}^3$, and the real-valued functions α_1, α_2, α_3 are its *Euclidean coordinate functions.* Thus we write $\alpha = (\alpha_1, \alpha, \alpha_3)$, meaning, of course, that

$$\alpha(t) = (\alpha_1(t), \alpha_2(t), \alpha_3(t)) \quad \text{for all } t \text{ in } I.$$

We define the function α to be *differentiable* provided its (real-valued) coordinate functions are differentiable in the usual sense.

4.1 Definition A *curve* in $\mathbf{R}^3$ is a differentiable function $\alpha\colon I \to \mathbf{R}^3$ from an open interval I into $\mathbf{R}^3$.

We shall give several examples of curves, which will be used in Chapter 2 to experiment with results on the geometry of curves.

4.2 Example (1) ***Straight line.*** A line is the simplest type of curve in Euclidean space; its coordinate functions are linear (in the sense $t \to at + b$, not in the homogeneous sense $t \to at$). Explicitly, the curve $\alpha\colon \mathbf{R} \to \mathbf{R}^3$ such that

$$\alpha(t) = \mathbf{p} + t\mathbf{q} = (p_1 + tq_1,\, p_2 + tq_2,\, p_3 + tq_3) \quad (\mathbf{q} \neq 0)$$

is the *straight line* through the point $\mathbf{p} = \alpha(0)$ in the $\mathbf{q}$ direction.

(2) ***Helix***. (Fig. 1.7). The curve $t \to (a\cos t, a\sin t, 0)$ travels around a circle of radius $a > 0$ in the xy plane of $\mathbf{R}^3$. If we allow this curve to rise (or fall) at a constant rate, we obtain a *helix* $\alpha\colon \mathbf{R} \to \mathbf{R}^3$, given by the formula

$$\alpha(t) = (a \cos t, a \sin t, bt)$$

where $a > 0$, $b \neq 0$.

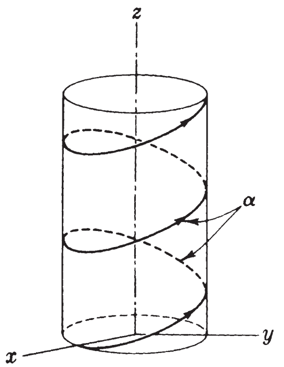

FIG. 1.7

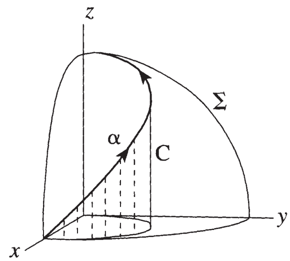

FIG. 1.8

(3) The curve

$$\alpha(t) = \left(1 + \cos t, \sin t, 2 \sin \frac{t}{2}\right) \quad \text{for all } t$$

has a noteworthy property: Let C be the cylinder in $\mathbf{R}^3$ over the circle in the xy plane with center at (1, 0, 0) and radius 1. Then α perpetually travels the route sliced from C by the sphere Σ with radius 2 and center at the origin. A segment of this route is shown in Fig. 1.8.

(4) The curve α: $\mathbf{R} \to \mathbf{R}^3$ such that

$$\alpha(t) = \left(e^t, e^{-t}, \sqrt{2}t\right)$$

shares with the helix in (2) the property of rising constantly. However, it lies over the hyperbola $xy = 1$ in the xy plane instead of over a circle.

(5) The 3-*curve* α: $\mathbf{R} \to \mathbf{R}^3$ is defined by

$$\alpha(t) = (3t - t^3, 3t^2, 3t + t^3).$$

If the coordinate functions of a curve are simple enough, its shape in $\mathbf{R}^3$ can be found, at least approximately, by plotting a few points. We could get a reasonable picture of curve α for $0 \leqq t \leqq 1$ by computing $\alpha(t)$ for $t = 0, 1/10, 1/2, 9/10, 1$.

If we visualize a curve α in $\mathbf{R}^3$ as a moving point, then at every time t there is a tangent vector at the point $\alpha(t)$ that gives the instantaneous velocity of α at that time. ◆

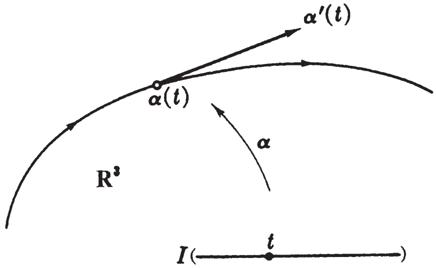

FIG. 1.9

4.3 Definition Let $\alpha\colon I \to \mathbf{R}^3$ be a curve in $\mathbf{R}^3$ with $\alpha = (\alpha_1, \alpha_2, \alpha_3)$. For each number t in I, the *velocity vector of* α *at* t is the tangent vector

$$\alpha'(t) = \left(\frac{d\alpha_1}{dt}(t), \frac{d\alpha_2}{dt}(t), \frac{d\alpha_3}{dt}(t)\right)_{\alpha(t)}$$

at the point $\alpha(t)$ in $\mathbf{R}^3$ (Fig. 1.9).

This definition can be interpreted geometrically as follows. The derivative at t of a real-valued function f on $\mathbf{R}$ is given by

$$\frac{df}{dt}(t) = \lim_{\Delta t \to 0} \frac{f(t + \Delta t) - f(t)}{\Delta t}.$$

This formula still makes sense if f is replaced by a curve $\alpha = (\alpha_1, \alpha_2, \alpha_3)$. In fact,

$$\frac{1}{\Delta t}(\alpha(t + \Delta t) - \alpha(t)) = \left(\frac{\alpha_1(t + \Delta t) - \alpha_1(t)}{\Delta t}, \frac{\alpha_2(t + \Delta t) - \alpha_2(t)}{\Delta t}, \frac{\alpha_3(t + \Delta t) - \alpha_3(t)}{\Delta t}\right).$$

This is the vector from $\alpha(t)$ to $\alpha(t + \Delta t)$, scalar multiplied by $1/\Delta t$ (Fig. 1.10).

Now, as Δt gets smaller, $\alpha(t + \Delta t)$ approaches $\alpha(t)$, and in the limit as $\Delta t \to 0$, we get a vector *tangent* to the curve α at the point $\alpha(t)$, namely,

$$\left(\frac{d\alpha_1}{dt}(t), \frac{d\alpha_2}{dt}(t), \frac{d\alpha_3}{dt}(t)\right).$$

As the figure suggests, the point of application of this vector must be the point $\alpha(t)$. Thus the standard limit operation for derivatives gives rise to our definition of the velocity of a curve.

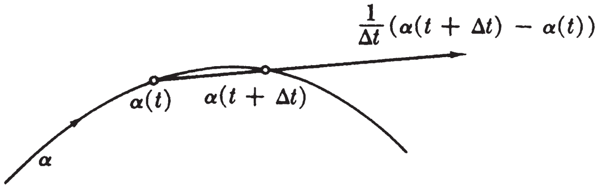

FIG. 1.10

An application of the identity

$$(v_1, v_2, v_3)_p = \sum v_i U_i(\mathbf{p})$$

to the velocity vector $\alpha'(t)$ at t yields the alternative formula

$$\alpha'(t) = \sum \frac{d\alpha_i}{dt}(t) U_i(\alpha(t)).$$

For example, the velocity of the straight line $\alpha(t) = \mathbf{p} + t\mathbf{q}$ is

$$\alpha'(t) = (q_1, q_2, q_3)_{\alpha(t)} = \mathbf{q}_{\alpha(t)}.$$

The fact that α is straight is reflected in the fact that all its velocity vectors are parallel; only the point of application changes as t changes.

For the helix

$$\alpha(t) = (a \cos t, a \sin t, bt),$$

the velocity is

$$\alpha'(t) = (-a \sin t, a \cos t, b)_{\alpha(t)}.$$

The fact that the helix rises constantly is shown by the constancy of the z coordinate of $\alpha'(t)$.

Given any curve, it is easy to construct new curves that follow the same route.

4.4 Definition Let $\alpha: I \to \mathbf{R}^3$ be a curve. If $h: J \to I$ is a differentiable function on an open interval J, then the composite function

$$\beta = \alpha(h): J \to \mathbf{R}^3$$

is a curve called a *reparametrization* of α by h.

For each $s \in J$, the new curve β is at the point $\beta(s) = \alpha(h(s))$ reached by α at $h(s)$ in I (Fig. 1.11). Thus β represents a different trip over at least part of the route of α.

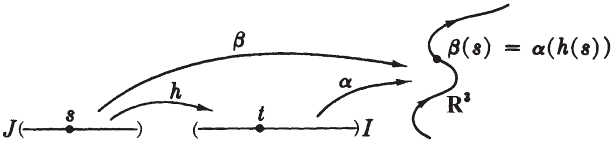

FIG. 1.11

To compute the coordinates of β, simply substitute $t = h(s)$ into the coordinates $\alpha_1(t)$, $\alpha_2(t)$, $\alpha_3(t)$ of α. For example, suppose

$$\alpha(t) = \left(\sqrt{t}, t\sqrt{t}, 1 - t\right) \quad on \quad I: 0 < t < 4.$$

If $h(s) = s^2$ on J: $0 < s < 2$, then the reparametrized curve is

$$\beta(s) = \alpha(h(s)) = \alpha(s^2) = (s, s^3, 1 - s^2).$$

The following lemma relates the velocities of a curve and of a reparametrization.

4.5 Lemma If β is the reparametrization of α by h, then

$$\beta'(s) = (dh/ds)(s)\alpha'(h(s)).$$

Proof. If $\alpha = (\alpha_1, \alpha_2, \alpha_3)$, then

$$\beta(s) = \alpha(h(s)) = (\alpha_1(h(s)), \alpha_2(h(s)), \alpha_3(h(s)).$$

Using the "prime" notation for derivatives, the chain rule for a composition of real-valued functions f and g reads $(g(f))' = g'(f) \cdot f'$. Thus, in the case at hand,

$$\alpha_i(h)'(s) = \alpha_i'(h(s)) \cdot h'(s).$$

By the definition of velocity, this yields

$$\begin{aligned}\beta'(s) &= \alpha(h)'(s) \\ &= (\alpha_1'(h(s)) \cdot h'(s), \alpha_2'(h(s)) \cdot h'(s), \alpha_3'(h(s)) \cdot h'(s)) \\ &= h'(s)\alpha'(h(s)).\end{aligned}$$ ◆

According to this lemma, to obtain the velocity of a reparametrization of α by h, first reparametrize α' by h, then scalar multiply by the derivative of h.

Since velocities are tangent vectors, we can take the derivative of a function with respect to a velocity.

4.6 Lemma Let α be a curve in $\mathbf{R}^3$ and let f be a differentiable function on $\mathbf{R}^3$. Then

$$\alpha'(t)[f] = \frac{d(f(\alpha))}{dt}(t).$$

Proof. Since

$$\alpha' = \left(\frac{d\alpha_1}{dt}, \frac{d\alpha_2}{dt}, \frac{d\alpha_3}{dt}\right)_\alpha,$$

we conclude from Lemma 3.2 that

$$\alpha'(t)[f] = \sum \frac{\partial f}{\partial x_i}(\alpha(t))\frac{d\alpha_i}{dt}(t).$$

But the composite function $f(\alpha)$ may be written $f(\alpha_1, \alpha_2, \alpha_3)$, and the chain rule then gives exactly the same result for the derivative of $f(\alpha)$. ◆

By definition, $\alpha'(t)[f]$ is the rate of change of f along the line through $\alpha(t)$ in the $\alpha'(t)$ direction. (If $\alpha'(t) \neq 0$, this is the tangent line to α at $\alpha(t)$; see Exercise 9.) The lemma shows that this rate of change is the same as that of f along the curve α itself.

Since a curve $\alpha\colon I \to \mathbf{R}^3$ is a function, it makes sense to say that α is one-to-one; that is, $\alpha(t) = \alpha(t_1)$ only if $t = t_1$. Another special property of curves is periodicity: A curve $\alpha\colon \mathbf{R} \to \mathbf{R}^3$ is *periodic* if there is a number $p > 0$ such that $\alpha(t + p) = \alpha(t)$ for all t—and the smallest such number p is then called the *period* of α.

From the viewpoint of calculus, the most important condition on a curve α is that it be *regular*, that is, have all velocity vectors different from zero. Such a curve can have no corners or cusps.

The following remarks about curves (offered without proof) describe another familiar way to formulate the concept of "curve." If f is a differentiable real-valued function on $\mathbf{R}^2$, let

$$C\colon f = a$$

be the set of all points $\mathbf{p}$ in $\mathbf{R}^2$ such that $f(\mathbf{p}) = a$. Now, if the partial derivatives $\partial f/\partial x$ and $\partial f/\partial y$ are never simultaneously zero at any point of C, then C consists of one or more separate "components," which we shall call *Curves.*† For example, $C\colon x^2 + y^2 = r^2$ is the circle of radius r centered at the

† The capital C distinguishes this notion from a (parametrized) curve $\alpha\colon I \to \mathbf{R}^2$.

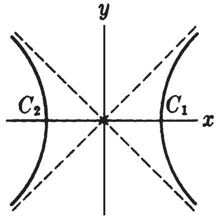

FIG. 1.12

origin of $\mathbf{R}^2$, and the hyperbola C: $x^2 - y^2 = r^2$ splits into two Curves ("branches") C_1 and C_2 as shown in Fig. 1.12.

Every Curve C is the route of many regular curves, called *parametrizations* of C. For example, the curve

$$\alpha(t) = (r \cos t, r \sin t)$$

is a well-known periodic parametrization of the circle given above, and for $r > 0$ the one-to-one curve

$$\beta(t) = (r \cosh t, r \sinh t)$$

parametrizes the branch $x > 0$ of the hyperbola.

Exercises

1. Compute the velocity vector of the curve in Example 4.2(3) for arbitrary t and for $t = 0$, $t = \pi/2$, $t = \pi$, visualizing those on Fig. 1.8.

2. Find the unique curve such that $\alpha(0) = (1, 0, 5)$ and $\alpha'(t) = (t^2, t, e^t)$.

3. Find the coordinate functions of the curve $\beta = \alpha(h)$, where α is the curve in Example 4.2(3) and $h(s) = \cos^{-1}(s)$ on J: $0 < s < 1$.

4. Reparametrize the curve α in Example 4.2(4) using $h(s) = \log s$ on J: $s > 0$. Check the equation in Lemma 4.5 in this case by calculating each side separately.

5. Find the equation of the straight line through the points $(1, -3, -1)$ and $(6, 2, 1)$. Does this line meet the line through the points $(-1, 1, 0)$ and $(-5, -1, -1)$?

6. Deduce from Lemma 4.6 that in the definition of directional derivative (Def. 3.1) the straight line $t \to \mathbf{p} + t\mathbf{v}$ can be replaced by any curve α with *initial velocity* $\mathbf{v}_p$, that is, such that $\alpha(0) = \mathbf{p}$ and $\alpha'(0) = \mathbf{v}_p$.

7. (*Continuation.*)

(a) Show that the curves with coordinate functions

$$(t, 1+t^2, t), \quad (\sin t, \cos t, t), \quad (\sinh t, \cosh t, t)$$

all have the same initial velocity $\mathbf{v}_p$.

(b) If $f = x^2 - y^2 + z^2$, compute $\mathbf{v}_p[f]$ by calculating $d(f(\alpha))/dt$ at $t = 0$, using each of three curves in (a).

8. Sketch the following Curves in $\mathbf{R}^2$, and find parametrizations for each.

(a) C: $4x^2 + y^2 = 1$, (b) C: $3x + 4y = 1$,

(c) C: $y = e^x$.

9. For a fixed t, the *tangent line* to a regular curve α at the point $\alpha(t)$ is the straight line $u \to \alpha(t) + u\alpha'(t)$, where we delete the point of application of $\alpha'(t)$. Find the tangent line to the helix $\alpha(t) = (2\cos t, 2\sin t, t)$ at the points $\alpha(0)$ and $\alpha(\pi/4)$.

1.5 1-Forms

If f is a real-valued function on $\mathbf{R}^3$, then in elementary calculus the differential of f is usually defined as

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz.$$

It is not always made clear exactly what this formal expression means. In this section we give a rigorous treatment using the notion of 1-form, and forms tend to appear at crucial moments in later work.

5.1 Definition A *1-form* ϕ on $\mathbf{R}^3$ is a real-valued function on the set of all tangent vectors to $\mathbf{R}^3$ such that ϕ is linear at each point, that is,

$$\phi(a\mathbf{v} + b\mathbf{w}) = a\phi(\mathbf{v}) + b\phi(\mathbf{w})$$

for any numbers a, b and tangent vectors $\mathbf{v}$, $\mathbf{w}$ at the same point of $\mathbf{R}^3$.

We emphasize that for every tangent vector $\mathbf{v}$, a 1-form ϕ defines a real number $\phi(\mathbf{v})$; and for each point $\mathbf{p}$ in $\mathbf{R}^3$, the resulting function ϕ_p: $T_p(\mathbf{R}^3) \to \mathbf{R}$ is linear. Thus at each point $\mathbf{p}$, ϕ_p is an element of the *dual space* of $T_p(\mathbf{R}^3)$. In this sense the notion of 1-form is dual to that of vector field.

The sum of 1-forms ϕ and ψ is defined in the usual pointwise fashion:

$$(\phi + \psi)(\mathbf{v}) = \phi(\mathbf{v}) + \psi(\mathbf{v}) \quad \text{for all tangent vectors } \mathbf{v}.$$

Similarly, if f is a real-valued function on $\mathbf{R}^3$ and ϕ is a 1-form, then $f\phi$ is the 1-form such that

$$(f\phi)(\mathbf{v}_p) = f(\mathbf{p})\phi(\mathbf{v}_p)$$

for all tangent vectors $\mathbf{v}_p$.

There is also a natural way to *evaluate a 1-form ϕ on a vector field V* to obtain a real-valued function $\phi(V)$: At each point $\mathbf{p}$ the value of $\phi(V)$ is the number $\phi(V(\mathbf{p}))$. Thus a 1-form may also be viewed as a machine that converts vector fields into real-valued functions. If $\phi(V)$ is differentiable whenever V is, we say that ϕ is *differentiable*. As with vector fields, we shall always assume that the 1-forms we deal with are differentiable.

A routine check of definitions shows that $\phi(V)$ is linear in both ϕ and V; that is,

$$\phi(fV + gW) = f\phi(V) + g\phi(W)$$

and

$$(f\phi + g\psi)(V) = f\phi(V) + g\psi(V),$$

where f and g are functions.

Using the notion of directional derivative, we now define a most important way to convert functions into 1-forms.

5.2 Definition If f is a differentiable real-valued function on $\mathbf{R}^3$, the *differential df* of f is the 1-form such that

$$df(\mathbf{v}_p) = \mathbf{v}_p[f] \quad \text{for all tangent vectors } \mathbf{v}_p.$$

In fact, df is a 1-form, since by definition it is a real-valued function on tangent vectors, and by (1) of Theorem 3.3 it is linear at each point $\mathbf{p}$. Clearly, df knows all rates of change of f in all directions on $\mathbf{R}^3$, so it is not surprising that differentials are fundamental to the calculus on $\mathbf{R}^3$.

Our task now is to show that these rather abstract definitions lead to familiar results when expressed in terms of coordinates.

5.3 Example 1-Forms on $\mathbf{R}^3$. (1) The differentials dx_1, dx_2, dx_3 of the natural coordinate functions. Using Lemma 3.2 we find

$$dx_i(\mathbf{v}_p) = \mathbf{v}_p[x_i] = \sum_j v_j \frac{\partial x_i}{\partial x_j}(\mathbf{p}) = \sum_j v_j \delta_{ij} = v_i,$$

where δ_{ij} is the Kronecker delta (0 if $i \neq j$, 1 if $i = j$). Thus *the value of* dx_i *on an arbitrary tangent vector* $\mathbf{v}_p$ *is the* ith *coordinate* v_i *of its vector part*—and does not depend on the point of application $\mathbf{p}$.

(2) The 1-form $\psi = f_1 dx_1 + f_2 dx_2 + f_3 dx_3$. Since dx_i is a 1-form, our definitions show that ψ is also a 1-form for any functions f_1, f_2, f_3. The value of ψ on an arbitrary tangent vector $\mathbf{v}_p$ is

$$\psi(\mathbf{v}_p) = (\sum f_i dx_i)(\mathbf{v}_p) = \sum f_i(\mathbf{p}) dx_i(\mathbf{v}_p) = \sum f_i(\mathbf{p}) v_i,$$

The first of these examples shows that the 1-forms dx_1, dx_2, dx_3 are the analogues for tangent vectors of the natural coordinate functions x_1, x_2, x_3 for points. Alternatively, we can view dx_1, dx_2, dx_3 as the "duals" of the natural unit vector fields U_1, U_2, U_3. In fact, it follows immediately from (1) above that the function $dx_i(U_j)$ has the constant value δ_{ij}.

We now show that every 1-form can be written in the concrete manner given in (2) above.

5.4 Lemma If ϕ is a 1-form on $\mathbf{R}^3$, then $\phi = \sum f_i dx_i$, where $f_i = \phi(U_i)$. These functions f_1, f_2, f_3 are called the *Euclidean coordinate functions* of ϕ.

Proof. By definition, a 1-form is a function on tangent vectors; thus ϕ and $\sum f_i dx_i$ are equal if and only if they have the same value on every tangent vector $\mathbf{v}_p = \sum v_i U_i(\mathbf{p})$. In (2) of Example 5.3 we saw that

$$(\sum f_i dx_i)(\mathbf{v}_p) = \sum f_i(\mathbf{p}) v_i.$$

On the other hand,

$$\phi(\mathbf{v}_p) = \phi(\sum v_i U_i(\mathbf{p})) = \sum v_i \phi(U_i(\mathbf{p})) = \sum v_i f_i(\mathbf{p})$$

since $f_i = \phi(U_i)$. Thus ϕ and $\sum f_i dx_i$ do have the same value on every tangent vector. ◆

This lemma shows that a 1-form on $\mathbf{R}^3$ is nothing more than an expression $f\,dx + g\,dy + h\,dz$, and such expressions are now rigorously defined as functions on tangent vectors. Let us now show that the definition of differential of a function (Definition 5.2) agrees with the informal definition given at the start of this section.

5.5 Corollary If f is a differentiable function on $\mathbf{R}^3$, then

$$df = \sum \frac{\partial f}{\partial x_i} dx_i.$$

Proof. The value of $\sum(\partial f/\partial x_i)dx_i$ on an arbitrary tangent vector $\mathbf{v}_p$ is $\sum(\partial f/\partial x_i)\,(\mathbf{p})v_i$. By Lemma 3.2, $df(\mathbf{v}_p) = \mathbf{v}_p[f]$ is the same. Thus the 1-forms df and $\sum(\partial f/\partial x_i)\,dx_i$ are equal. ◆

Using either this result or the definition of d, it is immediate that

$$d(f+g) = df + dg.$$

Finally, we determine the effect of d on *products* of functions and on *compositions* of functions.

5.6 Lemma Let fg be the product of differentiable functions f and g on $\mathbf{R}^3$. Then

$$d(fg) = g\,df + f\,dg.$$

Proof. Using Corollary 5.5, we obtain

$$\begin{aligned} d(fg) &= \sum \frac{\partial(fg)}{\partial x_i}dx_i = \sum\left(\frac{\partial f}{\partial x_i}g + f\frac{\partial g}{\partial x_i}\right)dx_i \\ &= g\left(\sum \frac{\partial f}{\partial x_i}dx_i\right) + f\left(\sum \frac{\partial g}{\partial x_i}dx_i\right) = g\,df + f\,dg. \end{aligned}$$ ◆

5.7 Lemma Let f: $\mathbf{R}^3 \to \mathbf{R}$ and h: $\mathbf{R} \to \mathbf{R}$ be differentiable functions, so the composite function $h(f)$: $\mathbf{R}^3 \to \mathbf{R}$ is also differentiable. Then

$$d(h(f)) = h'(f)\,df.$$

Proof. (The prime here is just the ordinary derivative, so $h'(f)$ is again a composite function, from $\mathbf{R}^3$ to $\mathbf{R}$.) The usual chain rule for a composite function such as $h(f)$ reads

$$\frac{\partial(h(f))}{\partial x_i} = h'(f)\frac{\partial f}{\partial x_i}.$$

Hence

$$d(h(f)) = \sum \frac{\partial(h(f))}{\partial x_i}dx_i = \sum h'(f)\frac{\partial f}{\partial x_i}dx_i = h'(f)\,df.$$ ◆

To compute df for a given function f it is almost always simpler to use these properties of d rather than substitute in the formula of Corollary 5.5. Then from df we immediately get the partial derivatives of f, and, in fact, *all its directional derivatives*. For example, suppose

$$f = (x^2 - 1)y + (y^2 + 2)z.$$

Then by Lemmas 5.6 and 5.7,

$$\begin{aligned} df &= (2x\,dx)y + (x^2 - 1)dy + (2y\,dy)z + (y^2 + 2)dz \\ &= \underbrace{2xy}_{\partial f/\partial x}\,dx + \underbrace{(x^2 + 2yz - 1)}_{\partial f/\partial y}dy + \underbrace{(y^2 + 2)}_{\partial f/\partial z}dz. \end{aligned}$$

Now use the rules above to evaluate this expression on a tangent vector $\mathbf{v}_p$. The result is

$$\mathbf{v}_p[f] = df(\mathbf{v}_p) = 2p_1p_2v_1 + (p_1^2 + 2p_2p_3 - 1)v_2 + (p_2^2 + 1)v_3.$$

Exercises

1. Let $\mathbf{v} = (1, 2, -3)$ and $\mathbf{p} = (0, -2, 1)$. Evaluate the following 1-forms on the tangent vector $\mathbf{v}_p$.

(a) $y^2\,dx$. (b) $z\,dy - y\,dz$.
(c) $(z^2 - 1)dx - dy + x^2\,dz$.

2. If $\phi = \sum f_i dx_i$ and $V = \sum v_i U_i$, show that the 1-form ϕ evaluated on the vector field V is the function $\phi(V) = \sum f_i v_i$.

3. Evaluate the 1-form $\phi = x^2\,dx - y^2\,dz$ on the vector fields
$V = xU_1 + yU_2 + zU_3$,
$W = xy\,(U_1 - U_3) + yz\,(U_1 - U_2)$, and $(1/x)V + (1/y)W$.

4. Express the following differentials in terms of df:

(a) $d(f^5)$. (b) $d\left(\sqrt{f}\right)$, where $f > 0$.
(c) $d(\log(1 + f^2))$.

5. Express the differentials of the following functions in the standard form $\sum f_i\,dx_i$.

(a) $(x^2 + y^2 + z^2)^{1/2}$. (b) $\tan^{-1}(y/x)$.

6. In each case compute the differential of f and find the directional derivative $\mathbf{v}_p[f]$, for $\mathbf{v}_p$ as in Exercise 1.

(a) $f = xy^2 - yz^2$. (b) $f = xe^{yz}$.
(c) $f = \sin(xy)\cos(xz)$.

7. Which of the following are 1-forms? In each case ϕ is the function on tangent vectors such that the value of ϕ on $(v_1, v_2, v_3)_p$ is

(a) $v_1 - v_3$.
(b) $p_1 - p_3$.
(c) $v_1 p_3 + v_2 p_1$.
(d) $\mathbf{v}_p[x^2 + y^2]$.
(e) 0.
(f) $(p_1)^2$.

In case ϕ is a 1-form, express it as $\sum f_i\, dx_i$.

8. Prove Lemma 5.6 directly from the definition of d.

9. A 1-form ϕ is *zero* at a point $\mathbf{p}$ provided $\phi(\mathbf{v}_p) = 0$ for all tangent vectors at $\mathbf{p}$. A point at which its differential df is zero is called a *critical point* of the function f. Prove that $\mathbf{p}$ is a critical point of f if and only if

$$\frac{\partial f}{\partial x}(\mathbf{p}) = \frac{\partial f}{\partial y}(\mathbf{p}) = \frac{\partial f}{\partial z}(\mathbf{p}) = 0.$$

Find all critical points of $f = (1 - x^2)y + (1 - y^2)z$.

(*Hint:* Find the partial derivatives of f by computing df.)

10. (*Continuation.*) Prove that the local maxima and local minima of f are critical points of f. (f has a *local maximum* at $\mathbf{p}$ if $f(\mathbf{q}) \leqq f(\mathbf{p})$ for all $\mathbf{q}$ near $\mathbf{p}$.)

11. It is sometimes asserted that df is the linear approximation of Δf.

(a) Explain the sense in which $(df)(\mathbf{v}_p)$ is a linear approximation of $f(\mathbf{p} + \mathbf{v}) - f(\mathbf{p})$.

(b) Compute exact and approximate values of $f(0.9, 1.6, 1.2) - f(1, 1.5, 1)$, where $f = x^2 y/z$.

1.6 Differential Forms

The 1-forms on $\mathbf{R}^3$ are part of a larger system called the *differential forms* on $\mathbf{R}^3$. We shall not give as rigorous an account of differential forms as we did of 1-forms since our use of the full system on $\mathbf{R}^3$ is limited. However, the *properties* established here are valid whenever differential forms are used.

Roughly speaking, a *differential form* on $\mathbf{R}^3$ is an expression obtained by adding and multiplying real-valued functions and the differentials dx_1, dx_2, dx_3 of the natural coordinate functions of $\mathbf{R}^3$. These two operations obey the usual associative and distributive laws; however, the multiplication is not commutative. Instead, it obeys the

$$\textit{alternation rule: } dx_i\, dx_j = -dx_j\, dx_i \quad (1 \leqq i, j \leqq 3).$$

This rule appears—although rather inconspicuously—in elementary calculus (see Exercise 9).

A consequence of the alternation rule is the fact that "repeats are zero," that is, $dx_i\,dx_i = 0$, since if $i = j$ the alternation rule reads

$$dx_i dx_i = -dx_i dx_i.$$

If each summand of a differential form contains p dx_i's ($p = 0, 1, 2, 3$), the form is called a *p-form*, and is said to have *degree p*. Thus, shifting to dx, dy, dz, we find

A 0-form is just a differentiable function f.

A 1-form is an expression $f\,dx + g\,dy + h\,dz$, just as in the preceding section.

A 2-form is an expression $f\,dx\,dy + g\,dx\,dz + h\,dy\,dz$.

A 3-form is an expression $f\,dx\,dy\,dz$.

We already know how to add 1-forms: simply add corresponding coefficient functions. Thus, in index notation,

$$\sum f_i dx_i + \sum g_i dx_i = \sum (f_i + g_i) dx_i.$$

The corresponding rule holds for 2-forms or 3-forms.

On three-dimensional Euclidean space, all p-forms with $p > 3$ are zero. This is a consequence of the alternation rule, for a product of more than three dx_i's must contain some dx_i twice, but repeats are zero, as noted above. For example, $dx\,dy\,dx\,dz = -dx\,dx\,dy\,dz = 0$, since $dx\,dx = 0$. As a reminder that the alternation rule is to be used, we denote this multiplication of forms by a *wedge* $\wedge$. (However, we do not bother with the wedge when only products of dx, dy, dz are involved.)

6.1 Example Computation of wedge products.

(1) Let

$$\phi = x\,dx - y\,dy \quad \text{and} \quad \psi = z\,dx + x\,dz.$$

Then

$$\begin{aligned} \phi \wedge \psi &= (x\,dx - y\,dy) \wedge (z\,dx + x\,dz) \\ &= xz\,dx\,dx + x^2\,dx\,dz - yz\,dy\,dx - yx\,dy\,dz. \end{aligned}$$

But $dx\,dx = 0$ and $dy\,dx = -dx\,dy$. Thus

$$\phi \wedge \psi = yz\,dx\,dy + x^2 dx\,dz - xy\,dy\,dz.$$

In general, the product of two 1-forms is a 2-form.

(2) Let ϕ and ψ be the 1-forms given above and let $\theta = z\,dy$. Then

$$\theta \wedge \phi \wedge \psi = yz^2\,dy\,dx\,dy + x^2z\,dy\,dx\,dz - xyz\,dy\,dy\,dz.$$

Since $dy\,dx\,dy$ and $dy\,dy\,dz$ each contain repeats, both are zero. Thus

$$\theta \wedge \phi \wedge \psi = -x^2z\,dx\,dy\,dz.$$

(3) Let ϕ be as above, and let η be the 2-form $y\,dx\,dz + x\,dy\,dz$. Omitting forms containing repeats, we find

$$\phi \wedge \eta = x^2\,dx\,dy\,dz - y^2\,dy\,dx\,dz = (x^2 + y^2)\,dx\,dy\,dz.$$

It should be clear from these examples that the wedge product of a p-form and a q-form is a $(p + q)$-form. Thus such a product is automatically zero whenever $p + q > 3$.

6.2 Lemma If ϕ and ψ are 1-forms, then

$$\phi \wedge \psi = -\psi \wedge \phi.$$

Proof. Write

$$\phi = \sum f_i\,dx_i, \quad \psi = \sum g_i\,dx_i.$$

Then by the alternation rule,

$$\phi \wedge \psi = \sum f_i g_j\,dx_i\,dx_j = -\sum g_j f_i\,dx_j\,dx_i = -\psi \wedge \phi. \qquad \blacklozenge$$

In the language of differential forms, the operator d of Definition 5.2 converts a 0-form f into a 1-form df. It is easy to generalize to an operator (also denoted by d) that converts a p-form η into a $(p + 1)$-form $d\eta$: One simply applies d (of Definition 5.2) to the coefficient functions of η. For example, here is the case $p = 1$.

6.3 Definition If $\phi = \sum f_i\,dx_i$ is a 1-form on $\mathbf{R}^3$, the *exterior derivative* of ϕ is the 2-form $d\phi = \sum df_i \wedge dx_i$.

If we expand the preceding definition using Corollary 5.5, we obtain the following interesting formula for the exterior derivative of

$$\phi = f_1dx_1 + f_2dx_2 + f_3dx_3:$$

$$d\phi = \left(\frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2}\right)dx_1dx_2 + \left(\frac{\partial f_3}{\partial x_1} - \frac{\partial f_1}{\partial x_3}\right)dx_1dx_3 + \left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3}\right)dx_2dx_3.$$

There is no need to memorize this formula; it is more reliable simply to apply the definition in each case. For example, suppose

$$\phi = xy\ dx + x^2\ dz.$$

Then

$$\begin{aligned} d\phi &= d(xy) \wedge dx + d(x^2) \wedge dz \\ &= (y\ dx + x\ dy) \wedge dx + (2x\ dx) \wedge dz \\ &= -x\ dx\ dy + 2x\ dx\ dz. \end{aligned}$$

It is easy to check that the general exterior derivative enjoys the same linearity property as the particular case in Definition 5.2; that is,

$$d(a\phi + b\psi) = a\ d\phi + b\ d\psi,$$

where ϕ and ψ are arbitrary forms and a and b are numbers.

The exterior derivative and the wedge product work together nicely:

6.4 Theorem Let f and g be functions, ϕ and ψ 1-forms. Then

(1) $d(fg) = df\ g + f\ dg$.
(2) $d(f\phi) = df \wedge \phi + f\ d\phi$.
(3) $d(\phi \wedge \psi) = d\phi \wedge \psi - \phi \wedge d\psi$.†

Proof. The first formula is just Lemma 5.6. We include it to show the family resemblance of all three formulas. The proof of (2) is a simpler version of that of (3), so we outline a proof of the latter—watching to see where the minus sign comes from.

It suffices to prove the formula when $\phi = f\ du$, $\psi = g\ dv$, where u and v are any of the coordinate functions x_1, x_2, x_3. In fact, every 1-form is a sum of such terms, so the general case will follow by the linearity of d and the algebra of wedge products.

For example, let us try the typical case $\phi = f\ dx$, $\psi = g\ dy$. Since repeats kill, there is no use writing down terms that are bound to be eliminated. Hence

$$d(\phi \wedge \psi) = d(fg\ dx\ dy) = \frac{\partial(fg)}{\partial z} dz\ dx\ dy = \left(f \frac{\partial g}{\partial z} + g \frac{\partial f}{\partial z} \right) dx\ dy\ dz. \quad (*)$$

† As usual, multiplication takes precedence over addition or subtraction, so this expression should be read as $(d\phi \wedge \psi) - (\phi \wedge d\psi)$.

Now,

$$d\phi \wedge \psi = d(f\ dx) \wedge g\ dy = \frac{\partial f}{\partial z}\ dz \wedge dx \wedge g\ dy = g\frac{\partial f}{\partial z}\ dx\ dy\ dz.$$

But

$$\phi \wedge d\psi = f\ dx \wedge d(g\ dy) = f\ dx \wedge \frac{\partial g}{\partial z}\ dz \wedge dy = -f\frac{\partial g}{\partial z}\ dx\ dy\ dz,$$

since $dx\ dz\ dy = -dx\ dy\ dz$. Thus we must *subtract* this last equation from its predecessor to get (∗). ◆

One way to remember the minus sign in equation (3) of the theorem is to treat d as if it were a 1-form. To reach ψ, d must change places with ϕ, hence the minus sign is consistent with the alternation rule in Lemma 6.2.

Differential forms, and the associated notions of wedge product and exterior derivative, provide the means of expressing quite complicated relations among the partial derivatives in a highly efficient way. The wedge product saves much useless labor by discarding, right at the start, terms that will eventually disappear. But the exterior derivative d is the key. Exercise 8 shows, for example, how it replaces all three of the differentiation operations of classical vector analysis.

Exercises

1. Let $\phi = yz\ dx + dz$, $\psi = \sin z\ dx + \cos z\ dy$, $\xi = dy + z\ dz$. Find the standard expressions (in terms of $dxdy, \ldots$) for

(a) $\phi \wedge \psi$, $\psi \wedge \xi$, $\xi \wedge \phi$. (b) $d\phi$, $d\psi$, $d\xi$.

2. Let $\phi = dx/y$ and $\psi = z\ dy$. Check the Leibnizian formula (3) of Theorem 6.4 in this case by computing each term separately.

3. For any function f show that $d(df) = 0$. Deduce that $d(f\ dg) = df \wedge dg$.

4. Simplify the following forms:

(a) $d(f\ dg + g\ df)$. (b) $d((f - g)\ (df + dg))$.
(c) $d(f\ dg \wedge g\ df)$. (d) $d(gf\ df) + d(f\ dg)$.

5. For any three 1-forms $\phi_i = \sum_j f_{ij} dx_j$ $(1 \leqq i \leqq 3)$, prove

$$\phi_1 \wedge \phi_2 \wedge \phi_3 = \begin{vmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{vmatrix} dx_1\ dx_2\ dx_3.$$

6. If r, ϑ, z are the cylindrical coordinate functions on $\mathbf{R}^3$, then $x = r\cos\vartheta$, $y = r\sin\vartheta$, $z = z$. Compute the *volume element* $dx\,dy\,dz$ of $\mathbf{R}^3$ in cylindrical coordinates. (That is, express $dx\,dy\,dz$ in terms of the functions r, ϑ, z, and their differentials.)

7. For a 2-form

$$\eta = f\,dx\,dy + g\,dx\,dz + h\,dy\,dz,$$

the *exterior derivative* $d\eta$ is defined to be the 3-form obtained by replacing f, g, and h by their differentials. Prove that for any 1-form ϕ, $d(d\phi) = 0$.

Exercises 3 and 7 show that $d^2 = 0$, that is, for any form ξ, $d(d\xi) = 0$. (If ξ is a 2-form, then $d(d\xi) = 0$, since its degree exceeds 3.)

8. Classical *vector analysis* avoids the use of differential forms on $\mathbf{R}^3$ by converting 1-forms and 2-forms into vector fields by means of the following one-to-one correspondences:

$$\sum f_i\,dx_i \overset{(1)}{\leftrightarrow} \sum f_i U_i \overset{(2)}{\leftrightarrow} f_3 dx_1 dx_2 - f_2 dx_1 dx_3 + f_1 dx_2 dx_3.$$

Vector analysis uses three basic operations based on partial differentiation:

Gradient of a function f:

$$\text{grad } f = \sum \frac{\partial f}{\partial x_i} U_i.$$

Curl of a vector field $V = \sum f_i U_i$:

$$\text{curl } V = \left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3}\right)U_1 + \left(\frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1}\right)U_2 + \left(\frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2}\right)U_3.$$

Divergence of a vector field $V = \sum f_i U_i$:

$$\text{div } V = \sum \frac{\partial f_i}{\partial x_i}.$$

Prove that all three operations may be expressed by exterior derivatives as follows:

(a) $df \overset{(1)}{\leftrightarrow} \text{grad } f$.

(b) If $\phi \overset{(1)}{\leftrightarrow} V$, then $d\phi \overset{(2)}{\leftrightarrow} \text{curl } V$.

(c) If $\eta \overset{(1)}{\leftrightarrow} V$, then $d\eta = (\text{div } V)\,dx\,dy\,dz$.

9. Let f and g be real-valued functions on $\mathbf{R}^2$. Prove that

$$df \wedge dg = \begin{vmatrix} \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} \\ \dfrac{\partial g}{\partial x} & \dfrac{\partial g}{\partial y} \end{vmatrix} dx\, dy.$$

This formula appears in elementary calculus; show that it implies the alternation rule.

1.7 Mappings

In this section we discuss functions from $\mathbf{R}^n$ to $\mathbf{R}^m$. If $n = 3$ and $m = 1$, then such a function is just a real-valued function on $\mathbf{R}^3$. If $n = 1$ and $m = 3$, it is a curve in $\mathbf{R}^3$. Although our results will necessarily be stated for arbitrary m and n, we are primarily interested in only three other cases:

$$\mathbf{R}^2 \to \mathbf{R}^2, \quad \mathbf{R}^2 \to \mathbf{R}^3, \quad \mathbf{R}^3 \to \mathbf{R}^3.$$

The fundamental observation about a function $F\colon \mathbf{R}^n \to \mathbf{R}^m$ is that it can be completely described by m real-valued functions on $\mathbf{R}^n$. (We saw this already in Section 4 for $n = 1$, $m = 3$.)

7.1 Definition Given a function $F\colon \mathbf{R}^n \to \mathbf{R}^m$, let $f_1, f_2, \ldots, f_m$, denote the real-valued functions on $\mathbf{R}^n$ such that

$$F(\mathbf{p}) = (f_1(\mathbf{p}), f_2(\mathbf{p}), \ldots, f_m(\mathbf{p}))$$

for all points $\mathbf{p}$ in $\mathbf{R}^n$. These functions are called the *Euclidean coordinate functions* of F, and we write $F = (f_1, f_2, \ldots, f_m)$.

The function F is *differentiable* provided its coordinate functions are differentiable in the usual sense. A differentiable function $F\colon \mathbf{R}^n \to \mathbf{R}^m$ is called a *mapping* from $\mathbf{R}^n$ to $\mathbf{R}^m$.

Note that the coordinate functions of F are the composite functions $f_i = x_i(F)$, where $x_1, \ldots, x_m$ are the coordinate functions of $\mathbf{R}^m$.

Mappings may be described in many different ways. For example, suppose $F\colon \mathbf{R}^3 \to \mathbf{R}^3$ is the mapping $F = (x^2, yz, xy)$. Thus

$$F(\mathbf{p}) = \left(x(\mathbf{p})^2, y(\mathbf{p})z(\mathbf{p}), x(\mathbf{p})y(\mathbf{p})\right) \quad \text{for all } \mathbf{p}.$$

Now, $\mathbf{p} = (p_1, p_2, p_3)$, and by definition of the coordinate functions,

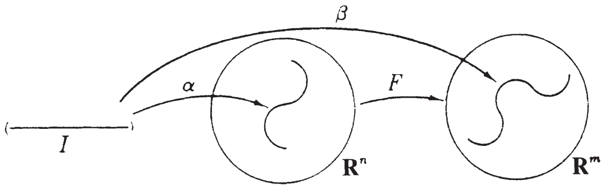

FIG. 1.13

$$x(\mathbf{p}) = p_1,\ y(\mathbf{p}) = p_2,\ z(\mathbf{p}) = p_3.$$

Hence we obtain the following *pointwise* formula for F:

$$F(p_1, p_2, p_3) = (p_1^2, p_2 p_3, p_1 p_2) \quad \text{for all } p_1, p_2, p_3.$$

Thus, for example,

$$F(1, -2, 0) = (1, 0, -2) \quad \text{and} \quad F(-3, 1, 3) = (9, 3, -3).$$

In principle, one could deduce the theory of curves from the general theory of mappings. But curves are reasonably simple, while a mapping, even in the case $\mathbf{R}^2 \to \mathbf{R}^2$, can be quite complicated. Hence we reverse this process and use curves, at every stage, to gain an understanding of mappings.

7.2 Definition If $\alpha: I \to \mathbf{R}^n$ is a curve in $\mathbf{R}^n$ and $F: \mathbf{R}^n \to \mathbf{R}^m$ is a mapping, then the composite function $\beta = F(\alpha): I \to \mathbf{R}^m$ is a curve in $\mathbf{R}^m$ called the *image of* α *under* F (Fig. 1.13).

7.3 Example Mappings. (1) Consider the mapping $F: \mathbf{R}^3 \to \mathbf{R}^3$ such that

$$F = (x - y, x + y, 2z).$$

In pointwise terms then,

$$F(p_1, p_2, p_3) = (p_1 - p_2, p_1 + p_2, 2p_3) \quad \text{for all } p_1, p_2, p_3.$$

Only when a mapping is quite simple can one hope to get a good idea of its behavior by merely computing its values at some finite number of points. But this function *is* quite simple —it is a *linear* transformation from $\mathbf{R}^3$ to $\mathbf{R}^3$.

Thus by a well-known theorem of linear algebra, F is completely determined by its values at three (linearly independent) points, say the *unit points*

$$\mathbf{u}_1 = (1, 0, 0), \quad \mathbf{u}_2 = (0, 1, 0), \quad \mathbf{u}_3 = (0, 0, 1).$$

(2) The mapping $F\colon \mathbf{R}^2 \to \mathbf{R}^2$ such that $F(u, v) = (u^2 - v^2, 2uv)$. (Here u and v are the coordinate functions of $\mathbf{R}^2$.) To analyze this mapping, we examine its effect on the curve $\alpha(t) = (r\cos t, r\sin t)$, where $0 \leqq t \leqq 2\pi$. This curve takes one counterclockwise trip around the circle of radius r with center at the origin. The image curve is

$$\begin{aligned}\beta(t) &= F(\alpha(t)) \\ &= F(r \cos t, r \sin t) \\ &= (r^2 \cos^2 t - r^2 \sin^2 t, 2r^2 \cos t \sin t)\end{aligned}$$

with $0 \leqq t \leqq 2\pi$. Using the trigonometric identities

$$\cos 2t = \cos^2 t - \sin^2 t, \quad \sin 2t = 2 \sin t \cos t,$$

we find for $\beta = F(\alpha)$ the formula

$$\beta(t) = (r^2 \cos 2t, r^2 \sin 2t),$$

with $0 \leqq t \leqq 2\pi$. This curve takes *two* counterclockwise trips around the circle of radius r^2 centered at the origin (Fig. 1.14).

Thus the effect of F is to wrap the plane $\mathbf{R}^2$ smoothly around itself twice—leaving the origin fixed, since $F(0, 0) = (0, 0)$. In this process, each circle of radius r is wrapped twice around the circle of radius r^2.

Generally speaking, differential calculus deals with approximation of smooth objects by linear objects. The best-known case is the approximation of a differentiable real-valued function f near x by the linear function $\Delta x \to f'(x)\,\Delta x$, which gives the tangent line at x to the graph of f. Our goal now is to define an analogous linear approximation for a mapping $F\colon \mathbf{R}^n \to \mathbf{R}^m$ near a point $\mathbf{p}$ of $\mathbf{R}^n$.

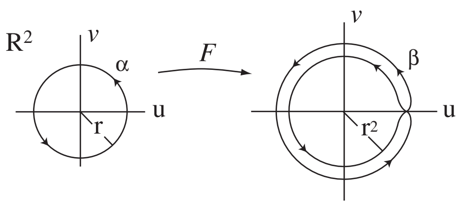

FIG. 1.14

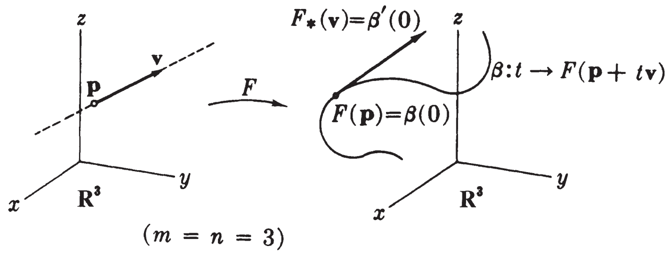

FIG. 1.15

Since $\mathbf{R}^n$ is filled by the radial lines $\alpha(t) = \mathbf{p} + t\mathbf{v}$ starting at $\mathbf{p}$, $\mathbf{R}^m$ is filled by their image curves $\beta(t) = F(\mathbf{p} + t\mathbf{v})$ starting at $F(\mathbf{p})$ (Fig. 1.15). So we approximate F near $\mathbf{p}$ by the map F_* that sends each initial velocity $\alpha'(0) = \mathbf{v}_p$ to the initial velocity $\beta'(0)$.

7.4 Definition Let F: $\mathbf{R}^n \to \mathbf{R}^m$ be a mapping. If $\mathbf{v}$ is a tangent vector to $\mathbf{R}^n$ at $\mathbf{p}$, let $F_*(\mathbf{v})$ be the initial velocity of the curve $t \to F(\mathbf{p} + t\mathbf{v})$. The resulting function F_* sends tangent vectors to $\mathbf{R}^n$ to tangent vectors to $\mathbf{R}^m$, and is called the *tangent map* of F.

The tangent map can be described explicitly as follows.

7.5 Proposition Let $F = (f_1, f_2, \ldots, f_m)$ be a mapping from $\mathbf{R}^n$ to $\mathbf{R}^m$. If $\mathbf{v}$ is a tangent vector to $\mathbf{R}^n$ at $\mathbf{p}$, then

$$F_*(\mathbf{v}) = (\mathbf{v}[f_1], \ldots, \mathbf{v}[f_m]) \quad \text{at } F(\mathbf{p}).$$

Proof. For definiteness, take $m = 3$. Then

$$\beta(t) = F(\mathbf{p} + t\mathbf{v}) = (f_1(\mathbf{p} + t\mathbf{v}), f_2(\mathbf{p} + t\mathbf{v}), f_3(\mathbf{p} + t\mathbf{v})).$$

By definition, $F_*(\mathbf{v}) = \beta'(0)$. To get $\beta'(0)$, we take the derivatives, at $t = 0$, of the coordinate functions of β (Definition 4.3). But

$$\frac{d}{dt}(f_i(\mathbf{p} + t\mathbf{v}))\Big|_{t=0} = \mathbf{v}[f_i].$$

Thus

$$F_*(\mathbf{v}) = (\mathbf{v}[f_1],\ \mathbf{v}[f_2],\ \mathbf{v}[f_3])\big|_{\beta(0)},$$

and $\beta(0) = F(\mathbf{p})$. ◆

Fix a point $\mathbf{p}$ of $\mathbf{R}^n$. The definition of tangent map shows that $F*$ sends tangent vectors at $\mathbf{p}$ to tangent vectors at $F(\mathbf{p})$. Thus for each $\mathbf{p}$ in $\mathbf{R}^n$, the function $F*$ gives rise to a function

$$F_{*p}\colon T_p(\mathbf{R}^n) \to T_{F(p)}(\mathbf{R}^m)$$

called the tangent map of F *at* $\mathbf{p}$. (Compare the analogous situation in elementary calculus where a function $f\colon \mathbf{R} \to \mathbf{R}$ has a derivative function $f'\colon \mathbf{R} \to \mathbf{R}$ that at each point t of $\mathbf{R}$ gives the derivative of f at t.)

7.6 Corollary If $F\colon \mathbf{R}^n \to \mathbf{R}^m$ is a mapping, then at each point $\mathbf{p}$ of $\mathbf{R}^n$ the tangent map $F_{*p}\colon T_p(\mathbf{R}^n) \to T_{F(p)}(\mathbf{R}^m)$ is a linear transformation.

Proof. We must show that for tangent vectors $\mathbf{v}$ and $\mathbf{w}$ at $\mathbf{p}$ and numbers a, b,

$$F_*(a\mathbf{v} + b\mathbf{w}) = aF_*(\mathbf{v}) + bF_*(\mathbf{w}).$$

This follows immediately from the preceding proposition by using the linearity in assertion (1) of Theorem 3.3. ◆

In fact, *the tangent map* F_{*p} at $\mathbf{p}$ *is the linear transformation that best approximates F near* $\mathbf{p}$. This idea is fully developed in advanced calculus, where it is used to prove Theorem 7.10.

Another consequence of the proposition is that *mappings preserve velocities of curves*. Explicitly:

7.7 Corollary Let $F\colon \mathbf{R}^n \to \mathbf{R}^m$ be a mapping. If $\beta = F(\alpha)$ is the image of a curve α in $\mathbf{R}^n$, then $\beta' = F*(\alpha')$.

Proof. Again, set $m = 3$. If $F = (f_1, f_2, f_3)$, then

$$\beta = F(\alpha) = (f_1(\alpha), f_2(\alpha), f_3(\alpha)).$$

Hence Theorem 7.5 gives

$$F_*(\alpha') = (\alpha'[f_1], \alpha'[f_2], \alpha'[f_3]).$$

But by Lemma 4.6,

$$\alpha'[f_i] = \frac{df_i(\alpha)}{dt}.$$

Hence

$$F_*(\alpha'(t)) = \left(\frac{df_1(\alpha)}{dt}(t), \frac{df_2(\alpha)}{dt}(t), \frac{df_3(\alpha)}{dt}(t)\right)_{\beta(t)} = \beta'(t). \qquad \blacklozenge$$

Let $\{U_j\}$ $(1 \le j \le n)$ and $\{\overline{U}_i\}$ $(1 \le i \le m)$ be the natural frame fields of $\mathbf{R}^n$ and $\mathbf{R}^m$, respectively (Def. 2.4). Then:

7.8 Corollary *If* $F = (f_1, \ldots, f_m)$ *is a mapping from* $\mathbf{R}^n$ *to* $\mathbf{R}^m$, *then*

$$F_*(U_j(\mathbf{p})) = \sum_{i=1}^{m} \frac{\partial f_i}{\partial x_j}(\mathbf{p})\overline{U}_i(F(\mathbf{p})) \quad (1 \le j \le n).$$

Proof. This follows directly from Proposition 7.5, since $U_j[f_i] = \dfrac{\partial f_i}{\partial x_j}$. $\blacklozenge$

The matrix appearing in the preceding formula,

$$\left(\frac{\partial f_i}{\partial x_j}(\mathbf{p})\right)_{1 \le i \le m,\ 1 \le j \le n}$$

is called the *Jacobian matrix* of F at $\mathbf{p}$. (When $m = n = 1$; it reduces to a single number: the derivative of F at $\mathbf{p}$.)

Just as the derivative of a function is used to gain information about the function, the tangent map F_* can be used in the study of a mapping F.

7.9 Definition A mapping F: $\mathbf{R}^n \to \mathbf{R}^m$ is *regular* provided that at every point $\mathbf{p}$ of $\mathbf{R}^n$ the tangent map F_{*p} is one-to-one.

Since tangent maps are linear transformations, standard results of linear algebra show that the following conditions are equivalent:

(1) F_{*p} is one-to-one.
(2) $F_*(\mathbf{v}_p) = 0$ implies $\mathbf{v}_p = 0$.
(3) The Jacobian matrix of F at $\mathbf{p}$ has rank n, the dimension of the domain $\mathbf{R}^n$ of F.

The following noteworthy property of linear transformations $T: V \to W$ will be useful in dealing with tangent maps. If the vector spaces V and W have the same dimension, then T is one-to-one if and only if it is onto, so either property is equivalent to T being a linear isomorphism.

A mapping that has a (differentiable) inverse mapping is called a *diffeomorphism*. The results of this section all remain valid when Euclidean spaces $\mathbf{R}^n$ are replaced by open sets of Euclidean spaces, so we can speak of a diffeomorphism from one open set to another.

We state without proof one of the fundamental results of advanced calculus.

7.10 Theorem Let $F: \mathbf{R}^n \to \mathbf{R}^n$ be a mapping between Euclidean spaces of the same dimension. If F_{*p} is one-to-one at a point $\mathbf{p}$, there is an open set $\mathscr{U}$ containing $\mathbf{p}$ such that F restricted to $\mathscr{U}$ is a diffeomorphism of $\mathscr{U}$ onto an open set $\mathscr{V}$.

This is called the *inverse function theorem* since it asserts that the restricted mapping $\mathscr{U} \to \mathscr{V}$ has a differentiable inverse mapping $\mathscr{V} \to \mathscr{U}$. Exercise 6 gives a suggestion of its importance.

Exercises

In the first four exercises F denotes the mapping $F(u, v) = (u^2 - v^2, 2uv)$ in Example 7.3.

1. Find all points $\mathbf{p}$ such that
(a) $F(\mathbf{p}) = (0, 0)$. (b) $F(\mathbf{p}) = (8, 6)$.
(c) $F(\mathbf{p}) = \mathbf{p}$.

2. (a) Sketch the horizontal line $v = 1$ and its image under F (a parabola).
(b) Do the same for the vertical $u = 1$.
(c) Describe the image of the unit square $0 \leqq u, v \leqq 1$ under F.

3. Let $\mathbf{v} = (v_1, v_2)$ be a tangent vector to $\mathbf{R}^2$ at $\mathbf{p} = (p_1, p_2)$. Apply Definition 7.4 directly to express $F_*(\mathbf{v})$ in terms of the coordinates of $\mathbf{v}$ and $\mathbf{p}$.

4. Find a formula for the Jacobian matrix of F at all points, and deduce that F_{*p} is a linear isomorphism at every point of $\mathbf{R}^2$ except the origin.

5. If $F: \mathbf{R}^n \to \mathbf{R}^m$ is a linear transformation, prove that $F_*(\mathbf{v}_p) = F(\mathbf{v})_{F(p)}$.

6. (a) Give an example to demonstrate that a one-to-one and onto mapping need not be a diffeomorphism. (*Hint:* Take $m = n = 1$.)

(b) Prove that if a one-to-one and onto mapping F: $\mathbf{R}^n \to \mathbf{R}^n$ is regular, then it is a diffeomorphism.

7. Prove that a mapping F: $\mathbf{R}^n \to \mathbf{R}^m$ preserves directional derivatives in this sense: If $\mathbf{v}_p$ is a tangent vector to $\mathbf{R}^n$ and g is a differentiable function on $\mathbf{R}^m$, then $F_*(\mathbf{v}_p)[g] = \mathbf{v}_p[g(F)]$.

8. In the definition of tangent map (Def. 7.4), the straight line $t \to \mathbf{p} + t\mathbf{v}$ can be replaced by any curve α with initial velocity $\mathbf{v}_p$.

9. Let F: $\mathbf{R}^n \to \mathbf{R}^m$ and G: $\mathbf{R}^m \to \mathbf{R}^p$ be mappings. Prove:
(a) Their composition GF: $\mathbf{R}^n \to \mathbf{R}^p$ is a (differentiable) mapping. (Take $m = p = 2$ for simplicity.)
(b) $(GF)_* = G_*F_*$. (*Hint:* Use the preceding exercise.)
This concise formula is the general *chain rule*. Unless dimensions are small, it becomes formidable when expressed in terms of Jacobian matrices.
(c) If F is a diffeomorphism, then so is its inverse mapping F^{-1}.

10. Show (in two ways) that the map F: $\mathbf{R}^2 \to \mathbf{R}^2$ such that $F(u, v) = (ve^u, 2u)$ is a diffeomorphism:
(a) Prove that it is one-to-one, onto, and regular;
(b) Find a formula for its inverse F^{-1}: $\mathbf{R}^2 \to \mathbf{R}^2$ and observe that F^{-1} is differentiable. Verify the formula by checking that both $F\,F^{-1}$ and $F^{-1}\,F$ are identity maps.

1.8 Summary

Starting from the familiar notion of real-valued functions and using linear algebra at every stage, we have constructed a variety of mathematical objects. The basic notion of tangent vector led to vector fields, which dualized to 1-forms—which in turn led to arbitrary differential forms. The notions of curve and differentiable function were generalized to that of a mapping F: $\mathbf{R}^n \to \mathbf{R}^m$.

Then, starting from the usual notion of the derivative of a real-valued function, we proceeded to construct appropriate differentiation operations for these objects: the directional derivative of a function, the exterior derivative of a form, the velocity of a curve, the tangent map of a mapping. These operations all reduced to (ordinary or partial) derivatives of real-valued coordinate functions, but it is noteworthy that in most cases the *definitions* of these operations did not involve coordinates. (This could be achieved in all cases.) Generally speaking, these differentiation operations all exhibited in one form or another the characteristic linear and Leibnizian properties of ordinary differentiation.

Most of these concepts are probably already familiar to the reader, at least in special cases. But we now have careful definitions and a catalogue of basic properties that will enable us to begin the exploration of differential geometry.