Some topics related to differential forms on $\mathbf{R}^n$

Stephen Semmes
Rice University

Preface

These informal notes are largely concerned with "calculus on manifolds", where "manifold" will normally mean something like a reasonably nice submanifold of $\mathbf{R}^n$ for some positive integer n, perhaps with some sort of reasonable nice boundary. The reader is expected to have some familiarity with basic analysis, at least on Euclidean spaces, as well as linear algebra. Of course, the reader should also be familiar with multivariable calculus, and my colleague Frank Jones' book [93] is a very helpful resource.

A basic question to be addressed here is

What is a *differential form*?

Normally we might be concerned with differential forms that depend on a point, and we may also be interested in continuity or differentiability properties of the form, as a function of the point. In order to answer the previous question, a more basic question to be addressed is

What is a differential form at a point?

This involves multilinear functions, which will be discussed starting in Chapter 1.

More precisely, differential forms involve *alternating multilinear functions*, and we shall look at *symmetric multilinear functions* some too. This is related to permutations on finite sets, and the notion of *even* and *odd* permutations in particular. A bit of group theory is relevant here, but the reader is not necessarily expected to be familiar with that already.

We shall be interested in *multiplying* differential forms as well. This is related to some other aspects of abstract algebra, and again the reader is not necessarily expected to be familiar with that already. Certain products of multilinear forms will be discussed in Chapter 2, corresponding to multiplying differential forms at a point in particular. Differential forms are formally defined in Chapter 3, along with related continuity and differentiability properties.

In Chapter 2, we shall also consider linear mappings between vector spaces, and their relationship with multilinear functions on these vector spaces. This will be used in Chapter 4 to *pull* differential forms *back* using differentiable mappings. This works for some other types of *tensor fields* too, as we shall see.

There is a particular type of derivative of a differential form known as the *exterior derivative*, which will be discussed in Chapter 4 as well. This includes exterior derivatives of products, and the relationship between the exterior derivative and pulling differential forms back using differentiable mappings.

We shall sometimes be concerned with n-dimensional volumes of reasonably nice subsets of $\mathbf{R}^n$, and integrals of continuous real-valued functions on them. The reader is not necessarily expected to be familiar with Lebesgue measure or integration, which permits one to deal with wider classes of sets and functions. See Frank Jones' book [92], for instance.

The *inverse function theorem* is reviewed in Chapter 5, without getting into the proof, except for a few basic points. This includes some preliminary remarks about matrix-valued functions and invertible matrices. *Diffeomorphisms* between open subsest of $\mathbf{R}^n$ are discussed as well. Some related aspects of analysis on Euclidean spaces are mentioned too, which will also be helpful later. Among the various topics are some constructions of smooth real-valued functions of one or more variables, and cells in $\mathbf{R}^n$.

Chapter 6 is largely concerned with integration of differential forms. This means integration of differential n-forms on $\mathbf{R}^n$, and on suitable n-dimensional objects in higher-dimensional Euclidean spaces. We shall also look at the role of the injectivity of the differential of a continuously-differentiable mapping from an open subset of $\mathbf{R}^n$ into $\mathbf{R}^m$ when $m > n$. The *implicit function theorem* is related to the surjectivity of the differential of a continuously-differentiable mapping from an open subset of $\mathbf{R}^m$ into $\mathbf{R}^n$ when $m > n$. We shall consider whether a diffeomorphism between open subsets of $\mathbf{R}^n$ preserves or reverses orientations as well.

Chapter 7 deals with *n-simplices* in $\mathbf{R}^m$ parameterized by various types of mappings, starting with affine mappings. We also consider *n-chains*, which correspond to formal sums of finitely many parameterized n-simplices with integer coefficients. The *boundary* of a parameterized n-simplex can be defined as an $(n-1)$-chain, for instance. Under suitable conditions, *Stokes' theorem* says that the integral of a differential $(n-1)$-form ω over the boundary of an n-chain Γ is equal to the integral of the exterior derivative of ω over Γ. This is discussed in the first five sections of Chapter 8.

Vector fields are discussed briefly in Chapter 4, and we return to this in Chapter 8. In particular, we consider *Lie brackets* of vector fields with suitable regularity on open sets in $\mathbf{R}^n$, which corresponds to commutators of the associated first-order differential operators. A version of the *Poincaré lemma* is discussed as well.

Lie derivatives of functions and vector fields are discussed in Section 8.8, and we continue with Lie derivatives of some other tensor fields in Chapter 9. More precisely, tensor fields of type $(0, k)$ were discussed in Sections 4.1 and 4.2, and vector fields may be considered as tensor fields of type $(1, 0)$, as in Section 4.3. We shall also consider tensor fields of type $(1, k)$ on open sets in $\mathbf{R}^n$. Interior products of multilinear forms on a vector space by elements of the vector space are discussed as well.

Some related topics involving linear and abstract algebra, Lie algebras, and complex analysis are discussed in the appendices. A number of books and other references may be found in the bibliography with more information and various perspectives. Of course, there are many other texts dealing with questions like these, and adjacent areas of mathematics.

There are also some places along the way with some elaborations or other matters that may not be neeeded for the moment, and that are related to the broader subject. The reader may want to skip some of these parts, at least temporarily, or simply not dwell on them too much at the beginning.

Contents

Chapter 1

Some helpful algebra

1.1 The binomial theorem

Let n be a positive integer, and let x, y be real numbers. It is well known that

$$(x+y)^n = \sum_{j=0}^{n} \binom{n}{j} x^j \, y^{n-j}, \tag{1.1.1}$$

where the *binomial coefficients* $\binom{n}{j}$ are given by

$$\binom{n}{j} = \frac{n!}{j! \, (n-j)!}. \tag{1.1.2}$$

Of course, if l is a positive integer, then

$$l! = \prod_{k=1}^{l} k, \tag{1.1.3}$$

which is interpreted as being equal to 1 when $l = 0$.

It is easy to see that $(x+y)^n$ can be expressed as in the right side of (1.1.1) for some coefficients $\binom{n}{j}$, which are positive integers. The expression (1.1.2) for these coefficients can be verified using induction. This can also be obtained using calculus. One can check that these are the only values of the coefficients for which (1.1.1) holds for all x, y. Note that $\binom{n}{j}$ is the same as the number of ways of choosing j elements from a set with n elements.

1.2 Polynomials on $\mathbf{R}^n$

The real line is denoted $\mathbf{R}$, as usual. Let n be a positive integer, and let $\mathbf{R}^n$ be the space of n-tuples $x = (x_1, \ldots, x_n)$ of real numbers. If $x, y \in \mathbf{R}^n$, then $x + y \in \mathbf{R}^n$ is defined by adding the coordinates of x and y, so that

$$x + y = (x_1 + y_1, \ldots, x_n + y_n). \tag{1.2.1}$$

Similarly, if $x \in \mathbf{R}^n$ and $t \in \mathbf{R}$, then $t\,x \in \mathbf{R}^n$ is defined by multiplying the coordinates of x by t, so that

$$t\,x = (t\,x_1, \ldots, t\,x_n). \tag{1.2.2}$$

It is well known that $\mathbf{R}^n$ is a vector space over the real numbers with respect to these operations of addition and scalar multiplication.

In order to talk about polynomials on $\mathbf{R}^n$, it will be helpful to introduce some additional notation and terminology. By a *multi-index* we mean an n-tuple $\alpha = (\alpha_1, \ldots, \alpha_n)$ of nonnegative integers. If α is a multi-index and $x \in \mathbf{R}^n$, then we put

$$x^\alpha = \prod_{j=1}^n x_j^{\alpha_j}. \tag{1.2.3}$$

Here $x_j^{\alpha_j}$ is interpreted as being equal to 1 when $\alpha_j = 0$, even if $x_j = 0$. We may refer to x^α as the *monomial* associated to α.

A *polynomial* on $\mathbf{R}^n$ may be expressed as a linear combination of finitely many monomials. This will be discussed further in the next section.

1.2.1 Homogeneous polynomials

Let d be a nonnegative integer, and suppose for the moment that

$$\sum_{j=1}^n \alpha_j = d. \tag{1.2.4}$$

Put

$$f_\alpha(x) = x^\alpha \tag{1.2.5}$$

for each $x \in \mathbf{R}^n$, for convenience. Observe that

$$f_\alpha(t\,x) = t^d\, f_\alpha(x) \tag{1.2.6}$$

for every $t \in \mathbf{R}$ and $x \in \mathbf{R}^n$. Thus f_α is *homogeneous of degree* d on $\mathbf{R}^n$. A polynomial on $\mathbf{R}^n$ is homogeneous of degree d if it can be expressed as a linear combination of monomials x^α corresponding to multi-indices α that satisfy (1.2.4).

It is well known that the number of multi-indices α satisfying (1.2.4) is equal to

$$\binom{n+d-1}{n-1}. \tag{1.2.7}$$

This is the same as the number of ways of choosing $n-1$ elements from a set of $n+d-1$ elements, as mentioned in the previous section. More precisely, suppose that we have a set of $n+d-1$ elements that is linearly ordered, like the set of positive integers from 1 to $n+d-1$. If we choose $n-1$ elements from this set, then the other d elements of the set can be partitioned into n subsets in a natural way, some of which may be empty. The numbers of the elements of these n subsets correspond exactly to a multi-index α that satisfies (1.2.4).

1.3 More on polynomials

Let n be a positive integer, and let N be a nonnegative integer. A *polynomial* on $\mathbf{R}^n$ of degree less than or equal to N with real coefficients may be expressed as

$$p(x) = \sum_{\alpha_1+\cdots+\alpha_n \le N} c_\alpha \, x^\alpha. \tag{1.3.1}$$

More precisely, the sum is taken over all multi-indices α with

$$\alpha_1 + \cdots + \alpha_n \le N, \tag{1.3.2}$$

and it is easy to see that there are only finitely many such multi-indices. The coefficient c_α is supposed to be a real number for each of these multi-indices α, so that (1.3.1) defines a real-valued function on $\mathbf{R}^n$.

It is well known that the coefficients c_α are uniquely determined by the corresponding function $p(x)$ on $\mathbf{R}^n$. This can be obtained by considering p and its derivatives at 0.

Let $\mathcal{P}(\mathbf{R}^n)$ be the space of all polynomials on $\mathbf{R}^n$ with real coefficients. This is a linear subspace of the space of all real-valued functions on $\mathbf{R}^n$, so that $\mathcal{P}(\mathbf{R}^n)$ may be considered as a vector space over the real numbers with respect to pointwise addition and scalar multiplication of functions. Of course, pointwise addition and scalar multiplication of polynomials on $\mathbf{R}^n$ corresponds exactly to termwise addition and scalar multiplications of sums as in the right side of (1.3.1).

1.3.1 More on homogeneous polynomials

Let d be a nonnegative integer, and suppose that a polynomial p on $\mathbf{R}^n$ can be expressed as

$$p(x) = \sum_{\alpha_1+\cdots+\alpha_n = d} c_\alpha \, x^\alpha, \tag{1.3.3}$$

where more precisely the sum is taken over all multi-indices α with

$$\alpha_1 + \cdots + \alpha_n = d, \tag{1.3.4}$$

and the coefficients c_α are real numbers. This implies that p is homogeneous of degree d on $\mathbf{R}^n$, in the sense that

$$p(t\,x) = t^d \, p(x) \tag{1.3.5}$$

for every $t \in \mathbf{R}$ and $x \in \mathbf{R}^n$, as in Subsection 1.2.1. Conversely, if a polynomial p on $\mathbf{R}^n$ is homogeneous of degree d, then one can check that p can be expressed as in (1.3.3).

Of course, homogeneous polynomials on $\mathbf{R}^n$ of degree 0 are the same as constant functions on $\mathbf{R}^n$. A homogeneous polynomial of degree 1 on $\mathbf{R}^n$ is the same as a *linear functional* on $\mathbf{R}^n$, which is to say a linear mapping from $\mathbf{R}^n$ into $\mathbf{R}$.

Let $\mathcal{P}_d(\mathbf{R}^n)$ be the space of polynomials on $\mathbf{R}^n$ with real coefficients that are homogeneous of degree d. This is a linear subspace of $\mathcal{P}(\mathbf{R}^n)$. The collection of monomials x^α, where α is a multi-index that satisfies (1.3.4), is a basis for $\mathcal{P}_d(\mathbf{R}^n)$, as a vector space over the real numbers. Thus

$$\text{the dimension of } \mathcal{P}_d(\mathbf{R}^n) \text{ is equal to (1.2.7),} \tag{1.3.6}$$

as a vector space over the real numbers.

Let p_1, p_2 be polynomials on $\mathbf{R}^n$ with real coefficients of degrees $\le N_1, N_2$, respectively. It is easy to see that their product

$$p_1(x)\, p_2(x) \tag{1.3.7}$$

is a polynomial on $\mathbf{R}^n$ of degree $\le N_1 + N_2$. If p_1, p_2 are homogeneous of degrees d_1, d_2, respectively, then

$$p_1(x)\, p_2(x) \text{ is homogeneous of degree } d_1 + d_2. \tag{1.3.8}$$

1.4 Even and odd permutations

If t is a real number, then let us use $\operatorname{sign}(t)$ to denote the sign of t, which is 1 when $t > 0$, -1 when $t < 0$, and 0 when $t = 0$. Let n be a positive integer, and let σ be a mapping from the set $\{1, \ldots, n\}$ of positive integers from 1 to n into itself. Put

$$\operatorname{sgn}(\sigma) = \prod_{1 \le j < l \le n} \operatorname{sign}(\sigma(l) - \sigma(j)), \tag{1.4.1}$$

which may be called the *signum* of σ, as on p319 of [20]. More precisely, the product is taken over all pairs of integers j, l satisfying the conditions indicated. If σ is not one-to-one on $\{1, \ldots, n\}$, then

$$\operatorname{sgn}(\sigma) = 0. \tag{1.4.2}$$

Suppose that σ is one-to-one on $\{1, \ldots, n\}$, so that

$$\operatorname{sgn}(\sigma) = 1 \text{ or } -1. \tag{1.4.3}$$

In this case, σ maps $\{1, \ldots, n\}$ onto itself, basically because $\{1, \ldots, n\}$ is a finite set. We say that σ is a *permutation* on $\{1, \ldots, n\}$, and the inverse mapping σ^{-1} is a permutation on $\{1, \ldots, n\}$ as well. One can verify that

$$\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1}). \tag{1.4.4}$$

Note that $\operatorname{sgn}(\sigma) = 1$ when σ is the identity mapping on $\{1, \ldots, n\}$.

Sometimes $\operatorname{sgn}(\sigma)$ is only defined for permutations σ, and this definition for all mappings σ from $\{1, \ldots, n\}$ is used in Definition 9.33 on p232 of [155]. This will also be helpful for considering some properties of related mappings. If σ, τ are mappings from $\{1, \ldots, n\}$ into itself, then their composition $\tau \circ \sigma$

is the mapping from $\{1, \ldots, n\}$ into itself that sends j to $\tau(\sigma(j))$. This is a permutation on $\{1, \ldots, n\}$ when σ and τ are both permuations. Note that

(1.4.5) there are n^n mappings from $\{1, \ldots, n\}$ into itself,

and

(1.4.6) there are $n!$ permutations on $\{1, \ldots, n\}$.

A permutation τ on $\{1, \ldots, n\}$ is said to be a *transposition* if it interchanges two elements of $\{1, \ldots, n\}$, and maps every other element of $\{1, \ldots, n\}$ to itself. One can check that

$$\operatorname{sgn}(\tau) = -1 \tag{1.4.7}$$

when τ is a transposition. A permutation σ on $\{1, \ldots, n\}$ is said to be *even* when $\operatorname{sgn}(\sigma) = 1$, and *odd* when $\operatorname{sgn}(\sigma) = -1$. It is well known that every permutation on $\{1, \ldots, n\}$ can be expressed as the composition of finitely many transpositions.

If σ, τ are mappings from $\{1, \ldots, n\}$ into itself, then it is well known that

$$\operatorname{sgn}(\tau \circ \sigma) = \operatorname{sgn}(\tau)\, \operatorname{sgn}(\sigma). \tag{1.4.8}$$

This is easy to see when either σ or τ is not one-to-one on $\{1, \ldots, n\}$, which implies that $\tau \circ \sigma$ is not one-to-one on $\{1, \ldots, n\}$, so that both sides of (1.4.8) are equal to 0. A proof of (1.4.8) when σ and τ are both permutations will be given in the next section.

1.5 More on permutations

Let n be a positive integer, and let σ be a mapping from $\{1, \ldots, n\}$ into itself. If $x \in \mathbf{R}^n$, then let $T_\sigma(x)$ be the element of $\mathbf{R}^n$ whose jth coordinate is equal to

$$(T_\sigma(x))_j = x_{\sigma(j)} \tag{1.5.1}$$

for each $j = 1, \ldots, n$. This defines a linear mapping from $\mathbf{R}^n$ into itself. If σ is not one-to-one on $\{1, \ldots, n\}$, then T_σ is not one-to-one on $\mathbf{R}^n$. If σ is a permutation on $\{1, \ldots, n\}$, then T_σ is a one-to-one mapping from $\mathbf{R}^n$ onto itself, with

$$(T_\sigma)^{-1} = T_{\sigma^{-1}}. \tag{1.5.2}$$

Let τ be another mapping from $\{1, \ldots, n\}$ into itself. If $x \in \mathbf{R}^n$, then the jth coordinate of $T_\sigma(T_\tau(x))$ is equal to

$$(T_\sigma(T_\tau(x)))_j = (T_\tau(x))_{\sigma(j)} = x_{\tau(\sigma(j))} = (T_{\tau\circ\sigma}(x))_j \tag{1.5.3}$$

for each $j = 1, \ldots, n$. This means that

$$T_\sigma \circ T_\tau = T_{\tau\circ\sigma}. \tag{1.5.4}$$

1.5.1 An interesting polynomial

Consider the polynomial $P(x)$ defined on $\mathbf{R}^n$ by

$$P(x) = \prod_{1 \le j < l \le n} (x_l - x_j). \tag{1.5.5}$$

This is a homogeneous polynomial of degree $n\,(n-1)/2$, which is the number of factors in the product on the right side. Observe that

$$P(T_\sigma(x)) = \prod_{1 \le j < l \le n} ((T_\sigma(x))_l - (T_\sigma(x))_j) = \prod_{1 \le j < l \le n} (x_{\sigma(l)} - x_{\sigma(j)}) \tag{1.5.6}$$

for every $x \in \mathbf{R}^n$. One can use this to get that

$$P(T_\sigma(x)) = \operatorname{sgn}(\sigma)\, P(x) \tag{1.5.7}$$

for every $x \in \mathbf{R}^n$. In particular, (1.5.6) is equal to 0 when σ is not one-to-one on $\{1, \ldots, n\}$.

On p153 of [20], a polynomial is defined which is the same as (1.5.5), except for a possible factor of -1, depending on the dimension. This possible extra factor is not important for (1.5.7). This polynomial is used to define even and odd permuations, and $\operatorname{sgn}(\sigma)$ is defined for a permutation using that on p319 of [20]. Note that $\operatorname{sgn}(\sigma)$ is uniquely determined by (1.5.7).

One can get (1.4.8) from (1.5.4) and (1.5.7). This corresponds to the properties of compositions of even and odd permuations in (12) on p153 of [20]. This also corresponds to Theorem 17 on p66 of [125], with somewhat different notation.

1.6 Multilinear forms on $\mathbf{R}^n$

Let k and n be positive integers. In this section, we shall be concerned with real-valued functions on the Cartesian product

$$\mathbf{R}^n \times \cdots \times \mathbf{R}^n, \tag{1.6.1}$$

with k factors of $\mathbf{R}^n$. This may also be expressed as $(\mathbf{R}^n)^k$, the space of k-tuples of elements of $\mathbf{R}^n$. A real-valued function on this space may be expressed as $\mu(x(1), \ldots, x(k))$, where $x(1), \ldots, x(k)$ are elements of $\mathbf{R}^n$. The jth coordinate of $x(l)$ may be expressed as $x_j(l)$ for each $j = 1, \ldots, n$ and $l = 1, \ldots, k$.

We say that $\mu(x(1), \ldots, x(k))$ is *multilinear* on $(\mathbf{R}^n)^k$ if it is linear in $x(l)$ for each $l = 1, \ldots, k$, when $x(r)$ is fixed for $r \ne l$. One may also say that μ is *k-linear* in this case, to emphasize the role of k. One may call μ a *multilinear form* or *k-linear form* on $\mathbf{R}^n$ as well, to emphasize the role of n and perhaps k too, and the fact that μ takes values in $\mathbf{R}$. If $k = 1$, then this is the same as a linear mapping from $\mathbf{R}^n$ into $\mathbf{R}$, which is to say a linear functional on $\mathbf{R}^n$. If $k = 2$, then this is the same as a *bilinear form* on $\mathbf{R}^n$.

Let $\lambda_1, \ldots, \lambda_k$ be linear functionals on $\mathbf{R}^n$. It is easy to see that

$$(x(1), \ldots, x(k)) \mapsto \prod_{l=1}^{k} \lambda_l(x(l)) \tag{1.6.2}$$

defines a k-linear form on $\mathbf{R}^n$. This k-linear form may be denoted $\lambda_1 \otimes \cdots \otimes \lambda_k$, so that

$$(\lambda_1 \otimes \cdots \otimes \lambda_n)(x(1), \ldots, x(k)) = \prod_{l=1}^{k} \lambda_l(x(l)). \tag{1.6.3}$$

Let

$$\mathcal{M}_k(\mathbf{R}^n) \tag{1.6.4}$$

be the space of all k-linear forms on $\mathbf{R}^n$. If $\mu, \nu \in \mathcal{M}_k(\mathbf{R}^n)$ and $t \in \mathbf{R}$, then it is easy to see that

$$\mu + \nu \in \mathcal{M}_k(\mathbf{R}^n) \tag{1.6.5}$$

and

$$t\,\mu \in \mathcal{M}_k(\mathbf{R}^n). \tag{1.6.6}$$

Thus $\mathcal{M}_k(\mathbf{R}^n)$ is a linear subspace of the space of all real-valued functions on $(\mathbf{R}^n)^k$, as a vector space over the real numbers with respect to pointwise addition and scalar multiplication. In particular, $\mathcal{M}_k(\mathbf{R}^n)$ is a vector space over the real numbers with respect to pointwise addition and scalar multiplication.

1.6.1 Characterizing multilinear forms on $\mathbf{R}^n$

Let $e_1, \ldots, e_n$ be the *standard basis* for $\mathbf{R}^n$, so that the jth coordinate of e_m is equal to 1 when $j = m$, and to 0 otherwise. Thus

$$x = \sum_{m=1}^{n} x_m\, e_m \tag{1.6.7}$$

for every $x \in \mathbf{R}^n$. Let μ be a k-linear form on $\mathbf{R}^n$. If $x(1), \ldots, x(k) \in \mathbf{R}^n$, then

$$x(l) = \sum_{m=1}^{n} x_m(l)\, e_m \tag{1.6.8}$$

for each $l = 1, \ldots, k$, as in (1.6.7). Using the linearity of $\mu(x(1), \ldots, x(k))$ in $x(l)$ for each $l = 1, \ldots, k$, we get that

$$\mu(x(1), \ldots, x(k)) = \sum_{m_1=1}^{n} \cdots \sum_{m_k=1}^{n} x_{m_1}(1) \cdots x_{m_k}(k)\, \mu(e_{m_1}, \ldots, e_{m_k}). \tag{1.6.9}$$

This shows that μ is uniquely determined on $(\mathbf{R}^n)^k$ by the real numbers

$$\mu(e_{m_1}, \ldots, e_{m_k}), \quad 1 \le m_1, \ldots, m_k \le n. \tag{1.6.10}$$

One can also get a k-linear form μ on $\mathbf{R}^n$ as in (1.6.9) for any family of real numbers as in (1.6.10). It follows that

$$\text{the dimension of } \mathcal{M}_k(\mathbf{R}^n) \text{ is equal to } n^k, \tag{1.6.11}$$

as a vector space over the real numbers.

1.6.2 A basis for $\mathcal{M}_k(\mathbf{R}^n)$

Let θ_j be the linear functional on $\mathbf{R}^n$ defined for each $j = 1, \ldots, n$ by

$$\theta_j(x) = x_j \tag{1.6.12}$$

for every $x \in \mathbf{R}^n$. If $j_1, \ldots, j_k$ are positive integers less than or equal to n, then

$$\theta_{j_1} \otimes \cdots \otimes \theta_{j_k} \tag{1.6.13}$$

defines a k-linear form on $\mathbf{R}^n$, as in (1.6.3). More precisely,

$$(\theta_{j_1} \otimes \cdots \otimes \theta_{j_k})(x(1), \ldots, x(k)) = \prod_{l=1}^{k} x_{j_l}(l). \tag{1.6.14}$$

Thus

$$\begin{aligned}(\theta_{j_1} \otimes \cdots \otimes \theta_{j_k})(e_{m_1}, \ldots, e_{m_k}) &= 1 \quad \text{when } j_l = m_l \text{ for each } l = 1, \ldots, k \\ &= 0 \quad \text{otherwise.}\end{aligned} \tag{1.6.15}$$

The collection of k-linear forms on $\mathbf{R}^n$ as in (1.6.13) is a basis for $\mathcal{M}_k(\mathbf{R}^n)$, as a vector space over the real numbers.

1.7 Bilinear forms on $\mathbf{R}^n$

Let n be a positive integer, and let $b(x, y)$ be a bilinear form on $\mathbf{R}^n$. If $e_1, \ldots, e_n$ are the standard basis vectors in $\mathbf{R}^n$, then

$$b(x, y) = \sum_{m_1=1}^{n} \sum_{m_2=1}^{n} x_{m_1}\, y_{m_2}\, b(e_{m_1}, e_{m_2}) \tag{1.7.1}$$

for all $x, y \in \mathbf{R}^n$, as in Subsection 1.6.1. Thus b is uniquely determined by the family of real numbers

$$b(e_{m_1}, e_{m_2}), \quad 1 \le m_1, m_2 \le n, \tag{1.7.2}$$

and any family of real numbers of this type determines a bilinear form on $\mathbf{R}^n$ as in (1.7.1).

If

$$b(x, y) = b(y, x) \tag{1.7.3}$$

for all $x, y \in \mathbf{R}^n$, then b is said to be a *symmetric bilinear form* on $\mathbf{R}^n$. It is easy to see that this happens if and only if

$$b(e_{m_1}, e_{m_2}) = b(e_{m_2}, e_{m_1}) \tag{1.7.4}$$

for all $m_1, m_2 = 1, \ldots, n$.

If

$$b(x, y) = -b(y, x) \tag{1.7.5}$$

for all $x, y \in \mathbf{R}^n$, then b is said to be an *antisymmetric bilinear form* on $\mathbf{R}^n$. This happens if and only if

$$b(e_{m_1}, e_{m_2}) = -b(e_{m_2}, e_{m_1}) \tag{1.7.6}$$

for all $m_1, m_2 = 1, \ldots, n$. In particular, this implies that

$$b(e_m, e_m) = 0 \tag{1.7.7}$$

for every $m = 1, \ldots, n$.

If b is any bilinear form on $\mathbf{R}^n$, then

$$b(x + y, x + y) = b(x, x) + b(x, y) + b(y, x) + b(y, y). \tag{1.7.8}$$

If b is antisymmetric, then

$$b(w, w) = 0 \tag{1.7.9}$$

for every $w \in \mathbf{R}^n$. Conversely, if (1.7.9) holds for every $w \in \mathbf{R}^n$, then one can check that b is antisymmetric, using (1.7.8).

If b is any bilinear form on $\mathbf{R}^n$ again, then

$$(1/2)\,(b(x, y) + b(y, x)) \tag{1.7.10}$$

is a symmetric bilinear form on $\mathbf{R}^n$, and

$$(1/2)\,(b(x, y) - b(y, x)) \tag{1.7.11}$$

is an antisymmetric bilinear form on $\mathbf{R}^n$. Of course, $b(x, y)$ is equal to the sum of (1.7.10) and (1.7.11). If b is both symmetric and antisymmetric, then

$$b(x, y) = 0 \tag{1.7.12}$$

for every $x, y \in \mathbf{R}^n$.

1.7.1 Associated quadratic polynomials

If b is any bilinear form on $\mathbf{R}^n$, then

$$P_b(x) = b(x, x) = \sum_{m_1=1}^{n} \sum_{m_2=1}^{n} x_{m_1}\, x_{m_2}\, b(e_{m_1}, e_{m_2}) \tag{1.7.13}$$

defines a homogeneous polynomial on $\mathbf{R}^n$ of degree 2, with real coefficients. It is easy to see that every homogeneous polynomial on $\mathbf{R}^n$ of degree 2 with real coefficients corresponds to a bilinear form b on $\mathbf{R}^n$ in this way.

Note that the polynomial corresponding to b as in (1.7.13) is the same as the polynomial corresponding to the symmetric part (1.7.10). This implies that every homogeneous polynomial on $\mathbf{R}^n$ of degree 2 with real coefficients corresponds to a symmetric bilinear form on $\mathbf{R}^n$ in the same way as before, which could also be verified more directly.

If b is a symmetric bilinear form on $\mathbf{R}^n$, then one can check that b is uniquely determined by P_b, using (1.7.8).

1.8 Multilinear forms and homogeneous polynomials

Let k and n be positive integers, and let μ be a k-linear form on $\mathbf{R}^n$. Under these conditions,

$$P_\mu(x) = \mu(x, \ldots, x) \tag{1.8.1}$$

defines a homogeneous polynomial on $\mathbf{R}^n$ of degree k. To be more precise, if $e_1, \ldots, e_n$ are the standard basis vectors in $\mathbf{R}^n$, then

$$P_\mu(x) = \sum_{m_1=1}^{n} \cdots \sum_{m_k=1}^{n} x_{m_1} \cdots x_{m_k}\, \mu(e_{m_1}, \ldots, e_{m_k}), \tag{1.8.2}$$

as in Subsection 1.6.1. Note that

$$\mu \mapsto P_\mu \tag{1.8.3}$$

is a linear mapping from the space $\mathcal{M}_k(\mathbf{R}^n)$ of all k-linear forms on $\mathbf{R}^n$ into the space $\mathcal{P}_k(\mathbf{R}^n)$ of all homogeneous polynomials on $\mathbf{R}^n$ of degree k with real coefficients.

Let $\lambda_1, \ldots, \lambda_k$ be k linear functionals on $\mathbf{R}^n$, and suppose for the moment that

$$\mu = \lambda_1 \otimes \cdots \otimes \lambda_k, \tag{1.8.4}$$

where the right side is as in Section 1.6. In this case,

$$P_\mu(x) = \prod_{l=1}^{k} \lambda_l(x) \tag{1.8.5}$$

for every $x \in \mathbf{R}^n$.

Let θ_j be the linear functional on $\mathbf{R}^n$ defined by $\theta_j(x) = x_j$ for each $j = 1, \ldots, n$, as before. Suppose for the moment again that

$$\mu = \theta_{j_1} \otimes \cdots \otimes \theta_{j_k}, \tag{1.8.6}$$

where $j_1, \ldots, j_k$ are positive integers less than or equal to n. This implies that

$$P_\mu(x) = \prod_{l=1}^{k} x_{j_k} \tag{1.8.7}$$

for every $x \in \mathbf{R}^n$.

Using this, it is easy to see that (1.8.3) maps $\mathcal{M}_k(\mathbf{R}^n)$ onto $\mathcal{P}_k(\mathbf{R}^n)$. This was mentioned in Subsection 1.7.1 when $k = 2$. If $k = 1$, then (1.8.3) is one-to-one.

1.9 Symmetric multilinear forms

If X is any nonempty set, then we let $\mathrm{Sym}(X)$ be the space of all one-to-one mappings from X onto itself. This is a group with respect to composition of mappings, which is known as the *symmetric group* on X. If k is a positive integer and $X = \{1, \ldots, k\}$, then we may use the notation $\mathrm{Sym}(k)$ for $\mathrm{Sym}(X)$.

Let k and n be positive integers, and let μ be a k-linear form on $\mathbf{R}^n$. If $\sigma \in \mathrm{Sym}(k)$, then let μ^σ be the k-linear form on $\mathbf{R}^n$ defined by

$$\mu^\sigma(x(1), \ldots, x(k)) = \mu(x(\sigma(1)), \ldots, x(\sigma(k))). \tag{1.9.1}$$

Of course, this is the same as μ when σ is the identity mapping on $\{1, \ldots, k\}$. Note that

$$\mu \mapsto \mu^\sigma \tag{1.9.2}$$

is a onto-to-one linear mapping from the space $\mathcal{M}_k(\mathbf{R}^n)$ of k-linear forms on $\mathbf{R}^n$ onto itself. If σ is any mapping from $\{1, \ldots, k\}$ into itself, then (1.9.1) defines a real-valued function on $(\mathbf{R}^n)^k$, but it is not necessarily k-linear.

Let τ be another element of $\mathrm{Sym}(k)$, so that $\tau \circ \sigma \in \mathrm{Sym}(k)$ as well. Observe that

$$\begin{aligned} \mu^{\tau \circ \sigma}(x(1), \ldots, x(k)) &= \mu(x(\tau(\sigma(1))), \ldots, x(\tau(\sigma(k)))) \\ &= \mu^\sigma(x(\tau(1)), \ldots, x(\tau(k))) = (\mu^\sigma)^\tau(x(1), \ldots, x(k)). \end{aligned} \tag{1.9.3}$$

We say that μ is a *symmetric* k-linear form on $\mathbf{R}^n$ if

$$\mu^\sigma = \mu \tag{1.9.4}$$

for every $\sigma \in \mathrm{Sym}(k)$. This holds trivially when $k = 1$, and it is equivalent to the definition of a symmetric bilinear form on $\mathbf{R}^n$ in Section 1.7 when $k = 2$. In order to verify that μ is symmetric for any k, it suffices to check that (1.9.4) holds when σ is a transposition on $\{1, \ldots, k\}$, because every element of $\mathrm{Sym}(k)$ may be expressed as a composition of transpositions, as mentioned in Section 1.4.

Let $e_1, \ldots, e_n$ be the standard basis vectors in $\mathbf{R}^n$, as usual. If (1.9.4) holds for some $\sigma \in \mathrm{Sym}(k)$, then

$$\mu(e_{m_{\sigma(1)}}, \ldots, e_{m_{\sigma(k)}}) = \mu(e_{m_1}, \ldots, e_{m_k}) \tag{1.9.5}$$

for all positive integers $m_1, \ldots, m_k \le n$. Conversely, if (1.9.5) holds for all such $m_1, \ldots, m_k$, then (1.9.4) holds, as in Subsection 1.6.1. Thus μ is symmetric if and only if (1.9.5) holds for all $\sigma \in \mathrm{Sym}(k)$ and positive integers $m_1, \ldots, m_k \le n$. As before, it suffices to check that (1.9.5) holds for all transpositions σ on $\{1, \ldots, k\}$ and $1 \le m_1, \ldots, m_k \le n$, to get that μ is symmetric.

1.9.1 Symmetrizing multilinear forms

If μ is any k-linear form on $\mathbf{R}^n$, then put

$$S(\mu) = \frac{1}{k!} \sum_{\sigma \in \mathrm{Sym}(k)} \mu^\sigma, \tag{1.9.6}$$

which is another k-linear form on $\mathbf{R}^n$. Note that S is a linear mapping from $\mathcal{M}_k(\mathbf{R}^n)$ into itself, which may also be denoted S_k, to indicate the role of k. If μ is symmetric, then

$$S(\mu) = \frac{1}{k!} \sum_{\sigma \in \mathrm{Sym}(k)} \mu = \mu, \tag{1.9.7}$$

because $\mathrm{Sym}(k)$ has $k!$ elements, as mentioned in Section 1.4. Let us check that

$$S(\mu) \text{ is symmetric} \tag{1.9.8}$$

for every k-linear form μ on $\mathbf{R}^n$. One may describe $S(\mu)$ as the *symmetrization* of μ.

If $\tau \in \mathrm{Sym}(k)$, then

$$S(\mu)^\tau = \frac{1}{k!} \sum_{\sigma \in \mathrm{Sym}(k)} (\mu^\sigma)^\tau = \frac{1}{k!} \sum_{\sigma \in \mathrm{Sym}(k)} \mu^{\tau \circ \sigma}, \tag{1.9.9}$$

using (1.9.3) in the second step. Observe that

$$\sigma \mapsto \tau \circ \sigma \tag{1.9.10}$$

is a one-to-one mapping from $\mathrm{Sym}(k)$ onto itself. It follows that

$$S(\mu)^\tau = S(\mu), \tag{1.9.11}$$

so that (1.9.8) holds.

Similarly,

$$S(\mu^\tau) = \frac{1}{k!} \sum_{\sigma \in \mathrm{Sym}(k)} (\mu^\tau)^\sigma = \frac{1}{k!} \sum_{\sigma \in \mathrm{Sym}(k)} \mu^{\sigma \circ \tau}. \tag{1.9.12}$$

We also have that

$$\sigma \mapsto \sigma \circ \tau \tag{1.9.13}$$

is a one-to-one mapping from $\mathrm{Sym}(k)$ onto itself. This implies that

$$S(\mu^\tau) = S(\mu) \tag{1.9.14}$$

for every $\tau \in \mathrm{Sym}(k)$.

1.10 More on symmetric multilinear forms

Let k and n be positive integers again, and let μ be a k-linear form on $\mathbf{R}^n$. Also let $P_\mu(x) = \mu(x, \ldots, x)$ be the homogeneous polynomial on $\mathbf{R}^n$ of degree k associated to μ as in Section 1.8. If $\sigma \in \mathrm{Sym}(k)$, then

$$P_{\mu^\sigma} = P_\mu. \tag{1.10.1}$$

This implies that

$$P_{S(\mu)} = \frac{1}{k!} \sum_{\sigma \in \mathrm{Sym}(k)} P_{\mu^\sigma} = \frac{1}{k!} \sum_{\sigma \in \mathrm{Sym}(k)} P_\mu = P_\mu. \tag{1.10.2}$$

Remember that every homogeneous polynomial on $\mathbf{R}^n$ of degree k with real coefficients is of the form P_μ for some k-linear form μ on $\mathbf{R}^n$, as in Section 1.8. Using this and (1.10.2), we get that every such polynomial on $\mathbf{R}^n$ is of the form P_μ for some symmetric k-linear form μ on $\mathbf{R}^n$.

Let $\{1, \ldots, n\}^k$ be the set of all k-tuples of positive integers less than or equal to n. If $(m_1, \ldots, m_k)$ is an element of this set, then put

$$\alpha_j = \#\{1 \le l \le k : m_l = j\}, \tag{1.10.3}$$

for each $j = 1, \ldots, n$, where more precisely the right side is the number of positive integers $l \le k$ such that $m_l = j$. This defines a multi-index $\alpha = (\alpha_1, \ldots, \alpha_n)$ with

$$\sum_{j=1}^{n} \alpha_j = k. \tag{1.10.4}$$

Note that every multi-index α that satisfies (1.10.4) corresponds to an element of $\{1, \ldots, n\}^k$ in this way.

Suppose that a multi-index α corresponds to $(m_1, \ldots, m_k) \in \{1, \ldots, n\}^k$ as in (1.10.3). If $\sigma \in \mathrm{Sym}(k)$, then

$$(m_{\sigma(1)}, \ldots, m_{\sigma(k)}) \tag{1.10.5}$$

is another element of $\{1, \ldots, n\}^k$, which corresponds to α in the same way. Conversely, every element of $\{1, \ldots, n\}^k$ that corresponds to α in this way is of the form (1.10.5) for some $\sigma \in \mathrm{Sym}(k)$.

Remember that a k-linear form μ on $\mathbf{R}^n$ is symmetric if and only if (1.9.5) holds for all $\sigma \in \mathrm{Sym}(k)$ and positive integers $m_1, \ldots, m_k \le n$. This condition holds if and only if

(1.10.6) $\mu(e_{m_1}, \ldots, e_{m_k})$ only depends on the multi-index α corresponding to $(m_1, \ldots, m_k)$,

by the remarks in the preceding paragraph.

1.10.1 The space $\mathcal{SM}_k(\mathbf{R}^n)$

Let

$$\mathcal{SM}_k(\mathbf{R}^n) \tag{1.10.7}$$

be the space of symmetric k-linear forms on $\mathbf{R}^n$. It is easy to see that this is a linear subspace of the space $\mathcal{M}_k(\mathbf{R}^n)$ of all k-linear forms on $\mathbf{R}^n$. One can check that the dimension of $\mathcal{SM}_k(\mathbf{R}^n)$, as a vector space over the real numbers, is equal to the number of multi-indices α that satisfy (1.10.4), using the remarks in the previous paragraphs.

This is the same as the dimension of the space $\mathcal{P}_k(\mathbf{R}^n)$ of all homogeneous polynomials on $\mathbf{R}^n$ of degree k with real coefficients, as mentioned in Subsection 1.3.1. It follows that

(1.10.8) $\mu \mapsto P_\mu$ defines a one-to-one mapping from $\mathcal{SM}_k(\mathbf{R}^n)$ onto $\mathcal{P}_k(\mathbf{R}^n)$.

1.11 Alternating multilinear forms

Let k and n be positive integers, and let μ be a k-linear form on $\mathbf{R}^n$ again. We say that μ is an *alternating* k-linear form if

$$\mu^\sigma = \operatorname{sgn}(\sigma)\,\mu \tag{1.11.1}$$

for every $\sigma \in \operatorname{Sym}(k)$. This condition holds trivially when $k = 1$, and it reduces to the definition of an antisymmetric bilinear form on $\mathbf{R}^n$ in Section 1.7 when $k = 2$. In order to show that μ is alternating for any k, it suffices to check that

$$\mu^\sigma = -\mu \tag{1.11.2}$$

when σ is a transposition on $\{1, \ldots, k\}$, because every element of $\operatorname{Sym}(k)$ may be expressed as a product of transpositions, as mentioned in Section 1.4.

If μ is an alternating k-linear form, then

$$\mu(x(1), \ldots, x(k)) = 0 \tag{1.11.3}$$

whenever

$$x(l_1) = x(l_2) \tag{1.11.4}$$

for some $1 \le l_1 < l_2 \le k$. This follows from (1.11.2), by taking σ to be the transposition that interchanges l_1 and l_2.

Conversely, if (1.11.3) holds when (1.11.4) holds, then (1.11.2) holds, with σ equal to the transposition just mentioned. To see this, let $x(l) \in \mathbf{R}^n$ be given for $l \ne l_1, l_2$, so that $\mu(x(1), \ldots, x(k))$ may be considered as a bilinear form on $\mathbf{R}^n$, as a function of $x(l_1)$ and $x(l_2)$. In this case, the hypothesis that (1.11.3) holds when (1.11.4) holds implies that this bilinear form is antisymmetric, as in Section 1.7.

If (1.11.3) holds when (1.11.4) holds for any $1 \le l_1 < l_2 \le n$, then we get that (1.11.2) holds for every transposition σ. This implies that μ is alternating, as before.

Let $e_1, \ldots, e_n$ be the standard basis vectors in $\mathbf{R}^n$ again, and let $\sigma \in \mathrm{Sym}(k)$ be given. Of course, (1.11.1) implies that

$$\mu(e_{m_{\sigma(1)}}, \ldots, e_{m_{\sigma(k)}}) = \mathrm{sgn}(\sigma)\, \mu(e_{m_1}, \ldots, e_{m_k}) \tag{1.11.5}$$

for all positive integers $m_1, \ldots, m_k \leq n$. Conversely, if (1.11.5) holds for all $1 \leq m_1, \ldots, m_k \leq n$, then (1.11.1) holds, as in Subsection 1.6.1. This means that μ is alternating if and only if (1.11.5) holds for all $\sigma \in \mathrm{Sym}(k)$ and positive integers $m_1, \ldots, m_k \leq n$.

Similarly, (1.11.2) holds for some transposition σ if and only if

$$\mu(e_{m_{\sigma(1)}}, \ldots, e_{m_{\sigma(k)}}) = -\mu(e_{m_1}, \ldots, e_{m_k}) \tag{1.11.6}$$

for all positive integers $m_1, \ldots, m_k \leq n$. It follows that μ is alternating if and only if (1.11.6) holds for all transpositions σ and $1 \leq m_1, \ldots, m_k \leq n$. In particular, if μ is alternating, then

$$\mu(e_{m_1}, \ldots, e_{m_k}) = 0 \tag{1.11.7}$$

when

$$m_{l_1} = m_{l_2} \tag{1.11.8}$$

for some $1 \leq l_1 < l_2 \leq k$, as in (1.11.3).

If $k > n$, then any family of k positive integers $m_1, \ldots, m_k \leq n$ satisfies (1.11.8) for some $1 \leq l_1 < l_2 \leq k$. If μ is alternating, then we get that $\mu = 0$ in this case, as in Subsection 1.6.1.

1.11.1 The space $\mathcal{AM}_k(\mathbf{R}^n)$

Let

$$\mathcal{AM}_k(\mathbf{R}^n) \tag{1.11.9}$$

be the space of alternating k-linear forms on $\mathbf{R}^n$. This is a linear subspace of the space $\mathcal{M}_k(\mathbf{R}^n)$ of all k-linear forms on $\mathbf{R}^n$. If $k > n$, then

$$\mathcal{AM}_k(\mathbf{R}^n) = \{0\}, \tag{1.11.10}$$

as in the preceding paragraph.

1.12 More on alternating multilinear forms

Let k and n be positive integers, and let $e_1, \ldots, e_n$ be the standard basis for $\mathbf{R}^n$, as usual. If μ is any k-linear form on $\mathbf{R}^n$, then μ is uniquely determined by the family of real numbers

$$\mu(e_{m_1}, \ldots, e_{m_k}), \tag{1.12.1}$$

where $m_1, \ldots, m_k$ are positive integers less than or equal to n, as in Subsection 1.6.1. If μ is alternating, then (1.12.1) is equal to 0 unless the m_l's are distinct, as in (1.11.7). In particular, this always happens when $k > n$, and so we suppose now that $k \le n$.

Thus, if μ is alternating, then μ is uniquely determined by the family of real numbers (1.12.1), where $m_1, \ldots, m_k \le n$ are distinct. We also have that (1.12.1) determines

$$\mu(e_{m_{\sigma(1)}}, \ldots, e_{m_{\sigma(k)}}) \tag{1.12.2}$$

for every $\sigma \in \mathrm{Sym}(k)$, as in (1.11.5). This implies that μ is uniquely determined by the family of real numbers (1.12.1), with

$$1 \le m_1 < \cdots < m_k \le n. \tag{1.12.3}$$

The family of real numbers (1.12.1) with $m_1, \ldots, m_k$ as in (1.12.3) for an alternating k-linear form μ may be arbitrary too. More precisely, if (1.12.1) is given when $m_1, \ldots, m_k$ satisfy (1.12.3), then one can determine (1.12.1) when $m_1, \ldots, m_k$ are distinct positive integers less than or equal to n using (1.11.5). If we set (1.12.1) equal to 0 when the m_l's are not distinct, then (1.12.1) has been defined for all positive integers $m_1, \ldots, m_k \le n$. This determines a k-linear form μ on $\mathbf{R}^n$ as in Subsection 1.6.1. One would like to check that

$$\mu \text{ is alternating} \tag{1.12.4}$$

under these conditions.

To do this, it suffices to verify that (1.11.5) holds for all $\sigma \in \mathrm{Sym}(k)$ and positive integers $m_1, \ldots, m_k \le n$. If the m_l's are not distinct, then the $m_{\sigma(l)}$'s are not distinct either, and both sides of (1.11.5) are equal to 0, by construction. If the m_l's satisfy (1.12.3), then (1.11.5) holds by construction again.

If $m_1, \ldots, m_k$ satisfy (1.12.3) and $\tau \in \mathrm{Sym}(k)$, then $m_{\tau(1)}, \ldots, m_{\tau(k)}$ are distinct positive integers less than or equal to k, and any sequence of k distinct positive integers less than or equal to n can be expressed in this way. If σ is an element of $\mathrm{Sym}(k)$, then we would like to have that

$$\mu(e_{m_{\sigma(\tau(1))}}, \ldots, e_{m_{\sigma(\tau(k))}}) = \mathrm{sgn}(\sigma)\, \mu(e_{m_{\tau(1)}}, \ldots, e_{m_{\tau(k)}}). \tag{1.12.5}$$

By construction, the left side is equal to

$$\mathrm{sgn}(\sigma \circ \tau)\, \mu(e_{m_1}, \ldots, e_{m_k}), \tag{1.12.6}$$

and the right side is equal to

$$\mathrm{sgn}(\sigma)\ \mathrm{sgn}(\tau)\, \mu(e_{m_1}, \ldots, e_{m_k}). \tag{1.12.7}$$

Thus (1.12.5) follows from the fact that $\mathrm{sgn}(\sigma \circ \tau) = \mathrm{sgn}(\sigma)\ \mathrm{sgn}(\tau)$, as in Section 1.4 and Subsection 1.5.1.

1.12.1 The dimension of $\mathcal{AM}_k(\mathbf{R}^n)$

This shows that the dimension of $\mathcal{AM}_k(\mathbf{R}^n)$, as a vector space over the real numbers, is equal to the number of subsets of $\{1, \ldots, n\}$ with exactly k elements. This means that

$$\dim \mathcal{AM}_k(\mathbf{R}^n) = \binom{n}{k} \tag{1.12.8}$$

when $k \le n$.

In particular,

$$\dim \mathcal{AM}_n(\mathbf{R}^n) = 1. \tag{1.12.9}$$

It is well known that the determinant of an $n \times n$ matrix corresponds to an element of $\mathcal{AM}_n(\mathbf{R}^n)$, which will be discussed further in Section 1.15. Indeed, the determinant is uniquely determined by this and the fact that the determinant of the identity matrix is equal to 1. Any other element of $\mathcal{AM}_n(\mathbf{R}^n)$ can be expressed as a real number times the element of $\mathcal{AM}_n(\mathbf{R}^n)$ corresponding to the determinant, because of (1.12.9).

1.13 Alternatizing multilinear forms

Let k and n be positive integers, and let μ be a k-linear form on $\mathbf{R}^n$. Remember that if $\sigma \in \mathrm{Sym}(k)$, then μ^σ may be defined as a k-linear form on $\mathbf{R}^n$ as in Section 1.9. Put

$$A(\mu) = \frac{1}{k!} \sum_{\sigma \in \mathrm{Sym}(k)} \mathrm{sgn}(\sigma)\, \mu^\sigma. \tag{1.13.1}$$

This defines a linear mapping from the space $\mathcal{M}_k(\mathbf{R}^n)$ of k-linear forms on $\mathbf{R}^n$ into itself, which may also be denoted A_k, to indicate the role of k.

If μ is alternating, then

$$A(\mu) = \frac{1}{k!} \sum_{\sigma \in \mathrm{Sym}(k)} \mathrm{sgn}(\sigma)^2\, \mu. \tag{1.13.2}$$

This implies that

$$A(\mu) = \frac{1}{k!} \sum_{\sigma \in \mathrm{Sym}(k)} \mu = \mu, \tag{1.13.3}$$

because $\mathrm{sgn}(\sigma)^2 = 1$ for every $\sigma \in \mathrm{Sym}(k)$, and $\mathrm{Sym}(k)$ has $k!$ elements.

Let μ be any k-linear form on $\mathbf{R}^n$ again, and let $\tau \in \mathrm{Sym}(k)$ be given. Observe that

$$A(\mu)^\tau = \frac{1}{k!} \sum_{\sigma \in \mathrm{Sym}(k)} \mathrm{sgn}(\sigma)\, (\mu^\sigma)^\tau = \frac{1}{k!} \sum_{\sigma \in \mathrm{Sym}(k)} \mathrm{sgn}(\sigma)\, \mu^{\tau \circ \sigma}, \tag{1.13.4}$$

where the second step is as in Section 1.9. This means that

$$A(\mu)^\tau = \frac{\mathrm{sgn}(\tau)}{k!} \sum_{\sigma \in \mathrm{Sym}(k)} \mathrm{sgn}(\tau \circ \sigma)\, \mu^{\tau \circ \sigma}, \tag{1.13.5}$$

because $\mathrm{sgn}(\tau \circ \sigma) = \mathrm{sgn}(\tau)\, \mathrm{sgn}(\sigma)$ and $\mathrm{sgn}(\tau)^2 = 1$. It follows that

$$A(\mu)^\tau = \mathrm{sgn}(\tau)\, A(\mu), \tag{1.13.6}$$

because $\sigma \mapsto \tau \circ \sigma$ is a one-to-one mapping from $\mathrm{Sym}(k)$ onto itself. Thus

$$A(\mu) \text{ is alternating,} \tag{1.13.7}$$

and $A(\mu)$ may be described as the *alternatization* of μ.

Similarly,

$$A(\mu^\tau) = \frac{1}{k!} \sum_{\sigma \in \mathrm{Sym}(k)} \mathrm{sgn}(\sigma)\, (\mu^\tau)^\sigma = \frac{1}{k!} \sum_{\sigma \in \mathrm{Sym}(k)} \mathrm{sgn}(\sigma)\, \mu^{\sigma \circ \tau}. \tag{1.13.8}$$

This implies that

$$A(\mu^\tau) = \frac{\mathrm{sgn}(\tau)}{k!} \sum_{\sigma \in \mathrm{Sym}(k)} \mathrm{sgn}(\sigma \circ \tau)\, \mu^{\sigma \circ \tau}, \tag{1.13.9}$$

for essentially the same reasons as before. Using this, we get that

$$A(\mu^\tau) = \mathrm{sgn}(\tau)\, A(\mu), \tag{1.13.10}$$

because $\sigma \mapsto \sigma \circ \tau$ is a one-to-one mapping from $\mathrm{Sym}(k)$ onto itself.

1.14 The multinomial theorem

Let n be a positive integer, and let α be a multi-index. If $\sum_{j=1}^n \alpha_j = k$, then put

$$\binom{k}{\alpha} = \binom{k}{\alpha_1 \cdots \alpha_n} = \frac{k!}{\alpha_1! \cdots \alpha_n!}, \tag{1.14.1}$$

which is the *multinomial coefficient* corresponding to k and α. If $n = 2$, then this is the same as the binomial coefficients $\binom{k}{\alpha_1}$ and $\binom{k}{\alpha_2}$.

If k is a positive integer and $x \in \mathbf{R}^n$, then it is well known then

$$(x_1 + \cdots + x_n)^k = \sum_{\alpha_1 + \cdots + \alpha_n = k} \binom{k}{\alpha} x^\alpha, \tag{1.14.2}$$

where the sum is taken over all multi-indices α with $\sum_{j=1}^n \alpha_j = k$. This is trivial when $n = 1$, and the $n = 2$ case is the same as the binomial theorem. The $k = 1$ case is also easy, for any n.

As for the binomial theorem, it is clear that the left side of (1.14.2) can be expanded into a linear combination of monomials x^α of degree k whose coefficients are positive integers. One can get (1.14.2) directly using induction on k, but if one already knows the binomial theorem, then it is easier to use induction on n. One can also obtain (1.14.2) using calculus, by taking derivatives of both sides of order k, which shows that these are the only coefficients such that (1.14.2) holds for every $x \in \mathbf{R}^n$.

The coefficient of x^α in (1.14.2) is the same as the number of ways of partitioning $\{1, \ldots, k\}$ into n pairwise-disjoint sets, where the jth set has exactly α_j elements for each j. Thus the number of these partitions is the same as the multinomial coefficient (1.14.1). Alternatively, any permutation on $\{1, \ldots, k\}$ sends a partition of this type to another such partition, and every partiition of this type may be obtained from any other partition of this type in this way. There are $k!$ permutations on $\{1, \ldots, k\}$, and the number of permutations that send a particular partition of this type to itself is

$$\alpha_1! \cdots \alpha_n!. \tag{1.14.3}$$

This is another way to see that the number of these partitions is equal to (1.14.1).

A partition of $\{1, \ldots, k\}$ into n pairwise-disjoint sets corresponds exactly to a mapping from $\{1, \ldots, k\}$ into $\{1, \ldots, n\}$. A mapping of this type also corresponds to a k-tuple $(m_1, \ldots, m_k)$ of positive integers less than or equal to n. Of course,

$$(x_1 + \cdots + x_n)^k = \sum_{m_1=1}^{n} \cdots \sum_{m_k=1}^{n} x_{m_1} \cdots x_{m_k} \tag{1.14.4}$$

for every $x \in \mathbf{R}^n$. A monomial of the form

$$x_{m_1} \cdots x_{m_k} \tag{1.14.5}$$

is of the form x^α exactly when

$$\alpha_j = \#\{1 \le l \le k : m_l = j\} \tag{1.14.6}$$

for each $j = 1, \ldots, n$.

Thus (1.14.1) is the same as the number of k-tuples $(m_1, \ldots, m_k)$ of positive integers less than or equal to n that satisfy (1.14.6) for each j. This is related to some of the remarks in Section 1.10.

If we take $x_j = 1$ for each j in (1.14.2), then we get that

$$n^k = \sum_{\alpha_1 + \cdots + \alpha_n = k} \binom{k}{\alpha}. \tag{1.14.7}$$

Of course, the set $\{1, \ldots, n\}^k$ of k-tuples of positive integers less than or equal to n has exactly n^k elements. The right side of (1.14.7) expresses this as the sum over all multi-indices α with $\sum_{j=1}^n \alpha_j = k$ of the number of k-tuples $(m_1, \ldots, m_k)$ in $\{1, \ldots, n\}^k$ that satisfy (1.14.6) for each j.

1.15 Matrices and determinants

Let k and n be positive integers, and let $x(1), \ldots, x(k)$ be k elements of $\mathbf{R}^n$. Thus $x_j(l)$ defines an $n \times k$ matrix of real numbers.

Let $u_1, \ldots, u_k$ be the standard basis for $\mathbf{R}^k$, so that the lth coordinate of u_r is equal to 1 when $l = r$, and to 0 when $l \neq r$. There is a unique linear mapping from $\mathbf{R}^k$ into $\mathbf{R}^n$ such that

$$T(u_l) = x(l) \tag{1.15.1}$$

for each $l = 1, \ldots, k$. More precisely, if $w \in \mathbf{R}^k$, then

$$w = \sum_{l=1}^{k} w_l \, u_l, \tag{1.15.2}$$

and

$$T(w) = \sum_{l=1}^{k} w_l \, T(u_l) = \sum_{l=1}^{k} w_l \, x(l). \tag{1.15.3}$$

This means that the jth coordinate of $T(w)$ is given by

$$(T(w))_j = \sum_{l=1}^{k} x_j(l) \, w_l \tag{1.15.4}$$

for each $j = 1, \ldots, n$. Of course, this is the same as saying that T corresponds to the matrix $x_j(l)$ in the usual way.

Suppose from now on in this section that $k = n$, so that $x_j(l)$ is an $n \times n$ matrix of real numbers. Put

$$\mu_{\det}(x(1), \ldots, x(n)) = \sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sgn}(\sigma) \prod_{j=1}^{n} x_j(\sigma(j)). \tag{1.15.5}$$

The right side is the same as the *determinant* of the matrix $x_j(l)$. Equivalently,

$$\mu_{\det}(x(1), \ldots, x(n)) = \sum_{\tau \in \mathrm{Sym}(n)} \mathrm{sgn}(\tau) \prod_{l=1}^{n} x_{\tau(l)}(l). \tag{1.15.6}$$

This corresponds to taking $\tau = \sigma^{-1}$ in (1.15.5), using the fact that $\mathrm{sgn}(\sigma) = \mathrm{sgn}(\sigma^{-1})$.

Clearly $\mu_{\det}$ is an n-linear form on $\mathbf{R}^n$. More precisely, $\mu_{\det}$ is an alternating n-linear form on $\mathbf{R}^n$, and in fact it can be expressed as an alternatization in a nice way. Put $\theta_j(y) = y_j$ for each $j = 1, \ldots, n$ and $y \in \mathbf{R}^n$, so that

$$(\theta_1 \otimes \cdots \otimes \theta_n)(y(1), \ldots, y(n)) = \prod_{j=1}^{n} y_j(j) \tag{1.15.7}$$

defines an n-linear form on $\mathbf{R}^n$, as in Section 1.6. If $\sigma \in \mathrm{Sym}(n)$, then

$$\begin{aligned} (\theta_1 \otimes \cdots \otimes \theta_n)^{\sigma}(x(1), \ldots, x(n)) &= (\theta_1 \otimes \cdots \otimes \theta_n)(x(\sigma(1)), \ldots, x(\sigma(n))) \\ &= \prod_{j=1}^{n} x_j(\sigma(j)), \end{aligned} \tag{1.15.8}$$

using the notation in Section 1.9 in the first step. This means that

$$\mu_{det} = \sum_{\sigma \in \mathrm{Sym}(n)} \mathrm{sgn}(\sigma) \, (\theta_1 \otimes \cdots \otimes \theta_n)^{\sigma} = n! \, A(\theta_1 \otimes \cdots \otimes \theta_n), \tag{1.15.9}$$

using the notation in Section 1.13 in the second step.

It is well known that the determinant of the identity matrix is equal to 1. This corresponds to

$$\mu_{\det}(e_1, \dots, e_n) = 1 \tag{1.15.10}$$

in the present notation, where $e_1, \dots, e_n$ are the standard basis vectors in $\mathbf{R}^n$. Indeed,

$$(\theta_1 \otimes \cdots \otimes \theta_n)(e_1, \dots, e_n) = 1, \tag{1.15.11}$$

and

$$(\theta_1 \otimes \cdots \otimes \theta_n)^\sigma(e_1, \dots, e_n) = (\theta_1 \otimes \cdots \otimes \theta_n)(e_{\sigma(1)}, \dots, e_{\sigma(n)}) = 0 \tag{1.15.12}$$

when $\sigma \in \mathrm{Sym}(n)$ is not the identity mapping on $\{1, \dots, n\}$. Note that (1.15.6) corresponds to the fact that the determinant of a matrix is equal to the determinant of its transpose. Some other properties of determinants will be discussed in Section 2.5.

Chapter 2

More on multilinear forms

2.1 Some direct sums

Let N be a positive integer, and let $V_1, \ldots, V_N$ be N vector spaces over the real numbers. Consider the Cartesian product

$$\prod_{j=1}^{N} V_j = V_1 \times \cdots \times V_N, \tag{2.1.1}$$

which is the set of N-tuples $(v_1, \ldots, v_N)$ with $v_j \in V_j$ for each $j = 1, \ldots, N$. This is also a vector space over the real numbers, with respect to coordinatewise addition and scalar multiplication. This is the *direct sum* of $V_1, \ldots, V_N$, which may be denoted

$$\bigoplus_{j=1}^{N} V_j \text{ or } V_1 \bigoplus \cdots \bigoplus V_N. \tag{2.1.2}$$

If $v_j \in V_j$ for $j = 1, \ldots, N$, then we may use the notation $v_1 \oplus \cdots \oplus v_N$ for the element of the direct sum corresponding to $(v_1, \ldots, v_N)$.

There is an obvious mapping ι_l from V_l into the direct sum for each $l = 1, \ldots, N$. Namely, if $v_l \in V_l$, then let $\iota_l(v_l)$ be the element of the direct sum whose lth coordinate is equal to v_l, and whose other coordinates are equal to 0. This is a one-to-one linear mapping from V_l into the direct sum for each l. Sometimes one may wish to identify V_l with its image $\iota_l(V_l)$ in the direct sum under ι_l. Every element of the direct sum corresponds to a sum of elements of $\iota_l(V_l)$, $1 \le l \le N$, in a unique way, by construction.

Let V be another vector space over the real numbers, and suppose for the moment that V_l is a linear subspace of V for each $l = 1, \ldots, N$. Suppose also that every element of V can be expressed in a unique way as a sum of elements of $V_1, \ldots, V_N$. This leads to a one-to-one linear mapping from the direct sum of $V_1, \ldots, V_N$ onto V. We may say that V corresponds to the direct sum of $V_1, \ldots, V_N$, as a vector space over the real numbers, under these conditions.

If V_j has finite dimension for each $j = 1, \ldots, N$, then the direct sum has finite dimension as well, with

$$\dim \bigoplus_{j=1}^{N} V_j = \sum_{j=1}^{N} \dim V_j. \tag{2.1.3}$$

This may be seen by choosing a basis for each V_l, and considering the images of the basis vectors in the direct sum under ι_l. The combination of the images of these bases of the V_l's is a basis for the direct sum.

2.1.1 Direct sums and linear mappings

Let W be another vector space over the real numbers, and suppose that T_l is a linear mapping from V_l into W for each $l = 1, \ldots, N$. Observe that

$$T(v_1 \oplus \cdots \oplus v_N) = \sum_{l=1}^{N} T_l(v_l) \tag{2.1.4}$$

defines a linear mapping from $\bigoplus_{j=1}^{N} V_j$ into W. Conversely, any linear mapping T from $\bigoplus_{j=1}^{N} V_j$ into W determines a unique linear mapping T_l from V_l into W for each $l = 1, \ldots, N$ such that (2.1.4) holds. More precisely, one can take

$$T_l = T \circ \iota_l \tag{2.1.5}$$

for each $l = 1, \ldots, n$.

2.2 Multilinear mappings

Let k be a positive integer, let $V_1, \ldots, V_k$ be k vector spaces over the real numbers, and let Z be another vector space over the real numbers. Also let ϕ be a function defined on the Cartesian product

$$\prod_{l=1}^{k} V_l = V_1 \times \cdots \times V_k \tag{2.2.1}$$

with values in Z. We say that ϕ is *multilinear* on (2.2.1) if $\phi(v_1, \ldots, v_k)$ is linear in v_l for each $l = 1, \ldots, k$, when v_r is fixed for $r \neq l$. This is the same as in Section 1.6 when $V_l = \mathbf{R}^n$ for each $l = 1, \ldots, k$ and $Z = \mathbf{R}$. As before, we may say that ϕ is *k-linear* to emphasize the role of k.

Of course, if $k = 1$, then this is the same as a linear mapping from V_1 into Z. If $k = 2$, then this is the same as a *bilinear mapping* from $V_1 \times V_2$ into Z.

Remember that (2.2.1) may be considered as a vector space over the real numbers with respect to coordinatewise addition and scalar multiplication, as in the previous section. Note that k-linear mappings on (2.2.1) are not the same as linear mappings when $k \geq 2$.

Suppose for the moment that $Z = \mathbf{R}$, considered as a vector space over itself. A real-valued multilinear function on (2.2.1) may be called a *multilinear form* or *k-linear form*, as in Section 1.6 again. This is the same as a *linear functional* on V_1 when $k = 1$, and it may be called a *bilinear form* when $k = 2$, as before.

2.2.1 Some basic properties and examples

Let λ_l be a linear functional on V_l for each $l = 1, \ldots, k$, which is to say a linear mapping from V_l into $\mathbf{R}$. Observe that

$$(\lambda_1 \otimes \cdots \otimes \lambda_k)(v_1, \ldots, v_k) = \prod_{l=1}^{k} \lambda_l(v_l) \tag{2.2.2}$$

defines a k-linear form on (2.2.1). This was mentioned in Section 1.6 when $V_l = \mathbf{R}^n$ for each $l = 1, \ldots, k$.

Let Z be any vector space over the real numbers again, and let $z \in Z$ be given. Also let ψ be a k-linear form on (2.2.1). Under these conditions,

$$\phi(v_1, \ldots, v_k) = \psi(v_1, \ldots, v_k)\, z \tag{2.2.3}$$

defines a k-linear mapping from (2.2.1) into Z.

If X is a nonempty set, then the space of all Z-valued functions on X is a vector space over the real numbers, with respect to pointwise addition and scalar multiplication of functions. In particular, the space of all Z-valued functions on (2.2.1) is a vector space over the real numbers in this way. It is easy to see that the space of all k-linear mappings from (2.2.1) into Z is a linear subspace of the space of all Z-valued functions on (2.2.1).

Let V be a vector space over the real numbers, and suppose that we take $V_l = V$ for each $l = 1, \ldots, k$. In this case, (2.2.1) is the same as the Cartesian product

$$V \times \cdots \times V, \tag{2.2.4}$$

with k factors of V. This is the same as the space V^k of k-tuples of elements of V. The space of all k-linear forms on V may be denoted

$$\mathcal{M}_k(V), \tag{2.2.5}$$

as in Section 1.6. This is a linear subspace of the space of all real-valued functions on V^k, as in the preceding paragraph.

2.3 Some compositions

Let V and W be vector spaces over the real numbers, let k be a positive integer, and let $T_1, \ldots, T_k$ be k linear mappings from V into W. If μ is a k-linear form on W, then

$$\mu(T_1(v_1), \ldots, T_k(v_k)) \tag{2.3.1}$$

defines a k-linear form on V.

If T is a linear mapping from V into W, then we can apply the preceding remark with $T_l = T$ for each l. We may use the notation

$$(T^*(\mu))(v_1, \ldots, v_k) = \mu(T(v_1), \ldots, T(v_k)) \tag{2.3.2}$$

for the k-linear form on V obtained in this way, or the notation $(T)^*_k(\mu)$ to indicate the role of k. Note that this defines a linear mapping from $\mathcal{M}_k(W)$ into $\mathcal{M}_k(V)$. The notation δT may be used for this linear mapping as well, as in 2.13 on p62 of [184]. If $k = 1$, then (2.3.2) is the same as saying that

$$T^*(\mu) = (T)^*_1(\mu) = \mu \circ T. \tag{2.3.3}$$

Let Y be another vector space over the real numbers, and let R be a linear mapping from W into Y. Also let ν be a k-linear form on Y, so that $R^*(\nu)$ defines a k-linear form on W, as before. Similarly, $T^*(R^*(\nu))$ and $(R \circ T)^*(\nu)$ define k-linear forms on V, using the fact that $R \circ T$ is a linear mapping from V into Y in the second case. If $v_1, \ldots, v_k \in V$, then

$$\begin{aligned} (T^*(R^*(\nu)))(v_1, \ldots, v_k) &= (R^*(\nu))(T(v_1), \ldots, T(v_k)) \\ &= \nu(R(T(v_1)), \ldots, R(T(v_k))) \\ &= ((R \circ T)^*(\nu))(v_1, \ldots, v_k). \end{aligned} \tag{2.3.4}$$

This means that

$$(R \circ T)^* = T^* \circ R^* \tag{2.3.5}$$

as linear mappings from $\mathcal{M}_k(Y)$ into $\mathcal{M}_k(V)$.

If T is a one-to-one linear mapping from V onto W, then it is easy to see that

$$\begin{aligned} &T^* = (T)^*_k \text{ is a one-to-one linear mapping} \\ &\text{from } \mathcal{M}_k(W) \text{ onto } \mathcal{M}_k(V). \end{aligned} \tag{2.3.6}$$

More precisely,

$$((T)^*_k)^{-1} = (T^{-1})^*_k, \tag{2.3.7}$$

as in (2.3.5).

If W has dimension n for some positive integer n, then there is a one-to-one linear mapping T from $\mathbf{R}^n$ onto W. In fact, if $e_1, \ldots, e_n$ is the standard basis for $\mathbf{R}^n$, then one can choose T so that $T(e_1), \ldots, T(e_n)$ is any basis for W.

2.4 Compositions and permutations

Let V and W be vector spaces over the real numbers again, and let T be a linear mapping from V into W. Also let k be a positive integer, and let μ be a k-linear form on W. If σ is a permutation on $\{1, \ldots, k\}$, then μ^σ may be defined as a k-linear form on W as in Section 1.9, which is to say that

$$\mu^\sigma(w_1, \ldots, w_k) = \mu(w_{\sigma(1)}, \ldots, w_{\sigma(k)}) \tag{2.4.1}$$

for all $w_1, \ldots, w_k \in W$. If τ is another permutation on $\{1, \ldots, k\}$, then

$$\mu^{\tau \circ \sigma} = (\mu^\sigma)^\tau, \tag{2.4.2}$$

as before. Of course, we can do the same with k-linear forms on V.

In particular, we can do this for $T^*(\mu)$. Observe that

$$\begin{aligned} (T^*(\mu))^\sigma(v_1, \ldots, v_k) &= (T^*(\mu))(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) \\ &= \mu(T(v_{\sigma(1)}), \ldots, T(v_{\sigma(k)})) \\ &= \mu^\sigma(T(v_1), \ldots, T(v_k)) = (T^*(\mu^\sigma))(v_1, \ldots, v_k) \end{aligned} \tag{2.4.3}$$

for all $v_1, \ldots, v_k \in V$. Thus

$$(T^*(\mu))^\sigma = T^*(\mu^\sigma), \tag{2.4.4}$$

as k-linear forms on V.

We say that μ is a *symmetric* k-linear form on W if

$$\mu^\sigma = \mu \tag{2.4.5}$$

for every permutation σ on $\{1, \ldots, k\}$, as in Section 1.9. Similarly, we say that μ is an *alternating* k-linear form on W if

$$\mu^\sigma = \operatorname{sgn}(\sigma)\, \mu \tag{2.4.6}$$

for every $\sigma \in \operatorname{Sym}(k)$, as in Section 1.11. In both cases, it suffices to check that the condition holds when σ is a transposition on $\{1, \ldots, k\}$, as before. The analogous conditions may be used for k-linear forms on V as well. If μ is a symmetric or alternating k-linear form on W, then $T^*(\mu)$ has the same property as a k-linear form on V, because of (2.4.4).

Let

$$\mathcal{SM}_k(W) \text{ and } \mathcal{AM}_k(W) \tag{2.4.7}$$

be the spaces of symmetric and alternating k-linear forms on W, respectively. These are linear subspaces of $\mathcal{M}_k(W)$, as in the case where $W = \mathbf{R}^n$ for some positive integer n, as in Subsections 1.10.1 and 1.11.1. Using the same notation for V, we have that

$$T^*(\mathcal{SM}_k(W)) \subseteq \mathcal{SM}_k(V) \tag{2.4.8}$$

and

$$T^*(\mathcal{AM}_k(W)) \subseteq \mathcal{AM}_k(V), \tag{2.4.9}$$

as in the preceding paragraph. If T is a one-to-one linear mapping from V onto W, then we can combine these statements with their analogous for T^{-1}, to get that

$$T^*(\mathcal{SM}_k(W)) = \mathcal{SM}_k(V) \tag{2.4.10}$$

and

$$T^*(\mathcal{AM}_k(W)) = \mathcal{AM}_k(V). \tag{2.4.11}$$

2.4.1 Symmetrizations and alternatizations

If μ is any k-linear form on W again, then the *symmetrization* and *alternatization* of μ are defined by

$$S(\mu) = S_k(\mu) = \frac{1}{k!} \sum_{\sigma \in \mathrm{Sym}(k)} \mu^\sigma \tag{2.4.12}$$

and

$$A(\mu) = A_k(\mu) = \frac{1}{k!} \sum_{\sigma \in \mathrm{Sym}(k)} \mathrm{sgn}(\sigma)\, \mu^\sigma, \tag{2.4.13}$$

respectively, as in Subsection 1.9.1 and Section 1.13. As before, if μ is already symmetric or alternating, then

$$S(\mu) = \mu \tag{2.4.14}$$

or

$$A(\mu) = \mu, \tag{2.4.15}$$

as appropriate. If μ is any k-linear form on W, then

$$S(\mu) \text{ is symmetric} \tag{2.4.16}$$

and

$$A(\mu) \text{ is alternating,} \tag{2.4.17}$$

for the same reasons as before. If $\tau \in \mathrm{Sym}(k)$, then

$$S(\mu^\tau) = S(\mu) \tag{2.4.18}$$

and

$$A(\mu^\tau) = \mathrm{sgn}(\tau)\, A(\mu), \tag{2.4.19}$$

as before. It is easy to see that

$$T^*(S(\mu)) = S(T^*(\mu)) \tag{2.4.20}$$

and

$$T^*(A(\mu)) = A(T^*(\mu)), \tag{2.4.21}$$

using (2.4.4) and the analogues of S and A for V on the right sides of the equations.

2.5 Compositions and determinants

Let n be a positive integer, let T be a linear mapping from $\mathbf{R}^n$ into itself, and let $e_1, \ldots, e_n$ be the standard basis for $\mathbf{R}^n$. The *determinant* of T is normally defined as the determinant of the matrix associated to T using $e_1, \ldots, e_n$. This means that

$$\det T = \mu_{\det}(T(e_1), \ldots, T(e_n)), \tag{2.5.1}$$

where $\mu_{\det}$ is the alternating n-linear form on $\mathbf{R}^n$ associated to the determinant as in Section 1.15.

If ν is any alternating n-form on $\mathbf{R}^n$, then

$$\nu = \nu(e_1, \ldots, e_n)\, \mu_{\det}. \tag{2.5.2}$$

To see this, remember that the space $\mathcal{AM}_n(\mathbf{R}^n)$ of alternating n-linear forms on $\mathbf{R}^n$ has dimension 1 as a vector space over the real numbers, as in Subsection 1.12.1. This implies that ν is equal to a real number times $\mu_{\det}$. This real number is equal to $\nu(e_1, \ldots, e_n)$, because $\mu_{\det}(e_1, \ldots, e_n) = 1$.

Note that $T^*(\mu_{\det})$ is an alternating n-linear form on $\mathbf{R}^n$, as in the previous section. We also have that

$$T^*(\mu_{\det}) = (T^*(\mu_{\det}))(e_1, \ldots, e_n)\, \mu_{\det} = (\det T)\, \mu_{\det}, \tag{2.5.3}$$

using (2.5.2) in the first step, and (2.5.1) in the second step. This implies that

$$T^*(\nu) = (\det T)\, \nu \tag{2.5.4}$$

for every alternating n-linear form ν on $\mathbf{R}^n$, because ν is a constant multiple of $\mu_{\det}$.

Let R be another linear mapping from $\mathbf{R}^n$ into itself, so that

$$\det R = \mu_{\det}(R(e_1), \ldots, R(e_n)), \tag{2.5.5}$$

as in (2.5.1). Similarly,

$$\det(T \circ R) = \mu_{\det}(T(R(e_1)), \ldots, T(R(e_n))). \tag{2.5.6}$$

Observe that

$$(T^*(\mu_{\det}))(R(e_1), \ldots, R(e_n)) = (\det T)\, \mu_{\det}(R(e_1), \ldots, R(e_n)), \tag{2.5.7}$$

by (2.5.3). Of course, the right side of (2.5.6) is the same as the left side of (2.5.7). It follows that

$$\det(T \circ R) = (\det T)\, (\det R), \tag{2.5.8}$$

by (2.5.5).

2.5.1 n-Dimensional vector spaces

Let W be a vector space over the real numbers of dimension n, and let L be a one-to-one linear mapping from $\mathbf{R}^n$ onto W. If M is a linear mapping from W into itself, then one can define the *determinant* of M by

$$\det M = \det(L^{-1} \circ M \circ L), \tag{2.5.9}$$

where the right side is the determinant of $L^{-1} \circ M \circ L$, as a linear mapping from $\mathbf{R}^n$ into itself. One can check that this does not depend on the choice of L, using (2.5.8).

Using L, we get a one-to-one linear mapping $L^* = (L)^*_k$ from $\mathcal{M}_k(W)$ onto $\mathcal{M}_k(\mathbf{R}^n)$ for each positive integer k, as in Section 2.3. We also get that $(L)^*_k$ sends $\mathcal{SM}_k(W)$, $\mathcal{AM}_k(W)$ onto $\mathcal{SM}_k(\mathbf{R}^n)$, $\mathcal{AM}_k(\mathbf{R}^n)$, respectively, as in the previous section. This implies that $\mathcal{M}_k(W)$, $\mathcal{SM}_k(W)$, and $\mathcal{AM}_k(W)$ have the same dimensions as $\mathcal{M}_k(\mathbf{R}^n)$, $\mathcal{SM}_k(\mathbf{R}^n)$, and $\mathcal{AM}_k(\mathbf{R}^n)$, respectively.

In particular,

$$\dim \mathcal{AM}_n(W) = 1. \tag{2.5.10}$$

One can verify that (2.5.4) holds for every linear mapping T from W into itself and alternating n-linear form ν on W, by reducing to the analogous statement for $\mathbf{R}^n$ using L.

2.6 Subgroups of $\mathrm{Sym}(k)$

Let X be a nonempty set, and remember that $\mathrm{Sym}(X)$ is the space of one-to-one mappings from X onto itself, as in Section 1.9. A subset Σ of $\mathrm{Sym}(X)$ is said to be a *subgroup* if it satisfies the following three conditions. First, the identity mapping on X is an element of Σ. Second, if $\sigma \in \Sigma$, then

$$\sigma^{-1} \in \Sigma. \tag{2.6.1}$$

Third, if $\sigma, \tau \in \Sigma$, then

$$\sigma \circ \tau \in \Sigma. \tag{2.6.2}$$

It is well known that $\mathrm{Sym}(X)$ is an example of a *group*, with respect to composition of mappings on X. The definition of a subgroup of $\mathrm{Sym}(X)$ in the preceding paragraph corresponds exactly to the definition of a subgroup of an arbitrary group. Some basic examples of subgroups of $\mathrm{Sym}(X)$ can be obtained by considering the permutations that map cerain elements of X to themselves, or certain subsets of X to themselves. Note that $\mathrm{Sym}(X)$ is a subgroup of itself, and that the subset of $\mathrm{Sym}(X)$ consisting only of the identity mapping on X is a subgroup of $\mathrm{Sym}(X)$ too.

Let k be a positive integer, and let us now take $X = \{1, \ldots, k\}$. Thus we let Σ be a subgroup of $\mathrm{Sym}(k)$. Also let W be a vector space over the real numbers, and let μ be a k-linear form on W. Remember that μ^σ may be defined as a k-linear form on W for every $\sigma \in \mathrm{Sym}(k)$ as in Section 2.4. Let us say that μ is *symmetric with respect to* Σ if

$$\mu^\sigma = \mu \tag{2.6.3}$$

for every $\sigma \in \Sigma$. Similarly, we say that μ is *alternating with respect to* Σ if

$$\mu^\sigma = \mathrm{sgn}(\sigma)\, \mu \tag{2.6.4}$$

for every $\sigma \in \Sigma$. If $\Sigma = \mathrm{Sym}(k)$, then these definitions are the same as the previous definitions of symmetric and alternating k-forms on W in Section 2.4.

2.6.1 Associated symmetrizations and alternatizations

The *symmetrization* and *alternatization with respect to* Σ of any k-linear form μ on W may be defined by

$$S_\Sigma(\mu) = \frac{1}{\#\Sigma} \sum_{\sigma \in \Sigma} \mu^\sigma \tag{2.6.5}$$

and

$$A_\Sigma(\mu) = \frac{1}{\#\Sigma} \sum_{\sigma \in \Sigma} \operatorname{sgn}(\sigma)\, \mu^\sigma, \tag{2.6.6}$$

respectively, where $\#\Sigma$ is the number of elements of Σ. These are the same as the usual symmetrization and alternatization of μ, as in Subsection 2.4.1, when $\Sigma = \operatorname{Sym}(k)$. It is easy to see that

$$S_\Sigma(\mu) = \mu \tag{2.6.7}$$

when μ is symmetric with respect to Σ, and that

$$A_\Sigma(\mu) = \mu \tag{2.6.8}$$

when μ is alternating with respect to Σ. One can check that

$$S_\Sigma(\mu) \text{ is symmetric with respect to } \Sigma \tag{2.6.9}$$

and

$$A_\Sigma(\mu) \text{ is alternating with respect to } \Sigma \tag{2.6.10}$$

for any k-linear form μ on W, using the same type of arguments as in Subsection 1.9.1 and Section 1.13. If $\tau \in \Sigma$, then

$$S_\Sigma(\mu^\tau) = S_\Sigma(\mu) \tag{2.6.11}$$

and

$$A_\Sigma(\mu^\tau) = \operatorname{sgn}(\tau)\, A_\Sigma(\mu), \tag{2.6.12}$$

for essentially the same reasons as before.

Observe that

$$\begin{aligned} S(S_\Sigma(\mu)) = S\Big(\frac{1}{\#\Sigma} \sum_{\tau \in \Sigma} \mu^\tau\Big) &= \frac{1}{\#\Sigma} \sum_{\tau \in \Sigma} S(\mu^\tau) \\ &= \frac{1}{\#\Sigma} \sum_{\tau \in \Sigma} S(\mu) = S(\mu), \end{aligned} \tag{2.6.13}$$

because $S(\mu^\tau) = S(\mu)$ for every $\tau \in \operatorname{Sym}(k)$, as in Subsection 2.4.1. Similarly,

$$\begin{aligned} A(A_\Sigma(\mu)) = A\Big(\frac{1}{\#\Sigma} \sum_{\tau \in \Sigma} \operatorname{sgn}(\tau)\, \mu^\tau\Big) &= \frac{1}{\#\Sigma} \sum_{\tau \in \Sigma} \operatorname{sgn}(\tau)\, A(\mu^\tau) \\ &= \frac{1}{\#\Sigma} \sum_{\tau \in \Sigma} \operatorname{sgn}(\tau)^2\, A(\mu) = A(\mu), \end{aligned} \tag{2.6.14}$$

because $A(\mu^\tau) = \operatorname{sgn}(\tau)\, A(\mu)$ for every $\tau \in \operatorname{Sym}(k)$, as before.

2.7 Permutations on subsets

Let X be a nonempty set, and let E be a nonempty subset of X. Consider the spaces $\mathrm{Sym}(X)$ and $\mathrm{Sym}(E)$ of one-to-one mappings from X and E onto themselves, respectively, as before. If $\sigma \in \mathrm{Sym}(E)$, then let $\widehat{\sigma}$ be the mapping from X into itself which is equal to σ on E, and equal to the identity mapping on $X \setminus E$. It is easy to see that $\widehat{\sigma}$ is an element of $\mathrm{Sym}(X)$, which is the identity mapping on X when σ is the identity mapping on E. We also have that

$$\widehat{(\sigma \circ \tau)} = \widehat{\sigma} \circ \widehat{\tau} \tag{2.7.1}$$

for every $\sigma, \tau \in \mathrm{Sym}(X)$.

This means that

$$\sigma \mapsto \widehat{\sigma} \tag{2.7.2}$$

is a *homomorphism* from $\mathrm{Sym}(E)$ into $\mathrm{Sym}(X)$, as groups with respect to compositions of mappings. Note that (2.7.2) is a one-to-one mapping from $\mathrm{Sym}(E)$ into $\mathrm{Sym}(X)$. Put

$$\Sigma_E = \{\widehat{\sigma} : \sigma \in \mathrm{Sym}(E)\}, \tag{2.7.3}$$

which is a subgroup of $\mathrm{Sym}(X)$, as in the previous section. This uses the fact that

$$\widehat{(\sigma^{-1})} = (\widehat{\sigma})^{-1} \tag{2.7.4}$$

for every $\sigma \in \Sigma$.

Now let k be a positive integer, and let us take $X = \{1, \ldots, k\}$ again. If $\sigma \in \mathrm{Sym}(E)$, then one can check that

$$\mathrm{sgn}(\sigma) = \mathrm{sgn}(\widehat{\sigma}). \tag{2.7.5}$$

More precisely, one can define $\mathrm{sgn}(\sigma)$ by identifying σ with an element of $\mathrm{Sym}(\#E)$, by listing the elements of E in order. In fact, one could list the elements of E in any order and get the same result for $\mathrm{sgn}(\sigma)$, but we do not really need that now.

It is easier to verify (2.7.5) when the elements of E are consecutive in $\{1, \ldots, k\}$, which holds in many examples of interest. Otherwise, there can be an even number of additional fectors of -1 in the definition of $\mathrm{sgn}(\widehat{\sigma})$.

Alternatively, if σ is a transposition on E, then $\widehat{\sigma}$ is a transposition on $\{1, \ldots, k\}$. Thus (2.7.5) is clear in this case, and otherwise one can express σ as the composition of finitely many transposition.

2.8 Products of multilinear forms

Let W be a vector space over the real numbers, and let r be a positive integer. Also let $k_1, \ldots, k_r$ be r positive integers, and put

$$k = \sum_{m=1}^{r} k_m. \tag{2.8.1}$$

If $1 \le m \le r$, then we put

$$L(m) = \sum_{i=1}^{m-1} k_i, \tag{2.8.2}$$

which is interpreted as being equal to 0 when $m = 1$. Note that every positive integer less than or equal to k can be expressed in a unique way as

$$L(m) + l, \tag{2.8.3}$$

where $1 \le m \le r$ and $1 \le l \le k_m$.

Suppose that μ_m is a k_m-linear form on W for each $m = 1, \ldots, r$. If $w_1, \ldots, w_k$ are k elements of W, then put

$$(\mu_1 \otimes \cdots \otimes \mu_r)(w_1, \ldots, w_k) = \prod_{m=1}^{r} \mu_m(w_{L(m)+1}, \ldots, w_{L(m)+k_m}). \tag{2.8.4}$$

This defines a k-linear form

$$\mu = \mu_1 \otimes \cdots \otimes \mu_r \tag{2.8.5}$$

on W. This is related to some remarks in Section 1.6 and Subsection 2.2.1 when $k_m = 1$ for each $m = 1, \ldots, r$. Note that

$$(\mu_1, \ldots, \mu_r) \mapsto \mu_1 \otimes \cdots \otimes \mu_r \tag{2.8.6}$$

defines an r-linear mapping from

$$\prod_{m=1}^{r} \mathcal{M}_{k_m}(W) = \mathcal{M}_{k_1}(W) \times \cdots \times \mathcal{M}_{k_r}(W) \tag{2.8.7}$$

into $\mathcal{M}_k(W)$.

If $r = 2$, then (2.8.4) is the same as saying that

$$(\mu_1 \otimes \mu_2)(w_1, \ldots, w_{k_1+k_2}) = \mu_1(w_1, \ldots, w_{k_1})\, \mu_2(w_{k_1+1}, \ldots, w_{k_1+k_2}). \tag{2.8.8}$$

Similarly, if $r = 3$, then (2.8.4) is the same as saying that

$$\begin{aligned} &(\mu_1 \otimes \mu_2 \otimes \mu_3)(w_1, \ldots, w_{k_1+k_2+k_3}) \\ &= \mu_1(w_1, \ldots, w_{k_1})\, \mu_2(w_{k_1+1}, \ldots, w_{k_1+k_2})\, \mu_3(w_{k_1+k_2+1}, \ldots, w_{k_1+k_2+k_3}). \end{aligned} \tag{2.8.9}$$

2.8.1 Polynomials, associativity, and linear mappings

Suppose for the moment that $W = \mathbf{R}^n$ for some positive integer n. Remember that a multilinear form ν on $\mathbf{R}^n$ corresponds to a homogeneous polynomial P_ν on $\mathbf{R}^n$ as in Section 1.8. If μ is as in (2.8.5), then it is easy to see that

$$P_\mu = \prod_{m=1}^{r} P_{\mu_m}. \tag{2.8.10}$$

Suppose for the moment again that $r = 3$. One can define $\mu_1 \otimes \mu_2$ as a $(k_1 + k_2)$-linear form on W as before, and use that to define

$$(\mu_1 \otimes \mu_2) \otimes \mu_3 \tag{2.8.11}$$

as a $(k_1 + k_2 + k_3)$-linear form on W. Similarly, one can define $\mu_2 \otimes \mu_3$ as a $(k_2 + k_3)$-linear form on W, and use that to define

$$\mu_1 \otimes (\mu_2 \otimes \mu_3) \tag{2.8.12}$$

as a $(k_1 + k_2 + k_3)$-linear form on W. It is easy to see that (2.8.11) and (2.8.12) are the same as

$$\mu_1 \otimes \mu_2 \otimes \mu_3. \tag{2.8.13}$$

Let V be another vector space over the real numbers, and let T be a linear mapping from V into W. It is easy to see that

$$T^*(\mu_1 \otimes \cdots \otimes \mu_r) = T^*(\mu_1) \otimes \cdots \otimes T^*(\mu_r), \tag{2.8.14}$$

as k-linear forms on V. Of course, this uses the mappings induced on multilinear forms by T as in Section 2.3.

2.8.2 Products and permutations

Let ϕ be a permutation on $\{1, \ldots, r\}$, so that $k_{\phi(1)}, \ldots, k_{\phi(r)}$ are r positive integers with

$$\sum_{m=1}^{r} k_{\phi(m)} = k. \tag{2.8.15}$$

Thus

$$\nu = \mu_{\phi(1)} \otimes \cdots \otimes \mu_{\phi(r)} \tag{2.8.16}$$

is a k-linear form on W, as before. If $W = \mathbf{R}^n$, then

$$P_\nu = \prod_{m=1}^{r} P_{\mu_{\phi(m)}} = P_\mu. \tag{2.8.17}$$

Put

$$L_\phi(m) = \sum_{i=1}^{m-1} k_{\phi(i)} \tag{2.8.18}$$

for each $m = 1, \ldots, r$, as before. Every positive integer less than or equal to k can be expressed in a unique way as

$$L_\phi(m) + l \tag{2.8.19}$$

for some $m = 1, \ldots, r$ and $l = 1, \ldots, k_{\phi(m)}$, as before.

Consider the mapping σ_ϕ from $\{1, \ldots, k\}$ into itself defined by

$$\sigma_\phi(L_\phi(m) + l) = L(\phi(m)) + l \tag{2.8.20}$$

for each $m = 1, \ldots, r$ and $l = 1, \ldots, k_{\phi(m)}$. More precisely, this defines a permutation on $\{1, \ldots, k\}$. One can check that

$$\nu = \mu^{\sigma_\phi}, \tag{2.8.21}$$

where the right side is defined as in Section 2.4.

If $r = 2$, $\phi(1) = 2$, and $\phi(2) = 1$, then

$$\nu = \mu_2 \otimes \mu_1. \tag{2.8.22}$$

Note that $L(1) = 0$, $L(2) = k_1$ and $L_\phi(1) = 0$, $L_\phi(2) = k_2$ in this case. This means that

$$\sigma_\phi(l) = k_1 + l \text{ for } l = 1, \ldots, k_2 = k_{\phi(1)} \tag{2.8.23}$$

and

$$\sigma_\phi(k_2 + l) = l \text{ for } l = 1, \ldots, k_1 = k_{\phi(2)}. \tag{2.8.24}$$

2.9 Symmetry and products

Let us continue with the same notation and hypotheses as in the previous section. If $\mu_1, \ldots, \mu_r$ are symmetric multilinear forms on W, then we may prefer to consider a product that is a symmetric multilinear form as well. In this case, we may put

$$\mu_1 \odot \cdots \odot \mu_r = S(\mu), \tag{2.9.1}$$

where the right side is as in Subsection 2.4.1. This defines an r-linear mapping

$$(\mu_1, \ldots, \mu_r) \mapsto \mu_1 \odot \cdots \odot \mu_r \tag{2.9.2}$$

from $\prod_{m=1}^r \mathcal{SM}_{k_m}(W) = \mathcal{SM}_{k_1}(W) \times \cdots \times \mathcal{SM}_{k_r}(W)$ into $\mathcal{SM}_k(W)$.

If $W = \mathbf{R}^n$, then

$$P_{S(\mu)} = P_\mu = \prod_{m=1}^r P_{\mu_m}, \tag{2.9.3}$$

where the first step is as in Section 1.10. This determines $S(\mu)$ uniquely, as in Subsection 1.10.1.

Let ϕ be a permutation on $\{1, \ldots, r\}$ again, and let ν be as in (2.8.16). Observe that

$$S(\nu) = S(\mu^{\sigma_\phi}) = S(\mu), \tag{2.9.4}$$

using (2.8.21) in the first step, and where the second step is as in Subsection 2.4.1. This means that

$$\mu_{\phi(1)} \odot \cdots \odot \mu_{\phi(r)} = \mu_1 \odot \cdots \odot \mu_r. \tag{2.9.5}$$

2.9.1 Some symmetrizations of products

Suppose for the moment that $r = 2$. We would like to check that

$$S_{k_1+k_2}(S_{k_1}(\mu_1) \otimes \mu_2) = S_{k_1+k_2}(\mu_1 \otimes \mu_2), \tag{2.9.6}$$

where S_{k_1}, $S_{k_1+k_2}$ are the symmetrization operators associated to k_1 and k_1+k_2 as in Subsection 2.4.1. If $\tau \in \mathrm{Sym}(k_1)$, then let $\widehat{\tau}$ be the element of $\mathrm{Sym}(k_1+k_2)$ that is equal to τ on $\{1, \ldots, k_1\}$, and is the identity mapping on $\{k_1+1, \ldots, k_1+k_2\}$, as in Section 2.7. It is easy to see that

$$\mu_1^\tau \otimes \mu_2 = (\mu_1 \otimes \mu_2)^{\widehat{\tau}}, \tag{2.9.7}$$

using the notation from Section 2.4.

Remember that

$$\Sigma = \{\widehat{\tau} : \tau \in \mathrm{Sym}(k_1)\} \tag{2.9.8}$$

is a subgroup of $\mathrm{Sym}(k_1+k_2)$, as in Sections 2.6 and 2.7. Of course, the number of elements of Σ is the same as the number of elements of $\mathrm{Sym}(k_1)$, which is $k_1!$. One can check that

$$S_{k_1}(\mu_1) \otimes \mu_2 = S_\Sigma(\mu_1 \otimes \mu_2), \tag{2.9.9}$$

where S_Σ is as in Subsection 2.6.1, using (2.9.7). This implies that

$$S_{k_1+k_2}(S_{k_1}(\mu_1) \otimes \mu_2) = S_{k_1+k_2}(S_\Sigma(\mu_1 \otimes \mu_2)) = S_{k_1+k_2}(\mu_1 \otimes \mu_2), \tag{2.9.10}$$

where the second step is as in Subsection 2.6.1.

Alternatively, if $W = \mathbf{R}^n$, then the polynomial associated to the right side of (2.9.6) is as in (2.9.3). One can verify that this is the same as the polynomial associated to the left side of (2.9.6).

Similarly,

$$S_{k_1+k_2}(\mu_1 \otimes S_{k_2}(\mu_2)) = S_{k_1+k_2}(\mu_1 \otimes \mu_2). \tag{2.9.11}$$

It follows that

$$\begin{aligned} S_{k_1+k_2}(S_{k_1}(\mu_1) \otimes S_{k_2}(\mu_2)) &= S_{k_1+k_2}(\mu_1 \otimes S_{k_2}(\mu_2)) \\ &= S_{k_1+k_2}(\mu_1 \otimes \mu_2). \end{aligned} \tag{2.9.12}$$

One could also look at this in terms of the subgroup of $\mathrm{Sym}(k_1 + k_2)$ corresponding to the product of $\mathrm{Sym}(k_1)$ and $\mathrm{Sym}(k_2)$, as in Section 2.13.

Suppose now that $r = 3$ again. Observe that

$$S_k(S_{k_1+k_2}(\mu_1 \otimes \mu_2) \otimes \mu_3) = S_k(\mu_1 \otimes \mu_2 \otimes \mu_3), \tag{2.9.13}$$

where $k = k_1 + k_2 + k_3$, as in (2.9.6). Similarly,

$$S_k(\mu_1 \otimes S_{k_2+k_3}(\mu_2 \otimes \mu_3)) = S_k(\mu_1 \otimes \mu_2 \otimes \mu_3), \tag{2.9.14}$$

as in (2.9.11). It follows that

$$((\mu_1 \odot \mu_2) \odot \mu_3) = \mu_1 \odot \mu_2 \odot \mu_3 = \mu_1 \odot (\mu_2 \odot \mu_3) \tag{2.9.15}$$

when μ_1, μ_2, μ_3 are symmetric, in the notation of (2.9.1).

2.10 Alternatizing products

Let W be a vector space over the real numbers, let k_1, k_2 be positive integers, and let μ_1, μ_2 be k_1, k_2-linear forms on W, respectively. We would like to check that

$$A_{k_1+k_2}(A_{k_1}(\mu_1) \otimes \mu_2) = A_{k_1+k_2}(\mu_1 \otimes \mu_2), \tag{2.10.1}$$

where A_{k_1} and $A_{k_1+k_2}$ are the alternatization operators associated to k_1 and $k_1 + k_2$ as in Subsection 2.4.1.

If $\tau \in \mathrm{Sym}(k_1)$, then let $\widehat{\tau} \in \mathrm{Sym}(k_1 + k_2)$ be as in Subsection 2.9.1, and let Σ be as in (2.9.8) again. Remember that $\mathrm{sgn}(\tau) = \mathrm{sgn}(\widehat{\tau})$ for every $\tau \in \mathrm{Sym}(k_1)$, as in Section 2.7. One can use this and (2.9.7) to get that

$$A_{k_1}(\mu_1) \otimes \mu_2 = A_\Sigma(\mu_1 \otimes \mu_2), \tag{2.10.2}$$

where A_Σ is as in Subsection 2.6.1. This implies that

$$A_{k_1+k_2}(A_{k_1}(\mu_1) \otimes \mu_2) = A_{k_1+k_2}(A_\Sigma(\mu_1 \otimes \mu_2)) = A_{k_1+k_2}(\mu_1 \otimes \mu_2), \tag{2.10.3}$$

where the second step is as in Subsection 2.6.1.

Similarly,

$$A_{k_1+k_2}(\mu_1 \otimes A_{k_2}(\mu_2)) = A_{k_1+k_2}(\mu_1 \otimes \mu_2). \tag{2.10.4}$$

Thus

$$\begin{aligned} A_{k_1+k_1}(A_{k_1}(\mu_1) \otimes A_{k_2}(\mu_2)) &= A_{k_1+k_2}(\mu_1 \otimes A_{k_2}(\mu_2)) \\ &= A_{k_1+k_2}(\mu_1 \otimes \mu_2). \end{aligned} \tag{2.10.5}$$

As before, one can look at this more directly in terms of the appropriate subgroup of $\mathrm{Sym}(k_1 + k_2)$, as in Section 2.13.

Now let k_1, k_2, and k_3 be positive integers, and put $k = k_1 + k_2 + k_3$ again. If μ_1, μ_2, and μ_3 are k_1, k_2, and k_3-linear forms on W, respectively, then

$$A_k(A_{k_1+k_2}(\mu_1 \otimes \mu_2) \otimes \mu_3) = A_k(\mu_1 \otimes \mu_2 \otimes \mu_3), \tag{2.10.6}$$

as in (2.10.1). Similarly,

$$A_k(\mu_1 \otimes A_{k_1+k_2}(\mu_2 \otimes \mu_3)) = A_k(\mu_1 \otimes \mu_2 \otimes \mu_3), \tag{2.10.7}$$

as in (2.10.4).

2.10.1 The $\wedge_0$ product

If μ_1, μ_2 are alternating k_1, k_2-linear forms on W, respectively, then put

$$\mu_1 \wedge_0 \mu_2 = A_{k_1+k_2}(\mu_1 \otimes \mu_2) \tag{2.10.8}$$

which is an alternating $(k_1 + k_2)$-linear form on W. This corresponds to

$$\mu_1 \wedge_\beta \mu_2 \tag{2.10.9}$$

in the notation used in (3) on p60 of [184], and we shall say more about this in the next section. Note that

$$(\mu_1, \mu_2) \mapsto \mu_1 \wedge_0 \mu_2 \tag{2.10.10}$$

defines a bilinear mapping from $\mathcal{AM}_{k_1}(W) \times \mathcal{AM}_{k_2}(W)$ into $\mathcal{AM}_{k_1+k_2}(W)$. If μ_3 is an alternating k_3-linear form on W, then we have that

$$(\mu_1 \wedge_0 \mu_2) \wedge_0 \mu_3 = \mu_1 \wedge_0 (\mu_2 \wedge_0 \mu_3) = A_{k_1+k_2+k_3}(\mu_1 \otimes \mu_2 \otimes \mu_3), \tag{2.10.11}$$

by (2.10.6) and (2.10.7).

Now let r be a positive integer, let $k_1, \dots, k_r$ be r positive integers, and put $k = \sum_{m=1}^r k_m$, as in Section 2.8. If μ_m is an alternating k_m-linear form on W for each $m = 1, \dots, r$, then

$$\mu_1 \wedge_0 \cdots \wedge_0 \mu_r = A_k(\mu_1 \otimes \cdots \otimes \mu_r) \tag{2.10.12}$$

is an alternating k-linear form on W, where the right side is as in Subsection 2.4.1. More precisely, the left side may be defined by combining alternating multilinear forms two at a time using $\wedge_0$ as in the preceding paragraph, and one gets the same answer however this is done, because of (2.10.11). One can also check that the result is equal to the right side of (2.10.12), using (2.10.11). This defines an r-linear mapping from

$$\prod_{m=1}^{r} \mathcal{AM}_{k_m}(W) = \mathcal{AM}_{k_1}(W) \times \cdots \times \mathcal{AM}_{k_r}(W) \tag{2.10.13}$$

into $\mathcal{AM}_k(W)$.

This is a nice way to "multiply" alternating multilinear forms on W, but it is customary to use a different normalization. This will be discussed in the next section.

2.11 The wedge product

Let W be a vector space over the real numbers, and let k_1, k_2 be positive integers. If μ_1, μ_2 are alternating k_1, k_2-linear forms on W, respectively, then put

$$\mu_1 \wedge \mu_2 = \frac{(k_1+k_2)!}{k_1!\,k_2!}\,(\mu_1 \wedge_0 \mu_2) = \frac{(k_1+k_2)!}{k_1!\,k_2!}\,A_{k_1+k_2}(\mu_1 \otimes \mu_2), \tag{2.11.1}$$

which is an alternating $(k_1 + k_2)$-linear form on W. This corresponds to

$$\mu_1 \wedge_\alpha \mu_2 \tag{2.11.2}$$

in the notation used in (4) on p60 of [184]. Of course,

$$(\mu_1, \mu_2) \mapsto \mu_1 \wedge \mu_2 \tag{2.11.3}$$

is a bilinear mapping from $\mathcal{AM}_{k_1}(W) \times \mathcal{AM}_{k_2}(W)$ into $\mathcal{AM}_{k_1+k_2}(W)$.

Let k_3 be another positive integer, and let μ_3 be an alternating k_3-linear form on W. Using (2.10.11), one can check that

$$\begin{aligned} (\mu_1 \wedge \mu_2) \wedge \mu_3 &= \mu_1 \wedge (\mu_2 \wedge \mu_3) \\ &= \frac{(k_1 + k_2 + k_3)!}{k_1!\, k_2!\, k_3!} A_{k_1+k_2+k_3}(\mu_1 \otimes \mu_2 \otimes \mu_3). \end{aligned} \tag{2.11.4}$$

Similarly, let r be a positive integer, let $k_1, \ldots, k_r$ be r positive integers, and put $k = \sum_{m=1}^{r} k_m$, as before. If μ_m is an alternating k_m-linear form on W for each $m = 1, \ldots, r$, then

$$\begin{aligned} \mu_1 \wedge \cdots \wedge \mu_r &= \frac{k!}{k_1! \cdots k_r!} (\mu_1 \wedge_0 \cdots \wedge_0 \mu_r) \\ &= \frac{k!}{k_1! \cdots k_r!} A_k(\mu_1 \otimes \cdots \otimes \mu_r) \end{aligned} \tag{2.11.5}$$

is an alternating k-linear form, where the second step is as in (2.10.12). As before, the left side may be defined using the wedge product of alternating multilinear forms two at a time, and one gets the same answer however this is done, because of (2.11.4). To get the first step in (2.11.5), one should verify the coefficient on the right side. Note that this defines an r-linear mapping from $\prod_{m=1}^{r} \mathcal{AM}_{k_m}(W)$ into $\mathcal{AM}_k(W)$.

Suppose for the moment that $k_m = 1$ for each $m = 1, \ldots, r$, so that $k = r$. In this case, (2.11.5) implies that

$$\begin{aligned} \mu_1 \wedge \cdots \wedge \mu_k &= k!\, A_k(\mu_1 \otimes \cdots \otimes \mu_k) \\ &= \sum_{\sigma \in \mathrm{Sym}(k)} \mathrm{sgn}(\sigma)\, (\mu_1 \otimes \cdots \otimes \mu_k)^{\sigma}, \end{aligned} \tag{2.11.6}$$

using the definition of A_k in Subsection 2.4.1 in the second step.

Suppose for the moment again that $W = \mathbf{R}^n$ for some positive integer n, and let $\theta_1, \ldots, \theta_n$ be the usual coordinate functions on $\mathbf{R}^n$, as in Subsection 1.6.2. Using (2.11.6), we get that

$$\theta_1 \wedge \cdots \wedge \theta_n = \mu_{det}, \tag{2.11.7}$$

where μ_{det} is as in Section 1.15.

2.12 More on wedge products

Let k_1, k_2 be positive integers, and let ψ_{k_1,k_2} be the mapping from

$$\{1, \ldots, k_1 + k_2\} \tag{2.12.1}$$

into itself defined by

$$\begin{aligned} \psi_{k_1,k_2}(l) &= l + k_1 \quad \text{when } 1 \le l \le k_2 \\ &= l - k_2 \quad \text{when } k_2 + 1 \le l \le k_1 + k_2, \end{aligned} \tag{2.12.2}$$

It is easy to see that this defines a permutation on (2.12.1). One can check that

$$\operatorname{sgn}(\psi_{k_1,k_2}) = (-1)^{k_1\,k_2}. \tag{2.12.3}$$

Let W be a vector space over the real numbers, and let μ_1, μ_2 be k_1, k_2-linear forms on W, respectively. One can verify that

$$\mu_2 \otimes \mu_1 = (\mu_1 \otimes \mu_2)^{\psi_{k_1,k_2}}, \tag{2.12.4}$$

where the right side is as defined in Section 2.4. This corresponds to a remark in Subsection 2.8.2.

This implies that

$$\begin{aligned} A_{k_1+k_2}(\mu_2 \otimes \mu_1) &= A_{k_1+k_2}((\mu_1 \otimes \mu_2)^{\psi_{k_1,k_2}}) \\ &= \operatorname{sgn}(\psi_{k_1,k_2})\, A_{k_1+k_2}(\mu_1 \otimes \mu_2) \\ &= (-1)^{k_1\,k_2}\, A_{k_1+k_2}(\mu_1 \otimes \mu_2), \end{aligned} \tag{2.12.5}$$

where the second step is as in Subsection 2.4.1. If μ_1, μ_2 are alternating k_1, k_2-linear forms on W, then this is the same as saying that

$$\mu_2 \wedge_0 \mu_1 = (-1)^{k_1\,k_2}\,(\mu_1 \wedge_0 \mu_2). \tag{2.12.6}$$

Equivalently, this means that

$$\mu_2 \wedge \mu_1 = (-1)^{k_1\,k_2}\,(\mu_1 \wedge \mu_2). \tag{2.12.7}$$

It is convenient to put

$$\mathcal{M}_0(W) = \mathbf{R}, \tag{2.12.8}$$

so that a real number is considered as a 0-linear form on W. We can extend the previous definition of products of multilinear forms in Section 2.8 to include 0-linear forms, using scalar multiplication of multilinear forms by real numbers, and ordinary multiplication on $\mathbf{R}$.

Let us consider 0-linear forms on W to be both symmetric and alternating, so that

$$\mathcal{AM}_0(W) = \mathcal{SM}_0(W) = \mathbf{R} \tag{2.12.9}$$

as well. The symmetrization and alternatization operators S_k and A_k defined in Subsection 2.4.1 may be extended to $k = 0$ using the identity mapping on $\mathbf{R}$. Wedge products of alternating multilinear forms may be extended to include 0-linear forms too, using scalar multiplication by real numbers again, and ordinary multiplication on $\mathbf{R}$.

2.12.1 Linear mappings and $\mathcal{AM}(W)$

Let V be another vector space over the real numbers, and let T be a linear mapping from V into W. If μ_1, μ_2 are alternating k_1, k_2-linear forms on W again, then

$$\begin{aligned} T^*(\mu_1 \wedge_0 \mu_2) &= T^*(A_{k_1+k_2}(\mu_1 \otimes \mu_2)) = A_{k_1+k_2}(T^*(\mu_1 \otimes \mu_2)) \\ &= A_{k_1+k_2}(T^*(\mu_1) \otimes T^*(\mu_2)) = T^*(\mu_1) \wedge_0 T^*(\mu_2). \end{aligned} \tag{2.12.10}$$

More precisely, the second step here is as in Subsection 2.4.1, and the third step is as in Subsection 2.8.1. It follows that

$$T^*(\mu_1 \wedge \mu_2) = T^*(\mu_1) \wedge T^*(\mu_2). \tag{2.12.11}$$

Suppose now that W has dimension n for some positive integer n, as a vector space over the real numbers. Remember that

$$\mathcal{AM}_k(W) = \{0\} \text{ when } k > n \tag{2.12.12}$$

in this case, as in Subsection 1.11.1 and Section 2.4. This implies that

$$\mu_1 \wedge \mu_2 = 0 \text{ when } k_1 + k_2 > n. \tag{2.12.13}$$

Let us define $\mathcal{AM}(W)$ initially as a vector space over the real numbers by

$$\mathcal{AM}(W) = \bigoplus_{k=0}^{n} \mathcal{AM}_k(W), \tag{2.12.14}$$

where the right side is as in Section 2.1. We shall normally identify elements of $\mathcal{AM}_l(W)$ with their images in $\mathcal{AM}(W)$ under the embeddings mentioned earlier, so that every element of $\mathcal{AM}(W)$ corresponds to a sum of elements of $\mathcal{AM}_l(W)$, $0 \le l \le n$.

The wedge product has a natural extension to a bilinear mapping from $\mathcal{AM}(W) \times \mathcal{AM}(W)$ into $\mathcal{AM}(W)$. This makes $\mathcal{AM}(W)$ into an associative algebra over the real numbers. Note that $\mathbf{R}$ corresponds to a subalgebra of $\mathcal{AM}(W)$, because of (2.12.9), and that 1 is also the multiplicative identity element of $\mathcal{AM}(W)$.

2.13 Some more subgroups of $\mathrm{Sym}(k)$

In this and the next two sections, we shall consider some additional subgroups of $\mathrm{Sym}(k)$, and related properties of k-linear forms. This will be used to give another way to look at wedge products in Subsection 2.15.2. The reader may not want to dwell on these three sections too much at first.

Let r be a positive integer, let $k_1, \ldots, k_r$ be r positive integers, and put $k = \sum_{m=1}^{r} k_m$, as before. If $1 \le m \le r$, then put $L(m) = \sum_{i=1}^{m-1} k_i$, which is interpreted as being equal to 0 when $m = 1$, as in Section 2.8. Also put

$$E_m = \{L(m) + 1, \ldots, L(m) + k_m\}, \tag{2.13.1}$$

and note that

$$\bigcup_{m=1}^{r} E_m = \{1, \ldots, k\}, \tag{2.13.2}$$

and that the E_m's are pairwise-disjoint, as before.

If $\sigma_m \in \mathrm{Sym}(k_m)$, then let $\widehat{\sigma}_m$ be the mapping from $\{1, \ldots, k\}$ into itself defined by

$$\widehat{\sigma}_m(L(m) + l) = L(m) + \sigma(l) \tag{2.13.3}$$

for $l = 1, \ldots, k_m$, and taking $\widehat{\sigma}$ to be the identity mapping on the complement of E_m. It is easy to see that this defines an element of $\mathrm{Sym}(k)$, and that

$$(\widehat{\sigma_m \circ \tau_m}) = \widehat{\sigma}_m \circ \widehat{\tau}_m \tag{2.13.4}$$

for every $\sigma_m, \tau_m \in \mathrm{Sym}(k_m)$. This basically corresponds to some remarks in Section 2.7, by identifying $\mathrm{Sym}(k_m)$ with $\mathrm{Sym}(E_m)$ using the mapping

$$l \mapsto L(m) + l \tag{2.13.5}$$

from $\{1, \ldots, k_m\}$ onto E_m.

As before,

$$\sigma_m \mapsto \widehat{\sigma}_m \tag{2.13.6}$$

is a one-to-one mapping from $\mathrm{Sym}(k_m)$ into $\mathrm{Sym}(k)$, and a group homomorphism. We also have that

$$\mathrm{sgn}(\sigma_m) = \mathrm{sgn}(\widehat{\sigma}_m) \tag{2.13.7}$$

for every $\sigma_m \in \mathrm{Sym}(k_m)$, as before.

Suppose for the moment that $1 \le m_1 \ne m_2 \le r$, so that

$$E_{m_1} \cap E_{m_2} = \emptyset. \tag{2.13.8}$$

If $\sigma_{m_1} \in \mathrm{Sym}(k_{m_1})$ and $\sigma_{m_2} \in \mathrm{Sym}(k_{m_2})$, then it is easy to see that

$$\widehat{\sigma}_{m_1} \circ \widehat{\sigma}_{m_2} = \widehat{\sigma}_{m_2} \circ \widehat{\sigma}_{m_1}. \tag{2.13.9}$$

Consider the subset Σ of $\mathrm{Sym}(k)$ defined by

$$\Sigma = \{\widehat{\sigma}_1 \circ \cdots \circ \widehat{\sigma}_r : \sigma_m \in \mathrm{Sym}(k_m) \text{ for each } m = 1, \ldots, r\}. \tag{2.13.10}$$

One can check that this is a subgroup of $\mathrm{Sym}(k)$, as in Section 2.6. Equivalently, Σ consists of the elements of $\mathrm{Sym}(k)$ that map E_m onto itself for each $m = 1, \ldots, r$.

Consider the Cartesian product

$$\prod_{m=1}^{r} \mathrm{Sym}(k_m) = \mathrm{Sym}(k_1) \times \cdots \times \mathrm{Sym}(k_r), \tag{2.13.11}$$

which is the set of r-tuples $(\sigma_1, \ldots, \sigma_r)$ with $\sigma_m \in \mathrm{Sym}(k_m)$ for each $m = 1, \ldots, r$. By construction,

$$(\sigma_1, \ldots, \sigma_r) \mapsto \widehat{\sigma}_1 \circ \cdots \circ \widehat{\sigma}_r \tag{2.13.12}$$

maps this Cartesion product onto Σ. One can check that this mapping is one-to-one as well. This implies that the number of elements of Σ is given by

$$\#\Sigma = \prod_{m=1}^{r} (k_m!). \tag{2.13.13}$$

In fact, (2.13.11) may be considered as a group, where the group operations are defined coordinatewise. This is an example of a *direct product* of groups. One can verify that (2.13.12) is a homomorphism from this group onto Σ, using (2.13.9). More precisely, (2.13.12) is an *isomorphism* from (2.13.11) onto Σ, because this mapping is one-to-one, as in the preceding paragraph.

2.14 Subgroups and products of forms

Let us continue with the same notation and hypotheses as in the previous section. Also let W be a vector space over the real numbers, and let μ_m be a k_m-linear form on W for each $m = 1, \ldots, r$. Thus $\mu = \mu_1 \otimes \cdots \otimes \mu_r$ defines a k-linear form on W, as in Section 2.8.

Let $\sigma_m \in \mathrm{Sym}(k_m)$ be given for each $m = 1, \ldots, r$, and put

$$\sigma = \widehat{\sigma}_1 \circ \cdots \circ \widehat{\sigma}_r. \tag{2.14.1}$$

It is easy to see that

$$\mu^\sigma = \mu_1^{\sigma_1} \otimes \cdots \otimes \mu_m^{\sigma_m}, \tag{2.14.2}$$

using the notation from Section 2.4. If μ_m is a symmetric k_m-linear form on W for each $m = 1, \ldots, r$, then it follows that μ is symmetric with respect to Σ, as in Section 2.6.

Suppose for the moment that μ_m is an alternating k_m-linear form on W for each $m = 1, \ldots, r$. This implies that

$$\mu^\sigma = \Big(\prod_{m=1}^{r} \mathrm{sgn}(\sigma_m) \Big) \, \mu, \tag{2.14.3}$$

because of (2.14.2). Observe that

$$\prod_{m=1}^{r} \mathrm{sgn}(\sigma_m) = \prod_{m=1}^{r} \mathrm{sgn}(\widehat{\sigma}_m) = \mathrm{sgn}(\sigma), \tag{2.14.4}$$

using (2.13.7) in the first step. Thus

$$\mu^\sigma = \mathrm{sgn}(\sigma) \, \mu, \tag{2.14.5}$$

so that μ is alternating with respect to Σ, as in Section 2.6.

Remember that $S_\Sigma(\mu)$ may be defined as in Subsection 2.6.1. If μ_m is any k_m-linear form on W for each m again, then one can check that

$$S_\Sigma(\mu) = S_{k_1}(\mu_1) \otimes \cdots \otimes S_{k_r}(\mu_r), \tag{2.14.6}$$

using (2.14.2).

Similarly, $A_\Sigma(\mu)$ may be defined as in Subsection 2.6.1. One can verify that

$$A_\Sigma(\mu) = A_{k_1}(\mu_1) \otimes \cdots \otimes A_{k_r}(\mu_r), \tag{2.14.7}$$

using (2.14.2) and (2.14.4).

2.15 Some related shuffles

Let us continue with the same notation and hypotheses as in the previous two sections. Let us say that $\tau \in \mathrm{Sym}(k)$ is a *shuffle* with respect to $k_1, \ldots, k_r$ if for each $m = 1, \ldots, r$,

$$\text{the restriction of } \tau \text{ to } E_m \text{ is increasing.} \tag{2.15.1}$$

This means that for each $m = 1, \ldots, r$ and $1 \le l_1 < l_2 \le k_m$, we have that

$$\tau(L(m) + l_1) < \tau(L(m) + l_2). \tag{2.15.2}$$

This is mentioned with $r = 2$ on p60 of [184]. If τ is a shuffle with respect to $k_1, \ldots, k_r$, then for each $m = 1, \ldots, r$, the restriction of τ to E_m is uniquely determined by $\tau(E_m)$.

Let $\rho \in \mathrm{Sym}(k)$ be given, and put

$$C_m = \rho(E_m) \tag{2.15.3}$$

for each $m = 1, \ldots, r$. Observe that

$$\#C_m = \#E_m \tag{2.15.4}$$

for each $m = 1, \ldots, r$,

$$\bigcup_{m=1}^{r} C_m = \{1, \ldots, k\}, \tag{2.15.5}$$

and

$$\text{the } C_m\text{'s are pairwise-disjoint.} \tag{2.15.6}$$

If $C_1, \ldots, C_r$ is any sequence of r subsets of $\{1, \ldots, k\}$ that satisfies these three conditions, then it is easy to see that there is a unique shuffle $\tau \in \mathrm{Sym}(k)$ with respect to $k_1, \ldots, k_r$ such that

$$C_m = \tau(E_m) \tag{2.15.7}$$

for every $m = 1, \ldots, r$. Put $\sigma = \tau^{-1} \circ \rho$, so that $\sigma \in \mathrm{Sym}(k)$ and

$$\rho = \tau \circ \sigma. \tag{2.15.8}$$

Clearly

$$\sigma(E_m) = E_m \tag{2.15.9}$$

for each $m = 1, \ldots, r$, by construction.

2.15.1 The set of shuffles

Remember that Σ is the subgroup of $\mathrm{Sym}(k)$ consisting of $\sigma \in \mathrm{Sym}(k)$ that satisfy (2.15.9), as in Section 2.13. Let Ξ be the set of $\tau \in \mathrm{Sym}(k)$ such that τ is a shuffle with respect to $k_1, \ldots, k_r$. Consider the Cartesian product

$$\Xi \times \Sigma, \tag{2.15.10}$$

which is the set of ordered pairs (τ, σ) with $\tau \in \Xi$ and $\sigma \in \Sigma$. Note that

$$(\tau, \sigma) \mapsto \tau \circ \sigma \tag{2.15.11}$$

defines a mapping from (2.15.10) into $\mathrm{Sym}(k)$. In fact, this is a one-to-one mapping from (2.15.10) onto $\mathrm{Sym}(k)$, by the remarks in the preceding paragraph.

We also have that

$$\#\Xi = \binom{k}{k_1 \cdots k_r} = \frac{k!}{k_1! \cdots k_r!}. \tag{2.15.12}$$

More precisely, the number of shuffles $\tau \in \mathrm{Sym}(k)$ with respect to $k_1, \ldots, k_r$ is the same as the number of ways of partitioning $\{1, \ldots, k\}$ into r subsets $C_1, \ldots, C_r$, where $\#C_m = k_m$ for each $m = 1, \ldots, r$, as before. The number of these partitions is the same as the multinomial coefficient in (2.15.12), as in Section 1.14. Alternatively,

$$(\#\Xi)\,(\#\Sigma) = \#\,\mathrm{Sym}(k), \tag{2.15.13}$$

as in the preceding paragraph. This implies (2.15.12), because $\#\,\mathrm{Sym}(k) = k!$ and $\#\Sigma = \prod_{m=1}^{r}(k_m!)$.

2.15.2 Symmetrization, alternatization, and shuffles

If ν is a k-linear form on W, then

$$S_k(\nu) = \frac{1}{k!} \sum_{\rho \in \mathrm{Sym}(k)} \nu^\rho = \frac{1}{k!} \sum_{\tau \in \Xi} \sum_{\sigma \in \Sigma} \nu^{\tau \circ \sigma} = \frac{1}{k!} \sum_{\tau \in \Xi} \sum_{\sigma \in \Sigma} (\nu^\sigma)^\tau, \tag{2.15.14}$$

where the first and third steps are as in Section 2.4, and the second step uses the earlier remarks about (2.15.11). This implies that

$$S_k(\nu) = \frac{1}{k!} \sum_{\tau \in \Xi} \Big(\sum_{\sigma \in \Sigma} \nu^\sigma \Big)^\tau = \frac{1}{\#\Xi} \sum_{\tau \in \Xi} S_\Sigma(\nu)^\tau, \tag{2.15.15}$$

using (2.15.13) and the definition of $S_\Sigma(\nu)$ in Subsection 2.6.1 in the second step. In particular, if ν is symmetric with respect to Σ, then

$$S_k(\nu) = \frac{1}{\#\Xi} \sum_{\tau \in \Xi} \nu^\tau. \tag{2.15.16}$$

Similarly,

$$
\begin{aligned}
A_k(\nu) &= \frac{1}{k!} \sum_{\rho \in \mathrm{Sym}(k)} \mathrm{sgn}(\rho)\, \nu^\rho = \frac{1}{k!} \sum_{\tau \in \Xi} \sum_{\sigma \in \Sigma} \mathrm{sgn}(\tau \circ \sigma)\, \nu^{\tau \circ \sigma} \\
&= \frac{1}{k!} \sum_{\tau \in \Xi} \sum_{\sigma \in \Sigma} \mathrm{sgn}(\tau)\ \mathrm{sgn}(\sigma)\, (\nu^\sigma)^\tau. \qquad (2.15.17)
\end{aligned}
$$

It follows that

$$
A_k(\nu) = \frac{1}{k!} \sum_{\tau \in \Xi} \mathrm{sgn}(\tau) \Big(\sum_{\sigma \in \Sigma} \mathrm{sgn}(\sigma)\, \nu^\sigma \Big)^\tau = \frac{1}{\#\Xi} \sum_{\tau \in \Xi} \mathrm{sgn}(\tau)\, A_\Sigma(\nu)^\tau. \qquad (2.15.18)
$$

If ν is alternating with respect to Σ, then we get that

$$
A_k(\nu) = \frac{1}{\#\Xi} \sum_{\tau \in \Xi} \mathrm{sgn}(\tau)\, \nu^\tau. \qquad (2.15.19)
$$

Suppose that μ_m is an alternating k_m-linear form on W for each $m = 1, \ldots, r$, so that $\mu = \mu_1 \otimes \cdots \otimes \mu_r$ is alternating with respect to Σ, as in the previous section. Observe that

$$
\mu_1 \wedge_0 \cdots \wedge_0 \mu_r = A_k(\mu) = \frac{1}{\#\Xi} \sum_{\tau \in \Xi} \mathrm{sgn}(\tau)\, \mu^\tau, \qquad (2.15.20)
$$

where the first step is as in Subsection 2.10.1, and the second step is as in (2.15.19). It follows that

$$
\mu_1 \wedge \cdots \wedge \mu_r = \sum_{\tau \in \Xi} \mathrm{sgn}(\tau)\, \mu^\tau, \qquad (2.15.21)
$$

using (2.15.12) and a remark in Section 2.11. This corresponds to (2) on p60 of [184] when $r = 2$.

Chapter 3

Some geometry and analysis

3.1 Metric spaces

Let X be a set. A nonnegative real-valued function $d(x, y)$ defined for $x, y \in X$ is said to be a *metric* on X if it satisfies the following three conditions. First,

$$d(x, y) = 0 \text{ is and only if } x = y. \tag{3.1.1}$$

Second,

$$d(x, y) = d(y, x) \tag{3.1.2}$$

for every $x, y \in X$. Third,

$$d(x, z) \le d(x, y) + d(y, z) \tag{3.1.3}$$

for every $x, y, z \in X$. This third condition is known as the *triangle inequality.* If $d(x, y)$ is a metric on X, then $(X, d(x, y))$ is called a *metric space*.

If t is a real number, then the *absolute value* $|t|$ of t is defined as usual by $|t| = t$ when $t \ge 0$, and $|t| = -t$ when $t \le 0$. It is easy to see that

$$|r + t| \le |r| + |t| \tag{3.1.4}$$

for all $r, t \in \mathbf{R}$. One can use this to check that

$$d(x, y) = |x - y| \tag{3.1.5}$$

defines a metric on the real line, which is the *standard Euclidean metric* on $\mathbf{R}$.

If X is any set, then the *discrete metric* on X is defined by

$$\begin{aligned} d(x, y) &= 1 \quad \text{when } x \ne y \\ &= 0 \quad \text{when } x = y. \end{aligned} \tag{3.1.6}$$

One can check that this defines a metric on X.

Let $(X, d(x,y))$ be a metric space, and let E be a subset of X. It is easy to see that

$$\text{the restriction of } d(x,y) \text{ to } x, y \in E \tag{3.1.7}$$

defines a metric on E.

We shall often be concerned here with metrics obtained from norms on vector spaces, as in the next section. More precisely, we shall often be concerned with norms obtained from inner products, as in Section 3.3.

3.2 Norms on vector spaces

Let V be a vector space over the real numbers. A nonnegative real-valued function N on V is said to be a *norm* if it satisfies the following three conditions. First,

$$N(v) = 0 \text{ if and only if } v = 0. \tag{3.2.1}$$

Second,

$$N(t\,v) = |t|\,N(v) \tag{3.2.2}$$

for every $t \in \mathbf{R}$ and $v \in V$. Third,

$$N(v + w) \le N(v) + N(w) \tag{3.2.3}$$

for every $v, w \in V$, which is the *triangle inequality* for a norm.

Let n be a positive integer, and remember that $\mathbf{R}^n$ is a vector space over the real numbers with respect to coordinatewise addition and scalar multiplication. The *standard Euclidean norm* is defined on $\mathbf{R}^n$ by

$$\|v\|_2 = \Big(\sum_{j=1}^{n} v_j^2\Big)^{1/2}, \tag{3.2.4}$$

using the nonnegative square root on the right side. It is easy to see that this satisfies the first two conditions in the definition of a norm, but the triangle inequality is more complicated when $n \ge 2$. This will be discussed further in the next section.

One can check directly that

$$\|v\|_1 = \sum_{j=1}^{n} |v_j| \tag{3.2.5}$$

defines a norm on $\mathbf{R}^n$. One can also verify that

$$\|v\|_\infty = \max_{1 \le j \le n} |v_j| \tag{3.2.6}$$

defines a norm on $\mathbf{R}^n$. More precisely, to get that this satisfies the triangle inequality, one can observe that

$$\|v + w\|_\infty = \max_{1 \le j \le n} |v_j + w_j| \le \max_{1 \le j \le n} (|v_j| + |w_j|) \le \|v\|_\infty + \|w\|_\infty \tag{3.2.7}$$

for every $v, w \in \mathbf{R}^n$.

Note that

$$\|v\|_\infty \le \|v\|_1,\ \|v\|_2 \tag{3.2.8}$$

for every $v \in \mathbf{R}^n$. We also have that

$$\|v\|_1 \le n\,\|v\|_\infty \tag{3.2.9}$$

and

$$\|v\|_2 \le \sqrt{n}\,\|v\|_\infty \tag{3.2.10}$$

for every $v \in \mathbf{R}^n$.

If $v \in \mathbf{R}^n$, then

$$\|v\|_2^2 = \sum_{j=1}^n v_j^2 \le \|v\|_1\,\|v\|_\infty \le \|v\|_1^2, \tag{3.2.11}$$

using (3.2.8) in the third step. This implies that

$$\|v\|_2 \le \|v\|_1. \tag{3.2.12}$$

If N is a norm on any vector space V over the real numbers, then it is easy to see that

$$d_N(v, w) = N(v - w) \tag{3.2.13}$$

defines a metric on V. The *standard Euclidean metric* on $\mathbf{R}^n$ is the metric

$$d_2(v, w) = \|v - w\|_2 \tag{3.2.14}$$

associated to the standard Euclidean norm $\|\cdot\|_2$ on $\mathbf{R}^n$. Similarly,

$$d_1(v, w) = \|v - w\|_1 \tag{3.2.15}$$

and

$$d_\infty(v, w) = \|v - w\|_\infty \tag{3.2.16}$$

define metrics on $\mathbf{R}^n$ too.

If W is a linear subspace of V and N is a norm on V, then

(3.2.17) the restriction of N to W defines a norm on W

as well.

3.3 Inner product spaces

Let V be a vector space over the real numbers, and let $\langle v, w\rangle$ be a symmetric bilinear form on V. If

$$\langle v, v\rangle > 0 \tag{3.3.1}$$

for every $v \in V$ with $v \neq 0$, then $\langle \cdot, \cdot \rangle$ is said to be an *inner product* on V. Under these conditions, $(V, \langle \cdot, \cdot \rangle)$ is called an *inner product space.*

If n is a positive integer, then the *standard inner product* on $\mathbf{R}^n$ is defined by

$$\langle v, w \rangle = \sum_{j=1}^{n} v_j \, w_j. \tag{3.3.2}$$

In this case,

$$\langle v, v \rangle = \sum_{j=1}^{n} v_j^2 = \|v\|_2^2 \tag{3.3.3}$$

for every $v \in \mathbf{R}^n$, which is strictly positive when $v \neq 0$.

Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space, and put

$$\|v\| = \langle v, v \rangle^{1/2} \tag{3.3.4}$$

for every $v \in V$, using the nonnegative square root on the right side. Note that

$$\|t\, v\| = \langle t\, v, t\, v \rangle^{1/2} = |t| \, \|v\| \tag{3.3.5}$$

for every $v \in V$ and $t \in \mathbf{R}$. If $v, w \in V$, then it is well known that

$$|\langle v, w \rangle| \leq \|v\| \, \|w\|, \tag{3.3.6}$$

which is known as the *Cauchy–Schwarz inequality.* This can be obtained from the fact that

$$\langle v - t\, w, v - t\, w \rangle \geq 0 \tag{3.3.7}$$

for every $t \in \mathbf{R}$.

One can use the Cauchy–Schwarz inequality to show that

$$\|v + w\| \leq \|v\| + \|w\| \tag{3.3.8}$$

for every $v, w \in V$. This implies that

$$\| \cdot \| \text{ is a norm on } V. \tag{3.3.9}$$

Of course, the standard Euclidean norm on $\mathbf{R}^n$ is the same as the norm associated to the standard inner product on $\mathbf{R}^n$.

If W is a linear subspace of V, then

(3.3.10) the restriction of $\langle v, w \rangle$ to $v, w \in W$ defines an inner product on W.

3.4 Continuous functions on $\mathbf{R}^n$

Let n be a positive integer, and let E be a nonempty subset of $\mathbf{R}^n$. Also let f be a real-valued function on E, and let x be an element of E. One can define what it means for f to be *continuous at* x as a function defined on E in a standard way. In fact, one can define what it means for a function defined on a subset of a metric space with values in another metric space to be continuous at a point in the set on which the function is defined in essentially the same way. Here we are using the standard Euclidean metrics on $\mathbf{R}^n$ and $\mathbf{R}$.

We say that f is *continuous on E* if f is continuous at every point in E. Let

$$C(E) = C(E, \mathbf{R}) \tag{3.4.1}$$

be the space of all continuous real-valued functions on E. It is well known that this is a linear subspace of the space of all real-valued functions on E. In particular, $C(E)$ is a vector space over the real numbers with respect to pointwise addition and scalar multiplication of functions on E.

If f, g are continuous real-valued functions on E, then it is well known that their product

$$f\,g \text{ is continuous on } E \tag{3.4.2}$$

too. Using this, it is easy to see that polynomials are continuous on $\mathbf{R}^n$. Thus the space $\mathcal{P}(\mathbf{R}^n)$ of all polynomials on $\mathbf{R}^n$ with real coefficients may be considered as a linear subspace of $C(\mathbf{R}^n)$.

If f is a continuous real-valued function on E, and $f(x) \neq 0$ for each $x \in E$, then it is well known that

$$1/f \text{ is continuous on } E. \tag{3.4.3}$$

This implies that rational functions are continuous on sets where the denominator is nonzero.

Let m be a positive integer, and let f be a function on E with values in $\mathbf{R}^m$. If $x \in E$, then $f(x)$ can be expressed as

$$f(x) = (f_1(x), \ldots, f_m(x)), \tag{3.4.4}$$

where $f_1(x), \ldots, f_m(x)$ are real numbers. Thus f corresponds to m real-valued functions $f_1, \ldots, f_m$ on E. If $x \in E$, then one can define what it means for f to be *continuous at x* in a standard way, using the standard Euclidean metric on $\mathbf{R}^m$. It is well known that this is equivalent to the continuity of $f_1, \ldots, f_m$ at x as real-valued functions on E.

As before, we say that f is *continuous on E* if f is continuous at every point in E. Let

$$C(E, \mathbf{R}^m) \tag{3.4.5}$$

be the space of all continuous functions on E with values in $\mathbf{R}^m$. The space of all $\mathbf{R}^m$-valued functions on E is a vector space over the real numbers, with respect to pointwise addition and scalar multiplication. It is easy to see that $C(E, \mathbf{R}^m)$ is a linear subspace of the space of all $\mathbf{R}^m$-valued functions on E. Thus $C(E, \mathbf{R}^m)$ is a vector space over the real numbers with respect to pointwise addition and scalar multiplication of functions on E.

Let A be a nonempty subset of $\mathbf{R}^m$, and suppose that $f(x) \in A$ for every $x \in E$, so that

$$f(E) \subseteq A. \tag{3.4.6}$$

Also let k be another positive integer, and let g be a function on A with values in $\mathbf{R}^k$. Thus the composition $g \circ f$ of f and g may be defined as a function on E with values in $\mathbf{R}^k$, with

$$(g \circ f)(x) = g(f(x)) \tag{3.4.7}$$

for every $x \in E$. If f is continuous at x, and g is continuous at $f(x)$, then it is well known that $g \circ f$ is continuous at x. In particular, if f is continuous on E, and g is continuous on A, then $g \circ f$ is continuous on E.

3.4.1 Open balls and open sets

If $x \in \mathbf{R}^n$ and r is a positive real number, then the *open ball* in $\mathbf{R}^n$ centered at x with radius r with respect to the standard Euclidean metric is defined as usual by

$$B(x, r) = \{y \in \mathbf{R}^n : \|x - y\|_2 < r\}, \tag{3.4.8}$$

where $\|\cdot\|_2 = \|\cdot\|_{2,\mathbf{R}^n}$ is the standard Euclidean norm on $\mathbf{R}^n$, as in Section 3.2. A subset U of $\mathbf{R}^n$ is said to be an *open set* with respect to the standard Euclidean metric if for every $x \in U$ there is an $r > 0$ such that

$$B(x, r) \subseteq U. \tag{3.4.9}$$

It is well known that open balls in $\mathbf{R}^n$ are open sets. This can be shown using the triangle inequality. More precisely, one can define open balls and open sets in any metric space, and one can show that open balls are open sets in the same way.

3.5 Partial derivatives

Let n be a positive integer, and let U be a nonempty open subset of $\mathbf{R}^n$, with respect to the standard Euclidean metric. Also let f be a real-valued function on U, and let $x \in U$ be given. The *partial derivative* of f at x in the jth variable, $1 \le j \le n$, may be denoted

$$\partial_j f(x) = \frac{\partial f}{\partial x_j}(x), \tag{3.5.1}$$

when it exists.

Similarly, let m be a positive integer, and suppose now that f is a function on U with values in $\mathbf{R}^m$. The partial derivative of f at x in the jth variable may be denoted as in (3.5.1) again, when it exists, which is an element of $\mathbf{R}^m$. This happens exactly when the partial derivative of each of the components $f_1, \ldots, f_m$ at x in the jth variable exists, in which case

$$\partial_j f(x) = (\partial_j f_1(x), \ldots, \partial_j f_m(x)). \tag{3.5.2}$$

3.5.1 Continuous differentiability

We say that f is *continuously differentiable* on U if

(3.5.3) $\partial_j f(x)$ exists at every point $x \in U$ for each $j = 1, \ldots, n$, and is continuous as an $\mathbf{R}^m$-valued function on U.

Equivalently, this means that each of the components $f_1, \ldots, f_m$ of f is continuously differentiable on U in the same sense as a real-valued function on U. It is well known that this implies that f is continuous on U. Of course, if $n = 1$, then continuity at a point is implied by the existence of the derivative at that point.

Let

$$C^1(U, \mathbf{R}^m) \tag{3.5.4}$$

be the space of continuously-differentiable $\mathbf{R}^m$-valued functions on U. This is a linear subspace of the space $C(U, \mathbf{R}^m)$ of all continuous $\mathbf{R}^m$-valued functions on U.

If f, g are continuously-differentiable real-valued functions on U, then one can check that

(3.5.5) $f\,g$ is continuously differentiable on U,

using the product rule. Note that polynomials are continuously differentiable on $\mathbf{R}^n$.

If f is a continuously-differentiable real-valued function on U and $f(x) \neq 0$ for every $x \in U$, then one can verify that

(3.5.6) $1/f$ is continuously differentiable on U.

It follows that rational functions are continuously differentiable on the complement of the set where the denominator is equal to 0.

3.5.2 Second derivatives

Let f be an $\mathbf{R}^m$-valued function on U again, and suppose that the partial derivatives of f exist at each point in U and in each variable. The partial derivative of $\partial_l f$ at x in the jth variable may be denoted

$$\partial_j \partial_l f(x) = \frac{\partial^2 f}{\partial x_j \, \partial x_l}(x), \tag{3.5.7}$$

when it exists. If

(3.5.8) the partial derivatives $\partial_1 f, \ldots, \partial_n f$ of f are all continuously differentiable on U,

then f is said to be *twice continuously differentiable* on U. This implies that the partial derivatives of f are continuous on U, and that f is continuous on U, as before. Note that f is twice continuously differentiable on U if and only if the components $f_1, \ldots, f_m$ of f are all twice continuously differentiable as real-valued functions on U.

If f is twice continuously differentiable on U, then it is well known that

$$\frac{\partial^2 f}{\partial x_j \, \partial x_l} = \frac{\partial^2 f}{\partial x_l \, \partial x_j} \tag{3.5.9}$$

on U for all $j, l = 1, \ldots, n$. This is normally stated for real-valued functions, which implies the analogouus result for $\mathbf{R}^m$-valued functions.

3.6 Additional regularity

Let m, n, and r be positive integers, and let U be a nonempty open subset of $\mathbf{R}^n$ again. If f is an $\mathbf{R}^m$-valued function on U, then let us say that f is *r-times continuously differentiable* on U if

(3.6.1) all of the derivatives of f of order up to r exist and are continuous on U.

This is equivalent to continuous differentiability of f on U when $r = 1$, and to twice continuous differentiability of f on U when $r = 2$. It is sometimes convenient to interpret this with $r = 0$ as meaning that f is continuous on U.

More precisely, f is r-times continuously differentiable on U if

(3.6.2) all of the derivatives of f of order up to r exist on U, and the derivatives of f of order r are continuous on U.

This corresponds to the previous definitions of continuous differentiability and twice continuous differentiability of f when $r = 1$ and 2, respectively. If $r \geq 2$, then this is the same as saying that

(3.6.3) all of the derivatives of f of order $r - 1$ are continuously differentiable on U.

In particular, this implies that all of the derivatives of f of order $r - 1$ are continuous on U, as before. One can repeat the process to get that the derivatives of f of all orders up to r are continuous on U.

Alternatively, if $r \geq 2$, then f is r-times continuously differentiable on U if and only if

(3.6.4) the partial derivatives $\partial_1 f, \ldots, \partial_n f$ exist on U and are $(r - 1)$-times continuously differentiable on U.

Note that f is r-times continuously differentiable on U if and only if each of the components $f_1, \ldots, f_m$ of f is r-times continuously differentiable as a real-valued function on U.

3.6.1 More on r-times continuous differentiability

Let

$$C^r(U, \mathbf{R}^m) \tag{3.6.5}$$

be the space of all $\mathbf{R}^m$-valued functions on U that are r-times continuously differentiable on U. This may be interpreted as being the space $C(U, \mathbf{R}^m)$ of all continuous $\mathbf{R}^m$-valued functions on U when $r = 0$, as before. It is easy to see that

$$C^r(U, \mathbf{R}^m) \text{ is a linear subspace of } C(U, \mathbf{R}^m) \tag{3.6.6}$$

for each r. We also have that

$$C^{r+1}(U, \mathbf{R}^m) \subseteq C^r(U, \mathbf{R}^m) \tag{3.6.7}$$

for each r.

Let $j_1, \ldots, j_r$ be r positive integers less than or equal to n. Suppose that the derivatives of f on U up to order r exist, and consider the rth-order derivative

$$\partial_{j_1} \cdots \partial_{j_r} f \tag{3.6.8}$$

of f on U. If σ is a permutation on $\{1, \ldots, r\}$, then

$$\partial_{j_{\sigma(1)}} \cdots \partial_{j_{\sigma(n)}} f \tag{3.6.9}$$

is an rth-order derivative of f on U as well. If f is r-times continuously differentiable on U, then it is well known that (3.6.8) is equal to (3.6.9). The $r = 2$ case was mentioned in Subsection 3.5.2.

If $r > 2$, then the previous case implies that (3.6.8) is equal to (3.6.9) when σ is a transposition of a pair of consecutive elements of $\{1, \ldots, r\}$. To get that this holds for all permutations σ on $\{1, \ldots, r\}$, one can show that every such permutation is the composition of finitely many transpositions of this type.

Let α be a multi-index with

$$\sum_{j=1}^{n} \alpha_j = r. \tag{3.6.10}$$

We may use the notation

$$\partial^\alpha f = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} f = \frac{\partial^r f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} \tag{3.6.11}$$

for the corresponding derivative of f of other r. Of course, this means that no derivative is taken in the jth variable when $\alpha_j = 0$. If f is r-times continuously differentiable on U, then the order in which the derivatives are taken does not matter, as in the previous two paragraphs, and every rth-order derivative of f is of this form for some α.

If f, g are real-valued functions on U that are r-times continuously differentiable for some r, then one can check that

$$f\, g \text{ is } r\text{-times continuously differentiable on } U. \tag{3.6.12}$$

If $f(x) \neq 0$ for every $x \in U$, then one can verify that

$$1/f \text{ is } r\text{-times continuously differentiable on } U. \tag{3.6.13}$$

3.7 Smooth functions

Let m and n be positive integers, let U be a nonempty open subset of $\mathbf{R}^n$, and let f be an $\mathbf{R}^m$-valued function on U again. Let us say that f is *infinitely differentiable* on U if

(3.7.1) $\quad f$ is r-times continuously differentiable on U for every $r \geq 1$.

Sometimes f is simply said to be *smooth* on U in this case. As before, f is infinitely differentiable on U if and only if each of the components $f_1, \ldots, f_m$ is infinitely differentiable as a real-valued function on U. The space of infinitely differentiable $\mathbf{R}^m$-valued functions on U is denoted

$$C^\infty(U, \mathbf{R}^m) = \bigcap_{r=1}^{\infty} C^r(U, \mathbf{R}^m), \tag{3.7.2}$$

which is a linear subspace of $C^r(U, \mathbf{R}^m)$ for each r.

If f and g are infinitely-differentiable real-valued functions on U, then

(3.7.3) $\quad f\,g$ is infinitely differentiable on U.

Of course, polynomials on $\mathbf{R}^n$ are infinitely differentiable. If $f(x) \neq 0$ for every $x \in U$, then

(3.7.4) $\quad 1/f$ is infinitely differentiable on U.

This implies that rational functions are infinitely differentiable on the complement of the set where the denominator is equal to 0.

3.7.1 Compositions with functions on R

Let V be an open set in the real line, and let ψ be a real-valued function on V. Suppose that

(3.7.5) $\quad f$ takes values in V on U,

so that the composition $\psi \circ f$ is defined as a real-valued function on U. If the partial derivative of f at $x \in U$ in the jth variable exists, and if

(3.7.6) $\quad \psi$ is differentiable at $f(x) \in V$,

then the partial derivative of $\psi \circ f$ in the jth variable exists at x, with

$$\frac{\partial(\psi \circ f)}{\partial x_j}(x) = \psi'(f(x))\,\frac{\partial f}{\partial x_j}(x). \tag{3.7.7}$$

This follows from the ordinary chain rule for functions of one variable, by considering f and $\psi \circ f$ as functions of the jth variable, with the other variables kept fixed.

If f is continuously differentiable on U, and ψ is continuously differentiable on V, then it follows that

(3.7.8) $\quad \psi \circ f$ is continuously differentiable on U.

If f is r-times continuously differentiable on U, and ψ is r-times continuously differentiable on V, then one can check that

$$\psi \circ f \text{ is } r\text{-times continuously differentiable on } U. \tag{3.7.9}$$

If f is infinitely differentiable on U, and ψ is infinitely differentiable on V, then we get that

$$\psi \circ f \text{ is infinitely differentiable on } U. \tag{3.7.10}$$

3.7.2 A smooth function on R

Consider the real-valued function defined on the real line by

$$\begin{aligned} \phi(x) &= \exp(-1/x) \quad \text{when } x > 0 \\ &= 0 \qquad\qquad\quad \text{when } x \le 0. \end{aligned} \tag{3.7.11}$$

It is easy to see that ϕ is infinitely differentiable when $x \neq 0$, using the remarks in the preceding paragraph when $x > 0$. It is well known and not too difficult to show that

$$\phi \text{ is infinitely differentiable on } \mathbf{R}, \tag{3.7.12}$$

with derivatives of all orders of ϕ at 0 equal to 0.

3.8 Differentiability

Let m and n be positive integers, let U be an open subset of $\mathbf{R}^n$, and let f be a mapping from U into $\mathbf{R}^m$. We say that f is *differentiable* at a point $x \in U$ if there is a linear mapping A from $\mathbf{R}^n$ into $\mathbf{R}^m$ such that

$$\lim_{h \to 0} \frac{\|f(x+h) - f(x) - A(h)\|_{2,\mathbf{R}^m}}{\|h\|_{2,\mathbf{R}^n}} = 0. \tag{3.8.1}$$

Here $\|\cdot\|_{2,\mathbf{R}^m}$ and $\|\cdot\|_{2,\mathbf{R}^n}$ are the standard Euclidean norms on $\mathbf{R}^m$ and $\mathbf{R}^n$, respectively, as in Section 3.2. Note that the quotient whose limit is being considered is defined for $h \in \mathbf{R}^n$ with $h \neq 0$ and $\|h\|_{2,\mathbf{R}^n}$ sufficiently small.

If f is differentiable at x, then one can check that

$$f \text{ is continuous at } x. \tag{3.8.2}$$

One can also check directly that the linear mapping A is unique if this case, and another way to see this will be mentioned in Subsection 3.8.1. If $n = 1$, then this definition of differentiability reduces to the usual notion of differentiability of a function of one variable, where A corresponds to multiplying a real number by the ordinary derivative of f at x.

If f is differentiable at x, then we may put

$$f'(x) = A, \tag{3.8.3}$$

where A is as in (3.8.1). This may be called the *differential* of f at x. Although this is similar to the notation for the usual derivative when $n = 1$, one should keep in mind that this is a linear mapping from $\mathbf{R}^n$ into $\mathbf{R}^m$. We may also use the notation

$$df_x \tag{3.8.4}$$

for the differential of f at x, when it exists.

3.8.1 Directional derivatives

If $v \in \mathbf{R}^n$, then

$$f(x + t\,v) \tag{3.8.5}$$

may be considered as an $\mathbf{R}^m$-valued function of t that is defined on an open set in the real line that contains 0. The *directional derivative* of f at x in the direction v is defined to be

$$\frac{d}{dt} f(x + t\,v) \text{ at } t = 0, \tag{3.8.6}$$

if the derivative exists, as usual. If f is differentiable at x and $v \in \mathbf{R}^n$, then one can check that the directional derivative of f at x in the direction v exists, and is equal to

$$f'(x)(v) = df_x(v). \tag{3.8.7}$$

In particular, this means that the partial derivative of f at x in the jth variable exists for each $j = 1, \ldots, n$, with

$$\frac{\partial f}{\partial x_j}(x) = f'(x)(e_j) = df_x(e_j), \tag{3.8.8}$$

where $e_1, \ldots, e_n$ are the usual standard basis vectors for $\mathbf{R}^n$. This can be used to get the uniqueness of the differential of f at x when it exists, as mentioned earlier.

3.8.2 Some properties of the differential

Of course, if f is constant on U, then f is differentiable at every $x \in U$, with

$$f'(x) = 0. \tag{3.8.9}$$

If f is a linear mapping from $\mathbf{R}^n$ into $\mathbf{R}^m$, then f is differentiable at every $x \in \mathbf{R}^n$, with differential equal to f. If f is any continuously-differentiable mapping from U into $\mathbf{R}^m$, then it is well known that

$$f \text{ is differentiable at every } x \in U. \tag{3.8.10}$$

If f is any mapping from U into $\mathbf{R}^m$ that is differentiable at $x \in U$, then

$$f'(x)(v) = df_x(v) = \sum_{j=1}^{n} v_j \, \frac{\partial f}{\partial x_j}(x) \tag{3.8.11}$$

for every $v \in \mathbf{R}^n$, by (3.8.8).

As usual, one can verify that f is differentiable at x if and only if its m components $f_1, \ldots, f_m$ are each differentiable at x as real-valued functions on U. In this case, if $v \in \mathbf{R}^n$, then

$$f_1'(x)(v), \ldots, f_m'(x)(v) \tag{3.8.12}$$

are the m components of $f'(x)(v)$.

If f is differentiable at x and $t \in \mathbf{R}$, then it is easy to see that $t\,f$ is differentiable at x too, as an $\mathbf{R}^m$-valued function on U, with

$$(t\,f)'(x) = t\,f'(x). \tag{3.8.13}$$

If g is another mapping from U into $\mathbf{R}^m$ that is differentiable at x, then one can verify that $f+g$ is differentiable at x as an $\mathbf{R}^m$-valued function on U, with

$$(f+g)'(x) = f'(x) + g'(x). \tag{3.8.14}$$

If f, g are real-valued functions on U that are differentiable at x, then it is not too difficult to show that $f\,g$ is differentiable at x, with

$$(f\,g)'(x) = g(x)\,f'(x) + f(x)\,g'(x). \tag{3.8.15}$$

3.9 The chain rule

Let m, n, and p be positive integers, let U be an open subset of $\mathbf{R}^n$, and let V be an open subset of $\mathbf{R}^m$. Also let f be a mapping from U into V, and let g be a mapping from V into $\mathbf{R}^p$. Thus the composition $g \circ f$ of f and g is defined as a mapping from U into $\mathbf{R}^p$. Suppose that

$$f \text{ is differentiable at } x \in U, \tag{3.9.1}$$

and that

$$g \text{ is differentiable at } f(x) \in V, \tag{3.9.2}$$

so that the differentials $f'(x)$ and $g'(f(x))$ of f and g at x and $f(x)$, respectively, are defined as linear mappings from $\mathbf{R}^n$ and $\mathbf{R}^m$ into $\mathbf{R}^m$ and $\mathbf{R}^p$, respectively. There is a version of the *chain rule* that states that

$$g \circ f \text{ is differentiable at } x, \tag{3.9.3}$$

with

$$(g \circ f)'(x) = g'(f(x)) \circ f'(x). \tag{3.9.4}$$

Equivalently, this means that

$$d(g \circ f)_x = dg_{f(x)} \circ df_x, \tag{3.9.5}$$

as a linear mapping from $\mathbf{R}^n$ into $\mathbf{R}^p$. It is easy to see that this reduces to the usual chain rule for real-valued functions of a single variable when $m = n = p = 1$.

Of course, a linear mapping from $\mathbf{R}^n$ into $\mathbf{R}^m$ corresponds to an $m \times n$ *matrix* of real numbers in a standard way. Similarly, a linear mapping from $\mathbf{R}^m$ into $\mathbf{R}^p$ corresponds to a $p \times m$ matrix. The composition of these two linear mappings corresponds to a $p \times n$ matrix, which can be expressed in terms of the other two matrices using *matrix multiplication*.

The entries of the matrices associated to $f'(x)$ and $g'(f(x))$ are given by the partial derivatives of the components of f and g at x and $f(x)$, respectively. The chain rule implies that

(3.9.6) the partial derivatives of the components of $g \circ f$ at x
can be expressed as a sum of products of the partial derivatives
of the components of f and g at x and $f(x)$, respectively,
using matrix multiplication.

If f and g are continuously differentiable on U and V, respectively, then one can use this to check that

$$g \circ f \text{ is continuously differentiable on } U. \tag{3.9.7}$$

More precisely, this uses the fact that

$$(\partial_l g)(f(x)) \text{ is continuous on } U \tag{3.9.8}$$

for each $l = 1, \ldots, m$, because $\partial_l g$ is continuous on V, and f is continuous on U. This also uses the fact that sums of products of continuous real-valued functions on U are continuous as well.

3.9.1 Additional regularity of compositions

Suppose that g is twice continuously differentiable on V, so that the partial derivatives of g are continuously differentiable on V. If f is continuously differentiable on U, then it follows that

$$(\partial_l g)(f(x)) \text{ is continuously differentiable on } U \tag{3.9.9}$$

for each $l = 1, \ldots, m$, as in the preceding paragraph. If f is twice continuously differentiable on U as well, then the partial derivatives of f are continuously differentiable on U. One can use this to get that

(3.9.10) the partial derivatives of $g \circ f$ are continuously differentiable on U,

because sums of products of continuously differentiable functions on U are continuously differentiable on U too. This means that $g \circ f$ is twice continuously differentiable on U.

Similarly, if f and g are r-times continuously differentiable on U and V, respectively, for some positive integer r, then

(3.9.11) $\qquad g \circ f$ is r-times continuously differentiable on U.

This is the same as saying that

(3.9.12) $\qquad$ the partial derivatives of $g \circ f$ are $(r-1)$-times continuously differentiable on U.

The partial derivatives of f and g are $(r-1)$-times continuously differentiable on U and V, respectively, by hypothesis. One can use induction to get that

(3.9.13) $\quad (\partial_l g)(f(x))$ is $(r-1)$-times continuously differentiable on U

for each $l = 1, \ldots, m$, because f is $(r-1)$-times continuously differentiable on U. This implies (3.9.12), because sums of products of $(r-1)$-times continuously differentiable functions on U are $(r-1)$-times continuously differentiable on U as well.

If f and g are infinitely differentiable on U and V, respectively, then we get that

(3.9.14) $\qquad g \circ f$ is infinitely differentiable on U.

This was mentioned in Subsection 3.7.1 when $m = p = 1$.

3.10 A basis for $\mathcal{AM}_k(\mathbf{R}^n)$

Let k and n be positive integers, with $k \le n$, and let $\theta_1, \ldots, \theta_n$ be the standard coordinate functions on $\mathbf{R}^n$, as in Subsection 1.6.2. Remember that the dimension of the space $\mathcal{AM}_k(\mathbf{R}^n)$ of alternating k-linear forms on $\mathbf{R}^n$ is equal to $\binom{n}{k}$, as in Subsection 1.12.1. We would like to describe a particular basis for $\mathcal{AM}_k(\mathbf{R}^n)$, as a vector space over the real numbers.

Let I be a subset of $\{1, \ldots, n\}$ with exactly k elements. Thus

$$I = \{j_1, \ldots, j_k\} \tag{3.10.1}$$

for some integers $j_1, \ldots, j_k$ with

$$1 \le j_1 < \cdots < j_k \le n. \tag{3.10.2}$$

Put

$$\theta^I = \theta_{j_1} \wedge \cdots \wedge \theta_{j_k}, \tag{3.10.3}$$

where the right side is as in Section 2.11. This means that

$$\theta^I = \sum_{\sigma \in \mathrm{Sym}(k)} \mathrm{sgn}(\sigma)\, (\theta_{j_1} \otimes \cdots \otimes \theta_{j_k})^\sigma, \tag{3.10.4}$$

as before. If $x(1), \ldots, x(k)$ are k elements of $\mathbf{R}^n$, then we get that

$$\begin{aligned}
\theta^I(x(1), \ldots, x(k)) &= \sum_{\sigma \in \mathrm{Sym}(k)} \mathrm{sgn}(\sigma)\, (\theta_{j_1} \otimes \cdots \otimes \theta_{j_k})^\sigma (x(1), \ldots, x(k)) \\
&= \sum_{\sigma \in \mathrm{Sym}(k)} \mathrm{sgn}(\sigma)\, (\theta_{j_1} \otimes \cdots \otimes \theta_{j_k})(x(\sigma(1)), \ldots, x(\sigma(k))) \qquad (3.10.5) \\
&= \sum_{\sigma \in \mathrm{Sym}(k)} \mathrm{sgn}(\sigma) \prod_{l=1}^{k} \theta_{j_l}(x(\sigma(l))) \\
&= \sum_{\sigma \in \mathrm{Sym}(k)} \mathrm{sgn}(\sigma) \prod_{l=1}^{k} x_{j_l}(\sigma(l)),
\end{aligned}$$

where the second step is as in Section 1.9, and the third step is as in Section 1.6.

Let $e_1, \ldots, e_n$ be the usual standard basis vectors for $\mathbf{R}^n$, and let $m_1, \ldots, m_k$ be k positive integers less than or equal to n. Note that

$$\theta^I(e_{m_1}, \ldots, e_{m_k}) = \sum_{\sigma \in \mathrm{Sym}(k)} \mathrm{sgn}(\sigma) \prod_{l=1}^{k} \theta_{j_l}(e_{m_{\sigma(l)}}), \tag{3.10.6}$$

as in (3.10.5). This implies that

$$\theta^I(e_{m_1}, \ldots, e_{m_k}) = 0 \tag{3.10.7}$$

unless there is a $\tau \in \mathrm{Sym}(k)$ such that

$$m_{\tau(l)} = j_l \text{ for each } l = 1, \ldots, k. \tag{3.10.8}$$

In particular, this means that $m_1, \ldots, m_k$ are distinct elements of $\{1, \ldots, n\}$, because $j_1, \ldots, j_k$ are distinct elements of $\{1, \ldots, n\}$. In this case, there is exactly one $\tau \in \mathrm{Sym}(k)$ with this property, and we get that

$$\theta^I(e_{m_1}, \ldots, e_{m_k}) = \mathrm{sgn}(\tau). \tag{3.10.9}$$

Of course, there is a $\tau \in \mathrm{Sym}(k)$ such that (3.10.8) holds if and only if

$$\{m_1, \ldots, m_k\} = \{j_1, \ldots, j_k\} = I. \tag{3.10.10}$$

We also have that $m_1, \ldots, m_k$ are k distinct positive integers less than or equal to n if and only if

$$\{m_1, \ldots, m_k\} \tag{3.10.11}$$

is a subset of $\{1, \ldots, n\}$ with exactly k elements.

Remember that a k-linear form μ on $\mathbf{R}^n$ is uniquely determined by the real numbers

$$\mu(e_{m_1}, \ldots, e_{m_k}), \tag{3.10.12}$$

$1 \le m_1, \ldots, m_k \le n$, and that every family of n^k real numbers corresponds to a k-linear form on $\mathbf{R}^n$ in this way, as in Subsection 1.6.1. If μ is an alternating k-linear form on $\mathbf{R}^n$, then (3.10.12) is equal to 0 unless the m_l's are distinct, as in Section 1.11. We also have that (3.10.12) determines

$$\mu(e_{m_{\sigma(1)}}, \ldots, e_{m_{\sigma(k)}}) \tag{3.10.13}$$

for every $\sigma \in \mathrm{Sym}(k)$ in this case, as before. This implies that an alternating k-linear form μ on $\mathbf{R}$ is uniqely determined by the family of real numbers (3.10.12) with

$$1 \le m_1 < \cdots < m_k \le n, \tag{3.10.14}$$

as in Section 1.12. We have seen that any family of real numbers of this type corresponds to an alternating k-linear form on $\mathbf{R}^n$ too.

One can use this to get that

$$\text{the collection of } \theta^I, \text{ with } I \subseteq \{1, \ldots, n\} \text{ and } \#I = k, \text{ is a basis for } \mathcal{AM}_k(\mathbf{R}^n). \tag{3.10.15}$$

More precisely, one can use the θ^I's to get that that there is an alternating k-linear form μ on $\mathbf{R}^n$ for which (3.10.12) is any real number when (3.10.14) holds.

Remember that we take

$$\mathcal{AM}_0(\mathbf{R}^n) = \mathbf{R}, \tag{3.10.16}$$

as in Section 2.12. Let us put

$$\theta^I = 1 \text{ when } I = \emptyset, \tag{3.10.17}$$

so that (3.10.15) holds when $k = 0$.

3.11 A basis for $\mathcal{AM}(\mathbf{R}^n)$

Let n be a positive integer, and let us contunue with the notation used in the previous section. As in Subsection 2.12.1, we put

$$\mathcal{AM}(\mathbf{R}^n) = \bigoplus_{k=0}^{n} \mathcal{AM}_k(\mathbf{R}^n), \tag{3.11.1}$$

and we shall normally identify elements of $\mathcal{AM}_l(\mathbf{R}^n)$ with their images in $\mathcal{AM}(\mathbf{R}^n)$ under the obvious embeddings for each $l = 0, \ldots, n$. Using this identification, we get that

$$\text{the collection of } \theta^I,\ I \subseteq \{1, \ldots, n\}, \text{ is a basis for } \mathcal{AM}(\mathbf{R}^n), \tag{3.11.2}$$

as a vector space over the real numbers.

Let k be a positive integer, and let $\lambda_1, \ldots, \lambda_k$ be k linear functionals on $\mathbf{R}^n$. If $\tau \in \mathrm{Sym}(k)$ and $x(1), \ldots, x(k) \in \mathbf{R}^n$, then

$$\begin{aligned}(\lambda_1 \otimes \cdots \otimes \lambda_k)^\tau(x(1), \ldots, x(k)) &= (\lambda_1 \otimes \cdots \otimes \lambda_k)(x(\tau(1)), \ldots, x(\tau(k))) \\ &= \prod_{l=1}^k \lambda_l(x(\tau(l))),\end{aligned} \tag{3.11.3}$$

using the notation in Section 1.9 in the first step, and the notation in Section 1.6 in the second step. It follows that

$$\begin{aligned}(\lambda_1 \otimes \cdots \otimes \lambda_k)^\tau(x(1), \ldots, x(k)) &= \prod_{p=1}^k \lambda_{\tau^{-1}(p)}(x(p)) \\ &= (\lambda_{\tau^{-1}(1)} \otimes \cdots \otimes \lambda_{\tau^{-1}(k)})(x(1), \ldots, x(k)).\end{aligned} \tag{3.11.4}$$

This means that

$$(\lambda_1 \otimes \cdots \otimes \lambda_k)^\tau = \lambda_{\tau^{-1}(1)} \otimes \cdots \otimes \lambda_{\tau^{-1}(k)}. \tag{3.11.5}$$

We can replace τ with τ^{-1} to get that

$$\lambda_{\tau(1)} \otimes \cdots \otimes \lambda_{\tau(k)} = (\lambda_1 \otimes \cdots \otimes \lambda_k)^{\tau^{-1}}. \tag{3.11.6}$$

This implies that

$$\begin{aligned}\lambda_{\tau(1)} \wedge \cdots \wedge \lambda_{\tau(k)} &= k!\, A_k(\lambda_{\tau(1)} \otimes \cdots \otimes \lambda_{\tau(k)}) \\ &= k!\, A_k((\lambda_1 \otimes \cdots \otimes \lambda_k)^{\tau^{-1}}),\end{aligned} \tag{3.11.7}$$

where the first step is as in Section 2.11. Thus

$$\begin{aligned}\lambda_{\tau(1)} \wedge \cdots \wedge \lambda_{\tau(k)} &= \mathrm{sgn}(\tau^{-1})\, k!\, A_k(\lambda_1 \otimes \cdots \otimes \lambda_k) \\ &= \mathrm{sgn}(\tau)\, \lambda_1 \wedge \cdots \wedge \lambda_k,\end{aligned} \tag{3.11.8}$$

where the first step is as in Section 1.13. The second step uses the fact that $\mathrm{sgn}(\tau^{-1}) = (\mathrm{sgn}(\tau))^{-1} = \mathrm{sgn}(\tau)$. In particular, if λ_1, λ_2 are linear functionals on $\mathbf{R}^n$, then

$$\lambda_2 \wedge \lambda_1 = -\lambda_1 \wedge \lambda_2, \tag{3.11.9}$$

which also follows from a remark in Section 2.12. Using this, we get that

$$\lambda \wedge \lambda = 0 \tag{3.11.10}$$

for every linear functional λ on $\mathbf{R}^n$.

Let I_1, I_2 be subsets of $\{1, \ldots, n\}$ with k_1, k_2 elements, respectively, and put $k = k_1 + k_2$. If

$$I_1 \cap I_2 \neq \emptyset, \tag{3.11.11}$$

then it is easy to see that

$$\theta^{I_1} \wedge \theta^{I_2} = 0, \tag{3.11.12}$$

using the remarks in the preceding paragraph. Note that (3.11.11) holds automatically when $k > n$. Of course, we also know that $\mathcal{AM}_k(\mathbf{R}^n) = \{0\}$ in this case, as in Subsection 1.11.1.

Suppose now that

$$I_1 \cap I_2 = \emptyset, \tag{3.11.13}$$

and put

$$I = I_1 \cup I_2. \tag{3.11.14}$$

Let $1 \le j_1 < \cdots < j_k \le n$ be a list of the elements of I, in order. There is a unique $\tau \in \mathrm{Sym}(k)$ such that

$$\tau(\{1, \ldots, k_1\}) = I_1, \ \tau(\{k_1 + 1, \ldots, k\}) = I_2, \tag{3.11.15}$$

and τ is increasing on each of $\{1, \ldots, k_1\}$ and $\{k_1 + 1, \ldots, k\}$. Note that τ is a shuffle, as in Section 2.15, with $r = 2$. We also have that

$$\theta^{I_1} \wedge \theta^{I_2} = \theta_{j_{\tau(1)}} \wedge \cdots \wedge \theta_{j_{\tau(k)}} = \mathrm{sgn}(\tau)\, \theta_{j_1} \wedge \cdots \wedge \theta_{j_k} = \mathrm{sgn}(\tau)\, \theta^I, \tag{3.11.16}$$

using (3.11.8) in the second step.

3.12 Differential forms

Let n be a positive integer, and let E be a nonempty subset of $\mathbf{R}^n$. A *differential form* on E is

$$\text{a function on } E \text{ with values in } \mathcal{AM}(\mathbf{R}^n). \tag{3.12.1}$$

Similarly, if k is a nonnegative integer, then a *differential k-form* on E is

$$\text{a function on } E \text{ with values in } \mathcal{AM}_k(\mathbf{R}^n). \tag{3.12.2}$$

This is automatically equal to 0 when $k > n$. A differential form on E may be considered as a sum of differential k-forms on E, $0 \le k \le n$. Note that a differential 0-form on E is the same as a real-valued function on E.

A differential k-form on E may be expressed as

$$\alpha = \sum_{\#I = k} \alpha_I\, \theta^I, \tag{3.12.3}$$

where the sum is taken over all subsets I of $\{1, \ldots, n\}$ with exactly k elements. This uses the fact that the θ^I's, $\#I = k$, form a basis for $\mathcal{AM}_k(\mathbf{R}^n)$, as in Section 3.10. The α_I's, $\#I = k$, are real-valued functions on E.

Similarly, a differential form on E may be expressed as

$$\beta = \sum_I \beta_I\, \theta^I, \tag{3.12.4}$$

where now the sum is taken over all subsets I of $\{1, \ldots, n\}$. The β_I's are real-valued functions on E, as before.

The space of differential forms on E is a vector space over the real numbers with respect to pointwise addition and scalar multiplication. The space of differential k-forms on E may be considered as a linear subspace of the space of differential forms on E. If α_1, α_2 are differential forms on E, then their wedge product

$$\alpha_1 \wedge \alpha_2 \tag{3.12.5}$$

is defined as a differential form on E as well, using the wedge product of alternating multilinear forms on $\mathbf{R}^n$ at each point. If α_1, α_2 are differential k_1, k_2-forms on E for some nonnegative integers k_1, k_2, respectively, then

$$\alpha_1 \wedge \alpha_2 \text{ is a differential } (k_1 + k_2)\text{-form on } E. \tag{3.12.6}$$

Of course, this is equal to 0 when

$$k_1 + k_2 > n. \tag{3.12.7}$$

3.12.1 Continuous differential forms

It is sometimes convenient to identify $\mathcal{AM}_k(\mathbf{R}^n)$ with $\mathbf{R}^{\binom{n}{k}}$ when $k \le n$, using the basis for $\mathcal{AM}_k(\mathbf{R}^n)$ consisting of the θ^I's with $\#I = k$. To do this, we should list the subsets of $\{1, \ldots, n\}$ with exactly k elements in a sequence, with $\binom{n}{k}$ terms. The order in which these subsets of $\{1, \ldots, n\}$ are listed will normally not really matter. In particular, this leads to a metric on $\mathcal{AM}_k(\mathbf{R}^n)$, corresponding to the standard Euclidean metric on $\mathbf{R}^{\binom{n}{k}}$. This metric does not depend on the order in which the subsets of $\{1, \ldots, n\}$ with k elements are listed, because the standard Euclidean metric on $\mathbf{R}^m$ is invariant under permutations of the coordinates for any positive integer m.

Similarly, we may identify $\mathcal{AM}(\mathbf{R}^n)$ with $\mathbf{R}^{2^n}$, using the basis for $\mathcal{AM}(\mathbf{R}^n)$ consisting of the θ^I's corresponding to all subsets I of $\{1, \ldots, n\}$. This involves listing the subsets of $\{1, \ldots, n\}$ in a sequence with 2^n terms, and the order in which the subsets of $\{1, \ldots, n\}$ are listed will not normally matter, as before. This leads to a metric on $\mathcal{AM}(\mathbf{R}^n)$, corresponding to the standard Euclidean metric on $\mathbf{R}^{2^n}$, that does not depend on the order in which the subsets of $\{1, \ldots, n\}$ are listed, as before.

Let us say that a differential form on E is *continuous* if it corresponds to a continuous mapping from E into $\mathbf{R}^{2^n}$ in this way. This does not depend on the order in which the subsets of $\{1, \ldots, n\}$ are listed, as usual. Equivalently, a differential form β on E as in (3.12.4) is continuous if and only if

$$\text{the coefficients } \beta_I \text{ are continuous as real-valued functions on } E \tag{3.12.8}$$

for all subsets I of $\{1, \ldots, n\}$. The space

$$C(E, \mathcal{AM}(\mathbf{R}^n)) = C\big(E, \mathbf{R}^{2^n}\big) \tag{3.12.9}$$

of all continuous differential forms on E is a linear subspace of the space of all differential forms on E.

Similarly, we say that a differential k-form on E is continuous if it corresponds to a continuous mapping from E into $\mathbf{R}^{\binom{n}{k}}$ when $k \le n$, which does not depend on the order in which the subsets of $\{1, \ldots, n\}$ with k elements are listed. This is the same as saying that a differential k-form α on E as in (3.12.3) is continuous if and only if

(3.12.10) the coefficients α_I are continuous as real-valued functions on E

for all subsets I of $\{1, \ldots, n\}$ with exactly k elements. The space

$$C(E, \mathcal{AM}_k(\mathbf{R}^n)) = C\big(E, \mathbf{R}^{\binom{n}{k}}\big) \tag{3.12.11}$$

of all continuous differential k-forms on E is a linear subspace of the space of all differential k-forms on E. This may be considered as a linear subspace of the space of all continuous differential forms on E.

If α_1, α_2 are continuous differential forms on E, then it is easy to see that

(3.12.12) $\alpha_1 \wedge \alpha_2$ is continuous as a differential form on E.

More precisely, if $\alpha_1 \wedge \alpha_2$ is expressed as in (3.12.4), then the corresponding coefficients of the θ^I's may be expressed in terms of sums of products of the analogous coefficients of α_1 and α_2.

3.13 Additional regularity of differential forms

Let n and r be positive integers, and let U be a nonempty open subset of $\mathbf{R}^n$. Let us say that a differential form on U is *r-times continuously differentiable* on U if it corresponds to an r-times continuously-differentiable mapping from U into $\mathbf{R}^{2^n}$, as in Subsection 3.12.1. This does not depend on the order in which the subsets of $\{1, \ldots, n\}$ are listed, as before. A differential form β on U as in (3.12.4) is r-times continuously differentiable if and only if

(3.13.1) the coefficients β_I are r-times continuously differentiable as real-valued functions on U

for all subsets I of $\{1, \ldots, n\}$. The space

$$C^r(U, \mathcal{AM}(\mathbf{R}^n)) = C^r\big(U, \mathbf{R}^{2^n}\big) \tag{3.13.2}$$

of all r-times continuously differentiable differential forms on U is a linear subspace of the space of all continuous differential forms on U.

Similarly, if k is a nonnegative integer less than or equal to n, then we say that a differential k-form on U is r-times continuously differentiable if it corresponds to an r-times continuously-differentiable mapping from U into $\mathbf{R}^{\binom{n}{k}}$. This does not depend on the order in which the subsets of $\{1, \ldots, n\}$ with k

elements are listed, as usual. A differential k-form α on U as in (3.12.3) is r-times continuously differentiable on U if and only if

$$\text{the coefficients } \alpha_I \text{ are } r\text{-times continuously differentiable as real-valued functions on } U \tag{3.13.3}$$

for all subsets I of $\{1, \ldots, n\}$ with exactly k elements. The space

$$C^r(U, \mathcal{AM}_k(\mathbf{R}^n)) = C^r\big(U, \mathbf{R}^{\binom{n}{k}}\big) \tag{3.13.4}$$

of all r-times continuously-differentiable differential k-forms on U is a linear subspace of the space of all continuous differential k-forms on U. This may be considered as a linear subspace of the space of all r-times continuously-differentiable differential forms on U.

Let us say that a differential form on U is *infinitely differentiable* or *smooth* if it corresponds to an infinitely differentiable mapping from U into $\mathbf{R}^{2^n}$, which does not depend on the order in which the subsets of $\{1, \ldots, n\}$ are listed, as usual. A differential form β on U as in (3.12.4) is infinitely differentiable on U if and only if

$$\text{the coefficients } \beta_I \text{ are infinitely differentiable as real-valued functions on } U \tag{3.13.5}$$

for all subsets I of $\{1, \ldots, n\}$. The space

$$C^\infty(U, \mathcal{AM}(\mathbf{R}^n)) = C^\infty\big(U, \mathbf{R}^{2^n}\big) \tag{3.13.6}$$

of all infinitely-differentiable differential forms on U is a linear subspace of (3.13.2) for each r.

Similarly, a diiferential k-form on U is said to be infinitely differentiable or smooth if it corresponds to an infinitely-differentiable mapping from U into $\mathbf{R}^{\binom{n}{k}}$, which does not depend on the order in which the subsets of $\{1, \ldots, n\}$ with k elements are listed. A differential k-form α on U as in (3.12.3) is infinitely differentiable on U if and only if

$$\text{the coefficients } \alpha_I \text{ are infinitely differentiable as real-valued functions on } U \tag{3.13.7}$$

for all subsets I of $\{1, \ldots, n\}$ with exactly k elements. The space

$$C^\infty(U, \mathcal{AM}_k(\mathbf{R}^n)) = C^\infty\big(U, \mathbf{R}^{\binom{n}{k}}\big) \tag{3.13.8}$$

of all infinite-differentiable differential k-forms on U is a linear subspace of (3.13.4) for each r. This may be considered as a linear subspace of the space of all infinitely-differentiable differential forms on U.

As before, it is sometimes convenient to interpret r-times continuous differentiability of a differential form as being the same as continuity when $r = 0$. If

α_1, α_2 are r-times continuously-differentiable differential forms on U, then one can check that

$$\alpha_1 \wedge \alpha_2 \text{ is } r\text{-times continuously differentiable} \tag{3.13.9}$$

as a differential form on U. If α_1, α_2 are infinitely differentiable on U, then it follows that

$$\alpha_1 \wedge \alpha_2 \text{ is infinitely differentiable on } U \tag{3.13.10}$$

as well.

3.14 Differentials as 1-forms

Let n be a positive integer, let U be a nonempty open subset of $\mathbf{R}^n$, and let f be a real-valued function on U. Let us say that f is *differentiable* on U if

$$f \text{ is differentiable at every } x \in U, \tag{3.14.1}$$

in the sense of Section 3.8. In this case, the differential df of f defines a differential 1-form on U, because it defines a linear functional df_x on $\mathbf{R}^n$ at every $x \in U$.

It is customary to consider x_j as the jth coordinate function on $\mathbf{R}^n$ for each $j = 1, \ldots, n$, which we have also denoted θ_j. The differential dx_j of x_j is the same as θ_j at every point in $\mathbf{R}^n$. Thus one often uses dx_j for the differential 1-form on $\mathbf{R}^n$ that is equal to θ_j at every point.

If f is differentiable on U, then the differential of f may be expressed as

$$df = \sum_{j=1}^n (\partial_j f)\, dx_j = \sum_{j=1}^n \frac{\partial f}{\partial x_j}\, dx_j, \tag{3.14.2}$$

using this notation. This means that

$$df_x = \sum_{j=1}^n (\partial_j f)(x)\, dx_j = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x)\, dx_j \tag{3.14.3}$$

for each $x \in U$. If $v \in \mathbf{R}^n$, then we get that

$$df_x(v) = \sum_{j=1}^n (\partial_j f)(x)\, v_j = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x)\, v_j. \tag{3.14.4}$$

If f is r-times continuously differentiable on U for some positive integer r, then

$$df \text{ is } (r-1)\text{-times continuously differentiable} \tag{3.14.5}$$

as a differential 1-form on U. If f is infinitely differentiable on U, then df is infinitely differentiable as a differential 1-form on U.

Let $I = \{j_1, \ldots, j_k\}$ be a subset of $\{1, \ldots, n\}$ with k elements, where the j_l's are strictly increasing. We may use the notation

$$dx^I = dx_{j_1} \wedge \cdots \wedge dx_{j_k}, \tag{3.14.6}$$

which is the differential k-form equal to θ^I at every point.

Thus a differential k-form on U may be expressed as

$$\alpha = \sum_{\#I=k} \alpha_I \, dx^I, \tag{3.14.7}$$

where the sum is taken over all subsets I of $\{1, \ldots, n\}$ with exactly k elements, and α_I is a real-valued function on U for each such I. Similarly, a differential form on U may be expressed as

$$\beta = \sum_I \beta_I \, dx^I, \tag{3.14.8}$$

where the sum is taken over all subsets I of $\{1, \ldots, n\}$, and β_I is a real-valued function on U for each such I.

3.15 Isometric mappings

Although the topics in this section may not be needed for the moment, they do fit nicely with the discussion of metrics, norms, and inner products at the beginning of the chapter. Of course, we are also concerned with various properties of mappings more broadly, and some related matters will be mentioned later.

Let (X, d_X) and (Y, d_Y) be metric spaces. A mapping f from X into Y is said to be an *isometry* if

$$d_Y(f(x), f(w)) = d_X(x, w) \tag{3.15.1}$$

for every $x, w \in X$. Note that this implies that f is one-to-one on X. If f also maps X onto Y, then it is easy to see that the inverse mapping f^{-1} is an isometry from Y onto X.

Let (Z, d_Z) be another metric space. If f is an isometry from X into Y and g is an isometry from Y into Z, then it is easy to see that

$$g \circ f \text{ is an isometry from } X \text{ into } Z. \tag{3.15.2}$$

Let V, W be vector spaces over the real numbers with norms N_V, N_W, respectively. A linear mapping T from V into W is said to be an *isometric linear mapping* if

$$N_W(T(v)) = N_V(v) \tag{3.15.3}$$

for every $v \in V$. Of course, this implies that

$$N_W(T(u) - T(v)) = N_W(T(u - v)) = N_V(u - v) \tag{3.15.4}$$

for all $u, v \in V$, so that T is an isometry from V into W with respect to the metrics associated to N_V, N_W, respectively.

If T is an isometric linear mapping from V onto W, then the inverse mapping T^{-1} is an isometric linear mapping from W onto V. If T is a one-to-one linear mapping from V into W, and V, W have the same finite dimension, then it is well known that T maps V onto W.

Let Z be another vector space over the real numbers with a norm N_Z. If T is an isometric linear mapping from V into W, and R is an isometric linear mapping from W into Z, then

(3.15.5) $R \circ T$ is an isometric linear mapping from V into Z.

Let $\langle \cdot, \cdot \rangle_V$, $\langle \cdot, \cdot \rangle_W$ be inner products on V, W, respectively. Observe that a linear mapping T from V into W is an isometric linear mapping with respect to the norms associated to these inner products if and only if

$$\langle T(v), T(v) \rangle_W = \langle v, v \rangle_V \tag{3.15.6}$$

for every $v \in V$. In this case, one can check that

$$\langle T(u), T(v) \rangle_W = \langle u, v \rangle_V \tag{3.15.7}$$

for every $u, v \in V$.

Some connections between isometries and volumes of sets will be mentioned in Section 4.12. We shall say a bit more about isometries between finite-dimensional inner product spaces in Subsection A.9.1.

Chapter 4

Pull-backs and exterior differentiation

4.1 Tensors of type $(0, k)$

Let k and n be positive integers, and let E be a nonempty subset of $\mathbf{R}^n$. A *tensor field of type* $(0, k)$ on E is

(4.1.1) $$\text{a function on } E \text{ with values in } \mathcal{M}_k(\mathbf{R}^n).$$

If $k = 1$, then this is the same as a differential 1-form on E. If $k = 0$, then this is interpreted as being a real-valued function on E, as before. One may also consider *tensor fields of type* (r, s) for arbitrary nonnegative integers r, s, as in Definition 2.15 on p63 of [184], but we shall not pursue this here.

If $j_1, \ldots, j_k$ are positive integers less than or equal to k, then

(4.1.2) $$\theta_{j_1} \otimes \cdots \otimes \theta_{j_k}$$

is an element of $\mathcal{M}_k(\mathbf{R}^n)$, as in Section 1.6. The collection of these elements of $\mathcal{M}_k(\mathbf{R}^n)$ forms a basis for $\mathcal{M}_k(\mathbf{R}^n)$, as a vector space over the real numbers, as before. A tensor field of type $(0, k)$ on E may be expressed as

(4.1.3) $$a = \sum_{\{1,\ldots,n\}^k} a_{j_1,\ldots,j_k}\, \theta_{j_1} \otimes \cdots \otimes \theta_{j_k},$$

where the sum is taken over the set $\{1, \ldots, n\}^k$ of all k-tuples of positive integers less than or equal to n. Here the coefficients $a_{j_1,\ldots,j_k}$ are real-valued functions on E for all such k-tuples. The space of tensor fields of type $(0, k)$ on E is a vector space over the real numbers with respect to pointwise addition and scalar multiplication.

4.1.1 Continuous tensor fields

It is sometimes convenient to identiy $\mathcal{M}_k(\mathbf{R}^n)$ with $\mathbf{R}^{n^k}$, using the basis mentioned in the preceding paragraph. This involves listing the elements of

$$\{1, \ldots, n\}^k \tag{4.1.4}$$

in a sequence with n^k terms, and the order in which the elements of (4.1.4) are listed will not normally matter, as usual. This leads to a metric on $\mathcal{M}_k(\mathbf{R}^n)$, corresponding to the standard Euclidean metric on $\mathbf{R}^{n^k}$. This metric on $\mathcal{M}_k(\mathbf{R}^n)$ does not depend on the order in which the elements of (4.1.4) are listed.

Let us say that a tensor field of type $(0, k)$ on E is *continuous* if it corresponds to a continuous mapping from E into $\mathbf{R}^{n^k}$ in this way. This does not depend on the order in which the elements of (4.1.4) are listed. Equivalently, a tensor field a of type $(0, k)$ on E as in (4.1.3) is continuous if and only if

(4.1.5) the coefficients $a_{j_1,\ldots,j_k}$ are continuous as real-valued functions on E

for all $j_1, \ldots, j_k$. The space

$$C(E, \mathcal{M}_k(\mathbf{R}^n)) = C\big(E, \mathbf{R}^{n^k}\big) \tag{4.1.6}$$

of all continuous tensor fields of type $(0, k)$ on E is a linear subspace of the space of all tensor fields of type $(0, k)$ on E.

If a, b are tensor fields of types $(0, k_1)$, $(0, k_2)$ on E, respectively, for some nonnegative integers k_1, k_2, then their product

$$a \otimes b \tag{4.1.7}$$

may be defined as a (k_1+k_2)-linear form on $\mathbf{R}^n$ at each point in E, as in Section 2.8. This defines (4.1.7) as a tensor field of type $(0, k_1 + k_2)$ on E. If a, b are continuous as tensor fields on E, then one can check that

(4.1.8) $a \otimes b$ is continuous as a tensor field on E.

More precisely, the coefficients of (4.1.7) with respect to the usual basis for $\mathcal{M}_{k_1+k_2}(\mathbf{R}^n)$ may be expressed in terms of products of the coefficients of a and b with respect to the usual bases for $\mathcal{M}_{k_1}(\mathbf{R}^n)$ and $\mathcal{M}_{k_2}(\mathbf{R}^n)$, respectively.

4.2 More on tensor fields

Let n be a positive integer, let U be a nonempty open subset of $\mathbf{R}^n$, and let k, r be nonnegative integers. A tensor field of type $(0, k)$ on U is said to be *r-times continuously differentiable* on U if it corresponds to an r-times continuously-differentiable mapping from U into $\mathbf{R}^{n^k}$ as in Subsection 4.1.1. This does not depend on the order in which the elements of $\{1, \ldots, n\}^k$ are listed, as usual. Remember that r-times continuous differentiability is interpreted as being the same as continuity when $r = 0$.

A tensor field a of type $(0, k)$ on U as in (4.1.3) is r-times continuously differentiable if and only if

$$\text{the coefficients } a_{j_1,\ldots,j_k} \text{ are } r\text{-times continuously differentiable as real-valued functions on } U \tag{4.2.1}$$

for all $j_1, \ldots, j_k$. The space

$$C^r(U, \mathcal{M}_k(\mathbf{R}^n)) = C^r\big(U, \mathbf{R}^{n^k}\big) \tag{4.2.2}$$

of all r-times continuously-differentiable tensor fields of type $(0, k)$ on U is a linear subspace of the space of all continuous tensor fields of type $(0, k)$ on U.

A tensor field of type $(0, k)$ on U is said to be *infinitely differentiable* or *smooth* if it corresponds to an infinitely-differentiable mapping from U into $\mathbf{R}^{n^k}$, which does not depend on the order in which the elements of $\{1, \ldots, n\}^k$ are listed. A tensor field a of type $(0, k)$ on U as in (4.1.3) is infinitely differentiable on U if and only if

$$\text{the coefficients } a_{j_1,\ldots,j_k} \text{ are infinitely differentiable as real-valued functions on } U \tag{4.2.3}$$

for all $j_1, \ldots, j_k$. The space

$$C^\infty(U, \mathcal{M}_k(\mathbf{R}^n)) = C^\infty\big(U, \mathbf{R}^{n^k}\big) \tag{4.2.4}$$

of all infinitely-differentiable tensor fields of type $(0, k)$ on U is a linear subspace of (4.2.2) for each r.

Let a, b be tensor fields of types $(0, k_1)$, $(0, k_2)$ on U, respectively, for some nonnegative integers k_1, k_2, so that their product $a \otimes b$ is defined as a tensor field of type $(0, k_1 + k_2)$ on U, as in Subsection 4.1.1. If a, b are r-times continuously differentiable on U, then one can verify that

$$a \otimes b \text{ is } r\text{-times continuously differentiable} \tag{4.2.5}$$

as a tensor field on U. If a, b are infinitely differentiable on U, then we get that

$$a \otimes b \text{ is infinitely differentiable} \tag{4.2.6}$$

as a tensor field on U.

If a real-valued function f is differentiable on U, then the differential df of f defines a tensor of type $(0, 1)$ on U, as in Section 3.14. Note that

$$dx_{j_1} \otimes \cdots \otimes dx_{j_k} \tag{4.2.7}$$

is the same as the tensor field of type $(0, k)$ equal to (4.1.2) at every point. Thus a tensor field a of type $(0, k)$ on U may be expressed as

$$a = \sum_{\{1,\ldots,n\}^k} a_{j_1,\ldots,j_k}\, dx_{j_1} \otimes \cdots \otimes dx_{j_k}, \tag{4.2.8}$$

as in (4.1.3).

4.3 Vector fields

Let n be a positive integer, and let E be a nonempty subset of $\mathbf{R}^n$. An $\mathbf{R}^n$-valued function on E may be called a *vector field* on E. This may be considered as a *tensor field of type* $(1,0)$ on E, as in Definition 2.15 on p63 of [184].

Let k be a positive integer, and let a be a tensor field of type $(0,k)$ on E. Suppose that a is as in (4.1.3), so that the value a_x of a at $x \in E$ is given by

$$a_x = \sum_{\{1,\ldots,n\}^k} a_{j_1,\ldots,j_k}(x)\, \theta_{j_1} \otimes \cdots \otimes \theta_{j_k}, \tag{4.3.1}$$

as a k-linear form on $\mathbf{R}^n$. Let $\xi_1, \ldots, \xi_k$ be k vector fields on E, and let $\xi_{l,j}$ be the jth component of ξ_l for $j = 1, \ldots, n$ and $l = 1, \ldots, k$. We can evaluate a at $\xi_1, \ldots, \xi_k$ at each point in E to get a real-valued function

$$a(\xi_1, \ldots, \xi_k) \tag{4.3.2}$$

on E. More precisely, the value of this function at $x \in E$ is

$$a_x(\xi_1(x), \ldots, \xi_k(x)) = \sum_{\{1,\ldots,n\}^k} a_{j_1,\ldots,j_k}(x) \prod_{l=1}^{k} \xi_{l,j_l}(x). \tag{4.3.3}$$

Suppose for the moment that a is continuous as a tensor field of type $(0,k)$ on E, as in Subsection 4.1.1. Suppose also that $\xi_1, \ldots, \xi_k$ are continuous vector fields on E, which is to say that they are continuous as $\mathbf{R}^n$-valued functions on E. Under these conditions,

(4.3.4) $a(\xi_1, \ldots, \xi_k)$ is continuous as a real-valued function on E.

4.3.1 Associated differential operators

Now let U be a nonempty open subset of $\mathbf{R}^n$. If ξ is a vector field on U, then

$$\mathcal{X}_\xi = \sum_{j=1}^{n} \xi_j \frac{\partial}{\partial x_j} \tag{4.3.5}$$

defines a first-order linear differential operator on U, where ξ_j is the jth component of ξ for each $j = 1, \ldots, n$. If f is a differentiable real-valued function on U, then

$$\mathcal{X}_\xi(f) = \sum_{j=1}^{n} \xi_j \frac{\partial f}{\partial x_j} \tag{4.3.6}$$

is a real-valued function on U, whose value

$$(\mathcal{X}_\xi(f))(x) = \sum_{j=1}^{n} \xi_j(x) \frac{\partial f}{\partial x_j}(x) \tag{4.3.7}$$

at $x \in U$ is the same as the directional derivative of f at x in the direction $\xi(x)$. Equivalently,

$$df(\xi) = \mathcal{X}_\xi(f), \tag{4.3.8}$$

where df is considered as a tensor field of type $(0,1)$ on U. This means that

$$df_x(\xi(x)) = (\mathcal{X}_\xi(f))(x) \tag{4.3.9}$$

for every $x \in U$.

If f is r-times continuously differentiable on U for some positive integer r, and ξ is $(r-1)$-times continuously differentiable on U, then

(4.3.10) $\mathcal{X}_\xi(f)$ is $(r-1)$-times continuously differentiable on U.

If ξ and f are infinitely differentiable on U, then

(4.3.11) $\mathcal{X}_\xi(f)$ is infinitely differentiable on U.

Let a be a tensor field of type $(0,k)$ on U, and let $\xi_1, \ldots, \xi_k$ be k vector fields on U. If a and $\xi_1, \ldots, \xi_k$ are r-times continuously differentiable on U, then

(4.3.12) $a(\xi_1, \ldots, \xi_k)$ is r-times continuously differentiable on U.

If a and $\xi_1, \ldots, \xi_k$ are infinitely differentiable on U, then

(4.3.13) $a(\xi_1, \ldots, \xi_k)$ is infinitely differentiable on U.

4.4 Pulling some tensors back

Let n be a positive integer, let U be a nonempty open subset of $\mathbf{R}^n$, and let k be a nonnegative integer. Also let n_1 be a positive integer, let U_1 be a nonempty open subset of $\mathbf{R}^{n_1}$, and let ψ be a mapping from U_1 into $\mathbf{R}^n$ such that

$$\psi(U_1) \subseteq U. \tag{4.4.1}$$

Suppose that ψ is *differentiable* on U_1, in the sense that ψ is differentiable at every point in U_1, as in Section 3.8. If a is a tensor field of type $(0,k)$ on U, then we would like to *pull a back* to get a tensor field

$$\psi^*(a) \tag{4.4.2}$$

of type $(0,k)$ on U_1. This may also be denoted $\delta\psi(a)$, as in Definition 2.22 on p68 of [184].

If $k = 0$, then a is simply a real-valued function on U. In this case, we put

$$\psi^*(a) = a \circ \psi, \tag{4.4.3}$$

which is a real-valued function on U_1.

Suppose that $k \geq 1$, and let a_x be the value of a at $x \in U$, which is a k-linear form on $\mathbf{R}^n$. Let $w \in U_1$ be given, so that the differential $d\psi_w$ of ψ at w is a linear mapping from $\mathbf{R}^{n_1}$ into $\mathbf{R}^n$. The value of $\psi^*(a)$ at w is defined to be the k-linear form on $\mathbf{R}^{n_1}$ given by

$$(\psi^*(a))_w = (d\psi_w)^*(a_{\psi(w)}), \tag{4.4.4}$$

where the right side is defined as in Section 2.3. Note that $\psi(w) \in U$ by hypothesis, so that $a_{\psi(w)}$ is defined as a k-linear form on $\mathbf{R}^n$. If $v_1, \ldots, v_k \in \mathbf{R}^{n_1}$, then

$$\begin{aligned} (\psi^*(a))_w(v_1, \ldots, v_k) &= ((d\psi_w)^*(a_{\psi(w)}))(v_1, \ldots, v_k) \\ &= a_{\psi(w)}(d\psi_w(v_1), \ldots, d\psi_w(v_k)), \end{aligned} \tag{4.4.5}$$

as in Section 2.3.

4.4.1 Some properties of pull-backs

Note that $\psi^*(a)$ is linear in a, as in Section 2.3. If a is a differential k-form on U, then a_x is an alternating k-linear form on $\mathbf{R}^n$ for each $x \in U$, as in Section 3.12. This implies that (4.4.4) is an alternating k-linear form on $\mathbf{R}^{n_1}$ for every $w \in U_1$, as in Section 2.4. This means that

$$\psi^*(a) \text{ is a differential } k\text{-form on } U_1 \tag{4.4.6}$$

in this case. Similarly, if a_x is a symmetric k-linear form on $\mathbf{R}^n$ for every $x \in U$, then

$$(\psi^*(a))_w \text{ is a symmetric } k\text{-linear form on } \mathbf{R}^{n_1} \tag{4.4.7}$$

for every $w \in U_1$.

Suppose that a, b are tensor fields of types $(0, k_1)$, $(0, k_2)$ on U for some nonnegative integers k_1, k_2, respectively. Observe that

$$\psi^*(a \otimes b) = \psi^*(a) \otimes \psi^*(b) \tag{4.4.8}$$

as tensor fields of type $(0, k_1 + k_2)$ on U_1. This follows from the analogous statement for multilinear forms in Subsection 2.8.1. Similarly, if a, b are differential k_1, k_2-forms on U, then

$$\psi^*(a \wedge b) = \psi^*(a) \wedge \psi^*(b), \tag{4.4.9}$$

as differential $(k_1 + k_2)$-forms on U_1. This follows frm the analogous statement for alternating multilinear forms in Subsection 2.12.1.

4.5 More on pull-backs

Let us continue with the same hypotheses and notation as in the previous section. Remember that dx_j is a differential 1-form on $\mathbf{R}^n$ for each $j = 1, \ldots, n$, whose value at each point is the jth standard coordinate function θ_j, as in Section 3.14. Equivalently, this may be considered as a tensor field of type $(0, 1)$ on $\mathbf{R}^n$, as in Sections 4.1 and 4.2.

Let ψ_j be the jth component of ψ for each $j = 1, \ldots, n$, which is a differentiable real-valued function on U_1. Of course, the differential of ψ_j corresponds to the jth component of the differential of ψ in a simple way for each $j = 1, \ldots, n$. One can use this to get that

$$\psi^*(dx_j) = d\psi_j \tag{4.5.1}$$

on U_1 for each $j = 1, \ldots, n$. More precisely, the right side is a differential 1-form on U_1, as in Section 3.14, and thus a tensor field of type $(0, 1)$. To get (4.5.1), one can take $a = dx_j$ in (4.4.5), so that $k = 1$.

Let a be a tensor field of type $(0, k)$ on U, expressed as in (4.2.8). Observe that

$$\begin{aligned}\psi^*(a) &= \sum_{\{1,\ldots,n\}^k} \psi^*(a_{j_1,\ldots,j_k})\, \psi^*(dx_{j_1}) \otimes \cdots \otimes \psi^*(dx_{j_k}) \\ &= \sum_{\{1,\ldots,n\}^k} (a_{j_1,\ldots,j_k} \circ \psi)\, d\psi_{j_1} \otimes \cdots \otimes d\psi_{j_k}\end{aligned} \tag{4.5.2}$$

on U_1.

Similarly, a differential k-form α on U may be expressed as

$$\alpha = \sum_{1 \le j_1 < \cdots < j_k \le n} \alpha_{j_1,\ldots,j_k}\, dx_{j_1} \wedge \cdots \wedge dx_{j_k}, \tag{4.5.3}$$

as in Section 3.14. More precisely, the sum on the right is taken over all k-tuples $(j_1, \ldots, j_k)$ of integers such that $1 \le j_1 < \cdots < j_k \le n$. In this case, we get that

$$\begin{aligned}\psi^*(\alpha) &= \sum_{1 \le j_1 < \cdots < j_k \le n} \psi^*(\alpha_{j_1,\ldots,j_k})\, \psi^*(dx_{j_1}) \wedge \cdots \wedge \psi^*(dx_{j_k}) \\ &= \sum_{1 \le j_1 < \cdots < j_k \le n} (\alpha_{j_1,\ldots,j_k} \circ \psi)\, d\psi_{j_1} \wedge \cdots \wedge d\psi_{j_k}\end{aligned} \tag{4.5.4}$$

on U_1.

4.5.1 Pull-backs and composiitons

Let n_0 be a positive integer, let U_0 be a nonempty open subset of $\mathbf{R}^{n_0}$, and let ϕ be a mapping from U_0 into $\mathbf{R}^{n_1}$ such that

$$\phi(U_0) \subseteq U_1. \tag{4.5.5}$$

Suppose that ϕ is differentiable on U_0 too. Thus $\psi \circ \phi$ is a mapping from U_0 into $\mathbf{R}^n$ with

$$(\psi \circ \phi)(U_0) = \psi(\phi(U_0)) \subseteq \psi(U_1) \subseteq U. \tag{4.5.6}$$

Remember that

$$\psi \circ \phi \text{ is differentiable on } U_0, \tag{4.5.7}$$

as in Section 3.9.

If a is a tensor field of type $(0,k)$ on U, then one can check that

$$(\psi \circ \phi)^*(a) = \phi^*(\psi^*(a)) \tag{4.5.8}$$

on U_0. This is very simple when $k = 0$. If $k \geq 1$, then this follows from the chain rule, and the analogous statement for multilinear forms on Section 2.3.

4.5.2 Regularity of pull-backs

Let a be a tensor field of type $(0,k)$ on U again, expressed as in (4.2.8). If a is continuous on U, and ψ is continuously differentiable on U_1, then

$$\psi^*(a) \text{ is continuous on } U_1. \tag{4.5.9}$$

More precisely, if $k = 0$, then this holds when ψ is continuous on U_1. If $k \geq 1$, then this can be obtained using (4.5.2).

Similarly, if a is r-times continuously differentiable on U for some positive integer r, and ψ is $(r+1)$-times continuously differentiable on U_1, then

$$\psi^*(a) \text{ is } r\text{-times continuously differentiable on } U_1. \tag{4.5.10}$$

If $k = 0$, then it is enough to ask that ψ be r-times continuously differentiable on U_1, as in Section 3.9. If a is infinitely differentiable on U, and ψ is infinitely differentiable on U_1, then it follows that

$$\psi^*(a) \text{ is infinitely differentiable on } U_1. \tag{4.5.11}$$

4.6 Exterior differentiation

Let n be a positive integer, and let U be a nonempty open subset of $\mathbf{R}^n$. If α is a continuously-differentiable differential form on U, then we would like to define the *exterior derivative*

$$d\alpha \tag{4.6.1}$$

of α as a continuous differential form on U. More precisely, if α is a continuously-differentiable differential k-form on U for some nonnegative integer k, then

$$d\alpha \text{ is a continuous differential } (k+1)\text{-form on } U. \tag{4.6.2}$$

In particular,

$$d\alpha = 0 \text{ when } k \geq n. \tag{4.6.3}$$

If α is a differential 0-form on U, then α is a real-valued function on U. If α is continuously differentiable on U, then we take

$$d\alpha \text{ to be the differential of } \alpha, \tag{4.6.4}$$

considered as a differential 1-form on U, as in Section 3.14.

Suppose that α is a differential k-form on U, with $1 \leq k \leq n$, which may be expressed as

$$\alpha = \sum_{\#I=k} \alpha_I \, dx^I, \tag{4.6.5}$$

as in Section 3.14. If α is continuously differentiable on U, then

$$\alpha_I \text{ is continuously differentiable} \tag{4.6.6}$$

as a real-valued function on U for each subset I of $\{1, \ldots, n\}$ with exactly k elements, as in Section 3.13. In this case, we put

$$d\alpha = \sum_{\#I=k} d\alpha_I \wedge dx^I. \tag{4.6.7}$$

One can check that this defines a continuous differential $(k+1)$-form on U. Note that this is equal to 0 when $k = n$, as before.

If α is any continuously-differentiable differential form on U, then α may be expressed as a sum of continuously-differentiable differential k-forms on U, $0 \leq k \leq n$. Under these conditions, $d\alpha$ is defined to be the sum of the exterior derivatives of these differential k-forms. This defines the exterior derivative as a linear mapping from $C^1(U, \mathcal{AM}(\mathbf{R}^n))$ into $C(U, \mathcal{AM}(\mathbf{R}^n))$.

If α is r-times continuously differentiable on U for some positive integer r, then one can verify that

$$d\alpha \text{ is } (r-1)\text{-times continuously differentiable on } U, \tag{4.6.8}$$

so that

$$d(C^r(U, \mathcal{AM}(\mathbf{R}^n))) \subseteq C^{r-1}(U, \mathcal{AM}(\mathbf{R}^n)). \tag{4.6.9}$$

If α is infinitely differentiable on U, then it follows that

$$d\alpha \text{ is infinitely differentiable on } U, \tag{4.6.10}$$

so that

$$d(C^\infty(U, \mathcal{AM}(\mathbf{R}^n))) \subseteq C^\infty(U, \mathcal{AM}(\mathbf{R}^n)). \tag{4.6.11}$$

4.7 The anti-derivation property

Let us continue with the same notation and hypotheses as in the previous section. Let α be a continuously-differentiable differential k-form on U for some $0 \leq k \leq n$, and let β be another continuously-differentiable differential form on U. We would like to show that

$$d(\alpha \wedge \beta) = (d\alpha) \wedge \beta + (-1)^k \, \alpha \wedge (d\beta) \tag{4.7.1}$$

on U. This is the *anti-derivation* property of the exterior derivative, as in Section 2.11 on p61 of [184], and Theorem 2.20 on p65 of [184]. Note that it suffices to verify this when β is a differential l-form on U for some $0 \leq l \leq n$.

Suppose for the moment that $k = 0$. If β is also a differential 0-form on U, then (4.7.1) is the same as saying that

$$d(\alpha\,\beta) = (d\alpha)\,\beta + \alpha\,(d\beta) \tag{4.7.2}$$

on U. Of course, this holds by the usual product rule for partial derivatives. If β is a differential l-form on U for some $l \geq 1$, then (4.7.1) is the same as saying that

$$d(\alpha\,\beta) = (d\alpha) \wedge \beta + \alpha\,(d\beta). \tag{4.7.3}$$

This can be verified directly from the definition of the exterior derivative using the product rule for partial derivatives again, or by reducing to the $l = 0$ case.

Thus we suppose from now on in this section that $k \geq 1$. If β is a differential 0-form on U, then (4.7.1) is the same as saying that

$$d(\alpha\,\beta) = (d\alpha)\,\beta + (-1)^k\,\alpha \wedge (d\beta). \tag{4.7.4}$$

This can be verified directly, as in the preceding paragraph, using basic properties of the wedge product, as in Section 2.12. One could also reduce to the previous case, using the fact that

$$\alpha \wedge (d\beta) = (-1)^k\,(d\beta) \wedge \alpha, \tag{4.7.5}$$

because $d\beta$ is a differential 1-form on U.

4.7.1 The cases where $k, l \geq 1$

This permits us to take β to be a differential l-form on U with $l \geq 1$. In fact, we may as well suppose that

$$\alpha = \alpha_I\,dx^I \tag{4.7.6}$$

and

$$\beta = \beta_L\,dx^L, \tag{4.7.7}$$

where I, L are subsets of $\{1, \ldots, n\}$ with exactly $k, l \geq 1$ elements, respectively, and α_I, β_L are continuously-differentiable real-valued functions on U. If

$$I \cap L \neq \emptyset, \tag{4.7.8}$$

then one can check directly that both sides of (4.7.1) are equal to 0.

Suppose now that $I \cap L = \emptyset$, and put

$$M = I \cup L. \tag{4.7.9}$$

In this case, we have that

$$dx^I \wedge dx^L = \operatorname{sgn}(\tau)\,dx^M \tag{4.7.10}$$

for a certain $\tau \in \operatorname{Sym}(k + l)$, as in Section 3.11. Thus

$$\alpha \wedge \beta = \alpha_I\,\beta_L\,dx^I \wedge dx^L = \alpha_I\,\beta_L\,\operatorname{sgn}(\tau)\,dx^M. \tag{4.7.11}$$

This implies that

$$d(\alpha \wedge \beta) = \mathrm{sgn}(\tau)\, d(\alpha_I\, \beta_L) \wedge dx^M, \tag{4.7.12}$$

by the definition of the exterior derivative. This means that

$$d(\alpha \wedge \beta) = \mathrm{sgn}(\tau)\, \beta_L\, d\alpha_I \wedge dx^M + \mathrm{sgn}(\tau)\, \alpha_I\, d\beta_L \wedge dx^M, \tag{4.7.13}$$

as in (4.7.2).

It follows that

$$d(\alpha \wedge \beta) = \beta_L\, d\alpha_I \wedge dx^I \wedge dx^L + \alpha_I\, d\beta_L \wedge dx^I \wedge dx^L, \tag{4.7.14}$$

because of (4.7.10). The second term on the right is the same as

$$(-1)^k\, \alpha_I\, dx^I \wedge d\beta_L \wedge dx^L, \tag{4.7.15}$$

because $d\beta_L$ is a differential 1-form on U, as in Section 2.12. This shows that (4.7.1) holds in this case, as desired.

4.8 The second exterior derivative

Let n be a positive integer, and let U be a nonempty open subset of $\mathbf{R}^n$. Also let

(4.8.1) α be a twice continuously-differentiable differential form on U,

so that $d\alpha$ is a continuously-differentiable differential form on U. We would like to show that

$$d(d\alpha) = 0 \tag{4.8.2}$$

on U.

If α is a differential 0-form on U, then

$$d\alpha = \sum_{l=1}^{n} (\partial_l \alpha)\, dx_l, \tag{4.8.3}$$

as in Section 3.14. This implies that

$$d(d\alpha) = \sum_{l=1}^{n} d(\partial_l \alpha) \wedge dx_l = \sum_{j=1}^{n} \sum_{l=1}^{n} (\partial_j \partial_l \alpha)\, dx_j \wedge dx_l. \tag{4.8.4}$$

Remember that

$$dx_j \wedge dx_l = -dx_l \wedge dx_j \tag{4.8.5}$$

for all $j, l = 1, \ldots, n$, as in Section 2.12. We also have that

$$\partial_j\, \partial_l\, \alpha = \partial_l\, \partial_j\, \alpha \tag{4.8.6}$$

on U for all $j, l = 1, \ldots, n$, as in Subsection 3.5.2. One can use this to get that (4.8.2) holds in this case.

Suppose now that α is as in (4.7.6) for some nonempty subset I of $\{1,\ldots,n\}$, where α_I is a twice continuously-differentiable real-valued function on U. Thus

$$d\alpha = d\alpha_I \wedge dx^I, \tag{4.8.7}$$

and

$$d(d\alpha) = d(d\alpha_I) \wedge dx^I, \tag{4.8.8}$$

as in the previous section. More precisely, this uses the fact that

$$d(dx^I) = 0, \tag{4.8.9}$$

by the definition of the exterior derivative. Of course, the right side of (4.8.8) is equal to 0, as in the preceding paragraph. It follows that (4.8.2) holds in this case, and for all twice continuously-differentiable differential forms on U.

If β is a continuously-differentiable differential form on U and

$$d\beta = 0 \tag{4.8.10}$$

on U, then β is said to be *closed* as a differential form on U. If γ is a continuously-differentiable differential form on U, then

$$d\gamma \tag{4.8.11}$$

is said to be *exact* as a differential form on U. If γ is twice continuously differentiable on U, then (4.8.11) is closed on U, as before.

4.9 Differentiating pull-backs

Let n, n_1 be positive integers, let U, U_1 be nonempty open subsets of $\mathbf{R}^n$, $\mathbf{R}^{n_1}$, respectively, and let

(4.9.1) ψ be a twice continuously-differentiable mapping from U_1 into $\mathbf{R}^n$

such that

$$\psi(U_1) \subseteq U. \tag{4.9.2}$$

If α is a continuously-differentiable differential form on U, then we would like to show that

$$d(\psi^*(\alpha)) = \psi^*(d\alpha), \tag{4.9.3}$$

as differential forms on U_1. Remember that

(4.9.4) $\psi^*(\alpha)$ is continuously differentiable on U_1

under these conditions, as in Subsection 4.5.2.

Suppose first that α is a differential 0-form on U, so that $\psi^*(\alpha) = \alpha \circ \psi$. Remember that $d\alpha$, $d(\alpha\circ\psi)$ are the same as the differentials of α, $\alpha\circ\psi$ considered as differential 1-forms on U, U_1, respectively, as in Section 3.14. If $w \in U_1$, then

$$d(\alpha \circ \psi)_w = d\alpha_{\psi(w)} \circ d\psi_w, \tag{4.9.5}$$

as linear functionals on $\mathbf{R}^{n_1}$, by the chain rule, as in Section 3.9. This means that

$$d(\alpha \circ \psi)_w = (d\psi_w)^*(d\alpha_{\psi(w)}), \tag{4.9.6}$$

where the right side is as in Section 2.3. This is the same as saying that (4.9.3) holds at w, as in Section 4.4.

More precisely, if α is a differential 0-form on U, then it suffices to ask that

(4.9.7) ψ be a continuously-differentiable mapping from U_1 into $\mathbf{R}^n$

with $\psi(U_1) \subseteq U$, instead of (4.9.1). Note that $\psi^*(\alpha) = \alpha \circ \psi$ is continuously differentiable on U_1 in this case, as in Section 3.9.

4.9.1 Some basic cases

Suppose next that

(4.9.8) f is a twice continuously-differentiable real-valued function on U,

and that

$$\alpha = df. \tag{4.9.9}$$

Thus

$$\psi^*(\alpha) = \psi^*(df) = d(f \circ \psi) \tag{4.9.10}$$

on U_1, as in the preceding paragraph. We also have that

$$d\alpha = d(df) = 0 \tag{4.9.11}$$

on U and

$$d(\psi^*(\alpha)) = d(d(f \circ \alpha)) = 0 \tag{4.9.12}$$

on U_1, as in the previous section. More precisely, this uses the fact that $f \circ \psi$ is twice continuously differentiable on U_1, because ψ and f are twice continuously differentiable by hypothesis, as in Subsection 3.9.1. It follows that both sides of (4.9.3) are equal to 0 under these conditions.

Let k be a positive integer, and suppose that

$$\alpha = dx_{j_1} \wedge \cdots \wedge dx_{j_k} \tag{4.9.13}$$

for some positive integers $j_1 < \cdots < j_k \le n$. Observe that

$$\psi^*(\alpha) = \psi^*(dx_{j_1}) \wedge \cdots \wedge \psi^*(dx_{j_k}), \tag{4.9.14}$$

as in Subsection 4.4.1. This implies that

$$\psi^*(\alpha) = d\psi_{j_1} \wedge \cdots \wedge d\psi_{j_k}, \tag{4.9.15}$$

where $\psi_1, \ldots, \psi_n$ are the components of ψ, as in the preceding paragraph. Of course,

$$d(d\psi_j) = 0 \tag{4.9.16}$$

on U_1 for each $j = 1, \ldots, n$, as in the previous section, because ψ_j is twice continuously differentiable on U_1, by hypothesis. It follows that

$$d(\psi^*(\alpha)) = 0 \tag{4.9.17}$$

on U_1, as in Section 4.7. Clearly $d\alpha = 0$ on U, by definition of the exterior derivative, in Section 4.6. This means that both sides of (4.9.3) are equal to 0 in this case too.

4.9.2 The rest of the proof

Let I be a nonempty subset of $\{1, \ldots, n\}$, and suppose that

$$\alpha = \alpha_I \, dx^I, \tag{4.9.18}$$

where α_I is a twice-continuously-differentiable real-valued function on U. Of course,

$$d\alpha = d\alpha_I \wedge dx^I, \tag{4.9.19}$$

by the definition of the exterior derivative. Observe that

$$\psi^*(\alpha) = (\alpha_I \circ \psi) \, \psi^*(dx^I). \tag{4.9.20}$$

This implies that

$$\begin{aligned} d(\psi^*(\alpha)) &= d(\alpha_I \circ \psi) \wedge \psi^*(dx^I) + (\alpha_I \circ \psi) \, d(\psi^*(dx^I)) \\ &= \psi^*(d\alpha_I) \wedge \psi^*(dx^I) = \psi^*(d\alpha), \end{aligned} \tag{4.9.21}$$

where the first step is as in Section 4.7. The second term on the right side of the first line is equal to 0, as in (4.9.17). The second step also uses the earlier remarks about differential 0-forms. The third step uses a remark in Subsection 4.4.1. This implies that (4.9.3) holds for differential k-forms on U when $k \geq 1$.

4.10 Connected sets

This section and the next deal with connectedness, path connectedness, and the relationship between them. Connectedness can be used to prove the intermediate value theorem, and it can also be helpful in calculus, as in Subsection 4.10.2. Another aspect of this will be mentioned in Section 6.14.

Let n be a positive integer, and let E be a subset of $\mathbf{R}^n$. The property of *connectedness* of E can be defined in a standard way, that we shall not discuss here. In fact, one can define connectedness for subsets of arbitrary metric spaces.

It is well known that a subset of a metric space is connected if and only if it is connected as a subset of itself, with respect to the restriction of the metric to that set. In particular,

$$E \text{ is connected as a subset of } \mathbf{R}^n, \tag{4.10.1}$$

with respect to the standard Euclidean metric, if and only if

$$E \text{ is connected as a subset of itself,} \tag{4.10.2}$$

with respect to the restriction of the standard Euclidean metric on $\mathbf{R}^n$ to E.

Suppose for the moment that $n = 1$. It is well known that E is connected if and only if for every $x, y \in E$ with $x < y$, we have that

$$(x, y) \subseteq E. \tag{4.10.3}$$

Let m be a positive integer, and let f be a continuous mapping from E into $\mathbf{R}^m$. If E is connected as a subset of $\mathbf{R}^n$, then it is well known that

$$f(E) \text{ is connected in } \mathbf{R}^m. \tag{4.10.4}$$

This uses the equivalence of (4.10.1) and (4.10.2).

4.10.1 Convex sets

We say that E is *convex* if for every $x, y \in E$ and $t \in \mathbf{R}$ with $0 \le t \le 1$, we have that

$$(1 - t)\, x + t\, y \in E. \tag{4.10.5}$$

It is well known that

$$\text{convex subsets of } \mathbf{R}^n \text{ are connected.} \tag{4.10.6}$$

This can be obtained from another result about connectedness that will be mentioned in the next section. Note that connected subsets of the real line are convex.

One can define convexity of a subset of any vector space V over the real numbers in the same way. If N is a norm on V, as in Section 3.2, then one can check that

$$\begin{aligned}&\text{open balls in } V \text{ with respect to the metric}\\ &\text{associated to } N \text{ are convex.}\end{aligned} \tag{4.10.7}$$

In particular, open balls in $\mathbf{R}^n$ with respect to the standard Euclidean metric are convex.

4.10.2 Locally constant functions

It is well known that an open set U in $\mathbf{R}^n$ is connected if and only if U cannot be expressed as the union of two nonempty open subsets of $\mathbf{R}^n$. More precisely, the analogous statement holds in any metric space.

Let f be a function defined on U with values in a set Y. Let us say that f is *locally constant* on U is for each $x \in U$ we have that

$$f(w) = f(x) \tag{4.10.8}$$

for all $w \in U$ that are sufficiently close to x with respect to the standard Euclidean metric on $\mathbf{R}^n$, depending on x. This is equivalent to saying that f is continuous as a mapping from U into Y when Y is equipped with the discrete metric, as in Section 3.1.

Alternatively, f is locally constant on U if and only if for each $y \in Y$,

$$f^{-1}(\{y\}) = \{w \in U : f(w) = y\} \tag{4.10.9}$$

is an open set in $\mathbf{R}^n$. This implies that

$$f^{-1}(Y \setminus \{z\}) = U \setminus f^{-1}(\{z\}) = \{w \in U : f(w) \neq z\} \tag{4.10.10}$$

for every $z \in Y$, because it is the union of $f^{-1}(\{y\})$ over $y \in Y \setminus \{z\}$. In this case, if U is connected, then one can check that

$$f \text{ is constant on } U. \tag{4.10.11}$$

If U is not connected, then U can be expressed as the union of two nonempty open sets, as before. One can use this to find locally constant functions on U that are not constant.

Suppose now that f is a real-valued function on U. If f is locally constant on U, then

$$\frac{\partial f}{\partial x_j} = 0 \tag{4.10.12}$$

at every point in U. Conversely, if the partial derivatives of f exist at every point in U and satisfy (4.10.12), then one can check that f is locally constant on U.

4.11 Path-connected sets

Let n be a positive integer again, and let E be a subset of $\mathbf{R}^n$. We say that E is *path connected* if every $x, y \in E$ can be connected by a continuous path in E. More precisely, this means that there are real numbers $a \leq b$ and a continuous mapping p from the closed interval $[a, b]$ in the real line into $\mathbf{R}^n$ such that $p(a) = x$, $p(b) = y$, and

$$p([a, b]) \subseteq E. \tag{4.11.1}$$

In this case, one can always take $a = 0$ and $b = 1$, although it is sometimes convenient to use different intervals. Note that this definition can also be used for subsets of arbitrary metric spaces.

It is well known that

$$\text{path-connected sets are connected,} \tag{4.11.2}$$

and indeed this works in arbitrary metric spaces. This uses the fact that the image of a continuous path is a connected set. In particular,

$$\text{convex sets in } \mathbf{R}^n \text{ are path connected,} \tag{4.11.3}$$

and thus connected, as mentioned in Subsection 4.10.1.

4.11.1 Path-connected components

If $x, y \in E$, then put

$$x \sim_E y \tag{4.11.4}$$

when x can be connected to y by a continuous path in E, as before. This holds for every $x, y \in E$ exactly when E is path connected. One can verify that

$$\sim_E \text{ defines an equivalence relation on } E. \tag{4.11.5}$$

This implies that E can be partitioned into the equivalence classes corresponding to $\sim_E$, where $x, y \in E$ are elements of the same equivalence class if and only if (4.11.4) holds.

These equivalence classes may be called the *path-connected components* of E. They are path connected sets, and maximal with respect to inclusion among path-connected subsets of E. This works in arbitrary metric spaces too.

Let U be an open subset of $\mathbf{R}^n$, and let $x \sim_U y$ be defined for $x, y \in U$ as before. If $x \in U$, then one can check that

$$x \sim_U w \text{ for all } w \in U \text{ that are sufficiently close to } x, \tag{4.11.6}$$

with respect to the standard Euclidean metric on $\mathbf{R}^n$. This implies that

$$\text{the path-connected components of } U \text{ are open sets.} \tag{4.11.7}$$

If U is connected, then one can use this to get that U has only one path-connected component, so that

$$U \text{ is path connected.} \tag{4.11.8}$$

It is not too difficult to show that the path connected components of U are maximal with respect to inclusion among connected subsets of U, so that they may also be called the connected components of U.

4.12 Volumes and determinants

These next few sections are concerned with n-dimensional volumes of subsets of $\mathbf{R}^n$ and other n-dimensional vector spaces over the real numbers. In this section, we look at the effect on volumes of mappings from $\mathbf{R}^n$ into itself.

Let n be a positive integer, and let T be a linear mapping from $\mathbf{R}^n$ into itself. If E is a reasonably nice subset of $\mathbf{R}^n$, then its n-dimensional volume $\mathrm{Vol}_n(E)$ can be defined in standard ways. In this case, it is well known that

$$\mathrm{Vol}_n(T(E)) = |\det T| \, \mathrm{Vol}_n(E). \tag{4.12.1}$$

If one is familiar with n-dimensional Lebesgue measure on $\mathbf{R}^n$, then this holds for all Lebesgue measurable subsets E of $\mathbf{R}^n$, but we shall not pursue this here.

Let N be a norm on $\mathbf{R}^n$, as in Section 3.2. Suppose for the moment that T is an isometry with respect to N, so that

$$N(T(v)) = N(v) \tag{4.12.2}$$

for every $v \in \mathbf{R}^n$, as in Section 3.15. Of course, this implies that T is one-to-one on $\mathbf{R}^n$, so that $T(\mathbf{R}^n) = \mathbf{R}^n$. One can use (4.12.1) to get that

$$|\det T| = 1, \tag{4.12.3}$$

by taking E to be the unit ball in $\mathbf{R}^n$ with respect to N. More precisely, this uses some additional well-known facts, such as that N and the standard Euclidean norm on $\mathbf{R}^n$ can each be bounded by a constant multiple of the other, and we shall not pursue this here.

Let $d_N(\cdot,\cdot)$ be the metric on $\mathbf{R}^n$ associated to N, as in Section 3.2. One can use this metric to define n-dimensional Hausdorff measure on $\mathbf{R}^n$, which is equal to a positive multiple of n-dimensional Lebesgue measure on $\mathbf{R}^n$. Hausdorff measures are automatically preserved by isometries. This gives another way to look at (4.12.3), which we shall not pursue in detail here.

4.12.1 Orthogonal transformations

Suppose now that T is an isometry with respect to the standard Euclidean norm on $\mathbf{R}^n$. Let $\langle \cdot,\cdot \rangle$ be the standard inner product on $\mathbf{R}^n$, as in Section 3.3. The condition that T be an isometry with respect to the standard Euclidean norm on $\mathbf{R}^n$ is equivalent to saying that

$$\langle T(u), T(v) \rangle = \langle u, v \rangle \tag{4.12.4}$$

for every $u, v \in \mathbf{R}^n$, as in Section 3.15. This is also equivalent to saying that T is a one-to-one linear mapping from $\mathbf{R}^n$ onto itself such that

$$\langle T(u), w \rangle = \langle u, T^{-1}(w) \rangle \tag{4.12.5}$$

for every $u, w \in \mathbf{R}^n$. Under these conditions, T is said to be an *orthogonal transformation* on $\mathbf{R}^n$.

Of course, linear mappings from $\mathbf{R}^n$ into itself correspond to $n \times n$ matrices of real numbers in a standard way. An invertible linear mapping T on $\mathbf{R}^n$ is an orthogonal transformation if and only if

(4.12.6) the matrix associated to T^{-1} is the same as the transpose of the matrix associated to T,

because this is equivalent to (4.12.5).

If T is any invertible linear mapping on $\mathbf{R}^n$, then

$$\det T^{-1} = (\det T)^{-1}. \tag{4.12.7}$$

If T is an orthogonal transformation, then

$$\det T^{-1} = \det T, \tag{4.12.8}$$

because the determinant of a matrix is the same as the determinant of its transpose. This implies that

$$(\det T)^2 = 1, \tag{4.12.9}$$

which is the same as (4.12.3).

4.13 Admissible volumes

On an arbitrary finite-dimensional vector space over the real numbers, it is not necessarily clear how to measure the volume of a set, without additional information. However, there is a nice family of ways of doing this, that are discussed in this section. One can use linear mappings to reduce to measuring volumes in Euclidean spaces, and different linear mappings can lead to different ways of measuring volumes.

Let n be a positive integer, and let W be an n-dimensional vector over the real numbers. Also let L be a one-to-one linear mapping from $\mathbf{R}^n$ onto W. If E is a reasonably nice subset of W, in the sense that $L^{-1}(E)$ is a reasonably nice subset of $\mathbf{R}^n$, then

$$\mathrm{Vol}_n(L^{-1}(E)) \tag{4.13.1}$$

may be considered as a type of n-dimensional volume of E.

Let L_1 be another one-to-one linear mapping from $\mathbf{R}^n$ onto W, so that $L_1^{-1} \circ L$ is a one-to-one linear mapping from $\mathbf{R}^n$ onto itself. If E is a reasonably nice subset of W, then

$$\begin{aligned} \mathrm{Vol}_n(L_1^{-1}(E)) &= \mathrm{Vol}_n((L_1^{-1} \circ L)(L^{-1}(E))) \\ &= |\det(L_1^{-1} \circ L)| \, \mathrm{Vol}_n(L^{-1}(E)), \end{aligned} \tag{4.13.2}$$

using (4.12.1) in the second step.

Let us say that Λ is an *admissible volume* on W if there is a nonnegative real number $c_{\Lambda,L}$ such that

$$\Lambda(E) = c_{\Lambda,L} \, \mathrm{Vol}_n(L^{-1}(E)) \tag{4.13.3}$$

for all reasonably nice subsets E of W. If

$$c_{\Lambda,L} > 0, \tag{4.13.4}$$

then we say that Λ is *nondegenerate* on W. If one uses a different one-to-one linear mapping from $\mathbf{R}^n$ onto W, then one gets the same familty of admissible volumes on W, because of (4.13.2).

4.13.1 Admissible volumes and linear mappings

Let R be a linear mapping from W into itself, so that $L^{-1} \circ R \circ L$ is a linear mapping from $\mathbf{R}^n$ into itself. Remember that

$$\det R = \det(L^{-1} \circ R \circ L), \tag{4.13.5}$$

basically by the definition of the determinant of a linear mapping from W into itself, as in Subsection 2.5.1. If E is a reasonably nice subset of W, then

$$\begin{aligned} \mathrm{Vol}_n(L^{-1}(R(E))) &= \mathrm{Vol}_n((L^{-1} \circ R \circ L)(L^{-1}(E))) \\ &= |\det(L^{-1} \circ R \circ L)| \,\mathrm{Vol}_n(L^{-1}(E)) \\ &= |\det R| \,\mathrm{Vol}_n(L^{-1}(E)), \end{aligned} \tag{4.13.6}$$

using (4.9.3) in the second step. If Λ is an admissible volume on W, then it follows that

$$\Lambda(R(E)) = |\det R| \,\Lambda(E). \tag{4.13.7}$$

4.14 Parallelepipeds

Let V be a vector space over the real numbers, let k be a positive integer, and let $v_1, \ldots, v_k$ be k elements of V. The *parallelepiped* in V associated to $v_1, \ldots, v_k$ is the subset of V defined by

$$\begin{aligned} P(v_1, \ldots, v_k) &= P_V(v_1, \ldots, v_k) \\ &= \Big\{ \sum_{l=1}^{k} t_l\, v_l : t_l \in \mathbf{R},\ 0 \le t_l \le 1 \text{ for each } l = 1, \ldots, k \Big\}. \end{aligned} \tag{4.14.1}$$

Let us say that this parallelepiped is *nondegenerate* when $v_1, \ldots, v_k$ are linearly independent in V.

Let W be another vector space over the real numbers, and let T be a linear mapping from V into W. Observe that

$$T(P_V(v_1, \ldots, v_k)) = P_W(T(v_1), \ldots, T(v_k)). \tag{4.14.2}$$

Let n be a positive integer, and let $e_1, \ldots, e_n$ be the standard basis vectors in $\mathbf{R}^n$. The corresponding parallelepiped in $\mathbf{R}^n$ is the same as the closed unit cube in $\mathbf{R}^n$,

$$P_{\mathbf{R}^n}(e_1, \ldots, e_n) = [0,1]^n. \tag{4.14.3}$$

Of course, the n-dimensional volume of this is equal to 1.

If B is any linear mapping from $\mathbf{R}^n$ into itself, then

$$P_{\mathbf{R}^n}(B(e_1), \ldots, B(e_n)) = B(P_{\mathbf{R}^n}(e_1, \ldots, e_n)) = B([0,1]^n), \tag{4.14.4}$$

where the first step is as in (4.14.2). Note that $B(e_1), \ldots, B(e_n)$ may be any n elements of $\mathbf{R}^n$. The relation between volumes and determinants mentioned in Section 4.12 says that

$$\mathrm{Vol}_n(P_{\mathbf{R}^n}(B(e_1), \ldots, B(e_n))) = |\det B|. \tag{4.14.5}$$

More precisely, this can be shown first, as in Theorem 8 on p329 of [20]. One can use this to obtain the previous relation for reasonably nice subsets of $\mathbf{R}^n$, as in Theorem 9 and its corollary on p329f of [20].

4.14.1 Parallelepipeds and admissible volumes

Suppose that W has dimension equal to n, and let $u_1, \ldots, u_n$ be a basis for W. If L is a one-to-one linear mapping from $\mathbf{R}^n$ onto W, then

$$L^{-1}(u_1), \ldots, L^{-1}(u_n) \tag{4.14.6}$$

is a basis for $\mathbf{R}^n$. In fact, we can choose L so that this is the standard basis for $\mathbf{R}^n$. In particular,

$$\mathrm{Vol}_n(L^{-1}(P_W(u_1, \ldots, u_n))) = \mathrm{Vol}_n(P_{\mathbf{R}^n}(L^{-1}(u_1), \ldots, L^{-1}(u_n))) \tag{4.14.7}$$

is strictly positive.

Let Λ be an admissible volume on W, as in (4.13.3). Thus

$$\Lambda(P_W(u_1, \ldots, u_n)) = c_{\Lambda,L}\, \mathrm{Vol}_n(P_{\mathbf{R}^n}(L^{-1}(u_1), \ldots, L^{-1}(u_n))). \tag{4.14.8}$$

by (4.14.7). We may choose $c_{\Lambda,L}$ so that this is any given nonnegative real number, because (4.14.7) is positive. More precisely, this determines $c_{\Lambda,L}$ uniquely, so that Λ is uniquely determined in this way as well.

4.14.2 Admissible volumes and alternating forms

If ν is an alternating n-linear form on W, then it follows that there is a unique admissible volume Λ_ν on W such that

$$\Lambda_\nu(P_W(u_1, \ldots, u_n)) = |\nu(u_1, \ldots, u_n)|. \tag{4.14.9}$$

If $w_1, \ldots, w_n$ are any n elements of W, then

$$\Lambda_\nu(P_W(w_1, \ldots, w_n)) = |\nu(w_1, \ldots, w_n)|. \tag{4.14.10}$$

To see this, let R be the linear mapping from W into itself such that $R(u_j) = w_j$ for each j. The left side of (4.14.10) is equal to

$$\begin{aligned} \Lambda_\nu(P_W(R(u_1), \ldots, R(u_n))) &= \Lambda_\nu(R(P_W(u_1, \ldots, u_n))) \\ &= |\det R|\, \Lambda(P_W(u_1, \ldots, u_n)), \end{aligned} \tag{4.14.11}$$

using (4.14.2) in the first step, and (4.13.7) in the second step. The right side of (4.14.10) is equal to

$$\begin{aligned} |\nu(R(u_1), \ldots, R(u_n))| &= |(R^*(\nu))(u_1, \ldots, u_n)| \\ &= |\det R|\, |\nu(u_1, \ldots, u_n)|, \end{aligned} \tag{4.14.12}$$

using the definition of $R^*(\nu)$ in Section 2.3 in the first step, and a remark from Subsection 2.5.1 in the second step.

4.15 Volumes and inner products

Let n be a positive integer, and let W be an n-dimensional vector space over the real numbers again. Also let $\langle \cdot, \cdot \rangle_W$ be an inner product on W, and let $\| \cdot \|_W$ be the associated norm on W. Similarly, let $\langle \cdot, \cdot \rangle_{\mathbf{R}^n}$ be the standard inner product on $\mathbf{R}^n$, as in Section 3.3, whose associated norm $\| \cdot \|_{\mathbf{R}^n}$ is the standard Euclidean norm on $\mathbf{R}^n$.

One can use an orthonormal basis for W to get an isometric linear mapping L from $\mathbf{R}^n$ onto W, as in Section A.8. This means that

$$\langle L(x), L(y) \rangle_W = \langle x, y \rangle_{\mathbf{R}^n} \tag{4.15.1}$$

for all $x, y \in \mathbf{R}^n$, as in Section 3.15. Using L, we get an admissible volume

$$\Lambda_W(E) = \mathrm{Vol}_n(L^{-1}(E)) \tag{4.15.2}$$

on W, as in Section 4.13.

If L_1 is another isometric linear mapping from $\mathbf{R}^n$ onto W, then $L_1^{-1} \circ L$ is an isometric linear mapping from $\mathbf{R}^n$ onto itself. Equivalently,

$$L_1^{-1} \circ L \text{ is an orthogonal transformation on } \mathbf{R}^n, \tag{4.15.3}$$

and in particular $L_1^{-1} \circ L$ preserves the n-dimensional volume of reasonably nice subsets of $\mathbf{R}^n$, as in Section 4.12. This implies that

$$\Lambda_W(E) = \mathrm{Vol}_n(L_1^{-1}(E)) \tag{4.15.4}$$

for all reasonable nice subsets E of W, as in Section 4.13.

Let $e_1, \ldots, e_n$ be the standard basis vectors in $\mathbf{R}^n$, as usual. Any sequence of n vectors in $\mathbf{R}^n$ may be expressed as

$$B(e_1), \ldots, B(e_n) \tag{4.15.5}$$

for some linear mapping B from $\mathbf{R}^n$ into itself. This is an orthonormal basis for $\mathbf{R}^n$ with respect to the standard inner product if and only if B is an orthogonal transformation on $\mathbf{R}^n$, as in Section A.8. In this case, we get that

$$\mathrm{Vol}_n(P_{\mathbf{R}^n}(B(e_1), \ldots, B(e_n))) = |\det B| = 1, \tag{4.15.6}$$

where the first step is as in (4.14.5), and the second step is as in Section 4.12.

If $w_1, \ldots, w_n$ are n elements of W, then

$$\Lambda_W(P_W(w_1, \ldots, w_n)) = \mathrm{Vol}_n(P_{\mathbf{R}^n}(L^{-1}(w_1), \ldots, L^{-1}(w_n))), \tag{4.15.7}$$

by (4.14.2) and (4.15.2). If

$$w_1, \ldots, w_n \text{ is an orthonormal basis for } W, \tag{4.15.8}$$

then

$$L^{-1}(w_1), \ldots, L^{-1}(w_n) \text{ is an orthonormal basis for } \mathbf{R}^n, \tag{4.15.9}$$

as in Section A.8. This implies that

$$\Lambda_W(P_W((w_1, \ldots, w_n))) = 1, \tag{4.15.10}$$

by (4.15.6).

Chapter 5

Diffeomorphisms and cells

5.1 Matrix-valued functions

Let m_1 and m_2 be positive integers, and let

$$M_{m_1,m_2}(\mathbf{R}) \tag{5.1.1}$$

be the space of $m_1 \times m_2$ matrices of real numbers. This is a vector space over the real numbers with respect to entrywise addition and scalar multiplication of matrices. It is sometimes convenient to identify $M_{m_1,m_2}(\mathbf{R})$ with $\mathbf{R}^{m_1\,m_2}$, by listing the entries of an $m_1 \times m_2$ matrix by a sequence of length $m_1\,m_2$.

The order in which the entries of an $m_1 \times m_2$ matrix are listed will normally not really matter, as long as we use the same listing for all such matrices. This leads to a metric on $M_{m_1,m_2}(\mathbf{R})$, corresponding to the standard Euclidean metric on $\mathbf{R}^{m_1\,m_2}$. In particular, this metric does not depend on the order in which the entries of an $m_1 \times m_2$ matrix are listed, as long as we use the same listing for all such matrices, because the standard Euclidean metric on $\mathbf{R}^k$ is invariant under permutations of the coordinates for any positive integer k.

Let n be a positive integer, and let E be a nonempty subset of $\mathbf{R}^n$. Also let f be a function on E with values in $M_{m_1,m_2}(\mathbf{R})$. Let us say that f is *continuous* on E if it corresponds to a continuous mapping from E into $\mathbf{R}^{m_1\,m_2}$ as in the previous two paragraphs. This does not depend on the order in which the entries of an $m_1 \times m_2$ matrix are listed, as long as we use the same listing for all such matrices, as before. Equivalently, f is continuous on E if and only if the $m_1\,m_2$ entries of f are continuous as real-valued functions on E, as in Section 3.4.

Of course, the space of all $M_{m_1,m_2}(\mathbf{R})$-valued functions on E is a vector space over the real numbers, with respect to pointwise addition and scalar multiplication of these functions. The space

$$C(E, M_{m_1,m_2}(\mathbf{R})) = C(E, \mathbf{R}^{m_1\,m_2}) \tag{5.1.2}$$

of all continuous $M_{m_1,m_2}(\mathbf{R})$-valued functions on E is a linear subspace of the space of all $M_{m_1,m_2}(\mathbf{R})$-valued functions on E.

5.1.1 Products of matrix-valued functions

If m_3 is another positive integer, then the product of an $m_1 \times m_2$ matrix of real numbers with an $m_2 \times m_3$ matrix of real numbers may be defined as an $m_1 \times m_3$ matrix of real numbers in the usual way. This defines a bilinear mapping from $M_{m_1,m_2}(\mathbf{R}) \times M_{m_2,m_3}(\mathbf{R})$ into $M_{m_1,m_3}(\mathbf{R})$.

If f and g are functions on E with values in $M_{m_1,m_2}(\mathbf{R})$ and $M_{m_2,m_3}(\mathbf{R})$, respectively, then their product $f\,g$ may be defined as a function on E with values in $M_{m_1,m_3}(\mathbf{R})$ using matrix multiplication at each point. If f, g are continuous on E, then one check that

$$f\,g \text{ is continuous on } E. \tag{5.1.3}$$

This is because the entries of $f\,g$ are given by sums of products of the entries of f and g.

5.1.2 C^r Matrix-valued functions

Let U be a nonempty open subset of $\mathbf{R}^n$, and let r be a positive integer. We say that a function f on U with values in $M_{m_1,m_2}(\mathbf{R})$ is *r-times continuously differentiable* on U if it corresponds to an r-times continuously-differentiable mapping from U into $\mathbf{R}^{m_1\,m_2}$ as before. This does not depend on the order in which the entries of an $m_1 \times m_2$ matrix are listed, as long as we use the same listing for all such matrices, as usual. Equivalently, this means that the $m_1\,m_2$ entries of f are r-times continuously differentiable as real-valued functions on U. If $r = 0$, then we may interpret this as meaning that f is continuous on U, as before.

Let us say that f is *infinitely differentiable* or *smooth* on U if f corresponds to an infinitely differentiable mapping from U into $\mathbf{R}^{m_1\,m_2}$. This does not depend on the order in which the entries of an $m_1 \times m_2$ matrix are listed, as long as we use the same listing for all such matrices, and indeed the smoothness of f is equivalent to the smoothness of the $m_1\,m_2$ entries of f as real-valued functions on U.

The space

$$C^r(U, M_{m_1,m_2}(\mathbf{R})) = C^r(U, \mathbf{R}^{m_1\,m_2}) \tag{5.1.4}$$

of r-times continuously-differentiable functions on U with values in $M_{m_1,m_2}(\mathbf{R})$ is a linear subspace of the space of continuous functions on U with values in $M_{m_1,m_2}(\mathbf{R})$ for each r. Similarly, the space

$$C^\infty(U, M_{m_1,m_2}(\mathbf{R})) = C^\infty(U, \mathbf{R}^{m_1\,m_2}) \tag{5.1.5}$$

of infinitely-differentiable functions on U with values in $M_{m_1,m_2}(\mathbf{R})$ is a linear subspace of (5.1.4) for each r.

Let f and g be functions on U with values in $M_{m_1,m_2}(\mathbf{R})$ and $M_{m_2,m_3}(\mathbf{R})$, so that $f\,g$ is defined as a function on U with values in $M_{m_1,m_3}(\mathbf{R})$, as before. If f, g are r-times continuously differentiable on U for some r, then one can verify that

$$f\,g \text{ is } r\text{-times continuously differentiable on } U. \tag{5.1.6}$$

If f, g are infinitely differentiable on U, then it follows that

(5.1.7) $$f\,g \text{ is infinitely differentiable on } U.$$

5.2 General linear groups

Let m be a positive integer, and consider the space $M_{m,m}(\mathbf{R})$ of $m \times m$ matrices of real numbers, as in the previous section. As before, it is sometimes convenient to identify $M_{m,m}(\mathbf{R})$ with $\mathbf{R}^{m^2}$, by listing the entries os an $m \times m$ matrix by a sequence of length m^2. The order in which the entries are listed will normally not really matter, as long as we use the same listing for all of these matrices, as usual.

Note that $M_{m,m}(\mathbf{R})$ is an associative algebra over the real numbers with respect to matrix multiplication. The *identity matrix* in $M_{m,m}(\mathbf{R})$ has diagonal entries equal to 1, and all other entries equal to 0. This is the multiplicative identity element in $M_{m,m}(\mathbf{R})$.

An element of $M_{m,m}(\mathbf{R})$ is said to be *invertible* if it has a multiplicative inverse in $M_{m,m}(\mathbf{R})$. Let

(5.2.1) $$GL(m, \mathbf{R})$$

be the set of invertible elements of $M_{m,m}(\mathbf{R})$. This is another example of a *group*, with respect to matrix multiplication in this case. This is known as the *general linear group* of $m \times m$ matrices with entries in $\mathbf{R}$.

5.2.1 $GL(m, \mathbf{R})$ And the determinant

Of course, the determinant defines a real-valued function on $M_{m,m}(\mathbf{R})$. More precisely, the determinant corresponds to a homogeneous polynomial on $\mathbf{R}^{m^2}$ of degree m. In particular, this defines a continuous function on $\mathbf{R}^{m^2}$.

It is well known that an element of $M_{m,m}(\mathbf{R})$ is invertible if and only if its determinant is not zero, because of Cramer's rule. One can use this and the continuity of the determinant to get that

(5.2.2) $$GL(m, \mathbf{R}) \text{ is an open set in } M_{m,m}(\mathbf{R}).$$

This means that $GL(m, \mathbf{R})$ corresponds to an open set in $\mathbf{R}^{m^2}$, with respect to the standard Euclidean metric.

Equivalently, the complement of $GL(m, \mathbf{R})$ in $M_{m,m}(\mathbf{R})$ consists of matrices whose determinant is equal to 0. This corresponds to a closed set in $\mathbf{R}^{m^2}$, because the determinant is continuous.

Cramer's rule implies that the inverse of an element of $GL(m, \mathbf{R})$ can be expressed in terms of the determinant of the matrix, and the determinants of its minors. In particular, the entries of the inverse are rational functions of the entries of the matrix, wth the determinant of the matrix as the denominator. This implies that

(5.2.3) the mapping from $GL(m, \mathbf{R})$ into itself that sends an element of $GL(m, \mathbf{R})$ to its multiplicative inverse in $M_{m,m}(\mathbf{R})$ is smooth.

Another way to look at this mapping will be mentioned in Section 5.4.

5.3 Some spaces of linear mappings

Let m_1 and m_2 be positive integers, and let

$$\mathcal{L}(\mathbf{R}^{m_2}, \mathbf{R}^{m_1}) \tag{5.3.1}$$

be the space of linear mappings from $\mathbf{R}^{m_2}$ into $\mathbf{R}^{m_1}$. This is a linear subspace of the space of all functions on $\mathbf{R}^{m_2}$ with values in $\mathbf{R}^{m_1}$, as a vector space over the real numbers with respect to pointwise addition and scalar multiplication. It is sometimes convenient to identify $\mathcal{L}(\mathbf{R}^{m_2}, \mathbf{R}^{m_1})$ with $M_{m_1,m_2}(\mathbf{R})$, using the usual correspondence between linear mappings from $\mathbf{R}^{m_2}$ into $\mathbf{R}^{m_1}$ with $m_1 \times m_2$ matrices of real numbers. Thus we may also identify $\mathcal{L}(\mathbf{R}^{m_2}, \mathbf{R}^{m_1})$ with $\mathbf{R}^{m_1 \, m_2}$, as in Section 5.1.

Let n be a positive integer, let E be a nonempty subset of $\mathbf{R}^n$, and let f be a function on E with values in $\mathcal{L}(\mathbf{R}^{m_2}, \mathbf{R}^{m_1})$. Let us say that f is *continuous* if f corresponds to a continuous function on E with values in $M_{m_1,m_2}(\mathbf{R})$, or equivalently $\mathbf{R}^{m_1,m_2}$, as in Section 5.1. Note that the space of functions on E with values in $\mathcal{L}(\mathbf{R}^{m_2}, \mathbf{R}^{m_1})$ is a vector space over the real numbers, with respect to pointwise addition and scalar multiplication of these functions. The space

$$C(E, \mathcal{L}(\mathbf{R}^{m_2}, \mathbf{R}^{m_1})) = C(E, M_{m_1,m_2}(\mathbf{R})) = C(E, \mathbf{R}^{m_1 \, m_2}) \tag{5.3.2}$$

of all continuous functions on E with values in $\mathcal{L}(\mathbf{R}^{m_2}, \mathbf{R}^{m_1})$ is a linear subspace of the space of all functions on E with values in $\mathcal{L}(\mathbf{R}^{m_2}, \mathbf{R}^{m_1})$.

Let m_3 be another positive integer, so that we may also consider linear mappings from $\mathbf{R}^{m_3}$ into $\mathbf{R}^{m_1}$ or $\mathbf{R}^{m_2}$. Of course, if T_2 is a linear mapping from $\mathbf{R}^{m_3}$ into $\mathbf{R}^{m_2}$, and T_1 is a linear mapping from $\mathbf{R}^{m_2}$ into $\mathbf{R}^{m_1}$, then their composition $T_1 \circ T_2$ is a linear mapping from $\mathbf{R}^{m_3}$ into $\mathbf{R}^{m_1}$. It is easy to see that

$$(T_1, T_2) \mapsto T_1 \circ T_2 \tag{5.3.3}$$

defines a bilinear mapping from

$$\mathcal{L}(\mathbf{R}^{m_2}, \mathbf{R}^{m_1}) \times \mathcal{L}(\mathbf{R}^{m_3}, \mathbf{R}^{m_2}) \tag{5.3.4}$$

into $\mathcal{L}(\mathbf{R}^{m_3}, \mathbf{R}^{m_1})$. This corresponds to multiplication of the matrices associated to T_1 and T_2, as usual.

Let f and g be functions on E with values in $\mathcal{L}(\mathbf{R}^{m_2}, \mathbf{R}^{m_1})$, $\mathcal{L}(\mathbf{R}^{m_3}, \mathbf{R}^{m_2})$, respectively. Their product $f\,g$ may be defined as a function on E with values in $\mathcal{L}(\mathbf{R}^{m_3}, \mathbf{R}^{m_1})$, using composition of linear mappings as multiplication at each point. This means that the value of $f\,g$ at $x \in E$ is equal to

$$(f\,g)(x) = f(x) \circ g(x), \tag{5.3.5}$$

where the right side is the linear mapping from $\mathbf{R}^{m_3}$ into $\mathbf{R}^{m_1}$ obtained by composing $g(x)$ with $f(x)$. If f and g are continuus on E, then

$$f\,g \text{ is continuous on } E, \tag{5.3.6}$$

as in Subsection 5.1.1.

5.3.1 C^r $\mathcal{L}(\mathbf{R}^{m_2}, \mathbf{R}^{m_1})$-Valued functions

Let U be a nonempty open subset of $\mathbf{R}^n$, let r be a positive integer, and let f be a function on U with values in $\mathcal{L}(\mathbf{R}^{m_2}, \mathbf{R}^{m_1})$. We say that f is *r-times continuously differentiable* on U if it corresponds to an r-times continuously differentiable function on U with values in $M_{m_1,m_2}(\mathbf{R})$, or equivalently in $\mathbf{R}^{m_1\,m_2}$, as in Subsection 5.1.2. If $r = 0$, then we may interpret this as meaning that f is continuous on U, as usual. Similarly, we say that f is *infinitely differentiabile* if f corresponds to an infinitely differentiable function on U with values in $M_{m_1,m_2}(\mathbf{R})$, or equivalently in $\mathbf{R}^{m_1\,m_2}$, as before.

The space

$$C^r(U, \mathcal{L}(\mathbf{R}^{m_2}, \mathbf{R}^{m_1})) = C^r(U, M_{m_1,m_2}(\mathbf{R})) = C^r(U, \mathbf{R}^{m_1\,m_2}) \tag{5.3.7}$$

of r-times continuously-differentiable $\mathcal{L}(\mathbf{R}^{m_2}, \mathbf{R}^{m_1})$-valued functions on U is a linear subspace of the space of continuous functions on U with values in $\mathcal{L}(\mathbf{R}^{m_2}, \mathbf{R}^{m_1})$ for each r. Similarly, the space

$$C^\infty(U, \mathcal{L}(\mathbf{R}^{m_2}, \mathbf{R}^{m_1})) = C^\infty(U, M_{m_1,m_2}(\mathbf{R})) = C^\infty(U, \mathbf{R}^{m_1\,m_2}) \tag{5.3.8}$$

of infinitely-differentiable functions on U with values in $\mathcal{L}(\mathbf{R}^{m_2}, \mathbf{R}^{m_1})$ is a linear subspace of (5.3.7) for each r.

Let g be a function on U with values in $\mathcal{L}(\mathbf{R}^{m_3}, \mathbf{R}^{m_2})$, so that $f\,g$ is defined as a function on U with values in $\mathcal{L}(\mathbf{R}^{m_3}, \mathbf{R}^{m_1})$, as in (5.3.5). If f and g are r-times continuously differentiable on U, then

$$f\,g \text{ is } r\text{-times continuously differentiable on } U, \tag{5.3.9}$$

as in Subsection 5.1.2. If f and g are infinitely differentiable on U, then we get that

$$f\,g \text{ is infinitely differentiable on } U. \tag{5.3.10}$$

5.4 More on linear mappings

Let m_1 and m_2 be positive integers, and let T be a linear mapping from $\mathbf{R}^{m_2}$ into $\mathbf{R}^{m_1}$. It is well known that there is a nonnegative real number $C(T)$, depending on T, such that

$$\|T(v)\|_{2,\mathbf{R}^{m_1}} \le C(T)\,\|v\|_{2,\mathbf{R}^{m_2}} \tag{5.4.1}$$

for every $v \in \mathbf{R}^{m_2}$. Here $\|\cdot\|_{2,\mathbf{R}^m}$ is the standard Euclidean norm on $\mathbf{R}^m$ for each m, as in Section 3.2, although one could consider other norms as well. In fact, one can take $C(T)$ to be the standard Euclidean norm of the $m_1 \times m_2$ matrix of real numbers associated to T, considered as an element of $\mathbf{R}^{m_1\,m_2}$, as on p211 of [155].

Now let m be a positive integer, and let T be a one-to-one linear mapping from $\mathbf{R}^m$ onto itself. There is a positive real number $C(T^{-1})$ such that

$$\|T^{-1}(u)\|_{2,\mathbf{R}^m} \le C(T^{-1})\, \|u\|_{2,\mathbf{R}^m} \tag{5.4.2}$$

for every $u \in \mathbf{R}^m$, as before. This is the same as saying that

$$C(T^{-1})^{-1}\, \|v\|_{2,\mathbf{R}^m} \le \|T(v)\|_{2,\mathbf{R}^m} \tag{5.4.3}$$

for every $v \in \mathbf{R}^m$, by taking $v = T^{-1}(u)$.

Let R be another linear mapping from $\mathbf{R}^m$ into itself, and suppose that R is fairly close to T, in the sense that

$$\|R(v) - T(v)\|_{2,\mathbf{R}^m} \le (1/2)\, C(T^{-1})^{-1}\, \|v\|_{2,\mathbf{R}^m} \tag{5.4.4}$$

for every $v \in \mathbf{R}^m$. If $v \in \mathbf{R}^m$, then we get that

$$\begin{aligned} C(T^{-1})^{-1}\|v\|_{2,\mathbf{R}^m} \le \|T(v)\|_{2,\mathbf{R}^m} &\le \|R(v)\|_{2,\mathbf{R}^m} + \|R(v) - T(v)\|_{2,\mathbf{R}^m} \\ &\le \|R(v)\|_{2,\mathbf{R}^m} + (1/2)\, C(T^{-1})^{-1}\, \|v\|_{2,\mathbf{R}^m}. \end{aligned} \tag{5.4.5}$$

This implies that

$$(1/2)\, C(T^{-1})^{-1}\, \|v\|_{2,\mathbf{R}^m} \le \|R(v)\|_{2,\mathbf{R}^m}. \tag{5.4.6}$$

In particular, (5.4.6) implies that the kernel of R is trivial. This means that R is a one-to-one mapping from $\mathbf{R}^m$ onto itself, by well-known results in linear algebra. We also get that

$$\|R^{-1}(u)\|_{2,\mathbf{R}^m} \le 2\, C(T^{-1})\, \|u\|_{2,\mathbf{R}^m} \tag{5.4.7}$$

for every $u \in \mathbf{R}^m$, by taking $v = R^{-1}(u)$.

5.4.1 Linear mappings on $\mathbf{R}^m$

Let

$$\mathcal{L}(\mathbf{R}^m) = \mathcal{L}(\mathbf{R}^m, \mathbf{R}^m) \tag{5.4.8}$$

be the space of linear mappings from $\mathbf{R}^m$ into itself. This is an associative algebra over the real numbers with respect to composition of mappings. The identity mapping on $\mathbf{R}^m$ is the multiplicative identity element in this algebra. It is sometimes convenient to identify this algebra with $M_{m,m}(\mathbf{R})$, as in the previous section.

Let

$$GL(\mathbf{R}^m) \tag{5.4.9}$$

be the set of one-to-one linear mappings from $\mathbf{R}^m$ onto itself, which are the same as the invertible elements of $\mathcal{L}(\mathbf{R}^m)$. This is a *group*, with respect to composition of mappings, which is known as the *general linear group* of $\mathbf{R}^m$. This corresponds exactly to $GL(m, \mathbf{R})$, using the usual correspondence between linear mappings from $\mathbf{R}^m$ into itself and $m \times m$ matrices of real numbers.

The fact that $GL(m, \mathbf{R})$ corresponds to an open set in $\mathbf{R}^{m^2}$, as in Subsection 5.2.1, can also be obtained from the remarks in this section. If $R.T \in GL(\mathbf{R}^m)$, then it is easy to see that

$$R^{-1} - T^{-1} = R^{-1} \circ (T - R) \circ T^{-1}, \tag{5.4.10}$$

as linear mappings on $\mathbf{R}^m$. One can use this to get continuity and other regularity properties of $R \mapsto R^{-1}$ on $GL(\mathbf{R}^m)$.

5.5 Continuity of the differential

Let m and n be positive integers, let W be a nonempty open subset of $\mathbf{R}^n$, and let f be a differentiable mapping from W into $\mathbf{R}^m$. This means that f is differentiable in the sense of Section 3.8 at every point in W, so that the differential

$$f'(x) = df_x \tag{5.5.1}$$

of f at $x \in W$ defines a function on W with values in $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$.

Suppose from now on in this section that f is continuously differentiable on W, which implies that f is differentiable on W, as mentioned in Subsection 3.8.2. In this case,

$$\begin{aligned}&\text{the differential of } f \text{ is continuous}\\ &\text{as a function on } W \text{ with values in } \mathcal{L}(\mathbf{R}^n, \mathbf{R}^m),\end{aligned} \tag{5.5.2}$$

because the entries of the corresponding matrices are given by the partial derivatives of the components of f. In fact, continuous differentiability may be defined equivalently in terms of the existence and continuity of the differential in this way, as in Definition 9.20 and Theorem 9.21 on p219 of [155].

Let us also suppose from now on in this section that $m = n$. One can check that

$$\{x \in W : f'(x) = df_x \in GL(\mathbf{R}^n)\} \tag{5.5.3}$$

is an open set in $\mathbf{R}^n$. This uses the continuity of the differential of f on W, the hypothesis that W be an open set in $\mathbf{R}^n$, and the fact that $GL(n, \mathbf{R})$ corresponds to an open set in $\mathbf{R}^{n^2}$, as in Subsection 5.2.1 and the previous section. Although one could use the definition of continuity directly here, it is convenient to use a well-known characterization of continuity in terms of inverse images of open sets, which will be reviewed in Section 5.9.

Alternatively, (5.5.3) is the same as

$$\{x \in W : \det f'(x) = \det df_x \neq 0\}. \tag{5.5.4}$$

Observe that

$$\det f'(x) = \det df_x \text{ is continuous as a real-valued function on } W. \tag{5.5.5}$$

because of (5.5.2), the continuity of the determinant as a real-valued function on $\mathcal{L}(\mathbf{R}^n)$, and the fact that compositions of continuous mappings are continuous, as in Section 3.4. As a slightly different approach, one can use the definition of the determinant to express

$$\det f'(x) = \det df_x \tag{5.5.6}$$

in terms of a sum of products of partial derivatives of the components of f, and use the fact that sums and products of continuous real-valued functions are continuous too. One can verify that (5.5.4) is an open set in $\mathbf{R}^n$, using (5.5.5) and the hypothesis that W be an open set in $\mathbf{R}^n$, as before. Equivalently,

$$\{x \in W : \det f'(x) = \det df_x = 0\} \tag{5.5.7}$$

is relatively closed in W, in the sense that is reviewed in Section 5.7.

We also have that

$$f'(x)^{-1} = (df_x)^{-1} \text{ is continuous on (5.5.3).} \tag{5.5.8}$$

More precisely, this means that the inverse of the differential of f at x, as a linear mapping from $\mathbf{R}^n$ onto itself, is continuous as a function of x on (5.5.3) with values in $\mathcal{L}(\mathbf{R}^n)$. This follows from the continuity of the differential of f, as a function on W with values in $\mathcal{L}(\mathbf{R}^n)$, and the continuity of the mapping that sends an element of $GL(\mathbf{R}^n)$ to its inverse, as in Subsections 5.2.1 and 5.4.1. This uses the fact that compositions of continuous mappings are continuous again too. Alternatively, the entries of the matrix associated to

$$f'(x)^{-1} = (df_x)^{-1} \tag{5.5.9}$$

can be expressed in terms of the partial derivatives of the components of f and (5.5.6) using Cramer's rule, and one can use this to get that the entries of the matrix are continuous as real-valued functions of x on W.

5.6 The inverse function theorem

Let n be a positive integer, let W be a nonempty open subset of $\mathbf{R}^n$, and let f be a continuously-differentiable mapping from W into $\mathbf{R}^n$. Also let $w \in W$ be given, and suppose that

$$f'(w) = df_w \in GL(\mathbf{R}^n). \tag{5.6.1}$$

Under these conditions, the *inverse function theorem* states that there are open subsets $U(w)$ and $V(w)$ of $\mathbf{R}^n$ with the following properties, that we shall describe in two parts.

In the first part, we have that

$$w \in U(w),\ U(w) \subseteq W, \text{ and } f(w) \in V(w). \tag{5.6.2}$$

We are also able to choose $U(w)$ and $V(w)$ so that

(5.6.3) the restriction of f to $U(w)$ is a one-to-one mapping onto $V(w)$

and

$$f'(x) = df_x \in GL(\mathbf{R}^n) \text{ for each } x \in U(w). \tag{5.6.4}$$

Let $g = g_w$ be the inverse of the restriction of f to $U(w)$, so that

$$g \text{ is a one-to-one mapping from } V(w) \text{ onto } U(w), \tag{5.6.5}$$

with

$$g(f(x)) = x \tag{5.6.6}$$

for every $x \in U(w)$. The second part of the inverse function theorem states that

$$g \text{ is continuously differentiable on } V(w), \tag{5.6.7}$$

and that

$$g'(y) = f'(g(y))^{-1} \tag{5.6.8}$$

for each $y \in V(w)$. Equivalently, this means that

$$dg_y = (df_{g(y)})^{-1}, \tag{5.6.9}$$

where the right side is the inverse of the differential of f at $g(y)$, as a linear mapping from $\mathbf{R}^n$ onto itself.

Note that

$$f(g(y)) = y \tag{5.6.10}$$

for each $y \in V(w)$, by construction. The differentiability of g at y implies that

$$f'(g(y)) \circ g'(y) = df_{g(y)} \circ dg_y \text{ is the identity mapping on } \mathbf{R}^n, \tag{5.6.11}$$

by the chain rule, as in Section 3.9. Of course, (5.6.8) or equivalently (5.6.9) follows from this. One can use this to get the continuity of the differential of g from the continuity of g and (5.5.8).

In this case, the restriction of f to $U(w)$ is a C^1 diffeomorphism from $U(w)$ onto $V(w)$, in the sense discussed in Section 5.11. Note that the mean value theorem and intermediate value theorem are very helpful for some properties like these when $n = 1$, as in Section 5.11. If f is r-times continuously differentiable for some positive integer r, or infinitely differentiable, then g has the same property, as we shall see in Section 5.12. This means that the restriction of f to $U(w)$ is a C^r or C^∞ diffeomorphism from $U(w)$ onto $V(w)$, as appropriate.

5.7 Closed sets and limit points

Let n be a positive integer, and let E be a subset of $\mathbf{R}^n$. A point $x \in \mathbf{R}^n$ is said to be a *limit point* of E with respect to the standard Euclidean metric if for every positive real number r there is a $w \in E$ such that $w \neq x$ and

$$\|x - w\|_2 < r, \tag{5.7.1}$$

where $\|\cdot\|_2 = \|\cdot\|_{2,\mathbf{R}^n}$ is the standard Euclidean norm on $\mathbf{R}^n$, as in Section 3.2.

We say that E is a *closed set* in $\mathbf{R}^n$ if for every $x \in \mathbf{R}^n$ such that x is a limit point of E with respect to the standard Euclidean metric, we have that

$$x \in E. \tag{5.7.2}$$

It is well known and not difficult to show that this happens if and only if the *complement*

$$\mathbf{R}^n \setminus E = \{z \in \mathbf{R}^n : z \notin E\} \tag{5.7.3}$$

of E in $\mathbf{R}^n$ is an open set with respect to the standard Euclidean metric.

Note that

$$\mathbf{R}^n \setminus (\mathbf{R}^n \setminus E) = E. \tag{5.7.4}$$

The previous statement is equivalent to saying that

$$U \subseteq \mathbf{R}^n \text{ is an open set if and only if } \mathbf{R}^n \setminus U \text{ is a closed set.} \tag{5.7.5}$$

The *closure* of E in $\mathbf{R}^n$ is defined to be the set

$$\overline{E} = \{x \in \mathbf{R}^n : x \in E \text{ or } x \text{ is a limit point of } E\}. \tag{5.7.6}$$

It is easy to see that E is a closed set in $\mathbf{R}^n$ if and only if

$$E = \overline{E}. \tag{5.7.7}$$

It is well known and not too difficult to show that for any subset E of $\mathbf{R}^n$,

$$\overline{E} \text{ is a closed set in } \mathbf{R}^n. \tag{5.7.8}$$

Let us say that a subset A of $E \subseteq \mathbf{R}^n$ is *relatively closed* in E if A contains all of its limit points in E. Equivalently, this means that

$$A = \overline{A} \cap E. \tag{5.7.9}$$

Note that E is automatically relatively closed as a subset of itself. If E is a closed set in $\mathbf{R}^n$, then $A \subseteq E$ is relatively closed in E if and only if A is a closed set in $\mathbf{R}^n$. This uses the fact that a limit point of A in $\mathbf{R}^n$ is also a limit point of E when $A \subseteq E$. If A_1 is a closed set in $\mathbf{R}^n$, then one can check that

$$A_1 \cap E \tag{5.7.10}$$

is relatively closed in E. Every relatively closed subset of E is of this form, as in (5.7.9).

Let us say that a subset U of $E \subseteq \mathbf{R}^n$ is *relatively open* if for every $x \in U$ there is an $r > 0$ such that

$$B(x, r) \cap E \subseteq U. \tag{5.7.11}$$

It is well known and not too difficult to show that this happens if and only if there is an open subset W of $\mathbf{R}^n$ such that

$$U = W \cap E. \tag{5.7.12}$$

Clearly E is automatically relatively open as a subset of itself. If E is an open set in $\mathbf{R}^n$, then $U \subseteq E$ is relatively open in E if and only if U is an open set in $\mathbf{R}^n$.

It is well known and not too difficult to show that $A \subseteq E$ is relatively closed in E if and only if

$$E \setminus A = \{x \in E : x \notin A\} \tag{5.7.13}$$

is relatively open in E. This means that $U \subseteq E$ is relatively open in E if and only if $E \setminus U$ is relatively closed in E. The relatively open and closed subsets of E are the same as the subsets of E that are open or closed when E is considered as a metric space, with respect to the restriction of the standard Euclidean metric on $\mathbf{R}^n$ to E.

Let us say that E is *bounded* in $\mathbf{R}^n$ if E is contained in a ball. If E is bounded, then one can check that

$$\overline{E} \text{ is bounded in } \mathbf{R}^n \tag{5.7.14}$$

too.

The *boundary* of E in $\mathbf{R}^n$ is defined by

$$\partial E = \overline{E} \cap \overline{(\mathbf{R}^n \setminus E)}. \tag{5.7.15}$$

Note that this is automatically a closed set in $\mathbf{R}^n$.

5.8 Compact sets

Let n be a positive integer, and let K be a subset of $\mathbf{R}^n$. We say that K is *compact* if every open covering of K in $\mathbf{R}^n$ can be reduced to a finite subcovering. It is well known that

$$K \text{ is compact in } \mathbf{R}^n \text{ if and only if } K \text{ is closed and bounded.} \tag{5.8.1}$$

More precisely, a compact subset of any metric space is closed and bounded, and the converse uses additional properties of $\mathbf{R}^n$.

Let E be a subset of $\mathbf{R}^n$, and suppose for the moment that $K \subseteq E$. It is well known and not too difficult to show that K is compact if and only if every covering of K by relatively open subsets of E can be reduced to a finite subcovering. This means that K is compact as a subset of $\mathbf{R}^n$ if and only if K is compact as a subset of E, with respect to the restriction of the standard Euclidean metric on $\mathbf{R}^n$ to E.

If K is compact and A is a closed subset of $\mathbf{R}^n$, then it is well known that

$$K \cap A \text{ is compact.} \tag{5.8.2}$$

Although this follows from (5.8.1), this can be verified more directly, with an argument that works in any metric space.

If every infinite subset of K has a limit point in K, then K is said to have the *limit point property*. If every sequence of elements of K has a subsequence that converges to an element of K, then K is said to be *sequentially compact*. It is well known that compactness, the limit point property, and sequential compactness are equivalent in any metric space.

Suppose that $U \subseteq \mathbf{R}^n$ is an open set, and that

$$K \subseteq U. \tag{5.8.3}$$

This implies that for every $x \in K$ there is a positive real number $r(x)$ such that

$$B(x, r(x)) \subseteq U. \tag{5.8.4}$$

Observe that the open balls $B(x, r(x)/2)$, $x \in K$, form an open covering of E in $\mathbf{R}^n$. If K is nonempty and compact, then there are finitely many elements $x_1, \ldots, x_l$ of K such that

$$K \subseteq \bigcup_{j=1}^{l} B(x_j, r(x_j)/2). \tag{5.8.5}$$

If we put

$$r_0 = \min_{1 \le j \le l} (r(x_j)/2), \tag{5.8.6}$$

then r_0 is a positive real number, and one can check that

$$B(x, r_0) \subseteq U \tag{5.8.7}$$

for every $x \in K$.

5.8.1 Continuity and compactness

Let m be a positive integer, and let f be a continuous mapping from E into $\mathbf{R}^m$. If K is compact in $\mathbf{R}^n$, and $K \subseteq E$, then it is well known that

$$f(K) = \{f(x) : x \in K\} \tag{5.8.8}$$

is compact in $\mathbf{R}^m$.

Let f be a continuous real-valued function on E. If K is compact in $\mathbf{R}^n$, $K \subseteq E$, and $K \neq \emptyset$, then the *extreme value theorem* states that

$$f \text{ attains its maximum and minimum on } K. \tag{5.8.9}$$

This can be obtained from the fact that $f(K)$ is closed and bounded in the real line, because $f(K)$ is compact.

A mapping f from E into $\mathbf{R}^m$ is said to be *uniformly continuous* if for every $\epsilon > 0$ there is a $\delta > 0$ such that for every $x, w \in E$ with

$$\|x - w\|_{2,\mathbf{R}^n} < \delta, \tag{5.8.10}$$

we have that

$$\|f(x) - f(w)\|_{2,\mathbf{R}^m} < \epsilon. \tag{5.8.11}$$

Here $\|\cdot\|_{2,\mathbf{R}^n}$ and $\|\cdot\|_{2,\mathbf{R}^m}$ are the standard Euclidean norms on $\mathbf{R}^n$ and $\mathbf{R}^m$, respectively, as in Section 3.2. It is easy to see that uniform continuity on E implies ordinary continuity on E. If f is continuous on E, and if E is compact in $\mathbf{R}^n$, then it is well known that

(5.8.12) $$f \text{ is uniformly continuous on } E.$$

5.9 More on continuous mappings

Let n and m be positive integers, and let W be an open subset of $\mathbf{R}^n$. It is well known that a mapping f from W into $\mathbf{R}^m$ is continuous if and only if for every open set $V \subseteq \mathbf{R}^m$,

(5.9.1) $$f^{-1}(V) = \{x \in W : f(x) \in V\}$$

is an open set in $\mathbf{R}^n$. If W is any subset of $\mathbf{R}^n$, then f is continuous on W if and only if for every open set $V \subseteq \mathbf{R}^m$,

(5.9.2) $$f^{-1}(V) \text{ is a relatively open set in } W.$$

We also have that f is continuous on W if and only if for every relatively open set $V_0 \subseteq f(W)$,

(5.9.3) $$f^{-1}(V_0) \text{ is relatively open in } W.$$

Similarly, it is well known that a mapping f from a closed set $E \subseteq \mathbf{R}^n$ into $\mathbf{R}^m$ is continuous if and only if for every closed set $A \subseteq \mathbf{R}^m$,

(5.9.4) $$f^{-1}(A) = \{x \in E : f(x) \in A\}$$

is a closed set in $\mathbf{R}^n$. If E is any subset of $\mathbf{R}^n$, then f is continuous on E if and only if for every closed set $A \subseteq \mathbf{R}^m$,

(5.9.5) $$f^{-1}(A) \text{ is a relatively closed set in } E.$$

Observe that

(5.9.6) $$f^{-1}(\mathbf{R}^m \setminus A) = E \setminus f^{-1}(A)$$

for every $A \subseteq \mathbf{R}^m$. Alternatively, f is continuous on E if and only if for every relatively closed set $A_0 \subseteq f(E)$,

(5.9.7) $$f^{-1}(A_0) \text{ is relatively closed in } E.$$

5.9.1 Homeomorphisms

A mapping f from a subset E of $\mathbf{R}^n$ into $\mathbf{R}^m$ is said to be a *homeomorphism* onto its image $f(E) = \{f(x) : x \in E\}$ if f is one-to-one and continuous, and if

(5.9.8) $$\text{the inverse } f^{-1} \text{ of } f \text{ is continuous}$$

as a mapping from $f(E)$ onto E.

If E is compact, and f is a one-to-one continuous mapping from E into $\mathbf{R}^m$, then it is well known that

$$f \text{ is a homeomorphism of } E \text{ onto } f(E). \tag{5.9.9}$$

Note that

$$f(E) \text{ is a compact subset of } \mathbf{R}^m \tag{5.9.10}$$

in this case, so that

$$f(E) \text{ is a closed set in } \mathbf{R}^m \tag{5.9.11}$$

in particular, as in the previous section. In order to show that f^{-1} is continuous on $f(E)$, it suffices to get that

$$(f^{-1})^{-1}(A) \text{ is a closed set in } \mathbf{R}^m \tag{5.9.12}$$

for every closed set $A \subseteq \mathbf{R}^n$, as before.

In this case, one can check that

$$(f^{-1})^{-1}(A) = f(A \cap E). \tag{5.9.13}$$

We also have that

$$A \cap E \text{ is compact in } \mathbf{R}^n, \tag{5.9.14}$$

because E is compact and A is a closed set, as in the previous section. This implies that

$$f(A \cap E) \text{ is compact in } \mathbf{R}^m, \tag{5.9.15}$$

because f is continuous on E. It follows that

$$f(A \cap E) \text{ is a closed set in } \mathbf{R}^m, \tag{5.9.16}$$

as mentioned in the previous section.

5.10 Homeomorphisms and open mappings

Let m and n be positive integers, let E be a subset of $\mathbf{R}^n$, and let f be a mapping from E into $\mathbf{R}^m$. Suppose for the moment that f is a homeomorphism from E onto $f(E)$, which means that

$$f^{-1} \text{ is a homeomorphism from } f(E) \text{ onto } E. \tag{5.10.1}$$

Let k be another positive integer, and let g be a mapping from $f(E)$ into $\mathbf{R}^k$. If g is a homeomorphism from $f(E)$ onto $g(f(E))$, then

$$g \circ f \text{ is a homeomorphism from } E \text{ onto } g(f(E)). \tag{5.10.2}$$

A mapping f from an open subset W of $\mathbf{R}^n$ into $\mathbf{R}^m$ is said to be an *open mapping* if for every open subset U of $\mathbf{R}^n$ with $U \subseteq W$, we have that

$$f(U) \text{ is an open set in } \mathbf{R}^m. \tag{5.10.3}$$

In particular, this implies that

$$f(W) \text{ is an open set in } \mathbf{R}^m. \tag{5.10.4}$$

If W is any subset of $\mathbf{R}^n$, then we may say that f is an *open mapping* as a mapping from W onto its image $f(W)$ in $\mathbf{R}^m$ if for every relatively open subset U of W,

$$f(U) \text{ is relatively open in } f(W). \tag{5.10.5}$$

Of course, (5.10.5) implies (5.10.3) when (5.10.4) holds. Similarly, if W is an open set in $\mathbf{R}^n$, then every relatively open subset U of W is an open set in $\mathbf{R}^n$, so that (5.10.3) implies (5.10.5).

If f is one-to-one on W, then

$$f \text{ is an open mapping as a mapping from } W \text{ onto } f(W) \tag{5.10.6}$$

if and only if

$$f^{-1} \text{ is a continuous mapping from } f(W) \text{ onto } W. \tag{5.10.7}$$

This follows from one of the characterizations of continuity mentioned in the previous section. In this case, f is a homeomorphism from W onto $f(W)$ if and only if f is continuous and an open mapping as a mapping from W onto $f(W)$. If W and $f(W)$ are open sets, then (5.10.6) is the same as saying that

$$f \text{ is an open mapping from } W \text{ into } \mathbf{R}^m, \tag{5.10.8}$$

as in the preceding paragraph.

5.10.1 A sufficient condition

Suppose that for every $w \in W$ there is an open subset $U(w)$ in $\mathbf{R}^n$ such that

$$w \in U(w),\ U(w) \subseteq W, \tag{5.10.9}$$

and

$$\begin{gathered} \text{the restriction of } f \text{ to } U(w) \\ \text{is an open mapping from } U(w) \text{ into } \mathbf{R}^m. \end{gathered} \tag{5.10.10}$$

Under these conditions, one can check that (5.10.8) holds. More precisely, if U is an open subset of $\mathbf{R}^n$ with $U \subseteq W$ and $w \in U$, then $U \cap U(w)$ is an open set in $\mathbf{R}^n$ that is contained in $U(w)$, so that

$$f(U \cap U(w)) \text{ is an open set in } \mathbf{R}^m, \tag{5.10.11}$$

by hypothesis. One can use this to verify that $f(U)$ is an open set in $\mathbf{R}^m$, directly from the definition of an open set. Indeed, every element of $f(U)$ is of the form $f(w)$ for some $w \in U$, and in this case $f(U \cap U(w))$ is an open set in $\mathbf{R}^m$ that contains $f(w)$ and is contained in $f(U)$.

Alternatively, observe that

$$U = \bigcup_{w \in U} (U \cap U(w)). \tag{5.10.12}$$

This implies that

$$f(U) = f\Big(\bigcup_{w \in U} (U \cap U(w))\Big) = \bigcup_{w \in U} f(U \cap U(w)). \tag{5.10.13}$$

We also have that

$$\bigcup_{w \in U} f(U \cap U(w)) \text{ is an open set in } \mathbf{R}^m, \tag{5.10.14}$$

because the union of any family of open sets is an open set as well. This means that $f(U)$ is an open set in $\mathbf{R}^m$, as desired.

5.11 C^1 Diffeomorphisms

Let n be a positive integer, let W be a nonempty open subset of $\mathbf{R}^n$, and let f be a mapping from W into $\mathbf{R}^n$.

Suppose for the moment that f is one-to-one on W, and that $f(W)$ is an open set in $\mathbf{R}^n$ too. If f is differentiable at $x \in W$, and f^{-1} is differentiable at $f(x)$, then the chain rule implies that

$$d(f^{-1})_{f(x)} \circ df_x = df_x \circ d(f^{-1})_{f(x)} = \text{ the identity mapping on } \mathbf{R}^n, \tag{5.11.1}$$

as in Section 3.9. This means that df_x is invertible as a linear mapping on $\mathbf{R}^n$, with

$$(df_x)^{-1} = d(f^{-1})_{f(x)}. \tag{5.11.2}$$

Suppose for the moment again that f is continuously differentiable on W, and that

$$df_x \in GL(\mathbf{R}^n) \tag{5.11.3}$$

for every $x \in W$. Under these conditions, one can check that

$$f \text{ is an open mapping from } W \text{ into } \mathbf{R}^n, \tag{5.11.4}$$

using the inverse function theorem, as in Section 5.6. This also uses the criterion for f to be an open mapping mentioned in Subsection 5.10.1. More precisely in this case, (5.10.10) corresponds to the fact that $V(w)$ is an open set in $\mathbf{R}^n$, and that g_w is a continuous mapping from $V(w)$ onto $U(w)$, in the notation of Section 5.6. In particular, this means that $f(W)$ is an open set in $\mathbf{R}^n$.

5.11.1 The definition of a diffeomorphism

Suppose that f is one-to-one on W again, and that f is continuously differentiable on W. If $f(W)$ is an open set in $\mathbf{R}^n$, and

$$f^{-1} \text{ is continuously differentiable on } f(W), \tag{5.11.5}$$

then we say that

$$f \text{ is a } C^1 \text{ diffeomorphism from } W \text{ onto } f(W). \tag{5.11.6}$$

This implies that (5.11.3) holds for every $x \in W$, as before. Of course, if (5.11.6) holds, then

$$f \text{ is a homeomorphism from } W \text{ onto } f(W). \tag{5.11.7}$$

Alternatively, if (5.11.3) holds for every $x \in W$, then $f(W)$ is an open set in $\mathbf{R}^n$, as before. If f is also one-to-one on W, then one can check that (5.11.5) holds, using the inverse function theorem. More precisely, if $w \in U$, then the restriction of f^{-1} to the open set $V(w)$ in $\mathbf{R}^n$ mentioned in Section 5.6 is the same as the function g_w mentioned there. This means that f^{-1} is continuously differentiable on $V(w)$. It follows that (5.11.5) holds, because $w \in U$ is arbitrary, and $V(w)$ is an open set that contains $f(w)$.

Suppose that f is a C^1 diffeomorphism from W onto $f(W)$, which implies that

$$f^{-1} \text{ is a } C^1 \text{ diffeomorphism from } f(W) \text{ onto } W. \tag{5.11.8}$$

If g is a mapping from $f(W)$ into $\mathbf{R}^n$ that is a C^1 diffeomorphism from $f(W)$ onto $g(f(W))$, then

$$g \circ f \text{ is a } C^1 \text{ diffeomorphism from } W \text{ onto } g(f(W)). \tag{5.11.9}$$

5.11.2 The $n = 1$ case

Suppose now that $n = 1$, and that f is a differentiable real-valued function on W. If $x \in W$, then we let $f'(x)$ be the usual derivative of f at x, which is a real number. Thus the differential of f at x, as a linear mapping from $\mathbf{R}$ into itself, corresponds to multiplication by $f'(x)$. In this case, (5.11.3) is the same as saying that

$$f'(x) \neq 0. \tag{5.11.10}$$

Suppose also that W is connected, so that W is basically an open interval in $\mathbf{R}$, which may be unbounded, like an open half-line, or the real line. If (5.11.10) holds for every $x \in W$, then either

$$f'(x) > 0 \tag{5.11.11}$$

for every $x \in W$, or

$$f'(x) < 0 \tag{5.11.12}$$

for every $x \in W$. This follows from the intermediate value theorem when f' is continuous on W. There is a version of the intermediate value theorem for the derivative of a differentiable real-valued function on an interval, even if the derivative is not continuous, as in Theorem 5.12 on p108 of [155]. One can use this to get the same conclusion without asking that f' be continuous on W.

If (5.11.11) holds on W, then

$$f \text{ is strictly increasing on } W, \tag{5.11.13}$$

by the mean-value theorem. Similarly, if (5.11.12) holds on W, then

$$f \text{ is strictly decreasing on } W. \tag{5.11.14}$$

In both cases, we get that f is one-to-one on W. It is easy to see that f is an open mapping on W in both cases too, using the intermediate value theorem. This uses the fact that f is continuous on W, because f is differentiable.

5.12 Diffeomorphisms with more regularity

Let n and r be positive integers, let W be a nonempty open subset of $\mathbf{R}^n$, and let f be an r-times continuously differentiable mapping from W into $\mathbf{R}^n$. Thus the differential df_x of f at $x \in W$ defines a function on W with values in the space $\mathcal{L}(\mathbf{R}^n)$ of linear mappings from $\mathbf{R}^n$ into itself. Observe that

$$df \in C^{r-1}(W, \mathcal{L}(\mathbf{R}^n)), \tag{5.12.1}$$

in the notation of Subsections 5.3.1 and 5.4.1, because the entries of the $n \times n$ matrix corresponding to df_x are given by the partial derivatives of the components of f at $x \in W$.

Using (5.12.1), we get that

$$\det df \in C^{r-1}(W, \mathbf{R}), \tag{5.12.2}$$

because the determinant is a smooth function on $\mathcal{L}(\mathbf{R}^n)$. Alternatively,

$$\det df \tag{5.12.3}$$

can be expressed in terms of sums of products of partial derivatives of the cmponents of f, each of which is $(r-1)$-times continuously differentiable on W, by hypothesis.

Suppose also now that $df_x \in GL(\mathbf{R}^n)$ for every $x \in W$. In this case, we get that

$$(df)^{-1} \in C^{r-1}(W, \mathcal{L}(\mathbf{R}^n)), \tag{5.12.4}$$

because the mapping that sends an element of $GL(\mathbf{R}^n)$ to its inverse is smooth, as in Subsections 5.2.1 and 5.4.1. This could be obtained from Cramer's rule too, using (5.12.2).

5.12.1 Regularity of the inverse mapping

Suppose that f is one-to-one on W as well, and remember that f^{-1} is continuously differentiable on W, as in Subsection 5.11.1, and that the differential of f^{-1} is as in (5.11.2). Equivalently, this means that

$$d(f^{-1})_y = (df_{f^{-1}(y)})^{-1} \tag{5.12.5}$$

for every $y \in f(W)$. Let us use this to get that

$$d(f^{-1}) \in C^{r-1}(f(W), \mathcal{L}(\mathbf{R}^n)), \tag{5.12.6}$$

which is the same as saying that

$$f^{-1} \in C^r(f(W), \mathbf{R}^n). \tag{5.12.7}$$

More precisely, suppose for the moment that

$$f^{-1} \in C^l(f(W), \mathbf{R}^n), \tag{5.12.8}$$

where $0 \leq l \leq r - 1$. Observe that

$$df_{f^{-1}(y)} \text{ is } l\text{-times continuously differentiable} \tag{5.12.9}$$

as a function of $y \in f(W)$ with values in $\mathcal{L}(\mathbf{R}^n)$, because of (5.12.1). This uses the fact that compositions of l-times continuously differentiable functions are l-times continuously differentiable too, as in Section 3.9. It follows that

$$(df_{f^{-1}(y)})^{-1} \text{ is } l\text{-times continuously differentiable} \tag{5.12.10}$$

as a function of $y \in f(W)$ with values in $\mathcal{L}(\mathbf{R}^n)$, as before.

This means that

$$d(f^{-1}) \in C^l(f(W), \mathcal{L}(\mathbf{R}^n)), \tag{5.12.11}$$

because of (5.12.5). This is the same as saying that

$$f^{-1} \in C^{l+1}(f(W), \mathbf{R}^n), \tag{5.12.12}$$

as before. One can repeat the process as needed to get (5.12.6), and thus (5.12.7).

5.12.2 C^r And C^∞ diffeomorphisms

Let us say that

$$f \text{ is a } C^r \text{ diffeomorphism from } W \text{ onto } f(W) \tag{5.12.13}$$

under these conditions. This implies that

$$f^{-1} \text{ is a } C^r \text{ diffeomorphism from } f(W) \text{ onto } W. \tag{5.12.14}$$

If g is a mapping from $f(W)$ into $\mathbf{R}^n$ that is a C^r diffeomorphism from $f(W)$ onto $g(f(W))$, then

$$g \circ f \text{ is a } C^r \text{ diffeomorphism from } W \text{ onto } g(f(W)). \tag{5.12.15}$$

Similarly, if f is infinitely differentiable on W, then

$$df \in C^\infty(W, \mathcal{L}(\mathbf{R}^n)), \tag{5.12.16}$$

and

$$\det df \in C^\infty(W, \mathbf{R}). \tag{5.12.17}$$

If $df_x \in GL(\mathbf{R}^n)$ for every $x \in W$, then it follows that

$$(df)^{-1} \in C^\infty(W, \mathcal{L}(\mathbf{R}^n)), \tag{5.12.18}$$

as before. If f is also one-to-one on W, then we get that

$$f^{-1} \in C^\infty(W, \mathbf{R}^n), \tag{5.12.19}$$

as in (5.12.7). In this case, we say that

$$f \text{ is a } C^\infty \text{ diffeomorphism from } W \text{ onto } f(W), \tag{5.12.20}$$

and we have that

$$f^{-1} \text{ is a } C^\infty \text{ diffeomorphism from } f(W) \text{ onto } W. \tag{5.12.21}$$

If g is a mapping from $f(W)$ into $\mathbf{R}^n$ that is a C^∞ diffeomorphism from $f(W)$ onto $g(f(W))$ as well, then

$$g \circ f \text{ is a } C^\infty \text{ diffeomorphism from } W \text{ onto } g(f(W)). \tag{5.12.22}$$

5.13 Cells and supports of functions

Let n be a positive integer, and let $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ be real numbers with $a_j \le b_j$ for each $j = 1, \ldots, n$. The Cartesian product

$$\mathcal{C} = \prod_{j=1}^n [a_j, b_j] \tag{5.13.1}$$

of the closed intervals $[a_j, b_j]$, $1 \le j \le n$, is called a *cell* in $\mathbf{R}^n$.

It is well known that

$$\text{cells in } \mathbf{R}^n \text{ are compact,} \tag{5.13.2}$$

with respect to the standard Euclidean metric. In order to show that subsets of $\mathbf{R}^n$ that are closed and bounded are also compact, one often starts by showing that cells are compact. Note that a continuous mapping from a cell in $\mathbf{R}^n$ into $\mathbf{R}^m$ for some positive integer m is uniformly continuous, as in Subsection 5.8.1.

Let $e_1, \ldots, e_n$ be the standard basis vectors in $\mathbf{R}^n$, and let $\alpha_1, \ldots, \alpha_n$ be real numbers. Consider the corresponding parallelepiped

$$P_{\mathbf{R}^n}(\alpha_1\, e_1, \ldots, \alpha_n\, e_n) \tag{5.13.3}$$

in $\mathbf{R}^n$, as in Section 4.14. This is the same as the cell

$$\prod_{j=1}^{n} I_j, \tag{5.13.4}$$

where $I_j = [0, \alpha_j]$ when $\alpha_j \geq 0$, and $I_j = [\alpha_j, 0]$ when $\alpha_j \leq 0$. Every cell in $\mathbf{R}^n$ is a translate of a cell of this type. Note that the n-dimensional volume of the cell $\mathcal{C}$ is given by

$$\mathrm{Vol}_n(\mathcal{C}) = \prod_{j=1}^{n} (b_j - a_j). \tag{5.13.5}$$

5.13.1 Integrals over cells

Let f be a continuous real-valued function on $\mathcal{C}$, which is in fact uniformly continuous on $\mathcal{C}$, as mentioned earlier. One can check that for each $l = 1, \ldots, n$,

$$\int_{a_l}^{b_l} f(x_1, \ldots, x_l, \ldots, x_n)\, dx_l \tag{5.13.6}$$

is uniformly continuous as a function of the $n - 1$ variables

$$x_1, \ldots, x_{l-1}, x_{l+1}, \ldots, x_n. \tag{5.13.7}$$

More precisely, (5.13.6) is uniformly continuous as a real-valued function on the cell in $\mathbf{R}^{n-1}$ corresponding to the Cartesian product of the closed intervals $[a_j, b_j]$ with $j \neq l$.

Of course,

$$\int_{\mathcal{C}} f(x)\, dx \tag{5.13.8}$$

may be defined as an n-dimensional Riemann integral over $\mathcal{C}$ in a standard way. It is well known that this is the same as the iterated integral

$$\int_{a_1}^{b_1} \cdots \Big(\int_{a_n}^{b_n} f(x_1, \ldots, x_n)\, dx_n \Big) \cdots dx_1. \tag{5.13.9}$$

In fact, it is well known that (5.13.8) is the same as iterated integrals like (5.13.9), where $x_1, \ldots, x_n$ are integrated in any order. This may be shown intially for $n = 2$, which can be used to get the analogous statement for any n.

5.13.2 Supports of functions

Let m be another positive integer, and let f be a function on $\mathbf{R}^n$ with values in $\mathbf{R}^m$. The *support* of f is defined the closure of the set of $x \in \mathbf{R}^n$ such that $f(x) \neq 0$, which may be denoted

$$\operatorname{supp} f = \overline{\{x \in \mathbf{R}^n : f(x) \neq 0\}}. \tag{5.13.10}$$

Note that the support of f is automatically a closed set in $\mathbf{R}^n$. If the set of $x \in \mathbf{R}^n$ such that $f(x) \neq 0$ is bounded, then the support of f is bounded too, and thus compact.

Let f be a continuous real-valued function on $\mathbf{R}^n$ with bounded support. This implies that there is a cell $\mathcal{C}$ in $\mathbf{R}^n$ such that

$$\operatorname{supp} f \subseteq \mathcal{C}. \tag{5.13.11}$$

In this case, the integral of f over $\mathbf{R}^n$ may be defined by

$$\int_{\mathbf{R}^n} f(x)\,dx = \int_{\mathcal{C}} f(x)\,dx, \tag{5.13.12}$$

as in (3) on p247 of [155]. One can check that this does not depend on the particular cell $\mathcal{C}$ in $\mathbf{R}^n$ that satisfies (5.13.11).

5.14 Some smooth functions on R

Let ϕ be an infinitely-differentiable real-valued function on the real line such that

$$\begin{aligned} \phi(t) &> 0 \quad \text{when } t > 0 \\ &= 0 \quad \text{when } t \leq 0, \end{aligned} \tag{5.14.1}$$

as in Subsection 3.7.2. If a, b are real numbers with $a < b$, then put

$$\phi_{a,b}(t) = \phi(t-a)\,\phi(b-t). \tag{5.14.2}$$

This is an infinitely-differentiable real-valued function on the real line such that

$$\begin{aligned} \phi_{a,b}(t) &> 0 \quad \text{when } a < t < b \\ &= 0 \quad \text{when } t \leq a \text{ and when } b \leq t. \end{aligned} \tag{5.14.3}$$

If $r \in \mathbf{R}$, then

$$\phi_{a,b}(((a+b)/2) + r) = \phi(((b-a)/2) + r)\,\phi(((b-a)/2) - r). \tag{5.14.4}$$

This implies that

$$\phi_{a,b}(((a+b)/2) + r) = \phi_{a,b}(((a+b)/2) - r). \tag{5.14.5}$$

Note that

$$\int_a^b \phi_{a,b}(u)\,du > 0. \tag{5.14.6}$$

Put

$$\psi_{a,b}(t) = \Big(\int_0^1 \phi_{a,b}(u)\,du\Big)^{-1} \phi_{a,b}(t) \tag{5.14.7}$$

for each $t \in \mathbf{R}$. This is an infinitely-differentiable real-valued function on the real line such that

$$\begin{aligned} \psi_{a,b}(t) &> 0 \quad \text{when } a < t < b \\ &= 0 \quad \text{when } t \le a \text{ and when } b \le t \end{aligned} \tag{5.14.8}$$

and

$$\int_a^b \psi_{a,b}(t)\,dt = 1. \tag{5.14.9}$$

We can integrate $\psi_{a,b}$ on $\mathbf{R}$ to get an infinitely-differentiable real-valued function $\eta_{a,b}$ on $\mathbf{R}$ such that

$$\eta'_{a,b} = \psi_{a,b} \tag{5.14.10}$$

on $\mathbf{R}$, and

$$\eta_{a,b}(t) = 0 \quad \text{when } t \le a. \tag{5.14.11}$$

This means that

$$\eta_{a,b}(t) = \int_a^t \psi_{a,b}(u)\,du \tag{5.14.12}$$

when $t \ge a$, so that

$$0 < \eta_{a,b}(t) \le 1 \tag{5.14.13}$$

when $t > a$, and

$$\eta_{a,b}(t) = 1 \tag{5.14.14}$$

when $t \ge b$. Note that $\eta_{a,b}$ is strictly increasing on $[a, b]$, by construction.

Note that $1 - \eta_{a,b}$ is an infinitely-differentiable real-valued function on $\mathbf{R}$ such that

$$1 - \eta_{a,b}(t) = 1 \tag{5.14.15}$$

when $t \le a$,

$$0 < 1 - \eta_{a,b}(t) < 1 \tag{5.14.16}$$

when $a < t < b$, and

$$1 - \eta_{a,b}(t) = 0 \tag{5.14.17}$$

when $t \ge b$. We also have that $1 - \eta_{a,b}$ is strictly decreasing on $[a, b]$. More precisely,

$$1 - \eta_{a,b}(t) = \int_t^b \psi_{a,b}(u)\,du \tag{5.14.18}$$

when $t \le b$, by (5.14.9) and (5.14.12). One can check that

$$1 - \eta_{a,b}(((a+b)/2) + r) = \eta_{a,b}(((a+b)/2) - r) \tag{5.14.19}$$

for every $r \in \mathbf{R}$. This uses (5.14.5), which implies that $\psi_{a,b}$ has the same property.

If c, d are real numbers with $b \le c < d$, then put

$$\xi(t) = \xi_{a,b,c,d}(t) = \eta_{a,b}(t)\,(1 - \eta_{c,d}(t)) \tag{5.14.20}$$

for each $t \in \mathbf{R}$. This is an infinitely-differentiable real-valued function on $\mathbf{R}$ with

$$\xi(t) = 0 \tag{5.14.21}$$

when $t \le a$ and when $d \le t$. Similarly,

$$0 < \xi(t) < 1 \tag{5.14.22}$$

when $a < t < b$ and when $c < t < d$, and

$$\xi(t) = 1 \tag{5.14.23}$$

when $b \le t \le c$. In fact, ξ is strictly increasing on $[a, b]$, and strictly decreasing on $[c, d]$.

5.15 Some smooth functions on $\mathbf{R}^n$

Let n be a positive integer, let $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ be real numbers with $a_j \le b_j$ for each $j = 1, \ldots, n$, and let $\mathcal{C} = \prod_{j=1}^n [a_j, b_j]$ be the corresponding cell in $\mathbf{R}^n$, as in Section 5.13. Also let $a_1', \ldots, a_n'$ and $b_1', \ldots, b_n'$ be real numbers with

$$a_j' < a_j \text{ and } b_j < b_j' \tag{5.15.1}$$

for each j. Note that

$$\mathcal{C} \subseteq \prod_{j=1}^n (a_j', b_j'), \tag{5.15.2}$$

and that the right side is an open set in $\mathbf{R}^n$.

As in the previous section, there is an infinitely-differentiable real-valued function ξ_j on $\mathbf{R}$ for each j such that

$$0 \le \xi_j \le 1 \tag{5.15.3}$$

on $\mathbf{R}$,

$$\xi_j(t) > 0 \text{ if and only if } a_j' < t < b_j', \tag{5.15.4}$$

and

$$\xi_j(t) = 1 \text{ if and only if } a_j \le t \le b_j. \tag{5.15.5}$$

Put

$$\zeta(x) = \prod_{j=1}^n \xi_j(x_j) \tag{5.15.6}$$

for each $x \in \mathbf{R}^n$, which defines an infinitely-differentiable real-valued function on $\mathbf{R}^n$ such that

$$0 \le \zeta(x) \le 1 \tag{5.15.7}$$

for every $x \in \mathbf{R}^n$. Observe that

$$\zeta(x) > 0 \text{ if and only if } x \in \prod_{j=1}^{n} (a_j', b_j'), \tag{5.15.8}$$

because of (5.15.4). Similarly,

$$\zeta(1) = 1 \text{ if and only if } x \in \mathcal{C}, \tag{5.15.9}$$

because of (5.15.5). Using (5.15.8), we get that the support of ζ is given by

$$\operatorname{supp} \zeta = \prod_{j=1}^{n} [a_j', b_j']. \tag{5.15.10}$$

Now let a, b be positive real numbers with $a < b$, so that $a^2 < b^2$, and let

$$\eta_{a^2,b^2} \tag{5.15.11}$$

be the infinitely-differentiable real-valued function on $\mathbf{R}$ corresponding to a^2 and b^2 as in the previous section. If $x \in \mathbf{R}^n$, then put

$$\rho_{a,b}(x) = 1 - \eta_{a^2,b^2}(\|x\|_2^2) = 1 - \eta_{a^2,b^2}\Big(\sum_{j=1}^{n} x_j^2\Big), \tag{5.15.12}$$

where $\|x\|_2$ is the standard Euclidean norm on $\mathbf{R}^n$, as in Section 3.2. This is an infinitely-differentiable real-valued function on $\mathbf{R}^n$, with

$$0 \le \rho_{a,b}(x) \le 1 \tag{5.15.13}$$

for every $x \in \mathbf{R}^n$. In fact, we have that

$$\rho_{a,b}(x) = 1 \tag{5.15.14}$$

when $\|x\|_2 \le a$,

$$0 < \rho_{a,b}(x) < 1 \tag{5.15.15}$$

when $a < \|x\|_2 < b$, and

$$\rho_{a,b}(x) = 0 \tag{5.15.16}$$

when $\|x\|_2 \ge b$. It follows that

$$\operatorname{supp} \rho_{a,b} = \{x \in \mathbf{R}^n : \|x\|_2 \le b\}. \tag{5.15.17}$$

Chapter 6

Integration and n-surfaces

6.1 Some partitions of unity

Let n be a positive integer, and let $\phi_1, \ldots, \phi_l$ be finitely many real-valued functions on $\mathbf{R}^n$, with

$$0 \le \phi_k \le 1 \tag{6.1.1}$$

on $\mathbf{R}^n$ for each $k = 1, \ldots, n$. Put $\psi_1 = \phi_1$, and

$$\psi_k = \Big(\prod_{p=1}^{k-1} (1 - \phi_p)\Big)\, \phi_k \tag{6.1.2}$$

for each $k = 2, \ldots, l$, as in (28) on p251 of [155]. Observe that $0 \le 1 - \phi_p \le 1$ for each $p = 1, \ldots, l$, so that

$$0 \le \psi_k \le 1 \tag{6.1.3}$$

on $\mathbf{R}^n$ for each $k = 1, \ldots, l$. We also have that

$$\psi_k(x) = 0 \text{ when } \phi_k(x) = 0, \tag{6.1.4}$$

so that

$$\operatorname{supp} \psi_k \subseteq \operatorname{supp} \phi_k \tag{6.1.5}$$

for every $k = 1, \ldots, l$.

Let us check that

$$\sum_{p=1}^{k} \psi_p = 1 - \prod_{p=1}^{k} (1 - \phi_p) \tag{6.1.6}$$

for each $k = 1, \ldots, l$, as in (29) on p251 of [155]. This follows from the definition of ψ_1 when $k = 1$, and we can use induction to go from k to $k + 1$ when $k < l$. This uses the fact that

$$\psi_{k+1} = \Big(\prod_{p=1}^{k} (1 - \phi_p)\Big)\, \phi_{k+1} = \prod_{p=1}^{k} (1 - \phi_p) - \prod_{p=1}^{k+1} (1 - \phi_p) \tag{6.1.7}$$

when $k < l$. It follows that

$$\sum_{p=1}^{l} \psi_p(x) = 1 \tag{6.1.8}$$

on the set

$$\bigcup_{q=1}^{l} \{x \in \mathbf{R}^n : \phi_q(x) = 1\}, \tag{6.1.9}$$

as on p251 of [155]. This is a type of *partition of unity* for the set (6.1.9), as on p251 of [155].

If the ϕ_k's are continuous on $\mathbf{R}^n$, then

$$\text{the } \psi_k\text{'s are continuous on } \mathbf{R}^n \tag{6.1.10}$$

too. Similarly, if the ϕ_k's are r-times continuously differentiable on $\mathbf{R}^n$ for some positive integer r, then

$$\text{the } \psi_k\text{'s are } r\text{-times continuously differentiable on } \mathbf{R}^n \tag{6.1.11}$$

as well. If the ϕ_k's are infinitely differentiable on $\mathbf{R}^n$, then it follows that

$$\text{the } \psi_k\text{'s are infinitely differentiable on } \mathbf{R}^n. \tag{6.1.12}$$

This is related to Exercise 6 on p289 of [155].

Let m be another positive integer, let E be a subset of $\mathbf{R}^n$, and let f be a function on E with values in $\mathbf{R}^m$. If E is contained in (6.1.9), then we get that

$$f(x) = \sum_{p=1}^{l} \psi_p(x)\, f(x) \tag{6.1.13}$$

for every $x \in E$, by (6.1.8). Similarly, if f is defined on all of $\mathbf{R}^n$, and the support of f is contained in (6.1.9), then (6.1.13) holds for every $x \in \mathbf{R}^n$.

6.2 The standard simplex in $\mathbf{R}^n$

Let n be a positive integer, and consider

$$Q^n = \Big\{x \in \mathbf{R}^n : x_j \geq 0 \text{ for each } j = 1, \ldots, n, \text{ and } \sum_{j=1}^{n} x_j \leq 1\Big\}. \tag{6.2.1}$$

This is known as the *standard simplex or standard n-simplex in* $\mathbf{R}^n$, as in Example 10.4 on p247 of [155], and as mentioned on p141 of [184]. Note that

$$Q^n \text{ is closed and bounded in } \mathbf{R}^n, \tag{6.2.2}$$

and thus compact.

In fact,

$$Q^n \subseteq [0,1]^n, \tag{6.2.3}$$

where the right side is the *unit cube* in $\mathbf{R}^n$, consisting of $x \in \mathbf{R}^n$ such that $0 \le x_j \le 1$ for each $j = 1, \ldots, n$. This is the same as the Cartesian product of n copies of the closed unit interval $[0,1]$ in the real line, which is a cell in $\mathbf{R}^n$ in particular. If $n = 1$, then we have equality in (6.2.3).

Let f be a continuous real-valued function on Q^n, so that f is bounded and uniformly continuous on Q^n, because Q^n is compact. The integral

$$\int_{Q^n} f(x)\,dx \tag{6.2.4}$$

of f over Q^n may be defined as an n-dimensional Riemann integral in a standard way. Another way to look at this is discussed in Example 10.4 on p247 of [155]. Of course, if $n = 1$, then this is the same as an ordinary Riemann integral on $[0,1]$, and so we suppose that $n \ge 2$.

Let Q^{n-1} be the standard simplex in $\mathbf{R}^{n-1}$, which is defined in the same way as before. If $x \in Q^n$, then

$$(x_1, \ldots, x_{n-1}) \in Q^{n-1}. \tag{6.2.5}$$

More precisely, if (6.2.5) holds, then

$$(x_1, \ldots, x_{n-1}, x_n) \in Q^n \tag{6.2.6}$$

if and only if

$$0 \le x_n \le 1 - \sum_{j=1}^{n-1} x_j, \tag{6.2.7}$$

and every element of Q^n occurs in this way.

If (6.2.5) holds, then put

$$f_{n-1}(x_1, \ldots, x_{n-1}) = \int_0^{1-\sum_{j=1}^{n-1} x_j} f(x_1, \ldots, x_{n-1}, x_n)\,dx_n. \tag{6.2.8}$$

One can check that f_{n-1} is bounded and uniformly continuous on Q^{n-1}, because f is bounded and uniformly continuous on Q^n. We also have that

$$f_{n-1}(x_1, \ldots, x_{n-1}) = 0 \text{ when } \sum_{j=1}^{n-1} x_j = 1, \tag{6.2.9}$$

by construction.

Let us extend f_{n-1} to a real-valued function on $[0,1]^{n-1}$ by putting

$$f_{n-1}(x_1, \ldots, x_{n-1}) = 0 \tag{6.2.10}$$

when $(x_1, \ldots, x_{n-1}) \in [0,1]^{n-1} \setminus Q^{n-1}$, which means that

$$\sum_{j=1}^{n-1} x_j > 1. \tag{6.2.11}$$

This defines a continuous function on $[0,1]^{n-1}$, because of (6.2.9).

Thus the integral of f_{n-1} over $[0,1]^{n-1}$ may be defined in the usual way, which is the same as the integral of f_{n-1} over Q^{n-1}. This is one way to look at the integral of f over Q^n.

6.2.1 More on integration over Q^n

We can also extend f to a real-valued function on $[0,1]^n$ by putting

$$f(x) = 0 \tag{6.2.12}$$

when $x \in [0,1]^n \setminus Q^n$, which means that

$$\sum_{j=1}^{n} x_j > 1. \tag{6.2.13}$$

However, this extension will not be continuous at $x \in [0,1]^n$ such that

$$\sum_{j=1}^{n} x_j = 1, \tag{6.2.14}$$

unless (6.2.12) holds. Note that the previous extension of f_{n-1} to $[0,1]^{n-1}$ may be expressed as

$$f_{n-1}(x_1, \ldots, x_{n-1}) = \int_0^1 f(x_1, \ldots, x_{n-1}, x_n)\, dx_n, \tag{6.2.15}$$

using this extension of f to $[0,1]^n$.

To deal with this, a continuous real-valued function $F = F_\delta$ on $[0,1]^n$ may be defined on $[0,1]^n$ for $0 < \delta < 1$, with

$$\begin{aligned} F(x) &= f(x) \quad \text{when } \sum_{j=1}^{n} x_j \le 1 - \delta \\ &= 0 \qquad\;\; \text{when } \sum_{j=1}^{n} x_j \ge 1, \end{aligned} \tag{6.2.16}$$

and

$$|F(x)| \le |f(x)| \quad \text{when } 1 - \delta < \sum_{j=1}^{n} x_j < 1, \tag{6.2.17}$$

as on p247 of [155]. More precisely, we also have that

$$|f(x) - F(x)| \leq |f(x)| \quad \text{when } 1 - \delta < \sum_{j=1}^{n} x_j < 1, \tag{6.2.18}$$

because $F(x)$ is defined so that it has the same sign as $f(x)$ in this case.

Let C be a nonnegative real number such that

$$|f(x)| \leq C \tag{6.2.19}$$

for every $x \in Q^n$. If $F_{n-1} = F_{\delta.n-1}$ is defined on $[0,1]^{n-1}$ as in (6.2.15), then one can check that

$$|f_{n-1}(x_1, \ldots, x_{n-1}) - F_{n-1}(x_1, \ldots, x_{n-1})| \leq C\,\delta \tag{6.2.20}$$

on $[0,1]^{n-1}$, as in (7) on p247 of [155]. In particular, this implies that the integral of f_{n-1} over $[0,1]^{n-1}$ is approximated by the integral of F over $[0,1]^{n-1}$.

Of course, we could start by integrating f in any of the variables $x_1, \ldots, x_n$, with the same properties as before. The result would be uniformly approximated by the integral of F in the same variable.

6.3 Jacobians and changes of variables

Let n be a positive integer, let W be a nonempty open subset of $\mathbf{R}^n$, and let ϕ be a mapping from W into $\mathbf{R}^n$. If ϕ is differentiable at a point $x \in W$, then the determinant

$$\det \phi'(x) = \det d\phi_x \tag{6.3.1}$$

of the differential $\phi'(x) = d\phi_x$ of ϕ at x is known as the *Jacobian* of ϕ at x. Note that if $n = 1$, then the Jacobian of ϕ at x is the same as the ordinary derivative of ϕ at x.

Suppose that ϕ is continously differentiable on W, so that the differential $\phi'(x)$ of ϕ at x is continuous as a function of $x \in W$ with values in the space $\mathcal{L}(\mathbf{R}^n)$ of linear mappings from $\mathbf{R}^n$ into itself, as in Section 5.5. This implies that

$$\det \phi'(x) \text{ is continuous} \tag{6.3.2}$$

as a real-valued function of $x \in W$, as before.

Suppose now that $\phi(W)$ is an open set in $\mathbf{R}^n$, and that

$$\phi \text{ is a } C^1 \text{ diffeomorphism from } W \text{ onto } \phi(W), \tag{6.3.3}$$

as in Subsection 5.11.1. Let K be a compact subset of $\mathbf{R}^n$ that is contained in W, and let f be a continuous real-valued function on $\phi(K)$. Note that $\phi(K)$ is a compact set in $\mathbf{R}^n$ too, as in Subsection 5.8.1. It is well known that

$$\int_{\phi(K)} f(y)\,dy = \int_K f(\phi(x))\,|\det \phi'(x)|\,dx, \tag{6.3.4}$$

at least if K is nice enough for the Riemann integral to be defined. There are versions of this using Lebesgue measure and integration with milder hypotheses, but we shall not pursue this here.

Remember that if T is a linear mapping from $\mathbf{R}^n$ into itself and E is a reasonably nice subset of $\mathbf{R}^n$, then the n-dimensional volume of $T(E)$ is equal to $|\det T|$ times the n-dimensional volume of E, as mentioned in Section 4.12. Basically (6.3.4) behaves approximately like this on sufficiently small sets. Of course, if ϕ is a linear mapping on $\mathbf{R}^n$, then its Jacobian is constant, and (6.3.4) is somewhat simpler. This is also somewhat simpler when $n = 1$, in which case it may normally be stated with milder hypotheses.

This version of the change of variables formula is essentially the same as the one mentioned in (1) in Section 4.4 on p141 of [184]. Another version is given in Theorem 10.9 on p252 of [155], which will be discussed on the next section.

6.4 Another version of changing variables

Let us continue with the same notation and hypotheses as in the previous section. Let f be a continuous real-valued function $\mathbf{R}^n$ with compact support and

$$\operatorname{supp} f \subseteq \phi(W). \tag{6.4.1}$$

This implies that

$$\phi^{-1}(\operatorname{supp} f) \text{ is a compact set in } \mathbf{R}^n, \tag{6.4.2}$$

as in Subsection 5.8.1, because ϕ^{-1} is continuous as a mapping from $\phi(W)$ into $\mathbf{R}^n$. Of course,

$$\phi^{-1}(\operatorname{supp} f) \subseteq W, \tag{6.4.3}$$

because $\phi^{-1}(\phi(W)) = W$.

Let f_ϕ be the real-valued function defined on $\mathbf{R}^n$ by

$$\begin{aligned} f_\phi(x) &= f(\phi(x)) && \text{when } x \in W \\ &= 0 && \text{when } x \in \mathbf{R}^n \setminus W. \end{aligned} \tag{6.4.4}$$

One can check that

$$f_\phi \text{ is continuous on } \mathbf{R}^n, \tag{6.4.5}$$

with

$$\operatorname{supp} f_\phi = \phi^{-1}(\operatorname{supp} f). \tag{6.4.6}$$

Let F_ϕ be the real-valued function on $\mathbf{R}^n$ defined by

$$\begin{aligned} F_\phi(x) &= f(\phi(x))\,|\det \phi'(x)| && \text{when } x \in W \\ &= 0 && \text{when } x \in \mathbf{R}^n \setminus W. \end{aligned} \tag{6.4.7}$$

One can verify that

$$F_\phi \text{ is continuous on } \mathbf{R}^n, \tag{6.4.8}$$

with

$$\operatorname{supp} F_\phi = \operatorname{supp} f_\phi = \phi^{-1}(\operatorname{supp} f). \tag{6.4.9}$$

Under these conditions, Theorem 10.9 on p252 of [155] states that

$$\int_{\mathbf{R}^n} f(y)\,dy = \int_{\mathbf{R}^n} F_\phi(x)\,dx, \tag{6.4.10}$$

where both sides of the equation are defined as in Subsection 5.13.2.

Part of the proof of (6.4.10) is to use partitions of unity to reduce to the case of

(6.4.11) functions f supported in sufficiently small balls in $\mathbf{R}^n$.

In that case, one can use another result on the local behavior of C^1 diffeomorphisms between open subsets of $\mathbf{R}^n$, in Theorem 10.7 on p249 of [155].

Of course, the change of variables formula is simpler when $n = 1$. One can use this to obtain the change of variables formula for any n when the diffeomorphism is "primitive" in the sense of Definition 10.5 on p248 of [155], as mentioned on p252 of [155]. A linear mapping from $\mathbf{R}^n$ onto itself is called a "flip" when it interchanges two coordinates of a point in $\mathbf{R}^n$ and does not change the other coordinates, as in Definition 10.6 on p249 of [155]. It is easy to see that the change of variables formula holds for these mappings, as mentioned on p252 of [155].

Theorem 10.7 on p249 of [155] says that locally,

(6.4.12) a C^1 diffeomorphism can be expressed as a composition of primitive mappings, flips, and a translation.

This is used to get that the change of variables formula holds for functions supported in sufficiently small balls, as before. Thus one would like to express f as a sum of finitely many continuous functions supported in sufficiently small balls of this type. This can obtained using a partition of unity, as in (6.1.13). This also uses the compactness of the support of f, to cover f by finitely many balls that are contained in balls that are a bit larger and sufficiently small too.

6.5 Some remarks about linear independence

Let n and m be positive integers with $n < m$, and let $u_1, \ldots, u_m$ be the standard basis vectors in $\mathbf{R}^m$. Suppose that

(6.5.1) $v_1, \ldots, v_n$ are n linearly independent vectors in $\mathbf{R}^m$.

Let L be the linear span of $v_1, \ldots, v_n$ in $\mathbf{R}^m$, so that

$$\dim L = n, \tag{6.5.2}$$

as a vector space over the real numbers.

Under these conditions, it is well known that there is a set $I_2 \subseteq \{1, \ldots, m\}$ with exactly $m - n$ elements such that

$$v_1, \ldots, v_n \text{ together with } u_k,\ k \in I_2, \text{ is a basis for } \mathbf{R}^m. \tag{6.5.3}$$

One way to see this is to take I_2 to be a maximal subset of $\{1, \ldots, m\}$ so that this collection of vectors is linearly independent. Alternatively, one can take I_2 to be a minimal subset of $\{1, \ldots, m\}$ so that the linear span of this collection of vectors is equal to $\mathbf{R}^m$.

Put

$$I_1 = \{1, \ldots, m\} \setminus I_2, \tag{6.5.4}$$

so that I_1 is a subset of $\{1, \ldots, m\}$ with exactly n elements. Thus

$$I_1 = \{l_1, \ldots, l_n\}, \tag{6.5.5}$$

where

$$1 \le l_1 < \cdots < l_n \le m. \tag{6.5.6}$$

Let P_1 be the linear mapping from $\mathbf{R}^m$ onto $\mathbf{R}^n$ such that for each $y \in \mathbf{R}^m$,

$$\text{the } j\text{th coordinate of } P_1(y) \text{ is equal to the } l_j\text{-th coordinate of } y \tag{6.5.7}$$

for each $j = 1, \ldots, n$. Note that

$$P_1(u_k) = 0 \text{ when } k \in I_2, \tag{6.5.8}$$

by construction. More precisely, the kernel of P_1 is spanned by the u_k's, $k \in I_2$.

Let $e_1, \ldots, e_n$ be the standard basis vectors in $\mathbf{R}^n$. Remember that there is a unique linear mapping T from $\mathbf{R}^n$ into $\mathbf{R}^m$ such that

$$T(e_j) = v_j \tag{6.5.9}$$

for each $j = 1, \ldots, n$. Of course,

$$L = T(\mathbf{R}^n), \tag{6.5.10}$$

by construction. The condition that the v_j's be linearly independent in $\mathbf{R}^m$ says exactly that

$$T \text{ is one-to-one on } \mathbf{R}^n. \tag{6.5.11}$$

Observe that

$$P_1(T(\mathbf{R}^n)) = P_1(L) = P_1(\mathbf{R}^m) = \mathbf{R}^n, \tag{6.5.12}$$

using (6.5.3) and (6.5.8) in the second step. This implies that

$$P_1 \circ T \text{ is one-to-one on } \mathbf{R}^n, \tag{6.5.13}$$

because $P_1 \circ T$ is a linear mapping from $\mathbf{R}^n$ into itself.

6.5.1 Considering L as a graph

Equivalently,

(6.5.14) $$P_1 \text{ is one-to-one on } L.$$

This means that if $v \in L$, then

(6.5.15) the coordinates of v corresponding to elements of I_2
are uniquely determined by
the coordinates of v corresponding to elements of I_1.

It follows that

(6.5.16) L corresponds to the graph of a linear mapping
from $\mathbf{R}^n$ into $\mathbf{R}^{m-n}$,

by arranging the coordinates appropriately.

More precisely, put

(6.5.17) $$R = T \circ (P_1 \circ T)^{-1},$$

where $(P_1 \circ T)^{-1}$ is the inverse of $P_1 \circ T$, as a one-to-one linear mapping from $\mathbf{R}^n$ onto itself. Thus R is a one-to-one linear mapping from $\mathbf{R}^n$ into $\mathbf{R}^m$, and

(6.5.18) $$R(\mathbf{R}^n) = T(\mathbf{R}^n) = L.$$

By construction, we also have that

(6.5.19) $$P_1 \circ R = P_1 \circ T \circ (P_1 \circ T)^{-1} = I_{\mathbf{R}^n},$$

where $I_{\mathbf{R}^n}$ is the identity mapping on $\mathbf{R}^n$. This basically means that L corresponds to the graph of a mapping from $\mathbf{R}^n$ into $\mathbf{R}^{m-n}$ obtained from the components of R corresponding to elements of I_2, with the coordinates arranged appropriately.

6.6 Some nice local graphs

Let us continue with the same notation and hypotheses as in the previous section. Also let W be an open set in $\mathbf{R}^n$, and let F be a continuously-differentiable mapping from W into $\mathbf{R}^m$. Suppose that $w \in W$ and

(6.6.1) $$dF_w \text{ is one-to-one on } \mathbf{R}^n,$$

as a linear mapping from $\mathbf{R}^n$ into $\mathbf{R}^m$. If we put

(6.6.2) $$v_j = dF_w(e_j)$$

for each $j = 1, \ldots, n$, then (6.5.1) holds, because of (6.6.1).

This leads to subsets I_1, I_2 of $\{1, \ldots, m\}$ as before, as well as the linear mapping P_1 from $\mathbf{R}^m$ onto $\mathbf{R}^n$. We also have that

(6.6.3) $P_1 \circ dF_w$ is a one-to-one mapping from $\mathbf{R}^n$ onto itself,

as in (6.5.12) and (6.5.13). Under these conditions, we would like to use the inverse function theorem to say more about F near w.

Put

$$f = P_1 \circ F, \tag{6.6.4}$$

which is a continuously-differentiable mapping from W into $\mathbf{R}^n$. Note that

$$df_w = P_1 \circ dF_w, \tag{6.6.5}$$

which is an easy instance of the chain rule that can be seen more directly. It follows that

$$df_w \in GL(\mathbf{R}^n), \tag{6.6.6}$$

by (6.6.3). The inverse function theorem implies that

(6.6.7) f has a local inverse near w that is continuously differentiable,

as in Section 5.6.

6.6.1 Using the local inverse

We would like to use this to get that near w,

(6.6.8) F takes values in the graph of a continuously-differentiable function defined on a neighborhood of $f(w)$ in $\mathbf{R}^n$ with values in $\mathbf{R}^{m-n}$,

with the coordinates arranged appropriately. More precisely, the inverse function theorem implies that there are open subsets $U(w)$ and $V(w)$ of $\mathbf{R}^n$ such that $w \in U(w)$, $U(w) \subseteq W$, $f(w) \in V(w)$, and

(6.6.9) the restriction of f to $U(w)$ is a one-to-one mapping onto $V(w)$.

In particular, the injectivity of f on $U(w)$ implies that

(6.6.10) F is one-to-one on $U(w)$,

by the definition (6.6.4) of f.

The inverse $g = g_w$ of the restriction of f to $U(w)$ is continuously differentiable on $V(w)$, as before. It follows that

(6.6.11) $F \circ g$ is continuously differentiable

as a mapping from $V(w)$ into $\mathbf{R}^m$. Observe that

(6.6.12) $P_1 \circ F \circ g = f \circ g$ is the identity mapping on $V(w)$, by construction. This implies that

$$F(U(w)) = (F \circ g)(V(w)) \tag{6.6.13}$$

corresponds to the graph of a continuously-differentiable function on $V(w)$ with values in $\mathbf{R}^{m-n}$, with the coordinates arranged appropriately.

If F is r-times continuously differentiable on W for some positive integer r, then f is r-times continuously differentiable on W, and

$$g \text{ is } r\text{-times continuously differentiable on } V(w) \tag{6.6.14}$$

too, as in Subsection 5.12.1. This implies that

$$F \circ g \text{ is } r\text{-times continuously differentiable on } V(w) \tag{6.6.15}$$

as well, as in Subsection 3.9.1.

Similarly if F is infinitely differentiable on W, then f is infinitely differentiable on W, and

$$g \text{ is infinitely differentiable on } V(w). \tag{6.6.16}$$

This means that

$$F \circ g \text{ is infinitely differentiable on } V(w) \tag{6.6.17}$$

too.

6.7 Injective linear mappings

Let n and m be positive integers with $n \le m$, and let $e_1, \ldots, e_n$ be the standard basis vectors in $\mathbf{R}^n$, as usual. Remember that a linear mapping T from $\mathbf{R}^n$ into $\mathbf{R}^m$ is one-to-one on $\mathbf{R}^n$ if and only if

$$T(e_1), \ldots, T(e_n) \text{ are linearly independent in } \mathbf{R}^m. \tag{6.7.1}$$

In this case, there is a linear mapping P_1 from $\mathbf{R}^m$ onto $\mathbf{R}^n$ such that

$$P_1 \circ T \in GL(\mathbf{R}^n), \tag{6.7.2}$$

as in Section 6.5. More precisely, if $m = n$, then one can simply take P_1 to the the identity mapping on $\mathbf{R}^n$. Conversely, if (6.7.2) holds for any linear mapping P_1 from $\mathbf{R}^m$ into $\mathbf{R}^n$, then T is one-to-one on $\mathbf{R}^n$.

If P_1 is any linear mapping from $\mathbf{R}^m$ into $\mathbf{R}^n$, then

$$T \mapsto P_1 \circ T \tag{6.7.3}$$

defines a linear mapping from the space $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$ of linear mappings from $\mathbf{R}^n$ into $\mathbf{R}^m$ into the space $\mathcal{L}(\mathbf{R}^n)$ of linear mappings from $\mathcal{L}(\mathbf{R}^n)$ of linear mappings from $\mathbf{R}^n$ into itself. This corresponds to a linear mapping from the space $M_{m,n}(\mathbf{R})$ of $m \times n$ matrices of real numbers into $M_{n,n}(\mathbf{R})$, as in Section 5.3. This may be identified with a linear mapping from $\mathbf{R}^{m\,n}$ into $\mathbf{R}^{n^2}$, as in Section 5.1. In particular, this mapping is continuous with respect to the standard Euclidean metrics on $\mathbf{R}^{m\,n}$ and $\mathbf{R}^{n^2}$.

One can use this to get that

$$\{T \in \mathcal{L}(\mathbf{R}^n, \mathbf{R}^m) : P_1 \circ T \in GL(\mathbf{R}^n)\} \tag{6.7.4}$$

is an open set in $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$, in the sense that it corresponds to an open set in $\mathbf{R}^{m\,n}$. This uses the fact that $GL(\mathbf{R}^n)$ corresponds to an open set in $\mathbf{R}^{n^2}$, as in Subsection 5.2.1 and Section 5.4. Note that (6.7.2) implies that

$$P_1(\mathbf{R}^m) = \mathbf{R}^n. \tag{6.7.5}$$

This means that (6.7.4) is the empty set unless (6.7.5) holds.

Of course, (6.7.4) is the same as

$$\{T \in \mathcal{L}(\mathbf{R}^n, \mathbf{R}^m) : \det(P_1 \circ T) \neq 0\}. \tag{6.7.6}$$

It is easy to see that

$$\det(P_1 \circ T) \tag{6.7.7}$$

is given by a polynomial in the entries of the matrix associated to T, with P_1 fixed. In particular, this corresponds to a continuous real-valued function on $\mathbf{R}^{m\,n}$, so that (6.7.6) corresponds to an open set in $\mathbf{R}^{m\,n}$, as before.

It follows that

$$\{T \in \mathcal{L}(\mathbf{R}^n, \mathbf{R}^m) : T \text{ is one-to-one on } \mathbf{R}^n\} \tag{6.7.8}$$

is an open set in $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$ too. More precisely, (6.7.8) is the same as the union of (6.7.4) over all linear mappings P_1 from $\mathbf{R}^m$ into $\mathbf{R}^n$. In fact, one can restrict one's attention to linear mappings P_1 that satisfy (6.7.5), because otherwise (6.7.4) is the empty set, as before. The fact that (6.7.8) is an open set in $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^n)$ could also be obtained from an argument analogous to one in Section 5.4. Of course, if $m = n$, then (6.7.8) is the same as $GL(\mathbf{R}^n)$.

6.8 Embeddings and immersions

Let n and m be positive integers with $n \leq m$ again, and let W be a nonempty open subset of $\mathbf{R}^n$. Also let F be a continuously-differentiable mapping from W into $\mathbf{R}^m$, so that the differential of F is continuous as a function on W with values in $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$, as in Section 5.5. One can check that

$$\{x \in W : F'(x) = dF_x \text{ is one-to-one on } \mathbf{R}^n\} \tag{6.8.1}$$

is an open set in $\mathbf{R}^n$, using the fact that (6.7.8) is an open set in $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$. This is the same as a remark in Section 5.5 when $m = n$.

Let us say that F is an *immersion* of W into $\mathbf{R}^m$ if

(6.8.2) $$F'(x) = dF_x \text{ is one-to-one on } \mathbf{R}^n \text{ for every } x \in W.$$

This corresponds to part (a) of Definition 1.27 on p22 of [184]. This implies that

(6.8.3) $$F \text{ is one-to-one on a neighborhood of every point in } W,$$

as in Subsection 6.6.1. If F is r-times continuously differentiable on W for some positive integer r, or infinitely differentiable on W, then one may say that F is a C^r or C^∞ immersion, as appropriate.

Suppose that F is an immersion of W into $\mathbf{R}^m$. If we also have that

(6.8.4) $$F \text{ is one-to-one on } W,$$

then (W, F) may be called a *submanifold* of $\mathbf{R}^m$, as in part (b) of Definition 1.27 on p22 of [184]. In this case, if $x \in W$, then the *tangent space* of this submanifold at $x \in W$ is the linear subspace

(6.8.5) $$dF_x(\mathbf{R}^n)$$

of $\mathbf{R}^m$. Note that

(6.8.6) $$\dim dF_x(\mathbf{R}^n) = n,$$

because of (6.8.2). If F is r-times continuously differentiable on W, or infinitely differentiable on W, then (W, F) may be called a C^r or C^∞ submanifold of $\mathbf{R}^m$, as appropriate.

If

(6.8.7) $$F \text{ is a homeomorphism from } W \text{ onto } F(W),$$

as in Subsection 5.9.1, then F may be called an *imbedding* of W into $\mathbf{R}^m$, as in part (c) of Definition 1.27 on p22 of [184]. However, some colleagues have mentioned that the term *embedding* is sometimes used for one-to-one immersions. In order to be precise, we may say that F is an *embedding in the strong sense* when (6.8.7) holds. As before, if F is r-times continuously differentiable on W, or infinitely differentiable on W, then one may say that F is a C^r or C^∞ embedding of W into $\mathbf{R}^m$, as appropriate.

6.8.1 Some remarks about manifolds

One can consider broader notions of *manifolds*, but we shall not pursue this in detail here. One can start with n-dimensional topological manifolds, which are topological spaces that are locally homeomorphic to open sets in $\mathbf{R}^n$, and which are often asked to satisfy some additional conditions. If r is a positive integer, then one can consider n-dimensional C^r manifolds, which are locally C^r diffeomorphic to open sets in $\mathbf{R}^n$ in a suitable sense. Similarly, one can consider n-dimensional C^∞ or smooth manifolds, which are locally C^∞ diffeomorphic to open sets in $\mathbf{R}^n$ in an analogous sense. These topics are discussed beginning on p5 of [184], for instance.

The notions of immersions and submanifolds can be extended to mappings between C^r and C^∞ manifolds, as on p22 of [184]. In particular, broader notions of manifolds are already helpful for considering submanifolds of Euclidean spaces.

Basically, a (reasonably nicely embedded) C^r or C^∞ n-dimensional submanifold of $\mathbf{R}^m$ should look locally like a C^r or C^∞ submanifold in the previous sense at each point, as appropriate. Equivalently, this means that it looks locally like the graph of a C^r or C^∞ mapping from an open set in $\mathbf{R}^n$ into $\mathbf{R}^{m-n}$, with the coordinates of $\mathbf{R}^m$ arranged in a suitable way, as in Subsection 6.6.1.

The unit sphere in $\mathbf{R}^{n+1}$ with respect to the standard Euclidean metric is the same as

$$S^n = \Big\{ y \in \mathbf{R}^{n+1} : \sum_{j=1}^{n+1} y_j^2 = 1 \Big\}. \tag{6.8.8}$$

This is a nice example of a smooth n-dimensional manifold, which is a submanifold of $\mathbf{R}^{n+1}$.

6.9 Surjective linear mappings

Let n and m be positive integers with $n < m$, and let $e_1, \ldots, e_n$ be the standard basis for $\mathbf{R}^n$, as usual. Also let T be a linear mapping from $\mathbf{R}^m$ into $\mathbf{R}^n$. If

$$T(\mathbf{R}^m) = \mathbf{R}^n, \tag{6.9.1}$$

then for each $j = 1, \ldots, n$ there is a $v_j \in \mathbf{R}^m$ such that

$$T(v_j) = e_j. \tag{6.9.2}$$

Conversely, this condition implies (6.9.1).

If $v_1, \ldots, v_n$ are any elements of $\mathbf{R}^m$, then there is a unique linear mapping R_0 from $\mathbf{R}^n$ into $\mathbf{R}^m$ such that

$$R_0(e_j) = v_j \tag{6.9.3}$$

for each $j = 1, \ldots, n$. In this case, (6.9.2) is the same as saying that

$$T(R_0(e_j)) = e_j. \tag{6.9.4}$$

It is easy to see that this holds if and only if

$$T \circ R_0 \text{ is the identity mapping on } \mathbf{R}^n. \tag{6.9.5}$$

Let R_1 be a linear mapping from $\mathbf{R}^n$ into $\mathbf{R}^m$. If

$$T(R_1(\mathbf{R}^n)) = \mathbf{R}^n, \tag{6.9.6}$$

then it is easy to see that (6.9.1) holds. Note that (6.9.6) holds if and only if

$$T \circ R_1 \in GL(\mathbf{R}^n), \tag{6.9.7}$$

because $T \circ R_1$ is a linear mapping from $\mathbf{R}^n$ into itself. This means that $T \circ R_1$ is one-to-one on $\mathbf{R}^n$, which implies that

$$R_1 \text{ is one-to-one on } \mathbf{R}^n. \tag{6.9.8}$$

Of course,

$$T \mapsto T \circ R_1 \tag{6.9.9}$$

defines a linear mapping from $\mathcal{L}(\mathbf{R}^m, \mathbf{R}^n)$ into $\mathcal{L}(\mathbf{R}^n)$, which corresponds to a linear mapping from $M_{n,m}(\mathbf{R})$ into $M_{n,n}(\mathbf{R})$, as in Section 5.3. This may be identified with a linear mapping from $\mathbf{R}^{n\,m}$ into $\mathbf{R}^{n^2}$, as in Section 5.1, which is continuous in particular, with respect to the standard Euclidean metrics on $\mathbf{R}^{n\,m}$ and $\mathbf{R}^{n^2}$.

It follows that

$$\{T \in \mathcal{L}(\mathbf{R}^m, \mathbf{R}^n) : T \circ R_1 \in GL(\mathbf{R}^n)\} \tag{6.9.10}$$

is an open set in $\mathcal{L}(\mathbf{R}^m, \mathbf{R}^n)$, in the sense that it corresponds to an open set in $\mathbf{R}^{n\,m}$, because $GL(\mathbf{R}^n)$ corresponds to an open set in $\mathbf{R}^{n^2}$, as in Subsection 5.2.1 and Section 5.4. This is the empty set unless (6.9.8) holds, as before.

Note that (6.9.10) is the same as

$$\{T \in \mathcal{L}(\mathbf{R}^m, \mathbf{R}^n) : \det(T \circ R_1) \neq 0\}. \tag{6.9.11}$$

One can check that

$$\det(T \circ R_1) \tag{6.9.12}$$

is a polynomial in the entries of the matrix associated to T, with R_1 fixed. This can be used as another way to get that (6.9.11) corresponds to an open set in $\mathbf{R}^{n\,m}$.

Let us check that

$$\{T \in \mathcal{L}(\mathbf{R}^m, \mathbf{R}^n) : T(\mathbf{R}^m) = \mathbf{R}^n\} \tag{6.9.13}$$

is an open set in $\mathcal{L}(\mathbf{R}^m, \mathbf{R}^n)$ as well. In fact, (6.9.13) is the same as the union of (6.9.10) over all linear mappings R_1 from $\mathbf{R}^n$ into $\mathbf{R}^m$. More precisely, we may restrict our attention to linear mappings R_1 that satisfy (6.9.8), because (6.9.10) is the empty set otherwise, as before.

6.10 The implicit function theorem

Let n and m be positive integers with $n < m$ again, and let W be a nonempty open subset of $\mathbf{R}^m$. Also let ϕ be a continuously-differentiable mapping from W into $\mathbf{R}^n$, so that the differential of ϕ is continuous as a mapping from W into $\mathcal{L}(\mathbf{R}^m, \mathbf{R}^n)$, as in Section 5.5. Observe that

$$\{x \in W : d\phi_x(\mathbf{R}^m) = \mathbf{R}^n\} \tag{6.10.1}$$

is an open set in $\mathbf{R}^m$, because (6.9.13) is an open set in $\mathcal{L}(\mathbf{R}^m, \mathbf{R}^n)$.

Suppose that $w \in W$ and

$$d\phi_w(\mathbf{R}^m) = \mathbf{R}^n. \tag{6.10.2}$$

We would like to add some components to ϕ to get a mapping Φ from W into $\mathbf{R}^m$ with

$$d\Phi_w \in GL(\mathbf{R}^m), \tag{6.10.3}$$

so that we can use the inverse function theorem. Note that

$$\dim(\ker d\phi_w) = m - n, \tag{6.10.4}$$

because of (6.10.2) and standard arguments from linear algebra. This implies that there is a linear mapping P_0 from $\mathbf{R}^m$ onto $\mathbf{R}^{m-n}$ such that

$$P_0 \text{ is one-to-one on } \ker d\phi_w, \tag{6.10.5}$$

as in Section 6.5. More precisely, we can take P_0 so that its components are the same as $m - n$ of the coordinates of an element of $\mathbf{R}^m$, as before.

We would like to define Φ so that its first n components are the same as the components of ϕ, and the other $m-n$ components are equal to the components of P_0. In this case, the first n components of $d\Phi_w$ are the same as the components of $d\phi_w$, and the other $m - n$ components of $d\Phi_w$ are equal to the components of P_0, so that

$$\ker d\Phi_w = (\ker d\phi_w) \cap (\ker P_0) = \{0\}, \tag{6.10.6}$$

using (6.10.5) in the second step. Thus (6.10.3) holds, so that we can use the inverse function theorem, as in Section 5.6. This means that there are open sets $U(w)$, $V(w)$ in $\mathbf{R}^m$ such that

$$w \in U(w),\ U(w) \subseteq W,\ \Phi(w) \in V(w), \tag{6.10.7}$$

and

(6.10.8) the restriction of Φ to $U(w)$ is a C^1 diffeomorphism onto $V(w)$,

as before. Let $\Psi = \Psi_w$ be the inverse of the restriction of Φ to $U(w)$, which is a C^1 diffeomorphism from $V(w)$ onto $U(w)$.

6.10.1 The zero set of ϕ

Suppose that

$$\phi(w) = 0, \tag{6.10.9}$$

and consider

$$U_0(w) = \{x \in U(w) : \phi(x) = 0\}. \tag{6.10.10}$$

Put

$$V_0(w) = \Phi(U_0(w)), \tag{6.10.11}$$

which is the same as the set of points in $V(w)$ whose first n coordinates are equal to 0. Of course,

$$\Phi(w) \in V_0(w), \tag{6.10.12}$$

because $w \in U_0(w)$, by (6.10.9). We can identify $V_0(w)$ with an open set in $\mathbf{R}^{m-n}$, using the last $m-n$ coordinates of the elements of $V_0(w)$. Note that

$$\Psi(V_0(w)) = U_0(w), \tag{6.10.13}$$

by construction.

This shows that $U_0(w)$ is an $(m-n)$-dimensional C^1 submanifold of $\mathbf{R}^m$. This is a version of the *implicit function theorem*. If $x \in U_0(w)$, then one can check that the tangent space to $U_0(w)$ at x is equal to

$$\ker d\phi_x. \tag{6.10.14}$$

This uses the fact that

$$d\Psi_{\Phi(x)} = (d\Phi_x)^{-1}, \tag{6.10.15}$$

as in Section 5.6.

If ϕ is r-times continuously differentiable on W for some positive integer r, then

(6.10.16) the restriction of Φ to $U(w)$ is a C^r diffeomorphism onto $V(w)$,

Ψ is a C^r diffeomorphism from $V(w)$ onto $U(w)$, and $U_0(w)$ is a C^r submanifold of $\mathbf{R}^m$. Similarly, if ϕ is infinitely differentiable on W, then

(6.10.17) the restriction of Φ to $U(w)$ is a C^∞ diffeomorphism onto $V(w)$,

Ψ is a C^∞ diffeomorphism from $V(w)$ onto $U(W)$, and $U_0(w)$ is a C^∞ submanifold of $\mathbf{R}^m$.

Put

$$W_0 = \{x \in W : \phi(x) = 0\}, \tag{6.10.18}$$

and suppose that

$$d\phi_x(\mathbf{R}^m) = \mathbf{R}^n \tag{6.10.19}$$

for every $x \in W_0$. In this case, the previous remarks imply that W_0 is an $(m-n)$-dimensional C^1 submanifold of $\mathbf{R}^m$. If ϕ is r-times continuously differentiable on W, or infinitely differentiable on W, then W_0 is a C^r or C^∞ submanifold of $\mathbf{R}^m$, as appropriate. This corresponds to Theorem 1.38 on p31 of [184].

6.11 Injectivity and alternating forms

Let W be a vector space over the real numbers, and let W_0 be a linear subspace of W. If μ is a k-linear form on W for some positive integer k, then let μ_0 be the k-linear form on W_0 obtained by restricting μ to elements of W_0. More precisely,

(6.11.1) μ_0 is the restriction of μ to the set $(W_0)^k$

of k-tuples of elements of W_0, considered as a subset of the set W^k of k-tuples of elements of W. Of course, if μ is alternating or symmetric on W, then μ_0 has the same property on W_0.

Suppose for the moment that

$$\dim W_0 < k. \tag{6.11.2}$$

If μ is an alternating k-linear form on W, then μ_0 is an alternating k-linear form on W_0, and thus

$$\mu_0 = 0, \tag{6.11.3}$$

as in Subsections 1.11.1 and 2.5.1.

Let ι_0 be the obvious inclusion mapping of W_0 into W, which sends every element of W_0 to itself, and is considered as a linear mapping from W_0 into W. Note that $\iota_0^*(\mu)$ may be defined as a k-linear form on W_0, as in Section 2.3. In this case,

$$\iota_0^*(\mu) = \mu_0. \tag{6.11.4}$$

Let V be another vector space over the real numbers, and let T be a linear mapping from V into W. Suppose that

$$T(V) \subseteq W_0, \tag{6.11.5}$$

and let T_0 be the same as T, considered as a linear mapping from V into W_0. Observe that

$$T = \iota_0 \circ T_0, \tag{6.11.6}$$

as a linear mapping from V into W. If $T^*(\mu)$ is the k-linear form on V as in Section 2.3, then we get that

$$T^*(\mu) = (\iota_0 \circ T_0)^*(\mu) = T_0^*(\iota_0^*(\mu)), \tag{6.11.7}$$

where the second step is as in Section 2.3. This means that

$$T^*(\mu) = T_0^*(\mu_0), \tag{6.11.8}$$

because of (6.11.4).

Of course, we can simply take $W_0 = T(V)$. Suppose for the moment that

$$\dim V = k. \tag{6.11.9}$$

If

$$T \text{ is not one-to-one on } V, \tag{6.11.10}$$

then

$$\dim T(V) < k. \tag{6.11.11}$$

If μ is an alternating k-linear form on W, then we get that

$$T^*(\mu) = 0, \tag{6.11.12}$$

by (6.11.3) and (6.11.8).

6.11.1 Injectivity of the differential

Let n and m be positive integers, and let U, V be nonempty open subsets of $\mathbf{R}^n$, $\mathbf{R}^m$, respectively. Also let ψ be a mapping from U into $\mathbf{R}^m$ with

$$\psi(U) \subseteq V, \tag{6.11.13}$$

and let α be a differential n-form on V. Suppose that ψ is differentiable on U, so that $\psi^*(\alpha)$ may be defined as a differential n-form on U, as in Section 4.4.

Let $x \in U$ be given, and let $\alpha_{\psi(x)}$ be the value of α at $\psi(x) \in V$, which is an alternating k-linear form on $\mathbf{R}^m$. Remember that the value of $\psi^*(\alpha)$ at x is the alternating k-linear form on $\mathbf{R}^n$ given by

$$(\psi^*(\alpha))_x = (d\psi_x)^*(\alpha_{\psi(x)}), \tag{6.11.14}$$

as in Section 4.4. Suppose that

$$d\psi_x \text{ is not one-to-one} \tag{6.11.15}$$

as a linear mapping from $\mathbf{R}^n$ into $\mathbf{R}^m$. Under these conditions, we get that

$$(\psi^*(\alpha))_x = 0 \tag{6.11.16}$$

as in (6.11.12).

6.12 Integrating n-forms on $\mathbf{R}^n$

Let n be a positive integer, let E be a nonempty subset of $\mathbf{R}^n$, and let α be a differential n-form on E. Remember that α may be expressed as

$$\alpha = \alpha_0\, dx_1 \wedge \cdots \wedge dx_n, \tag{6.12.1}$$

where α_0 is a real-valued function on E, as in Section 3.14.

Suppose that E is compact in $\mathbf{R}^n$, and that α is continuous on E. This means that α_0 is continuous as a real-valued function on E, as in Subsection 3.12.1. If E is reasonably nice, then the *integral of α over E* may be defined by

$$\int_E \alpha = \int_E \alpha_0(x)\, dx, \tag{6.12.2}$$

where the right side is an ordinary Riemann integral, as in (1) in Section 4.5 on p141 of [184]. In particular, we may do this when E is a cell in $\mathbf{R}^n$, as in Subsection 5.13.1, or the standard simplex in $\mathbf{R}^n$, as in Section 6.2. If one uses the Lebesgue integral, then the right side of (6.12.2) may be defned more broadly.

If β is a differential form on $\mathbf{R}^n$, then the *support*

$$\operatorname{supp} \beta \tag{6.12.3}$$

of β may be defined as in Subsection 5.13.2. This is the closure of the set of $x \in \mathbf{R}^n$ at which the value of β is not zero, as an element of $\mathcal{AM}(\mathbf{R}^n)$. Equivalentlyf, if we identify $\mathcal{AM}(\mathbf{R}^n)$ with $\mathbf{R}^{2^n}$, as in Subsection 3.12.1, then the support of β is the same as the support of the corresponding function on $\mathbf{R}^n$ with values in $\mathbf{R}^{2^n}$.

Let α be a differential n-form on $\mathbf{R}^n$ as in (6.12.1), and note that the support of α as a differential form on $\mathbf{R}^n$ is the same as the support of α_0 as a real-valued function on $\mathbf{R}^n$. If α is continuous, and the support of α is compact, then the *integral of* α *over* $\mathbf{R}^n$ may be defined by

$$\int_{\mathbf{R}^n} \alpha = \int_{\mathbf{R}^n} \alpha_0(x)\,dx, \tag{6.12.4}$$

where the right side is as in Subsection 5.13.2. If $\mathcal{C}$ is a cell in $\mathbf{R}^n$ that contains the support of α, then

$$\int_{\mathbf{R}^n} \alpha = \int_{\mathcal{C}} \alpha = \int_{\mathcal{C}} \alpha_0(x)\,dx, \tag{6.12.5}$$

as before.

6.12.1 A family of integrals

Let β be a differential $(n-1)$-form on $\mathbf{R}^n$. We may express β as

$$\beta = \sum_{j=1}^n \beta_j\,dx_1 \wedge \cdots \wedge dx_{j-1} \wedge dx_{j+1} \wedge \cdots \wedge dx_n, \tag{6.12.6}$$

as in Section 3.14, where β_j is a real-valued function on $\mathbf{R}^n$. Suppose that

$$\beta \text{ is continuously differentiable} \tag{6.12.7}$$

as a differential form on $\mathbf{R}^n$, which means that β_j is continuously differentiable as a real-valued function on $\mathbf{R}^n$ for each $j = 1, \ldots, n$, as in Section 3.13. Remember that the exterior derivative $d\beta$ of β is a continuous differential n-form on $\mathbf{R}^n$, as in Section 4.6.

Observe that

$$d\beta = \sum_{j=1}^n d\beta_j \wedge dx_1 \wedge \cdots \wedge dx_{j-1} \wedge dx_{j+1} \wedge \cdots \wedge dx_n, \tag{6.12.8}$$

by the definition of the exterior derivative. Remember that

$$d\beta_j = \sum_{l=1}^n (\partial_l \beta_j)\,dx_l \tag{6.12.9}$$

for each j, as in Section 3.14. Using this, we get that

$$d\beta = \sum_{j=1}^n (\partial_j \beta_j)\,dx_j \wedge dx_1 \wedge \cdots \wedge dx_{j-1} \wedge dx_{j+1} \wedge \cdots \wedge dx_n, \tag{6.12.10}$$

because $dx_l \wedge dx_l = 0$ for every l, as in Section 3.11. It follows that

$$d\beta = \sum_{j=1}^{n} (-1)^{j-1} \, (\partial_j \beta_j) \, dx_1 \wedge \cdots \wedge dx_n, \tag{6.12.11}$$

because $dx_j \wedge dx_l = -dx_l \wedge dx_j$ for each j, l, as before.

Suppose that

$$\beta \text{ also has compact support in } \mathbf{R}^n, \tag{6.12.12}$$

which means that β_j has compact support for each j. This implies that $\partial_j \beta_j$ has compact support for each j, so that $d\beta$ has compact support too. Observe that

$$\begin{aligned} \int_{\mathbf{R}^n} d\beta &= \int_{\mathbf{R}^n} \sum_{j=1}^{n} (-1)^{j-1} (\partial_j \beta_j)(x) \, dx \\ &= \sum_{j=1}^{n} (-1)^{j-1} \int_{\mathbf{R}^n} (\partial_j \beta_j)(x) \, dx. \end{aligned} \tag{6.12.13}$$

It is easy to see that

$$\int_{\mathbf{R}^n} (\partial_j \beta_j)(x) \, dx = 0 \tag{6.12.14}$$

for each j, by integrating in the jth variable first, and using the fundamental theorem of calculus. This means that

$$\int_{\mathbf{R}^n} d\beta = 0 \tag{6.12.15}$$

under these conditions.

6.13 More on changes of variables

Let m and n be positive integers, let W be a nonempty open subset of $\mathbf{R}^n$, and let ψ be a mapping from W into $\mathbf{R}^m$. Suppose that

$$\psi \text{ is differentiable on } W, \tag{6.13.1}$$

in the sense that ψ is differentiable at every point in W. Let E be a nonempty subset of W, and let α be a differential k-form on a subset of $\mathbf{R}^m$ that contains $\psi(E)$ for some positive integer k. Under these conditions, we can define the *pull-back*

$$\psi^*(\alpha) \tag{6.13.2}$$

of α with respect to ψ as a differential k-form on E, in essentially the same way as in Section 4.4. More precisely, if $x \in E$, and $\alpha_{\psi(x)}$ is the value of α at $\psi(x)$, as an alternating k-linear form on $\mathbf{R}^m$, then the value of $\psi^*(\alpha)$ at x, as an alternating k-linear form on $\mathbf{R}^n$, is defined by

$$(\psi^*(\alpha))_x = (d\psi_x)^*(\alpha_{\psi(x)}), \tag{6.13.3}$$

where the right side is as in Section 2.3.

Suppose now that $k = m = n$, and that α is as in (6.12.1) for some real-valued function α_0 defined on a subset of $\mathbf{R}^n$ that contains $\psi(E)$. In this case,

$$\psi^*(\alpha) = (\alpha_0 \circ \psi)\, \psi^*(dx_1 \wedge \cdots \wedge dx_n) = (\alpha_0 \circ \psi)\, d\psi_1 \wedge \cdots \wedge d\psi_n \tag{6.13.4}$$

on E, where the second step is as in Section 4.5. Remember that the components $\psi_1, \ldots, \psi_n$ of ψ are differentiable as real-valued functions on W.

The differential n-form

$$dx_1 \wedge \cdots \wedge dx_n \tag{6.13.5}$$

is equal to

$$\theta_1 \wedge \cdots \wedge \theta_n \tag{6.13.6}$$

at every point in $\mathbf{R}^n$, as an alternating n-linear form on $\mathbf{R}^n$, as in Section 3.14. Remember that $\mu_{\det}$ is the alternating n-linear form on $\mathbf{R}^n$ corresponding to the determinant, as in Section 1.15, and that this is equal to (6.13.6), as in Section 2.11. It follows that

$$\psi^*(dx_1 \wedge \cdots \wedge dx_n) = (\det d\psi)\, dx_1 \wedge \cdots \wedge dx_n \tag{6.13.7}$$

on W, as in Section 2.5. This means that

$$\psi^*(\alpha) = (\alpha_0 \circ \psi)\, (\det d\psi)\, dx_1 \wedge \cdots \wedge dx_n \tag{6.13.8}$$

on E, by (6.13.4).

6.13.1 Integrals and diffeomorphisms

Suppose that $\psi(W)$ is an open set in $\mathbf{R}^n$ too, and that

$$\psi \text{ is a } C^1 \text{ diffeomorphism from } W \text{ onto } \psi(W). \tag{6.13.9}$$

Suppose also that E is compact as a subset of $\mathbf{R}^n$, and that α is continuous on $\psi(E)$, so that α_0 is continuous as a real-valued function on $\psi(E)$. This implies that $\alpha_0 \circ \psi$ is continuous on E, so that

$$\psi^*(\alpha) \text{ is continuous on } E, \tag{6.13.10}$$

because of (6.13.8). Remember that

$$\det d\psi \text{ is continuous} \tag{6.13.11}$$

as a real-valued function on W, because ψ is continuously differentiable on W, as in Section 6.3.

If we also have that

$$\det \psi'(x) = \det d\psi_x > 0 \tag{6.13.12}$$

for every $x \in E$, then

$$\int_{\psi(E)} \alpha = \int_E \psi^*(\alpha), \tag{6.13.13}$$

at least if E is nice enough for the Riemann integral to be defined. This follows from (6.13.8) and the change of variables formula in Section 6.3. Similarly, if

$$\det \psi'(x) = \det d\psi_x < 0 \tag{6.13.14}$$

for every $x \in E$, then

$$\int_{\psi(E)} \alpha = -\int_E \psi^*(\alpha), \tag{6.13.15}$$

if E is nice enought for the Riemann integral to be defined. This corresponds to (2) in Section 4.5 on p141 of [184]. As before, there are versions of this using Lebesgue measure and integration with milder hypotheses.

6.14 Diffeomorphisms and orientations

Let us continue with the same notation and hypotheses as in the previous section. Remember that

$$d\psi_x \text{ is invertible as a linear mapping from } \mathbf{R}^n \text{ into itself} \tag{6.14.1}$$

for every $x \in W$, because ψ is a C^1 diffeomorphism, as in Subsection 5.11.1. Equivalently, this means that

$$\det d\psi_x \neq 0 \tag{6.14.2}$$

for every $x \in W$. This implies that

$$\operatorname{sign}(\det d\psi_x) \text{ is locally constant on } W, \tag{6.14.3}$$

as in Subsection 4.10.2, because of (6.13.11). If

$$W \text{ is connected,} \tag{6.14.4}$$

then it follows that

$$\operatorname{sign}(\det d\psi_x) \text{ is constant on } W, \tag{6.14.5}$$

as before.

We say that

$$\psi \text{ is orientation-preserving on } W \tag{6.14.6}$$

if (6.13.12) holds for every $x \in W$. Similarly,

$$\psi \text{ is orientation-reversing on } W \tag{6.14.7}$$

if (6.13.14) holds for every $x \in W$.

6.14.1 Another version

The version of the change of variables formula in Section 6.4 may be restated in terms of differential forms as well. Let α be a differential n-form on $\mathbf{R}^n$ that is continuous and has compact support, with

$$\operatorname{supp}\alpha \subseteq \psi(W). \tag{6.14.8}$$

Thus α may be expressed as in (6.12.1), where α_0 is a continuous real-valued function on $\mathbf{R}^n$ with compact support, and

$$\operatorname{supp}\alpha_0 = \operatorname{supp}\alpha \subseteq \psi(W). \tag{6.14.9}$$

It follows that

$$\psi^{-1}(\operatorname{supp}\alpha) = \psi^{-1}(\operatorname{supp}\alpha_0) \text{ is a compact set in } \mathbf{R}^n, \tag{6.14.10}$$

because ψ^{-1} is continuous as a mapping from $\psi(W)$ into $\mathbf{R}^n$, as in Subsection 5.8.1. Note that

$$\psi^{-1}(\operatorname{supp}\alpha) = \psi^{-1}(\operatorname{supp}\alpha_0) \subseteq W, \tag{6.14.11}$$

because $\psi^{-1}(\psi(W)) = W$.

Let β_0 be the real-valued function defined on $\mathbf{R}^n$ by

$$\begin{aligned} \beta_0(x) &= \alpha_0(\psi(x)) \det d\psi_x && \text{when } x \in W \\ &= 0 && \text{when } x \in \mathbf{R}^n \setminus W. \end{aligned} \tag{6.14.12}$$

One can check that

$$\beta_0 \text{ is continuous on } \mathbf{R}^n, \tag{6.14.13}$$

as in Section 6.4. We also have that

$$\operatorname{supp}\beta_0 = \psi^{-1}(\operatorname{supp}\alpha_0) = \psi^{-1}(\operatorname{supp}\alpha), \tag{6.14.14}$$

as before. Put

$$\beta = \beta_0\, dx_1 \wedge \cdots \wedge dx_n, \tag{6.14.15}$$

which is a differential n-form on $\mathbf{R}^n$ that is continuous, with

$$\operatorname{supp}\beta = \operatorname{supp}\beta_0. \tag{6.14.16}$$

Note that

$$\beta = \psi^*(\alpha) \text{ on } W, \tag{6.14.17}$$

by (6.13.8). It is easy to see that

$$\int_{\mathbf{R}^n} \alpha = \int_{\mathbf{R}^n} \beta \tag{6.14.18}$$

when ψ is orientation-preserving on W, using the remarks in Section 6.4. Similarly,

$$\int_{\mathbf{R}^n} \alpha = -\int_{\mathbf{R}^n} \beta \tag{6.14.19}$$

when ψ is orientation-reversing on W.

6.15 n-Surfaces in $\mathbf{R}^m$

Let m and n be positive integers, let U be a nonempty open subset of $\mathbf{R}^n$, and let ϕ be a continuously-differentiable mapping from U into $\mathbf{R}^m$. Also let K be a nonempty compact subset of U, which is reasonably nice, so that one can take Riemann integrals of continuous real-valued functions on K. One may take K to be a cell in $\mathbf{R}^n$ that is contained in U, for instance, or the standard simplex in $\mathbf{R}^n$, if that is contained in U. One could consider wider classes of sets K using Lebesgue measure and integration, as usual.

Let V be a nonempty open subset of $\mathbf{R}^m$. If

$$\phi(K) \subseteq V, \tag{6.15.1}$$

then the restriction of ϕ to K is considered to be

(6.15.2) an *n-surface* in V,

as a mapping from K into V. This corresponds to Definition 10.10 on p254 of [155]. We may use $\Phi = \Phi(K)$ to refer to this n-surface, which involves both K and the restriction of ϕ to K. The set K may be called the *parameter domain* of the n-surface, as in [155].

More precisely, we shall also use the differential of ϕ at points in K. If K is reasonably nice, then the differential of ϕ at points in K is uniquely determined by the restriction of ϕ to K. In particular, this holds when K is the closure of an open set in $\mathbf{R}^n$.

If ϕ is r-times continuously differentiable on U for some positive integer r, then we may say that

(6.15.3) Φ is a C^r *n-surface* in V.

Similarly, if ϕ is infinitely differentiable on U, then we may say that

(6.15.4) Φ is a C^∞ *n-surface* in V.

6.15.1 Integration over n-surfaces

Let α be a differential n-form on $\phi(K)$ that is continuous on $\phi(K)$. The *integral of α over the n-surface* Φ is defined by

$$\int_\Phi \alpha = \int_K \phi^*(\alpha). \tag{6.15.5}$$

More precisely, $\phi^*(\alpha)$ is a differential n-form on K that is continuous on K under these conditions, so that the right side may be defined as in Section 6.12. This basically corresponds to (35) on p254 of [155], although this version is closer to Theorem 10.24 on p264 of [155].

Remember that

$$\phi^{-1}(V) = \{x \in U : \phi(x) \in V\} \tag{6.15.6}$$

is an open subset of $\mathbf{R}^n$, because ϕ is continuous, as mentioned in Section 5.9. Note that

$$K \subseteq \phi^{-1}(V), \tag{6.15.7}$$

by (6.15.1). Of course,

$$\phi(\phi^{-1}(V)) \subseteq V \tag{6.15.8}$$

automatically. If α is a differential form on V, then $\phi^*(\alpha)$ may be defined as a differential form on $\phi^{-1}(V)$ in the usual way.

6.15.2 Changing variables in $\mathbf{R}^n$

Let U_0 be another nonempty open subset of $\mathbf{R}^n$, and let θ be a C^1 diffeomorphism from U_0 onto U. Put

$$\phi_0 = \phi \circ \theta, \tag{6.15.9}$$

which is a continuously-differentiable mapping from U_0 into $\mathbf{R}^m$. Also put

$$K_0 = \theta^{-1}(K), \tag{6.15.10}$$

so that K_0 is a compact subset of $\mathbf{R}^n$ that is contained in U_0, and

$$\phi_0(K_0) = \phi(K) \subseteq V. \tag{6.15.11}$$

Thus the restriction of ϕ_0 to K_0 is considered to be another n-surface in V, which may be expressed as $\Phi_0 = \Phi_0(K_0)$. More precisely, this also uses the differential of ϕ_0 at points in K_0, as before.

Of course,

$$\phi_0^*(\alpha) = \theta^*(\phi^*(\alpha)) \tag{6.15.12}$$

on K_0, as in Subsection 4.5.1. This implies that

$$\int_{\Phi_0} \alpha = \int_{K_0} \phi_0^*(\alpha) = \int_{K_0} \theta^*(\phi^*(\alpha)). \tag{6.15.13}$$

If θ is orientation-preserving on U_0, then it follows that

$$\int_{\Phi_0} \alpha = \int_K \phi^*(\alpha) = \int_\Phi \alpha, \tag{6.15.14}$$

where the first step is as in Subsection 6.13.1. This also uses the fact that

$$\theta(K_0) = K, \tag{6.15.15}$$

because $\theta(U_0) = U$, by hypothesis. Similarly, if θ is orientation-reversing on U_0, then

$$\int_{\Phi_0} \alpha = -\int_K \phi^*(\alpha) = -\int_\Phi \alpha. \tag{6.15.16}$$

6.15.3 Mapping V into $\mathbf{R}^p$

Let p be another positive integer, and let η be a continuously-differentiable mapping from V into $\mathbf{R}^p$. Put

$$\psi = \eta \circ \phi, \tag{6.15.17}$$

which is defined as a mapping from $\phi^{-1}(V)$ into $\mathbf{R}^p$. In fact, ψ is continuously differentiable on $\phi^{-1}(V)$, as in Section 3.9.

Let W be an open subset of $\mathbf{R}^p$ such that

$$\psi(K) = \eta(\phi(K)) \subseteq W. \tag{6.15.18}$$

The restriction of ψ to K may be considered as an n-surface in W, as before. We may use $\Psi = \Psi(K)$ to refer to this n-surface.

Let β be a differential n-form on $\psi(K)$ that is continuous on $\psi(K)$, so that $\psi^*(\beta)$ is a differential n-form on K that is continuous on K, as before. Similarly, $\eta^*(\beta)$ defines a differential n-form on $\phi(K)$ that is continuous on $\phi(K)$, and $\phi^*(\eta^*(\beta))$ defines a differential n-form on K that is continuous on K. Note that

$$\psi^*(\beta) = \phi^*(\eta^*(\beta)) \tag{6.15.19}$$

on K, as in Subsection 4.5.1. Under these conditions, we have that

$$\int_\Psi \beta = \int_K \psi^*(\beta) = \int_K \phi^*(\eta^*(\beta)) = \int_\Phi \eta^*(\beta). \tag{6.15.20}$$

This corresponds to Theorem 10.25 on p265 of [155].

Chapter 7

Simplices and chains

7.1 Affine mappings

Let V and W be vector spaces over the real numbers. A mapping f from V into W is said to be *affine* if

(7.1.1) $$f - f(0) \text{ is a linear mapping from } V \text{ into } W,$$

as in p266 of [155]. Equivalently, this means that there is a linear mapping T from V into W such that

(7.1.2) $$f(v) = f(0) + T(v)$$

for every $v \in V$. It is easy to see that the space of all affine mappings from V into W is a linear subspace of the space of all mappings from V into W, as a vector space over the real numbers with respect to pointwise addition and scalar multiplication of functions. The space $\mathcal{L}(V, W)$ of all linear mappings from V into W is a linear subspace of the space of all affine mappings from V into W.

Let Z be another vector space over the real numbers. If f is an affine mapping from V into W, and g is an affine mapping from W into Z, then one can check that

(7.1.3) $$g \circ f \text{ is an affine mapping from } V \text{ into } Z.$$

Let f be an affine mapping from V into W, and let T be as in (7.1.2). Observe that

(7.1.4) $$f \text{ is one-to-one on } V$$

if and only if

(7.1.5) $$T \text{ is one-to-one on } V.$$

Similarly,

(7.1.6) $$f(V) = W$$

if and only if

(7.1.7) $$T(V) = W.$$

Let n be a positive integer, and let $e_1, \ldots, e_n$ be the standard basis vectors in $\mathbf{R}^n$. If f is an affine mapping from $\mathbf{R}^n$ into W, then f is uniquely determined by

$$f(0), f(e_1), \ldots, f(e_n). \tag{7.1.8}$$

More precisely, if T is the corresponding linear mapping from $\mathbf{R}^n$ into W as in (7.1.2), then

$$T(e_j) = f(e_j) - f(0) \tag{7.1.9}$$

for each $j = 1, \ldots, n$, and this determines T uniquely. It is easy to see that (7.1.8) may be any $n+1$ elements of W, because $T(e_1), \ldots, T(e_n)$ may be any n elements of W.

Let m be another positive integer, and let f be an affine mapping from $\mathbf{R}^n$ into $\mathbf{R}^m$. If T is the linear mapping from $\mathbf{R}^n$ into $\mathbf{R}^m$ corresponding to f as in (7.1.2), then T is the same as the differential of f at every point in $\mathbf{R}^n$.

7.2 Some remarks about determinants

Let n be a positive integer, and let $\sigma \in \mathrm{Sym}(n)$ be given. Also let T_σ be the linear mapping from $\mathbf{R}^n$ into itself defined by

$$(T_\sigma(x))_j = x_{\sigma(j)} \tag{7.2.1}$$

for each $x \in \mathbf{R}^n$ and $j = 1, \ldots, n$, as in Section 1.5. If $e_1, \ldots, e_n$ is the standard basis for $\mathbf{R}^n$, then it is easy to see that

$$T_\sigma(e_l) = e_{\sigma^{-1}(l)} \tag{7.2.2}$$

for each $l = 1, \ldots, n$.

Remember that

$$\det T = \mu_{\det}(T(e_1), \ldots, T(e_n)) \tag{7.2.3}$$

for any linear mapping T from $\mathbf{R}^n$ into itself, as in Section 2.5. Here $\mu_{\det}$ is the alternating n-linear form on $\mathbf{R}^n$ associated to the determinant as in Section 1.15. Thus

$$\det T_\sigma = \mu_{\det}(e_{\sigma^{-1}(1)}, \ldots, e_{\sigma^{-1}(n)}). \tag{7.2.4}$$

One can use this to check that

$$\det T_\sigma = \mathrm{sgn}(\sigma). \tag{7.2.5}$$

Let us put

$$e_0 = 0, \tag{7.2.6}$$

for convenience. Let τ be a permutation on $\{0, 1, \ldots, n\}$, and let R_τ be the linear mapping from $\mathbf{R}^n$ into itself such that

$$R_\tau(e_l) = e_{\tau(l)} - e_{\tau(0)} \tag{7.2.7}$$

for each $l = 1, \ldots, n$. Note that $\operatorname{sgn}(\tau)$ can be defined in the same way as in Section 1.4, which corresponds to identifying τ with an element of $\operatorname{Sym}(n+1)$ in the obvious way.

Suppose for the moment that

$$\tau(0) = 0, \tag{7.2.8}$$

so that

$$R_\tau(e_l) = e_{\tau(l)} \tag{7.2.9}$$

for each $l = 1, \ldots, n$. Let τ_0 be the restriction of τ to $\{1, \ldots, n\}$, which is an element of $\operatorname{Sym}(n)$. Note that

$$R_\tau = T_{\tau_0^{-1}} \tag{7.2.10}$$

in this case. Thus

$$\det R_\tau = \det T_{\tau_0^{-1}} = \operatorname{sgn}(\tau_0^{-1}) = \operatorname{sgn}(\tau_0), \tag{7.2.11}$$

usng (7.2.5) in the second step. This implies that

$$\det R_\tau = \operatorname{sgn}(\tau), \tag{7.2.12}$$

because $\operatorname{sgn}(\tau) = \operatorname{sgn}(\tau_0)$, as in Section 2.7.

7.2.1 The case where $\tau(0) \neq 0$

Suppose now that

$$\tau(0) \neq 0, \tag{7.2.13}$$

so that $\tau^{-1}(0) \neq 0$ too. Let ρ be the mapping from $\{1, \ldots, n\}$ into itself defined by

$$\begin{aligned} \rho(l) &= \tau(l) \quad \text{when } l \neq \tau^{-1}(0) \\ &= \tau(0) \quad \text{when } l = \tau^{-1}(0). \end{aligned} \tag{7.2.14}$$

It is easy to see that $\rho \in \operatorname{Sym}(n)$. Observe that

$$\begin{aligned} R_\tau(e_l) &= e_{\rho(l)} - e_{\rho(\tau^{-1}(0))} \quad \text{when } l \neq \tau^{-1}(0) \\ &= -e_{\tau(0)} = -e_{\rho(l)} \quad \text{when } l = \tau^{-1}(0). \end{aligned} \tag{7.2.15}$$

Let L_ρ be the linear mapping from $\mathbf{R}^n$ into itself such that

$$\begin{aligned} L_\rho(e_l) &= e_{\rho(l)} \quad \text{when } l \neq \tau^{-1}(0) \\ &= -e_{\rho(l)} \quad \text{when } l = \tau^{-1}(0). \end{aligned} \tag{7.2.16}$$

One can check that

$$\det R_\tau = \det L_\rho. \tag{7.2.17}$$

More precisely, one can verify that

$$\mu_{\det}(R_\tau(e_1), \ldots, R_\tau(e_n)) = \mu_{\det}(L_\rho(e_1), \ldots, L_\rho(e_n)). \tag{7.2.18}$$

Let $T_{\rho^{-1}}$ be the linear mapping on $\mathbf{R}^n$ associated to ρ^{-1} as at the beginning of the section. It is easy to see that

$$\det L_\rho = -\det T_{\rho^{-1}}, \tag{7.2.19}$$

using (7.2.3) and the definition of L_ρ. This means that

$$\det L_\rho = -\operatorname{sgn}(\rho), \tag{7.2.20}$$

because of (7.2.5). It follows that

$$\det R_\tau = -\operatorname{sgn}(\rho), \tag{7.2.21}$$

by (7.2.17).

Let ρ_1 be the permutation on $\{0, 1, \ldots, n\}$ that is equal to ρ on $\{1, \ldots, n\}$, and satisfies $\rho_1(0) = 0$. Thus

$$\operatorname{sgn}(\rho_1) = \operatorname{sgn}(\rho), \tag{7.2.22}$$

as in Section 2.7 again. One can check that τ is the same as the composition of ρ_1 with the permutation on $\{0, 1, \ldots, n\}$ that interchanges 0 and $\tau(0)$, and leaves the other points fixed. This implies that

$$\operatorname{sgn}(\tau) = -\operatorname{sgn}(\rho_1). \tag{7.2.23}$$

It follows that

$$\det R_\tau = -\operatorname{sgn}(\rho_1) = \operatorname{sgn}(\tau), \tag{7.2.24}$$

so that (7.2.12) holds in this case as well.

7.3 Affine simplices

Let n be a positive integer, and let $e_1, \ldots, e_n$ be the standard basis in $\mathbf{R}^n$, as usual. Remember that Q^n denotes the standard simplex in $\mathbf{R}^n$, as in Section 6.2. This may be described equivalently as the set of points in $\mathbf{R}^n$ of the form

$$\sum_{j=1}^{n} t_j\, e_j, \tag{7.3.1}$$

where t_j is a nonnegative real number for each j, and $\sum_{j=1}^{n} t_j \le 1$, as on p266 of [155].

Let m be another positive integer, and let $p_0, p_1, \ldots, p_n$ be elements of $\mathbf{R}^m$. These elements of $\mathbf{R}^m$ are not required to be distinct. Consider the affine mapping

$$\phi = \phi_{p_0, p_1, \ldots, p_n} \tag{7.3.2}$$

from $\mathbf{R}^n$ into $\mathbf{R}^m$ defined by

$$\phi_{p_0,p_1,\ldots,p_n}(x) = p_0 + \sum_{j=1}^{n} x_j\,(p_j - p_0). \tag{7.3.3}$$

This is the unique affine mapping from $\mathbf{R}^n$ into $\mathbf{R}^m$ such that

$$\phi(0) = p_0 \text{ and } \phi(e_j) = p_j \text{ for each } j = 1, \ldots, n, \tag{7.3.4}$$

as in Section 7.1. Note that every affine mapping from $\mathbf{R}^n$ into $\mathbf{R}^m$ corresponds to elements $p_0, p_1, \ldots, p_n$ of $\mathbf{R}^m$ in this way.

Let

$$\Phi = \Phi(p_0, p_1, \ldots, p_n) = [p_0, p_1, \ldots, p_n] \tag{7.3.5}$$

be the n-surface in $\mathbf{R}^m$ associated to $\phi_{p_0,p_1,\ldots,p_n}$ with parameter domain Q^n, as in Section 6.15. Let us call this the

(7.3.6) *parameterized affine n-simplex* in $\mathbf{R}^m$ associated to $p_0, p_1, \ldots, p_n$.

This may also be called the *oriented affine n-simplex* in $\mathbf{R}^n$ associated to $p_0, p_1, \ldots, p_n$, as on p266 of [155]. Indeed, one is often primarily concerned with the way that the affine n-simplex is oriented, rather than the particular affine parameterization, as in Section 7.5. The points $p_0, p_1, \ldots, p_n$ may be called the *vertices* of this affine n-simplex.

Of course,

$$\phi_{p_0,p_1,\ldots,p_n}(Q^n) = \left\{p_0 + \sum_{j=1}^{n} x_j\,(p_j - p_0) : x \in Q^n\right\}. \tag{7.3.7}$$

Equivalently,

$$\phi_{p_0,p_1,\ldots,p_n}(Q^n) = \left\{\left(1 - \sum_{j=1}^{n} x_j\right) p_0 + \sum_{j=1}^{n} x_j\,p_j : x \in Q^n\right\}. \tag{7.3.8}$$

If $m = n$, $p_0 = 0$, and $p_j = e_j$ for each $j = 1, \ldots, n$, then $\phi_{p_0,p_1,\ldots,p_n}$ is the identity mapping on $\mathbf{R}^n$, so that $\phi_{p_0,p_1,\ldots,p_n}(Q^n) = Q^n$.

Alternatively, $\phi_{p_0,p_1,\ldots,p_n}(Q^n)$ consists of the points in $\mathbf{R}^m$ of the form

$$\sum_{j=0}^{n} t_j\,p_j, \tag{7.3.9}$$

where $t_0, t_1, \ldots, t_n$ are nonnegative real numbers such that $\sum_{j=0}^{n} t_j = 1$. This description shows that

(7.3.10) $\phi_{p_0,p_1,\ldots,p_n}(Q^n)$ does not depend on the order in which the vertices $p_0, p_1, \ldots, p_n$ are listed.

This will be discussed further in the next section.

7.4 Some affine mappings on $\mathbf{R}^n$

Let V and W be vector spaces over the real numbers, and let f be an affine mapping from V into W, as in Section 7.1. Suppose that $v_1, \ldots, v_k$ are finitely many elements of V, and thet $t_1, \ldots, t_k$ are real numbers such that

$$\sum_{j=1}^{k} t_j = 1. \tag{7.4.1}$$

Under these conditions, one can check that

$$f\Big(\sum_{j=1}^{k} t_j\, v_j\Big) = \sum_{j=1}^{k} t_j\, f(v_j). \tag{7.4.2}$$

In fact, one can verify that this property characterizes affine mappings, with $k = 2$.

Let n be a positive integer, let $e_1, \ldots, e_n$ be the standard basis in $\mathbf{R}^n$, and let us put $e_0 = 0$, as in Section 7.2. Note that Q^n is the set of points in $\mathbf{R}^n$ of the form

$$\sum_{j=0}^{n} t_j\, e_j, \tag{7.4.3}$$

where $t_0, t_1, \ldots, t_n$ are nonnegative real numbers with $\sum_{j=0}^{n} t_j = 1$. Let m be another positive integer, let $p_0, p_1, \ldots, p_n$ be elements of $\mathbf{R}^m$, and let $\phi_{p_0, p_1, \ldots, p_n}$ be as in the previous section. One can use (7.4.2) to get the description of the elements of $\phi_{p_0, p_1, \ldots, p_n}(Q^n)$ in (7.3.9).

7.4.1 Some affine mappings from permutations

Let τ be a permutation on $\{0, 1, \ldots, n\}$, and let R_τ be the linear mapping from $\mathbf{R}^n$ into itself such that

$$R_\tau(e_l) = e_{\tau(l)} - e_{\tau(0)} \tag{7.4.4}$$

for each $l = 1, \ldots, n$, as in Section 7.2. If $x \in \mathbf{R}^n$, then put

$$\xi_\tau(x) = e_{\tau(0)} + R_\tau(x), \tag{7.4.5}$$

This is the unique affine mapping from $\mathbf{R}^n$ into itself such that

$$\xi_\tau(e_l) = e_{\tau(l)} \tag{7.4.6}$$

for each $l = 0, 1, \ldots, n$. This is the same as the identity mapping on $\mathbf{R}^n$ when τ is the identity mapping on $\{0, 1, \ldots, n\}$, and

$$\xi_\tau = \phi_{e_{\tau(0)}, e_{\tau(1)}, \ldots, e_{\tau(n)}}, \tag{7.4.7}$$

with $m = n$, for any permutation τ on $\{0, 1, \ldots, n\}$. In particular,

$$\xi_\tau(Q^n) = Q^n \tag{7.4.8}$$

for each τ, as in the preceding paragraph.

The differential of ξ_τ at any point in $\mathbf{R}^n$ is equal to R_τ, as in Section 7.1. Remember that the determinant of R_τ is equal to $\mathrm{sgn}(\tau)$, as in Section 7.2. This means that

$$\xi_\tau \text{ is orientation-preserving on } \mathbf{R}^n \text{ when } \tau \text{ is an even permutation,} \tag{7.4.9}$$

and that

$$\xi_\tau \text{ is orientation-reversing on } \mathbf{R}^n \text{ when } \tau \text{ is an odd permutation,} \tag{7.4.10}$$

as in Section 6.14.

Observe that

$$\phi_{p_0,p_1,\ldots,p_n}(\xi_\tau(e_l)) = \phi_{p_0,p_1,\ldots,p_n}(e_{\tau(l)}) = p_{\tau(l)} \tag{7.4.11}$$

for each $l = 0, 1, \ldots, n$. This implies that

$$\phi_{p_0,p_1,\ldots,p_n} \circ \xi_\tau = \phi_{p_{\tau(0)},p_{\tau(1)},\ldots,p_{\tau(n)}}. \tag{7.4.12}$$

In particular,

$$\phi_{p_{\tau(0)},p_{\tau(1)},\ldots,p_{\tau(n)}}(Q^n) = \phi_{p_0,p_1,\ldots,p_n}(\xi_\tau(Q^n)) = \phi_{p_0,p_1,\ldots,p_n}(Q^n), \tag{7.4.13}$$

as mentioned in the previous section.

7.5 Integrals over affine simplices

Let us continue with the same notation and hypotheses as in the previous section. Let β be a differential n-form on Q^n that is continuous on Q^n, so that $\xi_\tau^*(\beta)$ is another differential n-form on Q^n that is continuous on Q^n. Observe that

$$\int_{Q^n} \xi_\tau^*(\beta) = \mathrm{sgn}(\tau) \int_{Q^n} \beta, \tag{7.5.1}$$

as in Subsection 6.13.1. More precisely, this uses the fact that the differential of ξ_τ at any point is equal to R_τ, whose determinant is equal to $\mathrm{sgn}(\tau)$, as before.

Similarly, let α be a differential n-form on $\phi_{p_0,p_1,\ldots,p_n}(Q^n)$ that is continuous on this set. Note that

$$\phi^*_{p_{\tau(0)},p_{\tau(1)},\ldots,p_{\tau(n)}}(\alpha) = \xi_\tau^*(\phi^*_{p_0,p_1,\ldots,p_n}(\alpha)) \tag{7.5.2}$$

on Q^n, because of (7.4.12). It follows that

$$\begin{aligned} \int_{\Phi(p_{\tau(0)},p_{\tau(1)},\ldots,p_{\tau(n)})} \alpha &= \int_{Q^n} \phi^*_{p_{\tau(0)},p_{\tau(1)},\ldots,p_{\tau(n)}}(\alpha) \\ &= \int_{Q^n} \xi_\tau^*(\phi^*_{p_0,p_1,\ldots,p_n}(\alpha)) \qquad (7.5.3) \\ &= \mathrm{sgn}(\tau) \int_{Q^n} \phi^*_{p_0,p_1,\ldots,p_n}(\alpha) = \mathrm{sgn}(\tau) \int_{\Phi(p_0,p_1,\ldots,p_n)} \alpha. \end{aligned}$$

This uses (7.5.1) in the third step. Here $\Phi(p_0, p_1, \ldots, p_n)$ and

$$\Phi(p_{\tau(0)}, p_{\tau(1)}, \ldots, p_{\tau(n)}) = [p_{\tau(0)}, p_{\tau(1)}, \ldots, p_{\tau(n)}] \tag{7.5.4}$$

are as in Section 7.3, and the integrals over them are as in Subsection 6.15.1. Of course, (7.5.3) may be considered as a version of the remarks in Subsection 6.15.2. This corresponds to Theorem 10.27 on p267 of [155].

7.5.1 Orientations of affine simplices

Let us put

$$\Phi(p_{\tau(0)}, p_{\tau(1)}, \ldots, p_{\tau(n)}) \simeq \operatorname{sgn}(\tau)\, \Phi(p_0, p_1, \ldots, p_n), \tag{7.5.5}$$

or equivalently

$$[p_{\tau(0)}, p_{\tau(1)}, \ldots, p_{\tau(n)}] \simeq \operatorname{sgn}(\tau)\, [p_0, p_1, \ldots, p_n]. \tag{7.5.6}$$

If τ is an even permutation, then this means that

$$\begin{aligned} &\Phi(p_0, p_1, \ldots, p_n) \text{ and } \Phi(p_{\tau(0)}, p_{\tau(1)}, \ldots, p_{\tau(n)}) \\ &\text{have the } \textit{same orientation}, \end{aligned} \tag{7.5.7}$$

even though they may be parameterized differently as n-surfaces in $\mathbf{R}^m$. Similarly, if τ is an odd permutation, then this means that

$$\begin{aligned} &\Phi(p_0, p_1, \ldots, p_n) \text{ and } \Phi(p_{\tau(0)}, p_{\tau(1)}, \ldots, p_{\tau(n)}) \\ &\text{have } \textit{opposite orientations}. \end{aligned} \tag{7.5.8}$$

This corresponds to some remarks on p267 of [155]. Of course, this notation is compatible with (7.5.3).

Remember that $\phi_{p_0, p_1, \ldots, p_n}$ is one-to-one on $\mathbf{R}^n$ if and only if the corresponding linear mapping

$$\phi_{p_0, p_1, \ldots, p_n}(x) - p_0 = \sum_{j=1}^{n} x_j\, (p_j - p_0) \tag{7.5.9}$$

is a one-to-one linear mapping from $\mathbf{R}^n$ into $\mathbf{R}^m$, as in Section 7.1. This happens if and only if

$$p_j - p_0,\ 1 \le j \le n, \text{ are linearly indpendent in } \mathbf{R}^m. \tag{7.5.10}$$

If $m = n$, then we may say that $\Phi(p_0, p_1, \ldots, p_n)$ is *positively oriented* when the determinant of (7.5.9) is positive, and *negatively oriented* when the determinant is negative, as on p267 of [155].

7.5.2 0-Simplices

An *oriented affine* 0*-simplex* in $\mathbf{R}^m$ is defined to be a point $p_0 \in \mathbf{R}^m$ with a sign attached to indicate the orientation, i.e., $+p_0$ or $-p_0$, as on p267 of [155]. This may be denoted

$$\Phi = \epsilon \, p_0, \tag{7.5.11}$$

with $\epsilon = \pm 1$. If f is a differential 0-form defined at p_0, which is to say a real-valued function, then we put

$$\int_\Phi f = \epsilon \, f(p_0) \tag{7.5.12}$$

in this case, as in [155].

Alternatively,

$$\text{the } \textit{standard } 0\textit{-simplex } Q^0 \tag{7.5.13}$$

may be defined as the set $\{0\}$ with 0 as its only element, as on p141 of [184]. A *parameterized* 0*-simplex* in $\mathbf{R}^m$ may be considered as a mapping from Q^0 into $\mathbf{R}^m$, as on p142, 191 of [184]. This is essentially the same as a point in $\mathbf{R}^m$, without including a sign, as in the preceding paragraph. If $p \in \mathbf{R}^m$, then we may use

$$\Phi(p) \tag{7.5.14}$$

for the corresponding parameterized 0-simplex in $\mathbf{R}^m$. The corresponding integral may be defined as in (7.5.12), without the extra factor of ϵ on the right side.

However, signs may also be included in chains, as in the next section. Using this, the difference mentioned in the previous paragraphs does not really matter.

7.6 Affine chains

Let m and n be positive integers, and let V be a nonempty open subset of $\mathbf{R}^m$. An *affine n-chain* Γ in V consists of finitely many parameterized affine n-simplices

$$\Phi_1, \ldots, \Phi_k \tag{7.6.1}$$

in V with signs

$$\epsilon_1, \ldots, \epsilon_k \tag{7.6.2}$$

attached. More precisely, for each $l = 1, \ldots, k$, Φ_l corresponds to an affine mapping ϕ_l from $\mathbf{R}^n$ into $\mathbf{R}^m$ such that

$$\phi_l(Q^n) \subseteq V, \tag{7.6.3}$$

as in Section 7.3, where Q^n is the standard simplex in $\mathbf{R}^n$, as in Section 6.2. This basically corresponds to the definition in Section 10.28 on p268 of [155].

In [155], the signs $\epsilon_l = \pm 1$ are not included directly in this way, and instead they correspond to using odd permutations of the vertices of the simplices. As in [155], the n-simplices are not required to be distinct, so that they may occur with multiplicity. We can also allow $n = 0$ here, using parameterized affine 0-simplices in $\mathbf{R}^m$, as in Subsection 7.5.2. This is equivalent to using oriented affine 0-simplices in $\mathbf{R}^m$, by including the signs directly here.

We shall consider "simplices" that are parameterized by mappings that are r-times continuously differentiable for some nonnegative integer r, or infinitely differentiable, in Section 7.9. We shall also consider corresponding notions of chains there. We shall say a bit more about some aspects of chains in Section 7.13.

7.6.1 Integrals over affine chains

Let α be a differential n-form on V that is continuous on V. If Γ is as before, then we put

$$\int_\Gamma \alpha = \sum_{l=1}^{k} \epsilon_l \int_{\Phi_l} \alpha, \tag{7.6.4}$$

as in (82) on p268 of [155]. Remember that the integrals on the right are defined as in Subsection 6.15.1 when $n \geq 1$, and as in Subsection 7.5.2 when $n = 0$.

An affine n-chain Γ as before is often expressed as a formal sum

$$\Gamma = \sum_{l=1}^{k} \epsilon_l \, \Phi_l, \tag{7.6.5}$$

as in (83), (84) on p268 of [155]. The commutativity of addition reflects the fact that we are not really concerned with the order in which the terms are listed, as in the right side of (7.6.4). Similarly, if $c_1, \ldots, c_k$ are integers, then the formal sum

$$\sum_{l=1}^{k} c_l \, \Phi_l \tag{7.6.6}$$

may be considered as an affine n-chain in V. If $c_l \neq 0$, then this means that $\mathrm{sign}(c_l) \, \Phi_l$ occurs with multiplicity $|c_l|$ in the chain.

Note that this type of formal sum is different from taking sums of affine mappings from $\mathbf{R}^n$ into $\mathbf{R}^m$ when $n \geq 1$, as mentioned on p268 of [155]. Similarly, if $n = 0$, then this type of formal sum is different from taking sums of elements of $\mathbf{R}^m$.

7.6.2 Defining $\simeq$ for affine chains

Let $\widetilde{\Phi}_1, \ldots, \widetilde{\Phi}_k$ be parameterized affine n-simplices in V with signs $\widetilde{\epsilon}_1, \ldots, \widetilde{\epsilon}_k$, so that

$$\widetilde{\Gamma} = \sum_{l=1}^{k} \widetilde{\epsilon}_l \, \widetilde{\Phi}_l \tag{7.6.7}$$

is an affine n-chain in V. Suppose that

$$\epsilon_l \, \Phi_l \simeq \widetilde{\epsilon}_l \, \widetilde{\Phi}_l \tag{7.6.8}$$

for each $l = 1, \ldots, k$, in the notation of Subsection 7.5.1. Under these conditions, we extend this notation by putting

$$\Gamma \simeq \widetilde{\Gamma}. \tag{7.6.9}$$

If α is as before, then (7.6.8) implies that

$$\epsilon_l \int_{\Phi_l} \alpha = \widetilde{\epsilon}_l \int_{\widetilde{\Phi}_l} \alpha \tag{7.6.10}$$

for each $l = 1, \ldots, k$, as in Section 7.5. It follows that

$$\int_\Gamma \alpha = \int_{\widetilde{\Gamma}} \alpha, \tag{7.6.11}$$

by summing over l, as in (7.6.4).

7.7 Boundaries of affine simplices

Let m and n be positive integers, and let $p_0, p_1, \ldots, p_n$ be elements of $\mathbf{R}^m$, so that $\Phi(p_0, p_1, \ldots, p_n)$ defines a parameterized affine n-simplex in $\mathbf{R}^m$, as in Section 7.3. The *boundary* of $\Phi(p_0, p_1, \ldots, p_n)$ is defined as an affine $(n-1)$-chain in $\mathbf{R}^m$ by

$$\partial \Phi(p_0, p_1, \ldots, p_n) = \sum_{l=0}^{n} (-1)^l \, \Phi(p_0, \ldots, p_{l-1}, p_{l+1} \ldots, p_n), \tag{7.7.1}$$

as in (85) on p269 of [155]. Note that

$$\begin{aligned} &\text{the image of } \Phi(p_0, \ldots, p_{j-1}, p_{j+1}, \ldots, p_n) \\ &\text{is contained in the image of } \Phi(p_0, p_1, \ldots, p_n) \end{aligned} \tag{7.7.2}$$

for each j, by a remark in Section 7.3. If $n = 1$, then we get that

$$\partial \Phi(p_0, p_1) = \Phi(p_1) - \Phi(p_0), \tag{7.7.3}$$

where $\Phi(p)$ is the parameterized 0-simplex corresponding to $p \in \mathbf{R}^m$ as in Subsection 7.5.2.

If $n = 2$, then

$$\partial \Phi(p_0, p_1, p_2) = \Phi(p_1, p_2) - \Phi(p_0, p_2) + \Phi(p_1, p_2), \tag{7.7.4}$$

as on p269 of [155]. Note that

$$\Phi(p_0, p_2) \simeq -\Phi(p_2, p_0), \tag{7.7.5}$$

in the notation of Subsection 7.5.1. We can extend this notation to affine n-chains, as in Subsection 7.6.2, to get that

$$\partial\Phi(p_0, p_1, p_2) \simeq \Phi(p_0, p_1) + \Phi(p_1, p_2) + \Phi(p_2, p_0), \tag{7.7.6}$$

as in [155].

If n is any positive integer and $1 \le l \le n$, then

$$\Phi(p_0, \ldots, p_{l-1}, p_{l+1}, \ldots, p_n) = [p_0, \ldots, p_{l-1}, p_{l+1}, \ldots, p_n] \tag{7.7.7}$$

is the parameterized affine $(n-1)$-simplex in $\mathbf{R}^m$ associated to these vertices as in Section 7.3. If $n \ge 2$, then this corresponds to the affine mapping

$$\phi_l = \phi_{p_0,\ldots,p_{l-1},p_{l+1},\ldots,p_n} \tag{7.7.8}$$

from $\mathbf{R}^{n-1}$ into $\mathbf{R}^m$ defined by

$$\phi_l(x) = p_0 + \sum_{j=1}^{l-1} x_j\,(p_j - p_0) + \sum_{j=l}^{n-1} x_j\,(p_{j+1} - p_0), \tag{7.7.9}$$

as on p269 of [155]. Remember that the parameter domain of (7.7.7) is the standard simplex Q^{n-1} in $\mathbf{R}^{n-1}$.

The $l = 0$ term on the right side of (7.7.1) uses the parameterized affine $(n-1)$-simplex

$$\Phi(p_1, \ldots, p_n) = [p_1, \ldots, p_n] \tag{7.7.10}$$

in $\mathbf{R}^m$. If $n \ge 2$, then this corresponds to the affine mapping

$$\phi_0 = \phi_{p_1,\ldots,p_n} \tag{7.7.11}$$

from $\mathbf{R}^{n-1}$ into $\mathbf{R}^m$ defined by

$$\phi_0(x) = p_1 + \sum_{j=1}^{n-1} x_j\,(p_{j+1} - p_1), \tag{7.7.12}$$

as on p269 of [155].

7.7.1 Boundaries of affine chains

Let V be a nonempty open subset of $\mathbf{R}^m$, and let Γ be an affine n-chain in V, as in (7.6.5). The *boundary* of Γ is defined as an affine $(n-1)$-chain in V by

$$\partial\Gamma = \sum_{l=1}^{k} \epsilon_l\,\partial\Phi_l. \tag{7.7.13}$$

This corresponds to (90) on p270 of [155], in the case of affine chains. The boundary of a parameterized affine 0-simplex may be interpreted as being equal to 0, so that the boundary of an affine 0-chain is interpreted as being equal to 0 as well.

7.8 Boundaries and permutations

Let n be a positive integer, and let l be an integer with $0 \le l \le n$. Put

$$E_{n,l} = \{0, \ldots, l-1, l+1, n\}, \tag{7.8.1}$$

which is a set with n elements. Let $\rho_{n,l}$ be the unique mapping from

$$\{0, 1, \ldots, n-1\} \tag{7.8.2}$$

onto $E_{n,l}$ that is strictly increasing, so that

$$\begin{aligned} \rho_{n,l}(j) &= j && \text{when } j \le l-1 \\ &= j+1 && \text{when } j \ge l. \end{aligned} \tag{7.8.3}$$

Equivalently, $\rho_{n,l}^{-1}$ is the unique mapping from $E_{n,l}$ onto (7.8.2) that is strictly increasing, with

$$\begin{aligned} \rho_{n,l}^{-1}(k) &= k && \text{when } k \le l-1 \\ &= k-1 && \text{when } k \ge l+1. \end{aligned} \tag{7.8.4}$$

If τ is a permutation on $\{0, 1, \ldots, n\}$, then the restriction of τ to $E_{n,l}$ defines a one-to-one mapping onto $E_{n,\tau(l)}$. Put

$$\tau_l = \rho_{n,\tau(l)}^{-1} \circ \tau \circ \rho_{n,l}, \tag{7.8.5}$$

which defines a permutation on (7.8.2). We can define $\operatorname{sgn}(\tau)$ and $\operatorname{sgn}(\tau_l)$ as in Section 1.4, by identifying these permutations with elements of $\operatorname{Sym}(n+1)$ and $\operatorname{Sym}(n)$, respectively, in the obvious way.

More precisely,

$$\operatorname{sgn}(\tau) = \prod_{0 \le j < k \le n} \operatorname{sign}(\tau(k) - \tau(j)), \tag{7.8.6}$$

as before. Similarly,

$$\operatorname{sgn}(\tau_l) = \prod_{0 \le p < q \le n-1} \operatorname{sign}(\tau_l(q) - \tau_l(p)). \tag{7.8.7}$$

It is easy to see that this is the same as saying that

$$\operatorname{sgn}(\tau_l) = \prod_{\substack{0 \le j < k \le n \\ j,k \ne l}} \operatorname{sign}(\tau(j) - \tau(k)) \tag{7.8.8}$$

in this case. It follows that

$$\operatorname{sgn}(\tau) = \operatorname{sgn}(\tau_l) \Big(\prod_{j=0}^{l-1} \operatorname{sign}(\tau(l) - \tau(j)) \Big) \Big(\prod_{k=l+1}^{n} \operatorname{sign}(\tau(k) - \tau(l)) \Big). \tag{7.8.9}$$

We would like to check that

$$\Big(\prod_{j=0}^{l-1} \operatorname{sign}(\tau(l) - \tau(j))\Big) \Big(\prod_{k=l+1}^{n} \operatorname{sign}(\tau(k) - \tau(l))\Big) = (-1)^{\tau(l)-l}. \tag{7.8.10}$$

Observe that

$$\begin{aligned} \tau(l) &= \#\{j = 0, \ldots, l-1 : \tau(j) < \tau(l)\} \\ &\quad + \#\{k = l+1, \ldots, n : \tau(k) < \tau(l)\}. \end{aligned} \tag{7.8.11}$$

Of course,

$$\begin{aligned} l &= \#\{j = 0, \ldots, l-1 : \tau(j) < \tau(l)\} \\ &\quad + \#\{j = 0, \ldots, l-1 : \tau(j) > \tau(l)\}. \end{aligned} \tag{7.8.12}$$

This implies that

$$\begin{aligned} \tau(l) - l &= \#\{k = l+1, \ldots, n : \tau(k) < \tau(l)\} \\ &\quad - \#\{j = 0, \ldots, l-1 : \tau(j) > \tau(l)\}. \end{aligned} \tag{7.8.13}$$

It is easy to use this to get (7.8.10). This means that

$$\operatorname{sgn}(\tau_l) = (-1)^{\tau(l)-l} \operatorname{sgn}(\tau). \tag{7.8.14}$$

7.8.1 Permuting vertices of affine simplices

Let m be another positive integer, let $p_0, p_1, \ldots, p_n$ be elements of $\mathbf{R}^m$, and let $\Phi(p_0, p_1, \ldots, p_n)$ be as in Section 7.3. Remember that

$$\Phi(p_{\tau(0)}, p_{\tau(1)}, \ldots, p_{\tau(n)}) \simeq \operatorname{sgn}(\tau)\, \Phi(p_0, p_1, \ldots, p_n), \tag{7.8.15}$$

in the notation of Subsection 7.5.1. Similarly, one can verify that

$$\begin{aligned} &\Phi(p_{\tau(0)}, \ldots, p_{\tau(l-1)}, p_{\tau(l+1)}, \ldots, p_{\tau(n)}) \\ &\quad \simeq \operatorname{sgn}(\tau_l)\, \Phi(p_0, \ldots, p_{\tau(l)-1}, p_{\tau(l)+1}, \ldots, p_n), \end{aligned} \tag{7.8.16}$$

as parameterized affine $(n-1)$-simplices in $\mathbf{R}^m$. Basically,

$$p_0, \ldots, p_{\tau(l)-1}, p_{\tau(l)+1}, \ldots, p_n \tag{7.8.17}$$

and

$$p_{\tau(0)}, \ldots, p_{\tau(l-1)}, p_{\tau(l+1)}, \ldots, p_{\tau(n)} \tag{7.8.18}$$

list the same set of vertices in $\mathbf{R}^m$, where the order in the second list corresponds to the restriction of τ to $E_{n,l}$.

This implies that

$$\begin{aligned} &\Phi(p_{\tau(0)}, \ldots, p_{\tau(l-1)}, p_{\tau(l+1)}, \ldots, p_{\tau(n)}) \\ &\quad \simeq (-1)^{\tau(l)-l} \operatorname{sgn}(\tau)\, \Phi(p_0, \ldots, p_{\tau(l)-1}, p_{\tau(l)+1}, \ldots, p_n), \end{aligned} \tag{7.8.19}$$

by (7.8.14). It follows that

$$
\begin{aligned}
(7.8.20) \qquad & \partial\Phi(p_{\tau(0)}, p_{\tau(1)}, \ldots, p_{\tau(n)}) \\
&= \sum_{l=0}^{n} (-1)^l \, \Phi(p_{\tau(0)}, \ldots, p_{\tau(l-1)}, p_{\tau(l+1)}, \ldots, p_n) \\
&\simeq \sum_{l=0}^{n} (-1)^{\tau(l)} \, \mathrm{sgn}(\tau) \, \Phi(p_0, \ldots, p_{\tau(l)-1}, p_{\tau(l)+1}, \ldots, p_n) \\
&= \mathrm{sgn}(\tau) \, \partial\Phi(p_0, p_1, \ldots, p_n),
\end{aligned}
$$

by the definition (7.7.1) of the boundary.

Let V be a nonempty open subset of $\mathbf{R}^m$, and let Γ, $\widetilde{\Gamma}$ be affine n-chains in V, as in (7.6.5) and (7.6.7), respectively. If (7.6.8) holds, then we get that

$$\partial\Gamma \simeq \partial\widetilde{\Gamma}, \tag{7.8.21}$$

because of (7.8.20).

7.9 C^r Simplices and chains

Let m, n, and r be positive integers, and let U, V be nonempty open subsets of $\mathbf{R}^n$, $\mathbf{R}^m$, respectively. Remember that Q^n is the standard simplex in $\mathbf{R}^n$, as in Section 6.2, and suppose that

$$Q^n \subseteq U. \tag{7.9.1}$$

Let ϕ be an r-times continuously-differentiable mapping from U into $\mathbf{R}^m$ with

$$\phi(Q^n) \subseteq V. \tag{7.9.2}$$

This defines a C^r n-surface Φ in V with parameter domain Q^n, as in Section 6.15.

Under these conditions,

(7.9.3) Φ is considered as a *parameterized C^r n-simplex* in V.

One may also call Φ an *oriented C^r n-simplex* in V, which basically corresponds to the definition at the beginning of Section 10.30 on p269 of [155] when $r = 2$.

Similarly, if ϕ is infinitely differentiable on U, then

(7.9.4) Φ is considered as a *parameterized C^∞ n-simplex* in V.

This corresponds to a definition in Section 4.6 on p141 in [184]. As before, Φ may be called an *oriented C^∞ n-simplex* in V in this case as well. Of course, a parameterized affine n-simplex in V is a parameterized C^∞ n-simplex in V in particular.

If $r = 0$, then one can simply ask that

(7.9.5) ϕ be a continuous mapping from Q^n into V.

This is mentioned on p142, 191 of [184], and is used in algebraic topology.

We may interpret a *parameterized C^r or C^∞ 0-simplex* in V to be the same as a parameterized affine 0-simplex in V, as in Subsection 7.5.2. Similarly, an *oriented C^r or C^∞ 0-simplex* may be interpreted as being the same as an oriented affine 0-simplex in V.

7.9.1 C^r chains

A *C^r or C^∞ n-chain* Γ in V consists of finitely many parameterized C^r or C^∞ n-simplices

$$\Phi_1, \ldots, \Phi_k \tag{7.9.6}$$

in V, respectively, with signs

$$\epsilon_1, \ldots, \epsilon_k \tag{7.9.7}$$

attached. This basically corresponds to the definition on p270 of [155] when $r = 2$. As in Section 7.6, the n-simplices are not required to be distinct. Note that an affine n-chain in V is a C^∞ n-chain in V in particular. Some aspects of these notions will be discussed further in Section 7.13.

As before, Γ is often expressed as a formal sum

$$\Gamma = \sum_{l=1}^{k} \epsilon_l \, \Phi_l. \tag{7.9.8}$$

Similarly, if $c_1, \ldots, c_k$ are integers, then the formal sum

$$\sum_{l=1}^{k} c_l \, \Phi_l \tag{7.9.9}$$

may be considered as a C^r or C^∞ n-chain in V, as appropriate. In an analogous definition on p142 of [184], one considers formal sums of parameterized C^∞ n-simplices in V with coefficients in $\mathbf{R}$. A related definition on p192 of [184] uses formal sums of continuous n-simplices in V with integer coefficients. We shall say more about this in Section 7.13 as well.

Let Γ be a C^1 n-chain in V as in (7.9.8), and let α be a differential n-form on V that is continuous on V. Under these conditions, we put

$$\int_\Gamma \alpha = \sum_{l=1}^{k} \epsilon_l \int_{\Phi_l} \alpha, \tag{7.9.10}$$

where the integrals on the right are defined as in Subsection 6.15.1 when $n \geq 1$, and as in Subsection 7.5.2 when $n = 0$. This corresponds to (87) on p270 of [155], and to (10) on p143 of [184].

7.10 Simplices and compositions

Let us continue with the same notation and hypotheses as in the previous section. If ϕ and U are as before, then

$$\phi^{-1}(V) = \{x \in U : \phi(x) \in V\} \tag{7.10.1}$$

is an open set in $\mathbf{R}^n$, because ϕ is continuous on U, as in Section 5.9. We also have that

$$Q^n \subseteq \phi^{-1}(V), \tag{7.10.2}$$

because of (7.9.1) and (7.9.2). Note that

$$\phi(\phi^{-1}(V)) \subseteq V. \tag{7.10.3}$$

Let q be another positive integer, and let η be a mapping from V into $\mathbf{R}^q$ that is r-times continuously differentiable or infinitely differentiable on V, as appropriate. Observe that

$$\psi = \eta \circ \phi \tag{7.10.4}$$

defines a mapping from $\phi^{-1}(V)$ into $\mathbf{R}^q$ that is r-times continuously differentiable or infinitely differentiable, as appropriate. Let W be a nonempty open subset of $\mathbf{R}^q$ with

$$\eta(V) \subseteq W, \tag{7.10.5}$$

so that

$$\psi(Q^n) = \eta(\phi(Q^n)) \subseteq \eta(V) \subseteq W. \tag{7.10.6}$$

Using ψ, we get a C^r or C^∞ n-surface Ψ in W, as appropriate, with parameter domain Q^n, as in Section 6.15. This may be considered as a parameterized C^r or C^∞ n-simplex in W, as appropriate.

Under these conditions, we may use the notation

$$\Psi = \eta \circ \Phi. \tag{7.10.7}$$

Of course, if ϕ and η are both affine mappings, then

$$\psi \text{ is affine} \tag{7.10.8}$$

too, as in Section 7.1. Sometimes we may be concerned with cases where ϕ is affine and η may not be, which is related to some of the remarks in Section 10.30 starting on p269 of [155].

7.10.1 Chains and compositions

Suppose that Γ is a C^r or C^∞ n-chain in V as in (7.9.8). Thus

$$\Psi_l = \eta \circ \Phi_l \tag{7.10.9}$$

is a parameterized C^r or C^∞ n-simplex in W for each $l = 1, \ldots, k$, as appropriate. Put

$$\eta \circ \Gamma = \sum_{l=1}^{k} \epsilon_l \, \Psi_l, \tag{7.10.10}$$

which is a C^r or C^∞ n-chain in W. This is related to some remarks on p270 of [155].

Let β be a differential n-form on W that is continuous on W, so that $\eta^*(\beta)$ is a differential n-form on V that is continuous on V. If Γ is a C^1 n-chain in V, then

$$\int_{\eta \circ \Gamma} \beta = \int_\Gamma \eta^*(\beta). \tag{7.10.11}$$

Of course, both sides of the equation are defined as in (7.9.10). This follows from the analogous statement for integrals over n-surfaces in Subsection 6.15.3.

7.11 Boundaries of C^r simplices

Let n be a positive integer, let $e_1, \ldots, e_n$ be the standard basis in $\mathbf{R}^n$, and put $e_0 = 0$, as before. Consider the parameterized affine n-simplex

$$\Phi(e_0, e_1, \ldots, e_n) = [e_0, e_1, \ldots, e_n] \tag{7.11.1}$$

in $\mathbf{R}^n$ associated to $e_0, e_1, \ldots, e_n$ as in Section 7.3. The corresponding affine mapping

$$\phi_{e_0, e_1, \ldots, e_n} \tag{7.11.2}$$

from $\mathbf{R}^n$ into itself is the identity mapping on $\mathbf{R}^n$. In particular, this maps the standard simplex Q^n in $\mathbf{R}^n$ onto itself.

We may consider (7.11.1) as the *standard parameterized affine n-simplex* in $\mathbf{R}^n$. Its boundary

$$\partial \Phi(e_0, e_1, \ldots, e_n) = \sum_{l=0}^{n} (-1)^l \, \Phi(e_0, \ldots, e_{l-1}, e_{l+1}, \ldots, e_n) \tag{7.11.3}$$

is an affine $(n-1)$-chain in $\mathbf{R}^n$, as in Section 7.7. This may be called

(7.11.4) the *positively oriented boundary* of Q^n,

as in Section 10.31 on p270 of [155].

7.11.1 C^r Simplices as compositions

Let m be another positive integer, and let U, V be nonempty open subsets of $\mathbf{R}^n$, $\mathbf{R}^m$, respectively, with $Q^n \subseteq U$. Also let ϕ be a mapping from U into $\mathbf{R}^m$ such that $\phi(Q^n) \subseteq V$, and suppose that ϕ is r-times continuously differentiable on U for some nonnegative integer r, or infinitely differentiable on U. This defines a parameterized C^r or C^∞ n-simplex Φ in V, as appropriate, as in Section 7.9. Observe that

$$\Phi = \phi \circ \Phi(e_0, e_1, \ldots, e_n), \tag{7.11.5}$$

using the notation in the previous section on the right side. More precisely, (7.11.2) corresponds to ϕ in the previous section, and ϕ here corresponds to η in the previous section.

Under these conditions, the *boundary* of Φ may be defined as a C^r or C^∞ $(n-1)$-chain in V, as appropriate, by

$$\begin{aligned} \partial\Phi &= \phi \circ (\partial\Phi(e_0, e_1, \ldots, e_n)) \\ &= \sum_{l=0}^{n} (-1)^l \, \phi \circ \Phi(e_0, \ldots, e_{l-1}, e_{l+1}, \ldots, e_n), \end{aligned} \tag{7.11.6}$$

where the right side is as in the previous section. This basically corresponds to the definition of the boundary in (89) on p270 of [155]. One can check that this also corresponds to the definition of the boundary in (4) on p142 of [184]. To see this, let us look a bit more closely at (7.11.3).

If $1 \le l \le n$, then the lth term on the right side of (7.11.3) uses the parameterized affine $(n-1)$-simplex

$$\Phi(e_0, \ldots, e_{l-1}, e_{l+1}, \ldots, e_n) = [e_0, \ldots, e_{l-1}, e_{l+1}, \ldots, e_n] \tag{7.11.7}$$

in $\mathbf{R}^n$. If $n \ge 2$, then this corresponds to the affine mapping

$$\phi_l = \phi_{e_0, \ldots, e_{l-1}, e_{l+1}, \ldots, e_n} \tag{7.11.8}$$

from $\mathbf{R}^{n-1}$ into $\mathbf{R}^n$ defined by

$$\begin{aligned} \phi_l(x) &= e_0 + \sum_{j=1}^{l-1} x_j \, (e_j - e_0) + \sum_{j=l}^{n-1} x_j \, (e_{j+1} - e_0) \\ &= \sum_{j=1}^{l-1} x_j \, e_j + \sum_{j=l}^{n-1} x_j \, e_{j+1}, \end{aligned} \tag{7.11.9}$$

where the first step is as in (7.7.9). Similarly, if $l = 0$, then the right side of (7.11.3) uses the parameterized affine $(n-1)$-simplex

$$\Phi(e_1, \ldots, e_n) = [e_1, \ldots, e_n] \tag{7.11.10}$$

in $\mathbf{R}^n$. If $n \ge 2$, then this corresponds to the affine mapping

$$\phi_0 = \phi_{e_1, \ldots, e_n} \tag{7.11.11}$$

from $\mathbf{R}^{n-1}$ into $\mathbf{R}^n$ defined by

$$\phi_0(x) = e_1 + \sum_{j=1}^{n-1} x_j \, (e_{j+1} - e_1), \tag{7.11.12}$$

as in (7.7.12). One can use (7.11.9) and (7.11.12) to verify that (7.11.6) corresponds to the definition of the boundary in (4) on p142 of [184].

If ϕ is an affine mapping from $\mathbf{R}^n$ into $\mathbf{R}^m$, then Φ is a parameterized affine n-simplex in $\mathbf{R}^m$. In this case, one can check that (7.11.6) is equivalent to the definition of the boundary of Φ as an affine $(n-1)$-chain in $\mathbf{R}^m$, as in Section 7.7. This corresponds to a remark on p270 of [155].

7.11.2 Boundaries of C^r chains

If $\Gamma = \sum_{l=1}^k \epsilon_l \, \Phi_l$ is a C^r or C^∞ n-chain in V, then the *boundary* of Γ may be defined as a C^r or C^∞ $(n-1)$-chain in V, as appropriate, by

$$\partial \Gamma = \sum_{l=1}^k \epsilon_l \, \partial \Phi_l, \tag{7.11.13}$$

as in Subsection 7.7.1. This corresponds to (90) on p270 of [155], and to a remark on p142 of [184].

7.12 Compositions and affine simplices

Let m and n be positive integers, and let $p_0, p_1, \ldots, p_n$ be elements of $\mathbf{R}^m$. Using this, we get an affine mapping $\phi_{p_0,p_1,\ldots,p_n}$ from $\mathbf{R}^n$ into $\mathbf{R}^m$ as in Section 7.3, and the corresponding parameterized affine n-simplex $\Phi(p_0, p_1, \ldots, p_n)$ in $\mathbf{R}^m$. Let V be a nonempty open subset of $\mathbf{R}^m$, and suppose that

$$\phi_{p_0,p_1,\ldots,p_n}(Q^n) \subseteq V, \tag{7.12.1}$$

so that $\Phi(p_0, p_1, \ldots, p_n)$ may be considered as a parameterized affine n-simplex in V.

Let q be another positive integer, and let η be a mapping from V into $\mathbf{R}^q$ that is r-times continuously differentiable for some nonnegative integer r, or infinitely differentiable. Thus

$$\psi = \eta \circ \phi_{p_0,p_1,\ldots,p_n} \tag{7.12.2}$$

defines a mapping from

$$\phi^{-1}_{p_0,p_1,\ldots,p_n}(V) \tag{7.12.3}$$

into $\mathbf{R}^q$ that is r-times continuously differentiable or infinitely differentiable, as appropriate. Let W be a nonempty open subset of $\mathbf{R}^q$ with $\eta(V) \subseteq W$, so that

$$\psi(Q^n) = \eta(\phi_{p_0,p_1,\ldots,p_n}(Q^n)) \subseteq \eta(V) \subseteq W. \tag{7.12.4}$$

Using ψ, we get a parameterized C^r or C^∞ n-simplex

$$\Psi = \eta \circ \Phi(p_0, p_1, \ldots, p_n) \tag{7.12.5}$$

in W, as appropriate, as in Section 7.10.

7.12.1 Defining $\simeq$ for C^r simplices

Let τ be a permutation on $\{0, 1, \ldots, n\}$, and remember that

$$\phi_{p_{\tau(0)}, p_{\tau(1)}, \ldots, p_{\tau(n)}}(Q^n) = \phi_{p_0, p_1, \ldots, p_n}(Q^n), \tag{7.12.6}$$

as in Sections 7.3 and 7.4. In particular,

$$\phi_{p_{\tau(0)}, p_{\tau(1)}, \ldots, p_{\tau(n)}}(Q^n) \subseteq V, \tag{7.12.7}$$

so that

$$\eta \circ \Phi(p_{\tau(0)}, p_{\tau(1)}, \ldots, p_{\tau(n)}) \tag{7.12.8}$$

may be considered as a parameterized C^r or C^∞ n-simplex in W, as appropriate, as before. Under these conditions, let us put

$$\eta \circ \Phi(p_{\tau(0)}, p_{\tau(1)}, \ldots, p_{\tau(n)}) \simeq \operatorname{sgn}(\tau)\,(\eta \circ \Phi(p_0, p_1, \ldots, p_n)). \tag{7.12.9}$$

This extends the notation in Subsection 7.5.1 from parameterized affine n-simplices to parameterized C^r and C^∞ n-simplices. If η is an affine mapping from $\mathbf{R}^m$ into $\mathbf{R}^q$, then (7.12.5) and (7.12.8) are parameterized affine n-simplices in $\mathbf{R}^q$, and (7.12.9) is equivalent to the analogous notation in Subsection 7.5.1.

7.12.2 Defining $\simeq$ for C^r chains

Suppose that

$$\Gamma = \sum_{l=1}^{k} \epsilon_l\, \Phi_l, \ \widetilde{\Gamma} = \sum_{l=1}^{k} \widetilde{\epsilon}_l\, \widetilde{\Phi}_l \tag{7.12.10}$$

are C^r or C^∞ n-chains in W, as in Subsection 7.9.1. If

$$\epsilon_l\, \Phi_l \simeq \widetilde{\epsilon}_l\, \widetilde{\Phi}_l \tag{7.12.11}$$

for each $l = 1, \ldots, k$, then we put

$$\Gamma \simeq \widetilde{\Gamma}, \tag{7.12.12}$$

as in Subsection 7.6.2. In this case, one can check that

$$\partial\Gamma \simeq \partial\widetilde{\Gamma}, \tag{7.12.13}$$

as in Subsection 7.8.1.

7.13 Some remarks about n-chains

Let m and n be positive integers, and let V be a nonempty open subset of $\mathbf{R}^m$. If r is a nonnegative integer, then

$$\text{the space of } C^r \ n\text{-chains in } V \tag{7.13.1}$$

may be defined as the free abelian group generated by the parameterized C^r n-simplices in V. This is often used when $r = 0$ in algebraic topology, as on p192 of [184]. Similarly, the spaces of affine and C^∞ n-chains in V may be defined as the free abelian groups generated by the parameterized affine and C^∞ n-simplices in V, respectively. Although the details of this will not be discussed here, we would like to mention some related aspects of this in this section.

One could also consider finite linear combinations of parameterized affine, C^r, or C^∞ n-chains in V with coefficients in the real numbers, as on p142 of [184], and as mentioned in Subsection 7.9.1. The spaces of these finite linear combinations may be defined as the free vector spaces over the real numbers generated by the parameterized affine, C^r, or C^∞ n-simplices, as appropriate. We shall not get into the details of this here either, aside from some related aspects in this section.

7.13.1 Linear functionals on differential forms

Let Φ be an n-surface in V, as in Section 6.15. If α is a differential n-form on V that is continuous on V, then the integral of α over Φ may be defined as a real number as in Subsection 6.15.1. Thus

$$\alpha \mapsto \int_\Phi \alpha \tag{7.13.2}$$

defines a real-valued function on the space

$$C(V, \mathcal{AM}_n(\mathbf{R}^m)) \tag{7.13.3}$$

of all differential n-forms on V that are continuous on V, as mentioned in Section 10.28 on p268 of [155]. More precisely, (7.13.4) defines a linear functional on (7.13.3), which is to say a linear mapping from (7.13.3) into $\mathbf{R}$.

Similarly, if Γ is a C^1 n-chain in V, then the integral of α over Γ may be defined as a real number as in Subsection 7.9.1. This means that

$$\alpha \mapsto \int_\Gamma \alpha \tag{7.13.4}$$

defines a real-valued function on (7.13.3) too, which is a linear functional as well. By construction, this is a linear combination of linear functionals on (7.13.3) associated to finitely many parameterized C^1 n-simplices in V as in (7.13.2).

7.13.2 Defining $\cong$ for C^1 n-chains

If Γ_1, Γ_2 are C^1 n-chains in V, then put

$$\Gamma_1 \cong \Gamma_2 \tag{7.13.5}$$

when

$$\int_{\Gamma_1} \alpha = \int_{\Gamma_2} \alpha \tag{7.13.6}$$

for every differential n-form α on V that is continuous on V. This is the same as saying that the functions on (7.13.3) associated to Γ_1 and Γ_2 as in (7.13.4) are the same.

One could simply look at a C^1 n-chain in V as a type of linear functional on (7.13.3), as in Section 10.28 on p268 of [155]. If the space of C^1 n-chains in V is considered as a commutative group with respect to addition, then we get a homomorphism from this group into the space of linear functionals on (7.13.3), as another commutative group with respect to addition. If the space of C^1 n-chains in V is defined using coefficients in $\mathbf{R}$, then we get a linear map from that space into the space of linear functionals on (7.13.3).

7.13.3 Comparing $\simeq$ and $\cong$

Let Φ, $\widetilde{\Phi}$ be parameterized C^1 n-simplices in V such that

$$\epsilon\,\Phi \simeq \widetilde{\epsilon}\,\widetilde{\Phi} \tag{7.13.7}$$

for some $\epsilon, \widetilde{\epsilon} = \pm 1$, as in Subsection 7.12.1. Under these conditions, one can check that

$$\epsilon\,\Phi \cong \widetilde{\epsilon}\,\widetilde{\Phi}. \tag{7.13.8}$$

This follows from a remark in Section 7.5 when Φ and $\widetilde{\Phi}$ are parameterized affine n-simplices in V. One can reduce to that case as in Subsection 6.15.3. If Γ_1, Γ_2 are C^1 n-chains in V that satisfy (7.12.12), then it follows that (7.13.5) holds as well.

7.14 An associativity property

Let m and n be positive integers again, and let U, V be nonempty open subsets of $\mathbf{R}^n$, $\mathbf{R}^m$, respectively, with $Q^n \subseteq U$. Also let ϕ be a mapping from U into $\mathbf{R}^m$ such that $\phi(Q^n) \subseteq V$ and ϕ is r-times continuously differentiable on U for some nonnegative integer r, or infinitely differentiable on U, so that we get a parameterized C^r or C^∞ n-simplex Φ in V, as appropriate, as in Section 7.9. Let q be another positive integer, let η be a mapping from V into $\mathbf{R}^q$ that is r-times continuously differentiable or infintely differentiable, as appropriate, and let W be a nonempty open subset of $\mathbf{R}^q$ with $\eta(V) \subseteq W$. Thus $\psi = \eta \circ \phi$ defines a mapping from $\phi^{-1}(V)$ into $\mathbf{R}^q$ that is r-times continuously differentiable or infinitely differentiable, as appropriate, which leads to a parameterized C^r or C^∞ n-simplex $\Psi = \eta \circ \Phi$ in W, as in Section 7.10.

Similarly, let q_1 be a positive integer, let η_1 be a mapping from W into $\mathbf{R}^{q_1}$ that is r-times continuously differentiable or infinitely differentiable, as appropriate, and let W_1 be a nonempty open subset of $\mathbf{R}^{q_1}$ with

$$\eta_1(W) \subseteq W_1. \tag{7.14.1}$$

Under these conditions,

$$\psi_1 = \eta_1 \circ \psi \tag{7.14.2}$$

defines a mapping from $\phi^{-1}(V)$ into $\mathbf{R}^{q_1}$ that is r-times continuously differentiable or infinitely differentiable, as appropriate, with

$$\psi_1(Q^n) = \eta_1(\psi(Q^n)) \subseteq \eta_1(W) \subseteq W_1. \tag{7.14.3}$$

As before, we can use ψ_1 to get a parameterized C^r or C^∞ n-simplex Ψ_1 in W_1, as appropriate, which may be expressed as

$$\Psi_1 = \eta_1 \circ \Psi. \tag{7.14.4}$$

Observe that

$$\zeta = \eta_1 \circ \eta \tag{7.14.5}$$

defines a mapping from V into $\mathbf{R}^{q_1}$ that is r-times continuously differentiable or infinitely differentiable, as appropriate, with

$$\zeta(V) \subseteq W_1. \tag{7.14.6}$$

This means that

$$\zeta \circ \phi \tag{7.14.7}$$

defines a mapping from $\phi^{-1}(V)$ into $\mathbf{R}^{q_1}$ that is r-times continuously differentiable or infinitely differentiable, as appropriate. This determines a parameterized C^r or C^∞ n-simplex $\zeta \circ \Phi$ in W_1, as in Section 7.10 again. Of course,

$$\zeta \circ \phi = \eta_1 \circ \eta \circ \phi = \eta_1 \circ \psi = \psi_1, \tag{7.14.8}$$

so that

$$\zeta \circ \Phi = \Psi_1. \tag{7.14.9}$$

7.14.1 Associativity and chains

Let Γ be a C^r or C^∞ n-chain in V, so that $\eta \circ \Gamma$ is a C^r or C^∞ n-chain in W, as appropriate, as in Subsection 7.10.1. Similarly $\eta_1 \circ (\eta \circ \Gamma)$ and $\zeta \circ \Gamma$ are C^r or C^∞ n-chains in W_1, as appropriate. It is easy to see that

$$\eta_1 \circ (\eta \circ \Gamma) = (\eta_1 \circ \eta) \circ \Gamma, \tag{7.14.10}$$

as in (7.14.9).

7.15 Compositions and boundaries

Let us return to the notation and hypotheses at the beginning of the previous section. In particular, $\Psi = \eta \circ \Phi$ is a parameterized C^r or C^∞ n-simplex in W, as appropriate. Also let $e_1, \ldots, e_n$ be the standard basis in $\mathbf{R}^n$, and put $e_0 = 0$, as before. Remember that

$$\partial \Phi = \phi \circ (\partial \Phi(e_0, e_1, \ldots, e_n)), \tag{7.15.1}$$

as in (7.11.6). Similarly,

$$\partial\Psi = \psi \circ (\partial\Phi(e_0, e_1, \ldots, e_n)). \tag{7.15.2}$$

This means that

$$\begin{aligned} \partial(\eta \circ \Phi) &= (\eta \circ \phi) \circ (\partial\Phi(e_0, e_1, \ldots, e_n)) \\ &= \eta \circ (\phi \circ (\partial\Phi(e_0, e_1, \ldots, e_n))) = \eta \circ (\partial\Phi), \end{aligned} \tag{7.15.3}$$

where the second step is as in (7.14.10). This is related to some of the remarks on p270 of [155].

If Γ is a C^r or C^∞ n-chain in V, then $\eta \circ \Gamma$ is a C^r or C^∞ n-chain in W, as appropriate, as in Subsection 7.10.1. Similarly, $\eta \circ (\partial\Gamma)$ is a C^r or C^∞ $(n-1)$-chain in W, as appropriate. Under these conditions, it is easy to see that

$$\partial(\eta \circ \Gamma) = \eta \circ (\partial\Gamma), \tag{7.15.4}$$

using (7.15.3).

7.15.1 Boundaries of boundaries

If $p_0, p_1, \ldots, p_n$ are elements of $\mathbf{R}^m$, then one can check that

$$\partial(\partial\Phi(p_0, p_1, \ldots, p_n)) = 0, \tag{7.15.5}$$

as in Exercise 16 on p291 of [155]. More precisely, this is trivial when $n \le 1$, because the boundary of a 0-simplex or 0-chain is interpreted as being equal to 0, as in Section 7.7. If ϕ and Φ are as before, then we get that

$$\begin{aligned} \partial(\partial\Phi) &= \partial(\phi \circ (\partial\Phi(e_0, e_1, \ldots, e_n))) \\ &= \phi \circ (\partial(\partial\Phi(e_0, e_1, \ldots, e_n))) = 0, \end{aligned} \tag{7.15.6}$$

using (7.11.6) in the first step, and (7.15.3) in the second step. This implies that

$$\partial(\partial\Gamma) = 0 \tag{7.15.7}$$

for every C^r or C^∞ n-chain Γ in $\mathbf{R}^m$. This corresponds to the second part of Exercise 16 on p291 of [155], and to (7) on p143 of [184].

Chapter 8

Stokes' theorem and other matters

8.1 Stokes' theorem

Let m and n be positive integers, and let V be a nonempty open set in $\mathbf{R}^m$. Also let

(8.1.1) ω be a continuously-differentiable differential $(n-1)$-form on V,

so that the exterior derivative $d\omega$ of ω is a differential n-form on V that is continuous on V, as in Section 4.6. If

$$\Gamma \text{ is a } C^2 \ n\text{-chain in } V, \tag{8.1.2}$$

then *Stokes' theorem* states that

$$\int_{\partial\Gamma} \omega = \int_\Gamma d\omega. \tag{8.1.3}$$

This corresponds to Theorem 10.33 on p272 of [155], and Theorem 4.7 on p144 of [184]. More precisely, if $n = 1$, then it suffices to ask that

$$\Gamma \text{ be a } C^1 \ 1\text{-chain in } V, \tag{8.1.4}$$

instead of (8.1.2).

In particular, if

$$\partial\Gamma = 0, \tag{8.1.5}$$

then

$$\int_\Gamma d\omega = 0, \tag{8.1.6}$$

by (8.1.4). Similarly, let Γ_1, Γ_2 be C^2 n-chains in V, or C^1 1-chains in V. If

$$\partial\Gamma_1 = \partial\Gamma_2, \tag{8.1.7}$$

then

$$\int_{\Gamma_1} d\omega = \int_{\Gamma_2} d\omega. \tag{8.1.8}$$

This corresponds to Remark (b) on p276 of [155].

Suppose now that ω is a closed $(n-1)$-form on V, so that

$$d\omega = 0 \tag{8.1.9}$$

on V. In this case, (8.1.3) reduces to

$$\int_{\partial\Gamma} \omega = 0. \tag{8.1.10}$$

This corresponds to Remark (c) on p276 of [155].

8.1.1 Stokes' theorem and $\cong$

Let Γ_1, Γ_2 be C^2 n-chains in V again, or C^1 1-chains in V. Suppose that $\Gamma_1 \cong \Gamma_2$, in the sense of Subsection 7.13.2. If ω is as in (8.1.1), then we have that (8.1.8) holds, by hypothesis. Under these conditions, Stokes' theorem implies that

$$\int_{\partial\Gamma_1} \omega = \int_{\partial\Gamma_2} \omega. \tag{8.1.11}$$

One can use this to get that

$$\partial\Gamma_1 \cong \partial\Gamma_2, \tag{8.1.12}$$

in the sense of Subsection 7.13.2 again. More precisely, (8.1.12) was defined to mean that (8.1.11) holds for all differential $(n-1)$-forms ω on V that are continuous on V, and not just when ω is continuously differentiable on V. Sometimes this type of property is defined using smooth differential forms on V, for which (8.1.11) could be used directly. One can also approximate continuous differential forms on V, uniformly on compact subsets of V, by differential forms that are continuously differentiable, or even smooth, using partitions of unity, for instance. Using this, one can get (8.1.11) when ω is continuous from the analogous statement when ω is continuously differentiable.

8.2 More on Stokes' theorem

Let us continue with the same notation and hypotheses as at the beginning of the previous section. Let U be a open subset of $\mathbf{R}^n$ with $Q^n \subseteq U$, and let ϕ be a mapping from U into $\mathbf{R}^m$ such that $\phi(Q^n) \subseteq V$ and

(8.2.1) ϕ is twice continuously differentible on U,

so that we get

(8.2.2) a parameterized C^2 n-simplex Φ in V, as in Section 7.9. In this case, Stokes' theorem says that

$$\int_{\partial\Phi} \omega = \int_{\Phi} d\omega, \tag{8.2.3}$$

as in (92) on p273 of [155], and (2) on p144 of [184]. If $n = 1$, then it suffices to ask that

$$\phi \text{ be continuously differentiable on } U, \tag{8.2.4}$$

as before. It is easy to see that this version of Stokes' theorem implies the previous one for n-chains in V, as in [155, 184].

Let $e_1, \ldots, e_n$ be the standard basis in $\mathbf{R}^n$, put $e_0 = 0$, and let

$$\Phi_0 = \Phi(e_0, e_1, \ldots, e_n) \tag{8.2.5}$$

be the standard parameterized affine n-simplex in $\mathbf{R}^n$. Thus

$$\Phi = \phi \circ \Phi_0 \tag{8.2.6}$$

and

$$\partial\Phi = \phi \circ (\partial\Phi_0), \tag{8.2.7}$$

as in Subsection 7.11.1. Remember that

$$U_0 = \phi^{-1}(V) \tag{8.2.8}$$

is an open set in $\mathbf{R}^n$ that contains Q^n, as in Section 7.10.

Of course, $\phi^*(\omega)$ is a differential $(n-1)$-form on $\phi^{-1}(V)$, as in Subsection 4.4.1. More precisely,

$$\phi^*(\omega) \text{ is continuously differentiable on } U_0, \tag{8.2.9}$$

as in Subsection 4.5.2, because ω is continuously differentiable on V, and ϕ is twice continuously differentiable on U, by hypothesis. If $n = 1$, then it suffices to ask that ϕ be continuously differentiable on U, as before.

Remember that

$$\int_{\Phi} d\omega = \int_{\Phi_0} \phi^*(d\omega), \tag{8.2.10}$$

as in Subsection 6.15.3. We also have that

$$\phi^*(d\omega) = d(\phi^*(\omega)) \tag{8.2.11}$$

on $\phi^{-1}(V)$, as in Section 4.9. Thus

$$\int_{\Phi} d\omega = \int_{\Phi_0} d(\phi^*(\omega)). \tag{8.2.12}$$

Similarly,

$$\int_{\partial\Phi} \omega = \int_{\partial\Phi_0} \phi^*(\omega), \tag{8.2.13}$$

because of (8.2.7). If $n = 1$, then this involves integrals over parameterized 0-simplices, as in Subsection 7.5.2.

This means that (8.2.3) is the same as saying that

$$\int_{\partial \Phi_0} \phi^*(\omega) = \int_{\Phi_0} d(\phi^*(\omega)). \tag{8.2.14}$$

Thus we can reduce to the case where $\Phi = \Phi_0$. This corresponds to (94) on p273 of [155], and to (3) on p144 of [184]. If $n = 1$, then this is the same as the fundamental theorem of calculus, as in [155, 184].

8.3 Stokes' theorem for Φ_0

Let $n \geq 2$ be an integer, let $e_1, \ldots, e_n$ be the standard basis for $\mathbf{R}^n$, put $e_0 = 0$, and let Φ_0 be the standard parameterized n-simplex in $\mathbf{R}^n$, as in (8.2.5). Also let V be an open set in $\mathbf{R}^n$ with

$$Q^n \subseteq V, \tag{8.3.1}$$

and let ω be a continuously-differentiable differential $(n-1)$-form on V. We would like to show that

$$\int_{\partial \Phi_0} \omega = \int_{\Phi_0} d\omega, \tag{8.3.2}$$

which corresponds to (8.2.3) with $m = n$ and $\Phi = \Phi_0$. This will imply Stokes' theorem, as in the previous section.

Let k be an integer with $1 \leq k \leq n$, and suppose that ω is of the form

$$\omega = \omega_k(x)\, dx_1 \wedge \cdots \wedge dx_{k-1} \wedge dx_{k+1} \wedge \cdots \wedge dx_n, \tag{8.3.3}$$

where

(8.3.4) $\omega_k(x)$ is a continuously-differentiable real-valued function on V.

Note that every continuously-differentiable differential $(n-1)$-form on V can be expressed a sum of differential forms of this type. Thus it suffices to show that (8.3.2) holds in this case, because of the linearity of the integrals on both sides of the equation. This corresponds to (96) on p273 of [155], and to a remark shortly after (5) on p144 of [184].

8.3.1 The integral over $\partial \Phi_0$

Remember that

$$\partial \Phi_0 = \sum_{l=0}^{n} (-1)^l \, \Phi(e_0, \ldots, e_{l-1}, e_{l+1}, \ldots, e_n), \tag{8.3.5}$$

as in Section 7.11. The parameterized affine $(n-1)$-simplex

$$\Phi(e_0, \ldots, e_{l-1}, e_{l+1}, \ldots, e_n) \tag{8.3.6}$$

in $\mathbf{R}^n$ corresponds to the affine mapping

$$\phi_l = \phi_{e_0,\ldots,e_{l-1},e_{l+1},\ldots,e_n} \tag{8.3.7}$$

from $\mathbf{R}^{n-1}$ into $\mathbf{R}^n$ for each $l = 0, 1, \ldots, n$, as in Section 7.3. We have seen that

$$\phi_0(y) = e_1 + \sum_{j=1}^{n-1} y_j \, (e_{j+1} - e_1) \tag{8.3.8}$$

for every $y \in \mathbf{R}^{n-1}$, as in Subsection 7.11.1. Similarly, if $1 \le l \le n$, then

$$\phi_l(y) = \sum_{j=1}^{l-1} y_j \, e_j + \sum_{j=l}^{n-1} y_j \, e_{j+1} \tag{8.3.9}$$

for every $y \in \mathbf{R}^{n-1}$, as before.

Observe that

$$\int_{\partial \Phi_0} \omega = \sum_{l=0}^{n} (-1)^l \int_{\Phi(e_0,\ldots,e_{l-1},e_{l+1},\ldots,e_n)} \omega = \sum_{l=0}^{n} (-1)^l \int_{Q^{n-1}} \phi_l^*(\omega), \tag{8.3.10}$$

where the first step is as in Subsection 7.6.1, and the second step is as in Subsection 6.15.1. If $l \ge 1$, then it is easy to see that

$$\phi_l^*(dx_l) = 0, \tag{8.3.11}$$

because the lth component of ϕ_l is equal to 0, as in (8.3.9). This implies that

$$\phi_l^*(\omega) = 0 \tag{8.3.12}$$

when $l \ne 0, k$ and ω is as in (8.3.3). Thus (8.3.10) reduces to

$$\int_{\partial \Phi_0} \omega = \int_{Q^{n-1}} \phi_0^*(\omega) + (-1)^k \int_{Q^{n-1}} \phi_k^*(\omega) \tag{8.3.13}$$

in this case. This basically corresponds to the first line in (102) on p274 of [155], and to part of (9) on p145 of [184].

8.3.2 The integral over Φ_0

If ω is as in (8.3.3), then

$$d\omega = d\omega_k \wedge dx_1 \wedge \cdots \wedge dx_{k-1} \wedge dx_{k+1} \wedge \cdots \wedge dx_n, \tag{8.3.14}$$

as in Section 4.6. This implies that

$$d\omega = (\partial_k \omega_k) \, dx_k \wedge dx_1 \wedge \cdots \wedge dx_{k-1} \wedge dx_{k+1} \wedge \cdots \wedge dx_n. \tag{8.3.15}$$

It follows that

$$d\omega = (-1)^{k-1} \, (\partial_k \omega_k) \, dx_1 \wedge \cdots \wedge dx_n. \tag{8.3.16}$$

This corresponds to remarks on p275 of [155], and on p144 of [184].

Using (8.3.16), we get that

$$\int_{\Phi_0} d\omega = (-1)^{k-1} \int_{Q^n} (\partial_k \omega_k)(x)\, dx, \tag{8.3.17}$$

as in Section 6.12. If $x \in Q^n$, then

$$(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) \tag{8.3.18}$$

is an element of Q^{n-1}, as in Section 6.2. Let an element of Q^{n-1} be given as in (8.3.18), and consider

$$\int_0^{1-x_1-\cdots-x_{k-1}-x_{k+1}-\cdots-x_n} (\partial_k \omega_k)(x_1, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_n)\, dx_k. \tag{8.3.19}$$

This is equal to the sum of

$$\omega_k(x_1, \ldots, x_{k-1}, 1 - x_1 - \cdots - x_{k-1} - x_{k+1} - \cdots - x_n, x_{k+1}, \ldots, x_n) \tag{8.3.20}$$

and

$$-\omega_k(x_1, \ldots, x_{k-1}, 0, x_{k+1}, \ldots, x_n), \tag{8.3.21}$$

by the fundamental theorem of calculus.

Observe that

$$\int_{Q^n} (\partial_k \omega_k)(x)\, dx \tag{8.3.22}$$

is equal to the integral of (8.3.19), as a function of (8.3.18), over Q^{n-1}. This is related to some of the remarks in Section 6.2. Thus (8.3.22) is equal to the sum of the integrals of (8.3.20) and (8.3.21) over Q^{n-1}, as in the preceding paragraph.

8.3.3 Matching a pair of terms

If ω is as in (8.3.3), then it is easy to see that

$$\phi_k^*(\omega) = (\omega_k \circ \phi_k)\, dy_1 \wedge \cdots \wedge dy_{n-1}, \tag{8.3.23}$$

using (8.3.9), with $l = k$. This means that

$$\begin{aligned} (-1)^k \int_{Q^{n-1}} \phi_k^*(\omega) \\ = (-1)^k \int_{Q^{n-1}} \omega_k(y_1, \ldots, y_{k-1}, 0, y_k, \ldots, y_{n-1})\, dy, \end{aligned} \tag{8.3.24}$$

as in Section 6.12. Of course, this is the same as $(-1)^{k-1}$ times the integral of (8.3.21) over Q^{n-1}.

8.3.4 Matching the other two terms

If $y \in \mathbf{R}^{n-1}$, then

$$\phi_0(y) = \Big(1 - \sum_{j=1}^{n-1} y_j\Big)\, e_1 + \sum_{j=1}^{n-1} y_j\, e_{j+1}, \tag{8.3.25}$$

as in (8.3.8). This implies that

$$\phi_0^*(dx_1) = -\sum_{j=1}^{n-1} dy_j \tag{8.3.26}$$

and

$$\phi_0^*(dx_l) = dy_{l-1} \tag{8.3.27}$$

for $l = 2, \ldots, n$. If ω is as in (8.3.3) and $k = 1$, then we get that

$$\phi_0^*(\omega) = (\omega_1 \circ \phi_0)\, dy_1 \wedge \cdots \wedge dy_{n-1}. \tag{8.3.28}$$

If $k \ge 2$, then we obtain that

$$\phi_0^*(\omega) =$$

$$-(\omega_k \circ \phi_0) \Big(\sum_{j=1}^{n-1} dy_j\Big) \wedge dy_1 \wedge \cdots \wedge dy_{k-2} \wedge dy_k \wedge \cdots \wedge dy_{n-1}. \tag{8.3.29}$$

This means that

$$\phi_0^*(\omega) = -(\omega_k \circ \phi_0)\, dy_{k-1} \wedge dy_1 \wedge \cdots \wedge dy_{k-2} \wedge dy_k \wedge \cdots \wedge dy_{n-1}, \tag{8.3.30}$$

because all but one of the terms in the sum on the right side of (8.3.29) leads to a wedge product that is equal to 0. Thus

$$\phi_o^*(\omega) = (-1)^{k-1}\, (\omega_k \circ \phi_0)\, dy_1 \wedge \cdots \wedge dy_{n-1}, \tag{8.3.31}$$

by rearranging the wedge product on the right side of (8.3.30) appropriately. Note that this is the same as (8.3.28) when $k = 1$.

It follows that

$$\int_{Q^{n-1}} \phi_0^*(\omega) = (-1)^{k-1} \int_{Q^{n-1}} \omega_k(\phi_0(y))\, dy. \tag{8.3.32}$$

If $k = 1$, then this corresponds exactly to $(-1)^{k-1}$ times the integral of (8.3.20) over Q^{n-1}. If $k \ge 2$, then we basically need an additional change of variables, as on p274 of [155], and p145 of [184]. More precisely, one can use a change of variables as in (10) on p145 of [184] to get that (8.3.32) is equal to $(-1)^{k-1}$ times the integral of (8.3.20) over Q^{n-1}. This is related to some of the remarks in Subsection 7.4.1.

On p274 of [155], one considers a permutation of the standard basis vectors so that the argument is like the $k = 1$ case before. This will be discussed further in the next section.

8.4 A related argument

Let us continue with the same notation and hypotheses as in the previous section. Remember that

$$\int_{Q^{n-1}} \phi_0^*(\omega) = \int_{\Phi(e_1,\ldots,e_n)} \omega, \tag{8.4.1}$$

as in Subsection 6.15.1, which corresponds to the $l = 0$ term in the second step in (8.3.10), and in the right side of (8.3.13). Let τ be the permutation on $\{1,\ldots,n\}$ defined by

$$\begin{aligned} \tau(l) &= k && \text{when } l = 1 \\ &= l-1 && \text{when } l = 2,\ldots,k \\ &= l && \text{when } l = k+1,\ldots,n. \end{aligned} \tag{8.4.2}$$

One can check that

$$\operatorname{sgn}(\tau) = (-1)^{k-1}. \tag{8.4.3}$$

As in Subsection 7.5.1, we have that

$$\Phi(e_{\tau(1)},\ldots,e_{\tau(n)}) \simeq \operatorname{sgn}(\tau)\,\Phi(e_1,\ldots,e_n). \tag{8.4.4}$$

More precisely, the n in Subsection 7.5.1 should be taken to be $n-1$ here, so that $\{0,1,\ldots,n\}$ before should be taken to be $\{0,1,\ldots,n-1\}$ now. However, here we are basically using $\{1,\ldots,n\}$ instead. Equivalently, we get that

$$\Phi(e_k,e_1,\ldots,e_{k-1},e_{k+1},\ldots,e_n) \simeq (-1)^{k-1}\,\Phi(e_1,\ldots,e_n). \tag{8.4.5}$$

This corresponds to a remark on p274 of [155].

We also have that

$$\int_{\Phi(e_k,e_1,\ldots,e_{k-1},e_{k+1},\ldots,e_n)} \omega = (-1)^{k-1} \int_{\Phi(e_1,\ldots,e_n)} \omega, \tag{8.4.6}$$

as in Section 7.5. Remember that the parameterized affine $(n-1)$-simplex

$$\Phi(e_k,e_1,\ldots,e_{k-1},e_{k+1},\ldots,e_n) \tag{8.4.7}$$

corresponds to the affine mapping

$$\psi_k = \phi_{e_k,e_1,\ldots,e_{k-1},e_{k+1},\ldots,e_n} \tag{8.4.8}$$

from $\mathbf{R}^{n-1}$ into $\mathbf{R}^n$ as in Section 7.3. In this case, we have that

$$\psi_k(y) = e_k + \sum_{j=1}^{k-1} y_j\,(e_j - e_k) + \sum_{j=k}^{n-1} y_j\,(e_{j+1} - e_k) \tag{8.4.9}$$

for every $y \in \mathbf{R}^{n-1}$, as in Section 7.7. This means that

$$\psi_k(y) = \Big(1 - \sum_{j=1}^{n-1} y_j\Big)\, e_k + \sum_{j=1}^{k-1} y_j\, e_j + \sum_{k=k}^{n-1} y_j\, e_{j+1} \tag{8.4.10}$$

for every $y \in \mathbf{R}^{n-1}$.

As in Subsection 6.15.1,

$$\int_{\Phi(e_k, e_1, \ldots, e_{k-1}, e_{k+1}, \ldots, e_n)} \omega = \int_{Q^{n-1}} \psi_k^*(\omega). \tag{8.4.11}$$

Observe that

$$\begin{aligned} \psi_k^*(dx_l) &= dy_l && \text{when } l = 1, \ldots, k-1 \\ &= -\sum_{j=1}^{n-1} dy_j && \text{when } l = k \\ &= dy_{l-1} && \text{when } l = k+1, \ldots, n. \end{aligned} \tag{8.4.12}$$

If ω is as in (8.3.3), then we get that

$$\psi_k^*(\omega) = (\omega_k \circ \psi_k)\, dy_1 \wedge \cdots \wedge dy_{n-1}. \tag{8.4.13}$$

This means that

$$\int_{Q^{n-1}} \psi_k^*(\omega) = \int_{Q^{n-1}} \omega_k(\psi_k(y))\, dy, \tag{8.4.14}$$

as in Section 6.12. This corresponds exactly to the integral of (8.3.20) over Q^{n-1}.

As before, we get that (8.3.32) corresponds to $(-1)^{k-1}$ times the integral of (8.3.20) over Q^{n-1}, because of the factor of $(-1)^{k-1}$ on the right side of (8.4.6). This shows that (8.3.13) is equal to (8.3.17), as desired.

8.5 Stokes' theorem on cells

Let n be a positive integer, let $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ be real numbers with $a_j < b_j$ for each $j = 1, \ldots, n$, and let $\mathcal{C} = \prod_{j=1}^n [a_j, b_j]$ be the corresponding cell in $\mathbf{R}^n$, as in Section 5.13. Also let U be an open set in $\mathbf{R}^n$ with

$$\mathcal{C} \subseteq U, \tag{8.5.1}$$

and let β be a continuously-differentiable differential $(n-1)$-form on U. Thus β may be expressed as

$$\beta = \sum_{j=1}^n \beta_j\, dx_1 \wedge \cdots \wedge dx_{j-1} \wedge dx_{j+1} \wedge \cdots \wedge dx_n \tag{8.5.2}$$

on U, as in Section 3.14, where β_j is a continuously-differentiable real-valued function on U for each j. The exterior derivative of β can be expressed as

$$d\beta = \sum_{j=1}^{n} (-1)^{j-1} (\partial_j \beta_j)\, dx_1 \wedge \cdots \wedge dx_n, \tag{8.5.3}$$

as in Subsection 6.12.1.

This means that

$$\begin{aligned}\int_{\mathcal{C}} d\beta &= \int_{\mathcal{C}} \sum_{j=1}^{n} (-1)^{j-1} (\partial_j \beta_j)(x)\, dx \\ &= \sum_{j=1}^{n} (-1)^{j-1} \int_{\mathcal{C}} (\partial_j \beta_j)(x)\, dx.\end{aligned} \tag{8.5.4}$$

In the jth term on the right side, we can integrate in the jth variable first, and use the fundamental theorem of calculus. Let us express the result in terms of integrals of β over suitable $(n-1)$-surfaces in $\mathbf{R}^n$. We suppose from now on in this section that $n \geq 2$, since otherwise one can simply use the fundamental theorem of calculus directly.

If $1 \leq l \leq n$, then put

$$\mathcal{C}_l = [a_1, b_1] \times \cdots \times [a_{l-1}, b_{l-1}] \times [a_{l+1}, b_{l+1}] \times \cdots \times [a_n, b_n], \tag{8.5.5}$$

which is a cell in $\mathbf{R}^{n-1}$. Consider the mappings ϕ_l and ψ_l from $\mathbf{R}^{n-1}$ into $\mathbf{R}^n$ defined by

$$\phi_l(w) = (w_1, \ldots, w_{l-1}, a_l, w_l, \ldots, w_{n-1}) \tag{8.5.6}$$

and

$$\psi_l(w) = (w_1, \ldots, w_{l-1}, b_l, w_l, \ldots, w_{n-1}) \tag{8.5.7}$$

for each $w = (w_1, \ldots, w_{n-1}) \in \mathbf{R}^{n-1}$. Using ϕ_l and ψ_l, we get $(n-1)$-surfaces

$$\Phi_l = \Phi_l(\mathcal{C}_l) \text{ and } \Psi_l = \Psi_l(\mathcal{C}_l) \tag{8.5.8}$$

in $\mathbf{R}^n$ with parameter domain $\mathcal{C}_l$, as in Section 6.15. More precisely, these are $(n-1)$-surfaces in U, because

$$\phi_l(\mathcal{C}_l),\ \psi_l(\mathcal{C}_l) \subseteq \mathcal{C} \subseteq U. \tag{8.5.9}$$

Observe that

$$\int_{\mathcal{C}} (\partial_l \beta_l)(x)\, dx = \int_{\mathcal{C}_l} \beta_l(\psi_l(w))\, dw - \int_{\mathcal{C}_l} \beta_l(\phi_l(w))\, dw, \tag{8.5.10}$$

by the fundamental theorem of calculus. We also have that $\phi_l^{-1}(U)$, $\psi_l^{-1}(U)$ are open subsets of $\mathbf{R}^{n-1}$, with

$$\mathcal{C}_l \subseteq \phi_l^{-1}(U),\ \psi_l^{-1}(U), \tag{8.5.11}$$

because of (8.5.9).

It is easy to see that

$$\phi_l^*(\beta) = (\beta_l \circ \phi_l)\, dw_1 \wedge \cdots \wedge dw_{n-1} \tag{8.5.12}$$

on $\phi_l^{-1}(U)$. This uses the fact that $\phi_l^*(dx_l) = 0$, so that only the $j = l$ term in sum on the right side of (8.5.2) can have a nonzero pull-back with respect to ϕ_l. Similarly,

$$\psi_l^*(\beta) = (\beta_l \circ \psi_l)\, dw_1 \wedge \cdots \wedge dw_{n-1} \tag{8.5.13}$$

on $\psi_l^{-1}(U)$.

It follows that

$$\int_{\mathcal{C}} (\partial \beta_l)(x)\, dx = \int_{\Psi_l} \beta - \int_{\Phi_l} \beta. \tag{8.5.14}$$

More precisely, the two integrals on the right side are defined as in Subsection 6.15.1, with n replaced with $n-1$.

Combining this with (8.5.4), we get that

$$\int_{\mathcal{C}} d\beta = \sum_{l=1}^{n} (-1)^{l-1} \Big(\int_{\Psi_l} \beta - \int_{\Phi_l} \beta \Big). \tag{8.5.15}$$

This is a version of Stokes' theorem for cells in $\mathbf{R}^n$. This version could be obtained from the previous one, by expressing integrals over cells as integrals over affine chains.

8.6 More on vector fields

Let n and r be positive integers, and let U be a nonempty open subset of $\mathbf{R}^n$. Also let ξ be an $(r-1)$-times continuously-differentiable vector field on U, which is to say an $(r-1)$-times continuously-differentiable function on U with values in $\mathbf{R}^n$, as in Section 4.3. Thus

$$\mathcal{X}_\xi(f) = \sum_{j=1}^{n} \xi_j \, \frac{\partial f}{\partial x_j} \tag{8.6.1}$$

defines a linear mapping from $C^r(U, \mathbf{R})$ into $C^{r-1}(U, \mathbf{R})$, as before. If ξ is infinitely differentiable on U, then this defines a linear mapping from $C^\infty(U, \mathbf{R})$ into itself.

Note that

$$\xi \mapsto \mathcal{X}_\xi \tag{8.6.2}$$

defines a linear mapping from $C^{r-1}(U, \mathbf{R}^n)$ into the space

$$\mathcal{L}(C^r(U, \mathbf{R}), C^{r-1}(U, \mathbf{R})) \tag{8.6.3}$$

of linear mappings from $C^r(U, \mathbf{R})$ into $C^{r-1}(U, \mathbf{R})$. Similarly, (8.6.2) defines a linear mapping from $C^\infty(U, \mathbf{R}^n)$ into the space

$$\mathcal{L}(C^\infty(U, \mathbf{R})) \tag{8.6.4}$$

of linear mappings from $C^\infty(U, \mathbf{R})$ into itself. If $f_l(x) = x_l$ on U for $l = 1, \ldots, n$, then

$$\mathcal{X}_\xi(f_l) = \xi_l. \tag{8.6.5}$$

In particular, this means that (8.6.2) is injective as a mapping from $C^{r-1}(U, \mathbf{R}^n)$ into (8.6.3), and from $C^\infty(U, \mathbf{R}^n)$ into (8.6.4).

8.6.1 Commutators and Lie brackets

Suppose that $r \geq 2$, and let η be another $(r-1)$-times continuously-differentiable vector field on U. Note that $\mathcal{X}_\xi$ and $\mathcal{X}_\eta$ define linear mappings from $C^r(U, \mathbf{R})$ into $C^{r-1}(U, \mathbf{R})$, and from $C^{r-1}(U, \mathbf{R})$ into $C^{r-2}(U, \mathbf{R})$. This implies that the compositions

$$\mathcal{X}_\xi \circ \mathcal{X}_\eta \text{ and } \mathcal{X}_\eta \circ \mathcal{X}_\xi \tag{8.6.6}$$

of $\mathcal{X}_\xi$ and $\mathcal{X}_\eta$, in either order, define linear maps from $C^r(U, \mathbf{R})$ into $C^{r-2}(U, \mathbf{R})$. Thus

$$[\mathcal{X}_\xi, \mathcal{X}_\eta] = \mathcal{X}_\xi \circ \mathcal{X}_\eta - \mathcal{X}_\eta \circ \mathcal{X}_\xi \tag{8.6.7}$$

defines a linear mapping from $C^r(U, \mathbf{R})$ into $C^{r-2}(U, \mathbf{R})$ as well. One can check that there is a unique $(r-2)$-times continuously-differentiable vector field ζ on U such that

$$[\mathcal{X}_\xi, \mathcal{X}_\eta] = \mathcal{X}_\zeta \tag{8.6.8}$$

on $C^r(U, \mathbf{R})$.

We may use the notation

$$\zeta = [\xi, \eta] \tag{8.6.9}$$

in this case. This may be called the *Lie bracket* of ξ and η, as in Definition 1.44 on p36 of [184]. This defines a bilinear mapping from

$$C^{r-1}(U, \mathbf{R}^n) \times C^{r-1}(U, \mathbf{R}^n) \tag{8.6.10}$$

into $C^{r-2}(U, \mathbf{R}^n)$. It is easy to see that this mapping is antisymmetric, in the sense that

$$[\xi, \eta] = -[\eta, \xi]. \tag{8.6.11}$$

Similarly if ξ and η are infinitely differentiable on U, then $\mathcal{X}_\xi \circ \mathcal{X}_\eta$ and $\mathcal{X}_\eta \circ \mathcal{X}_\xi$ define linear mappings from $C^\infty(U, \mathbf{R})$ into itself, so that their commutator (8.6.7) defines a linear mapping from $C^\infty(U, \mathbf{R})$ into itself too. As before, there is a unique smooth vector field ζ on U such that (8.6.8) holds on $C^\infty(U, \mathbf{R})$. This may also be denoted

$$\zeta = [\xi, \eta]_{C^\infty(U, \mathbf{R}^n)} \tag{8.6.12}$$

under these conditions. This defines an antisymmetric bilinear mapping from $C^\infty(U, \mathbf{R}^n) \times C^\infty(U, \mathbf{R}^n)$ into $C^\infty(U, \mathbf{R}^n)$.

8.7 Some more compositions

Let n and r be positive integers again, and let U be a nonempty open subset of $\mathbf{R}^n$. Suppose now that $r \geq 3$, and put

$$V_k = C^{r-k}(U, \mathbf{R}^n) \tag{8.7.1}$$

for $k = 0, 1, 2, 3$. Thus

$$V_3 \subseteq V_2 \subseteq V_1 \subseteq V_0, \tag{8.7.2}$$

as in Section B.3. If α is an $(r-1)$-times continuously-differentiable vector field on U, then $\mathcal{X}_\alpha$ defines linear mappings from $C^{r-k}(U, \mathbf{R})$ into $C^{r-k-1}(U, \mathbf{R})$ for $k = 0, 1, 2$. This means that

$$\mathcal{X}_\alpha \in \mathcal{L}_{3,2}(V_1, V_0), \tag{8.7.3}$$

in the notation of Section B.3.

If β is another $(r-1)$-times continuously-differentiable vector field on U, then

$$\mathcal{X}_\alpha \circ \mathcal{X}_\beta \in \mathcal{L}_3(V_2, V_0), \tag{8.7.4}$$

in the notation of Section B.3. Thus

$$[\mathcal{X}_\alpha, \mathcal{X}_\beta] \in \mathcal{L}_3(V_2, V_0). \tag{8.7.5}$$

If γ is an $(r-1)$-times continuously-differentiable vector field on U too, then

$$\mathcal{X}_\alpha \circ \mathcal{X}_\beta \circ \mathcal{X}_\gamma \in \mathcal{L}(V_3, V_0). \tag{8.7.6}$$

Similarly,

$$([\mathcal{X}_\alpha, \mathcal{X}_\beta]) \circ \mathcal{X}_\gamma,\ \mathcal{X}_\gamma \circ ([\mathcal{X}_\alpha, \mathcal{X}_\beta]) \in \mathcal{L}(V_3, V_0). \tag{8.7.7}$$

Thus

$$[[\mathcal{X}_\alpha, \mathcal{X}_\beta], \mathcal{X}_\gamma] = ([\mathcal{X}_\alpha, \mathcal{X}_\beta]) \circ \mathcal{X}_\gamma - \mathcal{X}_\gamma \circ ([\mathcal{X}_\alpha, \mathcal{X}_\beta]) \in \mathcal{L}(V_3, V_0). \tag{8.7.8}$$

In fact,

$$[[\mathcal{X}_\alpha, \mathcal{X}_\beta], \mathcal{X}_\gamma] + [[\mathcal{X}_\beta, \mathcal{X}_\gamma], \mathcal{X}_\alpha] + [[\mathcal{X}_\gamma, \mathcal{X}_\alpha], \mathcal{X}_\beta] = 0 \tag{8.7.9}$$

in $\mathcal{L}(V_3, V_0)$, as in Section B.3. This is a version of the *Jacobi identity*.

8.7.1 The Jacobi identity for brackets

Note that

$$[\alpha, \beta], [\beta, \gamma], \text{ and } [\gamma, \alpha] \tag{8.7.10}$$

may be defined as $(r-2)$-times continuously-differentiable vector fields on U as in (8.6.9). Similarly,

$$[[\alpha, \beta], \gamma] \tag{8.7.11}$$

and analogous expressions may be defined as vector fields on U that are $(r-3)$-times continuously differentiable. One can verify that

$$[[\alpha,\beta],\gamma]+[[\beta,\gamma],\alpha]+[[\gamma,\alpha],\beta]=0, \tag{8.7.12}$$

which is essentially the same as (8.7.9).

If α, β, and γ are infinitely differentiable on U, then $\mathcal{X}_\alpha$, $\mathcal{X}_\beta$, and $\mathcal{X}_\gamma$ define linear mappings from $C^\infty(U,\mathbf{R})$ into itself, so that their various compositions define linear mappings from $C^\infty(U,\mathbf{R})$ into itself too. In this case, the analogue of (8.7.9) on $C^\infty(U,\mathbf{R})$ is a bit simpler, as mentioned at the beginning of Section B.3. This implies that

$$C^\infty(U,\mathbf{R}^n) \text{ is a Lie algebra} \tag{8.7.13}$$

over the real numbers with respect to (8.6.12), as in Section B.1. This corresponds to a remark after Proposition 1.45 on p36 of [184], and to example 3.5 (a) on p84 of [184].

8.8 Lie derivatives of functions and vector fields

Let n and r be positive integers, and let U be a nonempty open subset of $\mathbf{R}^n$. If ξ and η are r-times continuously-differentiable vector fields on U, then

$$L_\xi(\eta)=[\xi,\eta] \tag{8.8.1}$$

is an $(r-1)$-times continuously-differentiable vector field on U, as in Subsection 8.6.1. This may be called the *Lie derivative of η with respect to ξ*. More precisely, this type of Lie derivative is discussed in Definition 2.24 on p69 of [184] for smooth vector fields. The fact that the Lie derivative may be expressed as in (8.8.1) corresponds to part (b) of Theorem 2.25 on p70 of [184].

Similarly, if ξ is an $(r-1)$-times continuously-differentiable vector field on U and f is an r-times continuously-differentiable real-valued function on U, then

$$L_\xi(f)=\mathcal{X}_\xi(f) \tag{8.8.2}$$

is an $(r-1)$-times continuously-differentiable real-valued function on U, where $\mathcal{X}_\xi$ is as in Section 8.6. This may be called the *Lie derivative of f with respect to ξ*. This type of Lie derivative may be considered as a particular case of those discussed on the top of p70 of [184] for smooth vector fields. The fact that the Lie derivative may be expressed in this way corresponds to part (a) of Theorem 2.25 of [184].

Let r be a nonnegative integer, let a be an r-times continuously-differentiable real-valued function on U, and let ξ be an r-times continuously-differentiable vector field on U. We can define $a\,\xi$ as an r-times continuously-differentiable vector field on U, using scalar multiplication on $\mathbf{R}^n$, so that

$$(a\,\xi)(x)=a(x)\,\xi(x) \tag{8.8.3}$$

for every $x \in U$. If f is an $(r+1)$-times continuously-differentiable real-valued function on U, then

$$\mathcal{X}_{a\,\xi}(f) = a\,\mathcal{X}_\xi(f) \tag{8.8.4}$$

on U.

Suppose that $r \geq 1$ again, let a, ξ, and η be as before, and let b be another r-times continuously-differentiable real-valued function on U. One can check that

$$[a\,\xi, b\,\eta] = a\,b\,[\xi, \eta] + a\,\mathcal{X}_\xi(b)\,\eta - b\,\mathcal{X}_\eta(a)\,\xi. \tag{8.8.5}$$

This corresponds to part (b) of Proposition 1.45 on p36 of [184]. Equivalently, this means that

$$L_{a\,\xi}(b\,\eta) = a\,b\,L_\xi(\eta) + a\,L_\xi(b)\,\eta - b\,L_\eta(a)\,\xi. \tag{8.8.6}$$

In particular, if $a \equiv 1$ on U, then

$$L_\xi(b\,\eta) = L_\xi(b)\,\eta + b\,L_\xi(\eta). \tag{8.8.7}$$

Suppose now that $r \geq 2$, let ξ and η be as before, and let $\widetilde{\eta}$ be another r-times continuously-differentiable vector field on U. Under these conditions, one can check that

$$L_\xi([\eta, \widetilde{\eta}]) = [L_\xi(\eta), \widetilde{\eta}] + [\eta, L_\xi(\widetilde{\eta})], \tag{8.8.8}$$

using the Jacobi identity (8.7.12). If ξ is infinitely differentiable on U, then L_ξ is a derivation on the Lie algebra of smooth vector fields on U, as in Section B.4.

8.9 ϕ-Related vector fields

Let n and m be positive integers, let U be a nonempty open subset of $\mathbf{R}^n$, and let V be a nonempty open subset of $\mathbf{R}^m$. Also let ϕ be a differentiable mapping from U into V, let α be a vector field on U, and let ξ be a vector field on V. This means that α is an $\mathbf{R}^n$-valued function on U, and ξ is an $\mathbf{R}^m$-valued function on V. We say that

$$\alpha \text{ and } \xi \text{ are } \phi\textit{-related} \tag{8.9.1}$$

if

$$d\phi_x(\alpha(x)) = \xi(\phi(x)) \tag{8.9.2}$$

for every $x \in U$. This corresponds to Definition 1.54 on p41 of [184].

Let f be a differentiable real-valued function on V, so that $f \circ \phi$ is differentiable on U, as in Section 3.9. Let us use

$$\mathcal{X}_\alpha \tag{8.9.3}$$

for the first-order differential operator associated to α on U as in Subsection 4.3.1 and Section 8.6, and let us use

$$\mathcal{Y}_\xi \tag{8.9.4}$$

for the analogous differential operator associated to ξ on V. One can check that

$$\mathcal{X}_\alpha(f \circ \phi) = (\mathcal{Y}_\xi(f)) \circ \phi \tag{8.9.5}$$

on U, using (8.9.2) and the chain rule. Conversely, one can get (8.9.2) from this property, using $f_l(y) = y_l$ on V for $l = 1, \ldots, m$.

8.9.1 Brackets of ϕ-related vector fields

Suppose now that ϕ is a twice continuously-differentiable mapping from U into V, α and β are continuously-differentiable vector fields on U, and ξ and η are continuously-differentiable vector fields on V. Note that $[\alpha, \beta]$ and $[\xi, \eta]$ are continuous vector fields on U and V, respectively, as in Subsection 8.6.1. Suppose that (8.9.1) holds, and that

$$\beta \text{ and } \eta \text{ are } \phi\text{-related} \tag{8.9.6}$$

too. We would like to show that

$$[\alpha, \beta] \text{ and } [\xi, \eta] \text{ are } \phi\text{-related} \tag{8.9.7}$$

under these conditions. This corresponds to Proposition 1.55 on p41 of [184].

Let f be a twice continuously-differentiable real-valued function on V. This implies that $f \circ \phi$ is twice continuously differentiable on U, as in Subsection 3.9.1, so that $\mathcal{X}_\alpha(f \circ \phi)$ and $\mathcal{X}_\beta(f \circ \phi)$ are continuously differentiable on U. Similarly, $\mathcal{Y}_\xi(f)$ and $\mathcal{Y}_\eta(f)$ are continuously differentiable on V, and their compositions with ϕ are continuously differentiable on U. Observe that

$$\mathcal{X}_{[\alpha,\beta]}(f \circ \phi) = ([\mathcal{X}_\alpha, \mathcal{X}_\beta])(f \circ \phi) = \mathcal{X}_\alpha(\mathcal{X}_\beta(f \circ \phi)) - \mathcal{X}_\beta(\mathcal{X}_\alpha(f \circ \phi)) \tag{8.9.8}$$

on U, as in Subsection 8.6.1. Using (8.9.1) and (8.9.6), we get that the right side is equal to

$$\mathcal{X}_\alpha(\mathcal{Y}_\eta(f) \circ \phi) - \mathcal{X}_\beta(\mathcal{Y}_\xi(f) \circ \phi), \tag{8.9.9}$$

as in (8.9.5).

This is equal to

$$(\mathcal{Y}_\xi(\mathcal{Y}_\eta(f))) \circ \phi - (\mathcal{Y}_\eta(\mathcal{Y}_\xi(f))) \circ \phi, \tag{8.9.10}$$

as in (8.9.5) again. This is the same as

$$(\mathcal{Y}_{[\xi,\eta]}(f)) \circ \phi, \tag{8.9.11}$$

as in Subsection 8.6.1. Thus we get that

$$\mathcal{X}_{[\alpha,\beta]}(f \circ \phi) = (\mathcal{Y}_{[\xi,\eta]}(f)) \circ \phi \tag{8.9.12}$$

on U. This implies (8.9.7), as before.

8.10 Polynomial vector fields

Let (A, b) be an algebra in the strict sense over the real numbers, as in Section A.2. Also let δ be a derivation on A with respect to b, as in Subsection A.2.1. One can check that the kernel of δ is a subalgebra of A. If A has a multiplicative identity element e with respect to b, then one can verify that $\delta(e) = 0$.

Let m and n be positive integers. Consider the space

$$\mathcal{P}(\mathbf{R}^n, \mathbf{R}^m) \tag{8.10.1}$$

of functions on $\mathbf{R}^n$ with values in $\mathbf{R}^m$ each of whose components is a polynomial on $\mathbf{R}^n$ with real coefficients. This is a linear subspace of the space $C^\infty(\mathbf{R}^n, \mathbf{R}^m)$ of infinitely-differentiable functions on $\mathbf{R}^n$ with values in $\mathbf{R}^m$.

Let us now take $m = n$. One may describe

$$\mathcal{P}(\mathbf{R}^n, \mathbf{R}^n) \tag{8.10.2}$$

as the space of *polynomial vector fields* on $\mathbf{R}^n$. This is a linear subspace of the space $C^\infty(\mathbf{R}^n, \mathbf{R}^n)$ of all smooth vector fields on $\mathbf{R}^n$. Let ξ be a polynomial vector field on $\mathbf{R}^n$, and let $\mathcal{X}_\xi$ be the corresponding first-order differential operator on $\mathbf{R}^n$, as in Section 8.6. It is easy to see that $\mathcal{X}_\xi$ defines a linear mapping from the space

$$\mathcal{P}(\mathbf{R}^n) = \mathcal{P}(\mathbf{R}^n, \mathbf{R}) \tag{8.10.3}$$

of polynomials on $\mathbf{R}^n$ with real coefficients into itself.

As before, $\xi \mapsto \mathcal{X}_\xi$ defines a linear mapping from (8.10.2) into the space

$$\mathcal{L}(\mathcal{P}(\mathbf{R}^n)) \tag{8.10.4}$$

of linear mappings from $\mathcal{P}(\mathbf{R}^n)$ into itself. This mapping is injective, for the same reasons as in Section 8.6. If ξ and η are polynomial vector fields on $\mathbf{R}^n$, then it is easy to see that

$$[\xi, \eta] \in \mathcal{P}(\mathbf{R}^n, \mathbf{R}^n). \tag{8.10.5}$$

This means that (8.10.2) is a Lie subalgebra of $C^\infty(\mathbf{R}^n, \mathbf{R}^n)$, with respect to the Lie bracket, as in Section B.2.

Of course, $\mathcal{P}(\mathbf{R}^n)$ is a commutative associative algebra over the real numbers with respect to ordinary multiplication. If ξ is a polynomial vector field on $\mathbf{R}^n$, then $\mathcal{X}_\xi$ is a derivation on $\mathcal{P}(\mathbf{R}^n)$. It is not too difficult to show that every derivation on $\mathcal{P}(\mathbf{R}^n)$ is of this type.

8.11 Cell-nice sets in $\mathbf{R}^n$

Let n be a positive integer, and let $x, y \in \mathbf{R}^n$ be given. Consider the cell

$$\mathcal{C}(x, y) = \prod_{j=1}^n [\min(x_j, y_j), \max(x_j, y_j)] \tag{8.11.1}$$

in $\mathbf{R}^n$, as in Section 5.13. Of course,

$$x, y \in \mathcal{C}(x, y), \tag{8.11.2}$$

by construction. If $\mathcal{C}$ is any cell in $\mathbf{R}^n$ such that

$$x, y \in \mathcal{C}, \tag{8.11.3}$$

then it is easy to see that

$$\mathcal{C}(x, y) \subseteq \mathcal{C}. \tag{8.11.4}$$

Let us say that a subset E of $\mathbf{R}^n$ is *cell-nice* about a point $w \in E$ if for every $x \in E$,

$$\mathcal{C}(w, x) \subseteq E. \tag{8.11.5}$$

Cells in $\mathbf{R}^n$ are cell-nice about each of their elements, for instance. Remember that the open ball $B(w, r)$ in $\mathbf{R}^n$ centered at $w \in \mathbf{R}^n$ with radius $r > 0$ with respect to the standard Euclidean metric may be defined as in Subsection 3.4.1. One can check that

$$B(w, r) \text{ is cell-nice about } w. \tag{8.11.6}$$

It is easy to see that unions and intersections of families of subsets of $\mathbf{R}^n$ that are cell-nice about w are cell-nice about w too.

Let p be a positive integer with $p \le n$, and let π_p be the obvious projection from $\mathbf{R}^n$ onto $\mathbf{R}^p$, so that

$$\pi_p(x) = (x_1, \ldots, x_p) \tag{8.11.7}$$

for each $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$. If $\mathcal{C}$ is a cell in $\mathbf{R}^n$, then

$$\pi_p(\mathcal{C}) \text{ is a cell in } \mathbf{R}^p. \tag{8.11.8}$$

If $E \subseteq \mathbf{R}^n$ is cell-nice about $w \in E$, then one can verify that

$$\pi_p(E) \text{ is cell-nice about } \pi_p(w), \tag{8.11.9}$$

as a subset of $\mathbf{R}^p$. If $A \subseteq \mathbf{R}^p$, $\pi_p(w) \in W$, and A is cell-nice about $\pi_p(w)$ in $\mathbf{R}^p$, then it is easy that $\pi_p^{-1}(A)$ is cell-nice about w in $\mathbf{R}^n$.

If $w \in \mathbf{R}^n$ and $r > 0$, then it is easy to see that

$$\pi_p(B(w, r)) \tag{8.11.10}$$

is the same as the open ball in $\mathbf{R}^p$ centered at $\pi_p(w)$ with radius r, with respect to the standard Euclidean metric on $\mathbf{R}^p$. If U is an open set in $\mathbf{R}^n$, then it follows that

$$\pi_p(U) \text{ is an open set in } \mathbf{R}^p. \tag{8.11.11}$$

This means that π_p is an open mapping from $\mathbf{R}^n$ onto $\mathbf{R}^p$, as in Section 5.10.

8.12 Integrating some functions

Let n and p be positive integers with $p \leq n$ again, and let W be an open set in $\mathbf{R}^n$ that is cell-nice about a point $w \in W$. Thus

$$V = \pi_p(W) \tag{8.12.1}$$

is an open set in $\mathbf{R}^p$ that is cell-nice about $\pi_p(w)$, where π_p is the obvious projection from $\mathbf{R}^n$ onto $\mathbf{R}^p$, as before. Also let f be a continuously-differentiable real-valued function on W, and suppose that

$$\partial_j f(x) = 0 \tag{8.12.2}$$

for each $x \in W$ and $j = p+1, \ldots, n$. Of course, this condition holds vacuously when $p = n$. One can check that this implies that $f(x)$ does not depend on x_j for $j = p+1, \ldots, n$, because W is cell-nice about w. More precisely, if $x \in W$ and $p < n$, then

$$(x_1, \ldots, x_p, w_{p+1}, \ldots, w_n) \in W, \tag{8.12.3}$$

because W is cell-nice about w. One can verify that

$$f(x) = f(x_1, \ldots, x_p, w_{p+1}, \ldots, w_n) \tag{8.12.4}$$

for every $x \in W$ under these conditions.

This means that f can be expressed as

$$f(x) = \phi(\pi_p(x)), \tag{8.12.5}$$

where ϕ is a real-valued function on V. In fact, ϕ is continuously differentiable on V, because it is the same as the right side of (8.12.4). If $y = (y_1, \ldots, y_p) \in V$, then

$$(y_1, \ldots, y_{p-1}, t) \in V \tag{8.12.6}$$

when

$$w_p \leq t \leq y_p \text{ or } y_p \leq t \leq w_p, \tag{8.12.7}$$

as appropriate, because V is cell-nice about $\pi_p(w)$ in $\mathbf{R}^p$. Consider the real-valued function Φ defined on V by

$$\begin{aligned} \Phi(y_1, \ldots, y_p) &= \int_{w_p}^{y_p} \phi(y_1, \ldots, y_{p-1}, t)\, dt \qquad \text{when } y_p \geq w_p \\ &= -\int_{y_p}^{w_p} \phi(y_1, \ldots, y_{p-1}, t)\, dt \quad \text{when } y_p \leq w_p. \end{aligned} \tag{8.12.8}$$

Note that

$$\Phi(y_1, \ldots, y_p) = 0 \text{ when } y_p = w_p. \tag{8.12.9}$$

One can check that

$$\Phi \text{ is continuous on } V. \tag{8.12.10}$$

This is straightforward when $p = 1$, and otherwise one can use the uniform continuity of ϕ on compact subsets of V.

Of course,

$$\partial_p \Phi = \phi \tag{8.12.11}$$

on V, by construction. If $p \geq 2$ and $j = 1, \ldots, p-1$, then $\partial_j \Phi$ can be obtained by differentiating under the integral sign. More precisely, the difference quotients of Φ used to define $\partial_j \Phi$ are equal to analogous integrals of difference quotients of ϕ. In order to differentiate under the integral sign, it is enough to have uniform convergence of the difference quoitents of ϕ along the appropriate closed segment. This can be obtained from the uniform continuity of $\partial_j \phi$ on compact subsets of V, and the mean-value theorem of the fundamental theorem of calculus.

If $p \geq 2$, then $\partial_j \Phi$ is continuous on V for each $j = 1, \ldots, p-1$. This follows from the continuity of $\partial_j \phi$ on V, in the same way as in (8.12.10). This means that

$$\Phi \text{ is continuous differentiable on } V, \tag{8.12.12}$$

because (8.12.11) is continuous on V, by hypothesis.

8.12.1 Some properties of these functions

Consider the real-valued function F defined on W by

$$F = \Phi \circ \pi_p. \tag{8.12.13}$$

Note that

$$F \text{ is continuously differentiable on } W, \tag{8.12.14}$$

because Φ is continuously differentiable on V. It is easy to see that

$$\partial_p F = f \tag{8.12.15}$$

on W, because of (8.12.11). If $p < n$, then $F(x)$ does not depend on x_j for $j = p+1, \ldots, n$, so that

$$\partial_j F(x) = 0 \tag{8.12.16}$$

for every $x \in W$.

This basically corresponds to Theorem 10.38 on p278 of [155], which is stated for convex open sets in $\mathbf{R}^n$. The argument here is essentially the same as in [155], with some simplifications in this case. This is also related to Exercise 29 on p297 of [155], as mentioned in [155].

Let r be a positive integer, and suppose now that f is r-times continuously differentiable on W, which implies that ϕ is r-times continuously differentiable on V. Under these conditions, one can check that

$$\Phi \text{ is } r\text{-times continuously differentiable on } V, \tag{8.12.17}$$

using the same type of arguments as before. This implies that

$$F \text{ is } r\text{-times continuously differentiable on } W. \tag{8.12.18}$$

If f is infinitely differentiable on W, then it follows that

(8.12.19) $$F \text{ is infinitely differentiable on } W.$$

8.13 A version of Poincaré's lemma

Let n be a positive integer, and let W be an open set in $\mathbf{R}^n$ that is cell-nice about a point $w \in W$ again. Also let

(8.13.1) $$\beta \text{ be an infinitely-differentiable differential } k\text{-form on } W$$

for some $k = 1, \ldots, n$, and suppose that

(8.13.2) $$\beta \text{ is a closed form on } W,$$

so that $d\beta = 0$ on W. We would like to find an infinitely-differentiable differential $(k-1)$-form α on W such that

(8.13.3) $$d\alpha = \beta$$

on W, so that β is exact on W.

This is a version of the famous *Poincaré lemma*. This basically corresponds to Theorem 10.39 in [155], which is stated for convex open subsets of $\mathbf{R}^n$.

More precisely, suppose that β is r-times continuously differentiable on W with

(8.13.4) $$r \geq n - k + 1,$$

instead of infinite differentiability. Under these conditions, the proof will show that we can find a differential $(k-1)$-form α on W such that

(8.13.5) $$\alpha \text{ is } (r - n + k)\text{-times continuously differentiable on } W,$$

and (8.13.3) holds on W.

8.13.1 Invariance under diffeomorphisms

Let U be a nonempty open subset of $\mathbf{R}^n$, and suppose that

(8.13.6) every infinitely-differentiable differential k-form on U that is closed on U can be expressed as the exterior derivative of an infinitely differentiable differential $(k-1)$-form on U.

If

(8.13.7) ϕ is a C^∞ diffeomorphism from U onto another open set $V \subseteq \mathbf{R}^n$,

as in Subsection 5.12.2, then one can check that V has the same property. This corresponds to Theorem 10.40 on p280 of [155].

8.14 Some initial steps

Let us continue with the same notation and hypotheses as at the beginning of the previous section. Let p and r be positive integers, with

$$k \le p \le n. \tag{8.14.1}$$

Consider the space

$$Y(k, p, r) \tag{8.14.2}$$

of r-times continuously-differentiable differential k-forms γ on W of the form

$$\gamma = \sum_{I \subseteq \{1,\ldots,p\}, \#I = k} \gamma_I \, dx^I. \tag{8.14.3}$$

More precisely, the sum on the right is taken over all subsets I of $\{1, \ldots, p\}$ with exactly k elements.

Suppose for the moment that

$$\gamma \in Y(k, k, r). \tag{8.14.4}$$

This means that

$$\gamma = f \, dx_1 \wedge \cdots \wedge dx_k \tag{8.14.5}$$

for some r-times continuously-differentiable real-valued function f on W. Observe that

$$d\gamma = 0 \tag{8.14.6}$$

on W if and only if

$$\partial_j f = 0 \tag{8.14.7}$$

on W for each $j = k + 1, \ldots, n$, which is vacuous when $k = n$.

In this case, there is an r-times continuously-differentiable real-valued function F on W such that

$$\partial_k F = f \tag{8.14.8}$$

on W, and

$$\partial_j F = 0 \tag{8.14.9}$$

on W for each $j \ge k + 1$, as in Subsection 8.12.1. Note that F does not depend on r. In particular, if f is infinitely differentiable on W, then F is infinitely differentiable on W too, as before.

Thus

$$(-1)^{k-1} F \, dx_1 \wedge \cdots \wedge dx_{k-1} \in Y(k - 1, k - 1, r). \tag{8.14.10}$$

One can check that

$$\begin{aligned} & d((-1)^{k-1} F \, dx_1 \wedge \cdots \wedge dx_{k-1}) \\ = \quad & (\partial_k F) \, dx_1 \wedge \cdots \wedge dx_{k-1} \wedge dx_k = \gamma \end{aligned} \tag{8.14.11}$$

uunder these conditions. Note that this implies the Poincaré lemma when $k = n$.

8.14.1 The induction hypothesis

Suppose now that

$$k + 1 \le p \le n, \tag{8.14.12}$$

so that $p \ge 2$ in particular. Our induction hypothesis is that

(8.14.13) if $r \ge p - k$ and an element of $Y(k, p-1, r)$ is closed as a differential form on W, then it can be expressed as the exterior derivative of an element of $Y(k-1, p-2, r-p+k+1)$.

More precisely, this element of $Y(k-1, p-2, r)$ should not depend on r, so that it is infinitely differentiable when the initial element of $Y(k, p-1, r)$ is infinitely differentiable. Note that

(8.14.14) (8.14.13) holds when $p - 1 = k$,

as in (8.14.11).

We would like to use this to show that

(8.14.15) if $r \ge p - k + 1$ and an element of $Y(k, p, r)$ is closed as a differential form on W, then it can be expressed as the exterior derivative of an element of $Y(k-1, p-1, r-p+k)$.

As before, this element of $Y(k-1, p-1, r-p+k)$ should not depend on r, so that it is smooth when the initial element of $Y(k, p, r)$ is smooth.

We shall discuss (8.14.15) further in the next section. Of course, (8.14.13) is the same as (8.14.15) with p replaced with $p - 1$.

8.15 The induction step

Let us continue with the same notation and hypotheses as in the previous section. Suppose that the induction hypothesis (8.14.13) holds, and let us show that (8.14.15) holds as well, as in [155]. Let γ be an element of $Y(k, p, r)$ as in (8.14.3), with

$$r \ge p - k + 1, \tag{8.15.1}$$

and suppose that γ is closed on W.

Observe that

$$\begin{aligned} d\gamma &= \sum_{I \subseteq \{1, \ldots, p\}, \#I = k} d\gamma_I \wedge dx^I \\ &= \sum_{I \subseteq \{1, \ldots, p\}, \#I = k} \sum_{j=1}^{n} (\partial_j \gamma_I)\, dx_j \wedge dx^I \end{aligned} \tag{8.15.2}$$

on W. This is equal to 0 on W, by hypothesis. It follows that

$$\sum_{I \subseteq \{1, \ldots, p\}, \#I = k} (\partial_j \gamma_I)\, dx_j \wedge dx^I = 0 \tag{8.15.3}$$

on W for $j = p+1, \ldots, n$. In fact, one can use this to get that

$$\partial_j \gamma_I = 0 \tag{8.15.4}$$

on W for every $I \subseteq \{1, \ldots, p\}$ with exactly k elements and $j = p+1, \ldots, n$.

Note that $r \geq 2$, because of (8.14.12) and (8.15.1). If $I \subseteq \{1, \ldots, p\}$ has exactly k elements, then there is an r-times continuously-differentiable real-valued function Γ_I on W such that

$$\partial_p \Gamma_I = \gamma_I \tag{8.15.5}$$

and

$$\partial_j \Gamma_I = 0 \tag{8.15.6}$$

on W for $j = p+1, \ldots, n$, as in Subsection 8.12.1. Remember that Γ_I does not depend on r, so that it is smooth on W when γ is smooth on W.

8.15.1 Using the induction hypothesis

Put

$$\xi = \sum_{I \subseteq \{1, \ldots, p-1\}, \#I = k} \gamma_I \, dx^I \tag{8.15.7}$$

and

$$\eta = \sum_{I \subseteq \{1, \ldots, p\}, p \in I, \#I = k} \gamma_I \, dx^I, \tag{8.15.8}$$

so that $\xi \in Y_{p-1}$ and

$$\gamma = \xi + \eta. \tag{8.15.9}$$

If $I \subseteq \{1, \ldots, p\}$ has exactly k elements and $p \in I$, then put

$$I(p) = I \setminus \{p\}, \tag{8.15.10}$$

so that $I(p) \subseteq \{1, \ldots, p-1\}$ has exactly $k-1$ elements.

Put

$$\zeta = \sum_{I \subseteq \{1, \ldots, p\}, p \in I, \#I = k} \Gamma_I \, dx^{I(p)}, \tag{8.15.11}$$

which is a differential $(k-1)$-form on W. We also have that

$$\begin{aligned} d\zeta &= \sum_{I \subseteq \{1, \ldots, p\}, p \in I, \#I = k} d\Gamma_I \wedge dx^{I(p)} \\ &= \sum_{I \subseteq \{1, \ldots, p\}, p \in I, \#I = k} \sum_{j=1}^{p-1} (\partial_j \Gamma_I) \, dx_j \wedge dx^{I(p)} \\ &\quad + \sum_{I \subseteq \{1, \ldots, p\}, p \in I, \#I = k} \gamma_I \, dx_p \wedge dx^{I(p)} \\ &= \sum_{I \subseteq \{1, \ldots, p\}, p \in I, \#I = k} \sum_{j=1}^{p-1} (\partial_j \Gamma_I) \, dx_j \wedge dx^{I(p)} + (-1)^{k-1} \, \eta. \end{aligned} \tag{8.15.12}$$

Thus

$$\gamma - (-1)^{k-1}\, d\zeta$$

$$(8.15.13) \qquad = \quad \xi - (-1)^{k-1} \sum_{I \subseteq \{1,\ldots,p\}, p \in I, \# I = k} \sum_{j=1}^{p-1} (\partial_j \Gamma_I)\, dx_j \wedge dx^{I(p)}.$$

This is $(r-1)$-times continuously differentiable on W, because γ and the Γ_I's are r-times continuously differentiable on W. It is easy to see that

$$(8.15.14) \qquad \gamma - (-1)^{k-1}\, d\zeta \in Y(k, p-1, r-1),$$

by construction. We also have that

$$(8.15.15) \qquad d(\gamma - (-1)^{k-1}\, d\zeta) = 0$$

on W, because ζ is closed on W, by hypothesis.

Of course,

$$(8.15.16) \qquad r - 1 \ge p - k,$$

because of (8.15.1). This permits us to apply our induction hypothesis (8.14.13), with r replaced with $r-1$. It follows that

(8.15.17) $\quad \gamma - (-1)^{k-1}$ can be expressed as the exterior derivative of an element of $Y(k-1, p-2, r-p+k)$.

Note that

$$(8.15.18) \qquad \zeta \in Y(k-1, p-1, r),$$

by construction. This implies that

(8.15.19) $\quad \gamma$ can be expressed as the exterior derivative of an element of $Y(k-1, p-1, r-p+k)$,

as in (8.14.15).

Chapter 9

Lie derivatives and tensor fields

9.1 Tensor fields and multilinearity

Let k and n be positive integers, and let E be a nonempty subset of $\mathbf{R}^n$. Also let a be a tensor field of type $(0, k)$ on E, as in Section 4.1. If $\eta_1, \ldots, \eta_k$ are vector fields on E, then

$$a(\eta_1, \ldots, \eta_k) \tag{9.1.1}$$

defines a real-valued function on E, as in Section 4.3. Thus

$$(\eta_1, \ldots, \eta_k) \mapsto a(\eta_1, \ldots, \eta_k) \tag{9.1.2}$$

may be considered as a mapping from the space of k-tuples of vector fields on E to the space of real-valued functions on E.

The spaces of real-valued functions on E and vector fields on E are vector spaces over the real numbers with respect to pointwise addition and scalar multiplication, as before. In fact, (9.1.2) may be considered as a multilinear mapping from the space of k-tuples of vector fields on E into the space of real-valued functions on E, as in Section 2.2.

Let f_l be a real-valued function on E for some positive integer $l \leq k$. If η_l is a vector field on E, then

$$f_l\, \eta_l \tag{9.1.3}$$

defines a vector field on E as well, whose value at $x \in E$ is $f_l(x)\, \eta_l(x)$. If $\eta_1, \ldots, \eta_k$ are vector fields on E, then

$$a(\eta_1, \ldots, \eta_{l-1}, f_l\, \eta_l, \eta_{l+1}, \ldots, \eta_k) = f_l\, a(\eta_1, \ldots, \eta_k) \tag{9.1.4}$$

on E. This may be considered as a stronger version of the multilinearity of (9.1.2).

The space of real-valued functions on E is a commutative associative algebra over the real numbers, as in Section A.2. The space of vector fields on E may be considered as a module over the algebra of real-valued functions with respect to multiplication as in (9.1.3), as in Subsections B.7.1 and B.7.2. One does not need to distinguish between left and right modules here, because the algebra is commutative. Note that the algebra of real-valued functions on E may be considered as a module over itself, as in Subsections B.7.1 and B.7.2.

This means that we may consider (9.1.2) as a mapping from the space of k-tuples of vector fields on E into the space of real-valued functions on E, where the spaces of vector fields on E and real-valued functions on E are both considered as modules over the algebra of real-valued functions on E. We may say that this mapping is *multilinear* over the algebra of real-valued functions on E, because of (9.1.4). This is related to some remarks in Section 2.18 beginning on p64 of [184].

9.1.1 Multilinearity over $C(E, \mathbf{R})$

Suppose that a is continuous as a tensor field of type $(0, k)$ on E, as in Subsection 4.1.1. If $\eta_1, \ldots, \eta_k$ are continuous vector fields on E, then (9.1.1) is a continuous real-valued function on E. The space $C(E, \mathbf{R})$ of continuous real-valued functions on E is a subalgebra of the space of all real-valued function on E, and in particular it is a commutative associative algebra over $\mathbf{R}$ as well. The space $C(E, \mathbf{R}^n)$ of continuous vector fields on E may be considered as a module over $C(E, \mathbf{R})$, with multiplication as in (9.1.3). We may consider $C(E, \mathbf{R})$ as a module over itself, as before.

Thus we may consider (9.1.2) as a mapping from the space $C(E, \mathbf{R}^n)^k$ of k-tuples of continuous vector fields on E into $C(E, \mathbf{R})$, where $C(E, \mathbf{R})$ and $C(E, \mathbf{R}^n)$ are both considered as modules over $C(E, \mathbf{R})$. This mpping is *multilinear over* $C(E, \mathbf{R})$, because of (9.1.4).

9.1.2 Multilinearity over $C^r(U, \mathbf{R})$

Let U be a nonempty open subset of $\mathbf{R}^n$, let r be a nonnegative integer, and suppose that a is an r-times continuously differentiable tensor field of type $(0, k)$ on U, as in Section 4.2. If $\eta_1, \ldots, \eta_k$ are r-times continuously-differentiable vector fields on U, then (9.1.1) is r-times continuously differentiable as a real-valued function on U, as in Subsection 4.3.1. The space $C^r(U, \mathbf{R})$ of r-times continuously-differentiable real-valued functions on U is a subalgebra of the space $C(U, \mathbf{R})$ of continuous real-valued functions on U, and in particular is a commutative associaitve algebra over $\mathbf{R}$. The space $C^r(U, \mathbf{R}^n)$ of r-times continuously-differentiable vector fields on U may be considered as a module over $C^r(U, \mathbf{R})$, with multiplication as in (9.1.3). We may consider $C^r(U, \mathbf{R})$ as a module over itself, as usual.

We may consider (9.1.2) as a mapping from $C^r(U, \mathbf{R}^n)^k$ into $C^r(U, \mathbf{R})$, where $C^r(U, \mathbf{R})$ and $C^r(U, \mathbf{R}^n)$ are considered as modules over $C^r(U, \mathbf{R})$. This mapping is *multilinear over* $C^r(U, \mathbf{R})$, because of (9.1.4).

Similarly, suppose that a is infinitely differentiable as a tensor field of type $(0, k)$ on U, as in Section 4.2. If $\eta_1, \ldots, \eta_k$ are infinitely differentiable vector fields on U, then (9.1.1) is infinitely-differentiable as a real-valued function on U, as in Subsection 4.3.1. The space $C^\infty(U, \mathbf{R})$ of infinitely-differerntiable real-valued functions on U is a subalgebra of $C^r(U, \mathbf{R})$ for each $r \geq 0$, and in particular $C^\infty(U, \mathbf{R})$ is a commutative associaitve algebra over $\mathbf{R}$. The space $C^\infty(U, \mathbf{R}^n)$ of infinitely-differentiable vector fields on U may be considered as a module over $C^\infty(U, \mathbf{R})$, with multiplication as in (9.1.3), and we may consider $C^\infty(U, \mathbf{R})$ as a module over itself. If we consider (9.1.2) as a mapping from $C^\infty(U, \mathbf{R}^n)^k$ into $C^\infty(U, \mathbf{R})$, then this mapping is *multilinear over* $C^\infty(U, \mathbf{R})$, because of (9.1.4).

9.2 Spaces of multilinear mappings

Let k be a positive integer, and let $V_1, \ldots, V_k$, Z be vector spaces over the real numbers. Consider the space

$$\mathcal{L}(V_1, \ldots, V_k; Z) \tag{9.2.1}$$

of multilinear mappings from $\prod_{l=1}^k V_l$ into Z, as in Section 2.2. This is a linear subspace of the space of all functions on $\prod_{l=1}^k V_k$ with values in Z. In particular, this is a vector space over the real numbers with respect to pointwise addition and scalar multiplication of functions.

Let V be a vector space over the real numbers, and suppose now that $V_l = V$ for each $l = 1, \ldots, k$. In this case, (9.2.1) is the same as the space of multilinear mappings from the space V^k of k-tuples of elements of V into Z. This space may be denoted

$$\mathcal{M}_k(V, Z), \tag{9.2.2}$$

and is a linear subspace of the space of all Z-valued functions on V^k, as before. This is the same as $\mathcal{M}_k(V)$ when $Z = \mathbf{R}$, considered as a one-dimensional vector space over itself.

Let μ be a multilinear mapping from V^k into Z, and let σ be a permutation on $\{1, \ldots, k\}$. Put

$$\mu^\sigma(v_1, \ldots, v_k) = \mu(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) \tag{9.2.3}$$

for all $v_1, \ldots, v_k \in V$, as in Section 2.4. This is another multilinear mapping from V^k into Z, as before. If τ is another permutation on $\{1, \ldots, k\}$, then

$$\mu^{\tau \circ \sigma} = (\mu^\sigma)^\tau, \tag{9.2.4}$$

as in Sections 1.9 and 2.4. Note that $\mu \mapsto \mu^\sigma$ is a one-to-one linear mapping from (9.2.2) onto itself, as before.

We say that μ is a *symmetric* multilinear mapping from V^k into Z if

$$\mu^\sigma = \mu \tag{9.2.5}$$

for every $\sigma \in \mathrm{Sym}(k)$, as in Section 1.9 and 2.4. Similarly, we say that μ is an *alternating* if

$$\mu^\sigma = \mathrm{sgn}(\sigma)\, \mu \tag{9.2.6}$$

for every $\sigma \in \mathrm{Sym}(k)$, as in Sections 1.11 and 2.4. As usual, it suffices to check that these conditions hold when σ is a transposition on $\{1, \ldots, k\}$. Let

$$\mathcal{SM}_k(V, Z) \text{ and } \mathcal{AM}_k(V, Z) \tag{9.2.7}$$

be the spaces of symmetric and alternating multilinear mappings from V^k into Z, respectively. It is easy to see that these are linear subspaces of $\mathcal{M}_k(V, Z)$.

9.2.1 Multilinear mappings and $\mathbf{R}^n$

Let n be a positive integer, and let $e_1, \ldots, e_n$ be the standard basis for $\mathbf{R}^n$, as usual. If μ is a multilinear mapping from $(\mathbf{R}^n)^k$ into Z, then μ is uniquely determined by its values

$$\mu(e_{m_1}, \ldots, e_{m_k}) \tag{9.2.8}$$

on k-tuples of standard basis vectors in $\mathbf{R}^n$, as in Subsection 1.6.1. These may be arbitrary elements of Z, for the same reasons as before. It is easy to see that μ is symmetric or alternating on $(\mathbf{R}^n)$ if and only if its values (9.2.8) on k-tuples of standard basis vectors in $\mathbf{R}^n$ has the analogous property, as in Sections 1.9 and 1.11.

If μ is a multilinear mapping from V^k into $\mathbf{R}^n$, then the jth component μ_j of μ is a k-linear form on V for each $j = 1, \ldots, n$. Of course, μ is symmetric or alternating on V^k if and only if μ_j is symmetric or alternating on V^k, as appropriate, for each $j = 1, \ldots, n$.

If μ is a multilinear mapping from $(\mathbf{R}^n)^k$ into $\mathbf{R}^n$, then

$$\mu_j(e_{m_1}, \ldots, e_{m_k}) \tag{9.2.9}$$

is a real number for each $j = 1, \ldots, n$ and $1 \le m_1, \ldots, m_k \le n$. We also have that μ is uniquely determined by this family of real numbers, and that these real numbers may be arbitrary. In particular, this implies that

$$\dim \mathcal{M}_k(\mathbf{R}^n, \mathbf{R}^n) = n^{k+1}. \tag{9.2.10}$$

It is sometimes convenient to identify $\mathcal{M}_k(\mathbf{R}^n, \mathbf{R}^n)$ with $\mathbf{R}^{n^{k+1}}$, by listing the elements of

$$\{1, \ldots, n\}^{k+1} \tag{9.2.11}$$

in a sequence with n^{k+1} terms. The order in which the elements of (9.2.11) will normally not matter, as before. In particular, this leads to a metric on $\mathcal{M}_k(\mathbf{R}^n, \mathbf{R}^n)$, corresponding to the standard Euclidean metric on $\mathbf{R}^{n^{k+1}}$, that does not depend on the order in which the elements of (9.2.11) are listed.

9.3 Tensor fields of type $(1, k)$

Let k and n be positive integers, and let E be a nonempty subset of $\mathbf{R}^n$. A *tensor field of type* $(1, k)$ on E is a function on E with values in the space

$\mathcal{M}_k(\mathbf{R}^n, \mathbf{R}^n)$ of multilinear mappings from $(\mathbf{R}^n)^k$ into $\mathbf{R}^n$. This corresponds to part of Definition 2.15 on p63 of [184]. The space of tensor fields of type $(1, k)$ on E is a vector space over the real numbers with respect to pointwise addition and scalar multiplication.

Let us say that a tensor field of type $(1, k)$ on E is *continuous* if it corresponds to a continuous mapping from E into $\mathbf{R}^{n^{k+1}}$ as in Subsection 9.2.1. The space

$$C(E, \mathcal{M}_k(\mathbf{R}^n, \mathbf{R}^n)) = C\big(E, \mathbf{R}^{n^{k+1}}\big) \tag{9.3.1}$$

of continuous tensor fields of type $(1, k)$ on E is a linear subspace of the space of all tensor fields of type $(1, k)$ on E.

Remember that one can get a multilinear mapping from $(\mathbf{R}^n)^k$ into $\mathbf{R}^n$ by multiplying an element of $\mathbf{R}^n$ by a k-linear form on $\mathbf{R}^n$, as in Subsection 2.2.1. Similarly, if a is a tensor field of type $(0, k)$ on E, as in Section 4.1, and ξ is a vector field on E, as in Section 4.3, then

$$a\,\xi \tag{9.3.2}$$

defines a tensor field of type $(1, k)$ on E. If a and ξ are continuous on E, then it is easy to see that (9.3.2) is continuous on E, as a tensor field of type $(1, k)$.

Let μ be a tensor field of type $(1, k)$ on E, so that μ_x is a multilinear mapping from $(\mathbf{R}^n)^k$ into $\mathbf{R}^n$ for each $x \in E$. If $\eta_1, \ldots, \eta_k$ are vector fields on E, then

$$\mu(\eta_1, \ldots, \eta_k) \tag{9.3.3}$$

defines another vector field on E. The value of this vector field at $x \in E$ is

$$\mu_x(\eta_1(x), \ldots, \eta_k(x)). \tag{9.3.4}$$

If μ and $\eta_1, \ldots, \eta_k$ are continuous on E, then (9.3.3) is a continuous vector field on E.

Let U be a nonempty open subset of $\mathbf{R}^n$, and let r be a nonnegative integer. A tensor field of type $(1, k)$ on U is said to be *r-times continuously differentiable* if it corresponds to an r-times continuously-differentiable mapping from U into $\mathbf{R}^{n^{k+1}}$ as before. This is interpreted as being the same as continuity when $r = 0$, as usual. The space

$$C^r(U, \mathcal{M}_k(\mathbf{R}^n, \mathbf{R}^n)) = C^r\big(U, \mathbf{R}^{n^{k+1}}\big) \tag{9.3.5}$$

of all r-times continuously-differentiable tensor fields of type $(1, k)$ on U is a linear subspace of the space of all continuous tensor fields of type $(1, k)$ on U.

A tensor field of type $(1, k)$ on U is said to be *infinitely differentiable* or *smooth* if it corresponds to an infinitely-differentiable mapping from U into $\mathbf{R}^{n^{k+1}}$ as before. The space

$$C^\infty(U, \mathcal{M}_k(\mathbf{R}^n, \mathbf{R}^n)) = C^\infty\big(U, \mathbf{R}^{n^{k+1}}\big) \tag{9.3.6}$$

of all infinitely-differentiable tensor fields of type $(1, k)$ on U is a linear subspace of (9.3.5) for each r.

Let μ be a tensor field of type $(1, k)$ on U, and let $\eta_1, \ldots, \eta_k$ be vector fields on U. If μ and $\eta_1, \ldots, \eta_k$ are r-times continuously differentiable on U, then

(9.3.7) $\mu(\eta_1, \ldots, \eta_k)$ is r-times continuously differentiable on U,

as a vector field on U. Similarly, if μ and $\eta_1, \ldots, \eta_k$ are infinitely differentiable on U, then

(9.3.8) $\mu(\eta_1, \ldots, \eta_k)$ is infinitely differentiable on U

as well.

9.4 Some more multilinearity conditions

Let k and n be positive integers, let E be a nonempty subset of $\mathbf{R}^n$, and let μ be a tensor field of type $(1, k)$ on E, as in the previous section. If $\eta_1, \ldots, \eta_k$ are vector fields on E, then (9.3.3) is another vector field on E, as before. In fact,

$$(\eta_1, \ldots, \eta_k) \mapsto \mu(\eta_1, \ldots, \eta_k) \tag{9.4.1}$$

defines a multilinear mapping from the space of k-tuples of vector fields on E into the space of vector fields on E. If f_l is a real-valued function on E for some positive integer $l \leq k$, then

$$\mu(\eta_1, \ldots, \eta_{l-1}, f_l\, \eta_l, \eta_{l+1}, \ldots, \eta_k) = f_l\, \mu(\eta_1, \ldots, \eta_k) \tag{9.4.2}$$

on E. This may be considered as a stronger version of multilinearity of (9.4.1), as in Section 9.1.

As before, the spaces of real-valued functions on E and vector fields on E may be considered as modules over the algebra of real-valued functions on E. We may say that (9.4.1) is multilinear over the algebra of real-valued functions on E, because of (9.4.2). This is related to some remarks on p65 of [184].

Suppose that

(9.4.3) μ is continuous on E,

and remember that the space $C(E, \mathbf{R}^n)$ of continuous vector fields on E may be considered as a module over $C(E, \mathbf{R})$, as in Subsection 9.1.1. If $\eta_1, \ldots, \eta_k$ are continuous on E, then (9.3.3) is continuous on E, as in the previous section. In this case, (9.4.1) is *multilinear over* $C(E, \mathbf{R})$, as a mapping from $C(E, \mathbf{R}^n)^k$ into $C(E, \mathbf{R}^n)$, because of (9.4.2).

Let U be a nonempty open subset of $\mathbf{R}^n$, let r be a nonnegative integer, and suppose that μ is an r-times continuously-differentiable tensor field of type $(1, k)$ on U, as in the previous section. The space $C^r(U, \mathbf{R}^n)$ of r-times continuously-differentiable vector fields on U may be considered as a module over $C^r(U, \mathbf{R})$, as in Subsection 9.1.2. If $\eta_1, \ldots, \eta_k$ are r-times continuously-differentiable vector fields on U, then (9.3.3) is an r-times continuously-differentiable vector field on U as well, as in (9.3.7). Under these conditions, (9.4.1) is *multilinear over* $C^r(U, \mathbf{R})$, as a mapping from $C^r(U, \mathbf{R}^n)^k$ into $C^r(U, \mathbf{R}^n)$, because of (9.4.2).

Similarly, suppose that μ is infinitely differentiable as a tensor field of type $(1, k)$ on U, as in the previous section. The space $C^\infty(U, \mathbf{R}^n)$ of infinitely-differentiable vector fields on U may be considered as a module over $C^\infty(U, \mathbf{R})$, as in Subsection 9.1.2. If $\eta_1, \ldots, \eta_k$ are infinitely-differentiable vector fields on U, then (9.3.3) is an infinitely-differentiable vector field on U too, as in (9.3.8). If we consider (9.4.1) as a mapping from $C^\infty(U, \mathbf{R}^n)^k$ into $C^\infty(U, \mathbf{R}^n)$, then it is *multilinear over* $C^\infty(U, \mathbf{R})$, because of (9.4.2).

9.5 Some more Lie derivatives

Let n be a positive integer, and let U be a nonempty open subset of $\mathbf{R}^n$. Also let k be a positive integer, and let a be a tensor field of type $(0, k)$ on U that is r-times continuously differentiable on U for some positive integer r, as in Sections 4.1 and 4.2. If ξ is an r-times continuously-dfferentiable vector field on U, then we would like to define the *Lie derivative*

$$L_\xi(a) \tag{9.5.1}$$

of a with respect to ξ as a tensor field of type $(0, k)$ on U that is $(r-1)$-times continuously differentiable on U. This type of Lie derivative is mentioned on p70 of [184].

If $\eta_1, \ldots, \eta_k$ are vector fields on U that are r-times continuously differentiable, then

$$a(\eta_1, \ldots, \eta_k) \tag{9.5.2}$$

is a real-valued function on U that is r-times continuously differentiable as well, as in Section 4.3. We would like to define (9.5.1) so that

$$\begin{aligned} L_\xi(a(\eta_1, \ldots, \eta_k)) &= (L_\xi(a))(\eta_1, \ldots, \eta_k) \\ &\quad + \sum_{l=1}^k a(\eta_1, \ldots, \eta_{l-1}, L_\xi(\eta_l), \eta_{l+1}, \ldots, \eta_k), \end{aligned} \tag{9.5.3}$$

where the Lie deriviatives of (9.5.2) and $\eta_1, \ldots, \eta_k$ are as in Section 8.8. This means that

$$\begin{aligned} \mathcal{X}_\xi(a(\eta_1, \ldots, \eta_k)) &= (L_\xi(a))(\eta_1, \ldots, \eta_k) \\ &\quad + \sum_{l=1}^k a(\eta_1, \ldots, \eta_{l-1}, [\xi, \eta_l], \eta_{l+1}, \ldots, \eta_k), \end{aligned} \tag{9.5.4}$$

where $\mathcal{X}_\xi$ and $[\xi, \eta]$ are as in Sections 4.3 and 8.6.

Thus we would like to define (9.5.1) so that

$$\begin{aligned} (L_\xi(a))(\eta_1, \ldots, \eta_k) &= \mathcal{X}_\xi(a(\eta_1, \ldots, \eta_k)) \\ &\quad - \sum_{l=1}^k a(\eta_1, \ldots, \eta_{l-1}, [\xi, \eta_l], \eta_{l+1}, \ldots, \eta_k) \end{aligned} \tag{9.5.5}$$

on U. This can be used to define (9.5.1) as a tensor field of type $(0,k)$ on U, by taking $\eta_1,\ldots,\eta_k$ to be constant vector fields on U equal to any combination of standard basis vectors in $\mathbf{R}^n$. More precisely,

$$L_\xi(a) \text{ is } (r-1)\text{-times continuously differentiable on } U, \tag{9.5.6}$$

because the right side of (9.5.5) is $(r-1)$-times continuously differentiable on U. Similarly, if a and ξ are infinitely differentiable on U, then

$$L_\xi(a) \text{ is infinitely differentiable on } U. \tag{9.5.7}$$

If (9.5.1) is defined as in (9.5.5) when $\eta_1,\ldots,\eta_k$ are constant vector fields on U, each of which is a standard basis vector in $\mathbf{R}^n$, then it is easy to see that (9.5.5) holds for any constant vector fields $\eta_1,\ldots,\eta_k$ on U, using multilinearity. In fact, (9.5.5) holds when $\eta_1,\ldots,\eta_k$ are any continuously-differentiable vector fields on U. This will be discussed further in the next section.

9.6 Using nonconstant vector fields

Let us continue with the same notation and hypotheses as in the previous section. Let $\eta_1,\ldots,\eta_k$ be continuously-differentiable vector fields on U, and let $f_1,\ldots,f_k$ be continuously-differentiable real-valued functions on U. Note that

$$f_1\,\eta_1,\ldots,f_k\,\eta_k \tag{9.6.1}$$

are continuously-differentiable vector fields on U as well, as in Section 8.8. Of course,

$$a(f_1\,\eta_1,\ldots,f_k\,\eta_k) = (f_1\cdots f_k)\,a(\eta_1,\ldots,\eta_k) \tag{9.6.2}$$

on U. We also have that

$$[\xi, f_l\,\eta_l] = \mathcal{X}_\xi(f_l)\,\eta_l + f_l\,[\xi,\eta_l] \tag{9.6.3}$$

on U for each $l=1,\ldots,k$, as in Section 8.8.

It follows from (9.6.2) that

$$\begin{aligned}\mathcal{X}_\xi(a(f_1\,\eta_1,\ldots,f_k\eta_k)) &= \mathcal{X}_\xi(f_1\cdots f_k)\,a(\eta_1,\ldots,\eta_k)\\ &\quad +(f_1\cdots f_k)\,\mathcal{X}_\xi(a(\eta_1,\ldots,\eta_k))\end{aligned} \tag{9.6.4}$$

on U. We can use (9.6.3) to get that

$$\begin{aligned}&a(f_1\,\eta_1,\ldots,f_{l-1}\,\eta_{l-1},[\xi,f_l\,\eta_l],f_{l+1}\,\eta_{l+1},\ldots,f_k\,\eta_k)\\ &\quad = (f_1\cdots f_{l-1}\,\mathcal{X}_\xi(f_l)\,f_{l+1}\cdots f_k)\,a(\eta_1,\ldots,\eta_k)\\ &\qquad +(f_1\cdots f_k)\,a(\eta_1,\ldots,\eta_{l-1},[\xi,\eta_l],\eta_{l+1},\ldots,\eta_k)\end{aligned} \tag{9.6.5}$$

on U for each $l = 1, \ldots, k$. Combining (9.6.4) and (9.6.5), we obtain that

$$\begin{aligned}(9.6.6)\quad & \mathcal{X}_\xi(a(f_1\,\eta_1, \ldots, f_k\,\eta_k)) \\ & \quad - \sum_{l=1}^{k} a(f_1\,\eta_1, \ldots, f_{l-1}\,\eta_{l-1}, [\xi, f_l\,\eta_l], f_{l+1}\,\eta_{l+1}, \ldots, f_k\,\eta_k) \\ = & \quad (f_1 \cdots f_k)\Big(\mathcal{X}_\xi(a(\eta_1, \ldots, \eta_k)) - \sum_{l=1}^{k} a(\eta_1, \ldots, \eta_{l-1}, [\xi, \eta_l], \eta_{l+1}, \ldots, \eta_k)\Big)\end{aligned}$$

on U.

Note that

$$(L_\xi(a))(f_1\,\eta_1, \ldots, f_k\,\eta_k) = (f_1 \cdots f_k)\,(L_\xi(a))(\eta_1, \ldots, \eta_k) \tag{9.6.7}$$

on U, as in (9.6.2). Suppose that $\eta_1, \ldots, \eta_k$ are constant vector fields on U, so that (9.5.5) holds on U, by construction. This means that the right side of (9.6.7) is equal to the right side of (9.6.6), so that

$$\begin{aligned}(L_\xi(a))(f_1\,\eta_1, \ldots, f_k\,\eta_k) = \mathcal{X}_\xi(a(f_1\,\eta_1, \ldots, f_k\,\eta_k)) \\ - \sum_{l=1}^{k} a(f_1\,\eta_1, \ldots, f_{l-1}\,\eta_{l-1}, [\xi, f_l\,\eta_l], f_{l+1}\,\eta_{l+1}, \ldots, f_k\,\eta_k)\end{aligned} \tag{9.6.8}$$

on U. This is the same as (9.5.5), with $\eta_1, \ldots, \eta_k$ replaced with $f_1\,\eta_1, \ldots, f_k\,\eta_k$. One can use this and multilinearity to get that (9.5.5) holds for all continuously-differentiable vector fields $\eta_1, \ldots, \eta_k$ on U.

9.7 Another family of Lie derivatives

Let k and n be positive integers, and let U be a nonempty open subset of $\mathbf{R}^n$. Also let μ be a tensor field of type $(1, k)$ on U, as in Section 9.3. Suppose that μ is r-times continuously differentiable on U for some positive integer r, and let ξ be an r-times continuously-differentiable vector field on U. We would like to define the *Lie derivative*

$$L_\xi(\mu) \tag{9.7.1}$$

of μ with respect to ξ as an $(r-1)$-times continuously-differentiable tensor field of type $(1, k)$ on U. This type of Lie deriviative is mentioned on p70 of [184] as well.

Let $\eta_1, \ldots, \eta_k$ be r-times continuously-differentiable vector fields on U, so that

$$\mu(\eta_1, \ldots, \eta_k) \tag{9.7.2}$$

is r-times continuously differentiable as a vector field on U, as before. We would like to define (9.7.1) so that

$$\begin{aligned} L_\xi(\mu(\eta_1, \ldots, \eta_k)) \quad = \quad & (L_\xi(\mu))(\eta_1, \ldots, \eta_k) \\ & + \sum_{l=1}^{k} \mu(\eta_1, \ldots, \eta_{l-1}, L_\xi(\eta_l), \eta_{l+1}, \ldots, \eta_k),\end{aligned} \tag{9.7.3}$$

using Lie derivatives of vector fields as in Section 8.8. This is the same as saying that

$$(9.7.4)\quad [\xi, \mu(\eta_1, \ldots, \eta_k)] = (L_\xi(\mu))(\eta_1, \ldots, \eta_k) + \sum_{l=1}^{k} \mu(\eta_1, \ldots, \eta_{l-1}, [\xi, \eta_l], \eta_{l+1}, \ldots, \eta_k),$$

where the brackets of vector fields are as in Section 8.6.1.

As in Section 9.5, we can define (9.7.1) by taking

$$(9.7.5)\quad (L_\xi(\mu))(\eta_1, \ldots, \eta_k) = [\xi, \mu(\eta_1, \ldots, \eta_k)] - \sum_{l=1}^{k} \mu(\eta_1, \ldots, \eta_{l-1}, [\xi, \eta_l], \eta_{l+1}, \ldots, \eta_k)$$

on U when $\eta_1, \ldots, \eta_k$ are constant vector fields on U equal to any combination of standard basis vectors in $\mathbf{R}^n$. Note that

$$(9.7.6)\quad L_\xi(\mu) \text{ is } (r-1)\text{-times continuously differentiable on } U,$$

because the right side of (9.7.5) is r-times continuously differentiable on U. Similarly, if μ and ξ are infinitely differentiable on U, then

$$(9.7.7)\quad L_\xi(\mu) \text{ is infinitely differentiable on } U.$$

It is easy to see that (9.7.5) also holds when $\eta_1, \ldots, \eta_k$ are any constant vector fields on U, using multilinearity. We would like to check that (9.7.5) when $\eta_1, \ldots, \eta_k$ are any continuously-differentiable vector fields on U too.

9.7.1 Using nonconstant vector fields again

Let $\eta_1, \ldots, \eta_k$ be continuously-differentiable vector fields on U, and let $f_1, \ldots, f_k$ be continuously-differentiable real-valued functions on U, so that

$$(9.7.8)\quad f_1\, \eta_1, \ldots, f_k\, \eta_k$$

are continuously-differentiable vector fields on U, as before. Note that

$$(9.7.9)\quad \mu(f_1\, \eta_1, \ldots, f_k\, \eta_k) = (f_1 \cdots f_k)\, \mu(\eta_1, \ldots, \eta_k),$$

as vector fields on U. This implies that

$$(9.7.10)\quad [\xi, \mu(f_1\, \eta_1, \ldots, f_k\, \eta_k)] = \mathcal{X}_\xi(f_1 \cdots f_k)\, \mu(\eta_1, \ldots, \eta_k) + (f_1 \cdots f_k)\, [\xi, \mu(\eta_1, \ldots, \eta_k)]$$

on U, as in Section 8.8.

We also have that

$$(9.7.11) \qquad \begin{aligned} &\mu(f_1\,\eta_1, \ldots, f_{l-1}\,\eta_{l-1}, [\xi, f_l\,\eta_l], f_{l+1}\,\eta_{l+1}, \cdots, f_k\,\eta_k) \\ &\quad = \quad (f_1 \cdots f_{l-1}\,\mathcal{X}_\xi(f_l)\,f_{l+1} \cdots f_k)\,\mu(\eta_1, \ldots, \eta_k) \\ &\qquad\quad +(f_1 \cdots, f_k)\,\mu(\eta_1, \ldots, \eta_{l-1}, [\xi, \eta_l], \eta_{l+1}, \ldots, \eta_k) \end{aligned}$$

on U for each $l = 1, \ldots, k$, as in the previous section. Combining this with (9.7.10), we get that

$$(9.7.12) \begin{aligned} &[\xi, \mu(f_1\,\eta_1, \ldots, f_k\eta_k)] \\ &\quad - \sum_{l=1}^{k} \mu(f_1\,\eta_1, \ldots, f_{l-1}\,\eta_{l-1}, [\xi, f_l\,\eta_l], f_{l+1}\,\eta_{l+1}, \ldots, f_k\,\eta_k) \\ &= \quad (f_1 \cdots f_k)\,\Big([\xi, \mu(\eta_1, \ldots, \eta_k)] - \sum_{l=1}^{n} \mu(\eta_1, \ldots, \eta_{l-1}, [\xi, \eta_l], \eta_{l+1}, \ldots, \eta_k)\Big) \end{aligned}$$

on U.

Of course,

$$(9.7.13) \qquad (L_\xi(\mu))(f_1\,\eta_1, \ldots, f_k\,\eta_k) = (f_1 \cdots f_k)\,(L_\xi(\mu))(\eta_1, \ldots, \eta_k)$$

on U, as in (9.7.9). If $\eta_1, \ldots, \eta_k$ are constant vector fields on U, then the right side of (9.7.12) is equal to the right side of (9.7.13), as in (9.7.5). This implies that

$$(9.7.14) \qquad \begin{aligned} &(L_\xi(\mu))(f_1\,\eta_1, \ldots, f_k\,\eta_k) = [\xi, \mu(f_1\,\eta_1, \ldots, f_k\,\eta_k)] \\ &\quad - \sum_{l=1}^{k} \mu(f_1\,\eta_1, \ldots, f_{l-1},\,\eta_{l-1}, [\xi, f_l\,\eta_l].f_{l+1}\,\eta_{l+1}, \ldots, f_k\,\eta_k) \end{aligned}$$

on U. This is the same as (9.7.5), with $\eta_1, \ldots, \eta_k$ replaced with $f_1\,\eta_, \ldots, f_k\,\eta_k$. One can use this to get that (9.7.5) holds for all continuously-differentiable vector fields $\eta_1, \ldots, \eta_k$ on U.

9.8 Lie derivatives as representations

Let n and r be positive integers, and let U be a nonempty open subset of $\mathbf{R}^n$. Also let ξ, ζ be r-times continuously-differentiable vector fields on U, so that $[\xi, \zeta]$ is an $(r-1)$-times continuously-differentiable vector field on U, as in Subsection 8.6.1. If f is an r-times continuously-differentiable real-valued function on U, then

$$(9.8.1) \qquad L_{[\xi,\zeta]}(f) = \mathcal{X}_{[\xi,\zeta]}(f)$$

is an $(r-1)$-times continuously=differentiable real-valued function on U, as in Section 8.8. If f is twice continuously differentiable on U, then

$$(9.8.2) \quad L_{[\xi,\zeta]}(f) = \mathcal{X}_\xi(\mathcal{X}_\zeta(f)) - \mathcal{X}_\zeta(\mathcal{X}_\xi(f)) = L_\xi(L_\zeta(f)) - L_\zeta(L_\xi(f))$$

on U, where the first step is as in Subsection 8.6.1.

If ξ is an infinitely-differentiable vector field on U, then $\mathcal{X}_\xi$ defines a linear mapping from the space $C^\infty(U, \mathbf{R})$ of infinitely-differentiable real-valued functions on U into itself, as in Subsection 4.3.1. Thus we may consider

$$\xi \mapsto L_\xi = \mathcal{X}_\xi \tag{9.8.3}$$

as a linear mapping from the space $C^\infty(U, \mathbf{R}^n)$ of infinitely-differentiable vector fields on U into the space

$$\mathcal{L}(C^\infty(U, \mathbf{R})) \tag{9.8.4}$$

of linear mappings from $C^\infty(U, \mathbf{R})$ into itself. Remember that $C^\infty(U, \mathbf{R}^n)$ is a Lie algebra over the real numbers with respect to the Lie bracket, as mentioned in Section 8.7.1. Using (9.8.2), we get that (9.8.3) defines a representation of $C^\infty(U, \mathbf{R}^n)$, as a Lie algebra over $\mathbf{R}$, on $C^\infty(U, \mathbf{R})$, as in Section B.7.

Suppose that $r \geq 2$, and let η be another r-times continuously-differentiable vector field on U. Note that

$$L_{[\xi,\zeta]}(\eta) = [[\xi, \zeta], \eta] \tag{9.8.5}$$

on U, as in Section 8.8. One can check that

$$L_{[\xi,\zeta]}(\eta) = [\xi, [\zeta, \eta]] - [\zeta, [\xi, \eta]] = L_\xi(L_\zeta(\eta)) - L_\zeta(L_\xi(\eta)) \tag{9.8.6}$$

on U, using the Jacobi identity in the first step, as in Subsection 8.7.1.

If ξ is an infinitely-differentiable vector field on U, then L_ξ defines a linear mapping from $C^\infty(U, \mathbf{R}^n)$ into itself, as in Subsection 8.6.1. In fact,

$$\xi \mapsto L_\xi = [\xi, \cdot] \tag{9.8.7}$$

defines a linear mapping from $C^\infty(U, \mathbf{R}^n)$ into the space

$$\mathcal{L}(C^\infty(U, \mathbf{R}^n)) \tag{9.8.8}$$

of linear mappings from $C^\infty(U, \mathbf{R}^n)$ into itself. This defines a representation of $C^\infty(U, \mathbf{R}^n)$, as a Lie algebra over $\mathbf{R}$, on itself, as a vector space over the real numbers. This is the same as the adjoint representation of $C^\infty(U, \mathbf{R}^n)$, as in Section B.8.

9.9 Some more representations

Let n, k, and r be positive integers, let U be a nonempty open subset of $\mathbf{R}^n$, and let a be a tensor field of type $(0, k)$ on U that is r-times continuously differentiable on U. Suppose that $r \geq 2$, and that ξ, ζ are r-times continuously-differentiable vector fields on U. Thus $[\xi, \zeta]$ is an $(r-1)$-times continuously-differentiable vector field on U, as in Subsection 8.6.1, and one can check that

$$L_{[\xi,\zeta]}(a) = L_\xi(L_\zeta(a)) - L_\zeta(L_\xi(a)) \tag{9.9.1}$$

on U, where these Lie derivatives of a are as in Section 9.5. This uses the analogous statements for Lie derivatives of functions and vector fields mentioned in the previous section.

Remember that $C^\infty(U, \mathcal{M}_k(\mathbf{R}^n))$ is the space of all infinitely-differentiable tensor fields of type $(0, k)$ on U, as in Section 4.2. If ξ is infinitely differentiable on U, then L_ξ defines a linear mapping from $C^\infty(U, \mathcal{M}_k(\mathbf{R}^n))$ into itself, as in Section 9.5. Note that

$$\xi \mapsto L_\xi \tag{9.9.2}$$

defines a linear mapping from $C^\infty(U, \mathbf{R}^n)$ into the space

$$\mathcal{L}(C^\infty(U, \mathcal{M}_k(\mathbf{R}^n))) \tag{9.9.3}$$

of linear mappings from $C^\infty(U, \mathcal{M}_k(\mathbf{R}^n))$ into itself. This is a representation of $C^\infty(U, \mathbf{R}^n)$, as a Lie algebra over the real numbers with respect to the Lie bracket, on $C^\infty(U, \mathcal{M}_k(\mathbf{R}^n))$, because of (9.9.1).

Similarly, let μ be a tensor field of type $(1, k)$ on U that is r-times continuously differeneitable on U. One can check that

$$L_{[\xi,\zeta]}(\mu) = L_\xi(L_\zeta(\mu)) - L_\zeta(L_\xi(\mu)) \tag{9.9.4}$$

on U, where these Lie derivatives are as in Section 9.7. This uses the analogous statement for vector fields in the previous section.

The space of all infinitely-differentiable tensor fields of type $(1, k)$ on U is denoted $C^\infty(U, \mathcal{M}_k(\mathbf{R}^n, \mathbf{R}^n))$, as in Section 9.3. If ξ is infinitely differentiable on this space, then L_ξ defines a linear mapping from this space into itself, as in Section 9.7. As before, $\xi \mapsto L_\xi$ defines a linear mapping from $C^\infty(U, \mathbf{R}^n)$ into the space

$$\mathcal{L}(C^\infty(U, \mathcal{M}_k(\mathbf{R}^n, \mathbf{R}^n))) \tag{9.9.5}$$

of linear mappings from $C^\infty(U, \mathcal{M}_k(\mathbf{R}^n, \mathbf{R}^n))$ into itself. This is a representation of $C^\infty(U, \mathbf{R}^n)$, as a lie algebra over $\mathbf{R}$ with respect to the Lie bracket, on $C^\infty(U, \mathcal{M}_k(\mathbf{R}^n, \mathbf{R}^n))$, by (9.9.4).

9.10 Some properties of Lie derivatives

Let n, k, and r be positive integers, let U be a nonempty open subset of $\mathbf{R}^n$, let a be a tensor field of type $(0, k)$ on U that is r-times continuously differentiable on U, and let ξ be an r-times continuously-differentiable vector field on U. If σ is a permutation on $\{1, \ldots, k\}$, then let a^σ be the tensor field of type $(0, k)$ defined on U by

$$(a^\sigma)_x = (a_x)^\sigma \tag{9.10.1}$$

for every $x \in U$. More precisely, if $x \in U$, then a_x is a k-linear form on $\mathbf{R}^n$, so that $(a_x)^\sigma$ may be defined as a k-linear form on $\mathbf{R}^n$ as in Section 1.9. It is easy to see that a^σ is r-times continuously differentiable on U as well, as in Section 4.2.

One can verify that

$$L_\xi(a^\sigma) = (L_\xi(a))^\sigma, \tag{9.10.2}$$

where these Lie derivatives are as in Section 9.5. If a_x is a symmetric k-linear form on $\mathbf{R}^n$ for every $x \in U$, then it follows that

(9.10.3) $(L_\xi(a))_x$ is a symmetric k-linear form on $\mathbf{R}^n$ for every $x \in U$.

Similarly, if a_x is an alternating k-linear form on $\mathbf{R}^n$ for every $x \in U$, then

(9.10.4) $(L_\xi(a))_x$ is an alternating k-linear form on $\mathbf{R}^n$ for every $x \in U$.

This means that

(9.10.5) if a is a differential k-form on U, then $L_\xi(a)$ is a differential k-form on U too.

Lie derivatives of smooth differential forms are discussed on p70 of [184]. The definition of the Lie derivative of a differential form used here corresponds to part (e) of Proposition 2.25 on p70 of [184].

9.10.1 Lie derivatives of products

Let f be an r-times continuously-differentiable real-valued function on U, and note that $f\,a$ defines an r-times continuously-differentiable tensor field of type $(0, k)$ on U. If $\eta_1, \ldots, \eta_k$ are vector fields on U, then

$$(f\,a)(\eta_1, \ldots, \eta_k) = f\,a(\eta_1, \ldots, \eta_k), \tag{9.10.6}$$

as real-valued functions on U. Remember that $L_\xi(f) = \mathcal{X}_\xi(f)$ on U, as in Section 8.8. One can check that

$$L_\xi(f\,a) = L_\xi(f)\,a + f\,L_\xi(a), \tag{9.10.7}$$

as tensor fields of type $(0, k)$ on U.

Let k' be another positive integer, and let b be a tensor field of type $(0, k')$ on U that is r-times continuously differentiable on U. Remember that $a \otimes b$ may be defined as a tensor field of type $(0, k_1 + k_2)$ on U that is r-times continuously differentiable on U, as in Sections 4.1 and 4.2. If $\eta_1 \ldots, \eta_{k+k'}$ are vector fields on U, then

$$(a \otimes b)(\eta_1, \ldots, \eta_{k+1}) = a(\eta_1, \ldots, \eta_k)\, b(\eta_{k+1}, \ldots, \eta_{k+k'}), \tag{9.10.8}$$

as real-valued functions on U. One can verify that

$$L_\xi(a \otimes b) = (L_\xi(a)) \otimes b + a \otimes (L_\xi(b)), \tag{9.10.9}$$

as tensor fields of type $(0, k + k')$ on U.

If a, b are differential k, k'-forms on U, respectively, then

$$L_\xi(a \wedge b) = (L_\xi(a)) \wedge b + a \wedge (L_\xi(b)), \tag{9.10.10}$$

as differential $(k + k')$-forms on U. This can be obtained from (9.10.9) using (9.10.2). This corresponds to part of part (c) of Proposition 2.25 on p70 of [184].

9.10.2 Lie derivatives and exterior differentiation

Let f be an $(r+1)$-times continuously-differentiable real-valued function on U, so that

$$a = df \tag{9.10.11}$$

is an r-times continuously-differentiable differential one-form on U, or equivalently a tensor field of type $(0,1)$ on U. If η is an r-times continuously-differentiable vector field on U, then

$$a(\eta) = (df)(\eta) = \mathcal{X}_\eta(f) \tag{9.10.12}$$

is an r-times continuously-differentiable real-valued function on U. Observe that

$$\begin{aligned} (L_\xi(a))(\eta) = \mathcal{X}_\xi(a(\eta)) - a([\xi,\eta]) &= \mathcal{X}_\xi(\mathcal{X}_\eta(f)) - \mathcal{X}_{[\xi,\eta]}(f) \\ &= \mathcal{X}_\eta(\mathcal{X}_\xi(f)) = (d(\mathcal{X}_\xi(f)))(\eta), \end{aligned} \tag{9.10.13}$$

where the first step is as in Section 9.5, and the second step is as in Subsection 8.6.1. This means that

$$L_\xi(df) = d(L_\xi(f)), \tag{9.10.14}$$

as differential one-forms on U.

Of course, $L_\xi(f) = \mathcal{X}_\xi(f)$ is an r-times continuously-differentiabl real-valued function on U. If $r \geq 2$, then it follows that

$$d(L_\xi(df)) = d(d(L_\xi(f))) = 0, \tag{9.10.15}$$

as in Section 4.8.

Let α be a differential k-form on U that is twice continuously differentiable on U, so that $d\alpha$ is a differential $(k+1)$-form on U that is continuously differentiable on U. If ξ is twice continuously differentiable on U, then one can check that

$$L_\xi(d\alpha) = d(L_\xi(\alpha)) \tag{9.10.16}$$

as differential $(k+1)$-forms on U. More precisely, if $k=0$, then this is the same as (9.10.14), and it is enough to ask that ξ be continuously differentiable. If $k \geq 1$, then this can be obtained using the earlier remarks about Lie derivatives of products. This corresponds to another part of part (c) of Proposition 2.25 on p70 of [184].

9.10.3 Tensors of type $(1,k)$

Let μ be a tensor field of type $(1,k)$ on U that is r-times continuously differentiable on U. If $\sigma \in \mathrm{Sym}(k)$, then let μ^σ be the tensor field of type $(1,k)$ on U defined by

$$(\mu^\sigma)_x = (\mu_x)^\sigma \tag{9.10.17}$$

for each $x \in U$, where the right side is as in Section 9.2. Note that μ^σ is r-times continuously differentiable on U as well.

One can check that

(9.10.18) $$L_\xi(\mu^\sigma) = (L_\xi(\mu))^\sigma,$$

where these Lie derivatives are as in Section 9.7. If μ_x is symmetric as a multilinear mapping from $(\mathbf{R}^n)^k$ into $\mathbf{R}^n$ for each $x \in U$, then it follows that

(9.10.19) $(L_\xi(\mu))_x$ is symmetric as a multilinear mapping from $(\mathbf{R}^n)^k$ into $\mathbf{R}^n$ for every $x \in U$.

Similarly, if μ_x is alternating as a multilinear mapping from $(\mathbf{R}^n)^k$ into $\mathbf{R}^n$ for each $x \in U$, then

(9.10.20) $(L_\xi(\mu))_x$ is alternating as a multilinear mapping from $(\mathbf{R}^n)^k$ into $\mathbf{R}^n$ for every $x \in U$.

Let a be a tensor field of type $(0, k)$ on U that is r-times continuously differentiable on U agagin, and let ζ be a vector field on U that is r-times continuously differentiable on U. Consider

(9.10.21) $$\mu = a\,\zeta,$$

as in Section 9.3, which is also r-times continuously differentiable on U. If $\eta_1, \ldots, \eta_k$ are vector fields on U, then

(9.10.22) $$\mu(\eta_1, \ldots, \eta_k) = a(\eta_1, \ldots, \eta_k)\,\zeta,$$

as vector fields on U. One can check that

(9.10.23) $$L_\xi(a\,\xi) = (L_\xi(a))\,\zeta + a\,L_\xi(\zeta),$$

as tensor fields of type $(1, k)$ on U.

9.11 Some interior products

Let W be a vector space over the real numbers, and let u be an element of W. Also let k and p be positive integers, with $p \le k$. If μ is a k-linear form on W, then let

(9.11.1) $$(i_{k,p}(u))(\mu)$$

be the $(k-1)$-linear form on W defined by

(9.11.2) $$((i_{k,p}(u))(\mu))(w_1, \ldots, w_{k-1}) = \mu(w_1, \ldots, w_{p-1}, u, w_p, \ldots, w_{k-1}).$$

This defines a linear mapping

(9.11.3) $$i_{k,p}(u)$$

from $\mathcal{M}_k(W)$ into $\mathcal{M}_{k-1}(W)$. This linear mapping may be described as the *pth interior product by u* on $\mathcal{M}_k(W)$.

If $k = 1$, then μ is a linear functional on W. In this case, $p = 1$, and

$$(i_{1,1}(u))(\mu) = \mu(u), \tag{9.11.4}$$

which is a real number. Remember that $\mathcal{M}_0(W) = \mathbf{R}$, as in Section 2.12.

Put

$$i_{k,tot}(u) = \sum_{p=1}^{k} i_{k,p}(u), \tag{9.11.5}$$

which is another linear mapping from $\mathcal{M}_k(W)$ into $\mathcal{M}_{k-1}(W)$. This may be described as the *total interior product by u* on $\mathcal{M}_k(W)$. It is convenient to take $i_{k,tot}(u)$ to be 0 on $\mathcal{M}_0(W)$ when $k = 0$.

9.11.1 Connection with products of forms

Let k_1 and k_2 be positive integers, and let μ_1, μ_2 be k_1, k_2-linear forms on W, respectively. Thus $\mu_1 \otimes \mu_2$ may be defined as a $(k_1 + k_2)$-linear form on W, as in Section 2.8. One can check that

$$\begin{aligned}(i_{k_1+k_2,p}(u))(\mu_1 \otimes \mu_2) &\\ = \quad & ((i_{k_1,p}(u))(\mu_1)) \otimes \mu_2 \qquad \text{when } 1 \leq p \leq k_1\\ = \quad & \mu_1 \otimes ((i_{k_2,p-k_1}(u))(\mu_2)) \quad \text{when } k_1 + 1 \leq p \leq k_1 + k_2.\end{aligned} \tag{9.11.6}$$

Using this, one can verify that

$$\begin{aligned}&(i_{k_1+k_2,tot}(u))(\mu_1 \otimes \mu_2)\\ &\quad = \quad ((i_{k_1,tot}(u))(\mu_1)) \otimes \mu_2 + \mu_1 \otimes ((i_{k_2,tot}(u))(\mu_2)).\end{aligned} \tag{9.11.7}$$

This also holds when k_1 or k_2 is 0, with suitable interpretations.

The algebra $\mathcal{M}(W)$ of multilinear forms on W may be defined as the direct sum of $\mathcal{M}_k(W)$ over all nonnegative integers k, as in Section A.6. One can define a linear mapping

$$i_{tot}(u) \tag{9.11.8}$$

from $\mathcal{M}(W)$ into itself, using $i_{k,tot}(u)$ on $\mathcal{M}_k(W)$ for each $k \geq 0$. This may be described as the *total interior product by u* on $\mathcal{M}(W)$. It is easy to see that

$$i_{tot}(u) \text{ is a derivation on } \mathcal{M}(W), \tag{9.11.9}$$

as in Subsection A.2.1.

9.12 Interior products and symmetrizations

Let W be a vector space over the real numbers, let u be an element of W, and let k be a positive integer. If μ is a symmetric k-linear form on W, then it is easy to see that

$$(i_{k,p}(u))(\mu) = (i_{k,q}(u))(\mu) \tag{9.12.1}$$

for all positive integers $p, q \le k$. This implies that

$$k\,(i_{k,p}(u))(\mu) = (i_{k,tot}(u))(\mu) \tag{9.12.2}$$

for each $p \le k$. Note that

$$(i_{k,p}(u))(\mu) \text{ is a symmetric } (k-1)\text{-linear form on } W \tag{9.12.3}$$

in this case. Thus the restriction of $i_{k,p}(u)$ to $\mathcal{SM}_k(W)$ defines a linear mapping into $\mathcal{SM}_{k-1}(W)$.

Suppose that $k \ge 2$, let μ be any k-linear form on W, and let τ be a permutation on $\{1, \ldots, k-1\}$. Observe that

$$\begin{aligned} ((i_{k,p}(u))(\mu))^\tau(w_1, \ldots, w_{k-1}) &= ((i_{k,p}(u))(\mu))(w_{\tau(1)}, \ldots, w_{\tau(k-1)}) \\ &= \mu(w_{\tau(1)}, \ldots, w_{\tau(p-1)}, u, w_{\tau(p)}, \ldots, w_{\tau(k-1)}) \end{aligned} \tag{9.12.4}$$

for each $p \le k$ and $w_1, \ldots, w_{k-1} \in W$, where the left side is as in Section 2.4. We also have that

$$S_{k-1}((i_{k,tot}(u))(\mu)) = \frac{1}{(k-1)!} \sum_{p=1}^{k} \sum_{\tau \in \mathrm{Sym}(k-1)} ((i_{k,p}(u))(\mu))^\tau, \tag{9.12.5}$$

where S_{k-1} is as in Subsection 2.4.1. It follows that

$$\begin{aligned} &S_{k-1}((i_{k,tot}(u))(\mu))(w_1, \ldots, w_{k-1}) \\ &\quad = \frac{1}{(k-1)!} \sum_{p=1}^{k} \sum_{\tau \in \mathrm{Sym}(k-1)} ((i_{k,p}(u))(\mu))^\tau(w_1, \ldots, w_{k-1}) \\ &\quad = \frac{1}{(k-1)!} \sum_{p=1}^{k} \sum_{\tau \in \mathrm{Sym}(k-1)} \mu(w_{\tau(1)}, \ldots, w_{\tau(p-1)}, u, w_{\tau(p)}, \ldots, w_{\tau(k-1)}) \end{aligned} \tag{9.12.6}$$

for all $w_1, \ldots, w_{k-1} \in W$.

One can check that this is equal to

$$k\,(S_k(\mu))(u, w_1, \ldots, w_{k-1}) \tag{9.12.7}$$

for all $w_1, \ldots, w_{k-1} \in W$. One can look at this in terms of shuffles, as in Section 2.15, with $r = 2$, $k_1 = 1$, and $k_2 = k-1$. This means that

$$S_{k-1}((i_{k,tot}(u))(\mu)) = k\,(i_{k,1}(u))(S_k(\mu)). \tag{9.12.8}$$

Equivalently,

$$S_{k-1}((i_{k,tot}(u))(\mu)) = (i_{k,tot}(u))(S_k(\mu)), \tag{9.12.9}$$

because of (9.12.2). Of course, this is the same as saying that

$$S_{k-1} \circ i_{k,tot}(u) = i_{k,tot}(u) \circ S_k \tag{9.12.10}$$

on $\mathcal{M}_k(W)$.

9.12.1 Connection with symmetric products

Let k_1, k_2 be positive integers, and let μ_1, μ_2 be symmetric k_1, k_2-linear forms on W, respectively. Remember that $\mu_1 \odot \mu_2$ may be defined as a symmetric $(k_1 + k_2)$-linear form on W as in Section 2.9. Observe that

$$\begin{aligned} (i_{k_1+k_2,tot}(u))(\mu_1 \odot \mu_2) &= (i_{k_1+k_2,tot}(u))(S_{k_1+k_2}(\mu_1 \otimes \mu_2)) \\ &= S_{k_1+k_2-1}((i_{k_1+k_2,tot}(u))(\mu_1 \otimes \mu_2)), \end{aligned} \tag{9.12.11}$$

using (9.12.9) in the second step. Combining this with (9.11.7), we get that

$$\begin{aligned} (i_{k_1+k_2,tot}(u))(\mu_1 \odot \mu_2) &= S_{k_1+k_2-1}(((i_{k_1,tot}(u))(\mu_1)) \otimes \mu_2) \\ &\quad +S_{k_1+k_2-1,tot}(\mu_1 \otimes ((i_{k_2,tot}(u))(\mu_2))) \\ &= ((i_{k_1,tot}(u))(\mu_1)) \odot \mu_2 + \mu_1 \odot ((i_{k_2,tot}(u))(\mu_2)). \end{aligned} \tag{9.12.12}$$

The algebra $\mathcal{SM}(W)$ of symmetric multilinear forms on W may be defined as the direct sum of $\mathcal{SM}_k(W)$ over all nonnegative integers k, as in Subsection A.6.1. This may be considered as a linear subspace of $\mathcal{M}(W)$. Note that the restriction of $i_{k,tot}(u)$ to $\mathcal{SM}_k(W)$ defines a linear mapping into $\mathcal{SM}_{k-1}(W)$ for each k, so that

$$i_{tot}(u) \text{ maps } \mathcal{SM}(W) \text{ into itself.} \tag{9.12.13}$$

In fact,

$$i_{tot}(u) \text{ is a derivation on } \mathcal{SM}(W), \tag{9.12.14}$$

as in Subsection A.2.1, because of (9.12.12).

9.13 Interior products and polynomials

Let k and n be positive integers, and let us now take $W = \mathbf{R}^n$. If μ is a k-linear form on $\mathbf{R}^n$, then $P_\mu(x) = \mu(x, \ldots, x)$ defines a homogeneous polynomial of degree k on $\mathbf{R}^n$, as in Section 1.8. Let $u \in \mathbf{R}^n$ be given, so that $(i_{k,q}(u))(\mu)$ is defined as a $(k-1)$-linear form on $\mathbf{R}^n$ as in Section 9.11 for each $q = 1, \ldots, k$. Thus

$$P_{(i_{k,q}(u))(\mu)} \tag{9.13.1}$$

defines a homogeneous polynomial of degree $k - 1$ on $\mathbf{R}^n$ for each q. Similarly, $(i_{k,tot}(u))(\mu)$ is a $(k-1)$-linear form on $\mathbf{R}^n$, which corresponds to the homogeneous polynomial

$$P_{(i_{k,tot}(u))(\mu)} = \sum_{q=1}^{k} P_{(i_{k,q}(u))(\mu)} \tag{9.13.2}$$

of degree $k - 1$ on $\mathbf{R}^n$.

One can check that

$$P_{(i_{k,tot}(u))(\mu)}(x) = (P'_\mu(x))(u) = d(P_\mu)_x(u) \tag{9.13.3}$$

for each $x \in \mathbf{R}^n$. This is the same as the directional derivative of P_μ at x in the direction u, as in Subsection 3.8.1.

Similarly, if $\mu \in \mathcal{M}(\mathbf{R}^n)$, then P_μ may be defined as a polynomial on $\mathbf{R}^n$ with real coefficients as in Subsection A.6.2. Remember that $(i_{tot}(u))(\mu)$ may be defined as an element of $\mathcal{M}(\mathbf{R}^n)$ as in Subsection 9.11.1. As before, we have that

$$P_{(i_{tot}(u))(\mu)}(x) = (P'_\mu(x))(u) = d(P_\mu)_x(u) \tag{9.13.4}$$

for every $x \in \mathbf{R}^n$.

9.14 Interior multiplication on $\mathcal{AM}(W)$

Let W be a vector space over the real numbers, let u be an element of W, and let k be a positive integer. If μ is an alternating k-linear form on W, then let

$$(i_k(u))(\mu) \tag{9.14.1}$$

be the $(k-1)$-linear form defined on W by

$$((i_k(u))(\mu))(w_1, \ldots, w_{k-1}) = \mu(u, w_1, \ldots, w_{k-1}). \tag{9.14.2}$$

It is easy to see that this is an alternating $(k-1)$-linear form on W. This defines a linear mapping from $\mathcal{AM}_k(W)$ into $\mathcal{AM}_{k-1}(W)$, which is known as *interior multiplication by u*. This corresponds to part of Section 2.11 on p61 of [184].

Equivalently, $i_k(u)$ is the same as the restriction of $i_{k,1}(u)$, as in Section 9.11, to $\mathcal{AM}_k(W)$. Similarly,

$$i_{k,p}(u) = (-1)^{p-1}\, i_k(u) \text{ on } \mathcal{AM}_k(W) \tag{9.14.3}$$

for each $p \le k$. It is convenient to take $i_k(u)$ to be 0 on $\mathcal{AM}_0(W) = \mathbf{R}$ when $k = 0$. We can define $i(u)$ as a linear mapping from $\mathcal{AM}(W)$ into itself, which corresponds to $i_k(u)$ on $\mathcal{AM}_k(W)$ for each $k \ge 0$. This uses the definition of $\mathcal{AM}(W)$ in Subsection 2.12.1 when W has finite dimension, or the definition in Section A.7 otherwise.

Let k_1 and k_2 be nonnegative integers, and let μ_1, μ_2 be alternating k_1, k_2-linear forms on W, respectively. We would like to show that

$$\begin{aligned}(i_{k_1+k_2}(u))&(\mu_1 \wedge \mu_2) \\ &= ((i_{k_1}(u))(\mu_1)) \wedge \mu_2 + (-1)^{k_1}\, \mu_1 \wedge ((i_{k_2}(u))(\mu_2)).\end{aligned} \tag{9.14.4}$$

This means that $i(u)$ is an *anti-derivation* on $\mathcal{AM}(W)$, as in Section 2.11 on p61 of [184]. This corresponds to Proposition 2.12 on p61 of [184]. It is easy to see that (9.14.4) holds when k_1 or k_2 is 0, and so we suppose now that $k_1, k_2 \ge 1$.

9.14.1 Checking the anti-derivation property

Put

$$E_1 = \{1, \ldots, k_1\} \text{ and } E_2 = \{k_1 + 1, \ldots, k_1 + k_2\}, \tag{9.14.5}$$

as in Section 2.13, with $r = 2$. Remember that $\tau \in \mathrm{Sym}(k_1 + k_2)$ is said to be a shuffle with respect to k_1, k_2 if the restriction of τ to each of E_1 and E_2 is increasing, as in Section 2.15. Let Ξ be the set of $\tau \in \mathrm{Sym}(k_1 + k_2)$ that are shuffles with respect to k_1, k_2, as in Subsection 2.15.1. Thus

$$\mu_1 \wedge \mu_2 = \sum_{\tau \in \Xi} \mathrm{sgn}(\tau)\, (\mu_1 \otimes \mu_2)^\tau, \tag{9.14.6}$$

as in Subsection 2.15.2. Of course, $(\mu_1 \otimes \mu_2)^\tau$ is defined as in Sections 2.4 and 2.8.

If $v_1, \ldots, v_{k_1+k_2} \in W$, then

$$\begin{aligned} (\mu_1 \otimes \mu_2)^\tau(v_1, \ldots, v_{k_1+k_2}) &= (\mu_1 \otimes \mu_2)(v_{\tau(1)}, \ldots, v_{\tau(k_1+k_2)}) \\ &= \mu_1(v_{\tau(1)}, \ldots, v_{\tau(k_1)})\, \mu_2(v_{\tau(k_1+1)}, \ldots, v_{\tau(k_1+k_2)}) \end{aligned} \tag{9.14.7}$$

for every $\tau \in \mathrm{Sym}(k_1 + k_2)$. If τ is a shuffle with respect to k_1, k_2, then it is easy to see that

$$\tau^{-1}(1) = 1 \text{ or } k_1 + 1. \tag{9.14.8}$$

This means that

$$\begin{aligned} \mu_1 \wedge \mu_2 &= \sum_{\tau \in \Xi, \tau^{-1}(1) = 1} \mathrm{sgn}(\tau)\, (\mu_1 \otimes \mu_2)^\tau \\ &\quad + \sum_{\tau \in \Xi, \tau^{-1}(1) = k_1 + 1} \mathrm{sgn}(\tau)\, (\mu_1 \otimes \mu_2)^\tau. \end{aligned} \tag{9.14.9}$$

In order to get (9.14.4), it suffices to verify that

$$\sum_{\tau \in \Xi, \tau^{-1}(1) = 1} \mathrm{sgn}(\tau)\, (i_{k_1+k_2}(u))((\mu_1 \otimes \mu_2)^\tau) = ((i_{k_1}(u))(\mu_1)) \wedge \mu_2 \tag{9.14.10}$$

and

$$\begin{aligned} &\sum_{\tau \in \Xi, \tau^{-1}(1) = k_1 + 1} \mathrm{sgn}(\tau)\, (i_{k_1+k_2}(u))((\mu_1 \otimes \mu_2)^\tau) \\ &\quad = (-1)^{k_1}\, \mu_1 \wedge ((i_{k_2}(u))(\mu_2)). \end{aligned} \tag{9.14.11}$$

This uses (9.14.7) with $v_1 = u$ in both cases.

Note that the wedge products on the right sides of (9.14.10) and (9.14.11) may be expressed as in Subsection 2.15.2 as well, using shuffles with respect to $k_1 - 1$, k_2 and k_1, $k_2 - 1$, respectively. More precisely, this works when $k_1 \geq 2$ or $k_2 \geq 2$, as appropriate, and otherwise these wedge products reduce to the product of a real number and an alternating multilinear form. If $\tau^{-1}(1) = 1$ and $k_1 \geq 2$, then τ corresponds to a shuffle with respect to $k_1 - 1$, k_2 in a simple way, which can be used to obtain (9.14.10). This is easier when $k_1 = 1$, so that $(i_{k_1}(u))(\mu_1)$ is a real number. Similarly, if $\tau^{-1}(1) = k_1 + 1$ and $k_2 \geq 2$, then τ corresponds to a shuffle with respect to k_1, $k_2 - 1$ in a nice way. This can be used to obtain (9.14.11), because the shuffle with respect to k_1, $k_2 - 1$ corresponding to τ has sign equal to $(-1)^{k_1}\,\mathrm{sgn}(\tau)$. This is easier when $k_2 = 1$, as before.

One may focus on finite-dimensional vector spaces, as on p54 of [184], which is often of primary interest anyway. One can also reduce to that case, by considering finite-dimensional linear subspaces of W.

9.15 More on interior multiplication

Let n, k, and r be positive integers, let U be a nonempty open subset of $\mathbf{R}^n$, let α be a differential k-form on U that is r-times continuously differentiable on U, and let ξ be an r-times continuously-differentiable vector field on U. We can define

$$(i_k(\xi))(\alpha) \tag{9.15.1}$$

as a differential $(k-1)$-form on U using interior multiplication pointwise on U. This means that for each $x \in U$,

$$((i_k(\xi))(\alpha))_x = (i_k(\xi(x)))(\alpha_x), \tag{9.15.2}$$

where the right side is defined as an alternating $(k-1)$-linear form on $\mathbf{R}^n$ as in the previous section. One can check that (9.15.1) is r-times continuously differentiable on U as well. It is convenient to interpret (9.15.1) as being equal to 0 when $k = 0$, as before.

9.15.1 Some related linear mappings

Put

$$L_{k,\xi}(\alpha) = (i_{k+1}(\xi))(d\alpha) + d((i_k(\xi))(\alpha)), \tag{9.15.3}$$

which is a differential k-form on U that is $(r-1)$-times continuously differentiable on U. If $k = 0$, then this is interpreted as being

$$L_{0,\xi}(\alpha) = (i_k(\xi))(d\alpha) = (d\alpha)(\xi) = \mathcal{X}_\xi(\alpha). \tag{9.15.4}$$

Let k' be another positive integer, and let β be a differential k'-form on U that is r-times continuously differentiable on U. One can check that

$$L_{k+k',\xi}(\alpha \wedge \beta) = (L_{k,\xi}(\alpha)) \wedge \beta + \alpha \wedge (L_{k',\xi}(\beta)), \tag{9.15.5}$$

as differential $(k + k')$-forms on U. This also works when k or k' is 0, with the usual interpretations. This uses the anti-derivation properties of interior multiplication and exterior differentiation. This corresponds to a remark on p72 of [184], related to the proof of part (d) of Proposition 2.25 on p70 of [184].

Suppose that α is $(r+1)$-times continuously differentiable on U, so that $d\alpha$ is a differential $(k+1)$-form on U that is r-times continuously differentiable on U. In this case,

$$L_{k+1,\xi}(d\alpha) = d((i_{k+1}(\xi))(d\alpha)) \tag{9.15.6}$$

on U, which also works when $k = 0$. If ξ is $(r + 1)$-times continuously differentiable on U too, then (9.15.1) is $(r + 1)$-times continuously differentiable on U as well, and

$$d(L_{k,\xi}(\alpha)) = d((i_{k+1}(\xi))(d\alpha)) \tag{9.15.7}$$

on U. This means that

$$d(L_{k,\xi}(\alpha)) = L_{k+1,\xi}(d\alpha) \tag{9.15.8}$$

on U under these conditions. This corresponds to another remark on p72 of [184].

9.15.2 Connections with Lie derivatives

If f is an $(r + 1)$-times continuously-differentiable real-valued function on U, then df is an r-times continuously-differentiable differential 1-form on U, and

$$L_{1,\xi}(df) = d((i_1(\xi))(df)) = d((df)(\xi)) = d(\mathcal{X}_\xi(f)) \tag{9.15.9}$$

on U. One can check that

$$L_{k,\xi}(\alpha) = L_\xi(\alpha) \tag{9.15.10}$$

on U, where the right side is the Lie derivative of α, as in Section 9.5. This is the same as (9.15.4) when $k = 0$, and it follows from (9.15.9) when $k = 1$ and $\alpha = df$, as in Subsection 9.10.2. One can reduce to these cases using (9.15.5) and the analogous statement for Lie derivatives, as in Subsection 9.10.1. This corresponds to part (d) of proposition 2.25 on p70 of [184].

Appendix A

Some linear and abstract algebra

A.1 Some remarks about vector spaces

Let V be a *vector space* over the real numbers. We shall not get into the abstract definition of a vector space too much here, but it is sometimes convenient to use the terminology. Basically V should be a set with a distinguished element called 0, and with operations of addition and scalar multiplication defined on V that satisfy some standard properties. These include commutativity and associativity of addition, and distributivity of scalar multiplication with respect to addition.

If n is a positive integer, then the space $\mathbf{R}^n$ of n-tuples of real numbers is a vector space over $\mathbf{R}$ with respect to coordinatewise addition and scalar multiplication. Similarly, if X is a nonempty set, then the space of all real-valued functions on X is a vector space over $\mathbf{R}$ with respect to pointwise addition and scalar multiplication.

If V is a vector space over the real numbers, then a *linear subspace* of V is a subset V_0 of V that satisfies the following three conditions. First,

$$0 \in V_0. \tag{A.1.1}$$

Second, if $u, v \in V_0$, then

$$u + v \in V_0. \tag{A.1.2}$$

Third, if $v \in V_0$ and $t \in \mathbf{R}$, then

$$t\, v \in V_0. \tag{A.1.3}$$

Note that (A.1.3) implies (A.1.1) when $V_0 \neq \emptyset$, by taking $t = 0$.

Under these conditions, V_0 is a vector space over the real numbers too, with respect to the restrictions of the operations of addition and scalar multiplication on V to V_0. Many of the vector spaces with which we shall be concerned here are linear subspaces of $\mathbf{R}^n$ for some n, or linear subspaces of the space of all real-valued functions on some nonempty set X.

Let X be a nonempty set again, and let W be a vector space over the real numbers. The space of all W-valued functions on X is a vector space over the real numbers with respect to pointwise addition and scalar multiplication. Thus linear subspaces of this space are vector spaces over $\mathbf{R}$ as well.

A.2 Algebras over the real numbers

Let A be a vector space over the real numbers, and let b be a bilinear mapping from $A \times A$ into A, as in Section 2.2. Under these conditions, (A, b) is said to be an *algebra in the strict sense* over the real numbers. This corresponds to the definition of an F-algebra, with $F = \mathbf{R}$, in Section 1.3 on p4 of [86]. This also corresponds to the definition of a k-algebra, with $k = \mathbf{R}$, on p2 of [160]. One may describe b as a *bilinear operation* on A as well, as in Definition 3.53 on p117 of [184].

If

$$b(x, y) = b(y, x) \tag{A.2.1}$$

for every $x, y \in A$, then (A, b) is said to be *commutative* as an algebra in the strict sense over $\mathbf{R}$. If

$$b(b(x, y), z) = b(x, b(y, z)) \tag{A.2.2}$$

for every $x, y, z \in A$, then (A, b) is said to be an *associative algebra* over the real numbers.

An element e of A is said to be a *multiplicative identity element* with respect to b if

$$b(e, x) = b(x, e) = x \tag{A.2.3}$$

for every $x \in A$. It is easy to see that this is uique when it exists.

Let A_0 be a linear subspace of A. If

$$b(x, y) \in A_0 \tag{A.2.4}$$

for every $x, y \in A$, then A_0 is called a *subalgebra* of A with respect to b. This means that A_0 is an algebra in the strict sense over the real numbers with respect to the restriction of $b(x, y)$ to $x, y \in A$.

A.2.1 Algebra homomorphisms

Let (A_1, b_1) and (A_2, b_2) be algebras over the real numbers in the strict sense, and let ϕ be a linear mapping from A_1 into A_2. We say that ϕ is a *homomorphism* from A_1 into A_2, as algebras in the strict sense with respect to b_1, b_2, respectively, if

$$\phi(b_1(x, y)) = b_2(\phi(x), \phi(y)) \tag{A.2.5}$$

for every $x, y \in A_1$.

If ϕ is a homomorphism from A_1 into A_2 with respect to b_1 and b_2, respectively, then it is easy to see that

(A.2.6) $$\phi(A_1) \text{ is a subalgebra of } A_2 \text{ with respect to } b_2.$$

If A_0 is a subalgebra of A with respect to b, then the obvious inclusion mapping from A_0 into A is a homomorphism, with respect to b on A and its restriction to A_0.

Let (A_3, b_3) be another algebra in the strict sense over the real numbers. If ϕ is a homomorphism from A_1 into A_2 with respect to b_1 and b_2, respectively, and ψ is a homomorphism from A_2 into A_3 with respect to b_2 and b_3, respectively, then one can check that

(A.2.7) $$\psi \circ \phi \text{ is a homomorphism from } A_1 \text{ into } A_3,$$

with respect to b_1 and b_3, respectively.

Suppose that ϕ is a homomorphism from A_1 into A_2 with respect to b_1 and b_2, respectively, again. If ϕ is a one-to-one mapping from A_1 onto A_2, then it is easy to see that the inverse mapping ϕ^{-1} is a homomorphism from A_2 into A_1 with respect to b_2 and b_1, respectively. In this case, ϕ is said to be an *isomorphism* from A_1 onto A_2 with respect to b_1 and b_2, respectively.

If ϕ is an isomorphism from A_1 onto A_2 with respect to b_1 and b_2, respectively, and ψ is an isomorphism from A_2 onto A_3 with respect to b_2 and b_3, respectively, then

(A.2.8) $$\psi \circ \phi \text{ is an isomorphism from } A_1 \text{ onto } A_3,$$

with respect to b_1 and b_3, respectively.

An isomorphism from A onto itself with respect to b is said to be an *automorphism* of A with respect to b. Note that the identity mapping on A is automatically an automorphism of A with respect to b. The set of automorphisms of A with respect to b is a subgroup of the group of all one-to-one mappings from A onto itself, as in Section 2.6.

A linear mapping δ from A into itself is said to be a *derivation* with respect to b if

(A.2.9) $$\delta(b(x, y)) = b(\delta(x), y) + b(x, \delta(y))$$

for every $x, y \in A$.

A.3 More on vector spaces

Let V and W be vector spaces over the real numbers, and let ϕ be a one-to-one linear mapping from V onto W. It is well known and easy to see that ϕ^{-1} is a linear mapping from W onto V in this case. One may say that ϕ is an *isomorphism* from V onto W as vector spaces over $\mathbf{R}$ under these conditions.

The space of one-to-one linear mappings from V onto itself is denoted $GL(V)$. This is a subgroup of the group of all one-to-one mappings from V onto itself, as in Section 2.6. In particular, $GL(V)$ is a group with respect to composition of mappings on V, which is known as the *general linear group* of V.

Let V^* be the space of all linear functionals on V, which is to say the space of all linear mappings from V into $\mathbf{R}$, as a vector space over itself. This is known as the *dual* of V, which is sometimes denoted V'. This is a linear subspace of the space of all real-valued functions on V, as a vector space over the real numbers with respect to pointwise addition and scalar multiplication. In particular, V^* is a vector space over the real numbers.

If V has finite dimension, then it is well known that

$$\dim V^* = \dim V. \tag{A.3.1}$$

If n is a positive integer, then the n standard coordinate functions on $\mathbf{R}^n$ form a basis for the dual of $\mathbf{R}^n$.

A.3.1 Dual linear mappings

Let T be a linear mapping from V into W. If λ is a linear functional on W, then

$$T^*(\lambda) = \lambda \circ T \tag{A.3.2}$$

is a linear functional on V. This defines a linear mapping T^* from W^* into V^*, which is the *dual linear mapping* corresponding to T. This corresponds to the $k = 1$ case of an analogous definition in Section 2.3. If the dual spaces are denoted V', W', then one may use the notation T' for the dual linear mapping associated to T.

Let Z be another vector space over the real numbers, and let R be a linear mapping from W into Z. One can check that

$$(R \circ T)^* = T^* \circ R^*, \tag{A.3.3}$$

as linear mappings from Z^* into V^*, as in Section 2.3. If T is a one-to-one linear mapping from V onto W, then T^* is a one-to-one linear mapping from W^* onto V^*, with

$$(T^*)^{-1} = (T^{-1})^*, \tag{A.3.4}$$

as before.

A.3.2 Direct sums and dual spaces

Let N be a positive integer, and let $V_1, \ldots, V_N$ be N vector spaces over the real numbers. The direct sum of the V_j's may be defined as a vector space over $\mathbf{R}$ as in Section 2.1. Similarly, the direct sum

$$\bigoplus_{j=1}^{n} V_j^* \tag{A.3.5}$$

of the corresponding dual spaces is defined as a vector space over $\mathbf{R}$. The dual

$$\Big(\bigoplus_{j=1}^{N} V_j\Big)^* \tag{A.3.6}$$

of $\bigoplus_{j=1}^{N} V_j$ is defined as a vector space over $\mathbf{R}$ as well. There is a natural isomorphism between (A.3.5) and (A.3.6), as vector spaces over the real numbers, as follows.

If λ_j is a linear functional on V_j for each $j = 1, \ldots, N$, then

$$v_1 \oplus \cdots \oplus v_V \mapsto \sum_{j=1}^{N} \lambda_j(v_j) \tag{A.3.7}$$

defines a linear functional on $\bigoplus_{j=1}^{N} V_j$. This defines a linear mapping from (A.3.5) into (A.3.6), which sends $\lambda_1 \oplus \cdots \oplus \lambda_N$ to the linear functional (A.3.7). One can check that this mapping is a bijection. More precisely, let λ be a linear functional on $\bigoplus_{j=1}^{N} V_j$, and let ι_l be the natural mapping from V_l into $\bigoplus_{j=1}^{N} V_j$ for each $l = 1, \ldots, N$. Thus

$$\lambda \circ \iota_l \tag{A.3.8}$$

is a linear functional on V_l for each $l = 1, \ldots, N$. This can be used to define a linear mapping from (A.3.6) into (A.3.5). It is easy to see that this is the inverse of the previous mapping.

A.4 Some direct sums and products

Let V_j be a vector space over the real numbers for each nonnegative integer j. Of course, one could restrict one's attention to positive integers here, or use any nonempty set of indices. Consider the Cartesian product

$$\prod_{j=0}^{\infty} V_j, \tag{A.4.1}$$

which is the set of infintie sequences $\{v_j\}_{j=0}^{\infty}$ with $v_j \in V_j$ for each $j \ge 0$. This is a vector space over the real numbers, with respect to coordinatewise addition and scalar multiplication. This is known as the *direct product* of the V_j's, $j \ge 0$.

The *direct sum* of the V_j's, $j \ge 0$, is defined by

$$\bigoplus_{j=0}^{\infty} V_j = \Big\{ \{v_j\}_{j=0}^{\infty} \in \prod_{j=0}^{\infty} V_j : v_j = 0 \text{ for all but finitely many } j \Big\}. \tag{A.4.2}$$

This is a linear subspace of the direct product. The direct sum of finitely many vector spaces over the real numbers, as in Section 2.1, may also be considered as a direct product.

If l is a nonnegative integer, then there is an obvious mapping ι_l from V_l into the direct sum (A.4.2). This mapping sends $w_l \in V_l$ to the sequence whose lth term is equal to w_l, and whose jth term in V_j is equal to 0 when $j \neq l$. This is a one-to-one linear mapping from V_l into (A.4.2) for each $l \geq 0$, and one may wish to identify V_l with its image in the direct sum for each l. The direct sum is the same as the linear subspace of the direct product generated by the images of the V_l's.

A.4.1 Sums, products, and dual spaces

Consider the direct product

$$\prod_{j=0}^{\infty} V_j^* \tag{A.4.3}$$

of the duals of the V_j's. Let us also consider the dual

$$\Big(\bigoplus_{j=0}^{\infty} V_j\Big)^* \tag{A.4.4}$$

of the V_j's. There is a natural isomorphism between (A.4.3) and (A.4.4), as vector spaces over the real numbers, which is analogous to the corresponding statement for finite direct sums in Subsection A.3.2. If λ_j is a linear functional on V_j for each j, $v_j \in V_j$ for each j, and $v_j = 0$ for all but finitely many j, then

$$\lambda_j(v_j) = 0 \text{ for all but finitely many } j. \tag{A.4.5}$$

This means that

$$\sum_{j=0}^{\infty} \lambda_j(v_j) \tag{A.4.6}$$

may be defined as a real number, which is the same as

$$\sum_{j=0}^{n} \lambda_j(v_j) \tag{A.4.7}$$

when n is sufficiently large. This defines a linear functional on $\bigoplus_{j=0}^{\infty} V_j$. Using this, we get a linear mapping from (A.4.3) into (A.4.4).

One can check that this mapping is a bijection. More precisely, if λ is a linear functional on $\bigoplus_{j=0}^{\infty} V_j$, then

$$\lambda \circ \iota_l \tag{A.4.8}$$

is a linear functional on V_l for each $l \geq 0$. This leads to a linear mapping from (A.4.4) into (A.4.3). This is the inverse of the mapping described in the preceding paragraph, as before.

A.5 Some remarks about linear mappings

If V and W are vector spaces over the real numbers, then let

$$\mathcal{L}(V, W) \tag{A.5.1}$$

be the space of all linear mappings from V into W. This is a linear subspace of the space of all W-valued functions on V, with respect to pointwise addition and scalar multiplication of functions. If V and W have finite dimension, then it is well known that the dimension of $\mathcal{L}(V, W)$ is given by

$$\dim \mathcal{L}(V, W) = (\dim V)\,(\dim W). \tag{A.5.2}$$

Indeed, one can use bases for V and W to describe linear mappings from V into W in terms of matrices to get (A.5.2).

Suppose for the moment that $W = \mathbf{R}$, considered as a one-dimensional vector space over itself. In this case, $\mathcal{L}(V, W) = \mathcal{L}(V, \mathbf{R})$ is the same as the dual space V^* of linear functionals on V, as in Section A.3.

A.5.1 Linear mappings and direct sums

Let $V_1, \ldots, V_N$ be finitely many vector spaces over the real numbers. If T_j is a linear mapping from V_j into W for each $j = 1, \ldots, N$, then one can get a linear mapping T from $\bigoplus_{j=1}^N V_j$ into W as in Subsection 2.1.1. Conversely, every linear mapping from $\bigoplus_{j=1}^N V_j$ into W corresponds to unique linear mappings from V_j into W, $1 \le j \le N$, in this way, as before. This leads to an isomorphism from

$$\bigoplus_{j=1}^N \mathcal{L}(V_j, W) \tag{A.5.3}$$

onto

$$\mathcal{L}\Big(\bigoplus_{j=1}^N V_j, W\Big), \tag{A.5.4}$$

as vector spaces over the real numbers.

Similarly, let $W_1, \ldots, W_M$ be finitely many vector spaces over the real numbers. If R_l is a linear mapping from V into W_l for each $l = 1, \ldots, M$, then

$$R(v) = R_1(v) \oplus \cdots \oplus R_M(v) \tag{A.5.5}$$

defines a linear mapping from V into $\bigoplus_{l=1}^M W_l$. Conversely, it is easy to see that every linear mapping R from V into $\bigoplus_{l=1}^M W_l$ corresponds to unique linear mappings from V into W_l, $1 \le l \le M$, in this way. This leads to an isomorphism from

$$\bigoplus_{l=1}^M \mathcal{L}(V, W_l) \tag{A.5.6}$$

onto

$$\mathcal{L}\Big(V, \bigoplus_{l=1}^{M} W_l\Big), \tag{A.5.7}$$

as vector spaces over the real numbers.

A.5.2 Linear mappings and direct products

Now let V_j be a vector space over the real numbers for each nonnegative integer j. If T_j is a linear mapping from V_j into W for each $j \geq 0$, then one can get a linear mapping T from $\bigoplus_{j=0}^{\infty} V_j$ into W in basically the same way as before. This defines a linear mapping from

$$\prod_{j=0}^{\infty} \mathcal{L}(V_j, W) \tag{A.5.8}$$

into

$$\mathcal{L}\Big(\bigoplus_{j=0}^{\infty} V_j, W\Big). \tag{A.5.9}$$

Conversely, if T is any linear mapping from $\bigoplus_{j=0}^{\infty} V_j$ into W, then one can get a linear mapping T_l from V_l into W for each $l \geq 0$, by composing the standard embedding of V_l into $\bigoplus_{j=0}^{\infty} V_j$ with T. This defines a linear mapping from (A.5.9) into (A.5.8), which is the inverse of the previous mapping.

Similarly, let W_l be a vector space over the real numbers for every nonnegative integer l. If R_l is a linear mapping from V into W_l for each $l \geq 0$, then we get a linear mapping R from V into $\prod_{l=0}^{\infty} W_l$, where the lth coordinate of $R(v)$ is equal to $R_l(v)$ for every $v \in V$ and $l \geq 0$. This defines a linear mapping from

$$\prod_{l=0}^{\infty} \mathcal{L}(V, W_l) \tag{A.5.10}$$

into

$$\mathcal{L}\Big(V, \prod_{l=0}^{\infty} W_l\Big). \tag{A.5.11}$$

Conversely, if R is any linear mapping from V into $\prod_{l=0}^{\infty} W_l$, then one can get a linear mapping R_k from V into W_k for each $k \geq 0$, by composing R with the kth standard coordinate projection from $\prod_{l=0}^{\infty} W_l$ onto W_k. This defines a linear mapping from (A.5.11) into (A.5.10), which is the inverse of the previous mapping.

A.5.3 Some more direct sums

Of course,

$$\bigoplus_{l=0}^{\infty} \mathcal{L}(V, W_l) \tag{A.5.12}$$

is a linear subspace of (A.5.10), as in the previous section. We also have that

$$\mathcal{L}\Big(V, \bigoplus_{l=0}^{\infty} W_l\Big) \tag{A.5.13}$$

is a linear subspace of (A.5.11), because $\bigoplus_{l=0}^{\infty} W_l$ is a linear subspace of $\prod_{l=0}^{\infty} W_l$. The restriction of the linear mapping from (A.5.10) into (A.5.11) defined in the preceding paragraph is a one-to-one linear mapping from (A.5.12) into (A.5.13).

Let R be a linear mapping from V into $\bigoplus_{l=0}^{\infty} W_l$, and let R_k be the composition of R with the kth standard coordinate projection from $\bigoplus_{l=0}^{\infty} W_l$ onto W_k for each $k \ge 0$. If $v \in V$, then

$$R_k(v) = 0 \text{ for all but finitely many } k \ge 0, \tag{A.5.14}$$

by the definition of $\bigoplus_{l=0}^{\infty} W_l$. If V has finite dimension, then one can check that

$$R_k = 0 \text{ for all but finitely many } k \ge 0. \tag{A.5.15}$$

This means that the linear mapping from (A.5.12) into (A.5.13) defined in the previous paragraph is surjective when V has finite dimension.

A.6 The algebra of multilinear forms

Let W be a vector space over the real numbers. Let us define $\mathcal{M}(W)$ initially as a vector space over the real numbers by

$$\mathcal{M}(W) = \bigoplus_{k=0}^{\infty} \mathcal{M}_k(W), \tag{A.6.1}$$

where the right side is as in Section A.4. We shall normally identify elements of $\mathcal{M}_l(W)$ with their images in $\mathcal{M}(W)$ under the embeddings defined before, so that every element of $\mathcal{M}(W)$ corresponds to a finite sum of elements of $\mathcal{M}_l(W)$, $l \ge 0$.

If $\mu, \nu \in \mathcal{M}(W)$, then we can define

$$\mu \otimes \nu \tag{A.6.2}$$

as an element of $\mathcal{M}(W)$ in an obvious way. More precisely, if $\mu \in \mathcal{M}_{l_1}(W)$ and $\nu \in \mathcal{M}_{l_2}(W)$ for some positive integers l_1, l_2, then (A.6.2) may be defined as an element of $\mathcal{M}_{l_1+l_2}(W)$ as in Section 2.8. The cases where l_1 or l_2 is equal to 0 may be included as in Section 2.12. Otherwise, μ and ν correspond to finite sums of multilinear l_1, l_2-linear forms on W, $l_1, l_2 \ge 0$, and (A.6.2) may be defined as the corresponding sum of finitely many $(l_1 + l_2)$-linear forms on W.

This defines a bilinear mapping from $\mathcal{M}(W) \times \mathcal{M}(W)$ into $\mathcal{M}(W)$. This makes $\mathcal{M}(W)$ into an associative algebra over the real numbers, because of the associativity property of products of multilinear forms on W mentioned in Subsection 2.8.1. Remember that we take $\mathcal{M}_0(W) = \mathbf{R}$, as in Section 2.12, which corresponds to a subalgebra of $\mathcal{M}(W)$. In fact, the real number 1 corresponds to the multiplicative identity element of $\mathcal{M}(W)$ in this way.

A.6.1 The definition of $\mathcal{SM}(W)$

Similarly, we can define $\mathcal{SM}(W)$ as a vector space over the real numbers by

$$\mathcal{SM}(W) = \bigoplus_{k=0}^{\infty} \mathcal{SM}_k(W), \tag{A.6.3}$$

which is a linear subspace of $\mathcal{M}(W)$. Using the symmetrization mapping S_k from $\mathcal{M}_k(W)$ onto $\mathcal{SM}_k(W)$ defined for each $k \geq 1$ in Subsection 2.4.1, we get a symmetrization mapping S from $\mathcal{M}(W)$ onto $\mathcal{SM}(W)$. More precisely, we can take S_k to be the identity mapping on $\mathcal{M}_0(W) = \mathcal{SM}_0(W) = \mathbf{R}$ when $k = 0$, as in Section 2.12.

If $\mu, \nu \in \mathcal{SM}(W)$, then we can define

$$\mu \odot \nu \tag{A.6.4}$$

in an obvious way, to get a bilinear mapping from $\mathcal{SM}(W) \times \mathcal{SM}(W)$ into $\mathcal{SM}(W)$. This uses the bilinear mapping from $\mathcal{SM}_{l_1}(W) \times \mathcal{SM}_{l_2}(W)$ into $\mathcal{SM}_{l_1+l_2}(W)$ defined in Section 2.9. This makes $\mathcal{SM}(W)$ into a commutative associative algebra over the real numbers, because of the analogous properties discussed in Section 2.9. Remember that $\mathcal{SM}_0(\mathbf{R}) = \mathbf{R}$, as in Section 2.12, which corresponds to a subalgebra of $\mathcal{SM}(W)$. The real number 1 corresponds to the multiplicative identity element of $\mathcal{SM}(W)$, as before.

If $\mu, \nu \in \mathcal{M}(W)$, then

$$S(\mu \otimes \nu) = S(S(\mu) \otimes S(\nu)) = S(\mu) \odot S(\nu). \tag{A.6.5}$$

The first step corresponds to a property of multilinear forms mentioned in Subsection 2.9.1, and the second step follows from the definition of the right side. This shows that S is an algebra homomorphism from $\mathcal{M}(W)$ onto $\mathcal{SM}(W)$.

A.6.2 Polynomials and $\mathcal{M}(\mathbf{R}^n)$

If $W = \mathbf{R}^n$ for some positive integer n and $\mu \in \mathcal{M}(\mathbf{R}^n)$, then we can define

$$P_\mu \tag{A.6.6}$$

as a polynomial on $\mathbf{R}^n$ with real coefficients, using the earlier definition for multilinear forms on $\mathbf{R}^n$. If $\nu \in \mathcal{M}(\mathbf{R}^n)$ too, then

$$P_{\mu \otimes \nu} = P_\mu \, P_\nu, \tag{A.6.7}$$

because of the analogous property of multilinear forms on $\mathbf{R}^n$ mentioned in Subsection 2.8.1. Thus $\mu \mapsto P_\mu$ defines an algebra homomorphism from $\mathcal{M}(\mathbf{R}^n)$ onto the space $\mathcal{P}(\mathbf{R}^n)$ of all polynomials on $\mathbf{R}^n$ with real coefficients, with respect to pointwise multiplication of polynomials on $\mathbf{R}^n$.

Similarly, the restriction of $\mu \to P_\mu$ to $\mathcal{SM}(\mathbf{R}^n)$ defines an algebra isomorphism onto $\mathcal{P}(\mathbf{R}^n)$, with respect to $\odot$ on $\mathcal{SM}(\mathbf{R}^n)$. This uses some of the remarks in Section 1.10.

A.7 More on $\mathcal{AM}(W)$

Let W be a vector space over the real numbers again. Let us define $\mathcal{AM}(W)$ initially as a vector space over the real numbers by

$$\mathcal{AM}(W) = \bigoplus_{k=0}^{\infty} \mathcal{AM}_k(W), \tag{A.7.1}$$

where the right side is as in Section A.4. If W has dimension n for some positive integer n, then $\mathcal{AM}_k(W) = \{0\}$ when $k > n$, and this is essentially the same as the definition of $\mathcal{AM}(W)$ in Subsection 2.12.1. We can use the alternatization mapping A_k from $\mathcal{M}_k(W)$ onto $\mathcal{AM}_k(W)$ defined for $k \geq 1$ in Subsection 2.4.1 to get an alternatization mapping A from $\mathcal{M}(W)$ onto $\mathcal{AM}(W)$. As in Section 2.12, we take A_0 to be the the identity mapping on $\mathcal{M}_0(W) = \mathcal{AM}_0(W) = \mathbf{R}$.

If $\mu, \nu \in \mathcal{AM}(W)$, then we can define

$$\mu \wedge_0 \nu \tag{A.7.2}$$

as an element of $\mathcal{AM}(W)$ in an obvious way, using the analogous definition for alternating multilinear forms on W in Subsection 2.10.1. This defines a bilinear mapping from $\mathcal{AM}(W) \times \mathcal{AM}(W)$ into $\mathcal{AM}(W)$. This makes $\mathcal{AM}(W)$ into an associative algebra over the real numbers, because of the associativity property of $\wedge_0$ mentioned in Subsection 2.10.1. As before, $\mathcal{AM}_0(W) = \mathbf{R}$ corresponds to a subalgebra of $\mathcal{AM}(W)$ with respect to $\wedge_0$, and the real number 1 corresponds to the multiplicative identity element in $\mathcal{AM}(W)$ with respect to $\wedge_0$.

If $\mu, \nu \in \mathcal{M}(W)$, then

$$A(\mu \otimes \nu) = A(A(\mu) \otimes A(\nu)) = A(\mu) \wedge_0 A(\nu), \tag{A.7.3}$$

where the first step corresponds to a property of multilinear forms mentioned in Section 2.10, and the second step follows from the definition of the right side. This means that A is an algebra homomorphism from $\mathcal{M}(W)$ onto $\mathcal{AM}(W)$, with respect to $\wedge_0$ on $\mathcal{AM}(W)$.

If $\mu, \nu \in \mathcal{AM}(W)$, then we can define

$$\mu \wedge \nu \tag{A.7.4}$$

as an element of $\mathcal{AM}(W)$ in an obvious way, using the analogous definition for alternating multilinear forms on Section 2.11. This defines a bilinear mapping from $\mathcal{AM}(W) \times \mathcal{AM}(W)$ into $\mathcal{AM}(W)$, which makes $\mathcal{AM}(W)$ into an associative algebra over the real numbers, because of the associativity property mentioned in Section 2.11. Note that $\mathcal{AM}_0(W) = \mathbf{R}$ also corresponds to a subalgebra of $\mathcal{AM}(W)$ with respect to $\wedge$, and that the real number 1 corresponds to the multiplicative identity element in $\mathcal{AM}(W)$ with respect to $\wedge$.

If $\mu \in \mathcal{M}_k(W)$ for some nonnegative integer k, then put

$$\widetilde{A}_k(\mu) = k!\, A_k(\mu). \tag{A.7.5}$$

Using these mappings, we get a linear mapping $\widetilde{A}$ from $\mathcal{M}(W)$ onto $\mathcal{AM}(W)$.

If μ, ν are k_1, k_2-linear forms on W for some nonnegative integers k_1, k_2, then $\mu \otimes \nu$ is a $(k_1 + k_2)$-linear form on W, and we have that

$$\begin{aligned} \widetilde{A}_{k_1+k_2}(\mu \otimes \nu) &= (k_1 + k_2)!\, A_{k_1+k_2}(\mu \otimes \nu) = (k_1 + k_2)!\, A_{k_1}(\mu) \wedge_0 A_{k_2}(\nu) \\ &= \frac{(k_1 + k_2)!}{k_1!\, k_2!}\, \widetilde{A}_{k_1}(\mu) \wedge_o \widetilde{A}_{k_2}(\nu) = \widetilde{A}_{k_1}(\mu) \wedge \widetilde{A}_{k_2}(\nu), \end{aligned} \tag{A.7.6}$$

using (A.7.3) in the second step, and the definition of $\wedge$ in the fourth step. This implies that

$$\widetilde{A}(\mu \otimes \nu) = \widetilde{A}(\mu) \wedge \widetilde{A}(\nu) \tag{A.7.7}$$

for all $\mu, \nu \in \mathcal{M}(W)$. Thus $\widetilde{A}$ is an algebra homomorphism from $\mathcal{M}(W)$ onto $\mathcal{AM}(W)$, with respect to $\wedge$ on $\mathcal{AM}(W)$.

A.8 More on inner products

Let W be a vector space over the real numbers with an inner product $\langle \cdot, \cdot \rangle_W$, and let $\| \cdot \|_W$ be the corresponding norm on W, as in Section 3.3. We say that $u, w \in W$ are *orthogonal* with respect to $\langle \cdot, \cdot \rangle_W$ if

$$\langle u, w \rangle_W = 0. \tag{A.8.1}$$

It is easy to see that this happens if and only if

$$\|u + w\|_W^2 = \|u\|_W^2 + \|w\|_W^2. \tag{A.8.2}$$

Let n be a positive integer, and let $w_1, \ldots, w_n$ be n pairwise-orthogonal vectors in W. We say that $w_1, \ldots, w_n \in W$ are *orthonormal* with respect to $\langle \cdot, \cdot \rangle_W$ if

$$\|w_j\|_W = 1 \tag{A.8.3}$$

for each $j = 1, \ldots, n$. It is well known and not difficult to check that this implies that $w_1, \ldots, w_n$ are linearly independent in W. Note that the standard basis vectors $e_1, \ldots, e_n$ in $\mathbf{R}^n$ are orthonormal with respect to the standard inner product $\langle \cdot, \cdot \rangle_{\mathbf{R}^n}$ on $\mathbf{R}^n$.

If W has finite dimension, then it is well known that W has an orthonormal basis. In fact, an orthonormal basis for W can be obtained from any basis for W using the Gram–Schmidt process.

Let V be a vector space over the real numbers with an inner product $\langle \cdot, \cdot \rangle_V$ and associated norm $\| \cdot \|_V$. Also let T be a linear mapping from V into W, and remember that T is an isometry with respect to these norms if and only if T preserves the inner products, as in Section 3.15. Suppose that V has dimension n, and let $v_1, \ldots, v_n$ be an orthonormal basis for V. One can check that T is an isometry if and only if

(A.8.4) $T(v_1), \ldots, T(v_n)$ are orthonormal in W with respect to $\langle \cdot, \cdot \rangle_W$.

Any sequence of n vectors in W may be realized as

$$L(e_1), \ldots, L(e_n), \tag{A.8.5}$$

where L is a linear mapping from $\mathbf{R}^n$ into W. These vectors are linearly independent in W exactly when L is one-to-one, and these vectors span W exactly when L maps $\mathbf{R}^n$ onto W. If W has dimension n, then we can use an orthonormal basis for W to get an isometric linear mapping from $\mathbf{R}^n$ onto W, with respect to the standard inner product and norm on $\mathbf{R}^n$.

A.9 Inner products and adjoints

Let V be a vector space over the real numbers with an inner product $\langle \cdot, \cdot \rangle_V$. If $v \in V$, then

$$\lambda_v(u) = \langle u, v \rangle_V \tag{A.9.1}$$

defines a linear functional on V. It is easy to see that

$$v \mapsto \lambda_v \tag{A.9.2}$$

defines a linear mapping from V into the dual space of linear functionals on V. Note that

$$\lambda_v(v) = \langle v, v \rangle_V \tag{A.9.3}$$

is equal to 0 if and only if $v = 0$. This implies that (A.9.2) is one-to-one.

If V has finite dimension, then it is well known that every linear functional on V may be expressed as in (A.9.1) for a unique $v \in V$. Equivalently, this means that (A.9.2) is a one-to-one linear mapping from V onto the dual space of V. This can be verified directly when $V = \mathbf{R}^n$ for some positive integer n, with the standard inner product. Otherwise, one can use an orthonormal basis for V to reduce to that case, or to essentially the same argument. Alternatively, one can use the fact that the dimension of the dual space of V is the same as the dimension of V when the dimension of V is finite, as in Section A.3.

Let W be another vector space over the real numbers with an inner product $\langle \cdot, \cdot \rangle_W$, and let T be a linear mapping from V into W. If $w \in W$, then

$$\mu_w(u) = \langle T(u), w \rangle_W \tag{A.9.4}$$

defines a linear functional on V. Suppose that V has finite dimension, so that there is a unique element $T^*(w)$ of V such that

$$\langle T(u), w \rangle_W = \langle u, T^*(w) \rangle_V \tag{A.9.5}$$

for every $u \in V$, as in the preceding paragraph. One can check that T^* is a linear mapping from W into V. This is called the *adjoint* of T with respect to the inner products on V and W.

This is not quite the same as the dual linear mapping associated to T as in Subsection A.3.1, although we are using the same notation here. Sometimes one may use other notation for dual linear spaces and the corresponding dual linear mappings, such as V', W', and T', to distinguish the dual linear mapping from the adjoint. If W also has finite dimension, then one can characterize linear functionals on W in terms of the inner product in the same way, so that the dual linear mapping corresponds more closely to the adjoint.

If W has finite dimension too, then one can use orthonormal bases for V and W to characterize T in terms of a matrix. In this case, T^* corresponds to the transpose of this matrix, with respect to the same orthonormal bases for V and W.

A.9.1 Some properties of the adjoint

Let Z be a third vector space over the real numbers with an inner product $\langle \cdot, \cdot \rangle_Z$, and let R be a linear mapping from W into Z. If V and W have finite dimension, then the adjoints R^*, $(R \circ T)^*$ of R, $R \circ T$ may be defined as linear mappings from Z into W, V, respectively, as before. In particular,

$$\langle R(w), z \rangle_Z = \langle w, R^*(z) \rangle_W \tag{A.9.6}$$

for every $w \in W$ and $z \in Z$. If $u \in V$ and $z \in Z$, then

$$\langle R(T(u)), z \rangle_Z = \langle T(u), R^*(z) \rangle_W = \langle u, T^*(R^*(z)) \rangle_V. \tag{A.9.7}$$

This implies that

$$(R \circ T)^* = T^* \circ R^*, \tag{A.9.8}$$

as linear mappings from Z into V.

Remember that T is an isometric linear mapping from V into W, with respect to the norms associated to the inner products, if and only if

$$\langle T(u), T(v) \rangle_W = \langle u, v \rangle_V \tag{A.9.9}$$

for every $u, v \in V$, as in Section 3.15. If V has finite dimension, then this is the same as saying that

$$\langle u, T^*(T(v)) \rangle_V = \langle u, v \rangle_V \tag{A.9.10}$$

for every $u, v \in V$. One can check that this holds if and only if

$$T^* \circ T = I_V, \tag{A.9.11}$$

where I_V is the identity mapping on V.

If V has finite dimension and T is a one-to-one linear mapping from V onto itself, then

$$T^* \circ (T^{-1})^* = (T^{-1} \circ T)^* = I_V{*} = I_V, \tag{A.9.12}$$

using (A.9.8) in the first step. Similarly,

$$(T^{-1})^* \circ T^* = (T \circ T^{-1})^* = I_W^* = I_W, \tag{A.9.13}$$

where I_W is the identity mapping on W. This implies that T^* is a one-to-one linear mapping from W onto V, with

$$(T^*)^{-1} = (T^{-1})^*, \tag{A.9.14}$$

as linear mappings from V into W, using (A.9.8). Note that T is an isometric linear mapping from V onto W if and only if

$$T^{-1} = T^*, \tag{A.9.15}$$

because of (A.9.11). This is related to a remark in Section 4.12.

A.10 Nonnegativity and self-adjointness

Let n be a positive integer, and let V be an n-dimensional vector space over the real numbers with an inner product $\langle \cdot, \cdot \rangle_V$ and its associated norm $\| \cdot \|_V$. Of course, this includes $\mathbf{R}^n$ with its standard inner product and norm, and one can often reduce to this case using an isometric linear mapping from $\mathbf{R}^n$ onto V. This basically corresponds to choosing an orthonormal basis for V, as before.

A linear mapping B from V into itself is said to be *self-adjoint* if

$$B^* = B. \tag{A.10.1}$$

This means that the matrix associated to B with respect to an orthonormal basis of V is *symmetric*, in the sense that it is equal to its transpose.

Under these conditions, it is well known that

(A.10.2) there is an orthonormal basis for V consisting of eigenvectors for B.

This is often stated for real symmetric matrices, or self-adjoint linear mappings from $\mathbf{R}^n$ into itself, and one can reduce to that case.

A self-adjoint linear mapping B from V into itself is said to be *nonnegative* if

$$\langle B(v), v \rangle_V \geq 0 \tag{A.10.3}$$

for every $v \in V$. One can check that this implies that the eigenvalues of B are nonnegative real numbers. Conversely, if the eigenvalues of B are nonnegative real numbers, then one can use diagonalization of B mentioned in the preceding paragraph to get that B is nonnegative in this sense. We also get that

$$\det B \geq 0 \tag{A.10.4}$$

when B is nonnegative, because the determinant of B is equal to the product of its eigenvalues.

A.10.1 A family of examples

Let W be another vector space over the real numbers with an inner product $\langle\cdot,\cdot\rangle_W$ and associated norm $\|\cdot\|_W$, and let T be a linear mapping from V into W. Let us check that

$$(T^*)^* = T, \tag{A.10.5}$$

as linear mappings from V into W. If $u \in V$ and $w \in W$, then

$$\langle T(u), w\rangle_W = \langle u, T^*(w)\rangle_V = \langle (T^*)^*(u), w\rangle_W, \tag{A.10.6}$$

using the definition of T^* in the first step, and the definition of $(T^*)^*$ in the second step. This implies (A.10.5).

Note that $T^* \circ T$ is a linear mapping from V into itself. This mapping is self-adjoint, because

$$(T^* \circ T)^* = T^* \circ (T^*)^* = T^* \circ T. \tag{A.10.7}$$

If $v \in V$, then

$$\langle (T^* \circ T)(v), v\rangle_V = \langle T^*(T(v)), v\rangle_V = \langle T(v), T(v)\rangle_W = \|T(v)\|_W^2. \tag{A.10.8}$$

Thus $T^* \circ T$ is nonnegative as a self-adjoint linear mapping from V into itself. This implies that

$$\det(T^* \circ T) \geq 0, \tag{A.10.9}$$

as before.

Of course, if $V = W$, then T and T^* are linear mappings from V into itself, and

$$\det(T^* \circ T) = (\det T^*)\,(\det T). \tag{A.10.10}$$

If $\langle\cdot,\cdot\rangle_V = \langle\cdot,\cdot\rangle_W$ too, then

$$\det T^* = \det T. \tag{A.10.11}$$

This is because the matrix associated to T^* with respect to an orthonormal basis for V is the same as the transpose of the matrix associated to T with respect to that basis. This implies that

$$\det(T^* \circ T) = (\det T)^2 \tag{A.10.12}$$

in this case. This is another way to get (A.10.9) under these conditions.

A.11 Adjoints and volumes

Let n be a positive integer, and let W be an n-dimensional vector space over the real numbers with an inner product $\langle\cdot,\cdot\rangle_W$ and associated norm $\|\cdot\|_W$. We can use an orthonormal basis for W to get an isometric linear mapping L from $\mathbf{R}^n$ onto W, as in Section A.8, with respect to the standard inner product $\langle\cdot,\cdot\rangle_{\mathbf{R}^n}$ and norm $\|\cdot\|_{\mathbf{R}^n}$ on $\mathbf{R}^n$. If E is a reasonably nice subset of W, then put

$$\Lambda_W(E) = \mathrm{Vol}_n(L^{-1}(E)), \tag{A.11.1}$$

as in Section 4.15. Remember that this depends only on the inner product on W, and not the particular choice of L.

Let T be a linear mapping from $\mathbf{R}^n$ into W. If E_0 is a reasonably nice subset of $\mathbf{R}^n$, then

$$\Lambda_W(T(E_0)) = \mathrm{Vol}_n(L^{-1}(T(E_0))) = |\det(L^{-1} \circ T)| \, \mathrm{Vol}_n(E_0), \tag{A.11.2}$$

where the second step is as in Section 4.12. In particular, if $e_1, \ldots, e_n$ is the standard basis in $\mathbf{R}^n$, then

$$P_W(T(e_1), \ldots, T(e_n)) = T(P_{\mathbf{R}^n}(e_1, \ldots, e_n)) = T([0,1]^n), \tag{A.11.3}$$

as in Section 4.14. If we take $E_0 = [0,1]^n$ in (A.11.2), then we get that

$$\Lambda_W(P_W(T(e_1), \ldots, T(e_n))) = |\det(L^{-1} \circ T)|. \tag{A.11.4}$$

Observe that

$$(L^{-1} \circ T)^* = T^* \circ (L^{-1})^* = T^* \circ L, \tag{A.11.5}$$

where the second step uses the fact that L is an isometry from $\mathbf{R}^n$ onto W. This means that

$$(L^{-1} \circ T)^* \circ (L^{-1} \circ T) = T^* \circ L \circ L^{-1} \circ T = T^* \circ T. \tag{A.11.6}$$

It follows that

$$\det(T^* \circ T) = \det((L^{-1} \circ T)^* \circ (L^{-1} \circ T)) = (\det(L^{-1} \circ T))^2, \tag{A.11.7}$$

where the second step is as in (A.10.12). This shows that

$$|\det(L^{-1} \circ T)| = (\det(T^* \circ T))^{1/2}. \tag{A.11.8}$$

If E_0 is a reasonably nice subset of $\mathbf{R}^n$, then we get that

$$\Lambda_W(T(E_0)) = (\det(T^* \circ T))^{1/2} \, \mathrm{Vol}_n(E_0), \tag{A.11.9}$$

by (A.11.2). In particular,

$$\Lambda_W(P_W(T(e_1), \ldots, T(e_n))) = (\det(T^* \circ T))^{1/2}, \tag{A.11.10}$$

by (A.11.4). This is very close to Theorem 7 on p328 of [20], and we shall return to that in a moment.

A.11.1 Another inner product space

Let Z be another vector space over the real numbers with an inner product $\langle \cdot, \cdot \rangle_Z$. Suppose that W is a linear subspace of Z, and that $\langle \cdot, \cdot \rangle_W$ is the same as the restriction of $\langle \cdot, \cdot \rangle_Z$ to W. Let T_1 be the same as T, considered as a linear mapping from $\mathbf{R}^n$ into Z. The adjoint T_1^* of T_1 may be defined as before, as a linear mapping from Z into $\mathbf{R}^n$. Let us check that

$$T_1^* \circ T_1 = T^* \circ T, \tag{A.11.11}$$

as linear mappings from $\mathbf{R}^n$ into itself.

If $u, v \in \mathbf{R}^n$, then

$$\begin{aligned} \langle (T^* \circ T)(u), v\rangle_{\mathbf{R}^n} &= \langle T^*(T(u)), v\rangle_{\mathbf{R}^n} = \langle T(u), T(v)\rangle_W \\ &= \langle T_1(u), T_1(v)\rangle_Z = \langle T_1^*(T_1(u)), v\rangle_{\mathbf{R}^n} \\ &= \langle (T_1^* \circ T_1)(u), v\rangle_{\mathbf{R}^n}. \end{aligned} \tag{A.11.12}$$

This implies (A.11.11). Of course, it follows that

$$\det(T_1^* \circ T_1) = \det(T^* \circ T). \tag{A.11.13}$$

In Theorem 7 on p328 of [20], one basically takes Z to be another Euclidean space, with the standard inner product. It is also stated in terms of the matrix associated to T_1.

More precisely, the volume of a parallelepiped in a Euclidean space is defined on p328 of [20], and Theorem 7 there expresses the square of the volume as a determinant. This is used to determine the effect of a linear mapping from $\mathbf{R}^n$ into itself on the n-dimensional volumes of parallelepipeds or other subsets of $\mathbf{R}^n$.

If W is an n-dimensional linear subspace of a Euclidean space, then the restriction of the standard inner product on the Euclidean space to W defines an inner product on W. Remember that an admissible volume on W is determined by its value at a parallelepiped corresponding to a basis for W, as in Subsection 4.14.1. If the n-dimensional volume of a parallelepiped in a Euclidean space has already been defined, then one can use that to determine Λ_W.

A.12 Traces and transposes of matrices

Let m and n be positive integers, and let $a = (a_{j,l})_{j,l=1}^{m,n}$ be an $m \times n$ matrix of real numbers. The *transpose* of a is defined to be the $n \times m$ matrix $a^t = (a^t_{l,j})_{l,j=1}^{n,m}$ of real numbers with

$$a^t_{l,j} = a_{j,l} \tag{A.12.1}$$

for each $j = 1, \ldots, m$ and $l = 1, \ldots, n$. If $m = n$, then the *trace* of a is defined as usual by

$$\operatorname{tr} a = \sum_{j=1}^{n} a_{j,j}. \tag{A.12.2}$$

This defines a linear functional on the space $M_{n,n}(\mathbf{R})$ of $n \times n$ matrices with entries in $\mathbf{R}$. Clearly

$$\operatorname{tr} a^t = \operatorname{tr} a \tag{A.12.3}$$

in this case.

Let r be another positive integer, and let b be an $n \times r$ matrix of real numbers. If a is an $m \times n$ matrix of real numbers again, then $a\,b$ is defined as an $m \times r$ matrix of real numbers, by matrix multiplication. It is well known and not difficult to check that

$$(a\,b)^t = b^t\,a^t, \tag{A.12.4}$$

as $r \times m$ matrices of real numbers.

Suppose that $m = r$, so that $a\,b$ is an $m \times m$ matrix of real numbers, and $b\,a$ is an $n \times n$ matrix of real numbers. Under these conditions, it is well known and not difficult to verify that

$$\operatorname{tr}(a\,b) = \operatorname{tr}(b\,a) = \sum_{j=1}^{m} \sum_{l=1}^{n} a_{j,l}\,b_{l,j}. \tag{A.12.5}$$

Let I_n be the *identity matrix* in $M_{n,n}(\mathbf{R})$, whose diagonal entries are equal to 1, and whose other entries are equal to 0. If $a \in M_{n,n}(\mathbf{R})$ and $t \in \mathbf{R}$, then put

$$p_a(t) = \det(I_n + t\,a), \tag{A.12.6}$$

which is a polynomial in t with real coefficients. It is well known and not too difficult to check that

$$p_a'(0) = \operatorname{tr} a. \tag{A.12.7}$$

A.12.1 Traces of linear mappings

Let V be a vector space over the real numbers of dimension n, and let $v_1, \ldots, v_n$ be a basis for V. If T is a linear mapping from V into itself, then T corresponds to an $n \times n$ matrix $c(T)$ of real numbers with respect to this basis for V in a standard way. The *trace* of T may be defined as a real number by

$$\operatorname{tr} T = \operatorname{tr} c(T). \tag{A.12.8}$$

It is well known that this does not depend on the basis for V, and we shall return to this in a moment.

Let W be a vector space over the real numbers of dimension m, and let $w_1, \ldots, w_m$ be a basis for W. One can use this basis to define the trace of a linear mapping from W into itself, as in the preceding paragraph. Let A be a linear mapping from V into W, and let B be a linear mapping from W into V. Using the bases for V and W, we can associate $m \times n$ and $n \times m$ matrices a and b of real numbers to A and B, respectively.

Of course, $A \circ B$ is a linear mapping from W into itself, and $B \circ A$ is a linear mapping from V into itself. These linear mappings correspond to the product matrices $a\,b$ and $b\,a$, respectively, with respect to the given bases on V and W. It follows that

$$\operatorname{tr}(A \circ B) = \operatorname{tr}(B \circ A), \tag{A.12.9}$$

because of (A.12.5).

In particular, this holds with $V = W$, and $v_j = w_j$. If R is a one-to-one linear mapping from V onto itself, then it follows that

$$\operatorname{tr}(R \circ T \circ R^{-1}) = \operatorname{tr} T. \tag{A.12.10}$$

One can use this to get that (A.12.8) does not depend on the particular basis for V, by standard arguments.

Appendix B

Lie algebras

B.1 The definition and some examples

Let A be a vector space over the real numbers, and let $[\cdot, \cdot] = [\cdot, \cdot]_A$ be a bilinear mapping from $A \times A$ into A. Thus, if $x, y \in A$, then $[x, y] = [x, y]_A$ is an element of A too, and this is linear in each of x and y. We say that A is a *Lie algebra* over the real numbers with respect to $[\cdot, \cdot]$ if the following two additional conditions hold, as in Definition 3.4 on p84 of [184]. The first condition is that

(B.1.1) $$[x, y] = -[y, x]$$

for every $x, y \in A$. The second condition is that

(B.1.2) $$[[x, y], z] + [[y, z], x] + [[z, x], y] = 0$$

for every $x, y, z \in A$, which is known as the *Jacobi identity*.

Sometimes the Jacobi identity is formulated as saying that

(B.1.3) $$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$$

for every $x, y, z \in A$. It is easy to see that this is equivalent to (B.1.2) when the first condition (B.1.1) holds.

Of course, (B.1.1) implies that

(B.1.4) $$[x, x] = 0$$

for every $x \in A$. Conversely, if (B.1.4) holds for every $x \in A$, then one can check that (B.1.1) holds for every $x, y \in A$, as in Section 1.7. The definition of a Lie algebra is often formulated using (B.1.4) instead of (B.1.1), as on p1 of [86], and Definition 1 on p2 of [160]. These formulations are used more broadly, where (B.1.4) implies (B.1.1), but the converse may not hold.

If A is a Lie algebra with respect to $[\cdot, \cdot]$, then $[\cdot, \cdot]$ may be called the *Lie bracket* on A. Of course, a Lie algebra over $\mathbf{R}$ is a particular type of algebra over $\mathbf{R}$ in the strict sense, as in Section A.2. In particular, the notions of *homomorphisms* and *isomorphisms* between Lie algebras may be defined in the same way as before.

B.1.1 Some examples of Lie algebras

Suppose for the moment that A is an associative algebra over the real numbers, where multiplication of $x, y \in A$ is expressed as $x\,y$. In this case, one can verify that A is a Lie algebra with respect to

$$[x, y] = x\,y - y\,x. \tag{B.1.5}$$

This corresponds to Example (iii) on p2 of [160].

Let V be a vector space over the real numbers, and let

$$\mathcal{L}(V) = \mathcal{L}(V, V) \tag{B.1.6}$$

be the space of linear mappings from V into itself. This is a vector space over the real numbers, as in Section A.5. In fact, $\mathcal{L}(V)$ is an associative algebra over $\mathbf{R}$, with composition of linear mappings as multiplication.

If $R, T \in \mathcal{L}(V)$, then put

$$[R, T] = [R, T]_{\mathcal{L}(V)} = R \circ T - T \circ R, \tag{B.1.7}$$

which is another element of $\mathcal{L}(V)$. As before, $\mathcal{L}(V)$ is a Lie algebra over $\mathbf{R}$ with respect to (B.1.7). This is known as the *general linear algebra* associated to V, and denoted

$$gl(V), \tag{B.1.8}$$

as on p2 of [86]. Of course, this is the same as $\mathcal{L}(V)$ as a vector space over $\mathbf{R}$.

Let n be a positive integer, and let $M_{n,n}(\mathbf{R})$ be the space of $n \times n$ matrices with entries in $\mathbf{R}$, as in Sections 5.1 and 5.2. This is a vector space over the real numbers with respect to entrywise addition and scalar multiplication, and an associative algebra over $\mathbf{R}$ with respect to matrix multiplication, as before. The *general linear algebra*

$$gl(n, \mathbf{R}) \tag{B.1.9}$$

of $n \times n$ matrices with entries in $\mathbf{R}$ is the same as $M_{n,n}(\mathbf{R})$ as a vector space over the real numbers, considered as a Lie algebra with respect to the corresponding commutator bracket (B.1.5). This is example 3.5 (c) on p84 of [184].

There is a standard one-to-one correspondence between linear mappings from $\mathbf{R}^n$ into itself, and $n \times n$ matrices of real numbers, using the standard basis for $\mathbf{R}^n$. This defines an isomorphism between $\mathcal{L}(\mathbf{R}^n)$ and $M_{n,n}(\mathbf{R})$, as vector spaces over $\mathbf{R}$, and associative algebras over $\mathbf{R}$. This also defines an isomorphism between $gl(\mathbf{R}^n)$ and $gl(n, \mathbf{R})$, as Lie algebras over $\mathbf{R}$.

If A is any vector space over the real numbers, then A is a Lie algebra with respect to the Lie bracket defined by putting

$$[x, y] = 0 \tag{B.1.10}$$

for every $x, y \in A$. In this case, we may say that A is *abelian* or *commutative* as a Lie algebra, as in Section 1.4 on p4 of [86], Example (ii) on p2 of [160], and example 3.5 (b) on p84 of [184].

If $(A, [\cdot, \cdot])$ is any Lie algebra over $\mathbf{R}$, then A is commutative as an algebra in the strict sense if and only if

$$[x, y] = [y, x] \tag{B.1.11}$$

for every $x, y \in A$, as in Section A.2. Of course, (B.1.1) and (B.1.11) imply (B.1.10). However, this does not always work when considering vector spaces over other fields, or modules over other commutative rings. In those cases, a Lie algebra is considered to be abelian or commutative if the corresponding Lie bracket is identically 0, and not merely symmetric. Thus a Lie algebra over other fields or commutative rings could be commutative as an algebra in the strict sense, and not commutative as a Lie algebra.

B.2 Subalgebras of Lie algebras

Let $(A, [\cdot, \cdot])$ be a Lie algebra over the real numbers. If A_0 is a linear subspace of A such that

$$[x, y] \in A_0 \tag{B.2.1}$$

for every $x, y \in A_0$, then A_0 is said to be a *Lie subalgebra* of A. This is the same as saying that A_0 is a subalgebra of A as an algebra in the strict sense over $\mathbf{R}$, as in Section A.2. It is easy to see that A_0 is also a Lie algebra with respect to the restriction of $[x, y]$ to $x, y \in A_0$ in this case.

Suppose for the moment that A is an associative algebra over $\mathbf{R}$, with multiplication of $x, y \in A$ expressed as $x\,y$. If A_0 is a subalgebra of A, then A_0 is also a subalgebra of A as a Lie algebra with respect to the corresponding commutator bracket (B.1.5).

If n is a positive integer, then the *special linear algebra*

$$sl(n, \mathbf{R}) \tag{B.2.2}$$

consists of the $n \times n$ matrices a with entries in $\mathbf{R}$ such that

$$\operatorname{tr} a = 0. \tag{B.2.3}$$

It is easy to see that this is a Lie subalgebra of $gl(n, \mathbf{R})$, using (A.12.5). More precisely, if a, b are any $n \times n$ matrices of real numbers, then

$$a\,b - b\,a \in sl(n, \mathbf{R}). \tag{B.2.4}$$

Similarly, if V is a finite-dimensional vector space over the real numbers, then the *special linear algebra*

$$sl(V) \tag{B.2.5}$$

associated to V consists of the linear mappings T from V into itself such that

$$\operatorname{tr} T = 0. \tag{B.2.6}$$

If R and T are any linear mappings from V into itself, then

$$R \circ T - T \circ R \in sl(V), \tag{B.2.7}$$

because of (A.12.9). In particular, $sl(V)$ is a Lie subalgebra of $gl(V)$, as on p2 of [86]. The isomorphism between $gl(\mathbf{R}^n)$ and $gl(n, \mathbf{R})$, as Lie algebras over the real numbers, mentioned in Subsection B.1.1 also maps $sl(\mathbf{R}^n)$ onto $sl(n, \mathbf{R})$.

The *orthogonal algebra*

$$o(n, \mathbf{R}) \tag{B.2.8}$$

consists of the $n \times n$ matrices a of real numbers that are *anti-symmetric*, in the sense that

$$a^t = -a. \tag{B.2.9}$$

One can check that this is a Lie subalgebra of $gl(n, \mathbf{R})$, using (A.12.4). More precisely, this is a Lie subalgebra of $sl(n, \mathbf{R})$.

Let V be a finite-dimensional vector space over the real numbers with an inner product $\langle \cdot, \cdot \rangle_V$. If T is a linear mapping from V into itself, then the adjoint T^* of T with respect to $\langle \cdot, \cdot \rangle_V$ may be defined as a linear mapping from V into itself as in Section A.9. Under these conditions, we say that T is *anti-self-adjoint* if

$$T^* = -T. \tag{B.2.10}$$

The *orthogonal algebra*

$$o(V) \tag{B.2.11}$$

associated to $\langle \cdot, \cdot \rangle_V$ on V consists of the anti-self-adjoint linear mappings from V into itself. One can check that this is a Lie subalgebra of $gl(V)$, using (A.9.8).

If $V = \mathbf{R}^n$ with the standard inner product, then $o(\mathbf{R}^n)$ corresponds to $o(n, \mathbf{R})$ under the Lie algebra isomorphism between $gl(\mathbf{R}^n)$ and $gl(n, \mathbf{R})$ mentioned in Subsection B.1.1. This follows from the fact that the adjoint of a linear mapping from $\mathbf{R}^n$ into itself with respect to the standard inner product corresponds to the transpose of the associated $n \times n$ matrix, as in Section A.9.

Let V be a vector space over the real numbers, and let b be a bilinear form on V. Let us say that a linear mapping T from V into itself is *anti-symmetric* with respect to b if

$$b(T(v), w) = -b(v, T(w)) \tag{B.2.12}$$

for every $v, w \in V$. One can check that the collection of these linear mappings is a Lie subalgebra of $gl(V)$. This corresponds to part of Theorem 3.56 on p119 of [184].

B.3 Some linear mappings

Let V_0 be a vector space over the real numbers. If R and T are linear mappings from V_0 into itself, then let $[R, T]$ put

$$[R, T] = R \circ T - T \circ R, \tag{B.3.1}$$

as in (B.1.7). If U is another linear mapping from V_0 into itself, then

$$[[R,T],U] + [[T,U],R] + [[U,R],T] = 0, \tag{B.3.2}$$

by the Jacobi identity, as in Section B.1.

Let V_1, V_2, V_3 be linear subspaces of V, with

$$V_3 \subseteq V_2 \subseteq V_1. \tag{B.3.3}$$

Consider the space

$$\mathcal{L}_{3,2}(V_1, V_0) \tag{B.3.4}$$

of linear mappings T from V_1 into V_0 such that

$$T(V_3) \subseteq V_2,\ T(V_2) \subseteq V_1. \tag{B.3.5}$$

This is a linear subspace of the space $\mathcal{L}(V_1, V_0)$ of all linear mappings from V_1 into V_0.

Similarly, let

$$\mathcal{L}_3(V_2, V_0) \tag{B.3.6}$$

be the space of linear mappings B from V_2 into V_0 such that

$$B(V_3) \subseteq V_1. \tag{B.3.7}$$

This is a linear subspace of $\mathcal{L}(V_2, V_0)$. If $R, T \in \mathcal{L}_{3,2}(V_1, 0)$, then

$$R \circ T \in \mathcal{L}_3(V_2, V_0). \tag{B.3.8}$$

In this case, the commutator (B.3.1) of R and T is defined as a linear mapping from V_2 into V_0, and in fact

$$[R,T] \in \mathcal{L}_3(V_2, V_0). \tag{B.3.9}$$

If $R, T, U \in \mathcal{L}_{3,2}(V_1, V_0)$, then

$$R \circ T \circ U \in \mathcal{L}(V_3, V_0). \tag{B.3.10}$$

Similarly,

$$([R,T]) \circ U,\ U \circ ([R,T]) \in \mathcal{L}(V_3, V_0), \tag{B.3.11}$$

so that

$$[[R,T],U] = ([R,T]) \circ U - U \circ ([R,T]) \in \mathcal{L}(V_3, V_0). \tag{B.3.12}$$

One can check that (B.3.2) holds in $\mathcal{L}(V_3, V_0)$ under these conditions.

B.4 Some more Lie subalgebras

Let (A, b) be an algebra in the strict sense over the real numbers, as in Section A.2. Thus A is a vector space over the real numbers in particular, so that the space $\mathcal{L}(A)$ of all linear mappings from A into itself is a vector space over $\mathbf{R}$, and an associative algebra over $\mathbf{R}$ with respect to composition of linear mappings. Similarly $gl(A)$ is a Lie algebra over $\mathbf{R}$ with respect to the usual commutator of linear mappings, as in Section B.2.

Let

$$\mathrm{Der}(A) = \mathrm{Der}(A, b) \tag{B.4.1}$$

be the space of derivations on A with respect to b, as in Subsection A.2.1. It is easy to see that this is a linear subspace of $\mathcal{L}(A)$. In fact, one can check that

$$\mathrm{Der}(A) \text{ is a subalgebra of } gl(A), \tag{B.4.2}$$

as a Lie algebra over $\mathbf{R}$. This corresponds to a remark on p4 of [86], to Example (iv) on p2 of [160], and to part of Theorem 3.54 on p117 of [184].

Suppose for the moment that A is an associative algebra, where multiplication of x, y in A is expressed as $x\,y$. If $a \in A$, then put

$$\delta_a(x) = [a, x] = a\,x - x\,a \tag{B.4.3}$$

for each $x \in A$. One can check that

$$\delta_a \in \mathrm{Der}(A). \tag{B.4.4}$$

More precisely, δ_a is called an *inner derivation* of A. Note that

$$a \mapsto \delta_a \tag{B.4.5}$$

defines a linear mapping from A into $\mathrm{Der}(A)$.

If $b \in A$ too, then one can verify that

$$\delta_a \circ \delta_b - \delta_b \circ \delta_a = \delta_{[a,b]}. \tag{B.4.6}$$

This implies that

$$\{\delta_a : a \in A\} \tag{B.4.7}$$

is a Lie subalgebra of $\mathrm{Der}(A)$.

Suppose now that $(A, [\cdot,\cdot]_A)$ is a Lie algebra over the real numbers. In particular, A is an algebra over $\mathbf{R}$ in the strict sense, so that the previous notions and remarks about algebras in the strict sense may be used in this case. Thus a linear mapping δ from A into itself is a *derivation* of A as a Lie algebra if

$$\delta([x, y]_A) = [\delta(x), y]_A + [x, \delta(y)]_A \tag{B.4.8}$$

for every $x, y \in A$, as in Subsection A.2.1.

If $x \in A$, then let $\mathrm{ad}\, x = \mathrm{ad}_x$ be the linear mapping from A into itself defined by

$$\mathrm{ad}_x(y) = [x, y]_A \tag{B.4.9}$$

for every $y \in A$. One can check that

$$\mathrm{ad}_x \in \mathrm{Der}(A), \tag{B.4.10}$$

using the Jacobi identity. This corresponds to a remark on p4 of [86], and the first part of Theorem 3 on p3 of [160]. In fact, ad_x is called an *inner derivation* of A, as on p4 of [86]. It is sometimes convenient to use the notation $\mathrm{ad}_A x = \mathrm{ad}_{A,x}$ for ad_x, to indicate the role of A.

Clearly

$$x \mapsto \mathrm{ad}_x \tag{B.4.11}$$

is a linear mapping from A into $\mathrm{Der}(A)$. One can check that

$$\mathrm{ad}_x \circ \mathrm{ad}_y - \mathrm{ad}_y \circ \mathrm{ad}_x = \mathrm{ad}_{[x,y]_A} \tag{B.4.12}$$

for every $x, y \in A$, using the Jacobi identity. This corresponds to a remark on p8 of [86], and to the second part of Theorem 3 on p3 of [160]. It follows that

$$\{\mathrm{ad}_x : x \in A\} \tag{B.4.13}$$

is a Lie subalgebra of $\mathrm{Der}(A)$.

B.5 Some multiplication operators

Let (A, b) be an algebra over the real numbers in the strict sense, as in Section A.2. If $a \in A$, then put

$$L_a(x) = b(a, x) \tag{B.5.1}$$

and

$$R_a(x) = b(x, a) \tag{B.5.2}$$

for every $x \in A$. These define linear mappings from A into itself, which are the *left and right multiplication operators* on A associated to a, respectively. Note that

$$a \mapsto L_a \tag{B.5.3}$$

and

$$a \mapsto R_a \tag{B.5.4}$$

define linear mappings from A into the space $\mathcal{L}(A)$ of all linear mappings from A into itself.

Suppose for the moment that A has a multipicative identity element e with respect to b. In this case,

$$L_a(e) = R_a(e) = a \tag{B.5.5}$$

for every $a \in A$. In particular, this means that (B.5.3) and (B.5.4) are one-to-one mappings from A into $\mathcal{L}(A)$.

Let δ be a derivation on A with respect to b. It is easy to see that

$$\delta \circ L_a - L_a \circ \delta = L_{\delta(a)} \tag{B.5.6}$$

for every $a \in A$. Similarly,

(B.5.7) $$\delta \circ R_a - R_a \circ \delta = R_{\delta(a)}$$

for every $a \in A$.

Of course, if A is commutative with respect to b, then L_a and R_a are the same. If $(A, [\cdot,\cdot]_A)$ is a Lie algebra over $\mathbf{R}$, then

(B.5.8) $$L_a = \mathrm{ad}_a$$

and

(B.5.9) $$R_a = -\mathrm{ad}_a$$

for every $a \in A$.

B.5.1 Multiplication operators on associative algebras

Suppose now that A is an associative algebra over the real numbers, where multiplication of $x, y \in A$ is expressed as $x\,y$. Thus

(B.5.10) $$L_a(x) = a\,x$$

and

(B.5.11) $$R_a(x) = a\,x$$

for all $a, x \in A$. If $a_1, a_2 \in A$, then

(B.5.12) $$L_{a_1\,a_2}(x) = (a_1\,a_2)\,x = a_1\,(a_2\,x) = L_{a_1}(L_{a_2}(x))$$

for every $x \in A$. This means that

(B.5.13) $$L_{a_1\,a_2} = L_{a_1} \circ L_{a_2}$$

on A.

Similarly,

(B.5.14) $$R_{a_1\,a_2}(x) = x\,(a_1\,a_2) = (x\,a_1)\,a_2 = R_{a_2}(R_{a_1}(x))$$

for every $x \in A$. This is the same as saying that

(B.5.15) $$R_{a_1\,a_2} = R_{a_2} \circ R_{a_1}$$

on A. Observe that

(B.5.16) $$L_{a_1}(R_{a_2}(x)) = a_1\,(x\,a_2) = (a_1\,x)\,a_2 = R_{a_2}(L_{a_1}(x))$$

for every $x \in A$. This shows that

(B.5.17) $$L_{a_1} \circ R_{a_2} = R_{a_2}\,L_{a_1}$$

on A.

If $a \in A$, then put

(B.5.18) $$\delta_a = L_a - R_a,$$

as in (B.4.3). If δ is a derivation on A, then

(B.5.19) $$\delta \circ \delta_a - \delta_a \circ \delta = \delta_{\delta(a)},$$

because of (B.5.6) and (B.5.7).

Suppose that A is commutative as well. If δ is a derivation on A, then it is easy to see that

(B.5.20) $$L_a \circ \delta \in \mathrm{Der}(A)$$

for every $a \in A$.

B.6 Ideals in algebras over R

Let (A, b) be an algebra over the real numbers in the strict sense again, as in Section A.2. A linear subspace A_0 of A is said to be a *left ideal* in A with respect to b if

(B.6.1) $$b(x, y) \in A_0$$

for every $x \in A$ and $y \in A_0$. Similarly, A_0 is said to be a *right ideal* in A if (B.6.1) holds for every $x \in A_0$ and $y \in A$. If A_0 is both a left and right ideal in A_0, then A_0 is said to be a *two-sided ideal* in A. Of course, if A is commutative with respect to b, then left and right ideals in A are the same.

If A_0 is a left or right ideal in A, then A_0 is a subalgebra of A. It is easy to see that the kernel of a homomorphism from A into another algebra in the strict sense over $\mathbf{R}$ is a two-sided ideal in A.

Let A_0 be a linear subspace of A. One can define the quotient space A/A_0 as a vector space over the real numbers in a standard way. This includes a quotient mapping q, which is a linear mapping from A onto A/A_0 with

(B.6.2) $$\ker q = A_0.$$

If A_0 is a two-sided ideal in A, then it is not too difficult to show that there is a unique bilinear mapping b_0 from $(A/A_0) \times (A/A_0)$ into A/A_0 such that

(B.6.3) $$q(b(x, y)) = b_0(q(x), q(y))$$

for every $x, y \in A$. This means that A/A_0 is an algebra in the strict sense over the real numbers with respect to b_0, and that q is an algebra homomorphism from A onto A/A_0. This corresponds to a remark on p2 of [160].

If A is commutative or associative, then it is easy to see that A/A_0 has the same property. If A has a multiplicative identity element e with respect to b, then $q(e)$ is the multiplicative identity element in A/A_0 with respect to b_0.

B.6.1 Ideals in Lie algebras

Let $(A, [\cdot,\cdot]_A)$ be a Lie algebra over the real numbers. A linear subspace A_0 of A is said to be an *ideal* in A if

$$[x, y]_A \in A_0 \tag{B.6.4}$$

for every $x \in A$ and $y \in A_0$, as on p6 of [86]. This is the same as being a left, right, or two-sided ideal in A, because of the antisymmetry of the Lie bracket. It is easy to see that A/A_0 is also a Lie algebra in this case, as on p7 of [86].

If n is a positive integer, then $sl(n, \mathbf{R})$ is an ideal in $gl(n, \mathbf{R})$. Similarly, if V is a finite-dimensional vector space over $\mathbf{R}$, then $sl(V)$ is an ideal in $gl(V)$.

If B is an associative algebra over $\mathbf{R}$, then the set $\mathrm{Der}(B)$ of derivations on B is a Lie subalgebra of $gl(B)$, as in Section B.4. We have also seen that the set of derivations on B defined by commutators with elements of B is a Lie subalgebra of $\mathrm{Der}(B)$. In fact, this is an ideal in $\mathrm{Der}(B)$, as a Lie algebra over $\mathbf{R}$, as in (B.5.19).

Similarly, the set $\mathrm{Der}(A)$ of derivations on A as a Lie algebra is a Lie subalgebra of $gl(A)$, and the set (B.4.13) of derivations on A of the form ad_x, $x \in A$, is a Lie subalgebra of $\mathrm{Der}(A)$. More precisely, (B.4.13) is an ideal in $\mathrm{Der}(A)$, because of (B.5.6). This corresponds to Exercise 1 on p9 of [86].

The *center* of an associative algebra B is defined to be the set of elements of B that commute witl all other elements of B. It is easy to see that this is a subalgebra of B. The *center* of A as a Lie algebra is defined by

$$Z(A) = \{x \in A : [x, y]_A = 0 \text{ for all } y \in A\}. \tag{B.6.5}$$

This is an ideal in A, as mentioned on p6 of [86].

B.7 Representations

Let $(A, [\cdot,\cdot]_A)$ be a Lie algebra over the real numbers, and let V be a vector space over $\mathbf{R}$. A *representation* of A on V is defined to be a homomorphism from A into the Lie algebra $gl(V)$ of linear mappings from V into itself. If ρ is such a representation, then we may use ρ_x for the linear mapping on V corresponding to $x \in A$. In order for ρ to be a representation of A as a Lie algebra on V, we should have that

$$x \mapsto \rho_x \tag{B.7.1}$$

is linear as a mapping from A into the space of linear mappings on V, and that

$$\rho_{[x,y]_A} = \rho_x \circ \rho_y - \rho_y \circ \rho_x, \tag{B.7.2}$$

as linear mappings from V into itself, for every $x, y \in A$.

Under these conditions, we may say that V is a *module* over A, as a Lie algebra over $\mathbf{R}$. We may also put

$$x \cdot v = \rho_x(v) \tag{B.7.3}$$

for each $x \in A$ and $v \in V$, which is another element of V. Using this notation, the condition that (B.7.1) be a linear mapping from A into the space of linear mappings on V is equivalent to asking that

$$(x, v) \mapsto x \cdot v \tag{B.7.4}$$

be a bilinear mapping from $A \times V$ into V. Similarly, (B.7.2) is the same as saying that

$$([x, y]_A) \cdot v = x \cdot (y \cdot v) - y \cdot (x \cdot v) \tag{B.7.5}$$

for every $x, y \in A$ and $v \in V$.

This corresponds to some remarks on p8, 25 of [86], p31 of [160], and p90 of [184].

B.7.1 Representations of associative algebras

Suppose now that A is an associative algebra over the real numbers, where multiplication of $x, y \in A$ is expressed as $x\,y$. A *representation* of A on a vector space V over the real numbers is defined to be a homomorphism from A into the algebra $\mathcal{L}(V)$ of linear mappings from V into itself. If ρ is such a representation, then we may use ρ_x for the linear mapping on V corresponding to $x \in A$, as before. In order for ρ to be a representation of A as an associative algebra on V, we should have that $x \mapsto \rho_x$ is linear as a mapping from A into $\mathcal{L}(V)$, and that

$$\rho_{x\,y} = \rho_x \circ \rho_y, \tag{B.7.6}$$

as linear mappings from V into itself, for every $x, y \in A$.

Remember that A may also be considered as a Lie algebra over $\mathbf{R}$ with respect to the corresponding commutator bracket, as in Subsection B.1.1. If ρ is a representation of A as an associative algebra on V, then ρ may also be considered as a representation of A as a Lie algebra on V.

If ρ is a representation of A, as an associative algebra over $\mathbf{R}$, on V, then we may say that V is a *left module* over A. We may use the notation $x \cdot v$ for $\rho_x(v)$ when $x \in A$ and $v \in V$, as before. With this notation, the linearity of $x \mapsto \rho_x$ as a mapping from A into $\mathcal{L}(V)$ is equivalent to the bilinearity of $x \cdot v$ as a mapping from $A \times V$ into V, as before. We also have that (B.7.6) is equivalent to the condition that

$$(x\,y) \cdot v = x \cdot (y \cdot v) \tag{B.7.7}$$

for every $x, y \in A$ and $v \in V$. If A has a multiplicative identity element e, then one may ask that

$$e \cdot v = v \tag{B.7.8}$$

for every $v \in V$ too, so that ρ_e is the identity mapping on V.

We may consider A as a left module over itself, using multiplication in the algebra. This corresponds to the mapping that sends $a \in A$ to the left multiplication operator L_a on A, as in Subsection B.5.1.

B.7.2 Right modules over associative algebras

Let V be a vector space over the real numbers again, and suppose that $v \cdot x$ is defined as an element of V for every $x \in A$ and $v \in V$. More precisely, we ask that

$$(v, x) \mapsto v \cdot x \tag{B.7.9}$$

be a bilinear mapping from $V \times A$ into V. If

$$(v \cdot x) \cdot y = v \cdot (x\, y) \tag{B.7.10}$$

for all $x, y \in A$ and $v \in V$, then we say that V is a *right module* over A. If A has a multiplicative identity element e, then one may ask that

$$v \cdot e = v \tag{B.7.11}$$

for every $v \in V$ as well.

It is easy to see that A may be considered as a right module over itself, using multiplication in the algebra.

Suppose that V is a right module over A, and put

$$x \cdot v = -v \cdot x \tag{B.7.12}$$

for every $x \in X$ and $v \in V$. One can check that V is a module over A, as a Lie algebra with respect to the corresponding commutator bracket, in this way.

B.8 More on representations

Let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over the real numbers, and let V be a vector space over the real numbers. The *trivial representation* sends every element of A to the linear mapping equal to 0 on V. In particular, $\mathbf{R}$ may be considered as a module over A with respect to the trivial representation. We shall normally only be concerned with trivial representations of Lie algebras on $\mathbf{R}$, unless otherwise indicated, as on p31 of [160].

As in Section B.4, we put $\mathrm{ad}_x(y) = [x, y]_A$ for every $x, y \in A$. This defines a representation of A on itself, as before. This is called the *adjoint representation* of A. Note that the kernel of

$$x \mapsto \mathrm{ad}_x \tag{B.8.1}$$

is the center $Z(A)$ of A as a Lie algebra, defined in Subsection B.6.1, as mentioned on p8 of [86].

B.8.1 Subrepresentations

Let ρ be any representation of A on V, and let W be a linear subspace of V. If

$$\rho_x(W) \subseteq W \tag{B.8.2}$$

for every $x \in A$, then the restriction of ρ_x to W defines a representation of A on W. This is called a *subrepresentation* of ρ. Equivalently, W is said to be a *submodule* of V, as a module over A. An ideal in A is the same as a submodule of A as a module over itself with respect to the adjoint representation.

Now let A be an associative algebra over $\mathbf{R}$, where multiplication of $x, y \in A$ is expressed as $x\,y$, and let ρ be a representation of A on V, as in Subsection B.7.1. If W is a linear subspace of V that satisfies (B.8.2) for every $x \in A$, then the restriction of ρ_x to W defines a representation of A on W, called a *subrepresentation* of ρ again. We may also call W a *submodule* of V, as a left module over A. A left ideal in A is the same as a submodule of A as a left module over itself, using multiplication on A, as in Subsection B.7.1.

Suppose that V is a right module over A, as in Subsection B.7.2. If

$$w \cdot x \in W \tag{B.8.3}$$

for every $w \in W$ and $x \in A$, then W is said to be a *submodule* of V, as a right module over A. A right ideal in A is the same as a submodule of A as a right module over itself, using multiplication on A, as in Subsection B.7.2.

B.8.2 Direct sums of representations

Let $V_1, \ldots, V_N$ be finitely many vector spaces over the real numbers, so that their direct sum $\bigoplus_{j=1}^N V_j$ may be defined as a vector space over the real numbers as in Section 2.1. If $(A, [\cdot,\cdot]_A)$ is a Lie algebra over $\mathbf{R}$ and V_j is a module over A for each j, then $\bigoplus_{j=1}^N V_j$ may be considered as a module over A in a natural way too, so that

$$V_l \text{ corresponds to a submodule of } \bigoplus_{j=1}^N V_j \tag{B.8.4}$$

for each l. Similarly, if A is an associative algebra over $\mathbf{R}$ and V_j is a left module over A for each j, then $\bigoplus_{j=1}^N V_j$ is a left module over A in a natural way. If V_j is a right module over A for each j, then $\bigoplus_{j=1}^N V_j$ is a right module over A in a natural way.

If V_j is a vector space over $\mathbf{R}$ for each nonnegative integer j, then the direct sum $\bigoplus_{j=0}^\infty V_j$ and direct product $\prod_{j=0}^\infty V_j$ of the V_j's may be defined as vector space over $\mathbf{R}$ as in Section A.4. If A is a Lie algebra again and A_j is a module over A for each j, then it is easy to see that $\prod_{j=0}^\infty V_j$ is a module over A too, where the action of $x \in A$ on an element of $\prod_{j=0}^\infty V_j$ is defined coordinatewise. In this case,

$$\bigoplus_{j=0}^\infty V_j \text{ is a submodule of } \prod_{j=0}^\infty V_j, \tag{B.8.5}$$

as a module over A. There are analogous statements for left and right modules over associative algebras, as before.

B.8.3 Representations and dual spaces

Let V be a vector space over the real numbers, and let V^* be the dual space of linear functionals on V, as in Section A.3. If T is a linear mapping from V into itself, then $T^*(\lambda) = \lambda \circ T$ defines a linear mapping from V^* into itself, as in Subsection A.3.1. If R is another linear mapping from V into itself, then we have seen that

$$(R \circ T)^* = T^* \circ R^*. \tag{B.8.6}$$

This implies that

$$(R \circ T - T \circ R)^* = -(R^* \circ T^* - T^* \circ R^*). \tag{B.8.7}$$

Note that $T \mapsto T^*$ defines a linear mapping from $\mathcal{L}(V)$ into itself.

Let A be an associative algebra over $\mathbf{R}$ again, and suppose for the moment that V is a left module over A. If $x \in A$ and $\lambda \in V^*$, then put

$$(\lambda \cdot x)(v) = \lambda(x \cdot v) \tag{B.8.8}$$

for every $v \in V$. This defines $\lambda \cdot x$ as a linear functional on V, and one can check that V^* is a right module over A in this way, using the remarks in the preceding paragraph. Equivalently, if T_x is the linear mapping from V into itself defined by

$$T_x(v) = x \cdot v, \tag{B.8.9}$$

then

$$\lambda \cdot x = T_x^*(\lambda) \tag{B.8.10}$$

for every $x \in A$ and $\lambda \in V^*$.

Similarly, suppose that V is a right module over A. If $x \in A$ and $v \in V$, then put

$$(x \cdot \lambda)(v) = \lambda(v \cdot x) \tag{B.8.11}$$

for every $v \in V$, which defines a linear functional on V. One can check that V^* is a left module over A in this way.

Let $(A, [\cdot,\cdot]_A)$ be a Lie algebra over $\mathbf{R}$, and let $\rho = \rho^V$ be a representation of A on V. If $x \in A$, then ρ_x is a linear mapping from V into itself, so that the dual linear mapping $(\rho_x)^*$ is defined on V^* as before. One can check that

$$\rho_x^{V^*} = -(\rho_x)^* \tag{B.8.12}$$

defines a representation of A on V. Equivalently, if $x \in A$ acts on V as in (B.7.3), and $\lambda \in V^*$, then put

$$(x \cdot \lambda)(v) = -\lambda(x \cdot v) \tag{B.8.13}$$

for each $v \in V$. One can verify that V^* is a module over A in this way, as before. This is the *dual representation* or *module* of A associated to V, as on p26 of [86]. This is related to some remarks on p31 of [160], which will be discussed further in the next section.

B.8.4 Commuting representations of Lie algebras

Let $(A, [\cdot,\cdot]_A)$ be a Lie algebra over $\mathbf{R}$ again, and let $\rho^1, \ldots, \rho^n$ be finitely many representations of A on V. Suppose that these representations commute with each other, in the sense that

$$\rho^j_x \circ \rho^l_y = \rho^l_y \circ \rho^j_x \tag{B.8.14}$$

on V for all $x, y \in A$ and $j, l = 1, \ldots, n$. Under these conditions, one can check that

$$\rho_x = \sum_{j=1}^n \rho^j_x \tag{B.8.15}$$

defines a representation of A on V.

B.9 Representations and multilinear mappings

Let $(A, [\cdot,\cdot]_A)$ be a Lie algebra over the real numbers, let k be a positive integer, and let $V_1, \ldots, V_k$ and Z be vector spaces over $\mathbf{R}$. Suppose that

$$\rho^{V_l} \tag{B.9.1}$$

is a representation of A on V_l for each $l = 1, \ldots, k$, and that

$$\rho^Z \tag{B.9.2}$$

is a representation of A on Z. It will be convenient to express the actions of $x \in A$ on $V_1, \ldots, V_k$ and Z as in (B.7.3).

Remember that

$$\mathcal{L}(V_1, \ldots, V_k; Z) \tag{B.9.3}$$

denotes the space of multilinear mappings from $\prod_{l=1}^k V_l$ into Z, as in Section 9.2. We would like to define representations

$$\rho^1, \ldots, \rho^k \text{ and } \rho^0 \tag{B.9.4}$$

on (B.9.3) corresponding to $\rho^{V_1}, \ldots, \rho^{V_k}$ and ρ^Z, respectively, as follows. If μ is a multilinear mapping from $\prod_{l=1}^k V_l$ into Z and $x \in A$, then let $\rho^0_x(\mu)$ be the multilinear mapping from $\prod_{l=1}^k V_l$ into Z defined by

$$(\rho^0_x(\mu))(v_1, \ldots, v_k) = x \cdot \mu(v_1, \ldots, v_k) \tag{B.9.5}$$

for each $v_1, \ldots, v_k$ in $V_1, \ldots, V_k$, respectively. Equivalently,

$$\rho^0_x(\mu) = \rho^Z_x \circ \mu \tag{B.9.6}$$

on $\prod_{l=1}^k V_l$. It is easy to see that this defines a representation ρ^0 of A on (B.9.3).

Similarly, for each $l = 1, \ldots, k$, let $\rho_x^l(\mu)$ be the multilinear mapping from $\prod_{l=1}^k V_l$ into Z defined by

(B.9.7)
$$\begin{aligned}(\rho_x^l(\mu))(v_1, \ldots, v_k) &= -\mu(v_1, \ldots, v_{l-1}, x \cdot v_l, v_{l+1}, \ldots, v_k)\\ &= -\mu(v_1, \ldots, v_{l-1}, \rho_x^{V_l}(v_l), v_{l+1}, \ldots, v_k)\end{aligned}$$

for every $v_1, \ldots, v_k$ in $V_1, \ldots, V_k$, respectively. One can check that this defines a representation ρ^l of A on (B.9.3) as well. This is related to the remarks in Subsection B.8.3.

It is easy to see that these representations on (B.9.3) commute with each other, as in (B.8.14). This implies that

(B.9.8)
$$\rho = \sum_{l=0}^k \rho^l$$

defines a representation of A on (B.9.3) too, as in (B.8.15). This corresponds to some remarks on p27 of [86], and on p31 of [160].

This is a bit simpler when $k = 1$, so that (B.9.3) is the same as the space

(B.9.9)
$$\mathcal{L}(V_1, Z)$$

of linear mappings from V_1 into Z. In this case,

(B.9.10)
$$\rho_x^1(\mu) = -\mu \circ \rho_x^{V_1}$$

for every $x \in A$ and linear mapping μ from V_1 into Z.

B.9.1 Representations on $\mathcal{M}_k(V, Z)$

Let V be a vector space over the real numbers, and suppose that

(B.9.11)
$$V_l = V$$

for each $l = 1, \ldots, k$. This means that (B.9.3) is the same as the space

(B.9.12)
$$\mathcal{M}_k(V, Z)$$

of multilinear mappings from V^k into Z, as in Section 9.2. Also let ρ^V be a representation of A on V, and suppose that

(B.9.13)
$$\rho^{V_l} = \rho^V$$

for each $l = 1, \ldots, k$.

Let μ be a multilinear mapping from V^k into Z, and let σ be a permutation on $\{1, \ldots, k\}$. Remember that μ^σ may be defined as another element of $\mathcal{M}_k(V, Z)$ as in Section 9.2. It is easy to see that

(B.9.14)
$$\rho_x^0(\mu^\sigma) = (\rho_x^0(\mu))^\sigma$$

for every $x \in A$. If $x \in A$ and $1 \le l \le k$, then

$$\begin{aligned}
& (\rho_x^l(\mu^\sigma))(v_1, \ldots, v_k) \\
\text{(B.9.15)} \quad = \; & -\mu^\sigma(v_1, \ldots, v_{l-1}, x \cdot v_l, v_{l+1}, \ldots, v_k) \\
= \; & -\mu(v_{\sigma(1)}, \ldots, v_{\sigma(\sigma^{-1}(l)-1)}, x \cdot v_l, v_{\sigma(\sigma^{-1}(l)+1)}, \ldots, v_{\sigma(k)}) \\
= \; & (\rho_x^{\sigma^{-1}(l)}(\mu))(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) = (\rho_x^{\sigma^{-1}(l)}(\mu))^\sigma(v_1, \ldots, v_k)
\end{aligned}$$

for every $v_1, \ldots, v_k \in V$. Thus

(B.9.16) $$\rho_x^l(\mu^\sigma) = (\rho_x^{\sigma^{-1}(l)}(\mu))^\sigma.$$

One can use (B.9.14), (B.9.16), and the definition (B.9.8) of ρ to get that

(B.9.17) $$\rho_x(\mu^\sigma) = (\rho_x(\mu))^\sigma.$$

If μ is symmetric or antisymmetric on V^k, as in Section 9.2, then it follows that $\rho_x(\mu)$ has the same property. This means that

(B.9.18) $$\rho_x(\mathcal{SM}_k(V, Z)) \subseteq \mathcal{SM}_k(V, Z)$$

and

(B.9.19) $$\rho_x(\mathcal{AM}_k(V, Z)) \subseteq \mathcal{SM}_k(V, Z),$$

where $\mathcal{AM}_k(V, Z)$ and $\mathcal{SM}_k(V, Z)$ are as in Section 9.2. Thus $\mathcal{SM}_k(V, Z)$ and $\mathcal{AM}_k(V, Z)$ are submodules of $\mathcal{M}_k(V, Z)$, as a module over A with respect to ρ, as in Subsection B.8.1.

B.10 Homomorphisms between representations

Let $(A, [\cdot, \cdot]_A)$ be a Lie algebra over the real numbers, let V, W be vector spaces over the real numbers, and let ρ^V, ρ^W be representations of A on V, W, respectively. A linear mapping ϕ from V into W is said to define a *homomorphism* between these representations if

(B.10.1) $$\phi \circ \rho_x^V = \rho_x^W \circ \phi$$

for every $x \in A$. Equivalently, this means that

(B.10.2) $$\phi(x \cdot v) = x \cdot \phi(v)$$

for every $x \in A$ and $v \in V$, where the actions of x on v and $\phi(v)$ are defined using ρ_x^V and ρ_x^W, as in Section B.7. We may also say that ϕ is a homomorphism from V into W as modules over A under these conditions, as on p25 of [86], and on p31 of [160].

Let W_0 be a linear subspace of W that is a submodule of W, as a module over A, as in Subsection B.8.1. In this case, the obvious inclusion mapping from W_0 into W is a module homomorphism.

If ϕ is a homomorphism from V into W, as modules over A, then it is easy to see that

$$\ker \phi \text{ is a submodule of } V, \tag{B.10.3}$$

as a module over A. Similarly,

$$\phi(V) \text{ is a submodule of } W, \tag{B.10.4}$$

as a module over A.

Let Z be another module over A, and let ψ be a homomorphism from W into Z, as modules over A. One can check that

$$\psi \circ \phi \text{ is a homomorphism from } V \text{ into } Z, \tag{B.10.5}$$

as modules over A.

If ϕ is a homomorphism from V into W, as modules over A, and ϕ is a one-to-one mapping from V onto W, then

$$\phi^{-1} \text{ is a homomorphism from } W \text{ onto } V, \tag{B.10.6}$$

as modules over A. Under these conditions, we say that ϕ is an *isomorphism* from V onto W, as modules over A, or equivalently that ϕ is an isomorphism between these representations of A. Of course, this means that ϕ^{-1} is an isomorphism from W onto V, as modules over A. If ψ is also an isomorphism from W onto Z, as modules over A, then

$$\psi \circ \phi \text{ is an isomorphism from } V \text{ onto } Z, \tag{B.10.7}$$

as modules over A.

Let V_0 be a linear subspave of V that is a submodule of V, as a module over A. The quotient space V/V_0 can be defined as a vector space over the real numbers in the usual way, with the corresponding quotient mapping q from V onto V/V_0 with kernel V_0. If $x \in A$ and $v \in V$, then we would like to put

$$x \cdot q(v) = q(x \cdot v). \tag{B.10.8}$$

It is easy to see that the right side depends only on x and $q(v)$, because V_0 is a submodule of V.

One can check that V/V_0 is a module over A with respect to (B.10.8). This may be described as a *quotient module* or *quotient representation.* Of course, q is a homomorphism from V onto V/V_0, as modules over A, by construction.

B.10.1 Modules over associative algebras

Now let A be an associative algebra over the real numbers, where multiplication of $x, y \in A$ is expressed as $x\,y$, let V, W be vector spaces over the real numbers, and let ϕ be a linear mapping from V into W. If V and W are left nodules over A, and ϕ intertwines the corresponding representations of A, then we say that ϕ is a *homomorphism* between these representations, or equivalent a homomorphism from V into W, as left modules over A. Similarly, if V and W are right modules over A, and if

$$\phi(v \cdot x) = \phi(v) \cdot x \tag{B.10.9}$$

for every $x \in A$ and $v \in V$, then we say that ϕ is a *homomorphism* from V into W, as right modules over A.

B.11 Invariant elements

Let A be an associative algebra over the real numbers, where multiplication of $x, y \in A$ is expressed as $x\,y$, and let V be a vector space over $\mathbf{R}$. If $v \in V$ and V is a left module over A, then it is easy to see that

$$\{x \in A : x \cdot v = 0\} \tag{B.11.1}$$

is a left ideal in A. Similarly, if V is a right module over A, then

$$\{x \in A : v \cdot x = 0\} \tag{B.11.2}$$

is a right ideal in A.

Suppose now that $(A, [\cdot, \cdot]_A)$ is a Lie algebra over $\mathbf{R}$, and that V is a module over A. If $v \in V$, then one can check that (B.11.1) is a subalgebra of A.

An element v of V is said to be *invariant* under the representation of A on V if

$$x \cdot v = 0 \tag{B.11.3}$$

for every $x \in A$, as on p31 of [160]. The collection of invariant elements of V is a linear subspace of V, and a submodule of V, as a module over A.

B.11.1 Invariant linear mappings

Let V and W be vector spaces over the real numbers that are also modules over A. The space $\mathcal{L}(V, W)$ of linear mappings from V into W may be considered as a module over A as well, as in Section B.9. More precisely, if $T \in \mathcal{L}(V, W)$ and $x \in A$, then

$$x \cdot T \in \mathcal{L}(V, W) \tag{B.11.4}$$

is defined by putting

$$(x \cdot T)(v) = x \cdot (T(v)) - T(x \cdot v) \tag{B.11.5}$$

for every $v \in V$, as before. It is easy to see that T is an invariant element of $\mathcal{L}(V, W)$ with respect to this representation of A if and only if T is a homomorphism from V into W, as modules over A, as in Example 1 on p31 of [160].

B.11.2 Invariant bilinear forms

Let V and W be modules over A again, and consider the space

$$\mathcal{L}(V, W; \mathbf{R}) \tag{B.11.6}$$

of all bilinear mappings from $V \times W$ into $\mathbf{R}$. This may be considered as a module over A too, as in Section B.9, using the trivial representation of A on $\mathbf{R}$. If $\beta \in \mathcal{L}(V, W; \mathbf{R})$ and $x \in A$, then

$$x \cdot \beta \in \mathcal{L}(V, W; \mathbf{R}) \tag{B.11.7}$$

is defined by

$$(x \cdot \beta)(v, w) = -\beta(x \cdot v, w) - \beta(v, x \cdot w) \tag{B.11.8}$$

for every $v \in V$ and $w \in W$. It follows that β is invariant with respect to this representation of A if and only if

$$\beta(x \cdot v, w) = -\beta(v, x \cdot w) \tag{B.11.9}$$

for every $x \in A$, $v \in V$, and $w \in W$, as in Example 2 on p31 of [160]. If $V = W = A$, considered as a module over itself with respect to the adjoint representation, then this corresponds to the condition that a bilinear form on A be *associative*, as on p21 of [86].

Appendix C

Complex numbers and complex analysis

C.1 Complex numbers

A *complex number* z can be expressed in a unique way as

$$z = x + y\, i, \tag{C.1.1}$$

with $x, y \in \mathbf{R}$, and where $i^2 = -1$. We may call x and y the *real* and *imaginary* of Z, respectively, and these may be denoted $\operatorname{Re} z$ and $\operatorname{Im} z$, respectively. The set of complex numbers is denoted $\mathbf{C}$, and addition and multiplication on $\mathbf{R}$ can be extended to $\mathbf{C}$ in a standard way. Note that addition and multiplication on $\mathbf{C}$ satisfy the usual commutativity, associativity, and distributivity properties.

If $z \in \mathbf{C}$ is as in (C.1.1), then the *complex conjugate* of z is defined by

$$\overline{z} = x - y\, i. \tag{C.1.2}$$

If w is another complex number, then one can check that

$$\overline{z + w} = \overline{z} + \overline{w} \tag{C.1.3}$$

and

$$\overline{z\, w} = \overline{z}\, \overline{w}. \tag{C.1.4}$$

The *absolute value* or *modulus* is defined by

$$|z| = (x^2 + y^2)^{1/2}, \tag{C.1.5}$$

using the nonnegative square root on the right side. If $z \in \mathbf{R}$, then this is the same as the absolute value of z as a real number.

Observe that

$$z\, \overline{z} = |z|^2 \tag{C.1.6}$$

for every $z \in \mathbf{C}$. One can use this and (C.1.4) to get that

$$|z\,w| = |z|\,|w| \tag{C.1.7}$$

for every $z, w \in \mathbf{C}$. If $z \neq 0$, then $|z| > 0$, one can check that

$$\frac{1}{z} = \frac{\overline{z}}{|z|^2} \tag{C.1.8}$$

is the multiplicative inverse of z in $\mathbf{C}$. This means that $\mathbf{C}$ is a *field*, although we shall not discuss this in detail here.

If $z \in \mathbf{C}$ is as in (C.1.1), then we may identify z with $(x, y) \in \mathbf{R}^2$. Note that addition of complex numbers corresponds to addition on $\mathbf{R}^2$, and that (C.1.5) is the same as the standard Euclidean norm of (x, y) in $\mathbf{R}^2$. Similarly, we may identify the complex plane $\mathbf{C}$ with $\mathbf{R}^2$ as a metric space, with respect to the standard Euclidean metric.

C.1.1 Complex vector spaces

One can define the notion of a *vector space over the complex numbers* in essentially the same way as for vector spaces over the real numbers, using complex numbers as the scalars instead of real numbers, although we shall not discuss this in detail here. If n is a positive integer, then the space $\mathbf{C}^n$ of n-tuples of complex numbers is a vector space over $\mathbf{C}$, with respect to coordinatewise addition and scalar multiplication. Similarly, if X is a nonempty set, then the space of all complex-valued functions on X is a vector space over $\mathbf{C}$, with respect to pointwise addition and scalar multiplication.

Note that any vector space over the complex numbers may be considered as a vector space over the real numbers, by simply not using multiplication by i. In particular, $\mathbf{C}$ may be considered as a two-dimensional vector space over $\mathbf{R}$, which may be identified with $\mathbf{R}^2$, as before.

If V and W are vector spaces over the complex numbers, then the notion of a *linear mapping* from V into W, as vector spaces over the complex numbers, may be defined in the usual way. We may also consider linear mappings from V into W as vector spaces over $\mathbf{R}$. In order to be precise, we may refer to a mapping T from V into W as being *real-linear* if T is linear as a mapping from V into W when they are considered as vector spaces over the real numbers, and we may say that T is *complex-linear* if T is linear as a mapping from V into W when they are considered as vector spaces over the complex numbers. Thus T is complex-linear if and only if T is real-linear and

$$T(i\,v) = i\,T(v) \tag{C.1.9}$$

for every $v \in V$. A complex-linear mapping from $\mathbf{C}$ into itself corresponds to multiplication by a complex number, for instance, and a real-linear mapping from $\mathbf{C}$ into itself is really the same as a linear mapping from $\mathbf{R}^2$ into itself.

C.2 Complex derivatives

Let U be a nonempty open subset of the complex plane, and let f be a complex-valued function on U. The *complex derivative* of f at a point $z \in U$ is defined by

$$f'(z) = \lim_{w \to z} \frac{f(w) - f(z)}{w - z}, \tag{C.2.1}$$

when the limit on the right exists. If the complex derivative of f exists at every $z \in U$, then f is said to be *complex analytic* or *holomorphic* on U.

A complex-valued function on U corresponds to a mapping from U into $\mathbf{R}^2$, using the usual identification between $\mathbf{C}$ and $\mathbf{R}^2$. If the complex derivative of f at $z \in U$ exists, then one can check that f is differentiable at z as a mapping from U into $\mathbf{R}^2$, as in Section 3.8. In this case, the differential of f at z is the linear mapping from $\mathbf{R}^2$ into itself that corresponds to

$$\text{multiplication by } f'(z) \text{ on } \mathbf{C}, \tag{C.2.2}$$

which is a complex-linear mapping from $\mathbf{C}$ into itself.

Conversely, suppose that f is differentiable at z as a mapping from U into $\mathbf{R}^2$, as in Section 3.8, and that

$$\text{the differential of } f \text{ at } z \text{ is complex-linear}, \tag{C.2.3}$$

when considered as a mapping from $\mathbf{C}$ into itself. This means that

$$\begin{gathered}\text{the differential of } f \text{ at } z \text{ corresponds to} \\ \text{multiplication by a complex number on } \mathbf{C},\end{gathered} \tag{C.2.4}$$

and one can verify that the complex derivative of f at z exists and is equal to that complex number.

The condition (C.2.3) is characterized by a first-order linear system of partial differential equations in the real and imaginary parts of f, which are known as the *Cauchy–Rieamann equations.*

C.2.1 More on holomorphic functions

If f is complex analytic on U, then it is well known that

$$\text{the complex derivatives of } f \text{ of all orders exist on } U. \tag{C.2.5}$$

In particular, this implies that f is infinitely differentiable on U, as a mapping from U into $\mathbf{R}^2$.

Let g be another complex-valued function on U. If the complex derivatives of f and g exist at $z \in U$, then one can show that the complex derivatives of $f + g$ and $f\,g$ exist at z, with

$$(f + g)'(z) = f'(z) + g'(z) \tag{C.2.6}$$

and

$$(f\,g)'(z) = f'(z)\,g(z) + f(z)\,g'(z). \tag{C.2.7}$$

If f and g are both holomorphic on U, then it follows that

$$f + g \text{ and } f\,g \text{ are holomorphic on } U \tag{C.2.8}$$

as well.

Of course, constant functions are holomorphic on $\mathbf{C}$, with complex derivative equal to 0. It is easy to see that the identity mapping on $\mathbf{C}$ is holomorphic, with complex derivative equal to 1. Using this and (C.2.8), we get that

(C.2.9) all polynomials in z with complex coefficients are holomorphic on $\mathbf{C}$.

Suppose that f is not equal to 0 at any point in U, so that $1/f$ defines a complex-valued function on U too. If the complex derivative of f exists at $z \in U$, then one can show that the complex derivative of $1/f$ exists at z, with

$$(1/f)'(z) = -\frac{f'(z)}{f(z)^2}. \tag{C.2.10}$$

This implies that

(C.2.11) $1/f$ is holomorphic on U when f is holomorphic on U.

One can use this to get that rational functions of z are holomorphic on the set where the denominator is not zero.

C.2.2 Some related differential operators

Let f be a complex-valued function on U again. If U is considered as an open set in $\mathbf{R}^2$, then the *partial derivatives*

$$\partial f/\partial x \text{ and } \partial f/\partial y \tag{C.2.12}$$

may be defined as complex numbers, when they exist, in the same way as for real-valued functions on U. This is the same as taking the partial derivatives of f as a function on U with values in $\mathbf{R}^2$, and interpreting the partial derivatives, when they exist, as complex numbers.

Put

$$\frac{\partial f}{\partial z} = \frac{1}{2}\Big(\frac{\partial f}{\partial x} - i\,\frac{\partial f}{\partial y}\Big) \tag{C.2.13}$$

and

$$\frac{\partial f}{\partial \overline{z}} = \frac{1}{2}\Big(\frac{\partial f}{\partial x} + i\,\frac{\partial f}{\partial y}\Big), \tag{C.2.14}$$

when the partial derivatives on the right sides of these equations exist. We may be particularly concerned with these expresseions at points where f is differentiable as a mapping from U into $\mathbf{R}^2$. One can check that

$$\frac{\partial f}{\partial \overline{z}} = 0 \tag{C.2.15}$$

is equivalent to the system of Cauchy–Riemann equations for the real and imaginary parts of f. If the complex derivative of f at a point $z \in U$ exists, then one can verify that

$$f'(z) = \frac{\partial f}{\partial z}(z). \tag{C.2.16}$$

C.3 Complex differential forms

Let k be a positive integer, and let V be a vector space over the real numbers. Consider the space

$$\mathcal{M}_k(V, \mathbf{C}) \tag{C.3.1}$$

of multilinear mappings from the space V^k of k-tuples of eleents of V into $\mathbf{C}$, as a vector space over the real numbers. This is the same as in Section 9.2, with $Z = \mathbf{C}$, as a vector space over the real numbers.

The elements of (C.3.1) may be described as *complex-valued k-linear forms* on V. Note that a complex-valued function on V^k is an element of (C.3.1) if and only if its real and imaginary parts are real-valued k-linear forms on V.

It is easy to see that (C.3.1) is a linear subspace of the space of all complex-valued functions on V^k, as a vector space over the complex numbers with respect to pointwise addition and scalar multiplication of functions. This means that (C.3.1) may be considered as a vector space over the complex numbers, with respect to pointwise addition and scalar multiplication on V^k.

Basically everything that we did before for ordinary differential forms can be extended to differential forms with complex-valued coefficients.

C.3.1 Complex differential forms on C

In particular,

$$dz = dx + i\,dy \text{ and } d\overline{z} = dx - i\,dy \tag{C.3.2}$$

may be considered as differential 1-forms on the the complex plane with complex coefficients. Observe that

$$dz \wedge dz = d\overline{z} \wedge d\overline{z} = 0 \tag{C.3.3}$$

and

$$dz \wedge d\overline{z} = -d\overline{z} \wedge dz = -2\,i\,dx \wedge dy. \tag{C.3.4}$$

Of course,

$$dx = (1/2)\,(dz + d\overline{z}) \text{ and } dy = (-i/2)\,(dz - d\overline{z}). \tag{C.3.5}$$

Thus any differential 1-form on a subset E of $\mathbf{C}$ with complex-valued coefficients may be expressed in a unique way as

$$a\,dz + b\,d\overline{z}, \tag{C.3.6}$$

where a and b are complex-valued functions on E.

Let U be a nonempty open set in $\mathbf{C}$, and let f be a complex-valued function on U. Suppose that f is continuously differentiable as an $\mathbf{R}^2$-valued function on U. One can check that

$$df = \frac{\partial f}{\partial z}\, dz + \frac{\partial f}{\partial \overline{z}}\, d\overline{z} \tag{C.3.7}$$

on U, as differential 1-forms on U with complex coefficients.

C.3.2 Some closed forms on C

Similarly,

$$d(f\, dz) = df \wedge dz = \frac{\partial f}{\partial \overline{z}}\, d\overline{z} \wedge dz = -\frac{\partial f}{\partial \overline{z}}\, dz \wedge d\overline{z} \tag{C.3.8}$$

on U, using (C.3.3) in the second step. This implies that $f\, dz$ is closed as a differential 1-form on U with complex coefficients if and only if f is holomorphic on U.

Let F be another complex-valued function on U, and suppose that F is continuously-differentiable as an $\mathbf{R}^2$-valued function on U. It is easy to see that

$$dF = f\, dz \tag{C.3.9}$$

on U if and only if F is holomorphic on U, with

$$F' = f \tag{C.3.10}$$

on U.

Bibliography

[1] E. Acosta G. and C. Delgado G., *Fréchet vs. Carathéodory*, American Mathematical Monthly **101** (1994), 332–338. Note that a reference to [168] was inadvertently omitted from this paper, as mentioned in the review of the paper in Mathematical Reviews (MR1270957).

[2] I. Adamson, *Transformations of integrals*, American Mathematical Monthly **65** (1958), 590–596.

[3] L. Ahlfors, *Complex Analysis — An Introduction to the Theory of Analytic Functions of One Complex Variable*, 3rd edition, AMS Chelsea, 2021.

[4] S. Althoen and J. Weidner, *Real division algebras and Dickson's construction*, American Mathematical Monthly **85** (1978), 368–371.

[5] M. Atiyah, *Bakerian lecturee, 1975: global geometry*, Proceeedings of the Royal Society **A347** (1976), 291–299; American Mathematical Monthly **111** (2004), 716–723.

[6] M. Atiyah, *Mathematics in the 20th century*, American Mathematical Monthly **108** (2001), 654–666; Bulletin of the London Mathematical Society **34** (2002), 1–15.

[7] M. Atiyah and I. Macdonald, *Introduction to Commutative Algebra*, Addison-Wesley, 1969.

[8] D. Bachman, *A Geometric Approach to Differential Forms*, Birkhäuser, 2006.

[9] J. Bak and D. Newman, *Complex Analysis*, 3rd edition, Springer, 2010.

[10] J. Bak and S. Popvassilev, *The evolution of Cauchy's closed curve theorem and Newman's simple proof*, American Mathematical Monthly **124** (2017), 217–231.

[11] J. Baker, *Isometries in normed spaces*, American Mathematical Monthly **78** (1971), 655–658.

[12] J. Baker, *Plane curves, polar coordinates and winding numbers*, Mathematics Magazine **64** (1991), 75–91.

[13] J. Baker, *Integration over spheres and the divergence theorem for balls*, American Mathematical Monthly **104** (1997), 36–47.

[14] Y. Baryshnikov and R. Ghrist, *Stokes's theorem, data, and the polar ice caps*, American Mathematical Monthly **125** (2018), 830–834.

[15] A. Bennett, *On the Definition of Determinants*, American Mathematical Monthly **31** (1924), 343–345.

[16] A. Bennett, *The Notion "Matrix of a Determinant"*, American Mathematical Monthly **32** (1925), 182–185.

[17] G. Berman, *The wedge product*, American Mathematical Monthly **68** (1961), 112–119.

[18] G. Berresford, *Cauchy's theorem*, American Mathematical Monthly **88** (1981), 741–744.

[19] P. Binding, *More on cross products*, American Mathematical Monthly **98** (1991), 850–851.

[20] G. Birkhoff and S. Mac Lane, *A Survey of Modern Algebra*, 4th edition, Macmillan, 1977.

[21] K. Bisshopp, *The Triple Vector Product* $A \times (B \times C)$, American Mathematical Monthly **72** (1965), 163–165.

[22] K. Bogart, *A fresh(man) treatment of determinants*, American Mathematical Monthly **96** (1989), 915–920.

[23] W. Boothby, *On two classical theorems of algebraic topology*, American Mathematical Monthly **78** (1971), 237–249.

[24] W. Boothby, *An Introduction to Differentiable Manifolds and Riemannian Geometry*, Academmic Press, 1975. A second edition was published in 1986.

[25] R. Bott and L. Tu, *Differential Forms in Algebraic Topology*, Springer-Verlag, 1982.

[26] L. Brand, *The vector potential of a solenoidal vector*, American Mathematical Monthly **57** (1950). 161–167.

[27] J. Britt, *The anatomy of low-dimensional stable singularities*, American Mathematical Monthly **92** (1985), 183–201.

[28] A. Browder, *Mathematical Analysis: An Introduction*, Springer-Verlag, 1996.

[29] A. Browder, *Topology in the complex plane*, American Mathematical Monthly **107** (2000), 393–401.

[30] A. Browder, *Complex numbers and the ham sandwich theorem*, American Mathematical Monthly **113** (2006), 935–937.

[31] A. Brown, *On transformation of multiple integrals*, American Mathematical Monthly **48** (1941), 29–33.

[32] A. Brown, *Line integrals in a multiply-connected region*, Mathematics Magazine **34** (1960/61), 230–232.

[33] J. Bruce and P. Giblin, *Generic geometry*, American Mathematical Monthly **90** (1983), 529–545.

[34] F. Callahan, *Pivotal role of the triple cross product*, American Mathematical Monthly **76** (1969), 649–651.

[35] J. Callahan, *Singularities and plane maps*, American Mathematical Monthly **81** (1974), 211–240.

[36] J. Callahan, *Singularities and plane maps II: Sketching catastrophes*, American Mathematical Monthly **84** (1977), 765–803.

[37] T. Cameron, *The determinant from signed volume to the Laplace expansion*, American Mathematical Monthly **126** (2019), 437–447.

[38] J. Cantarella, D. DeTurck, and H. Gluck, *Vector calculus and the topology of domains in* 3*-space*, American Mathematical Monthly **109** (2002), 409–442.

[39] S. Cater. *Scalar valued mappings of square matrices*, American Mathematical Monthly **70** (1963), 163–169.

[40] S.-S. Chern, *From triangles to manifolds*, American Mathematical Monthly **86** (1979), 339–349.

[41] S.-S. Chern, *What is geometry?*, American Mathematical Monthly **97** (1990), 679–686.

[42] C. Coe and G. Rainich, *Fundamental Identity of Vector Algebra*, American Mathematical Monthly **56** (1949), 175–176.

[43] A. Coffman and Y. Zhang, *Vector fields with continuous curl but discontinuous partial derivatives*, American Mathematical Monthly **127** (2020), 911–917.

[44] R. Cohen, *Set isometries and their extensions in the max metrix*, American Mathematical Monthly **97** (1990), 795–808.

[45] D. Conant and W. Beyer, *Generalized Pythagorean theorem*, American Mathematical Monthly (1974), 262–265.

[46] J. Conway, *Functions of One Complex Variable*, 2nd edition, Springer-Verlag, 1978.

[47] J. Conway, *A First Course in Analysis*, Cambridge University Press, 2018.

[48] C. Cullen and K. Gale, *A Functional Definition of the Determinant*, American Mathematical Monthly **72** (1965), 403–406.

[49] A. Devinatz, *Advanced Calculus*, Holt, Rinehart, and Winston, 1968.

[50] A. Dittmer, *Cross product identities in arbitrary dimension*, American Mathematical Monthly **101** (1994), 887–891.

[51] T. Drammatico, *A new proof of the equality of mixed partial derivatives*, American Mathematical Monthly **122** (2015), 602–603.

[52] J.-F. Dumais, *On the Curl of a Vector Field*, American Mathematical Monthly **89** (1982), 469–473.

[53] D. Dwyer, *On the Vector Triple Cross Product Identity*, American Mathematical Monthly **68** (1961), 910.

[54] W. Eberlein, *The Gauss–Green and Cauchy integral theorems*, American Mathematical Monthly **82** (1975), 625–629.

[55] L. Eifler and N. Rhee, *The n-dimensional Pythagorean theorem via the divergence theorem*, American Mathematical Monthly **115** (2008), 456–457.

[56] H. Ellis, *Directional differentiation in the plane and tangent vectors on C^r manifolds*, American Mathematical Monthly **82** (1975), 641–645.

[57] A. Fadell, *A Proof of the Chain Rule for Derivatives in n-Space*, American Mathematical Monthly **80** (1973), 1134–1135.

[58] D. Fearnley-Sander, *A characterization of the determinant*, American Mathematical Monthly **82** (1975), 838–840.

[59] D. Fearnley-Sander, *Hermann Grassmann and the creation of linear algebra*, American Mathematical Monthly **86** (1979), 809–817.

[60] D. Fearnley-Sander, *Hermann Grassmann and the prehistory of universal algebra*, American Mathematical Monthly **89** (1982), 161–166.

[61] W. Fleming, *Functions of Several Variables*, 2nd edition, Springer-Verlag, 1977.

[62] T. Flett, *On transformations in $\mathbf{R}^n$ and a theorem of Sard*, American Mathematical Monthly **71** (1964), 623-629.

[63] G. Folland, *How to integrate a polynomial over a sphere*, American Mathematical Monthly **108** (2001), 446–448.

[64] J. Fortney, *A Visual Introduction to Differential Forms and Calculus on Manifolds*, Birkhäuser / Springer, 2018.

[65] J. Frame, *Proofs of Vector Product Expansions by Determinants*, American Mathematical Monthly **47** (1940), 644–646.

[66] A. Galbis and M. Maestre, *Vector Analysis versus Vector Calculus*, Springer, 2012.

[67] T. Gamelin, *Complex Analysis*, Springer-Verlag, 2001.

[68] T. Gamelin and R. Greene, *Introduction to Topology*, 2nd edition, Dover, 1999.

[69] E. Gaughan, *An Isometry of the Plane*, American Mathematical Monthly **69** (1962), 546.

[70] C. Goffman, *Calculus of Several Variables*, Harper & Row, 1965.

[71] C. Goffman, *Introduction to Real Analysis*, Harper & Row, 1966.

[72] W. Gordon, *On the diffeomorphisms of Euclidean space*, American Mathematical Monthly **79** (1972), 755–759; addendum, ibid., **80** (1973), 674–675.

[73] W. Gordon, *An application of Hadamard's inverse function theorem to algebra*, American Mathematical Monthly **84** (1977), 28–29.

[74] J. Gray and S. Morris, *When is a function that satisfies the Cauchy–Riemann equations analytic?*, American Mathematical Monthly **85** (1978), 246–256.

[75] M. Greenberg, *Element of area via exterior algebra*, American Mathematical Monthly **83** (1976), 274–275.

[76] R. Greene and S. Krantz, *Function Theory of One Complex Variable*, 3rd edition, American Mathematical Society, 2006.

[77] C. Grimm, *A note on the vector product* $a \times (b \times c)$, Mathematics Magazine **43** (1970), 150.

[78] G. Gross and E. Meinrenken, *Manifolds, Vector Fields, and Differential Forms: An Introduction to Differential Geometry*, Springer, 2023.

[79] R. Gunning, *An Introduction to Analysis*, Princeton University Press, 2018.

[80] F. Halpern, *Using the Multivariable Chain Rule*, American Mathematical Monthly 92 (1985), 144–145.

[81] H. Hanche-Olsen, *On Goursat's proof of Cauchy's integral theorem*, American Mathematical Monthly **115**, 648–652.

[82] J. Hannah, *A geometric approach to determinants*, American Mathematical Monthly **103** (1996), 401–409.

[83] A. Hatcher, *Algebraic Topology*, Cambridge University Press, 2002.

[84] O. Hijab, *The volume of the unit ball in* $\mathbf{C}^n$, American Mathematical Monthly **107** (2000), 259.

[85] J.-B. Hiriart-Urruty, *At What Points is the Projection Mapping Differentiable?*, American Mathematical Monthly **89** (1982), 45–458.

[86] J. Humphreys, *Introduction to Lie Algebras and Representation Theory*, 2nd printing, revised, Springer-Verlag, 1978.

[87] S.-G. Hwang and K.-B. Nam, *A determinant proof of a generalized Pythagorean theorem*, American Mathematical Monthly **127** (2020), 269–272.

[88] N. Ivanov, *A differential forms perspective on the Lax proof of the change of variables formula*, American Mathematical Monthly **112** (2005), 799–806.

[89] N. Jacobson, *Lie Algebras*, Dover, 1979.

[90] I. James, editor, *History of Topology*, North–Holland, 1999.

[91] K. Jänich, *Vector Analysis*, translated from the second German (1993) edition by L. Kay, Springer-Verlag, 2001.

[92] F. Jones, *Lebesgue Integration on Euclidean Space*, Jones and Bartlett, 1993.

[93] F. Jones, *Honors Calculus III and IV*, available from his website at Rice University (https://mathweb.rice.edu/honors-calculus).

[94] H. Junghenn, *A Course in Real Analysis*, CRC Press, 2015.

[95] M. Khosravi and M. Taylor, *The wedge product and analytic geometry*, American Mathematical Monthly **115** (2008), 623–644.

[96] A. Knapp, *Basic Real Analysis*, Birkhäuser, 2005.

[97] A. Knapp, *Basic Algebra*, Birkhäuser, 2006.

[98] A. Knapp, *Stokes's Theorem and Whitney Manifolds: A Sequel to* Basic Real Analysis, Anthony W. Knapp, East Setauket, NY, 2021.

[99] A. Konnully, *A Proof of the Cauchy Integral Theorem*, American Mathematical Monthhly **74** (1967), 54–56.

[100] T. Körner, *Butterfly hunting and the symmetry of mixed partial derivatives*, American Mathematical Monthly **105** (1998), 756–758.

[101] T. Körner, *A Companion to Analysis: A Second First and First Second Course in Analysis*, American Mathematical Society, 2004.

[102] S. Krantz, *Function Theory of Several Complex Variables*, AMS Chelsea, 2001.

[103] S. Krantz, *A Guide to Complex Variables*, Mathematical Association of America, 2008.

[104] S. Krantz, *Foundations of Analysis*, CRC Press, 2015.

[105] S. Krantz, *Real Analysis and Foundations*, 4th edition, CRC Press, 2017.

[106] S. Krantz and H. Parks, *The Implicit Function Theorem: History, Theory, and Applications*, Birkhäuser / Springer, 2013.

[107] K. Kreith, *Geometric interpretation of the Implicit Function Theorem*, Mathematics Magazine **36** (1963), 64–64.

[108] I. Kriz and A. Pultr, *Introduction to Mathematical Analysis*, Birkhäuser / Springer, 2013.

[109] J. Lafontaine, *An Introduction to Differential Manifolds*, based on the 2010 French second edition, Springer, 2015.

[110] S. Lang, *Undergraduate Analysis*, 2nd edition, Springer-Verlag, 1997.

[111] S. Lang, *Complex Analysis*, 4th edition, Springer-Verlag, 1999.

[112] H. Lass, *A Note on Cauchy's Theorem*, American Mathematical Monthly **60** (1953), 110–112.

[113] J. Lasserre and K. Avrachenkov, *The multi-dimensional version of* $\int_a^b x^p\,dx$, American Mathematical Monthly **108** (2001), 151–154.

[114] R. Lawlor, *On the Invariance of the Divergence of a Vector Function*, American Mathematical Monthly **38** (1931), 28–29.

[115] P. Lax, *Change of variables in multiple integrals*, American Mathematical Monthly **106** (1999), 497–501.

[116] P. Lax, *Change of variables in multiple integrals II*, American Mathematical Monthly **108** (2001), 115–119.

[117] P. Lax, *The Cauchy integral theorem*, American Mathematical Monthly **114** (2007), 725–727.

[118] P. Ledden, *A Note on the Theorem of Green and Stokes*, American Mathematical Monthly **75** (1968), 1095–1098.

[119] J. Lee, *Introduction to Smooth Manifolds*, 2nd edition, Springer, 2013.

[120] C.-K. Li, *Norms, isometries, and isometry groups*, American Mathematical Monthly **107** (2000), 334–340.

[121] C.-K. Li and W. So, *Isometries of l_p norm*, American Mathematical Monthly **101** (1994), 452–453.

[122] P. Loeb, *A note on Dixon's proof of Cauchy's integral theorem*, American Mathematical Monthly **98** (1991), 242–244.

[123] P. Loeb, *A Further Simplification of Dixon's Proof of Cauchy's Integral Theorem*, American Mathematical Monthly **100** (1993), 680–681.

[124] W. Lovitt, *A Geometrical Interpretation of Green's Formula*, American Mathematical Monthly **22** (1915), 152–154.

[125] S. Mac Lane and G. Birkhoff, *Algebra*, 3rd edition, Chelsea, 1988.

[126] W. Massey, *Cross products of vectors in higher-dimensional Euclidean spaces*, American Mathematical Monthly **90** (1983), 697–701.

[127] W. Massey, *A Basic Course in Algebraic Topology*, Springer-Verlag, 1991.

[128] P. McGrath, *Another proof of Clairaut's theorem*, American Mathematical Monthly **121** (2014), 165–166.

[129] A. McInerney, *First Steps in Differential Geometry: Riemannian, Contact, Symplectic*, Springer, 2013.

[130] B. Mendelson, *Introduction to Topology*, 3rd edition, Dover, 1990.

[131] P. Mikusiński and M. Taylor, *An Introduction to Multivariable Analysis from Vector to Manifold*, Birkhäuser, 2002.

[132] G. Miller, *On the History of Determinants*, American Mathematical Monthly **37** (1930), 216–219.

[133] J. Milnor, *Topology through the centuries: low dimensional manifolds*, Bulletin of the American Mathematical Society (N.S.) **52** (2015), 545–584.

[134] T. Moran, S. Moran, and S. Moran, *Elementary derivations of the Euclidean Hurwitz algebras adapted from Gadi Moran's last paper*, American Mathematical Monthly **128** (2021), 726–736.

[135] F. Morgan, *What is a surface?*, American Mathematical Monthly **103** (1996), 369–376.

[136] F. Morgan, *Geometric Measure Theory: A Beginner's Guide*, 5th edition, Illustrated by J. Bredt, Elsevier / Academic Press, 2016.

[137] R. Mortini, *A short proof of the chain rule for differentiable mappings in* $\mathbf{R}^n$, mathematics Magazine **85** (2012), 136–141.

[138] J. Munkres, *Analysis on Manifolds*, Addison-Wesley, 1991.

[139] A. Nash, *A generalized parallelogram law*, American Mathematical Monthly **110** (2003), 52–57.

[140] T. Needham, *Visual Differential Geometry and Forms — A Mathematical Drama in Five Acts*, Princeton University Press, 2021.

[141] T. Needham, *Visual Complex Analysis*, 25th anniversary edition, Oxford University Press, 2023.

[142] E. Outerelo and J. Ruiz, *Mapping Degree Theory*, American Mathematical Society, Real Sociedad Matemática Española, 2009.

[143] J. Petro, *Real division algebras of dimension* > 1 *contain* $\mathbf{C}$, American Mathematical Monthly **94** (1987), 445–449.

[144] I. Pinelis, *A simple and more general approach to Stokes' theorem*, Mathematics Magazine **92** (2019), 345–351.

[145] G. Porter, *k-Volume in $\mathbf{R}^n$ and the generalized Pythagorean theorem*, American Mathematical Monthly **103** (1996), 252–256.

[146] C. Pugh, *Real Mathematical Analysis*, 2nd edition, Springer, 2015.

[147] H. Pyle, *Non-square determinants and multilinear vectors*, Mathematics Magazine **35** (1962), 65–69.

[148] H. Pyle, *Inner products of multilinear vectors*, Mathematics Magazine **37** (1964), 93–100.

[149] J.-P. Quadrat, J. Lasserre, and J.-B. Hiriart-Urruty, *Pythagoras' theorem for areas*, American Mathematical Monthly **108** (2001), 549–551.

[150] R. Redheffer, *The homotopy theorems of function theory*, American Mathematical Monthly **76** (1969), 778–787.

[151] D. Richmond, *Complex Numbers and Vector Algebra*, American Mathematical Monthly **58** (1951), 622–628.

[152] J. Roe, *Winding Around: The Winding Number in Topology, Geometry, and Analysis*, American Mathematical Society; Mathematics Advanced Study Seminars, University Park, PA, 2015.

[153] M. Rosenlicht, *Introduction to Analysis*, Dover, 1986.

[154] W. Rudin, *Real and Complex Analysis*, 2nd edition, McGraw-Hill, 1974. A third edition was published in 1987.

[155] W. Rudin, *Principles of Mathematical Analysis*, 3rd edition, McGraw-Hill, 1976.

[156] H. Samelson, *Differential forms, the early days; or the stories of Deahna's theorem and of Volterra's theorem*, American Mathematical Monthly **108** (2001), 522–530.

[157] D. Schattschneider, *The taxicab group*, American Mathematical Monthly **91** (1984), 423–428.

[158] P. Scherk and M. Kwizak, *What are tensors?*, American Methamtical Monthly **58** (1951), 297–305.

[159] M. Schreiber, *Differential Forms: A Heuristic Introduction*, Springer-Verlag, 1977.

[160] J.-P. Serre, *Lie Algebras and Lie Groups*, 1964 lectures given at Harvard University, corrected fifth printing of the second (1992) edition, Lecture Notes in Mathematics **1500**, Springer-Verlag, 2006.

[161] R. Shakarchi, *Problems and Solutions for* Undergraduate Analysis [110], Springer-Verlag, 1998.

[162] R. Shakarchi, *Problems and Solutions for* Complex Analysis [111], Springer-Verlag, 1999.

[163] M. Shanks, *Note on Vector Products*, American Mathematical Monthy **58** (1951), 256–257.

[164] A. Shastri, *Elements of Differential Topology*, CRC Press, 2011.

[165] R. Shaw and F. Yeadon, *On* $(a \times b) \times c$, American Mathematical Monthly **96** (1989), 623–629.

[166] S. Shirali and H. Vasudeva, *Multivariable Analysis*, Springer-Verlag, 2011.

[167] H. Siegberg, *Some historical remarks concerning degree theory*, American Mathematical Monthly **88** (1981), 125–139.

[168] M. Spivak, *Calculus on Manifolds: A Modern Approach to Classical Theorems of Advanced Calculus*, Benjamin, 1965.

[169] E. Stein and R. Shakarchi, *Complex Analysis*, Princeton University Press, 2003.

[170] P. Stein, *A note on the volume of a simplex*, American Mathematical Monthly **73** (1966), 299–301.

[171] P. Stein, *A proof of a classical theorem in multiple integration*, American Mathematical Monthly **75** (1968), 160–163.

[172] K. Stromberg, *Introduction to Classical Real Analysis*, AMS Chelsea, 1981.

[173] J. Taylor, *Foundations of Analysis*, American Mathematical Society, 2012.

[174] M. Taylor, *Introduction to Analysis in Several Variables — Advanced Calulus*, American Mathematical Society, 2020.

[175] A. Thompson, *Minkowski Geometry*, Cambridge University Press, 1996.

[176] D. Trahan, *The Mixed Partial Derivatives and the Double Derivative*, American Mathematical Monthly **76** (1969), 76–77.

[177] F. Trèves, *Topological Vector Spaces, Distributions, and Kernels*, Dover, 2006.

[178] L. Tu, *An Introduction to Manifolds*, 2nd edition, Springer, 2011.

[179] J. Väisälä, *A proof of the Mazur–Ulam theorem*, American Mathematical Monthly **110** (2003), 633–635.

[180] D. Varberg, *On differentiable transformations in* $\mathbf{R}^n$, American Mathematical Monthly **73** (1966), no. 4, part II, 111–114.

[181] D. Varberg, *Change of variables in multiple integrals*, American Mathematical Monthly **78** (1971), 42–45.

[182] R. Výborný, *On the use of a differentiable homotopy in the proof of the Cauchy theorem*, American Mathematical Monthly **86** (1979), 380–382.

[183] B. Walsh, *The scarcity of cross products on Euclidean spaces*, American Mathematical Monthly **74** (1967), 188–194.

[184] F. Warner, *Foundations of Differentiable Manifolds and Lie Groups*, Scott, Foresman, & Co., 1971. A corrected reprint was published by Springer-Verlag in 1983.

[185] S. Weintraub, *Differential Forms: A Complement to Vector Calculus*, Academic Press, 1997.

[186] S. Weintraub, *Differential Forms: Theory and Practice*, 2nd edition, Elsevier / Academic Press, 2014.

[187] J. Weston, *Volume in vector spaces*, American Mathematical Monthly **66** (1959), 575–577.

[188] J. Weston, *Some remarks about the curl of a vector field*, American Mathematical Monthly **68** (1961), 359–361.

[189] D. Widder, *A simplified approach to Cauchy's integral theorem*, American Mathematical Monthly **53** (1946), 359–363.

[190] M. Williams and F. Stein, *A triple product of vectors in four-space*, Mathematics Magazine **37** (1964), 230–235.

[191] R. Wrede, *A Proof of the Vector Triple Cross Product Identity*, American Mathematical Monthly **67** (1960), 574–576.

[192] O. Wyler, *The Cauchy Integral Theorem*, American Mathematical Monthly **72** (1965), 50–53.

[193] O. Wyler, *Invariant Definitions for Vector Calculus*, American Mathematical Monthly **75** (1968), 394–396.

[194] S. Yap, *The Poincaré lemma and an elementary construction of vector potentials*, American Mathematical Monthly **116** (2009), 261–267.

Index