This is page 1
Printer: Opaque this

CHAPTER 1

MULTILINEAR ALGEBRA

1.1 Background

We will list below some definitions and theorems that are part of the curriculum of a standard theory-based sophomore level course in linear algebra. (Such a course is a prerequisite for reading these notes.) A *vector space* is a set, V, the elements of which we will refer to as *vectors*. It is equipped with two vector space operations:
Vector space addition. Given two vectors, v_1 and v_2, one can add them to get a third vector, $v_1 + v_2$.
Scalar multiplication. Given a vector, v, and a real number, λ, one can multiply v by λ to get a vector, λv.

These operations satisfy a number of standard rules: associativity, commutativity, distributive laws, etc. which we assume you're familiar with. (See exercise 1 below.) In addition we'll assume you're familiar with the following definitions and theorems.

1. *The zero vector.* This vector has the property that for every vector, v, $v + 0 = 0 + v = v$ and $\lambda v = 0$ if λ is the real number, zero.

2. *Linear independence.* A collection of vectors, v_i, $i = 1, \ldots, k$, is *linearly* independent if the map

$$\mathbb{R}^k \to V\,, \quad (c_1, \ldots, c_k) \to c_1 v_1 + \cdots + c_k v_k \tag{1.1.1}$$

is $1 - 1$.

3. The *spanning property.* A collection of vectors, v_i, $i = 1, \ldots, k$, *spans* V if the map (1.1.1) is onto.

4. The notion of *basis.* The vectors, v_i, in items 2 and 3 are a basis of V if they span V and are linearly independent; in other words, if the map (1.1.1) is bijective. This means that every vector, v, can be written uniquely as a sum

$$v = \sum c_i v_i\,. \tag{1.1.2}$$

5. The *dimension* of a vector space. If V possesses a basis, v_i, $i = 1, \dots, k$, V is said to be *finite dimensional*, and k is, by definition, the *dimension* of V. (It is a theorem that this definition is legitimate: every basis has to have the same number of vectors.) In this chapter all the vector spaces we'll encounter will be finite dimensional.

6. A subset, U, of V is a *subspace* if it's vector space in its own right, i.e., for v, v_1 and v_2 in U and λ in $\mathbb{R}$, λv and $v_1 + v_2$ are in U.

7. Let V and W be vector spaces. A map, $A : V \to W$ is *linear* if, for v, v_1 and v_2 in V and $\lambda \in \mathbb{R}$

$$A(\lambda v) = \lambda A v \tag{1.1.3}$$

and

$$A(v_1 + v_2) = Av_1 + Av_2 \,. \tag{1.1.4}$$

8. The *kernel of* A. This is the set of vectors, v, in V which get mapped by A into the zero vector in W. By (1.1.3) and (1.1.4) this set is a subspace of V. We'll denote it by "$\operatorname{Ker} A$".

9. *The image of* A. By (1.1.3) and (1.1.4) the image of A, which we'll denote by "$\operatorname{Im} A$", is a subspace of W. The following is an important rule for keeping track of the dimensions of $\operatorname{Ker} A$ and $\operatorname{Im} A$.

$$\dim V = \dim \operatorname{Ker} A + \dim \operatorname{Im} A \,. \tag{1.1.5}$$

Example 1. The map (1.1.1) is a linear map. The v_i's span V if its image is V and the v_i's are linearly independent if its kernel is just the zero vector in $\mathbb{R}^k$.

10. *Linear mappings and matrices.* Let $v_1, \dots, v_n$ be a basis of V and $w_1, \dots, w_m$ a basis of W. Then by (1.1.2) Av_j can be written uniquely as a sum,

$$Av_j = \sum_{i=1}^{m} c_{i,j} w_i \,, \quad c_{i,j} \in \mathbb{R} \,. \tag{1.1.6}$$

The $m \times n$ matrix of real numbers, $[c_{i,j}]$, is the *matrix* associated with A. Conversely, given such an $m \times n$ matrix, there is a unique linear map, A, with the property (1.1.6).

11. An *inner product* on a vector space is a map

$$B : V \times V \to \mathbb{R}$$

having the three properties below.

(a) For vectors, v, v_1, v_2 and w and $\lambda \in \mathbb{R}$

$$B(v_1 + v_2, w) = B(v_1, w) + B(v_2, w)$$

and

$$B(\lambda v, w) = \lambda B(v, w)\,.$$

(b) For vectors, v and w,

$$B(v, w) = B(w, v)\,.$$

(c) For every vector, v

$$B(v, v) \geq 0\,.$$

Moreover, if $v \neq 0$, $B(v, v)$ is positive.

Notice that by property (b), property (a) is equivalent to

$$B(w, \lambda v) = \lambda B(w, v)$$

and

$$B(w, v_1 + v_2) = B(w, v_1) + B(w, v_2)\,.$$

The items on the list above are just a few of the topics in linear algebra that we're assuming our readers are familiar with. We've highlighted them because they're easy to state. However, understanding them requires a heavy dollop of that indefinable quality "mathematical sophistication", a quality which will be in heavy demand in the next few sections of this chapter. We will also assume that our readers are familiar with a number of more low-brow linear algebra notions: matrix multiplication, row and column operations on matrices, transposes of matrices, determinants of $n \times n$ matrices, inverses of matrices, Cramer's rule, recipes for solving systems of linear equations, etc. (See §1.1 and 1.2 of Munkres' book for a quick review of this material.)

Exercises.

1. Our basic example of a vector space in this course is $\mathbb{R}^n$ equipped with the vector addition operation

$$(a_1, \dots, a_n) + (b_1, \dots, b_n) = (a_1 + b_1, \dots, a_n + b_n)$$

and the scalar multiplication operation

$$\lambda(a_1, \dots, a_n) = (\lambda a_1, \dots, \lambda a_n)\,.$$

Check that these operations satisfy the axioms below.

(a) Commutativity: $v + w = w + v$.

(b) Associativity: $u + (v + w) = (u + v) + w$.

(c) For the zero vector, $0 = (0, \dots, 0)$, $v + 0 = 0 + v$.

(d) $v + (-1)v = 0$.

(e) $1v = v$.

(f) Associative law for scalar multiplication: $(ab)v = a(bv)$.

(g) Distributive law for scalar addition: $(a + b)v = av + bv$.

(h) Distributive law for vector addition: $a(v + w) = av + aw$.

2. Check that the standard basis vectors of $\mathbb{R}^n$: $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0)$, etc. *are* a basis.

3. Check that the standard inner product on $\mathbb{R}^n$

$$B((a_1, \dots, a_n), (b_1, \dots, b_n)) = \sum_{i=1}^{n} a_i b_i$$

is an inner product.

1.2 Quotient spaces and dual spaces

In this section we will discuss a couple of items which are frequently, but not always, covered in linear algebra courses, but which we'll need for our treatment of multilinear algebra in §§1.1.3 – 1.1.8.

The quotient spaces of a vector space

Let V be a vector space and W a vector subspace of V. A *W-coset* is a set of the form

$$v + W = \{v + w\,,\ w \in W\}\,.$$

It is easy to check that if $v_1 - v_2 \in W$, the cosets, $v_1 + W$ and $v_2 + W$, coincide while if $v_1 - v_2 \notin W$, they are disjoint. Thus the W-cosets decompose V into a *disjoint* collection of subsets of V. We will denote this collection of sets by V/W.

One defines a vector addition operation on V/W by defining the sum of two cosets, $v_1 + W$ and $v_2 + W$ to be the coset

$$v_1 + v_2 + W \tag{1.2.1}$$

and one defines a scalar multiplication operation by defining the scalar multiple of $v + W$ by λ to be the coset

$$\lambda v + W\,. \tag{1.2.2}$$

It is easy to see that these operations are well defined. For instance, suppose $v_1 + W = v_1' + W$ and $v_2 + W = v_2' + W$. Then $v_1 - v_1'$ and $v_2 - v_2'$ are in W; so $(v_1 + v_2) - (v_1' + v_2')$ is in W and hence $v_1 + v_2 + W = v_1' + v_2' + W$.

These operations make V/W into a vector space, and one calls this space the *quotient space* of V by W.

We define a mapping

$$\pi : V \to V/W \tag{1.2.3}$$

by setting $\pi(v) = v + W$. It's clear from (1.2.1) and (1.2.2) that π is a linear mapping, and that it maps V to V/W. Moreover, for every coset, $v + W$, $\pi(v) = v + W$; so the mapping, π, is onto. Also note that the zero vector in the vector space, V/W, is the zero coset, $0 + W = W$. Hence v is in the kernel of π if $v + W = W$, i.e., $v \in W$. In other words the kernel of π is W.

In the definition above, V and W don't have to be finite dimensional, but if they are, then

$$\dim V/W = \dim V - \dim W\,. \tag{1.2.4}$$

by (1.1.5).

The following, which is easy to prove, we'll leave as an exercise.

Proposition 1.2.1. *Let U be a vector space and $A : V \to U$ a linear map. If $W \subset \operatorname{Ker} A$ there exists a unique linear map, $A^{\#} : V/W \to U$ with property, $A = A^{\#} \circ \pi$.*

The dual space of a vector space

We'll denote by V^* the set of all linear functions, $\ell : V \to \mathbb{R}$. If ℓ_1 and ℓ_2 are linear functions, their sum, $\ell_1 + \ell_2$, is linear, and if ℓ is a linear function and λ is a real number, the function, $\lambda\ell$, is linear. Hence V^* is a vector space. One calls this space the *dual space* of V.

Suppose V is n-dimensional, and let $e_1, \dots, e_n$ be a basis of V. Then every vector, $v \in V$, can be written uniquely as a sum

$$v = c_1 e_1 + \cdots + c_n e_n \quad c_i \in \mathbb{R}\,.$$

Let

$$e_i^*(v) = c_i\,. \tag{1.2.5}$$

If $v = c_1 e_1 + \cdots + c_n e_n$ and $v' = c_1' e_1 + \cdots + c_n' e_n$ then $v + v' = (c_1 + c_1')e_1 + \cdots + (c_n + c_n')e_n$, so

$$e_i^*(v + v') = c_i + c_i' = e_i^*(v) + e_i^*(v')\,.$$

This shows that $e_i^*(v)$ is a linear function of v and hence $e_i^* \in V^*$.

Claim: e_i^*, $i = 1, \dots, n$ is a basis of V^*.

Proof. First of all note that by (1.2.5)

$$e_i^*(e_j) = \begin{cases} 1\,, & i = j \\ 0\,, & i \neq j \end{cases}\,. \tag{1.2.6}$$

If $\ell \in V^*$ let $\lambda_i = \ell(e_i)$ and let $\ell' = \sum \lambda_i e_i^*$. Then by (1.2.6)

$$\ell'(e_j) = \sum \lambda_i e_i^*(e_j) = \lambda_j = \ell(e_j)\,, \tag{1.2.7}$$

i.e., ℓ and ℓ' take identical values on the basis vectors, e_j. Hence $\ell = \ell'$.

Suppose next that $\sum \lambda_i e_i^* = 0$. Then by (1.2.6), $\lambda_j = (\sum \lambda_i e_i^*)(e_j) = 0$ for all $j = 1, \dots, n$. Hence the e_j^*'s are linearly independent. □

Let V and W be vector spaces and $A : V \to W$, a linear map. Given $\ell \in W^*$ the composition, $\ell \circ A$, of A with the linear map, $\ell : W \to \mathbb{R}$, is linear, and hence is an element of V^*. We will denote this element by $A^*\ell$, and we will denote by

$$A^* : W^* \to V^*$$

the map, $\ell \to A^*\ell$. It's clear from the definition that

$$A^*(\ell_1 + \ell_2) = A^*\ell_1 + A^*\ell_2$$

and that

$$A^*\lambda\ell = \lambda A^*\ell\,,$$

i.e., that A^* is linear.

Definition. A^* is the transpose of the mapping A.

We will conclude this section by giving a matrix description of A^*. Let $e_1, \ldots, e_n$ be a basis of V and $f_1, \ldots, f_m$ a basis of W; let $e_1^*, \ldots, e_n^*$ and $f_1^*, \ldots, f_m^*$ be the dual bases of V^* and W^*. Suppose A is defined in terms of $e_1, \ldots, e_n$ and $f_1, \ldots, f_m$ by the $m \times n$ matrix, $[a_{i,j}]$, i.e., suppose

$$Ae_j = \sum a_{i,j} f_i\,.$$

Claim. A^* is defined, in terms of $f_1^*, \ldots, f_m^*$ and $e_1^*, \ldots, e_n^*$ by the transpose matrix, $[a_{j,i}]$.

Proof. Let

$$A^* f_i^* = \sum c_{j,i} e_j^*\,.$$

Then

$$A^* f_i^*(e_j) = \sum_k c_{k,i} e_k^*(e_j) = c_{j,i}$$

by (1.2.6). On the other hand

$$A^* f_i^*(e_j) \quad = \quad f_i^*(Ae_j) = f_i^*\left(\sum a_{k,j} f_k\right) = \sum_k a_{k,j} f_i^*(f_k) = a_{i,j}$$

so $a_{i,j} = c_{j,i}$.

□

Exercises.

1. Let V be an n-dimensional vector space and W a k-dimensional subspace. Show that there exists a basis, $e_1, \ldots, e_n$ of V with the property that $e_1, \ldots, e_k$ is a basis of W. *Hint:* Induction on $n-k$. To start the induction suppose that $n-k=1$. Let $e_1, \ldots, e_{n-1}$ be a basis of W and e_n any vector in $V-W$.

2. In exercise 1 show that the vectors $f_i = \pi(e_{k+i})$, $i = 1, \ldots, n-k$ are a basis of V/W.

3. In exercise 1 let U be the linear span of the vectors, e_{k+i}, $i = 1, \ldots, n-k$.

Show that the map

$$U \to V/W\,, \quad u \to \pi(u)\,,$$

is a vector space isomorphism, i.e., show that it maps U bijectively onto V/W.

4. Let U, V and W be vector spaces and let $A : V \to W$ and $B : U \to V$ be linear mappings. Show that $(AB)^* = B^* A^*$.

5. Let $V = \mathbb{R}^2$ and let W be the x_1-axis, i.e., the one-dimensional subspace

$$\{(x_1, 0)\,;\, x_1 \in \mathbb{R}\}$$

of $\mathbb{R}^2$.

(a) Show that the W-cosets are the lines, $x_2 = a$, parallel to the x_1-axis.

(b) Show that the sum of the cosets, "$x_2 = a$" and "$x_2 = b$" is the coset "$x_2 = a + b$".

(c) Show that the scalar multiple of the coset, "$x_2 = c$" by the number, λ, is the coset, "$x_2 = \lambda c$".

6. (a) Let $(V^*)^*$ be the dual of the vector space, V^*. For every $v \in V$, let $\mu_v : V^* \to \mathbb{R}$ be the function, $\mu_v(\ell) = \ell(v)$. Show that the μ_v is a linear function on V^*, i.e., an element of $(V^*)^*$, and show that the map

$$\mu : V \to (V^*)^* \quad v \to \mu_v \tag{1.2.8}$$

is a linear map of V into $(V^*)^*$.

(b) Show that the map (1.2.8) is bijective. (*Hint:* $\dim (V^*)^* = \dim V^* = \dim V$, so by (1.1.5) it suffices to show that (1.2.8) is injective.) Conclude that there is a *natural* identification of V with $(V^*)^*$, i.e., that V and $(V^*)^*$ are two descriptions of the same object.

7. Let W be a vector subspace of V and let

$$W^\perp = \{\ell \in V^*\,,\ \ell(w) = 0 \text{ if } w \in W\}\,.$$

Show that $W^\perp$ is a subspace of V^* and that its dimension is equal to $\dim V - \dim W$. (*Hint:* By exercise 1 we can choose a basis, $e_1, \ldots, e_n$ of V such that $e_1, \ldots e_k$ is a basis of W. Show that $e^*_{k+1}, \ldots, e^*_n$ is a basis of $W^\perp$.) $W^\perp$ is called the *annihilator* of W in V^*.

8. Let V and V' be vector spaces and $A : V \to V'$ a linear map. Show that if W is the kernel of A there exists a linear map, $B : V/W \to V'$, with the property: $A = B \circ \pi$, π being the map (1.2.3). In addition show that this linear map is injective.

9. Let W be a subspace of a finite-dimensional vector space, V. From the inclusion map, $\iota : W^\perp \to V^*$, one gets a transpose map,

$$\iota^* : (V^*)^* \to (W^\perp)^*$$

and, by composing this with (1.2.8), a map

$$\iota^* \circ \mu : V \to (W^\perp)^*\,.$$

Show that this map is onto and that its kernel is W. Conclude from exercise 8 that there is a *natural* bijective linear map

$$\nu : V/W \to (W^\perp)^*$$

with the property $\nu \circ \pi = \iota^* \circ \mu$. In other words, V/W and $(W^\perp)^*$ are two descriptions of the same object. (This shows that the "quotient space" operation and the "dual space" operation are closely related.)

10. Let V_1 and V_2 be vector spaces and $A : V_1 \to V_2$ a linear map. Verify that for the transpose map: $A^* : V_2^* \to V_1^*$

$$\operatorname{Ker} A^* \quad = \quad (\operatorname{Im} A)^\perp$$

and

$$\operatorname{Im} A^* \quad = \quad (\operatorname{Ker} A)^\perp\,.$$

11. (a) Let $B : V \times V \to \mathbb{R}$ be an inner product on V. For $v \in V$ let

$$\ell_v : V \to \mathbb{R}$$

be the function: $\ell_v(w) = B(v, w)$. Show that ℓ_v is linear and show that the map

$$L : V \to V^* , \quad v \to \ell_v \tag{1.2.9}$$

is a linear mapping.

(b) Prove that this mapping is bijective. (*Hint:* Since $\dim V = \dim V^*$ it suffices by (1.1.5) to show that its kernel is zero. Now note that if $v \neq 0$ $\ell_v(v) = B(v, v)$ is a positive number.) Conclude that if V has an inner product one gets from it a *natural* identification of V with V^*.

12. Let V be an n-dimensional vector space and $B : V \times V \to \mathbb{R}$ an inner product on V. A basis, $e_1, \ldots, e_n$ of V is *orthonormal* is

$$B(e_i, e_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \tag{1.2.10}$$

(a) Show that an orthonormal basis exists. *Hint:* By induction let e_i, $i = 1, \ldots, k$ be vectors with the property (1.2.10) and let v be a vector which is not a linear combination of these vectors. Show that the vector

$$w = v - \sum B(e_i, v) e_i$$

is non-zero and is orthogonal to the e_i's. Now let $e_{k+1} = \lambda w$, where $\lambda = B(w, w)^{-\frac{1}{2}}$.

(b) Let $e_1, \ldots e_n$ and $e'_1, \ldots e'_n$ be two orthogonal bases of V and let

$$e'_j = \sum a_{i,j} e_i \, . \tag{1.2.11}$$

Show that

$$\sum a_{i,j} a_{i,k} = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases} \tag{1.2.12}$$

(c) Let A be the matrix $[a_{i,j}]$. Show that (1.2.12) can be written more compactly as the matrix identity

$$AA^t = I \tag{1.2.13}$$

where I is the identity matrix.

(d) Let $e_1, \dots, e_n$ be an orthonormal basis of V and $e_1^*, \dots, e_n^*$ the dual basis of V^*. Show that the mapping (1.2.9) is the mapping, $Le_i = e_i^*$, $i = 1, \dots n$.

1.3 Tensors

Let V be an n-dimensional vector space and let V^k be the set of all k-tuples, $(v_1, \dots, v_k)$, $v_i \in V$. A function

$$T : V^k \to \mathbb{R}$$

is said to be linear in its i^{th} variable if, when we fix vectors, $v_1, \dots, v_{i-1}$, $v_{i+1}, \dots, v_k$, the map

$$v \in V \to T(v_1, \dots, v_{i-1}, v, v_{i+1}, \dots, v_k) \tag{1.3.1}$$

is linear in V. If T is linear in its i^{th} variable for $i = 1, \dots, k$ it is said to be *k-linear*, or alternatively is said to be a *k-tensor*. We denote the set of all k-tensors by $\mathcal{L}^k(V)$. We will agree that 0-tensors are just the real numbers, that is $\mathcal{L}^0(V) = \mathbb{R}$.

Let T_1 and T_2 be functions on V^k. It is clear from (1.3.1) that if T_1 and T_2 are k-linear, so is $T_1 + T_2$. Similarly if T is k-linear and λ is a real number, λT is k-linear. Hence $\mathcal{L}^k(V)$ is a vector space. Note that for $k = 1$, "k-linear" just means "linear", so $\mathcal{L}^1(V) = V^*$.

Let $I = (i_1, \dots i_k)$ be a sequence of integers with $1 \le i_r \le n$, $r = 1, \dots, k$. We will call such a sequence *a multi-index* of length k. For instance the multi-indices of length 2 are the square arrays of pairs of integers

$$(i, j)\,,\ 1 \le i, j \le n$$

and there are exactly n^2 of them.

Exercise.

Show that there are exactly n^k multi-indices of length k.

Now fix a basis, $e_1, \dots, e_n$, of V and for $T \in \mathcal{L}^k(V)$ let

$$T_I = T(e_{i_1}, \dots, e_{i_k}) \tag{1.3.2}$$

for every multi-index I of length k.

Proposition 1.3.1. *The T_I's determine T, i.e., if T and T' are k-tensors and $T_I = T'_I$ for all I, then $T = T'$.*

Proof. By induction on n. For $n = 1$ we proved this result in § 1.1. Let's prove that if this assertion is true for $n-1$, it's true for n. For each e_i let T_i be the $(k-1)$-tensor

$$(v_1, \dots, v_{n-1}) \to T(v_1, \dots, v_{n-1}, e_i)\,.$$

Then for $v = c_1 e_1 + \cdots c_n e_n$

$$T(v_1, \dots, v_{n-1}, v) = \sum c_i T_i(v_1, \dots, v_{n-1})\,,$$

so the T_i's determine T. Now apply induction.

$\square$

The tensor product operation

If T_1 is a k-tensor and T_2 is an ℓ-tensor, one can define a $k+\ell$-tensor, $T_1 \otimes T_2$, by setting

$$(T_1 \otimes T_2)(v_1, \dots, v_{k+\ell}) = T_1(v_1, \dots, v_k) T_2(v_{k+1}, \dots, v_{k+\ell})\,.$$

This tensor is called *the tensor product* of T_1 and T_2. We note that if T_1 or T_2 is a 0-tensor, i.e., scalar, then tensor product with *it* is just scalar multiplication by *it*, that is $a \otimes T = T \otimes a = aT$ $(a \in \mathbb{R}\,,\ T \in \mathcal{L}^k(V))$.

Similarly, given a k-tensor, T_1, an ℓ-tensor, T_2 and an m-tensor, T_3, one can define a $(k+\ell+m)$-tensor, $T_1 \otimes T_2 \otimes T_3$ by setting

$$\begin{aligned} &T_1 \otimes T_2 \otimes T_3(v_1, \dots, v_{k+\ell+m}) \\ &\quad = T_1(v_1, \dots, v_k) T_2(v_{k+1}, \dots, v_{k+\ell}) T_3(v_{k+\ell+1}, \dots, v_{k+\ell+m})\,. \end{aligned} \tag{1.3.3}$$

Alternatively, one can define (1.3.3) by defining it to be the tensor product of $T_1 \otimes T_2$ and T_3 or the tensor product of T_1 and $T_2 \otimes T_3$. It's easy to see that both these tensor products are identical with (1.3.3):

$$(T_1 \otimes T_2) \otimes T_3 = T_1 \otimes (T_2 \otimes T_3) = T_1 \otimes T_2 \otimes T_3\,. \tag{1.3.4}$$

We leave for you to check that if λ is a real number

$$\lambda(T_1 \otimes T_2) = (\lambda T_1) \otimes T_2 = T_1 \otimes (\lambda T_2) \tag{1.3.5}$$

and that the left and right distributive laws are valid: For $k_1 = k_2$,

$$(T_1 + T_2) \otimes T_3 = T_1 \otimes T_3 + T_2 \otimes T_3 \tag{1.3.6}$$

and for $k_2 = k_3$

$$T_1 \otimes (T_2 + T_3) = T_1 \otimes T_2 + T_1 \otimes T_3 \,. \tag{1.3.7}$$

A particularly interesting tensor product is the following. For $i = 1, \dots, k$ let $\ell_i \in V^*$ and let

$$T = \ell_1 \otimes \cdots \otimes \ell_k \,. \tag{1.3.8}$$

Thus, by definition,

$$T(v_1, \dots, v_k) = \ell_1(v_1) \dots \ell_k(v_k) \,. \tag{1.3.9}$$

A tensor of the form (1.3.9) is called a *decomposable* k-tensor. These tensors, as we will see, play an important role in what follows. In particular, let $e_1, \dots, e_n$ be a basis of V and $e_1^*, \dots, e_n^*$ the dual basis of V^*. For every multi-index, I, of length k let

$$e_I^* = e_{i_1}^* \otimes \cdots \otimes e_{i_k}^* \,.$$

Then if J is another multi-index of length k,

$$e_I^*(e_{j_1}, \dots, e_{j_k}) = \begin{cases} 1\,, & I = J \\ 0\,, & I \neq J \end{cases} \tag{1.3.10}$$

by (1.2.6), (1.3.8) and (1.3.9). From (1.3.10) it's easy to conclude

Theorem 1.3.2. *The e_I^*'s are a basis of $\mathcal{L}^k(V)$.*

Proof. Given $T \in \mathcal{L}^k(V)$, let

$$T' = \sum T_I e_I^*$$

where the T_I's are defined by (1.3.2). Then

$$T'(e_{j_1}, \dots, e_{j_k}) = \sum T_I e_I^*(e_{j_1}, \dots, e_{j_k}) = T_J \tag{1.3.11}$$

by (1.3.10); however, by Proposition 1.3.1 the T_J's determine T, so $T' = T$. This proves that the e_I^*'s are a spanning set of vectors for $\mathcal{L}^k(V)$. To prove they're a basis, suppose

$$\sum C_I e_I^* = 0$$

for constants, $C_I \in \mathbb{R}$. Then by (1.3.11) with $T' = 0$, $C_J = 0$, so the e_I^*'s are linearly independent.

□

As we noted above there are exactly n^k multi-indices of length k and hence n^k basis vectors in the set, $\{e_I^*\}$, so we've proved

Corollary. $\dim \mathcal{L}^k(V) = n^k$.

The pull-back operation

Let V and W be finite dimensional vector spaces and let $A : V \to W$ be a linear mapping. If $T \in \mathcal{L}^k(W)$, we define

$$A^*T : V^k \to \mathbb{R}$$

to be the function

$$A^*T(v_1, \ldots, v_k) = T(Av_1, \ldots, Av_k)\,. \tag{1.3.12}$$

It's clear from the linearity of A that this function is linear in its i^{th} variable for all i, and hence is k-tensor. We will call A^*T the *pull-back* of T by the map, A.

Proposition 1.3.3. *The map*

$$A^* : \mathcal{L}^k(W) \to \mathcal{L}^k(V)\,, \quad T \to A^*T\,, \tag{1.3.13}$$

is a linear mapping.

We leave this as an exercise. We also leave as an exercise the identity

$$A^*(T_1 \otimes T_2) = A^*T_1 \otimes A^*T_2 \tag{1.3.14}$$

for $T_1 \in \mathcal{L}^k(W)$ and $T_2 \in \mathcal{L}^m(W)$. Also, if U is a vector space and $B : U \to V$ a linear mapping, we leave for you to check that

$$(AB)^*T = B^*(A^*T) \tag{1.3.15}$$

for all $T \in \mathcal{L}^k(W)$.

Exercises.

1. Verify that there are exactly n^k multi-indices of length k.

2. Prove Proposition 1.3.3.

3. Verify (1.3.14).

4. Verify (1.3.15).

5. Let $A : V \to W$ be a linear map. Show that if ℓ_i, $i = 1, \dots, k$ are elements of W^*

$$A^*(\ell_1 \otimes \cdots \otimes \ell_k) = A^*\ell_1 \otimes \cdots \otimes A^*\ell_k\,.$$

Conclude that A^* maps decomposable k-tensors to decomposable k-tensors.

6. Let V be an n-dimensional vector space and ℓ_i, $i = 1, 2$, elements of V^*. Show that $\ell_1 \otimes \ell_2 = \ell_2 \otimes \ell_1$ if and only if ℓ_1 and ℓ_2 are linearly dependent. (*Hint:* Show that if ℓ_1 and ℓ_2 are linearly independent there exist vectors, v_i, $i =, 1, 2$ in V with property

$$\ell_i(v_j) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}\,.$$

Now compare $(\ell_1 \otimes \ell_2)(v_1, v_2)$ and $(\ell_2 \otimes \ell_1)(v_1, v_2)$.) Conclude that if $\dim V \geq 2$ the tensor product operation isn't commutative, i.e., it's usually not true that $\ell_1 \otimes \ell_2 = \ell_2 \otimes \ell_1$.

7. Let T be a k-tensor and v a vector. Define $T_v : V^{k-1} \to \mathbb{R}$ to be the map

$$T_v(v_1, \dots, v_{k-1}) = T(v, v_1, \dots, v_{k-1})\,. \tag{1.3.16}$$

Show that T_v is a $(k-1)$-tensor.

8. Show that if T_1 is an r-tensor and T_2 is an s-tensor, then if $r > 0$,

$$(T_1 \otimes T_2)_v = (T_1)_v \otimes T_2\,.$$

9. Let $A : V \to W$ be a linear map mapping $v \in V$ to $w \in W$. Show that for $T \in \mathcal{L}^k(W)$, $A^*(T_w) = (A^*T)_v$.

1.4 Alternating k-tensors

We will discuss in this section a class of k-tensors which play an important role in multivariable calculus. In this discussion we will need some standard facts about the "permutation group". For those of you who are already familiar with this object (and I suspect most of you are) you can regard the paragraph below as a chance to re-familiarize yourselves with these facts.

Permutations

Let $\sum_k$ be the k-element set: $\{1, 2, \ldots, k\}$. A *permutation of order k* is a bijective map, $\sigma : \sum_k \to \sum_k$. Given two permutations, σ_1 and σ_2, their *product*, $\sigma_1\sigma_2$, is the composition of σ_1 and σ_2, i.e., the map,

$$i \to \sigma_1(\sigma_2(i)),$$

and for every permutation, σ, one denotes by σ^{-1} the inverse permutation:

$$\sigma(i) = j \Leftrightarrow \sigma^{-1}(j) = i.$$

Let S_k be the set of all permutations of order k. One calls S_k the *permutation group* of $\sum_k$ or, alternatively, the *symmetric group on k letters.*

Check:

There are $k!$ elements in S_k.

For every $1 \le i < j \le k$, let $\tau = \tau_{i,j}$ be the permutation

$$\begin{aligned} \tau(i) &= j \\ \tau(j) &= i \\ \tau(\ell) &= \ell, \quad \ell \neq i, j. \end{aligned} \tag{1.4.1}$$

τ is called a *transposition*, and if $j = i + 1$, τ is called an *elementary transposition.*

Theorem 1.4.1. *Every permutation can be written as a product of finite number of transpositions.*

Proof. Induction on k: "$k = 2$" is obvious. *The induction step: "$k-1$" implies "k":* Given $\sigma \in S_k$, $\sigma(k) = i \Leftrightarrow \tau_{ik}\sigma(k) = k$. Thus $\tau_{ik}\sigma$ is, in effect, a permutation of $\sum_{k-1}$. By induction, $\tau_{ik}\sigma$ can be written as a product of transpositions, so

$$\sigma = \tau_{ik}(\tau_{ik}\sigma)$$

can be written as a product of transpositions.

□

Theorem 1.4.2. *Every transposition can be written as a product of elementary transpositions.*

Proof. Let $\tau = \tau_{ij}$, $i < j$. With i fixed, argue by induction on j. Note that for $j > i + 1$

$$\tau_{ij} = \tau_{j-1,j}\tau_{i,j-1}\tau_{j-1,j} \,.$$

Now apply induction to $\tau_{i,j-1}$.

□

Corollary. *Every permutation can be written as a product of elementary transpositions.*

The sign of a permutation

Let $x_1, \ldots, x_k$ be the coordinate functions on $\mathbb{R}^k$. For $\sigma \in S_k$ we define

$$(1.4.2) \qquad (-1)^\sigma = \prod_{i<j} \frac{x_{\sigma(i)} - x_{\sigma(j)}}{x_i - x_j} \,.$$

Notice that the numerator and denominator in this expression are identical up to sign. Indeed, if $p = \sigma(i) < \sigma(j) = q$, the term, $x_p - x_q$ occurs once and just once in the numerator and one and just one in the denominator; and if $q = \sigma(i) > \sigma(j) = p$, the term, $x_p - x_q$, occurs once and just once in the numerator and its negative, $x_q - x_p$, once and just once in the numerator. Thus

$$(1.4.3) \qquad (-1)^\sigma = \pm 1 \,.$$

Claim:

For $\sigma, \tau \in S_k$

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

Proof. By definition,

$$(-1)^{\sigma\tau} = \prod_{i<j} \frac{x_{\sigma\tau(i)} - x_{\sigma\tau(j)}}{x_i - x_j} .$$

We write the right hand side as a product of

$$\prod_{i<j} \frac{x_{\tau(i)} - x_{\tau(j)}}{x_i - x_j} = (-1)^{\tau} \tag{1.4.5}$$

and

$$\prod_{i<j} \frac{x_{\sigma\tau(i)} - x_{\sigma\tau(j)}}{x_{\tau(i)} - x_{\tau(j)}} \tag{1.4.6}$$

For $i < j$, let $p = \tau(i)$ and $q = \tau(j)$ when $\tau(i) < \tau(j)$ and let $p = \tau(j)$ and $q = \tau(i)$ when $\tau(j) < \tau(i)$. Then

$$\frac{x_{\sigma\tau(i)} - x_{\sigma\tau(j)}}{x_{\tau(i)} - x_{\tau(j)}} = \frac{x_{\sigma(p)} - x_{\sigma(q)}}{x_p - x_q}$$

(i.e., if $\tau(i) < \tau(j)$, the numerator and denominator on the right *equal* the numerator and denominator on the left and, if $\tau(j) < \tau(i)$ are *negatives* of the numerator and denominator on the left). Thus (1.4.6) becomes

$$\prod_{p<q} \frac{x_{\sigma(p)} - x_{\sigma(q)}}{x_p - x_q} = (-1)^{\sigma} .$$

$\square$

We'll leave for you to check that if τ is a transposition, $(-1)^{\tau} = -1$ and to conclude from this:

Proposition 1.4.3. *If σ is the product of an odd number of transpositions, $(-1)^{\sigma} = -1$ and if σ is the product of an even number of transpositions $(-1)^{\sigma} = +1$.*

Alternation

Let V be an n-dimensional vector space and $T \in \mathcal{L}^*(v)$ a k-tensor. If $\sigma \in S_k$, let $T^\sigma \in \mathcal{L}^*(V)$ be the k-tensor

$$T^\sigma(v_1, \dots, v_k) = T(v_{\sigma^{-1}(1)}, \dots, v_{\sigma^{-1}(k)}) . \tag{1.4.7}$$

Proposition 1.4.4. *1. If $T = \ell_1 \otimes \cdots \otimes \ell_k$, $\ell_i \in V^*$, then $T^\sigma = \ell_{\sigma(1)} \otimes \cdots \otimes \ell_{\sigma(k)}$.*

2. The map, $T \in \mathcal{L}^k(V) \to T^\sigma \in \mathcal{L}^k(V)$ is a linear map.

3. $T^{\sigma\tau} = (T^\tau)^\sigma$.

Proof. To prove 1, we note that by (1.4.7)

$$\begin{aligned} &(\ell_1 \otimes \cdots \otimes \ell_k)^\sigma(v_1, \dots, v_k) \\ &= \ell_1(v_{\sigma^{-1}(1)}) \cdots \ell_k(v_{\sigma^{-1}(k)}) . \end{aligned}$$

Setting $\sigma^{-1}(i) = q$, the i^{th} term in this product is $\ell_{\sigma(q)}(v_q)$; so the product can be rewritten as

$$\ell_{\sigma(1)}(v_1) \dots \ell_{\sigma(k)}(v_k)$$

or

$$(\ell_{\sigma(1)} \otimes \cdots \otimes \ell_{\sigma(k)})(v_1, \dots, v_k) .$$

The proof of 2 we'll leave as an exercise.

Proof of 3: By item 2, it suffices to check 3 for decomposable tensors. However, by 1

$$\begin{aligned} (\ell_1 \otimes \cdots \otimes \ell_k)^{\sigma\tau} &= \ell_{\sigma\tau(1)} \otimes \cdots \otimes \ell_{\sigma\tau(k)} \\ &= (\ell_{\tau(1)} \otimes \cdots \otimes \ell_{\tau(k)})^\sigma \\ &= ((\ell_1 \otimes \cdots \otimes \ell)^\tau)^\sigma . \end{aligned}$$

Definition 1.4.5. *$T \in \mathcal{L}^k(V)$ is* alternating *if $T^\sigma = (-1)^\sigma T$ for all $\sigma \in S_k$.*

We will denote by $\mathcal{A}^k(V)$ the set of all alternating k-tensors in $\mathcal{L}^k(V)$. By item 2 of Proposition 1.4.4 this set is a vector subspace of $\mathcal{L}^k(V)$.

It is not easy to write down simple examples of alternating k-tensors; however, there is a method, called the *alternation operation*, for constructing such tensors: Given $T \in \mathcal{L}^*(V)$ let

$$\operatorname{Alt} T = \sum_{\tau \in S_k} (-1)^\tau T^\tau \,. \tag{1.4.8}$$

We claim

Proposition 1.4.6. *For $T \in \mathcal{L}^k(V)$ and $\sigma \in S_k$,*

1. $(\operatorname{Alt} T)^\sigma = (-1)^\sigma \operatorname{Alt} T$
2. *if* $T \in \mathcal{A}^k(V)$, $\operatorname{Alt} T = k!T$.
3. $\operatorname{Alt} T^\sigma = (\operatorname{Alt} T)^\sigma$
4. *the map*

$$\operatorname{Alt} : \mathcal{L}^k(V) \to \mathcal{L}^k(V)\,,\; T \to \operatorname{Alt}(T)$$

is linear.

Proof. To prove 1 we note that by Proposition (1.4.4):

$$\begin{aligned} (\operatorname{Alt} T)^\sigma &= \sum (-1)^\tau (T^{\sigma\tau}) \\ &= (-1)^\sigma \sum (-1)^{\sigma\tau} T^{\sigma\tau} \,. \end{aligned}$$

But as τ runs over S_k, $\sigma\tau$ runs over S_k, and hence the right hand side is $(-1)^\sigma \operatorname{Alt}(T)$.

□

Proof of 2. If $T \in \mathcal{A}^k$

$$\begin{aligned} \operatorname{Alt} T &= \sum (-1)^\tau T^\tau \\ &= \sum (-1)^\tau (-1)^\tau T \\ &= k!\, T \,. \end{aligned}$$

□

Proof of 3.

$$\begin{aligned} \operatorname{Alt} T^\sigma &= \sum (-1)^\tau T^{\tau\sigma} = (-1)^\sigma \sum (-1)^{\tau\sigma} T^{\tau\sigma} \\ &= (-1)^\sigma \operatorname{Alt} T = (\operatorname{Alt} T)^\sigma \,. \end{aligned}$$

□

Finally, item 4 is an easy corollary of item 2 of Proposition 1.4.4. □

We will use this alternation operation to construct a basis for $\mathcal{A}^k(V)$. First, however, we require some notation:

Let $I = (i_1, \ldots, i_k)$ be a multi-index of length k.

Definition 1.4.7. *1.* *I is* repeating *if $i_r = i_s$ for some $r \neq s$.*

2. *I is strictly increasing if $i_1 < i_2 < \cdots < i_r$.*

3. *For $\sigma \in S_k$, $I^\sigma = (i_{\sigma(1)}, \ldots, i_{\sigma(k)})$.*

Remark: If I is non-repeating there is a unique $\sigma \in S_k$ so that I^σ is strictly increasing.

Let $e_1, \ldots, e_n$ be a basis of V and let

$$e_I^* = e_{i_1}^* \otimes \cdots \otimes e_{i_k}^*$$

and

$$\psi_I = \mathrm{Alt}\,(e_I^*)\,.$$

Proposition 1.4.8. *1.* $\psi_{I^\sigma} = (-1)^\sigma \psi_I$.

2. *If I is repeating, $\psi_I = 0$.*

3. *If I and J are strictly increasing,*

$$\psi_I(e_{j_1}, \ldots, e_{j_k}) = \left\{ \begin{array}{ll} 1 & I = J \\ 0 & I \neq J \end{array} \right. .$$

Proof. To prove 1 we note that $(e_I^*)^\sigma = e_{I^\sigma}^*$; so

$$\mathrm{Alt}\,(e_{I^\sigma}^*) = \mathrm{Alt}\,(e_I^*)^\sigma = (-1)^\sigma \mathrm{Alt}\,(e_I^*)\,.$$

□

Proof of 2: Suppose $I = (i_1, \ldots, i_k)$ with $i_r = i_s$ for $r \neq s$. Then if $\tau = \tau_{i_r, i_s}$, $e_I^* = e_{I^r}^*$ so

$$\psi_I = \psi_{I^r} = (-1)^\tau \psi_I = -\psi_I\,.$$

□

Proof of 3: By definition

$$\psi_I(e_{j_1},\ldots,e_{j_k}) = \sum(-1)^\tau e^*_{I^\tau}(e_{j_1},\ldots,e_{j_k})\,.$$

But by (1.3.10)

$$e^*_{I^\tau}(e_{j_1},\ldots,e_{j_k}) = \begin{cases} 1 \text{ if } I^\tau = J \\ 0 \text{ if } I^\tau \neq J \end{cases}\,. \tag{1.4.9}$$

Thus if I and J are strictly increasing, I^τ is strictly increasing if and only if $I^\tau = I$, and (1.4.9) is non-zero if and only if $I = J$.

□

Now let T be in $\mathcal{A}^k$. By Proposition 1.3.2,

$$T = \sum a_J e^*_J\,, \quad a_J \in \mathbb{R}\,.$$

Since

$$\begin{aligned} k!T &= \operatorname{Alt}(T) \\ T &= \frac{1}{k!}\sum a_J \operatorname{Alt}(e^*_J) = \sum b_J \psi_J\,. \end{aligned}$$

We can discard all repeating terms in this sum since they are zero; and for every non-repeating term, J, we can write $J = I^\sigma$, where I is strictly increasing, and hence $\psi_J = (-1)^\sigma \psi_I$.

Conclusion:

We can write T as a sum

$$T = \sum c_I \psi_I\,, \tag{1.4.10}$$

with I's strictly increasing.

Claim.

The c_I's are unique.

Proof. For J strictly increasing

$$T(e_{j_1}, \dots, e_{j_k}) = \sum c_I \psi_I(e_{j_1}, \dots, e_{j_k}) = c_J \,. \tag{1.4.11}$$

By (1.4.10) the ψ_I's, I strictly increasing, are a spanning set of vectors for $\mathcal{A}^k(V)$, and by (1.4.11) they are linearly independent, so we've proved

Proposition 1.4.9. *The alternating tensors, ψ_I, I strictly increasing, are a basis for $\mathcal{A}^k(V)$.*

Thus $\dim \mathcal{A}^k(V)$ is equal to the number of strictly increasing multi-indices, I, of length k. We leave for you as an exercise to show that this number is equal to

$$\binom{n}{k} = \frac{n!}{(n-k)!k!} = \text{“ } n \text{ choose } k\text{”} \tag{1.4.12}$$

if $1 \le k \le n$.

□

Hint: Show that every strictly increasing multi-index of length k determines a k element subset of $\{1, \dots, n\}$ and vice-versa.

Note also that if $k > n$ every multi-index

$$I = (i_1, \dots, i_k)$$

of length k has to be repeating: $i_r = i_s$ for some $r \neq s$ since the i_p's lie on the interval $1 \le i \le n$. Thus by Proposition 1.4.6

$$\psi_I = 0$$

for all multi-indices of length $k > 0$ and

$$\mathcal{A}^k = \{0\} \,. \tag{1.4.13}$$

Exercises.

1. Show that there are exactly $k!$ permutations of order k. *Hint:* Induction on k: Let $\sigma \in S_k$, and let $\sigma(k) = i$, $1 \le i \le k$. Show that $\tau_{ik}\sigma$ leaves k fixed and hence is, in effect, a permutation of $\sum_{k-1}$.

2. Prove that if $\tau \in S_k$ is a transposition, $(-1)^\tau = -1$ and deduce from this Proposition 1.4.3.

3. Prove assertion 2 in Proposition 1.4.4.

4. Prove that $\dim \mathcal{A}^k(V)$ is given by (1.4.12).

5. Verify that for $i < j-1$

$$\tau_{i,j} = \tau_{j-1,j}\tau_{i,j-1}, \tau_{j-1,j} \,.$$

6. For $k = 3$ show that every one of the six elements of S_3 is either a transposition or can be written as a product of two transpositions.

7. Let $\sigma \in S_k$ be the "cyclic" permutation

$$\sigma(i) = i+1\,, \quad i = 1, \dots, k-1$$

and $\sigma(k) = 1$. Show explicitly how to write σ as a product of transpositions and compute $(-1)^\sigma$. *Hint:* Same hint as in exercise 1.

8. In exercise 7 of Section 3 show that if T is in $\mathcal{A}^k$, T_v is in $\mathcal{A}^{k-1}$. Show in addition that for $v, w \in V$ and $T \in \mathcal{A}^k$, $(T_v)_w = -(T_w)_v$.

9. Let $A : V \to W$ be a linear mapping. Show that if T is in $\mathcal{A}^k(W)$, A^*T is in $\mathcal{A}^k(V)$.

10. In exercise 9 show that if T is in $\mathcal{L}^k(W)$, $\mathrm{Alt}\,(A^*T) = A^*(\mathrm{Alt}\,(T))$, i.e., show that the "Alt" operation commutes with the pull-back operation.

1.5 The space, $\Lambda^k(V^*)$

In § 1.4 we showed that the image of the alternation operation, $\mathrm{Alt} : \mathcal{L}^k(V) \to \mathcal{L}^k(V)$ is $\mathcal{A}^k(V)$. In this section we will compute the kernel of Alt.

Definition 1.5.1. *A decomposable k-tensor $\ell_1 \otimes \cdots \otimes \ell_k$, $\ell_i \in V^*$, is redundant if for some index, i, $\ell_i = \ell_{i+1}$.*

Let $\mathcal{I}^k$ be the linear span of the set of reductant k-tensors.

Note that for $k = 1$ the notion of redundant doesn't really make sense; a single vector $\ell \in \mathcal{L}^1(V^*)$ can't be "redundant" so we decree

$$\mathcal{I}^1(V) = \{0\}\,.$$

Proposition 1.5.2. *If $T \in \mathcal{I}^k$, $\mathrm{Alt}\,(T) = 0$.*

Proof. Let $T = \ell_k \otimes \cdots \otimes \ell_k$ with $\ell_i = \ell_{i+1}$. Then if $\tau = \tau_{i,i+1}$, $T^\tau = T$ and $(-1)^\tau = -1$. Hence $\mathrm{Alt}\,(T) = \mathrm{Alt}\,(T^\tau) = \mathrm{Alt}\,(T)^\tau = -\mathrm{Alt}\,(T)$; so $\mathrm{Alt}\,(T) = 0$. □

To simplify notation let's abbreviate $\mathcal{L}^k(V)$, $\mathcal{A}^k(V)$ and $\mathcal{I}^k(V)$ to $\mathcal{L}^k$, $\mathcal{A}^k$ and $\mathcal{I}^k$.

Proposition 1.5.3. *If $T \in \mathcal{I}^r$ and $T' \in \mathcal{L}^s$ then $T \otimes T'$ and $T' \otimes T$ are in $\mathcal{I}^{r+s}$.*

Proof. We can assume that T and T' are decomposable, i.e., $T = \ell_1 \otimes \cdots \otimes \ell_r$ and $T' = \ell_1' \otimes \cdots \otimes \ell_s'$ and that T is redundant: $\ell_i = \ell_{i+1}$. Then

$$T \otimes T' = \ell_1 \otimes \cdots \ell_{i-1} \otimes \ell_i \otimes \ell_i \otimes \cdots \ell_r \otimes \ell_1' \otimes \cdots \otimes \ell_s'$$

is redundant and hence in $\mathcal{I}^{r+s}$. The argument for $T' \otimes T$ is similar. □

Proposition 1.5.4. *If $T \in \mathcal{L}^k$ and $\sigma \in S_k$, then*

$$T^\sigma = (-1)^\sigma T + S \tag{1.5.1}$$

where S is in $\mathcal{I}^k$.

Proof. We can assume T is decomposable, i.e., $T = \ell_1 \otimes \cdots \otimes \ell_k$. Let's first look at the simplest possible case: $k = 2$ and $\sigma = \tau_{1,2}$. Then

$$\begin{aligned} T^\sigma - (-)^\sigma T &= \ell_1 \otimes \ell_2 + \ell_2 \otimes \ell_1 \\ &= ((\ell_1 + \ell_2) \otimes (\ell_1 + \ell_2) - \ell_1 \otimes \ell_1 - \ell_2 \otimes \ell_2)/2\,, \end{aligned}$$

and the terms on the right are redundant, and hence in $\mathcal{I}^2$. Next let k be arbitrary and $\sigma = \tau_{i,i+1}$. If $T_1 = \ell_1 \otimes \cdots \otimes \ell_{i-2}$ and $T_2 = \ell_{i+2} \otimes \cdots \otimes \ell_k$. Then

$$T - (-1)^\sigma T = T_1 \otimes (\ell_i \otimes \ell_{i+1} + \ell_{i+1} \otimes \ell_i) \otimes T_2$$

is in $\mathcal{I}^k$ by Proposition 1.5.3 and the computation above.
The general case: By Theorem 1.4.2, σ can be written as a product of m elementary transpositions, and we'll prove (1.5.1) by induction on m.

We've just dealt with the case $m = 1$.
The induction step: "$m-1$" implies "m". Let $\sigma = \tau\beta$ where β is a product of $m-1$ elementary transpositions and τ is an elementary transposition. Then

$$\begin{aligned} T^\sigma = (T^\beta)^\tau &= (-1)^\tau T^\beta + \cdots \\ &= (-1)^\tau (-1)^\beta T + \cdots \\ &= (-1)^\sigma T + \cdots \end{aligned}$$

where the "dots" are elements of $\mathcal{I}^k$, and the induction hypothesis was used in line 2.

□

Corollary. *If $T \in \mathcal{L}^k$, the*

(1.5.2) $$\mathrm{Alt}\,(T) = k!T + W\,,$$

where W is in $\mathcal{I}^k$.

Proof. By definition $\mathrm{Alt}\,(T) = \sum (-1)^\sigma T^\sigma$, and by Proposition 1.5.4, $T^\sigma = (-1)^\sigma T + W_\sigma$, with $W_\sigma \in \mathcal{I}^k$. Thus

$$\begin{aligned} \mathrm{Alt}\,(T) &= \sum (-1)^\sigma (-1)^\sigma T + \sum (-1)^\sigma W_\sigma \\ &= k!T + W \end{aligned}$$

where $W = \sum (-1)^\sigma W_\sigma$.

□

Corollary. *$\mathcal{I}^k$ is the kernel of* Alt *.*

Proof. We've already proved that if $T \in \mathcal{I}^k$, $\mathrm{Alt}\,(T) = 0$. To prove the converse assertion we note that if $\mathrm{Alt}\,(T) = 0$, then by (1.5.2)

$$T = -\tfrac{1}{k!} W \,.$$

with $W \in \mathcal{I}^k$. □

Putting these results together we conclude:

Theorem 1.5.5. *Every element, T, of $\mathcal{L}^k$ can be written uniquely as a sum, $T = T_1 + T_2$ where $T_1 \in \mathcal{A}^k$ and $T_2 \in \mathcal{I}^k$.*

Proof. By (1.5.2), $T = T_1 + T_2$ with

$$T_1 \quad = \quad \tfrac{1}{k!}\mathrm{Alt}\,(T)$$

and

$$T_2 \quad = \quad -\tfrac{1}{k!} W \,.$$

To prove that this decomposition is unique, suppose $T_1 + T_2 = 0$, with $T_1 \in \mathcal{A}^k$ and $T_2 \in \mathcal{I}^k$. Then

$$0 = \mathrm{Alt}\,(T_1 + T_2) = k! T_1$$

so $T_1 = 0$, and hence $T_2 = 0$.

□

Let

$$\Lambda^k(V^*) = \mathcal{L}^k(V^*)/\mathcal{I}^k(V^*)\,, \tag{1.5.3}$$

i.e., let $\Lambda^k = \Lambda^k(V^*)$ be the quotient of the vector space $\mathcal{L}^k$ by the subspace, $\mathcal{I}^k$, of $\mathcal{L}^k$. By (1.2.3) one has a linear map:

$$\pi : \mathcal{L}^k \to \Lambda^k \,, \quad T \to T + \mathcal{I}^k \tag{1.5.4}$$

which is onto and has $\mathcal{I}^k$ as kernel. We claim:

Theorem 1.5.6. *The map, π, maps $\mathcal{A}^k$ bijectively onto Λ^k.*

Proof. By Theorem 1.5.5 every $\mathcal{I}^k$ coset, $T + \mathcal{I}^k$, contains a unique element, T_1, of $\mathcal{A}^k$. Hence for every element of Λ^k there is a unique element of $\mathcal{A}^k$ which gets mapped onto it by π.

□

Remark. Since Λ^k and $\mathcal{A}^k$ are isomorphic as vector spaces many treatments of multilinear algebra avoid mentioning Λ^k, reasoning that $\mathcal{A}^k$ is a perfectly good substitute for it and that one should, if possible, not make two different definitions for what is essentially the same object. This is a justifiable point of view (and is the point of view taken by Spivak and Munkres[1]). There are, however, some advantages to distinguishing between A^k and Λ^k, as we'll see in § 1.6.

Exercises.

1. A k-tensor, T, $\in \mathcal{L}^k(V)$ is *symmetric* if $T^\sigma = T$ for all $\sigma \in S_k$. Show that the set, $\mathcal{S}^k(V)$, of symmetric k tensors is a vector subspace of $\mathcal{L}^k(V)$.

2. Let $e_1, \ldots, e_n$ be a basis of V. Show that every symmetric 2-tensor is of the form

$$\sum a_{ij} e_i^* \otimes e_j^*$$

where $a_{i,j} = a_{j,i}$ and $e_1^*, \ldots, e_n^*$ are the dual basis vectors of V^*.

3. Show that if T is a symmetric k-tensor, then for $k \geq 2$, T is in $\mathcal{I}^k$. *Hint:* Let σ be a transposition and deduce from the identity, $T^\sigma = T$, that T has to be in the kernel of Alt.

4. *Warning:* In general $\mathcal{S}^k(V) \neq \mathcal{I}^k(V)$. Show, however, that if $k = 2$ these two spaces are equal.

5. Show that if $\ell \in V^*$ and $T \in \mathcal{I}^{k-2}$, then $\ell \otimes T \otimes \ell$ is in $\mathcal{I}^k$.

6. Show that if ℓ_1 and ℓ_2 are in V^* and T is in $\mathcal{I}^{k-2}$, then $\ell_1 \otimes T \otimes \ell_2 + \ell_2 \otimes T \otimes \ell_1$ is in $\mathcal{I}^k$.

7. Given a permutation $\sigma \in S_k$ and $T \in \mathcal{I}^k$, show that $T^\sigma \in \mathcal{I}^k$.

8. Let $\mathcal{W}$ be a subspace of $\mathcal{L}^k$ having the following two properties.

(a) For $S \in \mathcal{S}^2(V)$ and $T \in \mathcal{L}^{k-2}$, $S \otimes T$ is in $\mathcal{W}$.

(b) For T in $\mathcal{W}$ and $\sigma \in S_k$, T^σ is in $\mathcal{W}$.

---

[1] and by the author of these notes in his book with Alan Pollack, "Differential Topology"

Show that $\mathcal{W}$ has to contain $\mathcal{I}^k$ and conclude that $\mathcal{I}^k$ is the smallest subspace of $\mathcal{L}^k$ having properties a and b.

9. Show that there is a bijective linear map

$$\alpha : \Lambda^k \to \mathcal{A}^k$$

with the property

$$\alpha\pi(T) = \frac{1}{k!}\mathrm{Alt}\,(T) \tag{1.5.5}$$

for all $T \in \mathcal{L}^k$, and show that α is the inverse of the map of $\mathcal{A}^k$ onto Λ^k described in Theorem 1.5.6 (*Hint:* §1.2, exercise 8).

10. Let V be an n-dimensional vector space. Compute the dimension of $S^k(V)$. *Some hints:*

(a) Introduce the following symmetrization operation on tensors $T \in \mathcal{L}^k(V)$:

$$\mathrm{Sym}(T) = \sum_{\tau \in S_k} T^\tau\,.$$

Prove that this operation has properties 2, 3 and 4 of Proposition 1.4.6 and, as a substitute for property 1, has the property: $(\mathrm{Sym}T)^\sigma = \mathrm{Sym}T$.

(b) Let $\varphi_I = \mathrm{Sym}(e_I^*)$, $e_I^* = e_{i_1}^* \otimes \cdots \otimes e_{i_n}^*$. Prove that $\{\varphi_I,\ I$ non-decreasing$\}$ form a basis of $S^k(V)$.

(c) Conclude from (b) that $\dim S^k(V)$ is equal to the number of non-decreasing multi-indices of length k: $1 \le i_1 \le i_2 \le \cdots \le \ell_k \le n$.

(d) Compute this number by noticing that

$$(i_1, \ldots, i_n) \to (i_1 + 0, i_2 + 1, \ldots, i_k + k - 1)$$

is a bijection between the set of these non-decreasing multi-indices and the set of increasing multi-indices $1 \le j_1 < \cdots < j_k \le n + k - 1$.

1.6 The wedge product

The tensor algebra operations on the spaces, $\mathcal{L}^k(V)$, which we discussed in Sections 1.2 and 1.3, i.e., the "tensor product operation" and the "pull-back" operation, give rise to similar operations on the spaces, Λ^k. We will discuss in this section the analogue of the tensor product operation. As in § 4 we'll abbreviate $\mathcal{L}^k(V)$ to $\mathcal{L}^k$ and $\Lambda^k(V)$ to Λ^k when it's clear which "V" is intended.

Given $\omega_i \in \Lambda^{k_i}$, $i = 1, 2$ we can, by (1.5.4), find a $T_i \in \mathcal{L}^{k_i}$ with $\omega_i = \pi(T_i)$. Then $T_1 \otimes T_2 \in \mathcal{L}^{k_1+k_2}$. Let

$$\omega_1 \wedge \omega_2 = \pi(T_1 \otimes T_2) \in \Lambda^{k_1+k_2}\,. \tag{1.6.1}$$

Claim.

This wedge product is well defined, i.e., doesn't depend on our choices of T_1 and T_2.

Proof. Let $\pi(T_1) = \pi(T_1') = \omega_1$. Then $T_1' = T_1 + W_1$ for some $W_1 \in \mathcal{I}^{k_1}$, so

$$T_1' \otimes T_2 = T_1 \otimes T_2 + W_1 \otimes T_2\,.$$

But $W_1 \in \mathcal{I}^{k_1}$ implies $W_1 \otimes T_2 \in \mathcal{I}^{k_1+k_2}$ and this implies:

$$\pi(T_1' \otimes T_2) = \pi(T_1 \otimes T_2)\,.$$

A similar argument shows that (1.6.1) is well-defined independent of the choice of T_2.

□

More generally let $\omega_i \in \Lambda^{k_i}$, $i = 1, 2, 3$, and let $\omega_i = \pi(T_i)$, $T_i \in \mathcal{L}^{k_i}$. Define

$$\omega_1 \wedge \omega_2 \wedge \omega_3 \in \Lambda^{k_1+k_2+k_3}$$

by setting

$$\omega_1 \wedge \omega_2 \wedge \omega_3 = \pi(T_1 \otimes T_2 \otimes T_3)\,.$$

As above it's easy to see that this is well-defined independent of the choice of T_1, T_2 and T_3. It is also easy to see that this triple wedge product is just the wedge product of $\omega_1 \wedge \omega_2$ with ω_3 or, alternatively, the wedge product of ω_1 with $\omega_2 \wedge \omega_3$, i.e.,

$$\omega_1 \wedge \omega_2 \wedge \omega_3 = (\omega_1 \wedge \omega_2) \wedge \omega_3 = \omega_1 \wedge (\omega_2 \wedge \omega_3). \tag{1.6.2}$$

We leave for you to check:

For $\lambda \in \mathbb{R}$

$$\lambda(\omega_1 \wedge \omega_2) = (\lambda\omega_1) \wedge \omega_2 = \omega_1 \wedge (\lambda\omega_2) \tag{1.6.3}$$

and verify the two distributive laws:

$$(\omega_1 + \omega_2) \wedge \omega_3 \;=\; \omega_1 \wedge \omega_3 + \omega_2 \wedge \omega_3 \tag{1.6.4}$$

and

$$\omega_1 \wedge (\omega_2 + \omega_3) \;=\; \omega_1 \wedge \omega_2 + \omega_1 \wedge \omega_3\,. \tag{1.6.5}$$

As we noted in § 1.4, $\mathcal{I}^k = \{0\}$ for $k = 1$, i.e., there are no non-zero "redundant" k tensors in degree $k = 1$. Thus

$$\Lambda^1(V^*) = V^* = \mathcal{L}^1(V^*). \tag{1.6.6}$$

A particularly interesting example of a wedge product is the following. Let $\ell_i \in V^* = \Lambda^1(V^*)$, $i = 1, \ldots, k$. Then if $T = \ell_1 \otimes \cdots \otimes \ell_k$

$$\ell_1 \wedge \cdots \wedge \ell_k = \pi(T) \in \Lambda^k(V^*)\,. \tag{1.6.7}$$

We will call (1.6.7) a *decomposable element* of $\Lambda^k(V^*)$.

We will prove that these elements satisfy the following wedge product identity. For $\sigma \in S_k$:

$$\ell_{\sigma(1)} \wedge \cdots \wedge \ell_{\sigma(k)} = (-1)^\sigma \ell_1 \wedge \cdots \wedge \ell_k\,. \tag{1.6.8}$$

Proof. For every $T \in \mathcal{L}^k$, $T = (-1)^\sigma T + W$ for some $W \in I^k$ by Proposition 1.5.4. Therefore since $\pi(W) = 0$

$$\pi(T^\sigma) = (-1)^\sigma \pi(T)\,. \tag{1.6.9}$$

In particular, if $T = \ell_1 \otimes \cdots \otimes \ell_k$, $T^\sigma = \ell_{\sigma(1)} \otimes \cdots \otimes \ell_{\sigma(k)}$, so

$$\begin{aligned} \pi(T^\sigma) &= \ell_{\sigma(1)} \wedge \cdots \wedge \ell_{\sigma(k)} = (-1)^\sigma \pi(T) \\ &= (-1)^\sigma \ell_1 \wedge \cdots \wedge \ell_k\,. \end{aligned}$$

□

In particular, for ℓ_1 and $\ell_2 \in V^*$

$$\ell_1 \wedge \ell_2 = -\ell_2 \wedge \ell_1 \tag{1.6.10}$$

and for ℓ_1, ℓ_2 and $\ell_3 \in V^*$

$$\ell_1 \wedge \ell_2 \wedge \ell_3 = -\ell_2 \wedge \ell_1 \wedge \ell_3 = \ell_2 \wedge \ell_3 \wedge \ell_1 \,. \tag{1.6.11}$$

More generally, it's easy to deduce from (1.6.8) the following result (which we'll leave as an exercise).

Theorem 1.6.1. *If $\omega_1 \in \Lambda^r$ and $\omega_2 \in \Lambda^s$ then*

$$\omega_1 \wedge \omega_2 = (-1)^{rs} \omega_2 \wedge \omega_1 \,. \tag{1.6.12}$$

Hint: It suffices to prove this for decomposable elements i.e., for $\omega_1 = \ell_1 \wedge \cdots \wedge \ell_r$ and $\omega_2 = \ell'_1 \wedge \cdots \wedge \ell'_s$. Now make rs applications of (1.6.10).

Let $e_1, \ldots, e_n$ be a basis of V and let $e_1^*, \ldots, e_n^*$ be the dual basis of V^*. For every multi-index, I, of length k,

$$e_{i_1}^* \wedge \cdots e_{i_k}^* = \pi(e_I^*) = \pi(e_{i_1}^* \otimes \cdots \otimes e_{i_k}^*) \,. \tag{1.6.13}$$

Theorem 1.6.2. *The elements (1.6.13), with I strictly increasing, are basis vectors of Λ^k.*

Proof. The elements

$$\psi_I = \mathrm{Alt}\,(e_I^*) \,,\ I \text{ strictly increasing,}$$

are basis vectors of $\mathcal{A}^k$ by Proposition 3.6; so their images, $\pi(\psi_I)$, are a basis of Λ^k. But

$$\begin{aligned} \pi(\psi_I) &= \pi \sum (-1)^\sigma (e_I^*)^\sigma \\ &= \sum (-1)^\sigma \pi(e_I^*)^\sigma \\ &= \sum (-1)^\sigma (-1)^\sigma \pi(e_I^*) \\ &= k!\pi(e_I^*) \,. \end{aligned}$$

□

Exercises:

1. Prove the assertions (1.6.3), (1.6.4) and (1.6.5).

2. Verify the multiplication law, (1.6.12) for wedge product.

3. Given $\omega \in \Lambda^r$ let ω^k be the k-fold wedge product of ω with itself, i.e., let $\omega^2 = \omega \wedge \omega$, $\omega^3 = \omega \wedge \omega \wedge \omega$, etc.

(a) Show that if r is odd then for $k > 1$, $\omega^k = 0$.

(b) Show that if ω is decomposable, then for $k > 1$, $\omega^k = 0$.

4. If ω and μ are in Λ^{2r} prove:

$$(\omega + \mu)^k = \sum_{\ell=0}^{k} \binom{k}{\ell} \omega^\ell \wedge \mu^{k-\ell}\,.$$

Hint: As in freshman calculus prove this binomial theorem by induction using the identity: $\binom{k}{\ell} = \binom{k-1}{\ell-1} + \binom{k-1}{\ell}$.

5. Let ω be an element of Λ^2. By definition the *rank* of ω is k if $\omega^k \neq 0$ and $\omega^{k+1} = 0$. Show that if

$$\omega = e_1 \wedge f_1 + \cdots + e_k \wedge f_k$$

with $e_i, f_i \in V^*$, then ω is of rank $\leq k$. *Hint:* Show that

$$\omega^k = k! e_1 \wedge f_1 \wedge \cdots \wedge e_k \wedge f_k\,.$$

6. Given $e_i \in V^*$, $i = 1, \ldots, k$ show that $e_1 \wedge \cdots \wedge e_k \neq 0$ if and only if the e_i's are linearly independent. *Hint:* Induction on k.

1.7 The interior product

We'll describe in this section another basic product operation on the spaces, $\Lambda^k(V^*)$. As above we'll begin by defining this operator on the $\mathcal{L}^k(V)$'s. Given $T \in \mathcal{L}^k(V)$ and $\mathrm{v} \in V$ let $\iota_{\mathrm{v}} T$ be the be the $(k-1)$-tensor which takes the value

(1.7.1)
$$\iota_{\mathrm{v}} T(\mathrm{v}_1, \ldots, \mathrm{v}_{k-1}) = \sum_{r=1}^{k} (-1)^{r-1} T(\mathrm{v}_1, \ldots, \mathrm{v}_{r-1}, \mathrm{v}, \mathrm{v}_r, \ldots, \mathrm{v}_{k-1})$$

on the $k-1$-tuple of vectors, $\mathrm{v}_1, \ldots, \mathrm{v}_{k-1}$, i.e., in the r^{th} summand on the right, v gets inserted between v_{r-1} and v_r. (In particular the first summand is $T(\mathrm{v}, \mathrm{v}_1, \ldots, \mathrm{v}_{k-1})$ and the last summand is $(-1)^{k-1} T(\mathrm{v}_1, \ldots, \mathrm{v}_{k-1}, \mathrm{v})$.) It's clear from the definition that if $\mathrm{v} = \mathrm{v}_1 + \mathrm{v}_2$

$$\iota_{\mathrm{v}} T = \iota_{\mathrm{v}_1} T + \iota_{\mathrm{v}_2} T\,, \tag{1.7.2}$$

and if $T = T_1 + T_2$

$$\iota_{\mathrm{v}} T = \iota_{\mathrm{v}} T_1 + \iota_{\mathrm{v}} T_2\,, \tag{1.7.3}$$

and we will leave for you to verify by inspection the following two lemmas:

Lemma 1.7.1. *If T is the decomposable k-tensor $\ell_1 \otimes \cdots \otimes \ell_k$ then*

$$\iota_{\mathrm{v}} T = \sum (-1)^{r-1} \ell_r(\mathrm{v}) \ell_1 \otimes \cdots \otimes \widehat{\ell_r} \otimes \cdots \otimes \ell_k \tag{1.7.4}$$

where the "cap" over ℓ_r means that it's deleted from the tensor product ,

and

Lemma 1.7.2. *If $T_1 \in \mathcal{L}^p$ and $T_2 \in \mathcal{L}^q$*

$$\iota_{\mathrm{v}}(T_1 \otimes T_2) = \iota_{\mathrm{v}} T_1 \otimes T_2 + (-1)^p T_1 \otimes \iota_{\mathrm{v}} T_2\,. \tag{1.7.5}$$

We will next prove the important identity

$$\iota_{\mathrm{v}}(\iota_{\mathrm{v}} T) = 0\,. \tag{1.7.6}$$

Proof. It suffices by linearity to prove this for decomposable tensors and since (1.7.6) is trivially true for $T \in \mathcal{L}^1$, we can by induction assume (1.7.6) is true for decomposible tensors of degree $k-1$. Let $\ell_1 \otimes \cdots \otimes \ell_k$ be a decomposable tensor of degree k. Setting $T = \ell_1 \otimes \cdots \otimes \ell_{k-1}$ and $\ell = \ell_k$ we have

$$\begin{aligned} \iota_{\mathrm{v}}(\ell_1 \otimes \cdots \otimes \ell_k) &= \iota_{\mathrm{v}}(T \otimes \ell) \\ &= \iota_{\mathrm{v}} T \otimes \ell + (-1)^{k-1} \ell(v) T \end{aligned}$$

by (1.7.5). Hence

$$\begin{aligned} \iota_{\mathrm{v}}(\iota_{\mathrm{v}}(T \otimes \ell)) &= \iota_{\mathrm{v}}(\iota_{\mathrm{v}} T) \otimes \ell + (-1)^{k-2} \ell(\mathrm{v}) \iota_{\mathrm{v}} T \\ &\quad + (-1)^{k-1} \ell(v) \iota_{\mathrm{v}} T\,. \end{aligned}$$

But by induction the first summand on the right is zero and the two remaining summands cancel each other out.

□

From (1.7.6) we can deduce a slightly stronger result: For $\mathrm{v}_1, \mathrm{v}_2 \in V$

(1.7.7)
$$\iota_{\mathrm{v}_1} \iota_{\mathrm{v}_2} = -\iota_{\mathrm{v}_2} \iota_{\mathrm{v}_1}\,.$$

Proof. Let $\mathrm{v} = \mathrm{v}_1 + \mathrm{v}_2$. Then $\iota_{\mathrm{v}} = \iota_{\mathrm{v}_1} + \iota_{\mathrm{v}_2}$ so

$$\begin{aligned} 0 = \iota_{\mathrm{v}} \iota_{\mathrm{v}} &= (\iota_{\mathrm{v}_1} + \iota_{\mathrm{v}_2})(\iota_{\mathrm{v}_1} + \iota_{\mathrm{v}_2}) \\ &= \iota_{\mathrm{v}_1}\iota_{\mathrm{v}_1} + \iota_{\mathrm{v}_1}\iota_{\mathrm{v}_2} + \iota_{\mathrm{v}_2}\iota_{\mathrm{v}_1} + \iota_{\mathrm{v}_2}\iota_{\mathrm{v}_2} \\ &= \iota_{\mathrm{v}_1}\iota_{\mathrm{v}_2} + \iota_{\mathrm{v}_2}\iota_{\mathrm{v}_1} \end{aligned}$$

since the first and last summands are zero by (1.7.6).

□

We'll now show how to define the operation, ι_{v}, on $\Lambda^k(V^*)$. We'll first prove

Lemma 1.7.3. *If $T \in \mathcal{L}^k$ is redundant then so is $\iota_{\mathrm{v}} T$.*

Proof. Let $T = T_1 \otimes \ell \otimes \ell \otimes T_2$ where ℓ is in V^*, T_1 is in $\mathcal{L}^p$ and T_2 is in $\mathcal{L}^q$. Then by (1.7.5)

$$\begin{aligned} \iota_{\mathrm{v}} T &= \iota_{\mathrm{v}} T_1 \otimes \ell \otimes \ell \otimes T_2 \\ &\quad + (-1)^p T_1 \otimes \iota_{\mathrm{v}}(\ell \otimes \ell) \otimes T_2 \\ &\quad + (-1)^{p+2} T_1 \otimes \ell \otimes \ell \otimes \iota_{\mathrm{v}} T_2\,. \end{aligned}$$

However, the first and the third terms on the right are redundant and

$$\iota_{\mathrm{v}}(\ell \otimes \ell) = \ell(\mathrm{v})\ell - \ell(\mathrm{v})\ell$$

by (1.7.4). □

Now let π be the projection (1.5.4) of $\mathcal{L}^k$ onto Λ^k and for $\omega = \pi(T) \in \Lambda^k$ define

$$\iota_{\mathrm{v}}\omega = \pi(\iota_{\mathrm{v}} T)\,. \tag{1.7.8}$$

To show that this definition is legitimate we note that if $\omega = \pi(T_1) = \pi(T_2)$, then $T_1 - T_2 \in \mathcal{I}^k$, so by Lemma 1.7.3 $\iota_{\mathrm{v}} T_1 - \iota_{\mathrm{v}} T_2 \in \mathcal{I}^{k-1}$ and hence

$$\pi(\iota_{\mathrm{v}} T_1) = \pi(\iota_{\mathrm{v}} T_2)\,.$$

Therefore, (1.7.8) doesn't depend on the choice of T.

By definition ι_{v} is a linear mapping of $\Lambda^k(V^*)$ into $\Lambda^{k-1}(V^*)$. We will call this the *interior product operation*. From the identities (1.7.2)–(1.7.8) one gets, for $\mathrm{v}, \mathrm{v}_1, \mathrm{v}_2 \in V$ $\omega \in \Lambda^k$, $\omega_1 \in \Lambda^p$ and $\omega_2 \in \Lambda^2$

$$\iota_{(\mathrm{v}_1+\mathrm{v}_2)}\omega = \iota_{\mathrm{v}_1}\omega + \iota_{\mathrm{v}_2}\omega \tag{1.7.9}$$

$$\iota_{\mathrm{v}}(\omega_1 \wedge \omega_2) = \iota_{\mathrm{v}}\omega_1 \wedge \omega_2 + (-1)^p \omega_1 \wedge \iota_{\mathrm{v}}\omega_2 \tag{1.7.10}$$

$$\iota_{\mathrm{v}}(\iota_{\mathrm{v}}\omega) = 0 \tag{1.7.11}$$

and

$$\iota_{\mathrm{v}_1}\iota_{\mathrm{v}_2}\omega = -\iota_{\mathrm{v}_2}\iota_{\mathrm{v}_1}\omega\,. \tag{1.7.12}$$

Moreover if $\omega = \ell_1 \wedge \cdots \wedge \ell_k$ is a decomposable element of Λ^k one gets from (1.7.4)

$$\iota_{\mathrm{v}}\omega = \sum_{r=1}^{k} (-1)^{r-1} \ell_r(\mathrm{v}) \ell_1 \wedge \cdots \wedge \widehat{\ell}_r \wedge \cdots \wedge \ell_k\,. \tag{1.7.13}$$

In particular if $e_1, \ldots, e_n$ is a basis of V, $e_1^*, \ldots, e_n^*$ the dual basis of V^* and $\omega_I = e_{i_1}^* \wedge \cdots \wedge e_{i_k}^*$, $1 \le i_1 < \cdots < i_k \le n$, then $\iota(e_j)\omega_I = 0$ if $j \notin I$ and if $j = i_r$

$$\iota(e_j)\omega_I = (-1)^{r-1}\omega_{I_r} \tag{1.7.14}$$

where $I_r = (i_1, \ldots, \widehat{i}_r, \ldots, i_k)$ (i.e., I_r is obtained from the multi-index I by deleting i_r).

Exercises:

1. Prove Lemma 1.7.1.

2. Prove Lemma 1.7.2.

3. Show that if $T \in \mathcal{A}^k$, $i_v = kT_v$ where T_v is the tensor (1.3.16). In particular conclude that $i_v T \in \mathcal{A}^{k-1}$. (See §1.4, exercise 8.)

4. Assume the dimension of V is n and let Ω be a non-zero element of the one dimensional vector space Λ^n. Show that the map

$$\rho : V \to \Lambda^{n-1}\,, \quad v \to \iota_v \Omega\,, \tag{1.7.15}$$

is a bijective linear map. *Hint:* One can assume $\Omega = e_1^* \wedge \cdots \wedge e_n^*$ where $e_1, \ldots, e_n$ is a basis of V. Now use (1.7.14) to compute this map on basis elements.

5. (The cross-product.) Let V be a 3-dimensional vector space, B an inner product on V and Ω a non-zero element of Λ^3. Define a map

$$V \times V \to V$$

by setting

$$v_1 \times v_2 = \rho^{-1}(Lv_1 \wedge Lv_2) \tag{1.7.16}$$

where ρ is the map (1.7.15) and $L : V \to V^*$ the map (1.2.9). Show that this map is linear in v_1, with v_2 fixed and linear in v_2 with v_1 fixed, and show that $v_1 \times v_2 = -v_2 \times v_1$.

6. For $V = \mathbb{R}^3$ let e_1, e_2 and e_3 be the standard basis vectors and B the standard inner product. (See §1.1.) Show that if $\Omega = e_1^* \wedge e_2^* \wedge e_3^*$ the cross-product above is the standard cross-product:

$$\begin{aligned} e_1 \times e_2 &= e_3 \\ e_2 \times e_3 &= e_1 \\ e_3 \times e_1 &= e_2\,. \end{aligned} \tag{1.7.17}$$

Hint: If B is the standard inner product $Le_i = e_i^*$.

Remark 1.7.4. *One can make this standard cross-product look even more standard by using the calculus notation:* $e_1 = \widehat{i}$, $e_2 = \widehat{j}$ *and* $e_3 = \widehat{k}$

1.8 The pull-back operation on Λ^k

Let V and W be vector spaces and let A be a linear map of V into W. Given a k-tensor, $T \in \mathcal{L}^k(W)$, the *pull-back*, A^*T, is the k-tensor

$$A^*T(v_1, \dots, v_k) = T(Av_1, \dots, Av_k) \tag{1.8.1}$$

in $\mathcal{L}^k(V)$. (See § 1.3, equation 1.3.12.) In this section we'll show how to define a similar pull-back operation on Λ^k.

Lemma 1.8.1. *If $T \in \mathcal{I}^k(W)$, then $A^*T \in \mathcal{I}^k(V)$.*

Proof. It suffices to verify this when T is a redundant k-tensor, i.e., a tensor of the form

$$T = \ell_1 \otimes \cdots \otimes \ell_k$$

where $\ell_r \in W^*$ and $\ell_i = \ell_{i+1}$ for some index, i. But by (1.3.14)

$$A^*T = A^*\ell_1 \otimes \cdots \otimes A^*\ell_k$$

and the tensor on the right is redundant since $A^*\ell_i = A^*\ell_{i+1}$.

□

Now let ω be an element of $\Lambda^k(W^*)$ and let $\omega = \pi(T)$ where T is in $\mathcal{L}^k(W)$. We define

$$A^*\omega = \pi(A^*T)\,. \tag{1.8.2}$$

Claim:

The left hand side of (1.8.2) is well-defined.

Proof. If $\omega = \pi(T) = \pi(T')$, then $T = T' + S$ for some $S \in \mathcal{I}^k(W)$, and $A^*T' = A^*T + A^*S$. But $A^*S \in \mathcal{I}^k(V)$, so

$$\pi(A^*T') = \pi(A^*T)\,.$$

Proposition 1.8.2. *The map*

$$A^* : \Lambda^k(W^*) \to \Lambda^k(V^*)\,,$$

mapping ω to $A^*\omega$ is linear. Moreover, *(i)* *If* $\omega_i \in \Lambda^{k_i}(W)$, $i = 1, 2$, *then*

$$A^*(\omega_1 \wedge \omega_2) = A^*\omega_1 \wedge A^*\omega_2 \,. \tag{1.8.3}$$

(ii) *If* U *is a vector space and* $B : U \to V$ *a linear map, then for* $\omega \in \Lambda^k(W^*)$,

$$B^*A^*\omega = (AB)^*\omega \,. \tag{1.8.4}$$

We'll leave the proof of these three assertions as exercises. *Hint:* They follow immediately from the analogous assertions for the pull-back operation on tensors. (See (1.3.14) and (1.3.15).)

As an application of the pull-back operation we'll show how to use it to define the notion of *determinant* for a linear mapping. Let V be a n-dimensional vector space. Then $\dim \Lambda^n(V^*) = \binom{n}{n} = 1$; i.e., $\Lambda^n(V^*)$ is a *one-dimensional* vector space. Thus if $A : V \to V$ is a linear mapping, the induced pull-back mapping:

$$A^* : \Lambda^n(V^*) \to \Lambda^n(V^*) \,,$$

is just "multiplication by a constant". We denote this constant by $\det(A)$ and call it the *determinant* of A, Hence, by definition,

$$A^*\omega = \det(A)\omega \tag{1.8.5}$$

for all ω in $\Lambda^n(V^*)$. From (1.8.5) it's easy to derive a number of basic facts about determinants.

Proposition 1.8.3. *If* A *and* B *are linear mappings of* V *into* V, *then*

$$\det(AB) = \det(A)\det(B) \,. \tag{1.8.6}$$

Proof. By (1.8.4) and

$$\begin{aligned} (AB)^*\omega &= \det(AB)\omega \\ &= B^*(A^*\omega) = \det(B)A^*\omega \\ &= \det(B)\det(A)\omega \,, \end{aligned}$$

so, $\det(AB) = \det(A)\det(B)$.

□

Proposition 1.8.4. *If $I : V \to V$ is the identity map, $Iv = v$ for all $v \in V$,* $\det(I) = 1$.

We'll leave the proof as an exercise. *Hint:* I^* is the identity map on $\Lambda^n(V^*)$.

Proposition 1.8.5. *If $A : V \to V$ is not onto,* $\det(A) = 0$.

Proof. Let W be the image of A. Then if A is not onto, the dimension of W is less than n, so $\Lambda^n(W^*) = \{0\}$. Now let $A = I_W B$ where I_W is the inclusion map of W into V and B is the mapping, A, regarded as a mapping from V to W. Thus if ω is in $\Lambda^n(V^*)$, then by (1.8.4)

$$A^*\omega = B^* I_W^* \omega$$

and since $I_W^*\omega$ is in $\Lambda^n(W)$ it is zero.

□

□

We will derive by wedge product arguments the familiar "matrix formula" for the determinant. Let V and W be n-dimensional vector spaces and let $e_1, \dots, e_n$ be a basis for V and $f_1, \dots, f_n$ a basis for W. From these bases we get dual bases, $e_1^*, \dots, e_n^*$ and $f_1^*, \dots, f_n^*$, for V^* and W^*. Moreover, if A is a linear map of V into W and $[a_{i,j}]$ the $n \times n$ matrix describing A in terms of these bases, then the transpose map, $A^* : W^* \to V^*$, is described in terms of these dual bases by the $n \times n$ transpose matrix, i.e., if

$$Ae_j = \sum a_{i,j} f_i \,,$$

then

$$A^* f_j^* = \sum a_{j,i} e_i^* \,.$$

(See § 2.) Consider now $A^*(f_1^* \wedge \cdots \wedge f_n^*)$. By (1.8.3)

$$\begin{aligned} A^*(f_1^* \wedge \cdots \wedge f_n^*) &= A^* f_1^* \wedge \cdots \wedge A^* f_n^* \\ &= \sum (a_{1,k_1} e_{k_1}^*) \wedge \cdots \wedge (a_{n,k_n} e_{k_n}^*) \end{aligned}$$

the sum being over all $k_1, \dots, k_n$, with $1 \le k_r \le n$. Thus,

$$A^*(f_1^* \wedge \cdots \wedge f_n^*) = \sum a_{1,k_1} \dots a_{n,k_n} \; e_{k_1}^* \wedge \cdots \wedge e_{k_n}^* \,.$$

If the multi-index, $k_1, \dots, k_n$, is repeating, then $e^*_{k_1} \wedge \cdots \wedge e^*_{k_n}$ is zero, and if it's not repeating then we can write

$$k_i = \sigma(i) \quad i = 1, \dots, n$$

for some permutation, σ, and hence we can rewrite $A^*(f^*_1 \wedge \cdots \wedge f^*_n)$ as the sum over $\sigma \in S_n$ of

$$\sum a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} \quad (e^*_1 \wedge \cdots \wedge e^*_n)^\sigma .$$

But

$$(e^*_1 \wedge \cdots \wedge e^*_n)^\sigma = (-1)^\sigma e^*_1 \wedge \cdots \wedge e^*_n$$

so we get finally the formula

$$A^*(f^*_1 \wedge \cdots \wedge f^*_n) = \det[a_{i,j}] e^*_1 \wedge \cdots \wedge e^*_n \tag{1.8.7}$$

where

$$\det[a_{i,j}] = \sum (-1)^\sigma a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} \tag{1.8.8}$$

summed over $\sigma \in S_n$. The sum on the right is (as most of you know) the *determinant* of $[a_{i,j}]$.

Notice that if $V = W$ and $e_i = f_i$, $i = 1, \dots, n$, then $\omega = e^*_1 \wedge \cdots \wedge e^*_n = f^*_1 \wedge \cdots \wedge f^*_n$, hence by (1.8.5) and (1.8.7),

$$\det(A) = \det[a_{i,j}] . \tag{1.8.9}$$

Exercises.

1. Verify the three assertions of Proposition 1.8.2.

2. Deduce from Proposition 1.8.5 a well-known fact about determinants of $n \times n$ matrices: If two columns are equal, the determinant is zero.

3. Deduce from Proposition 1.8.3 another well-known fact about determinants of $n \times n$ matrices: If one interchanges two columns, then one changes the sign of the determinant.

Hint: Let $e_1, \dots, e_n$ be a basis of V and let $B : V \to V$ be the linear mapping: $Be_i = e_j$, $Be_j = e_i$ and $Be_\ell = e_\ell$, $\ell \neq i, j$. What is $B^*(e^*_1 \wedge \cdots \wedge e^*_n)$?

4. Deduce from Propositions 1.8.3 and 1.8.4 another well-known fact about determinants of $n \times n$ matrix. If $[b_{i,j}]$ is the inverse of $[a_{i,j}]$, its determinant is the inverse of the determinant of $[a_{i,j}]$.

5. Extract from (1.8.8) a well-known formula for determinants of 2×2 matrices:

$$\det \begin{bmatrix} a_{11}\,, & a_{12} \\ a_{21}\,, & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21}\,.$$

6. Show that if $A = [a_{i,j}]$ is an $n \times n$ matrix and $A^t = [a_{j,i}]$ is its transpose $\det A = \det A^t$. *Hint:* You are required to show that the sums

$$\sum (-1)^\sigma a_{1,\sigma(1)} \dots a_{n,\sigma(n)} \qquad \sigma \in S_n$$

and

$$\sum (-1)^\sigma a_{\sigma(1),1} \dots a_{\sigma(n),n} \qquad \sigma \in S_n$$

are the same. Show that the second sum is identical with

$$\sum (-1)^\tau a_{\tau(1),1} \dots a_{\tau(n),n}$$

summed over $\tau = \sigma^{-1} \in S_n$.

7. Let A be an $n \times n$ matrix of the form

$$A = \begin{bmatrix} B & * \\ 0 & C \end{bmatrix}$$

where B is a $k \times k$ matrix and C the $\ell \times \ell$ matrix and the bottom $\ell \times k$ block is zero. Show that

$$\det A = \det B \det C\,.$$

Hint: Show that in (1.8.8) every non-zero term is of the form

$$(-1)^{\sigma\tau} b_{1,\sigma(1)} \dots b_{k,\sigma(k)} c_{1,\tau(1)} \dots c_{\ell,\tau(\ell)}$$

where $\sigma \in S_k$ and $\tau \in S_\ell$.

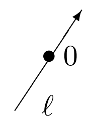
8. Let V and W be vector spaces and let $A : V \to W$ be a linear map. Show that if $Av = w$ then for $\omega \in \Lambda^p(w^*)$,

$$A^*\iota(w)\omega = \iota(v)A^*\omega\,.$$

(*Hint:* By (1.7.10) and proposition 1.8.2 it suffices to prove this for $\omega \in \Lambda^1(W^*)$, i.e., for $\omega \in W^*$.)

1.9 Orientations

We recall from freshman calculus that if $\ell \subseteq \mathbb{R}^2$ is a line through the origin, then $\ell - \{0\}$ has two connected components and an *orientation* of ℓ is a choice of one of these components (as in the figure below).



More generally, if $\mathbb{L}$ is a one-dimensional vector space then $\mathbb{L} - \{0\}$ consists of two components: namely if v is an element of $\mathbb{L} - [0\}$, then these two components are

$$\mathbb{L}_1 \quad = \quad \{\lambda v \ \ \lambda > 0\}$$

and

$$\mathbb{L}_2 \quad = \quad \{\lambda v, \ \lambda < 0\}\,.$$

An *orientation* of $\mathbb{L}$ is a choice of one of these components. Usually the component chosen is denoted $\mathbb{L}_+$, and called the *positive* component of $\mathbb{L} - \{0\}$ and the other component, $\mathbb{L}_-$, the *negative* component of $\mathbb{L} - \{0\}$.

Definition 1.9.1. *A vector, $v \in \mathbb{L}$, is positively oriented if v is in $\mathbb{L}_+$.*

More generally still let V be an n-dimensional vector space. Then $\mathbb{L} = \Lambda^n(V^*)$ is one-dimensional, and we define an *orientation* of V to be an orientation of $\mathbb{L}$. One important way of assigning an orientation to V is to choose a basis, $e_1, \dots, e_n$ of V. Then, if $e_1^*, \dots, e_n^*$ is the dual basis, we can orient $\Lambda^n(V^*)$ by requiring that $e_1^* \wedge \cdots \wedge e_n^*$ be in the positive component of $\Lambda^n(V^*)$. If V has already been assigned an orientation we will say that the basis, $e_1, \dots, e_n$, is *positively* oriented if the orientation we just described coincides with the given orientation.

Suppose that $e_1, \dots, e_n$ and $f_1, \dots, f_n$ are bases of V and that

$$e_j = \sum a_{i,j,} f_i\,. \tag{1.9.1}$$

Then by (1.7.7)

$$f_1^* \wedge \cdots \wedge f_n^* = \det[a_{i,j}] e_1^* \wedge \cdots \wedge e_n^*$$

so we conclude:

Proposition 1.9.2. *If* $e_1, \ldots, e_n$ *is positively oriented, then* $f_1, \ldots, f_n$ *is positively oriented if and only if* $\det[a_{i,j}]$ *is positive.*

Corollary 1.9.3. *If* $e_1, \ldots, e_n$ *is a positively oriented basis of* V*, the basis:* $e_1, \ldots, e_{i-1}, -e_i, e_{i+1}, \ldots, e_n$ *is negatively oriented.*

Now let V be a vector space of dimension $n > 1$ and W a subspace of dimension $k < n$. We will use the result above to prove the following important theorem.

Theorem 1.9.4. *Given orientations on* V *and* V/W*, one gets from these orientations a natural orientation on* W*.*

Remark What we mean by "natural' will be explained in the course of the proof.

Proof. Let $r = n - k$ and let π be the projection of V onto V/W . By exercises 1 and 2 of §2 we can choose a basis $e_1, \ldots, e_n$ of V such that $e_{r+1}, \ldots, e_n$ is a basis of W and $\pi(e_1), \ldots, \pi(e_r)$ a basis of V/W. Moreover, replacing e_1 by $-e_1$ if necessary we can assume by Corollary 1.9.3 that $\pi(e_1), \ldots, \pi(e_r)$ is a positively oriented basis of V/W and replacing e_n by $-e_n$ if necessary we can assume that $e_1, \ldots, e_n$ is a positively oriented basis of V. Now assign to W the orientation associated with the basis $e_{r+1}, \ldots, e_n$.

Let's show that this assignment is "natural" (i.e., doesn't depend on our choice of $e_1, \ldots, e_n$). To see this let $f_1, \ldots, f_n$ be another basis of V with the properties above and let $A = [a_{i,j}]$ be the matrix (1.9.1) expressing the vectors $e_1, \ldots, e_n$ as linear combinations of the vectors $f_1, \ldots f_n$. This matrix has to have the form

$$A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix} \tag{1.9.2}$$

where B is the $r \times r$ matrix expressing the basis vectors $\pi(e_1), \ldots, \pi(e_r)$ of V/W as linear combinations of $\pi(f_1), \ldots, \pi(f_r)$ and D the $k \times k$ matrix expressing the basis vectors $e_{r+1}, \ldots, e_n$ of W as linear combinations of $f_{r+1}, \ldots, f_n$. Thus

$$\det(A) = \det(B)\det(D)\,.$$

However, by Proposition 1.9.2, $\det A$ and $\det B$ are positive, so $\det D$ is positive, and hence if $e_{r+1}, \dots, e_n$ is a positively oriented basis of W so is $f_{r+1}, \dots, f_n$.

□

As a special case of this theorem suppose $\dim W = n - 1$. Then the choice of a vector $v \in V - W$ gives one a basis vector, $\pi(v)$, for the one-dimensional space V/W and hence if V is oriented, the choice of v gives one a natural orientation on W.

Next let V_i, $i = 1, 2$ be oriented n-dimensional vector spaces and $A : V_1 \to V_2$ a bijective linear map. A is *orientation-preserving* if, for $\omega \in \Lambda^n(V_2^*)_+$, $A^*\omega$ is in $\Lambda^n(V_+^*)_+$. For example if $V_1 = V_2$ then $A^*\omega = \det(A)\omega$ so A is orientation preserving if and only if $\det(A) > 0$. The following proposition we'll leave as an exercise.

Proposition 1.9.5. *Let V_i, $i = 1, 2, 3$ be oriented n-dimensional vector spaces and $A_i : V_i \to V_{i+1}$, $i = 1, 2$ bijective linear maps. Then if A_1 and A_2 are orientation preserving, so is $A_2 \circ A_1$.*

Exercises.

1. Prove Corollary 1.9.3.

2. Show that the argument in the proof of Theorem 1.9.4 can be modified to prove that if V and W are oriented then these orientations induce a natural orientation on V/W.

3. Similarly show that if W and V/W are oriented these orientations induce a natural orientation on V.

4. Let V be an n-dimensional vector space and $W \subset V$ a k-dimensional subspace. Let $U = V/W$ and let $\iota : W \to V$ and $\pi : V \to U$ be the inclusion and projection maps. Suppose V and U are oriented. Let μ be in $\Lambda^{n-k}(U^*)_+$ and let ω be in $\Lambda^n(V^*)_+$. Show that there exists a ν in $\Lambda^k(V^*)$ such that $\pi^*\mu \wedge \nu = \omega$. Moreover show that $\iota^*\nu$ is *intrinsically* defined (i.e., doesn't depend on how we choose ν) and sits in the positive part, $\Lambda^k(W^*)_+$, of $\Lambda^k(W)$.

5. Let $e_1, \dots, e_n$ be the standard basis vectors of $\mathbb{R}^n$. The *standard* orientation of $\mathbb{R}^n$ is, by definition, the orientation associated with this basis. Show that if W is the subspace of $\mathbb{R}^n$ defined by the

equation, $x_1 = 0$, and $v = e_1 \notin W$ then the natural orientation of W associated with v and the standard orientation of $\mathbb{R}^n$ coincide with the orientation given by the basis vectors, $e_2, \dots, e_n$ of W.

6. Let V be an oriented n-dimensional vector space and W an $n-1$-dimensional subspace. Show that if v and v' are in $V - W$ then $v' = \lambda v + w$, where w is in W and $\lambda \in \mathbb{R} - \{0\}$. Show that v and v' give rise to the same orientation of W if and only if λ is positive.

7. Prove Proposition 1.9.5.

8. A key step in the proof of Theorem 1.9.4 was the assertion that the matrix A expressing the vectors, e_i, as linear combinations of the vectors, f_i, had to have the form (1.9.2). Why is this the case?

9. (a) Let V be a vector space, W a subspace of V and $A : V \to V$ a bijective linear map which maps W onto W. Show that one gets from A a bijective linear map

$$B : V/W \to V/W$$

with property

$$\pi A = B\pi\,,$$

π being the projection of V onto V/W.

(b) Assume that V, W and V/W are compatibly oriented. Show that if A is orientation-preserving and its restriction to W is orientation preserving then B is orientation preserving.

10. Let V be a oriented n-dimensional vector space, W an $(n-1)$-dimensional subspace of V and $i : W \to V$ the inclusion map. Given $\omega \in \Lambda^b(V)_+$ and $v \in V - W$ show that for the orientation of W described in exercise 5, $i^*(\iota_v \omega) \in \Lambda^{n-1}(W)_+$.

11. Let V be an n-dimensional vector space, $B : V \times V \to \mathbb{R}$ an inner product and $e_1, \dots, e_n$ a basis of V which is positively oriented and orthonormal. Show that the “volume element”

$$\mathrm{vol} \;= e_1^* \wedge \cdots \wedge e_n^* \in \Lambda^n(V^*)$$

is intrinsically defined, independent of the choice of this basis. *Hint:* (1.2.13) and (1.8.7).

12. (a) Let V be an oriented n-dimensional vector space and B an inner product on V. Fix an oriented orthonormal basis, $e_1, \dots, e_n$, of V and let $A : V \to V$ be a linear map. Show that if

$$Ae_i = \mathrm{v}_i = \sum a_{j,i} e_j$$

and $b_{i,j} = B(\mathrm{v}_i, \mathrm{v}_j)$, the matrices $\mathcal{A} = [a_{i,j}]$ and $\mathcal{B} = [b_{i,j}]$ are related by: $\mathcal{B} = \mathcal{A}^+ \mathcal{A}$.

(b) Show that if ν is the volume form, $e_1^* \wedge \cdots \wedge e_n^*$, and A is orientation preserving

$$A^* \nu = (\det \mathcal{B})^{\frac{1}{2}} \nu \, .$$

(c) By Theorem 1.5.6 one has a bijective map

$$\Lambda^n(V^*) \cong \mathcal{A}^n(V) \, .$$

Show that the element, Ω, of $\mathcal{A}^n(V)$ corresponding to the form, ν, has the property

$$|\Omega(\mathrm{v}_1, \dots, \mathrm{v}_n)|^2 = \det([b_{i,j}])$$

where $\mathrm{v}_1, \dots, \mathrm{v}_n$ are any n-tuple of vectors in V and $b_{i,j} = B(\mathrm{v}_i, \mathrm{v}_j)$.

CHAPTER 2

DIFFERENTIAL FORMS

2.1 Vector fields and one-forms

The goal of this chapter is to generalize to n dimensions the basic operations of three dimensional vector calculus: div, curl and grad. The "div", and "grad" operations have fairly straight forward generalizations, but the "curl" operation is more subtle. For vector fields it doesn't have any obvious generalization, however, if one replaces vector fields by a closely related class of objects, differential forms, then not only does it have a natural generalization but it turns out that div, curl and grad are all special cases of a general operation on differential forms called *exterior differentiation*.

In this section we will review some basic facts about vector fields in n variables and introduce their dual objects: *one-forms*. We will then take up in §2.2 the theory of k-forms for k greater than one. We begin by fixing some notation.

Given $p \in \mathbb{R}^n$ we define the tangent space to $\mathbb{R}^n$ at p to be the set of pairs

$$T_p\mathbb{R}^n = \{(p, \mathrm{v})\}\,; \quad \mathrm{v} \in \mathbb{R}^n\,. \tag{2.1.1}$$

The identification

$$T_p\mathbb{R}^n \to \mathbb{R}^n\,, \quad (p, \mathrm{v}) \to \mathrm{v} \tag{2.1.2}$$

makes $T_p\mathbb{R}^n$ into a vector space. More explicitly, for v, v_1 and $\mathrm{v}_2 \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ we define the addition and scalar multiplication operations on $T_p\mathbb{R}^n$ by the recipes

$$(p, \mathrm{v}_1) + (p, \mathrm{v}_2) \quad = \quad (p, \mathrm{v}_1 + \mathrm{v}_2)$$

and

$$\lambda(p, \mathrm{v}) \quad = \quad (p, \lambda\mathrm{v})\,.$$

Let U be an open subset of $\mathbb{R}^n$ and $f : U \to \mathbb{R}^m$ a C^1 map. We recall that the derivative

$$Df(p) : \mathbb{R}^n \to \mathbb{R}^m$$

of f at p is the linear map associated with the $m \times n$ matrix

$$\left[\frac{\partial f_i}{\partial x_j}(p)\right] .$$

It will be useful to have a "base-pointed" version of this definition as well. Namely, if $q = f(p)$ we will define

$$df_p : T_p\mathbb{R}^n \to T_q\mathbb{R}^m$$

to be the map

$$df_p(p, \mathrm{v}) = (q, Df(p)\mathrm{v}) . \tag{2.1.3}$$

It's clear from the way we've defined vector space structures on $T_p\mathbb{R}^n$ and $T_q\mathbb{R}^m$ that this map is linear.

Suppose that the image of f is contained in an open set, V, and suppose $g : V \to \mathbb{R}^k$ is a C^1 map. Then the "base-pointed"" version of the chain rule asserts that

$$dg_q \circ df_p = d(f \circ g)_p . \tag{2.1.4}$$

(This is just an alternative way of writing $Dg(q)Df(p) = D(g \circ f)(p)$.)

In 3-dimensional vector calculus a *vector field* is a function which attaches to each point, p, of $\mathbb{R}^3$ a base-pointed arrow, $(p, \vec{\mathrm{v}})$. The n-dimensional version of this definition is essentially the same.

Definition 2.1.1. *Let U be an open subset of $\mathbb{R}^n$. A vector field on U is a function, v, which assigns to each point, p, of U a vector $v(p)$ in $T_p\mathbb{R}^n$.*

Thus a vector field is a vector-valued function, but its value at p is an element of a vector space, $T_p\mathbb{R}^n$ that itself depends on p.

Some examples.

1. Given a fixed vector, $\mathrm{v} \in \mathbb{R}^n$, the function

$$p \in \mathbb{R}^n \to (p, \mathrm{v}) \tag{2.1.5}$$

is a vector field. Vector fields of this type are *constant* vector fields.

2. In particular let $e_i, i = 1, \ldots, n$, be the standard basis vectors of $\mathbb{R}^n$. If $\mathrm{v} = e_i$ we will denote the vector field (2.1.5) by $\partial/\partial x_i$. (The reason for this "derivation notation" will be explained below.)

3. Given a vector field on U and a function, $f : U \to \mathbb{R}$ we'll denote by fv the vector field

$$p \in U \to f(p)v(p)\,.$$

4. Given vector fields v_1 and v_2 on U, we'll denote by $v_1 + v_2$ the vector field

$$p \in U \to v_1(p) + v_2(p)\,.$$

5. The vectors, (p, e_i), $i = 1, \dots, n$, are a basis of $T_p\mathbb{R}^n$, so if v is a vector field on U, $v(p)$ can be written uniquely as a linear combination of these vectors with real numbers, $g_i(p)$, $i = 1, \dots, n$, as coefficients. In other words, using the notation in example 2 above, v can be written uniquely as a sum

$$v = \sum_{i=1}^{n} g_i \frac{\partial}{\partial x_i} \tag{2.1.6}$$

where $g_i : U \to \mathbb{R}$ is the function, $p \to g_i(p)$.

We'll say that v is a $\mathcal{C}^\infty$ vector field if the g_i's are in $\mathcal{C}^\infty(U)$.

A basic vector field operation is *Lie differentiation*. If $f \in C^1(U)$ we define $L_v f$ to be the function on U whose value at p is given by

$$Df(p)\mathrm{v} = L_v f(p) \tag{2.1.7}$$

where $v(p) = (p, \mathrm{v})$. If v is the vector field (2.1.6) then

$$L_v f = \sum g_i \frac{\partial}{\partial x_i} f \tag{2.1.8}$$

(motivating our "derivation notation" for v).

Exercise.

Check that if $f_i \in C^1(U)$, $i = 1, 2$, then

$$L_v(f_1 f_2) = f_1 L_v f_2 + f_1 L_v f_2\,. \tag{2.1.9}$$

Next we'll generalize to n-variables the calculus notion of an "integral curve" of a vector field.

Definition 2.1.2. *A C^1 curve $\gamma : (a,b) \to U$ is an integral curve of v if for all $a < t < b$ and $p = \gamma(t)$*

$$\left(p, \frac{d\gamma}{dt}(t)\right) = v(p)$$

i.e., if v is the vector field (2.1.6) and $g : U \to \mathbb{R}^n$ is the function $(g_1, \ldots, g_n)$ the condition for $\gamma(t)$ to be an integral curve of v is that it satisfy the system of differential equations

$$\frac{d\gamma}{dt}(t) = g(\gamma(t))\,. \tag{2.1.10}$$

We will quote without proof a number of basic facts about systems of ordinary differential equations of the type (2.1.10). (A source for these results that we highly recommend is Birkhoff–Rota, *Ordinary Differential Equations*, Chapter 6.)

Theorem 2.1.3 (Existence)**.** *Given a point $p_0 \in U$ and $a \in \mathbb{R}$, there exists an interval $I = (a - T, a + T)$, a neighborhood, U_0, of p_0 in U and for every $p \in U_0$ an integral curve, $\gamma_p : I \to U$ with $\gamma_p(a) = p$.*

Theorem 2.1.4 (Uniqueness)**.** *Let $\gamma_i : I_i \to U$, $i = 1, 2$, be integral curves. If $a \in I_1 \cap I_2$ and $\gamma_1(a) = \gamma_2(a)$ then $\gamma_1 \equiv \gamma_2$ on $I_1 \cap I_2$ and the curve $\gamma : I_1 \cup I_2 \to U$ defined by*

$$\gamma(t) = \begin{cases} \gamma_1(t)\,, & t \in I_1 \\ \gamma_2(t)\,, & t \in I_2 \end{cases}$$

is an integral curve.

Theorem 2.1.5 (Smooth dependence on initial data)**.** *Let v be a $\mathcal{C}^\infty$-vector field, on an open subset, V, of U, $I \subseteq \mathbb{R}$ an open interval, $a \in I$ a point on this interval and $h : V \times I \to U$ a mapping with the properties:*

> *(i) $h(p, a) = p$.*
>
> *(ii) For all $p \in V$ the curve*
>
> $$\gamma_p : I \to U \qquad \gamma_p(t) = h(p, t)$$
>
> *is an integral curve of v.*
>
> *Then the mapping, h, is $\mathcal{C}^\infty$.*

One important feature of the system (2.1.11) is that it is an *autonomous* system of differential equations: the function, $g(x)$, is a function of x alone, it doesn't depend on t. One consequence of this is the following:

Theorem 2.1.6. *Let $I = (a, b)$ and for $c \in \mathbb{R}$ let $I_c = (a - c, b - c)$. Then if $\gamma : I \to U$ is an integral curve, the reparametrized curve*

$$\gamma_c : I_c \to U\,, \quad \gamma_c(t) = \gamma(t + c) \tag{2.1.11}$$

is an integral curve.

We recall that a C^1-function $\varphi : U \to \mathbb{R}$ is an *integral* of the system (2.1.11) if for every integral curve $\gamma(t)$, the function $t \to \varphi(\gamma(t))$ is constant. This is true if and only if for all t and $p = \gamma(t)$

$$0 = \frac{d}{dt}\varphi(\gamma(t)) = (D\varphi)_p\left(\frac{d\gamma}{dt}\right) = (D\varphi)_p(\mathrm{v})$$

where $(p, \mathrm{v}) = v(p)$. But by (2.1.6) the term on the right is $L_v\varphi(p)$. Hence we conclude

Theorem 2.1.7. *$\varphi \in C^1(U)$ is an integral of the system (2.1.11) if and only if $L_v\varphi = 0$.*

We'll now discuss a class of objects which are in some sense "dual objects" to vector fields. For each $p \in \mathbb{R}^n$ let $(T_p\mathbb{R})^*$ be the dual vector space to $T_p\mathbb{R}^n$, i.e., the space of all linear mappings, $\ell : T_p\mathbb{R}^n \to \mathbb{R}$.

Definition 2.1.8. *Let U be an open subset of $\mathbb{R}^n$. A* one-form *on U is a function, ω, which assigns to each point, p, of U a vector, ω_p, in $(T_p\mathbb{R}^n)^*$.*

Some examples:

1. Let $f : U \to \mathbb{R}$ be a C^1 function. Then for $p \in U$ and $c = f(p)$ one has a linear map

$$df_p : T_p\mathbb{R}^n \to T_c\mathbb{R} \tag{2.1.12}$$

and by making the identification,

$$T_c\mathbb{R} = \{c, \mathbb{R}\} = \mathbb{R}$$

df_p can be regarded as a linear map from $T_p\mathbb{R}^n$ to $\mathbb{R}$, i.e., as an element of $(T_p\mathbb{R}^n)^*$. Hence the assignment

$$p \in U \to df_p \in (T_p\mathbb{R}^n)^* \tag{2.1.13}$$

defines a one-form on U which we'll denote by df.

2. Given a one-form ω and a function, $\varphi : U \to \mathbb{R}$ the product of φ with ω is the one-form, $p \in U \to \varphi(p)\omega_p$.

3. Given two one-forms ω_1 and ω_2 their sum, $\omega_1 + \omega_2$ is the one-form, $p \in U \to \omega_1(p) + \omega_2(p)$.

4. The one-forms $dx_1, \ldots, dx_n$ play a particularly important role. By (2.1.12)

$$(dx_i)\left(\frac{\partial}{\partial x_j}\right)_p = \delta_{ij} \tag{2.1.14}$$

i.e., is equal to 1 if $i = j$ and zero if $i \neq j$. Thus $(dx_1)_p, \ldots, (dx_n)_p$ are the basis of $(T_p^*\mathbb{R}^n)^*$ *dual* to the basis $(\partial/\partial x_i)_p$. Therefore, if ω is *any* one-form on U, ω_p can be written uniquely as a sum

$$\omega_p = \sum f_i(p)(dx_i)_p\,, \quad f_i(p) \in \mathbb{R}$$

and ω can be written uniquely as a sum

$$\omega = \sum f_i\, dx_i \tag{2.1.15}$$

where $f_i : U \to \mathbb{R}$ is the function, $p \to f_i(p)$. We'll say that ω *is a* $\mathcal{C}^\infty$ *one-form* if the f_i's are $\mathcal{C}^\infty$.

Exercise.

Check that if $f : U \to \mathbb{R}$ is a $\mathcal{C}^\infty$ function

$$df = \sum \frac{\partial f}{\partial x_i}\, dx_i\,. \tag{2.1.16}$$

Suppose now that v is a vector field and ω a one-form on U. Then for every $p \in U$ the vectors, $v_p \in T_p\mathbb{R}^n$ and $\omega_p \in (T_p\mathbb{R}^n)^*$ can be paired to give a number, $\iota(v_p)\omega_p \in \mathbb{R}$, and hence, as p varies, an

$\mathbb{R}$-valued function, $\iota(v)\omega$, which we will call the *interior product* of v with ω. For instance if v is the vector field (2.1.6) and ω the one-form (2.1.15) then

$$\iota(v)\omega = \sum f_i g_i \,. \tag{2.1.17}$$

Thus if v and ω are $\mathcal{C}^\infty$ so is the function $\iota(v)\omega$. Also notice that if $\varphi \in \mathcal{C}^\infty(U)$, then as we observed above

$$d\varphi = \sum \frac{\partial \varphi}{\partial x_i} \frac{\partial}{\partial x_i}$$

so if v is the vector field (2.1.6)

$$\iota(v)\, d\varphi = \sum g_i \frac{\partial \varphi}{\partial x_i} = L_v \varphi \,. \tag{2.1.18}$$

Coming back to the theory of integral curves, let U be an open subset of $\mathbb{R}^n$ and v a vector field on U. We'll say that v is *complete* if, for every $p \in U$, there exists an integral curve, $\gamma : \mathbb{R} \to U$ with $\gamma(0) = p$, i.e., for every p there exists an integral curve that starts at p and *exists for all time*. To see what "completeness" involves, we recall that an integral curve

$$\gamma : [0, b) \to U \,,$$

with $\gamma(0) = p$, is called *maximal* if it can't be extended to an interval $[0, b')$, $b' > b$. (See for instance Birkhoff–Rota, §6.11.) For such curves it's known that either

i. $b = +\infty$
or

ii. $|\gamma(t)| \to +\infty$ as $t \to b$
or

iii. the limit set of

$$\{\gamma(t) \,, \quad 0 \leq t, b\}$$

contains points on the boundary of U.

Hence if we can exclude ii. and iii. we'll have shown that an integral curve with $\gamma(0) = p$ exists for all positive time. A simple criterion for excluding ii. and iii. is the following.

Lemma 2.1.9. *The scenarios ii. and iii. can't happen if there exists a proper C^1-function, $\varphi : U \to \mathbb{R}$ with $L_v\varphi = 0$.*

Proof. $L_v\varphi = 0$ implies that φ is constant on $\gamma(t)$, but if $\varphi(p) = c$ this implies that the curve, $\gamma(t)$, lies on the compact subset, $\varphi^{-1}(c)$, of U; hence it can't "run off to infinity" as in scenario ii. or "run off the boundary" as in scenario iii.

□

Applying a similar argument to the interval $(-b, 0]$ we conclude:

Theorem 2.1.10. *Suppose there exists a proper C^1-function, $\varphi : U \to \mathbb{R}$ with the property $L_v\varphi = 0$. Then v is complete.*

Example.

Let $U = \mathbb{R}^2$ and let v be the vector field

$$v = x^3 \frac{\partial}{\partial y} - y \frac{\partial}{\partial x}\,.$$

Then $\varphi(x, y) = 2y^2 + x^4$ is a proper function with the property above. Another hypothesis on v which excludes ii. and iii. is the following. We'll define the *support* of v to be the set

$$\operatorname{supp} v = \overline{q \in U\,, \quad v(q) \neq 0\}}\,,$$

and will say that v is compactly supported if this set is compact. We will prove

Theorem 2.1.11. *If v is compactly supported it is complete.*

Proof. Notice first that if $v(p) = 0$, the constant curve, $\gamma_0(t) = p$, $-\infty < t < \infty$, satisfies the equation

$$\frac{d}{dt}\gamma_0(t) = 0 = v(p)\,,$$

so it is an integral curve of v. Hence if $\gamma(t)$, $-a < t < b$, is any integral curve of v with the property, $\gamma(t_0) = p$, for some t_0, it has to coincide with γ_0 on the interval, $-a < t < a$, and hence has to be the constant curve, $\gamma(t) = p$, on this interval.

Now suppose the support, A, of v is compact. Then either $\gamma(t)$ is in A for all t or is in $U - A$ for some t_0. But if this happens, and

$p = \gamma(t_0)$ then $v(p) = 0$, so $\gamma(t)$ has to coincide with the constant curve, $\gamma_0(t) = p$, for all t. In neither case can it go off to ∞ or off to the boundary of U as $t \to b$.

□

One useful application of this result is the following. Suppose v is a vector field on U, and one wants to see what its integral curves look like on some compact set, $A \subseteq U$. Let $\rho \in \mathcal{C}_0^\infty(U)$ be a bump function which is equal to one on a neighborhood of A. Then the vector field, $w = \rho v$, is compactly supported and hence complete, but it is identical with v on A, so its integral curves on A coincide with the integral curves of v.

If v is complete then for every p, one has an integral curve, $\gamma_p : \mathbb{R} \to U$ with $\gamma_p(0) = p$, so one can define a map

$$f_t : U \to U$$

by setting $f_t(p) = \gamma_p(t)$. If v is $\mathcal{C}^\infty$, this mapping is $\mathcal{C}^\infty$ by the smooth dependence on initial data theorem, and by definition f_0 is the identity map, i.e., $f_0(p) = \gamma_p(0) = p$. We claim that the f_t's also have the property

$$f_t \circ f_a = f_{t+a} \,. \tag{2.1.19}$$

Indeed if $f_a(p) = q$, then by the reparametrization theorem, $\gamma_q(t)$ and $\gamma_p(t + a)$ are both integral curves of v, and since $q = \gamma_q(0) = \gamma_p(a) = f_a(p)$, they have the same initial point, so

$$\begin{aligned} \gamma_q(t) &= f_t(q) = (f_t \circ f_a)(p) \\ &= \gamma_p(t + a) = f_{t+a}(p) \end{aligned}$$

for all t. Since f_0 is the identity it follows from (2.1.19) that $f_t \circ f_{-t}$ is the identity, i.e.,

$$f_{-t} = f_t^{-1} \,,$$

so f_t is a $\mathcal{C}^\infty$ diffeomorphism. Hence if v is complete it generates a "one-parameter group", f_t, $-\infty < t < \infty$, of $\mathcal{C}^\infty$-diffeomorphisms.

For v not complete there is an analogous result, but it's trickier to formulate precisely. Roughly speaking v generates a one-parameter group of diffeomorphisms, f_t, but these diffeomorphisms are not defined on all of U nor for all values of t. Moreover, the identity (2.1.19) only holds on the open subset of U where both sides are well-defined.

We'll devote the remainder of this section to discussing some "functorial" properties of vector fields and one-forms. Let U and W be open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, and let $f : U \to W$ be a $\mathcal{C}^\infty$ map. If v is a $\mathcal{C}^\infty$-vector field on U and w a $\mathcal{C}^\infty$-vector field on W we will say that v and w are "f-related" if, for all $p \in U$ and $q = f(p)$

$$df_p(v_p) = \mathrm{w}_q \,. \tag{2.1.20}$$

Writing

$$v = \sum_{i=1}^n v_i \frac{\partial}{\partial x_i}\,, \quad v_i \in C^k(U)$$

and

$$\mathrm{w} = \sum_{j=1}^m \mathrm{w}_j \frac{\partial}{\partial y_j}\,, \quad \mathrm{w}_j \in C^k(V)$$

this equation reduces, in coordinates, to the equation

$$\mathrm{w}_i(q) = \sum \frac{\partial f_i}{\partial x_j}(p) v_j(p)\,. \tag{2.1.21}$$

In particular, if $m = n$ and f is a $\mathcal{C}^\infty$ diffeomorphism, the formula (3.2) defines a $\mathcal{C}^\infty$-vector field on W, i.e.,

$$\mathrm{w} = \sum_{j=1}^n \mathrm{w}_i \frac{\partial}{\partial y_j}$$

is the vector field defined by the equation

$$\mathrm{w}_i = \sum_{j=1}^n \left(\frac{\partial f_i}{\partial x_j} v_j \right) \circ f^{-1}\,. \tag{2.1.22}$$

Hence we've proved

Theorem 2.1.12. *If $f : U \to W$ is a $\mathcal{C}^\infty$ diffeomorphism and v a $\mathcal{C}^\infty$-vector field on U, there exists a unique $\mathcal{C}^\infty$ vector field, w, on W having the property that v and w are f-related.*

We'll denote this vector field by $f_* v$ and call it the *push-forward of v by f*.

I'll leave the following assertions as easy exercises.

Theorem 2.1.13. *Let U_i, $i = 1, 2$, be open subsets of $\mathbb{R}^{n_i}$, v_i a vector field on U_i and $f : U_1 \to U_2$ a $\mathcal{C}^\infty$-map. If v_1 and v_2 are f-related, every integral curve*

$$\gamma : I \to U_1$$

of v_1 gets mapped by f onto an integral curve, $f \circ \gamma : I \to U_2$, of v_2.

Corollary 2.1.14. *Suppose v_1 and v_2 are complete. Let $(f_i)_t : U_i \to U_i$, $-\infty < t < \infty$, be the one-parameter group of diffeomorphisms generated by v_i. Then $f \circ (f_1)_t = (f_2)_t \circ f$.*

Hints:

1. Theorem 4 follows from the chain rule: If $p = \gamma(t)$ and $q = f(p)$

$$df_p \left(\frac{d}{dt} \gamma(t) \right) = \frac{d}{dt} f(\gamma(t)) .$$

2. To deduce Corollary 5 from Theorem 4 note that for $p \in U$, $(f_1)_t(p)$ is just the integral curve, $\gamma_p(t)$ of v_1 with initial point $\gamma_p(0) = p$.

The notion of f-relatedness can be very succinctly expressed in terms of the Lie differentiation operation. For $\varphi \in \mathcal{C}^\infty(U_2)$ let $f^*\varphi$ be the composition, $\varphi \circ f$, viewed as a $\mathcal{C}^\infty$ function on U_1, i.e., for $p \in U_1$ let $f^*\varphi(p) = \varphi(f(p))$. Then

$$f^* L_{v_2} \varphi = L_{v_1} f^* \varphi . \tag{2.1.23}$$

(To see this note that if $f(p) = q$ then at the point p the right hand side is

$$(d\varphi)_q \circ df_p(v_1(p))$$

by the chain rule and by definition the left hand side is

$$d\varphi_q(v_2(q)) .$$

Moreover, by definition

$$v_2(q) = df_p(v_1(p))$$

so the two sides are the same.)

Another easy consequence of the chain rule is:

Theorem 2.1.15. *Let U_i, $i = 1, 2, 3$, be open subsets of $\mathbb{R}^{n_i}$, v_i a vector field on U_i and $f_i : U_i \to U_{i+1}$, $i = 1, 2$ a $\mathcal{C}^\infty$-map. Suppose that, for $i = 1, 2$, v_i and v_{i+1} are f_i-related. Then v_1 and v_3 are $f_2 \circ f_1$-related.*

In particular, if f_1 and f_2 are diffeomorphisms and $v = v_1$

$$(f_2)_*(f_1)_* v = (f_2 \circ f_1)_* v\,.$$

The results we described above have "dual" analogues for one-forms. Namely, let U and V be open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, and let $f : U \to V$ be a $\mathcal{C}^\infty$-map. Given a one-form, μ, on V one can define a "pull-back" one-form, $f^*\mu$, on U by the following method. For $p \in U$ let $q = f(p)$. By definition $\mu(q)$ is a linear map

$$\mu(q) : T_q\mathbb{R}^m \to \mathbb{R} \tag{2.1.24}$$

and by composing this map with the linear map

$$df_p : T_p\mathbb{R}^n \to T_q\mathbb{R}^n$$

we get a linear map

$$\mu_q \circ df_p : T_p\mathbb{R}^n \to \mathbb{R}\,,$$

i.e., an element $\mu_q \circ df_p$ of $T_p^*\mathbb{R}^n$.

Definition 2.1.16. *The one-form $f^*\mu$ is the one-form defined by the map*

$$p \in U \to (\mu_q \circ df_p) \in T_p^*\mathbb{R}^n$$

where $q = f(p)$.

Note that if $\varphi : V \to \mathbb{R}$ is a $\mathcal{C}^\infty$-function and $\mu = d\varphi$ then

$$\mu_q \circ df_p = d\varphi_q \circ df_p = d(\varphi \circ f)_p$$

i.e.,

$$f^*\mu = d\varphi \circ f\,. \tag{2.1.25}$$

In particular if μ is a one-form of the form, $\mu = d\varphi$, with $\varphi \in \mathcal{C}^\infty(V)$, $f^*\mu$ is $\mathcal{C}^\infty$. From this it is easy to deduce

Theorem 2.1.17. *If μ is any $\mathcal{C}^\infty$ one-form on V, its pull-back, $f^*\omega$, is $\mathcal{C}^\infty$. (See exercise 1.)*

Notice also that the pull-back operation on one-forms and the push-forward operation on vector fields are somewhat different in character. The former is defined for *all* $\mathcal{C}^\infty$ maps, but the latter is only defined for diffeomorphisms.

Exercises.

1. Let U be an open subset of $\mathbb{R}^n$, V an open subset of $\mathbb{R}^n$ and $f : U \to V$ a C^k map.

(a) Show that for $\varphi \in \mathcal{C}^\infty(V)$ (2.1.25) can be rewritten

$$f^*\, d\varphi = df^*\varphi\,. \tag{2.1.25$'$}$$

(b) Let μ be the one-form

$$\mu = \sum_{i=1}^m \varphi_i\, dx_i \qquad \varphi_i \in \mathcal{C}^\infty(V)$$

on V. Show that if $f = (f_1, \dots, f_m)$ then

$$f^*\mu = \sum_{i=1}^m f^*\varphi_i\, df_i\,.$$

(c) Show that if μ is $\mathcal{C}^\infty$ and f is $\mathcal{C}^\infty$, $f^*\mu$ is $\mathcal{C}^\infty$.

2. Let v be a complete vector field on U and $f_t : U \to U$, the one parameter group of diffeomorphisms generated by v. Show that if $\varphi \in C^1(U)$

$$L_v\varphi = \left(\frac{d}{dt} f_t^*\varphi\right)_{t=0}\,.$$

3. (a) Let $U = \mathbb{R}^2$ and let $\mathfrak{v}$ be the vector field, $x_1\partial/\partial x_2 - x_2\partial/\partial x_1$. Show that the curve

$$t \in \mathbb{R} \to (r\cos(t+\theta)\,,\, r\sin(t+\theta))$$

is the unique integral curve of $\mathfrak{v}$ passing through the point, $(r\cos\theta, r\sin\theta)$, at $t = 0$.

(b) Let $U = \mathbb{R}^n$ and let $\mathfrak{v}$ be the constant vector field: $\sum c_i \partial/\partial x_i$. Show that the curve

$$t \in \mathbb{R} \to a + t(c_1, \dots, c_n)$$

is the unique integral curve of $\mathfrak{v}$ passing through $a \in \mathbb{R}^n$ at $t = 0$.

(c) Let $U = \mathbb{R}^n$ and let $\mathfrak{v}$ be the vector field, $\sum x_i \partial/\partial x_i$. Show that the curve

$$t \in \mathbb{R} \to e^t(a_1, \dots, a_n)$$

is the unique integral curve of $\mathfrak{v}$ passing through a at $t = 0$.

4. Show that the following are one-parameter groups of diffeomorphisms:

(a) $f_t : \mathbb{R} \to \mathbb{R}\,, \quad f_t(x) = x + t$

(b) $f_t : \mathbb{R} \to \mathbb{R}\,, \quad f_t(x) = e^t x$

(c) $f_t : \mathbb{R}^2 \to \mathbb{R}^2\,, \quad f_t(x, y) = (\cos t\, x - \sin t\, y\,,\, \sin t\, x + \cos t\, y)$

5. Let $A : \mathbb{R}^n \to \mathbb{R}^n$ be a linear mapping. Show that the series

$$\exp tA = I + tA + \frac{t^2}{2!} A^2 + \frac{t^3}{3!} A^3 + \cdots$$

converges and defines a one-parameter group of diffeomorphisms of $\mathbb{R}^n$.

6. (a) What are the infinitesimal generators of the one-parameter groups in exercise 13?

(b) Show that the infinitesimal generator of the one-parameter group in exercise 14 is the vector field

$$\sum a_{i,j} x_j \frac{\partial}{\partial x_i}$$

where $[a_{i,j}]$ is the defining matrix of A.

7. Let v be the vector field on $\mathbb{R}$, $x^2 \frac{d}{dx}$ Show that the curve

$$x(t) = \frac{a}{a - at}$$

is an integral curve of v with initial point $x(0) = a$. Conclude that for $a > 0$ the curve

$$x(t) = \frac{a}{1 - at}\,, \quad 0 < t < \frac{1}{a}$$

is a maximal integral curve. (In particular, conclude that v isn't complete.)

8. Let U be an open subset of $\mathbb{R}^n$ and v_1 and v_2 vector fields on U. Show that there is a unique vector field, w, on U with the property

$$L_w\varphi = L_{v_1}(L_{v_2}\varphi) - L_{v_2}(L_{v_1}\varphi)$$

for all $\varphi \in \mathcal{C}^\infty(U)$.

9. The vector field w in exercise 8 is called the *Lie bracket* of the vector fields v_1 and v_2 and is denoted $[v_1, v_2]$. Verify that "Lie bracket" satisfies the identities

$$[v_1, v_2] = -[v_2, v_1]$$

and

$$[v_1[v_2, v_3]] + [v_2, [v_3, v_1]] + [v_3, [v_1, v_2]] = 0\,.$$

Hint: Prove analogous identities for L_{v_1}, L_{v_2} and L_{v_3}.

10. Let $v_1 = \partial/\partial x_i$ and $v_2 = \sum g_j \partial/\partial x_j$. Show that

$$[v_1, v_2] = \sum \frac{\partial}{\partial x_i} g_i \frac{\partial}{\partial x_j}\,.$$

11. Let v_1 and v_2 be vector fields and f a $\mathcal{C}^\infty$ function. Show that

$$[v_1, fv_2] = L_{v_1} f v_2 + f[v_1, v_2]\,.$$

12. Let U and V be open subsets of $\mathbb{R}^n$ and $f : U \to V$ a diffeomorphism. If w is a vector field on V, define the pull-back, f^*w of w to U to be the vector field

$$f^*w = (f_*^{-1} w)\,.$$

Show that if φ is a $\mathcal{C}^\infty$ function on V

$$f^* L_w \varphi = L_{f^*w} f^* \varphi\,.$$

Hint: (2.1.26).

13. Let U be an open subset of $\mathbb{R}^n$ and v and w vector fields on U. Suppose v is the infinitesimal generator of a one-parameter group of diffeomorphisms

$$f_t : U \to U \,, \quad -\infty < t < \infty \,.$$

Let $w_t = f_t^* w$. Show that for $\varphi \in \mathcal{C}^\infty(U)$

$$L_{[v,w]}\varphi \;=\; L_{\dot{w}}\varphi$$

where

$$\dot{w} \;=\; \frac{d}{dt} f_t^* w \,|_{t=0} \,.$$

Hint: Differentiate the identity

$$f_t^* L_w \varphi = L_{w_t} f_t^* \varphi$$

with respect to t and show that at $t = 0$ the derivative of the left hand side is

$$L_v L_w \varphi$$

by exercise 2 and the derivative of the right hand side is

$$L_{\dot{w}} + L_w(L_v\varphi) \,.$$

14. Conclude from exercise 13 that

$$[v,w] = \frac{d}{dt} f_t^* w \,|_{t=0} \,. \tag{2.1.26}$$

15. Let U be an open subset of $\mathbb{R}^n$ and let $\gamma : [a,b] \to U$, $t \to (\gamma_1(t), \ldots, \gamma_n(t))$ be a C^1 curve. Given $\omega = \sum f_i \, dx_i \in \Omega^1(U)$, define the *line integral* of ω over γ to be the integral

$$\int_\gamma \omega = \sum_{i=1}^n \int_a^b f_i(\gamma(t)) \frac{d\gamma_i}{dt} \, dt \,.$$

Show that if $\omega = df$ for some $f \in \mathcal{C}^\infty(U)$

$$\int_\gamma \omega = f(\gamma(b)) - f(\gamma(a)) \,.$$

In particular conclude that if γ is a closed curve, i.e., $\gamma(a) = \gamma(b)$, this integral is zero.

16. Let

$$\omega = \frac{x_1\,dx_2 - x_2\,dx_1}{x_1^2 + x_2^2} \in \Omega^1(\mathbb{R}^2 - \{0\}),$$

and let $\gamma : [0, 2\pi] \to \mathbb{R}^2 - \{0\}$ be the closed curve, $t \to (\cos t, \sin t)$. Compute the line integral, $\int_\gamma \omega$, and show that it's not zero. Conclude that ω can't be "d" of a function, $f \in \mathcal{C}^\infty(\mathbb{R}^2 - \{0\})$.

17. Let f be the function

$$f(x_1, x_2) = \begin{cases} \arctan \frac{x_2}{x_1}\,, x_1 > 0 \\ \frac{\pi}{2}\,,\ x_1 = 0\,,\ x_2 > 0 \\ \arctan \frac{x_2}{x_1} + \pi\,,\ x_1 < 0 \end{cases}$$

where, we recall: $-\frac{\pi}{2} < \arctan t < \frac{\pi}{2}$. Show that this function is $\mathcal{C}^\infty$ and that df is the 1-form, ω, in the previous exercise. Why doesn't this contradict what you proved in exercise 16?

2.2 k-forms

One-forms are the bottom tier in a pyramid of objects whose k^{th} tier is the space of *k-forms*. More explicitly, given $p \in \mathbb{R}^n$ we can, as in §1.5, form the k^{th} exterior powers

$$\Lambda^k(T_p^*\mathbb{R}^n)\,, \quad k = 1, 2, 3, \ldots, n \tag{2.2.1}$$

of the vector space, $T_p^*\mathbb{R}^n$, and since

$$\Lambda^1(T_p^*\mathbb{R}^n) = T_p^*\mathbb{R}^n \tag{2.2.2}$$

one can think of a one-form as a function which takes its value at p in the space (2.2.2). This leads to an obvious generalization.

Definition 2.2.1. *Let U be an open subset of $\mathbb{R}^n$. A k-form, ω, on U is a function which assigns to each point, p, in U an element $\omega(p)$ of the space (2.2.1) .*

The wedge product operation gives us a way to construct lots of examples of such objects.

Example 1.

Let ω_i, $i = 1, \dots, k$ be one-forms. Then $\omega_1 \wedge \cdots \wedge \omega_k$ is the k-form whose value at p is the wedge product

$$\omega_1(p) \wedge \cdots \wedge \omega_k(p)\,. \tag{2.2.3}$$

Notice that since $\omega_i(p)$ is in $\Lambda^1(T_p^*\mathbb{R}^n)$ the wedge product (2.2.3) makes sense and is an element of $\Lambda^k(T_p^*\mathbb{R}^n)$.

Example 2.

Let f_i, $i = 1, \dots, k$ be a real-valued $\mathcal{C}^\infty$ function on U. Letting $\omega_i = df_i$ we get from (2.2.3) a k-form

$$df_1 \wedge \cdots \wedge df_k \tag{2.2.4}$$

whose value at p is the wedge product

$$(df_1)_p \wedge \cdots \wedge (df_k)_p\,. \tag{2.2.5}$$

Since $(dx_1)_p, \dots, (dx_n)_p$ are a basis of $T_p^*\mathbb{R}^n$, the wedge products

$$(dx_{i_1})_p \wedge \cdots \wedge (dx_{1_k})_p\,, \quad 1 \le i_1 < \cdots < i_k \le n \tag{2.2.6}$$

are a basis of $\Lambda^k(T_p^*)$. To keep our multi-index notation from getting out of hand, we'll denote these basis vectors by $(dx_I)_p$, where $I = (i_1, \dots, i_k)$ and the I's range over multi-indices of length k which are *strictly increasing*. Since these wedge products are a basis of $\Lambda^k(T_p^*\mathbb{R}^n)$ every element of $\Lambda^k(T_p^*\mathbb{R}^n)$ can be written uniquely as a sum

$$\sum c_I (dx_I)_p\,, \quad c_I \in \mathbb{R}$$

and every k-form, ω, on U can be written uniquely as a sum

$$\omega = \sum f_I\, dx_I \tag{2.2.7}$$

where dx_I is the k-form, $dx_{i_1} \wedge \cdots \wedge dx_{i_k}$, and f_I is a real-valued function,

$$f_I : U \to \mathbb{R}\,.$$

Definition 2.2.2. *The k-form (2.2.7) is of class C^r if each of the f_I's is in $C^r(U)$.*

Henceforth we'll assume, unless otherwise stated, that *all the k-forms we consider are of class $\mathcal{C}^\infty$*, and we'll denote the space of these k-forms by $\Omega^k(U)$.

We will conclude this section by discussing a few simple operations on k-forms.

1. Given a function, $f \in \mathcal{C}^\infty(U)$ and a k-form $\omega \in \Omega^k(U)$ we define $f\omega \in \Omega^k(U)$ to be the k-form

$$p \in U \to f(p)\omega_p \in \Lambda^k(T_p^*\mathbb{R}^n)\,.$$

2. Given $\omega_i \in \Omega^k(U)$, $i = 1, 2$ we define $\omega_1 + \omega_2 \in \Omega^k(U)$ to be the k-form

$$p \in U \to (\omega_1)_p + (\omega_2)_p \in \Lambda^k(T_p^*\mathbb{R}^n)\,.$$

(Notice that this sum makes sense since each summand is in $\Lambda^k(T_p^*\mathbb{R}^n)$.)

3. Given $\omega_1 \in \Omega^{k_1}(U)$ and $\omega_2 \in \Omega^{k_2}(U)$ we define their *wedge product*, $\omega_1 \wedge \omega_2 \in \Omega^{k_1+k_2}(u)$ to be the $(k_1 + k_2)$-form

$$p \in U \to (\omega_1)_p \wedge (\omega_2)_p \in \Lambda^{k_1+k_2}(T_p^*\mathbb{R}^n)\,.$$

We recall that $\Lambda^0(T_p^*\mathbb{R}^n) = \mathbb{R}$, so a zero-form is an $\mathbb{R}$-valued function and a zero form of class $\mathcal{C}^\infty$ is a $\mathcal{C}^\infty$ function, i.e.,

$$\Omega^0(U) = \mathcal{C}^\infty(U)\,.$$

A fundamental operation on forms is the "d-operation" which associates to a function $f \in \mathcal{C}^\infty(U)$ the 1-form df. It's clear from the identity (2.1.10) that df is a 1-form of class $\mathcal{C}^\infty$, so the d-operation can be viewed as a map

$$d : \Omega^0(U) \to \Omega^1(U)\,. \tag{2.2.8}$$

We will show in the next section that an analogue of this map exists for every $\Omega^k(U)$.

Exercises.

1. Let $\omega \in \Omega^2(\mathbb{R}^4)$ be the 2-form, $dx_1 \wedge dx_2 + dx_3 \wedge dx_4$. Compute $\omega \wedge \omega$.

2. Let $\omega_i \in \Omega^1(\mathbb{R}^3)$, $i = 1, 2, 3$ be the 1-forms

$$\begin{aligned} \omega_1 &= x_2\,dx_3 - x_3\,dx_2 \\ \omega_2 &= x_3\,dx_1 - x_1\,dx_3 \end{aligned}$$

and

$$\omega_3 = x_1\,dx_2 - x_2\,dx_1\,.$$

Compute

(a) $\omega_1 \wedge \omega_2$.

(b) $\omega_2 \wedge \omega_3$.

(c) $\omega_3 \wedge \omega_1$.

(d) $\omega_1 \wedge \omega_2 \wedge \omega_3$.

3. Let U be an open subset of $\mathbb{R}^n$ and $f_i \in \mathcal{C}^\infty(U)$, $i = 1, \ldots, n$. Show that

$$df_1 \wedge \cdots \wedge df_n = \det \left[\frac{\partial f_i}{\partial x_j} \right] dx_1 \wedge \cdots \wedge dx_n \,.$$

4. Let U be an open subset of $\mathbb{R}^n$. Show that every $(n-1)$-form, $\omega \in \Omega^{n-1}(U)$, can be written uniquely as a sum

$$\sum_{i=1}^{n} f_i \; dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_n$$

where $f_i \in \mathcal{C}^\infty(U)$ and the "cap" over dx_i means that dx_i is to be deleted from the product, $dx_1 \wedge \cdots \wedge dx_n$.

5. Let $\mu = \displaystyle\sum_{i=1}^{n} x_i dx_i$. Show that there exists an $(n-1)$-form, $\omega \in \Omega^{n-1}(\mathbb{R}^n - \{0\})$ with the property

$$\mu \wedge \omega = dx_1 \wedge \cdots \wedge dx_n \,.$$

6. Let J be the multi-index $(j_1, \ldots, j_k)$ and let $dx_J = dx_{j_1} \wedge \cdots \wedge dx_{j_k}$. Show that $dx_J = 0$ if $j_r = j_s$ for some $r \neq s$ and show that if the j_r's are all distinct

$$dx_J = (-1)^\sigma \, dx_I$$

where $I = (i_1, \ldots, i_k)$ is the strictly increasing rearrangement of $(j_1, \ldots, j_k)$ and σ is the permutation

$$j_1 \to i_1, \ldots, j_k \to i_k \,.$$

7. Let I be a strictly increasing multi-index of length k and J a strictly increasing multi-index of length ℓ. What can one say about the wedge product $dx_I \wedge dx_J$?

2.3 Exterior differentiation

Let U be an open subset of $\mathbb{R}^n$. In this section we are going to define an operation

$$d : \Omega^k(U) \to \Omega^{k+1}(U)\,. \tag{2.3.1}$$

This operation is called *exterior differentiation* and is the fundamental operation in n-dimensional vector calculus.

For $k = 0$ we already defined the operation (2.3.1) in §2.1. Before defining it for the higher k's we list some properties that we will require to this operation to satisfy.

Property I. For ω_1 and ω_2 in $\Omega^k(U)$, $d(\omega_1 + \omega_2) = d\omega_1 + d\omega_2$.

Property II. For $\omega_1 \in \Omega^k(U)$ and $\omega_2 \in \Omega^\ell(U)$

$$d(\omega_1 \wedge \omega_2) = d\omega_1 \wedge \omega_2 + (-1)^k \omega_1 \wedge d\omega_2\,. \tag{2.3.2}$$

Property III. For $\omega \in \Omega^k(U)$

$$d(d\omega) = 0\,. \tag{2.3.3}$$

Let's point out a few consequences of these properties. First note that by Property III

$$d(df) = 0 \tag{2.3.4}$$

for every function, $f \in \mathcal{C}^\infty(U)$. More generally, given k functions, $f_i \in \mathcal{C}^\infty(U)$, $i = 1, \dots, k$, then by combining (2.3.4) with (2.3.2) we get by induction on k:

$$d(df_1 \wedge \cdots \wedge df_k) = 0\,. \tag{2.3.5}$$

Proof. Let $\mu = df_2 \wedge \cdots \wedge df_k$. Then by induction on k, $d\mu = 0$; and hence by (2.3.2) and (2.3.4)

$$d(df_1 \wedge \mu) = d(d_1 f) \wedge \mu + (-1)\, df_1 \wedge d\mu = 0\,,$$

as claimed.)

In particular, given a multi-index, $I = (i_1, \ldots, i_k)$ with $1 \le i_r \le n$

$$d(dx_I) = d(dx_{i_1} \wedge \cdots \wedge dx_{i_k}) = 0 \,. \tag{2.3.6}$$

Recall now that every k-form, $\omega \in \Omega^k(U)$, can be written uniquely as a sum

$$\omega = \sum f_I \, dx_I \,, \quad f_I \in \mathcal{C}^\infty(U)$$

where the multi-indices, I, are strictly increasing. Thus by (2.3.2) and (2.3.6)

$$d\omega = \sum df_I \wedge dx_I \,. \tag{2.3.7}$$

This shows that if there exists a "d" with properties I—III, it has to be given by the formula (2.3.7). Hence all we have to show is that the operator defined by this formula has these properties. Property I is obvious. To verify Property II we first note that for I strictly increasing (2.3.6) is a special case of (2.3.7). (Take $f_I = 1$ and $f_J = 0$ for $J \neq I$.) Moreover, if I is not strictly increasing it is either repeating, in which case $dx_I = 0$, or non-repeating in which case I^σ is strictly increasing for some permutation, $\sigma \in S_k$, and

$$dx_I = (-1)^\sigma \, dx_{I^\sigma} \,. \tag{2.3.8}$$

Hence (2.3.7) implies (2.3.6) for *all* multi-indices I. The same argument shows that for *any* sum over indices, I, for length k

$$\sum f_I dx_I$$

one has the identity:

$$d(\sum f_I \, dx_I) = \sum df_I \wedge dx_I \,. \tag{2.3.9}$$

(As above we can ignore the repeating I's, since for these I's, $dx_I = 0$, and by (2.3.8) we can make the non-repeating I's strictly increasing.)

Suppose now that $\omega_1 \in \Omega^k(U)$ and $\omega_2 \in \Omega^\ell(U)$. Writing

$$\omega_1 \;=\; \sum f_I \, dx_I$$

and

$$\omega_2 \;=\; \sum g_J \, dx_J$$

with f_I and g_J in $\mathcal{C}^\infty(U)$ we get for the wedge product

$$\omega_1 \wedge \omega_2 = \sum f_I g_J \, dx_I \wedge dx_J \tag{2.3.10}$$

and by (2.3.9)

$$d(\omega_1 \wedge \omega_2) = \sum d(f_I g_J) \wedge dx_I \wedge dx_J \,. \tag{2.3.11}$$

(Notice that if $I = (i_1, \cdots, i_k)$ and $J = (j_i, \ldots, i_\ell)$, $dx_I \wedge dx_J = dx_K$, K being the multi-index, $(i_1, \ldots, i_k, j_1, \ldots, j_\ell)$. Even if I and J are strictly increasing, K won't necessarily be strictly increasing. However in deducing (2.3.11) from (2.3.10) we've observed that this doesn't matter .) Now note that by (2.1.11)

$$d(f_I g_J) = g_J \, df_I + f_I \, dg_J \,,$$

and by the wedge product identities of §(1.6),

$$\begin{aligned} dg_J \wedge dx_I &= dg_J \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k} \\ &= (-1)^k \, dx_I \wedge dg_J \,, \end{aligned}$$

so the sum (2.3.11) can be rewritten:

$$\sum df_I \wedge dx_I \wedge g_J \, dx_J + (-1)^k \sum f_I \, dx_I \wedge \, dg_J \wedge \, dx_J \,,$$

or

$$\left(\sum df_I \wedge \, dx_I\right) \wedge \left(\sum g_J \, dx_J\right) + (-1)^k \left(\sum dg_J \wedge dx_J\right) \,,$$

or finally:

$$d\omega_1 \wedge \omega_2 + (-1)^k \omega_1 \wedge d\omega_2 \,.$$

Thus the "d" defined by (2.3.7) has Property II. Let's now check that it has Property III. If $\omega = \sum f_I \, dx_I$, $f_I \in \mathcal{C}^\infty(U)$, then by definition, $d\omega = \sum df_I \wedge dx_I$ and by (2.3.6) and (2.3.2)

$$d(d\omega) = \sum d(df_I) \wedge \, dx_I \,,$$

so it suffices to check that $d(df_I) = 0$, i.e., it suffices to check (2.3.4) for zero forms, $f \in \mathcal{C}^\infty(U)$. However, by (2.1.9)

$$df = \sum_{j=1}^n \frac{\partial f}{\partial x_j} \, dx_j$$

so by (2.3.7)

$$\begin{aligned} d(df) &= \sum_{j=1}^{n} d\left(\frac{\partial f}{\partial x_j}\right) dx_j \\ &= \sum_{j=1}^{n} \left(\sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} dx_i\right) \wedge dx_j \\ &= \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j} dx_i \wedge dx_j \,. \end{aligned}$$

Notice, however, that in this sum, $dx_i \wedge dx_j = -dx_j \wedge dx_i$ and

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

so the (i, j) term cancels the (j, i) term, and the total sum is zero.

□

A form, $\omega \in \Omega^k(U)$, is said to be *closed* if $d\omega = 0$ and is said to be *exact* if $\omega = d\mu$ for some $\mu \in \Omega^{k-1}(U)$. By Property III every exact form is closed, but the converse is not true even for 1-forms. (See §2.1, exercise 8). In fact it's a very interesting (and hard) question to determine if an open set, U, has the property: "For $k > 0$ every closed k-form is exact."[1]

Some examples of sets with this property are described in the exercises at the end of §2.5. We will also sketch below a proof of the following result (and ask you to fill in the details).

Lemma 2.3.1 (Poincaré's Lemma.)**.** *If ω is a closed form on U of degree $k > 0$, then for every point, $p \in U$, there exists a neighborhood of p on which ω is exact.*

(See exercises 5 and 6 below.)

Exercises:

1. Compute the exterior derivatives of the forms below.

[1]For $k = 0$, $df = 0$ *doesn't* imply that f is exact. In fact "exactness" doesn't make much sense for zero forms since there aren't any "-1" forms. However, if $f \in \mathcal{C}^\infty(U)$ and $df = 0$ then f is constant on connected components of U. (See § 2.1, exercise 2.)

(a) $x_1\, dx_2 \wedge dx_3$

(b) $x_1\, dx_2 - x_2\, dx_1$

(c) $e^{-f}\, df$ where $f = \sum_{i=1}^n x_i^2$

(d) $\sum_{i=1}^n x_i\, dx_i$

(e) $\sum_{i=1}^n (-1)^i x_i\, dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_n$

2. Solve the equation: $d\mu = \omega$ for $\mu \in \Omega^1(\mathbb{R}^3)$, where ω is the 2-form

(a) $dx_2 \wedge dx_3$

(b) $x_2\, dx_2 \wedge dx_3$

(c) $(x_1^2 + x_2^2)\, dx_1 \wedge dx_2$

(d) $\cos x_1\, dx_1 \wedge dx_3$

3. Let U be an open subset of $\mathbb{R}^n$.

(a) Show that if $\mu \in \Omega^k(U)$ is exact and $\omega \in \Omega^\ell(U)$ is closed then $\mu \wedge \omega$ is exact. *Hint:* The formula (2.3.2).

(b) In particular, dx_1 is exact, so if $\omega \in \Omega^\ell(U)$ is closed $dx_1 \wedge \omega = d\mu$. What is μ?

4. Let Q be the rectangle, $(a_1, b_1) \times \cdots \times (a_n, b_n)$. Show that if ω is in $\Omega^n(Q)$, then ω is exact.

Hint: Let $\omega = f\, dx_1 \wedge \cdots \wedge dx_n$ with $f \in \mathcal{C}^\infty(Q)$ and let g be the function

$$g(x_1, \ldots, x_n) = \int_{a_1}^{x_1} f(t, x_2, \ldots, x_n)\, dt\,.$$

Show that $\omega = d(g\, dx_2 \wedge \cdots \wedge dx_n)$.

5. Let U be an open subset of $\mathbb{R}^{n-1}$, $A \subseteq \mathbb{R}$ an open interval and (x, t) product coordinates on $U \times A$. We will say that a form, $\mu \in \Omega^\ell(U \times A)$ is *reduced* if it can be written as a sum

$$\mu = \sum f_I(x, t)\, dx_I\,, \tag{2.3.12}$$

(i.e., no terms involving dt).

(a) Show that every form, $\omega \in \Omega^k(U \times A)$ can be written uniquely as a sum:

$$\omega = dt \wedge \alpha + \beta \tag{2.3.13}$$

where α and β are reduced.

(b) Let μ be the reduced form (2.3.12) and let

$$\frac{d\mu}{dt} = \sum \frac{d}{dt} f_I(x,t)\, dx_I$$

and

$$d_U \mu = \sum_I \left(\sum_{i=1}^{n} \frac{\partial}{\partial x_i} f_I(x,t)\, dx_i \right) \wedge dx_I \,.$$

Show that

$$d\mu = dt \wedge \frac{d\mu}{dt} + d_U \mu \,.$$

(c) Let ω be the form (2.3.13). Show that

$$d\omega = dt \wedge d_U \alpha + dt \wedge \frac{d\beta}{dt} + d_U \beta$$

and conclude that ω is closed if and only if

$$\begin{aligned} \frac{d\beta}{dt} &= d_U \alpha \\ d\beta_U &= 0 \,. \end{aligned} \tag{2.3.14}$$

(d) Let α be a reduced $(k-1)$-form. Show that there exists a reduced $(k-1)$-form, ν, such that

$$\frac{d\nu}{dt} = \alpha \,. \tag{2.3.15}$$

Hint: Let $\alpha = \sum f_I(x,t)\, dx_I$ and $\nu = \sum g_I(x,t)\, dx_I$. The equation (2.3.15) reduces to the system of equations

$$\frac{d}{dt} g_I(x,t) = f_I(x,t) \,. \tag{2.3.16}$$

Let c be a point on the interval, A, and using freshman calculus show that (2.3.16) has a unique solution, $g_I(x,t)$, with $g_I(x,c) = 0$.

(e) Show that if ω is the form (2.3.13) and ν a solution of (2.3.15) then the form

$$\omega - d\nu \tag{2.3.17}$$

is reduced.

(f) Let

$$\gamma = \sum h_I(x,t)\, dx)I$$

be a reduced k-form. Deduce from (2.3.14) that if γ is closed then $\frac{d\gamma}{dt} = 0$ and $d_U\gamma = 0$. Conclude that $h_I(x,t) = h_I(x)$ and that

$$\gamma = \sum h_I(x)\, dx_I$$

is effectively a closed k-form on U. Now prove: If every closed k-form on U is exact, then every closed k-form on $U \times A$ is exact. *Hint:* Let ω be a closed k-form on $U \times A$ and let γ be the form (2.3.17).

6. Let $Q \subseteq \mathbb{R}^n$ be an open rectangle. Show that every closed form on Q of degree $k > 0$ is exact. *Hint:* Let $Q = (a_1, b_1) \times \cdots \times (a_n, b_n)$. Prove this assertion by induction, at the n^{th} stage of the induction letting $U = (a_1, b_1) \times \cdots \times (a_{n-1}, b_{n-1})$ and $A = (a_n, b_n)$.

2.4 The interior product operation

In §2.1 we explained how to pair a one-form, ω, and a vector field, v, to get a function, $\iota(v)\omega$. This pairing operation generalizes: If one is given a k-form, ω, and a vector field, v, both defined on an open subset, U, one can define a $(k-1)$-form on U by defining its value at $p \in U$ to be the interior product

$$\iota(v(p))\omega(p)\,. \tag{2.4.1}$$

Note that $v(p)$ is in $T_p\mathbb{R}^n$ and $\omega(p)$ in $\Lambda^k(T_p^*\mathbb{R}^n)$, so by definition of interior product (see §1.7), the expression (2.4.1) is an element of $\Lambda^{k-1}(T_p^*\mathbb{R}^n)$. We will denote by $\iota(v)\omega$ the $(k-1)-$form on U whose value at p is (2.4.1). From the properties of interior product on vector spaces which we discussed in §1.7, one gets analogous properties for this interior product on forms. We will list these properties, leaving their verification as an exercise. Let v and ω be vector fields, and ω_1

and ω_2 k-forms, ω a k-form and μ an ℓ-form. Then $\iota(v)\omega$ is linear in ω:

$$\iota(v)(\omega_1 + \omega_2) = \iota(v)\omega_1 + \iota(v)\omega_2\,, \tag{2.4.2}$$

linear in v:

$$\iota(v + w)\omega = \iota(v)\omega + z(w)\omega\,, \tag{2.4.3}$$

has the derivation property:

$$\iota(v)(\omega \wedge \mu) = \iota(v)\omega \wedge \mu + (-1)^k \omega \wedge \iota(v)\mu \tag{2.4.4}$$

satisfies the identity

$$\iota(v)(\iota(w)\omega) = -\iota(w)(\iota(v)\omega) \tag{2.4.5}$$

and, as a special case of (2.4.5), the identity,

$$\iota(v)(\iota(v)\omega) = 0\,. \tag{2.4.6}$$

Moreover, if ω is "decomposable" i.e., is a wedge product of one-forms

$$\omega \;=\; \mu_1 \wedge \cdots \wedge \mu_k\,, \tag{2.4.7}$$

then

$$\iota(v)\omega \;=\; \sum_{r=1}^{k} (-1)^{r-1} (\iota(v)\mu_r)\mu_1 \wedge \cdots \widehat{\mu}_r \cdots \wedge \mu_k\,. \tag{2.4.8}$$

We will also leave for you to prove the following two assertions, both of which are special cases of (2.4.8). If $v = \partial/\partial x_r$ and $\omega = dx_I = dx_{i_1} \wedge \cdots \wedge dx_{i_k}$ then

$$\iota(v)\omega \;=\; \sum_{r=1}^{k} (-1)^r \delta^i_{i_r}\, dx_{I_r} \tag{2.4.9}$$

where

$$\delta^i_{i_r} \;=\; \begin{cases} 1 & i = i_r \\ 0\,, & i \neq i_r \end{cases}.$$

and $I_r = (i_1, \dots, \widehat{i}_r, \dots, i_k)$ and if $v = \sum f_i \, \partial/\partial x_i$ and $\omega = dx_1 \wedge \cdots \wedge dx_n$ then

$$\iota(v)\omega = \sum (-1)^{r-1} f_r \, dx_1 \wedge \cdots \widehat{dx}_r \cdots \wedge dx_n \,. \tag{2.4.10}$$

By combining exterior differentiation with the interior product operation one gets another basic operation of vector fields on forms: the *Lie differentiation* operation. For zero-forms, i.e., for $\mathcal{C}^\infty$ functions, φ, we defined this operation by the formula (2.1.14). For k-forms we'll define it by the slightly more complicated formula

$$L_v \omega = \iota(v) \, d\omega + \, d\iota(v)\omega \,. \tag{2.4.11}$$

(Notice that for zero-forms the second summand is zero, so (2.4.11) and (2.1.14) agree.) If ω is a k-form the right hand side of (2.4.11) is as well, so L_v takes k-forms to k-forms. It also has the property

$$dL_v \omega = L_v \, d\omega \tag{2.4.12}$$

i.e., it "commutes" with d, and the property

$$L_v(\omega \wedge \mu) = L_v \omega \wedge \mu + \omega \wedge L_v \mu \tag{2.4.13}$$

and from these properties it is fairly easy to get an explicit formula for $L_v \omega$. Namely let ω be the k-form

$$\omega = \sum f_I \, dx_I \,, \quad f_I \in \mathcal{C}^\infty(U)$$

and v the vector field

$$\sum g_i \, \partial/\partial x_i \,, \quad g_i \in \mathcal{C}^\infty(U) \,.$$

By (2.4.13)

$$L_v(f_I \, dx_I) = (L_v f_I) \, dx_I + f_I (L_v \, dx_I)$$

and

$$L_v \, dx_I = \sum_{r=1}^{k} dx_{i_1} \wedge \cdots \wedge L_v \, dx_{i_r} \wedge \cdots \wedge dx_{i_k} \,,$$

and by (2.4.12)

$$L_v \, dx_{i_r} = dL_v x_{i_r}$$

so to compute $L_v\omega$ one is reduced to computing $L_v x_{i_r}$ and $L_v f_I$. However by (2.4.13)

$$L_v x_{i_r} \quad = \quad g_{i_r}$$

and

$$L_v f_I \quad = \quad \sum g_i \frac{\partial f_I}{\partial x_i} \, .$$

We will leave the verification of (2.4.12) and (2.4.13) as exercises, and also ask you to prove (by the method of computation that we've just sketched) the *divergence formula*

$$L_v(dx_1 \wedge \cdots \wedge dx_n) = \sum \left(\frac{\partial g_i}{\partial x_i} \right) dx_1 \wedge \cdots \wedge dx_n \, . \tag{2.4.14}$$

Exercises:

1. Verify the assertions (2.4.2)—(2.4.7).

2. Show that if ω is the k-form, dx_I and v the vector field, $\partial/\partial x_r$, then $\iota(v)\omega$ is given by (2.4.9).

3. Show that if ω is the n-form, $dx_1 \wedge \cdots \wedge dx_n$, and v the vector field, $\sum f_i \, \partial/\partial x_i$, $\iota(v)\omega$ is given by (2.4.10).

4. Let U be an open subset of $\mathbb{R}^n$ and v a $\mathcal{C}^\infty$ vector field on U. Show that for $\omega \in \Omega^k(U)$

$$dL_v\omega \quad = \quad L_v \, d\omega$$

and

$$\iota_v L_v \omega \quad = \quad L_v \iota_v \omega \, .$$

Hint: Deduce the first of these identities from the identity $d(d\omega) = 0$ and the second from the identity $\iota(v)(\iota(v)\omega) = 0\,$.)

5. Given $\omega_i \in \Omega^{k_i}(U)$, $i = 1, 2$, show that

$$L_v(\omega_1 \wedge \omega_2) = L_v\omega_1 \wedge \omega_2 + \omega_1 \wedge L_v\omega_2 \, .$$

Hint: Plug $\omega = \omega_1 \wedge \omega_2$ into (2.4.11) and use (2.3.2) and (2.4.4)to evaluate the resulting expression.

6. Let v_1 and v_2 be vector fields on U and let w be their Lie bracket. Show that for $\omega \in \Omega^k(U)$

$$L_w\omega = L_{v_1}(L_{v_2}\omega) - L_{v_2}(L_{v_1}\omega)\,.$$

Hint: By definition this is true for zero-forms and by (2.4.12) for exact one-forms. Now use the fact that every form is a sum of wedge products of zero-forms and one-forms and the fact that L_v satisfies the product identity (2.4.13).

7. Prove the divergence formula (2.4.14).

8. (a) Let $\omega = \Omega^k(\mathbb{R}^n)$ be the form

$$\omega = \sum f_I(x_1,\dots,x_n)\,dx_I$$

and $\mathfrak{v}$ the vector field, $\partial/\partial x_n$. Show that

$$L_{\mathfrak{v}}\omega = \sum \frac{\partial}{\partial x_n} f_I(x_1,\dots,x_n)\,dx_I\,.$$

(b) Suppose $\iota(\mathfrak{v})\omega = L_{\mathfrak{v}}\omega = 0$. Show that ω only depends on $x_1,\dots,x_{k-1}$ and $dx_1,\dots,dx_{k-1}$, i.e., is effectively a k-form on $\mathbb{R}^{n-1}$.

(c) Suppose $\iota(\mathfrak{v})\omega = d\omega = 0$. Show that ω is effectively a closed k-form on $\mathbb{R}^{n-1}$.

(d) Use these results to give another proof of the Poincaré lemma for $\mathbb{R}^n$. Prove by induction on n that every closed form on $\mathbb{R}^n$ is exact.

Hints:

i. Let ω be the form in part (a) and let

$$g_I(x_1,\dots,x_n) = \int_0^{x_n} f_I(x_1,\dots,x_{n-1},t)\,dt\,.$$

Show that if $\nu = \sum g_I\,dx_I$, then $L_{\mathfrak{v}}\nu = \omega$.

ii. Conclude that

$$\omega - d\iota(\mathfrak{v})\nu = \iota(\mathfrak{v})\,d\nu\,. \tag{*}$$

iii. Suppose $d\omega = 0$. Conclude from (*) and from the formula (2.4.6) that the form $\beta = \iota(\mathfrak{v})\,d\nu$ satisfies $d\beta = \iota(\mathfrak{v})\beta = 0$.

iv. By part c, β is effectively a closed form on $\mathbb{R}^{n-1}$, and by induction, $\beta = d\alpha$. Thus by (*)

$$\omega = d\iota(\mathfrak{v})\nu + d\alpha\,.$$

2.5 The pull-back operation on forms

Let U be an open subset of $\mathbb{R}^n$, V an open subset of $\mathbb{R}^m$ and $f : U \to V$ a $\mathcal{C}^\infty$ map. Then for $p \in U$ and $q = f(p)$, the derivative of f at p

$$df_p : T_p\mathbb{R}^n \to T_q\mathbb{R}^m$$

is a linear map, so (as explained in §7 of Chapter 1) one gets from it a pull-back map

$$df_p^* : \Lambda^k(T_q^*\mathbb{R}^m) \to \Lambda^k(T_p^*\mathbb{R}^n)\,. \tag{2.5.1}$$

In particular, let ω be a k-form on V. Then at $q \in V$, ω takes the value

$$\omega_q \in \Lambda^k(T_q^*\mathbb{R}^m)\,,$$

so we can apply to it the operation (2.5.1), and this gives us an element:

$$df_p^*\omega_q \in \Lambda^k(T_p^*\mathbb{R}^n)\,. \tag{2.5.2}$$

In fact we can do this for every point $p \in U$, so this gives us a function,

$$p \in U \to (df_p)^*\omega_q\,, \quad q = f(p)\,. \tag{2.5.3}$$

By the definition of k-form such a function is a k-form on U. We will denote this k-form by $f^*\omega$ and define it to be the *pull-back of ω by the map f*. A few of its basic properties are described below.

1. Let φ be a zero-form, i.e., a function, $\varphi \in \mathcal{C}^\infty(V)$. Since

$$\Lambda^0(T_p^*) = \Lambda^0(T_q^*) = \mathbb{R}$$

the map (2.5.1) is just the identity map of $\mathbb{R}$ onto $\mathbb{R}$ when k is equal to zero. Hence for zero-forms

$$(f^*\varphi)(p) = \varphi(q)\,, \tag{2.5.4}$$

i.e., $f^*\varphi$ is just the composite function, $\varphi \circ f \in \mathcal{C}^\infty(U)$.

2. Let $\mu \in \Omega^1(V)$ be the 1-form, $\mu = d\varphi$. By the chain rule (2.5.2) unwinds to:

$$(df_p)^*d\varphi_q = (d\varphi)_q \circ df_p = d(\varphi \circ f)_p \tag{2.5.5}$$

and hence by (2.5.4)

$$f^*\,d\varphi = df^*\varphi\,. \tag{2.5.6}$$

3. If ω_1 and ω_2 are in $\Omega^k(V)$ we get from (2.5.2)

$$(df_p)^*(\omega_1 + \omega_2)_q = (df_p)^*(\omega_1)_q + (df_p)^*(\omega_2)_q\,,$$

and hence by (2.5.3)

$$f^*(\omega_1 + \omega_2) = f^*\omega_1 + f^*\omega_2\,.$$

4. We observed in § 1.7 that the operation (2.5.1) commutes with wedge-product, hence if ω_1 is in $\Omega^k(V)$ and ω_2 is in $\Omega^\ell(V)$

$$df_p^*(\omega_1)_q \wedge (\omega_2)_q = df_p^*(\omega_1)_q \wedge df_p^*(\omega_2)_q\,.$$

In other words

$$f^*\omega_1 \wedge \omega_2 = f^*\omega_1 \wedge f^*\omega_2\,. \tag{2.5.7}$$

5. Let W be an open subset of $\mathbb{R}^k$ and $g : V \to W$ a $\mathcal{C}^\infty$ map. Given a point $p \in U$, let $q = f(p)$ and $w = g(q)$. Then the composition of the map

$$(df_p)^* : \Lambda^k(T_q^*) \to \Lambda^k(T_p^*)$$

and the map

$$(dg_q)^* : \Lambda^k(T_w^*) \to \Lambda^k(T_q^*)$$

is the map

$$(dg_q \circ df_p)^* : \Lambda^k(T_w^*) \to \Lambda^k(T_p^*)$$

by formula (1.7.4) of Chapter 1. However, by the chain rule

$$(dg_q) \circ (df)_p = d(g \circ f)_p$$

so this composition is the map

$$d(g \circ f)_p^* : \Lambda^k(T_w^*) \to \Lambda^k(T_p^*)\,.$$

Thus if ω is in $\Omega^k(W)$

$$f^*(g^*\omega) = (g \circ f)^*\omega\,. \tag{2.5.8}$$

Let's see what the pull-back operation looks like in coordinates. Using multi-index notation we can express every k-form, $\omega \in \Omega^k(V)$ as a sum over multi-indices of length k

$$\omega = \sum \varphi_I \, dx_I\,, \tag{2.5.9}$$

the coefficient, φ_I, of dx_I being in $\mathcal{C}^\infty(V)$. Hence by (2.5.4)

$$f^*\omega = \sum f^*\varphi_I f^*(dx_I)$$

where $f^*\varphi_I$ is the function of $\varphi \circ f$. What about $f^*\, dx_I$? If I is the multi-index, $(i_1, \ldots, i_k)$, then by definition

$$dx_I = dx_{i_1} \wedge \cdots \wedge dx_{i_k}$$

so

$$d^*\, dx_I = f^*\, dx_i \wedge \cdots \wedge f^*\, dx_{i_k}$$

by (2.5.7), and by (2.5.6)

$$f^*\, dx_i = df^* x_i = df_i$$

where f_i is the i^{th} coordinate function of the map f. Thus, setting

$$df_I = df_{i_1} \wedge \cdots \wedge df_{i_k}\,,$$

we get for each multi-index, I,

$$f^*\, dx_I = df_I \tag{2.5.10}$$

and for the pull-back of the form (2.5.9)

$$f^*\omega = \sum f^*\varphi_I\, df_I\,. \tag{2.5.11}$$

We will use this formula to prove that pull-back commutes with exterior differentiation:

$$d\, f^*\omega = f^*\, d\omega\,. \tag{2.5.12}$$

To prove this we recall that by (2.2.5), $d(\, df_I) = 0$, hence by (2.2.2) and (2.5.10)

$$\begin{aligned} d\, f^*\omega &= \sum d\, f^*\varphi_I \wedge df_I \\ &= \sum f^*\, d\varphi_I \wedge df^*\, dx_I \\ &= f^* \sum d\varphi_I \wedge dx_I \\ &= f^*\, d\omega\,. \end{aligned}$$

A special case of formula (2.5.10) will be needed in Chapter 4: Let U and V be open subsets of $\mathbb{R}^n$ and let $\omega = dx_1 \wedge \cdots \wedge dx_n$. Then by (2.5.10)

$$f^*\omega_p = (df_1)_p \wedge \cdots \wedge (df_n)_p$$

for all $p \in U$. However,

$$(df_i)_p = \sum \frac{\partial f_i}{\partial x_j}(p)(dx_j)_p$$

and hence by formula (1.7.7) of Chapter 1

$$f^*\omega_p = \det\left[\frac{\partial f_i}{\partial x_j}(p)\right](dx_1 \wedge \cdots \wedge dx_n)_p\,.$$

In other words

$$f^*\, dx_1 \wedge \cdots \wedge dx_n = \det\left[\frac{\partial f_i}{\partial x_j}\right] dx_1 \wedge \cdots \wedge dx_n\,. \tag{2.5.13}$$

We will outline in exercises 4 and 5 below the proof of an important topological property of the pull-back operation. Let U be an open subset of $\mathbb{R}^n$, V an open subset of $\mathbb{R}^m$, $A \subseteq \mathbb{R}$ an open interval containing 0 and 1 and $f_i : U \to V$, $i = 0, 1$, a $\mathcal{C}^\infty$ map.

Definition 2.5.1. *A $\mathcal{C}^\infty$ map, $F : U \times A \to V$, is a homotopy between f_0 and f_1 if $F(x, 0) = f_0(x)$ and $F(x, 1) = f_1(x)$.*

Thus, intuitively, f_0 and f_1 are *homotopic* if there exists a family of $\mathcal{C}^\infty$ maps, $f_t : U \to V$, $f_t(x) = F(x, t)$, which "smoothly deform f_0 into f_1". In the exercises mentioned above you will be asked to verify that for f_0 and f_1 to be homotopic they have to satisfy the following criteria.

Theorem 2.5.2. *If f_0 and f_1 are homotopic then for every closed form, $\omega \in \Omega^k(V)$, $f_1^*\omega - f_0^*\omega$ is exact.*

This theorem is closely related to the Poincaré lemma, and, in fact, one gets from it a slightly stronger version of the Poincaré lemma than that described in exercises 5–6 in §2.2.

Definition 2.5.3. *An open subset, U, of $\mathbb{R}^n$ is* contractable *if, for some point $p_0 \in U$, the identity map*

$$f_1 : U \to U\,, \quad f(p) = p\,,$$

is homotopic to the constant map

$$f_0 : U \to U\,, \quad f_0(p) = p_0\,.$$

From the theorem above it's easy to see that the Poincaré lemma holds for contractable open subsets of $\mathbb{R}^n$. If U is contractable every closed k-form on U of degree $k > 0$ is exact. (Proof: Let ω be such a form. Then for the identity map $f_0^*\omega = \omega$ and for the constant map, $f_0^*\omega = 0$.)

Exercises.

1. Let $f : \mathbb{R}^3 \to \mathbb{R}^3$ be the map

$$f(x_1, x_2, x_3) = (x_1 x_2, x_2 x_3^2, x_3^3)\,.$$

Compute the pull-back, $f^*\omega$ for

(a) $\omega = x_2\, dx_3$

(b) $\omega = x_1\, dx_1 \wedge dx_3$

(c) $\omega = x_1\, dx_1 \wedge dx_2 \wedge dx_3$

2. Let $f : \mathbb{R}^2 \to \mathbb{R}^3$ be the map

$$f(x_1, x_2) = (x_1^2, x_2^2, x_1 x_2)\,.$$

Complete the pull-back, $f^*\omega$, for

(a) $\omega = x_2\, dx_2 + x_3\, dx_3$

(b) $\omega = x_1\, dx_2 \wedge dx_3$

(c) $\omega = dx_1 \wedge dx_2 \wedge dx_3$

3. Let U be an open subset of $\mathbb{R}^n$, V an open subset of $\mathbb{R}^m$, $f : U \to V$ a $\mathcal{C}^\infty$ map and $\gamma : [a, b] \to U$ a $\mathcal{C}^\infty$ curve. Show that for $\omega \in \Omega^1(V)$

$$\int_\gamma f^*\omega = \int_{\gamma_1} \omega$$

where $\gamma_1 : [a, b] \to V$ is the curve, $\gamma_1(t) = f(\gamma(t))$. (See § 2.1, exercise 7.)

4. Let U be an open subset of $\mathbb{R}^n$, $A \subseteq \mathbb{R}$ an open interval containing the points, 0 and 1, and (x,t) product coordinates on $U \times A$. Recall (§ 2.2, exercise 5) that a form, $\mu \in \Omega^\ell(U \times A)$ is *reduced* if it can be written as a sum

$$\mu = \sum f_I(x,t)\, dx_I \tag{2.5.14}$$

(i.e., none of the summands involve "dt"). For a reduced form, μ, let $Q\mu \in \Omega^\ell(U)$ be the form

$$Q\mu = \left(\sum \int_0^1 f_I(x,t)\, dt\right) dx_I \tag{2.5.15}$$

and let $\mu_i \in \Omega^\ell(U)$, $i = 0,1$ be the forms

$$\mu_0 \;=\; \sum f_I(x,0)\, dx_I \tag{2.5.16}$$

and

$$\mu_1 \;=\; \sum f_I(x,1)\, dx_I\,. \tag{2.5.17}$$

Now recall that every form, $\omega \in \Omega^k(U \times A)$ can be written uniquely as a sum

$$\omega = \; dt \wedge \alpha + \beta \tag{2.5.18}$$

where α and β are reduced. (See exercise 5 of § 2.3, part a.)

(a) Prove

Theorem 2.5.4. *If the form (2.5.18) is closed then*

$$\beta_0 - \beta_1 = dQ\alpha\,. \tag{2.5.19}$$

Hint: *Formula (2.3.14).*

(b) Let ι_0 and ι_1 be the maps of U into $U \times A$ defined by $\iota_0(x) = (x,0)$ and $\iota_1(x) = (x,1)$. Show that (2.5.19) can be rewritten

$$\iota_0^*\omega - \iota_1^*\omega = \; dQ\alpha\,. \tag{2.5.20}$$

5. Let V be an open subset of $\mathbb{R}^m$ and $f_i : U \to V$, $i = 0,1$, $\mathcal{C}^\infty$ maps. Suppose f_0 and f_1 are homotopic. Show that for every closed form, $\mu \in \Omega^k(V)$, $f_1^*\mu - f_0^*\mu$ is exact. *Hint:* Let $F : U \times A \to V$ be a

homotopy between f_0 and f_1 and let $\omega = F^*\mu$. Show that ω is closed and that $f_0^*\mu = \iota_0^*\omega$ and $f_1^*\mu = \iota_1^*\omega$. Conclude from (2.5.20) that

$$f_0^*\mu - f_1^*\mu = dQ\alpha \tag{2.5.21}$$

where $\omega = dt \wedge \alpha + \beta$ and α and β are reduced.

6. Show that if $U \subseteq \mathbb{R}^n$ is a contractable open set, then the Poincaré lemma holds: every closed form of degree $k > 0$ is exact.

7. An open subset, U, of $\mathbb{R}^n$ is said to be *star-shaped* if there exists a point $p_0 \in U$, with the property that for every point $p \in U$, the line segment,

$$tp + (1-t)p_0\,, \quad 0 \le t \le 1\,,$$

joining p to p_0 is contained in U. Show that if U is star-shaped it is contractable.

8. Show that the following open sets are star-shaped:

(a) The open unit ball

$$\{x \in \mathbb{R}^n\,,\ \|x\| < 1\}\,.$$

(b) The open rectangle, $I_1 \times \cdots \times I_n$, where each I_k is an open subinterval of $\mathbb{R}$.

(c) $\mathbb{R}^n$ itself.

(d) Product sets

$$U_1 \times U_2 \subseteq \mathbb{R}^n = \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$$

where U_i is a star-shaped open set in $\mathbb{R}^{n_i}$.

9. Let U be an open subset of $\mathbb{R}^n$, $f_t : U \to U$, $t \in \mathbb{R}$, a one-parameter group of diffeomorphisms and v its infinitesimal generator. Given $\omega \in \Omega^k(U)$ show that at $t = 0$

$$\frac{d}{dt} f_t^*\omega = L_v\omega\,. \tag{2.5.22}$$

Here is a sketch of a proof:

(a) Let $\gamma(t)$ be the curve, $\gamma(t) = f_t(p)$, and let φ be a zero-form, i.e., an element of $\mathcal{C}^\infty(U)$. Show that

$$f_t^*\varphi(p) = \varphi(\gamma(t))$$

and by differentiating this identity at $t = 0$ conclude that (2.4.40) holds for zero-forms.

(b) Show that if (2.4.40) holds for ω it holds for $d\omega$. *Hint:* Differentiate the identity

$$f_t^*\, d\omega = df_t^*\omega$$

at $t = 0$.

(c) Show that if (2.4.40) holds for ω_1 and ω_2 it holds for $\omega_1 \wedge \omega_2$. *Hint:* Differentiate the identity

$$f_t^*(\omega_1 \wedge \omega_2) = f_t^*\omega_1 \wedge f_t^*\omega_2$$

at $t = 0$.

(d) Deduce (2.4.40) from a, b and c. *Hint:* Every k-form is a sum of wedge products of zero-forms and exact one-forms.

10. In exercise 9 show that for *all* t

$$\frac{d}{dt} f_t^*\omega = f_t^* L_v\omega = L_v f_t^*\omega\,. \tag{2.5.23}$$

Hint: By the definition of "one-parameter group", $f_{s+t} = f_s \circ f_t = f_r \circ f_s$, hence:

$$f_{s+t}^*\omega = f_t^*(f_s^*\omega) = f_s^*(f_t^*\omega)\,.$$

Prove the first assertion by differentiating the first of these identities with respect to s and then setting $s = 0$, and prove the second assertion by doing the same for the second of these identities.

In particular conclude that

$$f_t^* L_v\omega = L_v f_t^*\omega\,. \tag{2.5.24}$$

11. (a) By massaging the result above show that

$$\frac{d}{dt} f_t^*\omega \quad = \quad dQ_t\omega + Q_t\, d\omega \tag{2.5.25}$$

where

$$Q_t\omega \quad = \quad f_t^*\iota(v)\omega\,. \tag{2.5.26}$$

Hint: Formula (2.4.11).

(b) Let

$$Q\omega = \int_0^1 f_t^* \iota(v)\omega \, dt \,.$$

Prove the homotopy indentity

$$f_1^*\omega - f_0^*\omega = \, dQ\omega + Q \, d\omega \,. \tag{2.5.27}$$

12. Let U be an open subset of $\mathbb{R}^n$, V an open subset of $\mathbb{R}^m$, v a vector field on U, w a vector field on V and $f : U \to V$ a $\mathcal{C}^\infty$ map. Show that if v and w are f-related

$$\iota(v) f^* \omega = f^* \iota(w)\omega \,.$$

Hint: Chapter 1, §1.7, exercise 8.

2.6 Div, curl and grad

The basic operations in 3-dimensional vector calculus: grad, curl and div are, by definition, operations on *vector fields*. As we'll see below these operations are closely related to the operations

$$d : \Omega^k(\mathbb{R}^3) \to \Omega^{k+1}(\mathbb{R}^3) \tag{2.6.1}$$

in degrees $k = 0, 1, 2$. However, only two of these operations: grad and div, generalize to n dimensions. (They are essentially the d-operations in degrees zero and $n-1$.) And, unfortunately, there is no simple description in terms of vector fields for the other $n-2$ d-operations. This is one of the main reasons why an adequate theory of vector calculus in n-dimensions forces on one the differential form approach that we've developed in this chapter. Even in three dimensions, however, there is a good reason for replacing grad, div and curl by the three operations, (2.6.1). A problem that physicists spend a lot of time worrying about is the problem of *general covariance*: formulating the laws of physics in such a way that they admit as large a set of symmetries as possible, and frequently these formulations involve differential forms. An example is Maxwell's equations, the fundamental laws of electromagnetism. These are usually expressed as identities involving div and curl. However, as we'll explain below, there is an alternative formulation of Maxwell's equations based on the operations (2.6.1), and from the point of view of general covariance, this formulation is much more satisfactory: the only symmetries of $\mathbb{R}^3$ which preserve div and curl are translations and rotations, whereas the operations (2.6.1) admit all diffeomorphisms of $\mathbb{R}^3$ as symmetries.

To describe how grad, div and curl are related to the operations (2.6.1) we first note that there are two ways of converting vector fields into forms. The first makes use of the natural inner product, $B(v,w) = \sum v_i w_i$, on $\mathbb{R}^n$. From this inner product one gets by § 1.2, exercise 9 a bijective linear map:

$$L : \mathbb{R}^n \to (\mathbb{R}^n)^* \tag{2.6.2}$$

with the defining property: $L(v) = \ell \Leftrightarrow \ell(w) = B(v,w)$. Via the identification (2.1.2) B and L can be transferred to $T_p\mathbb{R}^n$, giving one an inner product, B_p, on $T_p\mathbb{R}^n$ and a bijective linear map

$$L_p : T_p\mathbb{R}^n \to T_p^*\mathbb{R}^n \,. \tag{2.6.3}$$

Hence if we're given a vector field, $\mathfrak{v}$, on U we can convert it into a 1-form, $\mathfrak{v}^\sharp$, by setting

$$\mathfrak{v}^\sharp(p) = L_p\mathfrak{v}(p) \tag{2.6.4}$$

and this sets up a one–one correspondence between vector fields and 1-forms. For instance

$$\mathfrak{v} = \frac{\partial}{\partial x_i} \Leftrightarrow \mathfrak{v}^\sharp = dx_i \,, \tag{2.6.5}$$

(see exercise 3 below) and, more generally,

$$\mathfrak{v} = \sum f_i \frac{\partial}{\partial x_i} \Leftrightarrow \mathfrak{v}^\sharp = \sum f_i \, dx_i \,. \tag{2.6.6}$$

In particular if f is a $\mathcal{C}^\infty$ function on U the vector field "grad f" is by definition

$$\sum \frac{\partial f}{\partial x_i} \frac{\partial}{\partial x_i} \tag{2.6.7}$$

and this gets converted by (2.6.8) into the 1-form, df. Thus the "grad" operation in vector calculus is basically just the operation, $d : \Omega^0(U) \to \Omega^1(U)$.

The second way of converting vector fields into forms is via the interior product operation. Namely let Ω be the n-form, $dx_1 \wedge \cdots \wedge dx_n$. Given an open subset, U of $\mathbb{R}^n$ and a $\mathcal{C}^\infty$ vector field,

$$v = \sum f_i \frac{\partial}{\partial x_i} \tag{2.6.8}$$

on U the interior product of v with Ω is the $(n-1)$-form

$$\iota(v)\Omega = \sum (-1)^{r-1} f_r dx_1 \wedge \cdots \wedge \widehat{dx}_r \cdots \wedge dx_n \,. \tag{2.6.9}$$

Moreover, every $(n-1)$-form can be written uniquely as such a sum, so (2.6.8) and (2.6.9) set up a one-one correspondence between vector fields and $(n-1)$-forms. Under this correspondence the d-operation gets converted into an operation on vector fields

$$v \to d\iota(v)\Omega \,. \tag{2.6.10}$$

Moreover, by (2.4.11)

$$d\iota(v)\Omega = L_v\Omega$$

and by (2.4.14)

$$L_v\Omega = \operatorname{div}(v)\Omega$$

where

$$\operatorname{div}(v) = \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i} \,. \tag{2.6.11}$$

In other words, this correspondence between $(n-1)$-forms and vector fields converts the d-operation into the divergence operation (2.6.11) on vector fields.

Notice that "div" and "grad" are well-defined as vector calculus operations in n-dimensions even though one usually thinks of them as operations in 3-dimensional vector calculus. The "curl" operation, however, is intrinsically a 3-dimensional vector calculus operation. To define it we note that by (2.6.9) every 2-form, μ, can be written uniquely as an interior product,

$$\mu = \iota(\mathfrak{w})\, dx_1 \wedge dx_2 \wedge dx_3 \,, \tag{2.6.12}$$

for some vector field $\mathfrak{w}$, and the left-hand side of this formula determines $\mathfrak{w}$ uniquely. Now let U be an open subset of $\mathbb{R}^3$ and $\mathfrak{v}$ a vector field on U. From $\mathfrak{v}$ we get by (2.6.6) a 1-form, $\mathfrak{v}^\sharp$, and hence by (2.6.12) a vector field, $\mathfrak{w}$, satisfying

$$d\mathfrak{v}^\sharp = \iota(\mathfrak{w})\, dx_1 \wedge dx_2 \wedge dx_3 \,. \tag{2.6.13}$$

The "curl" of $\mathfrak{v}$ is defined to be this vector field, in other words,

$$\operatorname{curl} \mathfrak{v} = \mathfrak{w} \,, \tag{2.6.14}$$

where $\mathfrak{v}$ and $\mathfrak{w}$ are related by (2.6.13).

We'll leave for you to check that this definition coincides with the definition one finds in calculus books. More explicitly we'll leave for you to check that if v is the vector field

$$v = f_1 \frac{\partial}{\partial x_1} + f_2 \frac{\partial}{\partial x_2} + f_3 \frac{\partial}{\partial x_3} \tag{2.6.15}$$

then

$$\operatorname{curl} v = g_1 \frac{\partial}{\partial x_1} + g_2 \frac{\partial}{\partial x_2} + g_3 \frac{\partial}{\partial x_3} \tag{2.6.16}$$

where

$$\begin{aligned} g_1 &= \frac{\partial f_2}{\partial x_3} - \frac{\partial f_3}{\partial x_2} \\ g_2 &= \frac{\partial f_3}{\partial x_1} - \frac{\partial f_1}{\partial x_3} \\ g_3 &= \frac{\partial f_1}{\partial x_2} - \frac{\partial f_2}{\partial x_1} \,. \end{aligned} \tag{2.6.17}$$

To summarize: the grad, curl and div operations in 3-dimensions are basically just the three operations (2.6.1). The "grad" operation is the operation (2.6.1) in degree zero, "curl" is the operation (2.6.1) in degree one and "div" is the operation (2.6.1) in degree two. However, to define "grad" we had to assign an inner product, B_p, to the next tangent space, $T_p\mathbb{R}^n$, for each p in U; to define "div" we had to equip U with the 3-form, Ω, and to define "curl", the most complicated of these three operations, we needed the B_p's *and* Ω. This is why diffeomorphisms preserve the three operations (2.6.1) but don't preserve grad, curl and div. The additional structures which one needs to define grad, curl and div are only preserved by translations and rotations.

We will conclude this section by showing how Maxwell's equations, which are usually formulated in terms of div and curl, can be reset into "form" language. (The paragraph below is an abbreviated version of Guillemin–Sternberg, *Symplectic Techniques in Physics*, §1.20.)

Maxwell's equations assert:

$$\operatorname{div} \mathfrak{v}_E = q \tag{2.6.18}$$

$$\operatorname{curl} \mathfrak{v}_E = -\frac{\partial}{\partial t}\mathfrak{v}_M \tag{2.6.19}$$

$$\operatorname{div} \mathfrak{v}_M = 0 \tag{2.6.20}$$

$$c^2 \operatorname{curl} \mathfrak{v}_M = \mathfrak{w} + \frac{\partial}{\partial t}\mathfrak{v}_E \tag{2.6.21}$$

where $\mathfrak{v}_E$ and $\mathfrak{v}_M$ are the *electric* and *magnetic* fields, q is the *scalar charge density*, $\mathfrak{w}$ is the *current density* and c is the velocity of light. (To simplify (2.6.25) slightly we'll assume that our units of space–time are chosen so that $c = 1$.) As above let $\Omega = dx_1 \wedge dx_2 \wedge dx_3$ and let

$$\mu_E = \iota(\mathfrak{v}_E)\Omega \tag{2.6.22}$$

and

$$\mu_M = \iota(\mathfrak{v}_M)\Omega\,. \tag{2.6.23}$$

We can then rewrite equations (2.6.18) and (2.6.20) in the form

$$d\mu_E = q\Omega \tag{2.6.18$'$}$$

and

$$d\mu_M = 0\,. \tag{2.6.20$'$}$$

What about (2.6.19) and (2.6.21)? We will leave the following "form" versions of these equations as an exercise.

$$d\mathfrak{v}_E^\sharp = -\frac{\partial}{\partial t}\mu_M \tag{2.6.19$'$}$$

and

$$d\mathfrak{v}_M^\sharp = \iota(\mathfrak{w})\Omega + \frac{\partial}{\partial t}\mu_E \tag{2.6.21$'$}$$

where the 1-forms, $\mathfrak{v}_E^\sharp$ and $\mathfrak{v}_M^\sharp$, are obtained from $\mathfrak{v}_E$ and $\mathfrak{v}_M$ by the operation, (2.6.4).

These equations can be written more compactly as differential form identities in $3+1$ dimensions. Let ω_M and ω_E be the 2-forms

$$\omega_M = \mu_M - \mathfrak{v}_E^\sharp \wedge dt \tag{2.6.24}$$

and

$$\omega_E = \mu_E - \mathfrak{v}_M^\sharp \wedge dt \tag{2.6.25}$$

and let Λ be the 3-form

$$\Lambda = q\Omega + \iota(\mathfrak{w})\Omega \wedge dt\,. \tag{2.6.26}$$

We will leave for you to show that the four equations (2.6.18) — (2.6.21) are equivalent to two elegant and compact (3+1)-dimensional identities

$$d\omega_M = 0 \tag{2.6.27}$$

and

$$d\omega_E = \Lambda\,. \tag{2.6.28}$$

Exercises.

1. Verify that the "curl" operation is given in coordinates by the formula (2.6.17).

2. Verify that the Maxwell's equations, (2.6.18) and (2.6.19) become the equations (2.6.20) and (2.6.21) when rewritten in differential form notation.

3. Show that in $(3+1)$-dimensions Maxwell's equations take the form (2.6.17)–(2.6.18).

4. Let U be an open subset of $\mathbb{R}^3$ and v a vector field on U. Show that if v is the gradient of a function, its curl has to be zero.

5. If U is simply connected prove the converse: If the curl of v vanishes, v is the gradient of a function.

6. Let $w = \operatorname{curl} v$. Show that the divergence of w is zero.

7. Is the converse statment true? Suppose the divergence of w is zero. Is $w = \operatorname{curl} v$ for some vector field v?

2.7 Symplectic geometry and classical mechanics

In this section we'll describe some other applications of the theory of differential forms to physics. Before describing these applications, however, we'll say a few words about the geometric ideas that are involved. Let $x_1, \dots, x_{2n}$ be the standard coordinate functions on $\mathbb{R}^{2n}$ and for $i = 1, \dots, n$ let $y_i = x_{i+n}$. The two-form

$$\omega = \sum_{i=1}^{n} dx_i \wedge jy_i \tag{2.7.1}$$

is known as the *Darboux* form. From the identity

$$\omega = -d\left(\sum y_i \, dx_i\right) . \tag{2.7.2}$$

it follows that ω is exact. Moreover computing the n-fold wedge product of ω with itself we get

$$\begin{aligned} \omega^n &= \left(\sum_{i_i=1}^{n} dx_{i_1} \wedge dy_{i_1}\right) \wedge \cdots \wedge \left(\sum_{i_n=1}^{n} dx_{i_n} \wedge dy_{i_n}\right) \\ &= \sum_{i_1,\dots,i_n} dx_{i_1} \wedge dy_{i_1} \wedge \cdots \wedge dx_{i_n} \wedge dy_{i_n} . \end{aligned}$$

We can simplify this sum by noting that if the multi-index, $I = i_1, \dots, i_n$, is repeating the wedge product

$$dx_{i_1} \wedge dy_{i_1} \wedge \cdots \wedge dx_{i_n} \wedge dx_{i_n} \tag{2.7.3}$$

involves two repeating dx_{i_1}'s and hence is zero, and if I is non-repeating we can permute the factors and rewrite (2.7.3) in the form

$$dx_1 \wedge dy_1 \wedge \cdots \wedge dx_n \wedge dy_n .$$

(See §1.6, exercise 5.) Hence since these are exactly $n!$ non-repeating multi-indices

$$\omega^n = n! \, dx_1 \wedge dy_1 \wedge \cdots \wedge dx_n \wedge dy_n$$

i.e.,

$$\frac{1}{n!}\omega^n = \Omega \tag{2.7.4}$$

where

$$\Omega = dx_1 \wedge dy_1 \wedge \cdots \wedge dx_n \wedge dy_n \tag{2.7.5}$$

is the *symplectic volume form* on $\mathbb{R}^{2n}$.

Let U and V be open subsets of $\mathbb{R}^{2n}$. A diffeomorphism $f : U \to V$ is said to be a *symplectic* diffeomorphism (or *symplectomorphism* for short) if $f^*\omega = \omega$. In particular let

$$f_t : U \to U\,, \quad -\infty < t < \infty \tag{2.7.6}$$

be a one-parameter group of diffeomorphisms and let v be the vector field generating (2.7.6). We will say that v is a *symplectic* vector field if the diffeomorphisms, (2.7.6) are symplectomorphisms, i.e., for all t,

$$f_t^*\omega = \omega\,. \tag{2.7.7}$$

Let's see what such vector fields have to look like. Note that by (2.5.23)

$$\frac{d}{dt} f_t^*\omega = f_t^* L_v\omega\,, \tag{2.7.8}$$

hence if $f_t^*\omega = \omega$ for all t, the left hand side of (2.7.8) is zero, so

$$f_t^* L_v\omega = 0\,.$$

In particular, for $t = 0$, f_t is the identity map so $f_t^* L_v\omega = L_v\omega = 0$. Conversely, if $L_v\omega = 0$, then $f_t^* L_v\omega = 0$ so by (2.7.8) $f_t^*\omega$ doesn't depend on t. However, since $f_t^*\omega = \omega$ for $t = 0$ we conclude that $f_t^*\omega = \omega$ for all t. Thus to summarize we've proved

Theorem 2.7.1. *Let $f_t : U \to U$ be a one-parameter group of diffeomorphisms and v the infinitesmal generator of this group. Then v is symplectic of and only if $L_v\omega = 0$.*

There is an equivalent formulation of this result in terms of the interior product, $\iota(v)\omega$. By (2.4.11)

$$L_v\omega = d\iota(v)\omega + \iota(v)\,d\omega\,.$$

But by (2.7.2) $d\omega = 0$ so

$$L_v\omega = d\iota(v)\omega\,.$$

Thus we've shown

Theorem 2.7.2. *The vector field v is symplectic if and only if $\iota(v)\omega$ is closed.*

If $\iota(v)\omega$ is not only closed but is exact we'll say that v is a *Hamiltonian* vector field. In other words v is Hamiltonian if

$$\iota(v)\omega = dH \tag{2.7.9}$$

for some $\mathcal{C}^\infty$ functions, $H \in \mathcal{C}^\infty(U)$.

Let's see what this condition looks like in coordinates. Let

$$v = \sum f_i \frac{\partial}{\partial x_i} + g_i \frac{\partial}{\partial y_i}\,. \tag{2.7.10}$$

Then

$$\begin{aligned} \iota(v)\omega &= \sum_{i,j} f_i \iota\left(\frac{\partial}{\partial x_i}\right) dx_j \wedge dy_j \\ &\quad + \sum_{i,j} g_i \iota\left(\frac{\partial}{\partial y_i}\right) dx_j \wedge dy_i\,. \end{aligned}$$

But

$$\iota\left(\frac{\partial}{\partial x_i}\right) dx_j = \begin{cases} 1 & i = i \\ 0 & i \neq j \end{cases}$$

and

$$\iota\left(\frac{\partial}{\partial x_i}\right) dy_j = 0$$

so the first summand above is

$$\sum f_i\, dy_i$$

and a similar argument shows that the second summand is

$$-\sum g_i\, dx_i\,.$$

Hence if v is the vector field (2.7.10)

$$\iota(v)\omega = \sum f_i\, dy_i - g_i\, dx_i\,. \tag{2.7.11}$$

Thus since

$$dH = \sum \frac{\partial H}{\partial x_i}\, dx_i + \frac{\partial H}{\partial y_i}\, dy_i$$

we get from (2.7.9)–(2.7.11)

$$f_i = \frac{\partial H}{\partial y_i} \text{ and } g_i = -\frac{\partial H}{\partial x_i} \tag{2.7.12}$$

so v has the form:

$$v = \sum \frac{\partial H}{\partial y_i}\frac{\partial}{\partial x_i} - \frac{\partial H}{\partial x_i}\frac{\partial}{\partial y_i}\,. \tag{2.7.13}$$

In particular if $\gamma(t) = (x(t)\,,\, y(t))$ is an integral curve of v it has to satisfy the system of differential equations

$$\begin{aligned} \frac{dx_i}{dt} &= \frac{\partial H}{\partial y_i}(x(t)\,,\, y(t)) \\ \frac{dy_i}{dt} &= -\frac{\partial H}{\partial x_i}(x(t)\,,\, y(t))\,. \end{aligned} \tag{2.7.14}$$

The formulas (2.7.10) and (2.7.11) exhibit an important property of the Darboux form, ω. Every one-form on U can be written uniquely as a sum

$$\sum f_i\, dy_i - g_i\, dx_i$$

with f_i and g_i in $\mathcal{C}^\infty(U)$ and hence (2.7.10) and (2.7.11) imply

Theorem 2.7.3. *The map, $v \to \iota(v)\omega$, sets up a* one-one correspondence *between vector field and one-forms.*

In particular for every $\mathcal{C}^\infty$ function, H, we get by correspondence a unique vector field, $v = v_H$, with the property (2.7.9).

We next note that by (1.7.6)

$$L_v H = \iota(v)\, dH = \iota(v)(\iota(v)\omega) = 0\,.$$

Thus

$$L_v H = 0 \tag{2.7.15}$$

i.e., H is an integral of motion of the vector field, v. In particular if the function, $H : U \to \mathbb{R}$, is proper, then by Theorem 2.1.10 the vector field, v, is complete and hence by Theorem 2.7.1 generates a one-parameter group of symplectomorphisms.

One last comment before we discuss the applications of these results to classical mechanics. If the one-parameter group (2.7.6) is a group of symplectomorphisms then $f_t^*\omega^n = f_t^*\omega \wedge \cdots \wedge f_t^*\omega = \omega^n$ so by (2.7.4)

$$f_t^*\Omega = \Omega \tag{2.7.16}$$

where Ω is the symplectic volume form (2.7.5).

The application we want to make of these ideas concerns the description, in Newtonian mechanics, of a physical system consisting of N interacting point-masses. The *configuration space* of such a system is

$$\mathbb{R}^n = \mathbb{R}^3 \times \cdots \times \mathbb{R}^3 \qquad (N \text{ copies})$$

with position coordinates, $x_1, \ldots, x_n$ and the *phase space* is $\mathbb{R}^{2n}$ with position coordinates $x_1, \ldots, x_n$ and momentum coordinates, $y_1, \ldots, y_n$. The *kinetic energy* of this system is a quadratic function of the momentum coordinates

$$\frac{1}{2}\sum \frac{1}{m_i} y_i^2 \,, \tag{2.7.17}$$

and for simplicity we'll assume that the potential energy is a function, $V(x_1, \ldots, x_n)$, of the position coordinates alone, i.e., it doesn't depend on the momenta and is time-independent as well. Let

$$H = \frac{1}{2}\sum \frac{1}{m_i} y_i^2 + V(x_1, \ldots, x_n) \tag{2.7.18}$$

be the *total energy* of the system. We'll show below that Newton's second law of motion in classical mechanics reduces to the assertion: *the trajectories in phase space of the system above are just the integral curves of the Hamiltonian vector field, v_H.*

Proof. For the function (2.7.18) the equations (2.7.14) become

$$\begin{aligned} \frac{dx_i}{dt} &= \frac{1}{m_i} y_i \\ \frac{dy_i}{dt} &= -\frac{\partial V}{\partial x_i} \,. \end{aligned} \tag{2.7.19}$$

The first set of equation are essentially just the definitions of momenta, however, if we plug them into the second set of equations we get

$$m_i \frac{d^2 x_i}{dt^2} = -\frac{\partial V}{\partial x_i} \tag{2.7.20}$$

and interpreting the term on the right as the force exerted on the i^{th} point-mass and the term on the left as mass times acceleration this equation becomes Newton's second law.

□

In classical mechanics the equations (2.7.14) are known as the Hamilton–Jacobi equations. For a more detailed account of their role in classical mechanics we highly recommend Arnold's book, *Mathematical Methods of Classical Mechanics*. Historically these equations came up for the first time, not in Newtonian mechanics, but in gemometric optics and a brief description of their origins there and of their relation to Maxwell's equations can be found in the bookl we cited above, *Symplectic Techniques in Physics*.

We'll conclude this chapter by mentioning a few implications of the Hamiltonian description (2.7.14) of Newton's equations (2.7.20).

1. *Conservation of energy.* By (2.7.15) the energy function (2.7.18) is constant along the integral curves of v, hence the energy of the system (2.7.14) doesn't change in time.

2. *Noether's principle.* Let $\gamma_t : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ be a one-parameter group of diffeomorphisms of phase space and w its infinitesmal generator. The γ_t's are called a *symmetry* of the system above if

(a) They preserve the function (2.7.18)
and
(b) the vector field w is Hamiltonian.

The condition (b) means that

$$\iota(w)\omega = dG \tag{2.7.21}$$

for some $\mathcal{C}^\infty$ function, G, and what Noether's principle asserts is that *this function is an integral of motion of the system* (2.7.14), i.e., satisfies $L_v G = 0$. In other words stated more succinctly: symmetries of the system (2.7.14) give rise to integrals of motion.

3. *Poincaré recurrence.* An important theorem of Poincaré asserts that if the function $H : \mathbb{R}^{2n} \to \mathbb{R}$ defined by (2.7.18) is proper then every trajectory of the system (2.7.14) returns arbitrarily close to its initial position at some positive time, t_0, and, in fact, does this not just once but does so infinitely often. We'll sketch a proof of this theorem, using (2.7.16), in the next chapter.

Exercises.

1. Let v_H be the vector field (2.7.13). Prove that $\mathrm{div}(v_H) = 0$.

2. Let U be an open subset of $\mathbb{R}^m$, $f_t : U \to U$ a one-parameter group of diffeomorphisms of U and v the infinitesmal generator of this group. Show that if α is a k-form on U then $f_t^*\alpha = \alpha$ for all t if and only if $L_v\alpha = 0$ (i.e., generalize to arbitrary k-forms the result we proved above for the Darboux form).

3. The harmonic oscillator. Let H be the function $\sum_{i=1}^n m_i(x_i^2 + y_i^2)$ where the m_i's are positive constants.

(a) Compute the integral curves of v_H.

(b) Poincaré recurrence. Show that if $(x(t), y(t))$ is an integral curve with initial point $(x_0, y_0) = (x(0), y(0))$ and U an arbitrarily small neighborhood of (x_0, y_0), then for *every* $c > 0$ there exists a $t > c$ such that $(x(t), y(t)) \in U$.

4. Let U be an open subset of $\mathbb{R}^{2n}$ and let H_i, $i = 1, 2$, be in $\mathcal{C}^\infty(U)_i$. Show that

$$[v_{H_1}, v_{H_2}] = v_H \tag{2.7.22}$$

where

$$H = \sum_{i=1}^n \frac{\partial H_1}{\partial x_i}\frac{\partial H_2}{\partial y_i} - \frac{\partial H_2}{\partial x_i}\frac{\partial H_1}{\partial y_i}\,. \tag{2.7.23}$$

5. The expression (2.7.23) is known as the *Poisson bracket* of H_1 and H_2 and is denoted by $\{H_1, H_2\}$. Show that it is anti-symmetric

$$\{H_1, H_2\} = -\{H_2, H_1\}$$

and satisfies Jacobi's identity

$$0 = \{H_1, \{H_2, H_3\}\} + \{H_2, \{H_3, H_1\}\} + \{H_3, \{H_1, H_2\}\}\,.$$

6. Show that

$$\{H_1, H_2\} = L_{v_{H_1}} H_2 = -L_{v_{H_2}} H_1\,. \tag{2.7.24}$$

7. Prove that the following three properties are equivalent.

(a) $\{H_1, H_2\} = 0$.

(b) H_1 is an integral of motion of v_2.

(c) H_2 is an integral of motion of v_1.

8. Verify Noether's principle.

9. Conservation of linear momentum. Suppose the potential, V in (2.7.18) is invariant under the one-parameter group of translations

$$T_t(x_1, \dots, x_n) = (x_1 + t, \dots, x_n + t)\,.$$

(a) Show that the function (2.7.18) is invariant under the group of diffeomorphisms

$$\gamma_t(x, y) = (T_t x, y)\,.$$

(b) Show that the infinitesmal generator of this group is the Hamiltonian vector field v_G where $G = \sum_{i=1}^n y_i$.

(c) Conclude from Noether's principle that this function is an integral of the vector field v_H, i.e., that "total linear moment" is conserved.

(d) Show that "total linear momentum" is conserved if V is the Coulomb potential

$$\sum_{i \neq j} \frac{m_i}{|x_i - x_j|}\,.$$

10. Let $R_t^i : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ be the rotation which fixes the variables, (x_k, y_k), $k \neq i$ and rotates (x_i, y_i) by the angle, t:

$$R_t^i(x_i, y_i) = (\cos t\, x_i + \sin t\, y_i\,,\ -\sin t\, x_i + \cos t\, y_i)\,.$$

(a) Show that R_t^i, $-\infty < t < \infty$, is a one-parameter group of symplectomorphisms.

(b) Show that its generator is the Hamiltonian vector field, v_{H_i}, where $H_i = (x_i^2 + y_i^2)/2$.

(c) Let H be the "harmonic oscillator" Hamiltonian in exercise 3. Show that the R_t^j's preserve H.

(d) What does Noether's principle tell one about the classical mechanical system with energy function H?

11. Show that if U is an open subset of $\mathbb{R}^{2n}$ and v is a symplectic vector field on U then for every point, $p_0 \in U$, there exists a neighborhood, U_0, of p_0 on which v is Hamiltonian.

12. Deduce from exercises 4 and 11 that if v_1 and v_2 are symplectic vector fields on an open subset, U, of $\mathbb{R}^{2n}$ their Lie bracket, $[v_1, v_2]$, is a Hamiltonian vector field.

13. Let α be the one-form, $\sum_{i=1}^n y_i\, dx_i$.

(a) Show that $\omega = -\, d\alpha$.

(b) Show that if α_1 is any one-form on $\mathbb{R}^{2n}$ with the property, $\omega = -d\alpha_1$, then

$$\alpha = \alpha_1 + F$$

for some $\mathcal{C}^\infty$ function F.

(c) Show that $\alpha = \iota(w)\omega$ where w is the vector field

$$-\sum y_i \frac{\partial}{\partial y_i}\,.$$

14. Let U be an open subset of $\mathbb{R}^{2n}$ and v a vector field on U. Show that v has the property, $L_v\alpha = 0$, if and only if

$$\iota(v)\omega = d\iota(v)\alpha\,. \tag{2.7.25}$$

In particular conclude that if $L_v\alpha = 0$ then v is Hamiltonian. *Hint:* (2.7.2).

15. Let H be the function

$$H(x,y) = \sum f_i(x) y_i\,, \tag{2.7.26}$$

where the f_i's are $\mathcal{C}^\infty$ functions on $\mathbb{R}^n$. Show that

$$L_{v_H}\alpha = 0\,. \tag{2.7.27}$$

16. Conversely show that if H is any $\mathcal{C}^\infty$ function on $\mathbb{R}^{2n}$ satisfying (2.7.27) it has to be a function of the form (2.7.26). *Hints:*

(a) Let v be a vector field on $\mathbb{R}^{2n}$ satisfying $L_v\alpha = 0$. By the previous exercise $v = v_H$, where $H = \iota(v)\alpha$.

(b) Show that H has to satisfy the equation

$$\sum_{i=1}^n y_i \frac{\partial H}{\partial y_i} = H\,.$$

(c) Conclude that if $H_r = \frac{\partial H}{\partial y_r}$ then H_r has to satisfy the equation

$$\sum_{i=1}^n y_i \frac{\partial}{\partial y_i} H_r = 0\,.$$

(d) Conclude that H_r has to be constant along the rays (x, ty), $0 \leq t < \infty$.

(e) Conclude finally that H_r has to be a function of x alone, i.e., doesn't depend on y.

17. Show that if $v_{\mathbb{R}^n}$ is a vector field

$$\sum f_i(x) \frac{\partial}{\partial x_i}$$

on configuration space there is a unique lift of $v_{\mathbb{R}^n}$ to phase space

$$v = \sum f_i(x) \frac{\partial}{\partial x_i} + g_i(x,y) \frac{\partial}{\partial y_i}$$

satisfying $L_v\alpha = 0$.

CHAPTER 3

INTEGRATION OF FORMS

3.1 Introduction

The change of variables formula asserts that if U and V are open subsets of $\mathbb{R}^n$ and $f : U \to V$ a C^1 diffeomorphism then, for every continuous function, $\varphi : V \to \mathbb{R}$ the integral

$$\int_V \varphi(y)\, dy$$

exists if and only if the integral

$$\int_U \varphi \circ f(x) |\det Df(x)|\, dx$$

exists, and if these integrals exist they are equal. Proofs of this can be found in [**?**], [**?**] or [**?**]. This chapter contains an alternative proof of this result. This proof is due to Peter Lax. Our version of his proof in §3.5 below makes use of the theory of differential forms; but, as Lax shows in the article [**?**] (which we strongly recommend as collateral reading for this course), references to differential forms can be avoided, and the proof described in§3.5 can be couched entirely in the language of elementary multivariable calculus.

The virtue of Lax's proof is that is allows one to prove a version of the change of variables theorem for other mappings besides diffeomorphisms, and involves a topological invariant, the *degree of a mapping*, which is itself quite interesting. Some properties of this invariant, and some topological applications of the change of variables formula will be discussed in §3.6 of these notes.

Remark 3.1.1. *The proof we are about to describe is somewhat simpler and more transparent if we assume that f is a $\mathcal{C}^\infty$ diffeomorphism. We'll henceforth make this assumption.*

3.2 The Poincaré lemma for compactly supported forms on rectangles

Let ν be a k-form on $\mathbb{R}^n$. We define the *support* of ν to be the closure of the set

$$\{x \in \mathbb{R}^n\,,\ \nu_x \neq 0\}$$

and we say that ν is *compactly* supported if $\operatorname{supp} \nu$ is compact. We will denote by $\Omega^k_c(\mathbb{R}^n)$ the set of all $\mathcal{C}^\infty$ k-forms which are compactly supported, and if U is an open subset of $\mathbb{R}^n$, we will denote by $\Omega^k_c(U)$ the set of all compactly supported k-forms whose support is contained in U .

Let $\omega = f\, dx_1 \wedge \cdots \wedge dx_n$ be a compactly supported n-form with $f \in \mathcal{C}^\infty_0(\mathbb{R}^n)$. We will define the integral of ω over $\mathbb{R}^n$:

$$\int_{\mathbb{R}^n} \omega$$

to be the usual integral of f over $\mathbb{R}^n$

$$\int_{\mathbb{R}^n} f\, dx\,.$$

(Since f is $\mathcal{C}^\infty$ and compactly supported this integral is well-defined.)

Now let Q be the rectangle

$$[a_1, b_1] \times \cdots \times [a_n, b_n]\,.$$

The Poincaré lemma for rectangles asserts:

Theorem 3.2.1. *Let ω be a compactly supported n-form, with* $\operatorname{supp}\omega \subseteq \operatorname{Int} Q$*. Then the following assertions are equivalent:*

a. $\int \omega = 0$.

b. *There exists a compactly supported $(n-1)$-form, μ, with* $\operatorname{supp}\mu \subseteq \operatorname{Int} Q$ *satisfying $d\mu = \omega$.*

We will first prove that (b)⇒(a). Let

$$\mu = \sum_{i=1}^n f_i\, dx_1 \wedge \ldots \wedge \widehat{dx_i} \wedge \ldots \wedge dx_n\,,$$

(the "hat" over the dx_i meaning that dx_i has to be omitted from the wedge product). Then

$$d\mu = \sum_{i=1}^{n} (-1)^{i-1} \frac{\partial f_i}{\partial x_i} dx_1 \wedge \ldots \wedge dx_n \,,$$

and to show that the integral of $d\mu$ is zero it suffices to show that each of the integrals

$$\int_{\mathbb{R}^n} \frac{\partial f}{\partial x_i}\, dx \tag{2.1$_i$}$$

is zero. By Fubini we can compute $(2.1)_i$ by first integrating with respect to the variable, x_i, and then with respect to the remaining variables. But

$$\int \frac{\partial f}{\partial x_i}\, dx_i = f(x) \Big|_{x_i=a_i}^{x_i=b_i} = 0$$

since f_i is supported on U.

We will prove that (a) $\Rightarrow$ (b) by proving a somewhat stronger result. Let U be an open subset of $\mathbb{R}^m$. We'll say that U has *property* P if every form, $\omega \in \Omega_c^m(U)$ whose integral is zero in $d\,\Omega_c^{m-1}(U)$.

We will prove

Theorem 3.2.2. *Let U be an open subset of $\mathbb{R}^{n-1}$ and $A \subseteq \mathbb{R}$ an open interval. Then if U has property P, $U \times A$ does as well.*

Remark 3.2.3. *It's very easy to see that the open interval A itself has property P. (See exercise 1 below.) Hence it follows by induction from Theorem 3.2.2 that*

$$\operatorname{Int} Q = A_1 \times \cdots \times A_n, \quad A_i = (a_i, b_i)$$

has property P, and this proves "(a) $\Rightarrow$ (b)".

To prove Theorem 3.2.2 let $(x,t) = (x_1, \ldots, x_{n-1}, t)$ be product coordinates on $U \times A$. Given $\omega \in \Omega_c^n(U \times A)$ we can express ω as a wedge product, $dt \wedge \alpha$ with $\alpha = f(x,t)\, dx_1 \wedge \cdots \wedge dx_{n-1}$ and $f \in \mathcal{C}_0^\infty(U \times A)$. Let $\theta \in \Omega_c^{n-1}(U)$ be the form

$$\theta = \left(\int_A f(x,t)\, dt \right) dx_1 \wedge \cdots \wedge dx_{n-1} \,. \tag{3.2.1}$$

Then

$$\int_{\mathbb{R}^{n-1}} \theta = \int_{\mathbb{R}^n} f(x,t)\, dx\, dt = \int_{\mathbb{R}^n} \omega$$

so if the integral of ω is zero, the integral of θ is zero. Hence since U has property P, $\beta = d\nu$ for some $\nu \in \Omega_c^{n-1}(U)$. Let $\rho \in \mathcal{C}^\infty(\mathbb{R})$ be a bump function which is supported on A and whose integral over A is one. Setting

$$\kappa = -\rho(t)\, dt \wedge \nu$$

we have

$$d\kappa = \rho(t)\, dt \wedge\, d\nu = \rho(t)\, dt \wedge \theta\,,$$

and hence

$$\omega - d\kappa \;=\; dt \wedge (\alpha - \rho(t)\theta) = dt \wedge u(x,t)\, dx_1 \wedge \cdots \wedge\, dx_{n-1}$$

where

$$u(x,t) \;=\; f(x,t) - \rho(t) \int_A f(x,t)\, dt$$

by (3.2.1). Thus

$$\int u(x,t)\, dt = 0\,. \tag{3.2.2}$$

Let a and b be the end points of A and let

$$v(x,t) = \int_a^t i(x,s)\, ds\,. \tag{3.2.3}$$

By (3.2.2) $v(a,x) = v(b,x) = 0$, so v is in $\mathcal{C}_0^\infty(U \times A)$ and by (3.2.3), $\partial v/\partial t = u$. Hence if we let γ be the form, $v(x,t)\, dx_1 \wedge \cdots \wedge dx_{n-1}$, we have:

$$d\gamma \;=\; u(x,t)\, dx \wedge \cdots \wedge dx_{n-1} = \omega - d\kappa$$

and

$$\omega \;=\; d(\gamma + \kappa)\,.$$

Since γ and κ are both in $\Omega_c^{n-1}(U \times A)$ this proves that ω is in $d\,\Omega_c^{n-1}(U \times A)$ and hence that $U \times A$ has property P.

Exercises for §3.2.

1. Let $f : \mathbb{R} \to \mathbb{R}$ be a compactly supported function of class C^r with support on the interval, (a,b). Show that the following are equivalent.

(a) $\int_a^b f(x)\,dx = 0$.

(b) There exists a function, $g : \mathbb{R} \to \mathbb{R}$ of class C^{r+1} with support on (a,b) with $\frac{dg}{dx} = f$.

Hint: Show that the function

$$g(x) = \int_a^x f(s)\,ds$$

is compactly supported.

2. Let $f = f(x,y)$ be a compactly supported function on $\mathbb{R}^k \times \mathbb{R}^\ell$ with the property that the partial derivatives

$$\frac{\partial f}{\partial x_i}(x,y)\,,\ i = 1,\ldots,k\,,$$

and are continuous as functions of x and y. Prove the following "differentiation under the integral sign" theorem (which we implicitly used in our proof of Theorem 3.2.2).

Theorem 3.2.4. *The function*

$$g(x) = \int f(x,y)\,dy$$

is of class C^1 *and*

$$\frac{\partial g}{\partial x_i}(x) = \int \frac{\partial f}{\partial x_i}(x,y)\,dy\,.$$

Hints: For y fixed and $h \in \mathbb{R}^k$,

$$f_i(x+h,y) - f_i(x,y) = D_x f_i(c)h$$

for some point, c, on the line segment joining x to $x+c$. Using the fact that $D_x f$ is continuous as a function of x and y and compactly supported, conclude:

Lemma 3.2.5. *Given* $\epsilon > 0$ *there exists a* $\delta > 0$ *such that for* $|h| \le \delta$

$$|f(x+h,y) - f(x,y) - D_x f(x,c)h| \le \epsilon |h|\,.$$

Now let $Q \subseteq \mathbb{R}^\ell$ be a rectangle with $\operatorname{supp} f \subseteq \mathbb{R}^k \times Q$ and show that

$$|g(x+h) - g(x) - \left(\int D_x f(x,y)\,dy\right) h| \le \epsilon \operatorname{vol}(Q)|h|\,.$$

Conclude that g is differentiable at x and that its derivative is

$$\int D_x f(x,y)\,dy\,.$$

3. Let $f : \mathbb{R}^k \times \mathbb{R}^\ell \to \mathbb{R}$ be a compactly supported continuous function. Prove

Theorem 3.2.6. *If all the partial derivatives of $f(x,y)$ with respect to x of order $\le r$ exist and are continuous as functions of x and y the function*

$$g(x) = \int f(x,y)\,dy$$

is of class C^r.

4. Let U be an open subset of $\mathbb{R}^{n-1}$, $A \subseteq \mathbb{R}$ an open interval and (x,t) product coordinates on $U \times A$. Recall (§2.2) exercise 5) that every form, $\omega \in \Omega^k(U \times A)$, can be written uniquely as a sum, $\omega = dt \wedge \alpha + \beta$ where α and β are *reduced*, i.e., don't contain a factor of dt.

(a) Show that if ω is compactly supported on $U \times A$ then so are α and β.

(b) Let $\alpha = \sum\limits_I f_I(x,t)\,dx_I$. Show that the form

$$\theta = \sum_I \left(\int_A f_I(x,t)\,dt\right) dx_I \tag{3.2.4}$$

is in $\Omega_c^{k-1}(U)$.

(c) Show that if $d\omega = 0$, then $d\theta = 0$. *Hint:* By (3.2.4)

$$\begin{aligned} d\theta &= \sum_{I,i} \left(\int_A \frac{\partial f_I}{\partial x_i}(x,t)\,dt\right) dx_i \wedge dx_I \\ &= \int_A (d_U \alpha)\,dt \end{aligned}$$

and by (**??**) $d_U \alpha = \dfrac{d\beta}{dt}$.

5. In exercise 4 show that if θ is in $d\,\Omega^{k-1}(U)$ then ω is in $d\,\Omega^k_c(U)$.
Hints:

(a) Let $\theta = d\nu$, with $\nu = \Omega^{k-2}_c(U)$ and let $\rho \in \mathcal{C}^\infty(\mathbb{R})$ be a bump function which is supported on A and whose integral over A is one. Setting $k = -\rho(t)\,dt \wedge \nu$ show that

$$\begin{aligned} \omega - d\kappa &= dt \wedge (\alpha - \rho(t)\theta) + \beta \\ &= dt \wedge (\sum_I u_I(x,t)\,dx_I) + \beta \end{aligned}$$

where

$$u_I(x,t) = f_I(x,t) - \rho(t) \int_A f_I(x,t)\,dt\,.$$

(b) Let a and b be the end points of A and let

$$v_I(x,t) = \int_a^t u_I(x,t)\,dt\,.$$

Show that the form $\sum v_I(x,t)\,dx_I$ is in $\Omega^{k-1}_c(U \times A)$ and that

$$d\gamma = \omega - d\kappa - \beta - d_U\gamma\,.$$

(c) Conclude that the form $\omega - d(\kappa + \gamma)$ is reduced.

(d) Prove: If $\lambda \in \Omega^k_c(U \times A)$ is reduced and $d\lambda = 0$ then $\lambda = 0$.
Hint: Let $\lambda = \sum g_I(x,t)\,dx_I$. Show that $d\lambda = 0 \Rightarrow \dfrac{\partial}{\partial t} g_I(x,t) = 0$ and exploit the fact that for fixed x, $g_I(x,t)$ is compactly supported in t.

6. Let U be an open subset of $\mathbb{R}^m$. We'll say that U has *property* P_k, for $k < n$, if every closed k-form, $\omega \in \Omega^k_c(U)$, is in $d\,\Omega^{k-1}_c(U)$. Prove that if the open set $U \subseteq \mathbb{R}^{n-1}$ in exercise 3 has property P_k then so does $U \times A$.

7. Show that if Q is the rectangle $[a_1, b_1] \times \cdots \times [a_n, b_n]$ and $U = \operatorname{Int} Q$ then u has property P_k.

8. Let $\mathbb{H}^n$ be the half-space

$$\{(x_1, \ldots, x_n)\,; \quad x_1 \leq 0\} \tag{3.2.5}$$

and let $\omega \in \Omega_c^n(\mathbb{R})$ be the n-form, $f\,dx_1 \wedge \cdots \wedge dx_n$ with $f \in \mathcal{C}_0^\infty(\mathbb{R}^n)$. Define:

$$\int_{\mathbb{H}^n} \omega = \int_{\mathbb{H}^n} f(x_1, \ldots, x_n)\,dx_1 \cdots dx_n \tag{3.2.6}$$

where the right hand side is the usual Riemann integral of f over $\mathbb{H}^n$. (This integral makes sense since f is compactly supported.) Show that if $\omega = d\mu$ for some $\mu \in \Omega_c^{n-1}(\mathbb{R}^n)$ then

$$\int_{\mathbb{H}^n} \omega = \int_{\mathbb{R}^{n-1}} \iota^* \mu \tag{3.2.7}$$

where $\iota : \mathbb{R}^{n-1} \to \mathbb{R}^n$ is the inclusion map

$$(x_2, \ldots, x_n) \to (0, x_2, \ldots, x_n)\,.$$

Hint: Let $\mu = \sum_i f_i\,dx_1 \wedge \cdots \widehat{dx_i} \cdots \wedge dx_n$. Mimicking the "(b) $\Rightarrow$ (a)" part of the proof of Theorem 3.2.1 show that the integral (3.2.6) is the integral over $\mathbb{R}^{n-1}$ of the function

$$\int_{-\infty}^0 \frac{\partial f_1}{\partial x_1}(x_1, x_2, \ldots, x_n)\,dx_1\,.$$

3.3 The Poincaré lemma for compactly supported forms on open subsets of $\mathbb{R}^n$

In this section we will generalize Theorem 3.2.1 to arbitrary connected open subsets of $\mathbb{R}^n$.

Theorem 3.3.1. *Let U be a connected open subset of $\mathbb{R}^n$ and let ω be a compactly supported n-form with* $\operatorname{supp}\omega \subset U$. *The the following assertions are equivalent,*

a. $\int \omega = 0$.

b. *There exists a compactly supported $(n-1)$-form, μ, with* $\operatorname{supp}\mu \subseteq U$ *and $\omega = d\mu$.*

Proof that (b) $\Rightarrow$ (a). The support of μ is contained in a large rectangle, so the integral of $d\mu$ is zero by Theorem 3.2.1.

Proof that (a) $\Rightarrow$ (b): Let ω_1 and ω_2 be compactly supported n-forms with support in U. We will write

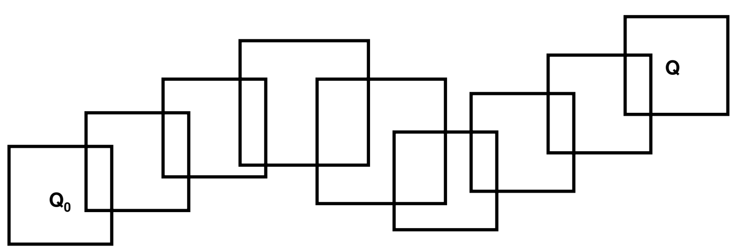
$$\omega_1 \sim \omega_2$$

as shorthand notation for the statement: "There exists a compactly supported $(n-1)$-form, μ, with support in U and with $\omega_1 - \omega_2 = d\mu$.", We will prove that (a) $\Rightarrow$ (b) by proving an equivalent statement: Fix a rectangle, $Q_0 \subset U$ and an n-form, ω_0, with $\operatorname{supp}\omega_0 \subseteq Q_0$ and integral equal to one.

Theorem 3.3.2. *If ω is a compactly supported n-form with $\operatorname{supp}\omega \subseteq U$ and $c = \int \omega$ then $\omega \sim c\omega_0$.*

Thus in particular if $c = 0$, Theorem 3.3.2 says that $\omega \sim 0$ proving that (a) $\Rightarrow$ (b).

To prove Theorem 3.3.2 let $Q_i \subseteq U$, $i = 1, 2, 3, \ldots$, be a collection of rectangles with $U = \cup \operatorname{Int} Q_i$ and let φ_i be a partition of unity with $\operatorname{supp}\varphi_i \subseteq \operatorname{Int} Q_i$. Replacing ω by the finite sum $\sum_{i=1}^m \varphi_i \omega$, m large, it suffices to prove Theorem 3.3.2 for each of the summands $\varphi_i \omega$. In other words we can assume that $\operatorname{supp}\omega$ is contained in one of the open rectangles, $\operatorname{Int} Q_i$. Denote this rectangle by Q. We claim that one can join Q_0 to Q by a sequence of rectangles as in the figure below.



Lemma 3.3.3. *There exists a sequence of rectangles, R_i, $i = 0, \ldots, N+1$ such that $R_0 = Q_0$, $R_{N+1} = Q$ and $\operatorname{Int} R_i \cap \operatorname{Int} R_{i+1}$ is non-empty.*

Proof. Denote by A the set of points, $x \in U$, for which there exists a sequence of rectangles, R_i, $i = 0, \ldots, N+1$ with $R_0 = Q_0$, with $x \in \operatorname{Int} R_{N+1}$ and with $\operatorname{Int} R_i \cap \operatorname{Int} R_{i+1}$ non-empty. It is clear that this set is open and that its complement is open; so, by the connectivity of U, $U = A$. $\square$

To prove Theorem 3.3.2 with $\operatorname{supp}\omega \subseteq Q$, select, for each i, a compactly supported n-form, ν_i, with $\operatorname{supp}\nu_i \subseteq \operatorname{Int} R_i \cap \operatorname{Int} R_{i+1}$ and with $\int \nu_i = 1$. The difference, $\nu_i - \nu_{i+1}$ is supported in $\operatorname{Int} R_{i+1}$, and its integral is zero; so by Theorem 3.2.1, $\nu_i \sim \nu_{i+1}$. Similarly, $\omega_0 \sim \nu_1$ and, if $c = \int \omega$, $\omega \sim c\nu_N$. Thus

$$c\omega_0 \sim c\nu_0 \sim \cdots \sim c\nu_N = \omega$$

proving the theorem.

3.4 The degree of a differentiable mapping

Let U and V be open subsets of $\mathbb{R}^n$ and $\mathbb{R}^k$. A continuous mapping, $f : U \to V$, is *proper* if, for every compact subset, B, of V, $f^{-1}(B)$ is compact. Proper mappings have a number of nice properties which will be investigated in the exercises below. One obvious property is that if f is a $\mathcal{C}^\infty$ mapping and ω is a compactly supported k-form with support on V, $f^*\omega$ is a compactly supported k-form with support on U. Our goal in this section is to show that if U and V are connected open subsets of $\mathbb{R}^n$ and $f : U \to V$ is a proper $\mathcal{C}^\infty$ mapping then there exists a topological invariant of f, which we will call its *degree* (and denote by $\deg(f)$), such that the "change of variables" formula:

$$\int_U f^*\omega = \deg(f) \int_V \omega \tag{3.4.1}$$

holds for all $\omega \in \Omega_c^n(V)$.

Before we prove this assertion let's see what this formula says in coordinates. If

$$\omega = \varphi(y)\, dy_1 \wedge \cdots \wedge dy_n$$

then at $x \in U$

$$f^*\omega = (\varphi \circ f)(x) \det(Df(x))\, dx_1 \wedge \cdots \wedge dx_n\,;$$

so, in coordinates, (3.4.1) takes the form

$$\int_V \varphi(y)\, dy = \deg(f) \int_U \varphi \circ f(x) \det(Df(x))\, dx\,. \tag{3.4.2}$$

Proof of 3.4.1. Let ω_0 be an n-form of compact support with $\operatorname{supp}\omega_0 \subset V$ and with $\int \omega_0 = 1$. If we set $\deg f = \int_U f^*\omega_0$ then (3.4.1) clearly holds for ω_0. We will prove that (3.4.1) holds for every compactly supported n-form, ω, with $\operatorname{supp}\omega \subseteq V$. Let $c = \int_V \omega$. Then by Theorem 3.1 $\omega - c\omega_0 = d\mu$, where μ is a completely supported $(n-1)$-form with $\operatorname{supp}\mu \subseteq V$. Hence

$$f^*\omega - cf^*\omega_0 = f^*\,d\mu = d\,f^*\mu\,,$$

and by part (a) of Theorem 3.1

$$\int_U f^*\omega = c\int f^*\omega_0 = \deg(f)\int_V \omega\,.$$

□

We will show in § 3.6 that the degree of f is always an integer and explain why it is a "topological" invariant of f. For the moment, however, we'll content ourselves with pointing out a simple but useful property of this invariant. Let U, V and W be connected open subsets of $\mathbb{R}^n$ and $f : U \to V$ and $g : V \to W$ proper $\mathcal{C}^\infty$ mappings. Then

$$\deg(g \circ f) = \deg(g)\deg(f)\,. \tag{3.4.3}$$

Proof. Let ω be a compactly supported n-form with support on W. Then

$$(g \circ f)^*\omega = g^*f^*\omega\,;$$

so

$$\begin{aligned}\int_U (g\circ f)^*\omega &= \int_U g^*(f^*\omega) = \deg(g)\int_V f^*\omega \\ &= \deg(g)\deg(f)\int_W \omega\,.\end{aligned}$$

□

From this multiplicative property it is easy to deduce the following result (which we will need in the next section).

Theorem 3.4.1. *Let A be a non-singular $n \times n$ matrix and $f_A : \mathbb{R}^n \to \mathbb{R}^n$ the linear mapping associated with A. Then* $\deg(f_A) = +1$ *if* $\det A$ *is positive and* -1 *if* $\det A$ *is negative.*

A proof of this result is outlined in exercises 5–9 below.

Exercises for §3.4.

1. Let U be an open subset of $\mathbb{R}^n$ and φ_i, $i = 1, 2, 3, \ldots$, a partition of unity on U. Show that the mapping, $f : U \to \mathbb{R}$ defined by

$$f = \sum_{k=1}^{\infty} k\varphi_k$$

is a proper $\mathcal{C}^\infty$ mapping.

2. Let U and V be open subsets of $\mathbb{R}^n$ and $\mathbb{R}^k$ and let $f : U \to V$ be a proper continuous mapping. Prove:

Theorem 3.4.2. *If B is a compact subset of V and $A = f^{-1}(B)$ then for every open subset, U_0, with $A \subseteq U_0 \subseteq U$, there exists an open subset, V_0, with $B \subseteq V_0 \subseteq V$ and $f^{-1}(V_0) \subseteq U_0$.*

Hint: Let C be a compact subset of V with $B \subseteq \operatorname{Int} C$. Then the set, $W = f^{-1}(C) - U_0$ is compact; so its image, $f(W)$, is compact. Show that $f(W)$ and B are disjoint and let

$$V_0 = \operatorname{Int} C - f(W)\,.$$

3. Show that if $f : U \to V$ is a proper continuous mapping and X is a closed subset of U, $f(X)$ is closed.

Hint: Let $U_0 = U - X$. Show that if p is in $V - f(X)$, $f^{-1}(p)$ is contained in U_0 and conclude from the previous exercise that there exists a neighborhood, V_0, of p such that $f^{-1}(V_0)$ is contained in U_0. Conclude that V_0 and $f(X)$ are disjoint.

4. Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be the translation, $f(x) = x + a$. Show that $\deg(f) = 1$.

Hint: Let $\psi : \mathbb{R} \to \mathbb{R}$ be a compactly supported $\mathcal{C}^\infty$ function. For $a \in \mathbb{R}$, the identity

$$\int \psi(t)\,dt = \int \psi(t-a)\,dt \tag{3.4.4}$$

is easy to prove by elementary calculus, and this identity proves the assertion above in dimension one. Now let

$$\varphi(x) = \psi(x_1)\ldots\varphi(x_n) \tag{3.4.5}$$

and compute the right and left sides of (3.4.2) by Fubini's theorem.

5. Let σ be a permutation of the numbers, $1, \dots, n$ and let $f_\sigma : \mathbb{R}^n \to \mathbb{R}^n$ be the diffeomorphism, $f_\sigma(x_1, \dots, x_n) = (x_{\sigma(1)}, \dots, x_{\sigma(n)})$. Prove that $\deg f_\sigma = \operatorname{sgn}(\sigma)$.

Hint: Let φ be the function (3.4.5). Show that if ω is equal to $\varphi(x)\, dx_1 \wedge \cdots \wedge dx_n$, $f^*\omega = (\operatorname{sgn} \sigma)\omega$.

6. Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be the mapping

$$f(x_1, \dots, x_n) = (x_1 + \lambda x_2, x_2, \dots, x_n).$$

Prove that $\deg(f) = 1$.

Hint: Let $\omega = \varphi(x_1, \dots, x_n)\, dx_1 \wedge \ldots \wedge dx_n$ where $\varphi : \mathbb{R}^n \to \mathbb{R}$ is compactly supported and of class $\mathcal{C}^\infty$. Show that

$$\int f^*\omega = \int \varphi(x_1 + \lambda x_2, x_2, \dots, x_n)\, dx_1 \dots dx_n$$

and evaluate the integral on the right by Fubini's theorem; i.e., by first integrating with respect to the x_1 variable and then with respect to the remaining variables. Note that by (3.4.4)

$$\int f(x_1 + \lambda x_2, x_2, \dots, x_n)\, dx_1 = \int f(x_1, x_2, \dots, x_n)\, dx_1\,.$$

7. Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be the mapping

$$f(x_1, \dots, x_n) = (\lambda x_1, x_2, \dots, x_n)$$

with $\lambda \neq 0$. Show that $\deg f = +1$ if λ is positive and -1 if λ is negative.

Hint: In dimension 1 this is easy to prove by elementary calculus techniques. Prove it in d-dimensions by the same trick as in the previous exercise.

8. (a) Let $e_1, \dots, e_n$ be the standard basis vectors of $\mathbb{R}^n$ and A, B and C the linear mappings

$$\begin{aligned} Ae_1 &= e, \quad Ae_i = \sum_j a_{j,i} e_j\,, \qquad i > 1 \\ Be_i &= e_i\,, \quad i > 1\,, \quad Be_1 = \sum_{j=1}^n b_j e_j \\ Ce_1 &= e_1\,, \quad Ce_i = e_i + c_i e_1\,, \quad i > 1\,. \end{aligned} \tag{3.4.6}$$

Show that

$$BACe_1 \;=\; \sum b_j e_j$$

and

$$BACe_i \;=\; \sum_{j}^{n} = (a_{j,i} + c_i b_j)e_j + c_i b_1 e_1$$

for $i > 1$.

(b)

$$Le_i = \sum_{j=1}^{n} \ell_{j,i} e_j\,, \quad i = 1, \dots, n\,. \tag{3.4.7}$$

Show that if $\ell_{1,1} \neq 0$ one can write L as a product, $L = BAC$, where A, B and C are linear mappings of the form (3.4.6).

Hint: First solve the equations

$$\ell_{j,1} = b_j$$

for $j = 1, \dots, n$, then the equations

$$\ell_{1,i} = b_1 c_i$$

for $i > 1$, then the equations

$$\ell_{j,i} = a_{j,i} + c_i b_j$$

for $i, j > 1$.

(c) Suppose L is invertible. Conclude that A, B and C are invertible and verify that Theorem 3.4.1 holds for B and C using the previous exercises in this section.

(d) Show by an inductive argument that Theorem 3.4.1 holds for A and conclude from (3.4.3) that it holds for L.

9. To show that Theorem 3.4.1 holds for an arbitrary linear mapping, L, of the form (3.4.7) we'll need to eliminate the assumption: $\ell_{1,1} \neq 0$. Show that for some j, $\ell_{j,1}$ is non-zero, and show how to eliminate this assumption by considering $f_\sigma \circ L$ where σ is the transposition, $1 \leftrightarrow j$.

10. Here is an alternative proof of Theorem 4.3.1 which is shorter than the proof outlined in exercise 9 but uses some slightly more sophisticated linear algebra.

(a) Prove Theorem 3.4.1 for linear mappings which are *orthogonal*, i.e., satisfy $L^t L = I$.

Hints:

i. Show that $L^*(x_1^2 + \cdots + x_n^2) = x_1^2 + \cdots + x_n^2$.
ii. Show that $L^*(dx_1 \wedge \cdots \wedge dx_n)$ is equal to $dx_1 \wedge \cdots \wedge dx_n$ or $-dx_1 \wedge \cdots \wedge dx_n$ depending on whether L is orientation preserving or orinetation reversing. (See § 1.2, exercise 10.)
iii. Let ψ be as in exercise 4 and let ω be the form

$$\omega = \psi(x_1^2 + \cdots + x_n^2)\, dx_1 \wedge \cdots \wedge dx_n\,.$$

Show that $L^*\omega = \omega$ if L is orientation preserving and $L^*\omega = -\omega$ if L is orientation reversing.

(b) Prove Theorem 3.4.1 for linear mappings which are *self-adjoint* (satisfy $L^t = L$). *Hint:* A self-adjoint linear mapping is diagonizable: there exists an intervertible linear mapping, $M : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$M^{-1} L M e_i = \lambda_i e_i\,, \quad i = 1, \ldots, n\,. \tag{3.4.8}$$

(c) Prove that every invertible linear mapping, L, can be written as a product, $L = BC$ where B is orthogonal and C is self-adjoint.

Hints:

i. Show that the mapping, $A = L^t L$, is self-adjoint and that it's eigenvalues, the λ_i's in 3.4.8, are positive.
ii. Show that there exists an invertible self-adjoint linear mapping, C, such that $A = C^2$ and $AC = CA$.
iii. Show that the mapping $B = LC^{-1}$ is orthogonal.

3.5 The change of variables formula

Let U and V be connected open subsets of $\mathbb{R}^n$. If $f : U \to V$ is a diffeomorphism, the determinant of $Df(x)$ at $x \in U$ is non-zero, and hence, since it is a continuous function of x, its sign is the same at every point. We will say that f is *orientation preserving* if this sign is positive and *orientation reversing* if it is negative. We will prove below:

Theorem 3.5.1. *The degree of f is $+1$ if f is orientation preserving and -1 if f is orientation reversing.*

We will then use this result to prove the following change of variables formula for diffeomorphisms.

Theorem 3.5.2. *Let $\varphi : V \to \mathbb{R}$ be a compactly supported continuous function. Then*

$$\int_U \varphi \circ f(x) |\det(Df)(x)| = \int_V \varphi(y)\, dy\,. \tag{3.5.1}$$

Proof of Theorem 3.5.1. Given a point, $a_1 \in U$, let $a_2 = -f(a_1)$ and for $i = 1, 2$, let $g_i : \mathbb{R}^n \to \mathbb{R}^n$ be the translation, $g_i(x) = x + a_i$. By (3.4.1) and exercise 4 of § 4 the composite diffeomorphism

$$g_2 \circ f \circ g_1 \tag{3.5.2}$$

has the same degree as f, so it suffices to prove the theorem for this mapping. Notice however that this mapping maps the origin onto the origin. Hence, replacing f by this mapping, we can, without loss of generality, assume that 0 is in the domain of f and that $f(0) = 0$.

Next notice that if $A : \mathbb{R}^n \to \mathbb{R}^n$ is a bijective linear mapping the theorem is true for A (by exercise 9 of § 3.4), and hence if we can prove the theorem for $A^{-1} \circ f$, (3.4.1) will tell us that the theorem is true for f. In particular, letting $A = Df(0)$, we have

$$D(A^{-1} \circ f)(0) = A^{-1} Df(0) = I$$

where I is the identity mapping. Therefore, replacing f by $A^{-1}f$, we can assume that the mapping, f, for which we are attempting to prove Theorem 3.5.1 has the properties: $f(0) = 0$ and $Df(0) = I$. Let $g(x) = f(x) - x$. Then these properties imply that $g(0) = 0$ and $Dg(0) = 0$.

□

Lemma 3.5.3. *There exists a $\delta > 0$ such that $|g(x)| \leq \frac{1}{2}|x|$ for $|x| \leq \delta$.*

Proof. Let $g(x) = (g_1(x), \ldots, g_n(x))$. Then

$$\frac{\partial g_i}{\partial x_j}(0) = 0\,;$$

so there exists a $\delta > 0$ such that

$$\left|\frac{\partial g_i}{\partial x_j}(x)\right| \leq \frac{1}{2}$$

for $|x| \leq \delta$. However, by the mean value theorem,

$$g_i(x) = \sum \frac{\partial g_i}{\partial x_j}(c)x_j$$

for $c = t_0 x$, $0 < t_0 < 1$. Thus, for $|x| < \delta$,

$$|g_i(x)| \leq \frac{1}{2} \sup |x_i| = \frac{1}{2} |x| ,$$

so

$$|g(x)| = \sup |g_i(x)| \leq \frac{1}{2} |x| .$$

$\square$

Let ρ be a compactly supported $\mathcal{C}^\infty$ function with $0 \leq \rho \leq 1$ and with $\rho(x) = 0$ for $|x| \geq \delta$ and $\rho(x) = 1$ for $|x| \leq \frac{\delta}{2}$ and let $\widetilde{f} : \mathbb{R}^n \to \mathbb{R}^n$ be the mapping

$$\widetilde{f}(x) = x + \rho(x)g(x) . \tag{3.5.3}$$

It's clear that

$$\widetilde{f}(x) = x \text{ for } |x| \geq \delta \tag{3.5.4}$$

and, since $f(x) = x + g(x)$,

$$\widetilde{f}(x) = f(x) \text{ for } |x| \leq \frac{\delta}{2} . \tag{3.5.5}$$

In addition, for all $x \in \mathbb{R}^n$:

$$|\widetilde{f}(x)| \geq \frac{1}{2} |x| . \tag{3.5.6}$$

Indeed, by (3.5.4), $|\widetilde{f}(x)| \geq |x|$ for $|x| \geq \delta$, and for $|x| \leq \delta$

$$\begin{aligned} |\widetilde{f}(x)| &\geq |x| - \rho(x)|g(x)| \\ &\geq |x| - |g(x)| \geq |x| - \frac{1}{2}|x| = \frac{1}{2}|x| \end{aligned}$$

by Lemma 3.5.3.

Now let $\mathcal{Q}_r$ be the cube, $\{x \in \mathbb{R}^n\,,\ |x| \leq r\}$, and let $\mathcal{Q}_r^c = \mathbb{R}^n - \mathcal{Q}_r$. From (3.5.6) we easily deduce that

$$\widetilde{f}^{-1}(\mathcal{Q}_r) \subseteq \mathcal{Q}_{2r} \tag{3.5.7}$$

for all r, and hence that $\widetilde{f}$ is *proper*. Also notice that for $x \in \mathcal{Q}_\delta$,

$$|\widetilde{f}(x)| \leq |x| + |g(x)| \leq \frac{3}{2}\,|x|$$

by Lemma 3.5.3 and hence

$$\widetilde{f}^{-1}(\mathcal{Q}^c_{\frac{3}{2}\delta}) \subseteq \mathcal{Q}^c_\delta\,. \tag{3.5.8}$$

We will now prove Theorem 3.5.1. Since f is a diffeomorphism mapping 0 to 0, it maps a neighborhood, U_0, of 0 in U diffeomorphically onto a neighborhood, V_0, of 0 in V, and by shrinking U_0 if necessary we can assume that U_0 is contained in $\mathcal{Q}_{\delta/2}$ and V_0 contained in $\mathcal{Q}_{\delta/4}$. Let ω be an n-form with support in V_0 whose integral over $\mathbb{R}^n$ is equal to one. Then $f^*\omega$ is supported in U_0 and hence in $\mathcal{Q}_{\delta/2}$. Also by (3.5.7) $\widetilde{f}^*\omega$ is supported in $\mathcal{Q}_{\delta/2}$. Thus both of these forms are zero outside $\mathcal{Q}_{\delta/2}$. However, on $\mathcal{Q}_{\delta/2}$, $\widetilde{f} = f$ by (3.5.5), so these forms are equal everywhere, and hence

$$\deg(f) = \int f^*\omega = \int \widetilde{f}^*\omega = \deg(\widetilde{f})\,.$$

Next let ω be a compactly supported n-form with support in $\mathcal{Q}^c_{3\delta/2}$ and with integral equal to one. Then $\widetilde{f}^*\omega$ is supported in $\mathcal{Q}^c_\delta$ by (3.5.8), and hence since $f(x) = x$ on $\mathcal{Q}^c_\delta$ $\widetilde{f}^*\omega = \omega$. Thus

$$\deg(\widetilde{f}) = \int f^*\omega = \int \omega = 1\,.$$

Putting these two identities together we conclude that $\deg(f) = 1$. Q.E.D.

If the function, φ, in Theorem 3.5.2 is a $\mathcal{C}^\infty$ function, the identity (3.5.1) is an immediate consequence of the result above and the identity (3.4.2). If φ is not $\mathcal{C}^\infty$, but is just continuous, we will deduce Theorem 3.5.2 from the following result.

Theorem 3.5.4. *Let V be an open subset of $\mathbb{R}^n$. If $\varphi : \mathbb{R}^n \to \mathbb{R}$ is a continuous function of compact support with* $\operatorname{supp}\varphi \subseteq V$*; then for every $\epsilon > 0$ there exists a $\mathcal{C}^\infty$ function of compact support, $\psi : \mathbb{R}^n \to \mathbb{R}$ with* $\operatorname{supp}\psi \subseteq V$ *and*

$$\sup |\psi(x) - \varphi(x)| < \epsilon\,.$$

Proof. Let A be the support of φ and let d be the distance in the sup norm from A to the complement of V. Since φ is continuous and compactly supported it is uniformly continuous; so for every $\epsilon > 0$ there exists a $\delta > 0$ with $\delta < \frac{d}{2}$ such that $|\varphi(x) - \varphi(y)| < \epsilon$ when $|x - y| \le \delta$. Now let Q be the cube: $|x| < \delta$ and let $\rho : \mathbb{R}^n \to \mathbb{R}$ be a non-negative $\mathcal{C}^\infty$ function with $\operatorname{supp}\rho \subseteq Q$ and

$$\int \rho(y)\,dy = 1\,. \tag{3.5.9}$$

Set

$$\psi(x) = \int \rho(y - x)\varphi(y)\,dy\,.$$

By Theorem 3.2.5 ψ is a $\mathcal{C}^\infty$ function. Moreover, if A_δ is the set of points in $\mathbb{R}^d$ whose distance in the sup norm from A is $\le \delta$ then for $x \notin A_\delta$ and $y \in A$, $|x - y| > \delta$ and hence $\rho(y - x) = 0$. Thus for $x \notin A_\delta$

$$\int \rho(y - x)\varphi(y)\,dy = \int_A \rho(y - x)\varphi(y)\,dy = 0\,,$$

so ψ is supported on the compact set A_δ. Moreover, since $\delta < \frac{d}{2}$, $\operatorname{supp}\psi$ is contained in V. Finally note that by (3.5.9) and exercise 4 of §3.4:

$$\int \rho(y - x)\,dy = \int \rho(y)\,dy = 1 \tag{3.5.10}$$

and hence

$$\varphi(x) = \int \varphi(x)\rho(y - x)\,dy$$

so

$$\varphi(x) - \psi(x) = \int (\varphi(x) - \varphi(y))\rho(y - x)\,dy$$

and

$$|\varphi(x) - \psi(x)| \quad \leq \quad \int |\varphi(x) - \varphi(y)|\, \rho(y-x)\, dy\,.$$

But $\rho(y-x) = 0$ for $|x-y| \geq \delta$; and $|\varphi(x)-\varphi(y)| < \epsilon$ for $|x-y| \leq \delta$, so the integrand on the right is less than

$$\epsilon \int \rho(y-x)\, dy\,,$$

and hence by (3.5.10)

$$|\varphi(x) - \psi(x)| \leq \epsilon\,.$$

□

To prove the identity (3.5.1), let $\gamma : \mathbb{R}^n \to \mathbb{R}$ be a $\mathcal{C}^\infty$ cut-off function which is one on a neighborhood, V_1, of the support of φ, is non-negative, and is compactly supported with $\operatorname{supp}\gamma \subseteq V$, and let

$$c = \int \gamma(y)\, dy\,.$$

By Theorem 3.5.4 there exists, for every $\epsilon > 0$, a $\mathcal{C}^\infty$ function ψ, with support on V_1 satisfying

$$|\varphi - \psi| \leq \tfrac{\epsilon}{2c}\,. \tag{3.5.11}$$

Thus

$$\begin{aligned} \left|\int_V (\varphi - \psi)(y)\, dy\right| \quad &\leq \quad \int_V |\varphi - \psi|(y)\, dy \\ &\leq \quad \int_V \gamma|\varphi - \psi|(xy)\, dy \\ &\leq \quad \tfrac{\epsilon}{2c}\int \gamma(y)\, dy \leq \tfrac{\epsilon}{2} \end{aligned}$$

so

$$\left|\int_V \varphi(y)\, dy - \int_V \psi(y)\, dy\right| \leq \tfrac{\epsilon}{2}\,. \tag{3.5.12}$$

Similarly, the expression

$$\left|\int_U (\varphi - \psi)\circ f(x) |\det Df(x)|\, dx\right|$$

is less than or equal to the integral

$$\int_U \gamma \circ f(x) |(\varphi - \psi) \circ f(x)| \, |\det Df(x)| \, dx$$

and by (3.5.11), $|(\varphi - \psi) \circ f(x)| \leq \frac{\epsilon}{2c}$, so this integral is less than or equal to

$$\tfrac{\epsilon}{2c} \int \gamma \circ f(x) |\det Df(x)| \, dx$$

and hence by (3.5.1) is less than or equal to $\frac{\epsilon}{2}$. Thus

(3.5.13)

$$\left| \int_U \varphi \circ f(x) \, |\det Df(x)| dx - \int_U \psi \circ f(x) |\det Df(x)| \ dx \right| \leq \tfrac{\epsilon}{2} \, .$$

Combining (3.5.12), (3.5.13) and the identity

$$\int_V \psi(y) \, dy = \int \psi \circ f(x) |\det Df(x)| \, dx$$

we get, for all $\epsilon > 0$,

$$\left| \int_V \varphi(y) \, dy - \int_U \varphi \circ f(x) |\det Df(x)| \, dx \right| \leq \epsilon$$

and hence

$$\int \varphi(y) \, dy = \int \varphi \circ f(x) |\det Df(x)| \, dx \, .$$

Exercises for §3.5

1. Let $h : V \to \mathbb{R}$ be a non-negative continuous function. Show that if the improper integral

$$\int_V h(y) \, dy$$

is well-defined, then the improper integral

$$\int_U h \circ f(x) |\det Df(x)| \, dx$$

is well-defined and these two integrals are equal.

Hint: If φ_i, $i = 1, 2, 3, \ldots$ is a partition of unity on V then $\psi_i = \varphi_i \circ f$ is a partition of unity on U and

$$\int \varphi_i h \, dy = \int \psi_i (h \circ f(x)) |\det Df(x)| \, dx \, .$$

Now sum both sides of this identity over i.

2. Show that the result above is true without the assumption that h is non-negative.

Hint: $h = h_+ - h_-$, where $h_+ = \max(h, 0)$ and $h_- = \max(-h, 0)$.

3. Show that, in the formula (3.4.2), one can allow the function, φ, to be a *continuous* compactly supported function rather than a $\mathcal{C}^\infty$ compactly supported function.

4. Let $\mathbb{H}^n$ be the half-space (**??**) and U and V open subsets of $\mathbb{R}^n$. Suppose $f : U \to V$ is an orientation preserving diffeomorphism mapping $U \cap \mathbb{H}^n$ onto $V \cap \mathbb{H}^n$. Show that for $\omega \in \Omega_c^n(V)$

$$\int_{U \cap \mathbb{H}^n} f^* \omega = \int_{V \cap \mathbb{H}^n} \omega \, . \tag{3.5.14}$$

Hint: Interpret the left and right hand sides of this formula as improper integrals over $U \cap \operatorname{Int} \mathbb{H}^n$ and $V \cap \operatorname{Int} \mathbb{H}^n$.

5. The boundary of $\mathbb{H}^n$ is the set

$$b\mathbb{H}^n = \{(0, x_2, \ldots, x_n) \, , \quad (x_2, \ldots, x_n) \in \mathbb{R}^n\}$$

so the map

$$\iota : \mathbb{R}^{n-1} \to \mathbb{H}^n \, , \quad (x_2, \ldots, x_n) \to (0, x_2, \ldots, x_n)$$

in exercise 9 in §3.2 maps $\mathbb{R}^{n-1}$ bijectively onto $b\mathbb{H}^n$.

(a) Show that the map $f : U \to V$ in exercise 4 maps $U \cap b\mathbb{H}^n$ onto $V \cap b\mathbb{H}^n$.

(b) Let $U' = \iota^{-1}(U)$ and $V' = \iota^{-1}(V)$. Conclude from part (a) that the restriction of f to $U \cap b\mathbb{H}^n$ gives one a diffeomorphism

$$g : U' \to V'$$

satisfying:

$$\iota \cdot g = f \cdot \iota \, . \tag{3.5.15}$$

(c) Let μ be in $\Omega_c^{n-1}(V)$. Conclude from (3.2.7) and (3.5.14):

$$\int_{U'} g^*\iota^*\mu = \int_{V'} \iota^*\mu \tag{3.5.16}$$

and in particular show that the diffeomorphism, $g : U' \to V'$, is orientation preserving.

3.6 Techniques for computing the degree of a mapping

Let U and V be open subsets of $\mathbb{R}^n$ and $f : U \to V$ a proper $\mathcal{C}^\infty$ mapping. In this section we will show how to compute the degree of f and, in particular, show that it is always an integer. From this fact we will be able to conclude that the degree of f is a topological invariant of f: if we deform f smoothly, its degree doesn't change.

Definition 3.6.1. *A point, $x \in U$, is a* critical point *of f if the derivative*

$$Df(x) : \mathbb{R}^n \to \mathbb{R}^n$$

fails to be bijective, i.e., if $\det(Df(x)) = 0$.

We will denote the set of critical points of f by C_f. It's clear from the definition that this set is a closed subset of U and hence, by exercise 3 in §3.4, $f(C_f)$ is a closed subset of V. We will call this image the set of *critical values* of f and the complement of this image the set of *regular values* of f. Notice that $V - f(U)$ is contained in $f - f(C_f)$, so if a point, $g \in V$ is not in the image of f, it's a regular value of f "by default", i.e., it contains no points of U in the pre-image and hence, a fortiori, contains no critical points in its pre-image. Notice also that C_f can be quite large. For instance, if c is a point in V and $f : U \to V$ is the constant map which maps all of U onto c, then $C_f = U$. However, in this example, $f(C_f) = \{c\}$, so the set of regular values of f is $V - \{c\}$, and hence (in this example) is an open dense subset of V. We will show that this is true in general.

Theorem 3.6.2. *(Sard's theorem.)*

If U and V are open subsets of $\mathbb{R}^n$ and $f : U \to V$ a proper $\mathcal{C}^\infty$ map, the set of regular values of f is an open dense subset of V.

We will defer the proof of this to Section 3.7 and, in this section, explore some of its implications. Picking a regular value, q, of f we will prove:

Theorem 3.6.3. *The set, $f^{-1}(q)$ is a finite set. Moreover, if $f^{-1}(q) = \{p_1, \dots, p_n\}$ there exist connected open neighborhoods, U_i, of p_i in Y and an open neighborhood, W, of q in V such that:*

i. *for $i \neq j$ U_i and U_j are disjoint;*

ii. $f^{-1}(W) = \bigcup U_i$,

iii. *f maps U_i diffeomorphically onto W.*

Proof. If $p \in f^{-1}(q)$ then, since q is a regular value, $p \notin C_f$; so

$$Df(p) : \mathbb{R}^n \to \mathbb{R}^n$$

is bijective. Hence by the inverse function theorem, f maps a neighborhood, U_p of p diffeomorphically onto a neighborhood of q. The open sets

$$\{U_p\,, \quad p \in f^{-1}(q)\}$$

are a covering of $f^{-1}(q)$; and, since f is proper, $f^{-1}(q)$ is compact; so we can extract a finite subcovering

$$\{U_{p_i}\,, \quad i = 1, \dots, N\}$$

and since p_i is the only point in U_{p_i} which maps onto q, $f^{-1}(q) = \{p_1, \dots, p_N\}$.

Without loss of generality we can assume that the U_{p_i}'s are disjoint from each other; for, if not, we can replace them by smaller neighborhoods of the p_i's which have this property. By Theorem 3.4.2 there exists a connected open neighborhood, W, of q in V for which

$$f^{-1}(W) \subset \bigcup U_{p_i}\,.$$

To conclude the proof let $U_i = f^{-1}(W) \cap U_{p_i}$.

□

The main result of this section is a recipe for computing the degree of f by counting the number of p_i's above, keeping track of orientation.

Theorem 3.6.4. *For each $p_i \in f^{-1}(q)$ let $\sigma_{p_i} = +1$ if $f : U_i \to W$ is orientation preserving and -1 if $f : U_i \to W$ is orientation reversing. Then*

$$\deg(f) = \sum_{i=1}^{N} \sigma_{p_i}\,. \tag{3.6.1}$$

Proof. Let ω be a compactly supported n-form on W whose integral is one. Then

$$\deg(f) = \int_U f^*\omega = \sum_{i=1}^{N} \int_{U_i} f^*\omega \,.$$

Since $f : U_i \to W$ is a diffeomorphism

$$\int_{U_i} f^*\omega = \pm \int_W \omega = +1 \text{ or } -1$$

depending on whether $f : U_i \to W$ is orientation preserving or not. Thus $\deg(f)$ is equal to the sum (3.6.1).

□

As we pointed out above, a point, $q \in V$ can qualify as a regular value of f "by default", i.e., by not being in the image of f. In this case the recipe (3.6.1) for computing the degree gives "by default" the answer zero. Let's corroborate this directly.

Theorem 3.6.5. *If* $f : U \to V$ *isn't onto,* $\deg(f) = 0$.

Proof. By exercise 3 of §3.4, $V - f(U)$ is open; so if it is non-empty, there exists a compactly supported n-form, ω, with support in $V - f(U)$ and with integral equal to one. Since $\omega = 0$ on the image of f, $f^*\omega = 0$; so

$$0 = \int_U f^*\omega = \deg(f) \int_V \omega = \deg(f)\,.$$

□

Remark: In applications the contrapositive of this theorem is much more useful than the theorem itself.

Theorem 3.6.6. *If* $\deg(f) \neq 0$ f *maps* U *onto* V.

In other words if $\deg(f) \neq 0$ the equation

$$f(x) = y \tag{3.6.2}$$

has a solution, $x \in U$ for *every* $y \in V$.

We will now show that the degree of f is a topological invariant of f: if we deform f by a "homotopy" we don't change its degree. To make this assertion precise, let's recall what we mean by a *homotopy* between a pair of $\mathcal{C}^\infty$ maps. Let U be an open subset of $\mathbb{R}^m$, V an open subset of $\mathbb{R}^n$, A an open subinterval of $\mathbb{R}$ containing 0 and 1, and $f_i : U \to V$, $i = 0, 1$, $\mathcal{C}^\infty$ maps. Then a $\mathcal{C}^\infty$ map $F : U \times A \to V$ is a homotopy between f_0 and f_1 if $F(x, 0) = f_0(x)$ and $F(x, 1) = f_1(x)$. (See Definition **??**.) Suppose now that f_0 and f_1 are proper.

Definition 3.6.7. *F is a* proper *homotopy between f_0 and f_1 if the map*

$$F^\sharp : U \times A \to V \times A \tag{3.6.3}$$

mapping (x, t) to $(F(x, t), t)$ is proper.

Note that if F is a proper homotopy between f_0 and f_1, then for every t between 0 and 1, the map

$$f_t : U \to V\,, \quad f_t(x) = F_t(x)$$

is proper.

Now let U and V be open subsets of $\mathbb{R}^n$.

Theorem 3.6.8. *If f_0 and f_1 are properly homotopic, their degrees are the same.*

Proof. Let

$$\omega = \varphi(y)\, d\,y_1 \wedge \cdots \wedge d\,y_n$$

be a compactly supported n-form on X whose integral over V is 1. The the degree of f_t is equal to

$$\int_U \varphi(F_1(x, t), \ldots, F_n(x, t)) \det D_x F(x, t)\, dx\,. \tag{3.6.4}$$

The integrand in (3.6.4) is continuous and for $0 \le t \le 1$ is supported on a compact subset of $U \times [0, 1]$, hence (3.6.4) is continuous as a function of t. However, as we've just proved, $\deg(f_t)$ is *integer* valued so this function is a constant. ☐

(For an alternative proof of this result see exercise 9 below.) We'll conclude this account of degree theory by describing a couple applications.

Application 1. The Brouwer fixed point theorem

Let B^n be the closed unit ball in $\mathbb{R}^n$:

$$\{x \in \mathbb{R}^n\,,\ \|x\| \leq 1\}\,.$$

Theorem 3.6.9. *If $f : B^n \to B^n$ is a continuous mapping then f has a fixed point, i.e., maps some point, $x_0 \in B^n$ onto itself.*

The idea of the proof will be to assume that there isn't a fixed point and show that this leads to a contradiction. Suppose that for every point, $x \in B^n$ $f(x) \neq x$. Consider the ray through $f(x)$ in the direction of x:

$$f(x) + s(x - f(x))\,, \quad 0 \leq s < \infty\,.$$

This intersects the boundary, S^{n-1}, of B^n in a unique point, $\gamma(x)$, (see figure 1 below); and one of the exercises at the end of this section will be to show that the mapping $\gamma : B^n \to S^{n-1}$, $x \to \gamma(x)$, is a continuous mapping. Also it is clear from figure 1 that $\gamma(x) = x$ if $x \in S^{n-1}$, so we can extend γ to a continuous mapping of $\mathbb{R}^n$ into $\mathbb{R}^n$ by letting γ be the identity for $\|x\| \geq 1$. Note that this extended mapping has the property

$$\|\gamma(x)\| \geq 1 \tag{3.6.5}$$

for all $x \in \mathbb{R}^n$ and

$$\gamma(x) = x \tag{3.6.6}$$

for all $\|x\| \geq 1$. To get a contradiction we'll show that γ can be approximated by a $\mathcal{C}^\infty$ map which has similar properties. For this we will need the following corollary of Theorem 3.5.4.

Lemma 3.6.10. *Let U be an open subset of $\mathbb{R}^n$, C a compact subset of U and $\varphi : U \to \mathbb{R}$ a continuous function which is $\mathcal{C}^\infty$ on the complement of C. Then for every $\epsilon > 0$, there exists a $\mathcal{C}^\infty$ function, $\psi : U \to \mathbb{R}$, such that $\varphi - \psi$ has compact support and $|\varphi - \psi| < \epsilon$.*

Proof. Let ρ be a bump function which is in $\mathcal{C}_0^\infty(U)$ and is equal to 1 on a neighborhood of C. By Theorem 3.5.4 there exists a function, $\psi_0 \in \mathcal{C}_0^\infty(U)$ such that $|\rho\varphi - \psi_0| < \epsilon$. Let $\psi = (1-\rho)\varphi + \psi_0$, and note that

$$\begin{aligned} \varphi - \psi &= (1-\rho)\varphi + \rho\varphi - (1-\rho)\varphi - \psi_0 \\ &= \rho\varphi - \psi_0\,. \end{aligned}$$

By applying this lemma to each of the coordinates of the map, γ, one obtains a $\mathcal{C}^\infty$ map, $g : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$\|g - \gamma\| < \epsilon < 1 \tag{3.6.7}$$

and such that $g = \gamma$ on the complement of a compact set. However, by (3.6.6), this means that g is equal to the identity on the complement of a compact set and hence (see exercise 9) that g is proper and has degree one. On the other hand by (3.6.8) and (3.6.6) $\|g(x)\| > 1-\epsilon$ for all $x \in \mathbb{R}^n$, so $0 \notin \operatorname{Im} g$ and hence by Theorem 3.6.4, $\deg(g) = 0$. Contradiction.

□

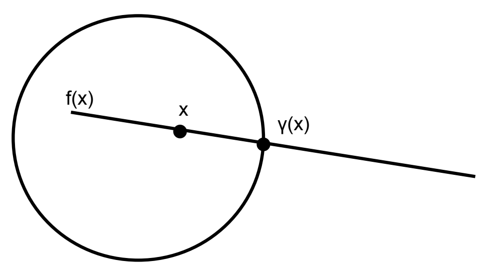


Figure 3.6.1.

Application 2. The fundamental theorem of algebra

Let $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$ be a polynomial of degree n with complex coefficients. If we identify the complex plane

$$\mathbb{C} = \{z = x + iy\,;\, x, y \in \mathbb{R}\}$$

with $\mathbb{R}^2$ via the map, $(x, y) \in \mathbb{R}^2 \to z = x + iy$, we can think of p as defining a mapping

$$p : \mathbb{R}^2 \to \mathbb{R}^2\,,\; z \to p(z)\,.$$

We will prove

Theorem 3.6.11. *The mapping, p, is proper and* $\deg(p) = n$.

Proof. For $t \in \mathbb{R}$

$$\begin{aligned} p_t(z) &= (1-t)z^n + tp(z) \\ &= z^n + t\sum_{i=0}^{n-1} a_i z^i \,. \end{aligned}$$

We will show that the mapping

$$g : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2 \,,\; z \to p_t(z)$$

is a proper homotopy. Let

$$C = \sup\{|a_i| \,,\; i = 0, \dots, n-1\} \,.$$

Then for $|z| \geq 1$

$$\begin{aligned} |a_0 + \cdots + a_{n-1}z^{n-1}| &\leq |a_0| + |a_1||z| + \cdots + |a_{n-1}|\,|z|^{n-1} \\ &\leq C|z|^{n-1} \,, \end{aligned}$$

and hence, for $|t| \leq a$ and $|z| \geq 2aC$,

$$\begin{aligned} |p_t(z)| &\geq |z|^n - aC|z|^{n-1} \\ &\geq aC|z|^{n-1} \,. \end{aligned}$$

If A is a compact subset of $\mathbb{C}$ then for some $R > 0$, A is contained in the disk, $|w| \leq R$ and hence the set

$$\{z \in \mathbb{C} \,,\; (p_t(z), t) \in A \times [-a, a]\}$$

is contained in the compact set

$$\{z \in \mathbb{C} \,,\; aC|z|^{n-1} \leq R\} \,,$$

and this shows that g is a proper homotopy. Thus each of the mappings,

$$p_t : \mathbb{C} \to \mathbb{C} \,,$$

is proper and $\deg p_t = \deg p_1 = \deg p = \deg p_0$. However, $p_0 : \mathbb{C} \to \mathbb{C}$ is just the mapping, $z \to z^n$ and an elementary computation (see exercises 5 and 6 below) shows that the degree of this mapping is n. □

In particular for $n > 0$ the degree of p is non-zero; so by Theorem 3.6.4 we conclude that $p : \mathbb{C} \to \mathbb{C}$ is surjective and hence has zero in its image.

Theorem 3.6.12. *(fundamental theorem of algebra)*
Every polynomial,

$$p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_0 ,$$

with complex coefficients has a complex root, $p(z_0) = 0$*, for some* $z_0 \in \mathbb{C}$*.*

Exercises for §3.6

1. Let W be a subset of $\mathbb{R}^n$ and let $a(x)$, $b(x)$ and $c(x)$ be real-valued functions on W of class C^r. Suppose that for every $x \in W$ the quadratic polynomial

$$(*) \qquad a(x)s^2 + b(x)s + c(x)$$

has two distinct real roots, $s_+(x)$ and $s_-(x)$, with $s_+(x) > s_-(x)$. Prove that s_+ and s_- are functions of class C^r.

Hint: What *are* the roots of the quadratic polynomial: as^2+bs+c?

2. Show that the function, $\gamma(x)$, defined in figure 1 is a continuous mapping of B^n onto S^{2n-1}. *Hint:* $\gamma(x)$ lies on the ray,

$$f(x) + s(x - f(x)) , \quad 0 \le s < \infty$$

and satisfies $\|\gamma(x)\| = 1$; so $\gamma(x)$ is equal to

$$f(x) + s_0(x - f(x))$$

where s_0 is a non-negative root of the quadratic polynomial

$$\|f(x) + s(x - f(x))\|^2 - 1 .$$

Argue from figure 1 that this polynomial has to have two distinct real roots.

3. Show that the Brouwer fixed point theorem isn't true if one replaces the closed unit ball by the open unit ball. *Hint:* Let U be the open unit ball (i.e., the interior of B^n). Show that the map

$$h : U \to \mathbb{R}^n , \quad h(x) = \frac{x}{1 - \|x\|^2}$$

is a diffeomorphism of U onto $\mathbb{R}^n$, and show that there are lots of mappings of $\mathbb{R}^n$ onto $\mathbb{R}^n$ which don't have fixed points.

4. Show that the fixed point in the Brouwer theorem doesn't have to be an interior point of B^n, i.e., show that it can lie on the boundary.

5. If we identify $\mathbb{C}$ with $\mathbb{R}^2$ via the mapping: $(x, y) \to z = x + iy$, we can think of a $\mathbb{C}$-linear mapping of $\mathbb{C}$ into itself, i.e., a mapping of the form

$$z \to cz , \quad c \in \mathbb{C}$$

as being an $\mathbb{R}$-linear mapping of $\mathbb{R}^2$ into itself. Show that the determinant of this mapping is $|c|^2$.

6. (a) Let $f : \mathbb{C} \to \mathbb{C}$ be the mapping, $f(z) = z^n$. Show that

$$Df(z) = nz^{n-1} .$$

Hint: Argue from first principles. Show that for $h \in \mathbb{C} = \mathbb{R}^2$

$$\frac{(z+h)^n - z^n - nz^{n-1}h}{|h|}$$

tends to zero as $|h| \to 0$.

(b) Conclude from the previous exercise that

$$\det Df(z) = n^2 |z|^{2n-2} .$$

(c) Show that at every point $z \in \mathbb{C} - 0$, f is orientation preserving.

(d) Show that every point, $w \in \mathbb{C} - 0$ is a regular value of f and that

$$f^{-1}(w) = \{z_1, \ldots, z_n\}$$

with $\sigma_{z_i} = +1$.

(e) Conclude that the degree of f is n.

7. Prove that the map, f, in exercise 6 has degree n by deducing this directly from the definition of degree. *Some hints:*

(a) Show that in polar coordinates, f is the map, $(r, \theta) \to (r^n, n\theta)$.

(b) Let ω be the two-form, $g(x^2+y^2)\, dx \wedge dy$, where $g(t)$ is a compactly supported $\mathcal{C}^\infty$ function of t. Show that in polar coordinates, $\omega = g(r^2) r\, dr \wedge d\theta$, and compute the degree of f by computing the integrals of ω and $f^*\omega$, in polar coordinates and comparing them.

8. Let U be an open subset of $\mathbb{R}^n$, V an open subset of $\mathbb{R}^m$, A an open subinterval of $\mathbb{R}$ containing 0 and 1, $f_i : U \to V$ $i = 0, 1$, a pair of $\mathcal{C}^\infty$ mappings and $F : U \times A \to V$ a homotopy between f_0 and f_1.

(a) In §2.3, exercise 4 you proved that if μ is in $\Omega^k(V)$ and $d\mu = 0$, then

$$f_0^*\mu - f_1^*\mu = d\nu \tag{3.6.8}$$

where ν is the $(k-1)$-form, $Q\alpha$, in formula (**??**). Show (by careful inspection of the definition of $Q\alpha$) that if F is a *proper* homotopy and $\mu \in \Omega_c^k(V)$ then $\nu \in \Omega_c^{k-1}(U)$.

(b) Suppose in particular that U and V are open subsets of $\mathbb{R}^n$ and μ is in $\Omega_c^n(V)$. Deduce from (3.6.8) that

$$\int f_0^*\mu = \int f_1^*\mu$$

and deduce directly from the definition of degree that degree is a proper homotopy invariant.

9. Let U be an open connected subset of $\mathbb{R}^n$ and $f : U \to U$ a proper $\mathcal{C}^\infty$ map. Prove that if f is equal to the identity on the complement of a compact set, C, then f is proper and its degree is equal to 1. *Hints:*

(a) Show that for every subset, A, of U, $f^{-1}(A) \subseteq A \cup C$, and conclude from this that f is proper.

(b) Let $C' = f(C)$. Use the recipe (1.6.1) to compute $\deg(f)$ with $q \in U - C'$.

10. Let $[a_{i,j}]$ be an $n \times n$ matrix and $A : \mathbb{R}^n \to \mathbb{R}^n$ the linear mapping associated with this matrix. Frobenius' theorem asserts: *If the $a_{i,j}$'s are non-negative then A has a non-negative eigenvalue.* In

other words there exists a $v \in \mathbb{R}^n$ and a $\lambda \in \mathbb{R}$, $\lambda \geq 0$, such that $Av = \lambda v$. Deduce this linear algebra result from the Brouwer fixed point theorem. *Hints:*

(a) We can assume that A is bijective, otherwise 0 is an eigenvalue. Let S^{n-1} be the $(n-1)$-sphere, $|x| = 1$, and $f : S^{n-1} \to S^{n-1}$ the map,

$$f(x) = \frac{Ax}{\|Ax\|} .$$

Show that f maps the set

$$Q = \{(x_1, \dots, x_n) \in S^{n-1} ; \quad x_i \geq 0\}$$

into itself.

(b) It's easy to prove that Q is homeomorphic to the unit ball B^{n-1}, i.e., that there exists a continuous map, $g : Q \to B^{n-1}$ which is invertible and has a continuous inverse. Without bothering to prove this fact deduce from it Frobenius' theorem.

3.7 Appendix: Sard's theorem

The version of Sard's theorem stated in §3.5 is a corollary of the following more general result.

Theorem 3.7.1. *Let U be an open subset of $\mathbb{R}^n$ and $f : U \to \mathbb{R}^n$ a $\mathcal{C}^\infty$ map. Then $\mathbb{R}^n - f(C_f)$ is dense in $\mathbb{R}^n$.*

Before undertaking to prove this we will make a few general comments about this result.

Remark 3.7.2. *If $\mathcal{O}_n$, $n = 1, 2,$ are open dense subsets of $\mathbb{R}^n$, the intersection*

$$\bigcap_n \mathcal{O}_n$$

*is dense in $\mathbb{R}^n$. (See [**?**], pg. 200 or exercise 4 below.)*

Remark 3.7.3. *If A_n, $n = 1, 2, \dots$ are a covering of U by compact sets, $\mathcal{O}_n = \mathbb{R}^n - f(C_f \cap A_n)$ is open, so if we can prove that it's dense then by Remark 3.7.2 we will have proved Sard's theorem. Hence since we can always cover U by a countable collection of closed cubes, it suffices to prove: for every closed cube, $A \subseteq U$, $\mathbb{R}^n - f(C_f \cap A)$ is dense in $\mathbb{R}^n$.*

Remark 3.7.4. *Let $g : W \to U$ be a diffeomorphism and let $h = f \circ g$. Then*

$$f(C_f) = h(C_h) \tag{3.7.1}$$

so Sard's theorem for g implies Sard's theorem for f.

We will first prove Sard's theorem for the set of *super-critical* points of f, the set:

$$C_f^\sharp = \{p \in U \,, \quad Df(p) = 0\} \,. \tag{3.7.2}$$

Proposition 3.7.5. *Let $A \subseteq U$ be a closed cube. Then the open set $\mathbb{R}^n - f(A \cap C_f^\sharp)$ is a dense subset of $\mathbb{R}^n$.*

We'll deduce this from the lemma below.

Lemma 3.7.6. *Given $\epsilon > 0$ one can cover $f(A \cap C_f^\sharp)$ by a finite number of cubes of total volume less than ϵ.*

Proof. Let the length of each of the sides of A be ℓ. Given $\delta > 0$ one can subdivide A into N^n cubes, each of volume, $\left(\frac{\ell}{N}\right)^n$, such that if x and y are points of any one of these subcubes

$$\left| \frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(y) \right| < \delta \,. \tag{3.7.3}$$

Let $A_1, \ldots, A_m$ be the cubes in this collection which intersect $C_f^\sharp$. Then for $z_0 \in A_i \cap C_f^\sharp$, $\frac{\partial f_i}{\partial x_j}(z_0) = 0$, so for $z \in A_i$

$$\left| \frac{\partial f_i}{\partial x_j}(z) \right| < \delta \tag{3.7.4}$$

by (3.7.3). If x and y are points of A_i then by the mean value theorem there exists a point z on the line segment joining x to y such that

$$f_i(x) - f_i(y) = \sum \frac{\partial f_i}{\partial x_j}(z)(x_j - y_j)$$

and hence by (3.7.4)

$$|f_i(x) - f_i(y)| \le \delta \sum |x_i - y_i| \le n\delta \frac{\ell}{N} \,. \tag{3.7.5}$$

Thus $f(C_f \cap A_i)$ is contained in a cube, B_i, of volume $\left(n\frac{\delta\ell}{N}\right)^n$, and $f(C_f \cap A)$ is contained in a union of cubes, B_i, of total volume less that

$$N^n n^n \frac{\delta^n \ell^n}{N^n} = n^n \delta^n \ell^n$$

so if w choose $\delta^n \ell^n < \epsilon$, we're done.

$\square$

Proof. To prove Proposition 3.7.5 we have to show that for every point $p \in \mathbb{R}^n$ and neighborhood, W, of p, $W - f(C_f^\sharp \cap A)$ is non-empty. Suppose

$$W \subseteq f(C_f^\sharp \cap A)\,. \tag{3.7.6}$$

Without loss of generality we can assume W is a cube of volume ϵ, but the lemma tells us that $f(C_f^\sharp \cap A)$ can be covered by a finite number of cubes whose total volume is *less* than ϵ, and hence by (3.7.6) W can be covered by a finite number of cubes of total volume less than ϵ, so its volume is less than ϵ. This contradiction proves that the inclusion (3.7.6) can't hold.

$\square$

To prove Theorem 3.7.1 let $U_{i,j}$ be the subset of U where $\frac{\partial f_i}{\partial x_j} \neq 0$. Then

$$U = \bigcup U_{i,j} \cup C_f^\sharp\,,$$

so to prove the theorem it suffices to show that $\mathbb{R}^n - f(U_{i,j} \cap C_f)$ is dense in $\mathbb{R}^n$, i.e., it suffices to prove the theorem with U replaced by $U_{i,j}$. Let $\sigma_i : \mathbb{R}^n \times \mathbb{R}^n$ be the involution which interchanges x_1 and x_i and leaves the remaining x_k's fixed. Letting $f_{\text{new}} = \sigma_i f_{\text{old}} \sigma_j$ and $U_{\text{new}} = \sigma_j U_{\text{old}}$, we have, for $f = f_{\text{new}}$ and $U = U_{\text{new}}$

$$\frac{\partial f_1}{\partial x_1}(p) \neq 0 \quad \text{for all } p \in U\} \tag{3.7.7}$$

so we're reduced to proving Theorem 3.7.1 for maps $f : U \to \mathbb{R}^n$ having the property (3.7.6). Let $g : U \to \mathbb{R}^n$ be defined by

$$g(x_1, \ldots, x_n) = (f_1(x), x_2, \ldots, x_n)\,. \tag{3.7.8}$$

Then

$$g^*x_1 = f^*x_1 \quad = \quad f_1(x_1,\dots,x_n) \tag{3.7.9}$$

and

$$\det(Dg) \quad = \quad \frac{\partial f_1}{\partial x_1} \neq 0\,. \tag{3.7.10}$$

Thus, by the inverse function theorem, g is locally a diffeomorphism at every point, $p \in U$. This means that if A is a compact subset of U we can cover A by a finite number of open subsets, $U_i \subset U$ such that g maps U_i diffeomorphically onto an open subset W_i in $\mathbb{R}^n$. To conclude the proof of the theorem we'll show that $\mathbb{R}^n - f(C_f \cap U_i \cap A)$ is a dense subset of $\mathbb{R}^n$. Let $h : W_i \to \mathbb{R}^n$ be the map $h = f \circ g^{-1}$. To prove this assertion it suffices by Remark 3.7.4 to prove that the set

$$\mathbb{R}^n - h(C_h)$$

is dense in $\mathbb{R}^n$. This we will do by induction on n. First note that for $n = 1$, $C_f = C_f^\sharp$, so we've already proved Theorem 3.7.1 in dimension one. Now note that by (3.7.8), $h^*x_1 = x_1$, i.e., h is a mapping of the form

$$h(x_1,\dots,x_n) = (x_1, h_2(x),\dots,h_n(x))\,. \tag{3.7.11}$$

Thus if we let W_c be the set

$$\{(x_2,\dots,x_n) \in \mathbb{R}^{n-1}\,;\, (c, x_2,\dots,x_n) \in W_i\} \tag{3.7.12}$$

and let $h_c : W_c \to \mathbb{R}^{n-1}$ be the map

$$h_c(x_2,\dots,x_n) = (h_2(c,x_2,\dots,x_n),\dots,h_n(c,x_2,\dots,x_n))\,. \tag{3.7.13}$$

Then

$$\det(Dh_c)(x_2,\dots,x_n) = \det(Dh)(c,x_2,\dots,x_n) \tag{3.7.14}$$

and hence

$$(c,x) \in W_i \cap C_h \Leftrightarrow x \in C_{h_c}\,. \tag{3.7.15}$$

Now let $p_0 = (c, x_0)$ be a point in $\mathbb{R}^n$. We have to show that every neighborhood, V, of p_0 contains a point $p \in \mathbb{R}^n - h(C_h)$. Let $V_c \subseteq \mathbb{R}^{n-1}$ be the set of points, x, for which $(c, x) \in V$. By induction V_c contains a point, $x \in \mathbb{R}^{n-1} - h_c(C_{h_c})$ and hence $p = (c, x)$ is in V by definition and in $\mathbb{R}^n - h(C_n)$ by (3.7.15).

Q.E.D.

Exercises for §3.7

1. (a) Let $f : \mathbb{R} \to \mathbb{R}$ be the map $f(x) = (x^2 - 1)^2$. What is the set of critical points of f? What is its image?

(b) Same questions for the map $f(x) = \sin x + x$.

(c) Same questions for the map

$$f(x) = \begin{cases} 0, & x \leq 0 \\ e^{-\frac{1}{x}}, & x > 0 \end{cases} .$$

2. Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be an affine map, i.e., a map of the form

$$f(x) = A(x) + x_0$$

where $A : \mathbb{R}^n \to \mathbb{R}^n$ is a linear map. Prove Sard's theorem for f.

3. Let $\rho : \mathbb{R} \to \mathbb{R}$ be a $\mathcal{C}^\infty$ function which is supported in the interval $\left(-\frac{1}{2}, \frac{1}{2}\right)$ and has a maximum at the origin. Let $r_1, r_2, \ldots,$ be an enumeration of the rational numbers, and let $f : \mathbb{R} \to \mathbb{R}$ be the map

$$f(x) = \sum_{i=1}^{\infty} r_i \rho(x - i) .$$

Show that f is a $\mathcal{C}^\infty$ map and show that the image of C_f is dense in $\mathbb{R}$. (The moral of this example: Sard's theorem says that the complement of C_f is dense in $\mathbb{R}$, but C_f can be dense as well.)

4. Prove the assertion made in Remark 3.7.2. *Hint:* You need to show that for every point $p \in \mathbb{R}^n$ and every neighborhood, V, of p, $\bigcap \mathcal{O}_n \cap V$ is non-empty. Construct, by induction, a family of closed balls, B_k, such that

(a) $B_k \subseteq V$

(b) $B_{k+1} \subseteq B_k$

(c) $B_k \subseteq \bigcap_{n \leq k} \mathcal{O}_n$

(d) radius $B_k < \frac{1}{k}$

and show that the intersection of the B_k's is non-empty.

5. Verify (3.7.1).