CHAPTER 5

Pulling back forms

Determinants

The determinant of a matrix is the oriented volume of the block spanned by its column vectors. It is therefore not surprising that differential forms are closely related to determinants. This section is a review of some fundamental facts concerning determinants. Let

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

be an $n \times n$-matrix with column vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$. Its determinant is variously denoted by

$$\det A = \det(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) = \det(a_{ij})_{1 \le i,j \le n} = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}.$$

Expansion on the first column

You have probably seen the following definition of the determinant:

$$\det A = \sum_{i=1}^{n} (-1)^{i+1} a_{i1} \det A_{i1}.$$

Here A_{ij} denotes the $(n-1) \times (n-1)$-matrix obtained from A by striking out the ith row and the jth column. This is a recursive definition, which reduces the calculation of any determinant to that of determinants of smaller size. (The recursion starts at $n = 1$; the determinant of a 1×1-matrix (a) is simply defined to be the number a.) It has two serious flaws: first, it is extremely inefficient computationally (except for matrices containing lots of zeroes), and second, it obscures the relationship with volumes of parallelepipeds.

Axioms

A far better definition is available. The determinant can be completely characterized by three simple laws, which make good sense in view of its geometrical significance and which comprise an efficient algorithm for calculating any determinant. To state this axiomatic definition let us temporarily write $f(A)$ instead of $\det A$.

5.1. DEFINITION. A *determinant* or *oriented volume function* is a function f which assigns to every n-tuple of vectors $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ a number $f(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ subject to the following axioms:

(i) f is multilinear (i.e. linear in each column):

$$f(\mathbf{a}_1, \mathbf{a}_2, \dots, c\mathbf{a}_i + c'\mathbf{a}'_i, \dots, \mathbf{a}_n) = cf(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) + c'f(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}'_i, \dots, \mathbf{a}_n)$$

for all scalars c, c' and all vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_i, \mathbf{a}'_i, \dots, \mathbf{a}_n$;

(ii) f is antisymmetric:

$$f(\mathbf{a}_1, \dots, \mathbf{a}_i, \dots, \mathbf{a}_j, \dots, \mathbf{a}_n) = -f(\mathbf{a}_1, \dots, \mathbf{a}_j, \dots, \mathbf{a}_i, \dots, \mathbf{a}_n)$$

for any $i \neq j$;

(iii) f is normalized: $f(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n) = 1$, where $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ are the standard basis vectors of $\mathbf{R}^n$.

We also write $f(A)$ instead of $f(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ if A is the matrix whose columns are $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$. Axiom (iii) lays down the value of $f(I)$. Axioms (i) and (ii) govern the behaviour of oriented volumes under the elementary column operations on matrices. Recall that these operations come in three types: adding a multiple of any column of A to any other column (type I); multiplying a column by a nonzero constant (type II); and interchanging any two columns (type III). Type I does not affect the determinant, type II multiplies it by the corresponding constant, and type III causes a sign change. This can be stated more precisely as follows.

5.2. LEMMA. *Suppose that f is a determinant function. If E is an elementary column operation, then $f(E(A)) = kf(A)$, where*

$$k = \begin{cases} 1 & \text{if } E \text{ is of type I} \\ \text{the value of the constant} & \text{if } E \text{ is of type II} \\ -1 & \text{if } E \text{ is of type III.} \end{cases}$$

The following result states that there is at most one determinant function. In other words, the axioms (i)–(iii) suffice to calculate *any* $n \times n$-determinant.

5.3. THEOREM (uniqueness of determinants). *Let f and f' be two determinant functions. Then*

$$f(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) = f'(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$$

for all $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$.

PROOF. Let A be the $n \times n$-matrix whose columns are $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$. Assume first that A is *not* invertible. Then the columns of A are linearly dependent. For simplicity let us say that the first column is a linear combination of the others: $\mathbf{a}_1 = c_2\mathbf{a}_2 + \dots + c_n\mathbf{a}_n$. Applying axioms (i) and (ii) we get

$$f(A) = \sum_{i=2}^{n} c_i f(\mathbf{a}_i, \mathbf{a}_2, \dots, \mathbf{a}_i, \dots, \mathbf{a}_n) = 0,$$

and for the same reason $f'(A) = 0$, so $f(A) = f'(A)$. Now assume that A is invertible. Then A is column equivalent to the identity matrix, i.e. it can be transformed to I by successive elementary column operations. Let $E_1, E_2, \dots, E_k$ be these elementary operations, so that $E_kE_{k-1}\cdots E_2E_1(A) = I$. Each operation E_i has the effect of multiplying the determinant by a factor c_i, so rule (iii) yields

$$1 = f(I) = f(E_kE_{k-1}\cdots E_2E_1(A)) = c_kc_{k-1}\cdots c_2c_1f(A).$$

Applying the same reasoning to $f'(A)$ we get $1 = c_kc_{k-1}\cdots c_2c_1f'(A)$. Hence $f(A) = 1/(c_1c_2\cdots c_k) = f'(A)$. QED

5.4. REMARK (change of normalization). In the setup of this theorem, suppose that f' is multilinear and antisymmetric but is normalized differently:

$$f'(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n) = c.$$

Then the same proof shows that $f'(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) = cf(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ for all $\mathbf{a}_1$, $\mathbf{a}_2, \dots, \mathbf{a}_n$.

This theorem begs the question whether there *exists* a determinant function. We could answer this question by showing directly that the expansion formula given above satisfies the axioms of Definition 5.1. We choose instead to display an explicit formula for the determinant of any $n \times n$-matrix. Unlike Definition 5.1, this formula is not very practical for the purpose of calculating large determinants, but it has other uses, notably in the theory of differential forms.

5.5. THEOREM (existence of determinants). *There exists a unique determinant function. It is denoted by* det *and is given by the formula*

$$\boxed{\det A = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}.}$$

This requires a little explanation. S_n stands for the collection of all *permutations* of the set $\{1, 2, \dots, n\}$. A permutation is a way of ordering the numbers $1, 2, \dots, n$. Permutations are usually written as row vectors containing each of the integers $\{1, 2, \dots, n\}$ exactly once. Thus for $n = 2$ there are only two permutations: $(1\,2)$ and $(2\,1)$. For $n = 3$ all possible permutations are

$$(1\,2\,3) \qquad (1\,3\,2) \qquad (2\,1\,3) \qquad (2\,3\,1) \qquad (3\,1\,2) \qquad (3\,2\,1).$$

For general n there are

$$n(n-1)(n-2)\cdots 3\cdot 2\cdot 1 = n!$$

permutations. An alternative way of thinking of a permutation is as a bijective (i.e. one-to-one and onto) map from the set $\{1, 2, \dots, n\}$ to itself. For example, for $n = 5$ a possible permutation is

$$(5\,3\,1\,2\,4),$$

and we think of this as a shorthand for the map σ given by $\sigma(1) = 5$, $\sigma(2) = 3$, $\sigma(3) = 1$, $\sigma(4) = 2$ and $\sigma(5) = 4$. The permutation $1, 2, 3, \dots, n-1, n$ then corresponds to the identity map on the set $\{1, 2, \dots, n\}$.

If σ is the identity permutation, then clearly $\sigma(i) < \sigma(j)$ whenever $i < j$. However, if σ is not the identity permutation, it cannot preserve the order in this way. An *inversion* in σ is any pair of numbers i and j such that $1 \le i < j \le n$ and $\sigma(i) > \sigma(j)$. The *length* of σ, denoted by $l(\sigma)$, is the number of inversions in σ. A permutation is called *even* or *odd* according to whether its length is even, resp.

odd. For instance, the permutation $(5\,3\,1\,2\,4)$ is even, because it has length 6. The *sign* of σ is

$$\operatorname{sign}(\sigma) = (-1)^{l(\sigma)} = \begin{cases} 1 & \text{if } \sigma \text{ is even;} \\ -1 & \text{if } \sigma \text{ is odd.} \end{cases}$$

Thus $\operatorname{sign}(5\,3\,1\,2\,4) = 1$. The permutations of $\{1, 2\}$ with their respective signs (written in parentheses) are

$$(1\,2)(+) \qquad (2\,1)(-),$$

whereas for $n = 3$ we have the table below.

σ	$l(\sigma)$	$\operatorname{sign}(\sigma)$
$(1\,2\,3)$	0	1
$(1\,3\,2)$	1	-1
$(2\,1\,3)$	1	-1
$(2\,3\,1)$	2	1
$(3\,1\,2)$	2	1
$(3\,2\,1)$	3	-1

Thinking of permutations in S_n as bijective maps from $\{1, 2, \ldots, n\}$ to itself, we can form the composition $\sigma\tau$ of any two permutations σ and τ in S_n. This is the permutation produced by first performing τ and then σ! For instance, if $\sigma = (5\,3\,1\,2\,4)$ and $\tau = (5\,4\,3\,2\,1)$, then

$$\tau\sigma = (1\,3\,5\,4\,2), \qquad \sigma\tau = (4\,2\,1\,3\,5).$$

A fundamental fact concerning signs, which we shall not prove here, is

$$\operatorname{sign}(\sigma\tau) = \operatorname{sign}(\sigma)\operatorname{sign}(\tau). \tag{5.1}$$

In particular, the product of two even permutations is even and the product of an even and an odd permutation is odd.

The determinant formula in Theorem 5.5 contains $n!$ terms, one for each permutation σ. Each term is a product which contains exactly one entry from each row and each column of A. For instance, for $n = 5$ to the permutation $5, 3, 1, 2, 4$ contributes the term $a_{15}a_{23}a_{31}a_{42}a_{54}$. For 2×2- and 3×3-determinants Theorem 5.5 gives the well-known formulæ

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{11}a_{33} + a_{12}a_{23}a_{31} \\ + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}.$$

OUTLINE OF PROOF OF THEOREM 5.5. We need to check that the right-hand side of the determinant formula in Theorem 5.5 satisfies axioms (i)–(iii) of Definition 5.1. Let us for the moment denote the right-hand side by $f(A)$.

Axiom (i) is checked as follows: each term in $f(A)$ contains exactly one entry from each row and each column in A, so if we multiply the i-th row of A by c, each term is multiplied by c. Therefore

$$f(\mathbf{a}_1, \mathbf{a}_2, \ldots, c\mathbf{a}_i, \ldots, \mathbf{a}_n) = cf(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_i, \ldots, \mathbf{a}_n).$$

Additivity is proved in a similar way. Axiom (ii) holds because if we interchange two columns in A, each term in $f(A)$ changes sign. (This follows from (5.1) plus the fact that a permutation that interchanges two numbers i and j and leaves all others fixed is odd; see Exercise 5.4.) Finally, rule (iii) is correct because if $A = I$, the only nonzero term in $f(A)$ occurs for the identity permutation, and this term is equal to 1. The upshot is that f is a determinant function. QED

Here are some further rules obeyed by determinants. Each can be deduced from Definition 5.1 or from Theorem 5.5.

5.6. THEOREM. *Let A and B be $n \times n$-matrices.*

(i) $\det(AB) = \det A \det B$;
(ii) $\det A^T = \det A$;
(iii) *(Expansion on the jth column)* $\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$ *for all* $j = 1, 2, \dots, n$. *Here A_{ij} denotes the $(n-1) \times (n-1)$-matrix obtained from A by striking out the ith row and the jth column.*
(iv) $\det A = a_{11} a_{22} \cdots a_{nn}$ *if A is upper triangular (i.e. $a_{ij} = 0$ for $i > j$).*

Volume change

We conclude this discussion with a slightly different geometric view of determinants. A square matrix A can be regarded as a linear map $A\colon \mathbf{R}^n \to \mathbf{R}^n$. Let C be the unit cube in $\mathbf{R}^n$,

$$C = \{\mathbf{x} \in \mathbf{R}^n : 0 \le x_i \le 1 \quad \text{for } i = 1, 2, \dots, n\}.$$

Then C has n-dimensional volume 1. Its image $A(C)$ under the map A is a parallelepiped with edges $A\mathbf{e}_1$, $A\mathbf{e}_2, \dots, A\mathbf{e}_n$, the columns of A. Hence $A(C)$ has n-dimensional volume $\operatorname{vol} A(C) = |\det A| = |\det A| \operatorname{vol} C$. This rule generalizes as follows: if X is a measurable subset of $\mathbf{R}^n$, then

$$\operatorname{vol} A(X) = |\det A| \operatorname{vol} X.$$

So $|\det A|$ can be interpreted as a volume change factor. (A set is *measurable* if it has a well-defined, finite or infinite, n-dimensional volume. Explaining exactly what this means is rather hard, but it suffices for our purposes to know that all open and all closed subsets of $\mathbf{R}^n$ are measurable.)

Pulling back forms

By substituting new variables into a differential form we obtain a new form of the same degree but possibly in a different number of variables.

5.7. EXAMPLE. Consider the 1-form on $\mathbf{R}^2 - \{0\}$

$$\alpha = \frac{-y\,dx + x\,dy}{x^2 + y^2}$$

and substitute $x = \cos\theta$ and $y = \sin\theta$. We obtain the following 1-form on $\mathbf{R}$:

$$\frac{-\sin\theta\,d\cos\theta + \cos\theta\,d\sin\theta}{\cos^2\theta + \sin^2\theta} = (-\sin^2\theta + \cos^2\theta)d\theta = d\theta.$$

We formalize this example as follows. Let U be an open set in $\mathbf{R}^n$ and V an open set in $\mathbf{R}^m$. Let $\phi\colon U \to V$ be a smooth map. (E.g. $U = \mathbf{R}$, $V = \mathbf{R}^2 - \{0\}$ and $\phi(\theta) = (\cos\theta, \sin\theta)$ in the example above.) For each k-form $\phi \in \Omega^k(V)$ we shall define a k-form $\phi^*\alpha \in \Omega^k(U)$ called the *pullback* of α *along* ϕ.

First consider the case of a 0-form, that is a smooth function f on V. Then we define

$$\phi^* f = f \circ \phi,$$

the composition of ϕ and f. We will often use $\phi^* f$ as an alternative notation for $f \circ \phi$. For example, if $f(x, y) = xy$ and $\phi(\theta) = (\cos\theta, \sin\theta)$, then $(\phi^* f)(\theta) = \cos\theta\sin\theta$.

To define pullbacks of arbitrary forms, first write ϕ in components as

$$\phi(\mathbf{x}) = \big(\phi_1(x_1, x_2, \dots, x_n), \phi_2(x_1, x_2, \dots, x_n), \dots, \phi_m(x_1, x_2, \dots, x_n)\big).$$

Each of the components ϕ_i is then a scalar-valued function on U. (Thus $\phi_1(\theta) = \cos\theta$ and $\phi_2(\theta) = \sin\theta$ for $\phi(\theta) = (\cos\theta, \sin\theta)$.) For clarity denote the coordinates on $\mathbf{R}^m$ by $(y_1, \dots, y_m)$ and those on $\mathbf{R}^n$ by $(x_1, \dots, x_n)$. Let α be a k-form on V. Then

$$\alpha = \sum_I f_I\, dy_I$$

for certain functions f_I defined on V. The pullback is obtained by substituting $y_i = \phi_i(x_1, \dots, x_n)$ for all i. More formally,

$$\phi^*\alpha = \sum_I \phi^* f_I\, \phi^* dy_I,$$

where $\phi^* dy_I$ is given by

$$\phi^* dy_I = \phi^*(dy_{i_1}\, dy_{i_2} \cdots dy_{i_k}) = d\phi_{i_1}\, d\phi_{i_2} \cdots d\phi_{i_k}.$$

Observe that the pullback operation turns k-forms on the target space V into k-forms on the source space U. Thus, while $\phi\colon U \to V$ is a map from U to V, ϕ^* is a map

$$\phi^*\colon \Omega^k(V) \to \Omega^k(U),$$

the opposite way from what you might naively expect. The property that ϕ^* "turns the arrow around" is called the *contravariance* of the pullback operation.

The formula for $\phi^*\alpha$ can be made more explicit. Let us do this first in degrees 1 and 2.

The pullback of a 1-form $\alpha = \sum_{i=1}^m f_i\, dy_i$ is

$$\phi^*\alpha = \sum_{i=1}^m \phi^* f_i\, \phi^* dy_i = \sum_{i=1}^m \phi^* f_i\, d\phi_i.$$

Now $\phi^* dy_i = d\phi_i = \sum_{j=1}^n \frac{\partial\phi_i}{\partial x_j} dx_j$ and so

$$\phi^*\alpha = \sum_{i=1}^m \left(\phi^* f_i \sum_{j=1}^n \frac{\partial\phi_i}{\partial x_j} dx_j\right) = \sum_{j=1}^n g_j\, dx_j,$$

with $g_j = \sum_{i=1}^m \phi^* f_i \frac{\partial\phi_i}{\partial x_j}$.

For a 2-form $\alpha = \sum_{1\le i<j\le m} f_{ij}\, dy_i\, dy_j$ we get

$$\phi^*\alpha = \sum_{1\le i<j\le m} \phi^* f_{ij}\, \phi^*(dy_i\, dy_j) = \sum_{1\le i<j\le m} \phi^* f_{ij}\, d\phi_i\, d\phi_j.$$

Observe that

$$\begin{aligned} d\phi_i d\phi_j &= \sum_{k,l=1}^{n} \frac{\partial \phi_i}{\partial x_k}\frac{\partial \phi_j}{\partial x_l} dx_k dx_l \\ &= \sum_{1 \le k < l \le n} \Big(\frac{\partial \phi_i}{\partial x_k}\frac{\partial \phi_j}{\partial x_l} - \frac{\partial \phi_i}{\partial x_l}\frac{\partial \phi_j}{\partial x_k} \Big) dx_k dx_l, \end{aligned}$$

where

$$\frac{\partial \phi_i}{\partial x_k}\frac{\partial \phi_j}{\partial x_l} - \frac{\partial \phi_i}{\partial x_l}\frac{\partial \phi_j}{\partial x_k} = \begin{vmatrix} \frac{\partial \phi_i}{\partial x_k} & \frac{\partial \phi_i}{\partial x_k} \\ \frac{\partial \phi_j}{\partial x_k} & \frac{\partial \phi_j}{\partial x_l} \end{vmatrix},$$

is the determinant of the 2×2-submatrix obtained from the Jacobi matrix by extracting rows i and j and columns k and l. So we get

$$\phi^*\alpha = \sum_{1 \le i < j \le m} \left(\phi^* f_{ij} \sum_{1 \le k < l \le n} \begin{vmatrix} \frac{\partial \phi_i}{\partial x_k} & \frac{\partial \phi_i}{\partial x_k} \\ \frac{\partial \phi_j}{\partial x_k} & \frac{\partial \phi_j}{\partial x_l} \end{vmatrix} dx_k dx_l \right) = \sum_{1 \le k < l \le n} g_{kl} dx_k dx_l$$

with

$$g_{kl} = \sum_{1 \le i < j \le m} \phi^* f_{ij} \begin{vmatrix} \frac{\partial \phi_i}{\partial x_k} & \frac{\partial \phi_i}{\partial x_k} \\ \frac{\partial \phi_j}{\partial x_k} & \frac{\partial \phi_j}{\partial x_l} \end{vmatrix}.$$

For an arbitrary k-form $\alpha = \sum_I f_I dy_I$ we obtain

$$\phi^*\alpha = \sum_I \phi^* f_I \, \phi^*(dy_{i_1} \cdots dy_{i_k}) = \sum_I \phi^* f_I \, d\phi_{i_1} \cdots d\phi_{i_k}.$$

Now

$$\begin{aligned} d\phi_{i_1} \cdots d\phi_{i_k} &= \sum_{j_1, j_2, \dots, j_k = 1}^{n} \frac{\partial \phi_{i_1}}{\partial x_{j_1}} \frac{\partial \phi_{i_2}}{\partial x_{j_2}} \cdots \frac{\partial \phi_{i_k}}{\partial x_{j_k}} dx_{j_1} dx_{j_2} \cdots dx_{j_k} \\ &= \sum_J \sum_{\sigma \in S_k} \operatorname{sign}(\sigma) \frac{\partial \phi_{i_1}}{\partial x_{j_{\sigma(1)}}} \frac{\partial \phi_{i_2}}{\partial x_{j_{\sigma(2)}}} \cdots \frac{\partial \phi_{i_k}}{\partial x_{j_{\sigma(k)}}} dx_J \\ &= \sum_J \det d\phi_{IJ} dx_J, \end{aligned}$$

where we have used formula of Theorem 5.5 for the determinant. Here the sum is over all increasing multi-indices J of degree k, and $(d\phi)_{IJ}$ denotes the IJ-*submatrix* of $d\phi$, that is the $k \times k$-matrix obtained from the Jacobi matrix by extracting rows $i_1, i_2, \dots, i_k$ and columns $j_1, j_2, \dots, j_k$. Therefore

$$\begin{aligned} \phi^*\alpha &= \sum_I \phi^* f_I \sum_J \det d\phi_{IJ} dx_J \\ &= \sum_J \Big(\sum_I \phi^* f_I \det d\phi_{IJ} \Big) dx_J. \end{aligned}$$

This proves the following result.

5.8. THEOREM. *Let $\phi \colon U \to V$ be a smooth map, where U is open in $\mathbf{R}^n$ and V is open in $\mathbf{R}^m$. Let $\alpha = \sum_I f_I dy_I$ be a k-form on V. Then $\phi^*\alpha$ is the k-form on U given by $\phi^*\alpha = \sum_J g_J dx_J$ with*

$$\boxed{g_J = \sum_I \phi^* f_I \det d\phi_{IJ}.}$$

This formula is seldom used to calculate pullbacks in practice. It is almost always easier to apply the definition of pullback directly. However, the formula has some important theoretical uses, one of which we record here.

Assume that $k = m = n$, that is to say, the number of new variables is equal to the number of old variables, and we are pulling back a form of top degree. Then

$$\alpha = f\,dy_1\,dy_2\cdots dy_n, \qquad \phi^*\alpha = \phi^* f\ \det d\phi\ dx_1\,dx_2\cdots dx_n.$$

If $f = 1$ then $\phi^* f = 1$, so we see that the Jacobi determinant $d\phi(\mathbf{x})$ can be interpreted as the ratio between the oriented volumes of an infinitesimal block positioned at $\mathbf{x}$ and its image under ϕ.

5.9. THEOREM. *Let $\phi\colon U \to V$ be a smooth map, where U and V are open in $\mathbf{R}^n$. Then the pullback of the volume form on V is equal to the Jacobi determinant times the volume form on U,*

$$\boxed{\phi^*(dy_1\,dy_2\cdots dy_n) = \det d\phi\ dx_1\,dx_2\cdots dx_n.}$$

Pulling back forms is nicely compatible with the other operations that we learned about (except the Hodge star).

5.10. PROPOSITION. *Let $\phi\colon U \to V$ be a smooth map, where U is open in $\mathbf{R}^n$ and V is open in $\mathbf{R}^m$. The pullback operation is*

(i) *linear:* $\phi^*(c_1\alpha_1 + c_2\alpha_2) = c_1\phi^*\alpha_1 + c_2\phi^*\alpha_2$;
(ii) *multiplicative:* $\phi^*(\alpha\beta) = (\phi^*\alpha)(\phi^*\beta)$;
(iii) *natural*: $\phi^*(\psi^*\alpha) = (\psi\circ\phi)^*\alpha$, *where $\psi\colon V \to W$ is a second smooth map with W open in $\mathbf{R}^k$ and α a form on W.*

The term "natural" in property (iii) is a mathematical catchword meaning that a certain operation (in this case the pullback) is well-behaved with respect to composition of maps.

OUTLINE OF PROOF. We skip the proof of (i). For the proof of (ii) consider two forms $\alpha = \sum_I f_I\,dy_I$ and $\beta = \sum_J g_J\,dy_J$. Then $\alpha\beta = \sum_{I,J} f_I g_J\,dy_I\,dy_J$, so

$$\phi^*(\alpha\beta) = \sum_{I,J} \phi^*(f_I g_J)\phi^*(dy_I\,dy_J).$$

Now

$$\phi^*(f_I g_J)(x) = f_I g_J(\phi(x)) = f_I(\phi(x))g_J(\phi(x)) = (\phi^* f_I)(\phi^* g_J)(x),$$

so $\phi^*(f_I g_J) = (\phi^* f_I)(\phi^* g_J)$. Furthermore,

$$\begin{aligned}\phi^*(dy_I\,dy_J) &= \phi^*(dy_{i_1}\cdots dy_{i_k}\,dy_{j_1}\cdots dy_{j_l})\\ &= d\phi_{i_1}\cdots d\phi_{i_k}\,d\phi_{j_1}\cdots d\phi_{j_l} = \phi^*(dy_I)(\phi^* dy_J),\end{aligned}$$

so

$$\begin{aligned}\phi^*(\alpha\beta) &= \sum_{IJ} (\phi^* f_I)(\phi^* g_J)\phi^*(dy_I)(\phi^* dy_J)\\ &= \Big(\sum_I (\phi^* f_I)\phi^*(dy_I)\Big)\Big(\sum_I (\phi^* g_J)(\phi^* dy_J)\Big) = (\phi^*\alpha)(\phi^*\beta),\end{aligned}$$

which establishes (ii).

For the proof of (iii) first consider a function f on W. Then

$$\begin{aligned}\phi^*(\psi^* f)(x) = (\psi^* f)(\phi(x)) = f(\psi(\phi(x))) &= f(\psi \circ \phi(x)) \\ &= (f \circ (\psi \circ \phi))(x) = (\psi \circ \phi)^* f(x),\end{aligned}$$

so $\phi^*(\psi^* f) = (\psi \circ \phi)^* f$. Next consider a 1-form $\alpha = dz_i$ on W, where $z_1, z_2, \dots, z_k$ are the variables on $\mathbf{R}^k$. Then $\psi^* \alpha = d\psi_i = \sum_{j=1}^m \frac{\partial \psi_i}{\partial y_j} dy_j$, so

$$\begin{aligned}\phi^*(\psi^* \alpha) &= \sum_{j=1}^m \phi^*\Big(\frac{\partial \psi_i}{\partial y_j}\Big) \phi^* dy_j \\ &= \sum_{j=1}^m \phi^*\Big(\frac{\partial \psi_i}{\partial y_j}\Big) d\phi_j \\ &= \sum_{j=1}^m \phi^*\Big(\frac{\partial \psi_i}{\partial y_j}\Big) \sum_{l=1}^n \frac{\partial \phi_j}{\partial x_l} dx_l \\ &= \sum_{l=1}^n \Big(\sum_{j=1}^m \phi^*\Big(\frac{\partial \psi_i}{\partial y_j}\Big) \frac{\partial \phi_j}{\partial x_l} \Big) dx_l \\ &= \sum_{l=1}^n \frac{\partial (\phi^* \psi_i)}{\partial x_l} dx_l \qquad \text{by chain rule} \\ &= d(\phi^* \psi_i) \\ &= d((\psi \circ \phi)_i) \\ &= (\psi \circ \phi)^* dz_i \\ &= (\psi \circ \phi)^* \alpha.\end{aligned}$$

Because every form on W is a sum of products of forms of type f and dz_i, the general case of (iii) follows from these two special cases. QED

The essential step in the proof of part (iii) relies on the chain rule in several variables in the following form. If $f(y_1, \dots, y_m)$ is a function of m variables and we substitute $y = \phi(x)$, i.e. $y_1 = \phi_1(x_1, \dots, x_n), \dots, y_m = \phi_m(x_1, \dots, x_n)$, then

$$\frac{\partial (f \circ \phi)}{\partial x_l}(x) = \sum_{j=1}^m \frac{\partial f}{\partial y_j}(\phi(x)) \frac{\partial \phi_j}{\partial x_l}(x). \tag{5.2}$$

Written in the language of pullbacks, this becomes

$$\frac{\partial (\phi^* f)}{\partial x_l} = \sum_{j=1}^m \phi^*\Big(\frac{\partial f}{\partial y_j}\Big) \frac{\partial \phi_j}{\partial x_l},$$

which is exactly what was used in the proof above (with $f = \psi_i$).

Another application of the chain rule yields the following important result.

5.11. THEOREM. *Let $\phi\colon U \to V$ be a smooth map, where U is open in $\mathbf{R}^n$ and V is open in $\mathbf{R}^m$. Then $\phi^*(d\alpha) = d(\phi^* \alpha)$ for $\alpha \in \Omega^k(V)$. In short*

$$\boxed{\phi^* d = d\phi^*.}$$

PROOF. First let f be a function. Then

$$\begin{aligned}
\phi^* df &= \phi^* \sum_{i=1}^{m} \frac{\partial f}{\partial y_i} dy_i \\
&= \sum_{i=1}^{m} \phi^*\Big(\frac{\partial f}{\partial y_i}\Big) d\phi_i \\
&= \sum_{i=1}^{m} \phi^*\Big(\frac{\partial f}{\partial y_i}\Big) \sum_{j=1}^{n} \frac{\partial \phi_i}{\partial x_j} dx_j \\
&= \sum_{j=1}^{n} \sum_{i=1}^{m} \phi^*\Big(\frac{\partial f}{\partial y_i}\Big) \frac{\partial \phi_i}{\partial x_j} dx_j \\
&= \sum_{j=1}^{n} \frac{\partial(\phi^* f)}{\partial x_j} dx_j \qquad\qquad \text{by chain rule} \\
&= d\phi^* f,
\end{aligned}$$

so the statement is true for functions. Next let $\alpha = f dy_I$. Then $d\alpha = df dy_I$, so

$$\phi^* d\alpha = \phi^*(df dy_I) = (\phi^* df)(\phi^* dy_I) = d(\phi^* f) d\phi_{i_1} \cdots d\phi_{i_k},$$

because $\phi^* df = d\phi^* f$. On the other hand,

$$\begin{aligned}
d\phi^* \alpha &= d\big((\phi^* f)(\phi^* dy_I)\big) \\
&= d\big((\phi^* f) d\phi_{i_1} \cdots d\phi_{i_k}\big) \\
&= d(\phi^* f) d\phi_{i_1} \cdots d\phi_{i_k} + (\phi^* f) d(d\phi_{i_1} \cdots d\phi_{i_k}) \\
&= d(\phi^* f) d\phi_{i_1} \cdots d\phi_{i_k}.
\end{aligned}$$

Here we have used the Leibniz rule for forms, Proposition 4.4(ii), plus the fact that the form $d\phi_{i_1} \cdots d\phi_{i_k}$ is always closed. (See Problem 4.9.) Comparing the two equations above we see that $\phi^* d\alpha = d\phi^* \alpha$. Any form is a linear combination of forms of type $f dy_I$, so the general case now follows. QED

Exercises

5.1. Deduce Theorem 5.6(iv) from Theorem 5.5.

5.2. Calculate the following determinants using column or row operations and Exercise 5.1.

$$\begin{vmatrix} 1 & 3 & 1 & 1 \\ 2 & 1 & 5 & 2 \\ 1 & -1 & 2 & 3 \\ 4 & 1 & -3 & 7 \end{vmatrix}, \qquad \begin{vmatrix} 1 & 1 & -2 & 4 \\ 0 & 1 & 1 & 3 \\ 2 & -1 & 1 & 0 \\ 3 & 1 & 2 & 5 \end{vmatrix}.$$

5.3. Tabulate all permutations in S_4 with their lengths and signs.

5.4. Determine the length and the sign of the following permutations.

(i) A permutation of the form $(1\,2\,\dots\,i-1\,j\,\dots\,j-1\,i\,\dots\,n)$ where $1 \le i < j \le n$. (Such a permutation is called a *transposition*. It interchanges i and j and leaves all other numbers fixed.)

(ii) $(n\,n-1\,n-2\,\dots\,3\,2\,1)$.

5.5. Find all permutations in S_n of length 1.

5.6. Calculate σ^{-1}, τ^{-1}, $\sigma\tau$ and $\tau\sigma$, where

(i) $\sigma = (3\,6\,1\,2\,5\,4)$ and $\tau = (5\,2\,4\,6\,3\,1)$;
(ii) $\sigma = (2\,1\,3\,4\,5\,\ldots\,n-1\,n)$ and $\tau = (n\,1\,2\,\ldots\,n-2\,n-1\,1)$ (i.e. the transpositions interchanging 1 and 2, resp. 1 and n).

5.7. Show that

$$dx_{i_{\sigma(1)}}\,dx_{i_{\sigma(2)}}\cdots dx_{i_{\sigma(k)}} = \operatorname{sign}(\sigma)\,dx_{i_1}\,dx_{i_2}\cdots dx_{i_k}$$

for any multi-index $(i_1, i_2, \ldots, i_k)$ and any permutation σ in S_k.

5.8. (i) Show that every permutation has the same length and sign as its inverse.
(ii) Deduce from Theorem 5.5 that $\det A^T = \det A$.

5.9. (i) Suppose that A has the shape

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ 0 & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \\ 0 & a_{n1} & \ldots & a_{nn} \end{pmatrix},$$

i.e. all entries below a_{11} are 0. Deduce from Theorem 5.5 that

$$\det A = a_{11} \begin{vmatrix} a_{22} & \ldots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \ldots & a_{nn} \end{vmatrix}.$$

(ii) Deduce from this the expansion rule, Theorem 5.6(iii).

5.10. Show that

$$\begin{vmatrix} 1 & 1 & \ldots & 1 \\ x_1 & x_2 & \ldots & x_n \\ x_1^2 & x_2^2 & \ldots & x_n^2 \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \ldots & x_n^{n-1} \end{vmatrix} = \prod_{i<j}(x_j - x_i)$$

for any numbers $x_1, x_2, \ldots, x_n$.

5.11. Define a map $P\colon \mathbf{R}^3 \to \mathbf{R}^3$ by $P(r,\phi,\theta) = (r\cos\phi\cos\theta, r\cos\phi\sin\theta, r\sin\phi)$ (spherical coordinates). Calculate $P^*\alpha$ for the following forms α:

$$dx, \quad dy, \quad dz, \quad dx\,dy, \quad dx\,dz, \quad dy\,dz, \quad dx\,dy\,dz.$$

5.12. Compute $f^*(x\,dy\,dz + y\,dz\,dx + z\,dx\,dy)$, where f is the stereographic projection defined in Exercise 2.8.