Advanced differential calculus

Hörmander, Lars

1994

ADVANCED
DIFFERENTIAL CALCULUS

Lars Hörmander

Department of Mathematics

Box 118, S-221 00 Lund, Sweden

ADVANCED DIFFERENTIAL CALCULUS

Lars Hörmander

Matematiska Institutionen

Box 118, 221 00 Lund

Lars Hörmander
Matematiska Institutionen
Lunds Universitet
Box 118
221 00 Lund
e-mail lvh@maths.lth.se

These lecture notes have been typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX.

PREFACE

During the academic year 1973–74 I gave a series of lectures entitled “Högre differentialkalkyl”. Lecture notes were published in Swedish with the assistance of Tomas Claesson and Arne Enqvist. From the preface of those notes I quote in a free translation: “The purpose of this series of lectures was to present basic facts on differential geometry and differential calculus on manifolds, with some topological applications, starting from elementary differential and integral calculus. Unfortunately the time did not suffice to carry out the plans. For example, Riemannian geometry and residue calculus in several complex variables are missing. Characteristic classes of complex vector bundles are defined, but their properties are not developed.”

These notes are essentially a translation with improved typography, a number of minor corrections and a few added explanations and references. The missing Riemannian geometry was discussed in a series of lectures during the academic year 1976–77, and a revised set of lecture notes was produced in 1990. However, residue calculus in several variables has been much advanced since 1973, and it is not possible to cover this topic adequately here.

Lund in August 1994

Lars Hörmander

Typeset by $\mathcal{AMS}$-TEX

CONTENTS

Typeset by $\mathcal{AMS}$-TEX

CHAPTER I

DIFFERENTIABLE FUNCTIONS

At first we shall only consider functions of *one* real variable but we allow them to take their values in a Banach space. Thus let $I \subset \mathbf{R}$ be an open interval, V a Banach space, and let f be a map $I \to V$. Then f is called *differentiable* at $x \in I$ if there is an element $f'(x) \in V$ such that

$$\|(f(x+h) - f(x))/h - f'(x)\| \to 0 \quad \text{when } h \to 0. \tag{1.1}$$

Equivalently we can write (1.1) in the form

$$\|f(x+h) - f(x) - f'(x)h\| = o(|h|), \quad \text{when } h \to 0. \tag{1.1$'$}$$

If $V = \mathbf{R}^n$ and we write $f = (f_1, \dots, f_n))$, this is obviously equivalent to differentiability of each component f_j of f. The mean value theorem is replaced by the following

THEOREM 1.1. *If $f : I \to V$ is differentiable at every point in the open interval I then*

$$\|f(x) - f(y)\| \le |x-y| \sup_{z \in [x,y]} \|f'(z)\|, \quad x, y \in I. \tag{1.2}$$

PROOF. Let $M > \sup_{z\in[x,y]} \|f'(z)\|$, and set

$$E = \{t; 0 \le t \le 1,\ \|f(x+t(y-x)) - f(x)\| \le Mt|y-x|\}.$$

For a fixed $s \in E$ with $s < 1$ we have if $t > s$ and $t - s$ is sufficiently small

$$\begin{aligned}\|f(x+t(y-x))-f(x)\| &\le \|f(x+t(y-x))-f(x+s(y-x))\|+\|f(x+s(y-x))-f(x)\| \\ &\le M|(t-s)(y-x)| + Ms|y-x| = Mt|y-x|.\end{aligned}$$

The set E is closed since f is continuous, and $0 \in E$. Hence the supremum of E belongs to E, and we have just proved that it is not in $[0,1)$, so $1 \in E$, that is,

$$\|f(x) - f(y)\| \le M|y-x|, \quad \text{if } M > \sup_{[x,y]} \|f'\|,$$

which proves (1.2).

Corollary 1.2. *Under the hypotheses in Theorem 1.1 we have*

$$(1.2)' \qquad \|f(y) - f(x) - f'(x)(y-x)\| \le |y-x| \sup_{z\in[x,y]} \|f'(z) - f'(x)\|.$$

Proof. The function $g(y) = f(y) - f(x) - f'(x)(y-x)$ is differentiable with $g'(y) = f'(y) - f'(x)$, and $(1.2)'$ follows if (1.2) is applied to g.

We shall now generalize the notion of differentiability to functions defined in an open subset Ω of another Banach space U, still with values in the Banach space V. We shall denote by $\mathcal{L}(U,V)$ the space of continuous linear maps $U \to V$ with the standard norm

$$\|T\| = \sup_{x\in U, \|x\|\le 1} \|Tx\|, \quad T \in \mathcal{L}(U,V).$$

Definition 1.3. A function $f : \Omega \to V$ where V is a Banach space and Ω is an open subset of another Banach space U is called differentiable at $x \in \Omega$ if there is an element $f'(x) \in \mathcal{L}(U,V)$ such that

$$(1.1)'' \qquad \|f(x+h) - f(x) - f'(x)h\| = o(\|h\|), \quad \text{when } h \to 0.$$

By $C^1(\Omega, V)$ we shall denote the set of continuously differentiable functions from Ω to V, that is, the set of functions which are differentiable at every point in Ω and for which $\Omega \ni x \mapsto f'(x) \in \mathcal{L}(U,V)$ is continuous.

Exercise 1.4. Prove using Corollary 1.2 that if $U = \mathbf{R}^n$ and $V = \mathbf{R}^m$ then $f = (f_1, \dots, f_m)$ is in $C^1(\Omega, V)$ if and only if the partial derivatives $\partial f_j(x)/\partial x_k$ exist for $x \in \Omega$, $j = 1, \dots, m$, $k = 1, \dots, n$, and are continuous in Ω.

Exercise 1.5. Let K be a compact subset of $\mathbf{R}^n$, let Ω be an open subset of $\mathbf{R}^n \times \mathbf{R}^m$, and let $\Omega \ni (x,y) \mapsto A(x,y) \in V$ be a continuous function such that the differential $A'_y(x,y)$ for fixed x exists for all $(x,y) \in \Omega$ and is a continuous function in Ω. Prove that

$$F = \{f \in C(K, \mathbf{R}^m); (x, f(x)) \in \Omega \ \forall x \in K\}$$

is an open subset of $C(K, \mathbf{R}^m)$, and that $F \ni f \mapsto A(\cdot, f)$ belongs to $C^1(F, C(K,V))$.

Here $C(K,V)$ denotes the space of continuous functions $f : K \to V$ with the norm $\sup_{x\in K} \|f(x)\|$.

Exercise 1.6. Let Ω be the set of invertible operators in $\mathcal{L}(U,V)$. Prove that Ω is open, that $\Omega \ni T \mapsto T^{-1}$ is in $C^1(\Omega, \mathcal{L}(V,U))$, and that the differential at T is

$$\mathcal{L}(U,V) \ni S \mapsto -T^{-1}ST^{-1}.$$

If $f \in \mathcal{L}(U,V)$ then f is differentiable and $f'(x) = f$ for every x. Let us more generally consider the space $\mathcal{L}(U_1, \dots, U_k; V)$ of continuous multilinear maps

$$U_1 \times \cdots \times U_k \ni (x_1, \dots, x_k) \mapsto f(x_1, \dots, x_k) \in V.$$

Continuity at the origin implies that

$$\|f\| = \sup_{\|x_j\|\le 1, j=1,\dots,k} \|f(x_1, \dots, x_k)\| < \infty,$$

and we leave as an exercise to prove that conversely this implies continuity everywhere. With this norm $\mathcal{L}(U_1, \dots, U_k; V)$ is a Banach space. The map

$$U_1 \oplus \cdots \oplus U_k \ni (x_1, \dots, x_k) \mapsto f(x_1, \dots, x_k) \in V$$

is differentiable for every x, and the differential is

$$\begin{aligned}
&(U_1 \oplus \cdots \oplus U_k) \ni (y_1, \dots, y_k) \\
&\qquad \mapsto f(y_1, x_2, \dots, x_k) + f(x_1, y_2, \dots, x_k) + \cdots + f(x_1, x_2, \dots, x_{k-1}, y_k).
\end{aligned}$$

The standard rules of differentiation are consequences of the preceding observations and the following discussion of composite maps.

Thus let $f : \Omega \to V$ where Ω is an open subset of U, assume that $f(\Omega) \subset \Omega'$ where Ω' is open in V, and let $g : \Omega' \to W$ where W is a third Banach space. If f is differentiable at a point $x \in \Omega$ and g is differentiable at $y = f(x) \in \Omega'$, then $h = g \circ f$ is differentiable at x and

$$h'(x) = g'(y)f'(x). \quad \text{(The chain rule.)} \tag{1.3}$$

The proof is obvious:

$$\begin{aligned}
h(x + x') = g(f(x + x')) &= g(y + f(x + x') - f(x)) \\
&= g(y) + g'(y)(f(x + x') - f(x)) + o(\|f(x + x') - f(x)\|) \\
&= g(y) + g'(y)f'(x)x' + o(\|x'\|).
\end{aligned}$$

From (1.3) it also follows that $h \in C^1$ if $g \in C^1$ and $f \in C^1$.

The differential f' can be viewed as a map

$$\Omega \times U \ni (x, \xi) \overset{f'}{\mapsto} (f(x), f'(x)\xi) \in \Omega' \times V,$$

which is linear along U.Then the chain rule states that if we have a commutative diagram

$$\begin{tikzcd}
 & \Omega' \arrow[dr, "g"] & \\
\Omega \arrow[ur, "f"] \arrow[rr, "h"] & & W
\end{tikzcd}$$

where $f, g \in C^1$, then $h \in C^1$ and we obtain a new commutative diagram

$$\begin{tikzcd}
 & \Omega' \times V \arrow[dr, "g'"] & \\
\Omega \times U \arrow[ur, "f'"] \arrow[r, "h'"] & W & \times W
\end{tikzcd}$$

EXERCISE 1.7. Prove that if f is differentiable at every point in the interval $I = [x, y] = \{tx + (1-t)y; 0 \le t \le 1\}$, then

$$\|f(x) - f(y)\| \le \|x - y\| \sup_I \|f'\|.$$

The notation df is often used instead of f', particularly when f is real valued. If f is defined in an open subset of $\mathbf{R}^n$ and we write $t = \sum_1^n t_j e_j$ where e_j is the jth unit vector, we obtain

$$(df)(t) = (df)\Big(\sum_1^n t_j e_j\Big) = \sum_1^n t_j df(e_j) = \sum_1^n \partial f/\partial x_j t_j.$$

Since $t_j = (dx_j)(t)$ we can write this equation in the form

$$df = \sum \partial f/\partial x_j dx_j.$$

According to the chain rule this formula remains valid if x_j are functions of $y \in U$ and both f and x_j are regarded as functions in U so that both sides are linear functions on U. This is called the *invariance of the differential.*

We can inductively define $C^k(\Omega, V)$ when k is an integer > 1 as the set of all $f \in C^1(\Omega, V)$ such that $f' \in C^{k-1}(\Omega, \mathcal{L}(U, V))$, and inductively we define

$$f^{(k)} \in C(\Omega, \mathcal{L}(U, \mathcal{L}(U, \dots, \mathcal{L}(U, V)))).$$

The vector space $\mathcal{L}(U, \mathcal{L}(U, \dots, \mathcal{L}(U, V)))$ is isomorphic as a Banach space to the space $\mathcal{L}(U, \dots, U; V)$ of k linear maps from U to V, for we have quite generally

$$\mathcal{L}(U, \mathcal{L}(U_1, \dots, U_j; V)) = \mathcal{L}(U, U_1, \dots, U_j; V). \tag{1.5}$$

In fact, every element T in the left-hand side gives rise to a multilinear map

$$U \times U_1 \times \cdots \times U_j \ni (x, x_1, \dots, x_j) \mapsto T(x)(x_1, \dots, x_j) \in V,$$

that is, an element in the right-hand side, and every such element can be obtained in this way with $T(x)$ defined by fixing the variable $x \in U$. It is obvious that the identification (1.5) is linear, and it is an easy exercise to verify that it is norm preserving.

We shall denote by $\mathcal{L}_s^k(U, V)$ the symmetric k linear maps (forms) from U to V such that the value at $(x_1, \dots, x_k)$ does not change if the variables $x_1, \dots, x_k$ are permuted. The following theorem states that the order of differentiation is irrelevant:

THEOREM 1.8. *If $f \in C^k(\Omega, V)$ then $f^{(k)}$ is a symmetric multilinear form in V.*

PROOF. It suffices to prove the statement when $k = 2$, for if a multilinear form is invariant for interchange of two ajacent variables it is invariant for arbitrary permutations. We shall prove that if $f \in C^2(\Omega, V)$ and $x \in \Omega$, $x_1, x_2 \in U$, then

$$f(x + x_1 + x_2) - f(x + x_1) - f(x + x_2) + f(x) = f''(x)(x_2, x_1) + o(\|x_1\|\|x_2\|) \tag{1.6}$$

when $x_1 \to 0$ and $x_2 \to 0$. Since the left-hand side is symmetrical in x_1 and x_2 we obtain by interchanging the vectors and subtracting that

$$f''(x)(x_1, x_2) - f''(x)(x_2, x_1) = o(\|x_1\|\|x_2\|).$$

Replacing x_1 and x_2 by εx_1 and εx_2 and dividing by ε^2 we conclude when $\varepsilon \to +0$ that $f''(x)(x_1, x_2) = f''(x)(x_2, x_1)$.

It remains to verify (1.6). The left-hand side can be written $g(x + x_1) - g(x)$ where $g(x) = f(x + x_2) - f(x)$. According to Exercise 1.7 we have in analogy with Corollary 1.2

$$\|g(x + x_1) - g(x) - g'(x)x_1\| \le \|x_1\| \sup_{0 \le t \le 1} \|g'(x + tx_1) - g'(x)\|.$$

Since $g'(x) = f'(x+x_2) - f'(x)$ another application of Exercise 1.7 gives for $0 \le t \le 1$

$$\|g'(x + tx_1) - f''(x)x_2)\| \le \sup_{0 \le s \le 1} \|f''(x + tx_1 + sx_2) - f''(x)\|\|x_2\| = o(\|x_2\|)$$

when $x_1 \to 0$ and $x_2 \to 0$. Hence

$$\|g(x + x_1) - g(x) - (f''(x)x_2)x_1\| = o(\|x_1\|\|x_2\|),$$

which completes the proof.

From (1.3) it follows at once by induction that $h = g \circ f \in C^k$ if $g, f \in C^k$. In fact, if $k > 1$ and this is proved with k replaced by $k - 1$, we have $g' \circ f \in C^{k-1}$ and $f' \in C^{k-1}$. Since the bilinear map $L(V, W) \times L(U, V) \to L(U, W)$ defined by multiplication of operators is continuous, hence infinitely differentiable, it follows from (1.3) that $h' \in C^{k-1}$, hence $h \in C^k$. It is clear that one can obtain a rather complicated formula for $h^{(k)}$. If $g' = \cdots = g^{(k-1)} = 0$ at $y_0 = f(x_0)$ it simplifies to

$$h^{(k)}(x_0)(t_1, \ldots, t_k) = g^{(k)}(f'(x)t_1, \ldots, f'(x)t_k). \tag{1.3$'$}$$

This is proved by entering $y = f(x) - f(x_0)$ in the Taylor expansion

$$g(y_0 + y) = g(y_0) + g'(y_0)y + \tfrac{1}{2}g''(y_0)(y, y) + \cdots + \tfrac{1}{k!}g^{(k)}(y_0)(y, \ldots, y) + o(\|y\|^k).$$

Notes. For a more detailed presentation of the topics discussed in this chapter we refer to the following books.

References

J. Dieudonné, *Foundations of modern analysis, Chap. VIII*, Academic Press, 1969.
L. Hörmander, *The analysis of linear partial differential operators, Chap. I*, Springer Verlag, 1983.
S. Lang, *Differentiable manifolds*, Interscience Publ., 1962.

CHAPTER II

INVERSE FUNCTIONS

The purpose of differential calculus is to reduce the study of general functions to the much simpler case of linear functions. The inverse function theorem is a basic result in this direction:

THEOREM 2.1. *Let $f \in C^k(\Omega, V)$ where $k \ge 1$ and Ω is open in U, and let $x_0 \in \Omega$, $f(x_0) = y_0$. In order that there shall exist a function $g \in C^k(\Omega', U)$ defined in a neighborhood Ω' of y_0 such that*

a) $f \circ g = \mathrm{Id}$ *in a neighborhood of y_0; or*
b) $g \circ f = \mathrm{Id}$ *in a neighborhood of x_0; or*
c) $f \circ g = \mathrm{Id}$ *in a neighborhood of y_0 and $g \circ f = \mathrm{Id}$ in a neighborhood of x_0;*

it is necessary and sufficient that there exists a linear transformation $A \in \mathcal{L}(V, U)$ such that respectively

a) $f'(x_0)A = \mathrm{Id}_V$*;*
b) $Af'(x_0) = \mathrm{Id}_U$*;*
c) $f'(x_0)A = \mathrm{Id}_V$ *and* $Af'(x_0) = \mathrm{Id}_U$.

The infinitesimal condition c) *is by Banach's theorem equivalent to bijectivity of $f'(x_0)$, and g is then uniquely determined in a neighborhood of x_0. If V or U is of finite dimension then* a) *(resp.* b)*) is equivalent to surjectivity (resp. injectivity) of $f'(x_0)$.*

Here Id denotes the identity (in a space indicated by a subscript), and U, V are Banach spaces. In case c) one calls f a *local diffeomorphism.*

PROOF. The necessity is an immediate consequence of the chain rule (1.3). To prove the sufficiency we first note that if $f \circ g_1 = \mathrm{Id}$ in a neighborhood of y_0 and $g_2 \circ f = \mathrm{Id}$ in a neighborhood of x_0, then $g_1 = g_2 \circ f \circ g_1 = g_2$ in a neighborhood of y_0, which proves uniqueness in case c), so we only have to prove existence in cases a) and b). Replacing f by $f \circ A$ resp. $A \circ f$ we find that it suffices to consider the case where $U = V$ and $f'(x_0) = \mathrm{Id}$. Choose $\delta > 0$ so that

$$\|f'(x) - \mathrm{Id}\,\| < \tfrac{1}{2} \quad \text{when } \|x - x_0\| \le \delta.$$

For $\|x_j - x_0\| \le \delta$, $j = 1, 2$, it follows that

$$\|(f(x_1) - f(x_2)) - (x_1 - x_2)\| \le \tfrac{1}{2}\|x_1 - x_2\|. \tag{2.1}$$

(See Exercise 1.7.) Hence f is injective in $\{x \in U; \|x - x_0\| \le \delta\}$. To solve the equation $f(x) = y$ when $\|y - y_0\| < \frac{1}{2}\delta$ we use the iteration scheme

$$x_k = x_{k-1} + y - f(x_{k-1}), \quad k = 1, 2, \dots \tag{2.2}$$

as long as it leads to points with $\|x_k - x_0\| \le \delta$. We have $\|x_1 - x_0\| = \|y - y_0\| \le \frac{1}{2}\delta$. If $k > 1$ and $\|x_j - x_0\| \le \delta$ and $\|x_j - x_{j-1}\| \le 2^{-j}\delta$ when $0 < j < k$, then the equation

$$x_k - x_{k-1} = x_{k-1} - f(x_{k-1}) - (x_{k-2} - f(x_{k-2}))$$

implies the estimate

$$\|x_k - x_{k-1}\| \le \tfrac{1}{2}\|x_{k-1} - x_{k-2}\| \le 2^{-k}\delta$$

by (2.1), so we obtain using the triangle inequality

$$\|x_k - x_0\| \le \delta \sum_1^k 2^{-j} < \delta.$$

Hence x_k is defined for every k and is a Cauchy sequence. If x is the limit then $\|x - x_0\| \le \delta$, and letting $k \to \infty$ in (2.2) we obtain $f(x) = y$.

To prove that the inverse $g(y) = x$ which is now defined when $\|y - y_0\| < \frac{1}{2}\delta$ is in C^1 we set

$$g(y) = x, \quad g(y+k) = x + h.$$

This means that $f(x+h) = y + k$ and that $f(x) = y$. Hence

$$k = f(x+h) - f(x) = f'(x)h + o(\|h\|).$$

From (2.1) it follows that $\|k - h\| \le \frac{1}{2}\|h\|$, hence

$$\tfrac{1}{2}\|h\| \le \|k\| \le \tfrac{3}{2}\|h\|.$$

When x is sufficiently close to x_0 we conclude that

$$h = f'(x)^{-1}k + o(\|k\|),$$

for $f'(x)^{-1}$ exists when $\|x - x_0\| \le \delta$ since $\|f'(x) - \operatorname{Id}\| < \frac{1}{2}$; we have

$$f'(x)^{-1} = \sum_0^\infty (\operatorname{Id} - f'(x))^j.$$

Thus $g'(y)$ exists and is equal to $f'(x)^{-1}$. If $f \in C^k$ then $f'(x)^{-1} \in C^{k-1}$ when $\|x - x_0\| \le \delta$ (see Exercise 1.6), and since $g'(y) = f'(g(y))^{-1}$ we conclude by induction that $g \in C^k$ if $f \in C^k$.

If $f'(x_0) : U \to V$ is surjective and U is finite dimensional, then V is finite dimensional. Whenever $f'(x_0)$ is surjective and V is finite dimensional a right inverse A of $f'(x_0)$ is obtained by taking a basis $v_1, \dots, v_n$ in V and defining $A \sum t_j v_j = \sum t_j u_j$ where $f'(x_0)u_j = v_j$. Similarly, if $f'(x_0)$ is injective and V is finite dimensional, then U is finite dimensional. Whenever U is finite dimensional and $f'(x_0)$ is injective then a left inverse of $f'(x_0)$ is obtained by composing the inverse of $f'(x_0) : U \to f'(x_0)U$ with a projection of V on the finite dimensional subspace $f'(x_0)U$. This completes the proof.

There are other iteration methods than (2.2) which are more advantageous from a numerical as well as a theoretical point of view, such as Newton's method

$$y - f(x_{k-1}) = f'(x_{k-1})(x_k - x_{k-1}). \tag{2.3}$$

Note that this method assumes differentiability at every point near x_0 whereas Picard's method used in the proof of Theorem 2.1 only assumed that $x \mapsto f(x) - x$ is a contraction in a neighborhood of x_0. If $f \in C^2$ in a neighborhood of x_0 and x_1, x_2 are sufficiently close to x_0, then

$$\begin{aligned}&\|f(x_1) - f(x_2) - f'(x_2)(x_1 - x_2)\| \\ &\qquad \le \|x_1 - x_2\| \sup_{0 \le t \le 1} \|f'(x_2 + t(x_1 - x_2)) - f'(x_2)\| \le C\|x_1 - x_2\|^2,\end{aligned}$$

which implies that

$$\|y - f(x_k)\| \le C\|x_k - x_{k-1}\|^2.$$

Hence (2.3) with k replaced by $k+1$ gives

$$\|x_{k+1} - x_k\| \le C'\|x_k - x_{k-1}\|^2.$$

If y is sufficiently close to y_0 we obtain a sequence x_k converging very rapidly to a solution of the equation $f(x) = y$.

Note that the necessity in Theorem 2.1 only refers to the existence of a continuously differentiable inverse. The function $f(x) = x^3$, for example, is bijective on $\mathbf{R}$ with inverse $g(y) = y^{\frac{1}{3}}$ although $f'(0) = 0$. Of course, g is not differentiable at the origin. We shall later on return to the question when the hypotheses are fulfilled and what conclusions are otherwise possible. However, we shall first discuss some examples starting with finite dimensional problems.

Definition 2.2. Let k be an integer ≥ 1. A subset F of $\mathbf{R}^n$ is called a C^k manifold of dimension ν if for every $x_0 \in F$ there is a C^k map κ, called a local parametrisation, from an open neighborhood $\omega \subset \mathbf{R}^\nu$ of a point $t_0 \in \mathbf{R}^\nu$ where $\kappa'(t_0)$ is injective, such that every neighborhood $\subset \omega$ of t_0 is mapped on a neighborhood of x_0 in F.

According to b) in Theorem 2.1 we can find a map $g \in C^k$ from a neighborhood of x_0 in $\mathbf{R}^n$ to a neighborhood of t_0 in $\mathbf{R}^\nu$ such that $g \circ \kappa = \mathrm{Id}$ in a neighborhood of t_0. Hence there is an open neighborhood $\omega_1 \subset \omega$ of t_0 where κ is injective. If κ_1 is another local parametrisation, defined in a neighborhood of $s_0 \in \mathbf{R}^\nu$ with $\kappa_1(s_0) = x_0$, we can therefore write $\kappa_1 = \kappa \circ \psi$ where ψ is continuous at s_0 and $\psi(s_0) = t_0$. Composition with g gives $\psi = g \circ \kappa_1 \in C^k$. In the same way we find that $\psi^{-1} \in C^k$, so ψ is a local diffeomorphism. Conversely, if ψ is a local diffeomorphism from a neighborhood of s_0 to a neighborhood of t_0, then $\kappa \circ \psi$ is a local parametrisation of F; thus local parametrisations can only differ by a local diffeomorphism. Since the equation $\kappa \circ \psi = \kappa_1$ implies $\kappa_1'(s_0) = \kappa'(\psi(s_0))\psi'(s_0)$ and $\psi'(s_0)$ is bijective, it follows that the vector spaces $\kappa_1'(s_0)\mathbf{R}^\nu$ and $\kappa'(t_0)\mathbf{R}^\nu$ are equal. This subspace of $\mathbf{R}^n$ of dimension ν is called the *tangent plane* of F at $x_0 = \kappa(t_0)$.

If h is a C^k map from a neighborhood of x_0 with values in $\mathbf{R}^\nu$ and $h'(x_0)\kappa'(t_0)$ is bijective, then $\psi = h \circ \kappa$ is a C^k local diffeomorphism and $\tilde{\kappa} = \kappa \circ \psi^{-1}$ becomes a local parametrisation with $h \circ \tilde{\kappa} = \mathrm{Id}$. In particular, we can always choose h as the projection

$$\mathbf{R}^n \ni (x_1, \dots, x_n) \mapsto (x_1, \dots, x_\nu) \in \mathbf{R}^\nu$$

if the coordinates have been labelled so that $\det(\partial\kappa_j/\partial t_i)_{j,i=1,\dots,\nu} \neq 0$ at t_0. Then $\tilde{\kappa}$ represents F in a neighborhood of x_0 as

$$x_j = \varphi_j(x_1, \dots, x_\nu), \quad j = \nu + 1, \dots, n,$$

where $\varphi_j \in C^k$. Thus F can be defined in a neighborhood of x_0 as the inverse image of 0 under the map

$$\mathbf{R}^n \ni (x_1, \dots, x_n) \mapsto (x_{\nu+1} - \varphi_{\nu+1}(x_1, \dots, x_\nu), \dots, x_n - \varphi_n(x_1, \dots, x_\nu)) \in \mathbf{R}^{n-\nu},$$

which has surjective differential at x_0. Conversely, we have:

Theorem 2.3. *Let f be a C^k function, $k \geq 1$, defined in a neighborhood of $x_0 \in \mathbf{R}^n$ with values in $\mathbf{R}^\mu$ such that $f'(x_0)$ is surjective. Then $\{x \in \Omega; f(x) = f(x_0)\}$ is a C^k manifold of dimension $n - \mu$ if Ω is a sufficiently small open neighborhood of x_0.*

Proof. Set $\nu = n - \mu$ and choose a C^k map $\psi : \mathbf{R}^n \to \mathbf{R}^\nu$, for example the projection on a suitable coordinate plane, such that the linear map $f'(x_0) \oplus \psi'(x_0) : \mathbf{R}^n \to \mathbf{R}^n$ is bijective. From c) in Theorem 2.1 it follows that the map $\Omega \ni x \mapsto (f(x), \psi(x)) \in \mathbf{R}^n$ has a C^k inverse F, defined in a neighborhood of $(f(x_0), \psi(x_0))$ if Ω is small enough. Then $t \mapsto F(f(x_0), t)$ gives a local parametrisation.

Remark. The proof also shows that if $f_1, \dots, f_\mu \in C^k$ have linearly independent differentials at x_0, then we can choose $f_{\mu+1}, \dots, f_n \in C^k$ so that $f = (f_1, \dots, f_n)$ is a local diffeomorphism at x_0.

The hypothesis in Theorem 2.3 that $f'(x_0)$ is surjective implies that the rank of $f'(x)$ is equal to μ for every x in a neighborhood of x_0. Hence the following theorem contains Theorem 2.3.

Theorem 2.4. *Let f be a C^k function from a neighborhood of $x_0 \in \mathbf{R}^n$ to $\mathbf{R}^\mu$ such that the rank of $f'(x)$ is equal to r for every x in a neighborhood of x_0. Then it follows that $M = \{x \in \Omega; f(x) = f(x_0)\}$ is a manifold of dimension $n - r$ if Ω is a sufficiently small neighborhood of x_0, and $f(M)$ is a manifold of dimension r.*

Proof. We can label the coordinates so that $f = (f_1, \dots, f_\mu)$ and $f_1, \dots, f_r$ have linearly independent differentials at x_0. By the remark after Theorem 2.3 we can then choose $g_{r+1}, \dots, g_n \in C^k$ so that

$$x \mapsto (f_1(x), \dots, f_r(x), g_{r+1}(x), \dots, g_n(x))$$

is a local diffeomorphism at x_0. Since the statement of the theorem is invariant under local diffeomorphisms we can compose with the inverse which reduces the proof to the case where

$$f_1(x) = x_1, \dots, f_r(x) = x_r.$$

Since $\partial f(x)/\partial x$ has rank r it follows that $\partial f_k(x)/\partial x_j = 0$ in a ball with center at x_0 if $j > r$, and we conclude that $f_k(x) = f_k(x_1, \dots, x_r)$ there. Hence the equation $f(x) = f(x_0)$ is there equivalent to $x_j = x_{0j}$ for $j = 1, \dots, r$, a linear manifold of codimension r. The range of f is defined by $x_k = f_k(x_1, \dots, x_r)$ when $k = r + 1, \dots, \mu$, which is a manifold of dimension r.

The statement and proof of Theorem 2.3 are very close to the following:

THEOREM 2.5 (THE IMPLICIT FUNCTION THEOREM). *Let f be a C^k function, $k \geq 1$, with values in $\mathbf{R}^p$ defined in a neighborhood of $(x_0, y_0) \in \mathbf{R}^n \oplus \mathbf{R}^m$, such that $f(x_0, y_0) = 0$ and $f'_y(x_0, y_0)$ is surjective. Here f'_y denotes the differential of $y \to f(x_0, y)$. Then there is a C^k function φ from a neighborhood of x_0 to $\mathbf{R}^m$ such that $\varphi(x_0) = y_0$ and $f(x, \varphi(x)) = 0$. If $f'_y(x_0, y_0)$ is bijective, then φ is uniquely determined and the equation $f(x, y) = 0$ is equivalent to $y = \varphi(x)$ when (x, y) is sufficiently close to (x_0, y_0).*

PROOF. The map $F(x, y) = (x, f(x, y))$ from a neighborhood of (x_0, y_0) to $\mathbf{R}^n \oplus \mathbf{R}^p$ has surjective (resp. bijective) differential at (x_0, y_0). Hence it follows from Theorem 2.1 that there is a C^k map G from a neighborhood of $(x_0, 0) \in \mathbf{R}^n \oplus \mathbf{R}^p$ to a neighborhood of (x_0, y_0) such that $F \circ G = \mathrm{Id}$ near $(x_0, 0)$. Thus $G(x, z) = (x, g(x, z))$ and $f(x, g(x, z)) = z$, so $\varphi(x) = g(x, 0)$ satisfies the conditions $f(x, \varphi(x)) = 0$ and $\varphi(x_0) = g(x_0, 0) = y_0$. In the bijective case G is also a left inverse, so $g(x, f(x, y)) = y$ for (x, y) close to (x_0, y_0). If $f(x, y) = 0$ this implies that $y = g(x, 0) = \varphi(x)$, which completes the proof.

EXERCISE 2.6. Prove that if $f \in C^k(\Omega, V)$ where $k \geq 1$, Ω is a neighborhood of x_0 in a Banach space U and V is another Banach space, and if condition a) in Theorem 2.1 is fulfilled, then there is a C^k map from a neighborhood of 0 in $\operatorname{Ker} f'(x_0)$ to a neighborhood of x_0 satisfying condition b) in Theorem 2.1 and with range equal to $\{x; f(x) = f(x_0)\}$ in a neighborhood of x_0.

We shall now study maps f at points where the hypotheses of Theorem 2.3 are not fulfilled.

DEFINITION 2.7. If $f \in C^1(\Omega, \mathbf{R}^m)$ where Ω is a neighborhood of $x \in \mathbf{R}^n$, then x is called a *critical point* and $f(x)$ is called a *critical value* if $f'(x)$ is not surjective. If y is not a critical value then y is called a *regular value.*

The following theorem describes the structure of the simplest kind of critical point for a scalar valued function:

THEOREM 2.8 (MORSE). *Let $f \in C^k(\Omega, \mathbf{R})$, $k > 2$, where Ω is a neighborhood of $x_0 \in \mathbf{R}^n$, and assume that $f'(x_0) = 0$ but $\det f''(x_0) \neq 0$. Then there is a local C^{k-2} diffeomorphism of a neighborhood of $0 \in \mathbf{R}^n$ to a neighborhood of x_0 such that $\psi(0) = x_0$, $\psi'(0) = \mathrm{Id}$ and*

$$f \circ \psi(t) = f(x_0) + \tfrac{1}{2} f''(x_0)(t, t).$$

PROOF. By Taylor's formula we have

$$f(x_0 + x) = f(x_0) + \int_0^1 (1 - s) f''(x_0 + sx)(x, x)\, ds = f(x_0) + \tfrac{1}{2}\langle B(x)x, x\rangle,$$

where B is the symmetric matrix

$$B(x) = 2\int_0^1 (1 - s) f''(x_0 + sx)\, ds \in C^{k-2}.$$

Set $\psi(t) = x_0 + R(t)t$ where R is a matrix to be determined so that $R(0) = \mathrm{Id}$, $R \in C^{k-2}$, and

$$R^* B(Rt) R = B(0) = f''(x_0).$$

For $t = 0$ and $R = \mathrm{Id}$, the differential of the map $R \mapsto R^*B(Rt)R$ is equal to

$$R \mapsto R^*B(0) + B(0)R.$$

This is a surjective map from the linear space of $n\times n$ matrices to the linear space of symmetric $n\times n$ matrices, for if C is a symmetric matrix we have $R^*B(0)+B(0)R = C$ if $R = \frac{1}{2}B(0)^{-1}C$. The existence of the function $R(t)$ is now a consequence of the implicit function theorem applied to the function $F(t,R) = R^*B(Rt)R - B(0)$ with values in the space of symmetric $n \times n$ matrices. The proof is complete.

EXERCISE 2.9. Prove that if $f \in C^k(\Omega, \mathbf{R})$, $k > 2$, where Ω is a neighborhood of x_0 in $\mathbf{R}^n$ and $f'(x_0) = 0$, then there is a C^{k-2} diffeomorphism ψ from a neighborhood of 0 to x_0 such that

$$f \circ \psi(t) = A(t') + g(t'')$$

where A is a non-degenerate quadratic form in $t' = (t_1, \dots, t_j)$ and g is a C^k function of $t'' = (t_{j+1}, \dots, t_n)$ with $g'(0) = g''(0) = 0$.

We shall now give a result on existence of solutions of a system of nonlinear equations where the implicit function theorem is not immediately applicable since second derivatives play an important role.

THEOREM 2.10. *Let $f \in C^k(\Omega, \mathbf{R}^m)$, where $k > 2$ and Ω is a neighborhood of x_0 in $\mathbf{R}^n$. For the existence of a function $x(t)$ from $\mathbf{R}$ to $\mathbf{R}^n$ with $x(0) = x_0$, $x'(0) = X$ and $f(x(t)) = 0$ it is then necessary that*

$$f(x_0) = 0, \quad f'(x_0)X = 0, \quad f''(x_0)(X,X) \in \operatorname{Im} f'(x_0), \tag{2.4}$$

and sufficient that in addition

$$\operatorname{Ker} f'(x_0) \ni Y \mapsto qf''(x_0)(X,Y) \in \operatorname{Coker} f'(x_0) \tag{2.5}$$

is surjective, if q is the natural map from $\mathbf{R}^m$ to Coker $f'(x_0)$.

PROOF. The differential of $q\circ f$ at x_0 is equal to 0, so the comments at the end of Chapter I prove that the conditions (2.4), (2.5) are invariant under composition with a diffeomorphism in $\mathbf{R}^n$. The necessity of (2.4) follows at once by differentiation of the equation $f(x(t)) = 0$. To prove the sufficiency of (2.4), (2.5) we can by composition with a diffeomorphism in $\mathbf{R}^n$ and a linear transformation in $\mathbf{R}^m$ achieve that with a splitting of the coordinates $x = (y, z)$ in $\mathbf{R}^n$ we have $x_0 = (0, 0)$ and $f(x) = (y, g(y, z))$ where $g'(0, 0) = 0$. Then the condition (2.4) means that

$$g(0,0) = 0, \quad X = (0, Z), \quad g''_{zz}(0,0)(Z,Z) = 0,$$

and (2.5) means that $\widetilde{Z} \mapsto g''_{zz}(0,0)(Z,\widetilde{Z})$ is surjective. The equation $f(x(t)) = 0$ means that $x(t) = (0, z(t))$ and that $g(0, z(t)) = 0$. To solve the equation $g(0, z(t)) = 0$ we set $z(t) = tw(t)$ Then the condition $x'(0) = X$ becomes $w(0) = Z$. We have

$$g(0, tw)/t^2 = h(w,t)$$

where $h \in C^{k-2}$ by Taylor's formula (see the proof of Theorem 2.8), and

$$h(w,0) = \tfrac{1}{2}g''_{zz}(0,0)(w,w).$$

The differential at Z is $w \mapsto g''_{zz}(0,0)(Z,w)$ which is surjective. For the equation $h(w,t) = 0$ the hypotheses of the implicit function theorem are thus fulfilled at $(Z,0)$, which completes the proof.

REMARK. By (2.4) $X \in \operatorname{Ker} f'(x_0)$, and in (2.5) X is mapped to 0 in $\operatorname{Coker} f'(x_0)$, by the last condition (2.4). If $\nu = \dim \operatorname{Ker} f'(x_0)$, thus $m-(n-\nu) = \dim \operatorname{Coker} f'(x_0)$, it follows that (2.5) requires that $m-(n-\nu) < \nu$, that is, $m < n$. When examining (2.5) one can also restrict Y to a hyperplane transversal to X. This observation is useful in the following exercise.

EXERCISE 2.11. Let F be a C^∞ function defined in a neighborhood of $(x_0, y_0) \in \mathbf{R} \oplus \mathbf{R}^m$. Assume that $F(x_0,y_0) = 0$, that $Z \in \mathbf{R}^m$ satisfies the conditions $F'_x(x_0,y_0) + F'_y(x_0,y_0)Z = 0$ and

$$F''_{xx}(x_0,y_0) + 2F''_{xy}(x_0,y_0)Z + F''_{yy}(x_0,y_0)(Z,Z) \in \operatorname{Im} F'_y(x_0,y_0),$$

and that $F''_{xy}(x_0,y_0) + F''_{yy}(x_0,y_0)Z$ induces a bijection $\operatorname{Ker} F'_y(x_0,y_0) \to \operatorname{Coker} F'_y(x_0,y_0)$. Prove that the equation $F(x,y) = 0$ has a C^∞ solution $y(x)$ with $y(x_0) = y_0$ and $y'(x_0) = Z$.

EXERCISE 2.12. Let $f \in C^k(\Omega, \mathbf{R}^m)$ where $k > 2$ and Ω is a neighborhood of $x_0 \in \mathbf{R}^{m+1}$. Assume that $f(x_0) = 0$, that $\dim \operatorname{Ker} f'(x_0) = 2$, and that the map

$$\operatorname{Ker} f'(x_0) \ni X \mapsto qf''(x_0)(X,X),$$

where q is the natural map $\mathbf{R}^m \to \operatorname{Coker} f'(x_0)$, is an indefinite quadratic form. (Note that $\dim \operatorname{Coker} f'(x_0) = m - \operatorname{rank} f'(x_0) = m - (m+1-2) = 1$.) Then there are precisely two C^{k-2} curves $t \mapsto x_j(t)$ with $x_j(0) = x_0$ and $f(x_j(t)) = 0$ which together give all solutions of the equation $f(x) = 0$ in a neighborhood of x_0. (Bifurcations of the solution x_0.)

EXAMPLE 2.13. Consider the nonlinear eigenvalue problem to find $u \in C^2([0,\pi])$ with

$$u'' + \lambda u + K(\lambda, u, u') = 0, \quad u(0) = u(\pi) = 0,$$

where $K \in C^3$ in a neighborhood of $\mathbf{R} \times \{0\} \times \{0\}$ and $K(\lambda,\xi,\eta) = O(\xi^2+\eta^2)$ when $(\xi,\eta) \to 0$. The Cauchy problem with initial data $u(0) = 0$, $u'(0) = h$ has for sufficiently small h a unique solution $u(x,\lambda,h)$ in C^3, and if we set $f(\lambda,h) = u(\pi,\lambda,h)$ the eigenvalue problem is equivalent to $f(\lambda,h) = 0$. By the uniqueness theorem for the Cauchy problem we have $f(\lambda,0) \equiv 0$, and $\partial f(\lambda,0)/\partial h = v(\pi,\lambda)$ where $v(x,\lambda) = \partial u(x,\lambda,0)/\partial h$ satisfies

$$v'' + \lambda v = 0, \quad v(0) = 0, \ v'(0) = 1.$$

Hence $v(x,\lambda) = (\sin\sqrt{\lambda}x)/\sqrt{\lambda}$ and $\partial f(\lambda,0)/\partial h = \sin(\sqrt{\lambda}\pi)/\sqrt{\lambda} = 0$ if and only if $\lambda = n^2$ where n is a positive integer. In that case $\partial^2 f(\lambda,0)/\partial\lambda^2 = 0$ and $\partial^2 f(\lambda,0)/\partial h\partial\lambda = \pi(-1)^n/2\lambda \neq 0$, so the hypotheses in Exercise 2.12 are fulfilled. Near $h = 0$ the solutions of the equation $f(\lambda,h) = 0$ therefore consist of the line $h = 0$ together with a C^1 curve intersecting the line transversally at $(n^2,0)$, for every positive integer n. Thus we have determined all solutions close to the zero solution.

As a simple example of the infinite dimensional case of Theorem 2.1 we shall now prove the existence theorem for ordinary differential equations assuming continuous differentiability instead of just a Lipschitz condition.

THEOREM 2.14. *Let $A(x,y)$ be a function with values in $\mathbf{R}^m$ defined in a neighborhood of $(x_0,y_0)\in\mathbf{R}\oplus\mathbf{R}^m$, such that $A(x,y)$ is differentiable with respect to y for fixed x and A and A'_y are continuous. For y sufficiently close to y_0 the Cauchy problem*

$$df(x)/dx = A(x,f(x)),\quad f(x_0)=y, \tag{2.6}$$

has one and only one solution in a neighborhood of x_0.

PROOF. If we make the change of variables $x=x_0+\varepsilon t$ and set $f(x)=g(t)$, the equations become

$$dg(t)/dt=\varepsilon A(x_0+\varepsilon t, g(t)),\quad g(0)=y.$$

We shall prove that for sufficiently small ε there is a unique solution $g\in C^1(I,\mathbf{R}^m)$ where $I=[-1,1]$. To do so we consider the map

$$C^1(I,\mathbf{R}^m)\oplus\mathbf{R}\ni(g,\varepsilon)\mapsto(dg/dt-\varepsilon A(x_0+\varepsilon\cdot,g),g(0),\varepsilon)\in C^0(I,\mathbf{R}^m)\oplus\mathbf{R}^m\oplus\mathbf{R}.$$

It maps $(g,\varepsilon)=(y_0,0)$ to $(0,y_0,0)$, and the differential at $(y_0,0)$ is (see Exercise 1.5)

$$(g,\varepsilon)\mapsto(dg/dt-\varepsilon A(x_0,y_0),g(0),\varepsilon).$$

It is bijective, for $dg/dt-\varepsilon A(x_0,y_0)=h\in C^0$ means that

$$g(t)=g(0)+\varepsilon tA(x_0,y_0)+\int_0^t h(s)\,ds\in C^1.$$

Hence the map has a C^1 inverse near $(0,y_0,0)$, and it maps $(0,y,\varepsilon)$ to the unique solution of the Cauchy problem.

From Theorem 2.14 it follows that (2.6) has a solution f with $|f(x)-y_0|\le M|x-x_0|$, $|x-x_0|\le a$, provided that $A(x,y)$ and $\partial A(x,y)/\partial y$ are continuous when $|x-x_0|\le a$ and $|y-y_0|\le b$, that $|A(x,y)|\le M$, and that $b\ge Ma$. In fact, (2.6) implies that $|f'(x)|\le M$ so $|f(x)-y_0|\le M|x-x_0|$ for $|x-x_0|\le\alpha$ if α is so small that the solution exists then. Let a_0 be the supremum of such α. Then f is continuous when $|x-x_0|\le a_0$, hence f is continuously differentiable and satisfies (2.6) then. If $a_0<a$ then Theorem 2.14 proves that a_0 is not maximal, so $a_0=a$. — The differentiability hypothesis in Theorem 2.14 can be removed without loss of the existence statement:

THEOREM 2.15 (PEANO). *If A is continuous and $|A|\le M$ in $R=\{(x,y);|x-x_0|\le a,\ |y-y_0|\le b\}$, and $b\ge Ma$, then the equation* (2.6) *has a solution when $|x-x_0|\le a$ with $|f(x)-y_0|\le M|x-x_0|$.*

PROOF. Decreasing a slightly we may assume that $b>Ma$. The regularisation

$$A_\varepsilon(x,y)=\int_0^1 A(x,y+\varepsilon t)\,dt=\int_0^\varepsilon A(x,y+t)\,dt/\varepsilon,$$
$$|x-x_0|\le a,\ |y-y_0|\le b-\varepsilon,\quad \varepsilon>0,$$

has a continuous derivative $\partial A_\varepsilon(x,y)/\partial y=A(x,y+\varepsilon)/\varepsilon-A_\varepsilon(x,y)/\varepsilon$ with respect to y. When ε is so small that $Ma\le b-\varepsilon$, we conclude that there exists a solution f_ε of the equation $f'_\varepsilon(x)=A_\varepsilon(x,f_\varepsilon(x))$ when $|x-x_0|\le a$, with $|f_\varepsilon(x)-y_0|\le M|x-x_0|$. Since $|f_\varepsilon(x)-f_\varepsilon(x')|\le M|x-x'|$ we can choose a sequence $\varepsilon_j\to0$ such that f_{ε_j} converges uniformly to a continuous function $f(x)$. Since $A_\varepsilon\to A$ uniformly it follows that $f'_{\varepsilon_j}(x)\to A(x,f(x))$ uniformly, so $f\in C^1$ and (2.6) holds.

Exercise 2.16. Prove that the solution in Theorem 2.15 is unique if A is Lipschitz continuous with respect to y.

Notes. Theorem 2.10 can be found in slightly less generality in [F], where also applications to the existence of periodic solutions of differential equations can be found. Exercise 2.12 can be extended to the infinite dimensional case; the result is called bifurcation theory. (See e.g. [CR].) The rapid convergence of the Newton method (2.3) can be used to give important improvements of the implicit function theorem where the hypotheses on the inverse of f' are weakened (see [M], [S] and [H]). One can also consult [S] for the material in the following two chapters.

References

[CR] M. Crandall and P. Rabinowitz, *Bifurcation, perturbation of simple eigenvalues and linearized stability*, Arch. Rat. Mech. Anal. **52** (1973), 161–180.

[F] K. Friedrichs, *Lectures on advanced ordinary differential equations*, Gordon and Breach, 1965.

[H] R. Hamilton, *The inverse function theorem of Nash and Moser*, Bull. Amer. Math. Soc. **7** (1982), 65–222.

[M] J. Moser, *A rapidly convergent iteration method and non-linear differential equations I–II*, Ann. Sc. Norm. Pisa **201** (1966), 265–315, 499–535.

[S] J. T. Schwartz, *Non-linear functional analysis*, Gordon and Breach, 1969.

CHAPTER III

THE MORSE-SARD THEOREM

The usefulness of the implicit function theorem is due to the fact that the hypotheses it makes on the differential are "usually" satisfied:

THEOREM 3.1 (MORSE-SARD). *Let $f \in C^\infty(\Omega, \mathbf{R}^m)$ where Ω is an open set in $\mathbf{R}^n$. Then the set of critical values assumed by f in Ω has Lebesgue measure* 0 *in $\mathbf{R}^m$.*

REMARK 1. A somewhat more complicated proof shows that it suffices to assume that $f \in C^k(\Omega, \mathbf{R}^m)$ where $k > \max(0, n-m)$, and an example constructed in [W] proves that this smoothness assumption cannot be weakened.

REMARK 2. If K is a compact subset of Ω, then the critical points in K form a compact set. It is mapped by f on a compact set, with Lebesgue measure 0 by Theorem 3.1. Thus it has even Jordan measure 0. The critical set is a countable union of such compact sets of measure 0, hence of the first category. However, it is convenient in the proof to use Lebesgue measure.

PROOF OF THEOREM 3.1. The proof is by induction so we assume that the theorem has already been proved with n replaced by $n-1$ if $n > 1$. We set

$$C_j = \{x \in \Omega; f'(x) = 0, \dots, f^{(j)}(x) = 0\}, \quad j = 1, 2, \dots,$$

and prove first that

$$m(f(C_j)) = 0 \quad \text{if} \quad (j+1)m > n. \tag{3.1}$$

It is sufficient to prove that $m(f(K \cap C_j)) = 0$ when K is a compact cube contained in Ω, with side A, say. If we divide the edges of the cube in ν equal parts we obtain ν^n cubes with side $\varepsilon = A/\nu$. Let $I_1, \dots, I_N$ denote the cubes so obtained which intersect C_j, and choose $x_k \in I_k \cap C_j$. By Taylor's formula and the definition of C_j

$$\|f(x + x_k) - f(x_k)\| \le B\|x\|^{j+1} \le B\varepsilon^{j+1}, \quad \text{if } x + x_k \in I_k,$$

and when $(j+1)m > n$ this implies that

$$m(f(I_k)) \le B^m \varepsilon^{(j+1)m} \le C\varepsilon^{n+1} = Cm(I_k)\varepsilon.$$

Hence

$$m(f(K \cap C_j)) \le \sum_1^N m(f(I_k)) \le C \sum_1^N m(I_k)\varepsilon \le CA^n\varepsilon,$$

which proves that $m(f(K \cap C_j)) = 0$.

The next step is to note that $E_k = C_k \setminus C_{k+1}$ is contained in a manifold S of dimension $n-1$ in a neighborhood V of any point in $x_0 \in E_k$, for there is some component g of $f^{(k)}$ with $g'(x_0) \neq 0$. If f has a critical point $x \in S$ then x is also a critical point of the restriction of f to S, which has therefore at least as many critical values. If the theorem is known with n replaced by $n-1$ it follows that

$$m(f(C_k \setminus C_{k+1})) = 0, \tag{3.2}$$

for we can cover E_k with countably many neighborhoods V. When $n = 1$ then S is discrete and the same conclusion is obvious.

It remains to prove that $m(f(C \setminus C_1)) = 0$ if C is the set of all critical points. By the chain rule the critical set is not changed if we replace f by $f \circ \psi$ where ψ is a local diffeomorphism; as above we can work locally. At a point in $C \setminus C_1$ where $\partial f_1/\partial x_1 \neq 0$, for example, we can choose ψ as the inverse of the map $x \mapsto (f_1(x), x_2, \dots, x_n)$ and obtain that $f \circ \psi(t) = (t_1, g(t))$ where g takes values in $\mathbf{R}^{m-1}$. If $n = 1$ and $m = 1$ then $f \circ \psi(t) = t_1$ has no critical point, and when $n = 1$ and $m > 1$ then the range of $f \circ \psi(t) = (t_1, g(t))$ is a curve, of measure 0 in $\mathbf{R}^m$. We may therefore assume that $n > 1$. It is then clear that $f \circ \psi$ has a critical point at $t = (t_1, t')$ if and only if $t' \mapsto g(t_1, t') \in \mathbf{R}^{m-1}$ has a critical point. For fixed t_1 the critical values of $f \circ \psi$ are therefore a set of measure 0 in the plane $\{t_1\} \times \mathbf{R}^{m-1}$. The set of critical values is a countable union of compact sets, hence Lebesgue measurable, so it follows from the Lebesgue-Fubini theorem that $m(f((C \setminus C_1) \cap V)) = 0$ for some neighborhood V of an arbitrary point in $C \setminus C_1$. The proof is complete.

Exercise 3.2. Prove that if $f \in C^\infty(\Omega, \mathbf{R})$ where Ω is an open subset of $\mathbf{R}^n$, then the critical points of $f_h(x) = f(x) - \langle x, h\rangle$ are non-degenerate for almost all $h \in \mathbf{R}^n$.

Notes. The original proofs of the Morse-Sard theorem, with minimal smoothness assumptions, can be found in [M] and [S]. A somewhat different proof is given in [St].

References

[M] A. P. Morse, *The behavior of a function on its critical set*, Ann. of Math. **40** (1939), 62–70.

[S] A. Sard, *The measure of the critical values of differentiable maps*, Bull. Amer. Math. Soc. **48** (1942), 883–890.

[St] S. Sternberg, *Lectures on differential geometry*, Prentice-Hall, 1964.

[W] H. Whitney, *A function not constant on a connected set of critical points*, Duke Math. J. **1** (1935), 514–517.

CHAPTER IV

THE DEGREE OF MAPPING

In this chapter we shall give sufficient conditions for solvability of an equation $f(x) = y$ where f is a continuous map between two Banach spaces. The method depends on approximation by infinitely differentiable functions between finite dimensional spaces where the Morse-Sard theorem will allow us to apply the implicit function theorem.

Let $f \in C^\infty(\mathbf{R}^n, \mathbf{R}^n)$ and let Ω be an *open, bounded* subset of $\mathbf{R}^n$, with closure $\overline{\Omega}$ and boundary $\partial\Omega$. By the Morse-Sard theorem

$$C = \{f(x); x \in \overline{\Omega}, \det f'(x) = 0\}$$

is a closed null set. If $y \notin C \cup f(\partial\Omega)$, then the equation $f(x) = y$ can only have a finite number of solutions $x \in \Omega$, and $\det f'(x) \neq 0$ for each of them. In fact, by the inverse function theorem the solutions in $\overline{\Omega}$ are isolated and there are no solutions in $\partial\Omega$ by assumption. We can therefore count the number of solutions with sign

$$d(f, \Omega, y) = \sum_{f(x)=y, x\in\Omega} \operatorname{sign} \det f'(x), \quad y \notin C \cup f(\partial\Omega). \tag{4.1}$$

The implicit function theorem implies also that $d(f, \Omega, y)$ is constant in a neighborhood of any point $\notin C \cup f(\partial\Omega)$, hence it is constant in any component of the complement of $C \cup f(\partial\Omega)$. To be able to extend the definition to more general f and y we must prove that this number is stable under perturbations of f and y. The following lemma is the most important step:

LEMMA 4.1. *Let $F \in C^\infty(\mathbf{R}^n \times I, \mathbf{R}^n)$ where I is an open interval $\subset \mathbf{R}$ containing $[0, 1]$, and set $f_j(x) = F(x, j)$, $j = 0, 1$. If $y \notin F(\partial\Omega \times [0, 1])$ and y is a regular value for f_0 and for f_1, then*

$$d(f_0, \Omega, y) = d(f_1, \Omega, y). \tag{4.2}$$

PROOF. Without changing either side of (4.2) we can replace y by a point nearby which is not in $F(\partial\Omega \times [0, 1])$ and is not a critical value for F in $\overline{\Omega} \times [0, 1]$. Then

$$E = \{(x, t); 0 \le t \le 1, x \in \overline{\Omega}, F(x, t) = y\}$$

is a C^∞ manifold of dimension 1 which is a subset of $\Omega \times [0, 1]$. If s is a local parameter on E then

$$\frac{\partial F_j}{\partial t}\frac{dt}{ds} + \sum_{k=1}^{n} \frac{\partial F_j}{\partial x_k}\frac{dx_k}{ds} = 0, \quad j = 1, \dots, n,$$

so Cramer's rule gives that $(dt/ds, dx/ds) = c(s)G(s)$ where $c(s) \neq 0$ and

$$G_0 = \det F'_x,$$
$$G_j = -\det(\partial F/\partial x_1, \dots, \partial F/\partial x_{j-1}, \partial F/\partial t, \partial F/\partial x_{j+1}, \dots, \partial F/\partial x_n),\ j = 1, \dots, n.$$

Replacing s by a primitive function of $c(s)$ we can make the constant factor equal to 1, which determines the parameter up to an additive constant. It is now easy to see that every component of E is either

a) a closed curve on which $0 < t < 1$; or
b) an arc which begins when $t = 0$ (resp. $t = 1$) and ends when $t = 1$ (resp. $t = 0$); or
c) an arc which begins and ends when $t = 0$ (resp. $t = 1$).

In fact, E is the union of finitely many arcs which do not contain more than one point where $t = 0$ or $t = 1$ since y is not a critical value for f_0 or f_1. On two overlapping arcs one can make the parameters agree and then join them. After a finite number of such steps one obtains either a closed curve (case a)) or an arc with end points where $t = 0$ or $t = 1$, that is, cases b) or c). The details are left for the reader. (See also [M]; another more analytical proof of Lemma 4.1 will be given in Chapter VIII.) Now recall that $dt/ds = \det F'_x$ on the curves. In case b) it follows that $\det F'_x$ has the same sign at both end points of the arc, and in case c) it follows that the signs are opposite. In the sums defining $d(f_0, \Omega, y)$ resp. $d(f_1, \Omega, y)$ the contributions from points joined by arcs of the type c) must therefore cancel, and the remaining terms are then equal when paired by the arcs of type b). This proves the lemma.

Example. Let $\Omega = (-2, 2)$ and $F(x, t) = x^2 + 1 - 2t$, $y = 0$. The equation $F(x, t) = 0$ defines a parabola connecting the solutions $x = \pm 1$ of the equation $F(x, 1) = 0$, and it is clear that these have opposite signs.

We can now prove that $d(f, \Omega, y)$ has the same value for all regular points y in the same component of $\complement f(\partial\Omega)$. To do so we choose $\varepsilon > 0$ smaller than the distance from y to $f(\partial\Omega)$ and let y' be a regular value of f with $|y - y'| < \varepsilon$. Then the hypotheses of Lemma 4.1 are satisfied by $F(x, t) = f(x) - t(y' - y)$, for $F(x, 1) = y$ means precisely that $f(x) = y'$, which by hypothesis is a regular value for $f = F(\cdot, 1)$. Hence $d(f, \Omega, y) = d(f, \Omega, y')$. We can join any two regular points in the same component of $\complement f(\partial\Omega)$ by a polygon such that the vertices are regular values of f and the edges are smaller than the distance from the polygon to $f(\partial\Omega)$. In fact, if we choose a polygon with these properties apart from the regularity of the vertices it follows from the Morse-Sard theorem that we can replace the vertices by regular values which are arbitrarily close.

Thus we can extend the definition of $d(f, \Omega, y)$ uniquely to all $y \in \complement f(\partial\Omega)$ so that the value is constant in each component. We note the following important properties of d:

(i) $d(f, \Omega, y)$ is defined in $\complement f(\partial\Omega)$ and is equal to an integer $d(f, \Omega, U)$ in every component U of $\complement f(\partial\Omega)$.
(ii) $d(f, \Omega, y) \neq 0$ implies that $y = f(x)$ for some $x \in \Omega$.
(iii) If $F \in C^\infty(\mathbf{R}^n \times I, \mathbf{R}^n)$ for a neighborhood of $I = [0, 1]$ and $y \notin F(\partial\Omega \times [0, 1])$ then $d(F(\cdot, t), \Omega, y)$ does not depend on $t \in [0, 1]$.

Using the stability property (iii) we shall now weaken the hypothesis that $f \in C^\infty$ to just continuity. More precisely, we now assume that $f \in C(\overline{\Omega}, \mathbf{R}^n)$ where Ω as before is open and bounded. For a given $\varepsilon > 0$ we can choose a finite open covering $O_1, \dots, O_N$ of $\overline{\Omega}$ such that $\|f(x) - f(x')\| < \varepsilon$ if $x, x' \in O_i \cap \overline{\Omega}$ for some i. Let $\varphi_1, \dots, \varphi_N$ be a corresponding partition of unity, that is, $\varphi_j \in C_0^\infty(O_j)$, $\varphi_j \ge 0$, $\sum_1^N \varphi_j \le 1$ with equality in $\overline{\Omega}$. Choose $x_j \in \overline{\Omega} \cap O_j$ and set $f_1 = \sum_1^N \varphi_j f(x_j)$. Then $f_1 \in C^\infty(\mathbf{R}^n, \mathbf{R}^n)$ and

$$\|f(x) - f_1(x)\| < \varepsilon, \quad x \in \overline{\Omega}, \tag{4.3}$$

for $f(x) - f_1(x) = \sum_1^N \varphi_j(x)(f(x) - f(x_j))$. If the distance from y to $f(\partial\Omega)$ is larger than ε then $y \notin f_1(\partial\Omega)$, and we set

$$D(f, \Omega, y) = d(f_1, \Omega, y). \tag{4.4}$$

This definition is independent of the choice of f_1, for if (4.3) is fulfilled with f_1 replaced by $f_0 \in C^\infty(\mathbf{R}^n, \mathbf{R}^n)$, then (4.3) is satisfied by $tf_1 + (1-t)f_0$ when $0 \le t \le 1$, so it follows from (iii) that

$$d(f_0, \Omega, y) = d(f_1, \Omega, y).$$

When $\varepsilon \to 0$ it follows that the properties (i)–(iii) are inherited by D, so we obtain:

Theorem 4.2. *For every $f \in C(\overline{\Omega}, \mathbf{R}^n)$ an integer $D(f, \Omega, y)$, called the degree of f in Ω at y, is uniquely defined when $y \notin f(\partial\Omega)$ by* (4.4) *where $f_1 \in C^\infty(\mathbf{R}^n, \mathbf{R}^n)$ satisfies* (4.3) *with ε smaller than the distance from y to $f(\partial\Omega)$. We have*

(i) *$D(f, \Omega, y)$ is a fixed integer $D(f, \Omega, U)$ in every component U of $\complement f(\partial\Omega)$.*
(ii) *$D(f, \Omega, y) \ne 0$ implies that $y = f(x)$ for some $x \in \Omega$.*
(iii) *If $F \in C(\overline{\Omega} \times [0,1], \mathbf{R}^n)$ and $y \notin F(\partial\Omega \times [0,1])$ then $D(F(\cdot, t), \Omega, y)$ is independent of $t \in [0,1]$.*

Note in particular that (iii) implies that $D(f, \Omega, y) = D(g, \Omega, y)$ if $f = g$ on $\partial\Omega$, for we can then take $F(x,t) = tf(x) + (1-t)g(x)$ which is independent of t when $x \in \partial\Omega$.

Corollary 4.3. *If $f \in C(\overline{\Omega}, \mathbf{R}^n)$ and $f(x) = x$ when $x \in \partial\Omega$, then $\overline{\Omega} \subset f(\overline{\Omega})$. More generally, if $g : \mathbf{R}^n \to \mathbf{R}^n$ is a homeomorphism and $f(x) = g(x)$ when $x \in \partial\Omega$, then $g(\overline{\Omega}) \subset f(\overline{\Omega})$.*

Proof. If $f(x) = x$ when $x \in \partial\Omega$ then $D(f, \Omega, y)$ is defined when $y \notin \partial\Omega$ and is not changed if f is replaced by the identity map, so $D(f, \Omega, y) = 1$ if $y \in \Omega$. Thus the first statement follows from (ii) in Theorem 4.2. To prove the second statement we note that $g^{-1} \circ f : \overline{\Omega} \to \mathbf{R}^n$ is equal to the identity on $\partial\Omega$, so $\overline{\Omega} \subset g^{-1} \circ f(\overline{\Omega})$ by the first statement, hence $g(\overline{\Omega}) \subset f(\overline{\Omega})$.

Corollary 4.4 (The Brouwer fixed point theorem). *If B is a compact convex subset of $\mathbf{R}^n$, then every continuous map $f : B \to B$ has a fixed point, that is, $f(x) = x$ for some $x \in B$.*

Proof. First assume that B is the Euclidean unit ball. If f has no fixed point we define $g(x) = x$ when $x \in \partial B$ and if $x \in B \setminus \partial B$ we let $g(x)$ be the intersection of ∂B with the line through x and $f(x)$ such that x lies between $g(x)$ and $f(x)$. This makes g a continuous map from B to ∂B which leaves every point on ∂B fixed, which contradicts Corollary 4.3 and proves that f must have a fixed point. For a general B we can assume that 0 is an interior point in B and define $\varphi : \mathbf{R}^n \to \mathbf{R}^n$ so that $\varphi(x) = q(x)x$ with $q(tx) = q(x)$ for $t > 0$ and $q(x) = 1/\|x\|$ when $x \in \partial B$, where $\|\cdot\|$ is the Euclidean norm. Then φ is a homeomorphism of B on the Euclidean unit ball, so $\varphi \circ f \circ \varphi^{-1}$ must have a fixed point in the unit ball, which means that f has a fixed point in B.

Note that Brower's fixed point theorem is a natural extension of the elementary fact that a continuous real valued function in an interval $[a, b]$ assumes all values in $[f(a), f(b)]$. This follows from the case where $a = f(a)$ and $b = f(b)$ which is the special case of Corollary 4.4 with $n = 1$. (It is an instructive exercise to determine $D(f, (a, b), y)$ when $y \notin \{a, b\}$.)

Brouwer's fixed point theorem states that the graph of f must intersect the diagonal in $B \times B$. The result remains valid for more general sets than graphs:

Corollary 4.4′ (Kakutani's fixed point theorem. *Let B be a compact convex subset of $\mathbf{R}^n$, and let $F \subset B \times B$ be a compact set such that $F(x) = \{y \in B; (x, y) \in F\}$ is convex and not empty for every $x \in B$. Then there is a point $x \in B$ with $x \in F(x)$, that is, $(x, x) \in F$.*

Proof. We may assume that B has interior points. For any $\varepsilon > 0$ we can choose a partition of unity $\{\varphi_j^\varepsilon\}$, $j = 1, \dots, N_\varepsilon$, in B such that the diameter of $\operatorname{supp} \varphi_j^\varepsilon$ is smaller than ε for every j. Choose $x_j^\varepsilon \in B \cap \operatorname{supp} \varphi_j^\varepsilon$ and $y_j^\varepsilon \in F(x_j^\varepsilon)$, and set

$$f_\varepsilon(x) = \sum_1^{N_\varepsilon} \varphi_j^\varepsilon(x) y_j^\varepsilon.$$

Then f_ε is a continuous map $B \to B$, since B is convex, so by Brouwer's fixed point theorem there is a fixed point $x_\varepsilon \in B$. The proof will be finished if we prove that all limit points of $(x_\varepsilon, x_\varepsilon)$ as $\varepsilon \to 0$ are in F. To do so we choose a point $(x^0, y^0) \in B \times B \setminus F$. Since $F(x^0)$ is convex, there is an open half space $H \subset \mathbf{R}^n$ with $y^0 \in H$ and $\overline{H} \cap F(x^0) = \emptyset$. Since F is closed it follows that $\overline{H} \cap F(x) = \emptyset$ when $\|x - x^0\| < 2\delta$, for some $\delta > 0$. If $\|x - x^0\| < \delta$ and $\varepsilon < \delta$ it follows that $y_j^\varepsilon \in \complement H$ when $\varphi_j(x) \neq 0$, for $\|x_j^\varepsilon - x^0\| < \delta + \varepsilon < 2\delta$; hence $f_\varepsilon(x) \in \complement H$. If $f_\varepsilon(x_\varepsilon) = x_\varepsilon$ and $\|x_\varepsilon - x^0\| < \delta$ it follows that $x_\varepsilon \in \complement H$, so (x^0, y^0) cannot be a limit point of $(x_\varepsilon, x_\varepsilon)$ as $\varepsilon \to 0$. The proof is complete.

After these applications we shall prove further properties of the function D defined in Theorem 4.2.

Theorem 4.5. *If $f \in C(\overline{\Omega}, \mathbf{R}^n)$ and Ω_1, Ω_2, … are disjoint open subsets of Ω, then*

$$D(f, \Omega, y) = \sum_\alpha D(f, \Omega_\alpha, y), \quad \textit{if } y \notin f(\overline{\Omega} \setminus \cup\Omega_\alpha), \tag{iv}$$

where only finitely many terms in the sum are $\neq 0$.

Proof. The terms in the sum are defined, for if $z \in \partial\Omega_\alpha$ then $z \in \overline{\Omega} \setminus \cup\Omega_\beta$, so $y \neq f(z)$ by assumption. If $D(f, \Omega_\alpha, y) \neq 0$ then $y \in f(\Omega_\alpha)$, that is $f(x_\alpha) = y$ for some $x_\alpha \in \Omega_\alpha$. Such points can only exist for finitely many values of α, for in a limit point $x \in \overline{\Omega}$ we would have $f(x) = y$, hence $x \in \Omega_\alpha$ for some α, and Ω_α is then a neighborhood of x. Hence it suffices to prove (iv) for finitely many Ω_α, and then the statement follows at once if we approximate f in $\overline{\Omega}$ by a C^∞ function and replace y by a regular value nearby.

Before proceeding we shall relax the hypothesis that Ω is bounded since we also want to discuss neighborhoods of infinity. From now on we just assume that Ω is open and that $f \in C(\overline{\Omega}, \mathbf{R}^n)$ is proper, that is, that $\|f(x)\| \to \infty$ as $x \to \infty$. This is sufficient to guarantee that $f(\partial\Omega)$ is closed. The properties (i)–(iv) of D remain valid with no change if we define $D(f, \Omega, y)$ when $y \notin f(\partial\Omega)$ as $D(f, \Omega_1, y)$ where the open bounded set $\Omega_1 \subset \Omega$ is chosen so large that $y \notin f(\overline{\Omega} \setminus \Omega_1)$. By Theorem 4.5 this definition is then independent of the choice of Ω_1.

THEOREM 4.6. *Let f and g be proper continuous maps $\mathbf{R}^n \to \mathbf{R}^n$, and let ω_α be the components of $\complement f(\partial\Omega)$. If $y \notin (g \circ f)(\partial\Omega)$ then*

$$D(g \circ f, \Omega, y) = \sum_\alpha D(f, \Omega, \omega_\alpha) D(g, \omega_\alpha, y), \tag{v}$$

where $D(g, \omega_\alpha, y) \neq 0$ only for finitely many α.

PROOF. Since $\partial\omega_\alpha \subset f(\partial\Omega)$ we have $y \notin g(\partial\omega_\alpha)$ so the terms in the right-hand side are defined. If $D(g, \omega_\alpha, y) \neq 0$ then $g(x_\alpha) = y$ for some $x_\alpha \in \omega_\alpha$, and all x_α belong to a compact set since g is proper. A limit point of such points would belong to $f(\partial\Omega)$, and since $g(x) = y$ this would contradict the hypothesis that $y \notin g \circ f(\partial\Omega)$, which proves the last statement. In the same way we see that y has a neighborhood V_y which only intersects $g(\omega_\alpha)$ for finitely many $\alpha_1, \dots, \alpha_k$. Choose now relatively compact open subsets ω'_j of ω_{α_j} so that $V_y \cap g(\omega_{\alpha_j} \setminus \omega'_j) = \emptyset$. The right-hand side of (v) is then equal to

$$\sum_1^k D(f, \Omega, \omega'_j) D(g, \omega'_j, y),$$

by property (iv) of D. Approximating f and g by C^∞ proper maps on a large compact set we conclude that it suffices to prove the theorem for such f and g when y is a regular value for g and for $g \circ f$. Then there are finitely many z_j with $g(z_j) = y$, and we have $z_j \notin \complement f(\partial\Omega)$, $\det g'(z_j) \neq 0$. For every j there are finitely many $x_{jk} \in \Omega$ with $f(x_{jk}) = z_j$, we have $\det f'(x_{jk}) \neq 0$, and

$$\operatorname{sign}\det(g \circ f)'(x_{jk}) = \operatorname{sign}\det g'(z_j) \operatorname{sign}\det f'(x_{jk})$$

by the chain rule. Hence

$$d(g \circ f, \Omega, y) = \sum_j \operatorname{sign}\det g'(z_j) \sum_k \operatorname{sign}\det f'(x_{jk}) = \sum_j \operatorname{sign}\det g'(z_j) d(f, \Omega, z_j).$$

Summation over all j with $z_j \in \omega_\alpha$ yields (v).

The Jordan-Brouwer theorem is an important consequence of Theorem 4.6:

THEOREM 4.7. *Let K and K' be compact subsets of $\mathbf{R}^n$ such that there exists a homeomorphism $f : K \to K'$. Then $\complement K$ and $\complement K'$ have equally many components. In particular there are two components if $K = S^{n-1}$.*

PROOF. We can extend f to a proper map $F \in C(\mathbf{R}^n, \mathbf{R}^n)$, and we can also choose a proper extension G of f^{-1}. Then $G \circ F$ and $F \circ G$ are equal to the identity on K and on K' respectively. Let D_α (resp. D'_α) be the components of $\complement K$ (resp. $\complement K'$). Then the matrices $D(F, D_\alpha, D'_\beta)$ and $D(G, D'_\beta, D_\alpha)$ are well defined, they have only finitely many elements $\neq 0$ in each column, and they are inverse to each other. The matrices are defined since $\partial D_\alpha \subset K$, hence $F(\partial D_\alpha) \subset F(K) = K'$ and similarly $G(\partial D'_\alpha) \subset K$. The finiteness follows from Theorem 4.6. If ω_j are the components of $\complement F(\partial D_\alpha)$ then

$$D(G \circ F, D_\alpha, y) = \sum_j D(F, D_\alpha, \omega_j) D(G, \omega_j, y), \quad y \in D_\beta,$$

by property (v). Since $F(\partial D_\alpha) \subset K'$ we have $D'_\gamma \subset \omega_j$ for some j, and

$$\overline{\omega}_j \setminus \bigcup_{D'_\gamma \subset \omega_j} D'_\gamma \subset K', \quad \text{hence } y \notin G(\overline{\omega}_j \setminus \bigcup_{D'_\gamma \subset \omega_j} D'_\gamma).$$

We can therefore use property (iv) and conclude since $G \circ F$ is the identity on ∂D_α that

$$\delta_{\alpha\beta} = \sum_\gamma D(F, D_\alpha, D'_\gamma) D(G, D'_\gamma, D_\beta), \tag{4.5}$$

where $\delta_{\alpha\beta}$ is the Kronecker symbol. For reasons of symmetry we can let the two matrices change places. If either the number of components D_α or the number of components D'_γ is finite, then the finiteness of the columns proves that one of the matrices has finite rank, and then they must both be finite, quadratic inverse matrices. This proves the theorem.

From (4.5) with $\alpha = \beta$ it follows that for every α there is some γ with $F(D_\alpha) \supset D'_\gamma$, for $D(F, D_\alpha, D'_\gamma) \neq 0$ for some γ. This leads to another theorem of Brouwer:

THEOREM 4.8 (INVARIANCE OF DOMAIN). *Let Ω be an open subset of $\mathbf{R}^n$ and let $f : \Omega \to \mathbf{R}^n$ be continuous and injective. Then it follows that $f(\Omega)$ is open and that f is a homeomorphism.*

PROOF. Let B be an open ball with $\overline{B} \subset \Omega$. Then f restricted to $\overline{B}$ is a homeomorphism $\overline{B} \to f(\overline{B})$, and $f(B)$ must contain a component of $\complement f(\partial B)$. Since B is connected it follows that $f(B)$, which is disjoint with $f(\partial B)$, is also connected which proves that $f(B)$ is equal to a component of $\complement f(\partial B)$, hence connected. This proves the theorem.

We shall now prove that the degree can also be defined for suitable maps in a Banach space B (or even in a Fréchet space). Let $\varphi : B \to B$ be a continuous map such that $\varphi(B)$ is contained in a compact set, and set $\Phi(x) = x - \varphi(x)$. Then we have:

LEMMA 4.9. *$\Phi(F)$ is closed if F is closed, and $\Phi^{-1}(C)$ is compact if C is compact.*

PROOF. Let $x_j \in F$ and $\Phi(x_j) = x_j - \varphi(x_j) \to y$. We can then choose a subsequence for which $\varphi(x_j)$ converges. The corresponding sequence x_j is then convergent, and if x is the limit we have $\Phi(x) = x - \varphi(x) = y$ since φ is continuous. If $\Phi(x_j) = x_j - \varphi(x_j) \in C$ then we can choose a subsequence for which $\varphi(x_j)$ and $x_j - \varphi(x_j)$ converge, which implies convergence of x_j and proves the second statement.

LEMMA 4.10. *Let K be a compact subset of the Banach space B, and let $\varepsilon > 0$. Then there exists a continuous map $\psi : K \to B_1$ where B_1 is a finite dimensional subspace of B, such that*

$$\|\psi(x) - x\| < \varepsilon, \quad x \in K,$$

and the range of ψ is contained in the convex hull of K.

PROOF. By the Borel-Lebesgue lemma we can find finitely many $x_j \in K$ such that K is covered by the balls $\{x \in B; \|x - x_j\| < \frac{1}{2}\varepsilon\}$. Let χ be a continuous non-negative function on $\mathbf{R}$ with $\chi(t) = 1$ for $|t| \le \frac{1}{2}$ and $\chi(t) = 0$ for $|t| \ge 1$, and set for $x \in K$

$$\varphi_j(x) = \chi(\|x - x_j\|/\varepsilon)/\Phi(x), \quad \Phi(x) = \sum_j \chi(\|x - x_j\|/\varepsilon).$$

Since $\Phi(x) > 0$ when $x \in K$ it is clear that φ_j is continuous on K, and $\sum_j \varphi_j(x) = 1$, $x \in K$. We have $\varphi_j(x) = 0$ when $\|x - x_j\| \ge \varepsilon$. The range of

$$\psi(x) = \sum_j \varphi_j(x)x_j, \quad x \in K,$$

is contained in the convex hull of $\{x_j\}$, hence in the finite dimensional vector space spanned by $\{x_j\}$, and since

$$\psi(x) - x = \sum_j \varphi_j(x)(x_j - x), \quad x \in K,$$

it follows that $\|\psi(x) - x\| \le \sum_j \varphi_j(x)\varepsilon = \varepsilon$ when $x \in K$. The lemma is proved.

We can now return to the continuous map φ with $\varphi(B) \subset K$. Recall that $\Phi(x) = x - \varphi(x)$. With ψ chosen according to Lemma 4.10 we set $\varphi_1 = \psi \circ \varphi$ and obtain $\|\varphi_1(x) - \varphi(x)\| < \varepsilon$ for every x. The range of φ_1 is contained in a finite dimensional vector space $B_1 \subset B$. If now Ω is an open subset of B and $y \notin \Phi(\partial\Omega)$, which is a closed set by Lemma 4.9, then we can choose $\varepsilon > 0$ smaller than the distance from y to $\Phi(\partial\Omega)$, which implies that $y \notin \Phi_1(\partial\Omega)$ if $\Phi_1(x) = x - \varphi_1(x)$. If B_1 has been chosen so that $y \in B_1$, we can define

$$D(\Phi, \Omega, y) = D(\Phi_1, \Omega \cap B_1, y), \tag{4.6}$$

but we have to prove that the definition is independent of the choice of Φ_1 and of B_1. Let Φ_0 and B_0 be another choice, and let B' be a finite dimensional vector space $\supset B_1 \cup B_0$. Then we have

$$D(\Phi_1, \Omega \cap B_1, y) = D(\Phi_1, \Omega \cap B', y). \tag{4.7}$$

For the proof we first assume that Φ_1 restricted to B' is in C^∞ and that y is a regular value for Φ_1 on $\Omega\cap B_1$. If $x \in B'$ and $\Phi_1(x) = y \in B_1$ then $x = y+\varphi_1(x) \in B_1$ since $\varphi_1(B') \subset B_1$. The determinant of $\Phi_1'(x)$, taken with respect to all the variables in B', is equal to the determinant when the differential is taken with respect to the variables in B_1, for $\Phi_1'(x)$ induces the identity in the quotient between the two tangent spaces. This proves (4.7), for the property $\varphi_1(B') \subset B_1$ which we have used is preserved by the approximation with C^∞ functions used in the definition of D in Theorem 4.2. The equality (4.7) is also valid for Φ_0. Now $\Phi_t = t\Phi_1+(1-t)\Phi_0$ is a homotopy with $y \notin \Phi_t(\partial\Omega)$, $0 \le t \le 1$, so it follows from the property (iii) of the degree that

$$D(\Phi_1, \Omega\cap B', y) = D(\Phi_0, \Omega\cap B', y). \tag{4.8}$$

Now (4.6) follows by combination of (4.8) with (4.7), also with Φ_1 and B_1 replaced by Φ_0 and B_0. This completes the definition of $D(\Phi, \Omega, y)$ when $\Phi(x) = x - \varphi(x)$ and φ is continuous with compact range, $y \notin \Phi(\partial\Omega)$. We leave as a simple but tedious exercise for the reader to verify that the properties (i)–(v) in Theorems 4.2, 4.5 and 4.6 are preserved by this extension. (In property (iii) it is assumed that the range of $F(x,t) - F(x,0)$ when $x \in \partial\Omega$ and $t \in [0,1]$ is compact.)

The proofs of Theorems 4.7 and 4.8 can now be repeated, which gives:

Theorem 4.7′. *Let F and F' be closed subsets of B, and assume that there exists a homeomorphism $\Phi : F \to F'$ such that $\{\Phi(x) - x; x \in F\}$ is relatively compact. Then it follows that $\complement F$ and $\complement F'$ have equally many components.*

Theorem 4.8′. *Let $\Omega \subset B$ be open, let $\Phi : \Omega \to B$ be continuous and injective, and assume that $\{\Phi(x) - x; x \in \Omega\}$ is relatively compact. Then it follows that $\Phi(\Omega)$ is open and that Φ is a homeomorphism.*

The proof of Theorem 4.7′ is just a repetition of that of Theorem 4.7 since the extension of the maps f and f^{-1} there can be made using the following lemma.

Lemma 4.11. *If φ is a continuous map from a closed set $F \subset B$ to a compact set $K \subset B$, then there is another compact set $\widetilde{K} \subset B$ and a continuous extension $\tilde\varphi$ of φ to B such that $\tilde\varphi(B) \subset \widetilde{K}$.*

Proof. By composing φ with a map given by Lemma 4.10 with $\varepsilon = 2^{-j}$ we can find a continuous map χ_j from F to a compact subset of a finite dimensional subspace B_j of B such that

$$\|\chi_j(x) - \varphi(x)\| < 2^{-j}, \quad x \in F.$$

Set $\varphi_1(x) = \chi_1(x)$ and $\varphi_j(x) = \chi_j(x) - \chi_{j-1}(x)$ when $j > 1$. Then we have $\|\varphi_j(x)\| < 2^{2-j}$ when $j > 1$ and $x \in F$, and

$$\varphi(x) = \sum_1^\infty \varphi_j(x), \quad x \in F.$$

Let $\|\varphi_1(x)\| < A$, $x \in F$. By Urysohn's theorem there is a continuous extension $\tilde\varphi_j$ of φ_j to B such that $\tilde\varphi_j(B) \subset B_j$ and $\|\tilde\varphi_j(x)\| < 2^{2-j}$ when $j > 1$ and $x \in B$, and $\|\tilde\varphi_1(x)\| < A$ when $x \in B$. Now $\tilde\varphi(x) = \sum_1^\infty \tilde\varphi_j(x)$ has the required properties, for

$$\widetilde{K} = \{\sum_1^\infty t_j; t_j \in B_j, \|t_j\| \le 2^{2-j} \text{ if } j > 1,\ \|t_1\| \le A\}$$

is compact. In fact, from a sequence $\sum_{j=1}^{\infty} t_j^\nu$ in $\tilde{K}$ we can use the Cantor diagonal procedure to select a subsequence such that $\lim_{\nu\to\infty} t_j^\nu = t_j$ exists for every j, and this implies that the sum converges in norm to $\sum_1^\infty t_j$.

PROOF OF THEOREM 4.8$'$. If F is a closed ball contained in Ω, then Φ is a homeomorphism $F \to \Phi(F)$. In fact, if $x_j \in F$ and $\Phi(x_j) \to y$ there is a subsequence such that $x_j - \Phi(x_j)$ has a limit, hence x_j has a limit $x \in F$. Then we have $\Phi(x) = y$, which by hypothesis determines x uniquely. Hence the full sequence x_j converges to x, for otherwise there would exist a subsequence such that $\|x - x_j\| > \delta > 0$, which is a contradiction. The rest of the proof is now a repetition of that of Theorem 4.8.

The following key result in the linear Fredholm theory is a special case of Theorem 4.8$'$:

THEOREM 4.12. *Let $T : B \to B$ be a compact linear map, that is, T maps bounded sets to relatively compact sets. If $\Phi(x) = x - Tx \neq 0$ when $0 \neq x \in B$, then Φ is invertible.*

PROOF. By hypothesis $T\Omega$ is relatively compact when $\Omega = \{x \in B; \|x\| < 1\}$ is the unit ball. By Theorem 4.8$'$ it follows that $\Phi(\Omega)$ is a neighborhood of the origin, which proves the theorem.

Before developing the linear Fredholm theory further by means of Theorem 4.12 and connecting it to the degree of mappings we shall prove two important fixed point theorems. The first of them depends only on the method used to define the degree.

THEOREM 4.13 (SCHAUDER'S FIXED POINT THEOREM). *Let K be a convex compact subset of B, and let $\Phi : K \to K$ be a continuous map. Then Φ has a fixed point, that is, $\Phi(x) = x$ for some $x \in K$.*

PROOF. If there is no fixed point we can choose $\varepsilon > 0$ so that $\|\Phi(x) - x\| > \varepsilon$ when $x \in K$. Choose ψ according to Lemma 4.10 with $\psi(K) \subset K$, and consider the map $\psi \circ \Phi : B_1 \cap K \to B_1 \cap K$. By Corollary 4.4 there exists some $x \in B_1 \cap K$ with $\psi(\Phi(x)) - x = 0$, hence $\|\Phi(x) - x\| = \|\Phi(x) - \psi(\Phi(x))\| < \varepsilon$. This is a contradiction which proves the theorem.

THEOREM 4.14 (LERAY-SCHAUDER'S FIXED POINT THEOREM). *Assume that the continuous map $\varphi : B \to B$ maps every bounded set to a relatively compact set, and set*

$$\Phi_t(x) = x - t\varphi(x).$$

If $\Phi_t(x) = 0$ implies $\|x\| \le C$ when $0 \le t \le 1$, then the equation $\Phi_t(x) = 0$ has a solution for every $t \in [0, 1]$.

PROOF. Let $\Omega = \{x; \|x\| < C + 1\}$, define

$$\tilde{\varphi}(x) = \begin{cases} \varphi(x), & \text{if } x \in \Omega, \\ \varphi(x(C+1)/\|x\|), & \text{if } x \notin \Omega, \end{cases}$$

and set $\widetilde{\Phi}_t(x) = x - t\tilde{\varphi}(x)$. Then $\tilde{\varphi}(B)$ is relatively compact, $0 \notin \widetilde{\Phi}_t(\partial\Omega)$ when $t \in [0, 1]$, and since $\widetilde{\Phi}_0(x) = x$ we have $D(\widetilde{\Phi}_0, \Omega, 0) = 1$. Hence it follows from property (iii) of the degree that $D(\widetilde{\Phi}_t, \Omega, 0) = 1$ for every $t \in [0,1]$, and the theorem follows from property (ii) of the degree.

In Chapter V we shall use Theorem 4.14 to prove existence of solutions of the stationary Navier-Stokes equations, but first we return to the linear Fredholm theory. Again we let T be a compact linear operator, that is, an operator mapping the unit ball to a relatively compact set. However, we no longer assume as in Theorem 4.12 that $I - T$ is injective when $I = \mathrm{Id}$ is the identity operator in B. To study the kernel we need the following classical lemma of F. Riesz:

LEMMA 4.15. *Let $B_1 \subset B_2 \subset \ldots$ be a sequence of finite dimensional subspaces of B such that $(I-T)B_k \subset B_{k-1}$, $k > 1$. Then there exists some m such that $B_j = B_m$ when $j \geq m$.*

PROOF. If all B_j are different we can for $j > 1$ choose $x_j \in B_j$ with $\|x_j\| = 1$ and $\|x_j - x\| \geq 1$ when $x \in B_{j-1}$. In fact, if $X \in B_j \setminus B_{j-1}$ and $Y \in B_{j-1}$ is a a point which minimizes $\|X - Y\|$ we can take $x_j = (X-Y)/\|X-Y\|$. When $j > k$ we have

$$Tx_j - Tx_k = x_j + (T-I)x_j - Tx_k \equiv x_j \mod B_{j-1},$$

which implies that $\|Tx_j - Tx_k\| \geq 1$. Hence the sequence Tx_j does not have any convergent subsequence which is a contradiction proving the lemma.

A first consequence of the lemma is that the vector space

$$N_k = \{x \in B; (I-T)^k x = 0\}$$

is finite dimensional, for $(I-T)^k = I - T_k$ where T_k is also compact. Since $(I-T)N_k \subset N_{k-1}$ we can also conclude that there is an integer m such that $N_j = N_m$ when $j \geq m$. Thus the finite dimensional subspace N_m of B contains all generalized eigenvectors of T with the eigenvalue 1. Moreover, $(I-T)^m B = B_m$ is a closed subspace. To prove this it suffices to prove that for some constant C

$$\|x\|_{B/N_m} \leq C\|(I-T)^m x\|, \quad x \in B.$$

If this were not true we could find a sequence x_j with

$$\|x_j\| = \|x_j\|_{B/N_m} = 1, \quad \|(I-T)^m x_j\| \to 0 \text{ as } j \to \infty.$$

We can pass to a subsequence such that Tx_j converges which implies that x_j has a limit x_0. Then we have $(I-T)^m x_0 = 0$, that is, $x_0 \in N_m$, which contradicts that the distance from x_j to N_m is ≥ 1.

We shall now prove that

$$B = B_m \oplus N_m,$$

that is, that every $x \in B$ has a unique decomposition $x = y + z$ with $y \in B_m$, $z \in N_m$, and $\|y\| + \|z\| \leq C\|x\|$. For any such decompsotion we must have

$$(I-T)^m x = (I-T)^m y.$$

Now $(I-T)^m = I - T_m$ restricts to an injective mapping $B_m \to B_m$, which has a continuous inverse $S : B_m \to B_m$ by Theorem 4.12. Thus

$$y = S(I-T)^m x, \quad z = x - S(I-T)^m x.$$

Conversely, if y and z are defined in this way then $y \in B_m$ and $(I-T)^m z = (I-T)^m x - (I-T)^m S(I-T)^m x = 0$, so $z \in N_m$, which proves the statement.

Summing up, B has a direct sum decomposition $B = B_m \oplus N_m$ such that

a) The restriction of $(I-T)$ to B_m is invertible.

b) N_m is a finite dimensional subspace with $TN_m \subset N_m$ and $(I-T)^m N_m = 0$.

For all λ in a neighborhood of 1 it follows that the restriction of $I - \lambda T$ to B_m is invertible while the determinant of the restriction to N_m is equal to $(1-\lambda)^\nu$ where $\nu = \dim N_m$. If Ω is an open bounded neighborhood of 0 we conclude when $\lambda - 1$ is sufficiently small but not 0 that

$$
\begin{aligned}
(4.9)\quad D(I-\lambda T, \Omega, 0) &= D(I-\lambda T, \Omega \cap B_m, 0) D(I-\lambda T, \Omega \cap N_m, 0)\\
&= D(I-\lambda T, \Omega \cap B_m, 0)(\operatorname{sign}(1-\lambda))^\nu.
\end{aligned}
$$

Here we have used the extension to the continuous case of the obvious formula

$$D(f_1 \times f_2, \Omega_1 \times \Omega_2, (y_1, y_2)) = D(f_1, \Omega_1, y_1) D(f_2, \Omega_2, y_2),$$

where $f_1 \in C^\infty(\mathbf{R}^n, \mathbf{R}^n)$, $f_2 \in C^\infty(\mathbf{R}^m, \mathbf{R}^m)$, $\Omega_1 \subset \mathbf{R}^n$, $\Omega_2 \subset \mathbf{R}^m$ and $y_j \notin f_j(\partial \Omega_j)$. The formula (4.9) means that $D(I-\lambda T, \Omega, 0)$ is defined and constant except when λ passes the isolated points for which $I - \lambda T$ is not injective. It is then multiplied by $(-1)^\nu$ where ν is the dimension of the corresponding space of generalized eigenvectors. When λ is sufficiently small the degree is equal to 1. Hence we have

Theorem 4.16. *If T is a compact operator in B and $I-T$ is injective, we have for every bounded open neighborhood Ω of* 0

$$D(I-T, \Omega, 0) = (-1)^\nu$$

where ν is the dimension of the space spanned by all solutions of $(I-\lambda T)^k x = 0$ when $0 < \lambda < 1$ and k is a positive integer.

Using Theorem 4.16 we can connect the degree of maps in Banach spaces to the definition which was our starting point in the smooth finite dimensional case.

Theorem 4.17. *Let $\Phi : B \to B$ be continuous and assume that $\{x - \Phi(x); x \in B\}$ is relatively compact. Assume that Φ is differentiable at x_0, that $\Phi'(x_0)$ is injective, and that $I - \Phi'(x_0)$ is compact. If Ω is a sufficiently small neighborhood of x_0 then*

$$D(\Phi, \Omega, \Phi(x_0)) = D(\Phi'(x_0), \Omega_0, 0)$$

where Ω_0 is a bounded open neighborhood of 0.

Proof. The differentiability means that

$$\|\Phi(x) - \Phi(x_0) - \Phi'(x_0)(x - x_0)\| = o(\|x - x_0\|).$$

If $\Phi_t(x) = t\Phi(x) + (1-t)(\Phi(x_0) + \Phi'(x_0)(x - x_0))$ it follows that

$$\|\Phi_t(x) - \Phi(x_0) - \Phi'(x_0)(x - x_0)\| = o(\|x - x_0\|), \quad t \in [0,1],$$

and since $\Phi'(x_0)$ is invertible it follows that $\Phi_t(x) \neq \Phi(x_0)$ when $t \in [0,1]$ and $0 < \|x - x_0\| < \delta$. Hence $D(\Phi_1, \Omega, \Phi(x_0)) = D(\Phi_0, \Omega, \Phi(x_0))$ when Ω is a sufficiently small open neighborhood of x_0, which proves the theorem.

Notes. The presentation of degree theory in this chapter has mainly followed [L]. However, references to algebraic topology in [L] have been replaced by differential calculus. The proof of the key Lemma 4.1 has been taken from [M].

References

[L] J. Leray, *La théorie des points fixes et ses applications en analyse*, Proc. Int. Congr. Math. Cambridge 1950, vol. II, pp. 202–207.

[M] J. Milnor, *Topology from a differentiable viewpoint*, University of Virginia Press, 1965.

CHAPTER V

STATIONARY SOLUTIONS OF THE NAVIER-STOKES EQUATIONS

As an application of the Leray-Schauder fixed point theorem we shall now sketch an existence proof for stationary solutions of the Navier-Stokes equations. We shall omit some technical details and will anticipate results from integration theory proved later on.

The physical problem is as follows. A bounded stationary obstacle $K \subset \mathbf{R}^3$ with C^2 boundary is immersed in a liquid with viscosity coefficient ν which moves with a prescribed speed $a = (a_1, a_2, a_3)$ at infinity. One wants to determine the velocity $v(x) = (v_1(x), v_2(x), v_3(x))$ and the pressure $p(x)$ when $x \notin K$ so that

$$-\nu\Delta v + \langle v, \partial\rangle v = -\operatorname{grad} p, \quad \operatorname{div} v = 0,$$

with the boundary conditions

$$v = 0 \quad \text{on } \partial K, \quad v(x) \to a \quad \text{as } x \to \infty.$$

(Here $\langle v, \partial\rangle v$ means $v'(v)$ with the notation in Chapter I.) We simplify the problem technically by taking a large open ball B containing K and look for a solution of the differential equations in the bounded open set $\Omega = B \setminus K$ such that

$$v = 0 \quad \text{on } \partial K, \quad v = a \quad \text{on } \partial B.$$

Assume that we have a solution with $v \in C^2(\overline{\Omega})$ and $p \in C^1(\overline{\Omega})$. If $\varphi = (\varphi_1, \varphi_2, \varphi_3)$ belongs to the space $\mathcal{H}(\Omega)$ of C^1 vector fields with

$$\operatorname{div} \varphi = 0 \quad \text{in } \Omega, \quad \varphi = 0 \quad \text{on } \partial\Omega, \tag{5.1}$$

then we obtain after scalar multiplication with φ and integration by parts

$$\nu \int_\Omega \sum_{j,k=1}^3 \partial\varphi_k/\partial x_j \partial v_k/\partial x_j \, dx - \int_\Omega \sum_{j,k=1}^3 v_j v_k \partial\varphi_k/\partial x_j \, dx = 0, \quad \varphi \in \mathcal{H}(\Omega). \tag{5.2}$$

The pressure p drops out since $\operatorname{div}\varphi = 0$ and another term where φ is not differentiated drops out since $\operatorname{div} v = 0$. Conversely, if $v \in C^2(\overline{\Omega})$ satisfies (5.2) and $\operatorname{div} v = 0$, then $-\nu\Delta v + \langle v, \partial\rangle v = P$ where P is continuous in $\overline{\Omega}$ with values in $\mathbf{R}^3$ and

$$\int \langle P, \varphi\rangle = 0 \quad \text{when } \varphi \in \mathcal{H}(\Omega).$$

This implies that $P = -\operatorname{grad} p$ for some $p \in C^1(\overline{\Omega})$. To prove this we just have to prove that if $[0,1] \ni t \mapsto x(t)$ is a closed polygonal curve contained in Ω, then $\int_0^1 \langle P(x(t)), dx(t)\rangle = 0$. It suffices to prove that if $\psi \in C_0^\infty(\mathbf{R}^3)$ has support in the unit ball and $\int \psi(y)\, dy = 1$, then

$$\int \psi_\varepsilon(y) \int_0^1 \langle P(x(t)+y), dx(t)\rangle = 0, \quad \psi_\varepsilon(y) = \psi(y/\varepsilon)/\varepsilon^3,$$

if ε is smaller than the distance from the curve to $\complement\Omega$, for the integral converges to $\int_0^1 \langle P(x(t)), dx(t)\rangle$ as $\varepsilon \to 0$. The double integral can be written

$$\iint_{0<t<1} \langle P(y), dx(t)\rangle \psi_\varepsilon(y - x(t))\, dy = \int \langle P(y), \varphi_\varepsilon(y)\rangle\, dy,$$
$$\varphi_\varepsilon(y) = \int_0^1 \psi_\varepsilon(y - x(t)) dx(t).$$

Since $\operatorname{supp} \varphi_\varepsilon$ is at distance $\leq \varepsilon$ from the curve, we have $\varphi_\varepsilon \in C_0^\infty(\Omega)$, and

$$\operatorname{div} \varphi_\varepsilon(y) = \int_0^1 \langle \psi_\varepsilon'(y - x(t)), dx(t)\rangle = -\int_0^1 d\psi_\varepsilon(y - x(t)) = 0$$

the claim has been proved. Thus it is reasonable to accept (5.2) and the condition $\operatorname{div} v = 0$ as a restatement of the original differential equations, with the pressure term removed.

If A is a vector field with $A = 0$ near ∂K and $A = a$ near ∂B, then the boundary conditions mean that $v - A = 0$ on $\partial\Omega$. It is easy to construct C^∞ vector fields $A(x)$ with $\operatorname{div} A = 0$ such that $A = 0$ near ∂K and $A = a$ near ∂B. In fact,

$$a_k = \tfrac{1}{2} \sum_{j=1}^{3} \partial(a_k x_j - a_j x_k)/\partial x_j, \quad k = 1,2,3,$$

so if $\varphi \in C^\infty$ is equal to 0 in a neighborhood of K and equal to $\frac{1}{2}$ in a neighborhood of $\complement B$ then

$$A_k(x) = \sum_{j=1}^{3} \partial(\varphi(x)(a_k x_j - a_j x_k))/\partial x_j, \quad k = 1,2,3$$

is a vector field with the desired properties. Note that for small $\delta > 0$ we can choose φ equal to 0 at distance $> \delta$ from ∂B and $\varphi' = O(\delta^{-1})$, which yields a vector field A^δ with the desired properties such that $A^\delta = 0$ at distance $> \delta$ from ∂B and $A^\delta = O(\delta^{-1})$. This will be needed later on.

We shall now extend the notion of solution somewhat. We make $\mathcal{H}(\Omega)$ into a prehilbert space by introducing the norm

$$\|\varphi\|_H^2 = \int_\Omega \sum_{j,k=1}^{3} |\partial\varphi_k/\partial x_j|^2\, dx.$$

Since

$$\sum_{k=1}^{3}\int_\Omega |\varphi_k|^2\,dx = -\int_\Omega \sum_{k=1}^{3} x_k\partial|\varphi_k|^2/\partial x_k\,dx \le C\|\varphi\|_H\Big(\sum_{k=1}^{3}\int_\Omega |\varphi_k|^2\,dx\Big)^{\frac12}$$

we obtain

$$\|\varphi\|_{L^2} \le C\|\varphi\|_H, \quad \varphi \in \mathcal{H}(\Omega). \tag{5.3}$$

The completion $H(\Omega)$ of $\mathcal{H}(\Omega)$ can therefore be regarded as a subset of $L^2(\Omega,\mathbf{R}^3)$. Our final formulation of the existence problem is now to *find* v *with* $v - A \in \mathcal{H}(\Omega)$ *so that* (5.2) *is valid for every* $\varphi \in \mathcal{H}(\Omega)$*.* The condition $v - A \in H(\Omega)$ implies $\operatorname{div} v = 0$ in the sense of distribution theory, and it is independent of the choice of A.[1] The equation (5.2) makes sense and is then true for all $\varphi \in H(\Omega)$ since Sobolev's inequalities give a stronger version of (5.3),

$$\|\varphi\|_{L^6} \le C\|\varphi\|_H, \quad u \in H(\Omega), \tag{5.4}$$

which guarantees that the second term in (5.2) is a continuous function of $\varphi \in H(\Omega)$ when $v - A \in H(\Omega)$. If we introduce $u = v - A$, the condition (5.2) can also be written

$$\begin{aligned}\nu\int_\Omega \sum_{j,k=1}^{3} (\partial u_k/\partial x_j + \partial A_k/\partial x_j)\partial\varphi_k/\partial x_j\,dx \\ -\int_\Omega \sum_{j,k=1}^{3} (u_j + A_j)(u_k + A_k)\partial\varphi_k/\partial x_j\,dx = 0,\ \varphi \in H(\Omega).\end{aligned} \tag{5.5}$$

If $b \in L^2(\Omega, M_3)$ where M is the space of real 3×3 matrices, we can define the projection $Pb \in H(\Omega)$ by

$$(\varphi, Pb)_H = \int_\Omega \sum_{j,k=1}^{3} \partial\varphi_k/\partial x_j b_{kj}\,dx, \quad \varphi \in H(\Omega). \tag{5.6}$$

In fact, the right-hand side can be estimated by $\|\varphi\|_H\|b\|_{L^2}$ and can therefore be written as the scalar product with a uniquely defined element $Pb \in H$. The linear operator $P : L^2 \to H$ is continuous with norm 1, and $P\partial a/\partial x = a$ if $a \in \mathcal{H}(\Omega)$.

We can now rewrite (5.5) in the form

$$\nu(u + P\partial A/\partial x) - Tu = 0, \quad u \in H(\Omega), \tag{5.7}$$

where T is the composition

$$H(\Omega) \to L^4(\Omega,\mathbf{R}^3) \xrightarrow{(\cdot+A)^2} L^2(\Omega, M_3) \xrightarrow{P} H(\Omega).$$

[1] When $\partial\Omega$ is sufficiently smooth one can prove that a solution in this sense is also smooth, so that it is a solution in the classical sense.

Here $(u + A)^2$ denotes the matrix $(u_j + A_j)(u_k + A_k)$, $j, k = 1, 2, 3$, when $u \in L^4(\Omega, \mathbf{R}^3)$. Each of the operators is continuous, and since $H(\Omega) \to L^2(\Omega)$ is compact while $H(\Omega) \to L^6(\Omega)$ is continuous, the embedding $H(\Omega) \to L^4(\Omega)$ is compact, for a sequence which is bounded in $L^6(\Omega)$ and converges to 0 in $L^2(\Omega)$ must converge to 0 in $L^p(\Omega)$ for $2 \le p < 6$.

We shall apply the Leray-Schauder fixed point theorem (Theorem 4.14) to the equation (5.7). This theorem guarantees the existence of a solution provided that there is a constant M such that

$$u \in H(\Omega),\ t \in [0, 1],\ u + P\partial A/\partial x - tTu/\nu = 0 \implies \|u\|_H \le M. \tag{5.8}$$

The equation $u + P\partial A/\partial x - tTu/\nu = 0$ means that (5.5) is valid with a factor t in front of the last term. We take $\varphi = u$ in the so modified equation (5.5) and note that

$$\int_\Omega \sum_{j,k=1}^{3} (A_j + u_j)u_k \partial u_k/\partial x_j \, dx = \tfrac{1}{2} \int_\Omega \sum_{j,k=1}^{3} (A_j + u_j)\partial u_k^2/\partial x_j \, dx = 0$$

since this is true for all $u \in \mathcal{H}(\Omega)$. For all u satisfying the hypotheses in (5.8) it follows that

$$\nu \|u\|_H^2 \le C_A \|u\|_H + \Big| \int_\Omega \sum_{j,k=1}^{3} u_j A_k \partial u_k/\partial x_j \, dx \Big|. \tag{5.9}$$

Assume now that (5.8) is not valid for any constant M. Then there is a sequence $u^{(n)}$, $t^{(n)}$ satisfying the hypotheses of (5.8) while $\|u^{(n)}\|_H \to \infty$. Set $w^{(n)} = u^{(n)}/\|u^{(n)}\|_H$. The bounded sequence $w^{(n)}$ has a subsequence converging weakly to a limit $w \in H(\Omega)$, which implies that it converges to w in L^2 norm. If we divide (5.9) by $\|u\|_H^2$ and take $u = u^{(n)}$, it follows when $n \to \infty$ through a subsequence that

$$\nu \le \Big| \int_\Omega \sum_{j,k=1}^{3} w_j A_k \partial w_k/\partial x_j \, dx \Big|. \tag{5.10}$$

We could from the beginning have worked with another divergence free vector field $\widetilde{A}$ with the appropriate boundary values; then $A - \widetilde{A}$ is an arbitrary element in $\mathcal{H}(\Omega)$. Since $P(\partial A/\partial x - \partial \widetilde{A}/\partial x) = A - \widetilde{A}$ it follows that (5.7) implies the corresponding equation with A replaced by $\widetilde{A}$ (which affects the definition of T), provided that u is replaced by $\tilde{u}$ and $\tilde{u} + \widetilde{A} = u + A$, for then we have $\tilde{u} + P\partial\widetilde{A}/\partial x = u + P\partial A/\partial x$. For a sequence $u^{(n)}$ such that $\|u^{(n)}\|_H \to \infty$ we have $\|\tilde{u}^{(n)}\|_H \to \infty$ for the corresponding sequence, so we conclude that A may be replaced by $\widetilde{A}$ in (5.10). Since the right-hand side of (5.10) must not vanish for any $\widetilde{A}$ this means in particular that

$$\int_\Omega \sum_{j,k=1}^{3} w_j a_k \partial w_k/\partial x_j \, dx = 0, \quad a \in \mathcal{H}(\Omega).$$

Assume for a moment that w is smooth in $\overline{\Omega}$. Then it follows that there is a function p such that

$$\sum_{j=1}^{3} w_j \partial w_k/\partial x_j = -\partial p/\partial x_k, \quad k = 1, 2, 3.$$

Since $w = 0$ on $\partial\Omega$ this implies that $\operatorname{grad} p = 0$ on $\partial\Omega$, hence that p is constant on ∂B. Hence

$$\int_\Omega \sum_{j,k=1}^{3} w_j A_k \partial w_k/\partial x_j \, dx = -\int_\Omega \sum_{k=1}^{3} A_k \partial p/\partial x_k \, dx$$
$$= -\int_\Omega \operatorname{div}(pA)\, dx = -\int_{\partial B} p\langle A, n\rangle \, dS = 0$$

since p is constant on ∂B and $\int_{\partial B} \langle A, n\rangle \, dS = \int_{\partial K} \langle A, n\rangle \, dS = 0$ by the Gauss-Green formula. This contradiction with (5.10) means from the point of view of physics that the Navier-Stokes equations with zero viscosity cannot be solved with the given boundary conditions.

To give a strict proof of a contradiction without any unjustified regularity assumptions we choose in (5.10) for A the vector fields A^δ constructed above. Let $B = \{x \in \mathbf{R}^3; |x| < R\}$ and set $B_\delta = \{x \in \mathbf{R}^3; R - \delta \le |x| < R\}$. Then (5.10) gives for small δ

$$\nu \le C\delta^{-1} \|w\|_{L^2(B_\delta)} \|\partial w/\partial x\|_{L^2(B_\delta)}. \tag{5.11}$$

If $f \in C^1(\overline{\Omega})$ and $f = 0$ on ∂B we have

$$\int_{B_\delta} f^2 \, dx \le \tfrac{1}{2} R_\delta^{-2} \int_{B_\delta} f^2 \sum_{j=1}^{3} x_j \partial(|x|^2 - R_\delta^2)/\partial x_j \, dx$$
$$= -\tfrac{1}{2} R_\delta^{-2} \int_{B_\delta} (|x|^2 - R_\delta^2) \sum_{j=1}^{3} \partial(x_j f^2)/\partial x_j \, dx$$
$$\le 2R^2 R_\delta^{-2} \delta \|f\|_{L^2(B_\delta)} \|\partial f/\partial x\|_{L^2(B_\delta)},$$

where $R_\delta = R - \delta$, and if $\delta < \frac{1}{2}R$ this implies

$$\|f\|_{L^2(B_\delta)} \le 4\delta \|\partial f/\partial x\|_{L^2(B_\delta)}.$$

We can apply this estimate to all $w \in \mathcal{H}(\Omega)$, hence to all $w \in H(\Omega)$, and using (5.11) we then obtain

$$\nu \le 4C \|\partial w/\partial x\|^2_{L^2(B_\delta)}. \tag{5.12}$$

This gives a contradiction when $\delta \to 0$ since the right-hand side converges to 0 then. This completes the proof of the existence of a solution of the stationary Navier-Stokes equation.

Notes. For a much more thorough discussion of the Navier-Stokes equations using the Leray-Schauder fixed point theorem we refer to [L], which we have followed here apart from the correction of a minor error. These methods can also be used to prove the existence of several solutions of the stationary Navier-Stokes equations under suitable conditions. The idea is to use Theorem 4.16 to determine the degree (index) of an isolated solution and prove that it is not equal to the total degree which is determined by means of a homotopy. Such results can be found in [V].

References

[L] O. A. Ladyzenskaja, *The mathematical theory of viscous incompressible flow*, Gordon and Breach, 1963.

[V] W. Velte, *Stabilität und Verzweigung stationärer Lösungen der Navier-Stokesschen Gleichungen beim Taylor Problem*, Arch. Rat. Mech. Anal. **22** (1966), 1–14.

CHAPTER VI

INTEGRATION IN A VECTOR SPACE

We assume that the reader is familiar with the basic properties of the Riemann integral; in particular that the integral $\int f\,dx$ is defined for all $f \in C_0(\mathbf{R}^n)$, the space of continuous functions with compact support, and that

$$
\begin{gathered}
\int (af+bg)\,dx = a\int f\,dx + b\int g\,dx, \quad f,g\in C_0,\ a,b\in\mathbf{R} \\
\int f\,dx \ge 0, \quad 0\le f\in C_0,
\end{gathered}
\tag{6.1}
$$

$$
\int f(x-h)\,dx = \int f(x)\,dx, \quad f\in C_0,\ h\in\mathbf{R}^n. \tag{6.2}
$$

These properties characterize the integral apart from a multiplicative constant which is determined by the condition that the volume of the unit cube shall be equal to 1. If $A:\mathbf{R}^n\to\mathbf{R}^n$ is a linear bijection and we set $A^*f(x)=f(Ax)$, then

$$
\int f\,dx = |\det A|\int A^*f\,dx, \quad f\in C_0(\mathbf{R}^n). \tag{6.3}
$$

In fact, the right-hand side has the properties (6.1) and (6.2) so it must be a constant depending on A times the left-hand side. This constant must be a multiplicative function of A which is equal to 1 for diagonal matrices and for orthogonal matrices, hence for all A.

We shall use the essential uniqueness to calculate some important integrals. First we note that if $x\mapsto\|x\|$ is an arbitrary norm in $\mathbf{R}^n$ then there is a constant C such that

$$
\int_{\mathbf{R}^n} f(\|x\|)\,dx = C\int_0^\infty f(t)t^{n-1}\,dt, \quad f\in C_0(\mathbf{R}). \tag{6.4}
$$

In the proof we may assume that $f=0$ in a neighborhood of 0 and rewrite (6.4) using the notation $f(t)t^n = g(\log t)$,

$$
\int_{\mathbf{R}^n} g(\log\|x\|)\|x\|^{-n}\,dx = C\int g(\log t)t^{-1}\,dt, \quad g\in C_0(\mathbf{R}). \tag{6.4$'$}
$$

Since the substitutions $x\mapsto e^a x$ and $t\mapsto e^a t$ show that both sides are unchanged if g is replaced by $g(\cdot+a)$, $a\in\mathbf{R}$, the two sides are functions of g satisfying (6.1) and (6.2), hence proportional. If we let f approach the characteristic function for

$[0,1]$ from above and below we conclude that C/n is the volume of the unit ball. With $f(t) = e^{-at^2}$ we also obtain

(6.5)
$$\int_{\mathbf{R}^n} e^{-a\|x\|^2}\,dx = C\int_0^\infty e^{-at^2}t^{n-1}\,dt = \tfrac12 C\int_0^\infty e^{-as}s^{\frac12 n-1}\,ds = \tfrac12 Ca^{-\frac12 n}\Gamma(\tfrac12 n)$$

where the Γ function is defined by

$$\Gamma(x) = \int_0^\infty e^{-s}s^{x-1}\,ds, \quad s>0. \tag{6.6}$$

Partial integration gives

$$x\Gamma(x) = \Gamma(x+1), \quad x>0, \tag{6.7}$$

hence

$$\Gamma(\nu) = (\nu-1)!, \quad \Gamma(\nu-\tfrac12) = (\nu-\tfrac32)(\nu-\tfrac52)\dots\tfrac12\Gamma(\tfrac12)$$

when ν is a positive integer.

Now we let $\|x\| = (x_1^2+\dots+x_n^2)^{\frac12}$ be the Euclidean norm. When $n=2$ it follows from (6.5) that

$$\int_{\mathbf{R}^2} e^{-a\|x\|^2}\,dx = 2\pi/2a = \pi/a,$$

for $C/n = C/2 = \pi$. Hence

$$\int_{\mathbf{R}^{2n}} e^{-a\|x\|^2}\,dx = (\pi/a)^n, \tag{6.8}$$

which implies that

$$\int_{\mathbf{R}^n} e^{-a\|x\|^2}\,dx = (\pi/a)^{\frac12 n}, \tag{6.9}$$

for the square of the left-hand side is the left-hand side of (6.8) since we can integrate over n variables at a time in (6.8). With $n=1$ in (6.5) we now obtain

$$\sqrt{\pi/a} = \Gamma(\tfrac12)/\sqrt{a},$$

hence $\Gamma(\frac12) = \sqrt{\pi}$. If we denote the volume of the unit ball in $\mathbf{R}^n$ by C_n and replace C in (6.5) by nC_n, it follows from (6.9) that

$$C_{2\nu} = \pi^\nu/\nu!, \quad C_{2\nu+1} = 2^{\nu+1}\pi^\nu/(2\nu+1)!!. \tag{6.10}$$

For every vector space V over $\mathbf{R}$ of dimension n there exist linear forms $C_0(V) \ni f \mapsto I(f)$ with the properties (6.1), (6.2), and they differ only by a constant factor. For if $\psi : \mathbf{R}^n \to V$ is a linear bijection we can choose $I(f) = \int \psi^* f\,dx$ where $\psi^* f = f\circ\psi \in C_0(\mathbf{R}^n)$, and the uniqueness apart from a constant factor follows from the same fact for $\mathbf{R}^n$. To fix the constant one needs some additional structure in V such as a bilinear form B on $V\times V$. Note that if $\psi : W \to V$ is a linear bijection then a bilinear form B on V induces another bilinear form ψ^*B on W defined by

$$(\psi^*B)(x,y) = B(\psi x, \psi y), \quad x, y \in W.$$

A bilinear form on $\mathbf{R}^n$ can always be written

$$B(x,y) = \sum_{j,k=1}^{n} B_{jk} x_j y_k,$$

and we associate with it the number $\det B = \det(B_{jk})$. If $\psi : \mathbf{R}^n \to \mathbf{R}^n$ is a linear bijection then

$$\det \psi^*B = (\det \psi)^2 \det B$$

since the matrix of ψ^*B is ${}^t\psi(B_{jk})\psi$. Hence it follows from (6.3) that

$$|\det B|^{\frac{1}{2}} \int f\,dx = |\det \psi^*B|^{\frac{1}{2}} \int \psi^*f\,dx, \quad f \in C_0(\mathbf{R}^n). \tag{6.11}$$

Let now V be a finite dimensional vector space with a given bilinear form B, and set with a linear bijection $\psi : \mathbf{R}^n \to V$

$$I_B(f) = |\det \psi^*B|^{\frac{1}{2}} \int \psi^*f\,dx, \quad f \in C_0(V). \tag{6.12}$$

This definition is independent of the choice of ψ, for if ψ_1 is another choice then $\psi_1 = \psi \circ \psi_2$ where $\psi_2 : \mathbf{R}^n \to \mathbf{R}^n$. If $\psi^*f = g$ we obtain $\psi_1^*f = \psi_2^*g$, and if $\psi^*B = C$ then $\psi_1^*B = \psi_2^*C$, so (6.11) gives

$$\begin{aligned} |\det \psi^*B|^{\frac{1}{2}} \int \psi^*f\,dx &= |\det C|^{\frac{1}{2}} \int g\,dx \\ &= |\det \psi_2^*C|^{\frac{1}{2}} \int \psi_2^*g\,dx = |\det \psi_1^*B|^{\frac{1}{2}} \int \psi_1^*g\,dx, \end{aligned}$$

which proves the statement. If W is another n dimensional vector space and ψ is a linear bijection $W \to V$, then the definition implies that

$$I_B(f) = I_{\psi^*B}(\psi^*f), \quad f \in C_0(V). \tag{6.13}$$

The form I_B is of course identically 0 if B is degenerate, that is, if there is some $x \neq 0$ in V with $B(x,V) = 0$ (or $B(V,x) = 0$). This does not happen if V is a Euclidean space, that is, V is provided with a positive definite quadratic form. Then we take for B the corresponding symmetric bilinear form with $B(x,x) = \|x\|^2$, for $B(x,x) > 0$ when $x \neq 0$. The normalisation of I_B is then also clear from the fact that by (6.9) and (6.13)

$$I_B(e^{-\pi\|\cdot\|^2}) = 1, \tag{6.14}$$

for we can find a map $\psi : \mathbf{R}^n \to V$ such that ψ^*B is the standard Euclidean scalar product.

Starting from (6.14) one can also define the *Gaussian measure*

$$f \mapsto I_B(fe^{-\pi\|\cdot\|^2})$$

when V is an arbitrary Hilbert space. For let $\mathcal{C}(V)$ denote the set of all cylindrical functions u, that is, bounded continuous functions u such that there is a closed subspace F of finite codimension with $u(x) = u(y)$ when $x - y \in F$. These functions form an algebra, for the intersection of subspaces of finite codimension has finite codimension. For $f \in \mathcal{C}(V)$ we can define $I(fe^{-\pi\|\cdot\|^2})$ as the integral over the orthogonal complement $F^\perp$ of F. This does not depend on the choice of F. In fact, if F_1 is another possible choice then $G = F \cap F_1$ has finite codimension and it is obviously sufficient to prove that the integrals over $F^\perp$ and $G^\perp$ are equal. But $G^\perp = F^\perp \oplus H$ where $H \subset F$ so that f is constant along H. This proves the statement since the integral of $e^{-\pi\|\cdot\|^2}$ over H is equal to 1. For the Gaussian measure now defined in $\mathcal{C}(V)$ we have (6.1) and

$$|I_B(f)| \leq \sup |f|.$$

One can now extend the domain of definition of I_B by the standard methods of Lebesgue (Daniell) integration theory, but we shall not pursue this direction further.

CHAPTER VII

THE EUCLIDEAN VOLUME ELEMENT ON A SUBMANIFOLD OF $\mathbf{R}^m$

In Chapter VI we only had to use linear changes of variables, but we shall now need a general result, which is proved by approximation with linear changes of variables:

THEOREM 7.1. *Let Ω_1, Ω_2 be open subsets of $\mathbf{R}^n$ and let $\psi : \Omega_1 \to \Omega_2$ be a C^1 diffeomorphism, that is, a C^1 bijection with C^1 inverse. Then we have*

$$\int_{\Omega_2} u\,dx = \int_{\Omega_1} (\psi^* u)|\det \psi'|\,dx, \quad u \in C_0(\Omega_2).$$

We shall now define the Euclidean volume element on a C^1 manifold $F \subset \mathbf{R}^m$ of dimension n (see Definition 2.2). For every $x_0 \in F$ we can choose a C^1 map from an open set $\omega \subset \mathbf{R}^n$ parametrising F in an open neighborhood Ω of x_0 in F. If $u \in C_0(\Omega)$ we would like to define

$$I(u) = \int \sqrt{\det({}^t\psi'(t)\psi'(t))}\psi^* u(t)\,dt, \tag{7.1}$$

for if ψ and therefore F is linear we have seen in Chapter VI that this gives the integral of u with respect to the Euclidean metric in $\mathbf{R}^m$, restricted to F. To justify (7.1) we first note that if $\psi_1 : \omega_1 \to F$ is another parametrisation and the support of u is contained in $\psi_1(\omega_1) \cap \psi(\omega)$, then

$$\int \sqrt{\det({}^t\psi_1'(s)\psi_1'(s))}\psi_1^* u\,ds = \int \sqrt{\det({}^t\psi'(t)\psi'(t))}\psi^* u\,dt. \tag{7.2}$$

In fact, $\psi_1 = \psi\circ\psi_2$ defines a diffeomorphism ψ_2 from a neighborhood of $\psi_1^{-1}(\operatorname{supp} u)$ in $\mathbf{R}^n$ to a neighborhood of $\psi^{-1}(\operatorname{supp} u)$ in $\mathbf{R}^n$, and $\psi_1^* u = \psi_2^* v$ where $v = \psi^* u$, hence

$$\int \sqrt{\det({}^t\psi'(t)\psi'(t))}v(t)\,dt = \int \sqrt{\psi_2^* \det({}^t\psi'\psi')}|\det \psi_2'|\psi_1^* u\,ds.$$

By the chain rule

$$\begin{aligned} \det({}^t\psi_1'(s)\psi_1'(s)) &= \det({}^t\psi_2'(s)\,{}^t\psi'(\psi_2(s))\psi'(\psi_2(s))\psi_2'(s)) \\ &= (\det \psi_2'(s))^2 \psi_2^* \det({}^t\psi'\psi'), \end{aligned} \tag{7.3}$$

which proves (7.2).

Now we generalise (7.1) further by writing

$$I(u) = \sum I_j(u_j), \quad \text{if } u = \sum u_j, \tag{7.4}$$

where the sum is finite, u_j is continuous with support covered by a local parametrisation ψ_j, and I_j is defined by (7.1) with $\psi = \psi_j$. Our next goal is to prove that (7.4) defines $I(u)$ uniquely for all $u \in C_0(F)$. We can choose a partition of unity $1 = \sum \varphi_k$ in a neighborhood of $\operatorname{supp} u$ so that the support of $\varphi_k \in C_0(F)$ is covered by a local parametrisation $\tilde{\psi}_k$. Then $u = \sum \varphi_k u$ is a decomposition for which we can apply (7.4). Given any decomposition as in (7.4) we can choose φ_k as above with $\sum \varphi_k = 1$ in a neighborhood of $\cup \operatorname{supp} u_j$ and obtain

$$\sum_j I_j(u_j) = \sum_j \sum_k I_j(\varphi_k u_j) = \sum_k \sum_j \widetilde{I}_k(\varphi_k u_j) = \sum_k \widetilde{I}_k(\varphi_k u)$$

where $\widetilde{I}_k$ is defined by means of the parametrisation $\tilde{\psi}_k$. This proves that (7.4) gives a unique definition of $I(u)$. It is obvious that I depends linearly on u and that $I(u) \geq 0$ when $u \geq 0$. Thus $I(u)$ is a positive measure on F; it is called the Euclidean volume measure (or area measure or arc measure depending on the dimension), and is often denoted by dS or $d\sigma$. Thus we write

$$I(u) = \int u \, d\sigma, \quad u \in C_0(F).$$

Let us now examine more closely what made it possible to obtain a unique global definition starting from (7.1). Assume that we are given a continuous function U on $F \times \mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$ with support in $K \times \mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$, where K is a compact subset of $\psi(\omega)$ and $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$ is the space of linear maps $\mathbf{R}^n \to \mathbf{R}^m$. The integral

$$\int U(\psi(t), \psi'(t)) \, dt \tag{7.5}$$

reduces to (7.1) if $U(x, A) = \sqrt{\det({}^tAA)}u(x)$. We can repeat the calculation which led to (7.2) to see if (7.5) is also independent of the choice of parametrisation. With the notation used in the proof of (7.2) this requires that

$$U(\psi_1(s), \psi'(\psi_2(s)))|\det \psi_2'(s)| = U(\psi_1(s), \psi'(\psi_2(s))\psi_2'(s)),$$

or with a change of notation

$$U(x, A)|\det B| = U(x, AB); \quad x \in F, \ A \in \mathcal{L}(\mathbf{R}^n, \mathbf{R}^m), \ B \in \mathcal{L}(\mathbf{R}^n, \mathbf{R}^n). \tag{7.6}$$

If we could restrict ourselves to parametrisations for which $\det \psi_2'(s) > 0$ then it would suffice to assume (7.6) when $\det B > 0$. It is geometrically plausible that this means using only parametrisations giving a consistent orientation. We shall give a precise meaning to this later on. Here we just note that with this reservation the integral (7.5) can be defined globally in a unique way if

$$U(x, A) \det B = U(x, AB); \quad x \in F, \ A \in \mathcal{L}(\mathbf{R}^n, \mathbf{R}^m), \ B \in \mathcal{L}(\mathbf{R}^n, \mathbf{R}^n). \tag{7.7}$$

An obvious advantage of this condition compared to (7.6) is that in (7.7) U can be a polynomial. We shall analyse the meaning of (7.7) in the following chapter.

CHAPTER VIII

EXTERIOR DIFFERENTIAL FORMS

Let V be a finite dimensional vector space. As a first step in the discussion of (7.7) we shall determine all real valued polynomials U defined in $\mathcal{L}(\mathbf{R}^n, V)$ such that

$$U(AB) = U(A)\det B; \quad A \in \mathcal{L}(\mathbf{R}^n, V),\ B \in \mathcal{L}(\mathbf{R}^n, \mathbf{R}^n). \tag{8.1}$$

Writing

$$Ax = \sum_1^n x_j v_j, \quad x = (x_1, \dots, x_n) \in \mathbf{R}^n,$$

where $v_j \in V$, we can identify $\mathcal{L}(\mathbf{R}^n, V)$ with $V \times \cdots \times V$ and consider U as a function $u(v_1, \dots, v_n)$, $\quad v_j \in V$. Then the condition (8.1) becomes

$$u\Big(\sum_{k_1=1}^n B_{k_1 1} v_{k_1}, \dots, \sum_{k_n=1}^n B_{k_n n} v_{k_n}\Big) = u(v_1, \dots, v_n)\det B. \tag{8.2}$$

When B is specialized to diagonal matrices we conclude that $u(v_1, \dots, v_n)$ must be homogeneous of degree 1 in every v_j, and since a homogeneous polynomial of degree 1 is linear, it follows that u is linear in every v_j, that is, $u \in \mathcal{L}^n(V)$, the set of all real valued n linear forms on V. If $\pi : \{1, \dots, n\} \to \{1, \dots, n\}$ is a permutation and we set $B_{jk} = 1$ for $j = \pi(k)$ and $B_{jk} = 0$ for $j \neq \pi(k)$ then (8.2) implies

$$u(v_{\pi(1)}, \dots, v_{\pi(n)}) = u(v_1, \dots, v_n)\operatorname{sgn}\pi. \tag{8.3}$$

Thus u is an alternating multilinear form. Conversely, it follows from (8.3) and the multilinearity that the left-hand side of (8.2) is equal to

$$\sum_{k_1, \dots, k_n=1}^n u(v_1, \dots, v_n)\operatorname{sgn}\begin{pmatrix} 1 \dots n \\ k_1 \dots k_n \end{pmatrix} B_{k_1 1} \cdots B_{k_n n} = u(v_1, \dots, v_n)\det B.$$

For a polynomial u the condition (8.2) is therefore equivalent to $u \in \mathcal{L}_a^n(V, \mathbf{R})$, the space of alternating n linear real valued forms on V. For $n = 1$ this is of course equal to the dual space V^* of V.

Example 8.1. Let $\theta_1, \dots, \theta_n \in V^*$. Then an element in $\mathcal{L}_a^n(V, \mathbf{R})$ is defined by

$$(v_1, \dots, v_n) \mapsto \det\langle v_k, \theta_j\rangle_{j,k=1}^n. \tag{8.4}$$

We shall now prove that (8.4) is not just an example but that the vector space $\mathcal{L}_a^n(V,\mathbf{R})$ is spanned by elements of this form. For the proof we select a basis $\{e_j\}$ in V and a dual basis $\{\theta_j\}$ in V^*, thus

$$\langle e_i, \theta_k\rangle = \delta_{ij}, \quad i,j = 1,\dots,\dim V,$$

and $v = \sum_j \langle v, \theta_j\rangle e_j$ if $v \in V$. If $u \in \mathcal{L}_a^n(V,\mathbf{R})$ we can now write

$$\begin{aligned} u(v_1,\dots,v_n) &= u\Big(\sum_{j_1}\langle v_1,\theta_{j_1}\rangle e_{j_1},\dots,\sum_{j_n}\langle v_n,\theta_{j_n}\rangle e_{j_n}\Big)\\ &= \sum_{j_1,\dots,j_n} u(e_{j_1},\dots,e_{j_n})\langle v_1,\theta_{j_1}\rangle\dots\langle v_n,\theta_{j_n}\rangle \\ &= \frac{1}{n!}\sum_{j_1,\dots,j_n} u(e_{j_1},\dots,e_{j_n})\det\langle v_k,\theta_{j_l}\rangle_{k,l=1}^n. \end{aligned} \tag{8.5}$$

The last equality follows since the last expression is equal to

$$\begin{aligned} &\frac{1}{n!}\sum_{j_1,\dots,j_n}\sum_{\pi} u(e_{j_1},\dots,e_{j_n})\operatorname{sgn}\pi\langle v_1,\theta_{j_{\pi(1)}}\rangle\dots\langle v_n,\theta_{j_{\pi(n)}}\rangle \\ &= \frac{1}{n!}\sum_{j_1,\dots,j_n}\sum_{\pi} u(e_{j_{\pi(1)}},\dots,e_{j_{\pi(n)}})\langle v_1,\theta_{j_{\pi(1)}}\rangle\dots\langle v_n,\theta_{j_{\pi(n)}}\rangle, \end{aligned}$$

where π varies over all permutations of $1,\dots,n$. We can also write (8.5) in the form

$$u(v_1,\dots,v_n) = \sum_{j_1<\dots<j_n} u(e_{j_1},\dots,e_{j_n})\det\langle v_k,\theta_{j_l}\rangle_{k,l=1}^n.$$

Conversely, if we set

$$u(v_1,\dots,v_n) = \sum_{j_1<\dots<j_n} a_{j_1\dots j_n}\det\langle v_k,\theta_{j_l}\rangle_{k,l=1}^n, \tag{8.6}$$

then

$$u(e_{i_1},\dots,e_{i_n}) = a_{i_1\dots i_n} \quad \text{if } i_1<\dots<i_n.$$

This proves that $\mathcal{L}_a^n(V)$ is a vector space of dimension $\binom{\dim V}{n}$, and a basis is displayed by (8.6).

It is natural to view the element in $\mathcal{L}_a^n$ defined by (8.4) as a product of the linear forms $\theta_1,\dots,\theta_n$. We shall denote it by $\theta_1\wedge\dots\wedge\theta_n$, and we shall prove that there is one and only one way in which one can define a bilinear multiplication $\wedge : \mathcal{L}_a^n(V)\times\mathcal{L}_a^m(V)\to L_a^{n+m}(V)$ so that

$$(\theta_1\wedge\dots\wedge\theta_n)\wedge(\sigma_1\wedge\dots\wedge\sigma_m) = \theta_1\wedge\dots\wedge\theta_n\wedge\sigma_1\wedge\dots\wedge\sigma_m,$$

for all $\theta_1,\dots,\theta_n,\sigma_1,\dots,\sigma_m \in V^*$. The uniqueness is clear since this defines the multiplication of basis elements in $\mathcal{L}_a^n$ and $\mathcal{L}_a^m$. Set $f = \theta_1\wedge\dots\wedge\theta_n$ and $g =$

$\sigma_1 \wedge \dots \wedge \sigma_m$. Then

$$(f \wedge g)(v_1, \dots, v_{n+m}) = \det \begin{pmatrix} \langle v_1, \theta_1 \rangle & \dots & \langle v_{n+m}, \theta_1 \rangle \\ \dots & \dots & \dots \\ \langle v_1, \theta_n \rangle & \dots & \langle v_{n+m}, \theta_n \rangle \\ \langle v_1, \sigma_1 \rangle & \dots & \langle v_{n+m}, \sigma_1 \rangle \\ \dots & \dots & \dots \\ \langle v_1, \sigma_m \rangle & \dots & \langle v_{n+m}, \sigma_m \rangle \end{pmatrix}$$

$$= \frac{1}{n!m!} \sum_\pi f(v_{\pi(1)}, \dots, v_{\pi(n)}) g(v_{\pi(n+1)}, \dots v_{\pi(m+n)}) \operatorname{sgn} \pi,$$

where π varies over all permutations of $1, \dots, n+m$. The division by $n!m!$ is required since we sum over ordered groups of columns and not over subsets of columns as in the standard Laplace expansion of determinants. Alternatively we could sum over permutations with $\pi(1) < \dots < \pi(n)$ and $\pi(n+1) < \dots < \pi(n+m)$ and avoid this division. Thus multiplication of general $f \in \mathcal{L}_a^n$ and $g \in \mathcal{L}_a^m$ must be defined by

$$(f \wedge g)(v_1, \dots, v_{n+m}) = \frac{1}{n!m!} \sum_\pi f(v_{\pi(1)}, \dots, v_{\pi(n)}) g(v_{\pi(n+1)}, \dots, v_{\pi(n+m)}) \operatorname{sgn} \pi.$$

The multiplication becomes associative since this is true when one just multiplies basis elements of the form (8.4).

Since $\mathcal{L}_a^n(V)$ is spanned by wedge products (exterior products) of elements in V^* we shall write

$$\mathcal{L}_a^n(V) = \wedge^n V^*.$$

The direct sum

$$\bigoplus_{n=0}^{\dim V} \wedge^n V^*$$

is therefore an algebra of dimension $\sum_0^{\dim V} \binom{\dim V}{n} = 2^{\dim V}$. It is called the *exterior algebra* over V^*. It is not commutative, for we have

$$f \wedge g = (-1)^{nm} g \wedge f, \quad \text{if } f \in \wedge^n V^*, \ g \in \wedge^m V^*.$$

We shall now return to (7.7) which requires us to study elements in the exterior algebra which may depend on a variable point. If $\Omega \subset V$ is open and k is an integer ≥ 0, we shall consider the space $C^k(\Omega, \wedge^n V^*)$. If $x_1, \dots, x_N$ is a coordinate system in V, thus $x_j = \langle x, \theta_j \rangle$ where θ_j is a basis in V^*, then we can write every element $u \in C^k(\Omega, \wedge^n V^*)$ in the form

$$x \mapsto \sum_I{}' u_I(x) \theta_{I_1} \wedge \dots \wedge \theta_{I_n},$$

where $u_I \in C^k(\Omega)$ and $I_1 < \dots < I_n$ in the sum. Now $\theta_j = dx_j$ so the sum can also be written in the conventional form

$$\sum_I{}' u_I(x) dx_{I_1} \wedge \dots \wedge dx_{I_n}.$$

This is called an *exterior differential form.* For the interpretation as n linear form one must of course always use (8.4).

If V_1 is another vector space and $\Omega_1 \subset V_1$ is an open set, $f \in C^{k+1}(\Omega_1, V)$ and $f(\Omega_1) \subset \Omega$, then f' is a map $\Omega_1 \times V_1 \to \Omega \times V$ which is linear along V_1. Since an element in $C^k(\Omega, \wedge^n V^*)$ can be viewed as a C^k map from $\Omega \times V \times \cdots \times V$ to $\mathbf{R}$ which is multilinear and alternating in the arguments from V, it is clear that composition with f' gives a map

$$f^* : C^k(\Omega, \wedge^n V^*) \to C^k(\Omega_1, \wedge^n V_1^*);$$

explicitly it is given by

$$(f^*u)(x; v_1, \dots, v_n) = u(f(x); f'(x)v_1, \dots, f'(x)v_n), \quad x \in \Omega_1,\ v_j \in V_1.$$

(The notation f'^* might be more adequate but the traditional notation is just f^*.) In particular, if u is an n form on Ω and $\psi \in C^{k+1}(\mathbf{R}^n, \Omega)$ then

$$\psi^* u = u(\psi(t); \partial\psi/\partial t_1, \dots, \partial\psi/\partial t_n)\, dt_1 \wedge \cdots \wedge dt_n.$$

It is clear that $f^*(u \wedge v) = (f^*u) \wedge (f^*v)$ and that $(f \circ g)^* = g^* \circ f^*$ when the composition is defined. The invariance of the differential implies that

$$f^* du = df^* u$$

when u is a 0 form, that is, a smooth function.

We shall now define the *differential* of a differential form. To simplify notation we just consider C^∞ forms and write $\lambda^n(\Omega) = C^\infty(\Omega, \wedge^n V^*)$. The differential of functions can be viewed as a linear map $d : \lambda^0(\Omega) \to \lambda^1(\Omega)$. If $f : \Omega_1 \to \Omega$ is a C^∞ map from an open set Ω_1 in another vector space V_1, the chain rule (the invariance of the differential) gives as already mentioned a commutative diagram

$$\begin{array}{ccc} \lambda^0(\Omega) & \xrightarrow{\ d\ } & \lambda^1(\Omega) \\ {\scriptstyle f^*}\big\downarrow & & {\scriptstyle f^*}\big\downarrow \\ \lambda^0(\Omega_1) & \xrightarrow{\ d\ } & \lambda^1(\Omega_1) \end{array} .$$

We shall now prove that it is possible to extend the definition of d to a linear operator $\lambda^n(\Omega) \to \lambda^{n+1}(\Omega)$ for all Ω and n so that the diagram

$$\begin{array}{ccccccc} \lambda^0(\Omega) & \xrightarrow{\ d\ } & \lambda^1(\Omega) & \xrightarrow{\ d\ } & \lambda^2(\Omega) & \xrightarrow{\ d\ } & \dots \\ {\scriptstyle f^*}\big\downarrow & & {\scriptstyle f^*}\big\downarrow & & {\scriptstyle f^*}\big\downarrow & & \\ \lambda^0(\Omega_1) & \xrightarrow{\ d\ } & \lambda^1(\Omega_1) & \xrightarrow{\ d\ } & \lambda^2(\Omega_1) & \xrightarrow{\ d\ } & \dots \end{array} \tag{8.7}$$

is always commutative. Moreover we shall prove that this determines d uniquely apart from a constant factor, depending on n, which is fixed if in addition we require Leibniz' rule

$$d(u \wedge v) = (du) \wedge v + u \wedge dv, \quad \text{if } u \in \lambda^0(\Omega),\ v \in \lambda^n(\Omega). \tag{8.8}$$

Finally we shall prove that

(8.9) $\quad d^2u = 0, \quad d(u \wedge v) = (du) \wedge v + (-1)^n u \wedge dv, \quad \text{if } u \in \lambda^n(\Omega), \ v \in \lambda^m(\Omega).$

Assume at first only that d is defined and that (8.7) is always commutative. Then it follows that

(i) $d(dx_1 \wedge \cdots \wedge dx_n) = 0$ in $\mathbf{R}^n$.

(ii) $d(x_1 dx_2 \wedge \cdots \wedge dx_{n+1}) = C_n dx_1 \wedge \cdots \wedge dx_{n+1}$ in $\mathbf{R}^{n+1}$.

(i) is obvious since every $n+1$ form in $\mathbf{R}^n$ is equal to 0. To prove (ii) we note that

$$d(x_1 dx_1 \wedge \cdots \wedge dx_{n+1}) = u(x)\, dx_1 \wedge \cdots \wedge dx_{n+1}$$

with $u \in C^\infty(\mathbf{R}^{n+1})$, for every $n+1$ form in $\mathbf{R}^{n+1}$ can be written in this way. If $\tau_a : x \mapsto x + a$ is a translation in $\mathbf{R}^{n+1}$, we get from (8.7) and (i) that

$$u(x+a) dx_1 \wedge \cdots \wedge dx_{n+1} = d((x_1 + a) dx_2 \wedge \cdots \wedge dx_{n+1}) = d(x_1 dx_2 \wedge \cdots \wedge dx_{n+1}).$$

Hence $u(x+a) = u(x)$, that is, $u(x) = C_n$. If (8.8) holds then $C_n = 1$.

If now $u \in \lambda^n(\Omega)$ where Ω is an open subset of V, and $x_1, \ldots, x_N$ is a coordinate system in V, we can write

$$u(x) = \sum_I u_I(x) dx_{I_1} \wedge \cdots \wedge dx_{I_n}$$

where $u \in C^\infty(\Omega)$. If we apply the commutative diagram (8.7) to the map $\Omega \ni x \mapsto (u_I(x), x_{I_1}, \ldots, x_{I_n}) \in \mathbf{R}^{n+1}$, it follows from (ii) that

$$d(u_I dx_{I_1} \wedge \cdots \wedge dx_{I_n}) = C_n du_I \wedge dx_{I_1} \wedge \cdots \wedge dx_{I_n},$$

and if $C_n = 1$ we obtain

$$du = \sum_I du_I \wedge dx_{I_1} \wedge \cdots \wedge dx_{I_n}. \tag{8.10}$$

This proves the uniqueness which we have claimed.

Let us now with a fixed coordinate system *define* du by (8.10). Then we have

$$d^2u = d \sum_{I,j} \partial u_I / \partial x_j dx_j \wedge dx_{I_1} \wedge \cdots \wedge dx_{I_n}$$
$$= \sum_{I,j,k} \partial^2 u_I / \partial x_j \partial x_k dx_k \wedge dx_j \wedge dx_{I_1} \wedge \cdots \wedge dx_{I_n} = 0,$$

for $\partial^2 u_I / \partial x_j \partial x_k = \partial^2 u_I / \partial x_k \partial x_j$. If $v = \sum_J v_J dx_{J_1} \wedge \cdots \wedge dx_{J_m}$ then it follows from the elementary product rule that (8.9) is valid, for moving the differential dv_J to the right of the differentials $dx_{I_1}, \ldots, dx_{I_n}$ introduces a factor $(-1)^n$. But conversely (8.9) implies that du must be given by (8.10). Since (8.9) does not depend on the choice of coordinates we conclude that (8.10) is also independent of how they are chosen. We also see from (8.9) that (8.10) is valid for arbitrary C^∞

functions x_i, whether they form a part of a coordinate system or not. To prove the commutativity of (8.7) it is now sufficient to note that

$$f^*u = \sum_I f^*u_I df_{I_1} \wedge \cdots \wedge df_{I_n},$$

so it follows from the chain rule that

$$df^*u = \sum_I (f^*du_I) \wedge df_{I_1} \wedge \cdots \wedge df_{I_n} = f^*du.$$

Summing up, we have now defined a first order differential operator d from n forms to $n+1$ forms and proved that (8.9) is valid and that (8.7) is commutative for every C^∞ map f. Of course it suffices to assume that f and u have one or two continuous derivatives, for this is all that these identities involve.

The first property in (8.9) gives important information about the equation

$$du = v, \tag{8.11}$$

where v is given in $\lambda^{n+1}(\Omega)$ and the unknown u is in $\lambda^n(\Omega)$: the equation $dv = 0$ is a *necessary* condition for solvability. When $n = 0$, for example, then (8.11) is equivalent to the system $\partial u/\partial x_j = v_j$, $j = 1, \ldots, N$, and $dv = 0$ is equivalent to $\partial v_j/\partial x_k - \partial v_k/\partial x_j = 0$ for $j, k = 1, \ldots, N$. We shall prove that the converse is true locally. Later on we shall see that the question whether the condition $dv = 0$ is sufficient also globally leads to basic notions of algebraic topology.

Theorem 8.1 (Poincaré's lemma). *Let v be a $n+1$ form with C^k coefficients, $k \geq 1$, in the open convex set $\Omega \subset V$. If $dv = 0$ then there exists an n form u in Ω with C^k coefficients such that $du = v$ in Ω.*

Proof. We may assume that $0 \in \Omega$. Then it follows that

$$\widetilde{\Omega} = \{(x,t) \in V \times \mathbf{R}; tx \in \Omega\}$$

is an open neighborhood of $\Omega \times [0,1]$. With $f(x,t) = f_t(x) = tx$ we form

$$f^*v = f_t^*v + dt \wedge w_t$$

where f_t^* is defined by regarding t as a parameter so that f_t^*v just as w_t is a differential form which contains no factor dt but only differentials of the coordinates in Ω. Since $df^*v = f^*dv = 0$ we have

$$0 = dt \wedge (\partial(f_t^*v)/\partial t - d_x w_t) + R$$

where R is a form which does not contain dt and $d_x w_t$ denotes the differential of w when t is regarded as a parameter. Hence

$$\partial(f_t^*v)/\partial t = d_x w_t,$$

and integration from $t = 0$ to $t = 1$ gives now

$$f_1^*v - f_0^*v = du, \quad u = \int_0^1 w_t\, dt.$$

But f_1 is the identity and f_0 maps Ω to 0 so this means that $v = du$ and the theorem is proved.

In this section we started from the problem of finding objects which allow integration over n dimensional submanifolds of a vector space V. If we look back at the arguments we are led to define the integral of a differential form $u = u(x)dx_1\wedge\cdots\wedge dx_n$ in $\mathbf{R}^n$ by

$$\int u = \int u(x)\, dx$$

where the integral in the right-hand side is the standard Riemann (or Lebesgue) integral. For an n form in V and a C^1 local parametrisation $\mathbf{R}^n \supset \omega \ni t \mapsto \psi(t) \in V$ of an n dimensional submanifold F of V we define

$$\int_F u = \int_\omega \psi^* u$$

when $F \operatorname{supp} u$ is a compact subset of $\psi(\omega)$. If $\omega_1 \ni s \mapsto \psi_1(s)$ is another parametrisation in a neighborhood of $F \cap \operatorname{supp} u$ and if $\det(\psi^{-1}\circ\psi_1)' > 0$, then the definition is not changed if we replace ψ by ψ_1, but the sign is changed if $\det(\psi^{-1}\circ\psi)' < 0$. If F is *oriented*, that is, F is provided with a system of parametrisations ψ_j covering F such that $\psi_j^{-1}\psi_k$ has a positive functional determinant for arbitrary j, k where it is defined, then the discussion in Chapter VII shows that $\int_F u$ can be uniquely defined for every n form u such that $F \cap \operatorname{supp} u$ is compact, if we only use parametrisations ψ with positive functional determinant for $\psi^{-1}\psi_j$, for every j. We shall return to this point and add greater precision and generality after introducing the general notion of a manifold.

As an application of the calculus of differential forms we shall end this section by giving an alternative derivation of the basic properties of the degree of mapping in Chapter IV.

Theorem 8.2. *Let $f \in C^\infty(\mathbf{R}^n, \mathbf{R}^n)$, and let Ω be an open bounded subset of $\mathbf{R}^n$. If $y \notin f(\partial\Omega)$ is a regular value for f and O is the component of y in $\complement f(\partial\Omega)$, then*

$$\int_\Omega f^* u = d(f, \Omega, y) \int_O u, \quad \text{if } u \in \lambda_0^n(O), \tag{8.12}$$

that is, u is an n form with compact support in O. Here $d(f, \Omega, y)$ is defined by (4.1).

In particular it follows that $d(f, \Omega, y)$ is independent of $y \in \Omega$, so we have a new proof of property (i) of d, which was an essential first step in the definition of the degree.

Proof. If $u = u(x)dx_1 \wedge \cdots \wedge dx_n$ is a continuous n form with support in a sufficiently small neighborhood of y, then $\Omega \cap \operatorname{supp} f^* u$ is contained in neighborhoods of the finitely many points $x \in \Omega$ with $f(x) = y$ where f is a diffeomorphism. By Theorem 7.1 we therefore obtain

$$\int_\Omega f^* u = \int_\Omega u(f(x)) \det f'(x)\, dx = d(f, \Omega, y) \int u(x)\, dx = d(f, \Omega, y) \int u.$$

Thus there are n forms with $\int u \neq 0$ for which (8.12) is valid.

In the general proof of (8.12) we shall use that if v is an $n-1$ form in C^1 with compact support and $dv = u$, then

(8.13)
$$\int u = 0.$$

The proof is obvious: we have

$$v = \sum_1^n v_j(x) dx_1 \wedge \cdots \wedge \widehat{dx_j} \wedge \cdots \wedge dx_n$$

where $\widehat{dx_j}$ means that dx_j should be omitted, and this gives

$$u = dv = \sum_1^n (-1)^{j-1} \partial v_j / \partial x_j dx_1 \wedge \cdots \wedge dx_n$$

so (8.13) follows immediately by integration. ((8.13) is a special case of Stokes' formula which will be discussed later on. Conversely, there is an addition to Theorem 8.1 (see Theorem 12.2 below) which states that (8.13) is sufficient for the existence of v with the stated properties. However, we shall at this time avoid the complete proof of this fact.) Since $dv = u$ implies $f^*u = f^*dv = d(f^*v)$ and $\operatorname{supp} f^*v \subset f^{-1}(\operatorname{supp} v)$, which is a closed set which does not intersect $\partial\Omega$, we also obtain

(8.14)
$$\int_\Omega f^*u \, dx = 0.$$

Let $\varphi \in C_0^1(\mathbf{R}^n)$ have support in a small neighborhood of 0, and assume that $0 \leq \varphi$ and $\int \varphi \, dx = 1$. Set for $a \in \mathbf{R}^n$

$$\varphi_a = \varphi(x-a) dx_1 \wedge \cdots \wedge dx_n.$$

Then $\int \varphi_a = 1$, and $\int_\Omega f^* \varphi_a$ is independent of a in $\{a; \operatorname{supp} \varphi_a \subset \complement f(\partial\Omega)\}$. In fact, the derivative of φ_a with respect to a_j is

$$-\partial\varphi(x-a)/\partial x_j dx_1 \wedge \cdots \wedge dx_n = (-1)^j d((\varphi(x-a) dx_1 \wedge \ldots \widehat{dx_j} \wedge \cdots \wedge dx_n),$$

so it follows that

$$\int \partial\varphi_a / \partial a_j = 0, \quad \int_\Omega f^* \partial\varphi_a / \partial a_j = 0.$$

If K is a compact subset of O and $\operatorname{supp} \varphi$ is sufficiently close to the origin we therefore obtain

$$\int f^* \varphi_a = d(f, \Omega, y), \quad a \in K.$$

If u is a continuous function with support in K we obtain by multiplication with $u(a)$ and integration that

$$\int_\Omega f^* \Big(\int u_a \varphi_a \, da \Big) = d(f, \Omega, y) \int u(a) \, da.$$

Here

$$\int u(a)\varphi_a\,da = \Big(\int u(a)\varphi(x-a)\,da\Big)dx_1\wedge\cdots\wedge dx_n.$$

When the support of φ tends to $\{0\}$ the integral in the right-hand side converges uniformly to $u(x)$. This completes the general proof of (8.12).

Remark. The deeper reason behind this proof is the homotopy invariance which will be proved in Chapter 11 (see Theorem 11.4).

Lemma 4.1 is an immediate consequence of Theorem 8.2. In fact, if we choose u with $\int u\,dx = 1$ and support in the component of y in $\complement F(\partial\Omega\times[0,1])$ then

$$d(f_t,\Omega,y) = \int_\Omega f_t^* u$$

where $f_t(x) = F(x,t)$. Here the left-hand side is an integer and the right-hand side is a continuous function of $t\in[0,1]$, so it must be a constant. The geometrical arguments in Lemma 4.1 have now been completely replaced by analytical proofs.

CHAPTER IX

MANIFOLDS AND VECTOR BUNDLES, STOKES' FORMULA

So far we have only considered submanifolds of a finite dimensional vector space. According to Definition 2.2 they are characterized by the existence of parametrisations. By isolating the properties of parametrisations which are independent of those of the embedding space we shall now define the general notion of manifold.

DEFINITION 9.1. A C^k atlas, $k \geq 1$, in a topological space X is a countable covering of X with open subsets X_i and for every i a homeomorphism $\psi_i : X_i' \to X_i$ where X_i' is open in $\mathbf{R}^n$, such that

$$\psi_i^{-1} \circ \psi_j \in C^k(\psi_j^{-1}(X_i \cap X_j), \mathbf{R}^n) \quad \text{for all } i \text{ and } j. \tag{9.1}$$

If U is open in X, then a function $f : U \to \mathbf{R}$ is said to be in $C^k(U)$ if

$$f \circ \psi_i \in C^k(\psi_i^{-1}(U \cap X_i)) \quad \text{for every } i.$$

Two different atlases give rise to the same definition of $C^k(U, \mathbf{R})$ for every U if and only if their union is an atlas. Then they are said to be equivalent, and an equivalence class of atlases is called a C^k structure on X. Their union is the maximal atlas defining the structure.

The difference between this definition and the definition of a C^k submanifold of $\mathbf{R}^N$ is that now it is *a priori* meaningless to talk about differentiability of the parametrisations ψ_i. However, it is possible to keep the condition that the maps $\psi_i^{-1}\psi_j$ comparing different parametrisations shall be in C^k. Later on we shall prove that every manifold can be realised as a manifold embedded in some $\mathbf{R}^N$, so the general notion of manifold is not really a generalisation but consists rather in disregarding the embedding.

Every C^k structure defines a C^j structure when $j \leq k$. For the sake of simplicity we shall usually consider C^∞ structures. The corresponding manifold is then said to be a(n infinitely) differentiable manifold.

If X and Y are two differentiable manifolds then we can define $C^k(X, Y)$ as the set of maps $f : X \to Y$ such that $\psi_j^{-1} \circ f \circ \varphi_i \in C^k$ where it is defined, ψ_j and φ_i denoting the parametrices in an atlas of Y and one of X.

The tangent space of an embedded manifold was defined in Chapter II as the image of $\mathbf{R}^n$ under the differential of the parametrisation. The definition has to be modified for a general manifold since we only have the differentials of the maps (9.1) available while the differential of ψ_i has no sense as yet. Denote the map (9.1) by ψ_{ij}. We can regard ψ_{ij}' as a C^∞ map

$$\psi_j^{-1}(X_i \cap X_j) \times \mathbf{R}^n \to \psi_i^{-1}(X_i \cap X_j) \times \mathbf{R}^n$$

which is linear in the variable in $\mathbf{R}^n$. Equivalently we can regard this as a map

$$g_{ij} : (X_i \cap X_j) \times \mathbf{R}^n \ni (x,t) \to (x, \psi'_{ij}(\psi_j^{-1}(x))t) \in (X_i \cap X_j) \times \mathbf{R}^n.$$

Since $\psi_{ik} = \psi_{ij} \circ \psi_{jk}$ where the right-hand side is defined, we have by the chain rule $\psi'_{ik} = \psi'_{ij} \circ \psi'_{jk}$, or more precisely

$$g_{ik} = g_{ij} g_{jk} \quad \text{in } (X_i \cap X_j \cap X_k) \times \mathbf{R}^n.$$

We can therefore collect the parametrised vector spaces $X_i \times \mathbf{R}^n$ to a *vector bundle* in the following way: In the set $\{(i,x,v); x \in X_i, v \in \mathbf{R}^n\}$ we set

$$(i,x,v) \sim (j,y,w) \quad \text{if } x = y \in X_i \cap X_j \text{ and } v = g_{ij}(x)w.$$

This is an equivalence relation. The set E of equivalence classes obtains a C^∞ structure from the atlas $X'_i \times \mathbf{R}^n \to X_i \times \mathbf{R}^n \to E$. By the definition of the equivalence relation we also have a C^∞ map $p : E \to X$, for equivalence of two triples requires that the two elements in X which occur are equal. Moreover, $E_x = p^{-1}(x)$ is for every $x \in X$ a vector space of dimension n, for all g_{ij} are linear in the $\mathbf{R}^n$ variables. One calls E the *tangent bundle* of X, and one writes $E = T(X)$, $E_x = T_x(X)$. This is a vector bundle in the sense of the following general definition:

Definition 9.2. If X is a C^∞ manifold then a C^∞ vector bundle over X with fiber of type $\mathbf{R}^n$ is a C^∞ manifold E with a C^∞ map $p : E \to X$ such that

(i) $p^{-1}(x) = E_x$ has a structure of n dimensional vector space over $\mathbf{R}$ if $x \in X$,
(ii) every $x \in X$ has a neighborhood U such that there is a diffeomorphism $p^{-1}(U) \to U \times \mathbf{R}^n$ which is linear on each fiber and commutes with the projections.

For an embedded manifold X one can of course identify the tangent bundle $T(X)$ just defined with the tangent space introduced in Chapter II. We shall also give an alternative definition of $T(X)$ which only depends on the class of C^∞ functions on X; this will confirm that the construction gives equivalent results for equivalent atlases. The new definition depends on the observation that when $U \subset \mathbf{R}^n$ is open and $x \in U$, $v \in \mathbf{R}^n$, then the map

$$C^\infty(U) \ni \varphi \mapsto \langle v, d\varphi \rangle = \varphi'(v) = L(\varphi) \tag{9.2}$$

is a linear form in φ such that

$$L(\varphi\psi) = \varphi(x)L(\psi) + L(\varphi)\psi(x). \tag{9.3}$$

Conversely, every linear form L satisfying (9.3) is of the form (9.2). In fact, by (9.3) we have $L(1) = 2L(1)$, hence $L(1) = 0$, and $L(\varphi) = 0$ if φ vanishes of second order at x, for then we can use Taylor's formula to write

$$\varphi(y) = \sum_{j,k=1}^{n} (y_j - x_j)(y_k - x_k)\varphi_{jk}(y)$$

where $\varphi_{jk} \in C^\infty$, and it follows from (9.3) that $L(\chi) = 0$ if χ is the product of two C^∞ functions vanishing at x. Hence $L(\varphi) = L(\varphi_1)$ if $\varphi_1(y) = \sum_1^n (y_j - x_j)\partial\varphi(x)/\partial x_j$, which proves the claim. The same conclusion is valid if we have a *derivation*, that is, a map satisfying (9.3), which is only defined on $C_0^\infty(U)$. The notion of derivation is also well defined at a point x in a manifold M, and since $C^\infty(M) \supset C_0^\infty(U)$ if U is an open neighborhood of $x \in M$, it follows that we can identify $T(X)$ with the space of all derivations; $T_x(X)$ is the space of derivations at x.

Starting from any one of these definitions of $T(X)$ it is also clear that if $f : X \to Y$ is a C^1 map then the differential f' maps $T(X)$ to $T(Y)$, and $f'(x)$ maps $T_x(X)$ linearly to $T_{f(x)}(Y)$ for every $x \in X$. Here the differential f' is defined using local coordinates as in Chapter I.

By a C^∞ *section* of a vector bundle E over X one means a C^∞ map $s : X \to E$ such that $p \circ s$ is the identity in X, that is, $s_x = s(x) \in E_x$ for every x. Let in particular s be a section of $T(X)$. If $u \in C^\infty(X)$ then $s \mapsto s_x u(x)$ is by our second interpretation of $T(X)$ a function defined in X. In local coordinates of an atlas s corresponds to functions $s^i \in C^\infty(X_i, \mathbf{R}^n)$, and if $u \circ \psi_i = u_i \in C^\infty(X_i')$ then

$$s_{\psi^i} u = u_i'(s^i) \quad \text{in } X_i', \tag{9.4}$$

which proves that $x \mapsto s_x u$ is in $C^\infty(X)$. That the right-hand side of (9.4) is a function in X expressed in the local coordinates is also clear since $u_j = u_i \circ \psi_{ij}$ in $\psi_j^{-1}(X_i \cap X_j)$, and

$$u_j'(s^j) = u_i'(\psi_{ij})(\psi_{ij}' s^j) = u_i'(s^i) \circ \psi_{ij}.$$

We can now also define a vector bundle $T^*(X)$ with fibers $T_x^*(X)$ equal to the dual spaces of the fibers $T_x(X)$ of $T(X)$. To do so we just replace the maps g_{ij} above by the maps $({}^t g_{ij})^{-1}$, that is, more explicitly

$$(X_i \cap X_j) \times \mathbf{R}^n \ni (x, \xi) \mapsto (x, {}^t\psi_{ji}'(\psi_i(x))\xi) \in (X_i \cap X_j) \times \mathbf{R}^n.$$

We leave as an exercise to prove the necessary transitivity conditions and to prove that the scalar product in $\mathbf{R}^n$ defines invariantly a scalar product of elements in $T_x(X)$ and $T_x^*(X)$. If $u \in C^\infty(X)$ then du becomes a section of $T^*(X)$. More generally, we define the vector bundle $\lambda^k T^*(X)$ for $k = 0, \dots, n$ by replacing $\mathbf{R}^n$ with $\wedge^k \mathbf{R}^n$ above. The sections of $\lambda^k T^*(X)$ are then in every local coordinate system precisely the differential forms in the local coordinates. Thus a section u of $\lambda^k T^*(X)$ corresponds to a differential form u^i in X_i' for every i such that $\psi_{ij}^* u_i = u_j$ in $\psi_j^{-1}(X_i \cap X_j)$ for all i, j. Conversely, such differential forms in X_i' define uniquely a differential form in X. (The advantage of the notion of vector bundle is essentially that we do not have to write down these transformation rules explicitly all the time.) The exterior differential d maps C^∞ sections of $\lambda^k(T^*(X))$ to sections of $\lambda^{k+1}(T^*(X))$. (Analogous statements are of course true for manifolds which are differentiable of finite order but we leave for the reader to consider how they should be formulated.)

We shall now return to the definition of integration of a differential forms on a manifold. First we need another definition:

DEFINITION 9.3. A manifold is called *oriented* if it is provided with an atlas such that all the maps (9.1) have a positive functional determinant (that is, preserve the orientation).

If X is an oriented manifold of dimension n, u is a continuous n form in X, and K is a compact subset of X, then the discussion in Chapters VII and VIII proves that $\int_K u$ is uniquely defined.

An orientation of a manifold X of dimension n is often specified by giving a continuous n form a on X which is $\neq 0$ at every point. If $\psi_i : X_i' \to X_i$ belongs to an atlas and X_i' is connected, then $\psi_i^* a$ is either positive in X_i' or else negative in X_i'. (We say that a n form f in an open subset of $\mathbf{R}^n$ is positive if it can be written $f(x)dx_1 \wedge \cdots \wedge dx_n$ with $f > 0$.) In the second case we replace ψ_i by the composition with the reflection in the plane $x_1 = 0$, for example. This gives a new atlas such that $\psi_i^* a$ is always positive. But then it defines an orientation for X; one says that X *is oriented by* $a > 0$. Conversely, if X is an oriented manifold it is easy to use a partition of unity to construct a n form a which is nowhere 0 so that the same orientation is defined by a.

A *submanifold* Y of a manifold is defined precisely as we defined a submanifold of $\mathbf{R}^N$. It is clear that Y itself is then a manifold. The embedding $i : Y \to X$ is injective and has injective differential, and the range is locally closed, that is, it has a closed intersection with some neighborhood of every point in Y. If Y is oriented and of dimension n and u is an n form in X, we can now for a compact subset K of Y define

$$\int_K u = \int_K i^* u.$$

If u is a k form with $k \neq n$ we define $\int_K u = 0$.

We can now state and prove Stokes' formula. Let X be a manifold and Y an oriented submanifold of dimension n. We assume given in Y a C^∞ function χ such that $\chi = 0$ implies that $d\chi \neq 0$, and $Y_- = \{u \in Y; \chi(y) \leq 0\}$ is compact. Then the boundary $\partial Y_- = \{y \in Y; \chi(y) = 0\}$ is also a C^1 manifold which can be oriented as follows: For every $y_0 \in \partial Y_-$ there is a neighborhood where we can find a parametrisation $\mathbf{R}^n \supset \omega \ni \psi(t) \mapsto \psi(t)$ of Y such that $\chi(\psi(t)) = t_1$. This is attained by taking an arbitrary parametrisation and replacing a suitable coordinate with χ, making sure that the functional determinant of the change of variables is not 0 at the point corresponding to y_0. To avoid a trivial exceptional case we assume that $n > 1$. If we choose ψ so that ψ is in the maximal atlas belonging to the orientation of Y, then the map $t' = (t_2, \ldots, t_n) \mapsto \psi(0, t')$ belongs to the mnaximal atlas defining ∂Y_-. If ψ_1 is another parametrisation with $\chi(\psi_1(s)) = s_1$ then

$$\psi_1^{-1} \circ \psi(t) = (t_1, g(t))$$

where $g(t) \in \mathbf{R}^{n-1}$ and $\det \partial g/\partial t' \neq 0$. Thus the maps $t' \mapsto \psi(0, t')$ define an atlas for ∂Y_- which orients ∂Y_-. The preceding calculation also proves that the orientation is independent of the choice of χ. Informally we can say that ∂Y_- is oriented by choosing positively oriented coordinate systems for Y which begin with χ and restricting them to ∂Y_-. We can now prove

THEOREM 9.4 (STOKES' FORMULA). *If Y is an oriented submanifold of X and*

u is a C^1 n form in a neighborhood of Y_- in X then

$$\int_{\partial Y_-} u = \int_{Y_-} du. \tag{9.5}$$

PROOF. By a partition of unity the proof is reduced to the case where the intersection of Y and $\operatorname{supp} u$ is contained in a coordinate patch of the form used to orient ∂Y_1 or a coordinate patch with closure contained in $Y_- \setminus \partial Y_-$. Let $v = \psi^* u$ where ψ is the corresponding parametrisation. Since $\psi^* du = d(\psi^* u) = dv$ we only have to prove that if v is a C^1 form of degree $n-1$ in $\mathbf{R}^n$ with compact support then

$$\int_{\mathbf{R}^n} dv = 0, \quad \int_{\{x \in \mathbf{R}^n; x_1 < 0\}} dv = \int_{\{x \in \mathbf{R}^n; x_1 = 0\}} v.$$

Writing

$$v = \sum_1^n (-1)^{j-1} v_j(x) dx_1 \wedge \cdots \wedge \widehat{dx_j} \wedge \cdots \wedge dx_n$$

we have

$$dv = \sum_1^n \partial v_j / \partial x_j dx_1 \wedge \cdots \wedge dx_n.$$

The integral of $\partial v_j/\partial x_j$ with respect to x_j over $\mathbf{R}$ is equal to 0, and the integral of $\partial v_1/\partial x_1$ when $x_1 < 0$ is equal to $v_1(0, \cdot)$, so we obtain

$$\int_{x_1<0} dv = \int v_1(0, x_2, \dots, x_n)\, dx_2 \dots dx_n = \int_{x_1=0} v$$

with the induced orientation. This completes the proof.

We shall now discuss some special cases. First we take $X = Y = \mathbf{R}^n$ and let $\Omega \subset \mathbf{R}^n$ be open and bounded with $\partial\Omega \in C^\infty$. Then the preceding calculation gives if v as above is a C^1 form of degree $n-1$ in a neighborhood of $\overline{\Omega}$

$$\int_\Omega \sum_1^n \partial v_j / \partial x_j \, dx = \int_{\partial\Omega} v.$$

To give the right-hand side a more familiar look we assume again that the support of v is contained in a neighborhood of a boundary point where we can introduce local parameters $x = x(t_1, \dots, t_n)$ with $t_1 = \chi(x)$. If we set $t' = (t_2, \dots, t_n)$ and $\psi(t') = x(0, t')$ then

$$\int_{\partial\Omega} v = \int \sum_1^n v_j(\psi(t')) N_j(t')\, dt', \quad N_j = (-1)^{j-1} \det(\partial\psi_i/\partial t_k)_{i=1,\dots,\hat{j},\dots,n; k=2,\dots,n}.$$

The definition of N means that

$$\langle N, w \rangle = \det(w, \partial\psi/\partial t_2, \dots, \partial\psi/\partial t_n), \quad w \in \mathbf{R}^n,$$

which proves that N is a Euclidean normal of $\partial\Omega$. Taking $w = N$ we obtain $\|N\|^2 = \det(N, \partial\psi/\partial t_2, \ldots, \partial\psi/\partial t_n)$, hence

$$\|N\|^4 = \det\begin{pmatrix} \langle N, N\rangle & 0 \\ 0 & (\langle\partial\psi/\partial t_j, \partial\psi/\partial t_k\rangle)\end{pmatrix} = \|N\|^2 \det({}^t\psi'(t')\psi'(t')),$$

which means that

$$\|N\|^2 = \det({}^t\psi'(t')\psi'(t')). \tag{9.6}$$

To determine the direction of N we note that since t is a positively oriented coordinate system we have

$$0 < \det(\partial x/\partial t) = \sum_1^n N_j \partial x_j/\partial t_1$$

and that differentiation of the equation $t_1 = \chi(x)$ with respect to t_1 gives

$$1 = \sum_1^n \partial\chi/\partial x_j \partial x_j/\partial t_1.$$

Since $\operatorname{grad}\chi$ is also a normal vector of $\partial\Omega$ it has also the same direction as N, which proves that N has the direction of the exterior normal. If $\nu = N/\|N\|$ denotes the exterior unit normal we obtain

$$\int_{\partial\Omega} v = \int \sum_1^n v_j\nu_j \sqrt{\det({}^t\psi'(t')\psi'(t'))}\, dt'.$$

Recalling the definition of the Euclidean surface measure $d\sigma$ in Chapter VII we have now proved that Stokes' formula contains the Gauss-Green formula

$$\int_\Omega \operatorname{div} v\, dx = \int_\Omega \sum_1^n \partial v_j \partial x_j\, dx = \int_{\partial\Omega} \langle v, \nu\rangle\, d\sigma, \quad v \in C^1(\overline{\Omega}, \mathbf{R}^n). \tag{9.7}$$

When $n = 3$ we shall now also derive the classical Stokes' formula. Let S be an oriented smooth surface in $\mathbf{R}^3$ and let Y be a relatively compact open subset of S with C^∞ boundary ∂Y, which is thus an oriented curve. If $v = \sum_1^3 v_j dx_j$ is a C^1 one form near Y then

$$dv = \sum_1^3 dv_j \wedge dx_j = w_1 dx_2\wedge dx_3 - w_2 dx_1 \wedge dx_3 + w_3 dx_2 \wedge dx_3, \quad \text{where}$$
$$w_1 = \partial v_3/\partial x_2 - \partial v_2/\partial x_3,\ w_2 = \partial v_1/\partial x_3 - \partial v_3/\partial x_1,\ w_3 = \partial v_2/\partial x_1 - \partial v_1/\partial x_2.$$

We can view (w_1, w_2, w_3) as a vector field, denoted by $\operatorname{rot} v$. Then Stokes' formula (9.5) yields

$$\int_{\partial Y} \sum_1^3 v_j\, dx_j = \int_{\partial Y} \langle v, t\rangle\, ds = \int_Y \langle \operatorname{rot} v, \nu\rangle\, d\sigma. \tag{9.8}$$

Here t is the unit tangent of the oriented curve ∂Y, ds is the arc measure on ∂Y, ν is the outer unit normal of S, and $d\sigma$ is the Euclidean surface measure on S. This follows at once from the calculations which led to (9.7). Formula (9.8) is the classical Stokes' formula.

Finally we shall use Stokes' formula to show how the degree of mapping can be expressed in terms of the *Kronecker form*

$$\omega = \tilde{\omega}/(\gamma_n \|x\|^n) \quad \text{in } \mathbf{R}^n \setminus \{0\}.$$

Here γ_n is the area of the Euclidean unit sphere $\{x \in \mathbf{R}^n; \|x\| = 1\}$, and

$$\tilde{\omega} = \sum_1^n (-1)^{j-1} x_j dx_1 \wedge \cdots \wedge \widehat{dx_j} \wedge \cdots \wedge dx_n,$$

so ω is a $n-1$ form in $\mathbf{R}^n \setminus \{0\}$. If g is a C^∞ function then

$$d(g\tilde{\omega}) = \sum_1^n \partial(g(x)x_j)/\partial x_j dx_1 \wedge \cdots \wedge dx_n = (ng(x) + \sum_1^n x_j \partial g(x)/\partial x_j) dx_1 \wedge \cdots \wedge dx_n.$$

When $g(x) = 1/(\gamma_n \|x\|^n)$ we conclude that $d\omega = 0$ in $\mathbf{R}^n \setminus \{0\}$. When $g = 1/\gamma_n$ we obtain by Stokes' formula that

$$\int_{\|x\|=1} \omega = \frac{1}{\gamma_n} \int_{\|x\|=1} \tilde{\omega} = \frac{n}{\gamma_n} \int_{\|x\| \le 1} dx_1 \wedge \cdots \wedge dx_n = 1. \tag{9.9}$$

More generally let $T : \mathbf{R}^n \to \mathbf{R}^n$ be a linear bijection. Then $T^*\omega$ is closed in $\mathbf{R}^n \setminus 0$, so two applications of Stokes' formula give for $\varepsilon > 0$

$$\begin{aligned}\int_{\|x\|=\varepsilon} T^*\omega &= \int_{\|Tx\|=1} T^*\omega = \int_{\|Tx\|=1} T^*\tilde{\omega}/\gamma_n = \int_{\|Tx\| \le 1} T^* d\tilde{\omega}/\gamma_n \\ &= \det T \frac{n}{\gamma_n} \int_{\|Tx\| \le 1} dx_1 \wedge \cdots \wedge dx_n = \det T/|\det T| = \operatorname{sign} \det T.\end{aligned}$$

It is now easy to prove

THEOREM 9.5. *Let $f \in C^\infty(\mathbf{R}^n, \mathbf{R}^n)$ and let Ω be an open bounded subset of $\mathbf{R}^n$ with smooth boundary. If $0 \notin f(\partial\Omega)$ is a regular value of f, then*

$$\int_{\partial\Omega} f^*\omega = d(f, \Omega, 0). \tag{9.10}$$

PROOF. Since $\overline{\Omega}$ is compact we can then choose a C^∞ function χ in $\mathbf{R}^n$ such that $\chi < 0$ in Ω, $\chi = 0$ and $\operatorname{grad} \chi \ne 0$ on $\partial\Omega$. The orientation of $\partial\Omega$ in (9.10) is defined as in Stokes' formula. Let now $x^1, \dots, x^N$ be the finitely many solutions in Ω of the equation $f(x) = 0$, and let U_i be a ball of small radius ε with center at x_i. Since $d(f^*\omega) = f^*(d\omega) = 0$ in $\overline{\Omega} \setminus (\cup_1^N U_i)$, we obtain from Stokes' formula for small ε

$$\int_{\partial\Omega} f^*\omega = \sum_1^N \int_{\partial U_i} f^*\omega.$$

By Taylor's formula $f(x) = f'(x_i)(x - x_i) + O(\varepsilon^2)$ and $f'(x) = f'(x_i) + O(\varepsilon)$ on ∂U_i, so it follows that

$$\int_{\partial U_i} f^*\omega \to \operatorname{sign}\det f'(x_i) \quad \text{when } \varepsilon \to 0,$$

which proves the theorem.

Remark. In Chapter XI we shall prove that the left-hand side of (9.10) only depends on the homotopy class of f which implies that it can also be defined for continuous maps f. Thus (9.10) holds for such f with $D(f, \Omega, 0)$ in the right-hand side.

CHAPTER X

EMBEDDING OF A MANIFOLD

Immediately after Definition 9.1 we pointed out that every manifold can be realised as a submanifold of $\mathbf{R}^N$ for some N. In this chapter we shall supply the proof of this statement.

Let X be a C^∞ manifold of dimension n. Recall that a map $f : X \to \mathbf{R}^N$ is called *proper* if for every compact set $K \subset \mathbf{R}^N$ the inverse image

$$f^{-1}(K) = \{x \in X; f(x) \in K\}$$

is compact. It is clear that $f(X)$ is closed if f is continuous and proper. Assume now that $f \in C^\infty(X, \mathbf{R}^N)$. One calls f an *immersion* if $f'(x)$ is injective for every $x \in X$. By the implicit function theorem this implies that a neighborhood of x in X is mapped bijectively on a submanifold of $\mathbf{R}^N$ of dimension n. One calls f a proper *embedding* if f is a proper injective immersion. The range $f(X)$ is then a closed C^∞ submanifold of $\mathbf{R}^N$, and $f : X \to f(X)$ is a diffeomorphism which identifies the abstract manifold X with the closed submanifold $f(X)$ of $\mathbf{R}^N$. Our goal in this chapter is to prove the existence of a proper embedding f of X.

Lemma 10.1. *For arbitrary $N \geq 1$ there exist proper maps $f \in C^\infty(X, \mathbf{R}^N)$.*

Proof. We may assume that $N = 1$. Choose countably many parametrisations $\psi_i : X_i' \to X_i \subset X$ where $X_i' \subset \mathbf{R}^n$, and compact sets $K_i \subset X_i$ such that $X = \cup K_i$. (We can repeat each parametrisation in an atlas countably many times and use that every open set in $\mathbf{R}^n$ is the union of countably many compact sets.) Now choose $\chi_i \in C_0^\infty(X_i)$ such that $0 \leq \chi_i \leq 1$ and $\chi_i = 1$ in a neighborhood of K_i. This can be done by taking a function in $C_0^\infty(X_i')$ equal to 1 in a neighborhood of $\psi_i^{-1}(K_i)$, composing it with ψ_i^{-1} and extending it by 0 in $X \setminus X_i$. Now

$$\varphi_j = \chi_j(1 - \chi_{j-1}) \dots (1 - \chi_1) \in C_0^\infty(X), \qquad \operatorname{supp} \varphi_j \subset X_j, \quad \varphi_j = 0 \text{ close to } K_1 \cup \dots \cup K_{j-1}.$$

(We define $\varphi_1 = \chi_1$.) Every point in X has then a neighborhood where $\varphi_j = 0$ except for finitely many j, and $\sum \varphi_j = 1$. Hence

$$f = \sum_1^\infty j\varphi_j \in C^\infty(X, \mathbf{R}),$$

and f has the required properties, for $f(x) \leq M$ implies that $\varphi_j(x) \neq 0$ for some $j \leq M$, hence that $x \in \cup_1^M \operatorname{supp} \varphi_j$.

To construct embeddings we first study the semiglobal problem to embed compact subsets of X. We keep the notation in the preceding proof and write $M_j = K_1 \cup \dots \cup K_j$.

LEMMA 10.2. *For every j one can find an integer N and $g \in C_0^\infty(X, \mathbf{R}^N)$ so that g is an injective immersion in a neighborhood of M_j.*

PROOF. Since $\psi_j : X_j' \to X_j$ is a parametrisation of X_j, the inverse ψ_j^{-1} is an embedding of X_j in $\mathbf{R}^n$. Set

$$g_j = (\chi_j \psi_j^{-1}, \chi_j) \in \mathbf{R}^{n+1} \text{ in } X_j, \quad g_j = 0 \text{ in } X \setminus X_j,$$

where χ_j is defined as in the proof of Lemma 10.1. Then g_j is an injective immersion of a neighborhood of K_j in $\mathbf{R}^{n+1}$, and

$$g = g_1 \oplus \cdots \oplus g_j$$

is an injective immersion of M_j in $\mathbf{R}^{(n+1)j}$.

REMARK. Note that if $K \subset X$ is compact and $K \cap M_j = \emptyset$ then all χ_j can be chosen equal to 0 in K, so one can choose $g = 0$ in K.

The weakness of Lemma 10.2 is that it gives no bound for the dimension N of the embedding space when j increases. To remove this flaw we shall now discuss how one can decrease the dimension of the embedding space. For $a \in \mathbf{R}^{N-1}$ we shall denote by π_a the projection

$$\mathbf{R}^N \ni x \mapsto \pi_a x = (x_1 - a_1 x_N, \dots, x_{N-1} - a_{N-1} x_N),$$

that is, the projection on the hyperplane where $x_N = 0$ along the vector $(a, 1)$.

LEMMA 10.3. *If $f \in C^\infty(X, \mathbf{R}^N)$ is an immersion of the compact set $K \subset X$ and $N > 2n$, then $\pi_a f$ is also an immersion of K except when a belongs to a closed null set in $\mathbf{R}^{N-1}$.*

PROOF. It is sufficient to prove the statement when K is contained in a coordinate patch, for every K can be written as a union of finitely many such compact sets. It is therefore no restriction to assume that $X \subset \mathbf{R}^n$. Let E be the set of all $a \in \mathbf{R}^{N-1}$ such that $\pi_a f$ is not an immersion on K. That $a \in E$ means that for some $x \in K$ and $\lambda \in \mathbf{R}^n$ with $|\lambda| = 1$

$$\sum_{k=1}^{n} \lambda_k (\partial f_j / \partial x_k - a_j \partial f_N / \partial x_k) = 0, \quad j = 1, \dots, N-1. \tag{10.1}$$

These equations define a closed subset of $\{(x, \lambda, a) \in K \times \mathbf{R}^n \times \mathbf{R}^{N-1}; |\lambda| = 1\}$, so the projection E in $\mathbf{R}^{N-1}$ is closed since it is proper. If we set $\mu = \sum_k \lambda_k \partial f_N / \partial x_k$ and $a_N = 1$, then (10.1) can also be written

$$\sum_{k=1}^{n} \lambda_k \partial f_j(x) / \partial x_k = \mu a_j, \quad j = 1, \dots, N, \tag{10.2}$$

which means that $(a, 1)$ is a tangent of $f(X)$ at $f(x)$. Since f is an immersion it follows from (10.2) that $\mu \neq 0$, so $(a, 1)$ is in the range of the map

$$\mathbf{R}^n \times K \ni (\lambda, x) \mapsto \sum_1^n \lambda_k \partial f(x) / \partial x_k \in \mathbf{R}^N. \tag{10.3}$$

By a rather trivial case of the Morse-Sard theorem (Theorem 3.1) the range is a null set when $N > 2n$, and since the intersection with the planes $\{a \in \mathbf{R}^N; a_N = \mu\}$ where $\mu \neq 0$ are homothetic they are all null sets by the Lebesgue-Fubini theorem. Hence E is a closed null set.

Lemma 10.4. *If $f \in C^\infty(X, \mathbf{R}^N)$ is an injective immersion on the compact set K and $N > 2n+1$ then $\pi_a f$ is also an injective immersion on K except when a belongs to a closed null set in $\mathbf{R}^{N-1}$.*

Proof. From Lemma 10.3 we already know that $\pi_a f$ is an immersion except when a belongs to a closed null set E. Let E' be the set of all $a \in \mathbf{R}^{N-1}$ such that $\pi_a f$ is not injective on K. To prove that $E \cup E'$ is closed we consider a sequence $a_j \in E'$, $a_j \to a$. Then we can find $x_j' \neq x_j''$ in K with $\pi_{a_j} f(x_j') = \pi_{a_j} f(x_j'')$ for $j = 1, 2, \dots$. Replacing the sequence by a subsequence we may assume that $x_j' \to x'$ and that $x_j'' \to x''$ for some $x', x'' \in K$. If $a \notin E$ then $\pi_{a_j} f$ is by the inverse function theorem an injective immersion on a fixed neighborhood of x' if j is sufficiently large, so we have $x' \neq x''$, and since $\pi_a f(x') = \pi_a f(x'')$ it follows that $a \in E'$. Hence $E \cup E'$ is closed. That $a \in E'$ means explicitly that

$$f_j(x') - a_j f_N(x') = f_j(x'') - a_j f_N(x''), \quad j = 1, \dots, N-1, \tag{10.1$'$}$$

for some $x', x'' \in K$ with $x' \neq x''$. With $a_N = 1$ and $\mu = f_N(x') - f_N(x'')$ we can write (10.1)$'$ in the form

$$f(x') - f(x'') = \mu a. \tag{10.2$'$}$$

Since f is injective on K we have $\mu \neq 0$, so $(a, 1)$ belongs to the range of the map

$$\mathbf{R}^n \times K \times K \ni (t, x', x'') \to t(f(x') - f(x'')) \in \mathbf{R}^N. \tag{10.3$'$}$$

Since $N > 2n+1$ the range of this map is a null set in $\mathbf{R}^N$, by Theorem 3.1, and by the homogeneity the intersection with the plane $a_N = 1$ is also a null set there. The lemma is proved.

Note that in Lemma 10.3 we just avoided projecting along a tangent of $f(K)$ whereas in Lemma 10.4 we also avoided projecting along chords. The set of forbidden directions depends on $2n-1$ and on $2n$ parameters respectively which explains the assumptions on the dimensions. We can now prove the main result in this chapter:

Theorem 10.5. *Let $f \in C^\infty(X, \mathbf{R}^N)$ be a proper map with $N \geq 2n+1$ where n is the dimension of X. For every positive continuous function ε on X there is a proper embedding $g \in C^\infty(X, \mathbf{R}^N)$ such that*

$$|g(x) - f(x)| \leq \varepsilon(x), \quad x \in X. \tag{10.4}$$

Proof. We may assume that ε is so small that $x \mapsto |f(x)| - \varepsilon(x) \in \mathbf{R}$ is proper and positive outside a compact set. Then (10.4) implies that g is proper. Let $M_1, M_2, \dots$ be a sequence of compact subsets of X, each contained in the interior of the following one, such that $\cup M_j = X$. We shall successively construct $g_j \in C^\infty(X, \mathbf{R}^N)$ so that $g_0(x) = f(x)$ and

$$|g_j(x) - g_{j-1}(x)| < \varepsilon(x)/2^j, \quad j = 1, 2, \dots, \ x \in X. \tag{10.5}$$

Moreover g_j shall be an injective immersion on M_j and $g_j = g_{j-1}$ on M_{j-1}. Suppose that g_{j-1} has already been chosen and that $j \geq 1$. By Lemma 10.2 we can choose

$h \in C_0^\infty(X, \mathbf{R}^\nu)$ for suitable ν so that $g_{j-1} \oplus h$ is an injective immersion of M_j in $\mathbf{R}^{N+\nu}$. By the remark after the proof of Lemma 10.2 we can take $h = 0$ on M_{j-1}. In fact, there is a neighborhood ω of M_{j-1} where g_{j-1} is an injective immersion, and we just have to choose h as an injective immersion in $M_j \setminus \omega$, which is a compact set disjoint with M_{j-1}. By repeated use of Lemma 10.4 we conclude that there exist linear transformations $T : \mathbf{R}^\nu \to \mathbf{R}^N$ with arbitrarily small norm such that

$$g_j = g_{j-1} + Th$$

is an injective immersion on M_j. If $\|T\|$ is small enough then (10.5) follows so g_j has all the desired properties.

It is now obvious that $g = \lim_{j\to\infty} g_j$ exists, and that $g \in C^\infty(X, \mathbf{R}^N)$. Since $g = g_j$ on M_j and g_j is an injective immersion on M_j, it follows that g is an injective immersion, and (10.5) implies (10.4). The proof is complete.

As an application of Theorem 10.5 we shall now discuss how one can approximate maps into X with C^∞ maps. We also have to prove that nearby maps f_0 and f_1 into X are homotopic. If X is realised as a closed submanifold of $\mathbf{R}^N$ it is natural to connect f_0 and f_1 by means of the linear homotopy $tf_1 + (1-t)f_0$, $0 \le t \le 1$. However, this does not stay in X so we need a C^∞ map from a neighborhood of X back to X which leaves X invariant (a *retraction*). We shall now construct a retraction by studying the neighborhood of a closed submanifold X of $\mathbf{R}^N$.

The tangent space of a closed submanifold X of $\mathbf{R}^N$ can be identified with $T(X) = \{(x, w) \in X \in \mathbf{R}^N; w \in \operatorname{Im} \psi'(t)\}$ where ψ is a local parametrisation in a neighborhood of $x = \psi(t)$. (See Chapter IX.) We define the normal bundle

$$N(X) = \{(x, w) \in X \times \mathbf{R}^N; \langle w, w' \rangle = 0 \text{ if } (x, w') \in T(X)\}.$$

If ψ is a local parametrisation then the condition on w is that ${}^t\psi'(t)w = 0$, which defines n coordinates of w as linear functions of the other $N - n$ coordinates. Thus the manifold $N(X)$ has dimension $n + N - n = N$. The map

$$F : N(X) \ni (x, w) \mapsto x + w \in \mathbf{R}^N \tag{10.6}$$

is a proper injection of $N_0 = \{(x, 0); x \in X\} \subset N(X)$ and has bijective differential at every point there. In fact, the range of the differential of F at $(x, 0)$ contains the range of the restriction to N_0, which is $T_x(X)$, and it also contains the range of the restriction to the fiber N_x, which is mapped linearly on the orthogonal space of $T_x(X)$. This implies that N_0 has an open neighborhood $\Omega \subset N(X)$ which is mapped diffeomorphically on a neighborhood $\widetilde{\Omega}$ of X in $\mathbf{R}^N$:

THEOREM 10.6. *If X is a closed submanifold of $\mathbf{R}^N$, then there is an open neighborhood $\Omega \subset N(X)$ of the zero section in the normal bundle and an open neighborhood $\widetilde{\Omega}$ of X in $\mathbf{R}^N$ such that* (10.6) *is a diffeomorphism $\Omega \to \widetilde{\Omega}$.*

The proof is left as an exercise in the following more general form:

EXERCISE 10.7. Let X and Y be two C^∞ manifolds of the same dimension, and assume that $f \in C^\infty(X, Y)$ has bijective differential at every point in a closed subset F of X. Prove that if $f|_F$ is proper and injective then there are open neighborhoods Ω_X and Ω_Y of F and $f(F)$ such that f is a diffeomorphism $\Omega_X \to \Omega_Y$.

COROLLARY 10.8. *For every closed C^∞ submanifold X of $\mathbf{R}^N$ there is a positive continuous function ε on X such that for every C^∞ manifold Y and arbitrary $f \in C(Y,X)$ and $f_0, f_1 \in C^\infty(Y,X)$ with $|f_j - f| < f^*\varepsilon$ there exists a function $g \in C^\infty(Y \times [0,1], X)$ with $g(\cdot,t) = f_t$ when $t = 0,1$.*

PROOF. With the notation in Theorem 10.6 we set

$$\varepsilon(x) = \sup\{\varepsilon; 0 < \varepsilon \le 1, |x - x'| < \varepsilon \implies x' \in \widetilde{\Omega}\}, \quad x \in X.$$

If $|f_j(y) - f(y)| < \varepsilon(f(y))$ it follows that

$$\tilde{f}(y,t) = tf_1(y) + (1-t)f_0(y) \in \widetilde{\Omega}, \quad 0 \le t \le 1.$$

With F defined by (10.6) and p equal to the projection $N(X) \to X$ we can therefore take $g = p \circ F^{-1} \circ \tilde{f}$.

We can also prove the existence of differentiable approximations:

COROLLARY 10.9. *Let X be a closed C^∞ submanifold of $\mathbf{R}^N$. If Y is a C^∞ manifold and $f \in C(Y,X)$ one can for every positive continuous function on X find $g \in C^\infty(Y,X)$ such that $|f(y) - g(y)| < \varepsilon(f(y))$, $y \in Y$.*

PROOF. We may assume that Y is a closed submanifold of $\mathbf{R}^{N'}$. By Theorem 10.6 there exist open neighborhoods Ω_X and Ω_Y of X and Y in $\mathbf{R}^N$ and $\mathbf{R}^{N'}$ and C^∞ retractions $R_X : \Omega_X \to X$ and $R_Y : \Omega_Y \to Y$. (With the notation in the proof of Corollary 10.8 we can take $R_X = p \circ F^{-1}$.) Choose a positive continuous function ε_1 on X such that

$$x \in X,\ x' \in \mathbf{R}^N,\ |x - x'| < \varepsilon_1(x) \implies x' \in \Omega_X \text{ and } |R_X x' - x| < \varepsilon(x),$$

and set $\tilde{f} = f \circ R_Y$, which is a continuous extension of f to Ω_Y. There is a positive C^∞ function δ on Y such that

$$y \in Y,\ y' \in \mathbf{R}^{N'},\ |y - y'| < \delta(y) \implies y' \in \Omega_Y \text{ and } |\tilde{f}(y') - f(y)| < \varepsilon_1(f(y)).$$

This is clear with a constant δ on any compact subset of Y, and we can use the proof of Lemma 10.1 to find $1/\delta$ increasing so rapidly that this is true. Choose $\varphi \in C_0^\infty(\mathbf{R}^{N'})$ with $\varphi \ge 0$, $\int \varphi(y)\,dy = 1$ and $\varphi(y) = 0$ when $|y| > \frac{1}{2}$. Then

$$g(y) = R_X(h(y)), \quad h(y) = \delta(y)^{-N'} \int \tilde{f}(y')\varphi((y - y')/\delta(y))dy'$$

has the required properties.

CHAPTER XI

DE RHAM COHOMOLOGY

Let X be a C^∞ manifold. In Chapter IX we defined the vector bundle $\lambda^p(T^*(X))$. For the sake of brevity we shall use the notation $\lambda^p(X)$ for its C^∞ sections, the C^∞ p forms on X. For every p the exterior differential gives a map

$$d : \lambda^p(X) \to \lambda^{p+1}(X),$$

and for every C^∞ map $f : X \to Y$ where Y is another C^∞ manifold we have a pullback map $f^* : \lambda^p(Y) \to \lambda^p(X)$ commuting with d. Thus the diagram

$$\begin{array}{ccc} \lambda^p(Y) & \xrightarrow{d} & \lambda^{p+1}(Y) \\ {\scriptstyle f^*}\downarrow & & {\scriptstyle f^*}\downarrow \\ \lambda^p(X) & \xrightarrow{d} & \lambda^{p+1}(X) \end{array} \tag{11.1}$$

is commutative. This was already discussed in Chapter VIII when X and Y are open subsets of $\mathbf{R}^m$ and $\mathbf{R}^n$. In view of (11.1) we can immediately take over the definitions from local coordinate systems. When $X \subset \mathbf{R}^n$ we have proved that $d^2 = 0$ for every p, and this remains valid for a manifold X.

Let us now recall some elementary definitions from algebra. Let $A_0, \dots, A_n$ be vector spaces, and assume that for $j = 0, \dots, n-1$ we are given a map $d_j : A_j \to A_{j+1}$ so that $d_{j+1}d_j = 0$. (We shall drop the subscript of d when it is obvious from the context.) Then we have a *complex* of vector spaces. By adding spaces $A_j = \{0\}$ when $j < 0$ or $j > n$ with $d_j = 0$ when $j < 0$ or $j \geq n$ we may assume that A_j is defined for every j which is notationally convenient.

The hypothesis that we have a complex means precisely that the range R_j of d_{j-1} is a subspace of the kernel N_j of d_j. If there is equality the complex is said to be *exact* at A_j. Thus exactness means that the range of d_{j-1} is described precisely as the kernel of d_j. This is usually not the case, and as a measure of how close the range R_j of d_{j-1} is to the kernel N_j of d_j one introduces the quotient space

$$H^j = N_j/R_j,$$

which is called the (co)homology of the complex at A_j, or of degree j. When $j = 0$ we obtain $H_0 = N_0$ while $H_n = A_n/R_n$ is the cokernel of d_{n-1}.

Let us also consider purely algebraically a situation similar to (11.1) where we have two complexes A and B and a homomorhism between them, that is, a linear map $T_i : A_i \to B_i$ for each i, such that the following diagram is commutative:

$$\begin{array}{ccccccccc} A_0 & \longrightarrow & A_1 & \longrightarrow & A_2 & \longrightarrow & A_3 & \longrightarrow & \dots \\ {\scriptstyle T_0}\downarrow & & {\scriptstyle T_1}\downarrow & & {\scriptstyle T_2}\downarrow & & {\scriptstyle T_3}\downarrow & & . \\ B_0 & \longrightarrow & B_1 & \longrightarrow & B_2 & \longrightarrow & B_3 & \longrightarrow & \dots \end{array}$$

Then it follows that the kernel of the operator $d_j : A_j \to A_{j+1}$ is mapped by T_j into the kernel of the operator $d_j : B_j \to B_{j+1}$, and similarly for the range of d_{j-1}. Hence a map $\widetilde{T}_j : H^j(A) \to H^j(B)$ is induced by T_j. If we have a third complex C and a map $B \to C$ of the same kind then the composition $H^j(A) \to H^j(B) \to H^j(C)$ is equal to the map induced by the composition $A \to C$.

If we apply the preceding discussion to the complex of p forms (called the *de Rham complex*) and the maps f^* defined by C^∞ maps $f : X \to Y$, we obtain the following situation:

Definition 11.1. If X is a C^∞ manifold then the de Rham cohomology groups $H^p(X)$ of X are the cohomology groups of the de Rham complex on X. We write $H^*(X)$ for the direct sum of all $H^p(X)$.

The groups $H^p(X)$ are of course vector spaces but the term group is used to accomodate related constructions.

Theorem 11.2. *For every C^∞ map $f : X \to Y$ the pullback of differential forms induces a linear map $f^* : H^p(Y) \to H^p(X)$. If $g : Y \to Z$ and $h = g \circ f$, then the diagram*

$$\begin{tikzcd} H^p(Z) \arrow[rr, "g^*"] \arrow[dr, "h^*"'] & & H^p(Y) \arrow[dl, "f^*"] \\ & H^p(X) & \end{tikzcd}$$

is commutative.

In Chapter VIII we also introduced a multiplication in $\lambda^*(X) = \oplus_{p \geq 0} \lambda^p(X)$ which makes it a ring with

$$u \wedge v = (-1)^{pq} v \wedge u, \quad u \in \lambda^p,\ v \in \lambda^q.$$

From (8.9) it follows that $d(u \wedge v) = 0$ if $du = 0$ and $dv = 0$. The class of $u \wedge v$ in $H^{p+q}(X)$ depends only on the classes of u and of v, for if the classes of u_0 and v_0 are 0, that is, $u_0 = du_1$ and $v_0 = dv_1$, then

$$\begin{aligned} (u + u_0) \wedge (v + v_0) &= u \wedge v + u_0 \wedge (v + v_0) + u \wedge v_0 \\ &= u \wedge v + d(u_1 \wedge (v + v_0)) + (-1)^p d(u \wedge v_1). \end{aligned}$$

Thus we obtain a multiplication

$$H^p(X) \times H^q(X) \to H^{p+q}(X)$$

which traditionally is denoted by $\cup$ and called the *cup product*. With this multiplication $H^*(X)$ is a ring with the same anticommutativity of the multiplication in $H^*(X)$ as in $\lambda^*(X)$. The following theorem is now trivial but quite important anyway:

Theorem 11.3. *If $f : X \to Y$ is a C^∞ map then the induced map $f^* : H^*(Y) \to H^*(X)$ is a ring homomorphism.*

The map f^* is stable under homotopic changes of f:

Theorem 11.4. *Let $F : X \times [0,1] \to Y$ be a C^∞ map, and set $f_t(x) = F(x,t)$ when $0 \le t \le 1$. Then f_t^* is independent of t.*

Proof. Let u be a differential form on Y with $du = 0$. We can write

$$F^*u = U_0 + dt \wedge U_1$$

where U_0 and U_1 are differential forms in X depending on t; we have of course $U_0 = f_t^* u$. That $dF^*u = F^*du = 0$ means in particular that "the coefficient of dt" is equal to 0, that is, that $\partial U_0/\partial t - d_X U_1 = 0$ where d_X denotes the exterior differential operator in X acting on U_1 as a form in X depending on the parameter t. Hence

$$f_1^*u - f_0^*u = U_0(1) - U_0(0) = \int_0^1 d_X U_1 \, dt = d_X \int_0^1 U_1 \, dt.$$

The right-hand side is the differential of a form in X, which proves the statement. (The proof is of course essentially a repetition of that of Theorem 8.1.)

Using the approximation method in Corollary 10.9 we can now conclude that f^* can be defined uniquely for an arbitrary continuous map $f : X \to Y$, for all sufficiently close C^∞ approximations g are homotopic by Corollary 10.8 so we can define $f^* = g^*$. This extended definition of f^* is also invariant for continuous homotopies, for a continuous homotopy $F : X \times [0,1] \to Y$ can first be extended to all of $X \times \mathbf{R}$ so that it is constant for $t \le 0$ and for $t \ge 1$ and then it can be approximated by C^∞ homotopies. Hence we have

Theorem 11.5. *Every continuous map $f : X \to Y$ defines a ring homomorphism $f^* : H^*(Y) \to H^*(X)$ with the properties in Theorem 11.3, and Theorem 11.4 remains valid for continuous homotopies.*

Corollary 11.6. *If the continuous map $f : X \to Y$ is a homotopy equivalence, that is, if there exists a continuous map $g : Y \to X$ such that $f \circ g$ and $g \circ f$ are homotopic to the identity in Y and X respectively, then f^* is a ring isomorphism.*

The map $f^* : \lambda^p(Y) \to \lambda^p(X)$ is in general neither injective nor surjective. When $p = 0$, for example, surjectivity requires that f is a proper embedding so that $f(X)$ can be regarded as a closed submanifold of Y. This is sufficient for any p:

Lemma 11.7. *If $f : X \to Y$ is a C^∞ proper embedding then $f^* : \lambda^*(Y) \to \lambda^*(X)$ is surjective.*

Proof. We may assume that Y is a closed submanifold of $\mathbf{R}^N$. Then $f(X)$ is also a closed submanifold of $\mathbf{R}^N$. We identify X with $f(X)$ and can then regard f as the natural injection $i : X \to Y$. The lemma states that every smooth differential form on X is then the restriction of a smooth form on Y. By Theorem 10.6 we can choose an open neighborhood $\widetilde{\Omega}$ of X in $\mathbf{R}^N$ and a C^∞ retraction $R : \widetilde{\Omega} \to X$, thus $R(x) = x$ when $x \in X$. Let $\varphi \in C^\infty(\mathbf{R}^N)$ be equal to 1 on X and equal to 0 in a neighborhood of $\complement\widetilde{\Omega}$. Then $\varphi R^* u$, defined as 0 in $\complement\widetilde{\Omega}$ is a form in $\mathbf{R}^N$ which restricted to X is equal to u, so the restriction to Y has the required property.

Under the assumptions of Lemma 11.7 the de Rham complexes in Y and X give a commutative diagram

$$\begin{array}{ccccccccc}
0 & \longrightarrow & A_0 & \longrightarrow & A_1 & \longrightarrow & A_2 & \longrightarrow & \dots \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & B_0 & \longrightarrow & B_1 & \longrightarrow & B_2 & \longrightarrow & \dots
\end{array},$$

where the vertical maps are surjective. If we denote their kernels by C_j we therefore obtain a commutative diagram with exact columns

$$\begin{array}{ccccccccc}
 & & 0 & & 0 & & 0 & & \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & C_0 & \longrightarrow & C_1 & \longrightarrow & C_2 & \longrightarrow & \dots \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & A_0 & \longrightarrow & A_1 & \longrightarrow & A_2 & \longrightarrow & \dots, \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & B_0 & \longrightarrow & B_1 & \longrightarrow & B_2 & \longrightarrow & \dots \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
 & & 0 & & 0 & & 0 & &
\end{array}$$

where the maps $C_j \to C_{j+1}$ are determined by the commutativity: If we go from C_j to A_j to A_{j+1} we obtain an element in C_{j+1}, for the image in B_{j+1} is equal to 0 since we could also have gone from A_j to B_{j+1} via B_j. Hence we obtain maps

$$H^j(C) \to H^j(A) \to H^j(B)$$

with composition 0. This complex is in fact exact. For let $a \in A_j$ represent a class in $H^j(A)$ mapped to 0. This means that the image of a_j in B_j comes from an element in B_{j-1}. By the assumed surjectivity it can be lifted to an element $a_{j-1} \in A_{j-1}$, and the commutativity implies that da_{j-1} and a_j have the same image in B_j. Hence $a_j - da_{j-1} \in C_j$, which means that the class in $H^j(A)$ represented by a_j is the image of a class in $H^j(C)$.

We shall now determine the range of the map $H^j(A) \to H^j(B)$. Let $b_j \in B_j$, $db_j = 0$. Then b_j is the image of an element $a_j \in A_j$, and da_j is mapped to 0 in B_{j+1}, so $da_j \in C_{j+1}$. We claim that the class of da_j in $H^{j+1}(C)$ only depends on the class of b_j in $H^j(B)$. To prove this consider the case where the class of b_j is 0, that is, $b_j = db_{j-1}$. Then we can lift b_{j-1} to an element $a_{j-1} \in A_{j-1}$ and obtain $a_j - da_{j-1} = c_j \in C_j$. Hence $da_j = dc_j$, so da_j defines the 0 class in $H^{j+1}(C)$. Thus we have defined a map

$$\delta_j : H^j(B) \to H^{j+1}(C).$$

We shall usually drop the subscript j. The sequence

$$H^j(A) \longrightarrow H^j(B) \xrightarrow{\delta_j} H^{j+1}(C)$$

now obtained is also an exact complex. In fact, the composition is 0 for if a class in $H^j(B)$ is the image of a class in $H^j(A)$ then one can choose a_j so that $da_j = 0$ with the notation above. On the other hand, if the class of b_j is mapped to 0 then $da_j = dc_j$ for some $c_j \in C_j$, so the element $a_j - c_j \in A_j$ defines a class in $H^j(A)$ mapped to the class of b_j.

Finally we claim that the sequence

$$H^j(B) \xrightarrow{\delta_j} H^{j+1}(C) \to H^{j+1}(A)$$

is an exact complex. That the composition is 0 follows at once from the definition of δ_j.On the other hand, consider an element in $H^{j+1}(C)$ with image 0 in $H^{j+1}(A)$, thus an element $c_{j+1} \in C_{j+1}$ with $c_{j+1} = da_j$ for some $a_j \in A_j$. If b_j is the image of a_j in B_j then δ_j maps the class of b_j to the class of c_{j+1}. Summing up, we have proved that

$$\begin{aligned} 0 \longrightarrow H^0(C) \longrightarrow H^0(A) \longrightarrow H^0(B) \longrightarrow H^1(C) \\ \longrightarrow H^1(A) \longrightarrow H^1(B) \longrightarrow H^2(C) \to \dots \end{aligned}$$

is an exact complex.

Remark. In the proof we never used that our objects are vector spaces. The arguments are therefore also valid for modules over a commutative ring.

We shall now apply the results to the inclusion $i : X \to Y$ of a closed submanifold X of Y. Thus we take $A_j = \lambda^j(Y)$, $B_j = \lambda^j(X)$, and have to introduce

$$C_j = \{u; u \in \lambda^j(Y),\ i^*u = 0\}.$$

However, it is better to use instead the subcomplex

$$C_j^0 = \{u \in \lambda^j(Y); u = 0 \text{ in a neighborhood of } X\}.$$

We claim that the map $H^j(C^0) \to H^j(C)$ is an isomorphism. For the proof we introduce the quotients C_j/C_j^0 and obtain the same algebraic situation as above, hence an exact sequence

$$\begin{aligned} 0 \longrightarrow H^0(C^0) \longrightarrow H^0(C) \longrightarrow H^0(C/C^0) \\ \longrightarrow H^1(C^0) \longrightarrow H^1(C) \longrightarrow H^1(C/C^0) \longrightarrow \dots . \end{aligned}$$

This proves that the claim is equivalent to $H^j(C/C^0) = 0$ for every j. To prove this we take $u \in C_j$ with $du \in C_{j+1}^0$. In a tubular neighborhood T of X where $du = 0$ we have a projection $p : T \to X$ which is homotopic to the identity on T (see Theorem 10.6), so it follows from Theorem 11.4 that $u - p^*u = dv$ in T where v is a form in T. Now $p = i \circ p$, so $p^*u = p^*i^*u = 0$ for $u \in C_j$. If $\varphi \in C^\infty(Y)$ is equal to 1 in a neighborhood of X and equal to 0 in a neighborhood of $\complement T$, then $u - d(\varphi v) \in C_j^0$. The proof of Theorem 11.4 shows that we can choose v with $i^*v = 0$, so it follows that the class of u in $H^j(C/C_0)$ is equal to 0.

Definition 11.8. If M is a C^∞ manifold then the de Rham cohomology groups $H_c^p(M)$ with compact support are defined as the cohomology of the complex of p forms of compact support. If X is a compact C^∞ submanifold of a compact C^∞ manifold Y we shall also write $H^p(Y, X)$ instead of $H_c^p(Y \setminus X)$ and refer to this as the relative cohomology group.

With this terminology we have now proved:

THEOREM 11.9. *If X is a compact C^∞ submanifold of a compact C^∞ manifold Y and $i : X \to Y$ denotes the inclusion of X in Y, then we have an exact sequence*

$$0 \longrightarrow H^0(Y,X) \longrightarrow H^0(Y) \xrightarrow{i^*} H^0(X) \xrightarrow{\delta} H^1(Y,X)$$
$$\longrightarrow H^1(Y) \xrightarrow{i^*} H^1(X) \xrightarrow{\delta} H^2(Y,X) \longrightarrow \dots .$$

Here the map $H^k(Y,X)$ is defined by the inclusion of differential forms with compact support in $Y \setminus X$ among arbitrary differential forms on Y while $\delta : H^k(X) \to H^{k+1}(Y,X)$ maps the class of i^*u to the class of du if u is a form on Y with $du = 0$ in a neighborhood of X.

We give a simple but important application:

THEOREM 11.10. *If X is a compact C^∞ submanifold of a compact C^∞ manifold Y and X is a retract of Y, that is, there exists a continuous map $f : Y \to X$ with $f(x) = x$ when $x \in X$, then*

$$0 \longrightarrow H^j(Y,X) \longrightarrow H^j(Y) \longrightarrow H^j(X) \longrightarrow 0$$

is exact for every j. We can therefore view $H^j(Y,X)$ as a subspace of $H^j(Y)$, and the quotient $H^j(Y)/H^j(Y,X)$ is isomorphic to $H^j(X)$.

PROOF. $f^* : H^*(X) \to H^*(Y)$ is a right inverse of i^* for $i^*f^* = (fi)^*$ is the identity on $H^*(X)$. Hence $H^j(Y) \to H^j(X)$ is surjective for every j, so the exact sequence in Theorem 11.9 gives that $H^{j+1}(Y,X) \to H^{j+1}(Y)$ is injective.

Using Theorem 11.9 we shall calculate $H^k(X)$ in Chapter XII for some important spaces X. We end the preparations here by a few comments on the cohomology $H^*_c(X)$ with compact support. Let $f \in C^\infty(X,Y)$. (We no longer assume that X and Y are compact.) If u is a form with compact support in Y, then $\operatorname{supp} f^*u \subset \{x; f(x) \in \operatorname{supp} u\}$, but this is not necessarily a compact set. An example is a constant map f when X is not compact.

If $f \in C^\infty(X,Y)$ is a proper map it is clear that f^*u has compact support, so then we get a map $f^* : H^*_c(Y) \to H^*_c(X)$. As before we can extend the definition to every continuous proper f. The map is not changed by a proper homotopy, that is, a proper map $X \times [0,1] \to Y$.

Note that $H^*_c(X)$ is not only a ring but even a $H^*(X)$ module, for the product of two forms has compact support if one of the forms has compact support. The ring structure is obtained from the module structure and the map $H^*_c(X) \to H^*(X)$.

CHAPTER XII

THE DE RHAM COHOMOLOGY FOR SOME IMPORTANT MANIFOLDS

We shall first calculate the cohomology of $\mathbf{R}^n$. In doing so we include the case $n = 0$, noting that the results of Chapter XI remain valid also in that case if we interpret $\mathbf{R}^0$ as a single point. In that case the constants are the only differential forms; thus $H^k(\mathbf{R}^0) = 0$ when $k \neq 0$ and $H^0(\mathbf{R}^0) = \mathbf{R}$. (By 0 we denote the vector space containing only the origin.)

THEOREM 12.1. *$H^*(\mathbf{R}^n)$ is for every n the ring of real numbers, that is, $H^k(\mathbf{R}^n) = 0$ when $k \neq 0$ and $H^0(\mathbf{R}^n) = \mathbf{R}$.*

PROOF. For the map $i : \mathbf{R}^0 \to \mathbf{R}^n$ mapping $\mathbf{R}^0$ to the origin and the map $p : \mathbf{R}^n \to \mathbf{R}^0$ the composition $p \circ i$ is the identity while $i \circ p$ is the projection from $\mathbf{R}^n$ to the origin. It is homotopic to the identity by the homotopy $\mathbf{R}^n \times [0, 1] \ni (x, t) \mapsto tx$. Hence i, p is a homotopy equivalence, so $H^*(\mathbf{R}^n)$ is isomorphic to $H^*(\mathbf{R}^0)$ (Corollary 11.6). — Note that the theorem is equivalent to Poincaré's lemma (Theorem 8.1), and the proof here is basically the same.

THEOREM 12.2. *$H_c^k(\mathbf{R}^n) = 0$ when $k \neq n$, and*

$$\lambda_c^n(\mathbf{R}^n) \ni u \mapsto \int u$$

induces an isomorphism between $H_c^n(\mathbf{R}^n)$ and $\mathbf{R}$.

By the theorem the product of arbitrary elements in $H_c^*(\mathbf{R}^n)$ is equal to 0, if $n \neq 0$.

PROOF. We can assume that $n > 0$. The statement is obvious when $k = 0$ since a closed 0 form is a constant which must be 0 if the support is compact. If $n = 1$ and $u = v(x)\,dx \in \lambda_c^1(\mathbf{R})$ then $u = dw$ where $w \in \lambda_c^0(\mathbf{R})$ means that

$$w(x) = \int_{-\infty}^{x} v(t)\,dt,$$

and w has compact support if and only if $\int_{\mathbf{R}} v(t)\,dt = \int u = 0$. For $n > 1$ it is convenient to combine the proof with the proof of the following:

THEOREM 12.3. *For the sphere $S^n \subset \mathbf{R}^{n+1}$ we have $H^k(S^n) = 0$ when $0 < k < n$, while $H^0(S^n) \cong \mathbf{R}$ as a ring and the map*

$$\lambda^n(S^n) \ni u \mapsto \int_{S^n} u$$

induces an isomorphism $H^n(S^n) \cong \mathbf{R}$.

PROOF. If P is a point in S^n then P is a retract of S^n, and since $S^n \setminus P$ is diffeomorphic to $\mathbf{R}^n$ we obtain from Theorem 11.10 an exact sequence

$$0 \longrightarrow H^j_c(\mathbf{R}^n) \longrightarrow H^j(S^n) \longrightarrow H^j(P) \longrightarrow 0.$$

Hence it follows from Theorem 12.1 that Theorems 12.2 and 12.3 are equivalent, so they are proved for $n = 1$. Assume now that $n > 1$ and that they are proved for lower dimensions. Stokes' formula implies that $\int_{S^n} u = 0$ if u is an exact differential form on S^n, so $\lambda^n(S^n) \ni u \mapsto \int_{S^n} u$ induces a linear form on $H^n(S^n)$ which is not identically 0. For example, the value is 1 on the Kronecker form discussed at the end of Chapter IX. Hence the dimension q of $H^n(S^n)$ is at least equal to 1. Now the intersection between $S^n \subset \mathbf{R}^{n+1}$ and a hyperplane in $\mathbf{R}^{n+1}$ is equal to S^{n-1}, and $S^n \setminus S^{n-1}$ has two components diffeomorphic to $\mathbf{R}^n$. This implies that

$$H^j_c(S^n \setminus S^{n-1}) \cong H^j_c(\mathbf{R}^n) \oplus H^j_c(\mathbf{R}^n) \cong H^j(S^n) \oplus H^j(S^n), \quad j > 0.$$

Hence the exact cohomology sequence in Theorem 11.9 gives the exact sequence

$$H^{n-1}(S^{n-1}) \longrightarrow H^n(S^n) \oplus H^n(S^n) \longrightarrow H^n(S^n) \to 0.$$

The range of the first map is of dimension $p \leq 1$ by the inductive hypothesis, and the exactness implies that $2q - p = q$, hence $q = p$. Since $p \leq 1 \leq q$ it follows that $p = q = 1$, so $H^n(S^n) \cong \mathbf{R}$ and the preceding map from $H^{n-1}(S^{n-1})$ is injective so the map $H^{n-1}(S^n) \oplus H^{(n-1)}(S^n) \to H^{n-1}(S^n)$ is surjective. From the exact cohomology sequence and the inductive hypothesis we now obtain the exactness of

$$0 \longrightarrow H^j(S^n) \oplus H^j(S^n) \longrightarrow H^j(S^n) \longrightarrow 0, \quad 1 \leq j \leq n-1,$$

where for $j = 1$ we also use that $H^0(S^n) \longrightarrow H^0(S^{n-1})$ is surjective. Hence $H^j(S^n) = 0$ for $0 < j < n$ which completes the proof.

The second part of the theorem is valid in much greater generality:

THEOREM 12.4. *If X is a compact connected C^∞ manifold of dimension n then $H^n(X) \cong \mathbf{R}$ if X is orientable and $H^n(X) = 0$ if X is not orientable. In the first case the isomorphism is induced by $\lambda^n(X) \ni u \mapsto \int_X u$ with an orientation chosen for X.*

PROOF. Assume first that X is oriented. We can write $X = \cup_1^J X_j$ where every X_j is diffeomorphic with $\mathbf{R}^n$. Let $1 = \sum_1^J \varphi_j$ where $\varphi_j \in C_0^\infty(X_j)$. If $u \in \lambda^n(X)$ and

$$\int \varphi_j u = 0, \quad j = 1, \dots, J,$$

then it follows from Theorem 12.2 that $\varphi_j u = dv_j$ where $v_j \in \lambda_c^{n-1}(X_j)$; hence $u = d(\sum_1^J v_j)$. If V is the linear subspace of $\mathbf{R}^J$ consisting of all $(\int_X \varphi_1 dv, \dots, \int_X \varphi_J dv)$ where $v \in \lambda^{n-1}(X)$, then it follows that $u = dv$ if and only if the vector $(\int_X \varphi_1 u, \dots, \int_X \varphi_n u) \in V$. Now V is defined by orthogonality to finitely many vectors $c = (c_1, \dots, c_J)$. Thus u is exact if and only if u satisfies finitely many equations of the form

$$\int_X \Big(\sum_1^J c_j\varphi_j\Big)u = 0.$$

Such an equation must always hold when $u = dv$ with $v \in \lambda^{n-1}(X)$. In every coordinate patch Ω diffeomorphic with $\mathbf{R}^n$ we conclude using Theorem 12.2 that $\sum_1^J c_j\varphi_j$ is a constant, for if $\operatorname{supp} u \subset \Omega$ then the relation must be a consequence of the equation $\int u = 0$. Since X is connected it follows that $\sum_1^J c_j\varphi_j$ is a constant C in all of X. Hence u is exact if (and only if) $\int_X u = 0$.

Assuming now that X is not orientable we introduce the double cover $\widetilde{X}$ of X consisting of points $x \in X$ with a chosen direction in the 1 dimensional fiber of $\lambda^n T^*(X)$ at x, that is, an n form $\neq 0$ there. Over every coordinate patch in X diffeomorphic to $\mathbf{R}^n$ we have two identical coordinate patches in $\widetilde{X}$, so it is clear that $\widetilde{X}$ is a manifold. The natural map $p : \widetilde{X} \to X$ is a local diffeomorphism such that $p^{-1}(x)$ consists of two points in $\widetilde{X}$ for every $x \in X$. For a component X_1 of $\widetilde{X}$ the set of points in X such that $X_1 \cap p^{-1}(x)$ consists of precisely j points must be open, and since X is connected it must be the empty set or X. If $\widetilde{X}$ is not connected it follows that some component X_1 has exactly one point over every point in X, but this is a contradiction since it would define an orientation in X. Hence $\widetilde{X}$ is connected.

Let $r : \widetilde{X} \to \widetilde{X}$ be the map exchanging the two points in $p^{-1}(x)$ for every $x \in X$. If $u \in \lambda^k(X)$ and $U = p^*u$ then $r^*U = U$ since $p \circ r = r$. Conversely, every $U \in \lambda^k(\widetilde{X})$ with $r^*U = U$ is equal to p^*u where $u \in \lambda^k(X)$ is obtained by pulling U back locally by a smooth map $s : X \to \widetilde{X}$ with $p \circ s$ equal to the identity; the pullback does not change if s is replaced by $r \circ s$. If $u \in \lambda^n(X)$ and $U = p^*u$ then $\int_{\widetilde{X}} U = 0$ since the integral over the two coordinate patches over a coordinate patch in X differ by a factor -1. Hence it follows from the first part of the theorem that $U = dV = d(r^*V) = dW$ where $W = \frac{1}{2}(V + r^*V)$, for $U = r^*U = d(r^*V)$. Since $r \circ r$ is the identity we have $r^*W = W$. Now $W = p^*w$ where $w \in \lambda^{n-1}(X)$ and we obtain $p^*(u - dw) = 0$, hence $u = dw$ since p is a local diffeomorphism. The theorem is proved.

If X and Y are compact, orientable, connected and have the same dimension n, and if $f : X \to Y$ is a continuous map, then the corresponding map $f^* : H^n(Y) \to H^n(X)$ can be viewed as a linear map $\mathbf{R} \to \mathbf{R}$, that is, multiplication by a number $D \in \mathbf{R}$. If $f \in C^1(X, Y)$ this means that

$$\int_X f^*u = D \int_Y u, \quad u \in \lambda^n(Y).$$

D is an integer called the *degree of* f. The reason is clear: if we choose $f_1 \in C^\infty(X, Y)$ in the homotopy class of f and take u with $\operatorname{supp} u$ close to a regular value $y \in Y$ for f, then D is the number of points $x \in X$ with $f(x) = y$ counted as $+1$ or -1 depending on whether f maps the orientation of X at x to the orientation of Y at y or the opposite one. We leave the details as an exercise for the reader, and as another exercise to prove that if $\Omega \subset \mathbf{R}^n$ is open and bounded with C^∞ boundary, $y \notin \partial f(\Omega)$, then $D(f, \Omega, y)$, defined in Section 4, is the sum of the degrees of the map $x \mapsto (f(x)-y)/|f(x)-f(y)| \in S^{n-1}$ on the different components of $\partial\Omega$, oriented as the boundary of $\partial\Omega$. (Cf. Chapter IX.)

Now that we know the cohomology of S^n we can turn the proof of Theorem 12.3 around:

THEOREM 12.5 (THE ALEXANDER DUALITY THEOREM). *Let X be a C^∞ compact submanifold of the sphere S^n. Then there are isomorphisms*

$$H^{k-1}(X) \cong H^k(S^n, X), \quad 1 < k < n,$$

and for $n > 1$ there is an exact sequence

$$0 \longrightarrow H^{n-1}(X) \longrightarrow H^n(S^n, X) \longrightarrow \mathbf{R} \longrightarrow 0,$$

which implies that $\dim H^n(S^n, X) = 1 + \dim H^{n-1}(X)$.

PROOF. In the exact sequence

$$H^{k-1}(S^n) \longrightarrow H^{k-1}(X) \longrightarrow H^k(S^n, X) \longrightarrow H^k(S^n)$$

the first and last spaces are 0 if $1 < k < n$, by Theorem 12.3, which proves the first statement. We have $H^n(X) = 0$ since $\dim X < n$. Hence the second statement follows if we take $k = n$ and continue the sequence with $\longrightarrow H^n(X)$.

If X is an orientable connected compact submanifold of S^n of dimension $n-1$, then $\dim H^{n-1}(X) = 1$, hence $\dim H^n(S^n, X) = 2$. From duality theorems proved later on we may then conclude that $S^n \setminus X$ has two components, so we have again proved the Jordan-Brouwer theorem (Theorem 4.7) although only in a very regular case.

We shall now study the real projective space $P^n_{\mathbf{R}}$, defined as the set of equivalence classes in $\mathbf{R}^{n+1} \setminus \{0\}$ if one identifies points which differ only by a real factor. For equivalence classes of points $x \in \mathbf{R}^{n+1}$ with $x_i \neq 0$ we can use x_j/x_i, $1 \le j \le n+1$, $j \neq i$, as local coordinates and conclude at once that $P^n_{\mathbf{R}}$ is a C^∞ manifold. We shall drop the subscript $\mathbf{R}$ for a moment.

The definition of P^n shows that we have a surjective map $p : S^n \to P^n$ such that $p^{-1}(x)$ consists of two points in S^n for every $x \in P^n$. If r is the antipodal map $x \mapsto -x$ on S^n then $p \circ r = p$. If u is a differential form on P^n then $U = p^*u$ is a form on S^n, and we have

$$r^*U = r^*p^*u = (p \circ r)^*u = p^*u = U.$$

Conversely, if U is a form on S^n with $r^*U = U$ then there is a form u on P^n with $U = p^*u$. The argument is the same as at the end of the proof of Theorem 12.4: Every point in P^n has a neighborhood Ω such that there is a smooth map $s : \Omega \to S^n$ with $p \circ s$ equal to the identity in Ω. There are two such maps, the other one is $r \circ s$. We can therefore define $u = s^*U = (r \circ s)^*U$ in Ω and obtain a globally defined form u on P^n with $p^*u = U$.

If u is a closed k form on P^n with $0 < k < n$ then we know from Theorem 12.3 that $p^*u = U = dV$ where V is a $k-1$ form on S^n. We have also $U = r^*U = dr^*V$, hence $U = dV_1$ where $V_1 = \frac{1}{2}(V + r^*V)$ has the property $r^*V_1 = V_1$, since $r \circ r$ is the identity. This means that $V_1 = p^*v$ where v is a $k-1$ form on P^n with $dv = u$, which proves that $H^k(P^n)=0$ when $0<k<n$. For $k=0$ we obtain $H^0(P^n)=\mathbf{R}$ since P^n is connected. To determine $H^n(P^n)$ we first consider the Kronecker form

$$\omega=\sum_{j=1}^{n+1}(-1)^{j-1}x_j dx_1\wedge\cdots\wedge\widehat{dx_j}\wedge\cdots\wedge dx_{n+1}$$

on S^n. We have $r^*\omega=(-1)^{n+1}\omega$, $d\omega=0$ and $\int_{S^n}\omega\neq 0$ by (9.9), so the class of ω generates $H^n(S^n)$. Hence the cohomology class of r^*U is equal to $(-1)^{n+1}$ times the cohomology class of U if U is a (closed) n form on S^n. But if $r^*U=U$ this implies that the cohomology class of U is 0 if n is *even*, and then it follows as above that $H^n(P^n)=0$. If n is *odd* then $\omega=p^*u$ where u is a (closed) n form on P^n with cohomology class $\neq 0$ since the class of ω is not 0. By the arguments above it follows that $H^n(P^n)=\mathbf{R}$, so we have proved

Theorem 12.6. *$H^k(P^n_{\mathbf{R}})=0$ when $0<k<n$, $H^0(P^n_{\mathbf{R}})=\mathbf{R}$, and $H^n(P^n_{\mathbf{R}})=\mathbf{R}$ when n is odd, $H^n(P^n_{\mathbf{R}})=0$ when n is even. Thus $P^n_{\mathbf{R}}$ is orientable if and only if n is odd.*

The preceding theorem is disappointing in the sense that it does not allow one to distinguish $P^n_{\mathbf{R}}$ from spheres and Euclidean spaces for all n. More satisfactory results are obtained by different definitions of cohomology which do not allow division by 2, which was essential above. However, the de Rham cohomology is quite satisfactory for *complex* projective spaces which we shall now discuss.

The complex projective space $P^n_{\mathbf{C}}$ is defined as $P^n_{\mathbf{R}}$ but with real numbers replaced by complex numbers. Thus it is the quotient of $\mathbf{C}^{n+1}\setminus\{0\}$ when elements differing by a complex factor are identified. The dimension of $P^n_{\mathbf{C}}$ as a real manifold is $2n$. When $n=1$ it is equal to S^2 so we know the cohomology in that case already.

Theorem 12.7. *$H^k(P^n_{\mathbf{C}})=0$ when k is odd and $H^k(P^n_{\mathbf{C}})=\mathbf{R}$ when k is even, $0\le k\le 2n$. More precisely: If $0\neq x\in H^2(P^n_{\mathbf{C}})$ then every element in $H^*(P^n_{\mathbf{C}})$ can be written in a unique way as a polynomial*

$$\sum_0^n a_j x^j$$

in x with real coefficients a_j. The ring operations in $H^*(P^n_{\mathbf{C}})$ correspond to the operations in the polynomial ring with the relation $x^{n+1}=0$; thus $H^*(P^n_{\mathbf{C}})$ is a truncated polynomial ring.

Proof. We have already proved the theorem for $n=1$. We shall prove it in general by induction. Note that $P^{n-1}_{\mathbf{C}}$ is embedded in $P^n_{\mathbf{C}}$ by the map

$$\mathbf{C}^n\setminus\{0\}\ni(z_1,\dots,z_n)\mapsto(z_1,\dots,z_n,0)\in\mathbf{C}^{n+1}\setminus\{0\}.$$

Then $P^n_{\mathbf{C}}\setminus P^{n-1}_{\mathbf{C}}$ consists of points with homogeneous coordinates such that $z_{n+1}\neq 0$. They can be uniquely represented by points for which $z_{n+1}=1$ which proves that $P^n_{\mathbf{C}}\setminus P^{n-1}_{\mathbf{C}}$ is diffeomorphic to $\mathbf{C}^n\cong\mathbf{R}^{2n}$. Hence we have an exact sequence

$$H^k_c(\mathbf{R}^{2n})\longrightarrow H^k(P^n_{\mathbf{C}})\longrightarrow H^k(P^{n-1}_{\mathbf{C}})\longrightarrow H^{k+1}_c(\mathbf{R}^{2n}).$$

When $k < 2n-1$ the first and last spaces are 0, hence $H^k(P^n_{\mathbf{C}}) \cong H^k(P^{n-1}_{\mathbf{C}})$ then. When $k = 2n-1$ the first and the third space are equal to 0, hence $H^{2n-1}(P^n_{\mathbf{C}}) = 0$. When $k = 2n$ we obtain by adding $H^{2n-1}(P^{n-1}_{\mathbf{C}}) \longrightarrow$ to the left that $H^{2n}(P^n_{\mathbf{C}}) \cong H^{2n}_c(\mathbf{R}^{2n}) \cong \mathbf{R}$, which completes the proof of the first part of the theorem. To prove the second one we shall write down explicitly a form defining x and verify that x^n defines an element $\neq 0$ in $H^{2n}(P^n_{\mathbf{C}})$. This implies that $x^j \neq 0$ when $1 \leq j \leq n$ which will complete the proof,

At this point it is convenient to make a digression concerning analytic functions and analytic manifolds.[1] If U and V are two vector spaces over $\mathbf{C}$, then a complex linear map $T : U \to V$ is of course also a linear map between U and V considered as vector spaces over $\mathbf{R}$. However, if T is linear with respect to real scalars then T is not complex linear unless $T(ix) = iTx$ for every $x \in U$. For a general real linear $T : U \to V$ we can write $T = T_1 + T_2$ where

$$T_1 x = \tfrac{1}{2}(Tx - iT(ix)), \quad T_2 x = \tfrac{1}{2}(Tx + iT(ix)).$$

Here $T_1(ix) = iT_1 x$ but $T_2(ix) = -iT_2 x$, that is,

$$T_1(ax) = aT_1 x, \quad T_2(ax) = \bar{a} T_2 x, \quad x \in U, \ a \in \mathbf{C}.$$

The decomposition is obviously unique; one calls T_1 complex linear and T_2 antilinear.

Now consider the case where $U = V = \mathbf{C}^n$. If T is complex linear we can form the determinant $\det_{\mathbf{C}} T$ of the $n \times n$ complex matrix for T. We can also consider T as a map $\mathbf{R}^{2n} \to \mathbf{R}^{2n}$ with a $2n \times 2n$ real matrix and form its determinant $\det_{\mathbf{R}} T$. Then

$$\det_{\mathbf{R}} T = |\det_{\mathbf{C}} T|^2. \tag{12.1}$$

Since every T can be written $U_1 D U_2$ where U_j are unitary and D has a diagonal matrix it suffices to prove the formula for the unitary and the diagonal case. It is obvious in the unitary case for both sides are then equal to 1. (The unitary group is connected so its elements considered as orthogonal transformations must have the determinant $+1$.) When T has a diagonal matrix it suffices to consider the case $n = 1$, thus $Tx = ax$, $x \in \mathbf{C}$. Complex multiplication by $a = a_1 + ia_2$ corresponds to the real matrix $\begin{pmatrix} a_1 & -a_2 \\ a_2 & a_1 \end{pmatrix}$ with determinant $a_1^2 + a_2^2 = |a|^2$. In particular (12.1) implies that $\det_{\mathbf{R}} T > 0$ if T is a bijection. This proves that in a complex vector space there is a natural orientation, corresponding to orienting $\mathbf{C}^n$ with complex coordinates $z_j = x_{2j-1} + ix_{2j}$, $j = 1, \dots, n$, by $dx_1 \wedge \dots \wedge dx_{2n} > 0$.

If now $f \in C^1(U, V)$ where Ω is open in a finite dimensional complex vector space U, then f is called analytic if f' is complex linear. Then it follows that $f \in C^\infty$. We can now define the notion of analytic manifold by demanding in Definition 9.1 that $X_i' \subset \mathbf{C}^n$ and that the maps (9.1) are analytic. By the observations above every analytic manifold has a natural orientation.

[1] For more details we refer to [C] and to [H].

The differential of any function $f \in C^1(\Omega, \mathbf{C})$ where Ω is an open subset of $\mathbf{C}^n$ can be split into a complex linear term ∂f and an antilinear term $\bar\partial f$,

$$\partial f = \sum_1^n \partial f/\partial z_j dz_j, \quad \bar\partial f = \sum_1^n \partial f/\partial \bar z_j d\bar z_j, \quad \text{where}$$
$$\partial/\partial z_j = \tfrac12(\partial/\partial x_{2j-1} - i\partial/\partial x_{2j}), \quad \partial/\partial \bar z_j = \tfrac12(\partial/\partial x_{2j-1} + i\partial/\partial x_{2j}).$$

The chain rule remains valid for analytic maps since they have complex linear differentials. Thus $\partial f^*u = f^*\partial u$ and $\bar\partial f^*u = f^*\bar\partial u$ when f is analytic.

Every differential form in Ω of degree k has a unique decomposition as a sum of forms of bidegree p, q where $p+q=k$, that is, forms which can be written

$$\sum_{|I|=p,|J|=q} u_{I,J} dz^I \wedge d\bar z^J. \tag{12.2}$$

The decomposition is invariant under pullback by analytic maps. The differential du of a form of bidegree p, q is the sum of a form ∂u of bidegree $p+1, q$ and a form $\bar\partial u$ of bidegree $p, q+1$, obtained by applying ∂ and $\bar\partial$ to the coefficients in (12.2). Since $d^2 = (\partial + \bar\partial)^2 = 0$ we have

$$\partial^2 = 0, \quad \bar\partial^2 = 0, \quad \partial\bar\partial + \bar\partial\partial = 0.$$

Note that it follows that $i\partial\bar\partial u$ is a real two form if u is a real valued C^2 function.

End of proof of Theorem 12.7. Let p be the natural map $\mathbf{C}^{n+1}\setminus\{0\} \to P^n_{\mathbf{C}}$. We shall prove that there is a real two form u on $P^n_{\mathbf{C}}$ such that

$$p^*u = (i/2\pi)\partial\bar\partial \log |z|^2.$$

It is obvious that this defines u uniquely. Set $F_i(p(z)) = \log |z/z_i|^2$ when $z \in \mathbf{C}^{n+1}$ and $z_i \ne 0$. This makes F_i well defined in the open subset U_i of $P^n_{\mathbf{C}}$ where $z_i \ne 0$ for the homogeneous coordinates. Since $p^*F_i = \log |z|^2 - \log |z_i|^2$ and $\partial\bar\partial \log |z_i|^2 = 0$ when $z_i \ne 0$ we have when $z_i \ne 0$

$$\partial\bar\partial \log |z|^2 = \partial\bar\partial p^* F_i = p^*\partial\bar\partial F_i.$$

In $U_i \cap U_j$ we have $F_i - F_j = \log |z_j/z_i|^2$, and since z_j/z_i is analytic in $U_i \cap U_j$ it follows that

$$\partial\bar\partial F_i = \partial\bar\partial F_j \quad \text{in } U_i \cap U_j.$$

Hence a closed real two form u is defined in $P^n_{\mathbf{C}}$ by

$$u = (i/2\pi)\partial\bar\partial F_i \quad \text{in } U_i,$$

and we have $p^*u = (i/2\pi)\partial\bar\partial \log |z|^2$.

We shall now calculate $\int u^n$ over $P^n_{\mathbf{C}}$. Since $P^n_{\mathbf{C}} \setminus U_{n+1}$ is a null set, it is equal to the integral over U_{n+1}, and the map $\mathbf{C}^n \ni z \mapsto p(z,1)$ is an analytic bijection $\mathbf{C}^n \to U_{n+1}$. The pullback of u under this map is $i/2\pi$ times

$$\partial\bar\partial \log(1+|z|^2) = (1+|z|^2)^{-1}\sum_1^n dz_j \wedge d\bar z_j - (1+|z|^2)^{-2}\sum_1^n \bar z_k dz_k \wedge \sum_1^n z_j d\bar z_j.$$

Since differential forms of even degree commute, we can calculate the nth power using the binomial theorem. The square of the second term is 0, so we can drop terms where it occurs to a power greater than 1. Since dz_j is everywhere else paired with $d\bar{z}_j$ we may also replace the second term by $(1+|z|^2)^{-2}\sum_1^n |z_j|^2 dz_j \wedge d\bar{z}_j$ without changing the nth power. Hence the pullback of u^n to $\mathbf{C}^n$ is

$$(i/2\pi)^n(n!(1+|z|^2)^{-n} - n(n-1)!(1+|z|^2)^{-n+1}|z|^2(1+|z|^2)^{-2})\tau$$
$$= (i/2\pi)^n n!(1+|z|^2)^{-n-1}\tau, \quad \tau = dz_1 \wedge d\bar{z}_1 \wedge \cdots \wedge dz_n \wedge d\bar{z}_n.$$

We integrate over $\mathbf{C}^n$ by replacing each complex coordinate by polar coordinates, noting that $dz_j \wedge d\bar{z}_j = -2i dx_{2j-1} \wedge dx_{2j}$, which gives

$$\int u^n = n!\pi^{-n}\int_{\mathbf{R}^{2n}} (1+|x|^2)^{-n-1}\, dx_1 \wedge \cdots \wedge dx_{2n}$$
$$= n!\pi^{-n}\int_0^\infty \cdots \int_0^\infty (1+r_1^2+\cdots+r_n^2)^{-n-1} d(\pi r_1^2)\ldots d(\pi r_n^2) = 1.$$

This completes the proof of Theorem 12.7. At the same time we have proved:

THEOREM 12.8. *If U is the differential form on $P^n_{\mathbf{C}}$ which pulled back to $\mathbf{C}^{n+1}\backslash 0$ is the form*

$$(i/2\pi)(|z|^{-2}\sum_1^{n+1} dz_j \wedge d\bar{z}_j - |z|^{-4}\sum_1^{n+1}\bar{z}_j dz_j \sum_1^{n+1} z_j d\bar{z}_j),$$

then

$$\int_{P^k} U^j = \delta_{jk}$$

for every subspace P^k of dimension k. Here δ is the Kronecker delta.

PROOF. The homotopy invariance shows that it suffices to prove this for the coordinate planes, and then it is a consequence of what we have just verified.

REMARK. Theorems 12.7 and 12.8 prove that every closed differential form v on $P^n_{\mathbf{C}}$ is cohomologous to

$$\sum_0^n a_k U^k \quad \text{where } a_k = \int_{P^k} v.$$

References

[C] S.-s. Chern, *Complex manifolds without potential theory*, Springer Verlag, 1979 (1968).
[H] L. Hörmander, *Introduction to complex analysis in several variables*, North Holland, 1990.

CHAPTER XIII

DE RHAM'S THEOREM AND POINCARÉ DUALITY

The de Rham cohomology is defined as the quotient of two infinite dimensional spaces. Nevertheless we found in Chapter XII that it is finite dimensional in all the cases discussed there. We shall now give a general method for computing the cohomology which will prove that this is a quite general fact.

Let X be a compact C^∞ manifold. We want to give a constructive method for deciding if a form $u \in \lambda^*(X)$ is *exact*, that is, if $u = dv$ for some v. A necessary condition is that u is *closed*, that is, that $du = 0$. As a motivation we first consider the case where $u \in \lambda^1(X)$ is a closed form. Let U_i, $i \in I$, be open connected and simply connected subsets of X which cover X. Then there is in each U_i a function v_i with $dv_i = u$, obtained by integrating u from a fixed point in U_i. In $U_i \cap U_j$ we have $d(v_i - v_j) = 0$. If $U_i \cap U_j$ is connected and $\neq \emptyset$ it follows that $v_i - v_j = c_{ij}$ in $U_i \cap U_j$, where c_{ij} are constants with

$$(13.1) \qquad c_{ij} + c_{jk} + c_{ki} = 0 \quad \text{if } U_i \cap U_j \cap U_k \neq \emptyset; \quad c_{ij} = -c_{ji} \quad \text{if } U_i \cap U_j \neq \emptyset.$$

Assume now that there exists a function v with $dv = u$. Then $v - v_j = c_j$ in U_j where c_j is a constant. Hence

$$(13.2) \qquad c_{ij} = c_j - c_i \quad \text{when } U_i \cap U_j \neq \emptyset.$$

Conversely, if there exist such numbers c_j then $v_i + c_i = v_j + c_j$ in $U_i \cap U_j$ so there is a function v defined in X such that $v = v_j + c_j$ in U_j for every j, which implies that $dv = u$ in X. Thus we have reduced the determination of $H^1(X)$ to the study of solvability of the system of equations (13.2) when (13.1) is valid.

If one repeats the procedure for a form $u \in \lambda^2(X)$ and Poincaré's lemma is valid for U_i, then the difference is just that c_{jk} and c_i become closed 1 forms in $U_j \cap U_k$ and U_i respectively. If $U_i \cap U_j$ is simply connected we can express the fact that $dc_{ij} = 0$ and $dc_i = 0$ by writing $c_{ij} = d\gamma_{ij}$ and $c_i = d\gamma_i$ with functions γ_{ij} and γ_i. We can take $\gamma_{ij} = -\gamma_{ji}$ when $U_i \cap U_j \neq \emptyset$ and obtain from (13.1) with constants δ_{ijk} that

$$\gamma_{ij} + \gamma_{jk} + \gamma_{ki} = \delta_{ijk}$$

if $U_i \cap U_j \cap U_k$ is connected and $\neq \emptyset$. Moreover, we have by (13.2) that $\delta_{ij} = \gamma_{ij} - \gamma_j + \gamma_i$ must be constant. We have

$$(13.1)' \qquad \delta_{jkl} - \delta_{ikl} + \delta_{ijl} - \delta_{ijk} = 0 \quad \text{if } U_i \cap U_j \cap U_k \cap U_l \neq \emptyset,$$

δ_{ijk} is antisymmetric in the indices, and we must find δ_{ij} antisymmetric in the indices when $U_i \cap U_j \neq 0$ so that

$$(13.2)' \qquad \delta_{ij} + \delta_{jk} + \delta_{ki} = \delta_{ijk} \quad \text{when } U_i \cap U_j \cap U_k \neq \emptyset.$$

This is again a finite system of linear equations, and it is easy to see that it captures everything one needs to determine $H^2(X)$.

To organize a general proof along these lines it is clear that one must introduce suitable notation. We shall do so as soon as we have proved the existence of coverings which satisfy all the conditions we needed in the course of the preliminary discussion above.

LEMMA 13.1. *Let X be a compact C^∞ submanifold of $\mathbf{R}^N$ of dimension n. If $\{U_i\}$ is a covering of X formed by the intersections of X and balls in $\mathbf{R}^N$ with radii $< \varepsilon$, then all intersections $U_{i_1\dots i_k} = U_{i_1} \cap \cdots \cap U_{i_k}$ are diffeomorphic to open convex subsets of $\mathbf{R}^n$ if $\varepsilon > 0$ is sufficiently small. In particular this implies that $H^j(U_{i_1\dots i_k}) = 0$ when $j > 0$.*

PROOF. Let $\psi : \omega \to X$ be a local parametrisation of X where ω is an open convex subset of $\mathbf{R}^n$, and let ω' be a convex open relatively compact subset of ω. The inverse image in ω' of the neighborhoods in the lemma are of the form

$$\omega'_{x_0,r} = \{t \in \omega'; |\psi(t) - x_0|^2 < r^2\}$$

where $x_0 \in \mathbf{R}^N$, the norm is Euclidean, and $r < \varepsilon$. To prove that $\omega'_{x_0,r}$ is convex we first observe that if $s, t \in \omega'_{x_0,r}$ then $|s-t| < Mr$ where M is a constant independent of r and x_0. Choose a fixed $t \in \omega'_{x_0,r}$. Then

$$|\psi(s) - x_0| \le |\psi(t) - x_0| + |\psi(s) - \psi(t)| \le M'r \quad \text{if } s \in \omega',\ |s-t| < Mr,$$

and this implies that for $a \in \mathbf{R}^n$

$$\tfrac{1}{2} \sum_{j,k=1}^{n} a_j a_k \partial^2 |\psi(s) - x_0|^2 / \partial s_j \partial s_k = |\psi'(s)a|^2 + O(r)|a|^2 \ge c|a|^2$$

with a constant $c > 0$, provided that r is small enough. Since

$$\omega'_{x_0,r} = \{s \in \omega'; |s - t| < Mr, |\psi(s) - x_0| < r\}$$

it follows that $\omega'_{x_0,r}$ is a convex set. Since we can cover X by finitely many coordinate patches $\psi(\omega')$ with the properties above, and all $U_{i_1}, \dots, U_{i_k}$ with $U_{i_1\dots i_k} \ne \emptyset$ must be completely covered by some $\psi(\omega')$, the lemma is proved.

Let $\{U_i\}_{i\in I}$ where I is a finite set be a fixed covering of X with the properties stated at the end of the lemma. (Such a covering is called *acyclic*.) Motivated by the introductory remarks in this chapter we introduce the following notation:

a) $C^p(\lambda^q)$ where $p \ge 0$, $q \ge 0$, the vector space of p cochains with values in λ^q, consists of all $|I|^{p+1}$ tuples $c = (c_s)$ where $s = (s_0, \dots, s_p) \in I^{p+1}$ and $c_s \in \lambda^q(U_s)$, $U_s = U_{s_0} \cap \cdots \cap U_{s_p}$, interpreted as 0 if $U_s = \emptyset$, such that

$$c_{\pi s} = \operatorname{sgn} \pi\, c_s$$

when π is a permutation of s. Thus it suffices to specify c_s when $s_0 < s_1 < \cdots < s_p$; in particular $c_s = 0$ when two indices are equal.

b) $C^p(\mathcal{Z}^q)$ is the set of cochains $c = (c_s) \in C^p(\lambda^q)$ with $dc_s = 0$ for every s. In particular, $C^p(\mathcal{Z}^0)$ consists of cochains with constant real coefficients.

These definitions are clearly motivated by (13.1) and (13.1)$'$. In view of (13.2) and (13.2)$'$ we also introduce

c) a map $\delta : C^p(\lambda^q) \to C^{p+1}(\lambda^q)$ defined by

$$(\delta c)_s = \sum_{j=0}^{p+1} (-1)^j c_{s_0,\dots,\widehat{s_j},\dots,s_{p+1}}, \quad \text{in } U_s,\ s \in I^{p+2},$$

where $\widehat{s_j}$ means omission of s_j. (The equation is vacuous when $U_s = \emptyset$.)

It is clear that δ maps $C^p(\mathcal{Z}^q)$ to $C^{p+1}(\mathcal{Z}^q)$. An elementary verification left for the reader shows that $\delta\delta c = 0$ when $c \in C^p(\lambda^q)$, so we have a complex

$$0 \longrightarrow C^0(\lambda^q) \xrightarrow{\ \delta\ } C^1(\lambda^q) \xrightarrow{\ \delta\ } \dots$$

which we denote by $\mathcal{C}(\lambda^q)$. (In the same way we define the subcomplex $\mathcal{C}(\mathcal{Z}^q)$.) The cohomology is easy to determine at C^0. In fact, if $c = (c_i) \in C^0(\lambda^q)$ and $\delta c = 0$ then $c_j - c_i = 0$ in $U_i \cap U_j$ which means that there is a form $u \in \lambda^q(X)$ with $u = c_i$ in U_i for every i. Hence we can identify $H^0(\mathcal{C}(\lambda^q))$ with $\lambda^q(X)$. Similarly $H^0(\mathcal{C}(\mathcal{Z}^q)) \cong \{u \in \lambda^q(X); du = 0\}$.

For $q \geq 0$ we now consider the commutative diagram

$$\begin{tikzcd}
 & 0 \arrow[d] & 0 \arrow[d] & \\
0 \arrow[r] & C^0(\mathcal{Z}^q) \arrow[r] \arrow[d] & C^1(\mathcal{Z}^q) \arrow[r] \arrow[d] & \dots \\
0 \arrow[r] & C^0(\lambda^q) \arrow[r] \arrow[d, "d"'] & C^1(\lambda^q) \arrow[r] \arrow[d, "d"'] & \dots \\
0 \arrow[r] & C^0(\mathcal{Z}^{q+1}) \arrow[r] \arrow[d] & C^1(\mathcal{Z}^{q+1}) \arrow[r] \arrow[d] & \dots \\
 & 0 & 0 &
\end{tikzcd}$$

The vertical maps are given by the inclusion of $\mathcal{C}(\mathcal{Z}^q)$ in $\mathcal{C}(\lambda^q)$ and the exterior differential operator $\mathcal{C}(\lambda^q) \xrightarrow{d} \mathcal{C}(\mathcal{Z}^{q+1})$. Since $H^{q+1}(U_s) = 0$ by hypothesis the columns are exact. Hence we obtain (see Chapter XI) an exact sequence

$$\begin{aligned}
0 \longrightarrow H^0(\mathcal{C}(\mathcal{Z}^q)) \longrightarrow H^0(\mathcal{C}(\lambda^q)) & \\
\xrightarrow{d} H^0(\mathcal{C}(\mathcal{Z}^{q+1})) \longrightarrow H^1(\mathcal{C}(\mathcal{Z}^q)) \longrightarrow H^1(\mathcal{C}(\lambda^q)) \dots \longrightarrow H^k(\mathcal{C}(\lambda^q)) & \\
\xrightarrow{d} H^k(\mathcal{C}(\mathcal{Z}^{q+1})) \longrightarrow H^{k+1}(\mathcal{C}(\mathcal{Z}^q)) \longrightarrow H^{k+1}(\mathcal{C}(\lambda^q)) \longrightarrow \dots . &
\end{aligned}$$

We shall prove later on:

LEMMA 13.2. *For every $k \geq 1$ we have $H^k(\mathcal{C}(\lambda^q)) = 0$.*

Accepting the lemma without proof for a moment we conclude from the first line in the exact sequence that

$$H^1(\mathcal{C}(\mathcal{Z}^q)) \cong H^0(\mathcal{C}(\mathcal{Z}^{q+1}))/dH^0(\mathcal{C}(\lambda^q)),$$

and since $H^0(\mathcal{C}(\mathcal{Z}^{q+1}))$ can be identified with closed $q+1$ forms in X and $H^0(\mathcal{C}(\lambda^q))$ can be identified with arbitrary q forms in X, we obtain

$$H^1(\mathcal{C}(\mathcal{Z}^q)) \cong H^{q+1}(X).$$

This conclusion is just an elaboration of the beginning of the motivating discussion at the beginning of the chapter. Furthermore we obtain for every $k \geq 1$, corresponding to the end of that discussion,

$$H^k(\mathcal{C}(\mathcal{Z}^{q+1})) \cong H^{k+1}(\mathcal{C}(\mathcal{Z}^q)).$$

Hence it follows that

$$H^{q+1}(X) \cong H^1(\mathcal{C}(\mathcal{Z}^q)) \cong H^2(\mathcal{C}(\mathcal{Z}^{q+1})) \cong \cdots \cong H^{q+1}(\mathcal{C}(\mathcal{Z}^0)), \quad q \geq 0.$$

But $\mathcal{Z}^0$ consists of cochains with real (constant) coefficients. Hence we have proved:

Theorem 13.3 (de Rham). *If $\{U_i\}$ is an acyclic covering of the compact C^∞ manifold X then $H^q(X)$ is equal to the cohomology in degree q of the complex of real cochains.*

Corollary 13.4. *If X is a compact C^∞ manifold then $H^q(X)$ is a finite dimensional vector space.*

Proof of Lemma 13.2. Let $c = (c_s) \in C^k(\lambda^q)$ and assume that $\delta c = 0$. Choose a partition of unity $\varphi_j \in C_0^\infty(U_j)$, thus $\sum_j \varphi_j = 1$, and set

$$c'_s = \sum_j \varphi_j c_{j,s}, \quad s \in I^k.$$

We define $\varphi_j c_{j,s}$ as 0 outside supp φ_j which gives a smooth form in U_s since $\varphi_j \in C_0^\infty(U_j)$, so we have defined a cochain $c' \in C^{k-1}(\lambda^q)$. For $s \in I^{k+1}$ we have in U_s

$$(\delta c')_s = \sum_j \sum_{i=0}^{k} \varphi_j (-1)^i c_{j,s_0,\dots,\widehat{s_i},\dots,s_k} = \sum \varphi_j c_s = c_s,$$

where we have used the hypothesis that $\delta c = 0$. This proves the lemma.

Remark. Even if X is not a manifold but for example an arbitrary compact space, and if R is an arbitrary ring, one can for open coverings $\{U_i\}$ of X form the complex of cochains with values in R corresponding to the covering. When the covering is refined indefinitely one can prove that the cohomology of this complex converges to a ring $H^*(X, R)$ and prove that the results in Chapter XI are valid for it with minor modifications. By suitable choice of R one can obtain more information about X than when $R = \mathbf{R}$. For example, $H^*(P^n_{\mathbf{R}}, \mathbf{Z}_2)$ where $\mathbf{Z}_2$ denotes the integers mod 2 is the polynomial ring in one variable with coefficients in $\mathbf{Z}_2$, truncated at degree n, in analogy to the cohomology $H^*(P^n_{\mathbf{C}}, \mathbf{R})$ determined in Theorem 12.7, while the de Rham cohomology $H^*(P^n_{\mathbf{R}}, \mathbf{R})$ was fairly uninteresting. For a general discussion of cohomology we refer to [G].

We shall now return to the study of the de Rham cohomology of a compact C^∞ manifold X. The proof of Theorem 13.3 will be used to prove some important facts on the solvability of the equation $dv = u$.

Theorem 13.4. *Let X be a compact, oriented C^∞ manifold, and let $u \in \lambda^*(X)$. Then $u = dv$ for some $v \in \lambda^*(X)$ if and only if*

$$\int u \wedge w = 0 \quad \textit{for every } w \in \lambda^*(X) \textit{ with } dw = 0.$$

Before the proof we shall give an equivalent statement of the theorem. If u and w are closed differential forms then $\int u \wedge w$ only depends on the cohomology classes a and b of u and w, for they determine the cohomology class of $u \wedge w$. Thus we can write

$$\int u \wedge w = \langle a, b \rangle$$

where $\langle a, b\rangle$ is a bilinear form on $H^*(X) \times H^*(X)$. With this notation Theorem 13.4 states that if $\langle a, b \rangle = 0$ for every $b \in H^*(X)$ then $a = 0$, that is, the bilinear form is non-degenerate so it establishes a duality between $H^k(X)$ and $H^{n-k}(X)$ for $0 \le k \le n$. We state this as a theorem:

Theorem 13.5 (Poincaré's duality theorem). *Let X be a compact oriented C^∞ manifold. Then the bilinear form*

$$\lambda^*(X) \times \lambda^*(X) \ni (u, w) \mapsto \int_X u \wedge w$$

induces a non-singular bilinear form on $H^*(X) \times H^*(X)$ which makes $H^k(X)$ and $H^{n-k}(X)$ dual spaces (hence of the same dimension). If $n = 2k$ then H^k is self dual, and $\langle a, b\rangle = (-1)^{k^2}\langle b, a\rangle$ when $a, b \in H^k(X)$. When k is even the form is thus symmetric and when k is odd it is skew symmetric, so $\dim H^k$ *is even then.*

Example. If $X = P^n_{\mathbf{C}}$ and $a = \sum_0^n a_j x^j$, $b = \sum_0^n b_j x^j$ where $x \in H^2(X)$ is the class defined by U in Theorem 12.8, then $\langle a, b\rangle = \sum_0^n a_j b_{n-j}$.

To prove Theorem 13.4 we shall verify inductively a closely related statement which easily gives Theorem 13.4. Using the same covering as in the proof of Theorem 13.3 we denote by $C^p(\lambda_c^{n-q})$ the space of p cochains $c = (c_s)$ where $c_s \in \lambda_c^{n-q}(U_s)$. If $u \in C^p(\lambda^q)$ and $v \in C^p(\lambda_c^{n-q})$ we define

$$\langle u, v\rangle = \sum_s \int_{U_s} u_s \wedge v_s.$$

If $u \in C^p(\lambda^{q-1})$ and $v \in C^p(\lambda_c^{n-q})$ then

$$\langle du, v\rangle = (-1)^q \langle u, dv\rangle,$$

for $\int (du_s) \wedge v_s + (-1)^{q-1} \int u_s \wedge dv_s = \int d(u_s \wedge v_s) = 0$. Moreover,

$$\langle \delta u, v \rangle = \langle u, \delta^* v\rangle, \quad u \in C^p(\lambda^q),\ v \in C^{p+1}(\lambda_c^{n-q}), \quad \text{where}$$
$$(\delta^* w)_s = \sum_j w_{j,s}, \quad s \in I^{p+1}.$$

The verification is left as an exercise. We can now formulate a variant of Theorem 13.4 which is adapted to the inductive proof of Theorem 13.3.

THEOREM 13.4′. *If* $u \in C^p(\mathcal{Z}^q)$, $p \geq 1$, *then* $u = \delta v$ *for some* $v \in C^{p-1}(\mathcal{Z}^q)$ *if and only if* $\langle u, w \rangle = 0$ *for all cochains* $w \in C^p(\lambda_c^{n-q})$ *with* $\delta^* w \in dC^{p-1}(\lambda_c^{n-q-1})$.

PROOF. The necessity is obvious, for if $u = \delta v$ then $\langle u, w \rangle = \langle v, \delta^* w \rangle$. If $q = 0$ then the sufficiency is also elementary, for $u = (u_s)$ is then a real cochain and the condition in the theorem means in view of Theorem 12.2 that

$$\sum u_s \int w_s = 0$$

for all (w_s) with $\int (\delta^* w)_s = 0$ for every s. But this means precisely that $\sum u_s W_s = 0$ for all real (W_s) such that this is true whenever $u = \delta v$.

We may now assume that $1 \leq q \leq n$ and that the theorem has been proved for smaller values of q. If we write $w = \delta^* w'$ where $w' \in C^{p+1}(\lambda_c^{n-q})$ the hypothesis gives $\langle \delta u, w' \rangle = 0$, hence $\delta u = 0$. Since $q > 0$ we can by Poincaré's lemma write $u = dv$ where $v \in C^p(\lambda^{q-1})$. Then we have $0 = \delta u = d\delta v$, so $\delta v \in C^{p+1}(\mathcal{Z}^{q-1})$. If $w \in C^{p+1}(\lambda_c^{n-q+1})$ and $\delta^* w = dW$, $W \in C^p(\lambda_c^{n-q})$, then

$$\langle \delta v, w \rangle = \langle v, \delta^* w \rangle = \langle v, dW \rangle = (-1)^q \langle dv, W \rangle = (-1)^q \langle u, W \rangle.$$

We have $d\delta^* W = \delta^* \delta^* w = 0$, and since $\delta^* W$ is a cochain of compactly supported differential forms of degree $< n$, it follows from Theorem 12.2 that $\delta^* W \in dC^{p-1}(\lambda_c^{n-q-1})$. Hence $\langle u, W \rangle = 0$ by hypothesis, so the inductive hypothesis gives that $\delta v = \delta u'$ where $u' \in C^p(\mathcal{Z}^{q-1})$. In view of Lemma 13.2 it follows that $v - u' = \delta V$ where $V \in C^{p-1}(\lambda^{q-1})$, which implies that $\delta dV = d\delta V = dv = u$, and since $dV \in C^{p-1}(\mathcal{Z}^q)$, the theorem is proved.

PROOF THAT THEOREM 13.4′ IMPLIES THEOREM 13.4. The proof of Theorem 13.4′ can be continued to prove that Theorem 13.4 follows from Theorem 13.4′. The necessity is obvious in Theorem 13.4 too. If u satisfies the conditions in the theorem we conclude that $du = 0$ by taking $w = dw'$. If u is a 0 form, this means that u is a constant in every component of X, and since every n form is closed we can then conclude that $u = 0$. Thus we may assume that $u \in \lambda^q(X)$ where $q > 0$. By Poincaré's lemma there is a $q-1$ form v_i in U_i such that $u = dv_i$ in U_i. These forms define $v \in C^0(\lambda^{q-1})$, and $d\delta v = \delta dv = 0$. We shall prove that δv satisfies the hypotheses of Theorem 13.4′. Let $q \in C^1(\lambda_c^{n-q+1})$ and assume that $\delta^* w = dW$ where $W \in C^0(\lambda_c^{n-q})$. Then we have

$$\langle \delta v, w \rangle = \langle v, \delta^* w \rangle = \langle v, dW \rangle = (-1)^q \langle dv, W \rangle = (-1)^q \int u \wedge \textstyle\sum_i W_i.$$

Since $dW_i = \sum_j w_{ji}$ and $w_{ji} = -w_{ij}$, we have $d(\sum_i W_i) = 0$, so the hypothesis in the theorem gives $\langle \delta v, w \rangle = 0$. By Theorem 13.4′ we can therefore find $u' \in C^0(\mathcal{Z}^{q-1})$ with $\delta u' = \delta v$. But then $v - u'$ defines a form $V \in \lambda^{q-1}(X)$ with $dV = u$, which completes the proof.

REMARK. Note that the proof of Theorem 13.4′ and the proof that Theorem 13.4 follows from Theorem 13.4′ are parallel to two parts of the proof of Theorem 13.3 — the reading of the exact sequence far away and at the beginning, respectively.

In Chapter XIV we shall have to study the solvability of the equation $dv = u$ when u depends on parameters, and we shall now make some preparations for that. Let ω be an open subset of $\mathbf{R}^\nu$. A form $u_t \in \lambda^*(X)$ depending on the parameter $t \in \omega$ is said to be a C^∞ function of the parameter if in every local coordinate system $(x_1, \ldots, x_n)$ on X one can write $u_t = \sum a_I(t, x) dx_{i_1} \wedge \cdots \wedge dx_{i_k}$ with $a_I(t, x) \in C^\infty$ as a function of (t, x).

THEOREM 13.6. *Let X be a compact C^∞ manifold. If $u_t \in \lambda^*(X)$ is a C^∞ function of the parameter $t \in \omega \subset \mathbf{R}^\nu$ and if for every $t \in \omega$ there is a form v_t such that $dv_t = u_t$, then one can choose v_t as a C^∞ function of $t \in \omega$.*

Just as in the proof of Theorem 13.4 we must first prove a related result adapted to the inductive proof of Theorem 13.3:

THEOREM 13.6$'$. *Let X be a compact C^∞ manifold. If $u_t \in C^p(\mathcal{Z}^q)$, $p \geq 1$, is a C^∞ function of $t \in \omega \subset \mathbf{R}^\nu$, and if $u_t = \delta v_t$ for some $v_t \in C^{p-1}(\mathcal{Z}^q)$ when $t \in \omega$, then v_t can be chosen as a C^∞ function of $t \in \omega$.*

The proof requires two simple lemmas:

LEMMA 13.7. *Let U be an open convex set in $\mathbf{R}^n$, and let $u_t \in \lambda^q(U)$, $q > 0$, be a C^∞ function of the parameter $t \in \omega$ with $du_t = 0$. Then $u_t = dv_t$ where v_t can be chosen as a C^∞ function of $t \in \omega$.*

PROOF. This follows at once from the construction of v_t in the proof of Theorem 8.1.

LEMMA 13.8. *Let $u_t \in C^p(\lambda^q)$, $p > 0$, be a C^∞ function of the parameter $t \in \omega$ such that $\delta u_t = 0$. Then we have $u_t = \delta v_t$ where $v_t \in C^{p-1}(\lambda^q)$ can be chosen as a C^∞ function of $t \in \omega$.*

PROOF OF THEOREM 13.6$'$. The theorem is true when $q = 0$, for if $A : V_1 \to V_2$ is a linear transformation between finite dimensional vector spaces, then there exists a linear map $B : V_2 \to V_1$ such that $ABu = u$ when $u \in \operatorname{Im} A$. The solution $v = Bu$ of the equation $Av = u$ is then a C^∞ function of u.

Now assume that $q > 0$ and that the theorem has already been proved for smaller values of q. By Lemma 13.7 we can write $u_t = dV_t$ where $V_t \in C^p(\lambda^{q-1})$ is a C^∞ function of $t \in \omega$. Since $d\delta V_t = \delta dV_t = \delta u_t = 0$ it follows that δV_t is a C^∞ function of $t \in \omega$ with values in $C^{p+1}(\mathcal{Z}^{q-1})$. Since u_t is cohomologous to 0 in the complex $\mathcal{C}(\mathcal{Z}^q)$ we know from the proof of Theorem 13.3 that δV_t is also cohomologous to 0 in the complex $\mathcal{C}(\mathcal{Z}^{q-1})$. By the inductive hypothesis it follows that $\delta V_t = \delta u'_t$ where $u'_t \in C^p(\mathcal{Z}^{q-1})$ is a C^∞ function of $t \in \omega$. By Lemma 13.8 we can now write $V_t - u'_t = \delta W_t$ where $W_t \in C^{p-1}(\lambda^{q-1})$ is a C^∞ function of $t \in \omega$. This implies that $\delta dW_t = dV_t = u_t$, so $v_t = dW_t$ has the desired properties.

PROOF THAT THEOREM 13.6$'$ IMPLIES THEOREM 13.6. We may assume that $u_t \in \lambda^q(X)$ where $q > 0$. By Lemma 13.7 we have $u_t = dV_{t,i}$ in U_i where $V_t \in C^0(\lambda^{q-1})$ is a C^∞ function of $t \in \omega$. As in the proof of Theorem 13.3 we obtain using Theorem 13.6$'$ that $\delta V_t = \delta u'_t$ where $u'_t \in C^0(\mathcal{Z}^{q-1})$ is a C^∞ function of $t \in \omega$. This implies that $V_t - u'_t$ defines a form $v_t \in \lambda^{q-1}(X)$ which is a C^∞ function of $t \in \omega$, and since $dv_t = u_t$ this completes the proof.

REMARK. The proof gives easily that there is a linear continuous map $E : d\lambda^*(X) \to \lambda^*(X)$ with $dEf = f$ when $f \in d\lambda^*(X)$. However, the notation becomes somewhat heavier in the proof so it will not be given here.

We shall now give two supplements to the Poincaré duality theorem which describe the modifications needed for manifolds which are not compact or not orientable. We shall begin with the second problem.

Let X be a compact C^∞ manifold. As in the proof of Theorem 12.4 we form a new manifold $\widetilde{X}$ consisting of pairs (x, o) where $x \in X$ and o is an orientation of the tangent space $T_x(X)$. We set $p(\tilde{x}) = x$ if $\tilde{x} = (x, o)$. For a neighborhood U_i as in Lemma 13.1 we obtain for the two choices of orientation two right inverses $U_i \to \widetilde{X}$ of p, and they define a differentiable structure in $\widetilde{X}$. For the map $p : \widetilde{X} \to X$ the set $p^{-1}(x)$ has precisely two points for every $x \in X$. If X is connected and orientable then $\widetilde{X}$ has two components diffeomorphic with X, but if X is connected but not orientable then $\widetilde{X}$ is connected. By the definition $\widetilde{X}$ is oriented in any case.

Let $r : \widetilde{X} \to \widetilde{X}$ be the map corresponding to a change of the orientation. Thus $r \circ r$ is the identity, $p \circ r = p$, and $r(\tilde{x}) \neq \tilde{x}$ for every $\tilde{x} \in \widetilde{X}$. If u is a differential form on X then $v = p^*u$ is a form on $\widetilde{X}$ with $r^*v = v$. Conversely, if v is a form on $\widetilde{X}$ with $r^*v = v$ then $v = p^*u$ where u is a form on X which is uniquely defined. This was explained in the proof of Theorem 12.4.

A closed form u on X is exact if and only if p^*u is exact. In fact, if $u = dv$ then $p^*u = dp^*v$. On the other hand, if $p^*u = dV$ then $p^*u = r^*p^*u = dr^*V = dV_1$ where $V_1 = \frac{1}{2}(V + r^*V)$. Since $r \circ r$ is the identity we have $r^*V_1 = V_1$, hence $V_1 = p^*v$ where v is a form on X with $u = dv$.

By Theorem 13.4 a form u on X is thus exact if and only if

$$\int_{\widetilde{X}} (p^*u) \wedge V = 0$$

for all closed forms V on $\widetilde{X}$. Now we have

$$\int_{\widetilde{X}} (p^*u) \wedge V = -\int r^*((p^*u) \wedge V) = -\int (p^*u) \wedge r^*V$$

since r reverses the orientation of $\widetilde{X}$. Hence it follows that

$$\int (p^*u) \wedge V = \int (p^*u) \wedge V' \quad \text{where } V' = \tfrac{1}{2}(V - r^*V), \text{ thus } r^*V' = -V'.$$

For V' we therefore have a symmetry which is opposite to that for forms lifted from X. Every form V on $\widetilde{X}$ can uniquely be written as $V = V^+ + V^-$ where $r^*V^+ = V^+$ and $r^*V^- = -V^-$, for this is equivalent to $V^+ = \frac{1}{2}(V + r^*V)$ and $V^- = \frac{1}{2}(V - r^*V)$. This decomposition commutes with the exterior differential operator d since d commutes with r^*.

Definition 13.9. By the twisted cohomology $H_t^*(X)$ for X one means the cohomology of the complex of odd forms V on $\widetilde{X}$, that is, forms with $r^*V = -V$.

There is an obvious injection $H_t^*(X) \to H^*(\widetilde{X})$ so $H_t^*(X)$ can be viewed as a subset of $H^*(\widetilde{X})$. If u is a form on X and V is a form on $\widetilde{X}$ with $r^*V = -V$ then the bilinear form

$$(u, V) \mapsto \tfrac{1}{2} \int (p^*u) \wedge V$$

induces a bilinear form on $H^*(X) \times H_t^*(X)$ which by the discussion above is non-singular. If X is orientable then we can identify $H_t^*(X)$ with $H^*(X)$ and we have just recovered the duality in Theorem 13.5. However, we have now proved with no hypothesis on orientability:

THEOREM 13.10. *If X is a compact C^∞ manifold then $H^*(X)$ and $H^*_t(X)$ are dual.*

Combining the result and its proof with Theorem 13.6 we obtain:

THEOREM 13.11. *Let X be a compact C^∞ manifold, and let $\alpha_1, \dots, \alpha_N$ be closed forms on X such that their cohomology classes form a basis for $H^*(X)$. If $u_t \in \lambda^*(X)$ is a C^∞ function of t and $du_t = 0$ for every t, then we can write*

$$u_t = \sum_1^N a_j(t)\alpha_j + dv_t$$

where $a_j \in C^\infty$ and $v_t \in \lambda^*(X)$ is a C^∞ function of t.

PROOF. If $A_1, \dots, A_N$ are odd closed forms on $\widetilde{X}$ with cohomology classes in $H^*_t(X)$ forming a basis which is biorthogonal to the classes of $\alpha_1, \dots, \alpha_N$ in $H^*(X)$, then we must have

$$a_j(t) = \tfrac{1}{2} \int_{\widetilde{X}} p^* u_t \wedge A_j,$$

which is a C^∞ function of t. Now $u_t - \sum_1^N a_j(t)\alpha_j$ is exact, so the theorem follows from Theorem 13.6.

We shall now discuss the cohomology of a non-compact C^∞ manifold X. For the sake of simplicity we assume that X is oriented and leave for the reader to state and prove an anlogue of Theorem 13.12 below for the case where X is not oriented. The case of main interest to us is that which occurred in Theorem 11.9, where there exists a compact C^∞ manifold Y and a compct submanifold Z such that X is diffeomorphic to $Y \setminus Z$. If T is a sufficiently small tubular neighborhood of Z then X is also diffeomorphic to $Y \setminus T$, by Theorem 10.6, which makes the following theorem applicable.

THEOREM 13.12. *Let X be diffeomorphic to a relatively compact open subset of an oriented manifold M and assume that the boundary ∂X is C^∞ and of codimension 1. Then $H^*_c(X)$ and $H^*(X)$ are dual with respect to the bilinear form induced by $\int_X u \wedge v$ when $u \in \lambda^*_c(X)$ and $v \in \lambda^*(X)$, $du = dv = 0$.*

Note that Theorem 12.2 is a consequence of Theorem 12.1 and Theorem 13.12. As just observed, the hypotheses of Theorem 13.12 are fulfilled if $X = Y \setminus Z$ where Z is a compact C^∞ submanifold of a compact oriented C^∞ manifold Y.

PROOF. An open neighborhood of ∂X is diffeomorphic to $\partial X \times (-1, 1)$ and we identify it with $\partial X \times (-1, 1)$. We may assume that $\partial X \times (-1, 0) \subset X$, that $\partial X \times 0 = \partial X$ and that $M = X \cup (\partial X \times [0, 1))$. We form a new manifold $\widetilde{M}$ consisting of two copies of M where we identify a point $(\xi, t) \in \partial X \times (-1, 1)$ in one copy with $(\xi, -t)$ in the other copy. Then we obtain an oriented manifold $\widetilde{M}$ where ∂X is a submanifold Y and $\widetilde{M} \setminus Y$ consists of two copies X^+ and X^- of X, with X^+ having the orientation of X and X^- the opposite orientation. On $\widetilde{M}$ we have a map $r : \widetilde{M} \to \widetilde{M}$ with $r \circ r$ equal to the identity which maps X^+ to X^- and leaves every point in Y fixed. If $p^+ : X^+ \to X$ and $p^- : X^- \to X$ are the natural maps then $p^+ \circ r = p^-$ and $p^- \circ r = p^+$. The map r reverses the orientation, so $\int_{\widetilde{M}} r^* u = -\int_{\widetilde{M}} u$ for $u \in \lambda^*(\widetilde{M})$.

If $u \in \lambda_c^*(X)$ we can define a form $U \in \lambda^*(M)$ by $U = p^{+*}u$ on X^+ and $U = -p^{-*}u$ on X^-. Then $r^*U = -U$, $U = 0$ in a neighborhood of Y, and $dU = 0$ if $du = 0$. If $\int_X u \wedge v = 0$ for every $v \in \lambda^*(X)$ with $dv = 0$ then $\int U \wedge V = 0$ for all $V \in \lambda^*(\widetilde{M})$ with $dV = 0$. By Poincaré's duality theorem it follows that $U = dV$ where V is a form on $\widetilde{M}$. Since

$$U = -r^*U = -r^*dV = d(-r^*V)$$

we have also $U = dV_1$ where $V_1 = \frac{1}{2}(V - r^*V)$, thus $r^*V_1 = -V_1$. If i is the embedding $Y \to \widetilde{M}$ then $r \circ i = i$, hence $i^*V_1 = i^*r^*V_1 = -i^*V_1$, so $i^*V_1 = 0$. But we proved before Definition 11.8 that this implies that there is a form $V_2 \in \lambda^*(\widetilde{M})$ vanishing in a neighborhood of Y such that $U = dV_2$. The pullback $v = (p^+)^{-1*}V_2$ is then in $\lambda_c^*(X)$ and $dv = u$, so we have proved that an element in $H_c^*(X)$ orthogonal to $H^*(X)$ must be 0.

Let us now consider an element in $H^*(X)$ which is orthogonal to $H_c^*(X)$. It is represented by a closed form $u \in \lambda^*(X)$ with $\int_X u \wedge v = 0$ for every $v \in \lambda_c^*(X)$ with $dv = 0$. Let $f : \widetilde{M} \to X$ be equal to $p^\pm$ in $\widetilde{M} \setminus (\partial X \times (-\frac{1}{2}, \frac{1}{2}))$ and define $f((y,t)) = (y, \varphi(t))$ when $y \in \partial X$ and $t \in (-1,1)$ where $\varphi \in C^\infty(\mathbf{R})$ is an increasing function of $|t|$, and $\varphi(t) = |t|$ when $|t| > \frac{1}{2}$, $\varphi(t) = \frac{1}{4}$ when $|t| < \frac{1}{5}$. Then we have $f \circ r = f$. Since u is a closed form on X it follows that f^*u is a closed form on $\widetilde{M}$, and we claim that

$$\int_{\widetilde{M}} (f^*u) \wedge V = 0 \tag{13.3}$$

for every closed form V on $\widetilde{M}$. Since f^* is homotopic to $p^\pm$ in $X^\pm$ we know that this is true for forms V which vanish in a neighborhood of ∂X. In view of the exact sequence

$$H_c^k(\widetilde{M} \setminus \partial X) \longrightarrow H^k(\widetilde{M}) \xrightarrow{i^*} H^k(\partial X)$$

we only have to prove that for every closed form V on $\widetilde{M}$ there is a closed form V' with $i^*V = i^*V'$ such that $\int_{\widetilde{M}} (f^*u) \wedge V' = 0$, for $V - V'$ is cohomologous to a form which vanishes in a neighborhood of Y. Since $i^*V = i^*r^*V$ we have $i^*V' = V$ if $V' = \frac{1}{2}(V + r^*V)$, and then it follows that

$$\int_{\widetilde{M}} (f^*u) \wedge V' = -\int_{\widetilde{M}} (r^*f^*u) \wedge r^*V' = -\int_{\widetilde{M}} (f^*u) \wedge V'$$

since r reverses the orientation and $f \circ r = f$. This proves (13.3), and it follows that f^*u is exact. Hence $(p^{+-1})^*f^*u$ is exact, and since $f \circ p^{+-1}$ is homotopic to the identity in X, it follows that u is exact. The proof is now complete.

In connection with de Rham's theorem we shall also discuss the definition of the Chern class of a complex line bundle. In analogy to Definition 9.2 a complex vector bundle L over X with fiber of type $\mathbf{C}^n$ is by definition a C^∞ manifold with a projection $p : L \to X$ such that

(i) $L_x = p^{-1}(x)$ is for every $x \in X$ a n dimensional vector space over $\mathbf{C}$;
(ii) every $x \in X$ has a neighborhood U such that there is a diffeomorphism $p^{-1}(U) \cong U \times \mathbf{C}^n$ which respects the projection and the vector structure of the fibers.

We shall only consider complex line bundles, that is, the case $n = 1$.

Choose an acyclic covering $\{U_i\}$ of X such that for every i there is a diffeomorphism

$$\psi_i : p^{-1}(U_i) \to U_i \times \mathbf{C}$$

with the properties (ii); they are called local trivialisations. The map

$$\psi_i \psi_j^{-1} : (U_i \cap U_j) \times \mathbf{C} \to (U_i \cap U_j) \times \mathbf{C}$$

consists of multiplication of the component in $\mathbf{C}$ by a function $g_{ij} \in C^\infty(U_i \cap U_j)$ with no zeros. We have

$$g_{ij} g_{jk} g_{ki} = 1 \quad \text{in } U_i \cap U_j \cap U_k; \; g_{ij} g_{ji} = 1 \quad \text{in } U_i \cap U_j; \; g_{ii} = 1 \quad \text{in } U_i.$$

The functions g_{ij} are called *transition functions*. To define a section of L is equivalent to defining in each U_i a complex valued function u_i such that $u_i = g_{ij} u_j$ in $U_i \cap U_j$ for all i, j.

We shall now examine if L is isomorphic to $X \times \mathbf{C}$, that is, if there exists a global trivialisation. In that case ψ_i could be viewed as multiplication by a function $g_i \in C^\infty(U_i)$ without zeros, and we would have $g_{ij} = g_i / g_j$. The problem is therefore to decide when there exist such functions. Since $U_j \cap U_k$ is simply connected we can choose C^∞ functions h_{jk} with $g_{jk} = \exp(-2\pi i h_{jk})$ and $h_{jk} = -h_{kj}$ when $U_j \cap U_k \neq \emptyset$, and obtain

$$h_{ij} + h_{jk} + h_{ki} = c_{ijk} \in \mathbf{Z}, \quad \text{if } U_i \cap U_j \cap U_k \neq \emptyset.$$

If the line bundle is trivial and we write $g_j = \exp(-2\pi i h_j)$ then we obtain $h_{ij} = h_i - h_j + c_{ij}$ with $c_{ij} \in \mathbf{Z}$, if $U_i \cap U_j \neq \emptyset$, and this implies

$$c_{ij} + c_{jk} + c_{ki} = c_{ijk}, \quad \text{if } U_i \cap U_j \cap U_k \neq \emptyset. \tag{13.4}$$

This means that (c_{ijk})must define the 0 cohomology class in $H^2(\mathcal{C}(\mathbf{R}))$. Conversely, if this is true and we choose a solution $c_{ij} \in \mathbf{R}$ of the preceding equations then $H_{ij} = h_{ij} - c_{ij}$ satisfies

$$H_{ij} + H_{jk} + H_{ki} = 0, \quad \text{if } U_i \cap U_j \cap U_k \neq \emptyset.$$

By Lemma 13.2 we can then write

$$H_{ij} = H_j - H_i, \quad \text{if } U_i \cap U_j \neq \emptyset,$$

where $H_i \in C^\infty(U_i)$. If $c_{ij} \in \mathbf{Z}$ this means that $g_{jk} = \exp(2\pi i H_j)/\exp(2\pi i H_k)$, so the line bundle is trivial. (In general we just conclude that the line bundle can be represented with the transition functions $\exp(-2\pi i c_{jk})$ which are constants with absolute value 1.) Apart from the distinction between integer and real solutions of (13.4) we conclude that the line bundle is trivial precisely when (c_{ijk}) defines the class 0 in $H^2(\mathcal{C}(\mathbf{R}))$. We shall now go back to the proof of de Rham's theorem and determine a differential form defining the corresponding cohomology class in $H^2(X)$. It is called the *Chern class* of the line bundle.

Choose a partition of unity $\varphi_i \in C_0^\infty(U_i)$. By the proof of Lemma 13.2 (c_{ijk}) considered as element in $C^2(\lambda^0)$ is equal to δ applied to the cochain in $C^1(\lambda^0)$ defined by

$$\sum_i \varphi_i c_{ijk} \in \lambda^0(U_j \cap U_k).$$

If we apply d to this cochain we obtain the cochain in $C^1(\mathcal{Z}^1)$

$$\sum_i c_{ijk} d\varphi_i \in \mathcal{Z}^1(U_j \cap U_k).$$

which defines the corresponding cohomology class in $H^1(\mathcal{C}(\mathcal{Z}^1))$. If we introduce the definition of c_{ijk} and use that $\sum_i d\varphi_i = 0$, the 1 cochain becomes

$$\sum_i (h_{ij} + h_{jk} + h_{ki}) d\varphi_i = \sum_i h_{ki} d\varphi_i - \sum_i h_{ji} d\varphi_i \in \mathcal{Z}^1(U_j \cap U_k),$$

which is δ applied to the cochain in $C^0(\lambda^1)$

$$\sum_i h_{ji} d\varphi_i \in \lambda^1(U_j).$$

The differential of this cochain

$$\sum_i dh_{ji} \wedge d\varphi_i \in \lambda^2(U_j)$$

is a cocycle in $C^0(\mathcal{Z}^2)$ (that is, annihilated by δ), so it is a globally defined closed two form in X. We can verify this directly, for

$$\sum_i (dh_{ji} - dh_{ki}) \wedge d\varphi_i = \sum_i dh_{jk} \wedge d\varphi_i = 0 \quad \text{in } U_j \cap U_k$$

since $\sum_i d\varphi_i = 0$. Now we have $dh_{ji} = -dh_{ij} = dg_{ij}/(2\pi i g_{ij})$, so the Chern class is defined by the differential form

$$\alpha = (2\pi i)^{-1} \sum_i (dg_{ij}/g_{ij}) \wedge d\varphi_i \quad \text{in } U_j. \tag{13.5}$$

The form (13.5) is defined for any covering of X by open sets U_i such that L is trivial in U_i for every i. We shall now prove that the cohomology class of α is independent of the choices of covering, trivialisations and partition of unity.

a) If $\{\psi_i\}$ is another partition of unity and

$$u_j = \sum_i (\varphi_i - \psi_i) dg_{ij}/(2\pi i g_{ij}) \in \lambda^1(U_j)$$

then we have

$$u_j - u_k = \sum_i (\varphi_i - \psi_i) dg_{kj}/(2\pi i g_{kj}) = 0 \quad \text{in } U_j \cap U_k$$

so the forms u_j define together a form $u \in \lambda(X)$. The the definitions of (13.5) using the two partitions of unity differ by the exact form du.

b) If we keep the covering and change the trivialisations, then they are just multiplied by functions $g_i \in C^\infty(U_i)$ with no zeros, that is, g_{ij} is replaced by $g_{ij}g_ig_j^{-1}$. This means adding to (13.5)

$$(2\pi i)^{-1}\sum_i (dg_i/g_i - dg_j/g_j)\wedge d\varphi_i = d\big((2\pi i)^{-1}\sum_i -\varphi_i dg_i/g_i\big)$$

which is an exact form.

c) Finally we shall study what happens if one refines the covering but keeps the trivialisations. Thus we take a new covering $\{V_j\}$ such that there is a map $j \to i(j)$ with $V_j \subset U_{i(j)}$ for every j. Choose a partition of unity $\chi_j \in C_0^\infty(V_j)$. Then

$$\varphi_i = \sum_{i(j)=i} \chi_j$$

is a partition of unity for the covering $\{U_i\}$, and with these partitions of unity we even get the same differential form (13.5) for the two coverings. For two arbitrary coverings one can choose a third covering which is a refinement of both, so the result is true for arbitrary coverings.

We have now completed the definition of the Chern class $c(L) \in H^2(X)$, which is defined by the form (13.5). The proof in c) above also shows that if $f : Y \to X$ is a C^∞ map then $c(f^*L) = f^*c(L)$.

Example. On the complex projective space $P^n_{\mathbf{C}}$ there is a natural complex line bundle L, for every point in $P^n_{\mathbf{C}}$ corresponds to a complex line in $\mathbf{C}^{n+1}$. For $i = 1, \dots, n+1$ let U_i be the set of points in $P^n_{\mathbf{C}}$ with the homogeneous coordinate $z_i \neq 0$. The restriction to U_i of L with the zero section removed can then be identified with $\{z \in \mathbf{C}^{n+1}; z_i \neq 0\}$, and we trivialise L in U_i by mapping it to $(p(z), z_i)$ where $p : \mathbf{C}^{n+1} \setminus \{0\} \to P^n_{\mathbf{C}}$ is the natural map. Then we obtain the transition functions z_i/z_j in $U_i \cap U_j$. The dual bundle with the reciprocal transition functions

$$g_{ij} = z_j/z_i \quad \text{in } U_i \cap U_j$$

is called the *Hopf bundle*, and we shall calculate its Chern class.

First assume that $n = 1$. Then $P^1_{\mathbf{C}}$ can be considered as $\mathbf{C}$ extended by a point at infinity. A point $z \in \mathbf{C}$ corresponds to the class of $p(z, 1)$, and infinity to $p(1, 0)$. Thus U_1 is the extended plane except 0, and U_2 is the finite plane. Choose $\varphi_2 \in C_0^\infty(\mathbf{C})$ equal to 1 in a neighborhood of 0. We regard φ_2 as a function in $C_0^\infty(U_2)$ and choose $\varphi_1 = 1 - \varphi_2 \in C_0^\infty(U_1)$. Then $g_{21} = z = 1/g_{12}$ so the Chern form (13.5) is given in the finite plane by

$$(2\pi i)^{-1}(dg_{12}/g_{12}) \wedge d\varphi_1 = (2\pi i)^{-1}(-dz/z) \wedge (-d\varphi_2)$$

which vanishes in a neighborhood of infinity. The integral is

$$(2\pi i)^{-1}\iint_{\mathbf{C}} \partial\varphi_2/\partial\bar{z} dz \wedge d\bar{z}/z = \varphi_2(0) = 1,$$

by Cauchy's integral formula. This proves that the Chern class is the natural generator of $H^2(P^1_{\mathbf{C}})$. We can extend this conclusion as follows:

THEOREM 13.13. *The Chern class of the Hopf bundle on $P^n_{\mathbf{C}}$ is the generator of $H^*(P^n_{\mathbf{C}})$ defined in Theorem 12.8.*

PROOF. The restriction to a one dimensional projective subspace is equal to the Chern class of its Hopf bundle, and we have just proved that its integral is equal to 1.

EXERCISE. Let X be a compact Riemann surface, that is, a compact one dimensional analytic manifold. For a covering of X with local analytic coordinate patches U_i we give for every i a meromorphic function g_i in U_i such that $g_{ij} = g_i/g_j$ is analytic in $U_i \cap U_j$, hence $\neq 0$ there. Prove that the Chern class of the corresponding line bundle is the generator of $H^2(X)$ times the number of zeros minus the number of poles of the given functions.

References

[G] R. Godement, *Topologie algébrique et théorie des faisceaux*, Hermann, Paris, 1958.

CHAPTER XIV

COHOMOLOGY OF PRODUCT AND FIBER SPACES

Let X and Y be compact C^∞ manifolds, not necessarily connected, but we assume that every component of X (resp. Y) has the same dimension. If $\alpha \in H^*(X)$ and $\beta \in H^*(Y)$ then we can define the *direct product* $\alpha \times \beta \in H^*(X \times Y)$ by

$$\alpha \times \beta = (\pi_X^* \alpha) \cup (\pi_Y^* \beta)$$

where $\pi_X : X \times Y \to X$ and $\pi_Y : X \times Y \to Y$ are the projections. Conversely we can recover the cup product from the direct product, for if $X = Y$ and δ denotes the diagonal map $X \ni x \mapsto (x, x) \in X \times X$, then

$$\delta^*(\alpha \times \beta) = ((\pi_X \circ \delta)^* \alpha) \cup ((\pi_Y \circ \delta)^* \beta) = \alpha \cup \beta.$$

Theorem 14.1 (Künneth's formula). *If $\alpha_1, \dots, \alpha_\mu$ is a basis for $H^*(X)$ and $\beta_1, \dots, \beta_\nu$ is a basis for $H^*(Y)$, then $\alpha_i \times \beta_j$, $i = 1, \dots, \mu$, $j = 1, \dots, \nu$, is a basis for $H^*(X \times Y)$.*

We postpone the proof until we have made some comments on this result. Since

$$(\alpha \times \beta) \cup (\alpha' \cup \beta') = (-1)^{pq} (\alpha \cup \alpha') \times (\beta \cup \beta')$$

if α' is of degree p and β is of degree q, it follows from Theorem 14.1 that the ring $H^*(X \times Y)$ is completely determined by $H^*(X)$ and $H^*(Y)$.

The *Poincaré polynomial* for X is defined as the polynomial

$$p(X, t) = \sum_0^{\dim X} t^j \dim H^j(X).$$

The degree is $\dim X$ if and only if X has an orientable component. The coefficients are called the *Betti numbers*. Theorem 14.1 implies:

Corollary 14.2. *If X and Y are compact C^∞ manifolds then*

$$p(X \times Y, t) = p(X, t) p(Y, t).$$

Example. For the sphere S^n we have $p(S^n, t) = 1 + t^n$. For a product

$$X = S^{n_1} \times \cdots \times S^{n_k}$$

it follows that $p(X, t) = (1 + t^{n_1}) \dots (1 + t^{n_k})$. We leave as a simple exercise for the reader to prove that two such products X are not homeomorphic unless the factors are the same apart from the order. — We have $p(P^n_{\mathbf{C}},t) = 1+t^2+\cdots+t^{2n} = (t^{2(n+1)}-1)/(t^2-1)$, hence

$$p(P^{n_1}_{\mathbf{C}}\times\cdots\times P^{n_k}_{\mathbf{C}},t) = \prod_1^k \big((t^{2(n_j+1)}-1)/(t^2-1)\big),$$

and it is easy to see that two such products are also different unless the factors are the same apart from the order.

By Poincaré's duality theorem the Poincaré polynomial $p(X,t)$ of any orientable compact manifold X is a reciprocal polynomial, that is,

$$t^{\dim X}p(X,1/t) = p(X,t),$$

as we saw in examples above.

Let a_j and b_k be closed differential forms in the classes α_j and β_k respectively. To prove Theorem 14.1 we must verify that every closed form on $X\times Y$ is cohomologous to a sum $\sum c_{ij}a_i \underset{\times}{\wedge} b_j$ with constant coefficients, and also prove that $c_{ij}=0$ for all i,j if the sum is cohomologous to 0. (If $a\in\lambda^*(X)$ and $b\in\lambda^*(Y)$ we denote by $a\underset{\times}{\wedge} Y$ the exterior product $\pi_X^*a\wedge\pi_Y^*b$. This is not a standard notation but it may clarify the formulas below.) The following lemma will allow us to prove both these facts at the same time. We use

Lemma 14.3. *Let $u\in\lambda^*(X\times Y)$ and assume that*

$$du = \sum_j A_j \underset{\times}{\wedge} b_j$$

for some $A_j\in\lambda^*(X)$. Then one can find $v\in\lambda^*(X\times Y)$ and $A'_j\in\lambda^*(X)$ such that

$$u = \sum A'_j \underset{\times}{\wedge} b_j + dv.$$

Proof. We may assume that $b_j\in\lambda^{p_j}(Y)$ for some p_j. Let $\{U_i\}$ be a finite covering of X with local coordinate patches. If x denotes the local coordinates in U_i, then we can write

$$u = \sum_I dx^I\wedge u_{I,x} \quad \text{in } U_i\times Y$$

where $I=(i_1,\dots,i_k)$ is increasing, $dx^I = dx_{i_1}\wedge\cdots\wedge dx_{i_k}$, and $u_{I,x}$ is a form on Y which is a C^∞ function of $x\in U_i$. We shall prove the lemma inductively under the hypothesis that the degree of $u_{I,x}$ is $\le\sigma$ for every U_i in the covering. This hypothesis is obviously independent of the choice of covering and local coordinates. The statement proves the lemma when $\sigma=\dim Y$, and it is vacuous when $\sigma=-1$. In the proof we may therefore assume that $\sigma\ge 0$ and that the statement has been proved for lower values of σ.

Let $u^\sigma_{I,x}$ be the part of $u_{I,x}$ of degree σ. If d_Y denotes the differential of $u_{I,x}$ when x is considered as a parameter then the only terms in du of degree $\sigma+1$ in the Y variables are

$$\sum_I(-1)^{|I|}dx^I\wedge d_Y u^\sigma_{I,x}.$$

Hence $d_Y u^\sigma_{I,x}$ must for fixed x be a linear combination of the forms b_j of degree $\sigma+1$. Since their cohomology classes are linearly independent and $d_Y u^\sigma_{I,x}$ is exact, the coefficients must be 0 so $d_Y u^\sigma_{I,x}=0$. If $\sigma=0$ it follows that $u^\sigma_{I,x}$ is a function of x, hence that u is a form in X lifted to $X\times Y$. The statement is true then, for 1 is cohomologous in Y to a linear combination of the forms b_j. If $\sigma>0$ it follows from Theorem 13.11 that

$$u^\sigma_{I,x}=\sum_j a_{I,j,x}b_j+dv_{I,x}$$

where b_j is of degree σ in the sum, and $a_{I,j,x}\in\mathbf{R}$ and $v_{I,x}\in\lambda^{\sigma-1}(Y)$ are C^∞ functions of $x\in U_i$. Set

$$v^i=\sum_I(-1)^{|I|}dx^I\wedge v_{I,x}$$

where the upper index i indicates that v^i is a form defined in $U_i\times Y$. It is a form of degree $\sigma-1$ with respect to the Y variables. Set

$$A^i_j=\sum_I a_{I,j,x}dx^I.$$

Then the difference

$$u-\sum_j A^i_j\underset{\times}{\wedge}b_j-dv^i\in\lambda^*(U^i\times Y)$$

has degree $\le\sigma-1$ with respect to the Y variables.

Let $\varphi_i\in C^\infty_0(U_i)$ be a partition of unity, thus $\sum_i\varphi_i=1$, and set

$$A'_j=\sum_i\varphi_iA^i_j\in\lambda^*(X),\quad v=\sum\varphi_iv^i\in\lambda^*(X\times Y),$$

where as usual a product by φ_i is defined as 0 outside $\operatorname{supp}\varphi_i$. Then

$$u_1=u-\sum_j A'_j\underset{\times}{\wedge}b_j-dv=\sum_i\varphi_i(u-\sum_j A^i_j\underset{\times}{\wedge}b_j-dv^i)-\sum_i(d\varphi_i)\wedge v^i$$

is of degree $\le\sigma-1$ with respect to the Y variables. We have

$$du_1=du-\sum_j(dA'_j)\underset{\times}{\wedge}b_j$$

so u_1 satisfies the hypotheses of the lemma. By the inductive hypothesis we conclude that

$$u_1=\sum_j A''_j\underset{\times}{\wedge}b_j+dv_1,$$

where $A''_j\in\lambda^*(X)$, and this gives

$$u=\sum_j(A'_j+A''_j)\underset{\times}{\wedge}b_j+d(v+v_1).$$

The proof is complete.

Proof of Theorem 14.1. If u is a closed form on $X \times Y$ it follows from Lemma 14.3 that u is cohomologous to a sum $\sum_j A_j \wedge b_j$ where $A_j \in \lambda^*(X)$. Since $\sum_j (dA_j) \underset{\times}{\wedge} b_j = 0$ we have $dA_j = 0$. But A_j is cohomologous to a sum $\sum_i c_{ij} a_i$ with constant coefficients, which proves that u is cohomologous to $\sum_{i,j} c_{ij} a_i \underset{\times}{\wedge} b_j$.

If $\sum_{i,j} c_{ij} a_i \underset{\times}{\wedge} b_j = dv$ it follows from Lemma 14.3 that

$$v = \sum_j A'_j \underset{\times}{\wedge} b_j + dw$$

where $A'_j \in \lambda^*(X)$ and $w \in \lambda^*(X \times Y)$. Hence

$$dv = \sum_j (dA'_j) \underset{\times}{\wedge} b_j = \sum_{i,j} c_{ij} a_i \underset{\times}{\wedge} b_j$$

which proves that $dA'_j = \sum_i c_{ij} a_i$. Since $\{a_i\}$ is a basis for the cohomology in X it follows that $dA'_j = 0$ and $c_{ij} = 0$, which completes the proof.

Lemma 14.3 was clearly the main point in the proof of Theorem 14.1, and in the proof of the lemma the spaces X and Y did not play symmetrical roles. We shall now see that the proof can actually be used to prove a generalisation of Theorem 14.1 where $X \times Y$ is replaced by a fiber space over X. This extension will be essential for the study of characteristic classes.

Definition 14.4. Let E, X be compact C^∞ manifolds, and let $\pi : E \to X$ be a C^∞ map. One calls E a *fiber space* with base X and projection π if the rank of π is everywhere equal to $\dim X$.

The definition implies that

$$E_x = \pi^{-1}(x) = \{e \in E; \pi e = x\}, \quad x \in X,$$

which is called the fiber of E over x, is a manifold for every $x \in X$. Every $x_0 \in X$ has an open neighborhood U such that $\pi^{-1}(U)$ is diffeomorphic to $U \times \pi^{-1}(x_0)$ under a fiber preserving diffeomorphism $\pi^{-1}(U) \ni e \mapsto (\pi e, \xi(e))$ where $\xi(e) \in \pi^{-1}(x_0)$. In fact, it follows from Theorem 10.6 that in a neighborhood of $\pi^{-1}(x_0)$ in E we have a projection p on $\pi^{-1}(x_0)$ with rank equal to the dimension of $\pi^{-1}(x_0)$. The map

$$e \mapsto (\pi e, p(e)) \in X \times \pi^{-1}(x_0)$$

is therefore locally bijective since $\dim E = \dim X + \dim \pi^{-1}(x_0)$. Since the restriction to $\pi^{-1}(x_0)$ is injective it follows that it is bijective in a neighborhood (see Exercise 10.7), so we have

Lemma 14.5. *If E is a compact C^∞ fiber space over the compact C^∞ manifold X, then there exists a finite covering $\{U_i\}$ of X such that $\pi^{-1}(U_i)$ for every i is diffeomorphic to $U_i \times \pi^{-1}(x_i)$ where $x_i \in U_i$, with a diffeomorphism preserving the fibers.*

If E is a fiber space over X with projection π, then we have a map $\pi^* : H^*(X) \to H^*(E)$, so we can view $H^*(E)$ as a $H^*(X)$ module.

Theorem 14.6 (Leray-Hirsch). *Let E be a C^∞ compact fiber space with the C^∞ compact manifold X as base and projection $\pi : E \to X$. Assume that there exist cohomology classes $\varepsilon_1, \dots, \varepsilon_N \in H^*(E)$ such that for every fiber $F = \pi^{-1}(x)$ the restrictions $i^*\varepsilon_1, \dots, i^*\varepsilon_N$, where i is the inclusion $F \to E$, form a basis for $H^*(F)$. Then it follows that $H^*(E)$ is the free $H^*(X)$ module generated by $\varepsilon_1, \dots, \varepsilon_N$. Thus every $\varepsilon \in H^*(E)$ has a unique representation of the form*

$$\varepsilon = \sum_1^N \pi^*\alpha_j \cup \varepsilon_j \quad \text{with } \alpha_j \in H^*(X),\ j = 1, \dots, N.$$

Corollary 14.7. *E and $X \times F$ have the same Poincaré polynomial if the hypotheses of the theorem are fulfilled.*

However, $H^*(E)$ and $H^*(X \times F)$ may have quite different structure as rings although the additive structures are the same. — Theorem 14.6 contains Theorem 14.1, for if $E = X \times Y$ we can take $\varepsilon_j = \pi_Y^*\beta_j$ where $\beta_1, \dots, \beta_N$ are a basis for $H^*(Y)$ and π_Y is the projection $E \to Y$.

It will clearly suffice to prove Theorem 14.6 when X is connected. Then it suffices to assume that there is *one* point $x_0 \in X$ such that $i^*\varepsilon_1, \dots, i^*\varepsilon_N$ is a basis for $H^*(F)$ if $F = \pi^{-1}(x_0)$ and i is the inclusion $F_0 \to X$. In fact, let $U \subset X$ be an open connected neighborhood of x_0 such that there is a fiber preserving diffeomorphism $\psi : U \times F_0 \to \pi^{-1}(U)$, and set $\tilde{\varepsilon}_j = \psi^*\varepsilon_j$. The maps $i_x : \xi \mapsto (x, \xi)$ from F_0 to $U \times F_0$ are homotopic when $x \in U$, so $i_x^*\tilde{\varepsilon}_j = (\psi \circ i_x)^*\varepsilon_j$ is independent of $x \in U$. Since $\psi \circ i_x$ is a diffeomorphism $F_0 \to \pi^{-1}(x)$ we conclude that $i_{x_0}^*\varepsilon_j$, $j = 1, \dots, N$, is a basis for $H^*(F_0)$, hence that $i_x^*\tilde{\varepsilon}_j$, $j = 1, \dots, N$, is a basis. This proves that the restrictions to $\pi^{-1}(x)$ of ε_j, $j = 1, \dots, N$, form a basis for $H^*(\pi^{-1}(x))$ for every $x \in U$. Since X is connected this is true for every $x \in X$.

Decomposing each ε_j in its homogeneous parts, that is, as a sum of elements in $H^n(E)$, $n = 0, 1, \dots, \dim E$, we obtain homogeneous cohomology classes $\hat{\varepsilon}_j$, $j = 1, \dots, \widehat{N}$, such that the restrictions to $\pi^{-1}(x)$ generate $H^*(\pi^{-1}(X))$ for every $x \in X$. For a fixed x_0 we can choose N of them which give a basis. From the preceding discussion it follows that they will then give a basis for every $x \in X$. Thus we can replace the basis $\varepsilon_1, \dots, \varepsilon_N$ in the hypothesis of Theorem 14.6 by another homogeneous basis, and this does not affect the hypothesis or the conclusion. We may therefore assume that $\varepsilon_j \in H^{n_j}(E)$ for some n_j when $j = 1, \dots, N$. Then the restrictions $i^*\varepsilon_j$ with $n(j) = \nu$ are for every ν a basis for $H^\nu(F)$ if i is the inclusion of a fiber F in E. For $j = 1, \dots, N$ we choose a form $e_j \in \lambda^{n_j}(E)$ in the cohomology class ε_j.

Lemma 14.3′. *Let $u \in \lambda^*(E)$ and assume that*

$$du = \sum_j \pi^*a_j \wedge e_j, \quad \text{where } a_j \in \lambda^*(X).$$

Then one can find $v \in \lambda^*(E)$ and $A_j \in \lambda^*(X)$ such that

$$u = \sum_j \pi^*A_j \wedge e_j + dv.$$

PROOF. Let $\{U_i\}$ be a finite covering of X with coordinate patches diffeomorphic to convex subsets of $\mathbf{R}^n$ which is so fine that with $F_i = \pi^{-1}(x_i)$, where $x_i \in U_i$, there is a fiber preserving diffeomorphism

$$\psi_i : U_i \times F_i \to \pi^{-1}(U_i), \quad \text{thus } \psi_i(x,\xi) = (x, \psi_i(x,\xi)) \quad \text{if } x \in U_i,\ \xi \in F_i.$$

If the local coordinates in U_i are denoted by x, then we can write

$$\psi_i^* u = \sum_I dx^I \wedge u_{I,x}$$

where $u_{I,x} \in \lambda^*(F_i)$ is a C^∞ function of x. We shall prove by induction over σ that the lemma is valid when the degree of $u_{I,x}$ is $\leq \sigma$ for every U_i in the covering. This hypothesis only depends on u and not on the choice of covering, local coordinates and trivialisations ψ_i. For let

$$\psi : U \times F_1 \ni (x,\xi) \mapsto (x, \psi(x,\xi)) \in U \times F_2$$

be a C^∞ map preserving the fibers. Then $\psi^* dx_i = dx_i$ whereas for a local coordinate η in F_2 the pullback $\psi^* d\eta = d(\eta(\varphi(x,\xi))$ is a sum of differentials along F_1 and along U. If V is a form on $U \times F_2$ of degree $\leq \sigma$ along F_2 it follows that $\psi^* V$ is of degree $\leq \sigma$ along F_1.

When $x \in U_i$ we denote the map $F_i \ni \xi \mapsto \psi_i(x,\xi)$ by $\psi_{i,x}$. By hypothesis $\psi_{i,x}^* e_1, \dots, \psi_{i,x}^* e_N$ is a basis for the cohomology in F_i, and by the homotopy invariance it is independent of x. Let u be of degree $\leq \sigma$ with respect to the fiber variables and write

$$\psi_i^* u = \sum_I dx^I \wedge u_{I,x},$$

as in the proof of Lemma 14.3, where $u_{I,x}$ is a form on F_i which is a C^∞ function of $x \in U_i$. Denote the part of $u_{I,x}$ of degree σ by $u_{I,x}^\sigma$. Using Theorem 13.11 it follows as in the proof of Lemma 14.3 that

$$u_{I,x}^\sigma = \sum_{1 \leq j \leq N, n_j = \sigma} a_{I,j,x} \psi_{I,x}^* e_j + d_{F_i} v_{I,x},$$

where $v_{I,x}$ is of form of degree $\leq \sigma - 1$ in Y which is a C^∞ function of $x \in U_i$, and $a_{I,j,x} \in C^\infty(U_i)$. Now we define v^i and A_j^i as in the proof of Lemma 14.3 and obtain that

$$\psi_i^* u - \sum_{j=1}^N \pi^* A_j^i \wedge \psi_i^* e_j - d_{F_i} v^i$$

is of degree $< \sigma$ with respect to the fiber variables. After pulling this form back to $\pi^{-1}(U_i)$ by the inverse of ψ_i we can use a partition of unity in X as at the end of Lemma 14.3 to complete the proof. The details are left for the reader.

PROOF OF THEOREM 14.6. The proof of Theorem 14.6 is essentially a repetition of that of Theorem 14.1, with Lemma 14.3 replaced by Lemma 14.3′, so it is left for the reader.

As an application of Theorem 14.6 we shall now define the Chern classes of a complex vector bundle V of fiber dimension n over X by determining the cohomology of the corresponding projective bundle. By the definition of a complex vector bundle, given in Chapter XIII, there is an open covering $X = \cup X_i$ such that $V|_{X_i}$ is diffeomorphic to $X_i \times \mathbf{C}^n$ with a diffeomorphism respecting the fibers and the vector operations in them. Now form the fiber space $P(V)$ with fiber over $x \in X$ consisting of the projective space defined by the fiber V_x. Thus $P(V)_x$ is $V_x \setminus \{0\}$ with elements differing by a complex factor identified. It is clear that $P(V)_{X_i} \cong X_i \times P^{n-1}_{\mathbf{C}}$, which proves at once that $P(V)$ is a fiber space over X. On $P(V)$ there is a natural line bundle H, for every element in $P(V)$ determines a complex line in the corresponding vector space. (In Chapter XIII we used the dual bundle but here we prefer not do so in order to get the desired sign.) The restriction of H to $P(V)_x$ is the natural line bundle on $P(V)_x$. If $c(H)$ is the Chern class of H, it follows from Theorems 12.7 and 13.13 that $c(H)_x{}^j$, $0 \le j < n$, is a basis for the cohomology in $P(V)_x$. Hence the theorem of Leray and Hirsch proves that every element in $H^*(P(V))$ can be written

$$\sum_0^{n-1} (p^* a_j) \cup c(H)^j, \quad a_j \in H^*(X),$$

where p is the projection $P(V) \to X$, and the representation is unique. In particular, $c(H)^n$ is of this form, so we have an equation

$$c(H)^n - c_1 c(H)^{n-1} + \cdots + (-1)^n c_n = 0, \quad c_j \in H^{2j}(X),$$

with uniquely determined coefficients $c_j(V)$. The ring structure of $H^*(P(V))$ can be calculated by means of them. It is clear that the coefficients c_j give important topological information on the vector bundle V. They are called the *Chern classes of the vector bundle*. For a line bundle V we have $P(V) = X$ and $H = V$, so $c_1(V)$ is then the Chern class defined in Chapter XIII.

Notes. For a thorough discussion of Chern classes we refer to [Hi]. For the definition of Chern classes by means of differential forms we refer to [BC] and to [Hö]. A systematic discussion of fiber bundles can be found in [Hu].

References

[BC] R. Bott and S. S. Chern, *Hermitian vector bundles and the equidistribution of the zeros of their holomorphic sections*, Acta Math. **114** (1965), 71–112.

[Hi] F. Hirzebruch, *Topological methods in algebraic geometry*, Springer Verlag, 1966.

[Hö] L. Hörmander, *Riemannian geometry*, Lecture Notes, Department of Mathematics, Lund 1990.

[Hu] D. Husemoller, *Fibre bundles*, Mc Graw Hill, 1966.

CHAPTER XV

DIRECT IMAGES OF COHOMOLOGY CLASSES

Let X and Y be compact C^∞ oriented manifolds, and let f be a continuous map $X \to Y$. Then we have a map$f^* : H^*(Y) \to H^*(X)$, which only depends on the homotopy class of f. By Theorem 13.5 (the Poincaré duality theorem) a non-singular bilinear form $(\alpha, \beta) \mapsto \langle \alpha, \beta \rangle$ on $H^*(X)$ is induced by the bilinear form $(u, v) \mapsto \int_X u \wedge v$ when $u, v \in \lambda^*(X)$ are in the cohomology classes α and β. The bilinear form $\langle \alpha, \beta \rangle$ identifies $H^*(X)$ with its dual space. Similarly we have a non-singular bilinear form on $H^*(Y)$. The adjoint of the map f^* can therefore be viewed as a map f_* from $H^*(X)$ to $H^*(Y)$. It is defined by

$$\langle f^*\alpha, \beta \rangle_Y = \langle \alpha, f_*\beta \rangle_X, \quad \alpha \in H^*(Y),\ \beta \in H^*(X). \tag{15.1}$$

On the left we have the bilinear form in $H^*(Y)$ and on the right the bilinear form in $H^*(X)$. The subscripts will often be omitted. (The notation f_* is meant to suggest that f_* goes in the same direction as f while the notation f^* is meant to indicate the opposite direction.) If $\beta \in H^j(X)$ it suffices to take $\alpha \in H^{\dim X - j}(Y)$ in (15.1), which implies that $f_*\beta \in H^{\dim Y - \dim X + j}(Y)$. (We assume tacitly that X and Y are connected or at least that every component has the same dimension.)

Example 1. Let us assume that X and Y are connected of the same dimension n. Then we have defined the degree m of f by

$$\langle f^*\alpha, 1_X \rangle = m\langle \alpha, 1_Y \rangle, \quad \alpha \in H^n(Y),$$

where 1_X (resp. 1_Y) is the 0 form which is identically 1 in X (resp. Y). By (15.1) this means that $f_*1_X = m \cdot 1_Y$.

Example 2. For the embedding $i : P_{\mathbf{C}}^{n-k} \to P_{\mathbf{C}}^n$ we have $i_*1 = \alpha^k$ where α is the cohomology class of the form U in Theorem 12.8. (Here 1 is the 0 form in $P_{\mathbf{C}}^{n-k}$ which is identically 1.) In fact, every cohomology class γ on $P_{\mathbf{C}}^n$ contains a polynomial $\sum_0^n a_j U^j$ with constant real coefficients, and we have

$$\langle i^*\gamma, 1 \rangle_{P_{\mathbf{C}}^{n-k}} = \int_{P_{\mathbf{C}}^{n-k}} \sum_0^n a_j U^j = a_{n-k} = \int_{P_{\mathbf{C}}^n} \Big(\sum_0^n a_j U^j\Big) \wedge U^k = \langle \gamma, \alpha^k \rangle_{P_{\mathbf{C}}^n}$$

by Theorem 12.8.

The map f_* is not multiplicative, but we can translate the multiplicative properties of f^* to a property of f_*. To do so let $\alpha, \beta \in H^*(Y)$ and $\gamma \in H^*(X)$. Then $f^*(\alpha \cup \beta) = (f^*\alpha) \cup (f^*\beta)$, hence

$$\langle f^*(\alpha \cup \beta), \gamma \rangle_X = \langle (f^*\alpha) \cup (f^*\beta), \gamma \rangle_X = \langle f^*\alpha, (f^*\beta) \cup \gamma \rangle_X = \langle \alpha, f_*((f^*\beta) \cup \gamma) \rangle_Y.$$

This implies that

$$\beta \cup f_*\gamma = f_*((f^*\beta) \cup \gamma), \quad \beta \in H^*(Y),\ \gamma \in H^*(X). \tag{15.2}$$

For $\gamma = 1_X$ we obtain in particular

$$f_*f^*\beta = \beta \cup f_*1_X, \quad \beta \in H^*(Y). \tag{15.3}$$

If X and Y are connected manifolds of the same dimension we saw in Example 1 that $f_*1_X = m \cdot 1_Y$ where m is the degree. Hence it follows then that

$$f_*f^*\beta = m\beta, \quad \beta \in H^*(Y),$$

which proves that f^* is injective and that f_* is surjective if $m \neq 0$. Thus we have proved:

THEOREM 15.1. *If X and Y are compact connected oriented manifolds of the same dimension and $f : X \to Y$ is a continuous map with degree $\neq 0$, then $f^* : H^*(Y) \to H^*(X)$ is injective, which implies that $H^*(Y)$ is a subring of $H^*(X)$. In particular, the Betti numbers of Y are at most equal to those of X. If they are equal then f^* is an isomorphism. This is true in particular if there also exists a map $g : Y \to X$ with degree $\neq 0$.*

EXAMPLE 3. If Y is the sphere S^n then $H^*(Y)$ consists of all pairs $(x_0, x_n) \in \mathbf{R}^2$ with componentwise addition and the multiplication

$$(x_0, x_n) \cup (y_0, y_n) = (x_0y_0, x_0y_n + x_ny_0).$$

This ring is a subring of $H^*(X)$ for every connected orientable X of dimension n, so the theorem gives no restrictions on manifolds which can be mapped with degree $\neq 0$ into the sphere of the same dimension. In fact, there is always such a map, for we can choose an open subset $X_1 \subset X$ which is diffeomorphic to $\mathbf{R}^n$, hence with $S^n \setminus P$ where P is a point in S^n. The diffeomorphism $X_1 \to S^n \setminus P$ extended to X by mapping $X \setminus X_1$ to P is a map $X \to S^n$ with degree 1.

However. if we take $X = S^n$ in Theorem 15.1 we find that $H^k(Y)$ must be 0 for $0 < k < n$, so $H^*(Y)$ must be isomorphic to $H^*(S^n)$ if S^n can be mapped into Y with degree $\neq 0$. For example, S^n cannot be mapped into the torus T^n with degree $\neq 0$ although the torus can be mapped into S^n with degree 1.

We can allow X and Y to be oriented but not necessarily compact manifolds provided that the duality theorem 13.12 is valid for them. For every continuous map $f : X \to Y$ we still have a linear map $f^* : H^*(Y) \to H^*(X)$, and we can define the adjoint map $f_* : H_c^*(X) \to H_c^*(Y)$. If f is *proper* we also have a map $f^* : H_c^*(Y) \to H_c^*(X)$ which has an adjoint map $f_* : H^*(X) \to H^*(Y)$. The map $f_* : H_c^*(X) \to H_c^*(Y)$ only depends on the homotopy class of f, and the map $f_* : H^*(X) \to H^*(Y)$ only depends on the proper homotopy class of f.

THEOREM 15.2. *Let X and Y be oriented C^∞ manifolds such that X is compact and Poincaré duality is valid for Y. If $f : X \to Y$ and $g : Y \to X$ define a homotopy equivalence, then $f_* : H^*(X) \to H_c^*(Y)$ is an isomorphism, and we have*

$$f_*\alpha = (g^*\alpha) \cup (f_*1_X), \quad \alpha \in H^*(X). \tag{15.4}$$

Proof. Since $f \circ g$ and $g \circ f$ are homotopic to the identity in Y and X, respectively, it follows that $f^* g^*$ and $g^* f^*$ are the identity in $H^*(X)$ and $H^*(Y)$ respectively. Hence f_* is an isomorphism $H^*(X) = H^*_c(X) \to H^*_c(Y)$. We obtain (15.4) by taking $\beta = g^*\alpha$ in (15.3) which remains valid under our present hypotheses.

We shall now discuss the special case where $f : X \to Y$ is an embedding. For every sufficiently small tubular neighborhood of X in Y the hypotheses of Theorem 15.2 are fulfilled with Y replaced by T and g equal to the projection $\pi : T \to X$. Hence we see that $H^*(X)$ is isomorphic to $H^*_c(T)$; the isomorphism is given by a map

$$H^*(X) \ni \alpha \mapsto (\pi^*\alpha) \cup \varepsilon \in H^*_c(T), \tag{15.5}$$

where $\varepsilon = f_* 1_X \in H^*_c(T)$. Note that if $e \in \lambda^*_c(T)$ is a form representing ε, then the definition of ε means that

$$\int_T u \wedge e = \int_X u,$$

when u is a closed form in T. In the particular case where Y is a real vector bundle — which is in fact the general case modulo diffeomorphisms — the class ε is called the *Euler class* of the bundle, and (15.5) is called the *Thom-Gysin isomorphism.*

For an embedding $f : X \to Y$ we shall now study the integral of $f_* 1_X$ over submanifolds of Y. Since $f_* 1_X$ can be represented by forms with support arbitrarily close to X, the integral is 0 over every manifold which does not intersect X; the integral can only depend on the intersections. (We identify X with $f(X)$.) When studying the intersections we shall also examine the more general case where f is not necessarily an embedding.

Definition 15.3. Let X_1, X_2, Y be oriented C^∞ manifolds such that X_1 and X_2 are compact, $\dim X_1 + \dim X_2 = \dim Y$, and let $f_j : X_j \to Y$, $j = 1, 2$, be C^∞ maps. One calls the maps *transversal* if for all $(x_1, x_2) \in X_1 \times X_2$ with $f_1(x_1) = f_2(x_2) = y$ the map

$$f_1'(x_1) \oplus f_2'(x_2) : T_{x_1}(X_1) \oplus T_{x_2}(X_2) \ni (t_1, t_2) \mapsto f_1'(x_1)t_1 + f_2'(x_2)t_2 \in T_y(Y)$$

is bijective. One calls (x_1, x_2) a positive (negative) intersection if this bijection preserves (reverses) the orientations. (A direct sum $A \oplus B$ of two finite dimensional oriented vector spaces is oriented by letting a positive basis for A followed by a positive basis in B be a positive basis in $A \oplus B$.)

Transversality is a very strong hypothesis:

Lemma 15.4. *Let f_1 and f_2 be transversal as in Definition 15.3. If $f_1(x_1) = f_2(x_2) = y$ there exists an open neighborhood U_j of x_j, $j = 1, 2$, and a diffeomorphism f of $U_1 \times U_2$ on an open neighborhood of y such that*

$$f(x_1, \xi_2) = f_2(\xi_2), \; f(\xi_1, x_2) = f_1(\xi_1), \quad \text{if } \xi_j \in U_j. \tag{15.6}$$

In particular, if $(\xi_1, \xi_2) \in U_1 \times U_2$, then $f_1(\xi_1) = f_2(\xi_2)$ implies that $(\xi_1, x_2) = (x_1, \xi_2)$, that is, $\xi_1 = x_1$ and $\xi_2 = x_2$, so (x_1, x_2) is the only intersection in $U_1 \times U_2$.

PROOF. Let $\psi : \mathbf{R}^n \to Y$ be a coordinate system in a neighborhood of y with $\psi(0) = y$. For (ξ_1, ξ_2) in a neighborhood of (x_1, x_2) we can define f by

$$f(\xi_1, \xi_2) = \psi\big(\psi^{-1}(f_1(\xi_1)) + \psi^{-1}(f_2(\xi_2))\big).$$

Then (15.5) is valid, and for $(\xi_1, \xi_2) = (x_1, x_2)$ we have

$$f' = \psi'(0)(\psi^{-1})'(y)(f_1'(x_1) \oplus f_2'(x_2)) = f_1'(x_1) \oplus f_2'(x_2).$$

Since this is a bijection it follows from the inverse function theorem that f is a diffeomorphism of some neighborhood $U_1 \times U_2$ of (x_1, x_2) on a neighborhood of y, which proves the lemma.

THEOREM 15.5. *Let X_1, X_2, Y be compact C^∞ oriented manifolds with $\dim X_1 + \dim X_2 = \dim Y$, and let $f_j : X_j \to Y$ be transversal C^∞ maps. Then $\langle f_{1*}1_{X_1}, f_{2*}1_{X_2}\rangle_Y$ is equal to the number of $(x_1, x_2) \in X_1 \times X_2$ with $f_1(x_1) = f_2(x_2)$ counted with the sign in Definition 15.3, that is, the number of signed intersections of $f_1(X_1)$ and $f_2(X_2)$.*

PROOF. We shall first assume that f_1 is an embedding. Then we know that the class $f_{1*}1_{X_1}$ contains forms e with support in an arbitrarily small tubular neighborhood T in Y of the manifold $f_1(X_1)$. For every $x_2 \in X_2$ we can find at most one $x_1 \in X_1$ with $f_1(x_1) = f_2(x_2)$. If we label the intersections (x_1^j, x_2^j), $j = 1, \dots, \nu$ the components x_2^j are therefore different. Choose for $j = 1, \dots, \nu$ open neighborhoods U_1^j and U_2^j of x_1^j and x_2^j according to Lemma 15.4 so that U_2^j are disjoint, and choose corresponding diffeomorphisms f^j from $U_1^j \times U_2^j$ to a neighborhood of $y^j = f_1(x_1^j) = f_2(x_2^j)$. We have

$$\langle f_{1*}1_{X_1}, f_{2*}1_{X_2}\rangle_Y = \langle f_2^* f_{1*}1_{X_1}, 1_{X_2}\rangle_{X_2} = \int_{X_2} f_2^* e.$$

If we choose the support of e sufficiently close to $f(X_1)$ then supp $f_2^* e \subset \cup_1^\nu U_2^j$, and we obtain

$$\int_{X_2} f_2^* e = \sum_{j=1}^{\nu} \int_{U_2^j} f_2^* e.$$

The definition of e means that

$$\int_{f(X_1)} u = \int_T u \wedge e \tag{15.7}$$

for arbitrary closed forms u in T. Choose U_1^j so small that $f^j(U_1^j \times \partial U_2^j) \cap Y = \emptyset$. If $v \in \lambda_c^{d_1}(U_1^j)$ where $d_1 = \dim X_1$, then $dv = 0$. Let u be the pullback of $v \underset{\times}{\wedge} 1 \in \lambda^{d_1}(U_1^j \times U_2^j)$ to $f^j(U_1^j \times U_2^j)$ by the inverse of f^j. Since u vanishes near the intersection of the boundary with T, we can set $u = 0$ in the the rest of T and obtain a form u in T with $du = 0$. From (15.7) it follows now that

$$\int_{U_1^j} v = \int_{f(X_1)} u = \int_T u \wedge e = \pm \int f^{j*}(u \wedge e) = \pm \int_{U_1^j \times U_2^j} (v \times 1) \wedge e'$$

where $e' = f^{j*}e$ and the sign is that in Definition 15.3. Since e' has degree $d_2 = \dim X_2$ the integral I of e' over U_2^j for fixed $\xi_1 \in U_1^j$ is a function $I(\xi_1)$ there, and we obtain

$$\int_{U_1^j} v = \pm \int_{U_1^j} Iv, \quad v \in \lambda_c^{d_1}(U_1^j),$$

which proves that $I(\xi_1) = \pm 1$. In particular we obtain when $\xi_1 = x_1^j$ that

$$\int_{U_2^j} f_2^{j*}e = \pm 1.$$

This completes the proof of the theorem when f_1 is an embedding. It is therefore also valid when f_2 is an embedding, for if we let f_1 and f_2 change places then both sides of the asserted inequality are multiplied by $(-1)^{\dim X_1 \dim X_2}$.

To prove the theorem without such restrictions on f_1 or f_2 we introduce the map

$$f : X = X_1 \times X_2 \ni (x_1, x_2) \mapsto (f_1(x_1), f_2(x_2)) \in Y \times Y,$$

and the diagonal map

$$\delta : Y \ni y \mapsto (y, y) \in Y \times Y.$$

The intersections of f and δ consist of all $(x_1, x_2, y) \in X \times Y$ with $f_1(x_1) = f_2(x_2) = y$, so they are in $1-1$ correspondence with the intersections of f_1 and f_2. The sign of an intersection of f and δ is the sign of the linear map

$$\begin{aligned}\varphi_1 : T_{x_1}(X_1) \otimes T_{x_2}(X_2) \otimes T_y(Y) &\ni (\xi_1, \xi_2, \eta) \\ &\mapsto (f_1'(x_1)\xi_1 + \eta, f_2'(x_2)\xi_2 + \eta) \in T_y(Y) \oplus T_y(Y),\end{aligned}$$

and the sign of the corresponding intersection of f_1 and f_2 is the sign of the map

$$\varphi_2 : T_{x_1}(X_1) \oplus T_{x_2}(X_2) \ni (\xi_1, \xi_2) \mapsto (f_1'(x_1)\xi_1 + f_2'(x_2)\xi_2) \in T_y(Y).$$

For the linar maps

$$\begin{aligned}\psi_1 : T_y(Y) \oplus T_y(Y) \ni (\eta_1, \eta_2) &\mapsto (\eta_1 - \eta_2, \eta_2) \in T_y(Y) \oplus T_y(Y), \\ \psi_2 : T_{x_1}(X_1) \oplus T_{x_2}(X_2) \ni (\xi_1, \xi_2) &\mapsto (\xi_1, -\xi_2) \in T_{x_1}(X_1) \oplus T_{x_2}(X_2)\end{aligned}$$

we have $\det \psi_1 = 1$ and $\det \psi_2 = (-1)^{\dim X_2}$, and

$$(\psi_1 \circ \varphi_1)(\xi_1, \xi_2, \eta) = (f_1'(x_1)\xi_1 - f_2'(x_2)\xi_2, f_2'(x_2)\xi_2 + \eta)$$

is bijective, so f and δ are transversal. The linear map

$$\begin{aligned}T_{x_1}(X_1) \oplus T_{x_2}(X_2) \oplus T_y(Y) &\ni (\xi_1, \xi_2, \eta) \\ &\mapsto (f_1'(x_1)\xi_1 - f_2'(x_2)\xi_2, t f_2'(x_2)\xi_2 + \eta) \in T_y(Y) \oplus T_y(Y)\end{aligned}$$

is bijective for every t. The sign is therefore independent of t, and when $t = 0$ it is equal to the sign of $\varphi_2 \circ \psi_2$. Hence the sign of φ_1 is $(-1)^{\dim X_2}$ times the sign of φ_2. Now we have

$$\begin{aligned}\langle f_{1*}1_{X_1}, f_{2*}1_{X_2}\rangle_Y &= \langle (f_{1*}1_{X_1}) \cup (f_{2*}1_{X_2}), 1_Y\rangle_Y \\ &= \langle \delta^*(f_{1*}1_{X_1} \times f_{2*}1_{X_2}), 1_Y\rangle_Y = \langle (f_{1*1_{X_1}}) \times (f_{2*}1_{X_2}), \delta_* 1_Y\rangle_{Y\times Y}.\end{aligned}$$

To calculate $(f_{1*}1_{X_1}) \times (f_{2*}1_{X_2})$ we take $\alpha_j \in H^{\dim X_j}(Y)$, $j = 1, 2$, and obtain

$$\begin{aligned}\langle \alpha_1 \times \alpha_2, (f_{1*}1_{X_1}) \times (f_{2*}1_{X_2})\rangle_{Y\times Y} &= (-1)^{(\dim X_2)^2}\langle \alpha_1, f_{1*}1_{X_1}\rangle_Y \langle \alpha_2, f_{2*}1_{X_2}\rangle_Y \\ = (-1)^{\dim X_2}\langle f_1^*\alpha_1, 1_{X_1}\rangle_{X_1}\langle f_2^*\alpha_2, 1_{X_2}\rangle_{X_2} &= (-1)^{\dim X_2}\langle (f_1^*\alpha_1)\times(f_2^*\alpha_2), 1_X\rangle_X \\ = (-1)^{\dim X_2}\langle f^*(\alpha_1\times\alpha_2), 1_X\rangle_X &= (-1)^{\dim X_2}\langle \alpha_1\times\alpha_2, f_*1_X\rangle_{Y\times Y}.\end{aligned}$$

(The first equality is clear if one thinks of the corresponding differential forms.) Hence

$$\langle f_{1*}1_{X_1}, f_{2*}1_{X_2}\rangle_Y = (-1)^{\dim X_2}\langle f_*1_X, \delta_*1_Y\rangle_{Y\times Y},$$

and since δ is an embedding this is $(-1)^{\dim X_2}$ times the number of intersections between f and δ which is precisely the number of intersections of f_1 and f_2. The proof is complete.

An important part of Theorem 15.5 remains valid for arbitrary continuous maps:

THEOREM 15.6. *Let X_1, X_2, Y be compact oriented C^∞ manifolds with*

$$\dim X_1 + \dim X_2 = \dim Y.$$

If $f_j : X_j \to Y$ are continuous maps, then $\langle f_{1*}1_{X_1}, f_{2*}1_{X_2}\rangle_Y$ is always an integer.

In view of Theorem 15.5 it is natural to call this number the *intersection number* between f_1 and f_2 or, with some suggestive abuse of language, between $f_1(X_1)$ and $f_2(X_2)$.

For the proof of Theorem 15.6 we first note that $\langle f_{1*}1_{X_1}, f_{2*}1_{X_2}\rangle_Y$ only depends on the homotopy classes of f_1 and of f_2, and by Corollary 10.9 every homotopy class contains C^∞ maps. Theorem 15.6 will follow from Theorem 15.5 and the following more precise fact:

THEOREM 15.7. *Let X_1, X_2, Y be compact C^∞ manifolds with*

$$\dim X_1 + \dim X_2 = \dim Y.$$

If $f_j : X_j \to Y$ are continuous maps, $j = 1, 2$, then one can find C^∞ transversal maps $g_j : X_j \to Y$, $j = 1, 2$, which are arbitrarily close to f_j, hence homotopic to f_j.

PROOF. We have just seen that it is no restriction to assume that $f_j \in C^\infty$. If Y is embedded in $\mathbf{R}^N$ we can for a sufficiently small open neighborhood U of $0 \in \mathbf{R}^N$ define

$$\varphi(y, t) = \pi(y + t), \quad y \in Y,\ t \in U,$$

where π is the projection on Y in a tubular neighborhood of Y. We have $\varphi \in C^\infty(Y \times U, Y)$, $\varphi(y, 0) = y$, $y \in Y$, and the differential of φ is surjective even for a fixed y. Now we claim that the C^∞ maps

$$g_1(x_1) = f_1(x_1), \quad g_2^t(x_2) = \varphi(f_2(x_2), t)$$

are transversal for all $t \in U \setminus C$ where C is a set of measure 0. To prove this we shall study the intersections

$$\Sigma = \{(x_1, x_2, t) \in X_1 \times X_2 \times U; f_1(x_1) = \varphi(f_2(x_2), t)\}.$$

If we prove that Σ is a manifold of dimension N and that g_1 and g_2^t are transversal if and only if t is not a critical value of the projection $\Sigma \to U$, then the theorem will follow from the Morse-Sard theorem (Theorem 3.1). Since these statements are local it suffices in the proof to consider points $(x_1^0, x_2^0, t^0) \in \Sigma$ such that $f_1(x_1^0) \in V_1$ where V_1 is an open subset of Y with a diffeomorphism $h : V_1 \to \mathbf{R}^{\dim Y}$. In a neighborhood of this point Σ is defined by the equation

$$F(x_1, x_2, t) = h(f_1(x_1)) - h(\varphi(f_2(x_2), t)) = 0.$$

Since F_t' is surjective this equation defines a manifold of dimension $\dim X_1 + \dim X_2 + N - \dim Y = N$. At a point in Σ with $f_2(x_2) = y$ the tangent space is given by

$$\{(\xi_1, \xi_2, \tau) \in T_{x_1}(X_1) \oplus T_{x_2}(X_2) \oplus \mathbf{R}^N; f_1'(x_1)\xi_1 - \varphi_y'(y,t) f_2'(x_2)\xi_2 - \varphi_t'(y,t)\tau = 0\},$$

and its projection on the component $\tau \in \mathbf{R}^N$ is therefore bijective if and only if $f_1'(x_1)\xi_1 - \varphi_y'(y,t)f_2'(x_2)\xi_2 = 0$ implies $(\xi_1, \xi_2) = 0$, which means precisely that g_1 and g_2^t are transversal. The proof is complete.

To explain the significance of Theorem 15.6 we shall give two examples.

Example 4. Let $f : X \to Y$ be a continuous map where X and Y are compact oriented connected C^∞ manifolds of the same dimension. Then $f_* 1_X = m \cdot 1_Y$ where m is the degree of f (see Example 1). If P is a point and $i : P \to Y$ is an embedding of P in Y, then

$$\langle f_* 1_X, i_* 1_P \rangle = m \langle 1_Y, i_* 1_P \rangle = m \langle i^* 1_Y, 1_P \rangle = m$$

so Theorem 15.6 proves again that m is an integer.

Example 5. Let $f : P_{\mathbf{C}}^n \to P_{\mathbf{C}}^n$ be continuous. For $0 \le k \le n$ we let $i_k : P_{\mathbf{C}}^k \to P_{\mathbf{C}}^n$ be an embedding of the k dimensional complex projective space as a subspace. Then $i_{k*} 1_{P_{\mathbf{C}}^k} = \alpha^{n-k}$ where α is the generator for $H^2(P_{\mathbf{C}}^n)$ constructed in Theorem 12.7 (see Example 2 above). Now Theorem 15.6 gives that

(15.8)
$$\langle f^* \alpha^k, \alpha^{n-k} \rangle_{P_{\mathbf{C}}^n} = \langle f^* i_{(n-k)*} 1_{P_{\mathbf{C}}^{n-k}}, i_{k*} 1_{P_{\mathbf{C}}^k} \rangle_{P_{\mathbf{C}}^n} = \langle i_{(n-k)*} 1_{P_{\mathbf{C}}^{n-k}}, (f \circ i_k)_* 1_{P_{\mathbf{C}}^k} \rangle_{P_{\mathbf{C}}^n}$$

is an integer. Since $H^k(P_{\mathbf{C}}^n) = \mathbf{R}\alpha^k$ we have $f^*\alpha^k = c_k \alpha^k$ where $c_k \in \mathbf{R}$. However, c_k is equal to the left-hand side of (15.8) so c_k is an integer. This is nothing new when $k = n$, for c_n is the degree of f. However, since $c_k = c_1^k$ we may now conclude that the degree is of the form m^n where m is an integer. This result is optimal. In fact, for every positive integer m we can choose f with the degree $(\pm m)^n$, for in terms of the homogeneous coordinates such a map is given by

$$(z_0, \dots, z_n) \mapsto (z_0^m, \dots, z_n^m) \quad \text{resp. } (z_0, \dots, z_n) \mapsto (\bar{z}_0^m., \dots, \bar{z}_n^m).$$

With X_1, X_2, Y as in Theorem 15.5 we shall now derive a method for calculating the intersection number between f_1 and f_2 when the intersections

$$S = \{(x_1, x_2) \in (X_1 \times X_2); f_1(x_1) = f_2(x_2)\}$$

form a finite set. This will be proved by rewriting the result in Theorem 15.5 in the transversal case so that it is still meaningful under this weaker assumption.

Let V be an open set in Y such that there exists a diffeomorphism $\psi : V \to \mathbf{R}^{\dim Y}$ preserving the orientation. Let U be an open subset of $X_1 \times X_2$ such that a neighborhood of $\overline{U}$ is diffeomorphic to an open subset of $\mathbf{R}^{\dim Y}$.

LEMMA 15.8. *Let $f_j \in C^\infty(X_j, Y)$ be transversal and assume that $f_j(x_j) \in V$ when $(x_1, x_2) \in \overline{U}$, and that $f_1(x_1) \neq f_2(x_2)$ when $(x_1, x_2) \in \partial U$. Then the number of intersections $(x_1, x_2) \in U$ of f_1 and f_2, counted with sign as in Theorem 15.5, is equal to $(-1)^{\dim X_2} D(\varphi, U, 0)$ where $\varphi : U \to \mathbf{R}^{\dim Y}$ is defined by*

$$\varphi(x_1, x_2) = \psi(f(x_1)) - \psi(f(x_2)). \tag{15.9}$$

Here the degree $D(\varphi, U, 0)$ was defined in Theorem 4.2.

PROOF. Since $\varphi(x_1, x_2) \neq 0$ when $(x_1, x_2) \in \partial U$, the degree $D(\varphi, U, 0)$ is defined. When $(x_1, x_2) \in U$ and $\varphi(x_1, x_2) = 0$, then $f(x_1) = f(x_2) = y$, and

$$\varphi'(x_1, x_2) = \varphi'(y)(f'(x_1) \oplus (-f'(x_2)),$$

so 0 is not a critical value of φ and the sign of $\det \varphi'(x_1, x_2)$ is equal to $(-1)^{\dim X_2}$ times the sign of the intersection of f_1 and f_2 at (x_1, x_2) according to Theorem 15.5. This proves the lemma.

We have given other expressions for $D(\varphi, U, 0)$ in Chapters VIII and IX and will not repeat them here.

THEOREM 15.9. *Let X_1, X_2, Y be compact C^∞ oriented manifolds with*

$$\dim X_1 + \dim X_2 = \dim Y.$$

If $f_j : X_j \to Y$ are continuous maps and the intersections

$$S = \{(x_1, x_2) \in X_1 \times X_2; f_1(x_1) = f_2(x_2)\}$$

form a finite set, then the intersection number between f_1 and f_2 is the sum over S of the indices of the intersections, defined as follows. For every $(x_1, x_2) \in S$ we take an orientation preserving diffeomorphism ψ of a neighborhood V of $y = f_1(x_1) = f_2(x_2)$ on $\mathbf{R}^{\dim Y}$ and a neighborhood U of $(x_1, x_2) \in X_1 \times X_2$ such that a neighborhood of $\overline{U}$ is diffeomorphic to an open bounded subset of $\mathbf{R}^{\dim Y}$ and $\overline{U} \cap S = \{(x_1, x_2)\}$. Then the index of (x_1, x_2) is $(-1)^{\dim X_2} D(\varphi, U, 0)$ where φ is defined by (15.9)*.*

PROOF. Choose neighborhoods $U_1, \dots, U_\nu$ of the intersections such that the closures are disjoint. If $g_j \in C^\infty(X_j, Y)$ are transversal and sufficiently close to f_j, then there are no intersections between g_1 and g_2 outside $\cup_1^\nu U_j$. By Lemma 15.8 the theorem is true for the intersection number between g_1 and g_2, which is equal to the intersection number between f_1 and f_2 when g_j is sufficiently close to f_j, $j = 1, 2$. The degrees at 0 of g_j and f_j are then also equal in $U_1, \dots, U_\nu$, which proves the theorem.

In Chapter XVI we shall give an important application of this theorem. We end this chapter with some comments on the intersections between $f_1 \in C^\infty(X_1, Y)$ and $f_2 \in C^\infty(X_2, Y)$ when $\dim X_1 + \dim X_2$ is not necessarily equal to $\dim Y$. Then f_1 and f_2 are called transversal if for arbitrary (x_1, x_2) in

$$X_3 = \{(x_1, x_2) \in X_1 \times X_2; f_1(x_1) = f_2(x_2)\}$$

the map

$$f_1'(x_1) \oplus f_2'(x_2) : T_{x_1}(X_1) \oplus T_{x_2}(X_2) \to T_y(Y)$$

is surjective. Then it follows from the implicit function theorem that X_3 is a C^∞ manifold of dimension $\dim X_1 + \dim X_2 - \dim Y$ (hence empty if this is a negative number). We have also a C^∞ map $f_3 : X_3 \to Y$ defined by $f_3(x_1, x_2) = f_1(x_1) = f_2(x_2)$ when $(x_1, x_2) \in X_3$. If X_1, X_2 and Y are oriented, then X_3 has a unique orientation such that

$$(f_{1*}1_{X_1}) \cup (f_{2*}1_{X_2}) = f_{3*}1_{X_3}.$$

One can extend Theorem 15.7 so that the restrictions on the dimensions are eliminated. The cohomology class $(f_{1*}1_{X_1}) \cup (f_{2*}1_{X_2})$ can therefore always be written in the form $f_{3*}1_{X_3}$ where X_3 is a compact manifold of dimension $\dim X_1 + \dim X_2 - \dim Y$ and f_3 is a C^∞ map $X_3 \to Y$, and one can regard $f_3(X_3)$ as a representative of the intersection between $f_1(X_1)$ and $f_2(X_2)$.

Notes. The intersection numbers studied here by means of de Rham cohomology are historically much older. They were the starting point of the duality in the classical homology theory of Poincaré. The theorems in this chapter can essentially be found in [H] although the formulations there are somewhat different.

References

[H] H. Hopf, *Zur algebra der Abbildungen von Mannigfaltigkeiten*, J. Reine Angew. Math. **163** (1930), 71–88.

CHAPTER XVI

THE FIXED POINT THEOREMS OF LEFSCHETZ AND HOPF

Let X be a compact oriented C^∞ manifold of dimension n, and let $f : X \to X$ be a continuous map. We want to study the fixed points of f, that is, the points $x \in X$ with $f(x) = x$. If G_f is the graph of f,

$$G_f = \{(x, f(x)); x \in X\} \subset X \times X,$$

then the fixed points are the intersections between G_f and the diagonal $\Delta \subset X \times X$, and we can calculate the intersection number using the methods in Chapter XV.

Let δ be the diagonal map $X \ni x \mapsto (x, x) \in X \times X$, and let F be the map $X \ni x \mapsto (x, f(x)) \in X \times X$. We choose a basis $\{\alpha_i\}$ for $H^*(X)$ consisting of homogeneous elements and denote by $\{\beta_j\}$ the biorthogonal basis, thus

$$\langle \alpha_i, \beta_j \rangle = \delta_{ij}.$$

The existence of the dual basis follows from the Poincaré duality theorem, and the cohomology classes β_j are also homogeneous. By Künneth's formula (Theorem 14.1) we have

$$F_* 1_X = \sum_{i,j} c_{ij} \alpha_i \times \beta_j, \quad c_{ij} = 0 \ \text{ if } \deg \alpha_i + \deg \beta_j \neq n,$$

so the intersection number between G_f and Δ is

$$\langle F_* 1_X, \delta_* 1_X \rangle_{X \times X} = \langle \delta^* F_* 1_X, 1_X \rangle_X = \sum_{i,j} c_{ij} \langle \alpha_i, \beta_j \rangle = \sum_i c_{ii}.$$

By the definition of F_* we have

$$\langle \beta_i \times \alpha_j, F_* 1_X \rangle_{X \times X} = \langle F^*(\beta_i \times \alpha_j), 1_X \rangle_X = \langle \beta_i, f^* \alpha_j \rangle_X.$$

When $i = j$ we obtain

$$\begin{aligned}(-1)^{\deg \alpha_i \deg \beta_i} \langle f^* \alpha_i, \beta_i \rangle_X &= \langle \beta_i \times \alpha_i, F_* 1_X \rangle_{X \times X} \\ &= c_{ii} \langle \beta_i \times \alpha_i, \alpha_i \times \beta_i \rangle = c_{ii} (-1)^{\deg \alpha_i (\deg \beta_i + \deg \alpha_i)} \langle \alpha_i, \beta_i \rangle_X \langle \alpha_i, \beta_i \rangle_X.\end{aligned}$$

(Note that for the corresponding differential forms one must let the form representing the last α_i commute through the form representing $\beta_i \times \alpha_i$.) Since $\deg \alpha_i - (\deg \alpha_i)^2$ is even it follows that

$$\sum_i c_{ii} = \sum_i (-1)^{\deg \alpha_i} \langle f^* \alpha_i, \beta_i \rangle = \sum_j (-1)^j \operatorname{Tr}(f^* | H^j(X)),$$

where $\operatorname{Tr}(F^* | H^j(X))$ denotes the trace of the linear map $f^* : H^j(X) \to H^j(X)$. Hence we have proved:

THEOREM 16.1 (LEFSCHETZ). *Let X be a compact oriented manifold, and let $f : X \to X$ be a continuous map. If F is the map $X \ni x \mapsto (x, f(x)) \to X \times X$ and δ is the diagonal map corresponding to the identity, then the intersection number between F and δ is equal to*

$$\sum_{j=0}^{\dim X} (-1)^j \operatorname{Tr}(f^*|H^j(X)), \tag{16.1}$$

which is called the Lefschetz number of f. It only depends on the homotopy class of f. When the Lefschetz number of f is not 0 then f must have a fixed point.

EXAMPLE 1. If f is homotopic to the identity map then the Lefschetz number is

$$\sum_{j=0}^{\dim X} (-1)^j \dim H^j(X)$$

which is called the *Euler characteristic* of X. Note that the Poincaré duality theorem implies that it is equal to 0 when n is odd.

Before giving other examples we shall write out explicitly the special cases of Theorem 16.1 when F and δ are transversal or just have finitely many intersections. Suppose first that $f \in C^\infty$ and that F and δ are transversal. This means that for a fixed point $x \in X$ the linear map

$$T_x(X) \oplus T_x(X) \ni (\xi_1, \xi_2) \mapsto (\xi_1 + \xi_2, f'(x)\xi_1 + \xi_2) \in T_x(X) \oplus T_x(X)$$

is bijective, and the sign of the intersection is the sign of this map. As in the second part of the proof of Theorem 15.5 we find that this is equivalent to the map

$$(\xi_1, \xi_2) \mapsto (\xi_1 - f'(x)\xi_1, f'(x)\xi_1 + \xi_2)$$

which is bijective if and only if $\xi_1 \neq f'(x)\xi_1$ when $\xi_1 \neq 0$. The sign is the same as for $\det(I - f'(x))$ where I is the identity in $T_x(X)$. The fixed point is said to be *non-degenerate* when this determinant is not equal to 0. As a special case of Theorem 16.1 we now obtain:

COROLLARY 16.2. *If X is a compact C^∞ (oriented) manifold and f is a C^∞ map $X \to X$ with no degenerate fixed point, then the number of fixed points x, counted with the sign of* $\det(I - f'(x))$*, is equal to the Lefschetz number* (16.1) *of f.*

Here the hypothesis that X is oriented has been put inside a parenthesis since it is not essential for the statement and not required for the result. However, we have only proved the theorem in the orientable case. We leave the extension to the reader and pass to a discussion of a continuous map $f : X \to X$ with only isolated fixed points. Let x_0 be a fixed point and identify a small open neighborhood of x_0 with a neighborhood of $0 \in \mathbf{R}^n$ by a diffeomorphism preserving the orientation. In order to apply Theorem 15.9 we must determine $(-1)^n D(\varphi, U \times U, 0)$ if U is a small convex neighborhood of 0 in $\mathbf{R}^n$ and

$$\varphi(x_1, x_2) = (x_1, f(x_1)) - (x_2, x_2).$$

For the homotopy

$$\varphi_t(x_1, x_2) = (tx_1 - x_2, f(x_1) - x_1 + t(tx_1 - x_2)), \quad (x_1, x_2) \in U \times U, \ t \in [0,1],$$

we have $\varphi_t(x_1, x_2) \neq 0$ unless $f(x_1) = x_1$ and $x_2 = tx_1$, hence $x_1 = x_2 = 0$. This implies that

$$D(\varphi, U \times U, 0) = D(\varphi_1, U \times U, 0) = D(\varphi_0, U \times U, 0) = (-1)^n D(I - f, U, 0),$$

since the determinant of the map $(y_1, y_2) \mapsto (-y_2, -y_1)$ in $\mathbf{R}^n \oplus \mathbf{R}^n$ is equal to $(-1)^n$. Thus we have proved:

Corollary 16.3. *Let X be a compact C^∞ (oriented) manifold and let $f : X \to X$ be a continuous map. If f has finitely many fixed points then the Lefschetz number* (16.1) *of f is equal to the sum of the indices of the fixed points. The index of a fixed point x_0 is defined by taking a diffeomorphism $\psi : \mathbf{R}^n \to U$ where U is a neighborhood of x_0 containing no other fixed point. If*

$$\varphi(x) = x - \psi^{-1} \circ f \circ \psi(x), \quad x \in \mathbf{R}^n,$$

then the index of x_0 is equal to $D(\varphi, \Omega, \psi^{-1}(x_0))$ where Ω is a bounded neighborhood of $\psi^{-1}(x_0)$.

Example 2. A continuous map $f : S^n \to S^n$ must have a fixed point unless the degree is equal to $(-1)^{n+1}$, which is the degree of the antipodal map $x \mapsto -x$ which obviously has no fixed point. In fact, the Lefschetz number is $1 + (-1)^n m$ where m is the degree.

Example 3. A continuous map $f : P^n_{\mathbf{C}} \to P^n_{\mathbf{C}}$ must have at least one fixed point unless n is odd and the degree is equal to -1. In fact, by Example 5 in Chapter V, given after Theorem 15.7, we know that if α is a generator for $H^2(P^n_{\mathbf{C}})$ then $f^*\alpha = c\alpha$, hence $f^*\alpha^k = c^k\alpha^k$, where c is an integer. (We do not need this information yet.) The Lefschetz number is therefore

$$1 + c + c^2 + \cdots + c^n = (c^{n+1} - 1)/(c - 1), \quad c \neq 1,$$

which has no real zero if n is even and only the real zero $c = -1$ when n is odd. If $c = -1$ and n is odd then the degree is $c^n = -1$. An example of a map without fixed point is defined in homogeneous coordinates by

$$(z_0, \dots, z_n) \mapsto (\bar{z}_1, -\bar{z}_0, \bar{z}_3, -\bar{z}_2, \dots, \bar{z}_n, -\bar{z}_{n-1})$$

when n is odd so that the number of homogeneous coordinates is even. For a fixed point the two $n+1$ tuples must be proportional, hence $z_0\bar{z}_0 + z_1\bar{z}_1 = 0, \dots$. This implies that $z_j = 0$ for every j, which does not define a point in $P^n_{\mathbf{C}}$.

As a special case of the Lefschetz formula (16.1) we shall now derive a theorem of Hopf on the zeros of a vector field on a manifold. Let v be a C^∞ vector field on the compact oriented manifold X, that is, a C^∞ section of the tangent bundle $T(X)$. By integration of the differential equation

$$df(x,t)/dt = v(f(x(t))), \quad f(x,0) = x, \tag{16.2}$$

we define a one parameter family (in fact a group) of C^∞ maps $f_t : X \ni x \mapsto f(t,x) \in X$ with f_0 equal to the identity. Since f_t is homotopic to the identity the Lefschetz number is equal to the Euler characteristic of X (see Example 1 above). On a compact set where $v \neq 0$ there is no fixed point for f_t when t is small, but a zero of v is a fixed point for every t. To determine when it is non-degenerate we choose local coordinates $x_1, \dots, x_n$ vanishing at a zero of v. Then the differential equations (16.2) for $f(x,t) = (f_1(x,t), \dots, f_n(x,t))$ have the form

$$\partial f_j(x,t)/\partial t = v_j(f(x,t)), \quad f_j(x,0) = x_j, \quad j = 0, \dots, n.$$

By Taylor's formula

$$f(x,t) = x + tv(x) + O(t^2);$$

more precisely, $(f(x,t) - x)/t \to v(x)$ in C^∞ in a neighborhood of $0 \in \mathbf{R}^n$ when $t \to 0$. If $\det \partial v(0)/\partial x \neq 0$, it follows from the implicit function theorem that there is a neighborhood of the origin where the equation $f(x,t) = x$ has only the zero $x = 0$ when t is small enough. We have

$$(I - f'_x(0,t))/t \to -\partial v(0)/\partial x \quad \text{when } t \to 0$$

which implies that

$$(-t)^n \det(I - f'_x(0,t)) \to \det \partial v/\partial x \quad \text{when } t \to 0.$$

Hence the fixed point is non-degenerate and the sign is equal to the sign of $\partial v/\partial x$ when n is even. When n is odd the sign also depends on the sign of t. This does not affect the following statement since the Euler characteristic is equal to 0 then.

Theorem 16.4 (Hopf). *Let X be a compact oriented C^∞ manifold, and let v be a C^∞ vector field on X such that $\det \partial v/\partial x \neq 0$ at every zero of v if $\partial v/\partial x$ is calculated in terms of local coordinates there. Then there are only finitely many zeros of v, and the number of zeros counted with the sign of $\det \partial v/\partial x$ is equal to the Euler characteristic*

$$\sum_{j=0}^{\dim X} (-1)^j \dim H^j(X).$$

We have of course a similar result for continuous vector fields with only finitely many zeros. The Euler characteristic is then equal to the sum of the indices of the zeros where the index of a zero is the degree of the map defined by the vector field in a small sphere around the zero in terms of local coordinates. The details of the statement and proof are left as an exercise.

Example 4. On a sphere of even dimension there are no vector fields without zeros since the Euler characteristic is equal to 2. However, such vector fields exist on all spheres of odd dimension (in fact, on all manifolds of odd dimension). An example on $S^{2n-1} \subset \mathbf{R}^{2n}$ is given by

$$v(x) = (-x_2, x_1, -x_3, x_2, \dots, -x_{2n}, x_{2n-1}), \quad x \in \mathbf{R}^{2n}.$$

Example 5. The Euler characteristic for $P^n_{\mathbf{C}}$ is equal to $n+1$, so vector fields on $P^n_{\mathbf{C}}$ have $n+1$ zeros "in general".

Notes. The fixed point theorems in this chapter were first proved in [H] and [L]. They are included in most textbooks on algebraic topology. A relatively recent extension of the Lefschetz fixed point theorems to complexes of differential operators other than the de Rham complex of exterior differential operators has been given in [AB]. Another proof can be found in [Hö, Section 19.4].

References

[AB] M. F. Atiyah and R. Bott, *The Lefschetz fixed-point theorem for elliptic complexes: I*, Ann. of Math. **86** (1967), 374–407; *II*, Ann. of Math. **88** (1968), 451–491.

[H] H. Hopf, *Abbildung geschlossener Manningfaltigkeiten auf Kugeln in n Dimensionen*, Jber. Deutsch. Math.-Verein. **34** (1925), 130–133.

[Hö] L. Hörmander, *The analysis of linear partial differential operators*, Springer Verlag, 1985.

[L] S. Lefschetz, *Intersections and transformations of complexes and manifolds*, Trans. Amer. Math. Soc. **28** (1926), 1–49.

Index