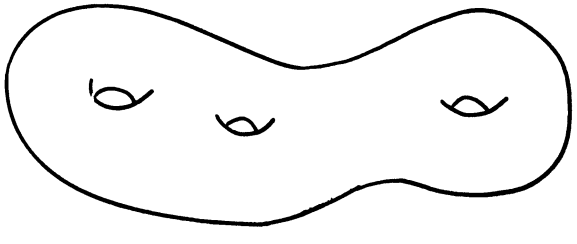
Introduction

The most intuitively evident topological invariant of a space is the number of connected pieces into which it falls. Over the past one hundred years or so we have come to realize that this primitive notion admits in some sense two higher-dimensional analogues. These are the *homotopy* and *cohomology groups* of the space in question.

The evolution of the higher homotopy groups from the component concept is deceptively simple and essentially unique. To describe it, let $\pi_0(X)$ denote the set of path components of X and if p is a point of X, let $\pi_0(X, p)$ denote the set $\pi_0(X)$ with the path component of p singled out. Also, corresponding to such a point p, let $\Omega_p X$ denote the space of maps (continuous functions) of the unit circle $\{z \in \mathbb{C} : |z| = 1\}$ which send 1 to p, made into a topological space via the compact open topology. The path components of this so-called *loop space* $\Omega_p X$ are now taken to be the elements of $\pi_1(X, p)$:

$$\pi_1(X, p) = \pi_0(\Omega_p X, \bar{p}).$$

The composition of loops induces a group structure on $\pi_1(X, p)$ in which the constant map $\bar{p}$ of the circle to p plays the role of the identity; so endowed, $\pi_1(X, p)$ is called the *fundamental group* or the *first homotopy group* of X at p. It is in general not Abelian. For instance, for a Riemann surface of genus 3, as indicated in the figure below:



R. Bott et al., *Differential Forms in Algebraic Topology*

$\pi_1(X, p)$ is generated by six elements $\{x_1, x_2, x_3, y_1, y_2, y_3\}$ subject to the single relation

$$\prod_{i=1}^{3} [x_i, y_i] = 1,$$

where $[x_i, y_i]$ denotes the commutator $x_i y_i x_i^{-1} y_i^{-1}$ and 1 the identity. The fundamental group is in fact sufficient to classify the closed oriented 2-dimensional surfaces, but is insufficient in higher dimensions.

To return to the general case, all the higher homotopy groups $\pi_k(X, p)$ for $k \geq 2$ can now be defined through the inductive formula:

$$\pi_{k+1}(X, p) = \pi_k(\Omega_p X, \bar{p}).$$

By the way, if p and p' are two points in X in the same path component, then

$$\pi_k(X, p) \simeq \pi_k(X, p'),$$

but the correspondence is not necessarily unique. For the Riemann surfaces such as discussed above, the higher π_k's for $k \geq 2$ are all trivial, and it is in part for this reason that π_1 is sufficient to classify them. The groups π_k for $k \geq 2$ turn out to be Abelian and therefore do not seem to have been taken seriously until the 1930's when W. Hurewicz defined them (in the manner above, among others) and showed that, far from being trivial, they constituted the basic ingredients needed to describe the homotopy-theoretic properties of a space.

The great drawback of these easily defined invariants of a space is that they are very difficult to compute. To this day not all the homotopy groups of say the 2-sphere, i.e., the space $x^2 + y^2 + z^2 = 1$ in $\mathbb{R}^3$, have been computed! Nonetheless, by now much is known concerning the general properties of the homotopy groups, largely due to the formidable algebraic techniques to which the "cohomological extension" of the component concept lends itself, and the relations between homotopy and cohomology which have been discovered over the years.

This cohomological extension starts with the dual point of view in which a component is characterized by the property that on it *every locally constant function is globally constant.* Such a component is sometimes called a connected component, to distinguish it from a *path component.* Thus, if we define $H^0(X)$ to be the vector space of real-valued *locally constant* functions on X, then dim $H^0(X)$ tells us the number of connected components of X. Note that on reasonable spaces where path components and connected components agree, we therefore have the formula

$$\text{cardinality } \pi_0(X) = \dim H^0(X).$$

Still the two concepts are dual to each other, the first using maps of the unit interval into X to test for connectedness and the second using maps of X

into $\mathbb{R}$ for the same purpose. One further difference is that the cohomology group $H^0(X)$ has, by fiat, a natural $\mathbb{R}$-module structure.

Now what should the proper higher-dimensional analogue of $H^0(X)$ be? Unfortunately there is no decisive answer here. Many plausible definitions of $H^k(X)$ for $k > 0$ have been proposed, all with slightly different properties but all isomorphic on "reasonable spaces". Furthermore, in the realm of differentiable manifolds, all these theories coincide with the *de Rham theory* which makes its appearance there and constitutes in some sense the most perfect example of a cohomology theory. The de Rham theory is also unique in that it stands at the crossroads of topology, analysis, and physics, enriching all three disciplines.

The gist of the "de Rham extension" is comprehended most easily when M is assumed to be an open set in some Euclidean space $\mathbb{R}^n$, with coordinates $x_1, \ldots, x_n$. Then amongst the C^∞ functions on M the locally constant ones are precisely those whose gradient

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

vanishes identically. Thus here $H^0(M)$ appears as the space of solutions of the differential equation $df = 0$. This suggests that $H^1(M)$ should also appear as the space of solutions of some natural differential equations on the manifold M. Now consider a 1-form on M:

$$\theta = \sum a_i\, dx_i,$$

where the a_i's are C^∞ functions on M. Such an expression can be integrated along a smooth path γ, so that we may think of θ as a function on paths γ:

$$\gamma \mapsto \int_\gamma \theta.$$

It then suggests itself to seek those θ which give rise to *locally constant* functions of γ, i.e., for which the integral $\int_\gamma \theta$ is left unaltered under small variations of γ—but keeping the endpoints fixed! (Otherwise, only the zero 1-form would be locally constant.) Stokes' theorem teaches us that these line integrals are characterized by the differential equations:

$$\frac{\partial a_i}{\partial x_j} - \frac{\partial a_j}{\partial x_i} = 0 \qquad \text{(written } d\theta = 0\text{)}.$$

On the other hand, the fundamental theorem of calculus implies that $\int_\gamma df = f(Q) - f(P)$, where P and Q are the endpoints of γ, so that *the gradients are trivially locally constant.*

One is here irresistibly led to the definition of $H^1(M)$ as the vector space of *locally constant line integrals modulo the trivially constant ones.* Similarly the higher cohomology groups $H^k(M)$ are defined by simply replacing line integrals with their higher-dimensional analogues, the *k-volume integrals.*

The Grassmann calculus of exterior differential forms facilitates these extensions quite magically. Moreover, the differential equations characterizing the locally constant k-integrals are seen to be C^∞ invariants and so extend naturally to the class of C^∞ manifolds.

Chapter I starts with a rapid account of this whole development, assuming little more than the standard notions of advanced calculus, linear algebra and general topology. A nodding acquaintance with singular homology or cohomology helps, but is not necessary. No real familiarity with differential geometry or manifold theory is required. After all, the concept of a manifold is really a very natural and simple extension of the calculus of several variables, as our fathers well knew. Thus for us a manifold is essentially a space constructed from open sets in $\mathbb{R}^n$ by patching them together in a smooth way. This point of view goes hand in hand with the "computability" of the de Rham theory. Indeed, the decisive difference between the π_k's and the H^k's in this regard is that if a manifold X is the union of two open submanifolds U and V:

$$X = U \cup V,$$

then the cohomology groups of U, V, $U \cap V$, and X are linked by a much stronger relation than the homotopy groups are. The linkage is expressed by the exactness of the following sequence of linear maps, *the Mayer–Vietoris sequence*:

$$\xrightarrow{d^*} H^{k+1}(X) \to H^k(X) \to H^k(U) \oplus H^k(V) \to H^k(U \cap V) \xrightarrow{d^*} \to H^{k-1}(U \cap V) \to \cdots \to H^0(X) \to 0$$

starting with $k = 0$ and extending up indefinitely. In this sequence every arrow stands for a linear map of the vector spaces and exactness asserts that the kernel of each map is precisely the image of the preceding one. The horizontal arrows in our diagram are the more or less obvious ones induced by restriction of functions, but the coboundary operator d^* is more subtle and uses the existence of a *partition of unity* subordinate to the cover $\{U, V\}$ of X, that is, smooth functions ρ_U and ρ_V such that the first has support in U, the second has support in V, and $\rho_U + \rho_V \equiv 1$ on X. The simplest relation imaginable between the H^k's of U, V, and $U \cup V$ would of course be that H^k behaves additively; the Mayer–Vietoris sequence teaches us that this is indeed the case if U and V are disjoint. Otherwise, there is a geometric feedback from $H^k(U \cap V)$ described by d^*, and one of the hallmarks of a topologist is a sound intuition for this d^*.

The exactness of the Mayer–Vietoris sequence is our first goal once the basics of the de Rham theory are developed. Thereafter we establish the

second essential property for the computability of the theory, namely that *for a smoothly* contractible manifold M,

$$H^k(M) = \begin{cases} \mathbb{R} & \text{for } k = 0, \\ 0 & \text{for } k > 0. \end{cases}$$

This *homotopy invariance* of the de Rham theory can again be thought of as having evolved from the fundamental theorem of calculus. Indeed, the formula

$$f(x)\,dx = d\int_0^x f(u)\,du$$

shows that every line integral (1-form) on $\mathbb{R}^1$ is a gradient, whence $H^1(\mathbb{R}^1) = 0$. The homotopy invariance is thus established for the real line. This argument also paves the way for the general case.

The two properties that we have just described constitute a verification of the *Eilenberg–Steenrod axioms* for the de Rham theory in the present context. Combined with a little geometry, they can be used in a standard manner to compute the cohomology of simple manifolds. Thus, for spheres one finds

$$H^k(S^n) = \begin{cases} \mathbb{R} & \text{for } k = 0 \text{ or } n \\ 0 & \text{otherwise,} \end{cases}$$

while for a Riemann surface X_g with g holes,

$$H^k(X_g) = \begin{cases} \mathbb{R} & \text{for } k = 0 \text{ or } 2 \\ \mathbb{R}^{2g} & \text{for } k = 1 \\ 0 & \text{otherwise.} \end{cases}$$

A more systematic treatment in Chapter II leads to the computability proper of the de Rham theory in the following sense. By a finite good cover of M we mean a covering $\mathfrak{U} = \{U_\alpha\}_{\alpha=1}^N$ of M by a finite number of open sets such that all intersections $U_{\alpha_1} \cap \cdots \cap U_{\alpha_k}$ are either vacuous or contractible. The purely combinatorial data that specify for each subset $\{\alpha_1, \ldots, \alpha_k\}$ of $\{1, \ldots, N\}$ which of these two alternatives holds are called the *incidence data* of the cover. The computability of the theory is the assertion that it can be computed purely from such incidence data. Along lines established in a remarkable paper by André Weil [1], we show this to be the case for the de Rham theory. Weil's point of view constitutes an alternate approach to the sheaf theory of Leray and was influential in Cartan's *théorie des carapaces*. The beauty of his argument is that it can be read both ways: either to prove the computability of de Rham or to prove the topological invariance of the combinatorial prescription.

To digress for a moment, it is difficult not to speculate about what kept Poincaré from discovering this argument forty years earlier. One has the feeling that he already knew every step along the way. After all, the homotopy invariance of the de Rham theory for $\mathbb{R}^n$ is known as the Poincaré

lemma! Nevertheless, he veered sharply from this point of view, thinking predominantly in terms of triangulations, and so he in fact was never able to prove either the computability of de Rham or the invariance of the combinatorial definition. Quite possibly the explanation is that the whole C^∞ point of view and, in particular, the partitions of unity were alien to him and his contemporaries, steeped as they were in real or complex analytic questions.

De Rham was of course the first to prove the topological invariance of the theory that now bears his name. He showed that it was isomorphic to the *singular cohomology*, which is trivially—i.e., by definition—topologically invariant. On the other hand, André Weil's approach relates the de Rham theory to the *Čech theory*, which is again topologically invariant.

But to return to the plan of our book, the bulk of Chapter I is actually devoted to explaining the fundamental symmetry in the cohomology of a compact oriented manifold. In its most primitive form this symmetry asserts that

$$\dim H^q(M) = \dim H^{n-q}(M).$$

Poincaré seems to have immediately realized this consequence of the locally Euclidean nature of a manifold. He saw it in terms of dual subdivisions, which turn the incidence relations upside down. In the de Rham theory the duality derives from the intrinsic pairing between differential forms of arbitrary and compact support. Indeed consider the de Rham theory of $\mathbb{R}^1$ with compactly supported forms. Clearly the only locally constant function with compact support on $\mathbb{R}^1$ is the zero function. As for 1-forms, not every 1-form $g\,dx$ is now a gradient of a *compactly supported* function f; this happens if and only if $\int_{-\infty}^{\infty} g\,dx = 0$. Thus we see that the compactly supported de Rham theory of $\mathbb{R}^1$ is given by

$$H_c^k(\mathbb{R}^1) = \begin{cases} 0 & \text{for } k = 0 \\ \mathbb{R} & \text{for } k = 1, \end{cases}$$

and is just the de Rham theory "upside down." This phenomenon now extends inductively to $\mathbb{R}^n$ and is finally propagated via the Mayer–Vietoris sequence to the cohomology of any compact oriented manifold.

One virtue of the de Rham theory is that the essential mechanism of this duality is via the familiar operation of integration, coupled with the natural ring structure of the theory: a p-form θ can be multiplied by a q-form ϕ to produce a $(p+q)$-form $\theta \wedge \phi$. This multiplication is "commutative in the graded sense":

$$\theta \wedge \phi = (-1)^{pq} \phi \wedge \theta.$$

(By the way, the commutativity of the de Rham theory is another reason why it is more "perfect" than its other more general brethren, which become commutative only on the cohomology level.) In particular, if ϕ has compact support and is of dimension $n - p$, where $n = \dim M$, then inte-

gration over M gives rise to a pairing

$$(\theta, \phi) \to \int_M \theta \wedge \phi,$$

which descends to cohomology and induces a pairing

$$H^p(M) \otimes H_c^{n-p}(M) \to \mathbb{R}.$$

A more sophisticated version of Poincaré duality is then simply that *the pairing above is dual*; that is, it establishes the two spaces as duals of each other.

Although we return to Poincaré duality over and over again throughout the book, we have not attempted to give an exhaustive treatment. (There is, for instance, no mention of Alexander duality or other phenomena dealing with relative, rather than absolute, theory.) Instead, we chose to spend much time bringing Poincaré duality to life by explicitly constructing the Poincaré dual of a submanifold N in M. The problem is the following. Suppose $\dim N = k$ and $\dim M = n$, both being compact oriented. Integration of a k-form ω on M over N then defines a linear functional from $H^k(M)$ to $\mathbb{R}$, and so, by Poincaré duality, must be represented by a cohomology class in $H^{n-k}(M)$. The question is now: how is one to construct a representative of this Poincaré dual for N, and can such a representative be made to have support arbitrarily close to N?

When N reduces to a point p in M, this question is easily answered. The dual of p is represented by any n-form ω with support in the component M_p of p and with total mass 1, that is, with

$$\int_{M_p} \omega = 1.$$

Note also that such an ω can be found with support in an arbitrarily small neighborhood of p, by simply choosing coordinates on M centered at p, say $x_1, \ldots, x_n$, and setting

$$\omega = \lambda(x) dx_1 \ldots dx_n$$

with λ a bump function of mass 1. (In the limit, thinking of Dirac's δ-function as the Poincaré dual of p leads us to de Rham's theory of currents.)

When the point p is replaced by a more general submanifold N, it is easy to extend this argument, provided N has a *product neighborhood* $D(N)$ in M in the sense that $D(N)$ is diffeomorphic to the product $N \times D^{n-k}$, where D^{n-k} is a disk of the dimension indicated. However, this need not be the case! Just think of the center circle in a Möbius band. Its neighborhoods are at best smaller Möbius bands.

In the process of constructing the Poincaré dual we are thus confronted by the preliminary question of how to measure the possible twistings of neighborhoods of N in M and to correct for the twist. This is a subject in its own right nowadays, but was initiated by H. Whitney and H. Hopf in just

the present context during the Thirties and Forties. Its trade name is *fiber bundle theory* and the cohomological measurements of the global twist in such "local products" as $D(N)$ are referred to as *characteristic classes*. In the last forty years the theory of characteristic classes has grown to such an extent that we cannot do it justice in our book. Still, we hope to have covered it sufficiently so that the reader will be able to see its ramifications in both differential geometry and topology. We also hope that our account could serve as a good introduction to the connection between characteristic classes and the global aspects of the gauge theories of modern physics.

That a connection between the equations of mathematical physics and topology might exist is not too surprising in view of the classical theory of electricity. Indeed, in a vacuum the electromagnetic field is represented by a 2-form in the (x, y, z, t)-space:

$$\omega = (E_x\, dx + E_y\, dy + E_z\, dz)dt + H_x\, dy\, dz - H_y\, dx\, dz + H_z\, dx\, dy,$$

and the form ω is locally constant in our sense, i.e., $d\omega = 0$. Relative to the Lorentz metric in $\mathbb{R}^4$ the *star* of ω is defined to be

$$*\omega = -(H_x\, dx + H_y\, dy + H_z\, dz)dt + E_x\, dy\, dz - E_y\, dx\, dz + E_z\, dx\, dz,$$

and Maxwell's equations simply assert that both ω and its star are closed: $d\omega = 0$ and $d*\omega = 0$. In particular, the cohomology class of $*\omega$ is a well defined object and is often of physical interest.

To take the simplest example, consider the Coulomb potential of a point charge q at rest in the origin of our coordinate system. The field ω generated by this charge then has the description

$$\omega = -qd\left(\frac{1}{r} \cdot dt\right)$$

with $r = (x^2 + y^2 + z^2)^{1/2} \neq 0$. Thus ω is defined on $\mathbb{R}^4 - \mathbb{R}_t$, where $\mathbb{R}_t$ denotes the t-axis. The de Rham cohomology of this set is easily computed to be

$$H^k(\mathbb{R}^4 - \mathbb{R}_t) = \begin{cases} \mathbb{R} & \text{for } k = 0, 2 \\ 0 & \text{otherwise.} \end{cases}$$

The form ω is manifestly cohomologically uninteresting, since it is d of a 1-form and so is trivially "closed", i.e., locally constant. On the other hand the $*$ of ω is given by

$$*\omega = \frac{q}{4\pi} \frac{x\, dy\, dz - y\, dx\, dz + z\, dx\, dy}{r^3},$$

which turns out to generate H^2. The cohomology class of $*\omega$ can thus be interpreted as the charge of our source.

In seeking differential equations for more sophisticated phenomena than electricity, the modern physicists were led to equations (the Yang–Mills) which fit perfectly into the framework of characteristic classes as developed by such masters as Pontrjagin and Chern during the Forties.

Having sung the praises of the de Rham theory, it is now time to admit its limitations. The trouble with it, is that it only tells part of the cohomology story and from the point of view of the homotopy theorists, only the simplest part. The de Rham theory ignores torsion phenomena. To explain this in a little more detail, recall that the homotopy groups do not behave well under the union operation. However, they behave very well under Cartesian products. Indeed, as is quite easily shown,

$$\pi_q(X \times Y) = \pi_q(X) \oplus \pi_q(Y).$$

More generally, consider the situation of a *fiber bundle* (twisted product). Here we are dealing with a space E mapped onto a space X with the fibers—i.e., the inverse images of points —all homeomorphic in some uniform sense to a fixed space Y. For fiber bundles, the additivity of π_q is stretched into an infinite exact sequence of Mayer-Vietoris type, however now going in the opposite direction:

$$\cdots \to \pi_q(Y) \to \pi_q(E) \to \pi_q(X) \to \pi_{q-1}(Y) \to \cdots.$$

This phenomenon is of course fundamental in studying the twist we talked about earlier, but it also led the homotopy theorists to the conjecture that in their much more flexible homotopy category, where objects are considered equal if they can be deformed into each other, every space factors into a twisted product of irreducible prime factors. This turns out to be true and is called the *Postnikov decomposition* of the space. Furthermore, the "prime spaces" in this context all *have nontrivial homotopy groups in only one dimension.* Now in the homotopy category such a prime space, say with nontrivial homotopy group π in dimension n, is determined uniquely by π and n and is denoted $K(\pi, n)$. These $K(\pi, n)$-spaces of Eilenberg and Mac-Lane therefore play an absolutely fundamental role in homotopy theory. They behave well under the standard group operations. In particular, corresponding to the usual decomposition of a finitely generated Abelian group:

$$\pi = \left(\bigoplus_p \pi^{(p)} \right) \oplus \mathbb{Z}^r$$

into p-primary parts and a free part (said to correspond to the prime at infinity), the $K(\pi, n)$ will factor into a product

$$K(\pi, n) = \left(\prod_p K(\pi^{(p)}, n) \right) \cdot K(\mathbb{Z}, n)^r.$$

It follows that in homotopy theory, just as in many questions of number theory, one can work one prime at a time. In this framework it is now quite easy to explain the shortcomings of the de Rham theory: the theory is sensitive only to the prime at infinity!

After having encountered the Čech theory in Chapter II, we make in Chapter III the now hopefully easy transition to cohomology with coefficients in an arbitrary Abelian group. This theory, say with coefficients in the

integers, is then sensitive to all the p-primary phenomena in homotopy theory.

The development sketched here is discussed in greater detail in Chapter III, where we also apply the ideas to the computation of some relatively simple homotopy groups. All these computations in the final analysis derive from Serre's brilliant idea of applying the spectral sequence of Leray to homotopy problems and from his coining of a sufficiently general definition of a twisted product, so that, as the reader will see, the Postnikov decomposition in the form we described it, is a relatively simple matter. It remains therefore only to say a few words to the uninitiated about what this "spectral sequence" is.

We remarked earlier that homotopy behaves additively under products. On the other hand, cohomology does not. In fact, neglecting matters of torsion, i.e., reverting to the de Rham theory, one has the *Künneth formula*:

$$H^k(X \times Y) = \sum_{p+q=k} H^p(X) \otimes H^q(Y).$$

The next question is of course how cohomology behaves for twisted products. It is here that Leray discovered some a priori bounds on the extent and manner in which the Künneth formula can fail due to a twist. For instance, one of the corollaries of his spectral sequence is that if X and Y have vanishing cohomology in positive dimensions less than p and q respectively, then however one twists X with Y, the Künneth formula will hold up to dimension $d < \min(p, q)$.

Armed with this sort of information, one can first of all compute the early part of the cohomology of the $K(\pi, n)$ inductively, and then deduce which $K(\pi, n)$ must occur in a Postnikov decomposition of X by comparing the cohomology on both sides. This procedure is of course at best ad hoc, and therefore gives us only fragmentary results. Still, the method points in the right direction and can be codified to prove the computability (in the logical sense) of any particular homotopy group, of a sphere, say. This theorem is due to E. Brown in full generality. Unfortunately, however, it is not directly applicable to explicit calculations—even with large computing machines.

So far this introduction has been written with a lay audience in mind. We hope that what they have read has made sense and has whetted their appetites. For the more expert, the following summary of the plan of our book might be helpful.

In Chapter I we bring out from scratch Poincaré duality and its various extensions, such as the Thom isomorphism, all in the de Rham category. Along the way all the axioms of a cohomology theory are encountered, but at first treated only in our restricted context.

In Chapter II we introduce the techniques of spectral sequences as an extension of the Mayer–Vietoris principle and so are led to A. Weil's Čech–de Rham theory. This theory is later used as a bridge to cohomology

in general and to integer cohomology in particular. We spend considerable time patching together the Euler class of a sphere bundle and exploring its relation to Poincaré duality. We also very briefly present the sheaf-theoretic proof of this duality.

In Chapter III we come to grips with spectral sequences in a more formal manner and describe some of their applications to homotopy theory, for example, to the computation of $\pi_5(S^3)$. This chapter is less self-contained than the others and is meant essentially as an introduction to homotopy theory proper. In the same spirit we close with a short account of Sullivan's rational homotopy theory.

Finally, in Chapter IV we use the Grothendieck approach towards characteristic classes to give a more or less self-contained treatment of Chern and Pontrjagin classes. We then relate them to the cohomology of the infinite Grassmannian.

Unfortunately there was no time left within the scope of our book to explain the functorial approach to classifying spaces in general and to make the connection with the Eilenberg–MacLane spaces. We had to relegate this material, which is most naturally explained in the framework of semi-simplicial theory, to a mythical second volume. The novice should also be warned that there are all too many other topics which we have not mentioned. These include generalized cohomology theories, cohomology operations, and the Adams and Eilenberg–Moore spectral sequences. Alas, there is also no mention of the truly geometric achievements of modern topology, that is, handlebody theory, surgery theory, and the structure theory of differentiable and piecewise linear manifolds. Still, we hope that our volume serves as an introduction to all this as well as to such topics in analysis as Hodge theory and the Atiyah–Singer index theorems for elliptic differenital operators.