though of course it would bring us no nearer to turning the conjecture into a theorem.

In this article we have seen only a glimpse of the modern theory of arithmetic geometry, and perhaps I have overemphasized mathematicians' successes at the expense of the much larger territory of questions, like Lang's conjecture above, about which we remain wholly ignorant. At this stage in the history of mathematics, we can confidently say that the schemes attached to Diophantine problems *have geometry*. What remains is to say as much as we can about *what this geometry is like*, and in this respect, despite the progress described here, our understanding is still quite unsatisfactory when compared with our knowledge of more classical geometric situations.

Further Reading

Dieudonné, J. 1985. *History of Algebraic Geometry*. Monterey, CA: Wadsworth.

Silverman, J., and J. Tate. 1992. *Rational Points on Elliptic Curves*. New York: Springer.

IV.6 Algebraic Topology

Burt Totaro

Introduction

Topology is concerned with the properties of a geometric shape that are unchanged when we continuously deform it. In more technical terms, topology tries to classify TOPOLOGICAL SPACES [III.90], where two spaces are considered the same if they are homeomorphic. Algebraic topology assigns numbers to a topological space, which can be thought of as the "number of holes" in that space. These holes can be used to show that two spaces are not homeomorphic: if they have different numbers of holes of some kind, then one cannot be a continuous deformation of the other. In the happiest cases, we can hope to show the converse statement: that two spaces with the same number of holes (in some precise sense) *are* homeomorphic.

Topology is a relatively new branch of mathematics, with its origins in the nineteenth century. Before that, mathematics usually sought to solve problems exactly: to solve an equation, to find the path of a falling body, to compute the probability that a game of dice will lead to bankruptcy. As the complexity of mathematical problems grew, it became clear that most problems would never be solved by an exact formula: a classic example is the problem, known as THE THREE-BODY PROBLEM [V.33], of computing the future movements of Earth, the Sun, and the Moon under the influence of gravity. Topology allows the possibility of making qualitative predictions when quantitative ones are impossible. For example, a simple topological fact is that a trip from New York to Montevideo must cross the equator at some point, although we cannot say exactly where.

1 Connectedness and Intersection Numbers

Perhaps the simplest topological property is one called *connectedness*. This can be defined in various ways, as we shall see in a moment, but once we have a notion of what it means for a space to be connected we can then divide a topological space up into connected pieces, called *components*. The number of these pieces is a simple but useful INVARIANT [I.4 §2.2]: if two spaces have different numbers of connected components, then they are not homeomorphic.

For nice topological spaces, the different definitions of connectedness are equivalent. However, they can be generalized to give ways of measuring the number of holes in a space; these generalizations are interestingly different and all of them are important.

The first interpretation of connectedness uses the notion of a *path*, which is defined to be a continuous mapping f from the unit interval $[0,1]$ to a given space X. (We think of f as a path from $f(0)$ to $f(1)$.) Let us declare two points of X to be equivalent if there is a path from one to the other. The set of EQUIVALENCE CLASSES [I.2 §2.3] is called the set of *path components* of X and is written $\pi_0(X)$. This is a very natural way of defining the "number of connected pieces" into which X breaks up. One can generalize this notion by considering mappings into X from other standard spaces such as spheres: this leads to the notion of homotopy groups, which will be the topic of section 2.

A different way of thinking about connectedness is based on functions from X to the real line rather than functions from a line segment into X. Let us assume that we are in a situation where it makes sense to differentiate functions on X. For example, X could be an open subset of some Euclidean space, or more generally a SMOOTH MANIFOLD [I.3 §6.9]. Consider all the real-valued functions on X whose derivative is everywhere equal to zero: these functions form a real VECTOR SPACE [I.3 §2.3], which we call $H^0(X,\mathbb{R})$ (the "zeroth cohomology group of X with real coefficients"). Calculus tells us that if a function defined on an interval has derivative zero, then it must be constant, but that is not true when the domain has several connected pieces: all we can say then is that the function is constant on each connected piece of X. The number of degrees of freedom of such a function is therefore equal to the number of connected pieces, so the dimension of the vector space $H^0(X, \mathbb{R})$ is another way to describe the number of connected components of X. This is the simplest example of a cohomology group. Cohomology will be discussed in section 4.

We can use the idea of connectedness to prove a serious theorem of algebra: every real polynomial of odd degree has a real root. For example, there must be some real number x such that $x^3 + 3x - 4 = 0$. The basic observation is that when x is a large positive number or a highly negative number, the term x^3 is much bigger (in absolute value) than the other terms of the polynomial. Since this top term is an odd power of x, we have $f(x) > 0$ for some positive number x and $f(x) < 0$ for some negative number x. If f were never equal to zero, then it would be a continuous mapping from the real line into the real line minus the origin. But the real line is connected, while the real line minus the origin has two connected components, the positive and negative numbers. It is easy to show that a continuous map from a connected space X to another space Y must map X into just one connected component of Y: in our case, this contradicts the fact that f takes both positive and negative values. Therefore f must be equal to zero at some point, and the proof is complete.

This argument can be phrased in terms of the "intermediate value theorem" of calculus, which is indeed one of the most basic topological theorems. An equivalent reformulation of this theorem states that a continuous curve that goes from the lower half-plane to the upper half-plane must cross the horizontal axis at some point. This idea leads to *intersection numbers*, one of the most useful concepts in topology. Let M be a smooth oriented manifold. (Roughly speaking, a manifold is oriented if you cannot continuously slide a shape about inside it and end up with a reflection of that shape. The simplest nonoriented manifold is a Möbius strip: to reflect a shape, slide it around the strip an odd number of times.) Let A and B be two closed oriented submanifolds of M with dimensions adding up to the dimension of M. Finally, suppose that A and B intersect transversely, so that their intersection has the "correct" dimension, namely 0, and is therefore a collection of separated points.

Now let p be one of these points. There is a way of assigning a weight of $+1$ or -1 to p, which depends

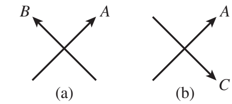

Figure 1 Intersection numbers: (a) $A \cdot B = 1$; (b) $A \cdot C = -1$.

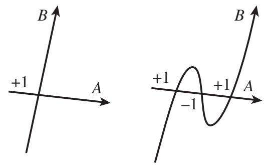

Figure 2 Moving a submanifold.

in a natural way on the relationship between the orientations of A, B, and M (see figure 1). For example, if M is a sphere, A is the equator of M, B is a closed curve, and appropriate directions are given to A and B, then the weight of p will tell you whether B crosses A upwards or downwards at p. If A and B intersect in only finitely many points, then we can define the intersection number of A and B, written $A \cdot B$, to be the sum of the weights ($+1$ or -1) at all the intersection points. In particular, this will happen if M is COMPACT [III.9] (that is, we can think of it as a closed bounded subset of $\mathbb{R}^N$ for some N).

The important point about the intersection number is that it is an *invariant*, in the following sense: if you move A and B about in a continuous way, ending up with another pair of transverse submanifolds A' and B', then the intersection number $A' \cdot B'$ is the same as $A \cdot B$, even though the number of intersection points can change. To see why this might be true, consider again the case where A and B are curves and M is two dimensional: if A and B meet at a point with weight 1, we can wiggle one of them to turn that point into three points with weights 1, -1, and 1, but the total contribution to the intersection number is unchanged. This is illustrated in figure 2. As a result, the intersection number $A \cdot B$ is defined for *any* two submanifolds of complementary dimension: if they do not intersect transversely, one can move them until they do and use the definition we have just given.

In particular, if two submanifolds have nonzero intersection number, then they can never be moved to

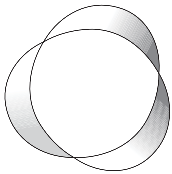

Figure 3 A surface bounded by a knot.

be disjoint from each other. This is another way to describe the earlier arguments about connectedness. It is easy to write down one curve from New York to Montevideo whose intersection number with the equator is equal to 1. Therefore, no matter how we move that curve (provided that we keep the endpoints fixed: more generally, if either A or B has a boundary, then that boundary should be kept fixed), its intersection number with the equator will always be 1, and in particular it must meet the equator in at least one point.

One of many applications of intersection numbers in topology is the idea of *linking numbers*, which comes from KNOT THEORY [III.44]. A *knot* is a path in space that begins and ends at the same point, or, more formally, a closed connected one-dimensional submanifold of $\mathbb{R}^3$. Given any knot K, it is always possible to find a surface S in $\mathbb{R}^3$ with K as its boundary (see figure 3). Now let L be a knot that is disjoint from K. The linking number of K with L is defined to be the intersection number of L with the surface S. The properties of intersection numbers imply that if the linking number of K with L is nonzero, then the knots K and L are "linked," in the sense that it is impossible to pull them apart.

2 Homotopy Groups

If we remove the origin from the plane $\mathbb{R}^2$, then we obtain a new space that is different from the plane in a fundamental way: it has a hole in it. However, we cannot detect this difference by counting components, since both the plane and the plane without the origin are connected. We begin this section by defining an invariant called the *fundamental group*, which does detect this kind of hole.

As a first approximation, one could say that the elements of the fundamental group of a space X are *loops*, which can be formally defined as continuous functions f from $[0,1]$ to X such that $f(0) = f(1)$. However, this is not quite accurate, for two reasons. The first reason, which is extremely important, is that two loops

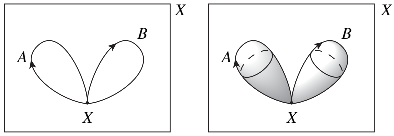

Figure 4 Multiplication in the fundamental group and in higher homotopy groups.

are regarded as equivalent if one can be continuously deformed to the other while all the time staying inside X. If this is the case, we say that they are *homotopic*. To be more formal about this, let us suppose that f_0 and f_1 are two loops. Then a *homotopy* between f_0 and f_1 is a collection of loops f_s in X, one for each s between 0 and 1, such that the function $F(s,t) = f_s(t)$ is a continuous function from $[0,1]^2$ to X. Thus, as s increases from 0 to 1, the loop f_s moves continuously from f_0 to f_1. If two loops are homotopic, then we count them as the same. So the elements of the homotopy group are not actually loops but equivalence classes, or *homotopy classes*, of loops.

Even this is not quite correct, because for technical reasons we need to impose an extra condition on our loops: that they all start from (and therefore end at) a given point, called the *base point*. If X is connected, it turns out not to matter what this base point is, but we need it to be the same for all loops. The reason for this is that it gives us a way to multiply two loops: if x is the base point and A and B are two loops that start and end at x, then we can define a new loop by going around A and then going around B. This is illustrated in figure 4. We regard this new loop as the product of the loops A and B. It is not hard to check that the homotopy class of this product depends only on the homotopy classes of A and B, and that the resulting binary operation turns the set of homotopy classes of loops into a GROUP [I.3 §2.1]. It is this group that we call the fundamental group of X. It is denoted $\pi_1(X)$.

The fundamental group can be computed for most of the spaces we are likely to encounter. This makes it an important way to distinguish one space from another. First of all, for any n the fundamental group of $\mathbb{R}^n$ is the trivial group with just one element, because any loop in $\mathbb{R}^n$ can be continuously shrunk to its base point. On the other hand, the fundamental group of $\mathbb{R}^2 \setminus \{0\}$, the plane with the origin removed, is isomorphic to the group $\mathbb{Z}$ of the integers. This tells us that we can associate with any loop in $\mathbb{R}^2 \setminus \{0\}$ an integer that does not change if we modify the loop in a continuous way. This integer is known as the *winding number*. Intuitively, the winding number measures the total number of times that the mapping goes around the origin, with counterclockwise circuits counting positively and clockwise ones negatively. Since the fundamental group of $\mathbb{R}^2 \setminus \{0\}$ is not the trivial group, $\mathbb{R}^2 \setminus \{0\}$ cannot be homeomorphic to the plane. (It is an interesting exercise to try to find an elementary proof of this result—that is, a proof that does not use, or implicitly reconstruct, any of the machinery of algebraic topology. Such proofs do exist, but it is tricky to find them.)

A classic application of the fundamental group is to prove THE FUNDAMENTAL THEOREM OF ALGEBRA [V.13], which states that every nonconstant polynomial with complex coefficients has a complex root. (The proof is sketched in the article just cited, though the fundamental group is not explicitly mentioned there.)

The fundamental group tells us about the number of "one-dimensional holes" that a space has. A basic example is given by the circle, which has fundamental group $\mathbb{Z}$, just as $\mathbb{R}^2 \setminus \{0\}$ does, and for essentially the same reason: given a path in the circle that begins and ends at the same point, we can see how many times it goes around the circle. In the next section we shall see some more examples.

Before we think about higher-dimensional holes, we first need to discuss one of the most important topological spaces: the n-dimensional sphere. For any natural number n, this is defined to be the set of points in $\mathbb{R}^{n+1}$ at distance 1 from the origin. It is denoted S^n. Thus, the 0-sphere S^0 consists of two points, the 1-sphere S^1 is the circle, and the 2-sphere S^2 is the usual sphere, like the surface of Earth. Higher-dimensional spheres take a little bit of getting used to, but we can work with them in the same way that we can with lower-dimensional spheres. For example, we can construct the 2-sphere from a closed two-dimensional disk by identifying all the points on the boundary circle with each other. In the same way, the 3-sphere can be obtained from a solid three-dimensional ball by identifying all the points on the boundary 2-sphere. A related picture is to think of the 3-sphere as being obtained from our familiar three-dimensional space $\mathbb{R}^3$ by adding one point "at infinity."

Now let us think about the familiar sphere S^2. This has trivial fundamental group, since any loop drawn on the sphere can be shrunk to a point. However, this does not mean that the topology of S^2 is trivial. It just means that in order to detect its interesting properties we need a different invariant. And it is possible to base such an invariant on the observation that even if loops can always be shrunk, there are other maps that cannot. Indeed, the sphere itself cannot be shrunk to a point. To say this more formally, the identity map from the sphere to itself is not homotopic to a map from the sphere to just one point.

This idea leads to the notion of higher-dimensional homotopy groups of a topological space X. The rough idea is to measure the number of "n-dimensional holes" in X, for any natural number n, by considering all the continuous mappings from the n-sphere to X. We want to see whether any of these spheres wrap around a hole in X. Once again, we consider two mappings from S^n to X to be equivalent if they are homotopic. And the elements of the nth homotopy group $\pi_n(X)$ are again defined to be the homotopy classes of these mappings.

Let f be a continuous map from $[0,1]$ to X with $f(0) = f(1) = x$. If we like we can turn the interval $[0,1]$ into the circle S^1 by "identifying" the points 0 and 1: then f becomes a map from S^1 to X, with one specified point in S^1 mapping to x. In order to be able to define a group operation for mappings from a higher-dimensional S^n, we similarly fix a point s in S^n and a base point x in X and look just at maps that send s to x.

Let A and B be two continuous mappings from S^n to X with this property. The "product" mapping $A \cdot B$ from S^n to X is defined as follows. First "pinch" the equator of S^n down to a point. When $n = 1$, the equator consists of just two points and the result is a figure eight. Similarly, for general n, we end up with two copies of S^n that touch each other, one made out of the northern hemisphere and one out of the southern hemisphere of the original unpinched copy of S^n. We now use the map A to map the bottom half into X and the map B to map the top half into X, with the equator mapping to the base point x. (For both halves, the pinched equator is playing the part of the point s.)

As in the one-dimensional case, this operation makes the set $\pi_n(X)$ into a group, and this group is the nth homotopy group of the space X. One can think of it as measuring how many "n-dimensional holes" a space has.

These groups are the beginning of "algebraic" topology: starting from any topological space, we construct an algebraic object, in this case a group. If two spaces are homeomorphic, then their fundamental groups (and higher homotopy groups) must be isomorphic. This is richer than the original idea of just measuring the *number* of holes, since a group contains more information than just a number.

Any continuous function from S^n into $\mathbb{R}^m$ can be continuously shrunk to a point in a straightforward way. This shows that all the higher homotopy groups of $\mathbb{R}^m$ are also trivial, which is a precise formulation of the vague idea that $\mathbb{R}^m$ has no holes.

Under certain circumstances one can show that two different topological spaces X and Y must have the same number of holes of all types. This is clearly true if X and Y are homeomorphic, but it is also true if X and Y are equivalent in a weaker sense, known as *homotopy equivalence*. Let X and Y be topological spaces and let f_0 and f_1 be continuous maps from X to Y. A homotopy from f_0 to f_1 is defined more or less as it was for spheres: it is a continuous family of continuous maps from X to Y that starts with f_0 and ends with f_1. As then, if such a homotopy exists, we say that f_0 and f_1 are homotopic. Next, a homotopy equivalence from a space X to a space Y is a continuous map $f : X \to Y$ such that there is another continuous map $g : Y \to X$ with the property that the composition $g \circ f : X \to X$ is homotopic to the identity map on X, and $f \circ g : Y \to Y$ is homotopic to the identity map on Y. (Notice that if we replaced the word "homotopic" with "equal," we would obtain the definition of a homeomorphism.) When there is a homotopy equivalence from X to Y, we say that X and Y are *homotopy equivalent*, and also that X and Y have the same *homotopy type*.

A good example is when X is the unit circle and Y is the plane with the origin removed. We have already observed that these have the same fundamental group, and commented that it was "for essentially the same reason." Now we can be more precise. Let $f : X \to Y$ be the map that takes (x, y) to (x, y) (where the first (x, y) belongs to the circle and the second to the plane). Let $g : Y \to X$ be the map that takes (u, v) to

$$\left(\frac{u}{\sqrt{u^2+v^2}}, \frac{v}{\sqrt{u^2+v^2}}\right).$$

(Note that u^2+v^2 is never zero because the origin is not contained in Y.) Then $g \circ f$ is easily seen to equal the identity on the unit circle, so it is certainly homotopic to the identity. As for $f \circ g$, it is given by the same formula as g itself. More geometrically, it takes the points on each radial line to the point where that line intersects the unit circle. It is not hard to show that this map is homotopic to the identity on Y. (The basic idea is to "shrink the radial lines down" to the points where they intersect the circle.)

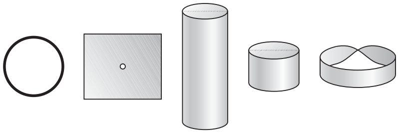

Figure 5 Some spaces that are homotopy equivalent to the circle.

Very roughly speaking, two spaces are homotopy equivalent if they have the same number of holes of all types. This is a more flexible notion of "having the same shape" than the notion of homeomorphism. For example, Euclidean spaces of different dimensions are not homeomorphic to each other, but they are all homotopy equivalent. Indeed, they are all homotopy equivalent to a point: such spaces are called *contractible*, and one thinks of them as the spaces that have no hole of any sort. The circle is not contractible, but it is homotopy equivalent to many other natural spaces: the plane $\mathbb{R}^2$ minus the origin (as we have seen), the cylinder $S^1 \times \mathbb{R}$, the compact cylinder $S^1 \times [0,1]$, and even the Möbius strip (see figure 5). Most invariants in algebraic topology (such as homotopy groups and cohomology groups) are the same for any two spaces that are homotopy equivalent. Thus, knowing that the fundamental group of the circle is isomorphic to the integers tells us that the same is true for the various homotopy equivalent spaces just mentioned. Roughly speaking, this says that all these spaces have "one basic one-dimensional hole."

3 Calculations of the Fundamental Group and Higher Homotopy Groups

To give some more feeling for the fundamental group, let us review what we already know and look at a few more examples. The fundamental group of the 2-sphere, or indeed of any higher-dimensional sphere, is trivial. The two-dimensional torus $S^1 \times S^1$ has fundamental group $\mathbb{Z}^2 = \mathbb{Z} \times \mathbb{Z}$. Thus, a loop in the torus determines two integers, which measure how many times it winds around in the meridian direction and how many in the longitudinal direction.

The fundamental group can also be non-Abelian; that is, we can have $ab \neq ba$ for some elements a and b of the fundamental group. The simplest example is a space X built out of two circles that meet at a single point (see figure 6). The fundamental group of X is the FREE GROUP [IV.10 §2] on two generators a and

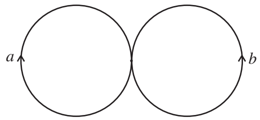

Figure 6 One-point union of two circles.

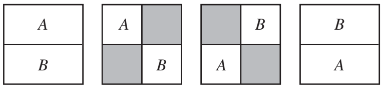

Figure 7 Proof that π_2 of any space is Abelian.

b. Roughly speaking, an element of this group is any product you can write down using the generators and their inverses, such as $abaab^{-1}a$, except that if a and a^{-1} or b and b^{-1} appear next to each other, you cancel them first. (So instead of $abb^{-1}bab^{-1}$ one would simply write $abab^{-1}$, for example.) The generators correspond to loops around each of the two circles. The free group is in a sense the most highly non-Abelian group. In particular, ab is not equal to ba, which in topological terms tells us that going around loop a and then loop b in the space X is not homotopic to the loop that goes around loop b and then loop a.

This space may seem somewhat artificial, but it is homotopy equivalent to the plane with two points removed, which appears in many contexts. More generally, the fundamental group of the plane with d points removed is the free group on d generators: this is a precise sense in which the fundamental group measures the number of holes.

In contrast with the fundamental group, the higher homotopy groups $\pi_n(X)$ are Abelian when n is at least 2. Figure 7 gives a "proof without words" in the case $n = 2$, the proof being the same for any larger n. In the figure, we view the 2-sphere as the square with its boundary identified to a point. So any elements A and B of $\pi_2(X)$ are represented by continuous maps of the square to X that map the boundary of the square to the base point x. The figure exhibits (several steps of) a homotopy from AB to BA, with the shaded regions and the boundary of the square all mapping to the base point x. The picture is reminiscent of the simplest nontrivial braid, in which one string is twisted around another; this is the beginning of a deep connection between algebraic topology and BRAID GROUPS [III.4].

The fundamental group is especially powerful in low dimensions. For example, every compact connected surface (or two-dimensional manifold) is homeomorphic to one of those on a standard list (see DIFFERENTIAL TOPOLOGY [IV.7 §2.3]), and we compute that all the manifolds on this list have different (nonisomorphic) fundamental groups. So, when you capture a closed surface in the wild, computing its fundamental group tells you exactly where it fits in the classification. Moreover, the geometric properties of the surface are closely tied to its fundamental group. The surfaces with a RIEMANNIAN METRIC [I.3 §6.10] of positive CURVATURE [III.13] (the 2-sphere and REAL PROJECTIVE PLANE [I.3 §6.7]) are exactly the surfaces with finite fundamental group; the surfaces with a metric of curvature zero (the torus and Klein bottle) are exactly the surfaces with a fundamental group that is infinite but "almost Abelian" (there is an Abelian subgroup of finite index); and the remaining surfaces, those that have a metric of negative curvature, have "highly non-Abelian" fundamental group, like the free group (see figure 8).

After more than a century of studying three-dimensional manifolds, we now know, thanks to the advances of Thurston and Perelman, that the picture is almost the same for these as it is for 2-manifolds: the fundamental group controls the geometric properties of the 3-manifold almost completely (see DIFFERENTIAL TOPOLOGY [IV.7 §2.4]). But this is completely untrue for 4-manifolds and in higher dimensions: there are many different *simply connected* manifolds, meaning manifolds with trivial fundamental group, and we need more invariants to be able to distinguish between them. (To begin with, the 4-sphere S^4 and the product $S^2 \times S^2$ are both simply connected. More generally, we can take the connected sum of any number of copies of $S^2 \times S^2$, obtained by removing 4-balls from these manifolds and identifying the boundary 3-spheres. These 4-manifolds are all simply connected, and yet no two of them are homeomorphic or even homotopy equivalent.)

An obvious way in which we might try to distinguish different spaces is to use *higher* homotopy groups, and indeed this works in simple cases. For example, π_2 of the connected sum of r copies of $S^2 \times S^2$ is isomorphic to $\mathbb{Z}^{2r}$. Also, we can show that the sphere S^n of any dimension is not contractible (although it is simply connected for $n \geqslant 2$) by computing that $\pi_n(S^n)$ is isomorphic to the integers (rather than the trivial group). Thus, each continuous map from the n-sphere to itself determines an integer, called the *degree* of the map,

Figure 8 A sphere, a torus, and a surface of genus 2.

which generalizes the notion of winding number for maps from the circle to itself.

In general, however, the homotopy groups are not a practical way of distinguishing one space from another, because they are amazingly hard to compute. A first hint of this was Hopf's 1931 discovery that $\pi_3(S^2)$ is isomorphic to the integers: it is clear that the 2-sphere has a two-dimensional hole, as measured by $\pi_2(S^2) \cong \mathbb{Z}$, but in what sense does it have a three-dimensional hole? This does not correspond to our naive view of what such a hole should be. The problem of computing the homotopy groups of spheres turns out to be one of the hardest in all of mathematics: some of what we know is shown in table 1, but despite massive efforts the homotopy groups $\pi_i(S^2)$, for example, are known only for $i \leqslant 64$. There are tantalizing patterns in these calculations, with a number-theoretic flavor, but it seems impossible to formulate a precise guess for the homotopy groups of spheres in general. And computing the homotopy groups for spaces more complex than spheres is even more complicated.

To get an idea of the difficulties involved, let us define the so-called *Hopf map* from S^3 to S^2, which turns out to represent a nonzero element of $\pi_3(S^2)$. There are in fact several equivalent definitions. One of them is to regard a point (x_1, x_2, x_3, x_4) in S^3 as a pair of complex numbers (z_1, z_2) such that $|z_1|^2 + |z_2|^2 = 1$. This we do by setting $z_1 = x_1 + \mathrm{i}x_2$ and $z_2 = x_3 + \mathrm{i}x_4$. We then map the pair (z_1, z_2) to the complex number z_1/z_2. This may not look like a map to S^2, but it is because z_2 may be zero, so in fact the image of the map is not $\mathbb{C}$ but the *Riemann sphere* $\mathbb{C} \cup \infty$, which can be identified with S^2 in a natural way.

Another way of defining the Hopf map is to regard points (x_1, x_2, x_3, x_4) in S^3 as unit quaternions. In the article on quaternions in this volume [III.76], it is shown that each unit quaternion can be associated with a rotation of the sphere. If we fix some point s in the sphere and map each unit quaternion to the image of s under the associated rotation, then we get a map from S^3 to S^2 that is homotopic to the map defined in the previous paragraph.

The Hopf map is an important construction, and will reappear more than once later in this article.

4 Homology Groups and the Cohomology Ring

Homotopy groups, then, can be rather mysterious and very hard to calculate. Fortunately, there is a different way to measure the number of holes in a topological space: homology and cohomology groups. The definitions are more subtle than the definition of homotopy groups, but the groups turn out to be easier to compute and are for this reason much more commonly used.

Recall that elements of the nth homotopy group $\pi_n(X)$ of a topological space X are represented by continuous maps from the n-sphere to X. Let X be a manifold, for simplicity. There are two key differences between homotopy groups and homology groups. The first is that the basic objects of homology are more general than n-dimensional spheres: *every* closed oriented n-dimensional submanifold A of X determines an element of the nth homology group of X, $H_n(X)$. This might make homology groups seem much bigger than homotopy groups, but that is not the case, because of the second major difference between homotopy and homology. As with homotopy, the elements of the homology groups are not the submanifolds themselves but equivalence classes of submanifolds, but the definition of the equivalence relation for homology makes it much easier for two of these submanifolds to be equivalent than it is for two spheres to be homotopic.

We shall not give a formal definition of homology, but here are some examples that convey some of its flavor. Let X be the plane with the origin removed and let A be a circle that goes around the origin. If we continuously deform this circle, we will obtain a new curve that is homotopic to the original circle, but with homology we can do more. For instance, we can start with a continuous deformation that causes two of its points to touch and turns it into a figure eight. One half of this figure eight will have to contain the origin, but we can leave

Table 1 The first few homotopy groups of spheres.

	S^1	S^2	S^3	S^4	S^5	S^6	S^7	S^8	S^9
π_1	$\mathbb{Z}$	0	0	0	0	0	0	0	0
π_2	0	$\mathbb{Z}$	0	0	0	0	0	0	0
π_3	0	$\mathbb{Z}$	$\mathbb{Z}$	0	0	0	0	0	0
π_4	0	$\mathbb{Z}/2$	$\mathbb{Z}/2$	$\mathbb{Z}$	0	0	0	0	0
π_5	0	$\mathbb{Z}/2$	$\mathbb{Z}/2$	$\mathbb{Z}/2$	$\mathbb{Z}$	0	0	0	0
π_6	0	$\mathbb{Z}/4 \times \mathbb{Z}/3$	$\mathbb{Z}/4 \times \mathbb{Z}/3$	$\mathbb{Z}/2$	$\mathbb{Z}/2$	$\mathbb{Z}$	0	0	0
π_7	0	$\mathbb{Z}/2$	$\mathbb{Z}/2$	$\mathbb{Z} \times \mathbb{Z}/4 \times \mathbb{Z}/3$	$\mathbb{Z}/2$	$\mathbb{Z}/2$	$\mathbb{Z}$	0	0
π_8	0	$\mathbb{Z}/2$	$\mathbb{Z}/2$	$\mathbb{Z}/2 \times \mathbb{Z}/2$	$\mathbb{Z}/8 \times \mathbb{Z}/3$	$\mathbb{Z}/2$	$\mathbb{Z}/2$	$\mathbb{Z}$	0
π_9	0	$\mathbb{Z}/3$	$\mathbb{Z}/3$	$\mathbb{Z}/2 \times \mathbb{Z}/2$	$\mathbb{Z}/2$	$\mathbb{Z}/8 \times \mathbb{Z}/3$	$\mathbb{Z}/2$	$\mathbb{Z}/2$	$\mathbb{Z}$
π_{10}	0	$\mathbb{Z}/3 \times \mathbb{Z}/5$	$\mathbb{Z}/3 \times \mathbb{Z}/5$	$\mathbb{Z}/8 \times \mathbb{Z}/3 \times \mathbb{Z}/3$	$\mathbb{Z}/2$	0	$\mathbb{Z}/8 \times \mathbb{Z}/3$	$\mathbb{Z}/2$	$\mathbb{Z}/2$

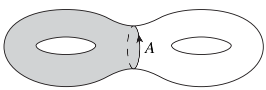

Figure 9 The circle A represents zero in the homology of the surface.

that still and slide the other part away. The result is then two closed curves, with the origin inside one and outside the other. This pair of curves, which together form a 1-manifold with two components, is equivalent to the original circle. It can be seen as a continuous deformation of a more general kind.

A second example shows how natural it is to include other manifolds in the definition of homology. This time let X be $\mathbb{R}^3$ with a circle removed, and let A be a sphere that contains the circle in its interior. Suppose that the circle is in the XY-plane and that both it and the sphere A are centered at the origin. Then we can pinch the top and bottom of A toward the origin until they just touch. If we do so, then we obtain a shape that looks like a torus, except that the hole in the middle has been shrunk to zero. But we can open up this hole with the help of a further continuous deformation and obtain a genuine torus, which is a "tube" around the original circle. From the point of view of homology, this torus is equivalent to the sphere A.

A more general rule is that if X is a manifold and B is a compact oriented $(n+1)$-dimensional submanifold of X with a boundary, then this boundary ∂B will be equivalent to zero (which is the same as saying that $[\partial B] = 0$ in $H_n(X)$): see figure 9.

The group operation is easy to define: if A and B are two disjoint submanifolds of X, giving rise to homology classes $[A]$ and $[B]$, then $[A] + [B]$ is the homology class of $[A \cup B]$. (More generally, the definition of homology allows us to add up any collection of submanifolds, whether or not they overlap.) Here are some simple examples of homology groups, which, unlike the fundamental group, are always Abelian. The homology groups of a sphere, $H_i(S^n)$, are isomorphic to the integers $\mathbb{Z}$ for $i = 0$ and for $i = n$, and 0 otherwise. This contrasts with the complicated homotopy groups of the sphere, and better reflects the naive idea that the n-sphere has one n-dimensional hole and no other holes. Note that the fundamental group of the circle, the group of integers, is the same as its first homology group. More generally, for any path-connected space, the first homology group is always the "Abelianization" of the fundamental group (which is formally defined to be its largest Abelian quotient). For example, the fundamental group of the plane with two points removed is the free group on two generators, while the first homology group is the free *Abelian* group on two generators, or $\mathbb{Z}^2$.

The homology groups of the two-dimensional torus $H_i(S^1 \times S^1)$ are isomorphic to $\mathbb{Z}$ for $i = 0$, to $\mathbb{Z}^2$ for $i = 1$, and to $\mathbb{Z}$ for $i = 2$. All of this has geometric meaning. The zeroth homology group of any space is isomorphic to $\mathbb{Z}^r$ for a space X with r connected components. So the fact that the zeroth homology group of the torus is isomorphic to $\mathbb{Z}$ means that the torus is connected. Any closed loop in the torus determines an element of the first homology group $\mathbb{Z}^2$, which measures how many times the loop winds around the meridian and longitudinal directions of the torus. And finally, the homology of the torus in dimension 2 is isomorphic to $\mathbb{Z}$ because the torus is a closed orientable manifold. That tells us that the whole torus defines an element of the second homology group of the torus, which is in fact a generator of that group. By contrast, the homotopy group

$\pi_2(S^1 \times S^1)$ is the trivial group: there are no interesting maps from the 2-sphere to the 2-torus, but homology shows that there are interesting maps from other closed 2-manifolds to the 2-torus.

As we have mentioned, calculating homology groups is much easier than calculating homotopy groups. The main reason for this is the existence of results that tell you the homology groups of a space that is built up from smaller pieces in terms of the homology groups of those pieces and their intersections. Another important property of homology groups is that they are "functorial" in the sense that a continuous map f from a space X to a space Y leads in a natural way to a homomorphism f_* from $H_i(X)$ to $H_i(Y)$ for each i: $f_*([A])$ is defined to be $[f(A)]$. In other words, $f_*([A])$ is the equivalence class of the image of A under f.

We can define the closely related idea of "cohomology" simply by a different numbering. Let X be a closed oriented n-dimensional manifold. Then we define the ith *cohomology group* $H^i(X)$ to be the homology group $H_{n-i}(X)$. Thus, one way to write down a cohomology class (an element of $H^i(X)$) is by choosing a closed oriented submanifold S of codimension i in X. (This means that the dimension of S is $n - i$.) We write $[S]$ for the corresponding cohomology class.

For more general spaces than manifolds, cohomology is not just a simple renumbering of homology. Informally, if X is a topological space, then we think of an element of $H^i(X)$ as being represented by a codimension-i subspace of X that can move around freely in X. For example, suppose that f is a continuous map from X to an i-dimensional manifold. If X is a manifold and f is sufficiently "well-behaved," then the inverse image of a "typical" point in the manifold will be an i-codimensional submanifold of X, and as we move the point about, this submanifold will vary continuously, and will do so in a way that is similar to the way that a circle became two circles and a sphere became a torus earlier. If X is a more general topological space, the map f still determines a cohomology class in $H^i(X)$, which we think of as being represented by the inverse image in X of any point in the manifold.

However, even when X is an oriented n-dimensional manifold, cohomology has distinct advantages over homology. This may seem odd, since the cohomology groups are the homology groups with different names. However, this renumbering allows us to give very useful extra algebraic structure to the cohomology groups of X: not only can we add cohomology classes, we can multiply them as well. Furthermore, we can do so in such a way that, taken together, the cohomology groups of X form a RING [III.81 §1]. (Of course, we could do this for the homology groups, but the cohomology groups form a so-called *graded* ring. In particular, if $[A] \in H^i(X)$ and $[B] \in H^j(X)$, then $[A] \cdot [B] \in H^{i+j}(X)$.)

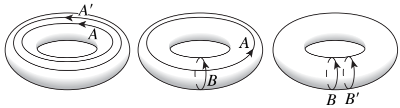

Figure 10 $A^2 = A \cdot A' = 0$, $A \cdot B =$ [point], and $B^2 = B \cdot B' = 0$.

The multiplication of cohomology classes has a rich geometric meaning, especially on manifolds: it is given by the *intersection* of two submanifolds. This generalizes our discussion of intersection numbers in section 1: there we considered zero-dimensional intersections of submanifolds, whereas we are now considering (cohomology classes of) higher-dimensional intersections. To be precise, let S and T be closed oriented submanifolds of X, of codimension i and j, respectively. By moving S slightly (which does not change its class in $H^i(X)$) we can assume that S and T intersect transversely, which implies that the intersection of S and T is a smooth submanifold of codimension $i + j$ in X. Then the product of the cohomology classes $[S]$ and $[T]$ is simply the cohomology class of the intersection $S \cap T$ in $H^{i+j}(X)$. (In addition, the submanifold $S \cap T$ inherits an orientation from S, T, and X: this is needed to define the associated cohomology class.)

As a result, to compute the cohomology ring of a manifold, it is enough to specify a basis for the cohomology groups (which, as we have already discussed, are relatively easy to determine) using some submanifolds and to see how these submanifolds intersect. For example, we can compute the cohomology ring of the 2-torus as shown in figure 10. For another example, it is not hard to show that the cohomology of the COMPLEX PROJECTIVE PLANE [III.72] $\mathbb{CP}^2$ has a basis given by three basic submanifolds: a point, which belongs to $H^4(\mathbb{CP}^2)$ because it is a submanifold of codimension 4; a complex projective line $\mathbb{CP}^1 = S^2$, which belongs to $H^2(\mathbb{CP}^2)$; and the whole manifold $\mathbb{CP}^2$, which is in $H^0(\mathbb{CP}^2)$ and represents the identity element 1 of the cohomology ring. The product in the cohomology ring is described by saying that $[\mathbb{CP}^1][\mathbb{CP}^1] = [\text{point}]$, because any two distinct lines $\mathbb{CP}^1$ in the plane meet transversely in a single point.

This calculation of the cohomology ring of the complex projective plane, although very simple, has several strong consequences. First of all, it implies Bézout's theorem on intersections of complex algebraic curves (see ALGEBRAIC GEOMETRY [IV.4 §6]). An algebraic curve of degree d in $\mathbb{CP}^2$ represents d times the class of a line $\mathbb{CP}^1$ in $H^2(\mathbb{CP}^2)$. Therefore, if two algebraic curves D and E of degrees d and e meet transversely, then the cohomology class $[D \cap E]$ equals

$$[D] \cdot [E] = (d[\mathbb{CP}^1])(e[\mathbb{CP}^1]) = de[\text{point}].$$

For complex submanifolds of a complex manifold, intersection numbers are always $+1$, not -1, and so this means that D and E meet in exactly de points.

We can also use the computation of the cohomology ring of $\mathbb{CP}^2$ to prove something about the homotopy groups of spheres. It turns out that $\mathbb{CP}^2$ can be constructed as the union of the 2-sphere and the closed four-dimensional ball, with each point of the boundary S^3 of the ball identified with a point in S^2 by the Hopf map, which was defined in the previous section.

A constant map from one space to another, or a map homotopic to a constant map, gives rise to the zero homomorphism between the homology groups H_i, at least when $i > 0$. The Hopf map $f : S^3 \to S^2$ also induces the zero homomorphism because the nonzero homology groups of S^3 and S^2 are in different dimensions. Nonetheless, we will show that f is not homotopic to the constant map. If it were, then the space $\mathbb{CP}^2$ obtained by attaching a 4-ball to the 2-sphere using the map f would be homotopy equivalent to the space obtained by attaching a 4-ball to the 2-sphere using a constant map. The latter space Y is the union of S^2 and S^4 identified at one point. But in fact Y is not homotopy equivalent to the complex projective plane, because their cohomology rings are not isomorphic. In particular, the product of any element of $H^2(Y)$ with itself is zero, unlike what happens in $\mathbb{CP}^2$ where $[\mathbb{CP}^1][\mathbb{CP}^1] = [\text{point}]$. Therefore f is nonzero in $\pi_3(S^2)$. A more careful version of this argument shows that $\pi_3(S^2)$ is isomorphic to the integers, and the Hopf map $f : S^3 \to S^2$ is a generator of this group.

This argument shows some of the rich relations between all the basic concepts of algebraic topology: homotopy groups, cohomology rings, manifolds, and so on. To conclude, here is a way to visualize the nontriviality of the Hopf map $f : S^3 \to S^2$. Look at the subset of S^3 that maps to any given point of the 2-sphere. These inverse images are all circles in the 3-sphere. To draw them, we can use the fact that S^3 minus a point is homeomorphic to $\mathbb{R}^3$; so these inverse images form a family of disjoint circles that fills up three-dimensional space, with one circle being drawn as a line (the circle through the point we removed from S^3). The striking feature of this picture is that any two of this huge family of circles have linking number 1 with each other: there is no way to pull any two of them apart (see figure 11).

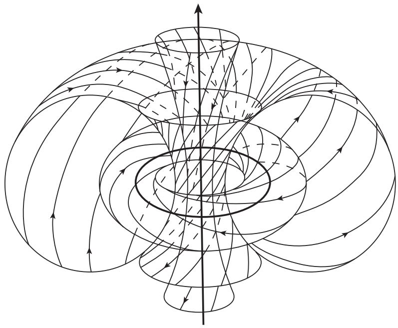

Figure 11 Fibers of the Hopf map.

5 Vector Bundles and Characteristic Classes

We now introduce another major topological idea: fiber bundles. If E and B are topological spaces, x is a point in B, and $p : E \to B$ is a continuous map, then the *fiber* of p *over* x is the subspace of E that maps to x. We say that p is a *fiber bundle*, with fiber F, if every fiber of p is homeomorphic to the same space F. We call B the *base space* and E the *total space*. For example, any product space $B \times F$ is a fiber bundle over B, called the trivial F-bundle over B. (The continuous map in this case is the map that takes (x, y) to x.) But there are many nontrivial fiber bundles. For example, the Möbius strip is a fiber bundle over the circle with fiber a closed interval. This example helps to explain the old name "twisted product" for fiber bundles. Another example: the Hopf map makes the 3-sphere the total space of a circle bundle over the 2-sphere.

Fiber bundles are a fundamental way to build up complicated spaces from simple pieces. We will focus on the most important special case: vector bundles. A *vector bundle* over a space B is a fiber bundle $p : E \to B$ whose fibers are all real vector spaces of some dimension n.

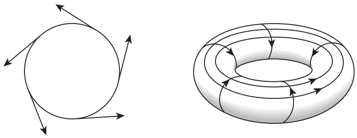

Figure 12 Trivializations of the tangent bundle for the circle and the torus.

This dimension is called the *rank* of the vector bundle. A *line bundle* means a vector bundle of rank 1; for example, we can view the Möbius strip (not including its boundary) as a line bundle over the circle S^1. It is a *nontrivial* line bundle; that is, it is not isomorphic to the trivial line bundle $S^1 \times \mathbb{R}$. (There are many ways of constructing it: one is to take the strip $\{(x, y) : 0 \leqslant x \leqslant 1\}$ and identify each point $(0, y)$ with the point $(1, -y)$. The base space of this line bundle is the set of all points $(x, 0)$, which is a circle since $(0, 0)$ and $(1, 0)$ have been identified.)

If M is a smooth manifold of dimension n, its *tangent bundle* $TM \to M$ is a vector bundle of rank n. We can easily define this bundle by considering M as a submanifold of some Euclidean space $\mathbb{R}^N$. (Every smooth manifold can be embedded into Euclidean space.) Then TM is the subspace of $M \times \mathbb{R}^N$ of pairs (x, v) such that the vector v is tangent to M at the point x; the map $TM \to M$ sends a pair (x, v) to the point x. The fiber over x then has the form of the set of all pairs (x, v) with v belonging to an affine subspace of $\mathbb{R}^N$ of dimension equal to that of M. For any fiber bundle, a *section* means a continuous map from the base space B to the total space E that maps each point x in B to some point in the fiber over x. A section of the tangent bundle of a manifold is called a *vector field*. We can draw a vector field on a given manifold by putting an arrow (possibly of zero length) at every point of the manifold.

In order to classify smooth manifolds, it is important to study their tangent bundles, and in particular to see whether they are trivial or not. Some manifolds, like the circle S^1 and the torus $S^1 \times S^1$, do have trivial tangent bundle. The tangent bundle of an n-manifold M is trivial if and only if we can find n vector fields that are linearly independent at every point of M. So we can prove that the tangent bundle is trivial just by writing down such vector fields; see figure 12 for the circle or the torus. But how can we show that the tangent bundle of a given manifold is nontrivial?

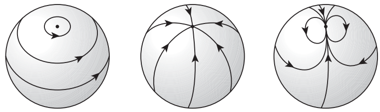

Figure 13 The hairy ball theorem.

One way is to use intersection numbers. Let M be a closed oriented n-manifold. We can identify M with the image of the "zero-section" inside the tangent bundle TM, the section that assigns to every point of M the zero vector at that point. Since the dimension of TM is precisely double that of M, the discussion of intersection numbers in section 1 gives a well-defined integer $M^2 = M \cdot M$, the self-intersection number of M inside TM; this is called the *Euler characteristic* $\chi(M)$. By the definition of intersection numbers, for any vector field v on M that meets the zero-section transversely, the Euler characteristic of M is equal to the number of zeros of v, counted with signs.

As a result, if the Euler characteristic of M is not zero, then every vector field on M must meet the zero-section; in other words, every vector field on M must equal zero somewhere. The simplest example occurs when M is the 2-sphere S^2. We can easily write down a vector field (for example, the one pointing toward the east along circles of latitude, which vanishes at the north and south poles) whose intersection number with the zero-section is 2. Therefore the 2-sphere has Euler characteristic 2, and so every vector field on the 2-sphere must vanish somewhere. This is a famous theorem of topology known as the "hairy ball theorem": it is impossible to comb the hair on a coconut (see figure 13).

This is the beginning of the theory of *characteristic classes*, which measure how nontrivial a given vector bundle is. There is no need to restrict ourselves to the tangent bundle of a manifold. For any oriented vector bundle E of rank n on a topological space X, we can define a cohomology class $\chi(E)$ in $H^n(X)$, the *Euler class*, which vanishes if the bundle is trivial. Intuitively, the Euler class of E is the cohomology class represented by the zero set of a general section of E, which (for example, if X is a manifold) should be a codimension-n submanifold of X, since X has codimension n in E. If X is a closed oriented n-manifold, then the Euler class of the tangent bundle in $H^n(X) = \mathbb{Z}$ is the Euler characteristic of X.

One of the inspirations for the theory of characteristic classes was the Gauss–Bonnet theorem, generalized to all dimensions in the 1940s. The theorem expresses the Euler characteristic of a closed manifold with a Riemannian metric as the integral over the manifold of a certain curvature function. More broadly, a central goal of differential geometry is to understand how the geometric properties of a Riemannian manifold such as its curvature are related to the topology of the manifold.

The characteristic classes for *complex* vector bundles (that is, bundles where the fibers are complex vector spaces) turn out to be particularly convenient: indeed, real vector bundles are often studied by constructing the associated complex vector bundle. If E is a complex vector bundle of rank n over a topological space X, the *Chern classes* of E are a sequence $c_1(E), \dots, c_n(E)$ of cohomology classes on X, with $c_i(E)$ belonging to $H^{2i}(X)$, which all vanish if the bundle is trivial. The top Chern class, $c_n(E)$, is simply the Euler class of E: thus, it is the first obstruction to finding a section of E that is everywhere nonzero. The more general Chern classes have a similar interpretation. For any $1 \leqslant j \leqslant n$, choose j general sections of E. The subset of X over which these sections become linearly dependent will have codimension $2(n+1-j)$ (assuming, for example, that X is a manifold). The Chern class $c_{n+1-j}(E)$ is precisely the cohomology class of this subset. Thus the Chern classes measure in a natural way the failure of a given complex vector bundle to be trivial. The *Pontryagin classes* of a real vector bundle are defined to be the Chern classes of the associated complex vector bundle.

A triumph of differential topology is Sullivan's 1977 theorem that there are only finitely many smooth closed simply connected manifolds of dimension at least 5 with any given homotopy type and given Pontryagin classes of the tangent bundle. This statement fails badly in dimension 4, as Donaldson discovered in the 1980s (see DIFFERENTIAL TOPOLOGY [IV.7 §2.5]).

6 *K*-Theory and Generalized Cohomology Theories

The effectiveness of vector bundles in geometry led to a new way of measuring the "holes" in a topological space X: looking at how many different vector bundles over X there are. This idea gives a simple way to define a cohomology-like ring associated to any space, known as K-theory (after the German word "Klasse," since the theory involves equivalence classes of vector bundles). It turns out that K-theory gives a very useful new angle by which to look at topological spaces. Some problems that could be solved only with enormous effort using ordinary cohomology became easy with K-theory. The idea was created in algebraic geometry by Grothendieck in the 1950s and then brought into topology by Atiyah and Hirzebruch in the 1960s.

The definition of K-theory can be given in a few lines. For a topological space X, we define an Abelian group $K^0(X)$, the *K-theory* of X, whose elements can be written as formal differences $[E] - [F]$, where E and F are any two complex vector bundles over X. The only relations we impose in this group are that $[E \oplus F] = [E] + [F]$ for any two vector bundles E and F over X. Here $E \oplus F$ denotes the *direct sum* of the two bundles; if E_x and F_x denote the fibers at a given point x in X, the fiber of $E \oplus F$ at x is simply $E_x \times F_x$.

This simple definition leads to a rich theory. First of all, the Abelian group $K^0(X)$ is in fact a ring: we multiply two vector bundles on X by forming the TENSOR PRODUCT [III.89]. In this respect, K-theory behaves like ordinary cohomology. The analogy suggests that the group $K^0(X)$ should form part of a whole sequence of Abelian groups $K^i(X)$, for integers i, and indeed these groups can be defined. In particular, $K^{-i}(X)$ can be defined as the subgroup of those elements of $K^0(S^i \times X)$ whose restriction to $K^0(\text{point} \times X)$ is zero.

Then a miracle occurs: the groups $K^i(X)$ turn out to be *periodic* of order 2: $K^i(X) = K^{i+2}(X)$ for all integers i. This is a famous phenomenon known as *Bott periodicity*. So there are really only two different K-groups attached to any topological space: $K^0(X)$ and $K^1(X)$.

This may suggest that K-theory contains less information than ordinary cohomology, but that is not so. Neither K-theory nor ordinary cohomology determines the other, although there are strong relations between them. Each brings different aspects of the shape of a space to the fore. Ordinary cohomology, with its numbering, shows fairly directly the way a space is built up from pieces of different dimensions. K-theory, having only two different groups, looks cruder at first (and is often easier to compute as a result). But geometric problems involving vector bundles often involve information that is subtle and hard to extract from ordinary cohomology, whereas this information is brought to the surface by K-theory.

The basic relation between K-theory and ordinary cohomology is that the group $K^0(X)$ constructed from the vector bundles on X "knows" something about all the even-dimensional cohomology groups of X. To be precise, the rank of the Abelian group $K^0(X)$ is the sum of the ranks of all the even-dimensional cohomology groups $H^{2i}(X)$. This connection comes from associating with a given vector bundle on X its Chern classes. The odd K-group $K^1(X)$ is related in the same way to the odd-dimensional ordinary cohomology.

As we have already hinted, the precise group $K^0(X)$, as opposed to just its rank, is better adapted to some geometric problems than ordinary cohomology. This phenomenon shows the power of looking at geometric problems in terms of vector bundles, and thus ultimately in terms of linear algebra. Among the classic applications of K-theory is the proof, by Bott, Kervaire, and Milnor, that the 0-sphere, the 1-sphere, the 3-sphere, and the 7-sphere are the only spheres whose tangent bundles are trivial. This has a deep algebraic consequence, in the spirit of the fundamental theorem of algebra: the only dimensions in which there can be a real division algebra (not assumed to be commutative or even associative) are 1, 2, 4, and 8. There are indeed division algebras of all four types: the real numbers, complex numbers, quaternions, and octonions (see QUATERNIONS, OCTONIONS, AND NORMED DIVISION ALGEBRAS [III.76]).

Let us see why the existence of a real division algebra of dimension n implies that the $(n-1)$-sphere has trivial tangent bundle. In fact, let us merely assume that we have a finite-dimensional real vector space V with a bilinear map $V \times V \to V$, which we call the "product," such that if x and y are vectors in V with $xy = 0$, then either $x = 0$ or $y = 0$. For convenience, let us also assume that there is an identity element 1 in V, so $1 \cdot x = x \cdot 1 = x$ for all $x \in V$; one can, however, do without this assumption. If V has dimension n, then we can identify V with $\mathbb{R}^n$. Then, for each point x in the sphere S^{n-1}, left multiplication by x gives a linear isomorphism from $\mathbb{R}^n$ to itself. By scaling the output to have length 1, left multiplication by x gives a diffeomorphism from S^{n-1} to itself which maps the point 1 (scaled to have length 1) to x. Taking the derivative of this diffeomorphism at the point 1 gives a linear isomorphism from the tangent space of the sphere at the point 1 to the tangent space at x. Since the point x on the sphere is arbitrary, a choice of basis for the tangent space of the sphere at the point 1 determines a trivialization of the whole tangent bundle of the $(n-1)$-sphere.

Among other applications, K-theory provides the best "explanation" for the low-dimensional homotopy groups of spheres, and in particular for the number-theoretic patterns that are seen there. Notably, denominators of Bernoulli numbers appear among those groups (such as $\pi_{n+3}(S^n) \cong \mathbb{Z}/24$ for n at least 5), and this pattern was explained using K-theory by Milnor, Kervaire, and Adams.

THE ATIYAH–SINGER INDEX THEOREM [V.2] provides a deep analysis of linear differential equations on closed manifolds using K-theory. The theorem has made K-theory important for gauge theories and string theories in physics. K-theory can also be defined for noncommutative rings, and is in fact the central concept in "noncommutative geometry" (see OPERATOR ALGEBRAS [IV.15 §5]).

The success of K-theory led to a search for other "generalized cohomology theories." There is one other theory that stands out for its power: *complex cobordism.* The definition is very geometric: the complex cobordism groups of a manifold M are generated by mappings of manifolds (with a complex structure on the tangent bundle) into M. The relations say that any manifold counts as zero if it is the boundary of some other manifold. For example, the union of two circles would count as zero if you could find a cylinder whose ends were those circles.

It turns out that complex cobordism is much richer than either K-theory or ordinary cohomology. It sees far into the structure of a topological space, but at the cost of being difficult to compute. Over the past thirty years, a whole series of cohomology theories, such as elliptic cohomology and Morava K-theories, have been constructed as "simplifications" of complex cobordism: there is a constant tension in topology between invariants that carry a lot of information and invariants that are easy to compute. In one direction, complex cobordism and its variants provide the most powerful tool for the computation and understanding of the homotopy groups of spheres. Beyond the range where Bernoulli numbers appear, we see deeper number theory such as MODULAR FORMS [III.59]. In another direction, the geometric definition of complex cobordism makes it useful in algebraic geometry.

7 Conclusion

The line of thought introduced by pioneering topologists like RIEMANN [VI.49] is simple but powerful. Try to translate any problem, even a purely algebraic one, into geometric terms. Then ignore the details of the geometry and study the underlying shape or topology of the problem. Finally, go back to the original problem and see how much has been gained. The fundamental topological ideas such as cohomology are used throughout mathematics, from number theory to string theory.

Further Reading

From the definition of topological spaces to the fundamental group and a little beyond, I like M. A. Armstrong's *Basic Topology* (Springer, New York, 1983). The current standard graduate textbook is A. Hatcher's *Algebraic Topology* (Cambridge University Press, Cambridge, 2002). Two of the great topologists, Bott and Milnor, are also brilliant writers. Every young topologist should read R. Bott and L. Tu's *Differential Forms in Algebraic Topology* (Springer, New York, 1982), J. Milnor's *Morse Theory* (Princeton University Press, Princeton, NJ, 1963), and J. Milnor and J. Stasheff's *Characteristic Classes* (Princeton University Press, Princeton, NJ, 1974).

IV.7 Differential Topology

C. H. Taubes

1 Smooth Manifolds

This article is about classifying certain objects called smooth manifolds, so I need to start by telling you what they are. A good example to keep in mind is the surface of a smooth ball. If you look at a small portion of it from very close up, then it looks like a portion of a flat plane, but of course it differs in a radical way from a flat plane on larger distance scales. This is a general phenomenon: a smooth manifold can be very convoluted, but must be quite regular in close-up. This "local regularity" is the condition that each point in a manifold belongs to a neighborhood that looks like a portion of standard Euclidean space in some dimension. If the dimension in question is d for every point of the manifold, then the manifold itself is said to have dimension d. A schematic of this is shown in figure 1.

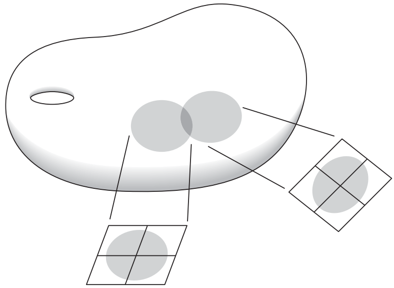

Figure 1 Small portions of a manifold resemble regions in a Euclidean space.

What does it mean to say that a neighborhood "looks like a portion of standard Euclidean space"? It means that there is a "nice" one-to-one map ϕ from the neighborhood into $\mathbb{R}^d$ (with its usual notion of distance). One can think of ϕ as "identifying" points in the neighborhood with points in $\mathbb{R}^d$: that is, x is identified with $\phi(x)$. If we do this, then the function ϕ is called a *coordinate chart* of the neighborhood, and any chosen basis for the linear functions on the Euclidean space is called a *coordinate system*. The reason for this is that ϕ allows us to use the coordinates in $\mathbb{R}^d$ to label points in the neighborhood: if x belongs to the neighborhood, then one can label it with the coordinates of $\phi(x)$. For example, Europe is part of the surface of a sphere. A typical map of Europe identifies each point in Europe with a point in flat, two-dimensional Euclidean space, that is, a square grid labeled with latitude and longitude. These two numbers give us a coordinate system for the map, which can also be transferred to a coordinate system for Europe itself.

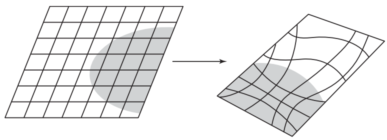

Figure 2 A transition function from a rectangular grid to a distorted rectangular grid.

Now, here is a straightforward but central observation. Suppose that M and N are two neighborhoods that intersect, and suppose that functions $\phi : M \to \mathbb{R}^d$ and $\psi : N \to \mathbb{R}^d$ are used to give them each a coordinate chart. Then the intersection $M \cap N$ is given *two* coordinate charts, and this gives us an identification between the open regions $\phi(M \cap N)$ and $\psi(M \cap N)$ of $\mathbb{R}^d$: given a point x in the first region, the corresponding point in the second is $\psi(\phi^{-1}(x))$. This composition of maps is called a *transition function*, and it tells you how the coordinates from one of the charts on the intersecting region relate to those of the other. The transition function is a HOMEOMORPHISM [III.90] between the regions $\phi(M \cap N)$ and $\psi(M \cap N)$.