# Some problems in topology

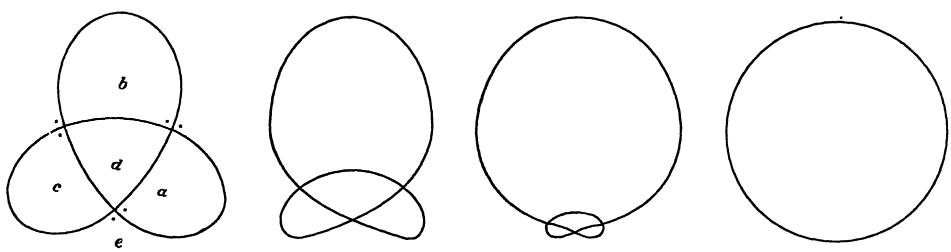
By J. W. Alexander, Princeton, N. J.

Broadly speaking, we may say that *analysis situs*, or *topology*, deals with the properties of geometrical figures that remain invariant when the figures are subjected to arbitrary continuous transformations. There are, however, several distinct kinds of analysis situs, because there are several distinct ways of interpreting the physical notion of continuity in mathematical language. For example, there is what we call *point theoretical* analysis situs which is different in spirit as well as in content from the sort of analysis situs originally proposed by Leibnitz. This branch of the science is essentially an outgrowth of function theory, whereas what Leibnitz had in mind was a new and independent type of mathematics, especially designed to avoid the complications of function theory and to deal directly with the purely qualitative aspects of geometrical problems. No doubt *combinatorial* analysis situs is more nearly a development of Leibnitz's original idea.

The vogue for point theoretical analysis situs seems to be due, in large part, to the predominating influence of analysis on mathematics in general. Nowadays, we tend, almost automatically, to identify physical space with the space of three real variables and to interpret physical continuity in the classical function theoretical manner. But the space of three real variables is not the only possible mathematical model of physical space, nor is it a perfectly satisfactory model for dealing with certain types of problems. Whenever we attack a topological problem by analytic methods it almost invariably happens that to the intrinsic difficulties of the problem, which we can hardly hope to avoid, there are added certain extraneous difficulties in no way connected with the problem itself, but apparently associated with the particular type of machinery used in dealing with it. Consider, for example, our old friend, the problem of the knotted string. In physical terms, what we have is this. We take an ordinary piece of string, tangle it up in an arbitrary manner, and seal its two ends together. We then ask ourselves whether we have really tied a knot in the string or whether the twists can all be disentangled without breaking the seal that holds the ends together. Now, to solve this problem it is evidently immaterial to know the exact shape of the string. We merely need to know something about the way in which the various branches lace over and under one another, all of which can be described with sufficient accuracy by a rough picture (fig. 1a), with some device (such as the system of dots shown in the figure), to distinguish the „upper" from the „lower" branch at each apparent crossing point. Moreover, all the really essential features of the knot picture may in turn be described by a table, or matrix: in this case, the matrix

| $a$ | $b$ | $c$ | $d$ | $e$ |
|---|---|---|---|---|
| 1 | −1 | 0 | 1 | −1 |
| 0 | 1 | −1 | 1 | −1 |
| −1 | 0 | 1 | 1 | −1, |

where each row is associated with an apparent crossing point, and where the signs of the terms in the row determine the ,,upper" and ,,under" branches at the point in question. In other words, the problem of the knotted string is essentially a combinatorial one, reducible to a problem in the manipulation of matrices. We might, however, be tempted to attack the problem analytically; for example, we might represent the string by a simple closed curve and ask ourselves whether or not the curve could be deformed into a circle without acquiring a multiple point at any stage of the process. Unfortunately, this would, at once, introduce all the complications associated with the mathematical notion of a simple, closed curve, some of which are so clearly brought into evidence by the well known curve of Antoine. Worse than this, unless special precautions were taken, the analytic procedure would lead to an incorrect mathematical formulation of the problem. For it is easy to see that every sufficiently smooth simple, closed curve can be deformed continuously into a circle without ever acquiring a multiple point, even though the curve may be one which




a b c d

Fig. 1

ought to be regarded as knotted. In figure 1, for example, we have a series of self-explanatory diagrams showing a trefoil knot in the process of being deformed into a circle. It will be noticed that the curve acquires no multiple point even at the moment when it becomes a circle. The moral of all this is that it is often better to avoid the artificial subtleties of analysis by using a more simple minded type of machinery.

In view of the above remarks, I am going to speak about three different kinds of analysis situs and to bring out, as far as I can, their relation to one another.

At one extreme, there will be *point theoretical* analysis situs, in which a space is regarded as a set of points, in which the structure of the space is expressed in terms of the notion of limit point, and in which a continuous transformation is merely a transformation preserving limit points. In this theory, two spaces are *homeomorphic* if there exists a one-one bi-continuous point transformation between them preserving limit points. At the other extreme will be *combinatorial* analysis situs, in which a space is not regarded as a set of points at all, but as something which may be cut up into a mosaic, or *complex*, of blocks called *cells*. These cells are not sets of points but primitive undefined entities. In the applications to physical geometry they will ordinarily represent „chunks“ of space. Two complexes are *congruent* if there exists a one-one correspondence between the cells of the first and the cells of the second such that the correspondence preserves incidence relations between pairs of cells. The notion of a continuous transformation between two complexes is arrived at in some such way as the following. Certain operations are defined which allow us, according to specified rules, of course, to replace a cell of a complex by a cluster of inter-related cells or a cluster of inter-related cells by a single cell[1]).

These operations represent the abstract formulation of the physical operation of cutting up a portion of space into smaller pieces or of amalgamating a number of smaller pieces to form a larger one. Two complexes $A$ and $B$ are *equivalent* if we can transform the complex $A$ into a complex $A'$ congruent to $B$ by a sequence of operations of the type allowed. The notion of equivalence plays a similar role here to the notion of homeomorphism in point theoretical analysis situs. Finally, there is a third type of analysis situs, which I shall call *flat* analysis situs, and which will serve as a sort of connecting link between the other two types. Here again, we shall be dealing with complexes of cells, but the cells, instead of being undefined abstractions, will be ordinary simplexes in the sense of analytic geometry. An allowable transformation of a complex will be one in which each simplex is cut up in an arbitrary manner into smaller ones, and two complexes $A$ and $B$ will be flat homeomorphic if they can be cut up into two complexes $A'$ and $B'$ such that these last two are congruent. The significance of flat homeomorphism in terms of general continuous transformations is obvious. If the complexes $A$ and $B$ are flat homeomorphic, the correspondence between the cells of $A'$ and $B'$ determines a one-one transformation between the points of $A$ and of $B$ which is not only continuous but *linear in patches* (more specifically, linear over each cell of $A'$ and $B'$). It is easy to verify that, conversely, if a one-one transformation, linear in patches, exists between the points of $A$ and of $B$ then $A$ and $B$ are flat homeomorphic.

[1]) In some formulations of the theory, operations of a more general type are allowed. Cf. Newman: „On the Foundations of Combinatory Analysis Situs“, Proc. of the Royal Acad. of Amsterdam, vol. 29, pp. 610—641.

Just a word, now, as to the relation between the three types of analysis situs described above. I say, first of all, that flat analysis situs, in which we are dealing with collections of ordinary flat simplexes, can be re-formulated in purely combinatorial terms, so as to become completely independent of analytic geometry. Suppose, for example, we are dealing with a complex of $n$-simplexes. Then, if we denote the vertices of the complex by the letters

$$a_i, \ (i = 1, \ 2, \ 3, \ \ldots),$$

we can conveniently denote the complex itself by a homogeneous form of degree $n + 1$ with unit coefficients,

$$(1) \qquad \Sigma\, a_{i_0} a_{i_1} \ldots a_{i_n},$$

such that each term of the form represents a simplex of the complex with vertices corresponding to the letters $a_i$ appearing in the term. Conversely, every such form in which there are neither repeated terms nor repeated vertex symbols within the same term represents a class of congruent complexes, one member of which may always be effectively constructed in a space of sufficiently many dimensions. Now, it would, no doubt, be very difficult to give a direct combinatorial formulation of the general operation of subdividing a simplex into smaller simplexes. There is one very special type of subdivision, however, which offers no such difficulty; namely, the one where we introduce a new vertex $b$ on an edge $a_i\, a_j$ and subdivide all simplexes of the complex with the edge $a_i\, a_j$ into pairs of simplexes with the edges $a_i\, b$ and $ba_j$ respectively. In this simple case the combinatorial analogue of the subdivision is merely to pick out all terms of the form (1) containing both $a_i$ and $a_j$ and to replace every such term by a pair of terms, in the first of which the letter $b$ replaces the letter $a_i$ and in the second the letter $a_j$. Let us call this symbolic operation and its inverse (when the inverse is possible) *elementary transformations*, and regard these elementary transformations as the allowable moves of combinatorial analysis situs. It is then possible to prove the following theorem which brings out at once the complete parallelism between the flat and combinatorial theories:

*A necessary and sufficient condition that two complexes of simplexes be flat homeomorphic is that their symbols be equivalent in the sense of combinatorial analysis situs.*

We now come to the much more difficult problem of the relation between point theoretical analysis situs and flat. One main branch of point theoretical analysis situs deals with the theory of abstract spaces. I shall have very little to say about abstract spaces, in general, beyond remarking that one of the clearest ways of dealing with a large class of these spaces is to regard them as limiting cases of complexes in the sense of flat analysis situs, as is done, for example, by Alexandroff. By far the most important spaces are the ones of the so called *manifold* type, and on these I shall concentrate my attention. An $n$-dimensional manifold arising in the course of an ordinary investigation will probably be given in some such form as this. We shall first obtain a portion of the manifold, say an $n$-cell $c_1$; then, by some process analogous to analytic extension, a second $n$-cell $c_2$ overlapping a part of $c_1$; then a third $n$-cell $c^3$ overlapping a part of $c_1 + c_2$, and so on. With each cell $c_i$ let us associate a coordinate system,

$$x_{i1}, \ x_{i2}, \ \ldots, \ x_{in}.$$

Then, if two cells $c_i$ and $c_j$ have a point $P$ in common, the coordinates of the points of $c_i$ neighboring to $P$ will be expressible in terms of the coordinates of the same points regarded as points of $c_j$ by relations of the form

$$(2) \qquad x_i = F_{ij}\,(x_{j1}, \ldots, x_{jn}), \ (i = 1, 2, \ldots, n),$$

where the functions $F_{ij}$ are continuous. Two questions now arise which it would be very desirable to answer: ($\alpha$) Can every manifold $M$ be cut up into a complex of cells, thus making it possible for us to regard it as the image of a complex $S$ consisting of flat simplexes? ($\beta$) If two manifolds $M_1$ and $M_2$ are the images of the simplicial complexes $S_1$ and $S_2$ respectively, can the complexes $S_1$ and $S_2$ (and, therefore, the manifolds $M_1$ and $M_2$) be homeomorphic without being flat homeomorphic? An affirmative answer to these two questions would enable us to reformulate the general theory of manifolds in terms of the flat theory and, therefore, ultimately, in terms of the combinatorial theory. If, by good fortune, the functions $F_{ij}$ in relations (2) are all analytic, or even if they are sufficiently smooth, it is merely a matter of honest toil to show that the manifold $M$ can be cut up into cells, thus answering the first question; but if the functions $F_{ij}$ are merely continous the problem is of quite another order of difficulty. I believe that questions ($\alpha$) and ($\beta$) both reduce, without too much difficulty, to the solution of the following hypothetical lemma:

*Consider an n-simplex $S$ and a one-one continuous transformation $\tau$ carrying the simplex $S$ into a region $R$ of Euclidean n-space. Then we can approximate the transformation $\tau$ as closely as we please by a one-one analytic transformation $\tau_1$ (or by a one-one continuous transformation $\tau_2$, linear in patches)*[2].

The essential difficulty in proving the lemma is to obtain a transformation $\tau_1$ (or $\tau_2$) which is one-one. Let

$$x_i' = F_i\,(x_1, \ldots, x_n), \ (i = 1, 2, \ldots, n),$$

be the equations of the transformation $\tau$. Then, by the classical methods of Weier-

[2]) The existence of the transformation $\tau_1$ can be shown to imply the existence of the transformation $\tau_2$, and vice-versa.

strass, we may approximate the functions $F_i$ as closely as we please by analytic functions $G_i$ and thus obtain an approximating analytic transformation

$$x_i' = G_i(x_1, x_2, \ldots, x_n).$$

Unfortunately, there is nothing to guarantee that the approximating transformation shall be one-one. When we confine our attention to manifolds of one and two dimensions the lemma can be solved without excessive difficulty; consequently, it is possible to reduce the general theory of these manifolds to combinatorial terms.

As soon as we go beyond two dimensions, even the combinatorial theory of manifolds is in a very incomplete state. I should, therefore, like to say a few words about some of the unsolved problems for the case of manifolds of dimensionality 3 – the simplest outstanding case. In the following discussion, it is always to be understood that we are dealing with figures that are so smooth as not to involve us in point theoretical difficulties of a pathological nature.

Suppose we have a closed surface $S$ of genus $p$ situated in ordinary 3-space and in the canonical shape of a sphere with $p$ exterior, tubular handles. The interior of the surface $S$ is then a region of a very special type, characterized by the property that it can be oriented and that it can be made simply connected by cutting it along $p$ suitably chosen 2-cells (2-cells cutting across the respective handles, for example). We shall call such a region a *canonical region* of genus $p$. The significance of canonical regions in the theory of manifolds has been brought out by Heegaard who has shown that in every orientable manifold (of the compact, non-bounded type) there is always at least one surface $S$ of suitable genus $p$ separating the manifold into two such regions. Let us call the surface $S$ a *canonical surface* of the manifold. Then, the minimum value of the genera $p$ of all canonical surfaces in the manifold is a theoretical invariant of the manifold which we are unfortunately unable to calculate effectively in our present state of knowledge. We shall call this invariant the *genus* of the manifold. If the genus of the manifold is zero, for example, it means that there exists a surface of genus 0 separating the manifold into a pair of 3-cells. In other words, the manifold is a 3-sphere. If the genus of the manifold is 1 it means that the manifold can be constructed by taking the interiors and boundaries of two anchor rings in 3-space and piecing the two domains together by setting their boundaries in one-one bi-continuous correspondence with one another. The manifolds of genus 1 are relatively simple in structure and their Poincaré groups are always of the cyclic type, yet no complete classification of them has ever been carried through. It has been shown that there are two distinct manifolds with the same cyclic groups of order 5 and, therefore, with the same Betti numbers and coefficients of torsion. There appear to be two distinct manifolds with cyclic groups of order 7, but the proof that they really are distinct has still to be given. When we come to manifolds of genus 2, examples of so-called *Poincaré spaces* arise. These are manifolds with the same Betti and torsion numbers as a 3-sphere, yet which are not simply connected.

One or two general remarks about the classification of manifolds according to Heegaard's program may, perhaps, be worth making. The problem divides itself naturally into two parts: (*i*) to determine in how many essentially different ways two canonical regions of genus $p$ can be matched together to form a manifold; (*ii*) to determine in how many essentially different ways a canonical region can be traced in a manifold. The first part of the problem does not seem hopelessly difficult; it is closely related to the problem of the number of essentially different one-one mappings of one surface of genus $p$ on another. As to the second part of the problem, I have a strong suspicion that if $S$ and $S^1$ are two canonical surfaces of the same genus in a manifold $M$ then there is always a continuous deformation of the manifold $M$ into itself carrying the surface $S$ into the surface $S^1$. It would be interesting to have a proof of this hypothetical theorem even for the case where the manifold $M$ is a hypersphere. The theorem for a general manifold $M$ seems to be reducible to this special case.

Poincaré once proposed the problem of determining whether the 3-sphere is the only 3-dimensional manifold such that its Poincaré group reduces to the identity – that is to say, such that every closed curve traced in the manifold can be deformed continuously into a point. This problem is a special case of the following more general one: is it always possible to map the universal covering space of a 3-dimensional manifold on a 3-sphere or on a portion of a 3-sphere? For the case of a 2-dimensional manifold, we know that the universal covering surface is always either a cell or a sphere. This simple theorem does not generalize, however, as the following example will show. We take the closed domain consisting of a region $R$ of 3-space bounded by two concentric spheres together with the spheres themselves, and form a manifold out of this domain by matching each point $P$ of the outer sphere with the point $P^1$ on the inner sphere in which the ray from the common centre of the spheres to $P$ cuts the inner sphere. It is then easy to see that the group of the manifold thus obtained is the infinite cyclic group and that the universal covering space is homeomorphic with a finite portion of 3-space bounded by two spheres, such as the region $R$ itself. The covering space may, however, be mapped on a portion of a 3-sphere, so that this example does not invalidate the theorem suggested above.

It would be interesting to have a complete census of all 3-dimensional manifolds with finite Poincaré groups. The covering space of a manifold of finite group is always a Poincaré space such that its group reduces to the identity. Very possibly this covering space is always a hypersphere. Let us assume that this is the case. Then the problem reduces to the determination of all the finite groups of transformations of a 3-sphere into itself such that no one of the transformations leaves invariant a point of the 3-spheres. If our assumption about the covering space is false we can, at all events, obtain in this way an important class of manifolds with finite Poincaré groups.

The theory of 3-dimensional manifolds seems to be closely dependent on the theory of knots and linkages in ordinary spherical 3-space. It can be shown, for example, that every orientable 3-dimensional manifold can be represented by a suitable $n$-sheeted Riemann space (in the sense of a generalized Riemann surface) with a finite number of simple, non-intersecting branch curves about which the various sheets of the space are permuted. What makes this mode of representation so complicated is the fact that the branch curves can be arbitrarily knotted and inter-linking. A number of useful knot invariants have been discovered in recent years which make it possible to distinguish between all the simpler knots and, in particular, to determine whether an alternating knot can be reduced to an un-

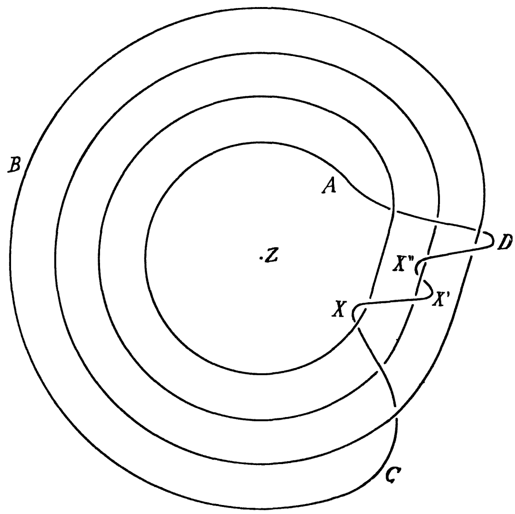



Fig. 2

knotted curve[3]). Every knot can be reduced to a simple semi-canonical form (fig. 2), which is unfortunately not unique. It consists of two arcs: an arc $A\,B\,C$ in the form of a plane spiral and an arc $C\,D\,A$ in the form of a 3-dimensional curve weaving over and under the turns of the spiral in such a way that as a variable point $x$ moves

[3]) For a connected account of modern knot theory the reader is referred to Reidemeister's genial monograph: Knotentheorie, Springer (1932).

along the arc from $C$ to $A$ the radius vector $ZX$ from the centre of the spiral to $X$ turns continually in the same direction around $Z$. If the arc $C\ D\ A$ goes straight across from $C$ to $A$ without doubling back on itself (as it doubles back, for instance, at $X, X', X''$ and $D$ in the figure) we can easily see that the curve is unknotted, no matter how $C\ D\ A$ weaves through the turns of the spiral. Perhaps it can be shown that if the arc $C\ D\ A$ *does* double back on itself, then, either the curve has a knot in it or else there is an obvious transformation which reduces the representation to one in which the arc $C\ D\ A$ no longer doubles back.

In conclusion, my main regret is that it has been so much easier to make up mathematical knots than to untie them.