A GEOMETRIC METHOD IN DIFFERENTIAL TOPOLOGY

BY ANDREW H. WALLACE

Introduction. My aim here is to explain in an informal way a technique which has proved itself useful in dealing with certain problems in differential topology. The title refers to the fact that the method relies principally on the geometric construction of spherical modifications and certain associated constructions. The general plan of this exposition is, firstly, to define the notion of spherical modification, secondly, to describe some operations which can be performed on these modifications, and thirdly, to give some applications. The article concludes with some remarks on related work. A more formal and detailed presentation of the material treated here appears in [3;4].

1. Definition. Let M_1 be a differentiable manifold of dimension n and let S^r be an r-sphere embedded in M_1. If S^r has a neighborhood in M_1 of the form $S^r \times E^{n-r}$, where E^{n-r} is an $(n-r)$-cell, then S^r is said to be directly embedded in M_1. If, in this case, $S^r \times E^{n-r}$ is removed from M_1, what is left is a manifold having a boundary of the form $S^r \times S^{n-r-1}$. The latter set is, however, also the boundary of $E^{r+1} \times S^{n-r-1}$. Thus, forming the union

$$\{ M_1 - (S^r \times E^{n-r}) \} \cup (E^{r+1} \times S^{n-r-1})$$

with the appropriate identification of boundaries, a new manifold M_2 is obtained, which can also be made differentiable by a suitable smoothing operation. The manifold M_2 is said to be obtained from M_1 by a spherical modification.

A familiar example of this process is obtained by taking M_1 to be the hyperboloid of two sheets $x^2 - y^2 - z^2 = 1$ and M_2 to be the hyperboloid of one sheet $x^2 - y^2 - z^2 = -1$ in Euclidean 3-space. S^r is to be taken as the 0-sphere consisting of the union of the two points $(-1, 0, 0)$ and $(1, 0, 0)$. In this case the set removed is a pair of discs and the set inserted is a circular cylinder.

More generally, M_i could be taken as the quadric hypersurface

$$\sum_{i=1}^{r+1} x_i^2 - \sum_{i=r+2}^{n+1} x_i^2 = 1$$

in $(n+1)$-space, and M_2 as the hypersurface whose equation is obtained from the above by changing the right hand side to -1. In this

example it is easily seen that one can make a continuous transition from M_1 to M_2, namely by passing through the family of quadrics

$$(1) \qquad \sum_{i=1}^{r+1} x_i^2 - \sum_{i=r+2}^{n+1} x_i^2 = t,$$

allowing t to vary from 1 to -1. All the quadrics in this family are nonsingular except when $t=0$, in which case the hypersurface has a nondegenerate quadratic singularity. Also, still referring to this example, the family of orthogonal trajectories F to the family (1) can be constructed. The family of curves F will have one singular point, namely the origin. All the members of F meeting the r-sphere S^r in M_1 obtained from (1) by setting $t=1$ and $x_i=0$ for $i \geq r+2$ or meeting the $(n-r-1)$-sphere S^{n-r-1} in M_2 obtained by setting $t=-1$ and $x_i=0$ for $i \leq r+1$ end at the origin. And all the members of F meeting neighborhoods of these spheres in M_1 and M_2 fill out an $(n+1)$-cell, while the union of the remaining members of F makes up the product $(M_1 - (S^r \times E^{n-r})) \times I$.

The description just given applies not only to this example, but to any transformation of a manifold M_1 to M_2 by a spherical modification. This follows from the fact that the whole operation of a spherical modification, as defined earlier in this section, depends only on a neighborhood of the corresponding r-sphere S^r. Thus also in the general case M_1 and M_2 form the boundary of a manifold M, which will be called the trace of the modification, and there is a family F of curves in M, one curve passing through each point with the exception of one singular point. These curves can be parametrized by a parameter t so that M_1 and M_2 are the sections $t=0$ and $t=1$, respectively, and so that all the sections $t=$ constant are nonsingular with one exception. This exceptional section will have one nondegenerate quadratic singularity P, namely the singularity of the family F. In the notation used earlier in this section, it will be noted that the members of the family F beginning on S^r all end at P, and those ending at a certain $(n-r-1)$-sphere S^{n-r-1} in M_2 all begin at P. For this reason it is convenient to think of the modification as shrinking S^r and introducing S^{n-r-1}.

More generally now, suppose that M_2 is reached from M_1 by performing some finite number of spherical modifications. Their traces can be put together to form a manifold M (to be called the trace of the set of modifications) with boundary $M_1 \cup M_2$. Thus M_1 and M_2 are cobounding manifolds. Also the families of curves corresponding as described above to the modifications can be put together to form one family in M, one curve passing through each point with a finite number of exceptions, one for each modification.

The converse of the result just stated is also true. Namely, if M_1 and M_2 are cobounding manifolds then there is a finite sequence of spherical modifications leading from M_1 to M_2. This can be proved by making a suitable embedding of the manifold M whose boundary is $M_1 \cup M_2$ in N-space so that M_1 and M_2 are the sections $x_N = 0$ and $x_N = 1$, respectively, while the sections of M by $x_N = $ constant are all nonsingular with a finite number of exceptions having nondegenerate quadratic singularities. Each singular section then corresponds to one modification.

2. **Rearrangement of modifications.** Some operations on modifications are now to be described. The first, which will be considered in this section, is the operation of switching the order of performing two modifications, under suitable conditions. Referring to Figure 1 for
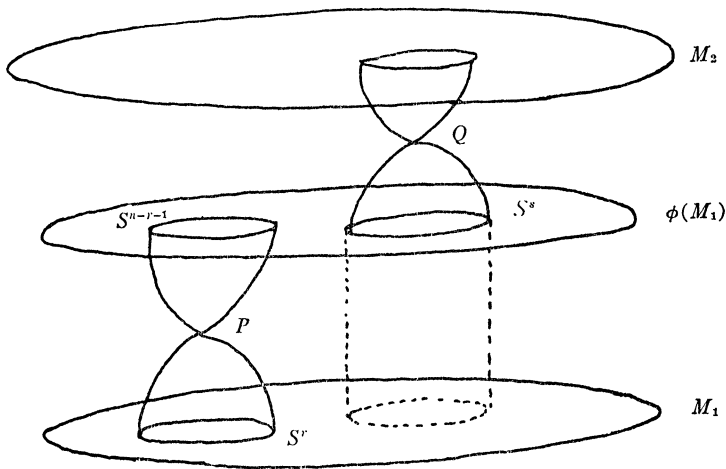

FIGURE 1.

guidance, let M_1 be transformed to M_2 by two modifications ϕ, ψ, the combined trace of these modifications being M. ϕ shrinks the sphere S^r to the singular point P of the curve family F associated with the modifications, and this modification introduces S^{n-r-1}. Let ψ, operating on $\phi(M_1)$, shrink the sphere S^s to the singular point Q of the family F and introduce S^{n-s-1}. Assume that $S^s \cap S^{n-r-1} = \varnothing$. Then S^s can be pulled back along the curves F into M_1. Thus the two modifications ϕ and ψ can both be thought of as operating on M_1. In the present arrangement, the modification ϕ is performed first, and so, as the parameter on the family F increases from 0 to 1, the singular point P is reached before Q. However the parameter can be changed

so that it increases faster on the curves of F starting at points near S^r. The result will be that, as one proceeds along the curves F, using the new parameter, the singular point Q is reached before P. Hence the modification ψ is performed before ϕ. That is to say, the order of carrying out the pair of modifications ϕ and ψ has been reversed.

The condition for the above reversal in the order of ϕ and ψ is that $S^s \cap S^{n-r-1} = \varnothing$. This can always be guaranteed if $s \leq r$. For in this case S^s can be made disjoint from S^{n-r-1} by a small isotopic displacement, and it is clear that the result of a modification is unaffected by such a displacement of the sphere to be shrunk.

The result just obtained can be applied systematically to any finite sequence of modifications leading from M_1 to M_2. When this is done, it can be arranged that modifications shrinking s-spheres are performed before those shrinking r-spheres whenever $s < r$. And of course modifications shrinking spheres of the same dimension can be commuted freely among themselves.

3. **Addition of modifications.** The title of this section is a convenient but somewhat elliptical form of description, since the things which are actually to be added are elements of homology or homotopy groups.
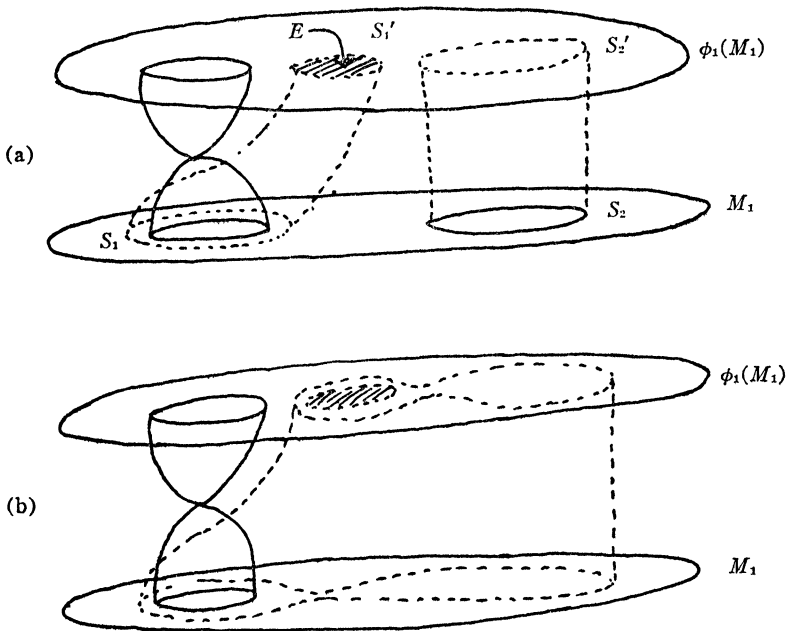

FIGURE 2. (b) shows S_2' deformed across E, then pulled back into M_1.

Let ϕ_1 and ϕ_2 be modifications shrinking r-spheres S_1, respectively S_2 in M_1, assumed disjoint, and let S_1, S_2 carry elements α_1, α_2, respectively, of $H_r(M_1)$. Suppose that the modification ϕ_1 is done first. Then the curve family F associated with ϕ_1 can be used to pull a displaced copy of S_1 up into $\phi_1(M_1)$, giving a sphere S_1' in that manifold. S_1' will, however, bound a cell E in $\phi_1(M_1)$. S_2, being disjoint from S_1, is carried into $\phi_1(M_1)$ by the curve family F, giving a sphere S_2' in $\phi_1(M_1)$. ϕ_2 is now to be performed on $\phi_1(M_1)$ by shrinking the sphere S_2'. But, as already remarked, the result of a modification is unaffected by an isotopic deformation of the sphere to be shrunk. In particular, ϕ_2 gives the same result if S_2' is first deformed across the cell E. But now, if the displaced S_2' is pulled back into M by means of the curve family F, a sphere is obtained which carries $\alpha_2 \pm \alpha_1$ instead of α_2, the sign depending on the way in which the deformation was made.

Hence the result of the pair of modifications ϕ_1, ϕ_2 is the same as that of a pair of modifications ψ_1, ψ_2 shrinking spheres carrying α_1, $\alpha_2 \pm \alpha_1$. Also the trace of the pair ψ_1, ψ_2 is the same as that of ϕ_1, ϕ_2.

Repeated application of this result shows that the pair of modifications ϕ_1, ϕ_2, shrinking spheres carrying α_1, α_2 in $H_r(M_1)$, gives the same result (and has the same trace) as a pair of modifications shrinking spheres carrying α_1 and $\alpha_2 + k\alpha_1$, where k is any integer.

Again this result may be applied repeatedly to a set of modifications ϕ_1, ϕ_2, $\cdots$, ϕ_h shrinking disjoint r-spheres carrying $\alpha_1, \alpha_2, \cdots$, α_h in $H_r(M_1)$. These modifications will then give the same result and have the same trace as a set of modifications $\psi_1, \psi_2, \cdots, \psi_h$ shrinking spheres carrying $\beta_1, \beta_2, \cdots, \beta_h$ in $H_r(M_1)$, where the β_i are obtained from the α_j by a unimodular linear transformation, for any such transformation can be expressed as a product of transformations of the type $(\alpha_1, \alpha_2) \to (\alpha_1, \alpha_2 + k\alpha_1)$.

The disjointness of the spheres involved in this result will usually arise automatically in applications, the most important cases being where $2r < \dim M_1$. The result is stated above in terms of homology, but a similar statement can be made in terms of homotopy groups, with suitable adjustments relative to the base point.

4. **Complementary modifications and cancellation.** Suppose that a modification ϕ on M_1 shrinks a sphere S^r which is the boundary of an $(r+1)$-cell E^{r+1}. The effect of the modification is to close up E^{r+1} to form an $(r+1)$-sphere S^{r+1} in $\phi(M_1)$. Now suppose that this sphere S^{r+1} is directly embedded and let ψ be a modification shrinking it. ψ will be said to be complementary to ϕ.

When two modifications are related in this way, it can be shown

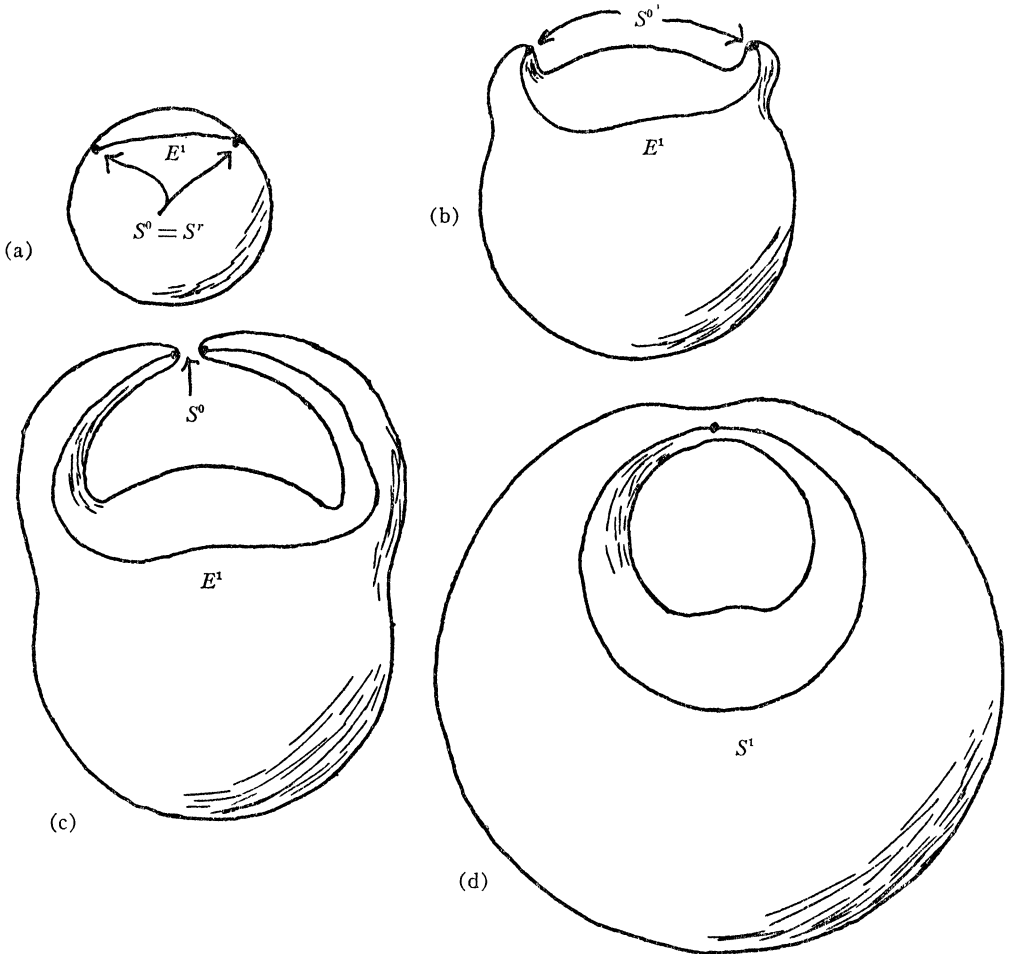that the final result $\psi\phi(M_1)$ is M_1 itself and that the trace of the pair ϕ, ψ is $M_1 \times I$.

FIGURE 3. (a) $S^r = S^0$ is marked on S^2; here $t = 0$.
     (b) t increases.
     (c) t increases further, and points forming S^0 are about to coalesce.
     (d) Modification ϕ is completed. Surface is now a torus, and stages (a), (b), (c) are supposed to be inside the solid which it bounds.

To illustrate this (Figure 3) take $M_1 = S^2$ and consider a modification ϕ shrinking a 0-sphere to give a torus for $\phi(M_1)$. The trace of ϕ is a solid bounded by a sphere and a torus, that is to say, a solid torus with a spherical hole in it. If the family F of curves correspond-

ing to ϕ is parametrized by t, the set $t=0$ is the inner spherical bound-
ary. As t increases from zero the set $t=$ constant will be a sphere in-
creasing in size, and growing horns around the points forming the 0-
sphere to be shrunk by ϕ. As t increases these horns meet and join
so that a torus is formed. The join of the two points forming S^r is a
1-cell which ϕ closes up to form a circle S^1 on the outer surface of the
solid torus. A second modification ψ is to shrink S^1. Reasoning as
above, it is clear that the trace of ψ will be a solid having the torus as
its inner boundary and a sphere as its outer boundary. Then the
traces of ϕ and ψ put together will form a solid sphere with a spherical
hole in it, that is to say $S^2 \times I$.

5. **Applications.** Suppose first that M_1 and M_2 form the boundary
of the manifold M, and that the corresponding set of modifications
leading from M_1 to M_2 consists of one, ϕ, shrinking an $(r-1)$-sphere
S^{r-1} and a number of modifications shrinking r-spheres. The results
of §2 imply that ϕ can be assumed to be done first. Assume also that
$\pi_r(M, M_1) = 0$. This implies in particular that the r-cell E_1 bounded in
M by S^{r-1} can be pulled back into M_1 giving a cell E_2 with boundary
S^{r-1} (cf. Figure 4). Then the modification ϕ shrinks S^{r-1} and so closes
up E_2 to form a sphere S^r. This is homotopic in M to $E_1 \cup E_2$ and so
bounds a cell in M. Now there is a deformation retraction, using the
curve family F, of M onto the union of $\phi(M_1)$ with a number of cells,
namely an $(n-r)$-cell bounded by the sphere introduced by ϕ and
$(r+1)$-cells bounded by the spheres to be shrunk by the other modi-
fications. Assuming here that $2r < \dim M_1$, S^r will thus carry an element
of $\pi_r(\phi(M_1))$ which is a linear combination of elements carried by
r-spheres to be shrunk by modifications leading to M_2. Provided that
$\dim M_1 \geqq 5$, the result of §3 can be applied to rearrange these modi-
fications so that one of them actually shrinks S^r. This modification
will be complementary to ϕ and so will cancel it out.

More generally the above cancellation can be carried out system-
atically for any M_1 and M_2 forming the boundary of a manifold M
and satisfying the conditions

$$\dim M_1 = \dim M_2 = 2m \text{ or } 2m - 1, \quad m \geqq 3,$$

$$\pi_i(M, M_1) = 0, \quad i \leqq m - 1.$$

It will then follow that the corresponding modifications can be re-
arranged so that all those shrinking spheres of dimension less than
$m-1$ cancel out.

If also $\pi_i(M, M_2) = 0$, $i \leqq m-1$, the same reasoning can be carried
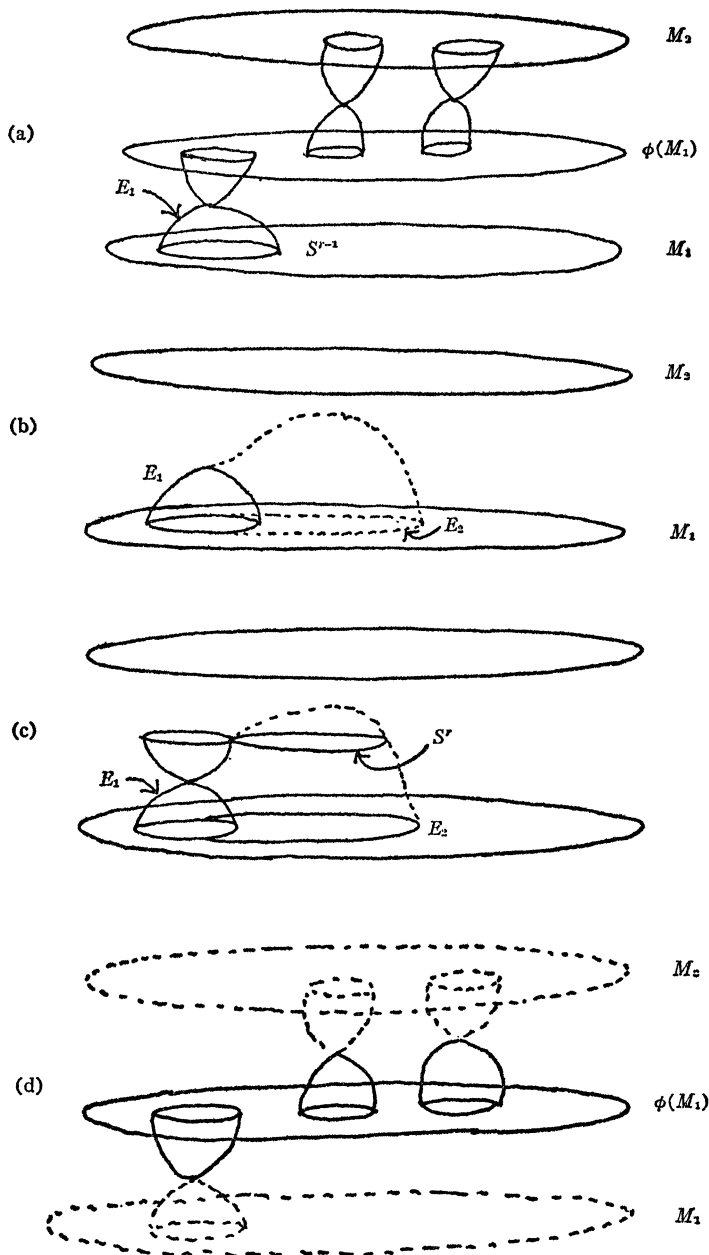out starting from M_2 and the result will be that all modifications will

FIGURE 4. (b) shows E_1 pulled back into M_1; (c) shows E_2 closed by ϕ to form S^r bounding a cell in M; (d) shows M retracted on $\phi(M_1)$ with one $(n-r)$-cell and several $(r+1)$-cells attached.

cancel out except one or two types. Indeed it can be shown that if, in addition, M_1 and M_2 are simply connected and $\pi_m(M, M_1) = \pi_m(M, M_2) = 0$, then, using a refinement of the above methods, all the modifications cancel out. Thus under these conditions $M_1 = M_2$ and $M = M_1 \times I$.

In particular let M be a simply connected homology sphere of dimension $n+1$. Remove two $(n+1)$-cells from M. The remainder M' has boundary $M_1 \cup M_2$, where the M_i are n-spheres, and $\pi_i(M', M_j) = 0$ for $j = 1$, 2 and $i \leq n$. The above result implies that all the corresponding modifications leading from M_1 to M_2 cancel out, and so M' is $M_1 \times I = S^n \times I$, for $n \geq 5$. Hence $M = S^{n+1}$. This solves the Poincaré problem for dimensions greater than 5.

6. **Remarks.** The operation called here a spherical modification appears elsewhere under various names, for example surgery, and χ-construction (cf. [1]). One of the original applications [1] was to the problem of killing the homotopy groups of a manifold. The idea is to apply modifications shrinking spheres carrying generators of homotopy groups, when these spheres are directly embedded. A systematic scheme for such a killing process is described in [4] and in the forthcoming papers [5].

The situation described here is closely related to the theory of critical points of functions on a manifold. In the present notation if the curve family F is parametrized by t, then the singular points of this family, corresponding to modifications leading from M_1 to M_2, are critical points of the function t on M. The rearrangement of a sequence of modifications so that spheres of lower dimension are shrunk first corresponds to Smale's construction of a "nice function" on M [2]. And in fact Smale's approach to the generalized Poincaré problem is closely related to the above treatment, the manipulation of modifications being replaced by the construction of functions with suitable properties.

The problem which first led to my interest in this subject was that of giving some sort of geometrical account of the relation of cobordism or cobounding between manifolds. The homotopy killing process mentioned above can be applied to any manifold to give a cobounding manifold with a specially simple homotopy structure. If this is assumed done in advance to all manifolds, then, when two manifolds belong to the same cobordism class, they are related by modifications of only one or two types [4]. In forthcoming papers [5] this will be further strengthened. However, some essential refinement of method seems necessary before the ultimate ideal is attained, namely of giving

in geometrical terms canonical representatives for the cobordism classes along with a geometrical algorithm for reducing a given manifold to the appropriate form.

REFERENCES

1. J. Milnor, *A procedure for killing homotopy groups of differentiable manifolds*, Differential Geometry, Proc. Sympos. Pure Math. Vol. 3, pp. 39–55, Amer. Math. Soc. Providence, R. I., 1961.

2. S. Smale, *Poincaré conjecture in higher dimensions*, Bull. Amer. Math. Soc. **66** (1960), 373–375.

3. A. H. Wallace, *Modifications and cobounding manifolds*, Canad. J. Math. **12** (1960), 503–528.

4. ———, *Modifications and cobounding manifolds*. II, J. Math. Mech. **10** (1961), 773–809.

5. ———, *Modifications and cobounding manifolds*. III, J. Math. Mech. **11** (1961); also IV. (to appear).

INDIANA UNIVERSITY