

On the structure of closed 3-manifolds

by

Jan Mycielski (Boulder, CO)

Abstract. We will show that for every irreducible closed 3-manifold M , other than the real projective space P^3 , there exists a piecewise linear map $f : S \rightarrow M$ where S is a non-orientable closed 2-manifold of Euler characteristic $\chi \equiv 2 \pmod{3}$ such that $|f^{-1}(x)| \leq 2$ for all $x \in M$, the closure of the set $\{x \in M : |f^{-1}(x)| = 2\}$ is a cubic graph G such that $S - f^{-1}(G)$ consists of $\frac{1}{3}(2 - \chi) + 2$ simply connected regions, $M - f(S)$ consists of two disjoint open 3-cells such that $f(S)$ is the boundary of each of them, and f has some additional interesting properties. The pair $(S, f^{-1}(G))$ fully determines M , and the minimal value of $\frac{1}{3}(2 - \chi)$ is a natural topological invariant of M . Given S there are only finitely many M 's for which there exists a map $f : S \rightarrow M$ with all those properties. Several open problems concerning the relationship between G and M are raised.

0. Introduction. An n -dimensional manifold, or briefly n -manifold, is called *closed* if it is compact, connected and boundaryless. In this paper we deal only with two- or three-dimensional manifolds, thus without loss of generality we can limit ourselves to the piecewise linear category.

Let us recall briefly the definition and basic results of the theory of connected sums. If M and N are two closed 3-manifolds, we can excavate from each of them the interior of a tame 3-cell and identify the resulting boundary spheres S^2 . The resulting manifold is called their *connected sum*, denoted $M \# N$. In general $M \# N$ depends upon the choice of one of the two non-homotopic identifications of those spheres. (If M or N is not orientable this choice does not matter.)

A closed 3-manifold M is called *irreducible* if every tame sphere $S^2 \subseteq M$ disconnects M and at least one of the two connected components of $M - S^2$ is an open 3-cell. And M is called *prime* if it is not a sphere S^3 and it has no decomposition into a connected sum of two closed 3-manifolds both different from S^3 . Hence every irreducible M , with the exception of the sphere S^3 , is prime. It is also known that all prime manifolds with two exceptions are irreducible. The exceptions are $S^2 \times S^1$ and $S^2 \tilde{\times} S^1$, where $\tilde{\times}$ is the skew

product (thus $S^2 \tilde{\times} S^1$ is not orientable). In 1929 H. Kneser showed that every M other than S^3 has a factorization into a connected sum of a finite sequence of prime 3-manifolds. Moreover, as shown by J. Milnor [M], if M is orientable then the Kneser decomposition is unique (up to the order of the factors), and if M is not orientable then the decomposition is also unique if we impose the additional condition that $S^2 \tilde{\times} S^1$ does not occur unless all the other prime factors are orientable and, in that case, it occurs only once. (If M is not orientable then the connected sums $(S^2 \tilde{\times} S^1) \# M$ and $(S^2 \times S^1) \# M$ are homeomorphic.) The non-orientable cases (and extensions of the above results to manifolds with boundary) are due to J. Hempel [H]. An error in [H] was corrected independently in [JP] and [Tr].

A closed 3-manifold M will be called *aprojective* if M does not contain the real projective plane P^2 . We point out an easy proposition: *if M is prime, orientable, and other than the real projective space P^3 , then M is aprojective.*

The purpose of this paper is to show that all irreducible closed 3-manifolds M , with the possible exception of P^3 , can be constructed in a certain way outlined above in the Abstract, and that the construction simplifies in many cases if M is aprojective. In Section 1, we will define the construction, show that it produces closed 3-manifolds (it seems that not all manifolds which it produces are irreducible), and prove a few properties of this construction. In Section 2, we will show that each irreducible 3-manifold other than P^3 can be constructed in this way. (It may be useful for the reader to jump now to Section 2, since the proofs which appear there are simple and they illustrate and motivate the definitions of Section 1.)

I am indebted to Randall Dougherty, Józef Przytycki and the referee for critical remarks which led to several improvements. Moreover, Dougherty solved Problem 3 (see below).

1. Definitions and results. We begin our construction as follows. Choose a positive integer s and let S be a non-orientable closed 2-manifold of Euler characteristic $\chi(S) = 2 - 3s$. Recall that a *cubic graph* G is a connected graph whose vertices are of valency 3 (multiple edges and loops are allowed). Choose a cubic graph G with $2s$ vertices (and hence $3s$ edges). [For example Kuratowski's graph $K_{3,3}$ and Petersen's graph are cubic with $s = 3$ and $s = 5$ respectively. See Fig. 1 for all cubic graphs with $s \leq 3$. According to A. T. Balaban [B] for $s = 1, 2, \dots, 5$ there are 2, 5, 17, 71 and 388 cubic graphs, respectively.] An *isthmus* is an edge of G whose interior disconnects G . Many cubic graphs have loops and each loop is connected to the rest of the graph by an isthmus.

tions. Let $H = h(^2G)$ and let the simple closed curves and the figure 8 curves of H which are the images of the doubled edges of G be called *basic curves of type 0 and type 8*, respectively. Thus if M is to be irreducible and aprojective there are no basic curves of type 8. We impose three conditions on the pair (S, H) :

(α) If two basic curves of type 0 intersect at a vertex v of H , then they cross each other in S at v . If C is a basic curve of type 8 at v , then the two loops of C kiss the basic curve of type 0 which they meet at v from opposite sides of O .

(β) $S - H$ consists of simply connected regions. (Since $\chi(S) = 2 - 3s$ and H has $2s$ vertices and $6s$ edges, there are $s + 2$ such regions.)

To state the third condition (γ) we need the following operations. We cut S along all the edges of H , thus getting $s + 2$ disjoint simply connected pieces (polygonal disks). Then, if possible, we reconstruct from these pieces some other surfaces in two different ways. Namely given any vertex v of H we glue some pairs of edges of those pieces which were meeting at v in one of two ways A and B shown in Fig. 2, which can be explained as follows.

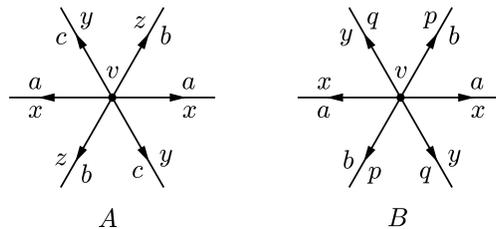


Fig. 2

In the case A , v doubles into two vertices of the reconstructed surface with the cycles of edges (a, b, c) and (x, y, z) at the new vertices. In the case B , v triples with the cycles of edges (a, b) , (p, q) and (x, y) at the new vertices. Notice that once we decide how to glue one pair of edges, say a to a as in A (or as in B), then the whole pattern of gluing all the $s + 2$ pieces is uniquely determined (since H is connected and only the patterns A or B are allowed at the vertices).

REMARK 1. If C is a basic curve of type 0 and v_1, v_2 are the two vertices of C , then the pieces meeting at v_1 and those meeting at v_2 are glued in the same way (both as in A or both as in B) iff C has an orientable open neighborhood in S . Likewise they are glued in different ways (one as in A and the other as in B) iff all open neighborhoods of C in S are non-orientable.

Notice that if our reconstructions are possible, then each yields a finite set of closed 2-manifolds. Our last condition on (S, H) is:

(γ) Both reconstructions are possible and each yields a single sphere S^2 .

Twelve examples of pairs (S, H) satisfying conditions (α), (β) and (γ) are shown in Figs. 3–13. The plane of each figure represents S with its 3s crosscaps represented by small empty circles. The large smooth closed curves are the basic curves of type O . Smaller type 8 basic curves appear in Figs. 4, 5, and 8. Three branches cross each other at every vertex and any (positive) number of basic curves may traverse the crosscaps. The vertices are split into two sets V_1, V_2 (marked 1 or 2) which will be explained below (in condition (γ_1)).

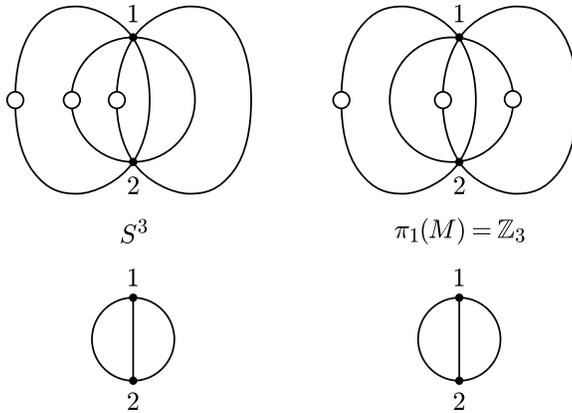


Fig. 3

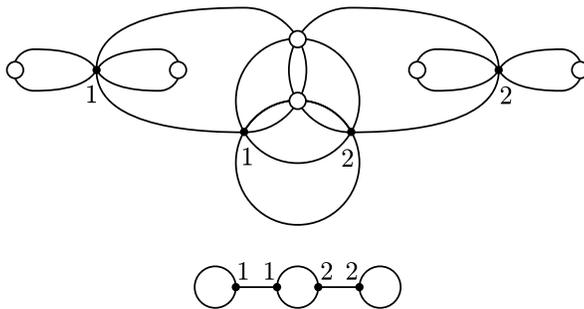


Fig. 4

PROBLEM 1. Which cubic graphs yield such constructions? (The first and fourth graph of Fig. 1 are not possible. Compare Problem 3 below.)

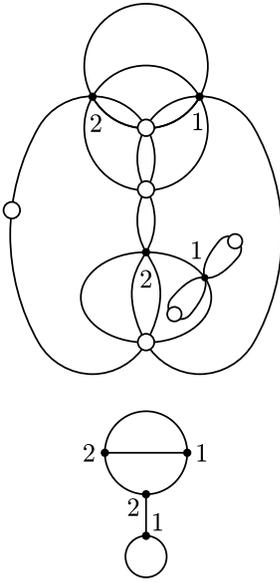


Fig. 5

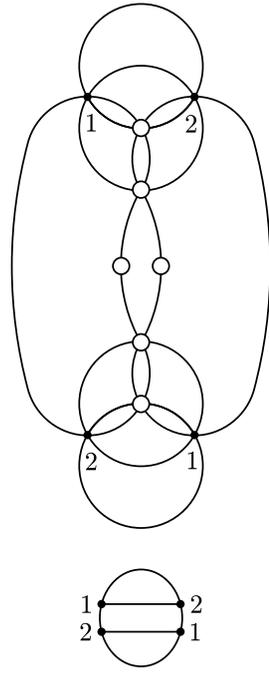


Fig. 6

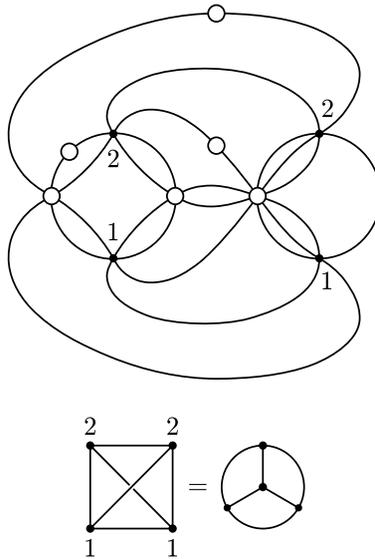


Fig. 7

REMARK 2. In order to check that a pair (S, H) satisfies condition (γ) it is useful to observe that (γ) is equivalent to the conjunction of the two

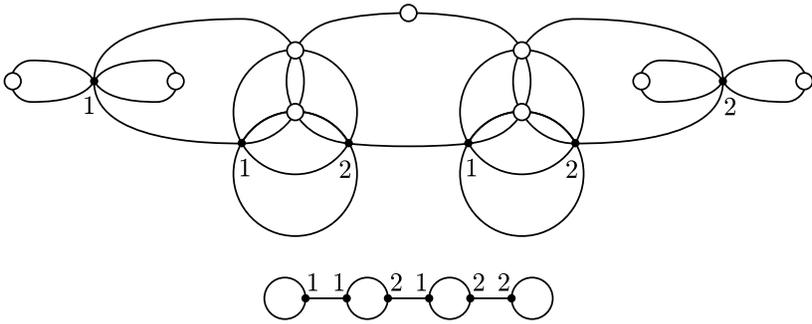


Fig. 8

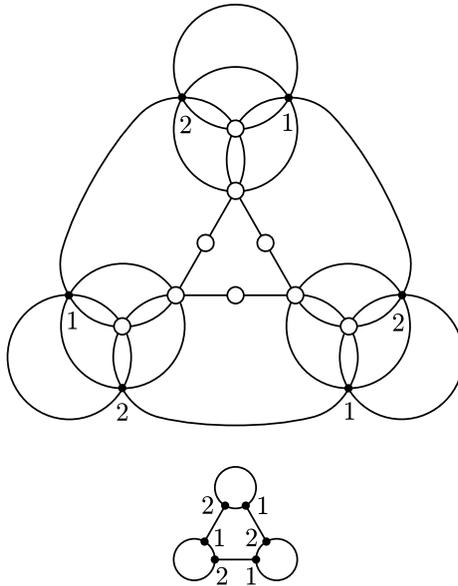


Fig. 9

conditions (γ_1) and (γ_2) below. (Figs. 3–13 were constructed with the help of this equivalence.)

(γ_1) The set of vertices of H can be split into two disjoint sets V_1 and V_2 , both of size s , such that every basic curve of type O whose vertices belong to the same V_i has an orientable open neighborhood in S , and every open neighborhood of every basic curve of type O whose vertices belong to different sets V_i is non-orientable. Finally, for each basic curve C of type 8, none of the loops of C have orientable open neighborhoods in S .

Notice that (γ_1) implies that both reconstructions are possible. By Remark 1, (γ) implies the first part of (γ_1) for basic curves of type O . It

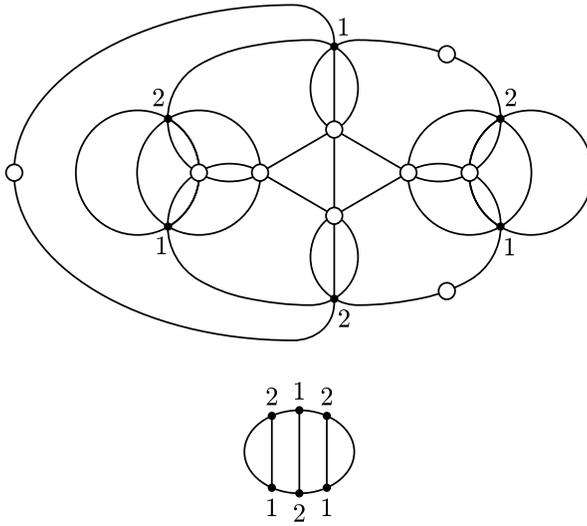


Fig. 10

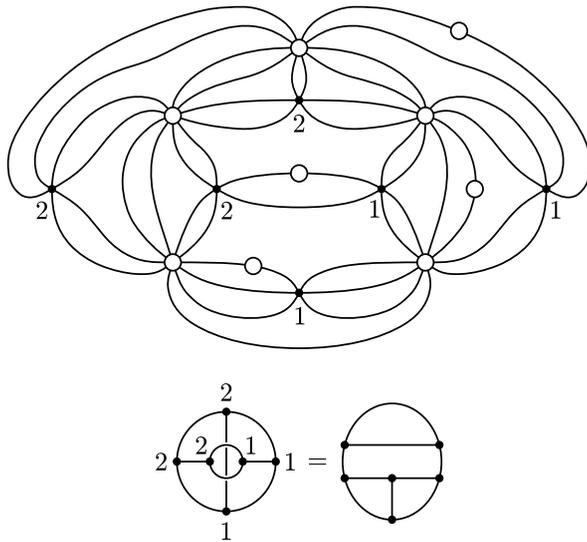


Fig. 11

is easy to check that it implies also the second part for basic curves of type 8.

(γ_2) Both reconstructions yield a single connected surface.

Observe that (γ_1) & (γ_2) implies (γ) . Indeed, by (γ_1) each reconstruction yields a set of closed 2-manifolds whose total number of vertices, edges and

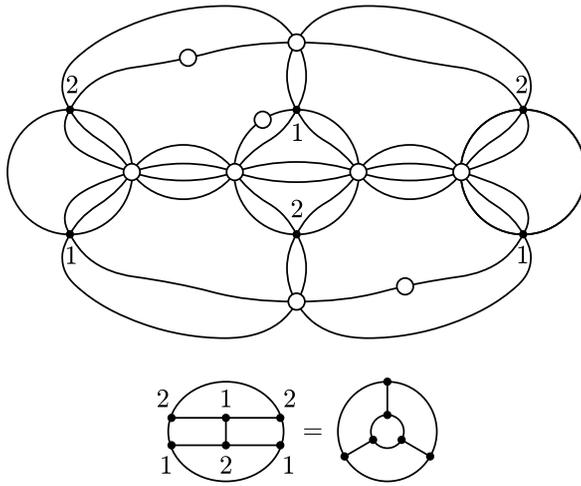


Fig. 12

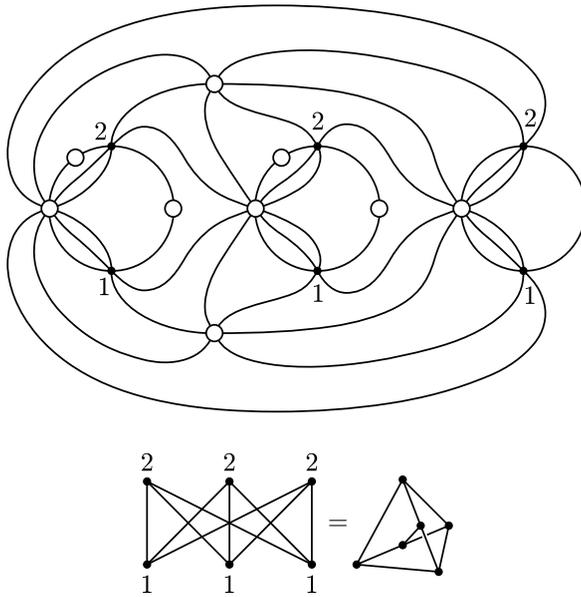


Fig. 13

regions is $5s$, $6s$ and $s + 2$ respectively. Thus the Euler characteristic of the whole set is 2. Hence by (γ_2) both sets are single spheres S^2 .

REMARK 3. The condition (γ_1) could not have been satisfied if S was orientable.

Finally, assuming (α) , (β) and (γ) , we will construct a certain space M : We attach to each of the spheres given by (γ) disjoint open 3-cells obtaining disjoint closed 3-cells C_1 and C_2 . Then we identify all pairs of edges in the boundary of each C_j in the way they were identified in S (prior to the reconstruction). Finally, M is obtained by identifying all pairs of 2-dimensional faces of those two complexes in the way they correspond to the 2-faces of (S, H) .

PROPOSITION. *The space M is a closed 3-manifold and M is fully defined by the pair (S, H) .*

Proof. The first part can be checked by inspecting the construction at all its vertices. The second part follows from Remark 1. ■

Our construction of M defines a piecewise linear map

$$f : S \rightarrow M$$

such that f identifies the pairs of arcs of the basic curves of H such that $f(H)$ is homeomorphic to the original cubic graph G . And $M - f(S)$ consists of two open 3-cells such that $f(S)$ is the boundary of each of them.

REMARK 4. Given the graph G , M still depends on the homeomorphism $h : {}^2G \rightarrow S$. Indeed the two pairs (S, H) of Fig. 3 yield different M 's. The first is the sphere S^3 and the second has a cyclic group of order 3 as its fundamental group.

The main result of this paper is the following

THEOREM 1. *Every irreducible closed 3-manifold, with the possible exception of the real projective space P^3 , can be constructed in the above way.*

The proof will be given in Section 2.

REMARK 5. Given the integer s there are only finitely many 3-manifolds M which can be obtained in the above way. [Indeed, there are only finitely many pairs (S, H) which can be constructed from $2s$ vertices, $6s$ edges and $s+2$ regions.] This finiteness distinguishes our construction from Heegaard's construction (by identifying the boundaries of two handlebodies of the same genus); see [H]. Indeed, already in the case of handlebodies of genus 1 Heegaard's construction yields infinitely many irreducible 3-manifolds.

PROBLEMS. 2. Can any of the manifolds P^3 , $S^1 \times S^2$ and $S^1 \tilde{\times} S^2$ be constructed in the above way?

3. Recall Petersen's theorem that every cubic graph without isthmus (i.e., no single point disconnects the graph) has a 1-factor. Suppose that G is isthmus free. Must the graph G of our construction have a 1-factor F such that no basic curve corresponding to an edge of F has an orientable open

neighborhood in S ? *Added in proof:* R. Dougherty found a counterexample in which the cubic graph G is the graph of edges of a cube.

4. Suppose again that G is isthmus free. Can our construction yield a composite (i.e., non-prime) manifold?

5. Can one choose one edge from each basic curve such that the union of those 3s edges does not disconnect S ?

REMARK 6. Since the complement of $f(S)$ in M consists of two open 3-cells, $f(S)$ is a retract of M twice punctured. Hence

$$\pi_1(M) = \pi_1(f(S)).$$

REMARK 7. Suppose that G has no loops. Then M is not simply connected iff S has a proper covering map $\phi : \tilde{S} \rightarrow S$ such that, for each basic curve C of S , the set $\phi^{-1}(C)$ splits into disjoint closed curves each of which is homeomorphically mapped by ϕ onto C . [This follows from Remark 6, since if there is such a map ϕ , we can collapse all the simple closed curves of $\phi^{-1}(C)$ (for each C) into arcs so as to obtain a non-trivial covering space of $f(S)$.]

THEOREM 2. *If the graph G (that is, $f(H)$) has an isthmus then M contains a projective plane P^2 or a sphere S^2 which properly crosses the interior of an isthmus of $f(H)$ at one point, and does not intersect $f(H)$ anywhere else.*

Proof. Let p be a point in the interior of an isthmus of G . Let C_j ($j = 1, 2$) be the two 3-cells in our construction of M whose boundaries are mapped onto $f(S)$. Let $\{p_{j1}, p_{j2}\}$ be the inverse image of p in C_j . Since p disconnects $f(H)$, $\{p_{j1}, p_{j2}\}$ disconnects the inverse image G_j of $f(H)$ in C_j . It follows that there exists a simple closed curve L_j in the boundary of the cell C_j , for $j = 1, 2$, such that $L_j \cap G_j = \{p_{j1}, p_{j2}\}$ and the two images of L_1 and of L_2 in $f(S)$ constitute the same figure 8 curve. Then we can span in the interiors of the two cells two or three open disks bounded by that figure 8 curve, which constitute a P^2 or an S^2 as required. ■

There are 4 edges of faces of $f(S)$ meeting along each of the 3s edges of $f(H)$. Therefore there are 3^{3s} ways of pairing those face edges and glueing each pair consistently with f . (One of them is the way they are glued in S , two other ways are those which appear in the two spheres S^2 of condition (γ) .)

THEOREM 3. *If M is simply connected then each of the 27^s ways yields a single closed 2-manifold.*

Proof. Suppose to the contrary that one of those ways of glueing gives more than one manifold and let S_0 be one of them. Then we have a piecewise

linear map

$$g : S_0 \rightarrow f(S)$$

consistent with f and such that $g(S_0)$ is a proper subset of $f(S)$. Hence $M - g(S_0)$ is connected. Then there exists a simple closed curve $C \subseteq M$ such that C properly crosses $g(S_0)$ just once. Since the parity of the number of crossings is invariant under piecewise linear homotopic deformations of C , C is not contractible in M . This contradicts the assumption of Theorem 3. ■

PROBLEM 6 (for s large enough). Is the conclusion of Theorem 3 sufficient for the vanishing of $H_1(M)$ or even the simple connectedness of M ? {If the answer was yes, this would yield an algorithm for checking if M is simply connected. I guess that the answer is no, and moreover that the latter is undecidable. [Undecidability would imply that the Poincaré Conjecture fails since, as shown by J. H. Rubinstein in 1992, the problem if M is homeomorphic to S^3 is decidable; see A. Thompson [T] and references therein.]}

COROLLARY. *If M is simply connected and K is a union of disjoint basic curves in S , then $S - K$ is connected.*

Proof. By Theorem 3. ■

REMARK 8. Let T be a finite tree in S such that $T \cap H$ is the set of vertices of H . Thus f restricted to T is a homeomorphism. We can contract T to a point p in S , and likewise contract $f(T)$ to $f(p)$ in M , and we get a new map

$$\tilde{f} : S \rightarrow M,$$

which has all the properties of f except that now there is only one point p in S at which \tilde{f} is not a local homeomorphism. However, the singularity of \tilde{f} at p is complicated, while the $2s$ singularities of f are all of the same type and they are very simple: Each vertex of $f(H)$ has a spherical neighborhood B such that $f(S) \cap B$ is homeomorphic to a cone over the curve shown in Fig. 14.

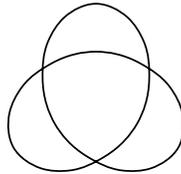


Fig. 14

REMARK 9. Let R be a spanning tree in the graph $f(H)$. Let us contract R in M to a single point. In the same way as in Remark 8 we get a new singular cellular decomposition of M with one vertex, $s+1$ edges (loops), $s+2$

faces and two 3-cells. Then the dual of this cellular decomposition has two vertices, $s+2$ regular edges, $s+1$ singular quadrilateral faces, and one 3-cell. As is well known, every closed 3-manifold can be obtained from a 3-cell B by identifying pairs of countries of a certain map Q on the boundary sphere of B . But in the case of irreducible manifolds M other than P^3 , the above dual singular decomposition of M yields a more regular construction of this kind. Namely, all the countries of Q are quadrilateral, the graph of vertices and edges of Q is 2-chromatic and the mapping identifying the countries is such that both color classes of vertices collapse into single vertices.

It remains to prove Theorem 1.

2. Proof of Theorem 1. We work in the piecewise linear category. Let Σ be the 2-dimensional skeleton of a triangulation of M . Our first task is to turn Σ into a more amenable object. We will do this by a series of contractions, i.e., modifications of the following kind. If X and Y are two closed subsets of M , we say that Y is a *contraction* of X if there exists a continuous surjection $g : M \rightarrow M$ such that $g(X) = Y$ and g restricted to $M - X$ is a homeomorphism. Then, for any subset of M containing X , its image under g will also be called a contraction. Later we will apply our modification of Σ to construct $f(H)$, $f(S)$ and f .

We pick a spanning tree of the 1-skeleton of Σ and we contract it in M to a single vertex v . Then each 2-face of the resulting singular complex Σ' is of one of the three types shown in Fig. 15.

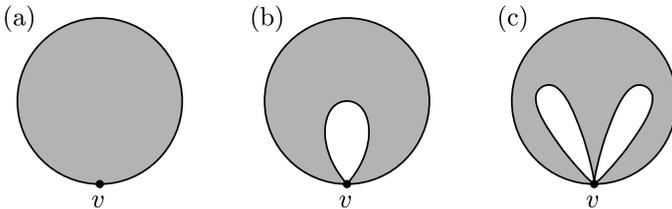


Fig. 15

Then we contract each face of type (a) to v , and repeat this operation as long as possible. So we get a complex Σ'' with 2-faces of types (b) or (c) only. Now we remove the interiors of some of the 2-faces of Σ'' until we get a complex Σ''' such that $M - \Sigma'''$ consists of one open 3-cell. This may have left us with some 2-faces having an edge which does not belong to any other 2-face. We contract those 2-faces to their remaining edges and repeat this operation as long as possible. Thus the remaining complex $\Sigma^{(4)}$ is such that $M - \Sigma^{(4)}$ is a single open 3-cell, $\Sigma^{(4)}$ has 2-faces of types (b) and (c) only. [But it may have edges (loops) belonging to none of the 2-faces.] Now we

modify $\Sigma^{(4)}$ to get rid of all edges belonging to more than three 2-faces. We do this by inserting additional 2-faces of type (b). So we get a singular complex $\Sigma^{(5)}$ such that each edge belongs to at most three 2-dimensional faces.

LEMMA 1. *If M is irreducible there exists a 2-dimensional singular complex $\Sigma_0 \subseteq M$ such that:*

- (1) Σ_0 has only one vertex v ;
- (2) all the 2-faces of Σ_0 are of type (b) or (c);
- (3) every edge of Σ_0 belongs to two or three 2-faces of Σ_0 ;
- (4) $M - \Sigma_0$ is an open 3-cell;
- (5) v does not locally disconnect Σ_0 , that is, v has arbitrarily small open neighborhoods B in M such that $(B \cap \Sigma_0) - \{v\}$ is connected.

Proof. Without using irreducibility we have already constructed a complex $\Sigma^{(5)}$ which has properties (1), (2), a weaker version of (3) where edges belonging to no faces are allowed, and (4). Notice that (5) implies that there are no such isolated edges. Thus it suffices to modify $\Sigma^{(5)}$ so as to satisfy (5) while preserving the other properties.

If $\Sigma^{(5)}$ violates (5), then there exists a 2-disk $D \subseteq M$ such that v is in the interior of D , $D \cap \Sigma^{(5)} = \{v\}$, and for every sufficiently small ball B around v both components of $B - D$ intersect $\Sigma^{(5)}$. And there also exists a disk $D' \subseteq M - \Sigma^{(5)}$ such that $D \cup D'$ is a sphere S^2 . Thus, by the irreducibility of M , this S^2 splits M into two components, and the closure of one of them is a 3-cell. We collapse this 3-cell in M to v . This removes some 2-faces of $\Sigma^{(5)}$ (which collapsed to v) and it is easy to check that the resulting 2-complex still has the properties (1)–(4) with the above weakening of (3). Repeating this operation as long as possible we obtain a Σ_0 which also satisfies (3) and (5). ■

Notice that if M is S^3 then Σ_0 can (but does not have to) be a single point (see the second part of Fig. 3).

Now we can define the graph G , and even the graph $f(H)$ in M . Namely, we choose one point in the interior of each edge of Σ_0 and one point in the interior of each 2-face of Σ_0 . In each 2-face we choose some arcs with disjoint interiors contained in the interior of the face connecting the chosen point of the interior of the 2-face with the points chosen on its edges. The union of all those arcs forms a connected graph whose vertices are the points chosen in the interiors of the faces of type (c) and those chosen in the interiors of edges which belong to three 2-faces. Clearly all those vertices are of valency 3. This is the graph $f(H)$ of our construction, which will also be the graph G . However, we must show that for all cases that matter G is not empty and does not reduce to a simple closed curve.

LEMMA 2. *If M is not homeomorphic to S^3 nor to P^3 , then G is a cubic graph.*

Proof. Since M is not S^3 , Σ_0 has some 2-faces, and hence G has edges. Since Σ_0 has no faces of type (a), it remains to show that G is not a simple closed curve. If that was the case then all 2-faces of Σ_0 would be of type (b), and Σ_0 would be a projective plane. Then, since $M - \Sigma_0$ is an open 3-cell, M would be homeomorphic to P^3 . But this is excluded by the assumption. Hence G has vertices of valency > 2 . By the structure of Σ_0 (conditions (2) and (3)) all those vertices are of valency 3. ■

In order to complete the proof of Theorem 1 it remains to construct S and f .

Proof of Theorem 1. The theorem is true for S^3 (see Remark 4). So let us assume that M is not S^3 . We choose $\Sigma_0 \subseteq M$ given by Lemma 1 and the cubic graph G (that is $f(H)$) given by Lemma 2. Now we imagine a sphere S^2 inflating around the single vertex v of Σ_0 . Let the inflation go faster along the edges of Σ_0 so that the inflating sphere engulfs these edges until it meets itself at all the points which were chosen in the interiors of the edges of Σ_0 . Then let it stop inflating along the edges but continue a faster inflation along the 2-faces of Σ_0 such that the sphere engulfs those faces until it meets itself along the arcs of $f(H)$ (which were chosen in the interiors of those faces). At this point the inflation stops. In this way our original S^2 turns into a 2-complex whose 1-skeleton is $f(H)$. This is the complex $f(S)$ required in Theorem 1. Notice that, by Lemma 1(5), the interiors of the 2-faces of $f(S)$ are simply connected, and that $M - f(S)$ consists of two open 3-cells.

In order to construct S it is enough to split $f(S)$ into its 2-faces and attach them back together by pairs along every (doubled) edge of $f(H)$ in the right way, that is, in such a way that we get a 2-manifold S and a map $f : S \rightarrow M$ which makes S cross itself along each edge. By Lemma 2 we can check all the properties of f described in Section 1. For example, since G is cubic, there is an s such that G has $2s$ vertices and $3s$ edges. Hence the number n of faces of $f(S)$ must satisfy the Euler–Poincaré equation for M , namely

$$2s - 3s + n - 2 = 0.$$

Therefore $n = s + 2$. And since 2G has $6s$ edges,

$$\chi(S) = 2s - 6s + (s + 2) = 2 - 3s.$$

Then it is easy to check that the graph H of edges of S is as required, that is, it satisfies conditions (α) , (β) and (γ) . And since (γ) implies (γ_1) it follows that S is not orientable.

This completes the proof of Theorem 1. ■

References

- [B] A. T. Balaban, *Chemical Applications of Graph Theory*, Academic Press, 1976.
- [H] J. Hempel, *3-Manifolds*, Princeton Univ. Press, 1976.
- [JP] W. Jakobsche and J. H. Przytycki, *Topology of 3-Manifolds*, Wydawnictwa Uniw. Warszawskiego, 1987 (in Polish, 2nd edition in preparation).
- [M] J. W. Milnor, *A unique decomposition theorem for 3-manifolds*, Amer. J. Math. 84 (1962), 1–7.
- [T] A. Thompson, *Algorithmic recognition of 3-manifolds*, Bull. Amer. Math. Soc. 35 (1998), 57–66.
- [Tr] B. Trace, *Two comments concerning the uniqueness of prime factorizations for 3-manifolds*, Bull. London Math. Soc. 19 (1987), 75–77.

Department of Mathematics
University of Colorado
Boulder, CO 80309-0395, U.S.A.
E-mail: jmyciel@euclid.colorado.edu

*Received 19 March 1999;
in revised form 14 April 2003*