SUTURED MANIFOLDS
AND GENERALIZED THURSTON NORMS

MARTIN SCHARLEMANN

Over the past several years, David Gabai has developed new and powerful machinery for the study of 3-manifolds ([3]–[7]). Among the long-mysterious questions he has answered are the Poenaru conjecture, the Property R conjecture, the superadditivity of knot genus, and property P for satellite knots. It is the intention here to give an account of these developments, starting at the very beginning.

This began as an expository paper. It first became apparent that a generalization of the Thurston norm would be useful in such an exposition when it provided a direct alternative to an awkward doubling argument in [3]. Eventually it became clear that the theory of sutured manifolds could be developed from a different perspective, a perspective which is combinatorial, uses no foliations, is perhaps more general (cf. the latitude in the definition of index in 7.4), and seems to recover and extend (cf. [12]) the principal topological results. The absence of foliations simplifies the proof at several points and is perhaps interesting in its own right.

All of the central results of the paper were proven first by Gabai, with the exception of superadditivity of knot genus. This last theorem we proved simultaneously (announcements crossed in the mail). Gabai's proof uses the machinery of [4] (§5 here) and a trick of A. Thompson [13]. The proof here is an illustration of the use of β-norms for β a 1-complex more complicated than a circle (in fact an "eyeglass").

The paper is organized as follows: §§0 and 1 establish notation and describe the generalized Thurston norm. §2 describes sutured manifold decompositions and show that β-taut decompositions can be built on any relative homology class. §3 is essentially a careful proof of [3, Lemma 3.5]. Gabai's original theorem is only sketchily proved and the hypotheses incomplete. Since this theorem is central to the development here (it substitutes for foliations), we give a complete statement and detailed proof. §4 proves that sutured manifold

hierarchies always exist, via the notion of complexity of a sutured manifold. Thus §§0–4 can be viewed as parallel to [3].

§5 shows that satellite knots of trivial winding number have property P. Since later we show that all satellite knots have property P this section is in some sense redundant. It is included because it is a beautiful illustration of the technique to be used, and is particularly simple because β is trivial and parametrizing surfaces are unnecessary. It is parallel to [4].

§6 contains a set of "reducing lemmas" which give criteria for simplifying the 1-complex β. §7 introduces "parametrizing surfaces" which are used to carry topological information down through a hierarchy.

§§8–10 are the main applications. §8 is a proof (somewhat generalized) of the superadditivity of knot genus under band sum. The same result is obtained in [6], though the proof is quite different. §9 is a proof that all satellite knots have property P. The combinatorics involved here (in some-what different form) originated with Gabai's proof of the Poenaru conjecture [5]. The application to property P is more natural, however, because of a far more obvious parametrizing surface. Finally §10 sketches how a parametrizing surface can be found sufficient to prove the Poenaru conjecture from the combinatorics of §9.

Throughout we have made no attempt to state results in their most general form, for these generalizations can be found in [3], [4], [5]. The emphasis here is rather to generalize technique and to pick particularly interesting applications as illustrations of the technique.

A reader with only a casual interest would do well to read §§0, 1, 2 (except 2.5) 3.5, 3.9, 4.18–4.19, and 5.

Topology seminars in Jerusalem and Austin and individual topologists too numerous to list have spotted errors and made suggestions on this manuscript. Particularly helpful was Cameron Gordon's discovery of a subtle combinatorial error in §9 whose correction simplified the argument quite a bit. I am grateful to the topologists at the Instituto de Matematicas, U.N.A.M., who provided a patient and attentive audience for the first exposition of this viewpoint in December 1985.

0. Preliminaries and notation

$(M, \partial M)$ will always be an oriented 3-manifold with boundary. β will denote a 1-complex contained in M. The collection of all valence 1 vertices of β is denoted $\partial\beta$ and the inclusion of β in M is proper, meaning $\partial M \cap \beta = \partial\beta$. We will henceforth reserve the term vertex of β for those vertices not in $\partial\beta$

(hence in the interior of M). The set of vertices of β are denoted $\mathscr{V}_\beta$ and the set of edges $\mathscr{E}_\beta$.

$\mathcal{N}(\beta)$ will denote a regular neighborhood of β, which we view as the union of 3-ball neighborhoods (0-handles) around the vertices, denoted $\mathcal{N}(\mathscr{V}_\beta)$, and 1-handles along the edges, denoted $\mathcal{N}(\mathscr{E}_\beta)$. In general, for any proper submanifold $(P, \partial P)$ of $(M, \partial M)$, $\mathcal{N}(P)$ denotes a relative regular neighborhood and $\dot{\mathcal{N}}(P)$ its boundary as a regular neighborhood. Similarly $\dot{\mathcal{N}}(\beta)$ denotes the standard handlebody boundary of $\mathcal{N}(\beta)$.

Each 1-handle in $\mathcal{N}(\mathscr{E}_\beta)$ can be viewed as a copy of $I \times D^2$. A disc $\{p\} \times D^2$ is called a *meridian disk*. Similarly $\{p\} \times \partial D^2$ is called a meridian circle.

Proper surfaces in $(M, \partial M)$ will always be regarded as oriented and in general position with each other and with β. In particular such proper surfaces are disjoint from $\mathcal{N}(\mathscr{V}_\beta)$ and intersect $\mathcal{N}(\mathscr{E}_\beta)$ only in meridian disks. Since our surfaces will always be oriented, by choosing an arbitrary orientation of an edge e of β it makes sense to say that two points of intersection of S with e have the same (or opposite) sign.

Denote by M_β the closed complement of $\mathcal{N}(\beta)$ in M. A proper surface $(Q, \partial Q) \subset (M_\beta, \partial M_\beta)$ will mean a surface in *normal position* with respect to the handle structure on $\mathcal{N}(\beta)$. That is, ∂Q intersects the 0-handle boundaries $\dot{\mathcal{N}}(\mathscr{V}_\beta)$ in arcs and circles which are essential in the complement of $\mathcal{N}(\mathscr{E}_\beta)$, and ∂Q intersects each 1-handle boundary $I \times \partial D^2 \subset \dot{\mathcal{N}}(\beta)$ in $I \times \{\text{points}\}$. In particular any component of ∂Q lying in $\dot{\mathcal{N}}(\mathscr{E}_\beta)$ and isotopic to a meridian circle is isotoped off $\dot{\mathcal{N}}(\mathscr{E}_\beta)$ so that it lies in either $\dot{\mathcal{N}}(\mathscr{V}_\beta)$ or in ∂M.
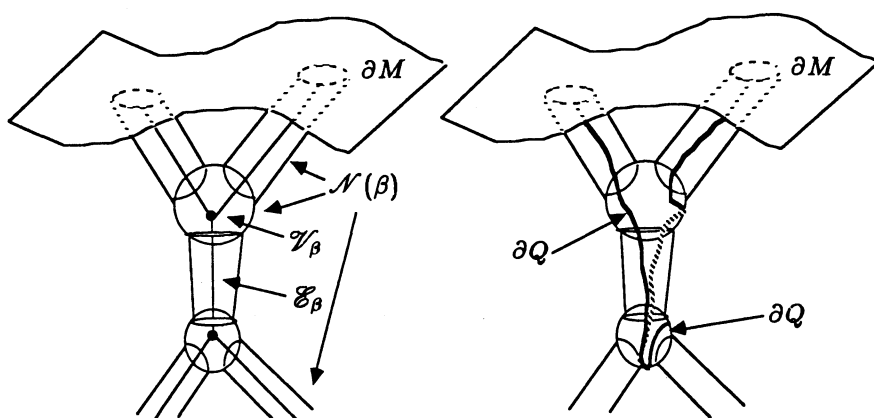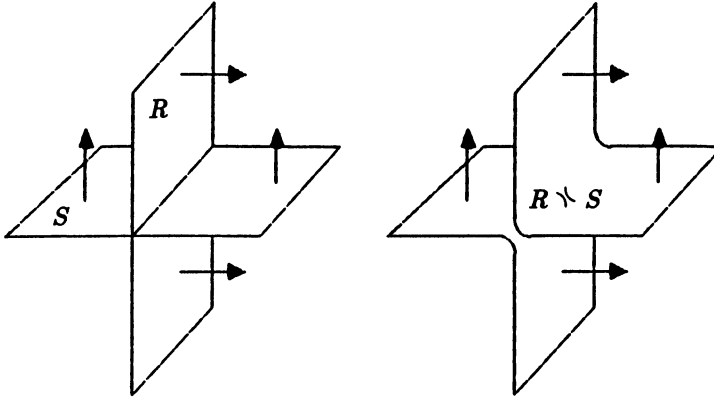

FIGURE 0.1

FIGURE 1.1. Local picture of double-curve sum

If P is a manifold (of dimension 0, 1, or 2 in practice), then $|P|$ denotes the number of components of P (thus $|P|$ is the cardinality of P if P is a finite collection of points).
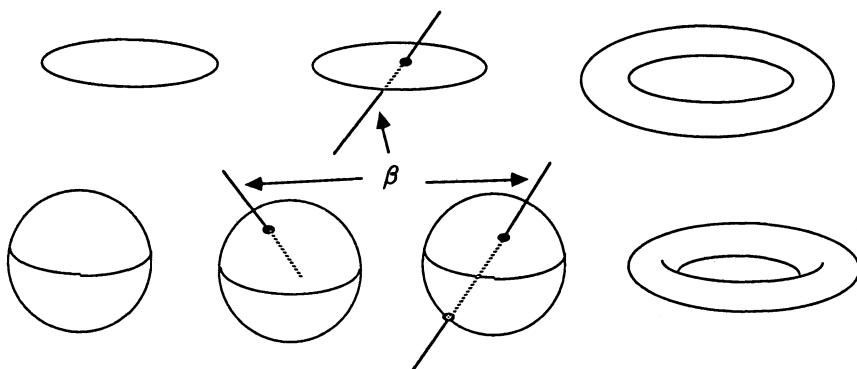
1.   Generalized Thurston norms

1.1 Definition.   Suppose $(S, \partial S)$ and $(R, \partial R)$ are two surfaces in general position in $(M, \partial M)$. Define $R \curlyvee S$ to be the "double-curve sum" of R and S (i.e., cut and paste along intersection curves and arcs to get an oriented surface representing the cycle $R + S$). $R \curlyvee S$ is an imbedded oriented surface coinciding with $R \cup S$ outside of a regular neighborhood of the double curve set.

Let $(M, \partial M)$ be a compact, oriented 3-manifold, N a subsurface of ∂M, and β a *properly imbedded* 1-*complex* in M. Let χ denote the Euler characteristic. For a connected surface $(S, \partial S) \subset (M, N)$ in general position with respect to β define

$$\chi_\beta(S) = \max(0, |S \cap \beta| - \chi(S)).$$

When S is not connected, define $\chi_\beta(S)$ to be the sum of $\chi_\beta(S_i)$ over all the components S_i of S. An alternative definition in terms of the Euler characteristic χ is:

$$\chi_\beta(S) = |S \cap \beta| - \chi(S) + |\text{disks in } S \text{ disjoint from } \beta|$$
$$+ |\text{spheres in } S \text{ intersecting } \beta \text{ in one point}|$$
$$+ 2 \cdot |\text{spheres in } S \text{ disjoint from } \beta|.$$

FIGURE 1.2. Examples of $\chi_\beta = 0$

Finally, define a function $x_\beta \colon H_2(M, N) \to \mathbf{Z}_+$ by

$$x_\beta(a) = \inf\{\chi_\beta(S) | S \text{ is an imbedded surface representing } a\}.$$

1.2 Definition.   Let $(S, \partial S)$ be a properly imbedded surface in $(M, \partial M)$.

(a) S is β-*minimizing* in $H_2(M, N)$ if $x_\beta[S, \partial S] = \chi_\beta(S)$.

(b) An arc or simple closed curve in S is β-*essential* if it is essential in the punctured surface $S - \beta$.

(c) If there is a β-essential simple closed curve C in S which bounds a disk in $M - S$ disjoint from β, then S is called β-*compressible*. Otherwise S is β-*incompressible*.

(d) S is β-*taut* if it is β-incompressible, β-minimizing in $H_2(M, \mathcal{N}(\partial S))$ and any given edge of β always intersects S with the same sign. If $\beta = \varnothing$ we say simply that S is *taut* (in the Thurston norm).

(e) M is β-*irreducible* if any 2-sphere disjoint from β in M bounds a 3-ball disjoint from β.

(f) S is N_β-*compressible* if there is a disk D in $M - \beta$ for which $D \cap (\partial M \cup S) = \partial D$ is the union of an arc in N and a β-essential arc in S.

1.3 Lemma.   (a) $\chi(R \searrow S) = \chi(R) + \chi(S)$.

(b) *If $(S, \partial S)$ is β-minimizing in $H_2(M, N)$, then it is β-minimizing in $H_2(M, \mathcal{N}(\partial S))$ but not vice versa.*

(c) $x_\beta(a + b) \leq x_\beta(a) + x_\beta(b)$.

(d) *Suppose $(R, \partial R) \subset (M, N)$ is β-incompressible and N_β-incompressible. Then any surface $(S, \partial S) \subset (M, N)$ can be modified by 2-surgery, without increasing $\chi_\beta(S)$, so that, for R' any collection of parallel copies (identically oriented) of components of R,*

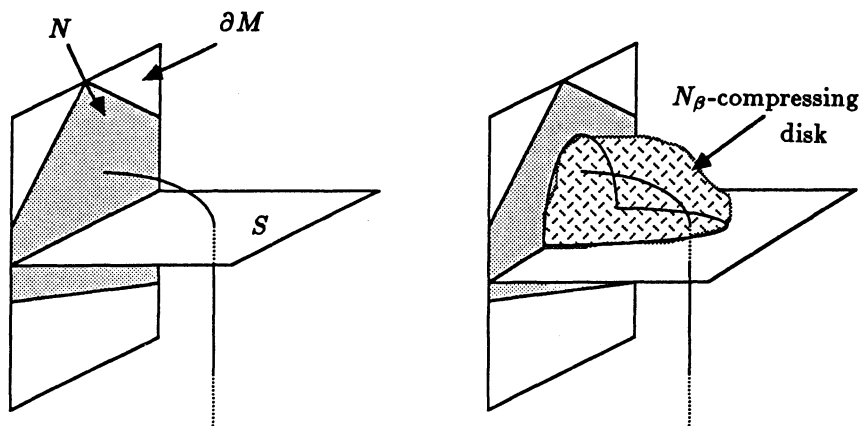$$\chi_\beta(S \searrow R') = \chi_\beta(S) + \chi_\beta(R').$$

FIGURE 1.3

(e) *Suppose $(R, \partial R) \subset (M, N)$ is β-taut. Then so is every collection of components of R.*

Proof. (a) $R \curlywedge S$ is obtained from the disjoint union of R and S by removing a collection of arcs and circles and sewing them back in differently. Hence χ is unchanged.

(b) Follows from the homomorphism $H_2(M, \mathscr{N}(\partial S)) \to H_2(M, N)$. Figure 1.4 is an example of a surface S β-minimizing in $H_2(M, \mathscr{N}(\partial S))$ but not in $H_2(M, N)$. In fact $\chi_\beta(S) = 1$ while $\chi_\beta(S') = 0$.
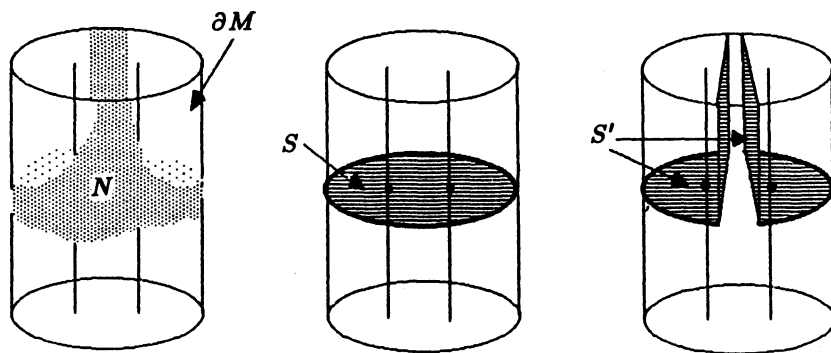

FIGURE 1.4

To prove (c) and (d), consider surfaces $(R, \partial R)$ and $(S, \partial S) \subset (M, N)$ in general position. Suppose no component of $R \cap S$ is β-inessential in R or S. Then it is easy to see that $\chi_\beta(R \curlywedge S) = \chi_\beta(R) + \chi_\beta(S)$.

To prove (c) choose β-minimizing representatives R and S of a and b. Repeatedly do 2-surgery on disks cut off by innermost β-inessential circles or outermost β-inessential arcs of $R \cap S$. This never increases χ_β. Eventually R and S intersect only in components β-essential in both. Then $x_\beta(a+b) \leq \chi_\beta(R \searrow S) = \chi_\beta(R) + \chi_\beta(S) = x_\beta(a) + x_\beta(b)$.

To prove (d) first do 2-surgery on S so that no circle of $R \cap S$ is β-inessential in R. Since R is β-incompressible, no circle of $R \cap S$ (hence $R' \cap S$) is β-inessential in S. Similarly alter S so no arc of $R \cap S$ is β-inessential in R (e.g. in Figure 1.3 think of S as the horizontal surface and the N_β-compressing disk as lying in R). Since R is N_β-incompressible, no arc of $R \cap S$ (hence $R' \cap S$) is β-inessential in S. Then $\chi_\beta(R' \searrow S) = \chi_\beta(R') + \chi_\beta(S)$.

(e) Suppose R is β-taut, let R' be a collection of components of R and $R'' = R - R'$. Clearly R' is β-incompressible. Suppose R' were not β-minimizing, and let T be a β-taut surface with $[T, \partial T] = [R', \partial R']$ in $H_2(M, \mathcal{N}(\partial R'))$ and ∂T parallel to but disjoint from $\partial R'$. Since R'' is β-incompressible, it is N_β-incompressible for N, the annuli $\mathcal{N}(\partial R'')$. Modify T as in (d) so that $\chi_\beta(T \searrow R'') = \chi_\beta(T) + \chi_\beta(R'') < \chi_\beta(R') + \chi_\beta(R'') = \chi_\beta(R)$. This contradicts the tautness of R.

1.4 Theorem.   (a) x_β *extends to a canonical continuous function* $x_\beta \colon H_2(M, N; R) \to R_+$ *which is convex, and is linear on rays through the origin.*

(b) *Any element of $H_2(M, N; \mathbb{Z})$ has a β-minimizing, β-incompressible representative $(S, \partial S)$ such that all intersections of S with a given edge of β have the same sign.*

Proof.   (a) This follows from Lemma 1(a) and [14, Lemma 1].

(b) Let $(S, \partial S)$ be a β-minimizing representative of a class in $H_2(M, N)$. Among all such β-minimizing representatives choose one with a minimal number of intersections with β. Then S intersects any given edge of β with the same orientations, for otherwise a 1-surgery on S would remove two points of $\partial S \cap \beta$.

Finally, β-compress S as much as possible. Neither operation increases $\chi_\beta(S)$.   q.e.d.

Let $N \subset \partial M$ and x_β be as in 1.1.

1.5 Corollary.   *For $a, b \in H_2(M, N)$ there is an $m \geq 0$ such that for all $k \geq 0$*

$$x_\beta((a + mb) + kb) = x_\beta(a + mb) + kx_\beta(b).$$

Proof.   Consider the function $\varphi \colon \mathbb{Z} \to \mathbb{Z}$ given by
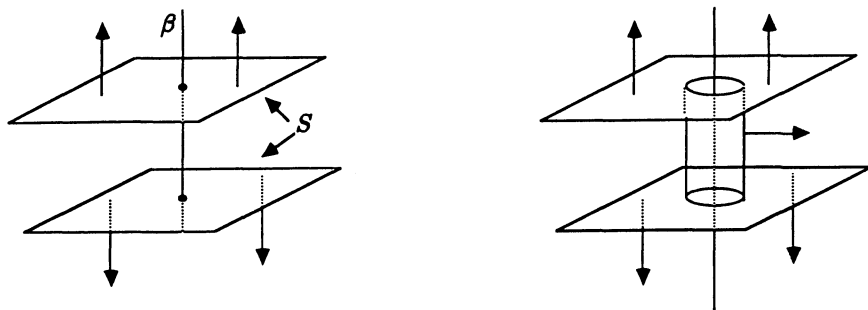
$$n \to x_\beta(a + (n+1)b) - x_\beta(a + nb).$$

FIGURE 1.5. Removing intersection points of opposite sign

Since $x_\beta(a + (n+1)b) + x_\beta(a + (n-1)b) \geq 2x_\beta(a + nb)$ by convexity, φ is a nondecreasing function into the integers no larger than $x_\beta(b)$. Hence eventually, say for $n \geq m$, φ is a constant c. Then for all $k \geq 0$

$$x_\beta((a + mb) + kb) = x_\beta(a + mb) + kc.$$

Moreover, $x_\beta(b) = \lim_{k \to \infty} [x_\beta((a + mb) + kb) - x_\beta(a + mb)]/k = c.$ q.e.d.

Now suppose that β^+ is the union of β and a collection δ of proper arcs in $(N, \partial N)$. (Regard δ as pushed slightly into M.)

1.6 Lemma.   *Suppose M is β-irreducible. If $(S, \partial S) \subset (M, N)$ is β^+-taut then it is β-taut.*

Proof.   First we show that S is β-incompressible. Suppose a disk $(D, \partial D)$ disjoint from β intersects S precisely in ∂D. Isotope D slightly in $S - \beta$ so that it lies entirely in the interior of M. Then D is disjoint from δ, hence from β^+. Since S is β^+-taut, ∂D bounds a disk in S disjoint from β^+, hence β.

Next we show that S is β-minimizing. Suppose $(S', \partial S')$ is β-minimizing for $[S, \partial S]$ in $H_2(M, \mathcal{N}(\partial S))$. By possibly capping off annuli in $\mathcal{N}(\partial S) - \partial S'$, which does not change $\chi_\beta(S')$, we can assume that $\partial S' = \partial S$ and, in particular, that the points of $(S' \cap \beta^+)$ are those of $(S' \cap \beta)$ together with points of $\partial S' \cap \delta = \partial S \cap \delta$. Now β-compress S' until it is β-incompressible. Then any component of $\partial S'$ bounding a disk in $M - \beta$ bounds a disk in S'.

Then since $\chi_{\beta+}(S') \geq \chi_{\beta+}(S)$ we have

$$\chi_\beta(S') = \chi_{\beta+}(S') - |\partial S' \cap \delta| + |\text{disks in } S' \text{ disjoint from } \beta$$
$$\text{whose boundary intersects } \delta|$$
$$\leq \chi_{\beta+}(S) - |\partial S \cap \delta| + |\text{disks in } S \text{ disjoint from } \beta$$
$$\text{whose boundary intersects } \delta|$$
$$= \chi_\beta(S).$$

Thus S is β-minimizing.

2. Sutured manifolds

2.1 Definition. A sutured manifold (M, γ, β) is a compact oriented 3-manifold M, containing the proper 1-complex β, with ∂M decomposed into the union along the boundary of two surfaces $\overline{R}_+$ and $\overline{R}_-$ oriented so that $\partial \overline{R}_+ = \partial \overline{R}_- = \gamma$ and so that $\partial M = \overline{R}_+ - \overline{R}_-$. Let $A(\gamma)$ denote a collection of disjoint annuli comprising a regular neighborhood of γ, let $T(\gamma)$ be some collection of torus components of $\overline{R}_+$ and $\overline{R}_-$ and define $R_\pm = \overline{R}_\pm - \overset{\circ}{A}(\gamma) - T(\gamma)$. Then $\partial M = R_+ \cup R_- \cup A(\gamma) \cup T(\gamma)$. We regard R_+ as the set of components of $\partial M - (A(\gamma) \cup T(\gamma))$ whose normal vectors point out of M, and R_- as those whose normal vectors point into M. The symbol $R_\pm$ will denote "R_+ or R_- respectively" while $R(\gamma)$ denotes $R_+ \cup R_-$. Frequently we will abbreviate (M, γ, β) to simply (M, γ).

2.2 Definition. (M, γ) is β-taut if
(a) the ends of β are all disjoint from $A(\gamma)$ and $T(\gamma)$,
(b) $T(\gamma)$ and $R(\gamma)$ (equivalently R_+ and R_-) are β-taut,
(c) M is β-irreducible.

Remarks. (b) and 1.2(c) imply that no edge in β has both ends in R_+ or both ends in R_-.

β-tautness is unaffected by the choice of which torus components of $R_\pm$ disjoint from β are counted as lying in $T(\gamma)$, since such components have trivial norm.

2.3 Definition. A d(ecomposing)-*surface* $(S, \partial S) \subset (M, \partial M)$ is an oriented properly imbedded surface which intersects each $T(\gamma)$ in coherently-oriented parallel essential circles and which intersects each component of $A(\gamma)$ either in circles parallel to γ (and oriented the same way) or in essential arcs (not necessarily oriented coherently). Then there is a natural sutured manifold structure on $M' = M - \mathcal{N}(S)$ (see [3, 3.2]). Denote by β' the 1-complex $\beta - \mathcal{N}(S)$. This process of constructing a sutured manifold structure on M' from one on M is called a (β-taut) *sutured manifold decomposition* (if $M(\gamma, \beta)$

is β-taut and (M', γ', β') is β'-taut). It is often helpful to continue to refer to β' as β. Note that in M', $R_\pm(\gamma')$ is parallel to $R_\pm(\gamma) \asymp S$. Figure 2.1 is a local picture near ∂M away from γ, while Figure 2.2 is one near $A(\gamma)$.
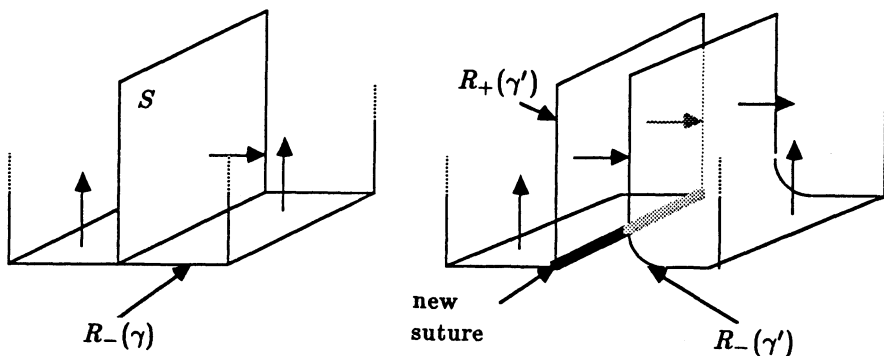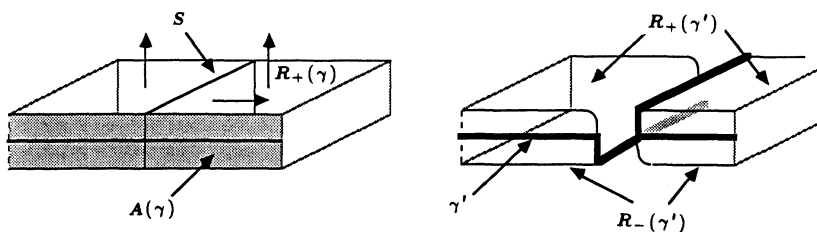

FIGURE 2.1

FIGURE 2.2

2.4 Definition. A *conditioned* surface $S \subset (M, \partial M)$ is a properly imbedded d-surface $(S, \partial S) \subset (M, \partial M)$ such that

(a) all arcs of ∂S in any annulus component of $A(\gamma)$ are oriented in the same direction,

(b) no collection of simple closed curves of $\partial S \cap R(\gamma)$ is trivial in $H_1(R(\gamma), \partial R(\gamma))$,

(c) any given edge e of β which intersects $S \cup R(\gamma)$ always intersects them with the same sign.

2.5 Theorem. *Given*:

(a) *a β-taut surface $(R, \partial R)$ in a β-irreducible 3-manifold $(M, \partial M)$,*

(b) *a properly imbedded family of C of oriented arcs and circles in $\partial M - \mathcal{N}(\partial R)$ which is in* ker: $H_1(\partial M, \mathcal{N}(\partial R)) \to H_1(M, \mathcal{N}(\partial R))$,

(c) *y in $H_2(M, \partial M)$ such that $\partial y = [C]$.*

Then there is a surface $(S, \partial S)$ *in* $(M, \partial M)$ *such that*

   (i) $\partial S - \mathcal{N}(\partial R) = C$,

  (ii) *for some integer* m, $[S, \partial S] = y + m[R, \partial R]$ *in* $H_2(M, \partial M)$,

 (iii) *for any collection* R' *of parallel copies of components of* R *(similarly oriented),* $S \setminus R'$ *is* β-*taut,*

 (iv) *any edge of* β *which intersects both* R *and* S *intersects them with the same sign.*

Proof. Let N be a regular neighborhood of $\partial R \cup C$ in ∂M. Let δ be a family of properly imbedded arcs in N so that $N - \delta$ deformation retracts to the disjoint union of ∂R and the circle components of C. Let δ^+ be δ pushed into M and $\beta^+ = \beta \cup \delta^+$.
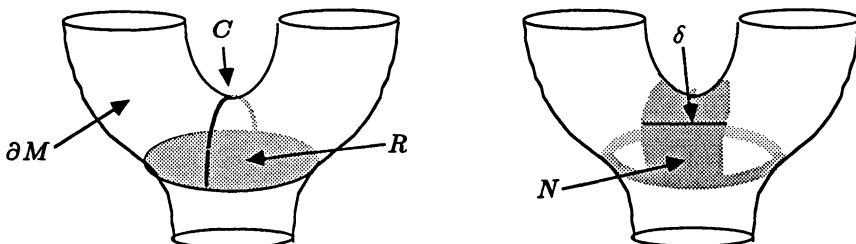

FIGURE 2.3

Claim 1. R is β^+-*incompressible,* N_β-*incompressible, and* β^+-*minimizing in* $H_2(M, N)$.

Proof of Claim 1. Since R is β-incompressible and disjoint from δ^+, it is clearly β^+-incompressible.

To show R is β^+-minimizing consider $(R', \partial R')$, a β^+-minimizing representative of $[R, \partial R]$ in $H_2(M, N)$. Since $[\partial R'] = [\partial R]$ in $H_1(N)$, it follows from 1.4(b) that we can assume R' and δ^+, hence $\partial R'$ and δ, are disjoint and, after possibly attaching annuli of $N - \partial R'$, that $\partial R' = \partial R$. Then since R is β-taut, $\chi_{\beta^+}(R) = \chi_\beta(R) \leq \chi_\beta(R') = \chi_{\beta^+}(R')$. Hence R is β^+-minimizing in $H_2(M, N)$.

To show R is N_{β^+}-incompressible, suppose there was an essential arc σ in $R - \delta^+$ which, together with an arc τ of N, bounded a disk disjoint from β^+. Since R is β^+-minimizing in $H_2(M, N)$, 2-surgery on the disk could not reduce $\chi_{\beta^+}(R)$ so the component R' of R which contains σ must be an annulus disjoint from β^+ or a disk intersecting β^+ once. The latter is impossible since σ is essential. But the former implies two boundary components of R' can be

connected by an arc in N, which is disjoint from δ. Since only one component of ∂R lies in each annular region of $N - \delta$, this, too, is impossible.

Claim 2.   There is a class z in $H_2(M, N)$ such that:

(a) $\partial z = [C]$ in the exact sequence of the triple $(M, \partial M, \mathscr{N}(\partial R))$,

(b) $x_{\beta^+}(z + [R', \partial R']) = x_{\beta^+}(z) + x_{\beta^+}([R', \partial R'])$ for any collection R' of parallel copies of components of R (similarly oriented),

(c) for some integer m, $z = y + m[R, \partial R]$ in $H_2(M, \partial M)$.

Proof of Claim 2.   Since N is defined so that $[C]$ is in $H_1(N, \mathscr{N}(\partial R))$, and y is in $H_2(M, N)$. Now apply 1.5 to y and $[R, \partial R]$ and let $z = y + m[R, \partial R]$. Choose k large enough so that $R' \subset kR$, and let $R'' = kR - R'$. Then the choice of z guarantees that $x_{\beta^+}(z) + kx_{\beta^+}([R, \partial R]) = x_{\beta^+}(z + k[R, \partial R])$. We have

$$
\begin{aligned}
x_{\beta^+}&(z) + kx_{\beta^+}([R, \partial R]) \\
&= x_{\beta^+}(z + k[R, \partial R]) \le x_{\beta^+}(z + [R', \partial R']) + x_{\beta^+}([R'', \partial R'']) \\
&\le x_{\beta^+}(z) + x_{\beta^+}([R', \partial R']) + x_{\beta^+}([R'', \partial R'']) \\
&= x_{\beta^+}(z) + kx_{\beta^+}([R, \partial R]),
\end{aligned}
$$

where the last equality follows from 1.3(e) and the fact that R is β^+-taut. Thus all the inequalities are equalities and

$$
x_{\beta^+}(z + [R', \partial R']) = x_{\beta^+}(z) + x_{\beta^+}([R', \partial R']).
$$

Now let $(S', \partial S') \subset (M, N)$ be a β^+-minimizing surface realizing z. If any pair of circle components of $S' \cap R$ are anti-parallel in R (i.e., bound an annulus in $R - \beta$), then by cutting and pasting along parallel copies of the annulus, S' may be modified to remove these circle components, without altering $\chi_{\beta^+}(S')$. Thus we can assume no circle components are anti-parallel. By Claim 1, we can further modify S' with respect to R as in 1.3(d) so that every component of $R \cap S'$ is β^+-essential in both R and S'. Finally, S' can be made β^+-incompressible by β^+-compressions disjoint from R, since R is β^+-incompressible and M is β-irreducible; hence β^+ irreducible. Since $[\partial S']$ restricts to $[C]$ in $H_1(N, \mathscr{N}(\partial R))$, we can assume (1.4(b)) that $\partial S' - \mathscr{N}(\partial R) = C$. Furthermore

$$
\begin{aligned}
\chi_{\beta^+}(S' \searrow R') &= \chi_{\beta^+}(S') + \chi_{\beta^+}(R') = x_{\beta^+}(z) + x_{\beta^+}([R', \partial R']) \\
&= x_{\beta^+}(z + [R', \partial R']) = x_{\beta^+}([S' \searrow R'])
\end{aligned}
$$

by Claim 2(b). Hence $S' \searrow R'$ is β^+-minimizing.

In fact $S' \searrow R'$ is also β^+-incompressible. For suppose T is a component of $S' \searrow R'$, properly isotoped in (M, N) so that ∂T has minimal intersection with ∂R. If T is β^+-compressible then, since T is β^+-minimizing, $\chi_{\beta^+}(T) = 0$. Since T is a component of $S' \searrow R'$, it is made up of components $\{U\}$ of $S' - R'$

and $R' - S'$ glued together along curves of $R' \cap S'$. By construction each such curve is β^+-essential in both R' and S', and no pair of closed curves are anti-parallel in $R' - \beta^+$. Since $\chi_{\beta^+}(T) = 0$, each component of $\{U\}$ must be a β^+-incompressible annulus disjoint from β^+ or a β^+-incompressible disk intersecting β^+ once. Possibly:

(1) T is a twice-punctured sphere or once-punctured disk. Then a β^+-compression of T would give a β^+-compression for the component of U containing a point of β^+, which is impossible.

(2) T is an annulus disjoint from β^+. Then if T is β^+-compressible so are the annular components of U it contains, hence the closed curves of $R \cap S$ which comprise their boundary. This is impossible since all such curves are β^+-essential.

(3) T is a β^+-compressible torus disjoint from β^+.

First note that T is contained in no 3-ball B^3 disjoint from β^+. If $T \subset R'$ or $T \subset S'$ this is obvious. If not, then suppose $T \subset B^3$ and let α be a circle component of $R' \cap S'$ in T. On the one hand α is β^+-essential in R' and S' by construction; on the other hand, any simple closed curve lying in $R' \cap B^3$ (including α) is β^+-inessential in R' since R' is β^+-incompressible.

Now view T as the double curve sum of copies of R' and S' pushed slightly off in the direction of their normal vectors, so any curve in $T \cap R'$ (resp. $T \cap S'$) is parallel in R' (resp. S') to a curve in $R' \cap S'$. Since M is β^+-irreducible, T bounds a solid torus W disjoint from β^+. Since R' is β^+-incompressible, each component of $R' \cap W$ is an annulus, with boundary components (oriented as $R' \cap T$) anti-parallel in R'. Since no two components of $R' \cap S'$ are anti-parallel in R' (by construction), it follows that R' is disjoint from W. Then any component of $\{U\}$ which came from R' would be an annulus whose boundary comes from anti-parallel curves of intersection with S', again contradicting the construction of S'. Finally, T cannot lie entirely in S' because S' is β^+-incompressible.

Choose k so large that any edge of β^+ which intersects $S' \searrow kR$ intersects it *algebraically* with the same sign that it intersects R. Tube together points of intersection of opposite sign on any given edge to get a β^+-minimizing surface S for which 2.5(i), (ii) and (iv) are satisfied. Now repeat the argument above on S instead of S', modifying S so that $S \searrow R'$ is β^+-taut. Then $S \searrow R'$ is β-taut by 1.6.

2.6 Theorem. *Let (M, γ) be a β-taut sutured manifold and let y be a nontrivial element of $H_2(M, \partial M)$. Then there exists a conditioned surface $(S, \partial S) \subset (M, \partial M)$ containing no closed separating component such that $[S, \partial S] = y$ in $H_2(M, \partial M)$, and the sutured manifold decomposition $(M, \gamma) \xrightarrow{S} (M', \gamma')$ is β-taut.*

Proof. Since (M, γ) is a β-taut sutured manifold, $R(\gamma)$ is β-taut. The class ∂y in $H_1(\partial M, A(\gamma))$ can be represented by a family of oriented arcs and circles $(C, \partial C)$ in $R(\gamma)$ so that all arcs intersect a given component of $\partial A(\gamma)$ with the same orientation, and no collection of circles is null-homologous in $R(\gamma)$. Let S'' be the surface from 2.5 applied to C and $R(\gamma)$. After possibly capping off oppositely oriented components of $\partial S''$ in $A(\gamma)$ and $T(\gamma)$ with annuli, S'' becomes a conditioned surface. Discard any closed separating components and call the result S. Since the fundamental class of $R_\pm(\gamma)$ is trivial in $H_2(M, \partial M)$, $[S, \partial S] = y$ in $H_2(M, \partial M)$. By 2.5 and 1.3(e) $R_+(\gamma) \searrow S$ and $R_-(\gamma) \searrow S$ are β-taut in M, hence they are β-taut in $M' = M - S$. This verifies 2.2(b) for $R(\gamma')$.

To verify 2.2(c) note that a sphere in $M - S$ disjoint from β bounds a 3-ball in M, which is disjoint from β. That 3-ball is disjoint from S since no component of S is null-homologous.

3. β-Tautness usually pulls back

Here we ask the following question: Suppose $(M, \gamma, \beta) \xrightarrow{S} (M', \gamma', \beta')$ is a sutured manifold decomposition for which (M', γ') is β'-taut. When is (M, γ) β-taut? Surprisingly, it almost always is.

3.1 Definition. For S a connected surface, $(\Lambda, \partial\Lambda) \subset (S, \partial S)$ a properly imbedded 1-manifold whose fundamental class is trivial in $H_1(S, \partial S)$, and x_0 a point in $S - \Lambda$ define the *potential* $\Xi_\Lambda: S - \Lambda \to \mathbb{Z}$ to be the function which assigns to any y in $S - \Lambda$ the algebraic intersection with Λ of a path from x_0 to y. A different choice of path will make no difference in Ξ_Λ since the algebraic intersection of a cycle with Λ is always trivial. The *variance* ξ_Λ of Λ is the maximum, over all y, z in $S - \Lambda$, of $|\Xi_\Lambda(y) - \Xi_\Lambda(z)|$ and is clearly independent of x_0.

3.2 Lemma. *Suppose $(T, \partial T)$ and $(S, \partial S)$ are properly imbedded surfaces in general position in the 3-manifold $(M, \partial M)$, and T is homologous rel ∂ to a surface in M, which is disjoint from the interior of S. Then*

(a) $\Lambda = T \cap S$ *has trivial fundamental class in $H_1(S, \partial S)$,*

(b) *for k the variance of Λ in S, $T \searrow kS$ is properly isotopic to a surface disjoint from S.*

Proof. (a) follows from standard homology theory: Λ is homologous in S to a family of arcs in ∂S.

To prove (b), let C_0 and C_k be respectively the components of $S - \Lambda$ of the lowest and highest potential. Let $\mathcal{N}(S)$ be a regular neighborhood of S in M, viewed as a bicollar of S. Then after isotoping the image of C_0 in $T \searrow S$

to one end of the bicollar $\mathcal{N}(S)$ and the image of C_k in $T \curlywedge S$ to the other (see Figure 3.1), the variance of the 1-manifold $(T \curlywedge S) \cap S$ will be $k - 1$. Repeat k times; then the variance is trivial, hence the surfaces are disjoint.
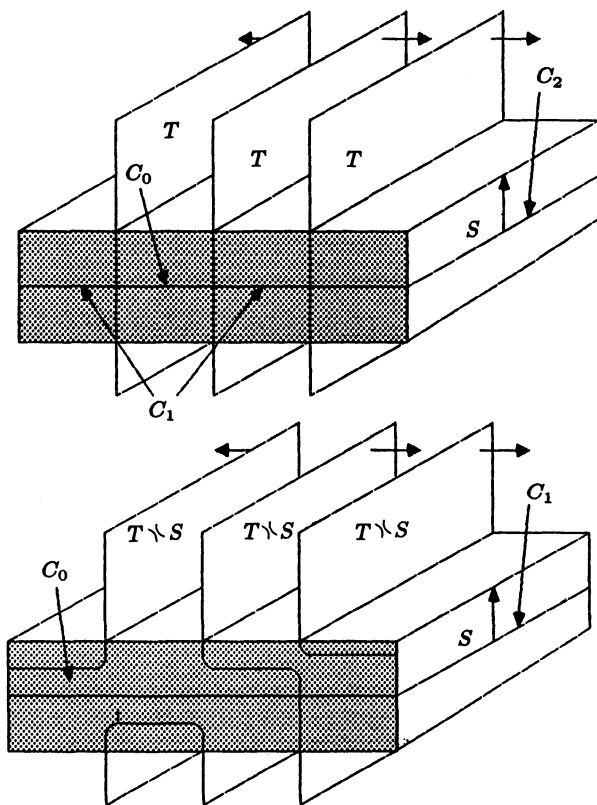

FIGURE 3.1

3.3 Corollary. *Suppose $(S, \partial S) \subset (M, \partial M)$ is a surface in general position with respect to the properly imbedded 1-complex $\beta \subset M$. Let M' be the complement in M of a bicollar $\mathcal{N}(S)$ of S with ends $S_\pm$, and $\beta' = \beta - \mathcal{N}(S)$. If S_+ and S_- are β'-taut in M', then S is β-taut in M.*

Proof. If S were β-compressible, then a compressing disk would have interior disjoint from S and S_+, or S_- would be β-compressible in M'. Hence S is β-incompressible.

If an edge e of β intersects S in points of opposite sign, then the subarc of e between those points would have both ends on S_+ (or both on S_-) contradicting the β'-tautness of S_+ (or S_-). Thus we can assume all points of intersection with S of any given edge of β have the same sign.

It remains to check that if S_+ and S_- are β-norm minimizing, so is S. Therefore suppose $(S', \partial S') \subset (M, \mathcal{N}(\partial S))$ is homologous to S in $H_2(M, \mathcal{N}(\partial S))$. We need to show that $\chi_\beta(S') \geq \chi_\beta(S)$.

Claim. If $S' \cap S = \varnothing$, then $\chi_\beta(S') \geq \chi_\beta(S)$.

Proof of Claim. We will assume S is closed, for the general case is similar. Let $\overline{S}$ denote the oriented surface $S \cup -S'$. Then $[\overline{S}]$ is trivial in $H_2(M)$. Suppose C is a component of $M - \overline{S}$ which can be oriented so that ∂C consists of a collection of components of $\overline{S}$ with their given orientation. With no loss, assume that $(\partial C \cap S)_+$ lies in C. Observe the following:

(a) If $\partial C \subset S'$, then removing ∂C from S' will not change $[S']$ and can only reduce $\chi_\beta(S')$. Thus it suffices to prove the claim with ∂C removed from S'.

(b) If $\partial C \subset S$, then because the orientation of ∂C agrees with that of S, C contains a null-homology of $(\partial C \cap S)_+$. Since S_+ is β'-norm-minimizing, we must have $\chi_\beta(\partial C) = 0$. Now alter S by reversing the orientation of all components of ∂C. The β-norm of S is unchanged, and both S_+ and S_- remain β'-norm-minimizing, since the change in their homology class has trivial norm. Hence it would suffice to prove the claim for the new S.

(c) If some components of ∂C are in S' and some in S then, because S_+ is β'-taut, $\chi_\beta(\partial C \cap S') \geq \chi_\beta(\partial C \cap S)$. Replacing $\partial C \cap S'$ with $(\partial C \cap S)_-$ does not alter $[S']$ nor raise $\chi_\beta(S')$. Thus it suffices to prove the claim for this altered S'.

To prove the claim, pick a base-point x_0 on ∂M and construct $M - \overline{S}$ a potential function just as in 3.1 (cf. also 3.4 below). Operations (a), (b), and (c) allow one to eliminate components of $M - \overline{S}$ whose potential is most distant from 0. Eventually all regions can be made to have potential 0 or 1. Moreover each region of potential one is either a collar between a component of S and a component of S' (arising from operation (c)) or has oriented boundary completely contained in S and of trivial χ_β, as in (b). Every component of $\overline{S}$ must lie on the boundary of exactly one region with potential one. So after the operations above, every component of S' is parallel to a component of S, and each component of S not parallel to one of S' has trivial χ_β. This proves the claim.

Proof of Corollary 3.3. We no longer assume that S' and S are disjoint.

Apply 1.3(d) with $N = \mathcal{N}(\partial S)$ to modify S' so that for all $k \geq 0$, $\chi_\beta(S' \searrow kS) = \chi_\beta(S') + k\chi_\beta(S)$. From 3.2, there is a $k \geq 0$ such that $S' \searrow kS \subset M'$. Moreover $S' \searrow kS$ represents $(k+1)[S]$ in $H_2(M, \mathcal{N}(\partial S))$, and so can be decomposed into $k+1$ disjoint pieces, each of them homologous to S in $H_2(M, \mathcal{N}(\partial S))$.

By the above claim, each piece has $\chi_\beta \geq \chi_\beta(S)$. Hence $\chi_\beta(S') + k\chi_\beta(S) \geq (k+1)\chi_\beta(S)$, i.e., $\chi_\beta(S') \geq \chi_\beta(S)$.

Now consider the notion of "potential" in 3.1 for complements of surfaces in 3-manifolds:

3.4 Definition. For $(T, \partial T) \subset (M, \partial M)$ a surface whose fundamental class is trivial in $H_2(M, \partial M)$ and x_0 in $M - T$, define the *potential* $\Psi_T \colon M - T \to \mathbb{Z}$ to be the function which assigns to any y in $M - T$ the algebraic intersection with T of an arc from x_0 to y. The *variance* ψ_T of T is the maximum, over all y, z in $M - T$, of $|\Psi_T(y) - \Psi_T(z)|$ and is independent of x_0.

3.5 Lemma. (a) *Suppose* $(T', \partial T') \subset (M, \partial M)$ *is a surface with* $\partial T' = \partial T$ *and* T' *is homologous to* T *in* $H_2(M, \partial T)$. *For* x_0 *in* $\partial M - \partial T$, $\Psi_T = \Psi_{T'}$ *on* $\partial M - \partial T$. (Note: $\Psi_T = \Xi_{\partial T}$ *perhaps only on the component of* ∂M *containing* x_0.)

(b) *Let* T' *be obtained from* T *by discarding null-homologous closed components. Then* $\chi_\beta(T') \leq \chi_\beta(T)$ *and the variance* $\psi_{T'}$ *is achieved on* $\partial M - T'$.

(c) *Let* T *be a submanifold of* ∂M, *with all components oriented either consistently or all inconsistently with the orientation of* ∂M. *If interior* (T) *is pushed into* M, *then* $\psi_T = 1$.

Proof. (a) Since T and T' are homologous, any path from x_0 to a point y in $\partial M - \partial T$ intersects T algebraically as often as it intersects T'.

(b) If the variance ψ_T is not achieved on $\partial M - T$, then either the region of highest or lowest potential in $M - T$ is disjoint from ∂M. Its boundary must then be a null-homologous closed component of T, which can be removed without increasing $\chi_\beta(T)$. The result follows by induction.

(c) Let x_0 be a point in $M - T$ not in the product region between T and ∂M. Then any other point in $M - T$ can be reached from x_0 by an arc which crosses T at most once, and always with the same sign.

3.6. Let $(S, \partial S) \subset (M, \partial M)$ be a d-surface in a sutured manifold (M, γ). Divide S into two surfaces: $S_2 = \{$components of S which β-compress $R(\gamma)\}$, $S_1 = S - S_2$. Then the decomposition of M by S can be factored through the decompositions

$$(M, \gamma) \xrightarrow{S_1} (M_1, \gamma_1) \xrightarrow{S_2} (M', \gamma').$$

Theorem. *Suppose* (M', γ') *is* β-taut. *Then either*

(a) $S_2 = \varnothing$ *and* (M, γ) *is* β-taut,

(b) *some component of* ∂S *is* β-inessential *in* $R(\gamma)$, *or*

(c) *every component of* M_1 *which contains a component of* S_2 *is a copy of* $D^2 \times S^1$, *disjoint from* γ_1 *and* β, *in which* S_2 *is a collection of coherently oriented meridian disks.*

Proof.   *Case* 1. $S_2 = \varnothing$, so $S = S_1$.

An edge of β with both ends in $R_\pm$ will produce an edge of β' with both ends in $R'_\pm = R_\pm(\gamma')$. Thus we can assume every edge of β' intersects $R(\gamma')$ with the same sign.

Next note that if there is a circle in $R(\gamma) \cap S$ which is inessential in $R(\gamma)$ *or* S, then (b) is satisfied. If the circle is in $R(\gamma)$ it is satisfied directly; if it is in S this follows from the assumption that $S_2 = \varnothing$ (S contains no β-compressing disks for $R(\gamma)$).

Claim 1. M is β-irreducible or (b) is satisfied.

Suppose M is β-reducible and let Q be a β-reducing 2-sphere with a minimal number of components of intersection with S. Then an innermost disk F in $Q - S$ cannot give a β-compression of $R(\gamma')$, so $\partial F \subset S$ bounds a disk D disjoint from β in say $R'_+ = R_+(\gamma')$. D cannot lie entirely in S, or Q would not have a minimal number of components of intersection with S. Hence some part of D lies in R_+, and some (including ∂D) in S. Then an innermost circle of $S \cap R_+$ in D bounds a disk in either R_+ or S, so (b) is satisfied.

Claim 2. $R(\gamma)$ and S are β-incompressible or (b) is satisfied.

We show that R_+ in β-incompressible. The proof for R_- and S is similar.

Let Q be a β-compressing disk for R_+ which, among all β-compressing disks, has a minimal number of components of intersection with S.

Subcase 1. $Q \cap S$ has a circle component.

Then the argument of Claim 1 applies.

Subcase 2. Q and S intersect only in arcs.

Then minimality of $|Q \cap S|$ and Subcase 1 imply that every arc of $\partial Q - \mathcal{N}(S)$ is β-essential in $R_+ - \mathcal{N}(S)$ and every arc of $Q \cap S$ is β-essential in S. Some component Q' of $Q - \mathcal{N}(S)$ has its boundary entirely in R'_+. Since $R(\gamma')$ is β'-incompressible, $\partial Q'$ also bounds a disk E in $R'_+ = R_+ \searrow S$. An outermost arc of $R_+ \cap S$ in E cuts off a disk F lying entirely in R_+ or entirely in S. But if F lay entirely in R_+, it would mean that $\partial Q - \mathcal{N}(S)$ is *not* β-essential in $R_+ - \mathcal{N}(S)$. If F lay entirely in S, then it would mean that *not* every arc of $Q \cap S$ is β-essential in S. Neither is possible, completing the contradiction.

Subcase 3. Q and S are disjoint.

Then ∂Q must bound a disk D in $R'_+ - \beta$, but not $R_+ - \beta$. Then an innermost closed curve of $S \cap R_+ = \partial S \cap R_+$ in D either bounds a disk in $R_+ - \beta$ or in S and so satisfies case (b).

Claim 3. $R(\gamma)$ is β-minimizing or (b) is satisfied.

Suppose R_+, say, is not β-minimizing in $H_2(M, \partial R_+)$. Let T be a β-taut surface with $[T] = [R_+, \partial R_+]$ in $H_2(M, \partial R_+)$.

By 3.5(c), $\psi_{R_+} = 1$. After discarding null-homologous components of T (including, by β-irreducibility, all sphere components disjoint from β) we have,
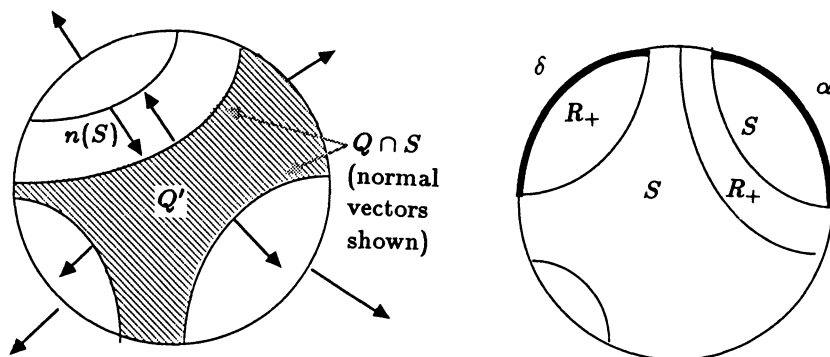
FIGURE 3.2

by 3.5(b), $\psi_T = 1$. Since R_+ is β-incompressible, and M is β-irreducible, any disk component of T which is disjoint from β is parallel to a component of R_+. We now proceed with a more delicate version of the proof of 1.3(d).

After possibly performing 2-surgeries to T, we can assume that no simple closed curve of $S \cap T$ bounds in S a disk disjoint from β. Then since $R(\gamma')$ is β-incompressible, either there is a component of ∂S bounding a disk in $R(\gamma)$ (case (b)) or no simple closed curve of $S \cap T$ bounds in T a disk disjoint from β. Hence any sphere component of $S \smallsetminus T$ either lies entirely in S, entirely in T or intersects β in at least two points.

Similarly, we may assume that any sphere component of R'_+ either lies in S, lies in R_+, or intersects β in at least two points.

Since $\partial T = \partial R_+$, $\partial(S \smallsetminus T) = \partial(R'_+)$. Since $\psi_T = 1$, we can assume that the variance of $\Lambda = S \cap T$ in S is also 1. Then by Lemma 3.2(b), $S \smallsetminus T$ can be isotoped into M'.

Since R'_+ is β-incompressible, if any component of $\partial(S \smallsetminus T)$ bounds a disk in M disjoint from β, it bounds such a disk in R'_+. Since T is taut, if a component of $\partial T = \partial R_+$ bounds a disk in M, it bounds a disk in T.

Letting

$$e = |\text{spheres of } S \smallsetminus T \text{ intersecting } \beta \text{ in a single point}|,$$
$$e' = |\text{spheres of } R'_+ \text{ intersecting } \beta \text{ in a single point}|,$$
$$d = |\text{disks in } S \smallsetminus T \text{ disjoint from } \beta|,$$
$$d'_\gamma = |\text{disks in } R'_+ \text{ disjoint from } \beta|,$$
$$d_S = |\text{disks in } S \text{ disjoint from } \beta|,$$
$$d_T = |\text{disks in } T \text{ disjoint from } \beta|,$$
$$d_\gamma = |\text{disks in } R_+ \text{ disjoint from } \beta|,$$

we thereby derive $d \leq d'_\gamma$, $d\gamma \leq d_T$, $e = |$spheres of S intersecting β in a single point$| + |$spheres of T intersecting β in a single point$|$, and similarly for e' and spheres of S and R_+.

Then

$$\chi_\beta(S \searrow T) = |(S \searrow T) \cap \beta| - \chi(S \searrow T) + d + e$$
$$= |S \cap \beta| + |T \cap \beta| - \chi(S) - \chi(T) + d + e$$
$$= \chi_\beta(S) + \chi_\beta(T) + d - d_S - d_T.$$

Since $\chi_\beta(T) < \chi_\beta(R_+)$,

$$\chi_\beta(S \searrow T) < \chi_\beta(S) + \chi_\beta(R_+) + d'_\gamma - d_S - d_\gamma$$
$$= |S \cap \beta| + |R_+ \cap \beta| - \chi(S) - \chi(R_+) + d'_\gamma + e' = \chi_\beta(R'_+).$$

Thus $S \searrow T \subset M'$ contradicts the β-tautness of R'_+.

Case 2. $S_2 \neq \varnothing$ (so (M, γ) is not taut).

With no loss of generality assume every component of M_1 contains a component of S_2. Every component D of S_2 produces two disks in M', one of them, denoted D_s, is bounded by a suture and the other, denoted D_n, is not. Construct an imbedded directed graph X in M_1 by choosing a vertex in every component of M', choosing an edge through each component of D with ends on the vertices in the components of M' on which D_s and D_n lie. Orient the edge to point in the direction from D_n to D_s.
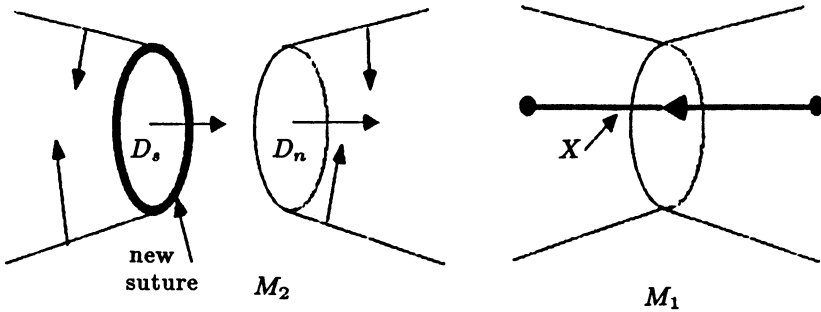

FIGURE 3.3

Claim 1. No component of M' has more than one edge directed in. If an edge is directed into a compact $\overline{M}$ of M', then $\overline{M}$ is a 3-ball disjoint from β and has at least one edge pointing out.

Proof of Claim 1. D_s is a disk in ∂M_2 disjoint from β and bounded by a suture. Since $\partial M_2 - \gamma$ is β-incompressible, the other component of $\partial M_2 - \gamma$ adjacent to that suture must also be a disk D' disjoint from β. Hence $\partial \overline{M}$ is

a sphere disjoint from β and any other suture. By β-irreducibility of M', $\overline{M}$ is a 3-ball disjoint from β.

If D' lies entirely in ∂M, then (b) is satisfied. In general, an innermost simple closed curve of ∂S lying in D' either bounds a disk in ∂M (case (b)) or bounds a disk D'' in S. Either $\partial D''$ bounds a disk in ∂M (again case (b)) or it is a compressing disk for ∂M and so a component of S_2. Then an edge of X points out through D''.

Claim 2. X is a disjoint union of cycles.

Suppose there are n vertices of X. Since every component of M_1 contains some disk in S_2, no vertex is isolated. By Claim 1, there is at least one edge pointing out of each vertex, and at most one pointing in. Hence there are exactly n edges in X, so X is the union of cycles.

By Claim 2, every vertex of X has an edge pointing in, so, by Claim 1, each component of M' is a 3-ball. Hence M_1 deformation retracts to X, completing Case 2.

3.7 Definition. For S a d-surface in a sutured manifold (M, γ), let $S_{k,l}$ denote $S \searrow kR_+ \searrow lR_-$.

3.8 Corollary. *If $(M, \gamma) \xrightarrow{S} (M', \gamma')$ is a β-taut sutured manifold decomposition and no component of ∂S bounds a disk in $R(\gamma) - \beta$, then (M, γ) $\xrightarrow{S_{k,l}} (M'', \gamma'')$ is a β-taut sutured manifold decomposition.*

Proof. With no loss of generality it suffices to consider the case $k = 1$, $l = 0$. Since (M, γ) is β-taut, no component of S is a β-compressing disk for $R(\gamma)$, hence no component of S is a disk disjoint from β with ∂S in $R(\gamma)$. There is a sutured manifold decomposition $(M'', \gamma'') \xrightarrow{T} (M', \gamma')$ in which T is disjoint from β and consists of product disks and annuli. A cross-section of T is shown in Figure 3.4.

We wish to show that no component of ∂T bounds a disk in $R(\gamma'') - \beta$ so that 3.6 applies. The complement of ∂T in R''_+ is naturally homeomorphic to the union of $R_+ \searrow S$ and of the complement of ∂S in R_+, with ∂T corresponding in both cases to $\partial S \cap R_+$.

A disk in $R_+ \searrow S - \beta$ with boundary in $\partial S \cap R_+$ would contain a disk with boundary in $R_+ \cap S$ and interior entirely contained in $R_+ - \beta$ interior entirely contained in $S - \beta$. The first is impossible by hypothesis, and the second is also because (M, γ) is β-incompressible.

By hypothesis no component of ∂T can bound a disk in R_+.

3.9 Corollary. *Suppose*

$$(M_0, \gamma_0) \xrightarrow{S_1} (M_1, \gamma_1) \xrightarrow{S_2} \cdots \xrightarrow{S_n} (M_n, \gamma_n)$$

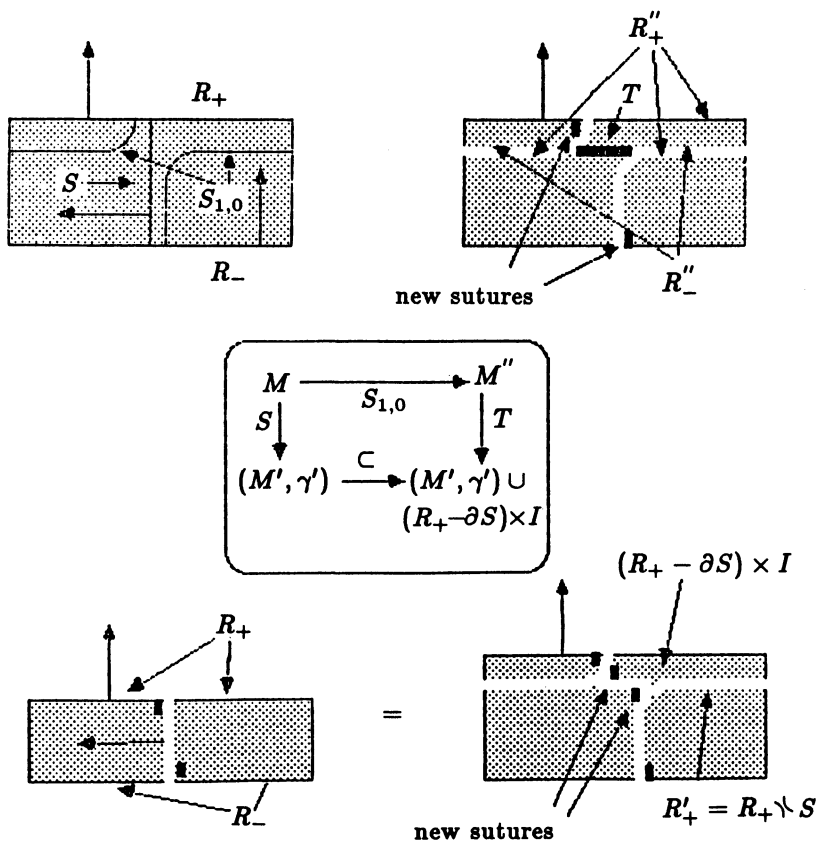*is a sequence of sutured manifold decompositions in which*

FIGURE 3.4

(a) *no component of M_0 is a solid torus disjoint from β and γ_0,*

(b) *each S_i is either a conditioned surface, a disk disjoint from β intersecting $A(\gamma_{i-1})$ twice, or an annulus disjoint from β with at least one boundary component β-essential in $R_{\pm}(\gamma_{i-1})$, and the other boundary component lying in $R_{\mp}(\gamma_{i-1})$,*

(c) *no closed component of any S_i separates.*

Then if (M_n, γ_n) is β-taut, so is every decomposition in the series.

Proof. The result follows from 3.6(a) once we verify that neither 3.6(b) nor 3.6(c) occurs anywhere in the sequence.

Suppose some decomposition $(M_{i-1}, \gamma_{i-1}) \xrightarrow{S_i} (M_i, \gamma_i)$ can be factored as in 3.6(c). That is, there is a collection of components $\hat{S}$ in S_i so that the decomposition $(M_{i-1}, \gamma_{i-1}) \xrightarrow{\hat{S}} (\hat{M}, \hat{\gamma})$ yields a component M' of $\hat{M}$ which is a solid torus disjoint from β_i and with boundary in $R_+(\hat{\gamma})$, say. Take, in fact,

S_i to be the first surface in the sequence admitting this factorization. Then there are three possibilities for $\partial M' \subset R_+(\gamma_{i-1}) \smallsetminus \hat{S}$:

(i) $\partial M'$ is a component of $\hat{S}$. This would contradict hypothesis (c).

(ii) $\partial M'$ is a component of $R_+(\gamma_{i-1})$. Then, since S_i was the first surface in the decomposition to admit such a factorization, in fact M' is a component of M_0, contradicting hypothesis (a).

(iii) $\partial M' - \hat{S}$ is a null-homology in $R_+(\gamma_{i-1})$ of $\partial \hat{S}$. This would contradict hypothesis (b) or 2.4(b).

Suppose 3.6(b) occurs and let S_i be its last occurrence. Then 3.6 applied to each term in the series

$$(M_i, \gamma_i) \xrightarrow{S_{i+1}} (M_{i+1}, \gamma_{i+1}) \cdots \xrightarrow{S_n} (M_n, \gamma_n)$$

shows that (M_i, γ_i) is β-taut. By hypothesis 3.9(b), S_i is an annulus disjoint from β with one end β-inessential in R_+, say, and the other end β-essential in R_-. There is a suture of γ_i at each end of S_i: γ_+ on the positive side S_i^+ and γ_- on the negative side S_i^-. Since the end of S_i in R_+ bounds a disk disjoint from β, γ_+ bounds a disk in $R_+(\gamma_i)$. Since (M_i, γ_i) is β-taut, γ_+ must bound a disk disjoint from β in $R_-(\gamma_i)$. This disk cannot contain S_i^- since it has the suture γ_- at an end. Hence the disk lies in $R(\gamma_{i-1})$, a contradiction.

4.   Complexity: Sutured manifold hierarchies

The goal of this section is to show that for any β-taut sutured manifold there is a finite sequence of β-taut decompositions which terminates in a manifold with trivial $H_2(M, \partial M)$. Let M_β denote $M - \overset{\circ}{\mathcal{N}}(\beta)$.

4.1 Definition.   An annulus $(S^1 \times I; S^1 \times \{0\}, S^1 \times \{1\}) \subset (M_\beta; R_+(\gamma),$ $R_-(\gamma))$ in the sutured manifold (M, γ, β) is called a *product annulus*, and a disk $(I \times I; I \times \{0\}, I \times \{1\}, \partial I \times I) \subset (M_\beta; R_+(\gamma), R_-(\gamma), A(\gamma))$ is called a *product disk*. A product annulus is *trivial* if it extends to an imbedding $(D^2 \times I; D^2 \times \{0\}, D^2 \times \{1\}) \subset (M_\beta; R_+(\gamma), R_-(\gamma))$.

A disk $(I \times I; I \times \{0\}, I \times \{1\}, \partial I \times I) \subset (M_\beta; R_+(\gamma), R_-(\gamma), \dot{\mathcal{N}}(\beta))$ is called an *amalgamating disk* because one can view the disk as amalgamating the two (or possibly one) arcs of β to which $\partial I \times I$ runs parallel. If both components of $\partial I \times I$ lie on the same component β_0 of β, the disk is a *self-amalgamating disk* at β_0.

A disk

$$(I \times I; I \times \{0\}, I \times \{1\}, \{0\} \times I, \{1\} \times I) \subset (M_\beta; R_+(\gamma), R_-(\gamma), \dot{\mathcal{N}}(\beta), A(\gamma))$$

is called a *cancelling disk* because one can view the disk as pushing the arc of β along which $\{0\} \times I$ runs into the structure along which $\{1\} \times I$ runs.

4.2 Lemma. *Suppose* $(M, \gamma) \xrightarrow{J} (M', \gamma')$ *is a sutured manifold decomposition with* J *either a product disk or product annulus. Suppose further that each component of* ∂J *is* β-*essential in* $R(\gamma)$. *Then* (M, γ, β) *is* β-*taut if and only if* (M', γ', β) *is* β-*taut.*

Proof. "If" follows from 3.6. So suppose (M, γ) is β-taut. $R(\gamma')$ is naturally homeomorphic to $R(\gamma) - \partial J$, so $\chi_\beta(R(\gamma')) = \chi_\beta(R(\gamma))$ or $\chi_\beta(R(\gamma)) - 2$ depending on whether J is a product annulus or a product disk. If $R(\gamma')$ is not β-minimizing, then attaching $\partial J \cap R(\gamma)$ to a β-minimizing surface would produce a surface homologous to $R(\gamma)$, with the same boundary, but lower β-norm. Hence $R(\gamma')$ is β-minimizing.

It is easy to check that $R'_\pm = R_\pm(\gamma')$ is β-minimizing. Since M is β-irreducible, so is $M' = M - J$.

Suppose Q is a β-compressing disk for $R(\gamma')$. Since $R(\gamma')$ can be regarded as $R(\gamma) - \partial J$, we can take $\partial Q \subset R(\gamma)$. Since $R(\gamma)$ is β-incompressible, ∂Q bounds a disk in $R(\gamma)$ disjoint from β. Since ∂J is β-essential in $R(\gamma)$, that disk persists in $R'_\pm \cong R_\pm - \partial J$, contradicting the fact that Q is a β-compressing disk.

4.3 Lemma. *Suppose* (M, γ) *is a* β-*taut sutured manifold,* Q *is a cancelling disk for the arc component* β_0 *of* β, *and* $\beta_- = \beta - \beta_0$. *Then* (M, γ) *is also* β_--*taut.*

Proof. **Claim 1.** M is β_--irreducible.

Proof of Claim 1. Suppose not. Among all spheres disjoint from β_-, not bounding balls disjoint from β_- and put in general position with respect to Q, let S have minimal number of components of intersection with Q. No component of $S \cap Q$ is a simple closed curve, for otherwise we could produce such a sphere with fewer components of intersection with Q by splitting S along an innermost disk of $Q - S$. If $S \cap Q \neq \varnothing$, then let α be an outermost arc in Q and F be a disk in $Q - S$ whose boundary consists of the union of α and a subarc of β_0. Then there is a regular neighborhood of F whose boundary S_0 intersects S in a regular neighborhood $\mathcal{N}$ of α. If we split S along $S_0 - \mathcal{N}$, the result is two spheres: S_0 and a sphere S_1 which intersects Q in fewer components than does S. Hence S_1 bounds a ball disjoint from β_-. Since F is disjoint from β_-, S_0 also bounds a ball disjoint from β_-, and finally so must S, a contradiction.

Claim 2. $R(\gamma)$ is β_--incompressible.

Proof of Claim 2. Suppose not, and among all β_--compressing disks for $R(\gamma)$, put in general position with respect to Q, let D have minimal number of components of intersection with Q. Then an argument quite similar to that of Claim 1, using also the β_--irreducibility of M, leads to a similar contradiction.
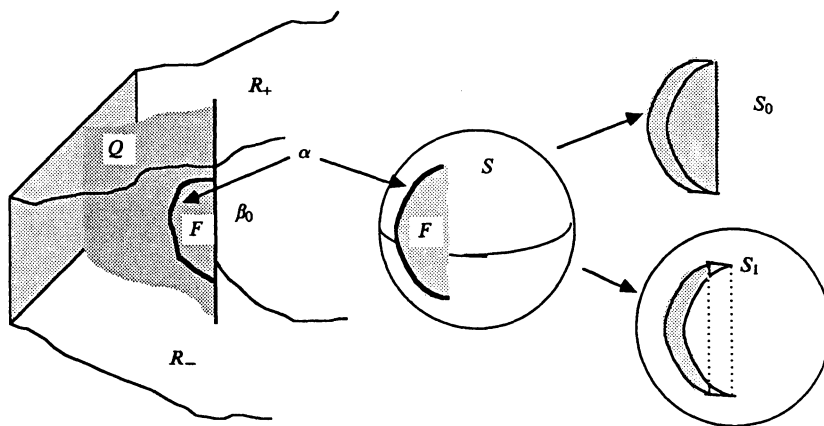
Claim 3. $R(\gamma)$ is β_--minimizing.

FIGURE 4.1

Proof of Claim 3. Let S be a β_--taut surface in M, with $\partial S = \gamma$ homologous to $R_\pm$ rel γ and, among all such, having minimal number of components of intersection with Q. Since S is β_--incompressible and M is β_--irreducible, there are no simple closed curves in $S \cap Q$. Since S is β-minimizing, there are no arcs of $S \cap Q$ with both ends in $\beta \subset \partial Q$. Since $\partial S = \gamma$, $\partial S \cap \partial Q$ is a single point, lying in $A(\gamma_0)$. Hence $S \cap Q$ consists of a single arc with one end in $\partial Q \cap A(\gamma_0)$ and one end in β. But then $\chi_{\beta_-}(S) = \chi_\beta(S) - 1 = \chi_\beta(R_\pm) - 1 = \chi_{\beta_-}(R_\pm)$, so $R_\pm$ are also β_--minimizing.

4.4 Lemma. *Suppose (M, γ) is a β-taut sutured manifold, and Q is an amalgamating disk for distinct arc components β_0 and β_1 of β. Let $\beta_- = \beta - \beta_0$. Then (M, γ) is also β_--taut.*

Proof. The proof is essentially the same as that of 4.3 above.

4.5 Lemma. *Suppose (M, γ) is a β-taut sutured manifold, and Q is a self-amalgamating disk at $\beta_0 \subset \beta$. Suppose the two arc components of $\partial Q \cap R(\gamma)$ are β-essential in $R(\gamma)$. Let $\beta_- = \beta - \beta_0$, and (M', γ') be the sutured manifold obtained by decomposing (M, γ, β_-) along the product annulus in $Q \cup \mathcal{N}(\beta_0)$. Then (M', γ') is β_--taut.*

Proof. This is a combination of 4.2 and 4.3.

4.6 Definition. Let $(D^2, \partial D^2) \subset (M_\beta, \partial M_\beta)$ be a disk in the sutured manifold (M, γ, β). A circuit around ∂D^2 will pass μ times across neighborhoods of edges of β, $\mathcal{V}$ times across neighborhoods of vertices of β, and ν times across components of $A(\gamma)$. Define the index

$$I(D) = \mu + \nu + 3\mathcal{V} - 2.$$

If ∂D^2 lies entirely in a neighborhood of a vertex or is a meridian circle of $\mathcal{N}(\beta)$ set $I(D) = 1$, and if ∂D^2 lies entirely in $R(\gamma)$ set $I(D) = -2$.

Remark. If $I(D) < 0$, then either $\partial D^2 \subset R(\gamma)$ or some edge of β has both ends in R_+ or both ends in R_- (and D compresses the edge into $R(\gamma)$).

If $I(D) = 0$, then D is either a cancelling, amalgamating or product disk, depending on whether $(\mu, \nu) = (2, 0), (0, 2)$ or $(1, 1)$.

4.7 Definition. For a collection $\{F\}$ of index-zero disks in M_β, the combination of decomposition along product disks and annuli $(M, \gamma, \beta) \rightarrow (M', \gamma', \beta)$ with cancellation and amalgamation of arcs in β to produce a complex β' in M' is called *elimination of* $\{F\}$. As in Lemma 4.5 we view elimination of a self-amalgamating disk as the combination of a decomposition along a product annulus with a cancellation.

4.8 Lemma. *Suppose* $(M, \gamma) \xrightarrow{S} (M', \gamma')$ *is a sutured manifold decomposition, and* $(D, \partial D) \subset (M_\beta, \partial M_\beta)$ *is a disk intersecting* S *only in arcs. Let* $D_1, \cdots, D_n$ *be the components of* $D - S$, *regarded as disks in the manifold* $M'_{\beta-S}$. *Then*

$$I(D_1) + \cdots + I(D_n) = I(D).$$

Proof. We need not consider the case in which ∂D lies entirely in the neighborhood of a vertex or is a meridian circle of $\mathcal{N}(\beta)$, for then D and S are disjoint.

It suffices to consider the case in which $S \cap D$ is a single arc α. The number of vertices crossed by $\partial D_1 \cup \partial D_2$ is clearly the same as for ∂D, since S, and hence α, are disjoint from the vertices of β. Then $\mathcal{V}$ of 4.3 is unchanged. If an end of α lies in β, then μ for $D_1 \cup D_2$ is 1 greater than μ for D. If an end of α lies in $A(\gamma)$, $T(\gamma)$, or $R_\pm$ (hence in general in ∂M), then one side or the other of the component of ∂S in which α terminates becomes a suture of γ'. Hence ν for $D_1 \cup D_2$ is 1 greater than ν for D. Thus in general $\mu + \nu$ for $D_1 \cup D_2$ is 2 greater than $\mu + \nu$ for D. But since the former contains two disks, the total index remains unchanged.

4.9 Definition. For M β-irreducible, a finite collection of disjoint ∂-reducing disks $(\{D_i\}, \{E_j\})$ in the irreducible 3-manifold M_β is called a set of *complexity disks* if

(a) $M_\beta - (\bigcup \mathcal{N}(D_i))$ is the disjoint union of m manifolds $\{P_\mu\}$ of the form (closed surface) $\times I$, k ∂-irreducible manifolds $\{M_\kappa\}$ not of that form, and l handlebodies $\{H_\lambda\}$,

(b) $(\bigcup E_j) \subset (\bigcup H_\lambda)$,

(c) each component of $(\bigcup H_\lambda) - (\bigcup \mathcal{N}(E_j))$ is a 3-ball whose boundary intersects closure$[(\bigcup \mathcal{N}(E_j)) \cup (\bigcup \mathcal{N}(D_i))]$ in ≤ 3 disks.

Let d_κ (resp. d'_μ) denote the number of disks $\{D_i\}$ which lie in ∂M_κ (resp. ∂P_μ).

Recall that the *length* (cf. [9, IV.13]) of a Haken manifold is the length of the longest partial hierarchy which can be constructed using only surfaces which are incompressible, boundary incompressible, but not disks.

4.10 Definition. For (M, γ, β) a connected β-irreducible sutured manifold and $(\{D_i\}, \{E_j\})$ a set of complexity disks, define the complexity $C(M, \{D_i\}, \{E_j\})$ of $(M, \{D_i\}, \{E_j\})$ to be the ordered 5-tuple $(C_1, C_2, C_3, C_4, C_5)$ where

$$C_1 = \sum (\text{length } M_\kappa), \qquad C_2 = \sum \{I(D_i)^2 | I(D_i) \geq 0\},$$

$$C_3 = \sum_{1 \leq \kappa \leq k} d_\kappa, \qquad C_4 = m + \sum_{1 \leq \mu \leq m} (d'_\mu)^2,$$

$$C_5 = \{I(D_i) > 0\} \cup \{I(E_j) > 0\} \quad \text{(repetitions retained)}.$$

4.11 Definition. For (M, γ, β) a connected sutured manifold define the *complexity* $C(M, \gamma, \beta)$ to be the minimum of the complexity of $(M, \{D_i\}, \{E_j\})$ taken over all sets $(\{D_i\}, \{E_j\})$ of complexity disks. The order is the dictionary order, with the order on the finite sets of integers C_5 given as follows (see [3, 4.3]): If A and B are two finite sets of integers, first order the elements of A and the elements of B in descending order, then compare A to B lexicographically.

For example, suppose the set B is obtained from A by replacing any positive element $\{a\}$ of A by the set $\{b_1, c_1, b_2, c_2, \cdots, b_n, c_n\}$ where each of b_i and c_i is positive and always $b_i + c_i = a$. Then under this ordering, no matter how large n is, always $B < A$.

4.12 Definition. For $\{F\}$ a set of disjoint index-zero disks in M_β, define $\hat{C}(\{F\})$ to be the complexity of the connected sutured manifold obtained by eliminating $\{F\}$ (cf. 4.7), but with C_5 altered to

$$C_5' = \{I(D_i) \geq 0\} \cup \{I(E_j) \geq 0\}.$$

Define the *reduced complexity* $\hat{C}(M, \gamma)$ to be the minimum of $\hat{C}(\{F\})$ taken over all sets $\{f\}$ of disjoint index-zero disks.

4.13 Remarks. (a) In fact in 4.12 we need only consider sets $\{F\}$ in which no disk is ∂-parallel, and no two disks are parallel in M_β. For example, if a disk D is boundary parallel in M_β, then a set of complexity disks for $M - D$ is a set of complexity disks for M.

(b) $\hat{C}(M, \gamma) \leq C(M, \gamma)$. Indeed, eliminating any set of index-zero disks never increases complexity, so we may assume that in the calculation of $\hat{C}(M, \gamma)$ all index-zero disks have been eliminated ($C_5' = C_5$).

(c) If M is β-taut, then all complexity disks have nonnegative index. In particular, if $f\hat{C}(M) = (0,0,0,0,\varnothing)$ then, after eliminating index zero disks, M_β becomes a handlebody ($C_1 = C_4 = 0$) admitting an empty set of complexity disks ($C_5' = \varnothing$). Thus M is a union of 3-balls and $\beta = \varnothing$.

4.14 Definition. For (M,γ,β) a sutured manifold and $(S,\partial S) \subset (M,\partial M)$ a d-surface, a *good disk* for S is a disk D disjoint from β with $D \cap (S \cup \partial M) = \partial D$ the union of an arc in S and an arc in $R_\pm$, and $I(D) \leq 0$, where index is computed in the manifold M_S obtained by decomposing M along S.

Remark. Seven examples of good disks are shown in [3, Figure 4.2].

4.15 Lemma. *Suppose (M,γ) is a sutured manifold, $(S,\partial S) \subset (M,\partial M)$ is a surface, and S' is S after a 2-surgery along a good disk. Let (M_S,γ_S) and $(M_{S'},\gamma_{S'})$ be the sutured manifolds obtained by decomposing along S and S' respectively. Then:*

(a) *$(M_{S'},\gamma_{S'})$ is obtained from (M_S,γ_S) by either β-compressing $R(\gamma_S)$ or decomposing it along a product disk,*

(b) *if S gives a β-taut sutured manifold decomposition, so does S'.*

Proof. (a) If $I(D) = -2$ then the former holds; if $I(D) = 0$ the latter holds.

(b) follows from (a) and 4.2.

4.16 Lemma. *Suppose $(S,\partial S) \subset (M,\partial M)$ is incompressible, and S' is obtained from S by 2-surgery along a disk whose boundary is the union of an arc in S and an arc of ∂M. If S' is ∂-parallel, so is S.*

Proof. Dually, S is obtained from S' by 1-surgery along an arc α in ∂M with $\partial \alpha \subset \partial S'$. (The 1-surgery looks like a "blistering" of ∂M near α.) If α lay inside the submanifold T of ∂M to which S' is parallel, then the disk $\alpha \times I$ in the product region between T and S' would be a compressing disk for S.

Hence α lies outside T. Then the product structure between S and T extends to a product structure between S' and $T \cup$ (neighborhood of α).

4.17 Theorem. *Suppose (M,γ) is a connected β-taut sutured manifold for which M_β has no index-zero ∂-compressing disks, $(S,\partial S) \subset (M,\partial M)$ is a connected β-incompressible surface, and the sutured manifold obtained by decomposing along S is β-taut. Let (M_S,γ_S) be any of its components. Then $C(M_S,\gamma_S) \leq \hat{C}(M,\gamma) = C(M,\gamma)$ and either*

(a) *$S - \mathcal{N}(\beta)$ is ∂-parallel in M_β and $[S,\partial S]$ is trivial in $H_2(M,\partial M)$, or*

(b) *$\hat{C}(M_S,\gamma_S) < \hat{C}(M,\gamma)$.*

Proof. Let S_β denote $S - \mathcal{N}(\beta)$. Choose a collection $(\{D_i\},\{E_j\})$ of complexity disks in M_β, which realize the complexity of (M,γ), and let $\{M_\kappa\}$ and $\{H_\lambda\}$ be as in 4.9(a). Modify S_β as follows:
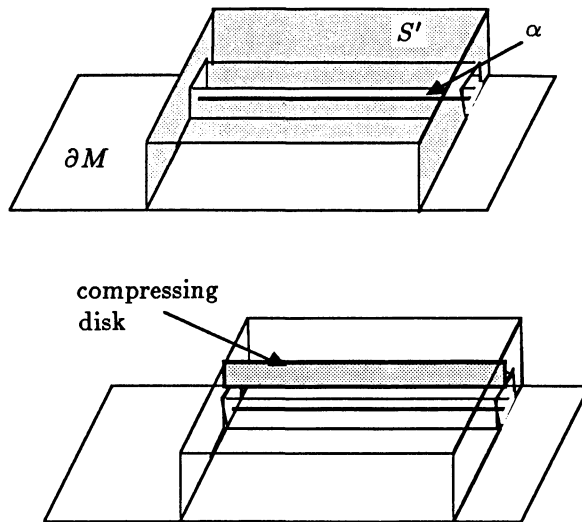
FIGURE 4.2

(a) Using β-irreducibility of M and β-incompressibility of S, isotope S_β to remove circles of intersection with complexity disks.

(b) Suppose there is an arc $\alpha \subset \beta$ such that α has one end in $R(\gamma)$, one end in S_β, and interior disjoint from S_β. Replace the meridian disk $\mathcal{N}(\alpha) \cap S$ by the annulus $\dot{\mathcal{N}}(\alpha)$. This alteration affects neither parallelism with ∂M_β, nor the β-tautness or reduced complexity of M_S.

(c) Suppose in a complexity disk Δ there is an outermost arc of $S_\beta \cap \Delta$, which cuts off a good disk in Δ. Then do 2-surgery on S_β via this good disk. By 4.15(a), this does not change $\hat{C}(M_S, \gamma_S)$ and, by Lemma 4.16, cannot make S_β ∂-parallel unless it was originally.

After enough such modifications, any outermost arc of intersection with S_β in any complexity disk has positive index, for operation (b) changes cancelling and amalgamating disks into product disks, and operation (c) eliminates product disks cut off by outermost arcs.

Case 1. For some M_κ, $S_\beta \cap M_\kappa$ has a component which is not boundary parallel. Then $M_\kappa - S_\beta$ has length $<$ length M_κ. This reduces C_1.

Henceforth then assume

A. For each M_κ, every component of $M_\kappa \cap S_\beta$ is boundary parallel.

Then each $M_\kappa - S_\beta$ is a union of handlebodies or (closed surfaces) $\times I$ and a copy of M_κ. Each $P_\mu - S_\beta$ is a union of handlebodies and (possibly) copies of P_μ. Each $H_\lambda - S_\beta$ is a union of handlebodies. In particular, the collection

$\{D'\}$ of all components of $\{D_i - S_\beta\}$ can be extended to a set of complexity disks $(\{D'\}, \{E'\})$ for $M - S$.

Case 2. For some disk D_i, $S_\beta \cap D_i \neq \varnothing$.

Since at least two disks of $D_i - S_\beta$ have positive index, it follows from 2.2 that a sutured decomposition along S reduces C_2 without increasing C_1.

Henceforth we therefore also assume

B. For every disk D_i, $S_\beta \cap D_i = \varnothing$.

Case 3. S is contained in some M_κ.

By assumption A, S_β is boundary parallel in M_κ. If, in fact, S_β is boundary parallel in M_β, then one component of $M_\beta - S_\beta$ is unchanged, and the other is either a (closed surface) $\times I$ or a handlebody, hence has lower complexity.

If S_β is not boundary parallel in M_β, the piece of the boundary of M_κ to which S_β is parallel must contain a D_i. Then decomposing along S_β does not increase C_1 and C_2 but reduces C_3.

We therefore assume

C. S_β and $(\bigcup M_\kappa)$ are disjoint.

Then decomposing along S cannot increase C_1, C_2, or C_3.

Recall that an incompressible, ∂-incompressible surface in a handlebody or in (surface) $\times I$ either is closed and ∂-parallel or decomposes the 3-manifold into handlebodies.

Case 4. S is contained in some P_μ.

If S_β is boundary parallel in M_β then, as in Case 3, the complexity of neither component of $M_\beta - S_\beta$ is larger than that of M_β.

If S_β is closed and boundary parallel in P_μ but not in M_β then, of the terms of C_4, m is increased by one, but d'_μ is partitioned into two positive integers, the sum of whose squares then must be no larger than $(d'_\mu)^2 - 2$. Hence C_4 is reduced.

If S_β has boundary, then m does not increase, but $\sum d'^2_\mu$ decreases and again C_4 is reduced.

We therefore assume

D. S_β is contained in some H_λ.

Case 5. S_β intersects some E_i.

$H_\lambda - S_\beta$ is a union of handlebodies, so C_1, C_2, C_3, C_4 are unchanged. We now consider C'_5. Let $\{E_\sigma\}$ denote the subcollection of $\{E_j\}$ which intersects S_β, and $\{E_\tau\}$ the subcollection which does not.

Claim. There is an isotopy of S_β rel γ and a collection of disks $\{E'_\sigma\}$, each parallel to some E_σ, such that, for $\{E'\}$ the set of disks in $M - S_\beta$ obtained by splitting $\{E'_\sigma\}$ along S_β, $(\{D_i\}, \{E_\tau\} \cup \{E'\})$, is a set of complexity disks for $M - S$.

Proof of Claim. By general position, S_β may be isotoped so that, outside $\mathcal{N}(\bigcup E_\sigma)$, each component of $H_\lambda - S$ is a 3-ball whose closure intersects $\mathcal{N}(\bigcup E_\sigma)$ in ≤ 3 disks (see [3, Figure 4.4]). Isotope S_β in $\mathcal{N}(\bigcup E_\sigma) = (\bigcup E_\sigma) \times I$ so that at most a single saddle occurs between distinct levels $0 = \alpha_1 < \cdots < \alpha_n = 1$. Then $\{E'_\sigma\} = \{E_\sigma\} \times \{\alpha_i\}$ satisfies the claim (see [3, Figure 4.1]).

To understand the effect on C'_5, consider a single E_σ. Each copy $E_\sigma \times \alpha_j$ is broken into disks by S_β, and at least two disks have positive index, since there are no good disks for S_β. From 4.8 the corresponding integers in $C'_5(M - S)$ are all strictly smaller than the entry $I(E_\sigma)$ in $C'_5(M)$, so C'_5 is decreased.

Case 6. S_β is disjoint from $\{E_i\}$.

Then S_β is parallel to a complexity disk. Obtain a set of complexity disks for $M - S$ by removing from $(\{D_i\}, \{E_j\})$ all those disks parallel to components of S_β. This does not increase $C_1, \cdots, C_4$, but reduces C'_5.

Finally S_β parallel in M_β implies that S has trivial fundamental class in $H_2(M_\beta, \partial M_\beta) \cong H_2(M, \partial M \cup \mathcal{N}(\beta))$. The exact sequence of the triple $(M, \partial M \cup \mathcal{N}(\beta), \partial M)$ then implies $[S, \partial S]$ is trivial in $H_2(M, \partial M)$.

4.18 Definition. A β-taut sutured manifold hierarchy is a finite sequence

$$(M_0, \gamma_0) \xrightarrow{S_1} (M_1, \gamma_1) \xrightarrow{S_2} \cdots \xrightarrow{S_n} (M_n, \gamma_n)$$

of β-taut sutured manifold decompositions for which

(a) each S_i is either a conditioned surface, a product disk, or a β-nontrivial product annulus,

(b) no closed component of any S_i separates,

(c) if either end of a product annulus

$$(S^1 \times I; S^1 \times 0, S^1 \times 1) \subset (M_i, R_+(\gamma_i), R_-(\gamma_i))$$

in S_{i+1} bounds a disk in $R(\gamma_i)$, then no component of β incident to the disk is an edge parallel to the annulus,

(d) $H_2(M_n, \partial M_n) = 0$, so in fact ∂M_n is a union of spheres. (Note, however, that because bits of β may lie in M_n, β-irreducibility of M is not enough to conclude here that M_n is a union of 3-balls.)

4.19 Theorem. *Any β-taut sutured manifold (M_0, γ_0) admits a β-taut sutured manifold hierarchy.*

Proof. Eliminate index-zero disks in (M, γ) so that the resulting sutured manifold (M_1, γ_1) has $C(M_1, \gamma_1) = \hat{C}(M_1, \gamma_1) = \hat{C}(M, \gamma)$ (4.13(b)). Doing self-amalgamations last guarantees the annuli satisfy 4.18(c). By 4.3–4.5 (M_1, γ_1) is still β_1-taut. If $H_2(M_1, \partial M_1) = 0$ then stop. Otherwise there is, by 2.6, a conditioned surface S_2 in some component of M_1 such that

(a) $[S_2, \partial S_2] \neq 0$ in $H_2(M_1, \partial M_1)$,

(b) S_2 decomposes M_1 to a β-taut sutured manifold (M_2, γ_2), and

(c) S_2 contains no separating closed component.

Then apply 4.17 first to a component S of S_2 for which $[S, \partial S] \neq 0$ in $H_2(M_1, \partial M_1)$ and then to the other components of S_2. This replaces the component of M_1 in which S_2 lies with components in M_2 all of which have lower reduced complexity. Continue in this manner until $H_2(M_n, \gamma_n)$ is trivial (perhaps because the reduced complexity is trivial as in 4.13(c)). This sequence, with all amalgamations and cancellations of arcs in β ignored, is the required series of decompositions.

5. Warm-up application: Satellite knots of winding number zero have property P

The first application is particularly easy because $\beta = \varnothing$. The principal result is also a consequence of the more general theorem (§9) that all satellite knots have property P.

Let M be an irreducible cobordism between tori T_0 and T_1 with the property that any torus $T' \subset M$ which separates T_0 from T_1 is ∂-parallel. (The more general criterion that M is "Sp-atoroidal" suffices, see [4].) Let M^φ be the 3-manifold obtained from M by attaching $D^2 \times S^1$ to T_1 via a framing $\varphi \colon \partial D^2 \times S^1 \to T_1$. Inclusion induces a natural homomorphism $\Phi \colon H_2(M, T_0) \to H_2(M^\varphi, T_0)$.

5.1 Theorem.   (a) *If some M_{φ_0} is reducible then for any $\varphi \neq \varphi_0$, M_φ is irreducible and the Thurston norm x satisfies $x(\Phi(\alpha)) = x(\alpha)$ for every $\alpha \in H_2(M, T_0)$.*

(b) *If M_φ is irreducible for every φ, then for any α in $H_2(M, T_0)$ there is a framing φ_α such that if $\varphi \neq \varphi_\alpha$ then $x(\Phi(\alpha)) = x(\alpha)$.*

Proof.   Regard $(M, \partial M)$ as a sutured manifold with T_0 a suture and T_1 in R_+. Let $(S, \partial S) \subset (M, T_0)$ be an incompressible surface for which $\chi_-(S) = x(\alpha)$ and $[S, \partial S] = \alpha$. With no loss of generality we can assume that ∂S consists of a possibly empty collection of coherently oriented parallel simple closed curves.

Construct a taut sutured manifold sequence

$$(M, T_0) \xrightarrow{S_1} (M_1, \gamma_1) \xrightarrow{S_2} \cdots \xrightarrow{S_i} (W_i, \gamma_i) \xrightarrow{S_{i+1}} \cdots$$

so that $S_1 = S$, each S_i is conditioned or a collection of product disks and annuli disjoint from T_1, and S_i contains no separating closed components. By 2.6 such a sequence can be extended so long as $H_2(W_i, \partial W_i - T_1) \neq 0$, and by 4.17 the extension can be made to decrease reduced complexity. Hence eventually the sequence terminates in (M_n, γ_n) with $H_2(M_n, \partial M_n - T_1) = 0$.

Then it is an easy consequence of duality that ∂M_n is the union of T_1, at most one other torus T' and a collection of spheres.

Since M was assumed irreducible, each sphere in ∂M_n bounds a ball in M_n, and there is a single component H of M_n with $\partial H = T' \cup T_1$. Since M was assumed atoroidal, either $H = T_1 \times I$ or T' is parallel to T_0 in M. The latter is impossible, since in that case S would necessarily have been in the collar between T_0 and T', and so null-homologous in $H_2(M, T_0)$. Therefore $H = T_1 \times I$.

As in the proof of 3.9, T' cannot lie entirely in $R(\gamma_n)$, so γ_n contains a nontrivial collection of essential curves in T'. Let H^φ be the manifold obtained by attaching $D^2 \times S^1$ to H via φ. Then $H^\varphi = D^2 \times S^1$ with $\partial H^\varphi = T'$. In particular, unless φ is attached by that framing φ_0 which makes $\gamma_n \cap T'$ meridians of H^φ, the induced sutured manifold structure on H^φ is taut.

For each $1 \leq i \leq n$ let M_i^φ denote the manifold obtained by attaching $D^2 \times S^1$ to M via φ. Then for every $\varphi \neq \varphi_0$ the S_i above induce a sutured manifold sequence

$$(M^\varphi, T_0) \xrightarrow{S_1} (M_1^\varphi, \gamma_1) \xrightarrow{S_2} \cdots \xrightarrow{S_n} (M_n^\varphi, \gamma_n)$$

in which the last term is taut. Then by 3.9 (M_1^φ, γ_1) is taut, and so, by 3.3, S_1 is taut in M^φ.

5.2 Corollary. *Suppose k is a knot in $S^1 \times D^2$ of winding number zero, and some nontrivial surgery on k produces a manifold N which is homeomorphic to $S^1 \times D^2$. Then k lies in a 3-ball.*

Proof. Case 1. Any torus in $M = S^1 \times D^2 - \mathscr{N}(k)$ separating $S^1 \times \partial D^2$ from $\mathscr{N}(k)$ is ∂-parallel in M.

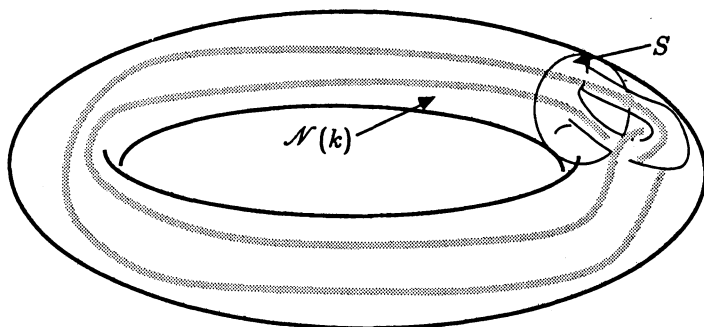
FIGURE 5.1

Proof of Case 1.   If M is ∂-reducible, then k lies in a 3-ball, so assume M is ∂-irreducible. Since K has winding number 0, there is a surface $(S, \partial S) \subset (M, S^1 \times \partial D^2)$ for which ∂S is a meridian circle $\{p\} \times \partial D^2$.

Since M is a ∂-irreducible, S cannot be a disk and so its homology class has nontrivial Thurston norm. But after filling in a solid torus at $\mathcal{N}(k)$ in two distinct ways, ∂S becomes a meridian of a solid torus. Hence $[S, \partial S]$ is representable by a meridian disk, and so has trivial norm, in both cases. This contradicts 5.1.

Case 2.   *The general case*: Suppose $T \subset M$ is any torus which separates $S^1 \times \partial D^2$ from $\mathcal{N}(k)$. Then T is compressible in $S^1 \times D^2$ and N. If there is a compressing disk in the component of $M - T$, which contains $S^1 \times \partial D^2$, then after the compression T bounds a 3-ball containing k. Assume then, on the contrary, that each torus $T \subset M$ which separates $S^1 \times \partial D^2$ from $\mathcal{N}(k)$ compresses in the component of $S^1 \times D^2 - T$ and of $N - T$ which contain $\partial \mathcal{N}(k)$, so that the component is in both cases a solid torus. Then a maximal collection of pairwise nonparallel tori in M bounds in $S^1 \times D^2$ a nested collection of solid tori $\mathcal{N}(k) = W_0 \subset W_1 \subset \cdots \subset W_n = S^1 \times D^2$. (They bound a similar nested collection in N.) Let w_i denote the winding number of W_i in W_{i+1}. Then the product of all w_i is 0, the winding number of k in $S^1 \times D^2$. Let $W_i \subset W_{i+1}$ be the innermost successive pair for which $w_i = 0$. Then $W_{i+1} - \overset{\circ}{W}_i$ is an irreducible, ∂-irreducible cobordism between two tori such that filling a solid torus on ∂W_i with two distinct framings gives a solid torus. (The framings are distinct because the winding number of k in W_i is nontrivial.) The proof then proceeds as in Case 1.

6.   Reduction lemmas

In 4.3 and 4.5 we showed that under certain conditions components of β can be removed without affecting β-tautness. The purpose of this section is to develop a few more examples of such reduction lemmas. The examples shown here are specifically designed for application in §8. Different applications may call for different reduction lemmas.

6.1 Definition.   Let $(Q, \partial Q)$ be a surface in (M_β, γ_β) in normal position. A *β-loop* in Q at point b in β is a proper arc α essential in $(Q, \partial Q \cap \mathcal{\dot{N}}(\beta))$ such that both ends of α lie on the meridian disk of $\mathcal{N}(\beta)$ at b. A disk D in M_β whose boundary is the union of the β-loop and an arc of the meridian at b is called a *β-loop compressor*.

6.2 Y-reduction.   Suppose (M, γ, β) is a sutured manifold for which a component β_0 of β is a tree with three ends: ε_1 and ε_2 in R_+ and a third in R_-.
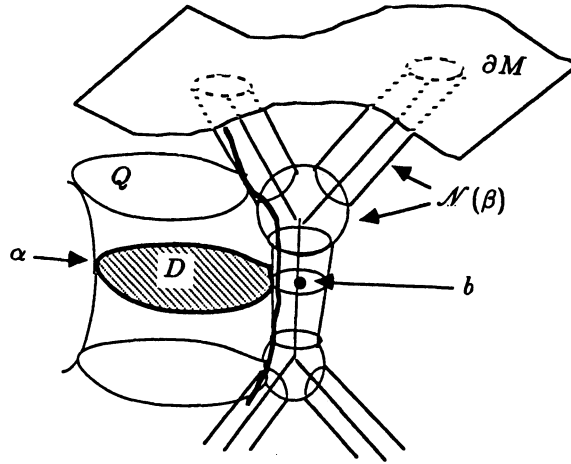
FIGURE 6.1

Suppose further that there is a surface $(Q, \partial Q) \subset (M_\beta, \partial M_\beta)$ with the properties:

    (i) Q is a disk.

    (ii) ∂Q lies entirely in $R_+ \cup \partial \mathscr{N}(\beta_0)$.

    (iii) Each component of $\partial Q \cap \partial \mathscr{N}(\beta_0)$ is an arc with one end at ε_1 and the other at ε_2.
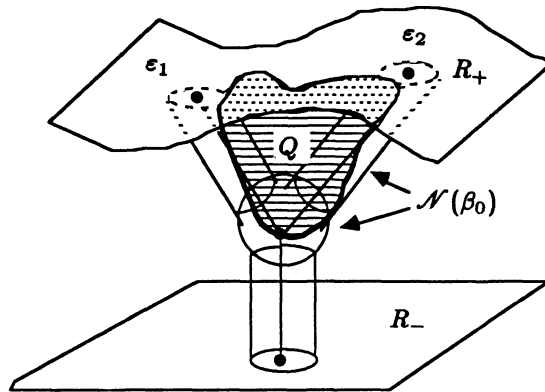

FIGURE 6.2

    Lemma.   *If (M, γ) is β-taut then either $\partial Q \subset R_+$, or there are compressible β-loops for Q at both ε_1 and ε_2.*

Proof. View β_0 as the union of an arc δ_+ with ends at ε_1 and ε_2 in R_+ and an arc δ_- with one end on δ_+ and the other on R_-. Then there is a tubular neighborhood $\eta \subset \mathcal{N}(\beta_0)$ of δ_+ which intersects δ_- (hence β) in a single point and whose complement in $\mathcal{N}(\beta_0)$ is a neighborhood of the arc $\delta_- - \eta$. Let S be the manifold obtained from R_+ by 1-surgery on η. Then $\chi_\beta(S) = \chi_\beta(R_+) + 1$. Since $\partial Q \subset R_+ \cup \partial\mathcal{N}(\beta_0)$, Q can be isotoped so that $\partial Q \subset S$.
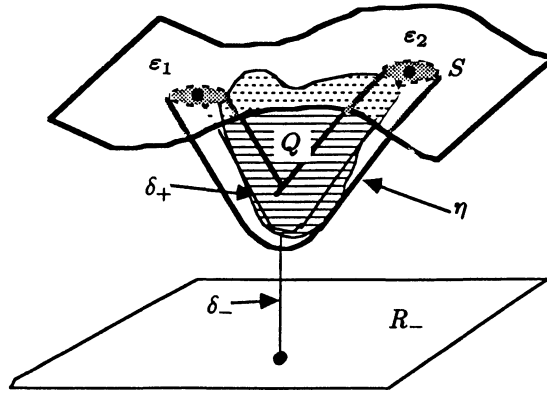

FIGURE 6.3

If ∂Q were essential in S or bounded a disk intersecting β more than once, then either (a) 2-surgery on S via Q would produce a surface S' with $\chi_\beta(S') = \chi_\beta(S) - 2 = \chi_\beta(R_+) - 1$, contradicting the β-tautness of R_+, or (b) $\chi_\beta(R_+) = 0$. In the latter case either S' is a sphere intersecting β in a single point or a component of $\partial Q \cap R_+$ is a compressible β-loop. So suppose ∂Q bounds a disk F in S with $|F \cap \beta| \leq 1$.

Let C denote the boundary of a meridian disk of $\mathcal{N}(\beta)$ at ε_1, say. Then some outermost arc of C in F cuts off a disk F' from F disjoint from β, and $\partial F' \cap \partial F$ is a β-loop of Q at ε_1 for which F' is a β-loop compressor.

6.3 Noose reduction. Suppose (M, γ) is a sutured manifold for which a component β_0 of β is a "noose", i.e., a circle at the end of an arc, with its end ε in R_+.

Suppose further that there is a surface $(Q, \partial Q) \subset (M_\beta, \partial M_\beta)$ with the properties:

    (i) Q is a disk.

    (ii) ∂Q lies entirely in $R_+ \cup \partial\mathcal{N}(\beta_0)$.

    (iii) Each component of $\partial Q \cap \partial\mathcal{N}(\beta_0)$ is an arc (with its end at ε).
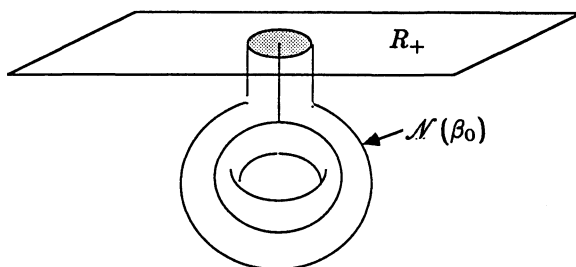
FIGURE 6.4

Lemma.   *If (M, γ) is β-taut, then either $\partial Q \subset R_+$ or there is a compressible β-loop for Q at any point in β.*

Proof.   The proof is an easy analogue of 6.2, where for η we take all of $\mathcal{N}(\beta_0)$.

6.4 H-reduction.   Suppose (M, γ, β) is a sutured manifold, with ∂M a union of spheres. Suppose that β is H-shaped; that is, β consists of three segments β_l, β_r and β_c with the following properties:

(a) β_l and β_r each have one endpoint (ε_l^+ and ε_r^+ respectively) in R_+ and one endpoint (ε_l^- and ε_r^- respectively) in R_-.

(b) β_c has one endpoint in β_l and one endpoint in β_r.

Suppose further that there is a surface $(Q, \partial Q) \subset (M_\beta, \partial M_\beta)$ consisting of three disk components Q_l, Q_r and Q_0 with the properties:

(i) Q_l (resp. Q_r) is a disk which would cancel β_l (resp. β_r) into a suture on γ, if β_l (resp. β_r) had been an arc component of β (see 4.1).

(ii) ∂Q_0 lies entirely in $R_+ \cup \partial \mathcal{N}(\beta)$.

(iii)   Each component of $\partial Q \cap \partial \mathcal{N}(\beta)$ is an arc with its ends at distinct ends of $\mathcal{N}(\beta)$.
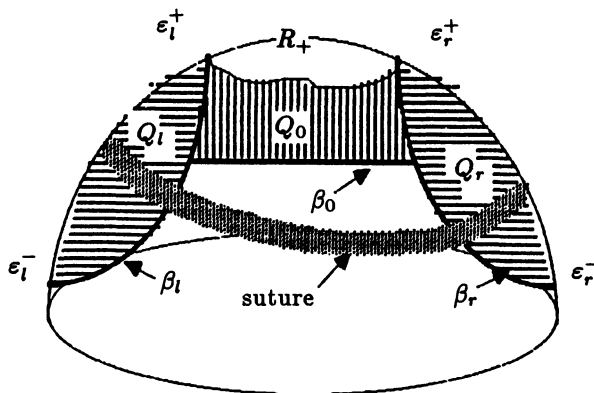

FIGURE 6.5

Lemma.   *If* (M, γ) *is* β-*taut, then either*

(a) $\partial Q_0 \subset R_+$,

(b) *there are compressible* β-*loops for* Q_0 *at* ε_l^+ *and* ε_r^+, *or*

(c) (M, γ) *is taut* (*in the Thurston norm*).

Proof.   **Claim 1.**   With no loss of generality, every disk component of $R(\gamma)$ contains an endpoint of β.

Proof of Claim 1.   If there is a disk component C of $R(\gamma)$ which is disjoint from β, then the other component C' of $R(\gamma)$ whose boundary lies on the same annular suture as ∂C must also be a disk disjoint from β, since $R(\gamma)$ is β-incompressible. Then since M is β-taut (hence β-irreducible), $C \cup C'$ bounds a ball in M disjoint from β, and so this component is taut in the Thurston norm.

Claim 2.   No arc of $\partial Q_0 \cap R_+$ has both ends at ε_l^+ or both at ε_r^+.

Proof of Claim 2.   Such an arc α at ε_l^+, say, would divide the component of ∂M in which it lies into two disks. If either disk is disjoint from β, then it is a β-loop compressor for α. Hence one disk D' contains two ends of β and the other disk D only one. Moreover, β_r cannot have one end in D and the other in D', since Q_r is disjoint from Q_0. Hence ε_l^- lies in D. Pushing D into M along the disk Q_l then produces a β-loop compressor for α.

Conclusion of Proof.   From Claim 2 it follows that every arc of $\partial Q_0 \cap R_+$ runs between an end of β_l and an end of β_r. There is a component C_0 of R_+ containing two ends of β and components C_l and C_r (possibly identical) containing respectively a single end of β_l and a single end of β_r. Each of the latter shares a common suture with C_0. By Claim 1 at least two of these components must be disks, and ∂M is a single 2-sphere. If C_0 is not a disk, then C_l and C_r are both disks, and C_0 is an annulus sharing a common suture with each of C_l and C_r. Then $\chi_\beta(R_-) = 0 \neq 2 = \chi_\beta(R_+)$, contradicting the tautness of R_+. Hence C_0 is a disk and, since either C_l or C_r is also a disk and shares a common suture with C_0, ∂M must have a single suture, and so is taut in the Thurston norm.

7.   Parametrizing surfaces

Here we show how to carry information about the topology of a sutured manifold down through a hierarchy. The idea is to start with a surface Q in M_β and follow it as it is cut up during a sutured manifold decomposition. The reducing lemmas of §§4 and 6 can then be applied to its remnants in the last term of the decomposition to eliminate components of β.

7.1 Definition. Let (M, γ, β) be a sutured manifold. A *parametrizing surface* $(Q, \partial Q) \subset (M_\beta, \partial M_\beta)$ is a surface such that no component of Q is a disk with boundary in R_+.

7.2 Definition. Let $(Q, \partial Q)$ be a parametrizing surface for a β-taut sutured manifold (M, γ) and suppose $(S, \partial S) \subset (M, \partial M)$ is a β-incompressible d-surface. Then S is *normalized* if the following hold:

(a) ∂S and ∂Q have been isotoped on $T(\gamma)$ and near $A(\gamma)$ so that the number of points in $\partial S \cap \partial Q$ lying in $A(\gamma) \cup T(\gamma)$ is as small as possible. (Thus, for example, no arc (circle) component of ∂Q in $A(\gamma)$ intersects any arc (resp. circle) component of ∂S in $A(\gamma)$.)

(b) Every simple closed curve of $S \cap Q$ is essential in Q.

7.3 Lemma. *Let $(Q, \partial Q)$ be a parametrizing surface for a β-taut sutured manifold (M, γ) and suppose $(S, \partial S) \subset (M, \partial M)$ is a β-incompressible d-surface. Then there is an isotopy normalizing S. Furthermore restricted to any component of $A(\gamma)$, $T(\gamma)$ or $R(\gamma)$, the isotopy of ∂S, is proper.*

Proof. (a) is immediate.

The proof of (b) is by induction on the number of simple closed curves in $S \cap Q$. It is a standard innermost disk argument, utilizing the β-incompressibility of S and the β-irreducibility of M.

7.4 Definition. Let Q be a parametrizing surface with ν the number of essential arcs in $\partial Q \cap A(\gamma)$, and μ the number of essential arcs in $\partial Q \cap \mathcal{N}$ (edges of β). Let κ be a function which assigns a real number to each arc λ in $\partial Q \cap \mathcal{N}(\mathcal{V}_\beta)$ and $\mathcal{K} = \sum_\lambda \kappa$. ($\kappa$ can be chosen arbitrarily, but a good choice can help exposition.) Define the index of Q as

$$I(Q) = \nu + \mu + \mathcal{K} - 2\chi(Q).$$

7.5 Lemma. *Let $(Q, \partial Q)$ be a parametrizing surface for a β-taut sutured manifold (M, γ) and suppose $(S, \partial S) \subset (M, \partial M)$ is a conditioned surface which decomposes (M, γ) into a β-taut sutured manifold. Let (M', γ', β') be the sutured manifold obtained by decomposing M along $S_{k,l}$ (cf. 3.7). For k and l sufficiently large and $S_{k,l}$ normalized with respect to Q, the surface $Q' = Q - \mathcal{N}(S_{k,l})$ is a parametrizing surface in (M', γ', β') with $I(Q') = I(Q)$.*

Proof. Note that (M', γ') is β'-taut by 3.8.

Claim 1. For k, l sufficiently large, and $S_{k,l}$ normalized with respect to Q, no component of Q' is a disk with boundary in R'_+.

Proof of Claim 1. Let $\varepsilon_+ \subset \partial Q$ denote the 1-manifold $R_+ \cap \partial Q$ and ε_- the 1-manifold $R_- \cap \partial Q$. Every component of $\varepsilon_\pm$ has a transverse orientation inherited from that of $R_\pm$. Choose in Q disjoint collar structures $\varepsilon_\pm \times I$ near $\varepsilon_\pm$ so that $S \cap (\varepsilon_\pm \times I)$ consists entirely of a collection of I fibers, and the normal field to $\varepsilon_\pm$ in the collar is always tangent to the I-fibers.
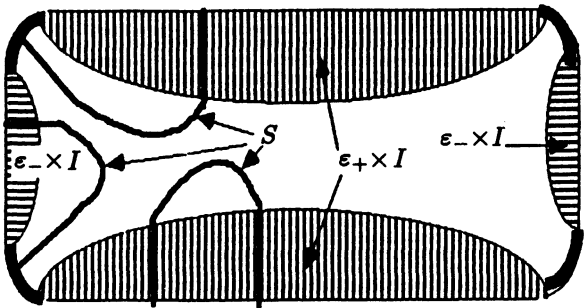
MARTIN SCHARLEMANN

FIGURE 7.1
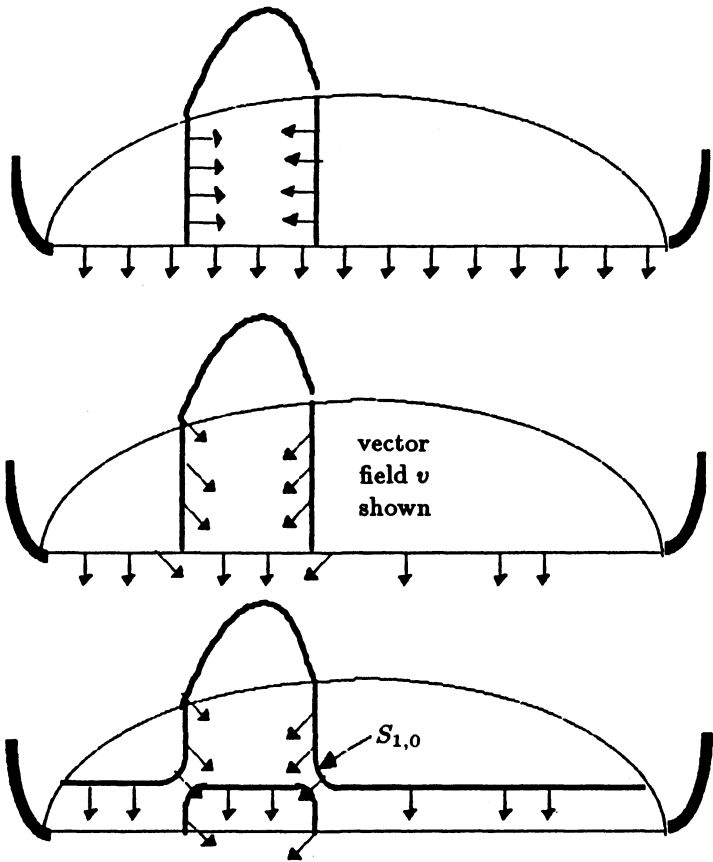

FIGURE 7.2

Using a partition of unity, average together near $S \cap (\varepsilon_\pm \times I)$ the vector field tangent to the I-fibers and the normal vector field to S in Q. The result is a nontrivial vector field v on $\varepsilon_\pm \times I$ which is transverse to $S_{k,l} \cap (\varepsilon_\pm \times I)$ for all $k, l \geq 0$. The double curve sums used in the construction of $S_{k,l}$ will be taken to lie within these collars, so that for all $k, l \geq 0$ $(S_{k,l} \cap Q) - (\varepsilon_\pm \times I) = (S \cap Q) - (\varepsilon_\pm \times I)$.

The proof of Claim 1 is now by induction on the number p of components of $(S \cap Q) - (\varepsilon_\pm \times I) = (S_{k,l} \cap Q) - (\varepsilon_\pm \times I)$. Suppose a component of Q' is a disk D with $\partial D \subset (R_+ \asymp S_{k,l})$, say. Then ∂D lies entirely inside of $S_{k+1,l}$ and so is removed when $S_{k+1,l}$ is normalized. Since ∂D cannot lie entirely in R_+, ∂D must contain an arc of $(S_{k,l} \cap Q) - (\varepsilon_\pm \times I)$. Then, after normalization, $|(S_{k+1,l} \cap Q) - (\varepsilon_\pm \times I)| < p$. Continue, if necessary, until $|(S_{k',l'} \cap Q) - (\varepsilon_\pm \times I)| = 0$. Then, if D is a disk in $Q - \mathcal{N}(S_{k',l'})$ the argument just given shows $\partial D \subset \varepsilon_\pm \times I$. It follows that $D \subset \varepsilon_\pm \times I$. But this is impossible, for v would then be a normal vector field to ∂D extending to a vector field over all of D.

Claim 2. $I(Q') = I(Q)$.

Proof of Claim 2.   In general $S_{k,l} \cap Q$ consists of a disjoint collection of arcs and circles in Q, but it suffices to assume $S_{k,l} \cap Q$ is connected.

If $S_{k,l} \cap Q$ is a circle, then no term in the definition of $I(Q)$ is changed when $\mathcal{N}(S_{k,l})$ is removed.

If $S_{k,l} \cap Q$ is an arc α then $\chi(Q') = \chi(Q) + 1$. The set of components λ of $\partial Q' \cap \mathcal{N}(\mathscr{V}_\beta)$ is the same as that in $\partial(Q) \cap \mathcal{N}(\mathscr{V}_\beta)$, since $S_{k,l}$ is disjoint from $\mathcal{N}(\mathscr{V}_\beta)$ so $\mathscr{K} = \sum_\lambda \kappa$ is unchanged. If an end of α lies in $\mathcal{N}(\beta)$ then μ for Q' is 1 greater than μ for Q. If an end of α lies in $A(\gamma)$, $T(\gamma)$ or $R_\pm$ (hence in general in ∂M), then one side or the other of the component of ∂S in which α terminates becomes a suture of γ'. Hence ν for Q' is 1 greater than ν for Q. Thus in general $\nu + \mu$ for Q' is 2 greater than $\nu + \mu$ for Q. By 7.5, the index I remains unchanged.

7.6.   Let $(Q, \partial Q)$ be a parametrizing surface for a β-taut sutured manifold (M, γ), and suppose $(S, \partial S) \subset (M, \partial M)$ is a product disk or β-incompressible product annulus normalized with respect to Q. Let Δ be the collection of arcs of $S \cap Q$ which have both ends in either $R_+(\gamma)$ or $R_-(\gamma)$. Do 2-surgeries on Q via outermost disks of S, and cut off by the arcs of Δ until $\Delta = \varnothing$; then remove all disks of Q whose boundaries lie entirely in $R_\pm(\gamma)$. Call the result Q_c. Let (M', γ', β') be the sutured manifold obtained by decomposing M along S.

Lemma.   *The surface* $Q' = Q_c - \mathcal{N}(S)$ *is a parametrizing surface in* (M', γ', β') *with* $I(Q') \leq I(Q)$.

Proof. Suppose there were a disk component D of Q' with its entire boundary in $R_+(\gamma')$. By construction ∂D must intersect S_+, the side of $\mathcal{N}(S)$ lying in $R_+(\gamma)'$. If it lay entirely in S_+, it would contradict the normalization of S with respect to Q. If only one part lay in S_+ then, since ∂D does not cross the suture lying at one end of S_+, it would contain an arc in S_+ with both ends on the other end of S_+, contradicting the construction of S_c. Hence Q' is a parametrizing surface.

Since Q is a parametrizing surface, no disk removed in the construction of Q_c was originally in Q. Hence the number of disks removed is no larger than Δ. Thus $\chi(Q_c) \geq \chi(Q)$, so $I(Q') \leq I(Q)$.

7.7 Definition. A β-taut sutured manifold hierarchy for (M, γ, β) *respects a parametrizing surface* Q if at each stage in the hierarchy either S_i is a product disk or nontrivial product annulus and Q_i is constructed from Q_{i-1} as in 7.6, or S_i is a conditioned surface and the surface $Q_i = Q_{i-1} \cap M_i$ is still a parametrizing surface. Notice that this guarantees that always $\partial Q_i \cap \mathcal{N}(\beta_i) = \partial Q \cap \mathcal{N}(\beta_i)$ and $I(Q_i) \leq I(Q)$.

7.8 Theorem. *If Q is a parametrizing surface for the β-taut sutured manifold (M, γ), then there is a β-taut sutured manifold hierarchy for (M, γ) which respects Q.*

Proof. The proof is similar to that of 4.19, with 3.8, 7.5, and 7.6 applied so that each successive Q_i remains a parametrizing surface.

8. Application II: Genus of knots is supperadditive under band sum

Let β be the planar "eyeglass" 1-complex consisting of two circles β_l and β_r (bounding disjoint disks in the plane) and an arc β_c joining them. Regard β as a complex in $\mathbb{R}^2 \subset \mathbb{R}^3$, and let P and W be regular neighborhoods of β in $\mathbb{R}^2$ and $\mathbb{R}^3$ respectively, so that P is a properly imbedded planar surface in the genus two handlebody W. $\partial P \subset \partial W$ has three components: Two of them, denoted c_l and c_R, are parallel in P to β_l and β_r respectively. Denote the third by c_0.

8.1. Suppose β is imbedded in a 3-manifold M for which ∂M is a torus. Regard W as a regular neighborhood of β in M. Suppose that M is β-irreducible and there is an incompressible, ∂-incompressible surface $(Q, \partial Q) \subset (M - \overset{\circ}{W}, \partial W)$ consisting of 3-components: Q_r, Q_l, and Q_0 with the following properties:

   (a) Q_r (Q_l) is an annulus for which one boundary component is $c_r \subset \partial W$ $(c_l \subset \partial W)$, and the other is an essential curve in the torus ∂M.

FIGURE 8.1

(b) Q_0 is a planar surface, all of whose ∂-components are parallel to c_0 in ∂W.

(c) There is no compressible β-loop (6.1) in Q at a point on β_l (β_r).

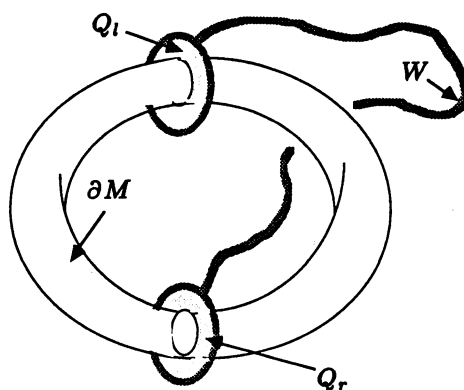(d) No sphere intersects β in a single point of β_c.

FIGURE 8.2

8.2 Theorem. *If $(S, \partial S) \subset (M, \partial M)$ is β-taut, then $S \cap \beta_c = \emptyset$ and S is taut in the Thurston norm.*

Proof. Let S be such a surface, and construct a β-taut sutured manifold hierarchy

$$(M, \partial M) \xrightarrow{S_1 = S} (M_1, \gamma_1) \xrightarrow{S_2} \cdots \longrightarrow (M_n, \gamma_n)$$

respecting Q. Following 2.4, the edges β_r, β_l and β_c can be oriented so that at any point of intersection with an S_i (hence with $R(\gamma_i)$) the orientation points in the direction of the normal vector to S_i. For each arc λ of $\partial Q \cap \mathcal{N}(\mathcal{V})$ define $\kappa(\lambda) = -1$ if λ passes between an edge of β pointing into the vertex and one pointing out, and define $\kappa(\lambda) = -2$ if λ passes between edges of β both pointing into the vertex or both pointing out.
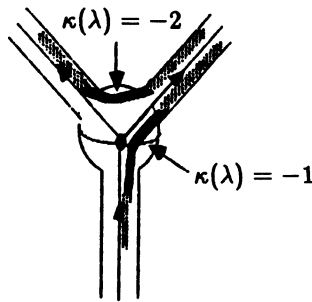

FIGURE 8.3

Note that the curve $c_0 \subset \partial W$ has, in the terminology of 7.4, $\nu = 0$, $\mu = 4$, and $\kappa = -6$, and the curves c_r and c_l each have $\nu = 0$, $\mu = 1$, $\mathcal{K} = -1$. Hence if Q_0 has p boundary components, $\chi(Q) = 2 - p$ and so $I(Q) = -2p + 2p - 4 = -4$.

Claim 1. $S_i \cap \beta_c = \varnothing$ for all i.

Suppose not, and let $k \geq 0$ be the smallest integer for which $S_k \cap \beta_c \neq \varnothing$. Denote by η_k the regular neighborhood $W \cap M_k$ of the 1-complex $\beta \cap M_k$. There are two components of η_k, which contain vertices. These components are either Y-shaped or noose-shaped (see §6), depending on whether or not β_l and β_r have been cut by some S_i. For the remainder of the argument we will assume they are Y-shaped, for if they are noose-shaped the only change would be the application of 6.3 instead of 6.2. Let η_+ (resp. η_-) denote the component with two ends in R_+ (resp. R_-) and one end in R_- (resp. R_+).

Every component of $\partial Q_0 \cap \dot{\eta}_\pm$ in M_k is an arc passing through two edges and one vertex, so for any such component $\mu + \kappa \geq 0$. Moreover, the inequality is strict unless $\kappa = -2$, i.e., both ends of the arc lie in either R_+ or R_-. Any component of ∂Q_0, which contains an arc in $\eta_\pm$ with one end in R_+ and one end in R_-, must contain either another such arc or an arc crossing a suture,

so for the component $\nu + \mu + \mathcal{K} \geq 2$. Since $I(Q_k) \leq I(Q) = -4$, it follows from 7.4 that some component q of Q_k is a disk in which ∂q never passes from R_+ and R_-. With no loss of generality we may then assume that ∂q is contained in $R_+ \cup \eta_+$.

Since q is a part of a parametrizing surface, ∂q does not lie in R_+ (7.1). Then 6.2 applied to η_+ and q produces a compressible β-loop for Q_k (hence for Q) at a meridian of β_l or β_r. This contradicts 8.1(c).

Claim 2. No product annuli coming from self-amalgamations need to be used in the above hierarchy.

Proof of Claim 2. Since β_c is left intact during the decompositions, the only arc components of β produced during the decomposition are pieces of β_l or β_r, but these are immediately cancellable via the remnants of Q_l or Q_r.

Claim 3. S is taut in the Thurston norm.

Proof of Claim 3. $\partial M_n \cap \beta \neq \varnothing$ since M is β-irreducible. Hence for some i, $S_i \cap \beta \neq \varnothing$. By the previous claim, S_i then intersects β_r or β_l hence Q_r or Q_l. It follows from the β-tautness of S_i that each arc of $S_i \cap Q_r$ and $S_i \cap Q_l$ must have one end in ∂W and one end in ∂M. Hence S_i intersects one of the circles $Q_r \cap \partial M$ or $Q_l \cap \partial M$. These circles must be parallel in the torus ∂M, since Q_r and Q_l are disjoint, so S_i must intersect *both* $Q_r \cap \partial M$ and $Q_l \cap \partial M$. Reversing the argument, it follows that S_i intersects both β_r and β_l, so β_n is a 1-complex which is the union of arcs parallel to sutures (via the remnants of Q_l and Q_r in M_n) and an H-shaped component as in 6.4.

Then, as in Claim 1, the component q of Q_n of negative index is a disk as in (ii) and (iii) of 6.4. Components of $Q_l \cap M_n$ and $Q_r \cap M_n$ satisfy 6.4(i). Again ∂q does not lie entirely in $R(\gamma_n)$. Then from 6.4 we conclude that (M_n, γ_n) is taut in the Thurston norm.

Since each S_i is either conditioned or a product disk, and contains no separating closed components, and M_1 is not a solid torus, it follows from 3.9 that (M_1, γ_1) is taut in the Thurston norm. Thus 3.3 (with $\beta = \varnothing$) implies S is taut in the Thurston norm.

8.3 Definition. For knots K_0 and K_1 in S^3, let the *band sum* of K_0 and K_1 be defined as follows: Imbed K_0 and K_1 into disjoint 3-balls in S^3. Let $b: I \times I \to S^3$ be an imbedding such that $b^{-1}(K_i) = I \times \{i\}$, $i = 0, 1$. Then join $K_0 - b(I \times \{0\})$ to $K_1 - b(I \times \{1\})$ by the arcs $b(\partial I \times I)$. The resulting knot is the band sum of K_0 and K_1 along b and is denoted $K_0 \#_b K_1$.

8.4 Theorem. $\mathrm{genus}(K_0 \#_b K_1) \geq \mathrm{genus}(K_0) + \mathrm{genus}(K_1)$.

Proof. Let λ be a cocore of the band b. It is an arc with both ends on $K = K_0 \#_b K_1$. A sphere S^2 separating K_0 from K_1 in S^3 intersects the band b in arcs parallel to λ. Choose S^2 so that $|S^2 \cap K|$, and hence the number of
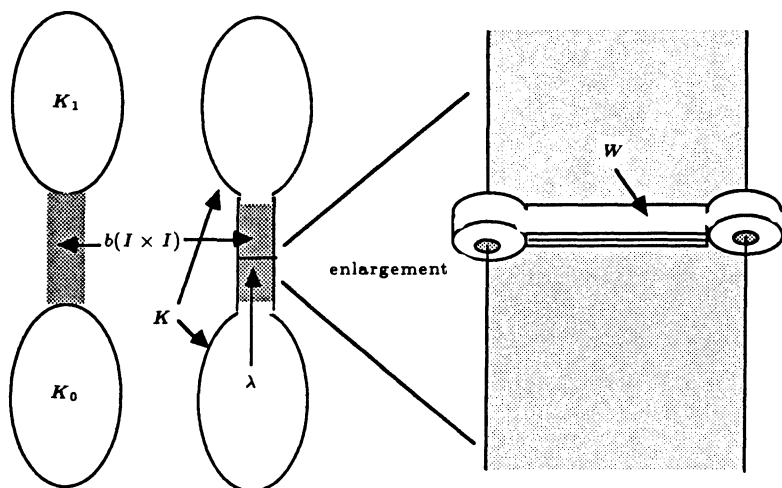
MARTIN SCHARLEMANN

FIGURE 8.4

these arcs is as small as possible among such separating spheres. (The number is odd, hence never trivial.)

Let η be a tubular neighborhood of K. Let $\beta_c = \lambda - \eta$ and β_l, β_r be disjoint meridians of η, which pass through the ends of β_c. The union of the band $b(I \times I) \cap M$ and disjoint collar neighborhoods of β_l and β_r is naturally homeomorphic to $\beta \times I$, and those collar neighborhoods may be chosen so that they contain $S^2 \cap \partial \eta$. Let W be a regular neighborhood of $\beta \times I$ and Q_0 be the planar surface obtained from S^2 by removing $\eta \cup W$. Let Q'_r and Q'_l be meridians of η. Now let η' be a tubular neighborhood of K contained in η, $M = S^3 - \eta'$, $Q_r = Q'_r - \eta'$ and $Q_l = Q'_l - \eta'$.

Let S be a β-taut Seifert surface for $K_0 \#_b K_1$ in M. Then by 8.2, S is disjoint from β_c, hence from the band, and is taut in the Thurston norm. The latter implies genus(S) = genus$(K_0 \#_b K_1)$. Since S is disjoint from the band, we may attach $b(I \times I)$ to S and obtain a surface S' with genus(S') = genus(S). Furthermore $\partial S'$ is the union of a longitude of K_0 and a longitude of K_1. Then a standard innermost disk argument shows that S' may be isotoped so that every circle of intersection with S^2 is essential in S'. After some compressions of S' it will be made disjoint from S^2 and so the union of disjoint Seifert surfaces for K_0 and K_1.

8.5 Remark. In case of equality, all the last compressions must be along parallel simple closed curves in S'. But in that case, we can assume all the compressing disks lie on one side of S'. Push the band, which lies in S', off

to the other side. This makes the band disjoint from minimal genus Seifert surfaces for K_0 and K_1.

8.6 Corollary ([11], [13]).   *If $K_0 \#_b K_1$ is the unknot, then each of K_0 and K_1 is the unknot, and the band sum is a connected sum.*

We wish to generalize 8.6 to a more general statement about links. This can be done via the combinatorial route of [11] but also follows with ease from the previous argument.

8.7 Definition.   A *complete system of spheres* Q in a 3-manifold M is an imbedded collection of incompressible, pairwise nonparallel spheres so that any other incompressible sphere in $M - Q$ is parallel to a component of Q (cf. [8, 3.14]). An example: If M is the connected sum of k irreducible 3-manifolds, then any complete system contains exactly $2k - 3$ components.

8.8 Definition.   For L an oriented link in S^3, a *band operation* on L is defined as follows: Let $b\colon I \times I \to S^3$ be an imbedding such that $b^{-1}(L) = I \times \partial I$ with the orientation inherited from $\partial(I \times I)$. Then $L - b(\overset{\circ}{I} \times I)$ is a link denoted by L_b.

8.9 Proposition.   *Let $L' = L_b$ for some band operation on the link L. There is either a complete system Q for $S^3 - L$ disjoint from the band b or a Seifert surface S for L containing $b(I \times I)$ for which $S - b(\overset{\circ}{I} \times I)$ is a taut Seifert surface for L'.*

Proof.   Choose the complete system Q to minimize $|b \cap Q|$.

Claim 1.   With no loss of generality, we can assume b intersects every component of Q.

Proof of Claim 1.   Let S' be a taut Seifert surface for L'. Let Q' be the union of all components of Q which are disjoint from b. Then $Q' \cap L' = \varnothing$. Since S' is taut, every component of $S' \cap Q'$ is inessential in S'. Thus S' can be altered by 2-surgery so that it is still taut and disjoint from Q'. Discard Q' and all components of L' which lie in the components of $S^3 - Q'$ not containing b. If the theorem is true for the remaining links, it will be true for the original links.

Let $M = S^3 - L'$, and let β be the eyeglass 1-complex imbedded in M as in the proof of 8.4.

Claim 2. M is β-irreducible.

Proof of Claim 2.   Suppose there were a sphere in M, disjoint from β, bounding no 3-ball disjoint from β. Among all such spheres, let P be one for which $|P \cap Q|$ is minimal. Let Q' be obtained from Q by 2-surgery along an innermost disk at P. Since no component of Q' is disjoint from β, neither is any component of Q'. Moreover, $|Q' \cap b| = |Q \cap b|$, and Q' contains a

complete system of spheres with one fewer component. This contradicts our assumption that Q minimizes $|Q \cap b|$.

Now the argument proceeds just as in the case of 8.4. Ultimately there is a Seifert surface S' for L', which is taut and disjoint from $b(\overset{\circ}{I} \times I)$. Then $S = S' \cup b(I \times I)$ is the required Seifert surface.

8.10 Corollary. *If $L' = L_b$ is the unlink, then either b is disjoint from some complete system of spheres for $S^3 - L$ or L bounds a collection of disks disjoint from b (i.e., L is the unlink and b is the trivial connected sum).*

Proof. If b fails to be disjoint from some complete system for L then, by 8.9, L' bounds a collection of disks S' which are disjoint from the band.

If the components of $b(\partial I \times I)$ lie on different components of L', then $S = S' \cup b(I \times I)$ is a collection of disks whose boundary is L. Thus L is the unlink and b clearly is disjoint from a complete system of spheres for L.

Suppose then the components of $b(\partial I \times I)$ lie on the same component l of L'. The union of $b(I \times I)$ and the component of S' that l bounds is an annulus A such that $A \cap L = \partial A$. If any component of $Q \cap A$ is a closed curve inessential in A, then 2-surgery on the disk bounded by an innermost such curve provides a collection of 2-spheres containing a complete system having no more intersections with b and fewer intersections with A. So we can assume $Q \cap A$ contains only essential curves in A. Then 2-surgery on the disk in Q bounded by an innermost circle of $Q \cap A$ changes $S = S' \cup b(I \times I)$ into a collection of disks, disjoint from b, whose boundary is L.

8.11 Corollary. *If $L' = L_b$, and both L and L' are the unlink, then b is a trivial band (i.e., after an isotopy, L, L', and b all lie in a plane).*

Proof. If the components of $b(I \times \partial I)$ lie on different components of L, then b intersects any complete system of spheres for $S^3 - L$. Thus by 8.10 the band is trivial.

If the components of $b(I \times \partial I)$ lie on the same component of L, then the components of $b(\partial I \times I)$ lie on different components of L'. Switch the roles of L and L', and this case follows from the previous one.

9.   Application III: Satellite knots have property P

Let $(M, \partial M)$ be a 3-manifold with ∂M a torus and $H_1(M)$ torsion free. Let β be a knot in M with tubular neighborhood $\mathcal{N}$. Suppose M is β-irreducible, and the manifold M^φ, obtained from M by performing some nontrivial surgery on β, also has $H_1(M^\varphi)$ torsion-free. Suppose finally that M^φ is ∂-reducible.

9.1 Theorem. *Any β-taut surface $(S, \partial S) \subset (M, \partial M)$ is taut in the Thurston norm.*

9.2 Corollary. *If k is a knot contained in the tubular neighborhood $K \times D^2$ of a knot K in S^3, then either*

(a) *k is contained in a 3-cell in $K \times D^2$,*

(b) *k is isotopic in $K \times D^2$ to $K \times \{0\}$,*

(c) *K is unknotted in S^3, or*

(d) *no manifold obtained from S^3 by surgery on k is a homotopy 3-sphere.*

Proof of 9.2 from 9.1. If the winding number $\omega(k) > 1$, then the corollary follows immediately from [2]. So henceforth suppose $\omega(k) = 1$ or 0.

Take for M the tubular neighborhood $K \times D^2$, let β be the knot k, and for S take a β-taut surface whose boundary is a meridian $\{\text{pt.}\} \times \partial D^2$ of $K \times \partial D^2$ and whose geometric intersection with β is its algebraic intersection. Suppose a manifold obtained by surgery on k is a homotopy 3-sphere. Then $K \times \partial D^2$ is compressible in the manifold, so either K is unknotted in S^3 (condition (c)) or M^φ is ∂-reducible and $H_1(M^\varphi)$ is torsion free. In the latter case the theorem applies, and S is taut in M, hence is a meridian disk. Then k intersects a meridian of $K \times D^2$ always with the same orientation. If $\omega(k) = 0$ then (a) holds. If $\omega(k) = 1$ then k is the knot sum of K and the arc $k - \mathcal{N}(S)$ in the 3-ball $(K \times D^2) - \mathcal{N}(S)$. If the latter is unknotted (b) holds; if it is knotted, k is a composite, so (d) holds [1].   q.e.d.

The proof of 9.1 relies on Lemmas 9.3 and 9.6 which give criteria for detecting homology torsion in 3-manifolds.

9.3. Let M be a β-irreducible 3-manifold, β a simple closed curve in M with neighborhood $\mathcal{N}(\beta)$, and $(Q, \partial Q) \subset (M_\beta, \partial M_\beta)$ a planar surface whose boundary is a collection of p nonmeridional essential curves in $\partial \mathcal{N}(\beta)$ and at most one other curve in ∂M.

Lemma. *Suppose there is a disk D in M, which intersects β nontrivially in the interior of D so that all points of intersection of β with D have the same orientation. Then either*

(a) *D contains a compressing or ∂-compressing disk for Q in M_β,*

(b) *Q is homologous (rel the curve $Q \cap \partial M$ if it exists) to a planar surface Q' with $\partial Q'$ a proper subset of ∂Q, and the manifold $\tilde{M}$ obtained by performing Dehn surgery on β with slope that of ∂Q contains a Lens-space summand, or*

(c) *$|Q \cap \partial D| \geq p$.*

Proof. Label the boundary components of ∂Q in $\partial \mathcal{N}(\beta)$ with the label $1, \cdots, p$ in order around the torus. We assume $Q \cap D$ contains no β-inessential circles in D, for if an innermost circle is also β-inessential in Q it can be removed by an isotopy of Q, and if it is essential (a) is satisfied.

Consider the graph Γ in D^2 whose vertices are the points of $\beta \cap D$ and whose edges are the arcs of $D \cap Q$ (where some ends of edges lie on ∂D, not just at the vertices). Each end of each arc on a vertex has a label in $\{1, \cdots, p\}$

corresponding to the component of ∂Q to which it belongs. All labels appear in the same order (always counterclockwise or always clockwise) some fixed number of times (depending on the slope of ∂Q on β) around each vertex, since β intersects D always with the same sign.

Either (c) or there is some label $x \in \{1, \cdots, p\}$ for which every edge of Γ with end labelled x has its other end lying on a vertex (i.e., not on ∂D). Then Γ contains a "great x-cycle" as in [2, 2.6]. Then the simple combinatorial argument of [2, 2.6.2] shows that among the disk components in the complement of Γ there is a disk which either ∂-compresses Q in M_β or winds algebraically more than once around the boundary of a solid torus in $\tilde{M}$. In the latter case there is a Lens-space summand in $\tilde{M}$, providing a cobordism between Q and the required planar surface Q' (cf. [2, 2.5.2]).   q.e.d.

The second lemma requires a bit of linear algebra:

9.4 Definition.   An $n \times n$ matrix A is *decomposable* if, for some permutation matrix P, $PAP^T = \left(\begin{smallmatrix} B & 0 \\ D & C \end{smallmatrix} \right)$, and the square matrix B is called a (proper) summand of A.

Let A_{ii} denote the (i, i)-minor of A, called a *principal* minor.

9.5 Lemma.   *Let A be an $n \times n$ integral matrix such that for all i*

$$0 \neq |a_{ii}| \geq \sum_{j \neq i} |a_{ij}|.$$

Then either

(1) *some summand B of A (possibly all of A) has $|b_{ii}| = \sum_{j \neq i} |b_{ij}|$ and $|\det(B_{ii})| \geq 2$ for every i,*

(2) $|\det A| \geq 2$, *or*

(3) *some $|a_{ii}| = 1$.*

Proof of 9.5.   The proof is by induction on n. Note that either (2) or (3) is satisfied when $n = 1$.

First suppose A decomposes as above for B and C. Then B and C each satisfy the theorem by inductive hypothesis. If either B or C satisfies (1) or (3), so does A. If both satisfy (2), so does A.

Next suppose A is indecomposable. Then every minor of A satisfies the theorem by inductive hypothesis. (1) cannot be satisfied for any principal minor A_{ii}, for such a summand of A_{ii} would be a (proper) summand of A. If (3) holds for any principal minor, it holds for A. So henceforth assume (2) holds for every principal minor: $|\det A_{ii}| \geq 2$.

Suppose A is indecomposable and $|a_{ii}| = \sum_{j \neq i} |a_{ij}|$ for all i. Since $|\det(A_{ii})| \geq 2$ for every i, (1) holds for A (taking $B = A$).

Suppose A is indecomposable and $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$ for at least one i. Then the real matrix A' obtained from A by replacing a_{ii} by $a_{ii} \pm (1/|\det A_{ii}|)$

continues to satisfy $|a'_{ii}| > \sum_{j \neq i} |a_{ij}|$, so, by [10, III 2.2.4],

$$|\det A| \neq 0 \neq |\det A'| = |(\det A) \pm 1|.$$

Hence $|\det A| \geq 2$ and (2) holds for A.   q.e.d.

Lemma 9.5 has the following topological consequence:

9.6 Lemma. *Let N be a manifold obtained by attaching n 2-handles* $q_1, \cdots, q_n$ *to a solid handlebody H of genus n. Regard H as obtained by attaching $(n+1)$ 1-handles to two 0-handles. Let $c_0, \cdots, c_n$ be the cocores of these 1-handles of H. Suppose the intersections satisfy the following conditions:*

(1) $|\partial q_i \cap c_i| \geq \sum_{0 \leq j \neq i} |\partial q_i \cap c_j|$.

(2) *For each $1 \leq i \leq n$, every intersection point of ∂q_i with c_i has the same orientation.*

Then either

(a) *some q_i intersects c_i and another c_k in a single point, and is disjoint from all other c_j, or*

(b) $N = N' \# W$, *where $H_1(W)$ has nontrivial torsion.*

Proof. Let A denote the $n \times n$ integral matrix whose entry a_{ij} is the algebraic intersection number $\langle \partial q_i, c_j \rangle$, $1 \leq i, j \leq n$. Condition (2) guarantees that $|a_{ii}|$ is also the geometric intersection number of $|\partial q_i \cap c_i|$. Hence (1) implies that for all $1 \leq i \leq n$, $|a_{ii}| \geq \sum_{j \neq i} |a_{ij}|$. Now apply 9.5:

In this context, 9.5(3) implies 9.6(a). Also 9.5(2) implies that ∂N is a sphere and $|H_1(N)| = |\det A|$, so $H_1(N)$ is finite and nontrivial.

So suppose 9.5(1) holds. With no loss of generality, the curves involved in the matrix B are $c_1, \cdots, c_r$ and $\partial q_1, \cdots, \partial q_r$ for some $1 \leq r \leq n$. From the equality $|b_{ii}| = \sum_{j \neq i} |b_{ij}|$ we conclude that each q_i, $1 \leq i \leq r$, is disjoint from c_0 and each c_j, $j > r$. Then the union of the 0-handles of H, the 1-handles $c_1, \cdots, c_r$, and the 2-handles $q_1, \cdots, q_{r-1}$ is a summand W of N with $|H_1(W)| = |\det(A_{rr})| \geq 2$. Hence $H_1(W)$ is finite and nontrivial.

Proof of 9.1. Regard M as a sutured manifold with ∂M a suture. Since M is β-irreducible, M is a β-taut sutured manifold. Let $(S, \partial S) \subset (M, \partial M)$ be β-taut. By capping off annuli in ∂M, modify S so that it is a d-surface (this modification does not change the Thurston norm). The restriction of a ∂-reducing disk in M^φ to $M - \mathcal{N}$ is a ∂-incompressible, incompressible planar surface Q in $M - \mathcal{N}$ with ∂Q the union of a single essential circle on ∂M and a family of p parallel essential circles of $\partial \mathcal{N}$. Hence Q is a parametrizing surface. Construct a β-taut sutured manifold hierarchy

$$(M, \partial M) \xrightarrow{S_1 = S} (M_1, \gamma_1) \xrightarrow{S_2} \cdots \xrightarrow{S_i} (M_i, \gamma_i) \xrightarrow{S_{i+1}} \cdots$$

which respects Q. The sequence terminates with a 3-manifold M_n with ∂M_n a collection of spheres. We will be focusing attention on (M_n, γ_n, β_n) and for simplicity continue to refer to β_n as β.

Let $q_1, \cdots, q_r$ be the components of Q_n, and for each $1 \leq \lambda \leq r$ let ν_λ and μ_λ denote, as in 7.4, the number of arcs in $\partial q_\lambda \cap A(\gamma_n)$ and $\partial q_\lambda \cap \dot{\mathcal{N}}$ (edges of β) respectively. Then, since there are no vertices in β, $I(q_\lambda) = (\nu_\lambda + \mu_\lambda) - 2\chi(q_\lambda)$. Since Q_n is a parametrizing surface, no q_λ is a disk with boundary disjoint from γ_n and β, so always $I(q_\lambda) \geq 0$. By 7.5 and 7.6, $\sum I(q_\lambda) \leq I(Q) = -2\chi(Q) = 2p - 2$.

Let $\{q'_\lambda\}$ be the collection of amalgamating and cancelling disks in Q_n. Let β' be the collection of arcs obtained from β by amalgamating and cancelling the β, using the q'_λ one at a time but never doing a self-amalgamation. Since self-amalgamations are never performed, this operation never uses product annuli, so (M_n, γ_n) is unchanged. By 4.3 and 4.4 (M_n, γ_n) is still β'-taut.

9.7 Main Lemma. $\beta' = \varnothing$.

Proof of 9.1 *from* 9.7. Then (M_n, γ_n) is taut in the Thurston norm. We wish to show that every decomposition $(M_{i-1}, \gamma_{i-1}) \xrightarrow{S_i} (M_i, \gamma_i)$ is a taut decomposition also in the Thurston norm. Suppose S_i is the last one which is not. Then from 4.18(a) and 3.9(b), S_i is a β-nontrivial product annulus in which both ends are inessential in $R(\gamma_{i-1})$. S_i thus lies in a sphere component of ∂M_i containing a single suture. Since M_i is taut, the sphere bounds a 3-ball $\hat{M}$. $\hat{M}$ contains arcs of β running between R_+ and R_- since S_i is not β-trivial. By 4.18(c) none of these arcs of β is parallel to the suture. No matter how $\hat{M}$ is further decomposed later in the sequence by product disks or annuli, a ball containing arcs of β not parallel to the suture will persist. The resulting ball $\hat{M}_n$ in M_n would contradict the claim that $\beta' = \varnothing$.

Proof of 9.7. By β'-irreducibility of M_n, any component of ∂M_n which is disjoint from β' bounds a ball disjoint from β', so, unless some component of ∂M_n intersects β', β' is empty and we are done. Let then N be a component of ∂M_n which intersects β'.

Before the amalgamations and cancellations, the points $\beta \cap N$, the sutures γ_n, and the arcs of $\partial Q_n \cap (N - \gamma_n)$ can be regarded as a type of graph Γ in N. In the following discussion a *vertex* of Γ will be a point of $\beta \cap N$, and an *edge* of Γ is an arc component of $\partial Q_n \cap (N - \gamma_n)$ with each end either at a vertex or at a suture of γ_n.

After the amalgamations and cancellations, the graph Γ is converted into a new graph Γ' whose vertices are the points $\beta' \cap N$ and edges come from those components of $\partial Q_n \cap (N - \gamma_n)$ which do not lie on amalgamating or cancelling disks. In fact, Γ' is obtained from Γ in two operations:

(a) Contract those edges of Γ which lie on amalgamating disks.

(b) Let e be an edge of Γ, which lies on a cancelling disk, and $v_\pm$ its two vertices. $v_\pm$ also are the ends of a component β_0 of β which is cancelled. Any edge of Γ (other than e) incident to v_+ is connected to an edge incident to v_- via an arc of $\partial Q_n \cap \mathcal{N}(\beta_0)$. Remove e and $v_\pm$, and in a band neighborhood of e connect ends of edges contiguous to v_+ to the corresponding ends of edges contiguous to v_-.

We will denote by $(Q', \partial Q') \subset (M_{\beta'}, \partial M_{\beta'})$ the components of Q_n which are not amalgamating or cancelling disks. Since these disks have index zero, still $I(Q') \leq 2p - 2$.

9.8 Definitions. For v a vertex of Γ', let the *star of v*, denoted by $\mathrm{st}(v)$, be the set of all edges of Γ' with a single end at v (thus loops are not included). Let the *valence* $\rho(v) = |\mathrm{st}(v)|$. A vertex v is *full* if $\rho(v) \geq p$.

Let δ be a suture in γ_n. Let $\mathrm{st}_\pm(\delta)$ be the set of all edges of Γ', which lie in $R_\pm$ and have a single end in δ. Similarly $\rho_\pm(\delta) = |\mathrm{st}_\pm(\delta)|$ and δ is $\pm$*full* if $\rho_\pm(\delta) \geq p$.

A loop in Γ' is *inessential* if it bounds in N a disk containing no other vertex of Γ' and no suture.

9.9 Remark. No suture δ of γ_n can bound a disk in N disjoint from β', for by β'-tautness both components of $R_\pm$ adjacent to δ would then be disks disjoint from β', contradicting our choice of N. Hence if a loop contains no vertices in its interior (i.e., the disk it bounds), then it is inessential.

Claim 1. If v is a vertex of Γ' on which no essential loops are based, then v is full.

Proof of Claim 1. Every loop on v is inessential by hypothesis. Let D_v be a regular neighborhood in N of the union of v and the interiors of all loops based at v. Then D_v contains no vertices but v, and ∂D_v intersects $\rho(v)$ edges on Γ. Thus D_v corresponds, before the amalgamation and cancellation of disks of Q_n, to a disk satisfying the hypotheses of 9.3, but not conclusions (a) or (b). But (c) implies that v is full.

Claim 2. Let δ be a suture in N, dividing N into two disks. Let D be the disk containing, say, that component of R_+ which is adjacent to δ. If the vertices of Γ' lying in D lie entirely in either R_+ or R_-, then δ is $+$full.

Proof of Claim 2. If there are no cancelling disks, then before amalgamation, the disk D satisfies the hypotheses of 9.3, but not conclusion (a) or (b). Conclusion (c) implies that δ is $+$full.

If there is a cancelling disk q_λ whose edge e in Γ has one or both vertices $v_\pm$ in D, then these vertices do not necessarily lie in R_+ (R_-), so the above argument does not work. If e crosses δ, then alter D by removing a regular neighborhood of e. This removes $v_\pm$ from D and further ensures that $\partial D \cap \Gamma$

is unaffected by the cancellation. If $e \subset \overset{\circ}{D}$, push D into M_n using a regular neighborhood of q_λ. The resulting disk does not contain v_+. After further pushes along amalgamating disks contiguous to q_λ, eventually D can be made to satisfy the hypotheses of 9.3 and to have its boundary intersect Γ before amalgamation and cancellation just as δ intersects Γ' afterwards.

Claim 3. If v and w are both vertices in $\Gamma' \cap R_+$ (or $\Gamma' \cap R_-$) and are both full, then v and w are adjacent in Γ'.

Proof of Claim 3. Suppose v and w are not adjacent, and consider a component α of $\partial Q'$ containing s edges of Γ' incident to v and t edges of Γ' incident to w, $s + t \geq 1$. Then α must contain at least $2(s + t)$ arcs of intersection with γ_n or $\mathcal{N}(\beta')$. That is, in the terminology of 7.4, $\mu_\alpha + \nu_\alpha \geq 2(s + t)$. This implies $\mu_\alpha + \nu_\alpha \geq s + t + 2$ unless $\mu_\alpha + \nu_\alpha = 2$ and $s + t = 1$.

More generally, for any component q_λ in Q' for which ∂q_λ contains s edges of Γ' incident to v and t edges of Γ' incident to w, $s + t \geq 1$, either $I(q_\lambda) \geq s + t$ or $I(q_\lambda) = 0$, q_λ is a disk, and s (or t) $= 1$. But in the latter case q_λ would be a cancelling or (nonself)amalgamating disk for the component of β' with end at v (or w). This is impossible by construction of β'. Hence always $I(q_\lambda) \geq s + t$.

Summing over all components of Q' we get $2p - 2 \geq I(Q') \geq \rho(v) + \rho(w) \geq 2p$, a contradiction.

Claim 4. If $v \in R_\pm$, and δ is a $\pm$full suture, then there is an edge in Γ' with ends in δ and v.

Proof of Claim 4. Essentially the same as that of Claim 3.

Claim 5. All essential loops in Γ' and all sutures are parallel.

Proof of Claim 5. By 9.9 any disk in N bounded by an essential loop or a suture contains vertices of Γ', therefore vertices on which no essential loop is based, and hence full vertices (Claim 1). If not all essential loops and sutures were parallel, there would then be at least three full vertices, no pair of which are adjacent, but two of which would both lie in either R_+ or R_-, contradicting Claim 3.

Definition. Following Claim 5, there are only two disk components of $R(\gamma_n)$. By 9.9 and Claim 3, one must be in R_+ and one in R_-, so we denote them $D_\pm$.

Claim 6. An innermost essential loop in Γ' contains sutures in its interior, so every vertex in $D_\pm$ is full.

Proof of Claim 6. If not, then by Claim 5 the loop lies in the interior of D_+, say. By Claim 2, ∂D_+ is a $+$full suture. Since the loop is essential, it contains in its interior a full vertex (Claim 1). Hence δ and v contradict Claim 4.

Claim 7. All vertices inside an innermost essential loop of Γ' lie in D_+ or D_-.

Proof of Claim 7. By Claim 5, the loop λ has, say, D_+ in its interior and D_- in its exterior. If there is a vertex v inside λ, then v either lies in D_+ or is adjacent to neither vertices in D_+ nor vertices in D_-, contradicting Claim 3.

Claim 8. There are no essential loops in Γ'.

Proof of Claim 8. Choose an innermost essential loop λ in Γ and let v be its vertex. If $|\operatorname{st}(v) \cap (\text{interior } \lambda)| < p$ then, by Claim 7, (interior λ) can be modified as D was in the proof of Claim 2 so that it corresponds, before the amalgamation and cancellation of disks of Q_n, to a disk satisfying the hypotheses of 9.3, but not the conclusion. If $|\operatorname{st}(v) \cap (\text{interior } \lambda)| \geq p$, then v is full and so, by Claim 3, must lie in $D_\pm$. This contradicts Claim 6.

Claim 9. Every vertex is in D_+ or D_-.

Proof of Claim 9. Every vertex is full by Claims 8 and 2, hence every vertex is adjacent to either a vertex of D_+ or a vertex of D_- by Claim 2.

Claim 10. N contains only one suture δ.

Proof of Claim 10. This follows immediately from Claims 2, 4 and 9.

Claim 11. For every component β_i of β' there is a disk component q_i of Q' such that ∂q_i crosses alternately from D_+ to D_- via β_i and from D_- back to D_+ via $\{\delta\} \cup \{\beta' - \beta_i\}$.

Proof of Claim 11. Let v be the end of β_i in D_+, say. For α any component of $\partial Q'$, $2|\alpha \cap \operatorname{st}(v)| \leq \mu_\alpha + \nu_\alpha$ since arcs of $\alpha \cap R(\gamma_n)$ lie in R_+ and R_- alternately. Let q_λ be a component of Q'. Sum over all components α of ∂q_λ to get

$$|\partial q_\lambda \cap \operatorname{st}(v)| \leq I(q_\lambda)/2 + \chi(q_\lambda).$$

If for all λ either the inequality is strict or $\chi(q_\lambda) \leq 0$, then always $|\partial q_\lambda \cap \operatorname{st}(v)| \leq I(q_\lambda)/2$. Summing over all q_λ we would thus get $p \leq p - 1$, a contradiction.

Hence there is a q_i which is a disk and for which $|\partial q_i \cap \operatorname{st}(v)| = I(q_\lambda)/2 + 1$. That is, for $\alpha = \partial q_i$, $|\alpha \cap \operatorname{st}(v)| = (\mu_\alpha + \nu_\alpha)/2$, so every arc of $\alpha \cap R_+$ lies in $\operatorname{st}(v)$. This proves the claim.

Now apply 9.6 to the punctured solid handlebody $(\partial N) \cup (\bigcup \mathcal{N}(\beta'))$, with each c_i, $i \geq 1$, corresponding to a meridian disk of $\mathcal{N}(\beta_i)$ and c_0 corresponding to δ. Since there are no amalgamating or cancelling disks among the q_i, 9.6(a) cannot occur. But conclusion 9.6(b) is also impossible since $W \subset M_n \subset M$. Hence the contradiction completes the proof.

10.   Afterword: The Poenaru conjecture is true

Can surgery on a knot in S^3 yield $S^1 \times S^2$? The so-called "Property R conjecture" is that the answer is no for all but the unknot. The Poenaru conjecture is stronger, and is best stated from the opposite direction: Can surgery on a reducible homology $S^1 \times S^2$ ever yield S^3? The conjecture is that the answer is no unless the manifold is $S^1 \times S^2$ and the knot is a standard $S^1 \times (\text{point}) \subset S^1 \times S^2$.

The proof of the conjecture follows with remarkable ease from the combinatorics of §9.

10.1 Theorem.   *Let M be a closed, reducible 3-manifold with $H_*(M) = H_*(S^1 \times S^2)$. Suppose there is a knot β in M such that some surgery on β produces S^3. Then $M = S^1 \times S^2$, and β is isotopic in M to $S^1 \times (\text{point})$.*

Proof.   It is easy to see that β generates $H_1(M)$. Any β-reducing sphere in M is then separating, and would cut off a summand of M which persists in S^3. Hence M is β-irreducible, and so β-taut. M viewed as a sutured manifold with $\beta = \varnothing$ is not taut in the Thurston norm since M is reducible. If there were an incompressible, ∂-incompressible, planar surface $(Q, \partial Q) \subset (M_\beta, \partial M_\beta)$ with ∂Q not meridian circles, then the argument of §9 could be applied verbatim to derive a contradiction.

We do not know that such a surface Q exists. But examination of the argument of §9 shows that in fact it is not necessary that Q be incompressible or ∂-incompressible, but only that no compressing or ∂-compressing disks lie in $R(\gamma_n) \subset \partial M_n$, the boundary of the last stage of the β-taut sutured manifold hierarchy (cf. 9.3(a)). That such a planar surface exists follows from the following remarkable lemma of Gabai (cf. [5, §4]). The knot k in S^3 is that obtained by surgery on β, and P is $\partial M_n - \mathcal{N}(\beta)$.

10.2 Lemma.   *Suppose k is a knot in S^3, and*

$$(P, \partial P) \subset (S^3 - \mathcal{N}(k), \dot{\mathcal{N}}(k))$$

is a properly imbedded surface, with ∂P a nonempty collection of nonmeridian essential curves. Then either k is the unknot or there is a planar surface $(Q, \partial Q) \subset (S^3 - \mathcal{N}(k), \dot{\mathcal{N}}(k))$ with ∂Q a nonempty collection of meridians of $\mathcal{N}(k)$ such that P contains no compressing or ∂-compressing disks for Q.

Before embarking on a proof we need some preliminary machinery.

10.3 Definition.   A knot k is in *generic position* if the projection $h \colon R^3 \to R$ has only nondegenerate critical points and no two critical points are at the same height. Suppose k has $2b$ critical points, and $Q_1, \cdots, Q_{2b-1}$ are level planes for h, one between each pair of critical levels. Let $t(k) = \sum_i |Q_i \cap k|$. A knot k with height function h is in *thin position* if $t(k)$ is minimal over all generic knots isotopic to k. Clearly $t(k) = 2$ if and only if k is the unknot.

Suppose k is a generic knot in S^3, and Q is a level plane for h. A $+disk$ (resp. $-disk$) is a disk in S^3 whose interior is disjoint from k and whose boundary is the union of an arc α_+ (α_-) in k lying above (below) Q and an arc δ_+ (δ_-) lying in Q.

10.4 Lemma.   *If for a generic knot k and a level plane Q there is a $+disk$ and a disjoint-disk, then k is not in thin position.*

Proof.   The two disks may be used to isotope $\alpha_\pm$ to $\delta_\pm$. Then push δ_+ slightly below Q, and δ_- slightly above. This decreases t by at least 2.

Proof of 10.2.   Put k in thin position. Isotope P so that all its critical points occur at different heights. Let Q_+ be a level plane located just below the highest critical point of k. Then $Q_+ \cap P$ contains arc components. If every arc component of $Q_+ \cap P$ were essential in P then, after some 2-surgeries on Q_+ along 2-disks in P, Q_+ would become the required planar surface. So we assume that some arc component δ of $Q_+ \cap P$ is inessential in P and so cuts off from P a disk D_+ with ∂D_+ the union of δ and an arc α of k. If α lies below Q_+, then clearly k is trivial. Hence we assume α lies above Q_+, so D_+ is a $+$disk.

Similarly, we can assume that a level plane Q_- just above the lowest critical point for k cuts off from P a $-$disk.
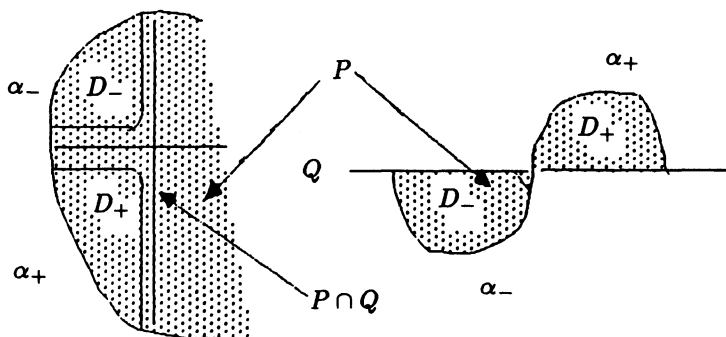

FIGURE 10.1

Now if Q is any plane between Q_+ and Q_-, then $Q \cap P$ contains arcs. If all are essential in P, then we are done as in the case of Q_+ above. If there are essential arcs cutting off both $+$disks and $-$disks, then by 10-4 k would not be in thin position. Hence somewhere there is a level plane Q between Q_+ and Q_-, so that a slightly higher plane cuts off a $+$disk D_+ but no $-$disk, and a slightly lower plane cuts off a $-$disk D_- but no $+$disk. Thus Q must

contain a saddle singularity of P, and α_+ and α_- are adjacent segments of $k - Q$. Hence D_+ and D_- can be used to isotope $\alpha_+ \cup \alpha_-$ into an arc near Q reducing the number of critical points of k (and hence $t(k)$) by at least 2. The contradiction proves the lemma.

References

[1] R. H. Bing & J. Martin, *Cubes with knotted holes*, Trans. Amer. Math. Soc. **155** (1971) 217–231.
[2] M. Culler, C. Gordon, J. Luecke & P. Shalen, *Dehn surgery on knots*, Ann. of Math. (2) **125** (1987) 237–300.
[3] D. Gabai, *Foliations and the topology of 3-manifolds*, J. Differential Geometry **18** (1983) 445–503.
[4] ____, *Foliations and the topology of 3-manifolds*. II, J. Differential Geometry **26** (1987) 461–478.
[5] ____, *Foliations and the topology of 3-manifolds*. III, J. Differential Geometry **26** (1987) 479–536.
[6] ____, *Genus is superadditive under band connected sum*, Topology **26** (1987) 209–210.
[7] ____, *Surgery on knots in solid tori*, Topology, to appear.
[8] J. Hempel, *3-manifolds*, Annals of Math. Studies No. 86, Princton University Press, Princeton, NJ, 1976.
[9] W. Jaco, *Lecture notes on 3-manifold topology*, CBMS Regional Conf. Ser. in Math., No. 43, Amer. Math. Soc., Providence, RI, 1980.
[10] M. Marcus & H. Minc, *A survey of matrix theory and matrix inequalities*, Allyn & Bacon, Boston, 1964.
[11] M. Scharlemann, *Smooth spheres in R^4 with four critical points are standard*, Invent. Math. **79** (1985) 121–141.
[12] M. Scharlemann & A. Thompson, *Unknotting number, genus and companion tori*, Math. Ann. **280** (1988) 191–205.
[13] A. Thompson, *Property P for the band-connect sum of two knots*, Topology **26** (1987) 205–208.
[14] W. Thurston, *Norm on the homology of three-manifolds*, Mem. Amer. Math. Soc. No. 339, 1986, 99–130.

UNIVERSITY OF CALIFORNIA, SANTA BARBARA