Crossing changes

Martin Scharlemann[1]

A naive but often useful way of thinking of a knot or link in R^3 is to generically project it onto a plane, keeping account of which part of the knot goes under and which part goes over at any given crossing. Think of laying the knot on a table and taking note of how it crosses itself whenever one part of it lies on top of another. It's natural to ask how the knot is changed by altering one of the crossings, reversing which arc goes over and which goes under at one of the crossing points.

This survey article is meant to explore this question. Though it's a naive question, it connects at a deep level to some of the most important ideas in modern low-dimensional topology.

1 Unknotting number, tunnel number, crossing number

Begin by making a distinction: Given a knot, one can consider what happens when one changes a crossing in a particular projection; or, more generally, one can ask what happens when one changes a crossing in some unspecified projection of the knot. The distinction is important, especially when special types of projections are being considered. For example, an *alternating knot* is a knot which has a projection (called an *alternating* projection) in which over- and under- crossings alternate as one travels a circuit around the knot. A crossing change of an alternating knot may or may not be one which can be realized in an alternating projection. A *minimal* projection of a knot is one which minimizes the number of crossings. Bleiler [1] has given an example of a knot for which two crossing changes will create the unknot, yet the two crossings cannot be realized in a minimal projection.

We will be primarily concerned with changes in a knot or link K which are crossing changes in *some* generic projection, $p : K \subset R^3 \to R^2$. There is an easy way to characterize such a change (see Figure 1). Imagine a small neighborhood of the projected crossing, looking like an X. Choose a small arc α through the center of the X and a disk D in its inverse image $p^{-1}(\alpha)$ that intersects K twice. Choose a small bi-collar neighborhood $D \times I \subset R^3$ of D. This neighborhood looks like a cylinder, which K intersects in two arcs, running from the top to the bottom of the cylinder. Now a full-twist of of the two strands of K inside the cylinder, if chosen in the right direction, changes the over-crossing to an under-crossing and vice-versa, i. e. it produces a crossing change. Note that there are, up to ambient isotopy, two choices of small arcs passing through the X and either one can make the above change. In the case that K is oriented, one of the choices will give rise to a disk

[1]Partially supported by an NSF grant

whose two intersection points with K have opposite orientation. This is the conventional choice of disk.

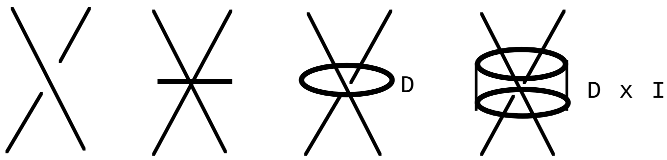

Fig. 1.

Conversely, suppose D is a disk intersecting a knot or link K transversally in two points, and $D \times I$ is a small bicollar neighborhood of D, intersecting K in two arcs. Isotope D very small, perpendicular to a projection plane, and so that both points of $D \cap K$ project to the same point. Then project the remainder, $K - D$, generically and disjoint from the projection of D. The result is a projection with a crossing in D, for which D plays the role described above. This leads to the following, projection independent, definition of a crossing change:

Definition 1.1 *Suppose K is an oriented knot or link in S^3 and D is a disk which intersects K in exactly two points, with opposite orientation. Then D is a* crossing disk *for K.*

Suppose D is a crossing disk and $D \times I \subset S^3$ is a bicollar which intersects K in two arcs with cylindrical coordinates $(\pm\pi/2, 1/2, t), t \in I$. Suppose the knot K_+ is obtained from K by replacing $K \cap (D \times I)$ with the arcs $(\pm\pi/2 + 2\pi t, 1/2, t), t \in I$. Then we say K_+ is obtained from K by a right handed crossing change (and, dually, K is obtained from K_+ by a left-handed crossing change).

The cylindrical coordinates here are understood to be orientation preserving. Note then that a change of orientation of I and hence D in $D \times I$ has no effect on the twist direction (right or left) of the crossing change. Furthermore, if both strands of $K \cap (D \times I)$ lie in the same component of K then the twist direction is independent of the orientation of K. On the other hand, if K is a link in which each of two components intersects D, and the orientation of one component is changed then D ceases to be a crossing disk. In the projection giving rise to a crossing corresponding to D, described above, the change in orientation generates a new crossing disk D', whose projection is perpendicular to the projection of D. (See Figure 2.) If K_+ is obtained from K by a right handed twist in the original orientation with crossing disk D then K_+ is obtained from K in the new orientation by a left handed twist with crossing disk D'.

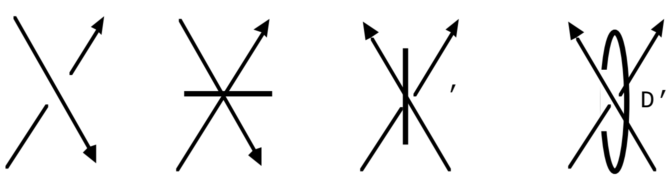

Fig. 2.

Any link K can be turned into the unlink by a sequence of crossing changes. To see this, first project K onto a plane. Order the components of K, then choose a base point in each, away from any crossing, then pick an orientation for each component. Order the points of the link (other than the base points) in the natural way: Given two points, first compare the order of the components that contain them. If they lie in the same component, compare when they are encountered when proceeding around that component from the base point with the given orientation. With this ordering now defined, lift the projected link to R^3 by requiring, at each crossing, that the overcrossing is that arc which occurs later in the order just described.

The resulting lift of the projection is the unlink, for one can also use the ordering to construct a set of unknotting disks: From each point p in the projected link draw a straight line to the base point of the component on which p lies. Lift the straight line to an arc in R^3 by having it overcross when encountering a point of the link ordered before p, and undercross when encountering a point ordered after p. The union of the lifted lines is a system of unlinking disks.

This discussion prompts two definitions and a corollary.

Definition 1.2 *The* unknotting number $u(K)$ *of a link* K *is the smallest number of crossing changes required to obtain the unlink.*

Definition 1.3 *The* crossing number $c(K)$ *of a link* K *is the smallest number of crossings, the minimum being taken over all regular projections of* K*.*

Lemma 1.4 *For any link* K*,* $u(K) \leq c(K)/2$*.*

Proof: Choose a regular projection with $c(K)$ crossings. In the process described above we created the unlink by reversing some crossings. By taking exactly the opposite ordering and opposite orientations we would have created the unlink reversing exactly the other crossings. So one choice or the other requires at most $c(K)/2$ crossing changes to create the unlink. ◇

On the other hand, it is fairly easy to show that unknotting number places no bound on crossing number. It is possible to construct infinitely many knots (hence knots of arbitrarily high crossing number) each of which has unknotting number one. To see this, first note that one way of getting unknotting number one knots is via the Whitehead doubling construction. Remove a regular neighborhood of a knot, no matter how complex, and replace it with the knot in a solid torus shown in Figure 3. With only one crossing change, this knot will bound a disk inside the solid torus. It is not hard to show that one can construct infinitely many distinct knots using this construction.

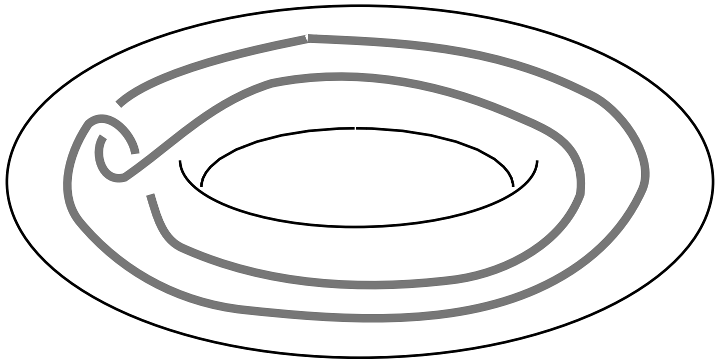

Fig. 3.

There may appear to be some ambiguity in the definition of unknotting number, since it's not clear if the relevant crossing changes are made sequentially or simultaneously. In fact, the distinction is unimportant:

Proposition 1.5 *Suppose the link K_1 is obtained from the link K by simultaneously changing n crossings, and K_2 is obtained from K_1 by changing one crossing. Then K_2 is obtained from K by simultaneously changing $n+1$ crossings.*

Proof: Let $D_1, D_2, ..., D_n$ be n crossing disks which determine the n crossing changes between K and K_1. Since the crossing changes are simultaneous, we may assume the disks are disjoint. These disks persist as crossing disks for K_1, which can be used to recreate K from K_1. For each $i, 1 \leq i \leq n$ let $\alpha_i \subset D_i$ be an arc connecting the two points $K_1 \cap D_i$. Let D_{n+1} be the crossing disk for K_1 whose corresponding crossing change makes it K_2 and let $\alpha_{n+1} \subset D_{n+1}$ be an arc connecting the two points of $K_1 \cap D_{n+1}$.

By general position we can assume that the arcs $\alpha_i, 1 \leq i \leq n+1$ are all disjoint in $S^3 - K_1$. By isotoping the D_i to lie in small regular neighborhoods of the α_i we may assume that all D_i are disjoint. Now undo the crossing changes along $D_1, D_2, ..., D_n$ to recover K and note that, since D_{n+1} is disjoint from these crossing disks, all crossings changes $D_1, \ldots D_{n+1}$ can then be performed simultaneously, changing K to K_2. $\diamond$

In view of the role played by the arcs α_i in the previous proposition, there may appear to be a connection between unknotting number and the *tunnel number* of a knot, that is the number of arcs which need to be attached to a link before the resulting 1-complex has a regular neighborhood which is an unknotted handlebody in S^3 (cf. [21]). Indeed, after adding such an arc α (or *tunnel*) to K, one can, just as in 1.5, take a disk containing α as a crossing disk, and so view α also as a tunnel for the knot K_+ obtained by doing the corresponding crossing change. Yet simple examples show there is no connection between unknotting number and tunnel number. All torus knots have tunnel number one, but their unknotting number is known (see 3.3), and can be arbitrarily high. On the other hand, Whitehead doubled knots, which have unknotting number one, can have arbitrarily high tunnel number. Indeed, given a tunnel number, there is a limit to how many non-isotopic disjoint essential tori can lie in the complement of any knot of that tunnel number [12], yet there is no limit on how many such tori can lie in the complement of Whitehead doubled knots. (This argument was pointed out to me by M. Sakuma.)

There are methods, simple and complex, for finding lower bounds for the unknotting number, and these can be used to establish unknotting number in many cases. Here is the simplest to describe, but see e. g. Section 3 and also [4], [20], [16] for more sophisticated methods.

Consider what a crossing change does to the 2-fold branched cover of a knot. The 2-fold branched cover of two unknotted arcs in a ball is the solid torus (since the 2-fold branched cover of two points in a disk is an annulus). The operation described before 1.2, which is the local picture of a crossing change, corresponds to a $1/2 \in \mathbf{Q}/\mathbf{Z}$ Dehn surgery on the solid torus, since the new meridian and the old intersect in 2 points (see Section 2 or [3]) . The 2-fold branched cover of S^3 along the unknot is S^3, so the 2-fold branched cover M of a knot K can be obtained from S^3 by $u(K)$ Dehn surgeries of slope $n/2, n \in \mathbf{Z}$. Each can change the rank of the homology of M by at most one. Thus the rank of the homology of the two-fold branched cover M is a lower bound for $u(K)$.

If K_1 and K_2 are two knots, then the crossing disks which do the unknotting for each knot can be combined to get such a set of crossing disks for $K_1 \# K_2$, so we have immediately:

Corollary 1.6 *For K_1 and K_2 two knots of unknotting number u_1 and u_2 respectively, $u(K_1 \# K_2) \leq u_1 + u_2$.*

It is possible that the inequality is in fact an equality – little is known about the question. The easiest case implied by the equality is known to be true, namely

Theorem 1.7 *If K has unknotting number one, then K is prime.*

There are several proofs (e. g. [22], [28]).

2 Connections with Dehn Surgery and Sutured Manifolds

One way of changing a 3-manifold M into a 3-manifold M' is by a process called *Dehn surgery* (see [3]). In this process a tubular neighborhood $\eta \cong S^1 \times D^2$ is removed from a knot in M and then reglued by a homeomorphism $h : \partial(S^1 \times D^2) \rightarrow \eta$. The construction is completely determined by how the meridian curve $\mu \cong (point) \times \partial D^2$ is reattached to $\partial\eta$. This in turn is determined by the algebraic intersection of $h(\mu)$ with μ (call this number q) and its intersection with the *longitude* $\lambda \cong S^1 \times point$ (call this number p). The numbers q and p, are relatively prime, so the surgery is determined by the ratio p/q. Ambiguity in the choice of λ, which depends on how $\eta \subset M$ is parameterized, means that in general p/q is only defined up to an integer, so $p/q \in \mathbf{Q}/\mathbf{Z}$. If the original manifold M is a homology sphere, then there is a natural choice of longitude: Take λ to be the unique curve in $\partial\eta$ which is null-homologous in the complement of the knot η. In this case we can take $p/q \in Q$.

A crossing change can be interpreted as a type of Dehn surgery. To see the connection, it is helpful to regard the product $D^2 \times I$ (eventually, the neighborhood of a crossing disk) to be the union of the smaller ball $\frac{1}{2}D^2 \times I = \{(r,\theta,t)|r \leq 1/2\}$ and the solid torus $\eta = \{(r,\theta,t)|1/2 \leq r \leq 1\}$. A meridian μ of η at $\theta = 0$ is the union of the two vertical arcs with $r = 1/2$ and $r = 1$ and of the two horizontal arcs $t = 0, 1$. A curve of slope $1/n$ in $\partial(\eta)$ is obtained by replacing, in μ, the vertical arc at $r = 1/2$ by the arc $(1/2, 2n\pi t, t), t \in I$. So the homeomorphism $H : \frac{1}{2}D^2 \times I \rightarrow \frac{1}{2}D^2 \times I$ given by $H(r,\theta,t) = (r, \theta + 2n\pi t, t)$, combined with $1/n$ surgery on η recreates $D^2 \times I$. The upshot is that $1/n$ surgery on $\eta \subset (D^2 \times I)$ creates a manifold still homeomorphic to $D^2 \times I$, but the homeomorphism converts a pair of vertical arcs to arcs which twist n times around each other. To summarize:

Proposition 2.1 *A crossing change for $K \subset S^3$ with crossing disk $D \subset S^3$ can be obtained by performing Dehn surgery on ∂D with slope ± 1. The crossing change is right handed (left handed) if the slope is $+1$ (resp. -1).*

An advantage of the Dehn surgery viewpoint is that machinery developed to understand the general problem of how manifolds change under Dehn surgery can be applied. For example, it was noted above (Theorem 1.7) that unknotting number one knots are prime. Stated in terms of Dehn surgery, this says the following, for D the crossing disk which unknots the knot K, and M the manifold $M - \eta(K)$: If ± 1 surgery on ∂D in M yields a ∂-reducible manifold, then no proper essential annulus in M has boundary a meridian of $\eta(K)$. The proof can be constructed combinatorially by studying how a ∂-reducing disk and the essential annulus intersect. (This was the way it was originally done in [22]). M. Eudave-Muñoz [5] has extended both the results and the combinatorial technique to prove generalized forms of the following.

Theorem 2.2 *If links K and K' are related by a crossing change, and both are composite, then the crossing change takes place within a proper summand.*

Theorem 2.3 *If links K and K' are related by a crossing change, and one is composite and the other is split, then the crossing change takes place within a proper summand.*

Theorem 2.4 *If links K and K' are related by a crossing change, and both are split, then the crossing change takes place within a splittand.*

For understanding crossing changes, however, there is a more recent and powerful approach to Dehn surgery, discovered by Gabai as part of his sutured manifold theory (cf. [7], [23]). Here is a version of the relevant theorem:

Theorem 2.5 *Suppose M is an irreducible compact orientable 3-manifold whose boundary is a union of tori, and suppose T_0 is a torus boundary component. Further suppose the technical condition that any torus in M which is rationally homology cobordant to T_0 is in fact parallel to it in M.*

Let $\alpha \in H_2(M, \partial M - T_0)$ be a homology class in M and $(S, \partial S) \subset (M, \partial M - T_0)$ be the essential surface of highest Euler characteristic representing α. Then there is at most one way of attaching a solid torus to T_0 so that either the resulting 3-manifold M is reducible or α can be represented by an essential surface of even higher Euler characteristic.

To explain how 2.5 is typically applied, the following definition is useful.

Definition 2.6 *Suppose K is an oriented link in a 3-manifold M and $S \subset M$ is an oriented surface whose boundary is K. Suppose $M - K$ is irreducible. Then S is* taut *if it is incompressible and, among all surfaces homologous to S in (M, K), S has maximal Euler characteristic.*

Now suppose that D is a crossing disk for a a crossing change of a link $K \subset S^3$ and that K does not split in the complement of D. Let $L = \partial D$ and take $M = S^3 - (\eta(K) \cup \eta(L))$ and $T_0 = \partial\eta(L)$. Because K has trivial algebraic intersection with D, K has a Seifert surface S which is disjoint from $L = \partial D$ and so lies entirely inside M. Suppose that S has been chosen to be taut in M. If a solid torus is attached to $\partial\eta(L)$ taking the new meridian to the old, i. e. just replacing $\eta(L)$, then S becomes a Seifert surface for K in S^3. According to 2.1, if instead a solid torus is attached by doing a $+1$ surgery on L then S is a Seifert surface for the knot K_+ obtained by a right handed crossing change for K via D. Theorem 2.5 says that for at most one of the two links K, K_+ there is a splitting sphere or a Seifert surface of greater Euler characteristic than S (i. e. lower genus, if K is a knot). Another way to say this:

Proposition 2.7 *At least one of K and K_+ is a non-split link for which S is a taut Seifert surface.*

Here is an easy application of this proposition. Say a crossing for K with crossing disk D is *nugatory* if ∂D bounds a disk disjoint from K. (Note then that the union of D and this disk form a sphere decomposing K into a connected sum, though of course a summand may be trivial.)

Proposition 2.8 *Suppose a crossing change in the unknot yields the unknot. Then the crossing is nugatory.*

Proof: With no loss, we suppose that the crossing change is right-handed. As above, let L be the boundary of a crossing disk for the unknot K, and let S be a taut Seifert surface for K. $S \cap D$ can be taken to be a single arc α.

If S is a disk, α cuts off from S a disk disjoint from L whose regular neighborhood has a boundary whose union is D and a disk disjoint from K. Thus the crossing is nugatory.

If S is not a disk, then S is not a taut Seifert surface for K in S^3 so, according to Proposition 2.7, S must be a taut Seifert surface for the knot K_+ obtained by $+1$ surgery on L. Hence K_+ is not the unknot. $\diamond$

The following subsections describe ways in which this sort of argument can be applied and extended.

2.1 Nullifying crossings and Conway skein trees

There is a second quite elementary move one can make at a crossing of a projection of a link K. Instead of changing the crossing, one can remove it by replacing the two crossed arcs by two uncrossed arcs (see Figure 4). The new link K_0 has either one more or one less component than K did. An equivalent way to obtain K_0, expressed in terms of a crossing disk D, is to replace the two strands $K \cap (D \times I)$ with strands in $D \times \partial I$. This process is called *nullifying* the crossing.

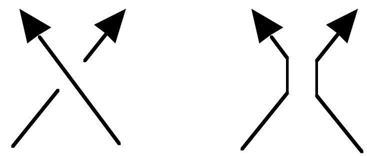

Fig. 4.

Suppose that S is a Seifert surface for K and S intersects the crossing disk D precisely in one arc. Then a Seifert surface S_0 for K_0 can be obtained from S by removing a band along the arc $S \cap D$. A fairly straightforward argument ([26, Theorem 1.4, Claim 2]) shows that S_0 is taut for K_0 if and only if S is taut for K in the manifold M_0 obtained by 0 surgery on ∂D. The central point is that Theorem 2.5 is relevant here too: S and S_0 fail to be taut for at most one of the relevant links K, K_+, K_0.

The central fact which allows for the rapid calculation of the Jones and other polynomials for an oriented link ([9]) is that they can be defined recursively by making precisely these moves $K \to K_\pm, K_0$ to crossings of the link. This is accomplished by the construction of a *Conway skein tree* for the link, a process we now describe.

Given a link K, project it so that it has minimal crossing number. As described in 1.4 , some crossing change in this projection will simplify the link, either because it can then be isotoped to have a projection with lower crossing number, or perhaps only because fewer crossings need be made in the present projection to unknot it. Nullifying that crossing also creates a link which is simpler because it has fewer crossings. So, assuming for concreteness that it is a positive crossing change which is made, of the three links K, K_+ and K_0, the latter two are simpler than K. Now continue the process on K_+ and K_0, iterating until all links have become the unlink. Because the polynomial for K can be expressed easily in terms of the polynomials for K_+ and K_0, and the polynomial for the unlink is immediate, this process gives a straightforward iterative process for calculating such knot polynomials.

What Theorem 2.5 tells us in this context is that during this process the maximal Euler characteristic of Seifert surfaces for K_+ and K_0 cannot *both* rise by much above that for K – indeed if it rises at all for K_+ then it can rise at most by one for K_0 (corresponding to a cut of the Seifert surface along the arc $D \cap S$). This puts a lower bound on how many iterations are needed for such a computation of a link polynomial.

The geometric picture of taut Seifert surfaces prompted by these considerations can also be useful. As a particular example, we have:

Corollary 2.9 *A non-trivial knot is a Whitehead doubled knot if and only if its genus and unknotting number are both* 1.

Proof: See [26, Cor. 3.2]. ◇

Similar geometric pictures arise in the more general context of rational homology spheres, and fibered links within them. For example:

Theorem 2.10 *Suppose that K is a fibered link in a rational homology 3-sphere M, that the link K' is obtained from K by a single crossing change, and that the Euler characteristic of a taut Seifert surface for K' is greater than that for K. Then some taut surface for K has a plumbed on Hopf band.*

Proof: See [11].

2.2 Deeper into sutured manifold theory: parameterizing surfaces

Here in a nutshell is a description of how sutured manifold theory works. Begin with an oriented compact connected 3-manifold M and a non-separating surface S within it. Construct a hierarchy for M, beginning with a oriented cutting surface S and keeping track of normal orientations of regions on the boundary at every stage in the hierarchy. Gabai shows that this can be done in such a way that useful information about S in the initial manifold can be recovered by careful examination of the appearance of normal orientations in the final stage of the hierarchy.

For example, here is a sketch of how Theorem 2.5 above is proven in [7]. Construct a hierarchy for M, beginning with the surface S and insisting that all later surfaces avoid T_0. Ultimately the process terminates in a 3-manifold which is a homology product, with one end T_0. The technical condition ensures that this is a real product, and the torus T_1 at the other end has some regions with normal orientation pointing out and some regions with normal orientations pointing in. The details of the construction ensure that the curves γ (called *sutures*) in T_1 which separate the two sorts of regions are essential curves. The deep part of the theorem says that the original surface S is taut even after filling in a solid torus at T_0 as long as these sutures are not meridians of the resulting solid torus in the final stage of the hierarchy. Hence at most one attachment slope destroys tautness of S.

A major advance in [8] is the discovery that other nice surfaces in M, not necessarily non-separating, can be useful in this process, because topological information can be gleaned from the remnants of the surface left at the last stage of the hierarchy. In particular, it can be arranged that such a surface (called a *parameterizing* surface) is cut in a way so as to preserve information about its Euler characteristic. This gives a way of integrating initial information about the existence of simple surfaces into the sutured manifold machine.

For example, in [25], sutured manifold theory is used to reprove and generalize Theorem 1.7. The generalization is to knots (like composite knots) which are satellite knots of non-trivial winding number. The relevant parameterizing surface is the companion torus. Suppose $K \subset S^3$ is such a satellite knot with non-trivial winding number with respect to the companion torus T, D is a crossing disk for K, and K' is the knot obtained by the crossing change. We have the following:

Theorem 2.11 *Either T is isotopic in $S^3 - K$ to a torus disjoint from ∂D or there is a minimal genus Seifert surface for K' intersecting ∂D in ≤ 2 points.*

Proof: See [25, 3.1] ◇

Corollary 2.12 *If genus*$(K') \leq$ *genus*$(K) - 2$ *then any companion torus of* K *can be isotoped to be disjoint from* ∂D.

Proof: A Seifert surface for K' intersecting ∂D in 2 points can be turned into a Seifert surface for K, of one higher genus, by tubing along an arc in ∂D. ◇

Corollary 2.13 *If* K *is a satellite knot of non-trivial winding number, then* $u(K) > 1$.

Proof: Such a satellite knot must have genus greater than one, and can't be unknotted by a crossing change inside the companion torus T, so this torus must intersect the boundary of the crossing disk. So a crossing change can't produce the unknot, because the unknot has genus zero. ◇

The last corollary implies 1.7 because a composite knot is a satellite of either of its summands, via the "follow-swallow" torus (see Figure 5).

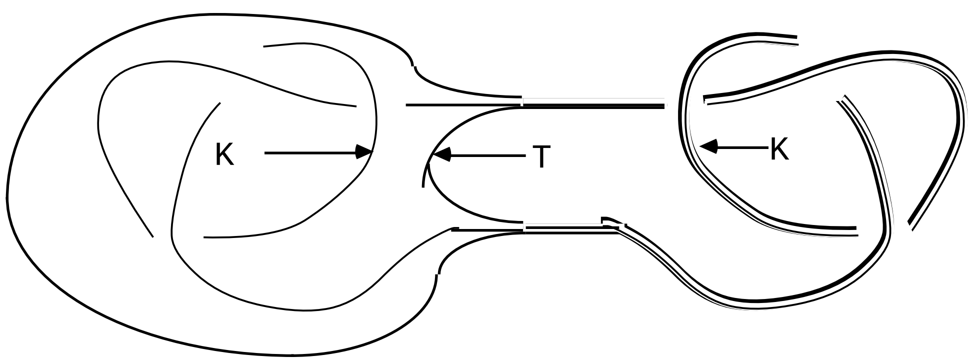

Fig. 5.

2.3 Generalized crossing changes

A *generalized* crossing change of *degree* q is one obtained by doing surgery of slope $1/q, q \in Z$ on the boundary of a crossing disk. Equivalently, two strands of the knot in a neighborhood of a crossing disk are given q full twists. An early example of such a generalization is T. Kobayashi's extension [10] of Theorem 2.11.

Recent work by Lackenby ([14], [15]) has dramatically extended to generalized crossing changes the range of the technique described in 2.2. Whereas Corollary 2.13 says that one cannot move from a certain type of satellite knot to the unknot by a simple crossing change, Lackenby shows:

Proposition 2.14 *For* K *any knot in* S^3, *there is a (computable) bound* $b(K)$ *so that any generalized crossing change that results in a knot of lower genus has degree less than* $b(K)$.

The fundamental idea is to regard boundary reducing disks (if any) for the complement of a taut Seifert surface for K as parameterizing surfaces. The argument is quite deep, for the argument requires new examination of what the parameterizing surfaces reveal in the final stage of the sutured manifold decomposition used to establish Theorem 2.5. He uses similar methods to establish some related theorems:

Theorem 2.15 *If K is a non-trivial fibered knot in S^3, then any generalized crossing change that reduces the genus of K has index ± 1.*

Theorem 2.16 *Suppose every minimal genus Seifert surface for a knot K has ∂-irreducible complement in S^3. Then no crossing change can lower its genus.*

The latter is also proven in [26], and examples of such knots are given in [25].

2.4 Crossing changes and strongly invertible knots

The theme of this section has been that results about Dehn surgery can be used to get information about crossing changes. There is another context in which the information flows the other way, based on an idea noted in the discussion following Corollary 1.4: crossing changes correspond to Dehn surgeries of slope $1/2 \in \mathbf{Q}/\mathbf{Z}$. This approach was originated by Montesinos [18].

Definition 2.17 *A knot K is* strongly invertible *if there is an orientation preserving involution of S^3 which carries K to itself, reversing its orientation.*

Waldhausen [27] has shown that such an involution is equivalent to a π-rotation of S^3 around an unknotted axis A intersecting K twice. The resulting quotient space is then also S^3, with A the branch set. The image of K becomes an arc α with both ends attached to A. A neighborhood $\eta(\alpha)$ of A is a ball intersecting A in two parallel unknotted strands, so we may view $\eta(\alpha)$ as the neighborhood of a crossing disk D for A in (the quotient) S^3. (See Figure 6.) Just as described above, a crossing change, using this crossing disk, is equivalent in the 2-fold branched cover (i. e. the original 3-sphere) to a Dehn surgery on the solid torus lift of $\eta(\alpha)$, that is the original knot K. So, for example, if $\pm 1/2$ surgery on K yields the unknot, then a crossing change of A along D also yields the unknot. It then follows from 2.8 that α is an unknotted arc, and so K is trivial.

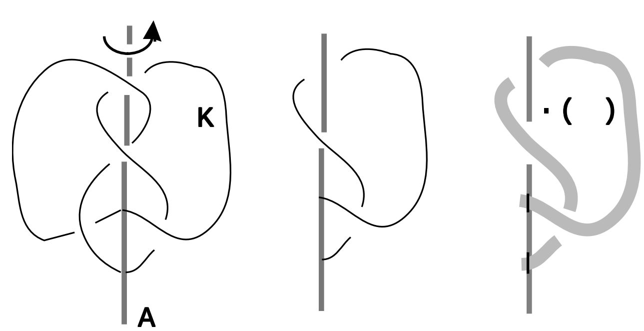

Fig. 6.

Variations of Montesinos' idea, combined with arguments about crossing changes and their generalizations, have given us very solid information about Dehn surgery on strongly invertible knots. In particular, it is known ([2] and [6] respectively) that such knots satisfy Property P (no Dehn surgery yields a homotopy 3-sphere) and the cabling conjecture (no Dehn surgery yields a reducible 3-manifold).

The discussion in this section has been limited to the changing of a single crossing, and this is no accident. Information about simultaneous Dehn surgery on a number of components is surprisingly sparse, so the approach outlined here has not proven of much use in understanding changes of more than one crossing. An exception is found in [24], but the results are modest.

3 Connections with 4-manifold topology

One surprising approach to finding lower bounds for unknotting number is to examine 4-manifolds whose boundary is the 3-sphere, and proper surfaces within the 4-manifold whose boundary in S^3 is the knot or link. The idea is fairly old [19], but a recent breakthrough in 4-manifold machinery has renewed the importance and interest in this approach.

Definition 3.1 *The* 4-ball genus $g^*(K)$ *of a knot* K *is the smallest genus of any properly imbedded surface* $(S, \partial S) \subset (B^4, S^3)$ *so that* $\partial S = K$.

Clearly the 4-ball genus is no greater than the genus of the knot, since the interior of a Seifert surface $S \subset S^3$ for K can be pushed into the interior of B^4. Less obvious is the following relation:

Proposition 3.2 *For any knot* K, $g^*(K) \leq u(K)$.

Proof: It is possible to construct a surface S in B^4, bounded by K, whose genus is the unknotting number of K. The idea is to construct a moving picture of the surface, starting with $K \subset \partial B^4$ and proceeding into the 3-ball.

Choose radii $0 < t_1 < t_2 < \ldots < t_n < 1$ in $S^3 \times (0, 1] \cong B^4 - \{0\}$. In $S^3 \times (t_n, 1]$ the surface will be just $K \times (t_n, 1]$. At level t_n, part of S will be horizontal. In fact, near the first crossing that is changed in unknotting K, define the surface $S \cap (S^3 \times \{t_n\})$ as shown in Figure 7. The surface S as so far defined is then of genus one and has boundary the union of $K \subset (S^3 \times \{1\})$ and $K_n \subset (S^3 \times \{t_n\})$, where K_n is the knot obtained from K by changing the crossing. We repeat the process in each successive $S^3 \times [t_i, t_{i-1}]$ until at level t_1 we arrive at the unknot. The unknot bounds a disk in $S^3 \times (0, t_n]$ which we attach. The result is a surface $S \subset B^4$ of genus n such that $\partial S = K \subset (S^3 \times \{1\})$. $\diamond$

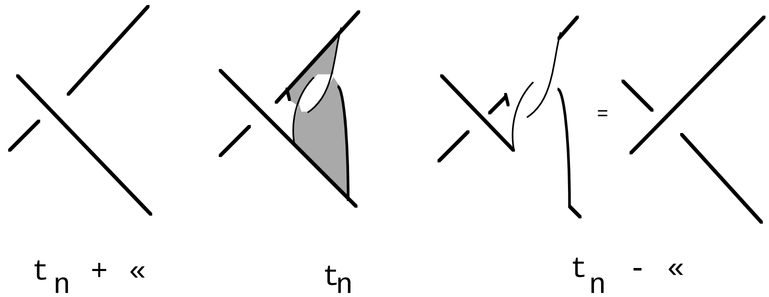

Fig. 7.

Of course this proposition is of little help without information about $g^*(K)$ and part of what Murasugi shows is that a directly computable knot invariant, namely the signature $\sigma(K)$ satisfies $\sigma(K)/2 \leq g^*(K)$. This effectively determines another lower bound for $u(K)$.

Cochran and Lickorish [4] take a somewhat different approach, inspired by groundbreaking new information about 4-manifolds that was developed in the mid-80's by Donaldson and others. Rather than using crossing changes to construct a surface in the 4-ball, they use them to construct a more complicated 4-manifold, bounded by S^3, in which the knot K bounds a disk. The fundamental construction is motivated by the discussion in Section 2, where it was shown that a change in crossing can be accomplished by ± 1 surgery on the boundary L of a crossing disk. Another way of viewing such a surgery is this: Begin with $S^3 \times I$ and attach the product $D^2 \times D^2$ to $\eta(L) \times \{1\} \subset S^3 \times \{1\}$ along $\partial D^2 \times D^2$. The attaching homeomorphism $\partial D^2 \times D^2 \rightarrow \eta(L)$ is determined by the surgery coefficient. The result is a 4-manifold which has two boundary components, each S^3, and which properly contains $K \times I$. In fact, the 4-manifold is a twice punctured complex projective plane (or conjugate projective plane) and $K \times \{1\}$ lies in the newly constructed boundary component as $K_{\pm}$. Continue the process until all crossing changes have been made, then cap off by a disk.

The end result is that K bounds a disk Δ in a connected sum W of copies of $CP(2)$ and $\overline{CP(2)}$, with a copy of the former for every right hand crossing change and a copy of the latter for every left hand crossing. The 2-fold branched cover of W along Δ is then a simply connected 4-manifold with easily calculated homology intersection pairing. Its boundary is the 2-fold branched cover of K. Donaldson theory puts limits on what 4-manifolds can occur. The possibilities excluded by Donaldson theory are sensitive to the sign of the intersection number, so the types of bounds achievable by this technique are typically lower bounds on how many crossings of a particular handedness can be used to unknot the knot.

We conclude with a recent and rather spectacular advance in this direction. Kronheimer and Mrowka [13] have exploited, at a very deep level, 4-manifold techniques coming out of gauge theory and algebraic geometry to get lower bounds on the genus of surfaces that can represent certain 2-dimensional homology classes in 4-manifolds. The results are applicable to knots which can be realized as links of complex singularities. Here is a sketch of what this means: Let V be the algebraic variety which is the solution set to a complex polynomial

$p : \mathbf{C^2} \to \mathbf{C}$. Any point in $V \subset \mathbf{C^2}$ at which p is not transverse to zero is a *singular point* of V. A neighborhood of a singular point in $\mathbf{C^2} = R^4$ is a 4-ball whose boundary S^3 intersects V in a knot. Torus knots occur in precisely this way.

The upshot is that there is a lower-bound to the 4-ball genus $g^*(K)$ for K a (p, q) torus knot. That lower bound is the genus of the surface $p^{-1}(\epsilon)$ for ϵ a generic point near 0, namely $\frac{1}{2}(p-1)(q-1)$. So by Proposition 3.2, this is a lower bound for $u(K)$. On the other hand, it's easy to construct an unknotting of the (p, q) torus knot with $\frac{1}{2}(p-1)(q-1)$ crossing changes, so this is also an upper bound. We conclude:

Theorem 3.3 *For K the p, q torus knot, $u(K) = \frac{1}{2}(p-1)(q-1)$.*

This had been a major unsolved conjecture, first made by Milnor in 1968 [17].

References

[1] S. Bleiler, A note on unknotting number, *Math. Proc. Cambridge Philos. Soc.* **96**, 469-471 (1984).

[2] S. Bleiler and M. Scharlemann, A projective plane in $\mathbf{R}^4$ with three critical points is standard. Strongly invertible knots have property P. *Topology* **27**, 519–540 (1988).

[3] S. Boyer, Dehn Surgery on Knots, (these articles).

[4] T. D. Cochran and W. B. R. Lickorish, Unknotting information from 4-manifolds. *Trans. Amer. Math. Soc.* **297**, 125-142 (1986).

[5] M. Eudave-Muñoz, Primeness and sums of tangles, *Trans. Am. Math. Soc.* **306**, 773-790 (1988).

[6] M. Eudave-Muñoz, Band sums of links which yield composite links. The cabling conjecture for strongly invertible knots. *Trans. Amer. Math. Soc.* **330** 463–501 (1992).

[7] D. Gabai, Foliations and the topology of 3-manifolds II, *Jour. Diff. Geom.* **26**, 557-614 (1987).

[8] D. Gabai, Foliations and the topology of 3-manifolds III, *Jour. Diff. Geom.* **26**, 479-536 (1987).

[9] P. Freyd, D. Yetter, J. Hoste, W. B. R. Lickorish, K. Millett, A. Ocneanu, A new polynomial invariant of knots and links. *Bull. Amer. Math. Soc.* **12**, 239-246 (1985).

[10] T. Kobayashi, Generalized unknotting operations and tangle decompositions, *Proc. Amer. Math. Soc.* **105**, 471-478 (1989).

[11] T. Kobayashi, Fibered links and unknotting operations, *Osaka J. Math.* **26**, 699-742 (1989).

[12] T. Kobayashi. Structures of full Haken manifolds. *Osaka J. Math.* **24**, 173–215 (1987).

[13] P. B. Kronheimer and T. S. Mrowka. Gauge theory for embedded surfaces. I, *Topology* **32**, 773–826 (1993).

[14] M. Lackenby, Surfaces, surgery and unknotting operations, preprint.

[15] M. Lackenby, Upper bounds in the theory of unknotting operations, preprint.

[16] W. B. R. Lickorish, The unknotting number of a classical knot. Combinatorial methods in topology and algebraic geometry (Rochester, N.Y., 1982), *Contemp. Math.*, **44**,117-121 (1985).

[17] J. Milnor, *Singular points of complex hypersurfaces* Ann. of Math. Studies, No. 61, Princeton Univ. Press, Princeton, N.J., 1968.

[18] J. M. Montesinos, Surgery on links and double branched covers of S^3, *Knots, groups, and* 3-*manifolds (Papers dedicated to the memory of R. H. Fox)*, pp. 227–259. Ann. of Math. Studies, No. 84, Princeton Univ. Press, Princeton, N.J., 1975.

[19] K. Murasugi, On a certain numerical invariant of link types, *Trans. Amer. Math. Soc.* **117**, 387-422 (1965).

[20] Y. Nakanishi, A note on unknotting number, *Math. Sem. Notes Kobe Univ.* **9** (1981), 99-108.

[21] M. Sakuma, The topology, geometry and algebra of unknotting tunnels, (these articles).

[22] M. G. Scharlemann, Unknotting number one knots are prime, *Invent. Math.* **82**, 37-55 (1985).

[23] M. G. Scharlemann, Sutured manifolds and generalized Thurston norms, *Jour. Diff. Geom.* **29**, 557-614 (1989).

[24] M. G. Scharlemann, Unlinking via simultaneous crossing changes, *Trans. Amer. Math. Soc.* **336** 855-868 (1993).

[25] M. G. Scharlemann and A. Thompson, Unknotting number, genus, and companion tori, *Math. Ann.* **280**, 191-205 (1988).

[26] M. G. Scharlemann and A. Thompson, Link genus and the Conway moves, *Comment. Math. Helvetici* **64**, 527-535 (1989).

[27] F. Waldhausen, Über Involutionen der 3-Sphäre, *Topology* **8**, 81-91 (1969).

[28] X. Zhang, Unknotting number one knots are prime: a new proof. *Proc. Amer. Math. Soc.* **113**, 611–612 (1991).

Department of Mathematics
University of California
Santa Barbara, CA 93106
USA