JOURNAL OF COMBINATORIAL THEORY, Series B 70, 134–156 (1997)
ARTICLE NO. TB971754

Making Curves Minimally Crossing by Reidemeister Moves

Maurits de Graaf*

Department of Mathematics, University of Amsterdam, Plantage Muidergracht 24, 1018 TV Amsterdam, The Netherlands

and

Alexander Schrijver

CWI, Kruislaan 413, 1098 SJ Amsterdam, The Netherlands; and Department of Mathematics, University of Amsterdam, Plantage Muidergracht 24, 1018 TV Amsterdam, The Netherlands

Received February 22, 1996

DEDICATED TO PROFESSOR W. T. TUTTE ON THE OCCASION OF HIS EIGHTIETH BIRTHDAY

Let $C_1, ..., C_k$ be a system of closed curves on a triangulizable surface S. The system is called *minimally crossing* if each curve C_i has a minimal number of self-intersections among all curves C'_i freely homotopic to C_i and if each pair C_i, C_j has a minimal number of intersections among all curve pairs C'_i, C'_j freely homotopic to C_i, C_j respectively ($i, j = 1, ..., k, i \neq j$). The system is called *regular* if each point traversed at least twice by these curves is traversed exactly twice, and forms a crossing.

We show that we can make any regular system minimally crossing by applying Reidemeister moves in such a way that at each move the number of crossings does not increase. It implies a finite algorithm to make a given system of curves minimally crossing by Reidemeister moves. © 1997 Academic Press

1. INTRODUCTION AND FORMULATION OF THE THEOREM

Let S be a surface. A *closed curve* on S is a continuous function $C: S^1 \to S$ (where S^1 is the unit circle in the complex plane). Two closed curves C and C' are *freely homotopic*, in notation: $C \sim C'$, if there exists a continuous function $\Phi: S^1 \times [0, 1] \to S$ such that $\Phi(z, 0) = C(z)$ and $\Phi(z, 1) = C'(z)$ for all $z \in S^1$.

* Present address: Dr. Neher Laboratorium, P.O. Box 421, 2260 AK Leidschendam, The Netherlands.

0095-8956/97 $25.00

For any closed curve C on S, the number of self-intersections (counting multiplicities) of C is denoted by $\mathrm{cr}(C)$. That is,

$$\mathrm{cr}(C) = \tfrac{1}{2}\,|\{(w, z) \in S^1 \times S^1 \mid C(w) = C(z),\ w \neq z\}|. \tag{1}$$

Moreover, $\mathrm{mincr}(C)$ denotes the minimum number of $\mathrm{cr}(C')$ where C' ranges over all closed curves freely homotopic to C. That is,

$$\mathrm{mincr}(C) = \min\{\mathrm{cr}(C') \mid C' \sim C\}. \tag{2}$$

For any pair of closed curves C, D on S, the number of intersections of C and D (counting multiplicities) is denoted by $\mathrm{cr}(C, D)$. That is,

$$\mathrm{cr}(C, D) = |\{(w, z) \in S^1 \times S^1 \mid C(w) = D(z)\}|. \tag{3}$$

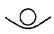
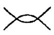
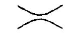
Moreover, $\mathrm{mincr}(C, D)$ denotes the minimum of $\mathrm{cr}(C', D')$ where C' and D' range over all closed curves freely homotopic to C and D, respectively. That is,

$$\mathrm{mincr}(C, D) = \min\{\mathrm{cr}(C', D') \mid C' \sim C,\ D' \sim D\}. \tag{4}$$

Let $C_1, ..., C_k$ be a system of closed curves on a surface S. We call $C_1, ..., C_k$ *minimally crossing* if

$$\begin{aligned} &\text{(i)} \quad \mathrm{cr}(C_i) = \mathrm{mincr}(C_i) \quad \text{for each } i = 1, ..., k; \\ &\text{(ii)} \quad \mathrm{cr}(C_i, C_j) = \mathrm{mincr}(C_i, C_j) \quad \text{for all } i, j = 1, ..., k \text{ with } i \neq j. \end{aligned} \tag{5}$$

We call $C_1, ..., C_k$ a *regular* system of curves if $C_1, ..., C_k$ have only a finite number of intersections (including self-intersections), each being a crossing of only two curve parts. That is, no point on S is traversed more than twice by $C_1, ..., C_k$ and each point of S traversed twice has a disk-neighborhood on which the curve parts are topologically two crossing straight lines. To such systems of curves we can apply the following four operations called *Reidemeister moves*:

0. replacing  by  (*type* 0);
I. replacing  by  (*type I*);
II. replacing  by  (*type II*);
III. replacing  by  (*type III*). (6)

The pictures here represent the intersection of the union of $C_1, ..., C_k$ with an open disk on S. So no other curve parts than the ones shown intersect such a disk.

Here and below we take all statements *topologically*. For instance, an open disk is any topological space homeomorphic to an open disk. Pictures are taken up to topological transformations. As an 'implicit' Reidemeister move we take shifting all curves simultaneously over the surface, by an isotopy $\Phi: S \to S$ (thus not changing the combinatorial structure of the system of curves).

The main result of this paper is:

THEOREM 1. *Let S be a triangulizable surface. Then any regular system of closed curves on S can be transformed to a minimally crossing system by a series of Reidemeister moves.*

This theorem will be used in a subsequent paper [4] to prove a theorem on decompositions of graphs and a homotopic circulation theorem.

It is important to note that the main content of Theorem 1 is that we do not need to apply the operations (6) in the reverse direction—otherwise the result would follow quite straigthforwardly with the classical techniques of simplicial approximation (as applied by Reidemeister [6]). Clearly, the reverse of a type III Reidemeister move is again a type III Reidemeister move; similarly for type 0. However, this does not hold for types I and II.

The theorem has as a consequence:

COROLLARY 1a. *There is a finite algorithm to transform a given regular system of closed curves on a surface, to a minimally crossing system of closed curves by Reidemeister moves.*

We can assume here that the system is given in a combinatorial way. That is, the curves are given by the graph formed by their embedding, and the surface by the faces made by that graph. For our purposes it only matters if a face is topologically a disk or not. This all can be described in a finite way.

The reason that our theorem gives a finite algorithm is that we can apply the Reidemeister moves without increasing the total number of crossings. So in a brute force way, we could enumerate all possible configurations that arise from the given system by any series of Reidemeister moves type III (there are only finitely many of them, since there are only finitely many graphs with a given number of vertices, and since for each graph there are only finitely many ways of attaching faces). Next we see if we can apply to any of these configurations a Reidemeister move of type 0, I or II. If so, we can continue with a simpler system; that is, with fewer crossings or with fewer closed curves (by removing a homotopically trivial closed curve). If not, our theorem says that the system is minimally crossing.

We can arrive at this conclusion by our theorem. If we would need to apply Reidemeister moves of type I or II also in the reverse direction, we would not obtain a finite procedure.

2. SOME FURTHER TERMINOLOGY AND NOTATION

Let S be a surface. A *curve* on S is a continuous function $C: I \to S$ where I is a connected subset of S^1. It is *closed* if $I = S^1$, *nonclosed* if $I \neq S^1$, and *simple* if it is one-to-one.

Let C be a curve on a surface S and let $A \subseteq S$. We call L a *chord on A of C* if $L = C|I$ for some connected component I of $C^{-1}[A]$. We call L a *chord on A of $C_1, ..., C_k$* if L is a chord on A of one of $C_1, ..., C_k$.

A closed curve C is called *nullhomotopic* if it is freely homotopic to a constant function. It is *orientation-preserving* if passing once around C does not change the meaning of 'left' and 'right'. Otherwise, C is *orientation-reversing.*

We will, if no confusion arise, identify a closed curve $C: S^1 \to S$ with its image $C[S^1]$. Moreover, we identify a closed curve C with any closed curve $C' = C \circ \phi$ if $\phi: S^1 \to S^1$ is a homeomorphism isotopic to the identity.

3. REDUCTION TO COMPACT SURFACES WITH A FINITE NUMBER OF HOLES

A *compact surface with a finite number of holes* is a surface arising from a compact surface by deleting a finite number of points. (So a compact surface with a finite number of holes need not be compact.)

We show that to prove Theorem 1 we may restrict ourselves to compact surfaces with a finite number of holes.

Let S be a surface and let $S' \subseteq S$. For closed curves C and D on S' denote the function mincr by mincr$'$ if it is with respect to S'. Clearly,

$$\text{mincr}'(C) \geqslant \text{mincr}(C) \qquad \text{and} \qquad \text{mincr}'(C, D) \geqslant \text{mincr}(C, D). \qquad (7)$$

PROPOSITION 1. *Let S be a triangulizable surface and $C_1, ..., C_k$ be a regular system of closed curves on S. Then S contains a compact surface S' with a finite number of holes such that S' contains $C_1, ..., C_k$ and such that* $\text{mincr}'(C_i) = \text{mincr}(C_i)$ *for each i and* $\text{mincr}'(C_i, C_j) = \text{mincr}(C_i, C_j)$ *for all $i, j (i \neq j)$.*

Proof. Consider a polygonal decomposition of S in which each vertex has degree 3. For all i, j with $1 \leqslant i < j \leqslant k$, let $\Delta_{i,j}$ be the set of all polygons traversed when shifting C_i and C_j to some closed curves C_i' and C_j' (respectively) satisfying $\text{cr}(C_i', C_j') = \text{mincr}(C_i, C_j)$. Similarly, for each $i = 1, ..., k$ let Δ_i be the set of all polygons intersected when shifting C_i to some closed curve C_i' satisfying $\text{cr}(C_i') = \text{mincr}(C_i)$. Note that each $\Delta_{i,j}$ and each Δ_i is finite. Let S' be the union of all $\Delta_{i,j}$ and Δ_i. Then S' is a compact bordered surface with a finite number of boundary components, and the proposition follows. ∎

Proposition 1 shows that in the sequel we may assume:

$$S \text{ is a compact surface with a finite number of holes.} \tag{8}$$

4. THE DISK

One important ingredient in our proof is a theorem of Ringel, and an extension of it, on shifting curves in a disk.

Let U be a closed disk. Consider systems of nonclosed curves $C_1, ..., C_k$ on U satisfying:

(9)
- (i) each C_i is simple and has end points on bd(U);
- (ii) if $i \neq j$, C_i and C_j have at most one intersection, being a crossing;
- (iii) each point of U traversed by at least two curves belongs to the interior of U and is a crossing of two curve parts, and is not traversed by any other curves.

Ringel [8] showed:

THEOREM 2 (Ringel's theorem). *Let U be a closed disk. Let $C_1, ..., C_k$ and $C'_1, ..., C'_k$ be systems of curves on U each satisfying* (9). *For each i, let C_i and C'_i have the same pair of end points. Then $C_1, ..., C_k$ can be moved to $C'_1, ..., C'_k$ by a series of Reidemeister moves of type III, each applied to the interior of U.*

Next consider systems of curves $C_1, ..., C_k$ on U satisfying:

(10)
- (i) each C_i is either closed and disjoint from bd(U) or is nonclosed and has two distinct end points on bd(U);
- (ii) each point p of U traversed by at least two curve parts belongs to the interior of U and is a crossing of the two curve parts while no other curve parts traverse p.

Call a system satisfying (10) *minimally crossing* if each curve is simple, and any two curves have at most one intersection. We derive from Ringel's theorem:

THEOREM 3. *Any system of curves on U satisfying* (10) *can be transformed to a minimally crossing system by a series of Reidemeister moves.*

Proof. Let $C_1, \ldots, C_k$ be a system of curves on U satisfying (10). We may assume that no series of Reidemeister moves decreases the number of (self-)crossings. We show that the system is minimally crossing, by induction on the number t of crossings (including self-crossings) of $C_1, \ldots, C_k$.

We first show that each of the C_i is simple. Suppose, say, C_1 is not simple. Then C_1 contains a simple 'loop' L—that is, there is an interval $I=[x, y]$ such that $C_1|I$ is one-to-one, except that $C_1(x)=C_1(y)$. Let U' be a disk in U containing L and its interior, except for a 'small' neighbourhood of $C_1(x)$. So U' contains less than t crossings, and hence, by the induction hypothesis, the chords of the C_i on U' are minimally crossing. Hence the chord $L \cap U'$ does not intersect any of the other chords. Therefore, all other chords are actually pairwise disjoint closed curves contained in the interior of L. With Reidemeister moves of type 0 they can be moved to the exterior of L. After that we can apply a Reidemeister move of type I to remove L, contradicting the minimality of the number of crossings.

We next show that any two of the C_i cross each other at most once. Suppose that, say, C_1 and C_2 cross each other more than once. Then there exist intervals $I_1=[x_1, y_1]$ and $I_2=[x_2, y_2]$ such that $C_1|I_1$ and $C_2|I_2$ are disjoint, except that $C_1(x_1)=C_2(x_2)$ and $C_1(y_1)=C_2(y_2)$. Let L be the digon formed by $C_1|I_1$ and $C_2|I_2$. Let U' be a disk on U containing L and its interior, except for a small neighbourhood of $C_1(x_1)$. So U' contains less than t crossings, and hence, by the induction hypothesis, the chords of the C_i on U' are minimally crossing. By Ringel's theorem (Theorem 2) we can apply Reidemeister moves so that the two chords formed by $C_1[I_1]$ and $C_2[I_2]$ have a crossing 'close' to $C_1(x_1)$, in such a way that the digon formed in the new situation does not contain any other curve parts. Hence it can be removed with a Reidemeister move of type II. This reduces the number of crossings, and hence contradicts the minimality of the number of crossing. ∎

5. PROPERTIES OF MINIMAL COUNTEREXAMPLES

With the help of the results of Section 4 we derive in this section some properties of 'minimal counterexamples' to Theorem 1. Let S be a triangulizable surface and let $C_1, \ldots, C_k$ be a regular system of closed curves on S. We call $C_1, \ldots, C_k$ a *minimal counterexample* if the following holds:

(i) the system $C_1, \ldots, C_k$ is not minimally crossing;

(ii) no series of Reidemeister moves decreases $\mathrm{cr}(C_i)$ for any $i \in \{1, \ldots, k\}$ or $\mathrm{cr}(C_i, C_j)$ for any $i, j \in \{1, \ldots, k\} (i \neq j)$; (11)

(iii) k is minimal (under (i) and (ii)).

It is obvious that any system obtained from a minimal counterexample by applying a series of Reidemeister moves of type III, is a minimal counterexample again (since such operations are reversible). Furthermore, we cannot apply a Reidemeister move of type 0, I, or II to any minimal counterexample.

PROPOSITION 2. *Let $C_1, ..., C_k$ be a minimal counterexample on S and let A be an open disk on S. Then the chords of $C_1, ..., C_k$ on A are minimally crossing, and none is a closed curve.*

Proof. Directly from Theorem 3 and (11)(ii). ∎

In particular:

PROPOSITION 3. *Let $C_1, ..., C_k$ be a minimal counterexample on S. Then there is no open disk containing any of the curves C_i for $i = 1, ..., k$.*

Proof. Directly from Proposition 2. ∎

Next we show:

PROPOSITION 4. *Let $C_1, ..., C_k$ be a minimal counterexample on S. Then $k \leqslant 2$ and if $k = 2$ then $\mathrm{cr}(C_i) = \mathrm{mincr}(C_i) (i = 1, 2)$.*

Proof. We first show for any regular system $C_1, ..., C_k$ of closed curves on S:

> if $C_1, ..., C_{k-1}$ can be transformed to closed curves $C'_1, ..., C'_{k-1}$ by a series of Reidemeister moves, then there exists a closed curve C'_k such that $C_1, ..., C_k$ can be transformed to $C'_1, ..., C'_k$ by a series of Reidemeister moves. (12)

To see this we may assume that $C'_1, ..., C'_{k-1}$ arise from $C_1, ..., C_{k-1}$ by one Reidemeister move. We assume this is a Reidemeister move of type III—the other types follow similarly.

Let P, Q, R be the three chords of $C_1, ..., C_{k-1}$ on an open disk $A \subset S$ to which the Reidemeister move is applied. Note that $C_1, ..., C_{k-1}$ do not have other chords on A, but C_k can have chords on A.

By Proposition 2 we know that the chords of $C_1, ..., C_k$ on A are minimally crossing, and by Theorem 2 we may assume that the triangle enclosed by P, Q and R does not intersect any of the chords of C_k on A. After this we can apply the Reidemeister move to P, Q, R and we obtain (12).

It implies:

> Let $C_1, ..., C_k$ be a minimal counterexample on S. Then for each $r \in \{1, ..., k\}$ the system $C_1, ..., C_{r-1}, C_{r+1}, ..., C_k$ is minimally crossing. (13)

For suppose that, say, $C_1, ..., C_{k-1}$ is not minimally crossing. By (11)(iii) there is a series of Reidemeister moves bringing $C_1, ..., C_{k-1}$ to $C'_1, ..., C'_{k-1}$ so that for some $i \in \{1, ..., k-1\}$, $\text{cr}(C'_i) < \text{cr}(C_i)$ or for some $i, j \in \{1, ..., k-1\}$, $\text{cr}(C'_i, C'_j) < \text{cr}(C_i, C_j)(i \neq j)$. By (12) there is a curve C'_k and a series of Reidemeister moves bringing $C_1, ..., C_{k-1}, C_k$ to $C'_1, ..., C'_{k-1}, C'_k$. This contradicts (11)(ii).

So we have (13), which gives the proposition. ∎

6. SPHERE, OPEN DISK, AND PROJECTIVE PLANE

We now have directly:

PROPOSITION 5. *Theorem* 1 *is true in case S is a sphere or an open disk.*

Proof. Directly from Proposition 3. ∎

PROPOSITION 6. *Theorem* 1 *is true in case S is the projective plane.*

Proof. Let $C_1, ..., C_k$ be a minimal counterexample on S. Let D be a simple closed nonnullhomotopic curve on S so that $D, C_1, ..., C_k$ is a regular system of curves and so that $\Sigma := \sum_{i=1}^{k} \text{cr}(D, C_i)$ is minimal. Let $A := S \backslash D$. So A is an open disk. We may assume that A is the unit open disk in $\mathbb{C}$ and that S is obtained from the closed unit disk K in $\mathbb{C}$ by identifying opposite points on the boundary of K. By Proposition 2 each chord of A is a simple path connecting two points on $\text{bd}(K)$ and each two chords intersect each other at most once. Moreover, by Ringel's theorem and Proposition 2 we may assume that all chords are straight line segments with endpoints on $\text{bd}(K)$.

Now if there is a chord l that does *not* connect two opposite points on $\text{bd}(K)$, then there is a straight line segment connecting two opposite points on $\text{bd}(K)$ and not intersecting l. This would give a nonnullhomotopic closed curve on S having fewer intersections with $C_1, ..., C_k$ than D—a contradiction.

So each chord connects two opposite points, and hence each chord corresponds to one nonnullhomotopic closed curve $C_i (i \in \{1, ..., k\})$. Hence the system $C_1, ..., C_k$ is minimally crossing, contradicting (11)(i). ∎

7. MINIMIZING THE CROSSING NUMBER OF PERMUTATIONS

Theorem 1 for the special cases of the annulus and the Möbius strip turns out to boil down to statements on permutations. These statements are basic also for our proof for more general surfaces.

Let π be a permutation of $\{1, ..., n\}$. A *crossing pair of* π is a pair $\{i, j\}$ with $(i-j)(\pi(i)-\pi(j))<0$. The *crossing number* $\mathrm{cr}(\pi)$ of π is the number of crossing pairs of π. (In Bourbaki [2] and Geck and Pfeiffer [3] the number $\mathrm{cr}(\pi)$ is called the *length* of the permutation π.)

Let $\mathrm{mincr}(\pi)$ denote the minimum of $\mathrm{cr}(\pi')$ taken over all conjugates π' of π. So $\mathrm{mincr}(\pi)$ only depends on the sizes of the orbits of π.

A *transposition* is any permutation $(k, k+1)$ for some $k \in \{1, ..., n-1\}$. Since each permutation σ is a product of transpositions $\tau_1, ..., \tau_m$, it is trivial to say that for each permutation π there exist transpositions $\tau_1, ..., \tau_m$ such that

$$\mathrm{cr}(\tau_m \cdots \tau_1 \pi \tau_1 \cdots \tau_m) = \mathrm{mincr}(\pi). \tag{14}$$

What however can be proved more strongly is:

THEOREM 4. *For each permutation* π *of* $\{1, ..., n\}$ *there exist transpositions* $\tau_1, ..., \tau_m$ *such that* (14) *holds and such that moreover*:

$$\mathrm{cr}(\tau_j \cdots \tau_1 \pi \tau_1 \cdots \tau_j) \leqslant \mathrm{cr}(\tau_{j-1} \cdots \tau_1 \pi \tau_1 \cdots \tau_{j-1}) \tag{15}$$

for each $j = 1, ..., m$.

That is, when going step by step to $\mathrm{mincr}(\pi)$ we never have to increase the number of crossings. In Section 9 we shall see that a similar statement also holds if we *maximize* the number of crossings.

We should remark here that Theorem 4 has been proved by Geck and Pfeiffer [3] for all Weyl groups (including the symmetric group). Its counterpart for maximizing, Theorem 5, is, according to our information, not known for Weyl groups. For completeness we give a proof of Theorem 4, for which we use the following proposition (which is also easy to derive with the theory developed in Bourbaki [2] (Chapter 4 Section 5) for the more general Coxeter groups).

PROPOSITION 7. *Let* π *be a permutation, let* τ *be the transposition* $(k, k+1)$, *and let* $\pi' := \tau\pi\tau$. *Then*:

$$\begin{gathered}\mathrm{cr}(\pi') \leqslant \mathrm{cr}(\pi) \quad \text{if and only if} \quad \pi' = \pi \\ \text{or} \quad \pi(k) > \pi(k+1) \quad \text{or} \quad \pi^{-1}(k) > \pi^{-1}(k+1).\end{gathered} \tag{16}$$

Proof. To see sufficiency, suppose $\mathrm{cr}(\pi') > \mathrm{cr}(\pi)$. Then clearly $\pi' \neq \pi$. Moreover, by parity, $\mathrm{cr}(\pi') \geqslant \mathrm{cr}(\pi) + 2$. Hence π' has a crossing pair $\{i, j\} \neq \{k, k+1\}$ such that $\{\tau(i), \tau(j)\}$ is not a crossing pair of π. We may assume that $i < j$, and hence $\tau(i) < \tau(j)$. So $\tau\pi\tau(i) > \tau\pi\tau(j)$ and $\pi\tau(i) < \pi\tau(j)$. Hence $\pi\tau(i) = k$ and $\pi\tau(j) = k+1$. So $\pi^{-1}(k) = \tau(i) < \tau(j) = \pi^{-1}(k+1)$.

One similarly shows that $\pi(k) < \pi(k+1)$ (since $\mathrm{cr}(\pi'^{-1}) = \mathrm{cr}(\pi') > \mathrm{cr}(\pi) = \mathrm{cr}(\pi^{-1})$).

To see necessity, suppose $\pi' \neq \pi$, $\pi(k) < \pi(k+1)$ and $\pi^{-1}(k) < \pi^{-1}(k+1)$. Then for each crossing pair $\{i, j\}$ of π, the pair $\{\tau(i), \tau(j)\}$ is a crossing pair of π'. Indeed, we may assume $i < j$; hence $\pi(i) > \pi(j)$. Since $\pi(k) < \pi(k+1)$ we know $\{i, j\} \neq \{k, k+1\}$. So $\tau(i) < \tau(j)$. If $\{\tau(i), \tau(j)\}$ is not a crossing pair of π', we have $\pi'(\tau(i)) < \pi'(\tau(j))$; that is, $\tau(\pi(i)) < \tau(\pi(j))$. So $\{\pi(i), \pi(j)\} = \{k, k+1\}$, and hence $\pi(i) = k+1$ and $\pi(j) = k$. So $\pi^{-1}(k+1) = i < j = \pi^{-1}(k)$, a contradiction.

Hence $\mathrm{cr}(\pi') \geqslant \mathrm{cr}(\pi)$. To show strict inequality, we show that $\{k, k+1\}$ is a crossing pair of π'. (Note that it is not a crossing pair of π.)

Suppose $\{k, k+1\}$ is not a crossing pair of π'. So $\pi'(k) < \pi'(k+1)$. That is, $\tau(\pi(k+1)) < \tau(\pi(k))$. As $\pi(k+1) > \pi(k)$, we know $\{\pi(k), \pi(k+1)\} = \{k, k+1\}$. But this would imply that $\pi' = \pi$, contradicting our assumption. ∎

We put $\pi' \preccurlyeq \pi$ if there exist permutations $\pi_0, ..., \pi_t$ such that $\pi_0 = \pi'$, $\pi_t = \pi$, and for each $i = 1, ..., t$, $\mathrm{cr}(\pi_{i-1}) \leqslant \mathrm{cr}(\pi_i)$ and there exists a transposition τ such that $\pi_i = \tau\pi_{i-1}\tau$. (Possibly $t = 0$.) So $\preccurlyeq$ is reflexive and transitive.

Proof of Theorem 4. We show that for each permutation π on $\{1, ..., n\}$ there exists a permutation $\pi' \preccurlyeq \pi$ such that $\pi' = (1, 2, ..., j_1)(j_1+1, ..., j_2) \cdots (j_{s-1}+1, ..., j_s)$ for some $j_1 < j_2 < \cdots < j_s = n$. This proves the theorem, since the number of crossing pairs of π' only depends on the sizes of the orbits.

Represent permutation π' as

$$\pi' = (k_1, ..., k_{j_1})(k_{j_1+1}, ..., k_{j_2}) \cdots (k_{j_{s-1}+1}, ..., k_{j_s}). \tag{17}$$

Choose π' and this representation so that $\pi' \preccurlyeq \pi$ and so that the vector $(k_1, ..., k_n)$ is lexicographically minimal. We may assume that $\pi' = \pi$.

We show that $k_j = j$ for $j = 1, ..., n$. Suppose this is not the case, and choose r satisfying $k_r \neq r$, with r as small as possible. So $k_j = j$ for all $j < r$, and $k_r > r$.

By the lexicographic minimality of representation (17), k_r is not the first of any of the orbits in this representation (otherwise we could choose r as the start of a new orbit). So $\pi^{-1}(k_r) = k_{r-1} = r-1$.

Define $\pi' := \tau\pi\tau$, where $\tau := (k_r - 1, k_r)$. Then $\pi^{-1}(k_r - 1) \in \{r, ..., n\}$, implying $\pi^{-1}(k_r - 1) \geqslant r > r-1 = \pi^{-1}(k_r)$. So by Proposition 7, $\mathrm{cr}(\pi') \leqslant \mathrm{cr}(\pi)$. This contradicts the lexicographic minimality of representation (17). ∎

Note that from the proof of Theorem 4 we also obtain:

$$\text{mincr}(\pi) = n - s \tag{18}$$

for any permutation π of $\{1, ..., n\}$ with s orbits.

8. THE ANNULUS

Theorem 4 implies Theorem 1 in case S is the annulus (the sphere with two points deleted).

PROPOSITION 8. *Theorem* 1 *is true in case* S *is the annulus.*

Proof. Let $C_1, ..., C_k$ be a minimal counterexample on S. We may assume that S is obtained from the square $K = [0, 1] \times (0, 1)$ by identifying $(0, x)$ and $(1, x)$ for each $x \in (0, 1)$. Let $A_i := i \times (0, 1)$, let A denote the curve on S arising after identifying A_0 and A_1, and let $U = (0, 1) \times (0, 1)$.

We may assume that we have chosen the representation so that $A, C_1, ..., C_k$ is regular and so that the number of crossings of A with $C_1, ..., C_k$ is as small as possible.

Then each chord of $C_1, ..., C_k$ on U connects A_0 and A_1 (when taking their closures in K). (Otherwise we could (with the help of Ringel's theorem) decrease the number of crossings of A with $C_1, ..., C_k$.) So we can orient each chord so that it runs on K from A_0 to A_1.

Let $x_1, ..., x_n$ be the crossing points of $C_1, ..., C_k$ with A, in order. So there is a permutation π of $\{1, ..., n\}$ such that the chord starting at x_i at A_0 ends at $x_{\pi(i)}$ at $A_1 (i = 1, ..., n)$. Note that $\text{cr}(\pi)$ is equal to the total number of crossings of $C_1, ..., C_k$.

Now we have the following:

> if τ is a transposition such that $\text{cr}(\tau\pi\tau) \leqslant \text{cr}(\pi)$, then we can apply Reidemeister moves to $C_1, ..., C_k$ such that the associated permutation becomes equal to $\tau\pi\tau$. (19)

Indeed, let $\tau = (m, m+1)$. By Proposition 7, we may assume that $\pi(m) > \pi(m+1)$. Hence the chords starting at x_m and at x_{m+1} cross. Therefore, by Ringel's theorem we can apply Reidemeister moves so that their crossing is the first in both of these chords. Then by a topological transformation we can shift the crossing beyond A. This makes that π is transformed to $\tau\pi\tau$. This shows (19).

Now if $k = 1$, π has one orbit. Let C'_1 be a closed curve on S freely homotopic to C_1 satisfying $\text{cr}(C'_1) = \text{mincr}(C_1)$. Then C'_1 gives similarly a permutation π'. As C'_1 is freely homotopic to C_1, π' is conjugate to π. As $\mathrm{cr}(C'_1) < \mathrm{cr}(C_i)$, we know that $\mathrm{cr}(\pi') < \mathrm{cr}(\pi)$.

So by Theorem 4 there exist transpositions $\tau_1, ..., \tau_m$ such that $\mathrm{cr}(\tau_j \cdots \tau_1 \pi \tau_1 \cdots \tau_j) \leqslant \mathrm{cr}(\tau_{j-1} \cdots \tau_1 \pi \tau_1 \cdots \tau_{j-1})$ for each $j = 1, ..., m$, with strict inequality for $j = m$. But this would give by (19) a series of Reidemeister moves so as to decrease the number of self-crossings of C_1—contradicting the fact that C_1 is a minimal counterexample.

If $k = 2$, then π has two orbits. Then we can consider similarly closed curves C'_1, C'_2 freely homotopic to C_1, C_2 respectively, satisfying $\mathrm{cr}(C'_1, C'_2) = \mathrm{mincr}(C_1, C_2)$. ∎

9. MAXIMIZING THE CROSSING NUMBER OF PERMUTATIONS

If we want to apply a similar technique to the Möbius strip, we have to consider *maximizing* the number of crossings of permutations. We define $\mathrm{maxcr}(\pi)$ to be the maximum of $\mathrm{cr}(\pi')$ taken over all permutations π' conjugate to π. Again trivially for any permutation π there exist transpositions $\tau_1, ..., \tau_m$ such that

$$\mathrm{cr}(\tau_m \cdots \tau_1 \pi \tau_1 \cdots \tau_m) = \mathrm{maxcr}(\pi). \tag{20}$$

Again this can be sharpened to:

THEOREM 5. *For each permutation π there exist transpositions $\tau_1, ..., \tau_m$ such that* (20) *holds and such that moreover*:

$$\mathrm{cr}(\tau_j \cdots \tau_1 \pi \tau_1 \cdots \tau_j) \geqslant \mathrm{cr}(\tau_{j-1} \cdots \tau_1 \pi \tau_1 \cdots \tau_{j-1}) \tag{21}$$

for each $j = 1, ..., m$.

We prove Theorem 5 directly only in case π has one or two orbits. The general case follows from Proposition 12 below.

We first show a few propositions. We define $\preccurlyeq$ as in the proof of Theorem 4.

Denote the sequence $1, n, 2, n-1, 3, n-2, ...$ by $a_1, a_2, a_3, a_4, a_5, ...$ So

$$\begin{aligned} a_r &= s && \text{if} \quad r = 2s - 1, \\ a_r &= n - s + 1 && \text{if} \quad r = 2s. \end{aligned} \tag{22}$$

Hence $a_n = \lfloor n/2 \rfloor + 1$.

Define permutation π_n of $\{1, ..., n\}$ by

$$\pi_n := (a_1, a_2, ..., a_n). \tag{23}$$

Moreover, if $h, k \geqslant 1$ with $h+k=n$, define permutation $\pi_{h,k}$ of $\{1, ..., n\}$ by

$$\pi_{h,k} := (a_1, ..., a_h)(a_{h+1}, ..., a_n). \tag{24}$$

So $\pi_{h,k}$ has orbits of sizes h and k.

PROPOSITION 9. *Let π be a permutation of $\{1, ..., n\}$.*

(i) *If π has one orbit then $\pi \preccurlyeq \pi_n$.*

(ii) *If π has two orbits, of size h and k, where 1 belongs to the orbit of size h, then $\pi \preccurlyeq \pi_{h,k}$.*

Proof. Write $\pi = (k_1, ..., k_n)$ (in case (i)) or $\pi = (k_1, ..., k_h)(k_{h+1}, ..., k_n)$ (in case (ii)), in such a way that $(k_1, -k_2, k_3, -k_4, ...)$ is lexicographically minimal.

We show that $k_j = a_j$ for $j = 1, ..., n$, thus proving the proposition. Suppose $k_r \neq a_r$ for some r, which we choose as small as possible. So $k_j = a_j$ for $j = 1, ..., r-1$ and $k_r \in \{a_{r+1}, ..., a_n\}$. Clearly, $k_1 = 1$, so $r \neq 1$. Moreover, in case (ii), $r \neq h+1$ (since otherwise $(k_1, ..., k_h) = (a_1, ..., a_h)$, so $a_r \in \{k_{h+1}, ..., k_n\}$, and we can put a_r in the position of k_{h+1}; this would contradict the lexicographic minimality assumption).

This implies

$$\pi^{-1}(k_r) = k_{r-1} = a_{r-1}. \tag{25}$$

Case 1. r is odd, say $r = 2s+1$. So $a_r = s+1$, $\{k_1, ..., k_{r-1}\} = \{a_1, ..., a_{2s}\} = \{1, ..., s\} \cup \{n-s+1, ..., n\}$ and

$$\{k_r, ..., k_n\} = \{s+1, ..., n-s\}. \tag{26}$$

By the choice of r we have that $k_r \neq a_r = s+1$, and so by (26), $s+2 \leqslant k_r \leqslant n-s$, and hence $k_r - 1 \in \{k_{r+1}, ..., k_n\}$. Therefore,

$$\pi^{-1}(k_r - 1) \in \{k_r, ..., k_n\} = \{s+1, ..., n-s\}. \tag{27}$$

Define $\tau := (k_r - 1, k_r)$ and $\pi' := \tau\pi\tau$. Then by (25) and since $k_r - 1, k_r \in \{s+1, ..., n-s\}$,

$$\begin{aligned} \pi'^{-1}(k_r - 1) &= \tau\pi^{-1}\tau(k_r - 1) = \tau\pi^{-1}(k_r) \\ &= \tau(k_{r-1}) = k_{r-1} = a_{r-1} = n-s+1. \end{aligned} \tag{28}$$

Moreover,

$$\pi'^{-1}(k_r) = \tau\pi^{-1}\tau(k_r) = \tau\pi^{-1}(k_r - 1) \in \{s+1, ..., n-s\}, \tag{29}$$

as $\pi^{-1}(k_r - 1) \in \{s+1, ..., n-s\}$ (by (27)) and as $k_r, k_r - 1 \in \{s+1, ..., n-s\}$.

By (28) and (29), $\pi'^{-1}(k_r) < \pi'^{-1}(k_r - 1)$, implying by Proposition 7 that $\mathrm{cr}(\pi') \geqslant \mathrm{cr}(\pi)$; so $\pi \preccurlyeq \pi'$.

This contradicts the lexicographic minimality assumption, since $\pi' = (k_1, ..., k_{r-1}, k_r - 1, ...)$.

Case 2. r is even, say $r = 2s$. So $a_r = n - s + 1$, $\{k_1, ..., k_{r-1}\} = \{1, ..., s\} \cup \{n - s + 2, ..., n\}$ and

$$\{k_r, ..., k_n\} = \{s + 1, ..., n - s + 1\}. \tag{30}$$

By the choice of r we have that $k_r \neq a_r = n - s + 1$, and so by (30), $s + 1 \leqslant k_r \leqslant n - s$, and hence $k_r + 1 \in \{k_{r+1}, ..., k_n\}$. Therefore,

$$\pi^{-1}(k_r + 1) \in \{k_r, ..., k_n\} = \{s + 1, ..., n - s + 1\}. \tag{31}$$

Define $\tau := (k_r, k_r + 1)$ and $\pi' := \tau\pi\tau$. Then by (25) and as $k_r, k_r + 1 \in \{s + 1, ..., n - s + 1\}$,

$$\pi'^{-1}(k_r + 1) = \tau\pi^{-1}\tau(k_r + 1) = \tau\pi^{-1}(k_r) = \tau(k_{r-1}) = k_{r-1} = a_{r-1} = s. \tag{32}$$

Moreover,

$$\pi'^{-1}(k_r) = \tau\pi^{-1}\tau(k_r) = \tau\pi^{-1}(k_r + 1) \in \{s + 1, ..., n - s + 1\}, \tag{33}$$

as $\pi^{-1}(k_r + 1) \in \{s + 1, ..., n - s + 1\}$ (by (31)) and as $k_r, k_r + 1 \in \{s + 1, ..., n - s + 1\}$.

By (32) and (33) $\pi'^{-1}(k_r) > \pi'^{-1}(k_r + 1)$, implying by Proposition 7 that $\mathrm{cr}(\pi') \geqslant \mathrm{cr}(\pi)$; so $\pi \preccurlyeq \pi'$. This again contradicts the lexicographic minimality assumption, since $\pi' = (k_1, ..., k_{r-1}, k_r + 1, ...)$. ∎

At this point we have shown Theorem 5 for permutations π with one orbit. It follows that for any permutation π of $\{1, ..., n\}$ with only one orbit one has

$$\mathrm{maxcr}(\pi) = \mathrm{cr}(\pi_n) = \binom{n}{2} - \left\lfloor \frac{n-1}{2} \right\rfloor. \tag{34}$$

Next:

PROPOSITION 10. *If h is even then* $\mathrm{cr}(\pi_{k,h}) \leqslant \mathrm{cr}(\pi_{h,k})$.

Proof. Observe that if $i, j \in \{k + 1, ..., n\}$ and $\{a_i, a_j\}$ is a crossing pair of $\pi_{k,h}$, then $\{a_{i-k}, a_{j-k}\}$ is a crossing pair of $\pi_{h,k}$. Similarly, if $i, j \in \{1, ..., k\}$ and $\{a_i, a_j\}$ is a crossing pair of $\pi_{k,h}$, then $\{a_{i+h}, a_{j+h}\}$ is a crossing pair of $\pi_{h,k}$.

Finally, each pair $\{a_i, a_j\}$ with $1 \leqslant i \leqslant h < j \leqslant n$, is a crossing pair of $\pi_{h,k}$. So we obtain the required inequality. ∎

Proposition 10 implies the theorem for permutations with two orbits of even size each. Indeed, by Proposition 9 we have that for each permutation π with two orbits, of even sizes h and k, one has $\pi \preccurlyeq \pi_{h,k}$ or $\pi \preccurlyeq \pi_{k,h}$. As by Proposition 10 one has $\mathrm{cr}(\pi_{h,k}) = \mathrm{cr}(\pi_{k,h})$, both $\pi_{h,k}$ and $\pi_{k,h}$ attain $\mathrm{maxcr}(\pi)$.

We are left to consider permutations with two orbits, at least one of them being odd. Then we have:

PROPOSITION 11. *Let h be odd and let k be such that k is even or $k \geqslant h$. Then $\pi_{h,k} \preccurlyeq \pi_{k,h}$.*

Proof. We may assume that $k \geqslant 2$ (otherwise $k=h=1$, and the claim is trivial).

By Proposition 9 it suffices to show that there exists a permutation π such that $\pi_{h,k} \preccurlyeq \pi$ and such that the orbit of π containing 1 has size k. To this end, it suffices to show that there exists a permutation π such that $\pi_{h,k} \preccurlyeq \pi$ and such that the orbit of π containing n has size k. This follows from the fact that if n belongs to the orbit of size k, then we may assume that $\pi(n)=1$, and hence 1 belongs to the orbit of size k.

Let $u := \lceil n/2 \rceil$. Consider permutations π such that $\pi_{h,k} \preccurlyeq \pi$ and such that

$$\pi = (1, k_2, ..., k_h)(k_{h+1}, ..., k_n) \tag{35}$$

where

$$\begin{aligned} &\text{(i)} \quad k_i + k_{i+1} = n+2 \quad \text{for each even} \quad i<n; \\ &\text{(ii)} \quad k_i < k_{i+2} \quad \text{for each odd} \quad i \leqslant n-2 \quad \text{with} \quad i \neq h; \\ &\text{(iii)} \quad k_i \leqslant u \quad \text{for each odd} \quad i \leqslant n. \end{aligned} \tag{36}$$

Such permutations π exist since (24) is of this form. Choose π such that $k_3 + k_5 + \cdots + k_h$ is as large as possible.

Note that condition (36)(iii) implies that

$$\{k_i \mid i \text{ odd}\} = \{1, 2, ..., u\}. \tag{37}$$

We first show:

$$\text{Let } k_j = k_i + 1 \text{ with } i, j \text{ odd and } 3 \leqslant i \leqslant h < j \leqslant n. \text{ Then } i<h \text{ and } j<n. \text{ Moreover, if } j \leqslant n-2, \text{ then } k_{j+2} > k_{i+2}. \tag{38}$$

Indeed, suppose to the contrary that $i=h$, or $j=n$, or $j \leqslant n-2$ and $k_{j+2} < k_{i+2}$. Then $\pi(k_i) < \pi(k_j)$. For if $i=h$ then $\pi(k_i) = 1 < \pi(k_j)$. If $i \leqslant h-2$ and $j=n$ then $k_{i+2} \geqslant k_i + 1 = k_j \geqslant k_{h+2}$, and hence $\pi(k_i) = k_{i+1} = n+2-k_{i+2} \leqslant n+2-k_{h+2} = k_{h+1} = \pi(k_j)$. If $i \leqslant h-2$ and $j \leqslant n-2$ and $k_{j+2} < k_{i+2}$, then $\pi(k_i) = k_{i+1} = n+2-k_{i+2} < n+2-k_{j+2} = k_{j+1} = \pi(k_j)$. So $\pi(k_i) < \pi(k_j)$.

Now let $\tau := (k_i, k_j)$ and $\pi' := \tau\pi\tau$. As $\pi(k_i) < \pi(k_j)$, we have $\pi'(k_i) > \pi'(k_j)$, and hence Proposition 7 gives $\mathrm{cr}(\pi') \geqslant \mathrm{cr}(\pi)$. So $\pi \preccurlyeq \pi'$.

Let $\tau' := (k_{i-1}, k_{j-1})$ and $\pi'' := \tau'\pi'\tau'$. Since $k_{i-1} = k_{j-1} + 1$ and $\pi'(k_{i-1}) = \tau\pi(k_{i-1}) = \tau(k_i) = k_j > k_i = \tau(k_j) = \tau\pi(k_{j-1}) = \pi'(k_{j-1})$, we know $\pi''(k_{i-1}) < \pi''(k_{j-1})$, and hence, again by Proposition 7, $\mathrm{cr}(\pi'') \geqslant \mathrm{cr}(\pi')$; so $\pi' \preccurlyeq \pi''$. Hence $\pi \preccurlyeq \pi''$.

However, the representation of π'' is obtained from that of π by interchanging k_i and k_j and by interchanging k_{i-1} and k_{j-1}. This contradicts the maximality of $k_3 + k_5 + \cdots + k_h$. Thus we have (38).

From this we derive that $k_3 \geqslant 3$, which finishes the proof, as it implies that $k_{h+2} = 2$ and hence $k_{h+1} = n$.

First we have $k_h = u$. For suppose $k_h < u$. Then by (37) there exists an odd $j \in \{h+1, ..., n\}$ such that $k_j = k_h + 1$, contradicting (38).

Next if k is even, then $k_{i+2} = k_i + 1$ for each odd i in $\{3 \leqslant i \leqslant h-2\}$. Otherwise, choose the largest odd i in $\{3, ..., h-2\}$ for which $k_{i+2} \geqslant k_i + 2$. Then there exists an odd $j \in \{h+2, ..., n\}$ such that $k_j = k_i + 1$. Then by (38), $j \leqslant n-1$, and hence (as n is odd), $j \leqslant n-2$. So by (38), $k_{j+2} > k_{i+2}$, contradicting the maximality of i (since $k_{i+2} < k_{j+2} < u = k_h$). Hence $k_3 = u - (h-3)/2 \geqslant 3$ (since $2u = n + 1 = h + k + 1 \geqslant h + 3$ as $k \geqslant 2$).

If k is odd, then n is even and $k \geqslant h$. Then $k_{i+2} \leqslant k_{i+2}$ for each odd i in $\{3 \leqslant i \leqslant h-2\}$. For suppose $k_{i+2} \geqslant k_i + 3$. Then there exists an odd $j \in \{h+2, ..., n-3\}$ such that $k_j = k_i + 1$ and $k_{j+2} = k_i + 2$. Then (38) implies $k_i + 2 = k_{j+2} > k_{i+2}$, a contradiction. Therefore, $k_3 \geqslant u - (h-3) \geqslant 3$ (since $2u = n = h + k \geqslant 2h$ as $k \geqslant h$). ∎

This finishes the proof of Theorem 5 for permutations with two orbits. Indeed, let π be a permutation with two orbits, of size h and k respectively, where h is odd and k is even or $k \geqslant h$. Then by Propositions 9 and 11, $\pi \preccurlyeq \pi_{k,h}$. So $\pi_{k,h}$ should attain a maximum number of crossings.

In fact, we obtain $\mathrm{maxcr}(\pi) = \mathrm{cr}(\pi_{h,k})$ for any permutation with two orbits of size h and k, where h is odd, and k is even or $k \geqslant h$. Concluding, for any permutation with two orbits, of sizes h and k:

$$\begin{aligned}
\mathrm{maxcr}(\pi) &= \binom{n}{2} - \left\lfloor \frac{h-1}{2} \right\rfloor - \left\lfloor \frac{k-1}{2} \right\rfloor - \min\{h, k\} && \text{if } h \text{ and } k \text{ are odd,} \\
\mathrm{maxcr}(\pi) &= \binom{n}{2} - \left\lfloor \frac{h-1}{2} \right\rfloor - \left\lfloor \frac{k-1}{2} \right\rfloor && \text{otherwise.}
\end{aligned} \tag{39}$$

10. THE MÖBIUS STRIP

Theorem 5 implies Theorem 1 in case S is the Möbius strip (the projective space with one point deleted) in the same way as Theorem 4 implies Theorem 1 in case S is the annulus as we saw in Section 8.

PROPOSITION 12. *Theorem 1 is true in case S is the Mőbius strip.*

Proof. Similar to the proof of Proposition 8. ∎

We should note here that a reverse derivation from Theorem 1 for the Mőbius strip implies Theorem 5 for permutations with any number of orbits.

11. GEODESICS ON HYPERBOLIC AND EUCLIDEAN SURFACES

All surfaces for which Theorem 1 remains to be proved are hyperbolic or Euclidean. It means that these surfaces can be equipped with a geometric structure, which gives 'geodesics' on the surface. Basic ingredient in our proof then is the fact that each nonnullhomotopic closed curve on such a surface can be brought arbitrarily close to a geodesic by a series of Reidemeister moves.

In order to give a more precise formulation and a proof of this statement we need some definitions and basic facts about surfaces and their universal covering surfaces, the background of which can be found in Baer [1], Koebe [5], Reinhart [7], and Stillwell [9].

Let U be the Euclidean or hyperbolic plane. There exists a metric dist on U such that for any three points x, y, z on U lie, in this order, on a line if and only if $\operatorname{dist}(x, z) = \operatorname{dist}(x, y) + \operatorname{dist}(y, z)$. An *isometry* on U is a homeomorphism $\phi: U \to U$ so that $\operatorname{dist}(\phi(x), \phi(y)) = \operatorname{dist}(x, y)$ for all $x, y \in U$. Thus, an isometry maps lines to lines.

Let S be any compact surface with a finite number of points deleted, with Euler characteristic $\chi(S) \leqslant 0$. If $\chi(S) = 0$, S is called *Euclidean* and if $\chi(S) < 0$, S is called *hyperbolic*. The Euclidean plane (if S is Euclidean) or the hyperbolic plane (if if S is hyperbolic) can be considered as a *universal covering surface* of S. That is, there exists a 'projection' function $\psi: U \to S$ with the following properties:

(i) for each $u \in U$ there is an open disk N containing u so that $\psi | N: N \to S$ is one-to-one;

(ii) if $u, u' \in U$ and $\psi(u) = \psi(u')$ then there exists an isometry $\phi: U \to U$ so that $\phi(u) = u'$ and $\psi \circ \phi = \psi$; (40)

(iii) for each closed curve $C: S^1 \to S$ and each $u \in \psi^{-1}[C(1)]$ there exists a unique continuous function $D: \mathbb{R} \to U$ such that $C'(0) = u$ and such that $\psi \circ D(x) = C(e^{2\pi i x})$ for all $x \in \mathbb{R}$. (D is a *lifting of C to U*.)

A closed curve J on S is called *geodesic* if any lifting of J to U is a line and if J has only a finite number of selfintersections. This last condition means that there is no closed curve K such that $J=K^n$ for some $n>1$.

Each nonnullhomotopic closed curve on S is freely homotopic to J^n for some geodesic J and some $n \geqslant 1$. If S is hyperbolic, then J and n are unique.

The projection function ψ transmits the distance function dist on U to a distance function dist_S on S given by:

$$\text{dist}_S(x, y) := \min\{\text{dist}(x', y') \mid x', y' \in U, \psi(x') = x, \psi(y') = y\} \quad (41)$$

for $x, y \in S$. Moreover, we can speak of a 'piecewise linear' curve C on S, of the length length(C) of such a curve, and of convex subsets of S (these are the subsets containing with any pair of points x, y also the shortest line segment connecting x and y). We may assume that each nonnullhomotopic piecewise linear function has length larger than 2.

We introduce a measure for the distance of a closed curve from a geodesic. Let $C: S^1 \to S$ be a piecewise linear closed curve on S, and let $D: \mathbb{R} \to U$ be a lifting of C to U. If C is nonnullhomotopic, the *deviation* dev(C) of C is equal to

$$\inf\{\varepsilon \mid D[\mathbb{R}] \subseteq B(L, \varepsilon) \text{ for some line } L\} \quad (42)$$

where $B(L, \varepsilon) := \{x \in U \mid \text{dist}(L, x) < \varepsilon\}$. If C is nullhomotopic, its *deviation* dev(C) is

$$\inf\{\varepsilon \mid D[\mathbb{R}] \subseteq B(u, \varepsilon) \text{ for some point } u\}. \quad (43)$$

PROPOSITION 13. *Let $C_1, ..., C_k$ be closed curves on S and let $\varepsilon > 0$. Then there exists a series of Reidemeister moves bringing $C_1, ..., C_k$ to $C'_1, ..., C'_k$ such that* $\text{dev}(C'_i) < \varepsilon$ *for each $i = 1, ..., k$.*

Proof. We introduce a second measure for the 'geodesicity' of a curve. Let $C: S^1 \to S$ be a closed curve. Let $C': \mathbb{R} \to U$ be any lifting of C to U. For any $t \in \mathbb{R}$, let I be the largest interval on $\mathbb{R}$ such that $t \in I$ and $C'[I] \subseteq B(C(t), 1)$. If I is bounded, let r and s be the end points of I. Define

$$\text{tort}_t(C') := \text{length}(C'[I]) - \text{dist}(C'(r), C'(s)). \quad (44)$$

If $I = \mathbb{R}$ (so C is nullhomotopic and C' is contained in a disk of radius 1), then $\text{tort}_t(C') := \text{length}(C')$. The 'tortuosity' of C is

$$\text{tort}(C) := \sup\{\text{tort}_t(C') \mid t \in \mathbb{R}\}. \quad (45)$$

Obviously, this number is independent of the choice of lifting C' of C.

The following relation between dev and tort is easy to see, by continuity:

For each L and each $\varepsilon > 0$ there exists a $\delta > 0$ such that each piecewise linear closed curve C on S with length$(C) \leqslant L$ and tort$(C) \leqslant \delta$ has dev$(C) < \varepsilon$. (46)

Now we prove Proposition 13. Let L be the maximum length of the C_i. Take δ as in (46). We consider the following operation applied to a point $u \in S$. Let $B(u, 1)$ be the ball with radius 1 around u. Replace each chord of $C_1, ..., C_k$ by the shortest curve on $B(u, 1)$ connecting the end points of that chord. If C_i is contained in $B(u, 1)$ we replace it by a closed curve of length close to 0.

This operation can be performed by Reidemeister moves (by Theorem 3). We perform this operation to any u, as long as the replacement reduces the length of at least one C_i by more than δ. So we can apply it only a finite number of times, and hence finally tort$(C_i) \leqslant \delta$ for each i. Therefore, by (46), dev$(C_i) < \varepsilon$ for each i. ∎

12. THE HYPERBOLIC SURFACES

Hyperbolic surfaces have the property that each nonnullhomotopic closed curve is freely homotopic to a unique geodesic—more precisely, to the power of a geodesic with a unique image. This is used to prove:

PROPOSITION 14. *Theorem 1 is true in case S is a hyperbolic surface.*

Proof. Let $C_1, ..., C_k$ be a minimal counterexample. By Proposition 4 we know that $k \leqslant 2$ and that if $k = 2$ then cr$(C_i) =$ mincr(C_i) for $i = 1, 2$. Moreover, from Propositions 2 and 13 we know that each C_i is nonnullhomotopic. Let J_i be a geodesic with $C_i \sim J_i^{n_i}$ for some $n_i \geqslant 1$. Let G_i be the image of J_i. So G_i is a graph embedded on S. As the J_i are geodesic, we know that if $G_i \neq G_{i'}$ then $G_i \cap G_{i'}$ is finite.

Let G be the graph $G_1 \cup \cdots \cup G_k$. Let V and E denote the vertex set and edge set of G. By introducing some extra vertices of degree 2, we may assume that G does not have loops or multiple edges. Moreover, we may assume that V is also the vertex set of each G_i. For each $v \in V$ and each $i = 1, ..., k$, let $d_{v,i}$ be half of the valency of v in G_i.

Now we consider a neighbourhood of G—in fact, we consider a polygonal decomposition of it. To this end we choose for each vertex v a convex polygon P_v containing v in its interior, and for each edge e a convex 4-gon P_e such that any edge $e = uv$ is contained in the interior of $P_u \cup P_e \cup P_v$. We can assume that the P_v are mutually disjoint and that the P_e are mutually disjoint, while P_v and P_e intersect if and only if v is incident with e. In that case, P_v and P_e intersect in a side both of P_e and of P_v. Moreover, each side of any P_v is equal to the intersection of P_v with P_e for some edge e incident with v. So, if e and e' are 'opposite' edges incident with vertex v, then P_e and $P_{e'}$ intersect P_v in opposite sides of P_v. We can also assume that if v and v' are the vertices incident with edge e, then P_v and $P_{v'}$ intersect P_e in opposite sides of P_e.

Choose $\varepsilon > 0$ such that for each edge $e = uv$, $B(e, \varepsilon)$ is contained in $P_u \cup P_e \cup P_v$. By Proposition 13 we may assume that we have applied Reidemeister moves to $C_1, ..., C_k$ so that $\text{dev}(C_i) < \varepsilon$ for each i. Hence the C_i are contained in the interior of the union of the P_v and P_e. We may assume moreover that no crossing of the C_i is on any side of any P_v, and that we have applied Reidemeister moves so as to minimize the number of intersections of the C_i with the sides of the P_v. By Proposition 2 the chords of the C_i on any P_v and on any P_e are minimally crossing.

This implies the following. Let J_i form the circuit $(v_0, e_1, v_1, ..., e_t, v_t)$ in G, with $v_0 = v_t$. Then C_i traverses $P_{v_0}, P_{e_1}, P_{v_1}, ..., P_{e_t}, P_{v_t}$, in this order, repeatedly—that is, n_i times. After entering a polygon at some side, it leaves the polygon at the opposite side. We may assume that any two chords of the C_i on any P_v cross each other only if they connect two different pairs of opposite sides.

First, suppose that $k = 1$. Choose an edge e_0 of G, with ends v_0 and v_1 say. Then we may assume that P_e does not contain any self-crossing of C_1, except if $e = e_0$. (This can be seen as follows. If e and e' are opposite edges of G incident with vertex v of G, then $P_e \cup P_v \cup P_{e'}$ forms a disk. So by Ringel's theorem (Theorem 2) we can 'move' crossings from P_e to $P_{e'}$.)

Let $R := P_{e_0} \cap P_{v_0}$. Let $n := n_1$. Let $p_1, ..., p_n$ be the crossing points of C_1 with R, in this order. Let $K_1, ..., K_n$ be the chords of $S \backslash R$, taking indices in such a way that each K_i, at the end traversing P_{e_0}, touches p_i. Then there is a permutation π of $\{1, ..., n\}$ such that $P_{\pi(i)}$ is the other end point of K_i.

If J_1 is orientation-preserving, the total number of self-crossings of C_1 is equal to

$$\text{cr}(\pi) + n^2 \sum_{v \in V} \binom{d_{v,1}}{2}. \tag{47}$$

Now if $\pi' \preccurlyeq \pi$ for some permutation π' then there exist Reidemeister moves changing C_1 so as to change π to π'. Since C_1 is a minimal counterexample, $\text{cr}(\pi)$ is as small as possible. Hence by Theorem 4, π is minimally crossing among all conjugates of π.

Now if C'_1 is a minimally self-crossing closed curve freely homotopic to C_1, and we would move C'_1 similarly close to G, we would obtain a permutation π' conjugate to π, and hence the number of self-crossings of C'_1

is not less than (47). Therefore, C_1 attains a minimum number of self-crossings.

If J_1 is orientation-reversing, the total number of self-crossings of C_1 is equal to

$$\binom{n}{2} - \text{cr}(\pi) + n^2 \sum_{v \in V} \binom{d_{v,1}}{2}. \tag{48}$$

Then we can proceed similarly to the orientation-preserving case, using Theorem 5.

Next, suppose that $k=2$ and that $G_1 \neq G_2$. Then

$$\text{cr}(C_1, C_2) = \sum_{v \in V} n_1 d_{v,1} n_2 d_{v,2}, \tag{49}$$

which number is also equal to $\text{mincr}(C_1, C_2)$ by Baer's theorem [1]. This contradicts the fact that C_1, C_2 is a minimal counterexample.

Finally, suppose that $k=2$ and $G_1 = G_2$. Then we may assume that $J_1 = J_2$. We can now proceed as in the case $k=1$. We obtain a permutation π of $\{1, ..., n\}$ with orbits of sizes n_1 and n_2 (with $n := n_1 + n_2$).

If J_1 is orientation-preserving, the total number of crossings (including self-crossings) of C_1 and C_2 is equal to (47). Like in the case $k=1$, it follows that C_1, C_2 is minimally crossing. (Note that if $\text{cr}(C'_1, C'_2) = \text{mincr}(C_1, C_2)$ for some $C'_1 \sim C_1$ and $C'_2 \sim C_2$, we can apply Reidemeister moves so as to obtain moreover that $\text{cr}(C'_1) = \text{mincr}(C_1)$ and $\text{cr}(C'_2) = \text{mincr}(C_2)$, since we have finished the case $k=1$ (using (12)).)

If J_1 is orientation-reversing, the total number of crossings (including self-crossings) of C_1 and C_2 is equal to (48). Then we can proceed similarly to the orientation-preserving case above. ∎

13. THE TORUS AND THE KLEIN BOTTLE

The only two surfaces for which we have not proved yet Theorem 1 are two Euclidean surfaces: the torus and the Klein bottle. The difference with the hyperbolic case is that on these surfaces there is not a unique geodesic freely homotopic to a given closed curve if it is orientation-preserving. However, in that case any two such geodesics can be moved in two essentially different ways to each other. This enables us to remove a point of the surface and to obtain a reduction to the hyperbolic case.

PROPOSITION 15. *Theorem 1 is true in case S is the torus or the Klein bottle.*

Proof. Let $C_1, ..., C_k$ form a minimal counterexample for the torus or the Klein bottle S. So $k=1$ or $k=2$. We may assume that if J is any geodesic freely homotopic to any C_i, and L and L' are two different liftings of J, then $\operatorname{dist}(L, L') > 1$. (Necessarily, L and L' are parallel lines.) By Proposition 13 we may assume that $\operatorname{dev}(C_i) < \frac{1}{4}$.

Then there exist geodesics $J_1, ..., J_k$ such that $C_i \sim J_i^{n_i}$ for some n_i and such that $\operatorname{dist}(D_i, L_i) < \frac{1}{4}$ for some liftings D_i and L_i of C_i and J_i respectively. Let $C'_i \sim C_i$ be such that $C'_1, ..., C'_k$ is minimally crossing. Again by Proposition 13, we may assume that there exist geodesics J'_i such that $C'_i \sim J'^{n_i}_i$ and such that $\operatorname{dist}(D'_i, L'_i) < \frac{1}{4}$ for some liftings D'_i and L'_i of C'_i and J'_i respectively. Since any two different liftings of any J_i are parallel line at least at distance 1 apart, and similarly for any two liftings of any J'_i, and since any liftings of J_i and J'_i are parallel lines for any fixed i, we can delete a point x from S such that no C_i and C'_i traverses x and such that for each i, C_i and C'_i are freely homotopic also in $S \backslash \{x\}$. As $S \backslash \{x\}$ is hyperbolic, Theorem 1 is reduced to the hyperbolic case. ∎

14. FORMULAS FOR CROSSING NUMBERS

As further consequences of the methods given above we give more explicit expressions for the minimal crossing number of closed curves on hyperbolic surfaces.

THEOREM 6. *Let C be a closed curve on a hyperbolic surface, and let J be the geodesic and n the natural number such that $C \sim J^n$. Then*:

(i) $\operatorname{mincr}(C) = n^2 \cdot \operatorname{cr}(J) + n - 1$ *if J is orientation-preserving*,

(ii) $\operatorname{mincr}(C) = n^2 \cdot \operatorname{cr}(J) + \lfloor n - 1/2 \rfloor$ *if J is orientation-reversing*.

Proof. We may assume that $\operatorname{cr}(C) = \operatorname{mincr}(C)$. In particular, no series of Reidemeister moves can decrease $\operatorname{cr}(C)$. Let G be the image of J, and let V and E denote the vertex set and edge set of G. For each $v \in V$, let d_v denote half of the valency of v in G.

We apply the same techniques as in the proof of Proposition 14 to move C close to G. By the fact that $\operatorname{cr}(J) = \sum_{v \in V} \binom{d_v}{2}$ and by (18), (34), (47), and (48), the formulas follow. ∎

THEOREM 7. *Let C, D be two closed curses on a hyperbolic surface, and let J, K be geodesics and m, n be natural numbers such that $C \sim J^m$ and $D \sim K^n$. Then*

(i) $\operatorname{mincr}(C, D) = 2mn \cdot \operatorname{cr}(J) + \min\{m, n\}$ *if $J \sim K$ and C and D are orientation-reversing*,

(ii) $\text{mincr}(C, D) = 2mn \cdot \text{cr}(J)$ *if* $J \sim K$ *and* C *or* D *is orientation-preserving*,

(iii) $\text{mincr}(C, D) = mn \cdot \text{cr}(J, K)$ *if* $J \not\sim K$.

Proof. Similar to the proof of Theorem 6, now using (18), (34), and (39). ∎

ACKNOWLEDGMENTS

We thank Andries E. Brouwer for helpful reference to work on Weyl groups.

REFERENCES

1. R. Baer, Kurventypen auf Flächen, *J. reine angewandte Math.* **156** (1927), 231–246.
2. N. Bourbaki, "Groupes et Algèbres de Lie," Hermann, Paris, 1968.
3. M. Geck and G. Pfeiffer, On the irreducible characters of Hecke algebras, *Advances in Mathematics* **102** (1993), 79–94.
4. M. de Graaf and A. Schrijver, Decomposition of graphs on surfaces, *J. Combin. Theory Ser. B* **70** (1997), 157–165.
5. P. Koebe, Riemannsche Mannigfaltigkeiten und nichteuklidische Raumformen (Vierte Mitteilung), *Sitzungsberichte Deutsche Akademie von Wissenschaften Berlin, mathematische-physikalische Klasse* (1929), 414–457.
6. K. Reidemeister, Elementare Begründung der Knotentheorie, *Abhandlungen aus dem mathematischen Seminar der Hamburgischen Universität* **5** (1926/1927), 24–32.
7. B. L. Reinhart, Algorithms for Jordan curves on compact surfaces, *Annals of Mathematics (2)* **75** (1962), 209–222.
8. G. Ringel, Teilungen der Ebene durch Geraden oder topologische Geraden, *Mathematische Zeitschrift* **64** (1955), 79–102.
9. J. Stillwell, "Geometry of Surfaces," Springer, New York, 1992.