On multiplicity of mappings between surfaces

SEMEON BOGATYI
JAN FRICKE
ELENA KUDRYAVTSEVA

Let M and N be two closed (not necessarily orientable) surfaces, and $f\colon M \to N$ a continuous map. By definition, the *minimal multiplicity* $\mathrm{MMR}[f]$ of the map f denotes the minimal integer k having the following property: f can be deformed into a map g such that the number $|g^{-1}(c)|$ of preimages of any point $c \in N$ under g is $\le k$. We calculate $\mathrm{MMR}[f]$ for any map f of positive absolute degree $A(f)$. The answer is formulated in terms of $A(f)$, $[\pi_1(N) : f_\#(\pi_1(M))]$, and the Euler characteristics of M and N. For a map f with $A(f)=0$, we prove the inequalities $2 \le \mathrm{MMR}[f] \le 4$.

54H25; 57M12, 55M20

In grateful memory of Heiner, his wonderful collaboration and friendship

1 Introduction

For a continuous map $f\colon X \to Y$ between topological spaces, we define the *multiplicity* of f as $\max_{y\in Y} |f^{-1}(y)|$, and the *minimal multiplicity* of f as the minimal multiplicity of maps homotopic to f, that is

$$\mathrm{MMR}[f] := \min_{g\simeq f} \max_{y\in Y} |g^{-1}(y)|.$$

From now on, $\simeq$ means that the mappings are homotopic. The problem of determining $\mathrm{MMR}[f]$ arises. This problem is closely related to the *self-intersection problem* of determining the *minimal self-intersection number* (see Bogatyi, Kudryavtseva and Zieschang [4; 3])

$$\mathrm{MI}[f] := \min_{g\simeq f} |\mathrm{Int}(g)|, \quad \mathrm{Int}(g) := \{(x,y) \in X\times X \mid x \ne y,\ g(x)=g(y)\}/\Sigma_2$$

(here Σ_2 is the symmetric group in two symbols, which acts on $X\times X$ by permutations of the coordinates), and to the problem of determining the *minimal (unordered) μ–tuple*

Published: 29 April 2008 DOI: 10.2140/gtm.2008.14.49

self-intersection number

$$\mathrm{MI}_\mu[f] := \min_{g\simeq f} |\mathrm{Int}_\mu(g)|, \quad \mathrm{Int}_\mu(g) := \{I \subset X \mid |I| = \mu,\ |g(I)| = 1\}, \quad \mu \ge 2.$$

Clearly, $\mathrm{MI}[f] = \mathrm{MI}_2[f]$, and one easily shows[1] that $\mathrm{MI}_{\mu+1}[f] \le (\mathrm{MI}_\mu[f])^2$, $\mu \ge 2$. The connection between $\mathrm{MMR}[f]$ and $\mathrm{MI}_\mu[f]$ is illustrated by the following properties:

$$\mathrm{MI}_\mu[f] = 0 \iff \mathrm{MMR}[f] < \mu \qquad \text{and} \qquad \mathrm{MI}_\mu[f] > 0 \iff \mathrm{MMR}[f] \ge \mu.$$

In particular, $\mathrm{MI}[f] = 0$ if and only if $\mathrm{MMR}[f] = 1$. The numbers $\mathrm{MMR}[f]$, $\mathrm{MI}[f]$, and $\mathrm{MI}_\mu[f]$, measure, in a sense, "complexity" of the self-intersection set $\mathrm{Int}(f)$.

It is natural to consider the above problem for maps $f\colon M^m \to N^n$ between closed connected (nonempty) smooth manifolds, where $m = \dim M$, $n = \dim N$. The problem is nontrivial for $0 < m \le n \le 2m$.

Hurewicz [14] proved that, if X is an m–dimensional compact metric space and $m+1 \le n \le 2m$, then any continuous map $f\colon X \to \mathbb{R}^n$ can be deformed, by means of an arbitrary small perturbation, to a map $g\colon X \to \mathbb{R}^n$ of multiplicity $\le [\frac{n}{n-m}]$. A similar assertion is also valid if the Euclidean space $\mathbb{R}^n$ is replaced by an arbitrary smooth manifold N^n. Thus, for $m < n \le 2m$, we have

$$\mathrm{MMR}[f] \le \left[\frac{n}{n-m}\right]. \tag{1.1}$$

This inequality follows by observing that, for a "generic" map $g\colon M \to N$, the set $\mathrm{Int}_{\mu+1}(g) \subset M$ has dimension $(\mu+1)m - \mu n$, which is negative (and, thus, $\mathrm{MMR}[f] \le \mu$) if $\mu > \frac{m}{n-m}$.

The special case $n = 2m$ is the classical self-intersection problem which gives rise to Whitney's work [21]. Here the estimation (**1.1**) gives $\mathrm{MMR}[f] \in \{1, 2\}$, and computing $\mathrm{MMR}[f]$ is equivalent to deciding whether $\mathrm{MI}[f] = 0$, ie whether the map f is homotopic to an embedding. Namely, we have $\mathrm{MMR}[f] = 1$ if $\mathrm{MI}[f] = 0$, and $\mathrm{MMR}[f] = 2$ if $\mathrm{MI}[f] > 0$. A useful tool for deciding whether $\mathrm{MI}[f] = 0$ is the *Nielsen self-intersection number* $\mathrm{NI}[f]$ of f [4; 3]. One can show by using the Whitney trick [21] that $\mathrm{MI}[f] = \mathrm{NI}[f]$ if $m \ge 3$. But, if $m \le 2$, one has only the inequality $\mathrm{MI}[f] \ge \mathrm{NI}[f]$ (see our papers with Zieschang [4; 3] for $m = 1$). For $m = 1$, there are several combinatorial and geometric methods for deciding whether a closed curve

[1](Indeed, take a map $g \simeq f$ such that $\mathrm{MI}_\mu[f] = |\mathrm{Int}_\mu(g)| =: \ell$. We can assume that $\ell < \infty$. Then $\ell = \sum_{i\ge\mu} \sum_{y\in Y,\ |g^{-1}(y)|=i} \binom{i}{\mu}$. Hence, for every nonvanishing summand in this sum, one has $\binom{i}{\mu} \le \ell$ and

$$\binom{i}{\mu+1} = \binom{i}{\mu}\frac{i-\mu}{\mu+1} < \binom{i}{\mu}\frac{i}{\mu} \le \binom{i}{\mu}^2 \le \ell\binom{i}{\mu}.$$

Therefore $\mathrm{MI}_{\mu+1}[f] \le |\mathrm{Int}_{\mu+1}(g)| = \sum_{i>\mu} \sum_{y\in Y,\ |g^{-1}(y)|=i} \binom{i}{\mu+1}$, which is at most $\ell \sum_{i>\mu} \sum_{y\in Y,\ |g^{-1}(y)|=i} \binom{i}{\mu} \le \ell^2$.)

on a surface is homotopic to a simple closed curve (see, for example, Gonçalves, Kudryavtseva and Zieschang [9] and references therein). An answer in terms of the Nielsen self-intersection number is given in Theorem 2.2. In the remaining case $m=2$, we only know that $\mathrm{NI}[f]>0$ implies $\mathrm{MI}[f]>0$ (and thus $\mathrm{MMR}[f]=2$), but the question whether $\mathrm{NI}[f]=0$ implies $\mathrm{MI}[f]=0$ is still open.

The present paper studies the number $\mathrm{MMR}[f]$ mainly in the case $m=n\le 2$. Here $\mathrm{MMR}[f]$ is closely related to the *absolute degree* $A(f)$ (as defined in Hopf [13] or Epstein [7]; see also Kneser [16], Olum [18] and Skora [19]) of the map f. A definition of the absolute degree is also given in Definition 3.7 in the paper by Gonçalves, Kudryavtseva and Zieschang [10] of this volume. Theorem 2.1 computes the number $\mathrm{MMR}[f]$ for a self-mapping f of a circle ($m=n=1$). In the case $m=n=2$ (mappings between closed surfaces), the following results are obtained. We calculate $\mathrm{MMR}[f]$ in terms of $A(f)$, $\ell(f) := [\pi_1(N) : f_\#\pi_1(M)]$, and the Euler characteristics of the surfaces, for any map $f\colon M\to N$ with $A(f)>0$ (Theorem 3.2 and Theorem 3.3). We also estimate $\mathrm{MMR}[f]$ for any map f with $A(f)=0$ (Theorem 4.2). In particular, we prove that

$$\begin{aligned}
&\mathrm{MMR}[f]\in\{A(f),A(f)+2\} && \text{if}\quad A(f)>0,\\
&\mathrm{MMR}[f]\in\{2,3,4\} && \text{if}\quad A(f)=0.
\end{aligned}$$

The authors do not know whether $\mathrm{MMR}[f]\ge A(f)$ if $m=n\ge 3$.

Acknowledgements This work was partially completed during the visit of the first and the third authors in June–July 2005 at the Fakultät für Mathematik, Universität Siegen, Deutschland. The visits have been supported by the FIGS-project "Niedrigdimensionale Topologie und geometrisch-topologische Methoden in der Gruppentheorie" (1st author), by the Grant of the President RF, project NSh–4578.2006.1, and by a PIMS Fellowship (3rd author).

2 Computing MMR[f] for mappings of a circle

Any map $f\colon S^1\to N$ with $\dim N\ge 3$ is homotopic to an embedding, thus $\mathrm{MMR}[f]=1$. Consider the cases $\dim N=1,2$.

Theorem 2.1 *For any self-map $f\colon S^1\to S^1$,*

$$\mathrm{MMR}[f]=\begin{cases} |\deg f|, & \deg f\ne 0,\\ 2, & \deg f=0.\end{cases}$$

Proof We will identify the circle S^1 with the unit circle in the complex plane $\mathbb{C}$. Consider the projection $p\colon \mathbb{R} \to S^1$, $p(r) = e^{2\pi i r}$, $r \in \mathbb{R}$, of the universal covering $\mathbb{R}$ of S^1 to S^1.

Suppose that $\deg f \neq 0$. Then f is homotopic to the map sending $z \mapsto z^{\deg f}$, $z \in S^1$. Thus all points have exactly $|\deg f|$ preimages, hence $\mathrm{MMR}[f] \leq |\deg f|$. Let us show that the number of preimages can not be reduced. Since $\deg f \neq 0$, for every point $s \in S^1$ there exists a point $t \in S^1$ such that $f(t) = s$. Let $r_0 \in \mathbb{R}$ be a point such that $p(r_0) = t$. Consider a lifting $\widetilde{f}\colon \mathbb{R} \to \mathbb{R}$ of $f\colon S^1 \to S^1$. Then $\widetilde{f}(r_0 + 1) = \widetilde{f}(r_0) + \deg f$, so by the Intermediate Value Theorem, there exist points $r_1, \dots, r_{|\deg f|-1} \in (r_0, r_0 + 1)$ such that $\widetilde{f}(r_i) = \widetilde{f}(r_0) + j \,\mathrm{sgn}(\deg f)$, $1 \leq j \leq |\deg f| - 1$. Thus $p(r_0), p(r_1), \dots, p(r_{|\deg f|-1})$ are different preimages of s under the mapping f. This shows $\mathrm{MMR}[f] \geq |\deg f|$.

Suppose that $\deg f = 0$. Let us show that there exists $g \simeq f$ with $|g^{-1}(s)| \leq 2$ for any $s \in S^1$. Indeed, take g to be the map given by the following rule: $g(z) = z$ if $\operatorname{Im} z \geq 0$, $g(z) = \overline{z}$ if $\operatorname{Im} z \leq 0$. It remains to show that for any $f\colon S^1 \to S^1$, $\deg f = 0$, there exists a point $s \in S^1$ with $|f^{-1}(s)| \geq 2$. Such a map f lifts to a map $\overline{f}\colon S^1 \to \mathbb{R}$, thus it is enough to show that $\overline{f}$ is not an embedding. This can be easily deduced by taking two points $s_0, s_1 \in S^1$ with $\overline{f}(s_0) = \min_{s \in S^1} \overline{f}(s)$, $\overline{f}(s_1) = \max_{s \in S^1} \overline{f}(s)$, and applying the Intermediate Value Theorem to the restriction of $\overline{f}$ to two segments in S^1 having endpoints at s_0, s_1. □

Consider a closed curve $f\colon S^1 \to N^2$ on a closed surface N^2. Then computing $\mathrm{MMR}[f]$ is equivalent to deciding whether the homotopy class $[f]$ of the curve f contains a simple closed curve. Namely, $\mathrm{MMR}[f] = 1$ if $[f]$ contains a simple curve, and $\mathrm{MMR}[f] = 2$ otherwise.

Theorem 2.2 [4; 3] *A closed curve $f\colon S^1 \to N^2$ on a closed surface N^2 is homotopic to a simple closed curve if and only if $\mathrm{NI}[f] = 0$ and one of the following conditions is fulfilled: the curve f is not homotopic to a proper power of any closed curve on N, or $f \simeq g^2$ for some orientation-reversing closed curve $g\colon S^1 \to N$.* □

An analogue of Theorem 2.2 was proved by Turaev and Viro [20, Corollary II], in terms of the intersection index introduced therein.

3 MMR(f) for maps of positive degree between surfaces

In the following, $M = M^2$ and $N = N^2$ are arbitrary connected closed surfaces, ie 2–dimensional manifolds. By $\chi(M)$, we denote the Euler characteristic of M. For a continuous mapping $f\colon M \to N$, $A(f)$ denotes its *absolute degree* (see Hopf [13], Epstein [7], Kneser [16], Olum [18], Skora [19] or Gonçalves, Kudryavtseva and Zieschang [8]). Denote the index of the image of the fundamental group of M in the fundamental group of N by $\ell(f) := [\pi_1(N, f(x_0)) : f_\#(\pi_1(M, x_0))]$ for some $x_0 \in M$. Actually the number $\ell(f)$ does not depend on the choice of the point x_0.

The following consequence of Kneser's inequality will be central in the proof of our main result.

Proposition 3.1 *If $f\colon M \to N$ has absolute degree $d = A(f) > 0$ then there are at most $d \cdot \chi(N) - \chi(M)$ points in N whose preimages have cardinality $\le d - 1$. Moreover, if pairwise different points $y_1, \dots, y_r$ of N have $\mu_1, \dots, \mu_r$ preimages, respectively, then*

$$d \cdot \chi(N) \ge \chi(M) + \sum_{i=1}^{r} (d - \mu_i).$$

Proof In the case when $r = 1$ and f is orientation-true, the latter inequality was proved in Theorem 2.5 (a) of [8]. In the general case, the inequality can be proved using the techniques in [1; 8; 11; 2], as follows.

If f is not orientation-true and $d = A(f) > 0$ then $d = \ell(f)$, due to the result of Kneser [15; 16]. On the other hand, one has $\mu_i \ge \ell(f)$, $1 \le i \le r$, since the map f admits a lifting $\widehat{f}\colon M \to \widehat{N}$ such that $f = p \circ \widehat{f}$, where $p\colon \widehat{N} \to N$ is an $\ell(f)$–fold covering corresponding to the subgroup $f_\#(\pi_1(M, x_0))$ of $\pi_1(N, f(x_0))$, and $A(\widehat{f}) = 1$, hence $\widehat{f}$ is surjective. Therefore $\sum_{i=1}^{r} (d - \mu_i) \le 0$. This, together with the Kneser inequality [16], $d \cdot \chi(N) \ge \chi(M)$, implies the desired inequality.

If f is orientation-true, one proceeds as in the proof of Proposition 2.5 (a) of [8], where one replaces the single point $y_0 \in N$ by the set of r points $y_1, \dots, y_r$. More specifically, by applying a suitable deformation, one can assume that there are small pairwise disjoint disks D_i, D_{ij}, $1 \le i \le r$, $1 \le j \le \mu_i$, around the points y_i of N and the points of $f^{-1}(y_i)$ such that $f^{-1}(\mathring{D}_i) = \bigcup_{j=1}^{\mu_i} \mathring{D}_{ij}$, and $f|_{D_{ij}}$ is a branched covering of type $z \mapsto z^{d_{ij}}$ for some positive integer d_{ij}. Therefore the complement of these open disks are two compact surfaces $F \subset M$, $G \subset N$ such that the restriction of f induces a proper map carrying the boundary into the boundary, $f|_F\colon (F, \partial F) \to (G, \partial G)$. By Proposition 1.6 of [8] (or by a more general Theorem 4.1 of [19]), $\chi(F) \le A(f) \cdot \chi(G)$. This, together with $\chi(F) = \chi(M) - \sum_{i=1}^{r} \mu_i$, $\chi(G) = \chi(N) - r$, gives the desired inequality. □

Theorem 3.2 *Suppose that $f\colon M \to N$ has absolute degree $d = A(f) > 0$. If $\ell(f) \ne d$, or $\ell(f) = d$ and $d \cdot \chi(N) = \chi(M)$, then* $\mathrm{MMR}[f] = d$.

Proof The inequality $\mathrm{MMR}[f] \geq A(f)$ follows from the first part of Proposition 3.1.

Let us show the converse inequality, $\mathrm{MMR}[f] \leq A(f)$. It follows from [6; 19; 16], respectively, that the mapping f is homotopic to a d–fold covering which is branched in the first case and unbranched in the second case. Thus, we found a mapping which is homotopic to f, and the preimage of any point of N has cardinality $\leq d$. □

Theorem 3.3 *Suppose that* $f\colon M \to N$ *has absolute degree* $d = A(f) > 0$. *If* $\ell(f) = d$ *and* $d \cdot \chi(N) \neq \chi(M)$, *then* $\mathrm{MMR}[f] = d + 2$.

Proof **Case 1** Suppose that $d = A(f) = 1$. It follows from [6; 19] that the mapping f is homotopic to a pinching map where the pinched subsurface $M' \subset M$, $\partial M' \simeq S^1$, is different from the 2–disk D^2 (here the natural projection $M \to M/M'$ is called a pinching map).

Let us show that such a pinching map is homotopic to a map g of multiplicity ≤ 3. For this, we construct a proper continuous map $g'\colon (M', \partial M') \to (D^2, \partial D^2)$ whose restriction to $\partial M'$ is a homeomorphism, and whose multiplicity equals 3. Such a map g' is shown in Figure 1. We may identify N with the surface which is obtained by gluing of $M \setminus \mathring{M}'$ and D^2 by means of the aforementioned homeomorphism of the boundary circles, where $\mathring{M}'$ denotes the interior of M'. Define $g\colon M \to N$ as $g|_{M \setminus M'} = \mathrm{id}_{M \setminus M'}$ and $g|_{M'} = g'$. Clearly, $f \simeq g$, since g' is homotopic relative boundary to a pinching map. In Case 2 below, we will use the following property of the constructed map g: its restriction to the preimage of the complement $N \setminus D^2$ of a disk is injective.

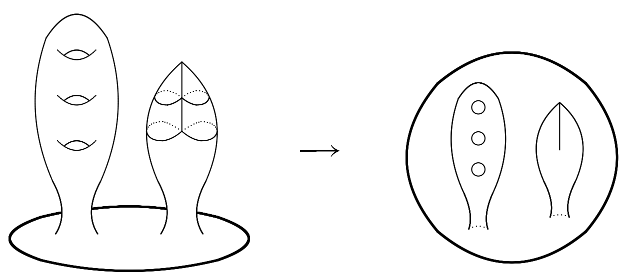


Figure 1: A proper map $g'\colon M' \to D^2$ of multiplicity 3

It follows from the inequality of Euler characteristics of M and N that f is not homotopic to an embedding. (Indeed, otherwise such an embedding g is a homeomorphism onto $g(M)$; it follows from Brouwer's Theorem on Invariance of Domain [5] that g

is surjective and, therefore, it is a homeomorphism.) Suppose that f is homotopic to a map $g\colon M \to N$ of multiplicity 2, we will show that this leads to a contradiction. Let $y \in N$ be a point with $g^{-1}(y) = \{x_1, x_2\}$. Then the local degree of g at each of the points x_1 and x_2 is defined modulo 2, and

$$\deg(g, x_1) + \deg(g, x_2) \equiv A(g) \equiv A(f) \equiv 1 \mod 2.$$

Without loss of generality, we may assume that $\deg(g, x_1) \neq 0$. This implies that the image of any neighbourhood of x_1 contains a neighbourhood of $y = g(x_1)$, since otherwise one could construct a map $F\colon D^2 \to S^1$ with $\deg(F|_{\partial D^2}) = \deg(g, x_1) \neq 0$. Therefore the restriction of g to an appropriate neighbourhood of x_2 is injective and, thus (by Brouwer's Theorem on Invariance of Domain [5]), is a homeomorphism onto a neighbourhood of y. This implies that $\deg(g, x_2) = \pm 1$. Similar arguments show that $\deg(g, x_1) = \pm 1$, a contradiction.

Case 2 Suppose that $d = A(f) = \ell(f) \geq 2$. Let us construct a map g which is homotopic to f and has multiplicity $A(f)+2$. Consider a covering $p\colon \widetilde{N} \to N$ which corresponds to the subgroup $f_\#(\pi_1(M, x_0))$ of $\pi_1(N, f(x_0))$. So, this is an $\ell(f)$–fold covering. Let $y \in N$ be an arbitrary point and D a small closed neighbourhood which is homeomorphic to the disk D^2. Let $D_1, \dots, D_d$ be the connected components of $p^{-1}(D)$.

Let $\widetilde{f}\colon M \to \widetilde{N}$ be a lifting of f. Then $A(\widetilde{f}) = \ell(\widetilde{f}) = 1$. By Case 1, there exists a map $\widetilde{g}\colon M \to \widetilde{N}$ which is homotopic to $\widetilde{f}$ and has multiplicity ≤ 3. Then the map $g := p \circ \widetilde{g}$ is homotopic to $f = p \circ \widetilde{f}$. By Case 1, we may also assume that $\widetilde{g}$ is injective on $\widetilde{g}^{-1}(\widetilde{N} \setminus D_1)$. Therefore the map g has multiplicity $\ell(f)+2 = A(f)+2$.

Let us show that the multiplicity of f is $\geq \ell(f)+2$. Let $\widetilde{f}\colon M \to \widetilde{N}$ be a lifting of f to this $\ell(f)$–fold covering, thus $A(\widetilde{f}) = \ell(\widetilde{f}) = 1$. By Case 1, there exists a point $\widetilde{y} \in \widetilde{N}$ whose preimage under $\widetilde{f}$ has cardinality ≥ 3. Since $A(\widetilde{f}) > 0$, every point of $p^{-1}(p(\widetilde{y}))$ has a nonempty preimage under $\widetilde{f}$. Therefore $f^{-1}(p(\widetilde{y}))$ has cardinality at least $\ell(f)+2 = A(f)+2$. $\square$

4 Estimates for MMR(f) if $A(f) = 0$

Suppose that M is a connected orientable closed surface of genus $g \geq 0$. Consider the standard presentation of the closed surface M as the boundary of a solid surface in $\mathbb{R}^3$ which is obtained from a closed 3-ball by attaching g solid handles; see Figure 2 (a). Choose a base point $x_0 \in M$ and consider a system of simple closed curves $\alpha_1, \beta_1, \dots, \alpha_g, \beta_g$ on M based at x_0 which form a *canonical system of cuts*; see

Figure 2 (a). Then the fundamental group $\pi_1(M,x_0)$ admits a canonical presentation

$$\pi_1(M,x_0)=\left\langle a_1,b_1,\dots,a_g,b_g \;\middle|\; \prod_{j=1}^{g}[a_j,b_j]\right\rangle,$$

where a_j,b_j are the homotopy classes of the based loops α_j,β_j, respectively. Denote by V_g the bouquet of g circles $\alpha_1\cup\ldots\cup\alpha_g$ if $g\geq 1$, $V_0:=\{x_0\}$ if $g=0$, and by ϱ a retraction $\varrho\colon M\to V_g$ which maps all loops β_j to the point x_0. We can assume that the curves $\alpha_1,\dots,\alpha_g$ are contained in the plane $\Pi\subset\mathbb{R}^3$ which is tangent to M at x_0. (In Figure 2, the plane Π is parallel to the plane of the picture.)

Let $i\colon M\to\mathbb{R}^3$ denote the inclusion, and $p_\Pi\colon\mathbb{R}^3\to\Pi$ the orthogonal projection. The following properties of the map $p=p_\Pi\circ i\colon M\to\Pi$ can be assumed without loss of generality, and will be used later:

(p1) The restriction of p to a neighbourhood U of the base point $x_0\in M$ is a homeomorphism onto a neighbourhood of the point $p(x_0)$ in Π. Moreover, $p|_{V_r}\colon V_r\to\Pi$ is an embedding, and all curves $p|_{\alpha_j}\colon\alpha_j\to\Pi$ are regular;

(p2) All curves $p|_{\beta_j}$ are contractible in $p(M)$;

(p3) $p(M)$ is a regular neighbourhood of the graph $p(V_r)$ in Π;

(p4) The map p has multiplicity 2.

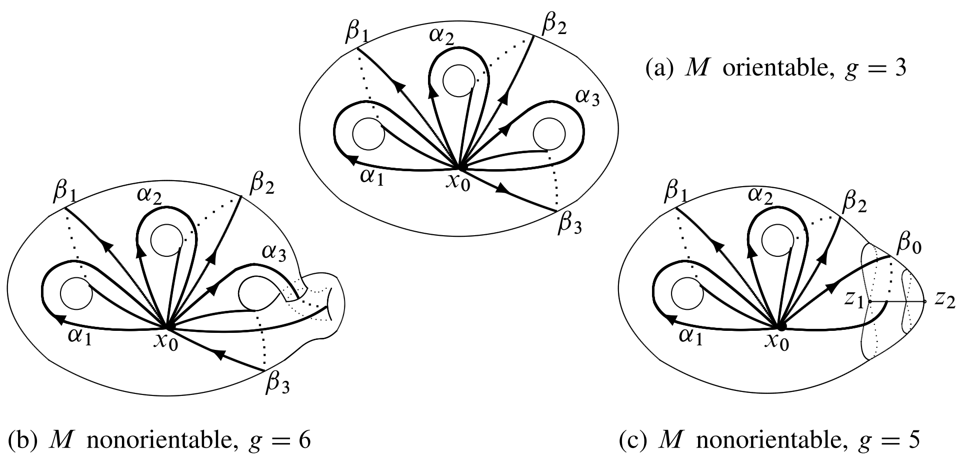


Figure 2: A canonical system of cuts on a closed surface M

Suppose that M is a connected nonorientable closed surface of genus $g \geq 1$. Choose a base point $x_0 \in M$. Then the fundamental group of M admits the following canonical presentation:

$$\pi_1(M, x_0) = \left\langle a_1, b_1, \dots, a_{g/2}, b_{g/2} \;\middle|\; \left(\prod_{j=1}^{g/2-1} [a_j, b_j] \right) \cdot [a_{g/2}, b_{g/2}]_- \right\rangle \quad \text{if } g \text{ is even,}$$

$$\pi_1(M, x_0) = \left\langle a_1, b_1, \dots, a_{[g/2]}, b_{[g/2]}, b_0 \;\middle|\; \left(\prod_{j=1}^{(g-1)/2} [a_j, b_j] \right) \cdot b_0^2 \right\rangle \qquad \text{if } g \text{ is odd,}$$

where we use the notation

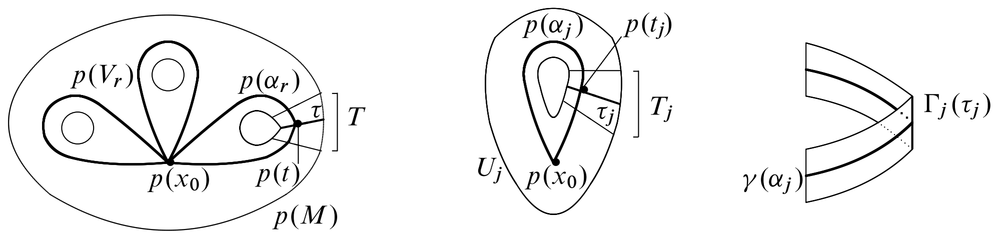
$$[x, y] = xyx^{-1}y^{-1}, \qquad [x, y]_- = xyx^{-1}y.$$

This presentation of the group $\pi_1(M, x_0)$ corresponds to a system of simple closed curves $\alpha_1, \beta_1, \dots, \alpha_{[g/2]}, \beta_{[g/2]}, \beta_0$ on M based at x_0, which form a *canonical system of cuts*; see Figure 2 (b), (c). Here the curve β_0 appears only if g is odd. Denote by V_r the bouquet of $r = [g/2]$ circles $\alpha_1 \cup \ldots \cup \alpha_{[g/2]}$ for $g \geq 2$, $V_0 = \{x_0\}$ for $g = 1$, and by ϱ a retraction $\varrho\colon M \to V_r$ which maps all loops β_j to the point x_0. We consider a realization of M in $\mathbb{R}^3$ via a map $i\colon M \to \mathbb{R}^3$ which is an immersion if g is even (see Figure 2 (b)), while, for g odd, the restriction $i|_{M \setminus \{z_1, z_2\}}$ to the complement of the set of two points $z_1, z_2 \in M \setminus \{x_0\}$ is an immersion; see Figure 2 (c). We can assume that $i|_{V_r}$ is an embedding with $i(V_r) \subset \Pi$, moreover Π coincides with the tangent plane to $i(M)$ at $i(x_0)$.

Let $p_\Pi\colon \mathbb{R}^3 \to \Pi$ denote the orthogonal projection. Without loss of generality, we may assume that the map $p = p_\Pi \circ i\colon M \to \Pi$ has the properties (p1), (p2), (p3) from above. Moreover, (p4) holds if g is odd, while the following property holds if g is even:

(p4$'$) The map p has multiplicity 4. Moreover, the set of all points of $p(M)$, whose preimage under p contains more than 2 points, lies in a regular neighbourhood T in $p(M)$ of a simple arc $\tau \subset p(M)$, where the endpoints of τ lie on the boundary of $p(M)$, τ intersects the graph $p(V_r)$ at the unique point $p(t)$, for some $t \in \alpha_r \setminus \{x_0\}$, and the intersection of τ and $p(\alpha_r)$ at the point $p(t)$ is transverse; see Figure 3 (a).

Proposition 4.1 *Suppose that M is an (orientable or nonorientable) closed surface of genus g, and $f\colon M \to N$ has absolute degree $A(f) = 0$. Then there exists a self-homeomorphism φ of M and a map $\gamma\colon V_r \to N$ such that $f \simeq \gamma \circ \varrho \circ \varphi$. Here $r = 2g$ if M is orientable, $r = [g/2]$ if M is nonorientable, and $\varrho\colon M \to V_r$ is the retraction defined above.*



(a) $T \subset p(M)$ (b) $T_j \subset U_j$ (c) $\Gamma_j(T_j) \subset N$

Figure 3: The strips T, T_j and "folding" of T_j via Γ_j

Proof Since $A(f) = 0$, it follows from [16] or [7] that f is homotopic to a map h which is not surjective; thus $h(M) \subset N^* = N \setminus \mathring{D}^2$ for an appropriate disk $D^2 \subset N$. Since the fundamental group of N^* is a free group, we obtain a homomorphism $h_\#\colon \pi_1(M) \to \pi_1(N^*)$ to the free group $\pi_1(N^*)$.

Suppose that M is orientable. It has been proved in Satz 2 of Zieschang [22] using the Nielsen method (see also Zieschang, Vogt and Coldewey [23], or Proposition 1.2 of Grigorchuk, Kurchanov and Zieschang [12]) that there is a sequence of "elementary moves" of the system of generators $a_1, b_1, \dots, a_g, b_g$ and the corresponding sequence of "elementary moves" of the system of cuts $\alpha_1, \beta_1, \dots, \alpha_g, \beta_g$ on M (see above), such that the resulting system of cuts $\widetilde{\alpha}_1, \widetilde{\beta}_1, \dots, \widetilde{\alpha}_g, \widetilde{\beta}_g$ is also canonical (this means there exists a self-homeomorphism φ of M such that $\alpha_j = \varphi(\widetilde{\alpha}_j)$, $\beta_j = \varphi(\widetilde{\beta}_j)$), and the loops $h|_{\widetilde{\beta}_j}\colon \widetilde{\beta}_j \to N^*$ are contractible in N^*. From this, using the fact that $\pi_2(N^*) = 0$, one can prove that $h \simeq \gamma \circ \varrho \circ \varphi$ where $\gamma := h|_{V_g}$.

Suppose that M is nonorientable. The method to prove Satz 2 of [22] can be successfully applied to construct a canonical system of cuts $\widetilde{\alpha}_1, \widetilde{\beta}_1, \dots, \widetilde{\alpha}_{[g/2]}, \widetilde{\beta}_{[g/2]}, \widetilde{\beta}_0$ on M (this means there exists a homeomorphism φ of M with $\alpha_j = \varphi(\widetilde{\alpha}_j)$, $\beta_j = \varphi(\widetilde{\beta}_j)$) such that the loops $h|_{\widetilde{\beta}_j}\colon \widetilde{\beta}_j \to N^*$ are contractible in N^*; see Ol'shanskiĭ [17] or Proposition 1.5 of [12]. (Again the curve β_0 is considered only if g is odd.) Similarly to the orientable case, this implies that $h \simeq \gamma \circ \varrho \circ \varphi$ where $\gamma := h|_{V_g}$. □

Theorem 4.2 *Suppose that $f\colon M \to N$ has absolute degree $A(f) = 0$. Then* $2 \le \mathrm{MMR}[f] \le 4$.

Proof Suppose that h is homotopic to f and has multiplicity 1. Then h is a homeomorphism onto $h(M)$. It follows from Brouwer's Theorem on Invariance of

Domain [5] that h is surjective and, therefore, it is a homeomorphism. Then $A(h)=1$, a contradiction. Therefore $\mathrm{MMR}[f]\geq 2$.

Let us prove the second inequality. Since $A(f)=0$, by Proposition 4.1, $f\simeq\gamma\circ\varrho\circ\varphi$ for a self-homeomorphism φ of M, the retraction $\varrho\colon M\to V_r$, and a map $\gamma\colon V_r\to N$, where $r=g$ if M is an orientable surface of genus g, $r=[g/2]$ if M is a nonorientable surface of genus g. Without loss of generality, we may assume that γ has the following properties:

($\gamma 1$) There exists a homeomorphism ψ of the neighbourhood U of x_0 in M onto a neighbourhood of $\gamma(x_0)$ in N such that $\gamma|_{V_r\cap U}=\psi|_{V_r\cap U}$. In other words, $\gamma|_{V_r\cap U}$ extends to an embedding $\psi\colon U\to N$;

($\gamma 2$) The restriction of γ onto each curve $\alpha_1,\dots,\alpha_r$ is an immersion $S^1\to N$. Moreover, γ has multiplicity ≤ 2, and it has only finitely many double points (ie pairs of distinct points of V_r having the same image).

Case 1 Suppose that the surface M is either orientable (thus $r=g$), or nonorientable with g odd (thus $r=(g-1)/2$). In both cases, the map $p=p_\Pi\circ i\colon M\to\Pi=\mathbb{R}^2$ of M to the plane Π has the properties (p1), (p2), (p3), (p4); see above.

Subcase 1 Suppose that N is orientable. Since every closed curve $\gamma|_{\alpha_j}$ is orientation-preserving, it follows from the properties $(\gamma 1)$, $(\gamma 2)$, (p1), (p3) that the map $\hat\gamma=\gamma\circ p^{-1}\colon p(V_r)\to N$ can be extended to an immersion $\Gamma\colon p(M)\to N$ of the regular neighbourhood $p(M)$ of $p(V_r)$ in the plane Π to N, such that Γ has multiplicity ≤ 2.

Consider the composition $\hat\varrho=p\circ\varrho\colon M\to\Pi$. Observe that the maps $\hat\varrho$ and p are homotopic as maps $M\to p(M)\subset\Pi$ with the target $p(M)$, due to $\hat\varrho|_{V_r}=p|_{V_r}$, (p2), and $\pi_2(p(M))=0$. From this and $\gamma=\Gamma\circ p|_{V_r}$, we have

$$f\simeq\gamma\circ\varrho\circ\varphi=\Gamma\circ p\circ\varrho\circ\varphi\simeq\Gamma\circ p\circ\varphi. \tag{4.3}$$

Since φ is bijective and each of Γ and p has multiplicity ≤ 2 (see (p4)), the multiplicity of the composition $\Gamma\circ p\circ\varphi$ is $\leq 2\cdot 2\cdot 1=4$.

Subcase 2 Suppose that N is nonorientable. So in general, the immersion $\hat\gamma\colon p(V_r)\to N$ can not be extended to an immersion of the regular neighbourhood $p(M)$ of $p(V_r)$ in $\Pi=\mathbb{R}^2$. However, due to $(\gamma 1)$, $(\gamma 2)$, and (p1), we can extend $\hat\gamma$ to an immersion $\widetilde\Gamma\colon p(D\cup V_r)\to N$, where $D\subset U$ is a small disk centred at x_0.

Now, for each curve α_j, we will extend the immersion $\widetilde\Gamma_j=\widetilde\Gamma|_{p(D\cup\alpha_j)}\colon p(D\cup\alpha_j)\to N$ to a regular neighbourhood $U_j\supset p(D)$ of $p(\alpha_j)$ in Π as follows. If the curve $\gamma|_{\alpha_j}$ is orientation-preserving then, similarly to Case 1, the immersion $\widetilde\Gamma_j\colon p(D\cup\alpha_j)\to N$

can be extended to an immersion $\Gamma_j\colon U_j \to N$. If the curve $\gamma|_{\alpha_j}$ is orientation-reversing, let us choose a point $t_j \in \alpha_j \setminus D$ such that t_j is the only preimage of the point $\gamma(t_j)$ under γ. Consider a simple arc $\tau_j \subset U_j \setminus p(D)$, which transversally intersects $p(\alpha_j)$ at the only point $p(t_j)$, and whose endpoints lie on the boundary of U_j. Let T_j be a regular neighbourhood of the arc τ_j in $U_j \setminus p(D)$, thus T_j is a "strip" in the annulus U_j; see Figure 3 (b). Outside the interior of the strip T_j, we extend $\widetilde{\Gamma}_j$ to an immersion $\overline{\Gamma}_j\colon (U_j \setminus T_j) \cup p(\alpha_j) \to N$ similarly to above. Now we extend the obtained immersion $\overline{\Gamma}_j$ to the whole annulus U_j, giving a map $\Gamma_j\colon U_j \to N$ which coincides with $\overline{\Gamma}_j$ outside $T_j \setminus p(\alpha_j)$ and has a "folding" along the arc $\tau_j \subset T_j$, as shown in Figure 3 (c).

Without loss of generality, we may assume that $U_j \subset p(M)$, and any two annuli U_j, U_k have only the disk $p(D)$ in common. Since the constructed mappings $\Gamma_j\colon U_j \to N$ agree on the common part $p(D)$, they determine an extension $\overline{\Gamma}\colon U \to N$ of the map $\widetilde{\Gamma}$, where $U = U_1 \cup \ldots \cup U_r$ is a regular neighbourhood of $p(V_r)$ in $\Pi = \mathbb{R}^2$. The above construction can be performed in such a way that the map $\overline{\Gamma}$ has multiplicity ≤ 2, due to $(\gamma 2)$ and the choice of the points $t_j \in \alpha_j$. Obviously, the map $\overline{\Gamma}$ can be extended to the regular neighbourhood $p(M)$ of $p(D \cup V_r)$ (see (p3)) and the extended map $\Gamma\colon p(M) \to N$ also has multiplicity ≤ 2.

Similarly to Subcase 1, the composition $\Gamma \circ p \circ \varphi$ has multiplicity $\le 2 \cdot 2 \cdot 1 = 4$, and (**4.3**) holds. This completes the proof in Case 1.

Case 2 Suppose that M is a nonorientable closed surface of even genus g, thus $r = g/2$, and the map $p = p_\Pi \circ i\colon M \to \Pi = \mathbb{R}^2$ of M to the plane Π has the properties (p1), (p2), (p3), (p4′); see above. We may assume, without loss of generality, that the map $\gamma\colon V_r \to N$ has the following additional property:

$(\gamma 3)$ The point $t \in \alpha_r$ considered in (p4′) is the only preimage of $\gamma(t)$ under γ, and the analogous property holds for any point $\tilde{t} \in \alpha_r \cap p^{-1}(T)$.

Subcase 1 Suppose that N is orientable. Similarly to Subcase 1 of Case 1, one shows using $(\gamma 1)$, $(\gamma 2)$, (p1), (p3) that the immersion $\hat{\gamma} = \gamma \circ p^{-1}\colon p(V_r) \to N$ extends to an immersion $\Gamma\colon p(M) \to N$ of multiplicity 2, and using (p2) that (**4.3**) holds. Taking into account (p4′) and $(\gamma 3)$, one can show that the multiplicity of $\Gamma \circ p \circ \varphi$ is ≤ 4.

Subcase 2 Suppose that N is nonorientable. We proceed as in Subcase 2 of Case 1. Namely, for those curves α_j whose image under γ is orientation-preserving, we extend the immersion $\widetilde{\Gamma}_j\colon p(D \cup \alpha_j) \to N$ to U_j, as in Case 1. For each of the remaining curves α_j, we choose a point $t_j \in \alpha_j \setminus D$ which is the only preimage of $\gamma(t_j)$ under γ, and we extend the corresponding immersion $\widetilde{\Gamma}_j$ to a map $\overline{\Gamma}_j\colon U_i \to N$ having a "folding" along an arc $\tau_j \subset T_j \subset U_j$, which transversally intersects $p(V_r)$ at the unique point $p(t_j)$; see Case 1. As above, this allows one to construct a map $\Gamma\colon\ p(M) \to N$ of multiplicity ≤ 2 which is an extension of $\hat{\gamma}$, and to show that (**4.3**) holds. Observe now that, if the curve $\gamma|_{\alpha_r}$ is orientation-reversing, we can choose the point $t_r \in \alpha_r$ in such a way that it is "far enough" from the point $t \in \alpha_r$ considered in (p4′). This, together with $(\gamma 3)$, shows that the above construction can be performed in such a way that the composition $\Gamma \circ p \circ \varphi$ has multiplicity ≤ 4. This completes the proof of Theorem 4.2. □

References

[1] **S A Bogatyĭ**, **D L Gonçalves**, **E Kudryavtseva**, **H Zieschang**, *The minimal number of preimages under mappings of surfaces*, Mat. Zametki 75 (2004) 13–19 MR2053145

[2] **S Bogatyi**, **D L Gonçalves**, **H Zieschang**, *The minimal number of roots of surface mappings and quadratic equations in free groups*, Math. Z. 236 (2001) 419–452 MR1821299

[3] **S A Bogatyi**, **E A Kudryavtseva**, **H Zieschang**, *On intersections of closed curves on surfaces*, preprint (2004)

[4] **S A Bogatyi**, **E A Kudryavtseva**, **H Zieschang**, *On the coincidence points of mappings of a torus into a surface*, Tr. Mat. Inst. Steklova 247 (2004) 15–34 MR2168160

[5] **L E J Brouwer**, *Beweis der Invarianz des n–dimensionalen Gebiets*, Math. Ann. 71 (1911) 305–313 MR1511658

[6] **A L Edmonds**, *Deformation of maps to branched coverings in dimension two*, Ann. of Math. (2) 110 (1979) 113–125 MR541331

[7] **D B A Epstein**, *The degree of a map*, Proc. London Math. Soc. (3) 16 (1966) 369–383 MR0192475

[8] **D L Gonçalves**, **E Kudryavtseva**, **H Zieschang**, *Roots of mappings on nonorientable surfaces and equations in free groups*, Manuscripta Math. 107 (2002) 311–341 MR1906200

[9] **D L Gonçalves**, **E Kudryavtseva**, **H Zieschang**, *An algorithm for minimal number of intersection points of curves on surfaces*, Proc. Seminar on Vector and Tensor Analysis 26 (2005) 139–167

[10] **D L Gonçalves**, **E Kudryavtseva**, **H Zieschang**, *Some quadratic equations in the free group of rank 2*, from: "The Zieschang Gedenkschrift", (M Boileau, M Scharlemann, R Weidmann, editors), Geom. Topol. Monogr. 14 (2008) 219–294

[11] **D L Gonçalves**, **H Zieschang**, *Equations in free groups and coincidence of mappings on surfaces*, Math. Z. 237 (2001) 1–29 MR1836771

[12] **R I Grigorchuk**, **P F Kurchanov**, **H Zieschang**, *Equivalence of homomorphisms of surface groups to free groups and some properties of* 3*–dimensional handlebodies*, from: "Proceedings of th eInternational Conference on Algebra, Part 1 (Novosibirsk, 1989)", Contemp. Math. 131, Amer. Math. Soc., Providence, RI (1992) 521–530 MR1175803

[13] **H Hopf**, *Zur Topologie der Abbildungen von Mannigfaltigkeiten*, Math. Ann. 102 (1930) 562–623 MR1512596

[14] **W Hurewicz**, *Über Abbildungen von endlichdimensionalen Räumen auf Teilmengen Cartesischer Räume*, Sitzungsberichte Preuss. Akad. Wiss., Phys.-Math. Kl. 24/25 (1933) 754–768 MR1362795

[15] **H Kneser**, *Glättung von Flächenabbildungen*, Math. Ann. 100 (1928) 609–617 MR1512504

[16] **H Kneser**, *Die kleinste Bedeckungszahl innerhalb einer Klasse von Flächenabbildungen*, Math. Ann. 103 (1930) 347–358 MR1512626

[17] **A Y Ol'shanskiĭ**, *Diagrams of homomorphisms of surface groups*, Sibirsk. Mat. Zh. 30 (1989) 150–171 MR1043443

[18] **P Olum**, *Mappings of manifolds and the notion of degree*, Ann. of Math. (2) 58 (1953) 458–480 MR0058212

[19] **R Skora**, *The degree of a map between surfaces*, Math. Ann. 276 (1987) 415–423 MR875337

[20] **V G Turaev**, **O Y Viro**, *Intersection of loops in two-dimensional manifolds. II. Free loops*, Mat. Sb. (N.S.) 121(163) (1983) 359–369 MR708001

[21] **H Whitney**, *The self-intersections of a smooth n–manifold in $2n$–space*, Ann. of Math. (2) 45 (1944) 220–246 MR0010274

[22] **H Zieschang**, *Alternierende Produkte in freien Gruppen*, Abh. Math. Sem. Univ. Hamburg 27 (1964) 13–31 MR0161901

[23] **H Zieschang**, **E Vogt**, **H-D Coldewey**, *Surfaces and planar discontinuous groups*, Lecture Notes in Mathematics 835, Springer, Berlin (1980) MR606743 Translated from the German by John Stillwell

SB, EK: *Department of Mathematics and Mechanics, Moscow State University*
Moscow 119992, Russia

JF: *Fachbereich 6 – Mathematik, Universität Siegen*
57068 Siegen, Germany

bogatyi@inbox.ru, fricke@math.uni-siegen.de, eakudr@mech.math.msu.su

Received: 29 April 2006