

*Pacific  
Journal of  
Mathematics*

NIELSEN ROOT THEORY AND HOPF DEGREE THEORY

ROBERT F. BROWN AND HELGA SCHIRMER

Volume 198 No. 1

March 2001



## NIELSEN ROOT THEORY AND HOPF DEGREE THEORY

ROBERT F. BROWN AND HELGA SCHIRMER

The Nielsen root number  $N(f; c)$  of a map  $f: M \rightarrow N$  at a point  $c \in N$  is a homotopy invariant lower bound for the number of roots at  $c$ , that is, for the cardinality of  $f^{-1}(c)$ . There is a formula for calculating  $N(f; c)$  if  $M$  and  $N$  are closed oriented manifolds of the same dimension. We extend the calculation of  $N(f; c)$  to manifolds that are not orientable, and also to manifolds that have non-empty boundaries and are not compact, provided that the map  $f$  is boundary-preserving and proper. Because of its connection with degree theory, we introduce the transverse Nielsen root number for maps transverse to  $c$ , obtain computational results for it in the same setting, and prove that the two Nielsen root numbers are sharp lower bounds in dimensions other than 2. We apply these extended root theory results to the degree theory for maps of not necessarily orientable manifolds introduced by Hopf in 1930. Thus we re-establish, in a new and modern treatment, the relationship of Hopf's Absolutgrad and the geometric degree with homotopy invariants of Nielsen root theory, a relationship that is present in Hopf's work but not in subsequent re-examinations of Hopf's degree theory.

### 1. Introduction.

The goal of this paper is two-fold. We will extend results from Nielsen root theory for maps between orientable  $n$ -manifolds so as to remove the orientability hypothesis. Then we will use the extended theory to re-establish the connection between Nielsen root theory and two variants of the degree of a map, namely, Hopf's Absolutgrad and the geometric degree. By using methods from present-day Nielsen theory, we will provide new ways of understanding some of the basic concepts of Hopf's theory as well as more direct proofs for some of the results. We next describe these goals in more detail.

If  $f: M \rightarrow N$  is a map between two manifolds and  $c \in N$ , then a *root* of  $f$  at  $c$  is a point in  $f^{-1}(c)$ . The Nielsen root number  $N(f; c)$  is a lower bound for the cardinality of  $f^{-1}(c)$ , and it is homotopy invariant. While it is possible to define  $N(f; c)$  even if  $M$  and  $N$  are not manifolds, it is usually not possible to compute it in such general settings. If, however,  $M$  and  $N$

are orientable  $n$ -manifolds, then a formula for computing  $N(f; c)$  is known. Following Hopf [H2], we write  $j$  to denote the cardinality of the coset space  $\pi_1(N)/f_\pi(\pi_1(M))$  and state the following partial version of Theorem 3.13 below, which is due to Hopf [H2, Satz VIIa] and Lin [L, Proposition 5].

**Theorem 1.1.** *If  $f: M \rightarrow N$  is a map of closed, connected oriented  $n$ -manifolds, then  $N(f; c) = 0$  if the degree of  $f$  is zero and  $N(f; c) = j$  if the degree is not zero.*

Our main extension of this theorem is to maps between closed  $n$ -manifolds that are not necessarily orientable, but we will also allow manifolds with boundary if  $f$  maps boundary to boundary, and non-compact manifolds if  $f$  is proper (see Theorem 3.11).

Nielsen root theory was used in the degree theory that Heinz Hopf initiated in 1930 [H2], and therefore his degree theory is quite different from others that existed at Hopf's time. The degree in Theorem 1.1, the classical degree due to Brouwer [Bw, p. 105], is usually defined in terms of the homomorphism of integer homology induced by  $f$ . The definition can be extended to proper boundary-preserving maps of orientable but not necessarily compact  $n$ -manifolds with boundary. But if at least one of  $M$  and  $N$  is non-orientable, then the homological degree can only be defined in terms of homology with coefficients in  $\mathbb{Z}/2$  and the resulting mod 2 degree  $\deg(f; 2)$  tells little about the map  $f$ , in particular about its geometric properties. To obtain such geometric information, and in particular to give an algebraic approach to the geometric degree which looks at counterimages of points, Hopf introduced a degree that he called the Absolutgrad (absolute degree). It does not require orientability and provides much better information about the map than does the mod 2 degree. Hopf's Absolutgrad may be viewed as a variant of the Nielsen root number, in fact it is precisely the "transverse Nielsen root number"  $N_{\pitchfork}(f; c)$  which is a lower bound for the cardinality of the set of roots of maps between  $n$ -manifolds that are transverse to  $c$  in a sense made precise in Definition 3.1 (see Theorem 5.3). In general,  $N_{\pitchfork}(f; c) \geq N(f; c)$  and equality need not hold. In Theorems 3.12 and 3.13, we compute  $N_{\pitchfork}(f; c)$  in the same setting in which we compute  $N(f; c)$  in Theorems 3.11 and 3.13.

An important reason for calculating the Nielsen root number  $N(f; c)$  of a map  $f$  of  $n$ -manifolds is that it contains geometric information: It is a sharp lower bound if  $n \neq 2$ , that is, there exists a map  $g$  homotopic to  $f$  such that  $g^{-1}(c)$  contains exactly  $N(f; c)$  points. The transverse Nielsen number is also sharp, even if  $n = 2$ , in the sense that there is a map  $g$  homotopic to  $f$  and transverse to  $c$  such that  $g^{-1}(c)$  contains exactly  $N_{\pitchfork}(f; c)$  points. This is equivalent to saying that  $N_{\pitchfork}(f; c)$  can be realized by a map  $g$  which has  $N_{\pitchfork}(f; c)$  as its geometric degree, and so it follows from the fact that  $N_{\pitchfork}(f; c)$

is sharp that the Absolutgrad of a map between  $n$ -manifolds is equal to its geometric degree. This is the property of the Absolutgrad that motivated Hopf's introduction of the concept in [H2]. Hopf was influenced in his 1930 study of the degree by work of Jakob Nielsen [N1, N2] on the subject of fixed points that was published a few years earlier and, in particular, the Nielsen root number (called "wesentliche Schichtenzahl") and the transverse root Nielsen number (called "Absolutgrad") appear for the first time in Hopf's paper.

Thus our paper may be viewed in part as a re-examination of Hopf's degree theory from a present-day mathematical standpoint. Ours is by no means the first updating and extension of Hopf's work, and in particular of Hopf's very novel concept of the Absolutgrad. The first such studies are contained in two important papers, by Olum [O] and Epstein [E]. In 1953, Olum [O] considered maps between closed but not necessarily orientable manifolds and used cohomology with local coefficients to introduce an algebraically defined "group ring degree" in a way which is more closely related to the definition of the classical Brouwer degree (but not to that of the geometric degree) than Hopf's definition of the Absolutgrad. Olum showed that it follows from his definition that Hopf's Absolutgrad equals the absolute value of the group ring degree, and he calculated the group ring degree in terms of a "twisted" global degree, introduced earlier in his paper, and the mod 2 degree [O, p. 478]. In an influential paper [E] that Epstein published in 1966, the calculations of Olum were interpreted in terms of cohomology degrees of lifts of the map  $f$  and these degrees were used by Epstein, and subsequently by other authors, as the definition of the Absolutgrad for maps between not necessarily orientable manifolds (see [E, (1.8) p. 371] and [Sk, Definition p. 416]). In the approach of Olum and Epstein, maps between  $n$ -manifolds are classified into three types and the Absolutgrad is defined separately for each type. This somewhat complicated definition makes the Absolutgrad more readily computable, but it obscures its meaning.

There have been several recent extensions of Hopf's work. In 1986, Lin [L] concentrated on the root theory component of Hopf's work and provided a modern definition of the multiplicity of a root class and a modern proof, with techniques that we also use in this paper, of the sharpness of the root Nielsen number in the special case that  $f$  is a map between closed orientable manifolds of dimension at least 3. But Lin did not consider non-orientable manifolds, nor did he re-establish the connection between Nielsen root theory and degree theory. In 1987, Skora [Sk] provided a modern geometric treatment of the connection between the Absolutgrad and the geometric degree for boundary-preserving maps between surfaces, and thus re-proved and extended results from [H2] which were proved even earlier by Kneser [Kn1, Kn2], but Skora did not connect his results to Nielsen root theory.

An extension of Nielsen root theory to proper maps  $f: M \rightarrow N$  between  $n$ -manifolds, where the point  $c \in N$  is replaced by a connected  $k$ -manifold of dimension  $0 \leq k \leq n$ , was obtained in 1992 by Yongwu Rong and Shicheng Wang [RW], and in the case  $k = 0$  the new and very geometric proof of their main result can be interpreted to show that the transverse root Nielsen number is sharp for proper maps, under the assumption that the manifolds  $M$  and  $N$  are closed and orientable and the homology degree of  $f$  (and hence  $N_{\uparrow}(f; c)$ , see Theorem 3.13 below) is non-zero. Their paper makes reference to Hopf [H2], but not to Hopf's Absolutgrad.

An important difference between the previous reinterpretations of Hopf's theory and this paper is that we re-establish the connection between Hopf's work and the ideas introduced by Nielsen. In particular, our approach is based on the concept of root class that Hopf used as the analogue, in the degree context, of Nielsen's central notion of fixed point class, and our methods are influenced by techniques of modern Nielsen theory. Most of our results are not new. Although the formulae for the two Nielsen root numbers in Theorems 3.11 and 3.12 are not due to Hopf, they can be obtained, by a careful inspection, from Olum [O]. The sharpness of both Nielsen root numbers was first proved by Hopf [H2], and an updated proof of the fact that the Absolutgrad equals the geometric degree in dimension  $\geq 3$ , and hence of the sharpness of the transverse Nielsen root number, was the goal of Epstein's paper [E]. On the other hand, some of our definitions and all our proofs are new and different from existing ones. We use local degree theory to define the integer-valued multiplicity of a root class and define the Nielsen root number to be the number of root classes with non-zero multiplicity. We introduce the transverse Nielsen root number, defined as the sum of the multiplicities of all the root classes. The calculation of  $N(f; c)$  and  $N_{\uparrow}(f; c)$  is obtained by combining results from local degree theory with the definition of root class in terms of a lift of  $f$ , and it also uses the various lifts of  $f$  employed by Epstein [E]. A form of the Whitney Lemma due to Jezierski [Je] and the theory of microbundle transversality are used to establish the sharpness results. We have also included many examples. By these means we obtain not only an extension of Nielsen root theory for maps between  $n$ -manifolds which need no longer be orientable, but also a foundation for Hopf's Absolutgrad and its connection to the geometric degree, which is in the spirit of Hopf but based in large part on developments in algebraic topology and manifold theory since the time of Hopf's work.

Our paper is organized as follows. We define the multiplicity of a root class in Section 2, and in Section 3 we define and compute the two Nielsen root numbers for maps of  $n$ -manifolds. Sharpness of the two Nielsen root numbers for such maps, for  $n \neq 2$ , is established in Section 4. In Section 5, we relate the Nielsen root theory developed in the previous sections to

Hopf's degree theory. Although the geometric results of Section 4 exclude maps of surfaces, much is known and we describe the Nielsen root theory and Hopf degree theory of maps of surfaces in Section 6.

A very readable introduction to Nielsen root theory can be found in Kiang's book [K $\mathbf{g}$ ]. For the background from Nielsen fixed point theory, see in addition [B $\mathbf{n1}$ ] and [J $\mathbf{g2}$ ].

We thank Albrecht Dold for his assistance with Remark 2.5, Ron Stern for help constructing Example 3.16 and Jerzy Jezierski for his useful comments.

## 2. The multiplicity of a root class.

Throughout this paper,  $M$  and  $N$  will be connected topological  $n$ -manifolds with (possibly empty) boundary. The manifolds are not necessarily compact and they can be orientable or not (a manifold with boundary is called *orientable* if its interior is an orientable manifold [D $\mathbf{o}$ , p. 257]). A map  $f: M \rightarrow N$  is *boundary-preserving* if it is a map of pairs  $f: (M, \partial M) \rightarrow (N, \partial N)$ . The map  $f$  is *proper* if  $K \subset N$  is compact implies that  $f^{-1}(K)$  is a compact subset of  $M$ . All homotopies in this paper are understood to be boundary-preserving, and so a *proper homotopy* is a proper map  $f: (M, \partial M) \times I \rightarrow (N, \partial N)$ .

We will be concerned with a proper map  $f: (M, \partial M) \rightarrow (N, \partial N)$  of manifolds of the same dimension and with the set  $f^{-1}(c)$ , for a point  $c \in \text{int } N$ , when that set is non-empty. We choose  $c$  as the basepoint for  $N$  and some  $x_0 \in f^{-1}(c)$  as the basepoint for  $M$ , so  $f$  induces a homomorphism  $f_\pi: \pi_1(M, x_0) \rightarrow \pi_1(N, c)$ .

A proper map  $f: (M, \partial M) \rightarrow (N, \partial N)$  induces a homomorphism  $f^*: \check{H}^n(N, \partial N) \rightarrow \check{H}^n(M, \partial M)$  of Čech cohomology with compact supports and integer coefficients (see [D $\mathbf{o}$ , 6.26, p. 290]). Let  $W = M - f^{-1}(\partial N)$  which is an open subset of  $\text{int } M$ . If  $M$  and  $N$  are oriented manifolds, then there is a class  $[N] \in \check{H}^n(N, \partial N)$  corresponding to the fundamental class  $[\text{int } N] \in \check{H}^n(\text{int } N)$  and a class  $[M] \in \check{H}^n(M, \partial M)$  corresponding to  $[W] \in \check{H}^n(W)$  obtained from a component of  $W$  by restricting the orientation of  $\text{int } M$ . The *cohomological degree* of  $f$ , denoted  $\text{deg}(f)$ , is defined by  $f^*[N] = \text{deg}(f)[M]$ .

All manifolds are orientable with respect to  $\mathbb{Z}/2$  coefficients so, just as in the previous paragraph, we may always obtain classes  $[N] \in \check{H}^n(N, \partial N; \mathbb{Z}/2)$  and  $[M] \in \check{H}^n(M, \partial M; \mathbb{Z}/2)$  and define the *mod 2 cohomological degree*, denoted  $\text{deg}(f, 2)$ , by setting  $f^*[N] = \text{deg}(f, 2)[M]$ .

Maps between not necessarily orientable manifolds are classified in the following manner. A map  $f: M \rightarrow N$  is called *orientation-true* if it maps orientation-preserving loops in  $M$  to orientation-preserving loops in  $N$  and orientation-reversing loops in  $M$  to orientation-reversing loops in  $N$ , otherwise it is called *not orientation-true*. The class of maps that are not orientation-true is subdivided to produce the following classification:

**Definition 2.1.** Let  $f: M \rightarrow N$  be a map of manifolds. Then three *types* of maps are defined as follows.

- (1) *Type I:*  $f$  is orientation-true.
- (2) *Type II:*  $f$  is not orientation-true but does not map an orientation-reversing loop in  $M$  to a contractible loop in  $N$ .
- (3) *Type III:*  $f$  maps an orientation-reversing loop in  $M$  to a contractible loop in  $N$ .

Further, a map  $f$  is defined to be *orientable* if it is of Type I or II, and *non-orientable* otherwise.

The term *orientable map* is sometimes used for Type I, that is orientation-true, maps; see [Do, Exercise 6, p. 271]. The characterisation of the three types is based on Olum [O, p. 475] (see also [Sk, p. 416]). An equivalent characterisation in terms of the orientability of covering spaces of  $M$  and  $N$ , which we will use in §3, is given by Epstein [E, p. 371]. In essence, the characterisation of the three types of maps is already contained in Hopf's paper [H2]. We shall see that maps of the first two types share many properties with regard to Nielsen root theory and, as Hopf was well aware of this, he considered maps of these first two types together and therefore introduced the concept of an orientable or non-orientable map [H2, Definition V, p. 579].

The following examples illustrate the three types of maps.

**Example 2.2** (Type I). (a) If  $M$  and  $N$  are orientable manifolds, then all maps  $f: M \rightarrow N$  are orientation-true. (b) For  $N$  a non-orientable manifold and  $M$  its orientable covering, the covering map  $p: M \rightarrow N$  is of Type I. (The case of  $N$  the projective plane is mentioned by Hopf [H2, p. 584].) (c) The identity map of a non-orientable manifold is an example of a Type I map between non-orientable manifolds.

**Example 2.3** (Type II). Let  $M_+$  be the Möbius band and let  $p: M_+ \rightarrow S^1$  be the fibration obtained by retracting  $M_+$  to its central circle. Let  $i: S^1 \rightarrow S^1 \times I = N_+$  be defined by setting  $i(x) = (x, 0)$ , then  $f_+ = i \circ p$  is a boundary-preserving map from the Möbius band to the annulus. Let  $f = 2f_+: M = 2M_+ \rightarrow 2N_+ = N$  be the double of the map  $f_+$ , so  $M$  is the Klein bottle and  $N$  is the torus. The loops representing elements in the kernel of the induced homomorphism  $f_\pi: \pi_1(M, x_0) \rightarrow \pi_1(N, c)$  are orientation-preserving, so  $f$  is not Type III. Since a map from a non-orientable manifold to an orientable manifold cannot be orientation-true, we conclude that  $f$  is Type II.

**Example 2.4** (Type III). (a) For  $M$  a non-orientable manifold, a constant map  $f: M \rightarrow N$  is obviously of Type III. (b) For an example of a Type III map of  $M$  onto  $N$ , let  $T^2$  denote the torus and  $P^2$  the projective plane. Let  $D$  be a disc in  $T^2$  and let  $id: T^2 - \text{int } D \rightarrow T^2 - \text{int } D$  be the identity. Extend  $id$  to  $f: T^2 \# P^2 \rightarrow T^2$  by extending the identity map on  $\partial D$  in  $P^2 - \text{int } D$

as a map from  $P^2 - \text{int } D$  to  $D \subset T^2$ . Since the generator of  $\pi_1(T^2 \# P^2, x_0)$  represented by an orientation-reversing loop is mapped into the contractible set  $D$ , we see that  $f$  is of Type III.

The remainder of this section will be devoted to defining, for  $f: (M, \partial M) \rightarrow (N, \partial N)$  a proper map of any type, the multiplicity of a root class of  $f$  at  $c \in \text{int } N$ .

Points  $x_1, x_2 \in f^{-1}(c)$  are in the same *root class* of  $f$  at  $c$  if there is a path  $w: I \rightarrow M$  from  $x_1$  to  $x_2$  such that  $f \circ w$  is a contractible loop at  $c$  (see [K $\mathbf{g}$ , Chapter V.B]). (This definition goes back to Hopf [H $\mathbf{2}$ , Definition V, p. 575], where a root class is called a ‘‘Schicht’’.) Since  $f: (M, \partial M) \rightarrow (N, \partial N)$  is proper, the root classes are compact subsets of  $M$  and there are only finitely many of them. Let  $V \subset \text{int } N$  be a contractible neighborhood of  $c$ . Since  $f$  is boundary-preserving,  $f^{-1}(V)$  is contained in  $\text{int } M$ . Let  $R$  be a root class of  $f$  at  $c$  and let  $U$  be an open subset of  $f^{-1}(V)$  such that  $U \cap f^{-1}(c) = R$ . Since  $U$  is an open subset of  $M$ , it is a manifold, that is a space locally homeomorphic to  $\mathbf{R}^n$ , but it is not necessarily connected.

We shall first assume that  $U$  is an oriented manifold. If  $M$  is itself an oriented manifold, then the orientation of  $U$  is the restriction of the orientation of  $M$ . The neighborhood  $V$  is contractible, so it is an orientable manifold and we choose an orientation for it, selecting the restriction of the orientation of  $N$  if that manifold is oriented. The integer-valued *local degree* of  $f|U: U \rightarrow V$  over  $c$  is defined; it is denoted by  $\text{deg}_c(f|U)$  [D $\mathbf{o}$ , Definition 4.2, p. 267]. If  $U_0 \subset U$ , an open subset containing  $R$ , is oriented by restricting the orientation of  $U$ , then  $\text{deg}_c(f|U_0) = \text{deg}_c(f|U)$ . Consequently, if  $U_1$  and  $U_2$  are open subsets of  $f^{-1}(V)$  containing  $R$  that are oriented so that their orientations agree on  $U_1 \cap U_2$ , then  $\text{deg}_c(f|U_1) = \text{deg}_c(f|U_2)$ . Moreover, if  $V_0 \subset V$  is also a contractible neighborhood of  $c$ , and  $U \subset f^{-1}(V_0)$  then, if  $V_0$  is oriented by restricting the orientation of  $V$ , it follows that  $\text{deg}_c(f|U)$  has the same value if we view  $f|U$  as a map into  $V_0$  as it does if we view  $f|U$  as a map into  $V$ .

The following remark describes the relationship between the cohomological degree and the local degree.

**Remark 2.5.** The definition above of the cohomological degree  $\text{deg}(f)$  of a proper map  $f: (M, \partial M) \rightarrow (N, \partial N)$  of oriented manifolds made use of fundamental classes  $[\text{int } N] \in \check{H}^n(\text{int } N)$  and  $[W] \in \check{H}^n(W)$ , where  $W = M - f^{-1}(\partial N)$ . Duality [D $\mathbf{o}$ , Prop. 7.14, p. 297] gives us corresponding elements of singular homology  $\{\text{int } N\} \in H_0(\text{int } N)$  and  $\{W\} \in H_0(W)$  so that, for the homology transfer homomorphism  $f_!$  [D $\mathbf{o}$ , Equation 10.7, p. 310], we have  $f_!\{\text{int } N\} = \text{deg}(f)\{W\}$ . Consequently, for  $f_*: H_0(W) \rightarrow H_0(\text{int } N)$ , we see that  $f_*f_!\{\text{int } N\} = \text{deg}(f)\{\text{int } N\}$ . On the other hand, [D $\mathbf{o}$ , Prop. 10.10, p. 312] implies that  $f_*f_!\{\text{int } N\} = \text{deg}_c(f|W)\{\text{int } N\}$ .

We conclude that if  $f: (M, \partial M) \rightarrow (N, \partial N)$  is a proper map of oriented manifolds, then  $\deg_c(f|W) = \deg(f)$ .

If  $M$  and  $N$  are oriented manifolds, since the orientations of  $V$  and of  $U = U_R$  for each root class  $R$  were chosen to be restrictions of those orientations and  $f^{-1}(c)$  is the union of the root classes, the additivity property of the local degree [Do, Prop. 4.7, p. 269] implies that

$$\sum (\deg_c(f|U_R): R \text{ is a root class of } f) = \deg_c\left(f\left|\bigcup U_R\right.\right).$$

By excision  $\deg_c(f|\bigcup U_R) = \deg_c(f|W)$  and therefore Remark 2.5 implies that

$$\sum (\deg_c(f|U_R): R \text{ is a root class of } f) = \deg(f).$$

We have assumed that  $U$  is orientable because the definition of  $\deg_c(f|U)$  required that the open subset  $U$  of  $f^{-1}(V)$  containing the root class  $R$  be an oriented manifold. Now we no longer assume that  $U$  is orientable. Then there may not be any open subset of  $f^{-1}(V)$  containing  $R$  that is orientable, for instance if  $M$  is a closed non-orientable manifold and  $f$  is the constant map to  $c$ . However, we will now show in our Orientation Procedure (2.6) that if  $f$  is an orientable map, then  $U$  is also orientable, and we will describe an orientation procedure that we will always use to orient  $U$  when  $f$  is an orientable map.

**(2.6) Orientation Procedure.** We first note that if  $f: (M, \partial M) \rightarrow (N, \partial N)$  is a proper orientable map and  $U \subset f^{-1}(V)$  is an open set containing a root class  $R$ , then  $U$  is an orientable manifold. For every loop in  $U$  maps to the contractible space  $V$ , and hence every loop must be orientable since  $f$  is an orientable map. Thus the fundamental group of each component of  $U$  is generated by orientable loops, so each component of  $U$  is an orientable manifold and therefore  $U$  is orientable. We shall always use the following Orientation Procedure to orient  $U$ : If  $M$  is an oriented manifold, we orient  $U$  by restricting the orientation of  $M$ . Otherwise, we choose some  $x_R \in R$  and an orientation of  $U$  at  $x_R$ . Let  $x \in R$  be any point, then there is a path  $w$  in  $M$  from  $x_R$  to  $x$  such that  $f \circ w$  is a contractible loop in  $N$  based at  $c$ . Orient  $U$  at  $x$  by extending the chosen orientation of  $x_R$  along  $w$ . The orientation is independent of the choice of the path  $w$  as follows. Let  $w'$  be another path in  $M$  from  $x_R$  to  $x$  such that  $f \circ w'$  is a contractible loop in  $N$  based at  $c$ . The loop  $w^{-1} \cdot w'$  at  $x$  is orientable because  $f \circ (w^{-1} \cdot w')$  is contractible and the map  $f$  is orientable. Therefore, the orientation of  $U$  at  $x$  is independent of the choice of the path  $w$  as we claimed. Since each component of  $U$  is orientable, in this way the chosen orientation of  $x_R$  determines an orientation of each component of  $U$  that intersects  $R$ . Choosing orientations of the remaining components arbitrarily, we make  $U$  an oriented manifold.

Unless  $M$  and  $N$  are oriented manifolds, there is no criterion for choosing orientations of  $U$  and  $V$ , however,  $\deg_c(f|U)$  is determined up to sign. This can be seen as follows: Suppose that  $U$  is orientable and that we fix the orientation of  $U$  for now. Having chosen an orientation for  $V$ , we have defined  $\deg_c(f|U)$ . Letting  $\deg'_c(f|U)$  denote the degree using the opposite orientation of  $V$ , we have  $\deg_c(f|U) = -\deg'_c(f|U)$ . Now fix the orientation of  $V$ . If  $f$  is an orientable map, choosing the opposite orientation of  $U$  at  $x_R$ , the Orientation Procedure (2.6) changes the orientation of each component of  $U$  that intersects  $R$ . The degree over  $c$  of the restriction of  $f$  to each of these components thus changes sign. Since  $R$  is compact, the number of components of  $U$  that intersect  $R$  is finite. The degree over  $c$  of the restriction of  $f$  to each of the components that fails to intersect  $R$  is zero, so applying the additivity property of the local degree to  $f|U: U \rightarrow V$ , which is a map of oriented manifolds, we see that changing the orientation of  $U$  at  $x_R$  reverses the sign of  $\deg_c(f|U)$  in this case also. If the map  $f$  is not orientable,  $U$  may still be an orientable manifold, for instance if  $R$  is finite and  $U$  is taken to be a union of euclidean neighborhoods. However, if  $f$  is not orientable, the orientation of  $U$  obtained by the Orientation Procedure (2.6) would depend not only on the orientation of  $U$  at  $x_R$  but also on the choice of paths between  $x_R$  and the other points of  $R$ . The facts that, (1) for orientable maps the integer-valued degree of  $f|U$  over  $c$  is defined only up to sign and (2) for maps that are not orientable the integer-valued degree may not be defined at all, motivate the following:

**Definition 2.7.** Let  $f: (M, \partial M) \rightarrow (N, \partial N)$  be a proper map, let  $c \in \text{int } N$  and let  $R$  be a root class of  $f$  at  $c$ . Let  $U$  be any open subset of  $f^{-1}(V)$  containing  $R$  but no other roots of  $f$  at  $c$ , where  $V \subset \text{int } N$  is a contractible neighborhood of  $c$ . Then  $|m(R)|$ , the *multiplicity* of  $R$ , is defined by

$$|m(R)| = |\deg_c(f|U)|,$$

where, for orientable (that is, Type I and II) maps,  $\deg_c(f|U)$  is the local degree with coefficients in  $\mathbb{Z}$  and  $U$  is oriented according to the Orientation Procedure (2.6) and, for non-orientable (that is, Type III) maps,  $\deg_c(f|U)$  is the local degree with coefficients in  $\mathbb{Z}/2$ .

The definition is independent of choices because, up to sign,  $\deg_c(f|U)$  is, as we noted in the paragraph preceding Remark 2.5.

If  $M$  and  $N$  are oriented manifolds without boundary, the multiplicity of Definition 2.7 agrees with the one used by Lin [L, p. 201] (see also [BS2, §3]) up to sign. Lin defined the ‘‘multiplicity’’  $m(R)$  of a root class  $R$  as  $m(R) = \deg_c(f|U)$ , where  $U$  is any open set of  $M$  which contains  $R$  but does not contain any roots of  $f$  that do not lie in  $R$ . In the oriented case it is not necessary to use the absolute value sign in order to obtain a well-defined homotopy invariant. Hopf was aware of the fact that in general an absolute

value has to be used, for his definition of the multiplicity (called “Beitrag”) of a root is essentially the value of  $|m(R)|$  in the following example. (See [H2, Definition VIIa, p. 581] for the definition of the “Beitrag”, and [H1, §5] for Hopf’s definition of the degree as the sum of the degree of  $f$  on euclidean neighborhoods.)

**Example 2.8.** Suppose  $f: (M, \partial M) \rightarrow (N, \partial N)$  is an orientable map such that  $f^{-1}(c)$  is finite. For a root class  $R = \{x_1, x_2, \dots, x_k\}$ , let  $U = \bigcup_{\ell=1}^k U_\ell$ , where the  $U_\ell$  are disjoint euclidean neighborhoods of  $x_\ell$  in  $f^{-1}(V)$  that contain no other roots of  $f$  at  $c$ , and orient  $U$  as in the Orientation Procedure (2.6). The additivity property of the local degree, applied to the map of oriented manifolds  $f|U: U \rightarrow V$ , implies that

$$\deg_c(f|U) = \sum (\deg_c(f|U_\ell): 1 \leq \ell \leq k).$$

We conclude that, for an orientable map  $f$  with  $f^{-1}(c)$  finite, we have the following formula for the multiplicity of a root class

$$|m(R)| = \left| \sum (\deg_c(f|U_\ell): 1 \leq \ell \leq k) \right|.$$

If  $f$  is non-orientable, then the same formula applies, but with coefficients in  $\mathbb{Z}/2$ .

If  $f: M \rightarrow N$  is a Type I map of manifolds without boundary, then the multiplicity  $|m(R)|$  of a root class at  $c$  is the absolute value of the coincidence index  $\text{ind}(c, f; R)$  of Gonçalves and Jezierski [GJ, Definition 5.1, p. 19] for  $c: M \rightarrow N$  the constant map. This can be proved by modifying the argument of Lemma 5.2 of [GJ]. Consequently, by Theorem 5.5 of [GJ],  $|m(R)| = |\text{ind}(c, f; R)|$ , the semi-index [GJ, p. 19] (see also [DJ] and [Je]).

### 3. The two Nielsen root numbers and their computation.

We now use the multiplicity of the previous section to define the Nielsen root number and the transverse Nielsen root number in our setting. The theory of these two numbers is closely linked, and the two numbers were already studied jointly by Hopf [H2]. The first, namely the Nielsen root number, is well-established [Kg, Definition 4.3, p. 129]. We shall see in §5 that the second, the transverse Nielsen root number, is closely related to degree theory, and it is for this reason that the number is introduced.

We define a root class  $R$  of a proper map  $f: (M, \partial M) \rightarrow (N, \partial N)$  at  $c \in \text{int } N$  to be *essential* if its multiplicity  $|m(R)| \neq 0$ , and let the Nielsen root number  $N(f; c)$  be the number of essential root classes of  $f$ . Hence  $N(f; c)$  is a lower bound for the number of roots of  $f$  at  $c$ . We write  $f \sim g$  if  $f$  is (properly) homotopic to  $g$ , and we will show in Corollary 3.14 that  $N(f; c)$  is a proper homotopy invariant. Let  $MR[f; c]$  denote the minimum

number of roots in the proper homotopy class of  $f$ , that is,

$$MR[f; c] = \min\{\#\text{root}(g; c) : g \sim f\}.$$

Corollary 3.14 will thus establish the inequality  $N(f; c) \leq MR[f; c]$ .

We next introduce a minimum number for roots under the condition that the map  $f$  covers a neighborhood of  $c$  geometrically by local homeomorphisms, and thus has no multiple roots (that is, roots with a multiplicity  $\neq \pm 1$ ). More precisely, we will require that  $f$  is transverse to  $c$ , for  $c \in \text{int}N$ , according to the following definition which is used by Epstein [E, p. 375], and is equivalent to the concept of “glatt” used by Hopf [H1, p. 599].

**Definition 3.1.** A proper map  $f: (M, \partial M) \rightarrow (N, \partial N)$  is *transverse to  $c$* , where  $c \in \text{int}N$ , if there exists a euclidean neighborhood  $V$  of  $c$  in  $N$  so that  $f^{-1}(V)$  consists of finitely many euclidean neighborhoods in  $\text{int}M$ , and each of them is mapped by  $f$  homeomorphically onto  $V$ .

For  $f: (M, \partial M) \rightarrow (N, \partial N)$  a proper map, we define the *minimum number of transverse roots* in the proper homotopy class of  $f$  by

$$(3.2) \quad MR_{\uparrow}[f; c] = \min\{\#\text{root}(g; c) : g \sim f \text{ and } g \text{ is transverse to } c\}.$$

We define  $N_{\uparrow}(f; c)$ , the *transverse Nielsen root number* by setting

$$(3.3) \quad N_{\uparrow}(f; c) = \sum (|m(R)| : R \text{ is a root class of } f).$$

Definition (3.3) has the same structure as the transversal Nielsen number that was introduced for the fixed point setting in [Sm]. If  $f$  is transverse to  $c \in \text{int}N$ , then clearly  $f^{-1}(c)$  is finite, and in Example 2.8 each summand  $\deg_c(f|_{U_\ell}) = \pm 1$ . So a root class  $R$  of such a map must contain at least  $|m(R)|$  roots. Thus  $N_{\uparrow}(f; c)$  is a lower bound for the number of roots of  $f$  at  $c \in \text{int}N$  if  $f$  is transverse to  $c$ , and we will show in Corollary 3.14 that it is also a proper homotopy invariant, so  $N_{\uparrow}(f; c) \leq MR_{\uparrow}[f; c]$ . Further, we will see in Corollary 3.14 that both the Nielsen root number  $N(f; c)$  and the transverse Nielsen root number  $N_{\uparrow}(f; c)$  are independent of the choice of  $c \in \text{int}N$ .

We will prove in the next section that if the dimension of the manifolds is different from two, then these Nielsen numbers are sharp, which means that  $N(f; c) = MR[f; c]$  and  $N_{\uparrow}(f; c) = MR_{\uparrow}[f; c]$ . Further, we shall see in Theorem 5.3 that the numbers defined in (3.2) and (3.3) can be interpreted as the geometric degree and the Absolutgrad, respectively, of maps between not necessarily orientable manifolds. Thus it is important to be able to calculate  $N(f; c)$  and  $N_{\uparrow}(f; c)$  and, for this purpose, we next investigate the multiplicity  $|m(R)|$ .

Let  $\widehat{q}: \widehat{N} \rightarrow N$  be the covering space corresponding to  $f_\pi(\pi_1(M, x_0))$  in  $\pi_1(N, c)$ . The space  $\widehat{N}$  is a space of equivalence classes of paths in  $N$  based at  $c$ . Let  $\widehat{f}: (M, \partial M) \rightarrow (\widehat{N}, \partial \widehat{N})$  be the lift of  $f$  that takes  $x_0$  to the class of the constant path at  $c$ . For each  $\widehat{c} \in \widehat{q}^{-1}(c)$ , either  $\widehat{f}^{-1}(\widehat{c})$  is empty or  $R = \widehat{f}^{-1}(\widehat{c})$  is a root class of  $f$  at  $c$  ([Bk2, Lemma 2], see also [H2, Satz III, p. 576]). We will next describe  $|m(R)|$  in terms of the lift  $\widehat{f}$ .

Let  $V \subset \text{int } N$  be a contractible neighborhood of  $c$  that is an elementary neighborhood for the covering space  $\widehat{q}: \widehat{N} \rightarrow N$ . Let  $R = \widehat{f}^{-1}(\widehat{c})$  be a root class and let  $\widehat{V}$  be the component of  $\widehat{q}^{-1}(V)$  containing  $\widehat{c}$ , so  $\widehat{q}|_{\widehat{V}}$  is a homeomorphism onto  $V$ . Since  $f$  is boundary-preserving,  $\widehat{f}$  has the same property and therefore  $\widehat{f}^{-1}(\widehat{V})$  is an open subset of  $\text{int } M$ , moreover,  $\widehat{f}^{-1}(\widehat{V}) \cap f^{-1}(c) = R$ . Let  $U$  be an open subset of  $\widehat{f}^{-1}(\widehat{V})$  containing  $R$ . If  $f$  is an orientable map, and therefore  $U$  can be oriented by the Orientation Procedure 2.6, then  $\text{deg}_{\widehat{c}}(\widehat{f}|U)$ , the integer-valued local degree of  $\widehat{f}|U: U \rightarrow \widehat{V}$  over  $\widehat{c}$  is defined when an orientation is chosen for  $\widehat{V}$ . If the map  $f$  is not orientable, then we still have the local degree  $\text{deg}_{\widehat{c}}(\widehat{f}|U)$  defined if we use  $\mathbb{Z}/2$  coefficients. Choose an orientation of  $V$ . If the homeomorphism  $\widehat{q}|_{\widehat{V}}: \widehat{V} \rightarrow V$  is orientation-preserving, then  $\text{deg}_{\widehat{c}}(\widehat{f}|U) = \text{deg}_c(f|U)$  whereas if  $\widehat{q}|_{\widehat{V}}$  is an orientation-reversing homeomorphism, then  $\text{deg}_{\widehat{c}}(\widehat{f}|U) = -\text{deg}_c(f|U)$ . Thus we have the following alternate description of the multiplicity of a root class.

**Theorem 3.4.** *Let  $f: (M, \partial M) \rightarrow (N, \partial N)$  be a proper map, let  $c \in \text{int } N$  and let  $\widehat{c} \in \widehat{q}^{-1}(c)$  such that  $R = \widehat{f}^{-1}(\widehat{c})$  is non-empty and thus a root class of  $f$  at  $c$ . Let  $U$  be any open subset of  $\widehat{f}^{-1}(\widehat{V})$  containing  $R$ , where  $\widehat{V}$  is the component of  $\widehat{q}^{-1}(V)$  containing  $\widehat{c}$  for  $V \subset \text{int } N$  a contractible elementary neighborhood of  $c$ . Then*

$$|m(R)| = |\text{deg}_{\widehat{c}}(\widehat{f}|U)|,$$

where, for orientable maps,  $\text{deg}_{\widehat{c}}(\widehat{f}|U)$  is the local degree with coefficients in  $\mathbb{Z}$  and  $U$  is oriented according to the Orientation Procedure (2.6) and, for non-orientable maps,  $\text{deg}_{\widehat{c}}(\widehat{f}|U)$  is the local degree with coefficients in  $\mathbb{Z}/2$ .

We now have the tool we need to prove:

**Theorem 3.5.** *Let  $R$  and  $R'$  be root classes of a proper map  $f: (M, \partial M) \rightarrow (N, \partial N)$  at  $c \in \text{int } N$ , then  $|m(R)| = |m(R')|$ .*

*Proof.* Let  $\widehat{c}, \widehat{c}' \in \widehat{N}$  so that  $R = \widehat{f}^{-1}(\widehat{c})$  and  $R' = \widehat{f}^{-1}(\widehat{c}')$ . If  $M$  and  $N$ , and therefore  $\widehat{N}$ , are orientable manifolds, choose orientations for them so that  $\widehat{q}: \widehat{N} \rightarrow N$  is an orientation-preserving map. Since  $\widehat{f}: M \rightarrow \widehat{N}$  is a map of oriented manifolds and  $\widehat{N}$  is connected,  $\text{deg}_{\widehat{c}}(\widehat{f}) = \text{deg}_{\widehat{c}'}(\widehat{f})$  by [Do, Proposition 4.5, p. 268]. Let  $V \subset \text{int } N$  be a contractible elementary neighborhood

of  $c$  and let  $\widehat{V}$  and  $\widehat{V}'$  be the components of  $\widehat{q}^{-1}(V)$  containing  $\widehat{c}$  and  $\widehat{c}'$ , respectively. Let  $U = \widehat{f}^{-1}(\widehat{V})$  and  $U' = \widehat{f}^{-1}(\widehat{V}')$  and orient all of  $U, U', \widehat{V}, \widehat{V}'$  and  $V$  by restricting the orientations of the manifolds in which they lie, then  $\deg_{\widehat{c}}(\widehat{f}|U) = \deg_{\widehat{c}'}(\widehat{f}|U')$  and therefore  $\deg_c(f|U) = \deg_c(f|U')$ , so Definition 2.7 tells us that  $|m(R)| = |m(R')|$ . In particular, since all manifolds are orientable with respect to  $\mathbb{Z}/2$  coefficients, if  $f: (M, \partial M) \rightarrow (N, \partial N)$  is a non-orientable map, then  $|m(R)| = |m(R')|$ .

Thus we now assume that at least one of the manifolds  $M$  and  $N$  is non-orientable and  $f$  is an orientable map, and we define  $\widehat{c}, \widehat{c}' \in \widehat{N}$  as before. Let  $E$  be a euclidean subset in  $\text{int } \widehat{N}$  and choose  $\widehat{b}, \widehat{b}' \in E$ . Using the construction of the homeomorphism in [V, p. 133-134], we obtain a homeomorphism  $h: \widehat{N} \rightarrow \widehat{N}$  such that  $h(\widehat{b}) = \widehat{c}$  and  $h(\widehat{b}') = \widehat{c}'$ . Thus  $S = h(E)$  is an open subset of  $\text{int } \widehat{N}$  homeomorphic to euclidean space such that  $S$  contains both  $\widehat{c}$  and  $\widehat{c}'$ . Since  $f$  is an orientable map, so also is  $\widehat{f}$  and therefore, since  $S$  is simply-connected,  $\widehat{f}^{-1}(S)$  is an orientable submanifold of  $\text{int } M$  (compare the proof in the Orientation Procedure 2.6). We extend the Orientation Procedure 2.6 to orient  $\widehat{f}^{-1}(S)$  in the following manner. If there is no component of  $\widehat{f}^{-1}(S)$  that intersects both  $R$  and  $R'$ , choose any points  $x_R \in R$  and  $x_{R'} \in R'$ , choose orientations at  $x_R$  and  $x_{R'}$ , orient the components of  $\widehat{f}^{-1}(S)$  that intersect  $R$  or that intersect  $R'$  by means of 2.6, and orient the remaining components arbitrarily. Otherwise, let  $C_0$  be a component of  $\widehat{f}^{-1}(S)$  that intersects both  $R$  and  $R'$ , choose an orientation for  $C_0$ , and choose  $x_R$  and  $x_{R'}$  both in  $C_0$ . We orient the components of  $\widehat{f}^{-1}(S)$  that intersect at least one of  $R$  and  $R'$  by means of 2.6 using the orientations at  $x_R$  and  $x_{R'}$  obtained from the orientation of  $C_0$ , and orient the other components arbitrarily. We will show that this procedure is well-defined, that is, if  $C$  is a component of  $\widehat{f}^{-1}(S)$  that intersects both  $R$  and  $R'$ , then it has the same orientation from 2.6 whether we use  $x_R$  or  $x_{R'}$  to orient it. Let  $x \in R \cap C$  and  $x' \in R' \cap C$  and let  $\zeta$  and  $\zeta'$  be paths in  $\text{int } M$  from  $x_R$  to  $x$  and from  $x_{R'}$  to  $x'$ , respectively, such that  $f \circ \zeta$  and  $f \circ \zeta'$  are contractible loops in  $\text{int } N$  at  $c$ . Let  $\alpha$  be a path in  $C_0$  from  $x_R$  to  $x_{R'}$  and let  $\beta$  be a path in  $C$  from  $x$  to  $x'$ . Since the orientations at  $x_R$  and  $x_{R'}$  are determined by the orientation of  $C_0$ , extending the orientation at  $x_R$  to  $x_{R'}$  along  $\alpha$  agrees with the chosen orientation at  $x_{R'}$ . According to 2.6, the orientations at  $x$  and  $x'$  are obtained by extending the orientations at  $x_R$  and  $x_{R'}$  along  $\zeta$  and  $\zeta'$ , respectively. Now suppose that the orientation of  $C$  determined according to 2.6 from the orientation at  $x$  is not the same as the orientation of  $C$  determined according to 2.6 from the orientation at  $x'$ . Then extending the orientation by means of 2.6 at  $x$  along  $\beta$  would not agree with the orientation at  $x'$  obtained by means of 2.6 and therefore the loop  $\lambda$  at  $x'$  defined by  $\lambda = \zeta'^{-1} \cdot \alpha^{-1} \cdot \zeta \cdot \beta$  would be a non-orientable

loop in  $M$ . Since  $f \circ \zeta$  and  $f \circ \zeta'$  are contractible loops at  $c$  in  $N$ , then  $\widehat{f} \circ \zeta$  and  $\widehat{f} \circ \zeta'$  (and consequently  $\widehat{f} \circ \zeta'^{-1}$ ) are contractible loops at  $\widehat{c}$  and  $\widehat{c}'$  respectively. Therefore, the loop  $\widehat{f} \circ \lambda$  is homotopic to  $(\widehat{f} \circ \alpha^{-1}) \cdot (\widehat{f} \circ \beta)$ , a loop in the contractible set  $S$ , so  $\widehat{f} \circ \lambda$  is a contractible loop in  $\widehat{N}$ . Since  $\widehat{f}$  is an orientable map, the loop  $\lambda$  must therefore be orientable and we conclude that if the Orientation Procedure 2.6 is used, then the orientation of  $C$  obtained from the orientation at  $x$  is equal to the orientation of  $C$  obtained from the orientation at  $x'$ , as we claimed.

Now choose an orientation for  $S$  then, for the map of oriented manifolds  $\widehat{f}: \widehat{f}^{-1}(S) \rightarrow S$ , we again have  $\deg_{\widehat{c}}(\widehat{f}) = \deg_{\widehat{c}'}(\widehat{f})$  by [Do, Proposition 4.5, p. 268]. Let  $S_0$  and  $S'_0$  be neighborhoods in  $S$  of  $\widehat{c}$  and  $\widehat{c}'$  respectively such that  $\widehat{q}|_{S_0}$  and  $\widehat{q}|_{S'_0}$  are homeomorphisms onto their images. Let  $V$  be a euclidean elementary neighborhood of  $c$  in  $\widehat{q}(S_0) \cap \widehat{q}(S'_0)$ , let  $\widehat{V}$  and  $\widehat{V}'$  be the components of  $\widehat{q}^{-1}(V)$  containing  $\widehat{c}$  and  $\widehat{c}'$ , respectively, and let  $U = \widehat{f}^{-1}(\widehat{V})$  and  $U' = \widehat{f}^{-1}(\widehat{V}')$ . Note that we have chosen  $V$  so that  $\widehat{V} \cup \widehat{V}' \subset S$  and therefore  $U \cup U' \subset \widehat{f}^{-1}(S)$ . Orienting  $U$  and  $U'$  by restricting the orientation of  $\widehat{f}^{-1}(S)$  we obtained using 2.6 and orienting  $\widehat{V}$  and  $\widehat{V}'$  by restricting the orientation of  $S$ , for the maps  $\widehat{f}|_U: U \rightarrow \widehat{V}$  and  $\widehat{f}|_{U'}: U' \rightarrow \widehat{V}'$  we then have  $\deg_{\widehat{c}}(\widehat{f}|_U) = \deg_{\widehat{c}'}(\widehat{f}|_{U'})$ . By 3.4, we conclude that  $|m(R)| = |m(R')|$ .  $\square$

In view of Theorem 3.5, when we calculate  $|m(R)|$  for a root class  $R$  of a map  $f$  we understand that the calculation is valid for all the root classes of the map. Next we will present, in Lemmas 3.6 and 3.7, two cases in which  $|m(R)|$  is easy to compute.

All manifolds, in particular  $M$  and  $\widehat{N}$ , are orientable with respect to  $\mathbb{Z}/2$  coefficients, so the mod 2 cohomological degree  $\deg(\widehat{f}, 2)$  of the map  $\widehat{f}: (M, \partial M) \rightarrow (\widehat{N}, \partial \widehat{N})$  is defined as in §2. If  $f$  is a proper Type III map, excision together with the argument of Remark 2.5 utilizing  $\mathbb{Z}/2$  coefficients implies that  $\deg_{\widehat{c}}(\widehat{f}|_U) = \deg(\widehat{f}, 2)$ , for  $U \subset W = M - f^{-1}(\partial N)$  as in Theorem 3.4. Consequently we have:

**Lemma 3.6.** *If  $f: (M, \partial M) \rightarrow (N, \partial N)$  is a proper Type III map, then  $|m(R)| = \deg(\widehat{f}, 2) \in \mathbb{Z}/2$ .*

Following Hopf [H2, Definition 2, p. 573], we write  $j$  to denote the cardinality of  $\widehat{q}^{-1}(c)$ , that is, the cardinality of the set of cosets  $\pi_1(N, c)/f_\pi(\pi_1(M, x_0))$ .

**Lemma 3.7.** *If  $f: (M, \partial M) \rightarrow (N, \partial N)$  is a proper map such that  $j$  is infinite then  $|m(R)| = 0$ .*

*Proof.* If  $|m(R)| \neq 0$  for some  $R = \widehat{f}^{-1}(\widehat{c})$  then  $\deg_{\widehat{c}}(\widehat{f}|_U) \neq 0$  by Theorem 3.4 and thus, for any other  $\widehat{c}' \in \widehat{q}^{-1}(c)$ , the proof of Theorem 3.5 shows that  $\deg_{\widehat{c}'}(\widehat{f}|_{U'}) \neq 0$  so  $\widehat{f}^{-1}(\widehat{c}')$  is non-empty. Thus, for every  $\widehat{c} \in \widehat{q}^{-1}(c)$  there

is a root class  $R = \widehat{f}^{-1}(\widehat{c})$ , which is open in  $f^{-1}(c)$ . Since the cardinality of  $\widehat{q}^{-1}(c)$  is  $j$  and  $f$  proper implies that  $f^{-1}(c)$  is compact, there cannot be infinitely many root classes, so  $j$  is finite.  $\square$

As a consequence of Lemmas 3.6 and 3.7 we can now restrict our attention to the remaining cases, in which  $f$  is an orientable map and  $j$  is finite.

For the proof of the next lemma, we will use the following set of covering spaces of  $M$  and  $N$ , and of basepoint-preserving lifts of  $f$  to these covering spaces; compare [E, p. 370].

$$\begin{array}{ccc}
 M', x'_0 & \xrightarrow{f'} & N', c' \\
 p' \downarrow & & \downarrow q' \\
 \widetilde{M}, \widetilde{x}_0 & \xrightarrow{\widetilde{f}} & \widetilde{N}, \widetilde{c} \\
 \widetilde{p} \downarrow & & \downarrow \widetilde{q} \\
 M, x_0 & \xrightarrow{\widehat{f}} & \widehat{N}, \widehat{c}
 \end{array}$$

The spaces and maps in this diagram are obtained in the following manner.

- (1)  $\widetilde{p}: \widetilde{M} \rightarrow M$  is the orientable covering of  $M$ . Hence  $\widetilde{M} = M$  if  $M$  is orientable, but  $\widetilde{M}$  is a 2-sheeted covering of  $M$  if  $M$  is not orientable. This covering space corresponds to the subgroup of  $\pi_1(M, x_0)$  which is generated by the orientation-preserving loop classes of  $M$ .
- (2)  $\widetilde{q}: \widetilde{N} \rightarrow \widehat{N}$  is the minimal covering space of  $\widehat{N}$  with the property that  $\widehat{f} \circ \widetilde{p}: \widetilde{M} \rightarrow \widehat{N}$  has a lift  $\widetilde{f}: \widetilde{M} \rightarrow \widetilde{N}$ . Hence  $\widetilde{N}$  corresponds to the subgroup of  $\pi_1(\widehat{N}, \widehat{c})$  which is generated by the images under  $f$  of all orientation-preserving loop classes of  $M$ , and the homomorphism  $\widetilde{f}_\pi: \pi_1(\widetilde{M}, \widetilde{x}_0) \rightarrow \pi_1(\widetilde{N}, \widetilde{c})$  is onto.
- (3)  $q': N' \rightarrow \widetilde{N}$  is the orientable covering of  $\widetilde{N}$ . Hence  $N'$  corresponds to the subgroup of  $\pi_1(\widetilde{N}, \widetilde{c})$  which is generated by the images under  $f$  of all orientation-preserving loop classes of  $M$  which have an orientation-preserving image in  $N$ .
- (4)  $p': M' \rightarrow \widetilde{M}$  is the minimal cover of  $\widetilde{M}$  with the property that  $\widetilde{f} \circ p': M' \rightarrow \widetilde{N}$  has a lift to  $f': M' \rightarrow N'$ . Hence  $M'$  corresponds to the subgroup of  $\pi_1(\widetilde{M}, \widetilde{x}_0)$  which is generated by all orientation-preserving loop classes of  $M$  which have an orientation-preserving image in  $N$  under  $f$ , and the homomorphism  $f'_\pi: \pi_1(M', x'_0) \rightarrow \pi_1(N', c')$  is onto. Equivalently,  $p': M' \rightarrow \widetilde{M}$  is the pullback of  $q': N' \rightarrow \widetilde{N}$  over  $\widetilde{M}$  by means of  $\widetilde{f}$ .

Note that all the spaces in the diagram are  $n$ -manifolds with (possibly empty) boundary because  $M$  and  $N$  are, and all the maps are proper and boundary-preserving because  $f$  is proper and boundary-preserving. It is easy to see from the fundamental groups to which the covering spaces correspond that each of these covering spaces is either 2-sheeted or the identity.

By construction, not only  $\widetilde{M}$ , but also  $M'$  and  $N'$  are always orientable, and so the integer-valued cohomological degree  $\deg(f')$  of the map  $f' : (M', \partial M') \rightarrow (N', \partial N')$  is always defined.

**Lemma 3.8.** *If  $f : (M, \partial M) \rightarrow (N, \partial N)$  is a proper orientable map and  $j$  is finite, then  $|m(R)| = |\deg(f')|$ .*

*Proof.* The orientation of  $U$  by the Orientation Procedure (2.6) corresponds to a cross-section  $s : U \rightarrow \widetilde{M}$  and we set  $\widetilde{U} = s(U)$ . Define  $\widetilde{R} = s(R)$  and let  $\widetilde{x}_R = s(x_R)$ . Set  $\widetilde{f}(\widetilde{x}_R) = \widetilde{c} \in \widetilde{q}^{-1}(\widehat{c})$ . We may assume that  $\widehat{V}$  in Theorem 3.4 has been chosen so that  $\widehat{V}$  is an elementary neighborhood of the covering space  $\widetilde{p} : \widetilde{N} \rightarrow \widehat{N}$ . Let  $\widetilde{V}$  be the component of  $\widetilde{q}^{-1}(\widehat{V})$  that contains  $\widetilde{c}$ , then the restriction of  $\widetilde{q}$  to  $\widetilde{V}$  is a homeomorphism. We note that  $\widetilde{f}(\widetilde{U}) \subset \widetilde{V}$  because  $\widehat{f}(U) \subset \widehat{V}$ . Letting  $\widetilde{S} = \widetilde{f}^{-1}(\widetilde{c})$ , we claim that  $\widetilde{S} = \widetilde{R}$ .

To verify the claim, we first show that  $\widetilde{R} \subseteq \widetilde{S}$ . We have  $\widetilde{x}_R \in \widetilde{S}$  by definition. Let  $\widetilde{x} \in \widetilde{R}$  be any other point. We obtained  $\widetilde{x}$  by taking a path  $w$  in  $M$  from  $x_R$  to  $x = \widetilde{p}(\widetilde{x})$  in  $M$  such that  $f \circ w$  is a contractible loop in  $N$  and lifting  $w$  to a path  $\widetilde{w}$  in  $\widetilde{M}$  starting at  $\widetilde{x}_R$ . Since  $\widetilde{f} \circ \widetilde{w}$  is a lifting of a contractible loop in  $N$ , it is a contractible loop and we conclude that  $\widetilde{f}(\widetilde{x}) = \widetilde{f}(\widetilde{x}_R) = \widetilde{c}$  so  $\widetilde{x} \in \widetilde{S}$ .

In order to prove that  $\widetilde{S} \subseteq \widetilde{R}$  we will assume that there exists  $\widetilde{x} \in \widetilde{S}$  that is not in  $\widetilde{R}$  and arrive at a contradiction. As  $\widehat{f} \circ \widetilde{p}(\widetilde{x}) = \widetilde{q} \circ \widetilde{f}(\widetilde{x}) = \widehat{c}$ , it follows that  $x = \widetilde{p}(\widetilde{x}) \in R$ . Thus, there is a path  $w$  from  $x_R$  to  $x$  such that  $f \circ w$  is contractible in  $N$ . Consider the path  $w^{-1}$  defined by  $w^{-1}(t) = w(1-t)$  and let  $\widetilde{w}^{-1}$  be the path in  $\widetilde{M}$  obtained by lifting  $w^{-1}$  to a path that starts at  $\widetilde{x}$ , then  $\widetilde{x}'_R = \widetilde{w}^{-1}(1)$  is in  $\widetilde{p}^{-1}(x_R)$  but  $\widetilde{x}'_R \neq \widetilde{x}_R$  because otherwise lifting  $w$  to a path that starts at  $\widetilde{x}'_R = \widetilde{x}_R$  would show that  $\widetilde{x} \in \widetilde{R}$ . However,  $\widetilde{x}'_R \in \widetilde{S}$  because lifting the contractible loop  $f \circ w$  we see that  $\widetilde{f} \circ \widetilde{w}$  is a contractible loop and therefore, in particular,  $\widetilde{f}(\widetilde{x}'_R) = \widetilde{c}$ . As  $\widetilde{x}'_R \in \widetilde{S}$ , there exists a path  $\widetilde{v}$  in  $\widetilde{M}$  from  $\widetilde{x}_R$  to  $\widetilde{x}'_R$  so that  $\widehat{f} \circ \widetilde{p} \circ \widetilde{v}$  is a contractible loop at  $\widehat{c}$  in  $\widehat{N}$ . But the loop  $v = \widetilde{p} \circ \widetilde{v}$  at  $x_R$  in  $M$  does not lift to a loop at  $\widetilde{x}_R$  in  $\widetilde{M}$ , and so it is orientation-reversing. As  $f = \widehat{q} \circ \widehat{f}$  is orientable, it cannot map the orientation-reversing loop  $v$  to the contractible loop  $\widehat{q} \circ \widehat{f} \circ v$ , and so we have a contradiction. Hence  $\widetilde{R} = \widetilde{S}$ .

Since we have homeomorphisms  $\widetilde{p}|\widetilde{U} : (\widetilde{U}, \widetilde{U} - \widetilde{R}) \rightarrow (U, U - R)$  and  $\widetilde{q}|\widetilde{V} : (\widetilde{V}, \widetilde{V} - \widetilde{c}) \rightarrow (\widehat{V}, \widehat{V} - \widehat{c})$  such that  $(\widetilde{f}|U) \circ (\widetilde{p}|\widetilde{U}) = (\widetilde{q}|\widetilde{V}) \circ (\widetilde{f}|\widetilde{U})$ , we conclude, using Theorem 3.4, that  $|m(R)| = |\deg_{\widehat{c}}(\widetilde{f}|U)| = |\deg_{\widetilde{c}}(\widetilde{f}|\widetilde{U})|$ .

We recall that  $p': M' \rightarrow \widetilde{M}$  is the pullback of the orientable cover  $q': N' \rightarrow \widetilde{N}$  over  $\widetilde{M}$  by  $\widetilde{f}$ , so

$$M' = \{(\widetilde{x}, y') \in \widetilde{M} \times N' \mid \widetilde{f}(\widetilde{x}) = q'(y')\}$$

and  $p'$  is projection  $p'(\widetilde{x}, y') = \widetilde{x}$ . Furthermore,  $\widetilde{f}$  can be lifted to  $f' : (M', \partial M') \rightarrow (N', \partial N')$  by setting  $f'(\widetilde{x}, y') = y'$ . Choosing an orientation for the contractible set  $\widetilde{V}$  defines a cross-section on  $\widetilde{V}$  to  $N'$  and we let  $V'$  be the image of the cross-section. Let  $c'_R = q'^{-1}(\widetilde{c}) \cap V'$ . We define an open subset  $U'$  of  $M'$  by letting

$$U' = \{(\widetilde{x}, y') \in M' \mid \widetilde{x} \in \widetilde{U} \text{ and } y' \in V'\}.$$

The restriction of  $p$  to  $U'$  is a homeomorphism  $h$  of  $U'$  onto  $\widetilde{U}$  with inverse  $h^{-1}: \widetilde{U} \rightarrow U'$  given by  $h^{-1}(\widetilde{x}) = (\widetilde{x}, q'^{-1}(\widetilde{f}(\widetilde{x})) \cap V')$ . Let

$$R' = \{(\widetilde{x}, y') \in M' \mid \widetilde{x} \in \widetilde{R} \text{ and } y' \in V'\}.$$

Letting  $S' = f'^{-1}(c')$ , we claim that  $R' = S'$ . If  $(\widetilde{x}, y') \in R'$  then  $\widetilde{x} \in \widetilde{R} = \widetilde{S}$ , that is,  $\widetilde{f}(\widetilde{x}) = \widetilde{c}$ . Therefore  $f'(\widetilde{x}, y') = y' \in q'^{-1}(\widetilde{c})$  but since  $y' \in V'$  it must be that  $f'(\widetilde{x}, y') = c'$  and we see that  $R' \subseteq S'$ . On the other hand,  $(\widetilde{x}, y') \in S'$  means that  $y' = c'$  and since  $(\widetilde{x}, y') \in M'$  we know that  $\widetilde{f}(\widetilde{x}) = q'(y') = \widetilde{c}$ , that is,  $\widetilde{x} \in \widetilde{R}$ , so  $(\widetilde{x}, y') \in R'$  and we have established the claim.

We have homeomorphisms  $h = p'|U': (U', U' - R') \rightarrow (\widetilde{U}, \widetilde{U} - \widetilde{R})$  and  $q'|V': (V', V' - c') \rightarrow (\widetilde{V}, \widetilde{V} - \widetilde{c})$  such that  $(\widetilde{f}|_{\widetilde{U}}) \circ (p'|U') = (q'|V') \circ (f'|U')$  so we conclude that  $|\deg_{\widetilde{c}}(\widetilde{f}|_{\widetilde{U}})| = |\deg_{c'}(f'|U')|$ . On the other hand, since  $M'$  and  $N'$  are orientable manifolds and  $R' = f'^{-1}(c')$ , then the fact that  $\deg_{c'}(f'|U') = \deg(f')$  follows from Remark 2.5, so we have proved that  $|m(R)| = |\deg(f')|$ .  $\square$

If  $f$  is a map of Type I, then  $\widetilde{M} = M'$  and  $\widetilde{N} = N'$ , and hence  $f' = \widetilde{f}$ . Thus Lemma 3.8 tells us that:

**Lemma 3.9.** *Let  $f: (M, \partial M) \rightarrow (N, \partial N)$  be a proper Type I map and let  $j$  be finite, then  $|m(R)| = |\deg(\widetilde{f})|$ .*

The following lemma is due to Epstein [E, Lemma 3.3]. We include the brief proof for the convenience of the reader.

**Lemma 3.10.** *Let  $f: (M, \partial M) \rightarrow (N, \partial N)$  be a proper Type II map, then  $\deg(f') = 0$  and therefore  $|m(R)| = 0$ .*

*Proof.* As  $f$  is not orientation-true, there exists loops  $v$  in  $M$  and  $w = f \circ v$  in  $N$  such that one of  $v$  and  $w$  is an orientation-preserving loop and the other is an orientation-reversing loop. Let  $T_v$  and  $T_w$  be the covering transformations of  $M'$  and  $N'$  that are induced by  $v$  and  $w$  on these covering spaces of  $M$  and  $N$  respectively, then one of these is an orientation-preserving homeomorphism and the other is orientation-reversing. Now  $f' \circ$

$T_v$  and  $T_w \circ f'$  are lifts of  $f$  that agree at one point, so they are equal and  $\deg(f') \deg(T_v) = \deg(T_w) \deg(f')$ . Since  $\deg(T_v) \deg(T_w) = -1$ , we have  $\deg(f') = -\deg(f')$ , and the lemma follows.  $\square$

We have now finished the calculation of the multiplicity  $|m(R)|$  for all cases. To summarize the results of Lemmas 3.6, 3.7, 3.9 and 3.10: If  $j$  is infinite or  $f$  is of Type II then  $|m(R)| = 0$ . If  $f$  is of Type I then  $|m(R)| = |\deg(\tilde{f})|$  whereas if  $f$  is of Type III then  $|m(R)| = \deg(\hat{f}, 2)$ . Since Theorem 3.5 tells us that all root classes of a map have the same multiplicity, the Nielsen root numbers can be calculated from their definitions as follows.

**Theorem 3.11.** *Let  $f: (M, \partial M) \rightarrow (N, \partial N)$  be a proper map between two  $n$ -dimensional manifolds. If  $f$  is a Type I map, then  $N(f; c) = j$  if  $\deg(\tilde{f}) \neq 0$  and  $j$  is finite. If  $f$  is a Type III map, then  $N(f; c) = j$  if  $\deg(\hat{f}, 2) = 1$  and  $j$  is finite. In all other cases,  $N(f; c) = 0$ .*

**Theorem 3.12.** *Let  $f: (M, \partial M) \rightarrow (N, \partial N)$  be a proper map between two  $n$ -dimensional manifolds. If  $f$  is a Type I map, then  $N_{\uparrow}(f; c) = j \cdot |\deg(\tilde{f})|$  if  $j$  is finite. If  $f$  is a Type III map, then  $N_{\uparrow}(f; c) = j \cdot \deg(\hat{f}, 2)$  if  $j$  is finite. In all other cases  $N_{\uparrow}(f; c) = 0$ .*

If both  $M$  and  $N$  are orientable, then all maps are of Type I and the results of Theorems 3.11 and 3.12 can be simplified by using  $\deg(f)$  rather than  $\deg(\tilde{f})$ . For compact manifolds  $M$  and  $N$ , the formula for  $N(f; c)$  given in the next theorem can also be obtained from [BS2, Theorems 3.1, 3.4, 3.12 and 4.8]. For closed manifolds, this theorem can be found in [L, Proposition 5].

**Theorem 3.13.** *Let  $f: (M, \partial M) \rightarrow (N, \partial N)$  be a proper map between two  $n$ -dimensional orientable manifolds. Then*

$$N(f; c) = \begin{cases} j, & \text{if } \deg(f) \neq 0 \\ 0, & \text{if } \deg(f) = 0, \end{cases}$$

and  $N_{\uparrow}(f; c) = |\deg(f)|$ .

*Proof.* We first assume that  $j$  is finite. It is easy to see that then  $\hat{q}$  is proper, and so we can use [Do, Propositions 4.5, p. 268 and 4.7, p. 269] to show that  $\deg(\hat{q}) = j > 0$ . As  $M$  and  $N$  are orientable, we have  $\tilde{f} = \hat{f}$ , and therefore  $|\deg(f)| = |\deg(\hat{q})| \cdot |\deg(\hat{f})| = j \cdot |\deg(\hat{f})|$ , and thus Theorem 3.13 follows from Theorems 3.11 and 3.12.

Now we assume that  $j$  is infinite. We claim that in this case  $\deg(f) = 0$ , and that therefore Theorem 3.13 follows again from Theorems 3.11 and 3.12. To verify our claim, we use Lemma 3.7 which states that for infinite  $j$  all root

classes have multiplicity  $|m(R)| = 0$ , and so  $\deg_c(f|U) = 0$  also. Therefore  $\deg(f) = \sum(\deg_c(f|U_R): R \text{ is a root class of } f) = 0$  as claimed.  $\square$

Maps that are homotopic by a proper homotopy induce the same homomorphism of Čech cohomology with compact supports [Do, p. 290], so the cohomological degree is a proper homotopy invariant. Therefore, Theorems 3.11 and 3.12 imply

**Corollary 3.14.** *Let  $f: (M, \partial M) \rightarrow (N, \partial N)$  be a proper map, then  $N(f; c)$  and  $N_{\uparrow}(f; c)$  are proper homotopy invariants. Moreover, the values of both  $N(f; c)$  and  $N_{\uparrow}(f; c)$  are independent of the choice of  $c \in \text{int } N$ .*

We conclude this section with some examples of maps with non-zero Nielsen root numbers.

**Example 3.15.** Let  $f: S^n \rightarrow S^n$  be a map of degree  $d \neq 0$  between two  $n$ -spheres, where  $n \geq 2$ . As  $j = 1$ , Theorem 3.13 shows that  $N(f; c) = 1$  and  $N_{\uparrow}(f; c) = |d|$ , and so  $N_{\uparrow}(f; c) \not\cong N(f; c)$  if  $|d| > 1$ . More generally, it follows from Theorem 3.13 that  $N_{\uparrow}(f; c) = |d|$  is strictly greater than  $N(f; c) = 1$  for any map  $f: (M, \partial M) \rightarrow (N, \partial N)$  between two orientable  $n$ -manifolds if  $f_{\pi}$  is an epimorphism and  $|\deg(f)| = |d| > 1$ . Examples with  $N_{\uparrow}(f; c) \neq N(f; c)$  for maps of non-orientable manifolds can readily be constructed by using cartesian products. To be specific, let  $f: P^2 \times S^2 \rightarrow P^2 \times S^2$  be the product map  $f = f_1 \times f_2$  with  $f_1$  the identity and  $f_2$  of degree  $d$  with  $|d| > 1$ . Then  $f$  is of Type I,  $j = 1$ ,  $f = \hat{f}$ ,  $\tilde{M} = \tilde{N} = S^2 \times S^2$  and  $\tilde{f} = \tilde{f}_1 \times \tilde{f}_2$  where  $\tilde{f}_1$  is the identity and  $\tilde{f}_2 = f_2$ . Now  $|\deg(\tilde{f})| = |d| \neq 0$  so  $N(f; c) = 1$  by Theorem 3.11 whereas Theorem 3.12 implies that  $N_{\uparrow}(f; c) = |d| > 1$ .

**Example 3.16.** A different type of example of a map  $f$  of non-orientable manifolds for which  $N_{\uparrow}(f; c) \neq N(f; c)$  is illustrated by the following. Represent the Klein bottle by  $K = P^2 \# S^2 \# P^2$ , then a rotation of  $S^2$  interchanging the copies of  $P^2$  defines an action of  $\mathbb{Z}/2$  on  $K$  with two fixed points. Therefore, the homomorphism of fundamental groups induced by the quotient map  $f: K \rightarrow P^2$  is onto [Bd, Cor. 6.3, p. 91]. Thus  $\hat{P}^2 = P^2$  and  $\hat{f}$  is the lift  $\hat{f}: T^2 \rightarrow S^2$  of  $f$  to the oriented covers. By inspection,  $N_{\uparrow}(\hat{f}; \tilde{c}) = 2$ , and so we obtain from Theorem 3.13 that  $\hat{f}$  is a map of degree  $\pm 2$ . Since the map  $f$  is of Type I, Theorems 3.11 and 3.12 imply that  $N(f; c) = 1$  and  $N_{\uparrow}(f; c) = 2$ .

**Example 3.17.** As in Example 2.2(b), let  $f: M \rightarrow N$  be the covering map of the orientable cover of a non-orientable manifold  $N$ . Since  $j = 2$  and  $\deg(\hat{f}) = 1$ , Theorem 3.11 implies that  $N(f; c) = 2$  and Theorem 3.12 implies that  $N_{\uparrow}(f; c) = 2$  also.

**Example 3.18.** Let  $f: K \rightarrow K$  be a map of the Klein bottle then, by [BO, Prop. 6.4], there are integers  $b, d$  and  $e$  such that  $f_\pi(\alpha) = \alpha^b \beta^d$  and  $f_\pi(\beta) = \beta^e$  for appropriately chosen generators  $\alpha$  and  $\beta$  of  $\pi_1(K, x_0)$ . If  $e \neq 0$  then  $b$  is odd and  $N(f; c) = |be| \neq 0$ . Thus Theorem 3.11 tells us that  $j = |be|$  and that  $f$  is not of Type II. From the proof of [BO, Prop. 6.4] we learn that  $f_\pi$  is a monomorphism in this case, so  $f$  cannot be of Type III and we conclude that  $f$  is of Type I. We claim that  $|\deg(\tilde{f})| = 1$  and therefore, by Theorem 3.12, that  $N_{\uparrow\uparrow}(f; c) = |be|$ . The map  $f$  is Type I and  $M = K$  is non-orientable so  $\widehat{N}$  is non-orientable (see [E, page 371]) and it follows that the fiber of the covering space  $q = \tilde{q} \circ \hat{q}: \widehat{N} \rightarrow N$  has cardinality  $2j = 2|be|$ . Let  $r: N^o \rightarrow N = K$  be the oriented cover and let  $q^o: (\tilde{N}, \tilde{c}) \rightarrow (N^o, c^o)$  be the covering space such that  $r \circ q^o = q$ . Since  $q^o$  is a covering map between closed oriented manifolds and the cardinality of the fiber is  $|be|$ , we conclude that  $|\deg(q^o)| = |be|$ . The Type I map  $f$  lifts to a map  $f^o: (\tilde{M}, \tilde{x}_0) \rightarrow (N^o, c^o)$  of the oriented covers and  $f^o = q^o \circ \tilde{f}$  by uniqueness of lifts. We see from the proof of Proposition 6.4 of [BO] that  $\deg(f^o) = be$ , so  $|\deg(\tilde{f})| = 1$  as we claimed.

**Example 3.19.** Let  $A$  be a  $2 \times 2$  integer matrix with determinant  $d$  odd and let  $f': T^2 \rightarrow T^2$  be the corresponding map of the torus. Let  $f'': T^2 \# P^2 \rightarrow T^2$  be the map of Example 2.4(b) and define  $f = f' \circ f'': T^2 \# P^2 \rightarrow T^2$ . The map  $f$  is of Type III because  $f''$  is. Since  $d$  is odd,  $\deg(f', 2) = 1$ . Now  $\deg(f'', 2) = 1$  also, so  $\deg(f, 2) = 1$  and therefore  $\deg(\tilde{f}, 2) = 1$ . Noting that  $f_\pi(\pi_1(T^2 \# P^2, x_0)) = f'_\pi(f''_\pi(\pi_1(T^2 \# P^2, x_0))) = f'_\pi(\pi_1(T^2, f''(x_0)))$ , we have  $j = |d|$ . Therefore, Theorems 3.11 and 3.12 tell us that  $N(f; c) = N_{\uparrow\uparrow}(f; c) = |d|$ .

#### 4. Sharpness.

In this section, we will prove the sharpness of the two Nielsen root numbers defined in §3, that is, that  $N(f; c) = MR[f; c]$  and  $N_{\uparrow\uparrow}(f; c) = MR_{\uparrow\uparrow}[f; c]$ , using results from [Je]. As in [Je], we will work in the general setting of topological manifolds, and hence we use the microbundle transversality from [KS, Essay III, p. 84]. Recall that a *microbundle*  $\xi = \xi^n$  over a space  $X$  consists of a total space  $E(\xi) \supset X$  together with a retraction  $r: E(\xi) \rightarrow X$  that is a submersion near  $X$ , and which has the property that, for all  $x \in X$ , the fibres  $\xi_x = r^{-1}(x)$  are  $n$ -manifolds without boundary. Now consider a pair  $(Y, P)$  of topological spaces, where  $P$  is closed in  $Y$  and is equipped with a normal microbundle  $\eta$ , which means that the total space of  $\eta$  is an open subset of  $Y$  containing  $P$ . If  $X$  is a topological manifold without boundary, then a map  $h: X \rightarrow Y$  is called *topologically transverse* to  $\eta$  if  $h^{-1}(P)$  is a topological submanifold of  $X$  admitting a normal microbundle  $\xi$  such that,

for every  $x \in h^{-1}(P)$ , a neighborhood of  $x$  in the fibre  $\xi_x$  is mapped by  $h$  homeomorphically onto a neighborhood of  $h(x)$  in the fibre  $\eta_{h(x)}$ . See [KS, p. 84] and [Je, p. 167].

In order to apply microbundle transversality to a proper map  $f: (M, \partial M) \rightarrow (N, \partial N)$  between two topological  $n$ -manifolds with (possibly empty) boundaries, we will have to assume that  $f$  has the additional property that  $f(\text{int } M) \subset \text{int } N$ . We write  $f^o$  for the restriction of  $f$  to  $\text{int } M$ , select  $c \in \text{int } N$ , and we will in the remainder of this section choose  $X = \text{int } M$ ,  $Y = \text{int } M \times \text{int } N$ ,  $P = \text{int } M \times \{c\}$ , and  $\eta$  as the normal microbundle of  $P$  in  $Y$  which has  $E(\eta) = Y$  as its total space and the projection  $r: Y \rightarrow P$  given by  $r(x, y) = (x, c)$  as its retraction. Let  $h: \text{int } M \rightarrow \text{int } M \times \text{int } N$  be defined by  $h = (e, f^o)$ , where  $e: \text{int } M \rightarrow \text{int } M$  is any map, then clearly  $h^{-1}(P) = \text{root}(f; c)$ . According to Definition 3.1, a proper map  $f: (M, \partial M) \rightarrow (N, \partial N)$  is transverse to  $c$ , where  $c \in \text{int } N$ , if there exists a euclidean neighborhood  $V$  of  $c$  in  $N$  so that  $f^{-1}(V)$  consists of finitely many euclidean neighborhoods in  $\text{int } M$ , and each of them is mapped by  $f$  homeomorphically onto  $V$ . Transversality of  $h$  to  $\eta$  and transversality of  $f$  to  $c$  are intimately related: A proper map of the form  $f: (M, \partial M, \text{int } M) \rightarrow (N, \partial N, \text{int } N)$  is transverse to  $c \in \text{int } N$  if and only if there exists a map  $e: \text{int } M \rightarrow \text{int } M$  so that the map  $h = (e, f^o): \text{int } M \rightarrow \text{int } M \times \text{int } N$  is topologically transverse to  $\eta$ . To see that this is true, note that if  $h = (e, f^o)$  is topologically transverse to  $\eta$ , then the map  $e$  must be a map that is constant on a neighborhood of each of the points in  $h^{-1}(P)$ . Conversely, given a proper map  $f$  transverse to  $c$ , such a map  $e$  can always be found in order to construct a map  $h$  which is topologically transverse to  $\eta$ .

The most important step in the proof of the sharpness of Nielsen numbers consists in uniting two points in the same Nielsen class whenever possible, and for this step we want to apply results concerning Nielsen classes from [Je] to a map  $h = (e, f^o)$ . This is possible as, according to the definition of the Nielsen relation in [Je, p. 168], two points of  $h^{-1}(P)$  are in such a relation if and only if they belong to the same root class of  $f$ . We will use a Whitney type lemma to unite roots, and for this we need local orientations. By a *local orientation*  $O(x)$  of  $M$  at  $x \in \text{int } M$  we mean a generator  $O(x) \in H_n(M, M - x)$  and by a *local orientation* of  $N$  at  $y \in \text{int } N$  we mean a generator  $O(y) \in H_n(N, N - y)$ , where the homology groups have coefficients in  $\mathbb{Z}$  (see [Do, Definition 2.1, p. 252]).

If  $f: (M, \partial M) \rightarrow (N, \partial N)$  is a map which is transverse to  $c \in \text{int } N$  and if  $x$  is a root of  $f$ , then  $f$  defines, by restriction, a homeomorphism of a euclidean neighborhood  $U_x$  of  $x$  in  $\text{int } M$  onto a euclidean neighborhood  $V$  of  $c$  in  $\text{int } N$ . Using the excision isomorphisms  $H_n(M, M - x) \cong H_n(U_x, U_x - x)$  and  $H_n(N, N - c) \cong H_n(V, V - c)$  we see that, given a (local) orientation  $O(x)$  of  $M$  at  $x$ , the restriction of  $f$  to  $U_x$  defines a corresponding orientation of  $N$  at  $c$  which, by abuse of notation, we denote by  $f_*(O(x))$ . As in [Je, p. 168]

we define a local orientation of the microbundle  $\eta$  at  $(x, c) \in \text{int } M \times \text{int } N$  to be a generator of  $H_n(x \times N, x \times (N - c))$ . Then it is straightforward to check that the  $\mathcal{R}$ -relation of [Je, Definition (1.2)], which characterises points of  $h^{-1}(P)$  that can be removed by a homotopy, takes in our case the following form: Let  $f: (M, \partial M, \text{int } M) \rightarrow (N, \partial N, \text{int } N)$  be a proper map that is transverse to  $c$ , where  $c \in \text{int } N$ , and let  $x_1, x_2 \in \text{root}(f; c)$ . Then  $x_1, x_2$  are  $\mathcal{R}$ -related with respect to a map  $h = (e, f^\circ): \text{int } M \rightarrow \text{int } M \times \text{int } N$  which is transverse to  $\eta$  if and only if there exists a path  $w: I \rightarrow \text{int } M$  from  $x_1$  to  $x_2$  such that  $f \circ w$  is a contractible loop at  $c$  and  $f_*(O(x_1)) = -f_*(O(x_2))$ , where  $O(x_1)$  is a local orientation of  $M$  at  $x_1$  and  $O(x_2)$  is the local orientation of  $M$  at  $x_2$  that is obtained from  $O(x_1)$  by continuation along  $w$ . Note that this characterisation of the  $\mathcal{R}$ -relation is independent of the choice of the map  $e$  and the orientation  $O(x_1)$ .

The following lemma gives information about the number of points in  $\mathcal{R}$ -relation to one another. Its proof helps to explain why the different behavior of orientable and non-orientable maps leads to different minimal root sets. The technique in the proof of “flipping” the local orientation when dealing with non-orientable maps was used by Hopf in the elimination of inessential root classes in [H2, p. 601].

**Lemma 4.1.** *Let  $f: (M, \partial N, \text{int } M) \rightarrow (N, \partial N, \text{int } N)$  be a proper map between two  $n$ -dimensional manifolds and let  $c \in \text{int } N$ . Let  $f$  be transverse to  $c$ , and let  $R = \{x_1, \dots, x_k\}$  be a root class of  $f$  at  $c$ . If  $k > |m(R)|$ , then there exist two points in  $R$  which are  $\mathcal{R}$ -related with respect to any map  $h: \text{int } M \rightarrow \text{int } M \times \text{int } N$  given by  $h = (e, f^\circ)$  that is transverse to  $\eta$ .*

*Proof.* We first assume that  $f$  is an orientable map. Since  $f^{-1}(c)$  is finite, as in Example 2.8 we may choose  $U$  to be the disjoint union of euclidean neighborhoods  $U_\ell$ , each containing one point  $x_\ell \in R$ . Furthermore, since  $f$  is transverse to  $c$ , we may choose the  $U_\ell$  so that each restriction  $f|U_\ell$  is a homeomorphism onto the euclidean neighborhood  $V$  of  $c$ . If we orient  $U$  as in the Orientation Procedure 2.6, then we see from Example 2.8 that

$$|m(R)| = \left| \sum (\deg_c(f|U_\ell): 1 \leq \ell \leq k) \right|,$$

where the local degree is defined with integer coefficients. Since each  $f|U_\ell$  is a homeomorphism, the corresponding summand equals  $\pm 1$  and therefore  $k > |m(R)|$  implies that not all local degrees can be equal. So, without loss of generality, we assume that  $\deg_c(f|U_1) = -\deg_c(f|U_2)$ , where according to the Orientation Procedure 2.6 the orientation of  $U_2$  is obtained from the orientation of  $U_1$  by continuation along a path  $w$  from  $x_1$  to  $x_2$  in  $\text{int } M$  that is chosen so that  $f \circ w$  is a contractible loop in  $N$  at  $c$ . The orientations of  $U_1$  and  $U_2$  define local orientations  $O(x_1)$  and  $O(x_2)$  of  $M$  at  $x_1$  and  $x_2$  and, according to our selection of orientations for  $U_1$  and  $U_2$ , the local orientation  $O(x_2)$  is obtained from  $O(x_1)$  by continuation along  $w$ . From

$\deg_c(f|U_1) = -\deg_c(f|U_2)$  it follows that  $f_*(O(x_1)) = -f_*(O(x_2))$ , and therefore  $x_1$  and  $x_2$  are  $\mathcal{R}$ -related with respect to a map  $h = (e, f^o)$  that is transverse at  $\eta$ .

Now let  $f$  be non-orientable, that is, of Type III. We can choose the open set  $U$  as before and obtain, by using Example 2.8 with coefficients in  $\mathbb{Z}/2$ ,

$$|m(R)| = \begin{cases} 1, & \text{if } k \text{ is odd} \\ 0, & \text{if } k \text{ is even.} \end{cases}$$

We first assume that  $k > |m(R)| = 1$ , and so there exist at least two roots in  $R$ , say  $x_1$  and  $x_2$ , and a path  $w$  in  $\text{int } M$  from  $x_1$  to  $x_2$  so that  $f \circ w$  is a contractible loop at  $c$  in  $N$ . If the continuation of the local orientation  $O(x_1)$  from  $x_1$  to  $x_2$  along  $w$  leads to a local orientation  $O(x_2)$  so that  $f_*(O(x_1)) = -f_*(O(x_2))$ , then it follows as before that  $x_1$  and  $x_2$  are  $\mathcal{R}$ -related. Otherwise, we select an orientation reversing loop  $\ell: I \rightarrow \text{int } M$  so that  $f \circ \ell$  is a contractible loop in  $N$ , and a path  $v: I \rightarrow \text{int } M$  from  $x_1$  to  $\ell(0)$ . Then  $w' = v \cdot \ell \cdot v^{-1} \cdot w$  is a path from  $x_1$  to  $x_2$  so that its image under  $f$  in  $N$  is a contractible loop at  $c$ . As the continuation of a local orientation along  $w'$  gives the opposite value as the continuation of this local orientation along the path  $w$ , it follows once again that  $x_1$  and  $x_2$  are  $\mathcal{R}$ -related with respect to  $h$ . This completes our proof if  $k > |m(R)| = 1$ . If  $|m(R)| = 0$  then  $k$  is even, and so  $k > |m(R)|$  implies  $k \geq 2$  and two  $\mathcal{R}$ -related roots can be found in the same way.  $\square$

In the next theorem we prove the sharpness of the transverse Nielsen root number with respect to proper homotopies. This result is essentially due to Hopf (see Theorem 5.4).

**Theorem 4.2.** *Let  $f: (M, \partial M) \rightarrow (N, \partial N)$  be a proper map between two  $n$ -dimensional manifolds and let  $c \in \text{int } N$ . If  $n \neq 2$ , then there exists a map  $g: (M, \partial M) \rightarrow (N, \partial N)$  which is homotopic to  $f$  by a proper homotopy, is transverse to  $c$ , and has precisely  $N_{\uparrow}(f; c)$  roots at  $c$ . Hence  $N_{\uparrow}(f; c)$  is sharp, that is  $N_{\uparrow}(f; c) = MR_{\uparrow}[f; c]$ .*

*Proof.* The theorem is obviously true if  $n = 1$ , and so we assume that  $n \geq 3$ . As  $c \in \text{int } N$ , all roots of  $f$  lie in  $\text{int } M$ . Using the construction in [BS1], there is a map  $f^+: (M, \partial M, \text{int } M) \rightarrow (N, \partial N, \text{int } N)$  that has the same root set as  $f$ , and it is easy to see from the construction that there exists a proper homotopy from  $f$  to  $f^+$  relative to the complement of the interior of a collar of  $\partial M$ .

As before, we choose  $X = \text{int } M$ ,  $Y = \text{int } M \times \text{int } N$ ,  $P = \text{int } M \times \{c\}$ , and let  $\eta$  be the microbundle of  $P$  in  $W$  given by the projection  $r(x, y) = (x, c)$ . If we define the map  $h: \text{int } M \rightarrow \text{int } M \times \text{int } N$  by  $h(x) = (x, f^+(x))$ , we can use the microbundle transversality of [KS] (see also [Je, Lemma 1.1]) to deform  $h$  on a compact neighborhood of root  $(f; c)$  in  $\text{int } M$ , and relative

to the boundary of this neighborhood, and then extend it over  $\text{int } M$  as the identity, to obtain a map  $h': \text{int } M \rightarrow \text{int } M \times \text{int } N$  that is transverse to  $\eta$ . The map  $h' = (h'_1, h'_2)$  defines a map  $f': (M, \partial M, \text{int } M) \rightarrow (N, \partial N, \text{int } N)$  by setting  $f'(c) = h'_2(x)$  for all  $x \in \text{int } M$  and  $f'(x) = f(x)$  for all  $x \in \partial M$ , and we have  $h'^{-1}(P) = \text{root}(f'; c)$ . As the map  $h'$  is transverse to  $\eta$ , the map  $f'$  is transverse to  $c$ .

Now let  $R$  be any root class of  $f'$  which consists of  $k > |m(R)|$  points. Then Lemma 4.1 implies that at least one pair of these roots is  $\mathcal{R}$ -related with respect to the map  $h'$ . Let this pair be  $x_1, x_2$ . As we can choose the path  $w$  which establishes the  $\mathcal{R}$ -relation to be a flat arc in  $\text{int } M$ , there exists an  $n$ -dimensional ball  $D \subset \text{int } M$  that contains  $w$  but no root of  $f'$  other than  $x_1, x_2$ . Therefore we can use the Whitney Type Lemma 3.1 from [Je] to obtain a homotopy of  $h'|_D$ , relative to the boundary of  $D$ , to a map from  $D$  into  $(\text{int } M \times \text{int } N) - P$ . We keep  $h'$  fixed on the complement of  $D$ .

We repeat this process as many times as possible to remove pairs of points from  $h'^{-1}(P)$  that belong to the same root class of  $f'$ , and so construct a map  $h'' = (h''_1, h''_2)$ . The map  $h''_2: \text{int } M \rightarrow \text{int } N$  extends to a map  $g: (M, \partial M) \rightarrow (N, \partial N)$  by setting  $g(x) = h''_2(x)$  for all  $x \in \text{int } M$ , and  $g(x) = f(x)$  for all  $x \in \partial M$ . It follows from Lemma 4.1 that each root class of  $g$  consists of exactly  $|m(R)|$  points. Hence  $g$  has exactly  $N_{\cap}(f; c)$  roots at  $c$  and, as  $h''$  is transverse to  $\eta$ , the map  $g$  is transverse to  $c$ . Each modification of  $h$  has been made on a compact subset of  $\text{int } M$ , so the sequence of the constructed homotopies defines a homotopy from  $f$  to  $g$  that is boundary-preserving and has a compact carrier, that is, it is constant outside a compact set. Therefore this homotopy from  $f$  to  $g$  is proper.  $\square$

With regard to  $N(f; c)$ , sharpness was first proved (under slightly different assumptions than those used here) in [H2, Satz XIIIb]. Techniques are now available that allow us to give a short proof of this fact, as follows.

**Theorem 4.3.** *Let  $f: (M, \partial N) \rightarrow (N, \partial N)$  be a proper map between two  $n$ -dimensional manifolds and let  $c \in \text{int } N$ . If  $n \neq 2$ , then there exists a map  $g: (M, \partial M) \rightarrow (N, \partial N)$  that is homotopic to  $f$  by a proper homotopy and has precisely  $N(f; c)$  roots at  $c$ . Hence  $N(f; c)$  is sharp, that is  $N(f; c) = MR[f; c]$ .*

*Proof.* Again we can assume that  $n \geq 3$ . We proceed as in the proof of Theorem 4.2 to obtain a proper boundary-preserving map which has  $N_{\cap}(f; c)$  roots at  $c$  and is transverse to  $c$ . The proof is completed by using the Creating and Cancelling Procedures which were developed to minimize the number of fixed points on differentiable manifolds by [Jg1] and used to minimize the number of roots on orientable PL manifolds by [L, §3, proof of Theorem B]. As these procedures consist of local changes that occur in  $\text{int } M$ , and

use neither a global orientation nor a PL or differentiable structure on the manifolds, they can be applied without change.  $\square$

Here are some examples to illustrate Theorems 4.2 and 4.3.

**Example 4.4.** Let  $f: S^n \rightarrow S^n$  be a map between two  $n$ -spheres, let  $c$  be any point of  $S^n$ , and let  $|\deg(f)| = |d| > 1$ . If  $n \geq 3$ , then we see from Example 3.15 and Theorem 4.3 that  $f$  can be homotoped to a map  $g$  with only one root at  $c$ , and it is easy to see directly that such a map  $g$  exists even if  $n = 2$ . But  $N_{\uparrow}(f; c) = |d| > 1$ , and so the map  $g$  cannot be transverse to  $c$  as any such map must have at least  $|d|$  roots at  $c$ . Theorem 4.2 and Theorem 6.1 below show that there exists in fact a map that is homotopic to  $f$ , is transverse to  $c$  and has  $|d|$  roots at  $c$ . Similarly,  $f: P^2 \times S^2 \rightarrow P^2 \times S^2$  from Example 3.15 is a map of 4-manifolds with  $N(f; c) = 1$  and  $N_{\uparrow}(f; c) = |d| > 1$  so it has the property, by Theorem 4.3, that there is a map  $g$  homotopic to  $f$  with just one root at  $c$  whereas any map homotopic to  $f$  and transverse to  $c$  must have at least  $|d|$  roots at  $c$ . Theorem 4.2 tells us that there is a map  $g$  homotopic to  $f$  and transverse to  $c$  with exactly  $|d|$  roots at  $c$ .

**Example 4.5.** Let  $f = f' \times f'': M = K \times S^1 \rightarrow K \times S^1 = N$  where  $f': K \rightarrow K$  is the map of Example 3.18 and  $f'': S^1 \rightarrow S^1$  is a map of degree  $r \neq 0$ . Since  $K \times S^1$  is a closed aspherical manifold and  $f_\pi: \pi_1(M, x_0) \rightarrow \pi_1(N, c)$  is a monomorphism then, as in the proof of [BO, Prop. 6.4], the Nielsen number  $N(f; c)$  is the absolute value of the degree of a lift  $f^o$  of  $f$  to  $T \times S^1$ , the orientable covering of  $K \times S^1$ . In the notation of Example 3.18, that degree equals  $ber$  in this case, and so  $N(f; c) = |ber|$ . Thus, Theorem 3.11 implies that  $f$  is of Type I and, as in Example 3.18, we can see that  $|\deg(\hat{f})| = 1$ . Therefore Theorem 3.12 shows that  $N_{\uparrow}(f; c) = |ber|$ . Theorem 4.3 tells us that there is a map  $g$  homotopic to  $f$  with exactly  $|ber|$  roots at  $c$  and, by Theorem 4.2, the map  $g$  can be chosen so that it is transverse to  $c$ .

**Example 4.6.** Let  $n \geq 4$  be an even integer. Let  $A$  be an  $n \times n$  integer matrix with determinant  $d$  odd and let  $f': T^n \rightarrow T^n$  be the corresponding map of the  $n$ -torus. We modify Example 2.4(b) to define a map  $f'': T^n \# P^n \rightarrow T^n$ , where  $P^n$  is  $n$ -dimensional real projective space. That is,  $f''$  is the identity on  $T^n - B^n$ , where  $B^n$  is an  $n$ -ball, and it maps  $P^n - B^n$  to  $B^n$ . Since  $n$  is even,  $P^n$  is a non-orientable manifold and, as in Example 2.4(b), we see that  $f''$  is a Type III map, so  $f = f' \circ f''$  is also Type III. We have  $\deg(f, 2) = 1$  (compare Example 3.19) and therefore  $N(f; c) = N_{\uparrow}(f; c) = |d|$  by Theorems 3.11 and 3.12. If  $c \notin B^n$ , it is evident that  $f$  is transverse to  $c$  and has exactly  $|d|$  roots at  $c$ . Theorem 4.3 tells us

that if  $c \in B^n$ , there is still a map  $g$  homotopic to  $f$  with exactly  $|d|$  roots at  $c$  and, by Theorem 4.2, we may require that  $g$  be transverse to  $c$  as well.

**Remark 4.7.** A root class  $R$  of a map  $f$  was defined in [Bk1, p. 24] to be *essential* if  $R$  cannot be removed by a homotopy of  $f$  (see [BB, p. 556] or [Bn2]). It was shown in [BS2, Remark 3.2], that if  $f: M \rightarrow N$  is a map of closed orientable  $n$ -manifolds, then a root class  $R$  of  $f$  is essential in this sense if and only if it has non-zero multiplicity. It follows from Theorem 4.3 that this is still true for boundary-preserving maps of not necessarily orientable manifold of dimension  $\neq 2$ , and from Theorem 6.1 below that is true for maps of closed surfaces. The reason is that Lemmas 3.6, 3.7, 3.9 and 3.10 establish the fact that the multiplicity  $|m(R)|$  of a root class is a proper homotopy invariant and therefore a root class with a non-zero multiplicity must be preserved by proper homotopies. Now Theorems 4.3 and 6.1 imply that, if their assumptions are satisfied, then there exists a proper homotopy from  $f$  to a map  $g$  that has no root classes of multiplicity zero and hence a root class with zero multiplicity cannot be preserved by all proper homotopies.

## 5. Relations between Nielsen root numbers and the degree of a map.

The purpose of this section is to interpret some of the concepts and results of the previous sections in the language of Hopf's degree theory. If  $f: (M, \partial M) \rightarrow (N, \partial N)$  is a proper map between  $n$ -manifolds, then the classical integer-valued cohomological degree  $\deg(f)$  only exists if both  $M$  and  $N$  are orientable (that is if their interiors are orientable). But if one or both of  $M$  and  $N$  are non-orientable, then one can only define the cohomological mod 2 degree  $\deg(f; 2)$  by using coefficients in  $\mathbb{Z}/2$ , and this degree generally provides little information about the geometric properties of the map  $f$ . In particular, it does not relate to the intuitive geometric concept of the degree, namely, the number of times the image  $f(M)$  of  $f$  covers the range  $N$ . This lack of an algebraic invariant which can characterise the geometric concept of the degree in all cases was drawn to the attention of Hopf by P. Alexandroff. That is the reason Hopf developed Nielsen root theory; he wanted to obtain algebraic, homotopy invariant information about the least number of "nice" (i.e., transverse) counter-images of points, and thus about the numbers of times  $f(M)$  covers  $N$ , and relate his algebraic invariant to the geometric degree (see Definition 5.2 below). Thus he proposed in [H2] a very different kind of degree that he called the "Absolutgrad", which provides geometric information even if one or both of the manifolds  $M$ ,  $N$  are not orientable. If  $M$  and  $N$  are both orientable and  $n \neq 2$ , then the Absolutgrad agrees with  $|\deg(f)|$ , the absolute value of the cohomological

degree (see Theorems 5.4 and 5.5 below). Here is the original definition of Hopf's degree.

**Definition 5.1** ([H2, Definition VIIc, p. 582]). Let  $f: (M, \partial M) \rightarrow (N, \partial N)$  be a proper map between two  $n$ -manifolds and let  $c \in \text{int } N$ . Then the *Absolutgrad* or *absolute degree*  $\mathcal{A}(f)$  of  $f$  is the sum of the multiplicities, in the sense of Definition 2.7, of its root classes.

The reason for Hopf's introduction of the absolute degree is that it provides an algebraic homotopy invariant that is closely linked to the very concrete concept of the geometric degree. This geometric interpretation of the absolute degree is based on the equality of the "algebraic" and "geometric" degrees which, as Hopf explained in the introduction to [H2] (see page 563), was the goal of that paper. (For maps of euclidean spaces, the equality had already been established in [H1, Satz IX, p. 590].) The following definition of the geometric degree is taken from [E, p. 372] and [Sk, p. 416]. It is a restatement of what Hopf understood by a geometric degree.

**Definition 5.2.** Let  $f: (M, \partial M) \rightarrow (N, \partial N)$  be a proper map between two  $n$ -manifolds. Then the *geometric degree*  $\mathcal{G}(f)$  of  $f$  is the least non-negative integer for which there exists a closed  $n$ -ball  $B^n \subset \text{int } N$  and a proper map  $g: (M, \partial M) \rightarrow (N, \partial N)$ , which is homotopic to  $f$  under a proper homotopy, such that  $g^{-1}(B^n)$  has  $\mathcal{G}(f)$  components, and each component is mapped by  $g$  homeomorphically onto  $B^n$ .

Note that for proper maps such an integer always exists, as  $\mathcal{G}(f)$  is bounded above by the number of roots of any transverse map homotopic to it.

The absolute and the geometric degree of a map are in essence concepts from Nielsen root theory. This fact follows immediately from the definition of these degrees in 5.1 and 5.2 and from the definition of the minimum number of transverse roots and the transverse Nielsen root number in (3.2) and (3.3), but we state it explicitly in the next theorem in order to emphasize and clarify the connection between Nielsen root theory and Hopf degree theory.

**Theorem 5.3.** *Let  $f: (M, \partial M) \rightarrow (N, \partial N)$  be a proper map between two  $n$ -manifold and let  $c$  be any point in  $\text{int } N$ . Then  $\mathcal{A}(f) = N_{\uparrow}(f; c)$  and  $\mathcal{G}(f) = MR_{\uparrow}(f; c)$ , that is, the absolute degree equals the transverse root Nielsen number and the geometric degree equals the least number of roots which are transverse at  $c$  for all maps in the proper homotopy class of  $f$ .*

One can see from the introduction of his paper [H2] that Hopf was very well aware of the fact brought out by Theorem 5.3, that the problem of finding the geometric degree is of a similar nature to the problem of finding

the least number of fixed points in the homotopy class of a map. Hopf clearly explains that his definition of the absolute degree uses an extension to root theory of concepts that had been used quite recently by J. Nielsen to study minimal sets of fixed points [N1, N2]. However, this motivation for Hopf's introduction of the absolute degree  $\mathcal{A}(f)$  is not mentioned in later studies and applications of  $\mathcal{A}(f)$ . As we mentioned in the introduction, Epstein in [E] interpreted calculations of Olum [O] in terms of degrees of maps of lifts of  $f$  to covering spaces, the covering spaces we described in Section 3. Olum obtained the values for  $N_{\uparrow}(f; c)$  computed here in Theorem 3.12, but in a very different way. Epstein made use of a classification of maps into types that is equivalent to Definition 2.1 and then defined the "absolute degree"  $\mathcal{A}(f)$  separately for each type [E, (1.8), p. 371]. Epstein acknowledged that  $\mathcal{A}(f)$  has a "complicated definition" [E, p. 372] but pointed out that it is justified by its geometric significance from the equality between the geometric and absolute degrees proved by Hopf. Epstein's paper is the one usually cited, so the connection with Nielsen root theory, that unified Hopf's treatment of the Absolutgrad, has not been preserved in Hopf degree theory.

The equality of the Absolutgrad and the geometric degree for maps of  $n$ -manifolds,  $n \neq 2$ , was first proved in [H2, Satz IV, p. 607]. A new proof of this equality was the aim of the paper of Epstein [E, Theorem 4.1, p. 376]. Because of the identifications described in Theorem 5.3, we see that Theorem 4.2 can be restated as the same result:

**Theorem 5.4.** *Let  $f: (M, \partial M) \rightarrow (N, \partial N)$  be a proper map between two  $n$ -manifolds. If  $n \neq 2$ , then the absolute degree  $\mathcal{A}(f)$  equals the geometric degree  $\mathcal{G}(f)$ .*

If  $M$  and  $N$  are both orientable, then Hopf degree theory can be related to the integer-valued cohomological degree that is defined in this case. From Theorems 3.13, 5.3 and 5.4 we have:

**Theorem 5.5.** *Let  $f: (M, \partial M) \rightarrow (N, \partial N)$  be a proper map between two  $n$ -manifolds. If  $M$  and  $N$  are orientable, then  $\mathcal{A}(f) = |\deg(f)|$ . If, further,  $n \neq 2$ , then  $\mathcal{G}(f) = |\deg(f)|$  also.*

## 6. Nielsen root numbers and degree of maps of surfaces.

Theorems 4.2, 4.3, 5.4 and 5.5, that are concerned with the sharpness of the two Nielsen root numbers for maps of  $n$ -manifolds, exclude the case  $n = 2$ . We will now discuss sharpness results for maps of surfaces. It was not known at the time of Hopf's work, but is now well known, that the Nielsen number for fixed points  $N(f)$  can be realized as the minimal set of fixed points for all maps in the homotopy class of  $f$  if  $f$  is a selfmap of a manifold of dimension  $\neq 2$ , but that this is often not possible for selfmaps of surfaces. So it is not surprising that the two Nielsen root numbers  $N_{\uparrow}(f; c)$  and  $N(f; c)$  are

not sharp for surface maps in general. But it is quite surprising that, as we shall see,  $N_{\uparrow}(f; c)$  may well be sharp when  $N(f; c)$  is not. In this section we describe what is known about sharpness for maps of surfaces, interpret it in the light of the results of this paper and add some consequences which follow from the results of prior sections.

Various authors have found conditions for the equality  $\mathcal{A}(f) = \mathcal{G}(f)$  to hold, and Theorem 5.3 shows that this equality is equivalent to the sharpness of  $N_{\uparrow}(f; c)$ . The basic result of this kind, which is stated as the next theorem, concerns closed surfaces.

**Theorem 6.1** (Kneser, Hopf, Skora). *Let  $f: M \rightarrow N$  be a map between two closed surfaces. Then  $N_{\uparrow}(f; c)$  is sharp and hence the geometric degree equals the absolute degree.*

Theorem 6.1 tells us, for instance, that there is a selfmap of the Klein bottle homotopic to the map of Example 3.18 that is transverse to  $c$  and has  $|be|$  roots. In other words, both the absolute and the geometric degree of the map  $f$  in Example 3.18 are equal to  $|be|$ . It further tells us that the map from the Klein bottle to the projective plane constructed in Example 3.16 has both absolute and geometric degree equal to 2.

Theorem 6.1 was first proved by H. Kneser [Kn1], [Kn2]. Hopf [H2, Satz XIVc, p. 605] showed that its assumptions can be weakened, as it is not necessary to assume that  $N$  is closed. Both Kneser and Hopf proved more general results concerning the fact that the map  $g$  which realizes  $\mathcal{A}(f) = N_{\uparrow}(f; c)$  can be constructed in such a way that it realizes the absolute degree as the geometric degree not only at  $c$ , but at all points in an open subset of  $N$  which is everywhere dense [H2, Satz XIVb, p. 605]. A modern proof of Theorem 6.1 was given by Skora in [Sk, Corollary 2.2].

Skora also found extensions of Theorem 6.1 to boundary-preserving maps. But additional assumptions are needed if the boundaries of the surfaces are not empty. One of these requires that  $f|\partial M$  be *allowable*, which means that  $f|\partial M$  is a  $(\mathcal{G}(f))$ -fold covering. It is a somewhat awkward assumption as  $\mathcal{G}(f)$  may not be known. In [Sk, Theorem 2.1 and 2.5], Skora proved the following extension of Theorem 6.1.

**Theorem 6.2** (Skora). *Let  $f: M \rightarrow N$  be a map between two surfaces with  $f^{-1}(\partial N) = \partial M$ . Then  $N_{\uparrow}(f; c)$  is sharp, and hence  $\mathcal{A}(f) = \mathcal{G}(f)$ , given that either  $f|\partial M$  is allowable and  $M$  and  $N$  are compact, or that  $f$  is orientation-true and proper.*

In addition, Skora showed that Theorem 6.2 is not true for  $n = 2$  without additional assumptions. In [Sk, §3], he constructed, for every  $d \geq 2$ , a boundary-preserving and proper but neither orientation-true nor allowable map from the twice punctured projective plane to the annulus such that

$\mathcal{G}(f) = d$  but  $\mathcal{A}(f) = 0$  if  $d$  is even and  $\mathcal{A}(f) = 1$  if  $d$  is odd. The computation of the geometric degree for all surface maps is an open problem.

The sharpness of the Nielsen root number  $N(f; c)$  is a more delicate property as the conclusions of Theorems 6.1 and 6.2 do not hold with  $N(f; c)$  in place of  $N_{\uparrow\uparrow}(f; c)$ . This was already known to Hopf who constructed examples, for every  $j > 4$  as well as for  $j = 1$ , of a map  $f$  from the double torus to the torus such that  $N(f; c) = j$  but  $f$  cannot be homotoped to a map with only  $N(f; c)$  roots (see [H2, Satz XVa, p. 610 and Satz XVb, p. 623]). Using results of Hopf, Lin [L, §4] constructed a map from the double torus to the torus with  $N(f; c) = 3$  and  $MR[f; c] = 4$ . For maps between orientable closed surfaces, necessary and sufficient conditions for the sharpness of the root Nielsen number have recently been found by D.L. Gonçalves and H. Zieschang [GZ1] and [GZ2]. They proved:

**Theorem 6.3** (Gonçalves and Zieschang). *Let  $M$  and  $N$  be closed orientable surfaces of genus  $h$  and  $k$  respectively. Let  $f: M \rightarrow N$  be a map such that  $j$  is finite, then  $N(f; c)$  is sharp if and only if*

$$|\deg(f)| \leq \frac{2h - 2 + j}{2k - 1}.$$

Some new results concerning the sharpness of  $N(f; c)$  for maps of surfaces can easily be obtained from our calculations of the two Nielsen root numbers in §3. We can use the fact that if  $N(f; c) = N_{\uparrow\uparrow}(f; c)$  and  $N_{\uparrow\uparrow}(f; c)$  is sharp then  $N(f; c)$  must also be sharp. So an inspection of Theorems 3.11, 3.12 and 6.1 immediately yields:

**Theorem 6.4.** *If  $f$  is a map between two closed surfaces which is not orientation-true, then  $N(f; c)$  is sharp.*

The map  $f: T^2 \# P^2 \rightarrow T^2$  of Example 3.19 has exactly  $|d|$  roots at  $c$  if  $c \notin D$ . Theorem 6.4 implies that if  $c \in D$ , then there is a map homotopic to  $f$  that has  $|d|$  roots at  $c$  (compare Example 4.6).

Theorems 6.3 and 6.4 give necessary and sufficient conditions for the sharpness of  $N(f; c)$  for all maps between closed surfaces, with the exception of the case of orientation-true maps between non-orientable surfaces. The Nielsen root number may be sharp in this case also. For instance, for any  $c \in P^2$ , the map  $f: K \rightarrow P^2$  of Example 3.16 is homotopic to a map  $g$  with one root at  $c$  because  $f$  is one-to-one at two points and  $P^2$  is homogeneous. However, it is likely that, for orientation-true maps between non-orientable manifolds,  $N(f; c)$  will only be sharp for some maps, as is true in the orientable case. We contribute a partial solution to this problem in the following theorem, which again follows immediately from Theorems 3.11, 3.12 and 6.1.

**Theorem 6.5.** *Let  $f$  be an orientation-true map between two non-orientable closed surfaces. Then  $N(f; c)$  is sharp if either  $j$  is infinite or if  $|\deg(\tilde{f})| \leq 1$ .  $\square$*

Finally, Theorems 3.11, 3.12 and 6.2 yield some results concerning the sharpness of  $N(f; c)$  in the case where the boundaries of the surfaces  $M, N$  are non-empty. We omit the details.

## References

- [Bd] G. Bredon, *An Introduction to Compact Transformation Groups*, Academic Press, New York, 1972.
- [Bk1] R. Brooks, *Coincidences, roots and fixed points*, Doctoral Dissertation, Univ. of California, Los Angeles, 1967.
- [Bk2] ———, *On the sharpness of the  $\Delta_2$  and  $\Delta_1$  Nielsen numbers*, J. Reine Angew. Math., **259** (1973), 101-108.
- [BB] R. Brooks and R. Brown, *A lower bound for the  $\Delta$ -Nielsen number*, Trans. Amer. Math. Soc., **143** (1969), 555-564.
- [BO] R. Brooks and C. Odenthal, *Nielsen numbers for roots of maps of aspherical manifolds*, Pacific J. Math., **170** (1995), 405-420.
- [Bw] L.E.J. Brouwer, *Über Abbildung von Mannigfaltigkeiten*, Math. Ann., **71** (1911), 97-115.
- [Bn1] R. Brown, *The Lefschetz Fixed Point Theorem*, Scott-Foresman, 1971.
- [Bn2] ———, *A middle-distance look at root theory*, in ‘Nielsen Theory and Reidemeister Torsion’, Banach Center Publications, **49** (1999), 29-41.
- [BS1] R. Brown and H. Schirmer, *Correction to “Nielsen coincidence theory and coincidence-producing maps for manifolds with boundary”*, Top. Appl., **67** (1995), 233-234.
- [BS2] ———, *Nielsen theory of roots of maps of pairs*, Top. Appl., **92** (1999), 247-274.
- [DJ] R. Dobrenko and J. Jezierski, *The coincidence Nielsen theory on non-orientable manifolds*, Rocky Mt. J. Math., **23** (1993), 67-85.
- [Do] A. Dold, *Lectures on Algebraic Topology*, 2nd edition, Springer-Verlag, Berlin, 1980.
- [E] D.B.A. Epstein, *The degree of a map*, Proc. London Math. Soc., **16** (1966), 369-383.
- [GJ] D. Goncalves and J. Jezierski, *Lefschetz coincidence formula on non-orientable manifolds*, Fund. Math., **153** (1997), 1-23.
- [GZ1] D. Goncalves and H. Zieschang, *Equations in free groups and coincidences of mappings on surfaces*, preprint.
- [GZ2] ———, *Equations in free groups and coincidences of mappings on surfaces*, II, preprint.
- [H1] H. Hopf, *Zur Topologie der Abbildungen von Mannigfaltigkeiten, Erster Teil*, Math. Ann., **100** (1928), 579-608.
- [H2] ———, *Zur Topologie der Abbildungen von Mannigfaltigkeiten, Zweiter Teil*, Math. Ann., **102** (1930), 562-623.

- [Je] J. Jezierski, *The Nielsen coincidence theory on topological manifolds*, Fund. Math., **143** (1993), 167-178.
- [Jg1] B. Jiang, *Fixed point classes from a differentiable viewpoint*, Fixed Point Theory (Proceedings, Sherbrooke, Quebec, 1980), Lecture Notes in Mathematics, **886**, Springer-Verlag, Berlin, (1981), 163-170.
- [Jg2] ———, *Lectures on Nielsen Fixed Point Theory*, Contemporary Math., **14** (1983).
- [Kg] T. Kiang, *The Theory of Fixed Point Classes*, Springer-Verlag and Science Press, Berlin and Beijing, 1989.
- [KS] R. Kirby and L. Siebenmann, *Foundational Essays on Topological Manifolds*, Annals of Math. Studies, **88**, Princeton U. Press, 1977.
- [Kn1] H. Kneser, *Glättung von Flächenabbildungen*, Math. Ann., **100(100)** (1928), 609-616.
- [Kn2] ———, *Die kleinste Bedeckungszahl innerhalb einer Klasse von Flächenabbildungen*, Math. Ann., **103** (1930), 347-358.
- [L] X.-S. Lin, *On the root classes of mapping*, Acta Math. Sinica N. S., **2** (1986), 199-206.
- [M] W. Massey, *Algebraic Topology: An Introduction*, Harcourt, Brace and World, New York, 1967.
- [N1] J. Nielsen, *Untersuchungen zur Topologie der geschlossene zweiseitigen Fläche*, Acta Math., **50** (1927), 189-358.
- [N2] ———, *Untersuchungen zur Topologie der geschlossene zweiseitigen Fläche*, II, Acta Math., **53** (1929), 1-76.
- [O] P. Olum, *Mappings of manifolds and the notion of degree*, Ann. Math., **58** (1953), 458-480.
- [RW] Y. Rong and S. Wang, *The preimages of submanifolds*, Math. Proc. Camb. Phil. Soc., **112** (1992), 271-279.
- [Sm] H. Schirmer, *Nielsen theory of transversal fixed point sets*, Fund. Math., **141** (1992), 31-59.
- [Sk] R. Skora, *The degree of a map between surfaces*, Math. Ann., **276** (1987), 415-423.
- [V] J. Vick, *Homology Theory*, Academic Press, New York, 1973.

Received May 10, 1999.

UNIVERSITY OF CALIFORNIA  
LOS ANGELES, CA 90095-1555  
*E-mail address:* rfb@math.ucla.edu

CARLETON UNIVERSITY  
OTTAWA, ONTARIO K1S 5B6  
CANADA  
*E-mail address:* schirmer@carleton.ca