

Birationality of étale maps via surgery

Scott Nollet

Department of Mathematics
Texas Christian University
Fort Worth, TX 76129

Laurence R. Taylor

Department of Mathematics
University of Notre Dame
Notre Dame, IN 46556

Frederico Xavier *

Department of Mathematics
University of Notre Dame
Notre Dame, IN 46556

Abstract

We use a counting argument and surgery theory to show that if D is a sufficiently general algebraic hypersurface in \mathbb{C}^n , then any local diffeomorphism $F : X \rightarrow \mathbb{C}^n$ of simply connected manifolds which is a d -sheeted cover away from D has degree $d = 1$ or $d = \infty$ (however all degrees $d > 1$ are possible if F fails to be a local diffeomorphism at even a single point). In particular, any étale morphism $F : X \rightarrow \mathbb{C}^n$ of algebraic varieties which covers away from such a hypersurface D must be birational.

1 Introduction

Polynomial maps $F : \mathbb{C}^n \rightarrow \mathbb{C}^n$ with nonvanishing jacobian have received much attention over the last 25 years, since the excellent expository article of Bass, Connell and Wright [1] pointed out that the Jacobian Conjecture is still unsolved. The map F has a finite degree $d > 0$ and one formulation of the conjecture asks whether necessarily $d = 1$. It is well-known that there

*Work partially supported by NSF grant DMS02-03637.

is a Zariski open set $U \subset \mathbb{C}^n$ for which the restriction $F^{-1}(U) \rightarrow U$ is a d -sheet covering map [8, Prop. 3.17]. More recently, Jelonek has shown that if $D \subset \mathbb{C}^n$ is the closed set over which F is not a d -cover, then either $D = \emptyset$ or D is a uniruled hypersurface [6]. He also gives a sharp degree bound for the equation of D and explains how to compute its equation with a Gröbner basis. Of course if $D = \emptyset$, then F is a covering map and simple connectivity of \mathbb{C}^n implies that $d = 1$ as conjectured, so one may study the problem in terms of the hypersurface D .

Kulikov [7] considers the more general question of whether an étale morphism $F : X \rightarrow \mathbb{C}^n$ of simply connected varieties which is surjective modulo codimension two must be birational, i.e. have degree $d = 1$. In terms of the divisor D above, he observes that it is equivalent to ask whether any subgroup of $\pi_1(\mathbb{C}^n - D)$ generated by geometric generators (loops about D with winding number one) must have index one. The answer is yes for $n = 1$, but Kulikov constructs a counterexample for $n = 2$ of degree $d = 3$ by taking a quartic curve $D \subset \mathbb{C}^2$ with three cusps and producing a subgroup $G \subset \pi_1(\mathbb{C}^2 - D)$ of index three defined by a geometric generator [7, §3]. We will show by contrast that for general hypersurfaces $D \subset \mathbb{C}^n$, F must indeed be birational.

A hypersurface $D \subset \mathbb{C}^n$ is given by an equation $f(z_1, z_2, \dots, z_n) = 0$. A theorem of Verdier says that the corresponding map $f : \mathbb{C}^n \rightarrow \mathbb{C}$ is a locally trivial fibration away from a finite subset of \mathbb{C} [17]. The smallest such set, the bifurcation locus $B_f \subset \mathbb{C}$, contains the image of the critical values of f , but may also contain the images of critical points at infinity: since f is not proper, one cannot apply the Ehresmann fibration theorem. Generalizing work of Hà Huy Vui and Lê Dũng Tráng [5], Parusiński shows that if f has only isolated singularities at infinity, then a regular value t_0 for f is not in B_f if and only if the Euler characteristic of $f^{-1}(t_0)$ is locally constant at t_0 [12]. If f is the polynomial of minimal degree defining D , we will say that D is *non-bifurcated* if 0 is not a bifurcation value for f (i.e. $0 \notin B_f$, meaning f is a trivial fibration in a neighborhood of $D = f^{-1}(0)$). More information on the behavior of polynomials can be found in [2] and [16].

Theorem 1.1. *Fix $n > 1$ and let $D \subset \mathbb{C}^n$ be a smooth connected non-bifurcated hypersurface. If $F : X \rightarrow \mathbb{C}^n$ is a local diffeomorphism of simply connected manifolds which is a d -fold covering map away from D , then $d = 1$ or $d = \infty$.*

For $D \subset \mathbb{C}^n$ as in 1.1, we highlight two algebro-geometric cases:

Corollary 1.2. *Let $D \subset \mathbb{C}^n$ be a smooth connected non-bifurcated hypersurface. If $F : X \rightarrow \mathbb{C}^n$ is an étale morphism with X simply connected and $\#F^{-1}(q) = \deg F$ for $q \notin D$, then F is birational.*

Corollary 1.3. *Let $D \subset \mathbb{C}^n$ be a smooth connected non-bifurcated hypersurface. If $F : \mathbb{C}^n \rightarrow \mathbb{C}^n$ is a polynomial map with nonvanishing jacobian and $\#F^{-1}(q) = \deg F$ for $q \notin D$, then $F \in \text{Aut}(\mathbb{C}^n)$.*

The conclusion of 1.2 also holds if D has at worst simple normal crossings away from a set of codimension ≥ 3 and meets the hyperplane at infinity transversely [11]. It would be interesting to try to weaken the hypothesis of Theorem 1.1, for example what happens if D merely smooth and connected? Or if D is a non-bifurcation hypersurface but reducible?

To state our main tool for obtaining these results, we need a definition. An $(n - 1)$ -submanifold $A \subset \mathbb{R}^n$ **nicely bounds** a closed subset $D \subset \mathbb{R}^n$ if

1. $D = \partial A = \bar{A} - A$.
2. Each connected component of \bar{A} contains exactly one connected component of D .
3. D is the closure of an $(n - 2)$ -submanifold $D_0 \subset \mathbb{R}^n$ with singular locus $\Sigma = D - D_0$ of codimension ≥ 4 in \mathbb{R}^n and \bar{A} is locally diffeomorphic to a half space along D_0 .

Remark 1.4. (a) For typical hypersurfaces $D \subset \mathbb{C}^n$, the existence of a nicely bounding manifold A is natural. Suppose that D is defined by the equation $P(z_1, z_2, \dots, z_n) = 0$ and $S \subset \mathbb{C}$ is a finite set for which the restriction $P : \mathbb{C}^n - P^{-1}(S) \rightarrow \mathbb{C} - S$ is a locally trivial fibration with fibre f as in Verdier's theorem [17]. If f is smooth and connected, then in choosing an open ray $l \subset \mathbb{C}$ emanating from 0 which misses S , the contractibility of l implies that $A = P^{-1}(l) \cong l \times f$ is an oriented manifold and it is clear that A nicely bounds D . If D is a non-bifurcation hypersurface, then $D \cong f$ and $\bar{A} \cong \bar{l} \times D$, hence we obtain an isomorphism $\pi_1(D) \rightarrow \pi_1(\bar{A})$.

(b) If $F : X \rightarrow \mathbb{C}^n$ is a d -fold cover away from D (with equation $P = 0$), we can arrange the situation in (a) by enlarging D as follows. Replace P with PQ , where Q is general and $\deg(PQ)$ is prime. Then the hypothesis on F is preserved, D is connected and P cannot be written $h(g(z_1, \dots, z_m))$ for $h(t) \in \mathbb{C}[t]$ of degree > 1 , which implies that the general fibre $f = P^{-1}(t)$ is smooth and connected by a Bertini theorem [14, Cor. 1 to Thm. 37]. Of course this divisor D may be bifurcated.

The degree of such maps often satisfies a certain trichotomy (Thm. 2.7):

Theorem 1.5. *Let $F : X \rightarrow \mathbb{R}^n$ be a local diffeomorphism of connected manifolds with $H_1(X, \mathbb{Z}) = 0$ and let $D \subset \mathbb{R}^n$ be a closed set for which the restriction $X - F^{-1}(D) \rightarrow \mathbb{R}^n - D$ is a d -sheeted covering map. If D is nicely bounded by a connected $(n - 1)$ -submanifold $A \subset \mathbb{R}^n$ with $\mathbb{R}^n - \overline{A}$ simply connected, then $d = 1, 2$ or ∞ .*

The conclusion can be false if F fails to be a local diffeomorphism at even a *single point* (visualize the d th power map on \mathbb{C} for $d > 2$).

Remark 1.6. *If the nicely bounding manifold A is oriented, but possibly disconnected, then $d = 1$ or $d = \infty$ (Theorem 2.5).*

Remark 1.7. Theorem 1.5 is sharp as follows.

(a) All three values of d occur. The identity map realizes $d = 1$ (for example take $D = \mathbb{R}^{n-2} \times (0, 0)$ and $A = \mathbb{R}^{n-2} \times (0, \infty) \times 0$). The value $d = 2$ is realized by a map $F : X \rightarrow \mathbb{R}^4$ with $D \subset \mathbb{R}^4$ an embedded Klein bottle which is nicely bounded by $M \times [-2, 2]$, where $M \subset \mathbb{R}^3$ is a Moebius band (see Example 2.8). The value $d = \infty$ is achieved by the complex exponential map $F : \mathbb{C} \rightarrow \mathbb{C}$ with $D = \{0\}$.

(b) The vanishing $H_1(X, \mathbb{Z}) = 0$ is necessary, for example the d th power map $F : \mathbb{C} - \{0\} \rightarrow \mathbb{C}$ is a d -sheet covering map.

Corollary 1.8. *Let $F : X \rightarrow \mathbb{R}^2$ be a local diffeomorphism. Assume that X is connected, $H_1(X, \mathbb{Z}) = 0$ and the restriction $F : X - f^{-1}(D) \rightarrow \mathbb{R}^2 - D$ is a d -sheeted cover for a finite subset $D \subset \mathbb{R}^2$. Then $d = 1$ or $d = \infty$.*

Corollary 1.8 follows from Remark 1.6 by taking A to be the union of disjoint open rays emanating from points in D , when $\mathbb{R}^2 - \overline{A}$ is contractible. For hypersurfaces $D \subset \mathbb{C}^n$ with $n > 1$ it is unlikely that the submanifold A produced in Remark 1.4 will satisfy $\pi_1(\mathbb{C}^n - \overline{A}) = 0$, hence Theorem 1.5 does not apply. In typical situations surgery theory [13] can be used to kill homotopy classes of maps and we can modify A to arrive at a suitable manifold B . For this we need more conditions on the nicely bounding manifold A .

1. D has finitely many connected components D_i contained in corresponding components $A_i \subset A$.
2. If $\Sigma \subset D$ is the singular locus, then $\pi_1(D_i - \Sigma) \rightarrow \pi_1(\overline{A}_i - \Sigma)$ is onto.

3. Each component A_i is orientable with $\pi_1(A_i)$ finitely generated.

Theorem 1.9. *Suppose that D is nicely bounded by a submanifold $A \subset \mathbb{R}^n$ satisfying the conditions 1-3 above and $n \geq 6$. Then there is another such B nicely bounding D such that $\pi_1(\mathbb{R}^n - \overline{B}) = 0$.*

Now Theorem 1.1 follows easily. For $n \geq 3$ and $F : X \rightarrow \mathbb{C}^n$, we produce an orientable submanifold $A \subset \mathbb{C}^n$ nicely bounding D as in Remark 1.4. The hypersurface D has finitely generated fundamental group because it is a finite CW complex [2]. Since there is only one connected component and Σ is empty, Remark 1.4(a) gives conditions 1-3 above. Theorem 1.9 allows us to replace nicely bounding A with B satisfying $\pi_1(\mathbb{C}^n - \overline{B}) = 0$, at which point Remark 1.6 gives the result. If $n = 2$, we consider $F \times \text{Id} : X \times \mathbb{C} \rightarrow \mathbb{C}^n \times \mathbb{C}$ and replace D with $D \times \mathbb{C}$.

We expect the surgery technique to work even if the hypersurface D is bifurcated, provided that the singularities are sufficiently mild. These problems fit in with our general program of trying to understand local versus global injectivity [9, 10, 11, 18]. Indeed, for such maps with finite fibres, we are really asking when a local immersion is injective. As to the structure of the paper, Theorem 1.5 is proven in § 2 and Theorem 1.9 is proven in § 3. We thank Francis Connolly for useful conversations.

2 A Counting Argument

In this section we prove Theorem 1.5 from the introduction. The dichotomy in the cases $d = 1$ and $d = \infty$ depends on the nature of the pre-image $F^{-1}(A) \subset X$. We will show that if no component of $F^{-1}(A)$ is closed, then $d = 1$, while if at least one component is closed, then $d = \infty$.

Proposition 2.1. *Let $f : X \rightarrow \mathbb{R}^n$ be a local diffeomorphism of connected manifolds which is a d -sheet cover away from a closed set D . Further assume that there is a $(n - 1)$ -submanifold $A \subset \mathbb{R}^n$ such that*

1. D is nicely bounded by A .
2. $\mathbb{R}^n - \overline{A}$ is simply connected.
3. No connected component of $f^{-1}(A)$ is closed in X .

Then $d = 1$.

Proof. It will suffice to show that $X - f^{-1}(\overline{A})$ is path-connected, for then $X - f^{-1}(\overline{A}) \rightarrow \mathbb{R}^n - \overline{A}$ is a d -sheeted covering map with $\mathbb{R}^n - \overline{A}$ simply connected, whence $d = 1$ [15, Ch. 2]. In particular, f is injective and identifies X with a dense open subset of \mathbb{R}^n .

To see that $X - f^{-1}(\overline{A})$ is path-connected, consider $a \neq b \in X - f^{-1}(\overline{A})$ and consider a path $\tau : [0, 1] \rightarrow X$ from a to b . Deforming τ , we may assume that

- (a) τ avoids $f^{-1}(D)$.
- (b) τ meets $f^{-1}(A)$ transversely a finite number of times.

Suppose that τ meets a connected component E of $f^{-1}(A)$, say $\tau(t_0) \in E$. Since E is not closed by hypothesis, there is a point $q \in \overline{E} - E \subset f^{-1}(D)$ which necessarily satisfies $f(q) \in \partial A = D$. Taking an open neighborhood U about q which maps diffeomorphically onto the neighborhood $f(U)$ about $f(q) \in D$, we have a diffeomorphism of pairs $(\overline{E} \cap U, U) \cong (\overline{A} \cap f(U), f(U))$ and since $\Sigma \subset D$ has positive codimension, we can choose $q \in U$ so that $f(q)$ avoids Σ and \overline{E} looks locally like a half-space at q . With this in mind, consider a path σ from $\tau(t_0)$ to the point $q \in \overline{E} - E$. Cover the image of σ with small open balls to obtain a tubular open neighborhood N of σ . A ball about q is not disconnected by \overline{E} and it follows that $N - \overline{E}$ itself is connected. We may thus replace a segment of the path τ through $\tau(t_0)$ with a path contained in $N - \overline{E}$. Continuing this way with each intersection, we obtain a new path from a to b which avoids $f^{-1}(\overline{A})$, hence the conclusion. \square

Remark 2.2. That f be a local diffeomorphism at each point is critical in the result above. For example, the squaring map $f : \mathbb{C} \rightarrow \mathbb{C}$ by $f(z) = z^2$ is a local diffeomorphism everywhere except at $z = 0$. Taking $D = \{0\}$ and A to be the positive real axis, the other hypotheses of Proposition 2.1 hold, but $d = 2$ and the conclusion fails. The proof given above breaks down because $f^{-1}(A)$ is obtained from removing the origin from the real axis, when it is clear that $\mathbb{C} - f^{-1}(\overline{A})$ is not connected: we can't connect the upper half plane and lower half plane by going around the corner at $z = 0$ precisely because f is not a local diffeomorphism there.

Now we deal with the case where some of the connected components of $f^{-1}(A)$ are closed. We begin with the following separation lemma. The example $X = S^1$ shows the necessity of the vanishing $H_1(X) = 0$.

Lemma 2.3. *Let X be a connected n -manifold satisfying $H_1(X, \mathbb{Z}) = 0$. If E_1, \dots, E_k are k disjoint closed connected $(n - 1)$ -submanifolds of X , then $X - \bigcup_{i=1}^k E_i$ has $k + 1$ connected components U_1, \dots, U_{k+1} .*

Proof. This result is probably well-known to experts (the case $k = 1$ can be found in [3]), but we include a proof for lack of suitable reference. The vanishing $H_1(X, \mathbb{Z}) = 0$ implies that X is orientable [3, VIII, 2.12]. Let $Y = \bigcup E_i$. If $\overline{H}_c^q(A, B)$ denotes the Alexander cohomology module with compact supports and coefficients in \mathbb{Z} [15, §6.4], the corresponding long exact cohomology sequence contains the fragment

$$\overline{H}_c^{n-1}(X) \rightarrow \overline{H}_c^{n-1}(Y) \rightarrow \overline{H}_c^n(X, Y) \rightarrow \overline{H}_c^n(X) \rightarrow \overline{H}_c^n(Y).$$

The last group vanishes due to dimension and the Alexander duality theorem [15, 6.9.11] yields $\overline{H}_c^{n-1}(X) \cong H_1(X) = 0$, $\overline{H}_c^{n-1}(Y) \cong H_0(Y) \cong \mathbb{Z}^k$, $\overline{H}_c^n(X) \cong H_0(X) \cong \mathbb{Z}$. It follows that $H_0(X - Y) = \overline{H}_c^n(X, Y) \cong \mathbb{Z}^{k+1}$, hence $U = X - Y$ has $k + 1$ connected components. \square

Proposition 2.4. *Let $f : X \rightarrow \mathbb{R}^n$ be a local diffeomorphism of connected manifolds which is a d -sheet cover away from a closed set D and assume that $H_1(X, \mathbb{Z}) = 0$. Assume that there is an oriented $(n - 1)$ -submanifold $A \subset \mathbb{R}^n$ such that*

1. D is nicely bounded by A .
2. $f^{-1}(A)$ has a connected component which is closed in X .

Then $d = \infty$.

Proof. We assume that $d < \infty$ and draw a contradiction from the counting argument that gives this section its name. Let E be a connected component of $f^{-1}(A)$ which is closed in X . Then $f(E)$ is contained in some connected component A_1 of A and the covering property shows that $f(E) = A_1$. The closed sub-manifold $E \subset X$ separates X into two connected components U_1 and U_2 by Lemma 2.3. Clearly the induced maps

$$U_i - f^{-1}(\overline{A}) \rightarrow \mathbb{R}^n - \overline{A}$$

have the covering property with sheet numbers adding to d , but $U_i - f^{-1}(\overline{A})$ may be disconnected if $f^{-1}(A)$ has closed components other than E .

Choose $q \in A_1$ and write $f^{-1}(q) = \{p_1, \dots, p_d\}$ with $p_1, \dots, p_r \in E$ for some $1 \leq r \leq d$. If B is a sufficiently small ball centered at q , then by the covering property we may write

$$f^{-1}(B) = B_1 \cup B_2 \cup \dots \cup B_d$$

disjointly with each B_i mapping homeomorphically onto B and $p_i \in B_i$. Thus B_1, \dots, B_r meet E and each B_{r+1}, \dots, B_d avoids E , hence is contained in either U_1 or U_2 . We may relabel the indices so that $B_{r+1} \dots B_s$ are contained in U_1 and $B_{s+1} \dots B_d$ are contained in U_2 .

Shrinking B if necessary, we may further assume that $B \cap A_1$ splits B into two half balls B^+ and B^- and that B does not meet A otherwise (in particular it intersects no other components of A). For $1 \leq i \leq r$, each B_i is separated by E into two connected pieces B_i^+ and B_i^- , with B_i^+ (resp. B_i^-) mapping homeomorphically onto B^+ (resp. B^-).

Since E separates X , either $B_1^+ \subset U_1$ or U_2 . We suppose that $B_1^+ \subset U_1$ and $B_1^- \subset U_2$, but the proof is similar if it were the other way around. From the orientability of A_1 , it follows that the pre-images of the half balls are coherently arranged in the sense that $B_i^+ \subset U_1$ (resp. $B_i^- \subset U_2$) for all $1 \leq i \leq r$. To see this, consider a path $\gamma(t)$ from p_1 to p_i for $0 \leq t \leq 1$. If $N(t)$ is a vector field along γ which points into U_1 , then the vector field $df(N)$ along the loop $f \circ \gamma$ points into B^+ at time $t = 0$. By invariance of local orientations under continuations along a closed path, $df(N)$ also points into B^+ at time $t = 1$, which implies that $B_i^+ \subset U_1$. Similarly the B_i^- are contained in U_2 .

Now we arrive at a contradiction by counting pre-images. A point in B^+ has s pre-images in U_1 : r of these come from the B_i^+ with $1 \leq i \leq r$ and the remaining $s - r$ come from $B_i \subset U_1$ for $r < i \leq s$. Similarly a point in B^- has $s - r$ pre-images in U_1 , all coming from B_i^- with $r < i \leq s$. In particular the cardinality of the fibres of $U_1 - f^{-1}(\bar{A}) \rightarrow \mathbb{R}^n - \bar{A}$ is not constant ($s \neq s - r$), a contradiction as this is a covering over a connected base. \square

Theorem 2.5. *Let X be a connected manifold satisfying $H_1(X, \mathbb{Z}) = 0$ and let $F : X \rightarrow \mathbb{R}^n$ be a local diffeomorphism which is a d -fold cover away from a closed codimension two subset $D \subset \mathbb{R}^n$. Suppose that there is an oriented $(n - 1)$ -submanifold $A \subset \mathbb{R}^n$ such that*

1. D is nicely bounded by A .
2. $\mathbb{R}^n - \bar{A}$ is simply connected.

Then $d = 1$ or $d = \infty$.

Proof. This follows from Propositions 2.1 and 2.4, since either $f^{-1}(A)$ has a closed connected component or it does not. \square

Remark 2.6. If D is given by the vanishing of $P(z_1, \dots, z_n)$ and the hypersurface $P(z_1, \dots, z_n) = c$ is simply connected for general $c \in \mathbb{C}$, then we can construct a suitable submanifold A as follows. There is a finite point set $B = \{b_1, \dots, b_s\} \subset \mathbb{C}$ such that $\mathbb{C}^n - P^{-1}(B) \rightarrow \mathbb{C} - B$ is a locally trivial fibration with fibre F . Enlarge B to assume that $0 \in B$. Now choose disjoint open rays l_i emanating from the b_i . Set $L = \bigcup l_i$ and $A = P^{-1}(L)$. Then $D = P^{-1}(B)$ has s connected components and is nicely bounded by A . Furthermore, since $\mathbb{C}^n - P^{-1}(\overline{A}) \rightarrow \mathbb{C} - \overline{A}$ is a locally trivial fibration over a contractible base with simply connected fibre F , it follows that $\mathbb{C}^n - P^{-1}(\overline{A})$ is simply connected. Thus Theorem 2.5 applies and $d = 1$ or $d = \infty$.

Our methods also work when A is not orientable, but connected.

Theorem 2.7. *Let X be a connected manifold satisfying $H_1(X) = 0$ and let $F : X \rightarrow \mathbb{R}^n$ be a local diffeomorphism which is a d -fold cover away from a closed subset $D \subset \mathbb{R}^n$. Suppose that there is a connected $(n - 1)$ -submanifold $A \subset \mathbb{R}^n$ such that*

1. D is nicely bounded by A .
2. $\mathbb{R}^n - \overline{A}$ is simply connected.

Then $d = 1, 2$ or ∞ .

Proof. If no connected component of $f^{-1}(A)$ is closed, then Proposition 2.1 still applies and we obtain $d = 1$. If there is a closed such component, we must modify the proof of Proposition 2.4. Here we consider *all* the closed connected components E_1, E_2, \dots, E_m of $f^{-1}(A)$. By Lemma 2.3 these separate X into $m + 1$ connected components V_1, V_2, \dots, V_{m+1} . Now each $V_i - f^{-1}(\overline{A}) \rightarrow \mathbb{R}^n - \overline{A}$ has the covering property. Moreover, we may argue as in Prop. 2.1 that each $V_i - f^{-1}(\overline{A})$ is connected: that f is a local diffeomorphism allows us to “go around the edge” in the same way. Since the base is 1-connected, we conclude that these maps are homeomorphisms and $m + 1 = d$.

Letting d_i be the covering degree of the restriction $f : E_i \rightarrow A$, we have $\sum_{i=1}^{d-1} d_i \leq d$. If $d > 2$, this forces $d_i = 1$ for some i (since $d_i \geq 1$ already), say $i = 1$. Then $E_1 \cong A$ via f and we run the proof of Proposition 2.4 with

$E = E_1$. Things proceed as before, though now $r = 1$. We don't need to use the orientability of A to check that the B_i^+ can be coherently arranged for $1 \leq i \leq r$, since there is only one. Thus the counting contradiction runs as before, as the preimage of a point in B^+ has s pre-images in U_1 while a point in B^- has $s - r = s - 1$ such pre-images. Hence $2 < d < \infty$ gives a contradiction, leaving $d = 1, 2$ or ∞ . \square

Remark 2.8. Comparing to Theorem 2.5, we see that when A is not orientable the new possibility $d = 2$ arises. We give an example in which this new possibility is realized.

Fix an unknotted circle $C_0 \subset \mathbb{R}^3$ and consider its normal disk bundle, which is a solid torus. Inside the solid torus there is a Moebius band M which makes one half twist as it goes around. The boundary $C_\partial = \partial M$ of this Moebius band is a 2-1 torus knot which is known to be the unknot. We will show that $M \times (-2, 2) \subset \mathbb{R}^4$ nicely bounds a closed set K , which can be smoothed to obtain a Klein bottle K_1 for which $\pi_1(\mathbb{R}^4 - K) \cong \mathbb{Z}/2\mathbb{Z}$. In particular, the composition $X \rightarrow \mathbb{R}^4 - K \hookrightarrow \mathbb{R}^4$ of the universal covering map with the inclusion has degree $d = 2$.

To verify the claims above, let $H_+ = \mathbb{R}^3 \times [0, \infty) \subset \mathbb{R}^4$ be the 4-dimensional half space. For each integer $n > 0$ we set $t_n^\pm = 2 \pm \frac{1}{n}$ and define the manifold $W_n \subset H_+$ by

$$W_n = ((\mathbb{R}^3 - C_\partial) \times [0, t_n^-]) \cup (\mathbb{R}^3 - M) \times [t_n^-, t_n^+] \cup (\mathbb{R}^3 \times [t_n^+, \infty)).$$

Notice that $W = \bigcup_{n=1}^{\infty} W_n = H_+ - (C_\partial \times [0, 2] \cup M \times \{2\})$ is just $H - +$ minus a Moebius band. It is a bit angular, but we will round off the corners later. More importantly, the inclusion $W_n \subset W_{n+1}$ is a deformation retract and hence a homotopy equivalence: thus $\pi_1(W)$, as the limit of the $\pi_1(W_n)$, is isomorphic to $\pi_1(W_n)$ for any n .

We compute $\pi_1(W_1)$ as follows. Apply Van Kampen's theorem to

$$V = ((\mathbb{R}^3 - C_\partial) \times [0, 1]) \cup ((\mathbb{R}^3 - M) \times [1, 3]).$$

The intersection of the two pieces is $(\mathbb{R}^3 - M) \times 1$ and the two inclusions $(\mathbb{R}^3 - M) \times 1 \subset (\mathbb{R}^3 - M) \times [1, 3]$ and $(\mathbb{R}^3 - C_\partial) \times 0 \subset (\mathbb{R}^3 - C_\partial) \times [0, 1]$ are homotopy equivalences.

Since C_∂ is unknotted, $\pi_1(\mathbb{R}^3 - C_\partial) \cong \mathbb{Z}$ and a generator is any small circle L_∂ linking C_∂ . Since $\pi_1(\mathbb{R}^3 - C_\partial)$ is Abelian, base points are not important. Recall $C_0 \subset M$ is the middle circle in the Moebius and note that

the inclusion $\mathbb{R}^3 - M \subset \mathbb{R}^3 - C_0$ is a homotopy equivalence. Since C_0 is unknotted, $\pi_1(\mathbb{R}^3 - M) \cong \mathbb{Z}$. One choice of generator is to pick a meridian circle to the solid torus containing M and push it out just a bit until it no longer intersects the solid torus, call this circle L_0 . Since $\pi_1(\mathbb{R}^3 - M)$ is abelian, base points are not important here.

Finally compute the map $\pi_1(\mathbb{R}^3 - M) \rightarrow \pi_1(\mathbb{R}^3 - C_\partial)$. The obvious disk bounding the generator of $\pi_1(\mathbb{R}^3 - M)$ intersects C_∂ in two points, both with the same sign, hence the map is multiplication by ± 2 : $L_0 = \pm 2L_\partial$. Thus $\pi_1(V) \cong \mathbb{Z}$ with generator L_∂ .

To finish, we apply Van Kampen's theorem to

$$W_1 = V \cup (\mathbb{R}^3 \times [3, \infty)).$$

The intersection is $(\mathbb{R}^3 - M) \times 3$. Since $\pi_1(\mathbb{R}^3 \times [3, \infty)) = 0$, it follows that $\pi_1(W_1) = \pi_1(V)/\pi_1((\mathbb{R}^3 - M) \times 3)$ which by the above calculation is $\mathbb{Z}/2\mathbb{Z}$. We conclude that $\pi_1(W) = \pi_1(W_1) = \mathbb{Z}/2\mathbb{Z}$ generated by a circle in $(\mathbb{R}^3 - C_\partial) \times 0$ linking C_∂ .

Having produced $W = W_+ \subset H_+$ above, we mirror the construction to obtain $W_- \subset H_- = \mathbb{R}^3 \times (-\infty, 0]$ whose complement is a Moebius band. Gluing H_+ to H_- along $\mathbb{R}^3 \times 0$ gives \mathbb{R}^4 which contains $M \times [-2, 2]$: let $\iota: M \times [-2, 2] \hookrightarrow \mathbb{R}^4$ be the embedding. The boundary of $M \times [-2, 2]$ is a non-orientable surface of Euler characteristic 0, hence a Klein bottle. Let $K = \iota(\partial(M \times [-2, 2]))$. By construction $H_\pm - K = W_\pm$ and again Van Kampen's theorem gives $\pi_1(\mathbb{R}^4 - K) = \mathbb{Z}/2\mathbb{Z}$.

This K is certainly not smooth but by rounding corners we can make it smooth. A direct approach in this example is to find a C^∞ function $f: M \rightarrow [\frac{1}{2}, 1]$ which is $\frac{1}{2}$ on a neighborhood of ∂M and 1 on a neighborhood of the core. Let $F: M \times [-2, 2] \rightarrow M \times [-2, 2]$ be defined by $F(m, t) = (m, f(m) \cdot t)$. Let $W = F(M \times [-2, 2]) \subset M \times \mathbb{R}$ be the image of F . The boundary of W is still a Klein bottle and W is a C^∞ submanifold of $M \times \mathbb{R}$. Then the composition $\bar{\iota}: W \subset M \times \mathbb{R} \xrightarrow{\iota} \mathbb{R}^4$ is C^∞ . Hence $K_1 = \bar{\iota}(\partial W)$ is a smoothly embedded Klein bottle, bounding a smoothly embedded 3-manifold with $\pi_1(\mathbb{R}^4 - K_1) = \mathbb{Z}/2\mathbb{Z}$.

3 Construction of bounding manifolds

In this section we prove Theorem 1.9 from the introduction. The main point is Lemma 3.6, which uses surgery to annihilate elements in the fundamental

group of the bounding manifold. We will use the following general set up.

Hypotheses 3.1. *Let M be an n -manifold with $(n - 1)$ -submanifold $A \subset M$ and closed subset D . The triple (M, D, A) satisfies Hypothesis 3.1 if the following hold.*

- (a) M and A are orientable.
- (b) D has finitely many components and is nicely bounded by A . In particular D is the closure of an $(n - 2)$ -submanifold D_0 with singular set $\Sigma = D - D_0$ of codimension ≥ 4 in M .
- (c) Each component of A has finitely generated fundamental group.
- (d) For each $x_0 \in D - \Sigma$, the map $\pi_1(D - \Sigma, x_0) \rightarrow \pi_1(\overline{A} - \Sigma, x_0)$ is onto.

Remarks 3.2. Since $\overline{A} - \Sigma$ is a manifold with boundary, $A \subset \overline{A} - \Sigma$ is a homotopy equivalence. Since Σ has codimension 2, $\pi_1(D - \Sigma, x_0) \rightarrow \pi_1(D, x_0)$ is onto and $\pi_1(M - \Sigma, x_0) \rightarrow \pi_1(M, x_0)$ is an isomorphism. If M is connected, so is $M - \overline{A}$.

We will need the following definition.

Definition 3.3. If A and B are submanifolds of M , we say that A and B are equal at ∞ provided there exists a compact set $C \subset M$ such that $A - C$ and $B - C$ are equal as subsets of $M - C$.

Theorem 3.4. *If the triple (\mathbb{R}^n, D, A) satisfies 3.1 and $n \geq 6$, then there exists a submanifold $B \subset \mathbb{R}^n$ such that (\mathbb{R}^n, D, B) also satisfies 3.1, B is equal to A at ∞ and*

$$\pi_1(\mathbb{R}^n - \overline{B}) = 0.$$

Proof. We reduce the proof to the two lemmas which follow. Remove the singular set Σ to obtain the new triple $(M, D_0, A) = (\mathbb{R}^n - \Sigma, D - \Sigma, A)$, which clearly satisfies 3.1. Lemma 3.6 allows us to replace A with a new submanifold A' which is equal to A at infinity and such that $\pi_1(A')$ requires one fewer generators than $\pi_1(A)$ (by which we mean the free product of $\pi_1(A_i, x_i)$, where one basepoint x_i is chosen from each connected component A_i , as A may be disconnected). Because A' and A are equal at infinity, the new triple (M, D_0, A') satisfies 3.1 and $\mathbb{R}^n - \overline{A'} = M - \overline{A'}$. Note the closures

take place in different sets: the left-hand closure is equal to the right-hand closure union Σ . Repeated application reduces to the case that $\pi_1(A)$ is trivial, at which point Lemma 3.5 shows that $\pi_1(M - \overline{A}) = 0$. \square

Lemma 3.5. *Suppose that (M, D_0, B) satisfies 3.1 with $\Sigma = \emptyset$, $\pi_1(M) = 0$ and $\pi_1(B) = 0$. Then $\pi_1(M - \overline{B}) = 0$.*

Proof. Let $e: S^1 \rightarrow M - \overline{B}$ represent an element of $\pi_1(M - \overline{B})$. It suffices to show that e extends to a map of the disk $D^2 \rightarrow M - \overline{B}$.

By hypothesis we can extend e to a map $e: D^2 \rightarrow M$ and we may assume that e is an embedding since $n \geq 6$. The set \overline{B} is a manifold with boundary D_0 , so we may further assume that $e: D^2 \rightarrow M$ is transverse to \overline{B} . Now $D^2 \cap e^{-1}\overline{B}$ is a closed, compact 1-manifold, hence a disjoint union of circles and closed arcs in the interior of D^2 . The transverse condition also guarantees that $D^2 \cap e^{-1}B$ consists of circles and open arcs and $D^2 \cap e^{-1}(D_0)$ consists entirely of points which are the end points of the arcs: no image of a circle is tangent to D_0 .

Each circle in D^2 divides the disk into two components, an inside (homeomorphic to a disk) and an outside (homeomorphic to an annulus). Pick an outer-most circle $C \subset D^2$ from $e^{-1}(B)$, not contained on the inside of any other circle of $e^{-1}(B)$. The embedded circle $e(C)$ in B is null homotopic in B by hypothesis. Since $\dim B \geq 5$ and $e(D^2) \cap B$ consists of a finite number of open arcs and circles, this circle bounds an embedded disk $\Delta \subset B$ such that $\Delta \cap e(D^2) = e(C)$.

The normal bundle ν of B in M is a real line bundle which restricted to Δ is trivial. The normal bundle to C in D^2 is also trivial and picking the direction into the outside annulus gives a trivialization of ν restricted to $e(C) = \partial\Delta$. This trivialization extends across Δ and it is now possible to cut out a neighborhood of the inside of C and glue in a copy of Δ pushed off of B (in the direction of ν) so that the new embedded disk has fewer circles. Since there are only finitely many circles, continuing this process will result in a new embedding $e: D^2 \rightarrow M$ for which $D^2 \cap e^{-1}(\overline{B})$ consists entirely of arcs in the interior of D^2 .

Let $\alpha \in \overline{B}$ be the image of an arc, so the interior of the arc is in B and the two endpoints are in D_0 . Both endpoints are in the same component of B and so by 3.1(b), they are in the same component of $D_0 = \partial\overline{B}$, hence they may be joined by an arc in $\partial\overline{B}$. The resulting circle in \overline{B} is null homotopic because $B \subset \overline{B}$ is a homotopy equivalence. Since $n \geq 6$ there is an embedded disk $\Delta \subset \overline{B}$ so that $\partial\Delta \cap \partial\overline{B}$ is an arc in $\partial\overline{B}$ and $\partial\Delta \cap B$ is the open interior

of the arc α . We can further assume that $\Delta \cap e(D^2) = \alpha$, so that no part of $e(D^2)$ goes near Δ except for the arc under discussion.

Enlarge Δ to a disk Δ_1 with Δ in its interior so that $\Delta_1 \cap \partial\bar{B}$ is a single arc, which must extend the arc $\partial D \cap \partial\bar{B}$. This arc divides Δ_1 into two pieces, one in B and one in $M - \bar{B}$.

The normal bundle to α in $e(D^2)$ is trivial, so pick a trivialization. By transversality this picks out a nonvanishing section of ν restricted to α , which extends to a section of ν restricted to $\Delta_1 \cap \bar{B}$. Viewing this as a section of the normal bundle to $\Delta_1 \cap \bar{B}$ in M , we may extend it to obtain a section τ of the normal bundle μ of Δ_1 in M which is normal to \bar{B} along the common intersection.

Now choose an isotopy of Δ_1 under which Δ ends up in the piece of Δ_1 in $M - \bar{B}$. Extend this to an isotopy of M , using the section τ of μ along Δ_1 to make the movement locally parallel to \bar{B} . When we apply this isotopy to $e(D^2)$, we remove the arc α and create no further intersections, thus the new embedding of the disk has the same boundary and intersects \bar{B} in no circles and one fewer arc. Continue until there are no arcs (or circles). \square

Lemma 3.6. *Suppose that (M, D_0, A) satisfies 3.1 with $\Sigma = \emptyset$ and $\pi_1(A)$ is generated by m elements. Then there is a submanifold $B \subset M$ such that (M, D_0, B) satisfies 3.1, A and B are equal at ∞ and $\pi_1(B)$ is generated by at most $m - 1$ elements.*

Proof. Let $g_1, g_2, \dots, g_m \in \pi_1(A)$ be generators. Since $\dim A \geq 5$, the element g_1 can be represented by an embedded circle γ on some component of A .

The normal bundle ν to \bar{A} in M is a trivial line bundle. Use this trivialization to embed an annulus in M intersecting \bar{A} precisely in γ . The embedded circle γ' at the other end of the annulus is null homotopic in M and extends to an embedded disk $\Delta \subset M$ because M has dimension ≥ 6 . Clearly $\Delta \cap \bar{A}$ contains γ by construction, but it may contain more. We may assume that Δ meets \bar{A} transversely away from γ , hence $\Delta \cap \bar{A}$ consists of γ along with a finite number of circles and arcs in the interior of Δ .

This time we deal with the arcs first. Let α be an arc in $\Delta \cap \bar{A}$. It lies in a component of \bar{A} , so both endpoints are in a component of $D_0 = \partial\bar{A}$ and can be connected by an arc τ in D_0 . Choosing one endpoint x_0 as basepoint, condition 3.1(d) gives a loop ρ based at x_0 in D_0 such that $\rho = \alpha \cup \tau$ in $\pi_1(\bar{A}, x_0)$. Replacing τ with $\alpha' = \tau\rho^{-1}$ gives an arc in $\partial\bar{A}$ for which the embedded circle $\theta = \alpha \cup \alpha'$ is null homotopic in \bar{A} . We may further assume that α' intersects the finite set $\partial\bar{A} \cap \Delta$ only in the two endpoints of α .

Consider the open annulus $X = \Delta - \gamma - \alpha$: $\Delta - \gamma$ is an open 2-disk and removing the arc γ is topologically the same as removing a point. Then $X \cap \bar{A}$ is a finite union of circles and closed arcs so θ is null homotopic in $\bar{A} - (X \cap \bar{A})$ (we have removed all the circles and arcs *except* γ and α). Since the dimension of \bar{A} is ≥ 5 , θ bounds an embedded disk $\Gamma \subset \bar{A}$ such that $\Gamma \cap \Delta = \alpha$.

Just as in the last part of the proof of Lemma 3.5, we extend Γ to a larger embedded disk Γ' such that $\Gamma' \cap \Delta$ is an arc containing α . There is a neighborhood N of Γ' such that $N \cap \Delta$ is the original annulus. The usual Whitney trick isotopy [13, § 7.3] produces a new embedded disk Δ' so that $\Delta' \cap \bar{A} = (\Delta \cap \bar{A}) - \alpha$. Repeating this process for each arc, we may assume that our disk Δ intersects \bar{A} in a finite number of circles: all arcs are gone.

So far we have not altered A , but to get rid of the circles we will. Begin by locating an inner-most circle γ' , one whose interior contains no other circles. Then γ' bounds an embedded disk Δ' such that $\Delta' \cap \bar{A} = \gamma'$.

We perform an operation known as *adding a 2-handle* [13, §2.4 and §5.4]. Specifically, let ν be the normal bundle to γ' in A . By transversality this is also the restriction of ν' , the normal bundle to Δ' in M . Since ν' is a bundle on a disk, it is trivial so fix a trivialization. This trivializes ν . Form the union $X = A \cup E(\nu')$ where $E(\nu')$ is the closed disk bundle associated to ν' . If x_0 is a basepoint in $A \cap E(\nu') = E(\nu) \cong \gamma' \times \bar{D}^{n-2}$, Van Kampen's theorem produces a pushout diagram

$$\begin{array}{ccc} \pi_1(E(\nu), x_0) = \langle \gamma' \rangle & \rightarrow & \pi_1(E(\nu'), x_0) = \{1\} \\ \downarrow & & \downarrow \\ \pi_1(A, x_0) & \rightarrow & \pi_1(X, x_0) \end{array}$$

from which we see that $\pi_1(X, x_0) \cong \pi_1(A, x_0) / \langle \gamma' \rangle$, where $\langle \gamma' \rangle$ is the normal closure of γ' .

Now form B_1 from X by removing the open disk bundle. The result is not a differentiable manifold as there are corners where A meets $\partial E(\nu')$, but let us note several properties of B_1 . Since $\Delta \cap B_1 = \Delta \cap A - \gamma$ by construction, the new intersection has one less circle. Since the construction takes place within a compact set, A and B_1 are equal at ∞ and it is clear that (M, D_0, B_1) satisfies 3.1 because $A = B_1$ near D_0 .

Now $\pi_1(B_1) \rightarrow \pi_1(X)$ is an isomorphism (by Van Kampen's theorem, for example) and we write $\pi_1(B_1) \cong \pi_1(A) / \langle \gamma' \rangle$ so that $\pi_1(B_1)$ is still generated

by $g_1 \dots g_m$. The commutative diagram

$$\begin{array}{ccc} \pi_1(D_0) & \rightarrow & \pi_1(\overline{B_1}) \\ \downarrow & & \downarrow \\ \pi_1(\overline{A}) & \rightarrow & \pi_1(\overline{X}) \end{array}$$

shows that the top map is onto. Indeed, the left vertical map is onto by hypothesis, the bottom map is onto because $\pi_1(A) \rightarrow \pi_1(X)$ is (the inclusions $A \subset \overline{A}$ and $X \subset \overline{X}$ are homotopy equivalences) and the right vertical map is an isomorphism because $\pi_1(B_1) \rightarrow \pi_1(X)$ is.

Finally we round the corners of B_1 to obtain a differentiable manifold. This process is a homeomorphism (though not a diffeomorphism) so we retain all the properties of B_1 above. By the same process we remove an inner-most circle of $B_1 \cap \Delta$ to get B_2 and so on until we find a B_k such that $B_k \cap \Delta = \gamma$, (M, D_0, B_k) satisfy 3.1 and $B_k = A$ at infinity.

Now we can trade the 2-handle Δ to get B where A and B are equal at ∞ , (M, D_0, B) satisfy 3.1 and $\pi_1(A)$ maps onto $\pi_1(B)$ with γ definitely in the kernel, hence $\pi_1(B)$ is generated by g_2, \dots, g_m . \square

References

- [1] H. Bass, E. Connell and D. Wright, The Jacobian conjecture: reduction of degree and formal expansion of the inverse, *Bull. A.M.S.* **7** (1982) 287-330.
- [2] A. Dimca, *Singularities and topology of hypersurfaces*, Universitext, Springer-Verlag, New York 1992.
- [3] A. Dold, *Lectures on Algebraic Topology*, Grundlehren der mathematischen Wissenschaften **200**, Springer-Verlag, 1972.
- [4] A. van den Essen, *Polynomial Automorphisms and the Jacobian Conjecture*, Progress in Mathematics **190**, Birkhäuser, 2000.
- [5] Hà Huy Vui and Lê Dũng Tráng, Sur la topologie des polynômes complexes, *Acta Math. Vietnamica* **9** (1984) 21–32.
- [6] Z. Jelonek, The set of points at which a polynomial map is not proper, *Ann. Polon. Math.* **58** (1993) 259–266.

- [7] V. Kulikov, Generalized and local Jacobian problems, Russian Acad. Sci. Izv. Math. **41** (1993), no. 2, 351–365.
- [8] D. Mumford, *Algebraic Geometry I: Complex Projective Varieties*, Grundlehren der mathematischen Wissenschaften **221**, Springer-Verlag, 1976.
- [9] S. Nollet and F. Xavier, Global Inversion via the Palais-Smale Condition, Discrete and Continuous Dynamical Systems A **8** (2002) 17–28.
- [10] S. Nollet and F. Xavier, Holomorphic Injectivity and the Hopf map, Geom. and Funct. Anal. **14** (2004) 1339–1351.
- [11] S. Nollet and F. Xavier, On Kulikov’s Problem, to appear in Archiv der Mathematik.
- [12] A. Parusiński, On the bifurcation set of complex polynomial with isolated singularities at infinity, Compositio Math. **97** (1995) 369–384.
- [13] A. Ranicki, Algebraic and geometric surgery, Clarendon Press, Oxford, 2002
- [14] A. Schinzel, *Polynomials with Special Regard to Reducibility*, Encyclopedia of Mathematics 77, Cambridge University Press, 2000.
- [15] E. Spanier, *Algebraic Topology*, McGraw-Hill, Inc. 1966.
- [16] M. Tibăr, *Polynomials and Vanishing Cycles*, Cambridge Tracts in Mathematics 170, Cambridge University Press, 2007.
- [17] J. Verdier, Stratifications de Whitney et theoreme de Bertini-Sard, Inventiones Math. **36** (1976) 295–312.
- [18] F. Xavier, Rigidity of the identity, to appear in Communications in Contemporary Mathematics.