Journal *of* Singularities
Volume 23 (2021), 205-235

received: 22 April 2021
in revised form: 16 November 2021
DOI: 10.5427/jsing.2021.23l

WHAT IS THE DEGREE OF A SMOOTH HYPERSURFACE?

ANTONIO LERARIO AND MICHELE STECCONI

Abstract. We deal with the problem of the algebraic approximation of type-W singularities of smooth functions on a closed n-disk, namely the set of points in the disk where the jet extension of the function meets a given semialgebraic subset W of the jet space; examples of sets arising in this way are the zero set of a function, or the set of its critical points.

We prove that the type-W singularity defined by a smooth function, satisfying a transversality condition, is isotopic to the one defined by a polynomial whose degree is explicitly bounded in terms of the distance of the original function from the set of functions which do not satisfy the transversality condition. Ultimately, the bound depends on the second derivatives of the jet of the function. The estimate on the degree of the approximating polynomial implies an estimate on the Betti numbers of the original singularity. Using more refined and recently developed tools, we prove a second bound, directly for Betti numbers, that is of lower order than that implied by the previous one, in that it depends only on the first derivatives of the jet.

These results specialize to the case of a smooth compact hypersurface, resulting in a control of the minimal degree of its algebraic realization (from which the title of the paper) and of its Betti numbers. As a corollary we prove an upper bound on the number of isotopy classes of compact hypersurfaces satisfying a certain quantitative transversality condition. Moreover, we show that in this case the second estimate on the Betti numbers is asymptotically sharp. Finally, we relate the two estimates - the one for the degree of a polynomial realization and the one for the Betti numbers - with the geometric data of the hypersurface, independent from its defining equation, showing that the bounds can be given in terms of the reach and the diameter.

Overview

Let D be a disk in $\mathbb{R}^n$ and $f \in C^{r+2}(D, \mathbb{R}^k)$. We deal with the problem of the algebraic approximation of the set $j^r f^{-1}(W)$ consisting of the points in the disk D where the r-th jet extension of f meets a given semialgebraic set $W \subset J^r(D, \mathbb{R}^k)$. We call such sets *type–W singularities*; examples of sets arising in this way are the zero set of f, or the set of its critical points.

Under some transversality conditions, we prove that f can be approximated with a polynomial map $p : D \to \mathbb{R}^k$ such that the corresponding singularity is diffeomorphic to the original one, and such that the degree of this polynomial map can be controlled by the C^{r+2} data of f. More precisely, denoting by $\Delta_W \subset C^{r+1}(D, \mathbb{R}^k)$ the set of maps whose r-th jet extension *is not* transverse to W, we show that there exists a polynomial p such that:

$$j^r p^{-1}(W) \sim j^r f^{-1}(W) \quad \text{and} \quad \deg(p) \leq O\left(\frac{\|f\|_{C^{r+2}(D,\mathbb{R}^k)}}{\mathrm{dist}_{C^{r+1}}(f, \Delta_W)}\right). \tag{0.1}$$

(Here "$\sim$" means "ambient isotopic" and the implied constant depends on the size of the disk). The estimate on the degree of p implies an estimate on the Betti numbers of the singularity, however, using more refined tools introduced in [Ste21], we prove a similar estimate, but involving only the C^{r+1} data of f.

For a given $W \subset J^r(D,\mathbb{R}^k)$, we introduce a notion of C^ℓ condition number $\kappa_W^{(\ell)}(f,D)$ of a type–W singularity of a map $f \in C^\ell(D,\mathbb{R}^k)$, $\ell \geq r+1$, and we show that our estimates, as (0.1), can be equivalently stated using this notion.

These results specialize to the case of zero sets of $f \in C^2(D,\mathbb{R})$, and give a way to approximate a smooth hypersurface defined by the equation $f = 0$ with an algebraic one, with controlled degree (from which the title of the paper). In this case we deal both with the approximation of $f = 0$ inside the disk D and with its global approximation in $\mathbb{R}^n$. As a corollary we prove an upper bound on the number of diffeomorphism classes of compact hypersurfaces with a bounded condition number.

Moreover, we also deal with the more basic problem of producing an actual equation $f = 0$ for a given compact hypersurface $Z \subset \mathbb{R}^n$, and we control the condition number of this equation with the geometric data of Z (its reach and its diameter). We prove that the Betti numbers of the hypersurface can be estimated with the $\kappa^{(1)}$ condition number of the defining equation and we show that the order of this estimate is sharp.

In particular, combining these results, we show that a compact hypersurface $Z \subset D \subset \mathbb{R}^n$ with positive reach $\rho(Z) > 0$ is isotopic to the zero set in D of a polynomial p of degree

$$\deg(p) \leq c(D) \cdot 2\left(1 + \frac{1}{\rho(Z)} + \frac{5n}{\rho(Z)^2}\right),$$

where $c(D) > 0$ is a constant depending on the size of the disk D (and in particular on the diameter of Z).

1. Introduction

1.1. **On the constants of the paper.** Many constants will appear in the paper. The objects on which they depend will be indicated according to the following convention. If $D \subset \mathbb{R}^n$ is a disk, then by writing $c = c(D)$ we mean that the constant c depends on n, on the center and on the diameter of D. Similarly, given a subset $W \subset J^r(D,\mathbb{R}^k)$ of the jet space (see the next paragraph) the notation $c = c(W)$ implies that the constant c depends also on n, r, k, D as well as on all the parameters that are mentioned in the definition of W.

We will keep the notation $c(D)$, although often the dependence on the position of the disk can be dropped, as this in that case will be obvious from the nature of the statements (see for instance subsection 1.7). The only exception is Theorem 1, in which case we make an explicit comment.

1.2. **Polynomial approximation of singularities of smooth maps.** In this paper we deal with the following problem: given a smooth (i.e. sufficiently regular) function $f : D \to \mathbb{R}$ defined on a disk $D \subseteq \mathbb{R}^n$ and whose zero set $Z(f)$ is a smooth compact manifold, what is the smallest degree of a polynomial p whose zero set $Z(p)$ is diffeomorphic to $Z(f)$?

More generally, we will consider the problem of the polynomial approximation of nondegenerate singularities of smooth maps: given a closed and stratified subset W of the jet space[1] $J^r(D,\mathbb{R}^k)$ and a smooth map $f : D \to \mathbb{R}^k$ transverse to all the strata of W, what is the smallest degree of a polynomial map $p : D \to \mathbb{R}^k$ such that the two pairs $(D, j^r f^{-1}(W))$ and $(D, j^r p^{-1}(W))$ are diffeomorphic?

Besides the case of hypersurfaces, which corresponds to the choice of

$$W = D \times \{0\} \subset J^0(D,\mathbb{R}) = D \times \mathbb{R},$$

[1]Here $J^r(D,\mathbb{R}^k)$ denotes the r-th jet bundle of maps $f : D \to \mathbb{R}^k$ and, given $f : D \to \mathbb{R}^k$ of class C^r, $j^r f : D \to \mathbb{R}^k$ denotes its r-th jet extension (see 2.1 for more details).

other examples of special interests covered by this framework are: systems of smooth inequalities, corresponding to the case of $W = D \times C \subset J^0(D, \mathbb{R}^k) = D \times \mathbb{R}^k$, where C is a closed polyhedral cone; critical points of a function $f : D \to \mathbb{R}$, corresponding to the choice

$$W = D \times \mathbb{R} \times \{0\} \subset J^1(D, \mathbb{R}) = D \times \mathbb{R} \times \mathbb{R}^n;$$

critical points of a smooth map $f : D \to \mathbb{R}^n$, with the choice

$$W = D \times \mathbb{R}^k \times \{\det = 0\} \subset J^1(D, \mathbb{R}^n) = D \times \mathbb{R}^n \times \mathbb{R}^{n\times n}.$$

In general we will call the set $j^r f^{-1}(W)$ a *type–W singularity*.

In order to answer the above questions, we shall adopt first a geometric approach. We make the assumption, verified in all cases of practical interest, that $W \subseteq J^r(D, \mathbb{R}^k)$ is a semialgebraic set. Note that this does not mean that the singularity is semialgebraic, but rather that it is given by semialgebraic conditions on the derivatives of a *smooth* map. Given such W, we denote by $\Delta_W \subset C^{r+1}(D, \mathbb{R}^k)$ the set:

$$\Delta_W = \left\{ f \in C^{r+1}(D, \mathbb{R}^k) \text{ such that } j^r f : D \to J^r(D, \mathbb{R}^k) \textit{ is not } \text{transverse to } W \right\}.$$

Here transversality means with respect to all the strata of a given fixed Whitney stratification of W, both for $j^r f$ and $(j^r f)|_{\partial D}$. The set Δ_W acts as a discriminant for our problem, and the jet of a map $f \in C^{r+1}(D, \mathbb{R}^k)$ is transverse to W if and only if $\mathrm{dist}_{C^{r+1}}(f, \Delta_W) > 0$. When both $j^r f$ and $(j^r f)|_{\partial D}$ are transverse to W, we will simply write $j^r f \pitchfork W$. In this case the set $j^r f^{-1}(W) \subseteq D$ is a Whitney stratified subcomplex[2] of the disk and we will refer to it as a *nondegenerate singularity*; for instance, if W is a smooth submanifold, then so is $j^r f^{-1}(W)$.

Given subcomplexes K_0 and K_1 of the disk, we will say that the two pairs (D, K_0) and (D, K_1) are *isotopic*, and write $(D, K_0) \sim (D, K_1)$, if there exists a continuous family of diffeomorphisms $\varphi_t : D \to D$, with $t \in [0, 1]$, such that $\varphi_0 = \mathrm{id}_D$ and $\varphi_1(K_0) = K_1$. With this notation, our first result is the following.

Theorem 1. *Let $W \subseteq J^r(D, \mathbb{R}^k)$ be closed and semialgebraic. For every $f \in C^{r+2}(D, \mathbb{R}^k)$ with $j^r f \pitchfork W$ there exists a polynomial map $p = (p_1, \ldots, p_k)$ with each $p_i \in \mathbb{R}[x_1, \ldots, x_n]$ with*

$$\deg(p_i) \leq c_1(r, D) \max \left\{ r + 1, \frac{\|f\|_{C^{r+2}(D,\mathbb{R}^k)}}{\mathrm{dist}_{C^{r+1}}(f, \Delta_W)} \right\} \tag{1.1}$$

and such that:

$$(D, j^r f^{-1}(W)) \sim (D, j^r p^{-1}(W)). \tag{1.2}$$

(Here $c_1(r, D)$ is a constant depending only on the size of the disk D and on r, n).

The transversality assumption in the previous statement is necessary to prevent pathological situations. For instance, every closed set in D is the zero set of a smooth function, thus there exists a smooth f such that $f^{-1}(0)$ equals the Cantor set; however in this case the pair $(D, f^{-1}(0))$ cannot be diffeomorphic to a pair $(D, p^{-1}(0))$ with p a polynomial.

From the previous result, using standard techniques from real algebraic geometry, one can immediately produce an upper bound on the topological complexity of a nondegenerate singularity, measured by the sum of its Betti numbers.

Corollary 2. *Given $W \subseteq J^r(D, \mathbb{R}^k)$ closed and semialgebraic there exists $c_2(W) > 0$ such that for every $f \in C^{r+2}(D, \mathbb{R}^k)$ with $j^r f \pitchfork W$ we have:*

$$b(j^r f^{-1}(W)) \leq c_2(W) \cdot \max \left\{ r + 1, \frac{\|f\|_{C^{r+2}(D,\mathbb{R}^k)}}{\mathrm{dist}_{C^{r+1}}(f, \Delta_W)} \right\}^n. \tag{1.3}$$

[2] Meaning that the closed disk admits a Whitney stratification such that all the strata intersecting W are contained in W and, together, form a Whitney stratification of W.

1.3. C^{r+1} **bound for Betti numbers.** Note that in (1.3), the C^{r+2} norm of the function f appears, even if the transversality condition would require only looking at the $(r+1)$-th jet of f. This comes from the fact that the error estimate for the polynomial approximation of f in the C^{r+1} topology involves its C^{r+2} norm (this condition can be slightly relaxed when working with Lipschitz derivatives). However, if one is only interested in bounding the topology of $j^r f^{-1}(W)$, it turns out that an estimate on the C^{r+1} norm of f suffices, at least in the case W is smooth, as we will prove in Theorem 3 below. In the case of hypersurfaces, we will discuss this in more details in Section 1.7.

Theorem 3. *Let $W \subseteq J^r(D, \mathbb{R}^k)$ be a smooth, compact and semialgebraic submanifold with the property that $W \pitchfork J^r_z(D, \mathbb{R}^k)$ for all $z \in D$. There exists a constant $c_3(W) > 0$ such that for every $f \in C^{r+1}(D, \mathbb{R}^k)$ with $j^r f \pitchfork W$ and $j^r f^{-1}(W) \cap \partial D = \emptyset$ we have:*

$$b(j^r f^{-1}(W)) \le c_3(W) \cdot \max\left\{ r, \frac{\|f\|_{C^{r+1}(D,\mathbb{R}^k)}}{\hat{\delta}_W(f, D)} \right\}^n . \tag{1.4}$$

The quantity $\hat{\delta}_W(f, D)$ is defined in Section 3.3. In analogy with the C^{r+1} distance from the discriminant appearing in the estimate (1.3), it depends only on W and on the $(r+1)$-th jet of the map f. In particular we have $\hat{\delta}_W(f, D) > 0$ if and only if $j^r f \pitchfork W$ and $j^r f^{-1}(W) \cap \partial D = \emptyset$. Notice the different type of boundary condition: the transversality assumption "$(j^r f)|_{\partial D} \pitchfork W$" in Theorem 2 is replaced here with the stronger requirement that there is no point $z \in \partial D$ such that $j^r f(z) \in W$.

What is interesting about Theorem 3 is that the right hand side of (1.4) does not depend on the size (nor even on the existence) of the derivatives of f of order higher than $r + 1$. This is obtained by means of Theorem 21, a recent result from [LS19] that says that given a function f such that $j^r f \pitchfork W$, the Betti numbers of a singularity $(j^r g)^{-1}(W)$ are bigger than those of $(j^r f)^{-1}(W)$, provided that g is close enough to f in the C^r topology. In this way we can control directly the Betti numbers of $j^r g^{-1}(W)$ and relax the requirements on the needed polynomial approximation, since we can bypass the isotopy condition (1.2) of Theorem 1.

1.4. **The condition number of a singularity.** It is often more convenient to substitute the distance from the discriminant with a more explicit quantity $\delta_W(f, D)$, defined as follows. Denote by $\Sigma_{W,z} \subset J^{r+1}_z(D, \mathbb{R}^k)$ the set of all possible jets of maps which are not transverse to W at z. This is a closed and semialgebraic subset of $J^{r+1}_z(D, \mathbb{R}^k)$ and we define the number:

$$\delta_W(f, D) = \min_{z \in D} \operatorname{dist}(j^{r+1} f(z), \Sigma_{W,z}).$$

Observe that when $z \in \partial D$ the set $\Sigma_{W,z}$ consists of "two pieces": in fact if $z \in \operatorname{int}(D)$ we only have to consider the transversality of $j^r f$, while if $z \in \partial D$ one needs also to take into account the transversality of $(j^r f)|_{\partial D}$, which involves a restriction of the jet, as a multilinear map, to $T_z(\partial D) \simeq \{z\}^\perp$.

The transversality of $j^r f$ in most cases has a simple geometric interpretation, and for $z \in \operatorname{int}(D)$ the set $\Sigma_{W,z}$ can be easily described. For example: in the case of hypersurfaces, $\Sigma_{W,z} = \{z\} \times \{0\} \times \{0\} \subset D \times \mathbb{R} \times \mathbb{R}^n$ and $\operatorname{dist}(j^1 f(z), \Sigma_{W,z}) = \left(|f(z)|^2 + \|\nabla f(z)\|^2\right)^{1/2}$; in the case of critical points of a smooth function $f : D \to \mathbb{R}$, we have that

$$\Sigma_{W,z} = \{z\} \times \mathbb{R} \times \{0\} \times \{\det = 0\} \subset D \times \mathbb{R} \times \mathbb{R}^n \times \operatorname{Sym}(n, \mathbb{R})$$

and $\operatorname{dist}(j^2 f(z), \Sigma_{W,z}) = \left(\|\nabla f(z)\|^2 + (\sigma_1(\operatorname{He}(f)(z)))^2\right)^{1/2}$, where $\operatorname{He}(f)(z)$ denotes the Hessian of f at z and σ_1 denotes the smallest singular value.

The quantity $\delta_W(f,D)$ vanishes exactly when $f \in \Delta_W$, and $j^r f \pitchfork W$ if and only if $\delta_W(f,D) > 0$ (Lemma 19 below). The key property, that allows to translate estimates in terms of this quantity into the above geometric framework, is the fact that the distance to the discriminant and the quantity $\delta_W(f,D)$ have the same order of magnitude. In fact, in Proposition 15, we show that there exists a constant $C = C(n,k,r) > 0$ such that:

$$\delta_W(f,D) \le \mathrm{dist}_{C^{r+1}}(f,\Sigma_W) \le C \cdot \delta_W(f,D). \tag{1.5}$$

The first inequality follows from a characterization of the set of functions not transverse to W at z as those with jet belonging to $\Sigma_{W,z}$; the second inequality uses a classical result of Whitney on the norm of a map with prescribed jet at one point.

Remark 4. The constant $C > 0$ in (1.5) depends also on the choice of the pointwise norm on the jet bundles. In general we have $C > 1$, as shown in Example 16.

Using this notation, we introduce the notion of condition number for a map with respect to a closed, stratified $W \subseteq J^r(D,\mathbb{R}^k)$ (Definition 18). Given $f \in C^\ell(D,\mathbb{R}^k)$ with $\ell \ge r+1$, we set:

$$\kappa_W^{(\ell)}(f,D) = \frac{\|f\|_{C^\ell(D,\mathbb{R}^k)}}{\delta_W(f,D)} \ge \frac{\|f\|_{C^\ell(D,\mathbb{R}^k)}}{\mathrm{dist}_{C^{r+1}}(f,\Delta_W)}. \tag{1.6}$$

Notice that the numerator in (1.6) depends on the topology we are considering on the space of functions, while the denominator does not. Transversality is synonymous of bounded (i.e. non-infinite) condition number. Because of the inequality on the right hand side of (1.6), the estimates (1.1) and (1.3) can also be stated using $\kappa_W^{(r+2)}(f,D)$.

1.5. **The condition number of a defining equation for a hypersurface.** Given a smooth and compact hypersurface Z contained in $\mathrm{int}(D)$, the existence of a polynomial p such that $(D,Z) \sim (D,Z(p))$ is guaranteed by a classical result of Seifert [Sei36]. This problem is also known as the "algebraic approximation problem", and has several generalizations, culminating with the celebrated Nash-Tognoli Theorem [Nas52, Tog73]: every smooth and compact manifold M is diffeomorphic to an algebraic set in $\mathbb{R}^n$. (Our problem concerns with the special case M is already a hypersurface in $\mathbb{R}^n$).

The proof of Seifert's Theorem (clearly explained in [Kol17]) consists in first realizing Z as the regular zero set of a function $f : \mathbb{R}^n \to \mathbb{R}$ and then approximating f on the disk D in the C^1 topology with a polynomial. In the very first step, one uses the fact that every smooth and compact hypersurface in $\mathbb{R}^n$ is the zero set of a smooth function. This result is well-known, see for instance [DFN85, Theorem 7.2.3], and a natural question in our framework is to produce an estimate on the condition number of a defining equation for Z, in terms of some metric data of the embedding $Z \hookrightarrow \mathbb{R}^n$.

To this end, given $Z \subset \mathbb{R}^n$ of class C^1 we define the *reach* of Z as:

$$\rho(Z) = \sup_{r>0} \{\mathrm{dist}(x,Z) < r \implies \exists! z \in Z \,|\, \mathrm{dist}(x,Z) = \mathrm{dist}(x,z)\}.$$

The reach of a C^1 manifold doesn't need to be positive, as shown in [KP81, Example 4], where for every $0 < \epsilon < 2$ an example of a $C^{2-\epsilon}$ compact curve with zero reach in $\mathbb{R}^2$ is constructed[3]; however $\rho(Z) > 0$ if Z is of class C^2. We prove the following result[4].

[3]Near the origin this is in fact just the set $Z \subset \mathbb{R}^2$ defined by $Z = \{y = x^{2-\epsilon}\}$

[4]In the special case of hypersurfaces, i.e. when $W = D \times \{0\}$, in the notation of the various involved quantities we omit the dependence from W in the subscripts.

Theorem 5. *Given a compact hypersurface $Z \subset \mathbb{R}^n$ of class C^1 with $\rho(Z) > 0$, there exists a C^1 function $f : \mathbb{R}^n \to \mathbb{R}$ whose zero set is Z and such that for every disk D containing Z such that* $\operatorname{dist}(Z, \partial D) > \rho(Z)$, *we have:*

$$\kappa^{(1)}(f, D) \leq 2\left(1 + \frac{1}{\rho(Z)}\right). \tag{1.7}$$

If moreover Z is of class C^2, then the function f can be chosen of class C^2 and satisfying:

$$\kappa^{(2)}(f, D) \leq 2\left(1 + \frac{1}{\rho(Z)} + \frac{5n}{\rho(Z)^2}\right). \tag{1.8}$$

1.6. **What is the degree of a smooth hypersurface?** Once a compact C^2 hypersurface $Z \subset \mathbb{R}^n$ is given, as we have seen with Theorem 5, we can produce a defining equation with C^2 condition number controlled with a function of the reach $\rho(Z) > 0$ (recall that the C^2 regularity assumption implies that the reach is nonzero). Therefore, assuming $Z = Z(f)$ with $f \in C^2$, we come back to our original question: what is the smallest degree of a polynomial p whose zero set $Z(p)$ is diffeomorphic to $Z(f)$? Here we should make a choice: whether we want to approximate Z only inside the disk D, or we want to get a global approximation in $\mathbb{R}^n$. In fact Nash-Tognoli Theorem also consists of two separate results: for a compact M, Nash first proved that M is diffeomorphic to a component of a real algebraic set, and then Tognoli proved that we can choose this algebraic set to be connected.

We deal with the problem of the global approximation in Appendix A, where we prove an explicit (but rather unpleasant) bound on the degree of the global approximating polynomial in terms of the C^2 data of the defining function. Concerning the approximation of $Z(f)$ inside the disk D, specializing Theorem 1, we see that there exists a polynomial p with:

$$\deg(p) \leq c_1(0, D) \cdot \max\left\{1, \kappa^{(2)}(f, D)\right\} \tag{1.9}$$

such that $(D, Z(f)) \sim (D, Z(p))$.

Remark 6. Recall that the constant $c_1(0, D)$ coming from Theorem 1 depends only on the dimension n and on the radius of the disk, but not on its position. The same holds for all the constants c_i appearing in the results of the present subsection and of subsection 1.7 regarding the case of hypersurfaces. As this fact is obvious from the nature of the statements, we won't mention it again.

Combining this with Theorem 5 we get the following result, which uses only the available geometry of Z.

Corollary 7. *Given a compact hypersurface $Z \subset \mathbb{R}^n$ of class C^2 and a disk D containing it such that* $\operatorname{dist}(Z, \partial D) > \rho(Z)$, *there exists a polynomial p with*

$$\deg(p) \leq c_1(0, D) \cdot 2\left(1 + \frac{1}{\rho(Z)} + \frac{5n}{\rho(Z)^2}\right)$$

such that

$$(D, Z) \sim (D, Z(p)).$$

Another interesting consequence of (1.9) is Theorem 8 below, which provides a bound on the number $\#(\kappa, D)$ of isotopy classes (and in particular on the number of diffeomorphism classes) of compact hypersurfaces $Z \subset \operatorname{int}(D)$ defined by a regular equation $Z = Z(f)$ with bounded C^2 condition number.

Theorem 8. *There exist two constants C_1, C_2 (depending on D) such that the number $\#(\kappa, D)$ of rigid isotopy classes of pairs $(D, Z(f))$ with $Z(f) \subset \mathrm{int}(D) \subset \mathbb{R}^n$ and with $\kappa^{(2)}(f, D) \le \kappa$ is bounded by:*

$$\#(\kappa, D) \le \min\left\{C_1, \kappa^{\left(C_2 \kappa^{n+1}\right)}\right\}$$

In the case of the approximation of a hypersurface Z inside the disk D, we introduce in Section 4.2 the notion of *isotopy degree* of Z in D, denoted by $\deg_{\mathrm{iso}}(Z, D)$: this is the smallest degree of a polynomial p such that $(D, Z(f)) \sim (D, Z(p))$. The idea here is that, as soon as you can move from the smooth category to the semialgebraic one, a complete new set of tools becomes available (for instance the use of Thom-Milnor for controlling the topology of the zero set) and Theorem 1 allows to make this transition quantitative.

An interesting observation is that if f is itself a polynomial of degree d, then $\deg_{\mathrm{iso}}(Z(f), D)$ might be smaller than d. In fact, for most polynomials f of degree d one has

$$\deg_{\mathrm{iso}}(Z(f), D) \le O(\sqrt{d \log d}),$$

see Remark 33.

1.7. **Semicontinuity of Betti numbers.** If we are interested in providing an upper bound for the sum of the Betti numbers of a compact C^2 hypersurface, we can in principle use Corollary 2 and get:

$$b(Z(f)) \le c_2 \max\left\{1, \kappa^{(2)}(f, D)\right\}^n,$$

for some constant $c_2 = c_2(W) > 0$. However, with some extra work, it is possible to provide a bound on the Betti numbers of $Z(f)$ only using $\kappa^{(1)}(f, D)$, i.e. only using C^1-information on f, as next Theorem shows.

Theorem 9. *Let $f \in C^1(D, \mathbb{R})$ be given such that the equation $f = 0$ is regular and all of its solutions belong to the interior of D.*

(1) There exists a constant $c_4 = c_4(D) > 0$ such that the total Betti number of $Z(f)$ is bounded by:

$$b(Z(f)) \le c_4 \cdot (\kappa^{(1)}(f, D))^n. \tag{1.10}$$

(2) There exists a bounded *sequence $\{f_m\}_{m \in \mathbb{N}} \subset C^1(D, \mathbb{R})$ with*

$$\lim_{m \to \infty} \kappa^{(1)}(f_m, D) = +\infty$$

and a constant $c_5 = c_5(D) > 0$ such that for every $m \in \mathbb{N}$ the zero set $Z(f_m) \subset \mathrm{int}(D)$ is regular and

$$b(Z(f_m)) \ge c_5 \cdot (\kappa^{(1)}(f_m, D))^n.$$

Let us comment this result. First, we observe that it is possible to deduce a result similar to (1.10) also from the work of Yomdin [Yom85], where bounds on the Betti numbers of $Z(f)$ are stated in terms of the distance from zero to the the set of critical values of f (Remark 17 below compares the two quantities). See also [Yom84] for an other result of Yomdin, concerning the Hausdorff measure of $Z(f)$.

Here (1.10) is a consequence of a semicontinuity result for the Betti numbers of the zero set of f under small perturbations in the C^0 topology (rather than in the C^1 topology, which is what one would expect to need). This is discussed in Section 1.7.

The second part of the statement shows that the bound in (1.10) is "sharp": as we get close to the discriminant Δ, the complexity of f can actually increase as the reciprocal of the distance from Δ.

Combining Theorem 9 with Theorem 5 we get the following estimate for the Betti numbers of a hypersurface in terms of its reach, if nonzero. (A similar bound can be obtained using the results from [NSW08], see Remark 34 below).

Corollary 10. *For any disk $D \subset \mathbb{R}^n$ there exists a constant $c_6 = c_6(D) > 0$ such that, for every compact hypersurface $Z \subset D$ of class C^1 with $\rho(Z) > 0$, we have:*

$$b(Z) \le c_6 \left(1 + \frac{1}{\rho(Z)}\right)^n .$$

1.8. **Related work, open questions and comments.** We conclude this Introduction by pointing out some related work and open problems.

To start with, we point out that one can formulate a more general question: given a compact manifold M, can we estimate ℓ and d such that M is diffeomorphic to an algebraic set Z of degree d in $\mathbb{R}^\ell$? Clearly, the existence of d, ℓ is guaranteed by Nash-Tognoli. Approaching this problem requires looking at the proof of Nash and Tognoli and making it quantitative – but using which data? That has to be understood. For instance, if M is a 3–manifold, the data could be some triangulation or the Heegaard decomposition. We plan to investigate this direction in a forthcoming work.

This paper, in some sense, deals with the problem of "immersed Nash-Tognoli for complete intersections". The regularity plays a crucial role for us: we essentially covers the C^2 case. The C^1 case is harder, and more fascinating. In fact, as a first step, one would need to solve the following problem: let M be a compact C^1 hypersurface in $\mathbb{R}^\ell$; find a regular smooth equation $f = 0$ for M with bounded condition number. Here is a toy model for this problem. Minimize the C^r norm of a function $f : [-1, 1] \to \mathbb{R}$ with prescribed derivatives up to order r at the origin. Here, by C^r norm we mean:

$$\|f\|_{C^r} := \max_{x\in[-1,1]} \sum_{k=0}^{r} |f^{(k)}(x)|.$$

It is a result of Whitney that $\|f\|_{C^r} = O(\|j^r f(0)\|)$. Several papers of Fefferman (see for instance [Fef10]) deal with this problem, but where the fiberwise norm is the sup-norm (this makes the computation simpler). Now, one can interpret the previous problem as follows: let $M \subset \mathbb{R}^n$ be a C^1 hypersurface and assign the one-jet of a function $f : \mathbb{R}^n \to \mathbb{R}$ along M to be: f vanishes on M and its gradient has norm 1 on M. Estimate the C^1 norm of the function f on a ball containing M. All this is related to the existence of regular neighborhoods, on which the only quantitative results that we are aware of is for the smooth case. In the C^2 case one has the reach as an estimator of the size of these regular neighborhoods, but what happens in the C^1 case? We point out that a related problem has recently been studied by Yomdin in [Yom20, Yom21b, Yom21a], in the context of *smooth rigidity.* A smooth rigidity inequality is an estimate that gives a lower bound for the $d+1$ derivatives of a function $f : D \to \mathbb{R}$ if f has a certain geometrical behaviour, forbidden for polynomials of degree d. An example of such behaviour is: the topology of the zero set of f exceeds Thom-Milnor's bound. Interestingly, [Yom21b, Theorem 5.2] uses our results and produces a lower bound for the C^2 norm of $f : D \to \mathbb{R}$, in terms of the forbidden degree d and the distance to the discriminant $\delta(f, D) > 0$. (This bound shows that the norm grows at least linearly with d.) We believe this is just the first step of a deeper connection between our approach, rigidity inequalities and extensions problem, which we hope will be investigated further.

1.9. **Acknowledgements.** The authors wish to thank Daouda Niang Diatta for many interesting discussions, and Pierre Lairez and Yosef Yomdin for their comments on the first draft of this paper. The authors also thank the anonymous referee for her/his helpful comments.

2. Preliminaries

2.1. **Jet bundles.** We assume that the reader is already familiar with notion of jet space, referring to the textbooks [Hir94, VA85] for more details. We simply recall that the jet space $J^r(D,\mathbb{R}^k)$ can be naturally identified with $D\times J^r(n,k)$, where

$$J^r(n,k)=\left(\bigoplus_{i=1}^{r}\mathbb{R}^{\binom{n+i}{i}}\right)^k$$

and $\mathbb{R}^{\binom{n+i}{i}}$ is the space of homogeneous polynomials of degree i in n variables (see [Hir94]). (The identification of a jet with an element of $J^r(n,k)$ is made through the list of the partial derivatives). We endow $J^r(n,k)$ with the standard euclidean structure, allowing to compute distances between elements in $J^r_z(D,\mathbb{R}^k)$.

Using this notation, we can define the C^r norm of $f\in C^r(D,\mathbb{R}^k)$ as

$$\|f\|_{C^r(D,\mathbb{R}^k)}=\max_{z\in D}\|j^rf(z)\|. \tag{2.1}$$

More explicitly, given $f=(f_1,\dots,f_k)\in C^r(D,\mathbb{R}^k)$ its C^r norm is:

$$\|f\|_{C^r(D,\mathbb{R}^k)}=\sup_{z\in D}\left(\sum_{i=1}^{k}\sum_{|\alpha|\le r}\left|\frac{\partial^\alpha f_i}{\partial x^\alpha}(z)\right|^2\right)^{1/2}.$$

We observe that the definition of C^r norm is sensitive to the choice of the norm in the fibers of the jet bundle. For instance, if one replaces in (2.1) the quantity $\|j^rf(z)\|$ with the sup of the value of each partial derivative of f order at most r at z, one gets an equivalent C^r norm. The choice of another fiberwise norm results in different constants appearing in the theorems below, which would have similar statements.

2.2. **Stratifications.** Recall from [BCR98, Section 9.7] that given two disjoint and connected Nash submanifolds (i.e. smooth and semialgebraic) X and Y in $\mathbb{R}^n$, such that $Y\subset \operatorname{clos}(X)$, we say that they satisfy the Whitney condition (a) if for every point $y\in Y$ and for every sequence $\{x_n\}_{n\ge 0}\subset X$ such that $\lim x_n=y$ and $\lim_n T_{x_n}X=\tau\in G(\dim(X),n)$, then τ contains T_yY.

Given a semialgebraic set $W\subseteq\mathbb{R}^p$ we will say that this set is *Whitney stratified* if it is stratified as $W=\coprod W_j$ with each stratum a Nash submanifold such that, whenever $W_{j_1}\subset\operatorname{clos}(W_{j_2})$ for some pair of strata W_{j_1} and W_{j_2}, then they satisfy Whitney condition (a). Recall that, by [BCR98, Theorem 9.7.11], every semialgebraic stratification can be refined to a Whitney stratification.

Definition 11. *Let $W\subseteq J^r(D,\mathbb{R}^k)$ be closed, semialgebraic and Whitney stratified:*

$$W=\coprod_{j=1}^{s}W_j. \tag{2.2}$$

Given $f\in C^{r+1}(D,\mathbb{R}^k)$ we will say that j^rf is transverse to W, and write $j^rf\pitchfork W$, if both j^rf and $j^rf|_{\partial D}$ are transverse to all the strata of the stratification (2.2).

Proposition 12. *Given $W\subseteq J^r(D,\mathbb{R}^k)$ closed, semialgebraic and Whitney stratified, there exists $\Sigma_W\subset J^{r+1}(D,\mathbb{R}^k)$, closed and semialgebraic, such that for every $f\in C^{r+1}(D,\mathbb{R}^k)$ we have that j^rf* is not *transverse to W at z if and only if $j^{r+1}f(z)\in\Sigma_W$.*

Proof. Let $W = \coprod_j W_j$ be the given Whitney stratification of W. The condition that $j^r f : D \to J^r(D, \mathbb{R}^k)$ is not transverse to W at x means that $j^r f(x) \in W_j$ for some $j = 1, \dots, s$ and that:

$$\mathrm{im}\,(d_x(j^r f)) + T_{j^r f(x)} W_j \neq T_{j^r f(x)} J^r(D, \mathbb{R}^k) \simeq \mathbb{R}^p. \tag{2.3}$$

For every stratum W_j let d_j denote its dimension and let $\tau_j : W_j \to \mathbb{R}^{d_j \times p}$ be a semialgebraic map such that for every $w \in W_j$ the columns of $\tau_j(w)$ are an orthonormal basis for $T_w W_j$ – we can assume that τ_j is continuous, possibly after refining the stratification. Then the condition (2.3) can be written as:

$$\mathrm{rk}(d_x j^r f, \tau_j(j^r f(x))) \leq p - 1.$$

In particular, since both $d_x j^r f$ and $j^r f(x)$ are linear images of $j^{r+1} f(x)$, i.e.

$$d_x j^r f = \pi_1(j^{r+1} f(x)) \quad \text{and} \quad j^r f(x) = \pi_2(j^{r+1} f(x))$$

for some semialgebraic maps

$$\pi_1 : J^{r+1}(D, \mathbb{R}^k) \to J^1(D, J^r(d, \mathbb{R}^k)) \quad \text{and} \quad \pi_2 : J^{r+1}(D, \mathbb{R}^k) \to J^r(D, \mathbb{R}^k),$$

it follows that the condition that $j^r f$ is not transverse to W can be written as: there exists $x \in D$ such that $\pi_2(j^{r+1} f(x)) \in W_j$ for some $j = 1, \dots, s$ and $\mathrm{rk}(\pi_1(j^{r+1} f(x)), \tau_j(\pi_2(j^{r+1} f(x))) \leq p - 1$.

In other words, defining the semialgebraic set:

$$\begin{aligned} \Sigma_W^1 &= \bigcup_{j=1}^{s} \pi_2^{-1}(W_j) \cap \{\mathrm{rk}\,(\pi_1(\cdot), \tau_j(\pi_2(\cdot))) \leq p - 1\} \\ &= \pi_2^{-1}(W) \cap \left(\bigcup_{j=1}^{s} \{\mathrm{rk}\,(\pi_1(\cdot), \tau_j(\pi_2(\cdot))) \leq p - 1\} \right), \end{aligned} \tag{2.4}$$

we see that $j^r f$ is not transverse to W at x if and only if $j^{r+1} f(x) \in \Sigma_W^1$. Analogously we can define a semialgebraic set Σ_W^2 such that $j^r f|_{\partial D}$ is not transverse to W if and only if $j^{r+1} f(x) \in \Sigma_W^2$ for some $x \in \partial D$. The set Σ_W is defined as the union of Σ_W^1 and Σ_W^2.

Let us now prove that Σ_W is closed. We will show that Σ_W^1 is closed, the proof for Σ_W^2 is similar. Since W is closed, it is enough to show that the set in the parenthesis in (2.4) is closed, and this follows immediately from the fact that the stratification for W satisfies Whitney condition (a). □

2.3. **Distance to the discriminant.**

Definition 13. *Let $W \subseteq J^r(D, \mathbb{R}^k)$ closed, semialgebraic and Whitney stratified. We define the quantity:*

$$\delta_W(f, D) = \inf_{z \in D} \mathrm{dist}(j^{r+1} f(z), \Sigma_{W,z}), \tag{2.5}$$

where we set $\Sigma_{W,z} = \Sigma_W \cap J_z^{r+1}(D, \mathbb{R}^k)$ and we adopt the convention that if $\Sigma_{W,z} = \emptyset$ then $\mathrm{dist}(j^{r+1} f(z), \Sigma_{W,z}) = \infty$.

We introduce now the set of all maps which are not transverse to W.

Definition 14. *Let $W \subseteq J^r(D, \mathbb{R}^k)$ be closed, semialgebraic and Whitney stratified. We denote by $\Delta_W \subset C^{r+1}(D, \mathbb{R}^k)$ the set*

$$\Delta_W = \{g \in C^{r+1}(D, \mathbb{R}^k) \,|\, g \text{ is not transverse to } W\},$$

and, for $z \in D$, by $\Delta_{W,z}(D)$ the set

$$\Delta_{W,z} = \{g \in C^{r+1}(D, \mathbb{R}^k) \,|\, g \text{ is not transverse to } W \text{ at } z\}.$$

Observe that, since transversality is an open condition, the set Δ_W is closed and

$$\Delta_W = \bigcup_{z\in D} \Delta_{W,z}. \tag{2.6}$$

Moreover, by Proposition 12, we have:

$$\Delta_{W,z} = \{g \in C^{r+1} \mid j^{r+1}g(z) \in \Sigma_{W,z}\}.$$

Next result relates $\delta_W(f, D)$ with the distance from Δ_W in the space of C^{r+1} maps.

Proposition 15. *There exists a constant $C = C(r, k, n)$ such that, for every $W \subseteq J^r(D, \mathbb{R}^k)$ closed, semialgebraic and Whitney stratified, and for every $f \in C^{r+1}(D, \mathbb{R}^k)$ we have:*

$$\delta_W(f, D) \le \mathrm{dist}_{C^{r+1}(D,\mathbb{R}^k)}(f, \Delta_W) \le C \cdot \delta_W(f, D).$$

Proof. Let us prove the first part of the statement:

$$\begin{aligned}
\mathrm{dist}_{C^{r+1}(D,\mathbb{R}^k)}(f, \Delta_W) &= \inf_{s\in\Delta_W} \|f - s\|_{C^{r+1}(D,\mathbb{R}^k)} \\
&= \inf_{z\in D} \inf_{s\in\Delta_{W,z}} \|f - s\|_{C^{r+1}(D,\mathbb{R}^k)} && \text{using (2.6)}\\
&= \inf_{z\in D} \inf_{s\in\Delta_{W,z}} \sup_{x\in D} \|j^{r+1}f(x) - j^{r+1}s(x)\| \\
&\ge \inf_{z\in D} \inf_{s\in\Delta_{W,z}} \|j^{r+1}f(z) - j^{r+1}s(z)\| \\
&\ge \inf_{z\in D} \inf_{s\in\Delta_{W,z}} \mathrm{dist}\left(j^{r+1}f(z), \Sigma_{W,z}\right) && \text{by Proposition 12}\\
&= \inf_{z\in D} \mathrm{dist}\left(j^{r+1}f(z), \Sigma_{W,z}\right) \\
&= \delta_W(f, D).
\end{aligned}$$

For the second part of the statement we use a classical result by Whitney on the infimum of the C^{r+1} norm of a function with prescribed $(r + 1)$-th jet at *one point* (see [Fef10, Theorem 1] for a modern reference): there exists $C = C(r, k, n) > 0$ such that, given $P \in J^{r+1}_z(D, \mathbb{R}^k)$, we have

$$\inf\left\{\|g\|_{C^{r+1}(D,\mathbb{R}^k)} \mid j^{r+1}g(z) = P\right\} \le C \cdot \|P\|. \tag{2.7}$$

Observe now that:

$$\begin{aligned}
\inf_{s\in\Delta_{W,z}} \|f - s\|_{C^{r+1}(D,\mathbb{R}^k)} &= \inf\left\{\|f - s\|_{C^{r+1}(D,\mathbb{R}^k)} \mid j^{r+1}s(z) \in \Sigma_{W,z}\right\} \\
&= \inf_{\sigma(z)\in\Sigma_{W,z}} \inf_{j^{r+1}s(z)=\sigma(z)} \|f - s\|_{C^{r+1}(D,\mathbb{R}^k)} \\
&= \inf_{\sigma(z)\in\Sigma_{W,z}} \left(\inf\left\{\|f - s\|_{C^{r+1}(D,\mathbb{R}^k)} \mid j^{r+1}s(z) = \sigma(z)\right\}\right) \\
&= \inf_{\sigma(z)\in\Sigma_{W,z}} \left(\inf\left\{\|g\|_{C^{r+1}(D,\mathbb{R}^k)} \mid j^{r+1}g(z) = j^{r+1}f(z) - \sigma(z)\right\}\right) \\
&\le \inf_{\sigma(z)\in\Sigma_{W,z}} \left(C \cdot \|j^{r+1}f(z) - \sigma(z)\|\right) \\
&= C \cdot \mathrm{dist}(j^{r+1}f(z), \Sigma_{W,z}), \qquad (2.8)
\end{aligned}$$

where in the first line we have used Proposition 12, and in the fifth line we have used Whitney's result (2.7). Taking now the infimum over $z \in D$ on both sides of (2.8) concludes the proof. □

Example 16. In this example we show that in general $C > 1$. Let us consider $D = [-1, 1]$ and the "critical points" singularity $W \subset J^1(D, \mathbb{R})$ defined by:

$$W = D \times \mathbb{R} \times \{0\}.$$

In this way, given $f \in C^1(D,\mathbb{R})$, we have that $j^1f^{-1}(W)$ consists of the set of critical points for f. The pointwise discriminant $\Sigma_{W,z} \subset J^2_z(D,\mathbb{R})$ is

$$\Sigma_{W,z} = \{z\} \times \mathbb{R} \times \{0\} \times \{0\}.$$

Let $f(x) = x$. We claim that $\mathrm{dist}_{C^2}(f, \Delta_W) > \delta_W(f, D)$. To start with, observe that:

$$\begin{aligned}\delta_W(f, D) &= \inf_{z\in D} \mathrm{dist}(j^2f(z), \Sigma_{W,z})\\ &= \inf_{z\in D} \mathrm{dist}((z, z, 1, 0), \{z\} \times \mathbb{R} \times \{0\} \times \{0\})\\ &= \inf_{z\in D} \|(z, z, 1, 0) - (z, z, 0, 0)\| = 1.\end{aligned}$$

We now assume that in this case $C = 1$, which implies $\mathrm{dist}_{C^2}(f, \Delta_W) \le 1$, and prove that this leads to a contradiction.

If $C = 1$, then for every $\epsilon > 0$ there is $g \in \Delta_W$ such that:

$$\|f - g\|_{C^2} < 1 + \epsilon.$$

Such a function g, being in Δ_W, has a degenerate critical point at $x_0 \in D$ and therefore:

$$\|f - g\|_{C_2} \ge \|j^2f(x_0) - j^2g(x_0)\| = \Big((x_0 - g(x_0))^2 + 1\Big)^{1/2}.$$

Together with this, the inequality $\mathrm{dist}_{C^2}(f, \Delta_W) \le 1$ implies now that:

$$|x_0 - g(x_0)|^2 < \epsilon^2 + 2\epsilon.$$

Let us now call $\phi = f - g$. This function satisfies:

$$|\phi(x_0)| < \sqrt{\epsilon^2 + 2\epsilon}, \quad \phi'(x_0) = 1, \quad \phi''(x_0) = 0 \quad \text{and} \quad \|\phi\|_{C^2} < 1 + \epsilon. \tag{2.9}$$

We show that the existence for every $\epsilon > 0$ of a C^2 function satisfying the list of conditions (2.9) leads to a contradiction. To this end, consider the following differential inequality:

$$y^2 + (y')^2 \le a^2, \quad y(0) = 1, \quad y'(0) = 0.$$

Then it is easy to see (by direct integration) that:

$$y(x) \ge y(0)\cos x - \sqrt{a^2 - y(0)}\sin x.$$

Letting now $y = \phi'$ and $a = 1 + \epsilon$, we get

$$\phi'(x) \ge \cos x - \sqrt{(1 + \epsilon)^2 - 1}\sin x$$

and, integrating

$$\phi(x) \ge \sin x + \sqrt{\epsilon^2 + 2\epsilon}(\cos x - 1) - \sqrt{\epsilon^2 + 2\epsilon}.$$

This means that

$$(1 + \epsilon)^2 \ge \|j^2\phi(x)\|^2 \ge 1 + 2(\epsilon^2 + 2\epsilon)(1 - \cos x) - 2\sqrt{\epsilon^2 + 2\epsilon}\sin x + \phi''(x)^2 - O(\epsilon^{1/2}).$$

In turn this implies:

$$|\phi''(x)| \le c\epsilon^{1/4},$$

and consequently

$$\phi'(x) \ge 1 - c\epsilon^{1/4}x \quad \text{and} \quad \phi(x) \ge x - c\epsilon^{1/4}x^2/2.$$

In particular:

$$\|j^2\phi(1)\|^2 \ge 2 - O(\epsilon^{1/4}),$$

which for $\epsilon > 0$ small enough is a contradiction.

Remark 17. For a given $f : \mathbb{R}^n \to \mathbb{R}$ of regularity class C^1, we can introduce the quantity:

$$\gamma(f, D) = \mathrm{dist}_{\mathbb{R}}(0, f(\mathrm{Crit}(f|_D)),$$

i.e. the distance from zero to the critical values of $f|_D$. Denoting as above by $\Delta = \Delta_W$ the set of functions for which $f = 0$ is not regular on D, for every $f \in C^1(D, \mathbb{R})$ we have:

$$\delta(f, D) \le \mathrm{dist}_{C^1(D,\mathbb{R})}(f, \Delta) \le \gamma(f, D)$$

The first inequality is just Proposition 15. For the second inequality, let $\gamma(f, D) = \gamma_0 > 0$ (otherwise $f \in \Delta$ and the inequality is trivial) and pick $\xi \in \mathbb{R}$ with $|\xi| = \gamma_0$ and ξ a critical value of $f|_D$. Then $f - \xi \in \Delta$ and $f - t\xi \notin \Delta$ if $|t| < 1$. Therefore:

$$\mathrm{dist}_{C^1(D,\mathbb{R})}(f, \Delta) \le \|f - (f - \xi)\|_{C^1(D,\mathbb{R})} = \|\xi\|_{C^1(D,\mathbb{R})} = |\xi| = \gamma_0.$$

Definition 18. *Let $W \subseteq J^r(D, \mathbb{R}^k)$ closed, semialgebraic and Whitney stratified and let $f \in C^\ell(D, \mathbb{R}^k)$ with $\ell \ge r + 1$. We define the C^ℓ-condition number of f on the disk D with respect to W as:*

$$\kappa_W^{(\ell)}(f, D) = \frac{\|f\|_{C^\ell(D,\mathbb{R}^k)}}{\delta_W(f, D)}.$$

Lemma 19. *Let $W \subseteq J^r(D, \mathbb{R}^k)$ closed and semialgebraic, Whitney stratified. For every $f \in C^{r+1}(D, \mathbb{R}^k)$, we have that $j^r f$ is transverse to W if and only if $\delta_W(f, D) > 0$, and in this case $j^r f^{-1}(W)$ is a Whitney stratified set. In particular, if W is smooth, then $j^r f^{-1}(W)$ is a smooth neat[5] submanifold of D.*

Proof. If $j^r f$ is transverse to W, then for no point $z \in D$ we have $j^{r+1} f(z) \in \Sigma_{W,z}$ by Proposition 12; this means that for every $z \in W$ we have $\mathrm{dist}(j^{r+1} f(z), \Sigma_{W,z}) > 0$, and since the function $z \mapsto \mathrm{dist}(j^{r+1} f(z), \Sigma_{W,z})$ is continuous on the compact set D, then the infimum in (2.5) is a minimum and this minimum must be positive.

Vice versa, if $\delta_W(f, D) > 0$ then $\mathrm{dist}(j^{r+1} f(z), \Sigma_{W,z}) > 0$ for every $z \in D$ and $j^r f$ is transverse to W by Proposition 12.

When $\delta_W(f, D) > 0$, the fact that $j^r f^{-1}(W)$ is a stratified set is a standard application of the transversality theorems; similarly in the case W is smooth. □

2.4. **Quantitative transversality.**

Lemma 20. *Let $W \subseteq J^r(D, \mathbb{R}^k)$ be closed and semialgebraic, Whitney stratified and let $f \in C^{r+1}(D, \mathbb{R}^k)$ be such that $j^r f$ is transverse to W (in particular $\delta_W(f, D) > 0$). For every $g \in C^{r+1}(D, \mathbb{R}^k)$ such that*

$$\|f - g\|_{C^{r+1}(D,\mathbb{R}^k)} < \mathrm{dist}_{C^{r+1}}(f, \Delta_W),$$

the two pairs $(D, j^r f^{-1}(W))$ and $(D, j^r g^{-1}(W))$ are isotopic. In particular, the result holds if $\|f - g\|_{C^{r+1}(D,\mathbb{R}^k)} < \delta_W(f, D)$.

Proof. By assumption on $\|f - g\|_{C^{r+1}}$, the homotopy $f_t = f + t(g - f) : D \to \mathbb{R}^k$, for $t \in [0, 1]$ is all disjoint from Δ_W. In particular the corresponding homotopy $j^r f_t : D \to J^r(D, \mathbb{R}^k)$ is all transverse to W and the result follows from Thom's Isotopy Lemma [Tho69, Théorème 2.D.2]. As for the second part of the statement: by Proposition 15 we have $\delta_W(f, D) \le \mathrm{dist}_{C^{r+1}}(f, \Delta_W)$ and the conclusion follows from the first part. □

[5]A *neat* submanifold W of a manifold with boundary D is a manifold with boundary such that $\partial W \subset \partial D$ and for any point $x \in \partial W$ the following condition holds: $T_x W \not\subset T_x(\partial D)$, see [Hir94].

2.5. **Semicontinuity of Betti numbers.** The key ingredient in the proof of the topological bound of Theorem 3 is the following result. Applied to the case of $W \subset J^r(D, \mathbb{R}^k) \subset \mathbb{R}^p$, $\varphi = j^r f$ and $\psi = j^r g$, it allows to control, from below, the Betti numbers of the singularity $Z = (j^r g)^{-1}(W)$, provided that g is close enough in the $\mathcal{C}^r$ topology to a known function f, such that $j^r f \pitchfork W$. In this sense, it plays a role analogous to that played by Lemma 20 in the proof of Theorem 1, with the significant difference that it requires only to know the C^r distance between f and g.

Theorem 21 (Theorem 2 from [LS19])**.** *Let $\varphi\colon D \to \mathbb{R}^p$ be a C^1 function and let $W \subset \mathbb{R}^p$ be a smooth closed submanifold. Assume that $\varphi \pitchfork W$ and that the set $Z = \varphi^{-1}(W)$ is contained in the interior of D. Let $E \subset E_1 \subset int(D)$ be tubular neighborhoods of Z with the property that $\overline{E} \subset E_1$[6].*

(1) Let the space $C^0(D, \mathbb{R}^p)$ be topologized by the standard C^0 norm. Define the set $\mathcal{U}_{E,\varphi}$ as the connected component containing φ of the set

$$\mathcal{U}_E = \left\{\psi \in \mathcal{C}^0(D, \mathbb{R}^p)\colon \psi^{-1}(W) \subset E\right\}. \tag{2.10}$$

Then $\mathcal{U}_{E,\varphi}$ is open in $C^0(D, \mathbb{R}^p)$.

(2) Let $\check{b}(Z)$ denote the sum of the Čech Betti numbers of Z. For any $g \in \mathcal{U}_{E,\varphi}$, we have

$$b(Z) \leq \check{b}\left(\psi^{-1}(W)\right). \tag{2.11}$$

Remark 22. The statement reported in [LS19] is weaker than the one above, in the sense that it doesn't specifies the neighborhood $\mathcal{U}_{E,\varphi}$ and it requires that $\psi \pitchfork W$. The proof of this stronger form can be found in the second author's PhD thesis [Ste21]. In the proof of [LS19, Theorem 2], the condition on $\psi \pitchfork W$ is needed in order to guarantee that a neighborhood of the zero set retracts to it. In this context, let us notice that we will use Theorem 21 in the proof of Theorem 9 and of Theorem 3, where ψ will be a polynomial and W a semialgebraic subset. In this case, being the preimage of a semialgebraic set via a polynomial map semialgebraic, the condition of having a neighborhood which retracts on it is already verified; moreover, for the same reason, on the right hand side of (2.11) one can take ordinary Betti numbers (as Čech cohomology coincides with the singular one). Finally, the description of the neighborhood as in (2.10), even if not stated in the original theorem [LS19, Theorem 2], is evident from its proof.

3. Polynomial approximation of singularities

3.1. **Proof of Theorem 1.** In the sequel we will need the following quantitative versions of Weierstrass' Approximation Theorem from [BBL02].

Theorem 23 (Theorem 2 from [BBL02])**.** *For every $r \geq 0$ there exists a constant $a_r(D) > 0$ such that for every $f \in C^r(D, \mathbb{R})$ and for every $d \geq 0$ there is a polynomial $p_d(f) \in \mathbb{R}[x_1, \ldots, x_n]$ of degree at most d such that for every $\ell \leq \min\{r, d\}$:*

$$\|f - p_d(f)\|_{C^\ell(D,\mathbb{R})} \leq \frac{a_r(D)}{d^{r-\ell}} \cdot \|f\|_{C^r(D,\mathbb{R})}.$$

Corollary 24. *For every $r \geq 0$ there exists a constant $a_{r+2}(D) > 0$ such that for every $f \in C^{r+2}(D, \mathbb{R})$ and for every $\varepsilon > 0$ there is a polynomial $p_d(f) \in \mathbb{R}[x_1, \ldots, x_n]$ of degree at most d such that $\|f - p_d(f)\|_{C^{r+1}(D,\mathbb{R})} < \varepsilon$ and:*

$$d \leq \max\left\{r + 1, \frac{\|f\|_{C^{r+2}(D,\mathbb{R})}}{\varepsilon} \cdot a_{r+2}(D)\right\}.$$

[6]Only the *existence* of E_1 is needed.

With this result available, we prove now Theorem 1.

Let $f = (f_1, \ldots, f_k)$ and use Theorem 23 to get for every $d \in \mathbb{N}$ a polynomial map $p_d(f) = (p_d(f_1), \ldots, p_d(f_k))$ such that:

$$\|f - p_d(f)\|_{C^{r+1}(D,\mathbb{R}^k)} = \left(\sum_{i=1}^{k} \|f_i - p_d(f_i)\|^2_{C^{r+1}(D,\mathbb{R})} \right)^{1/2} \leq \frac{a_{r+2}(D)}{d} \cdot \|f\|_{C^{r+2}(d,\mathbb{R}^k)}. \tag{3.1}$$

Choose now the approximating degree to be

$$d = \max \left\{ r+1, \left\lfloor \frac{\|f\|_{C^{r+2}(D,\mathbb{R}^k)}}{\mathrm{dist}_{C^{r+1}}(f, \Delta_W)} \cdot 2a_{r+2}(D) \right\rfloor \right\}. \tag{3.2}$$

Then

$$a_{r+2}(D) \cdot \frac{\|f\|_{C^{r+2}(D,\mathbb{R}^k)}}{\mathrm{dist}_{C^{r+1}}(f, \Delta_W)} < 2a_{r+2}(D) \cdot \frac{\|f\|_{C^{r+2}(D,\mathbb{R}^k)}}{\mathrm{dist}_{C^{r+1}}(f, \Delta_W)} \leq d,$$

and the inequality (3.1) becomes:

$$\|f - p_d(f)\|_{C^{r+1}(D,\mathbb{R}^k)} < \mathrm{dist}_{C^{r+1}}(f, \Delta_W).$$

The conclusion follows now from Lemma 20. Lemma 25 below ensures that the inequality (3.2) can be put in the slightly different form stated in the theorem, with the constant outside of $\max\{\ldots\}$.

Lemma 25. *Let $r, f, a > 0$ be positive real numbers and define $c := \max\{1, a\}$. Then*

$$\max\{r, af\} \leq c \cdot \max\{r, f\}.$$

Proof. Let $L := \max\{r, af\}$ and $R := \max\{r, f\}$. Depending on the order of r, f, a, there are 4 possible cases for the value of the couple (L, R), listed on the left column. On the rightmost column we write the ratio $\frac{R}{L}$.

(L, R)	$\frac{L}{R} \leq \ldots$
(r, r)	1
(af, r)	a
(r, f)	1
(af, f)	a

In all cases, the inequality $L \leq c \cdot R$ is satisfied, so the lemma is proved. ◻

Remark 26. From Lemma 25 we also deduce that the constant $c_1(r, D)$ can be defined as

$$c_1(r, D) := \max\{1, 2a_{r+2}(D)\}.$$

Notice that this number is independent from the position of the disk, since the result of Theorem 23 for a given disk $D \subset \mathbb{R}^n$ implies the same result for every other disk $D + x$, for all $x \in \mathbb{R}^n$.

3.2. **Proof of Corollary 2.** Let $p_d(f) : D \to \mathbb{R}^k$ be the polynomial map given by Theorem 1, with:

$$d = \deg(p_d(f)) \leq c_1(r, D) \cdot \max \left\{ r+1, \frac{\|f\|_{C^{r+2}(D,\mathbb{R}^k)}}{\mathrm{dist}_{C^{r+1}}(f, \Delta_W)} \right\}$$

and such that

$$(D, j^r f^{-1}(W)) \sim (D, j^r p_d(f)^{-1}(W)).$$

Observe that $j^r p_d(f) : D \to D \times J^r(n, k)$ is also a polynomial map, with each component of degree bounded by d. The set $W \subset J^r(D, \mathbb{R}^k)$ can be described by a quantifier-free Boolean formula without negations, involving a family of polynomials $\mathcal{Q} = \{q_1, \dots, q_s\}$ with $s = s(W)$ and $\deg(q_i) \le a_1(W)$, whose atoms are of the form $q_i \le 0$, $q_i \ge 0$ or $q_i = 0$. In particular $j^r p_d(f)^{-1}(W)$ can also be described by a quantifier-free Boolean formula without negations, involving a family of s polynomials of degrees bounded by $d \cdot a_1(W)$, whose atoms are of the form $q_i \circ j^r p_d(f) \le 0$, $q_i \circ j^r p_d(f) \ge 0$ or $q_i \circ j^r p_d(f) = 0$. We are therefore in the position of applying [Bas99, Theorem 1] and we get the existence of a constant $a_2 > 0$ such that:

$$\begin{aligned} b(j^r f^{-1}(W)) &= b(j^r p_d(f)^{-1}(W)) \\ &\le a_2^n s(W)^n (a_1(W) d)^n \\ &\le (a_2 s(W) a_1(W))^n \left(c_1(r, D) \cdot \max\left\{ r+1, \frac{\|f\|_{C^{r+2}(D,\mathbb{R}^k)}}{\mathrm{dist}_{C^{r+1}}(f, \Delta_W)} \right\} \right)^n \\ &\le c_2(W) \cdot \max\left\{ r+1, \frac{\|f\|_{C^{r+2}(D,\mathbb{R}^k)}}{\mathrm{dist}_{C^{r+1}}(f, \Delta_W)} \right\}^n . \end{aligned}$$

3.3. **Proof of Theorem 3.** Let us make the identification

$$J^{r+1}(D, \mathbb{R}^k) \cong D \times \mathbb{R}^p \times \mathbb{R}^q,$$

in such a way that the canonical projection $j_z^{r+1} f \to j_z^r f$ corresponds to the map $(z, \xi, \eta) \mapsto (z, \xi)$ and the so called "source map" $j_z^{r+1} f \mapsto z$ corresponds to the projection on the first factor.

Definition 27. *Let $W \subset J^r(D, \mathbb{R}^k) \cong D \times \mathbb{R}^p$ be a smooth closed semialgebraic submanifold and let $\theta \in W$.*

$$C(T_\theta W) := \left\{ \alpha \in \mathbb{R}^{n \times (n+p)} \,\middle|\, \dim\left(\mathrm{im}(\alpha) + T_\theta W\right) < n + p \right\}$$

The set $C(T_\theta W)$ is the semialgebraic subset of the space $\mathbb{R}^{n\times(n+p)}$ consisting of all the linear maps $\mathbb{R}^n \to \mathbb{R}^{n+p}$ that are not transverse to the vector space $T_\theta W \subset \mathbb{R}^{n+p}$.

Let W be compact. Suppose that $W \pitchfork J_z^r(D, \mathbb{R}^k)$ for all $z \in D$, where in our identification $J_z^r(D, \mathbb{R}^k) = \{z\} \times \mathbb{R}^p$, and define $W_z = W \cap J_z^r(D, \mathbb{R}^k)$. Then the set

$$B_\varepsilon(W) := \left\{ (z, \xi) \in J^r(D, \mathbb{R}^k) \,\middle|\, \mathrm{dist}\left((z, \xi), W_z\right) < \varepsilon \right\} \tag{3.3}$$

is a smooth open tubular neighborhood of W, for small enough ε whose fibers are contained in the spaces $J_z^r(D, \mathbb{R}^k)$ (here we are using the transversality assumption). Let us denote by $\varepsilon_0(W)$ the supremum of the set of all ε with such property.

Definition 28. *Let W be compact and such that $W \pitchfork J_z^r(D, \mathbb{R}^k)$ for all $z \in D$, so that $W_z = W \cap J^r(D, \mathbb{R}^k)$ is a closed submanifold. For any $f \in C^{r+1}(D, \mathbb{R}^k)$ let*

$$\pi_W(f, z) = \{\theta \in W_z \colon \mathrm{dist}(j^r f(z), W) = |j^r f(z) - \theta|\}$$

and define

$$\hat{\delta}_W(f, z) := \left(\mathrm{dist}(j^r f(z), W_z)^2 + \min_{\theta \in \pi_W(f,z)} \mathrm{dist}\left(d_z(j^r f), C(T_\theta W)\right)^2 \right)^{\frac{1}{2}};$$

$$\hat{\delta}_W(f, D) := \min\left\{ \inf_{z \in D} \hat{\delta}_W(f, z); \inf_{z \in \partial D} \mathrm{dist}(j^r f, W_z); \varepsilon_0(W) \right\};$$

Remark 29. Notice that $\hat{\delta}_W(f, D) > 0$ if and only if $j^r f$ is transverse to W at any point $z \in \mathrm{int}(D)$ and moreover $j^r f^{-1}(W) \cap \partial D = \emptyset$.

Lemma 30. *Let $W \subset D \times \mathbb{R}^p$ be a compact smooth submanifold and assume that $W \pitchfork \{z\} \times \mathbb{R}^p$ for all $z \in D$. Let $(z, \xi) \in \partial B_\varepsilon(W)$, with $0 < \varepsilon < \varepsilon_0(W)$ and let $(z, w) \in W$ be such that $|\xi - w| = \varepsilon$. Then*

$$T_{(z,w)}W \subset T_{(z,\xi)}\partial B_\varepsilon(W).$$

Proof. Let $\xi(t), w(s) \in \mathbb{R}^p$ be two smooth curves such that $\xi(0) = \xi$, $w(0) = w$, $(z, w(s)) \in W$ and $(z, \xi(t)) \in \partial B_\varepsilon(W)$. Then for all s, t we have

$$|\xi(t) - w(s)| \geq \operatorname{dist}((z, \xi(t)), W_z) = \varepsilon = |\xi(0) - w(0)|,$$

so that differentiating the function $(s, t) \mapsto |\xi(t) - w(s)|^2$ at the point $(0, 0)$, we get that $\langle (\xi - w), \dot{w} \rangle = 0$ and $\langle (\xi - w), \dot{\xi} \rangle = 0$. The arbitrariness in the choice of the two curves implies that

$$T_{(z,\xi)}W \cap (\{0\} \times \mathbb{R}^p) \subset \{0\} \times (\xi - w)^\perp = T_{(z,\xi)}\partial B_\varepsilon(W) \cap (\{0\} \times \mathbb{R}^p);$$

here on the right we have an equality instead than an inclusion for dimensional reasons.

Now let $(z(t), w(t)) \in W$ be any smooth curve in W such that $z(0), w(0) = z, w$. Given the tubular neighborhood structure of $\partial B_\varepsilon(W)$ (see the discussion after equation (3.3)), there also exists a smooth curve $(z(t), \xi(t)) \in \partial B_\varepsilon(W)$ with $\xi(0) = \xi$ and of course $(\dot{z}, \dot{\xi}) \in T_{(z,\xi)}\partial B_\varepsilon(W)$. Thus

$$|\xi(t) - w(t)| \geq \operatorname{dist}\left(\xi(t), W_{z(t)}\right) = \varepsilon = |\xi(0) - w(0)|,$$

so that differentiating at $t = 0$ we get that $\dot{\xi} - \dot{w} \in (w - \xi)^\perp$ meaning that the vector $(0, \dot{\xi} - \dot{w})$ belongs to the tangent space $T_{(z,\xi)}\partial B_\varepsilon(W)$. We conclude that $(\dot{z}, \dot{w}) \in T_{(z,\xi)}\partial B_\varepsilon(W)$. Since the latter construction can be repeated for any tangent vector $(\dot{z}, \dot{w}) \in T_{(z,w)}W$, this concludes the proof. □

Proof of Theorem 3. Consider the set $B_\varepsilon(W)$ defined as in (3.3) and let us prove that

$$j^r f \pitchfork \partial B_\varepsilon(W), \qquad \forall \varepsilon \in (0, \hat{\delta}_W(f, D)). \tag{3.4}$$

By contradiction assume that there is a point $z \in D$ such that $j^r f(z) \in \partial B_\varepsilon(W)$, but $j^r f$ is not transverse to $\partial B_\varepsilon(W)$ at z i.e.

$$\operatorname{im}(d_z(j^r f)) \subset T_{j^r f(z)}\partial B_\varepsilon(W). \tag{3.5}$$

Since $\varepsilon < \varepsilon_0(W)$, there exists a (unique) point $\theta \in W_z$ such that $\|j^r f - \theta\| = \varepsilon$ i.e. $\{\theta\} = \pi_W(j^r f(z))$. By Lemma 30 we have that $T_{j^r f(z)}\partial B_\varepsilon(W) \supset T_\theta W$ and this, together with (3.5), implies that the linear map $d_z(j^r f) \colon \mathbb{R}^n \to \mathbb{R}^{n+p}$ belongs to the critical set $C(T_\theta W)$. This leads to a contradiction:

$$\hat{\delta}_W(f, z) \leq \operatorname{dist}(j^r f(z), W_z) = \varepsilon < \hat{\delta}_W(f, z),$$

proving (3.4).

Fix $0 < \varepsilon < \hat{\delta}_W(f, D)$. From the transversality condition (3.4) it follows that the set $E = j^r f^{-1}(B_\varepsilon(W))$ is a smooth tubular neighborhood of the submanifold $Z = j^r f^{-1}(W)$ in the smooth manifold $\operatorname{int}(D)$. The fact that E is contained in the interior of D is due to the inequality

$$\varepsilon < \hat{\delta}_W(f, D) \leq \inf_{z \in \partial D} \operatorname{dist}(j^r f, W_z).$$

Now let $g \in \mathbb{R}[z_1, \ldots, z_n]^k$ be a polynomial such that $\|f - g\|_{C^r(D, \mathbb{R}^k)} \leq \varepsilon$. Then $j^r g^{-1}(W) \subset E$ indeed if $j^r g(z) \in W$, then

$$\operatorname{dist}(j^r f(z), W_z) \leq \operatorname{dist}(j^r f(z), j^r g(z)) + \operatorname{dist}(j^r g(z), W_z) \leq \varepsilon + 0,$$

hence $j^r f(z) \in B_\varepsilon(W)$. The same is true for all maps $g_t = tg + (1-t)f$, therefore the map g belongs to the set $\mathcal{U}_{E,j^r f} \subset C^0(D, \mathbb{R}^{n+p})$ defined in the statement of Theorem 21 and thus

$$b(Z) \leq b\left(j^r g^{-1}(W)\right). \tag{3.6}$$

(Since $j^r g$ is a polynomial, we can use the singular homology Betti numbers.) By Theorem 23 we can assume that each component of g has degree smaller than an integer number $d \leq a_{r+1}(D)\|f\|_{C^{r+1}(D,\mathbb{R}^k)}\varepsilon^{-1}$ or, if the latter is less than r, $d \leq r$ so that since W is semi-algebraic and the map $j^r g$ is polynomial there exists a constant $a(W) > 0$ depending only on W such that

$$b\left(j^r g^{-1}(W)\right) \leq a(W) d^n \leq a(W) \max\left\{r, a_{r+1}(D)\|f\|_{C^{r+1}(D,\mathbb{R}^k)}\varepsilon^{-1}\right\}^n. \tag{3.7}$$

Combining (3.6) and (3.7) and from the arbitrariness of ε, we obtain the thesis. □

4. What is the degree of a smooth hypersurface?

In this section we consider the case in which $j^r f^{-1}(W)$ is just the zero set of the function $f\colon D \to \mathbb{R}$. In this case $W = W_0$ is the subset of the space $J^0(D, R) \cong D \times \mathbb{R}$ corresponding to $W_0 = D \times \{0\}$.

In this particular case we have that $\delta_{W_0}(f, D) = \mathrm{dist}(f, \Delta_{W_0})$ (see Proposition 31 below) and for the rest of the current section we will denote this quantity by $\delta(f, D)$.

Notice that under the identification

$$J^1(D, \mathbb{R}) \cong D \times \mathbb{R} \times \mathbb{R}^n,$$

we have that $\Sigma_{W_0,z} = \{z\} \times \{0\} \times \{0\}$ for every $z \in \mathrm{int}(D)$, while $\Sigma_{W_0,z} = \{z\} \times \{0\} \times (T_z\partial D)^\perp$ if $z \in \partial D$, therefore the distance of f from the discriminant W_0 can be expressed by the following formula.

$$\delta(f, D) = \min\left\{\inf_{z\in D}\left(|f(z)|^2 + |\nabla f(z)|^2\right)^{\frac{1}{2}}, \inf_{z\in\partial D}\left(|f(z)|^2 + |\nabla(f|_{\partial D})(z)|^2\right)^{\frac{1}{2}}\right\}.$$

Proposition 31. $\delta_{W_0}(f, D) = \mathrm{dist}_{C^1}(f, \Delta_{W_0})$.

Proof. Let us make the identification $J^1(D, \mathbb{R}) \cong D \times \mathbb{R} \times \mathbb{R}^n$. By definition, there exists a point $x \in D$ and a jet $j_x g = (x, r, v) \in J^1_x(D, \mathbb{R})$ such that $j^1(f+g)(x) \in \Sigma_{W_0,x}$ and

$$\delta_{W_0}(f, D) = \|(r, v)\|.$$

Fix $\varepsilon > 0$ and let $\rho_\varepsilon\colon \mathbb{R} \to [-\varepsilon, \varepsilon]$ be a smooth function such that $\max_{t\in\mathbb{R}}|\rho'_\varepsilon(t)| = \rho'_\varepsilon(0) = 1$. For instance ρ can be constructed as follows:

$$\rho_\varepsilon(t) = \begin{cases} t & \text{if } -\frac{\varepsilon}{2} \leq t \leq \frac{\varepsilon}{2} \\ \text{is increasing and convex} & \text{if } t \leq -\frac{\varepsilon}{2} \\ \text{is increasing and concave} & \text{if } \frac{\varepsilon}{2} \leq t \\ \varepsilon & \text{if } |t| \geq 2\varepsilon. \end{cases}$$

Define now a new function $h \in C^1(D, \mathbb{R})$ such that $h(z) = r + \rho_\varepsilon(v^T(z-x))$. Then the function $f + h$ belongs to the discriminant set D_{W_0} since $j^1(f+h)(x) = j^1(f+g)(x) \in \Sigma_{W_0,x}$, therefore

$$\begin{aligned} \mathrm{dist}_{C^1}(f, \Delta_{W_0}) &\leq \|h\|_{C^1(D,\mathbb{R})} \\ &\leq \sup_{z\in D}\left(|r + \rho_\varepsilon(z)|^2 + |\rho'_\varepsilon(v^T(z-x))v|^2\right)^{\frac{1}{2}} \\ &\leq \left(|r+\varepsilon|^2 + |v|^2\right)^{\frac{1}{2}}. \end{aligned}$$

Thus, from the arbitrariness of $\varepsilon > 0$, we conclude that

$$\mathrm{dist}_{C^1}(f, \Delta_{W_0}) \le (r^2 + |v|^2)^{\frac{1}{2}} = \delta_{W_0}(f, D).$$

Combining this with the general inequality (1.5) we obtain the thesis.

□

4.1. **Proof of Theorem 5.** Denote by $\rho = \rho(Z)$. We consider the function $\mathrm{d}^*_Z : \mathbb{R}^n \to \mathbb{R}$ defined to be the *signed* distance from Z. By [Foo84, Remark 2], if Z is of class C^1 and with positive reach $\rho(Z) > 0$, the function d^*_Z is C^1 on the set $\{\mathrm{d}^*_Z < \rho(Z)\}$.

We need to consider also an auxiliary function $g : \mathbb{R} \to \mathbb{R}$ of class C^2 such that $g(t) = -g(-t)$ for all $t \in \mathbb{R}$ and:

$$g(t) = \begin{cases} t & \text{if } 0 \le t \le \frac{1}{2} \\ \text{is increasing and concave} & \text{if } \frac{1}{2} \le t \le \frac{7}{8} \\ \frac{3}{4} & \text{if } t \ge \frac{7}{8} \end{cases}$$

The existence of such a function is elementary; it can be taken, for instance, to be piecewise polynomial. Denoting by g_ρ the function $t \mapsto \rho \cdot g(t/\rho)$, we set:

$$f(x) = g_\rho(\mathrm{d}^*_Z(x)). \tag{4.1}$$

Let us start by estimating $\delta(f, D)$, for a disk D_R with $Z \subset \mathrm{int}D_{R-\rho}$. Notice that this condition implies that, by construction, the function $f \equiv \frac{3}{4}\rho$ on $\mathbb{R}^n \backslash \mathrm{int}(D)$ and in particular:

$$\delta(f, D) = \min\left\{ \inf_{z\in D} \left(|f(z)|^2 + \|\nabla f(z)\|^2\right)^{1/2}, \frac{3}{4}\rho \right\}. \tag{4.2}$$

Observe now that for every x such that $t = \mathrm{d}^*_Z(x) < \rho$ we have:

$$\begin{aligned} |f(x)|^2 + \|\nabla f(x)\|^2 &= |g_\rho(\mathrm{d}^*_Z(x))|^2 + |g'_\rho(\mathrm{d}^*_Z(x))|^2 \cdot \|\nabla \mathrm{d}^*_Z(x)\|^2 \\ &= |g_\rho(\mathrm{d}^*_Z(x))|^2 + |g'_\rho(\mathrm{d}^*_Z(x))|^2 \\ &= \rho^2 |g(t/\rho)|^2 + |g'(t/\rho)|^2, \end{aligned} \tag{4.3}$$

where in the second line we have used the fact that $\|\nabla \mathrm{d}^*_Z\| \equiv 1$. In particular, partitioning the domain of definition of the function g_ρ, it follows that:

$$\begin{aligned} \inf_{z\in D} |f(z)|^2 + \|\nabla f(z)\|^2 &= \inf \left\{ 1 + \frac{\rho^2}{4}, \rho^2(3/4)^2, \inf_{\frac{1}{2}\rho \le t \le \frac{7}{8}t} \rho^2 |g(t/\rho)|^2 + |g'(t/\rho)|^2 \right\} \\ &\ge \inf \left\{ 1 + \frac{\rho^2}{4}, \rho^2(3/4)^2, \inf_{\frac{1}{2}\rho \le t \le \frac{7}{8}t} \rho^2 |g(t/\rho)|^2 + \inf_{\frac{1}{2}\rho \le t \le \frac{7}{8}t} |g'(t/\rho)|^2 \right\} \\ &= \inf \left\{ 1 + \frac{\rho^2}{4}, \rho^2(3/4)^2, \frac{\rho^2}{4} \right\} \\ &= \frac{\rho^2}{4}. \end{aligned}$$

Together with (4.2), this gives:

$$\delta(f, D) \ge \frac{\rho}{2}. \tag{4.4}$$

Let us now estimate $\kappa(f, D)$. Again partitioning the domain and using (4.3) and the fact that $|g'(t)| \le 1$ for all t, we immediately get:

$$\|f\|_{C^1(D,\mathbb{R})} \le 1 + \frac{3}{4}\rho. \tag{4.5}$$

Combining (4.4) with (4.5) gives (1.7).

In order to get (1.8), we will work with the further assumption that Z is of class C^2. Under this assumption the function d_Z^* is C^2 on $\{\mathrm{d}_Z^* < \rho(Z)\}$ and, using (4.1) we get:

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \begin{cases} 0 & \text{if } \mathrm{d}_Z^*(x) > \frac{7}{8}\rho \\ g_\rho''(\mathrm{d}_Z^*(x))\partial_i \mathrm{d}_Z^*(x)\partial_i \mathrm{d}_Z^*(x) + g_\rho'(\mathrm{d}_Z^*(x))\partial_{ij}^2 \mathrm{d}_Z^*(x) & \text{otherwise} \end{cases}. \tag{4.6}$$

Since the function g is fixed with $\|g'\| \le 1$ and $\|g''\| \le a_1$ (for some constant $a_1 > 0$), we have

$$\|g_\rho'\| = \|g'\| \le 1 \quad \text{and} \quad \|g_\rho''\| = \frac{1}{\rho}\|g''\| \le \frac{a_1}{\rho}.$$

In particular, in order to estimate the absolute value of (4.6), we need an estimate for the Hessian of d_Z^*. We use now the following fact [GT01, Lemma 14.17]: given $x \in \{\mathrm{d}_Z^* < \rho\}$ with $\mathrm{d}_Z^*(x) = t$, the Hessian of d_Z^* at x has eigenvalues:

$$\beta_1(t) = \frac{-\lambda_1}{1-\lambda_1 t}, \dots, \beta_{n-1}(t) = \frac{-\lambda_{n-1}}{1-\lambda_{n-1}t}, \quad \beta_n(t) = 0$$

where $\lambda_1, \dots, \lambda_{n-1}$ are the eigenvalues of the Weingarten map of the hypersurface Z at $z(x) = \mathrm{argmin}_{z\in Z}\mathrm{dist}(z,x)$, i.e. the principal curvatures of Z in $\mathbb{R}^n$.

The modulus of each of these eigenvalues can be estimated by

$$|\lambda_i| \le \frac{1}{\rho(Z)},$$

and each ratio $|1-\lambda_i t|^{-1}$ is smaller than 1 if $\lambda_1 \le 0$ and, if $\lambda_i > 0$, smaller than its value at $t = \frac{7}{8}\rho$ (the extremum of the interval where we have to estimate the function), which is:

$$\frac{1}{|1-\lambda_i t|} \le \frac{1}{\left|1-\lambda_i \frac{7}{8}\rho\right|} \le 8,$$

where we have used $0 < \lambda_i \le \frac{1}{\rho}$. Going back to (4.6), we have

$$\left|\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right| \le |g_\rho''(t)| + |g_\rho'(t)| \cdot |\partial_{ij}^2 \mathrm{d}_Z^*(x)|.$$

Observe now that for the construction of the function g we need:

$$g''(1/2) = g''(7/8) = 0 \quad \text{and} \quad \int_{1/2}^{7/8} g''(s)ds = -1.$$

Since $7/8 - 1/2 = 3/8 > 1/3$, we can choose the function g such that $|g''| \le 3$, which implies $|g_\rho''| \le 3/\rho$.

From this we immediately deduce that

$$\begin{aligned}
\|f\|_{C^2(D,\mathbb{R})} &\le \|f\|_{C^1(D,\mathbb{R})} + \left(\sum_{i,j} |\partial^2_{ij} f|^2\right)^{1/2} \\
&\le 1 + \frac{3}{4}\rho + \left(\sum_{i,j} (|g''_\rho| + |\partial^2_{ij} \mathrm{d}^*_Z|)^2\right)^{1/2} \\
&\le 1 + \frac{3}{4}\rho + \left(\sum_{i,j} (3/\rho + |\partial^2_{ij} \mathrm{d}^*_Z|)^2\right)^{1/2} \\
&\le 1 + \frac{3}{4}\rho + \left(\sum_{i,j} 2((3/\rho)^2 + |\partial^2_{ij} \mathrm{d}^*_Z|^2)\right)^{1/2} \\
&\le 1 + \frac{3}{4}\rho + \sqrt{2}\left(n^2\frac{9}{\rho^2} + \sum_{i,j} |\partial^2_{ij} \mathrm{d}^*_Z|^2\right)^{1/2} \\
&= 1 + \frac{3}{4}\rho + \sqrt{2}\left(n^2\frac{9}{\rho^2} + \|\partial^2_{ij} \mathrm{d}^*_Z\|_F^2\right)^{1/2} \\
&= 1 + \frac{3}{4}\rho + \sqrt{2}\left(n^2\frac{9}{\rho^2} + \sum_{i=1}^n \lambda_i(\partial^2_{ij} \mathrm{d}^*_Z)^2\right)^{1/2} \\
&\le 1 + \frac{3}{4}\rho + \sqrt{2}\left(n^2\frac{9}{\rho^2} + \frac{n}{\rho^2}\right)^{1/2} \\
&\le 1 + \frac{3}{4}\rho + \frac{5n}{\rho}.
\end{aligned}$$

Together with (4.4) this gives (1.8).

4.2. **The isotopy degree of a smooth hypersurface.**

Definition 32. *Let $Z \subset D$ be a smooth, compact submanifold without boundary. We define the number* $\deg_{\mathrm{iso}}(Z, D)$*, the* isotopy degree *of Z in D, as the minimum degree of a polynomial p such $(D, Z) \sim (D, Z(p))$.*

Remark 33. We observe that the isotopy degree of a hypersurface defined by a polynomial of degree d can be smaller than d. In fact, following [DL18], one can prove that *for most* polynomials p of degree d we have

$$\deg_{\mathrm{iso}}(Z(p)) = O\left(\sqrt{d \log d}\right),$$

in the following sense. First, we can put a gaussian measure μ on the space of polynomials of degree d by defining for every open set A in $\mathbb{R}[x_1, \ldots, x_n]$ its measure by

$$\mu(A) = \frac{\int_A e^{-\|p\|^2_{\mathrm{FS}}} \mathrm{d}\lambda}{\int_{\mathbb{R}[x_1,\ldots,x_n]} e^{-\|p\|^2_{\mathrm{FS}}} \mathrm{d}\lambda},$$

where λ is the Lebesgue measure on $\mathbb{R}[x_1, \ldots, x_n] \simeq \mathbb{R}^N$ (the identification is made through the list of coefficients) and $\|\cdot\|_{\mathrm{FS}}$ denotes the Fubini-Study norm: this norm is induced by a scalar product for which

$$\left\{\sqrt{\frac{d!}{\alpha_1!\cdots\alpha_n!(d-|\alpha|)!}}\cdot x_1^{\alpha_1}\cdots x_n^{\alpha_n}\right\}_{|\alpha|\le d}$$

form an orthonormal basis. Let now $p\in\mathbb{R}[x_1,\dots,x_n]$ be a polynomial of degree d and denote by $x=(x_0,x_1,\dots,x_n)$ and by $h\in\mathbb{R}[x_0,\dots,x_n]$ the homogenization of p, the new variable being x_0. Then we can decompose h into its spherical harmonics part

$$h=h_d+\|x\|^2\cdot h_{d-2}+\|x\|^4\cdot h_{d-4}+\cdots.$$

Denote by $\tilde h$ the projection of h on the space of harmonics of degree at most $\sqrt{bd\log d}$, where $b>0$ is a positive constant (defined in [DL18, Proposition 6]):

$$\begin{aligned}\tilde h(x_0,\dots,x_n)&=\sum_{\ell\le\sqrt{bd\log d}}\|x\|^{d-\ell}h_\ell(x_0,\dots,x_n)\\&=\|x\|^{d-\lfloor\sqrt{bd\log d}\rfloor}\sum_{\ell\le\sqrt{bd\log d}}\|x\|^{\lfloor\sqrt{bd\log d}\rfloor-\ell}h_\ell(x_0,\dots,x_n)\\&=\|x\|^{d-\lfloor\sqrt{bd\log d}\rfloor}\cdot q(x_0,\dots,x_n).\end{aligned}$$

Here q is a homogeneous polynomial of degree bounded by $O(\sqrt{d\log d})$. Then, by [DL18, Theorem 7], there is a set $S_d\subset\mathbb{R}[x_0,\dots,x_n]$ of homogeneous polynomials of degree d such that $\mu(S_d)\to 1$ as $d\to\infty$ (i.e. with almost full measure) with the property that for every $h\in S_d$

$$(S^n,Z(h))\sim(S^n,Z(\tilde h))=(S^n,Z(q)).$$

When we set $x_0=1$ in q we obtain a polynomial $\tilde p$ on $\mathbb{R}^n\simeq\{x_0=1\}$, whose zero set is isotopic to $Z(p)$ and with degree $O(\sqrt{d\log d})$.

4.3. **Proof of Theorem 8.** We will first prove the following preliminary estimate.

For $d,n>0$ denote by $\mathcal{Z}_{d,n}$ the set of rigid isotopy classes of pairs $(\mathbb{R}^n,Z(p))$ with $Z(p)\subset\mathbb{R}^n$ regular zero set of a polynomial $p\in\mathbb{R}[x_1,\dots,x_n]$ of degree at most d. We claim that:

$$\#\mathcal{Z}_{d,n}\le 2T(2T-1)^{\ell-1}, \tag{4.7}$$

where $T=(n+1)(d-1)^n$ and $\ell=\binom{n+d+1}{n+1}$.

The cardinality of $\mathcal{Z}_{d,n}$ is bounded by the number of connected components of the complement of a discriminant in the space of polynomials. More precisely, denoting by

$$\Delta_{d,n}\subset\mathbb{R}[x_1,\dots,x_n]\simeq\mathbb{R}^\ell$$

the discriminant for $Z(p)$ being nonsingular, the number of rigid isotopy classes of $(\mathbb{R}^n,Z(p))$ is bounded by $b_0(\mathbb{R}^\ell\backslash\Delta_{d,n})$. Denoting by $\widehat{\Delta}_{d,n}\subset S^N$ the one point compactification of $\Delta_{d,n}$, we have (using Alexander duality):

$$\begin{aligned}\#\mathcal{Z}_{d,n}&\le b_0(\mathbb{R}^\ell\backslash\Delta_{d,n})\\&=b_0(S^\ell\backslash\widehat{\Delta}_{d,n})\\&=\tilde b_0(S^\ell\backslash\widehat{\Delta}_{d,n})+1\\&\le\tilde b(S^\ell\backslash\widehat{\Delta}_{d,n})+1\\&=\tilde b(\hat{\Delta}_{d,n})+1\\&=b(\widehat{\Delta}_{d,n}).\end{aligned}$$

In particular, in order to estimate $\#\mathcal{Z}_{d,n}$ it is enough to estimate the total Betti number of $\widehat{\Delta}_{d,n}$. To this end we will use a Mayer-Vietoris argument and write:

$$\widehat{\Delta}_{d,n} = \Delta_{d,n} \cup \left(U_\infty \cap \widehat{\Delta}_{d,n}\right),$$

where $U_\infty \subset S^\ell$ is an open ball centered at the point at infinity in S^ℓ. Observe that $U_\infty \cap \widehat{\Delta}_{d,n}$ is contractible and

$$\Delta_{d,n} \cap \left(U_\infty \cap \widehat{\Delta}_{d,n}\right) \sim Z(\mathrm{disc}_{d,n}, \|\cdot\|^2 = R),$$

meaning that the two spaces are homotopy equivalent; here $\mathrm{disc}_{d,n}$ is a polynomial on $\mathbb{R}^\ell$ whose zero set is $\Delta_{d,n}$ and $\|\cdot\|^2 = R$ defines a sphere on the same space (the boundary of U_∞, viewed as a subset of $\mathbb{R}^\ell$). In particular both $\Delta_{d,n}$ and $\Delta_{d,n} \cap \left(U_\infty \cap \widehat{\Delta}_{d,n}\right)$ are described in $\mathbb{R}^N$ by polynomial equations of degree bounded by [EH16, Proposition 7.4]:

$$\deg(\mathrm{disc}_{d,n}) = (n+1)(d-1)^n = T.$$

Using Mayer-Vietoris and [Mil64], it follows that:

$$\begin{aligned}
b(\widehat{\Delta}_{d,n}) &\le b(\Delta_{d,n}) + b\left(U_\infty \cap \widehat{\Delta}_{d,n}\right) + b\left(\Delta_{d,n} \cap \left(U_\infty \cap \widehat{\Delta}_{d,n}\right)\right) \\
&\le T(2T-1)^{\ell-1} + 1 + T(2T-1)^{\ell-1} \\
&\le 2T(2T-1)^{\ell-1},
\end{aligned}$$

which proves the claim (4.7).

For $d, n > 0$ denote now by $\mathcal{C}_{d,n}$ the set of rigid isotopy classes of pairs (D, C) with $C \subset \mathrm{int}(D) \subset \mathbb{R}^n$ a smooth component of the zero set $Z(p)$ of a polynomial $p \in \mathbb{R}[x_1, \ldots, x_n]$ of degree at most d. We claim now that:

$$\#\mathcal{C}_{d,n} \le d(2d-1)^{n-1}\#\mathcal{Z}_{d,n}. \tag{4.8}$$

In order to see this, observe first that if C is a smooth component of $Z(p)$ we can slightly perturb p within the space of polynomials of the same degree, without changing the rigid isotopy class of (D, C) and making the whole zero set $Z(p)$ smooth (i.e. we can assume $Z(p)$ is smooth already). Also, notice that the inclusion $(D, C) \hookrightarrow (\mathbb{R}^n, C)$ gives a correspondence of rigid isotopy classes and we can work in the whole $\mathbb{R}^n$ instead of the interior of the disk D (being the two ambient spaces diffeomorphic). To every rigid isotopy class of pairs $(\mathbb{R}^n, Z(p))$ corresponds at most $b_0(Z(p))$ rigid isotopy classes of pairs $(\mathbb{R}^n, C)$ with C a connected component of $Z(p)$ and this, together with the fact that $b_0(Z(p)) \le d(2d-1)^{n-1}$, gives (4.8).

By part Theorem 1 we know that, given a pair $(D, Z(f))$ with $f = 0$ regular, $Z(f) \subset \mathrm{int}(D)$ and $\kappa_2(f, D) \le \kappa$, there exists a polynomial p with

$$d = \deg(p) \le (1 + 2a_2(D) \cdot \kappa) := k$$

such that the pairs $(D, Z(f))$ and $(D, Z(p))$ are rigidly isotopic. In particular, $\#(\kappa, D)$ is bounded by $\#\mathcal{C}_{\lfloor 1+2a_2(D)\cdot\kappa\rfloor, n}$. Combining (4.8) and (4.7) we get:

$$\begin{aligned}
\#(\kappa, D) &\le d(2d-1)^{n-1} 2T(2T-1)^{\ell-1} \\
&\le 2^n d^n (n+1)(d-1)^n \left(2(n+1)d^n\right)^{\frac{(d+n+1)^{n+1}}{(n+1)!}} \\
&\le c(n)^{\frac{(k+n+1)^{n+1}}{(n+1)!}} k^{2n + \frac{(k+n+1)^{n+1}}{(n+1)!}} =: (*),
\end{aligned}$$

so that as $k \to +\infty$ we get:

$$
\begin{aligned}
(*) &\leq k^{3\frac{(2k)^{n+1}}{(n+1)!}} \\
&\leq k^{c'(n)k^{n+1}} \\
&= (1+2a_2(D)\cdot\kappa)^{c'(n)(1+2a_2(D)\cdot\kappa)^{n+1}} \\
&\leq (c''(n)\kappa)^{c''(n)\kappa^{n+1}} \\
&\leq \kappa^{c'''(n)\kappa^{n+1}}.
\end{aligned}
$$

By the continuity of the expression $(*)$ with respect to κ, we conclude that there is a constant $C_1(n)$ such that if $\kappa > C_1(n)$ then $\#(\kappa, D) \leq \kappa^{C_2(n)\kappa^{n+1}}$, where $C_2(n) = c'''(n)$ is the constant found before. This concludes the proof.

4.4. **Proof of Theorem 9.** For what regards the first part of the theorem, we will show that

$$b(Z(f)) \leq \left(a_1(D)\kappa^{(1)}(f, D) + 1\right)^n,$$

where $a_1(D)$ is the constant given by Theorem 23. This implies (1.10), since $\kappa^{(1)}(f, D) \geq 1$, by definition.

Fix $\varepsilon > 0$. First, observe that if $\varepsilon < \delta(f, D)$, then f and $f|_{\partial D}$ have no critical value in the interval $(-\varepsilon, \varepsilon)$, from which it follows that the set $E = f^{-1}(-\varepsilon, \varepsilon)$ is entirely contained in the interior of D. Moreover E is a tubular neighborhood of $Z(f)$, since by Morse theory $f^{-1}(-\varepsilon, \varepsilon)$ is diffeomorphic to $Z(f) \times (-\varepsilon, \varepsilon)$. To see that E satisfies the hypotheses of theorem 21, define $E_1 = f^{-1}(-\varepsilon_1, \varepsilon_1)$, where $\varepsilon < \varepsilon_1 < \delta(f, D)$ and notice that then $E_1 \subset \mathrm{int}(D)$ is a tubular neighborhood of $Z(f)$ such that $\overline{E} \subset E_1$.

Let $p \in \mathbb{R}[x_1, \ldots, x_n]$ be a polynomial such that $\|f - p\|_{C^0(D,\mathbb{R})} < \varepsilon$. By Theorem 23 we can assume that its degree d satisfies the bound

$$d - 1 \leq a_1(D)\|f\|_{C^1(D,\mathbb{R})}\frac{1}{\varepsilon} \tag{4.9}$$

(take $p = p_d(f)$, where d is the biggest positive integer such that (4.9) is true). Let $F_t = f + t(p - f)$ and call F_t its restriction to $M = \mathrm{int}(D)$. Consider the set $\mathcal{U}_E$ defined as in (2.10), where $N = \mathbb{R}$ and $Y = \{0\}$. Suppose that $F_t \in \mathcal{U}_E$ for every $t \in [0, 1]$, then we could apply Theorem 21 to deduce that

$$b(Z(f)) \leq b(Z(p)) \leq \left(\frac{a_1(D)\|f\|_{C^1(D,\mathbb{R})}}{\varepsilon} + 1\right)^n$$

where the second inequality is due to the Milnor-Thom bound [Mil64] and to (4.9). The thesis now would follow by the arbitrariness of ε.

Thus to conclude the proof it is sufficient to show that $F_t^{-1}(0) \subset E$. To see this, let $x \in D$ such that $F_t(x) = 0$ and observe that then

$$|f(x)| = |F_t(x) - t(p(x) - f(x))| \leq \|f - p\|_{C^0(D,\mathbb{R})} < \varepsilon.$$

Let us turn to the second statement of the theorem. Assume for simplicity that D is the standard unit disk in $\mathbb{R}^n$. We will show that for any given compact hypersurface $Z \subset D$ defined by a regular C^1 equation $f = 0$ such that $f = 1$ near ∂D, there exists a sequence of smooth functions f_m such that

$$b(Z(f_m)) \geq \frac{b(Z)}{\kappa^{(1)}(f, D)^n} h(D)\kappa^{(1)}(f_m, D)^n,$$

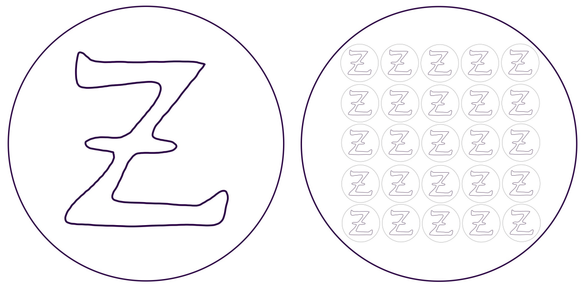

FIGURE 1. $Z(f_m)$, on the right, is the disjoint union of many copies of Z.

where $h(D)$ is the infimum among the numbers $N\varepsilon^n$, such that there exists a collection of N disjoint n-dimensional disks of radius ε contained in D.

Extend f to the whole space $\mathbb{R}^n$, by setting $f(x) = 1$ for all $x \notin D$. Let $m \in \mathbb{N}$ and define $f_{m,z} \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R})$ to be the function

$$f_{m,z}(x) = f\left(m(x - z)\right),$$

so that the submanifold $Z(f_{m,z})$ is contained in the interior of the disk of radius m^{-1} centered in the point z and it is diffeomorphic to Z. Moreover we can observe that, with $m \geq 1$, we have the inequalities

$$\delta(f_{m,z}, D) = \delta(f_{m,0,D}) = \inf_{x \in D} \left(|f(x)|^2 + m^2\|\nabla f(x)\|^2\right)^{1/2} \geq \delta(f, D);$$
$$\|f_{k,z}\|_{C^1(D,\mathbb{R})} = \|f_{k,0}\|_{C^1(D,\mathbb{R})} = \sup_{x \in D} \left(|f(x)|^2 + m^2\|\nabla f(x)\|^2\right)^{1/2} \leq m\|f\|_{C^1(D,\mathbb{R})};$$

therefore $\kappa^{(1)}(f_{m,z}, D) \leq m\kappa^{(1)}(f, D)$.

For any $m \in \mathbb{N}$, choose a finite family I_m of points $z_{m,i} \in D$, such that the disks $D_{m,i}$, centered in $z_{m,i}$ and with radius m^{-1}, are disjoint. Since the (Hausdorff) dimension of D^n is n, we can assume that the number of points in such a family is $\#(I_m) \geq h(D)m^n$. Define the function $f_m \in \mathcal{C}^1(D, \mathbb{R})$ as

$$f_m = 1 - \#(I_m) + \sum_{i \in I_m} f_{m,z_i},$$

so that f_m coincides with $f_{m,i}$ on the disk $D_{m,i}$ and it is constantly equal to 1 outside the union of all disks.

It follows that $\kappa^{(1)}(f_m, D) \leq m\kappa^{(1)}(f, D)$ and that the sequence $\kappa^{(1)}(f_m, D)$ is divergent as $m \to +\infty$. Moreover, the zero set of f_m is homeomorphic to a disjoint union of $\#(I_m)$ copies of Z, hence

$$b(Z(f_m)) \geq \#(I_m)b(Z) \geq h(D)b(Z)m^n.$$

Putting these two observations together we conclude that

$$b(Z(f_m)) \geq h(D)b(Z)\left(\frac{\kappa^{(1)}(f_m, D)}{\kappa^{(1)}(f, D)}\right)^n.$$

4.5. **Proof of Corollary 10.**

Remark 34. Combining Theorem 9 with Theorem 5, one can obtain the estimate for the Betti numbers of Z from Corollary 10. We observe that it is possible to obtain a similar estimate also from the work [NSW08], let us sketch how.

To start with, [NSW08, Proposition 3.1] claims that there exists $\epsilon > 0$ with

$$\epsilon < \sqrt{\frac{3}{5}}\rho(Z) \tag{4.10}$$

such that, if $x_1, \dots, x_m \in \mathbb{R}^n$ is a collection of points which are $\frac{\epsilon}{2}$–dense in Z, then the inclusion

$$Z \hookrightarrow \bigcup_{i=1}^{m} B(x_1, \epsilon) =: U$$

is a homotopy equivalence. Since each ball is contractible, the set U is also homotopy equivalent to the nerve complex C of the cover $\{B(p_i, \epsilon), i = 1, \dots, m\}$, and in particular $b(Z) = b(C)$.

The number of vertices of this complex, i.e. m, can be chosen to be of the order:

$$m \le a_1(n)\frac{\mathrm{vol}(Z)}{\epsilon^{n-1}}.$$

Moreover we can chose the $\frac{\epsilon}{2}$–net to also satisfy the following: every ball $B(p_i, \epsilon)$ intersects at most $a_2(n) > 0$ other balls from the family. In particular, each vertex of C belongs to at most $a_2(n)$ cells, and:

$$b(Z) = b(C) \le a_2(n)m \le a_3(n)\frac{\mathrm{vol}(Z)}{\epsilon^{n-1}}. \tag{4.11}$$

Now, by Weyl's tube formula:

$$\mathrm{vol}(\mathcal{U}(Z, \rho(Z))) = \mathrm{vol}(Z)\rho(Z) \le \mathrm{vol}(D_{R+\rho}),$$

from which we get that

$$\mathrm{vol}(Z) \le \frac{\mathrm{vol}(D_{R+\rho})}{\rho}.$$

Choose now $\epsilon = c_4\rho(Z)$ such that (4.10) is satisfied. Then, combining (4.11) with (4.10) we get:

$$b(Z) \le a_3(n)\frac{\mathrm{vol}(Z)}{\epsilon^{n-1}} = a_5(n)\frac{\mathrm{vol}(Z)}{\rho(Z)^{n-1}} \le a_6(n)\frac{\mathrm{vol}(D_{R+\rho})}{\rho(Z)^n}.$$

Proof of Corollary 10. Observe that $\rho(Z)$ cannot be greater than the radius of D unless Z is empty, in which case there is nothing to prove. Define D' to be the disk with a double radius than that of D, so that Z and D' satisfy the hypotheses of Theorem 5, thus there exists a function f such that

$$b(Z) = b\left(Z(f)\right) \le \left(c_4(D') \cdot \kappa_1(f, D')\right)^n \le c_4(D') \cdot 2^n \left(1 + \frac{1}{\rho(Z)}\right)^n,$$

where the first inequality is implied by Theorem 9. Taking $c_6(D) = 2^n c_4(D')$ we conclude the proof. $\square$

Appendix A. Global polynomial approximation of hypersurfaces

The aim of this Section is to prove Theorem 36, which gives a quantitative bound for the degree of a polynomial approximating a smooth hypersurface on the whole $\mathbb{R}^n$. We need first the following Lemma.

Lemma 35. *Let $D \subset \mathbb{R}^n$ be a euclidean disk of radius R. For every $\ell, n, d \geq 0$ there exists a constant $c_\ell(n,d) > 0$ such that if $p \in \mathbb{R}[x_1, \ldots, x_n]$ of degree $\deg(p) = d$, then:*

$$|p(x)| \leq c_\ell(n,d)\, \|p\|_{C^\ell(D,\mathbb{R})} \left(\frac{\|x\|}{R}\right)^d (1 + R^\ell) \quad \forall \|x\| \geq R > 0.$$

Proof. Let $V_{n,d} \subset \mathbb{R}[x_1, \ldots, x_n]$ be the space of polynomials of degree at most d. Since $V_{n,d}$ is a finite-dimensional vector space the norm $\|\cdot\|_{C^\ell(D_1,\mathbb{R})}|_{V_{n,d}}$ (here D_1 is the unit disk) and the norm $\|\cdot\|_{\text{coeff}}$, given by the "maximum of the modulus of the coefficients", are equivalent and there exists a constant $a_1(n,d) > 0$ such that:

$$\|f\|_{\text{coeff}} \leq a_1(n,d) \|f\|_{C^\ell(D_1,\mathbb{R})}.$$

Let now $p_R(y) = p(Ry)$. Then for every $\|y\| \geq 1$:

$$\begin{aligned} |p_R(y)| &\leq \dim(V_{n,d})\, \|p_R\|_{\text{coeff}}\, \|y\|^d \\ &\leq \dim(V_{n,d})\, a_1(n,d)\, \|p_R\|_{C^\ell(D_1,\mathbb{R})}\, \|y\|^d. \end{aligned}$$

On the other hand we have:

$$\begin{aligned} \|p_R\|_{C^\ell(D_1,\mathbb{R})} &= \max_{\|y\|\leq 1} \left(\sum_{|\alpha|\leq \ell} |\partial^\alpha p_R(y)|^2 \right)^{1/2} \\ &= \max_{\|y\|\leq 1} \left(\sum_{|\alpha|\leq \ell} \left|\partial^\alpha p(Ry) R^{|\alpha|}\right|^2 \right)^{1/2} \\ &\leq \max_{\|Ry\|\leq R} \left(\sum_{|\alpha|\leq \ell} |\partial^\alpha p(Ry)|^2 \right)^{1/2} (R^\ell + 1) \\ &= \|p\|_{C^\ell(D_1,\mathbb{R})}(R^\ell + 1). \end{aligned}$$

This gives:

$$|p(x)| = |p_R(x/R)| \leq \dim(V_{n,d})\, a_1(n,d)\, (R^\ell + 1) \|p\|_{C^\ell(D_R,\mathbb{R})} \left(\frac{\|x\|}{R}\right)^d.$$

Defining the constant $c_\ell(n,d) := \dim(V_{n,d})\, a_1(n,d)$ gives the claim. $\square$

Theorem 36. *Let D be a disk of radius $R > 0$ centered at the point z_0 and consider $f \in C^2(D,\mathbb{R})$ such that the equation $f = 0$ is regular in D and $Z(f) \subset \operatorname{int}(D)$. Let also $\tau = \tau(f,D) > 0$ be such that:*

$$e^{-3\tau} R = \max_{z \in Z(f)} \|z - z_0\|$$

and set $r = e^{-2\tau} R < R$. Denote by D_r the disk with the same center of D and with radius r. Let $c_1(n,d) > 0$ be the constant from Lemma 35 and define:

$$\tilde{\kappa}^{(\ell)} = \max\{\kappa^{(\ell)}(f,D), \kappa^{(\ell)}(f,D_r)\}.$$

There exists a polynomial $p \in \mathbb{R}[x_1, \dots, x_n]$ *with*

$$\deg(p) \le \max\left\{ r+1, \tilde{k}^{(2)} \cdot 2a_2(D), \frac{\log \tilde{k}^{(1)} + \log \frac{1}{\tau} + \log\left(8 + \frac{8}{Re^{-\tau}} + 4Re^{-\tau} + 4\tau\right) + \log\left(c_1(n,d)\right)}{\tau} \right\}$$

such that:

$$(\mathbb{R}^n, Z(p)) \sim (\mathbb{R}^n, Z(f)).$$

Remark 37. The estimate above is more interesting when $\tau \to 0$ and $\tilde{k}^{(1)} \to +\infty$, in which case it implies the simpler inequality

$$\deg(p) \le C \frac{\tilde{k}^{(2)}}{\tau^2}$$

Proof of Theorem 36. Assume that D is centered at 0. The first step of the proof is to argue as in the proof of Theorem 1 and find a polynomial $p_0 \in \mathbb{R}[x_1, \dots, x_n]$ such that

$$\|p_0 - f\|_{C^1(D,\mathbb{R})} < \frac{1}{2} \min\left\{\delta(f, D), \delta(f, D_r)\right\} =: \delta,$$

so that $(D, Z(p_0) \cap D) \cong (D, Z(f))$ and, at the same time, $Z(p_0) \cap D \subset \operatorname{int}(D_r)$. Thus we can assume that $p_0(D \setminus D_r) > 0$, since p_0 has no zeroes in that region. Moreover, thanks to Theorem 23 we can estimate the degree d of p_0:

$$\begin{aligned} d &\le \max\left\{ r+1, \frac{\|f\|_{C^2(D,\mathbb{R})}}{\min\left\{\delta(f,D), \delta(f,D_r)\right\}} \cdot 2a_{r+2(D)} \right\} \\ &= \max\left\{ r+1, \max\left\{ \frac{\|f\|_{C^2(D,\mathbb{R})}}{\delta(f,D)}, \frac{\|f\|_{C^2(D,\mathbb{R})}}{\delta(f,D_r)} \right\} \cdot 2a_{r+2(D)} \right\} \\ &= \max\left\{ r+1, \tilde{k} \cdot 2a_{r+2}(D) \right\}. \end{aligned} \tag{A.1}$$

Now we have to modify p_0 in order to eliminate those components of its zero set $Z(p_0)$ that are not contained in D. To this end we define a new polynomial $p \in \mathbb{R}[x_1, \dots, x_n]$ such that

$$p = p_0 + a\left(\frac{|x|^2}{s^2}\right)^{\ell},$$

for some $r < s < R$, $a > 0$ and $\ell \in \mathbb{N}$. We need to take ℓ and a so big that

(1) $p(x) > 0$ for every $x \notin \operatorname{int}(D)$. This ensures that $Z(p) = Z(p) \cap D$.
(2) $\|p - p_0\|_{C^1(D,\mathbb{R})} < \frac{1}{2}\delta(f, D)$, so that $\|p - f\|_{C^1(D,\mathbb{R})} < \delta(f, D)$ and thus $(D, Z(p) \cap D) \cong (D, Z(f))$ by Lemma 20.

Combined with the fact that $Z(f) \subset \operatorname{int}(D)$, the two conditions above imply that

$$(\mathbb{R}^n, Z(p)) \cong (\mathbb{R}^n, Z(f)).$$

Thus it remains to estimate the degree of such a polynomial p.

Let $x \notin D$. By Lemma 35 we have

$$\begin{aligned} p(x) &\ge a\left(\frac{|x|}{s}\right)^{2\ell} - |p_0(x)| \\ &\ge a\left(\frac{|x|}{s}\right)^{2\ell} - c_1(n,d)(1+s)\|p_0\|_{C^1(D_s,\mathbb{R})}\left(\frac{|x|}{s}\right)^{d}, \end{aligned}$$

where D_s is the disk of radius s. Therefore condition (1) is certainly satisfied for all $\ell > \frac{1}{2}d$ and

$$\begin{aligned} a &\geq c_1(n,d)(1+s)\cdot 2\|f\|_{C^1(D,\mathbb{R})} \\ &\geq c_1(n,d)(1+s)\cdot \left(\|f\|_{C^1(D_s,\mathbb{R})}+\delta\right) \\ &\geq c(n,d)(1+s^2)\cdot \|p_0\|_{C^1(D_s,\mathbb{R})}. \end{aligned} \tag{A.2}$$

Lemma 38. *For any $\rho > 1$ and $\ell \in \mathbb{N}$,*

$$\ell \log \rho \leq \rho^\ell - 1.$$

Proof. The function $\ell \mapsto \varphi(\ell) = \ell \log \rho - \rho^\ell + 1$ takes the value $\varphi(0) = 0$ at $\ell = 0$ and has negative derivative for all $\ell > 0$:

$$\varphi'(\ell) = \log \rho - \log \rho \cdot \rho^\ell = -(\rho^\ell - 1)\log \rho \leq 0.$$

$\square$

Applying the previous Lemma with $\rho = \frac{s}{r} > 1$, we obtain the inequality

$$\ell \leq \left(\left(\frac{s}{r}\right)^\ell - 1\right)\frac{1}{\log\left(\frac{s}{r}\right)}.$$

Therefore

$$\begin{aligned} \|p-p_0\|_{C^2(D_r,\mathbb{R})} &\leq a\left(\frac{r}{s}\right)^{2\ell} + a2\ell\frac{r^{2\ell-1}}{s^{2\ell}} \\ &= a\left(\frac{r}{s}\right)^{2\ell}\left(1+\frac{2\ell}{r}\right) \\ &\leq a\left(\frac{r}{s}\right)^{2\ell}\left(1+2\frac{\left(\frac{s}{r}\right)^{2\ell}-1}{r\log\left(\frac{s}{r}\right)}\right) \\ &\leq a\left(\left(\frac{r}{s}\right)^{2\ell}+\frac{2}{r\log\left(\frac{s}{r}\right)}\right) \\ &\leq a\left(1+\frac{2}{r\log\left(\frac{s}{r}\right)}\right)\left(\frac{r}{s}\right)^{2\ell}. \end{aligned}$$

To ensure that condition (2) is satisfied it is thus sufficient to assume that

$$a\left(1+\frac{2}{r\log\left(\frac{s}{r}\right)}\right)\left(\frac{r}{s}\right)^{2\ell} \leq \delta,$$

which, combined with (A.2), becomes

$$\begin{aligned} \left(\frac{r}{s}\right)^{2\ell} &\leq \left(1+\frac{2}{r\log\left(\frac{s}{r}\right)}\right)^{-1}\frac{\delta}{2c_1(n,d)(1+s)\|f\|_{C^1(D,\mathbb{R})}}, \\ \text{i.e.}\quad \left(\frac{s}{r}\right)^{2\ell} &\geq \left(\frac{8}{r\log\left(\frac{s}{r}\right)}+4\right)c_1(n,d)(1+s)\tilde{k}^{(1)}, \\ \text{i.e.}\quad 2\ell &\geq \frac{\log \tilde{k}^{(1)} + \log\left(c_1(n,d)(1+s)\left(\frac{8}{r\log\left(\frac{s}{r}\right)}+4\right)\right)}{\log\left(\frac{s}{r}\right)}. \end{aligned}$$

By taking $s = e^{-\tau}R \in (r, R)$, we get that $\frac{s}{r} = e^{\tau}$ and obtain the final formula formula

$$2\ell \geq \frac{\log \tilde{k}^{(1)} + \log\left(c_1(n,d)\right) + \log\left((1+e^{-\tau}R)\left(\frac{8}{Re^{-\tau}\tau}+4\right)\right)}{\tau}$$
$$= \frac{\log \tilde{k}^{(1)} + \log\frac{1}{\tau} + \log\left(8 + \frac{8}{Re^{-\tau}} + 4Re^{-\tau} + 4\tau\right) + \log\left(c_1(n,d)\right)}{\tau}.$$

We conclude that if we take $2\ell \geq d$ and such that the previous inequality holds, then p satisfies conditions (1) and (2). Combining this fact with the estimate (A.1) on d we conclude the proof. □

References

[Bas99] S. Basu. On bounding the Betti numbers and computing the Euler characteristic of semi-algebraic sets. *Discrete Comput. Geom.*, 22(1):1–18, 1999. DOI: 10.1007/PL00009443

[BBL02] T. Bagby, L. Bos, and N. Levenberg. Multivariate simultaneous approximation. *Constr. Approx.*, 18(4):569–577, 2002. DOI: 10.1007/s00365-001-0024-6

[BCR98] Jacek Bochnak, Michel Coste, and Marie-Françoise Roy. *Real algebraic geometry*, volume 36 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1998. Translated from the 1987 French original, Revised by the authors.

[DFN85] B. A. Dubrovin, A. T. Fomenko, and S. P. Novikov. *Modern geometry—methods and applications. Part II*, volume 104 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1985. The geometry and topology of manifolds, Translated from the Russian by Robert G. Burns.

[DL18] D. N. Diatta and A. Lerario. Low degree approximations of random polynomials, 2018. arχiv: 1812.10137

[EH16] D. Eisenbud and J. Harris. 3264 and all that: A second course in algebraic geometry. *Cambridge University PRess*, 2016. DOI: 10.1017/CBO9781139062046

[Fef10] Charles Fefferman. The C^m norm of a function with prescribed jets I. *Rev. Mat. Iberoam.*, 26(3):1075–1098, 2010. DOI: 10.4171/RMI/628

[Foo84] Robert L. Foote. Regularity of the distance function. *Proc. Amer. Math. Soc.*, 92(1):153–155, 1984. DOI: 10.1090/S0002-9939-1984-0749908-9

[GT01] David Gilbarg and Neil S. Trudinger. *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition. DOI: 10.1007/978-3-642-61798-0

[Hir94] Morris W. Hirsch. *Differential topology*, volume 33 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1994. Corrected reprint of the 1976 original.

[Kol17] János Kollár. Nash's work in algebraic geometry. *Bull. Amer. Math. Soc. (N.S.)*, 54(2):307–324, 2017. DOI: 10.1090/bull/1543

[KP81] Steven G. Krantz and Harold R. Parks. Distance to C^k hypersurfaces. *J. Differential Equations*, 40(1):116–120, 1981. DOI: 10.1016/0022-0396(81)90013-9

[LS19] Antonio Lerario and Michele Stecconi. Maximal and typical topology of real polynomial singularities, 2019. arχiv: 1906.04444

[Mil64] J. Milnor. On the Betti numbers of real varieties. *Proc. Amer. Math. Soc.*, 15:275–280, 1964. DOI: 10.1090/S0002-9939-1964-0161339-9

[Nas52] John Nash. Real algebraic manifolds. *Ann. of Math. (2)*, 56:405–421, 1952. DOI: 10.2307/1969649

[NSW08] Partha Niyogi, Stephen Smale, and Shmuel Weinberger. Finding the homology of submanifolds with high confidence from random samples. *Discrete Comput. Geom.*, 39(1-3):419–441, 2008. DOI: 10.1007/s00454-008-9053-2

[Sei36] H. Seifert. Algebraische Approximation von Mannigfaltigkeiten. *Math. Z.*, 41(1):1–17, 1936. DOI: 10.1007/BF01180402

[Ste21] Michele Stecconi. Random differential topology, 2021. arχiv: 2110.15694

[Tho69] R. Thom. Ensembles et morphismes stratifiés. *Bull. Amer. Math. Soc.*, 75:240–284, 1969. DOI: 10.1090/S0002-9904-1969-12138-5

[Tog73] A. Tognoli. Su una congettura di Nash. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3)*, 27:167–185, 1973. NUMDAM: ASNSP_1973_3_27_1_167_0

[VA85] A.N. Varchenko V.I. Arnold, S.M. Gusein-Zade. *Singularities of Differentiable Maps*. Monographs in Mathematics. Birhäser, 1985. DOI: 10.1007/978-1-4612-5154-5, DOI: 10.1007/978-1-4612-3940-6

[Yom84] Y. Yomdin. The set of zeroes of an "almost polynomial" function. *Proceedings of the American Mathematical Society*, 90(4):538–542, 1984. DOI: 10.2307/2045026

[Yom85] Y. Yomdin. Global bounds for the Betti numbers of regular fibers of differentiable mappings. *Topology*, 24(2):145–152, 1985. DOI: 10.1016/0040-9383%2885%2990051-5

[Yom20] Yosef Yomdin. Smooth rigidity and Remez-type inequalities, 2020. arχiv: 2009.13994

[Yom21a] Y. Yomdin. Higher derivatives of functions vanishing on a given set, 2021. arχiv: 2108.02459

[Yom21b] Yosef Yomdin. Smooth rigidity and Remez inequalities via topology of level sets, 2021. arχiv: 2106.06961

Antonio Lerario, SISSA, Via Bonomea 265, 34136 Trieste, Italy
Email address: `lerario@sissa.it`

Michele Stecconi, Laboratoire de Mathématiques Jean Leray, 2 rue de la Houssinière, 44300, Nantes, France
Email address: `stecconi.michele@gmail.com`