ANNALES
POLONICI MATHEMATICI
87 (2005)

Some quantitative results in singularity theory

by Y. Yomdin (Rehovot)

To the memory of S. Łojasiewicz

Abstract. The classical singularity theory deals with singularities of various mathematical objects: curves and surfaces, mappings, solutions of differential equations, etc. In particular, singularity theory treats the tasks of recognition, description and classification of singularities in each of these cases.

In many applications of singularity theory it is important to sharpen its basic results, making them "quantitative", i.e. providing explicit and effectively computable estimates for all the important parameters involved. This opens new possibilities for applications in analysis, geometry, differential equations, dynamics, and, last not least, in computations.

Application of the results of singularity theory in numerical data processing with finite accuracy stresses another important requirement: the "normalizing transformations" must be explicitly computable. The most natural interpretation of this requirement is in terms of the "jet calculus": given the Taylor polynomials of the input data, we should be able to produce explicitly the Taylor polynomials of the output normalizing transformations.

This papers provides a sample of initial results in these directions.

1. Introduction. In this paper we give several sample results in singularity theory and their "quantitative" counterparts. Some of these quantitative statements are well known in one form or another (like "quantitative implicit function theorem" and "quantitative Sard theorem"), other, like "quantitative Morse theorem" are apparently new. Obviously, these results represent only a small part of the body of the modern singularity theory. (See [60], [1, 4, 20, 25, 39], [40]–[47], covering the main parts of the classical theory. We are not aware of recent general books in this field.)

Let us describe briefly what we expect from a "quantitative" result in singularity theory.

A singularity occurs when in the process of solution of a mathematical problem some denominator vanishes. Very frequently this problematic de-

2000 *Mathematics Subject Classification*: 58Kxx, 58C15.

Key words and phrases: singularity, normal form, normalizing transformation, Taylor polynomial, jet calculus, explicit bounds.

nominator is a determinant of a matrix. Basically in this case singularity theory tells us that in order to solve the problem in a robust way we have to perform the following steps:

1. Consider the higher order approximation of our problem, and in the appropriate "jet space" consider the behavior of our initial data with respect to the "singular stratum" where the initial determinant vanishes.

2. If this behavior is non-degenerate ("transversal"), then in many cases we can bring our singularity to a "normal form" by appropriate "normalizing coordinate transformations" (see Section 2 below). To find explicitly these normalizing coordinate transformations we have to invert a certain *non-degenerate matrix*, representing the non-degenerate transversal behavior of our initial data with respect to the singular stratum.

In order to make this approach applicable in finite accuracy computations it is not enough to know that our determinants are non-zero. We have to know *how well they are separated from zero*. Consequently, a "quantitative" result has to assume and provide explicit bounds on the "measure of non-degeneracy" of the data involved. As the normalizing coordinate transformations are concerned, a quantitative result has to provide explicit lower bounds on the size of the neighborhood where these transformations are defined, as well as explicit upper bounds on their derivatives.

On the other hand, as the applications of the results of singularity theory in numerical data processing with finite accuracy are concerned, the explicit computation of the "normalizing transformations" is required. The most natural interpretation of this requirement is in terms of the "jet calculus": given the Taylor polynomials of the input data, we should be able to produce explicitly the Taylor polynomials of the output normalizing transformations.

Below in each of the examples considered we answer both these requirements: the bounds on the size of the neighborhood and on the derivatives of the normalizing transformations are given, together with the explicit jet calculus formulae, producing the Taylor polynomials of these transformations from the Taylor polynomials of the input data. We give these formulae explicitly, although mostly they are rather lengthy: indeed, this is, in a sense, the output of singularity theory, ready for use in high order computations ([3, 17, 18]).

Of course, many of the classical works on singularity theory answer the above requirements and provide quantitative information. The importance of the explicit quantitative bounds has been stressed in the work of the founders of the theory (see [57], [40–47]. Especially this concerns the work of S. Łojasiewicz (see e.g. [35–38]). His results in singularity theory and in semi-analytic geometry always stressed the role of quantitative geometric information. In particular, the research interests of the author have been formed under strong influence of discussions with S. Łojasiewicz and of his work.

Recently a new approach to the metric bounds in real algebraic geometry (also inspired by Łojasiewicz's work) has been found in [8], as well as in [32–34] and in other publications of these authors. One can expect this approach to provide new quantitative results in the study of singularities. In particular, by this method explicit and pretty accurate bounds in the quantitative Sard theorem of Section 3 below have been obtained ([8]).

Other important approaches to making the results of singularity theory more applicable have recently been developed. We mention here only some of the relevant publications: [9–13], [26, 51, 52, 61].

Quantitative information about geometric and analytic structure of singularities and near-singularities, their distribution and behavior is important in many problems of analysis, geometry, differential equations, and dynamics. In particular, in differential dynamics, a number of "quantitative" problems have been posed by M. Gromov in the early eighties. These concerned a quantitative behavior of periodic points, estimates for the volume growth and entropy etc. (see [21–24]). A "quantitative Kupka–Smale theorem", bounding a typical quantitative behavior of periodic points and conjectured by M. Gromov, has been obtained in [64]. Very recently striking results in this direction have been obtained by Kaloshin [29–31].

Important applications of quantitative transversality in symplectic geometry appeared in papers of S. K. Donaldson ([14–16], see also [56]). These results have been further extended in [2, 28] and other publications.

The quantitative Sard theorem has recently been applied in [52] to the study of the exponential stability of the motion in near-integrable Hamiltonian dynamics. In particular, a version of the quantitative Morse theorem, rather similar to Theorem 4.1 below, has been obtained in [52]. We briefly discuss the approach of L. Niederman in Sections 3 and 4 below.

The detailed proofs of the results of Section 2 and of Theorem 4.2 and Proposition 4.1, and all the jet calculus formulae of Sections 2 and 4 below have been obtained by D. Cohen (see [7]).

2. Implicit function theorem. The first result of singularity theory was obtained long before this name has appeared: this is the implicit function theorem, and its special case—the inverse function theorem. These results provide, in particular, *a normal form of a differential mapping at its regular point.*

Let us recall that a "normal form" is the simplest form to which a given object can be brought by the allowed "normalizing transformations". Of course, in each specific case this informal definition is replaced by an appropriate formal one. Lists of normal forms are among the main outputs of singularity theory, and the quantitative version of normal forms plays an important role in our approach. In the present paper we do not stress the notion of a "quantitative normal form", although it appears implicitly in several results below.

PROPOSITION 2.1. *Let* $f : \mathbb{R}^n \to \mathbb{R}^m$, $n \geq m$, *be a* C^k*-mapping,* $k \geq 1$, *given in coordinate form by* $y_1 = f_1(x_1, \dots, x_n)$, $\dots$, $y_m = f_m(x_1, \dots, x_n)$, *and let the differential* $df(0)$ *of* f *at the origin* $0 \in \mathbb{R}^n$ *have the maximal rank* m. *Then there is a new* C^k*-coordinate system* $w_1, \dots, w_n$ *near the origin in* $\mathbb{R}^n$ *in which the mapping* f *is written as* $y_1 = w_1, \dots, y_m = w_m$.

This may look almost a tautology: we simply *take* $y_1, \dots, y_m$ as the first m new coordinates. The key point is that the existence of the inverse coordinate transformation allows us *to express the old coordinates through the new ones*. Indeed, using these expressions we can parametrize the solutions of the system of equations $f_i = 0$, as is done in the more standard versions of the implicit function theorem.

Numerous versions of this result are scattered in the literature (see [5, 6] for a very particular sample). Many of these versions are "quantitative" in the sense that they provide explicit estimates for the size of the new coordinate neighborhood and for the inverse transformation.

Below we prove some quantitative versions of the implicit function theorem, restricting ourselves to the cases where the proof is really simple and illustrative. As explained above, besides obtaining explicit estimates for the size of the new coordinate neighborhood and for the inverse transformations, we also insist on the requirement that *the Taylor coefficients of the "output" functions be given by the explicit formulae through the Taylor coefficients of the "input" functions.* Because of this requirement, and in order to have simple estimates of the "truncation error" (as we replace the function by its Taylor polynomial of a relatively small degree) we assume all our functions to be *real analytic* and work with their infinite Taylor series. In the proofs we mostly work in the *complex* domain.

It is important to stress here that all the results below can be proved under finite smoothness assumptions, using the appropriate order Taylor formula with one form or another of the remainder term.

Let us start with the inverse function theorem in one variable. The result of Theorem 2.1 below is well known (see, for example, [27]) but we give a different proof which we use later on.

THEOREM 2.1. *Let $y = f(x)$ be a real analytic function with $f(0) = 0$, represented by a convergent power series*

$$f(x) = \sum_{k=1}^{\infty} a_k x^k \tag{2.1}$$

with $|a_1| = a > 0$ and $|a_k| \le M(1/R)^k$, $k = 2, 3, \ldots$. Then the inverse function $x = g(y) = f^{-1}(y)$ exists and is analytic in the disk $D_{R_1} = \{|y| < R_1\}$, and it is represented there by the convergent power series

$$g(y) = f^{-1}(y) = \sum_{k=1}^{\infty} b_k y^k. \tag{2.2}$$

The coefficients of (2.2) *satisfy the inequality $|b_k| \le M_1(1/R_1)^k$, $k = 1, 2, \ldots$. Here*

$$M_1 = \min\left(R/2, \frac{R^2 a}{64M}\right), \qquad R_1 = \frac{3a}{4} M_1.$$

Proof. 1. First of all, let us show that inside the disk $D_{R/2}$ the second derivative $f''(x)$ is bounded by the constant $M_2 = 16M(1/R)^2$. Indeed, by (2.1) we have

$$f''(x) = \sum_{k=2}^{\infty} k(k-1) a_k x^{k-2},$$

and therefore

$$|f''(x)| \le \sum_{k=2}^{\infty} k(k-1)|a_k|\,|x|^{k-2} \le M\left(\frac{1}{R}\right)^2 \sum_{k=2}^{\infty} k(k-1)\left(\frac{|x|}{R}\right)^{k-2}.$$

Since $x \in D_{R/2}$ the last inequality gives us

$$|f''(x)| \le M\left(\frac{1}{R}\right)^2 \sum_{k=2}^{\infty} k(k-1)\left(\frac{1}{2}\right)^{k-2}.$$

But the series

$$\sum_{k=2}^{\infty} k(k-1)\left(\frac{1}{2}\right)^{k-2}$$

is just the value of the second derivative $\left(\frac{1}{1-x}\right)'' = \frac{2}{(1-x)^3}$ at $x = \frac{1}{2}$, i.e. 16. Finally, we get $|f''(x)| \le 16M(1/R)^2 = M_2$.

2. Now we fix $R' \le R/2$ in such a way that for x inside the disk $D_{R'}$ the derivative $f'(x)$ is close enough to $f'(0) = a_1$. More accurately, write $f'(x) - f'(0) = \int_0^x f''(t)\,dt$. Hence $|f'(x) - f'(0)| = |f'(x) - a_1| \le M_2|x|$. Now fix $R' = a/4M_2$. For $x \in D_{R'}$ we get $|f'(x) - a_1| \le a/4$. Thus, for any such x the derivative $f'(x)$ belongs to the disk D of radius $a/4$ around $f'(0) = a_1$. Notice that D is at distance exactly $3a/4$ from the origin.

3. Consider now any two points x_1, x_2 inside $D_{R'}$. We have

$$f(x_2) - f(x_1) = \int_{x_1}^{x_2} f'(x)\,dx = \int_0^1 f'(x_1 + t(x_2 - x_1))(x_2 - x_1)\,dt$$
$$= (x_2 - x_1)\int_0^1 f'(x_1 + t(x_2 - x_1))\,dt.$$

Now the integral $A = \int_0^1 f'(x_1 + t(x_2 - x_1))\,dt$ is a convex combination of the derivatives $f'(x_1 + t(x_2 - x_1))$ and since by step 2 each of these derivatives belongs to the convex disk D, we find that $A \in D$. Hence for any two points $x_1, x_2 \in D_{R'}$ we have $|A| \geq 3a/4$, and finally

$$|f(x_2) - f(x_1)| = |A(x_2 - x_1)| \geq \frac{3a}{4}\,|x_2 - x_1|. \tag{2.3}$$

4. Now we conclude from (2.3) that f is one-to-one on the disk $D_{R'}$, and $f(D_{R'})$ contains a disk D_{R_1} of radius $R_1 = (3a/4)R'$. Indeed, (2.3) with $x_1 = 0$ shows that the circle of radius R' is mapped by f into the curve outside the disk D_{R_1} which makes exactly one turn around this disk. By Rouché's theorem, $D_{R_1} \subset f(D_{R'})$.

5. Consequently, f^{-1} is defined in D_{R_1} and $f^{-1}(D_{R_1})$ is contained in $D_{R'}$. Hence $f^{-1}(y)$ is bounded in absolute value by R' for $y \in D_{R_1}$. Now f^{-1} *is analytic* in D_{R_1}. This follows by the direct computation of the first derivative $\frac{df^{-1}(y)}{dy} = \frac{1}{f'(x)}$. By the Cauchy formula we now get

$$|b_k| = \left| \int_{\partial D_{R_1}} \frac{f^{-1}(y)}{y^k}\,dy \right| \leq R'\left(\frac{1}{R_1}\right)^k = M_1\left(\frac{1}{R_1}\right)^k, \qquad k = 1, 2, \ldots.$$

This completes the proof of the theorem.

To get explicit expressions for the first Taylor coefficients b_k of the inverse function $g(y)$ through the Taylor coefficients a_k of the input function $f(x)$, we use a recurrence relation given in Proposition 2.2 below. It can be produced by substituting (2.1) into (2.2) and comparing the coefficients of the corresponding powers of x.

Proposition 2.2. *The coefficients b_l satisfy the equations*

$$b_1 a_1 = 1, \qquad \sum_{m=1}^{l} b_m \Big(\sum_{\sum j_n = l} \prod_{n=1}^{m} a_{j_n} \Big) = 0, \qquad l \geq 2.$$

These equations yield, in turn, the recurrence relation

$$b_l = -\frac{1}{a_1^l} \sum_{m=1}^{l-1} b_m \Big(\sum_{\sum j_n = l} \prod_{n=1}^{m} a_{j_n} \Big), \qquad l \geq 2.$$

The explicit formulae for $b_k,\ k=1,\dots,4$, *are as follows*:

$$\begin{aligned}
b_1 &= \frac{1}{a_1},\\
b_2 &= -\frac{a_2}{a_1^3},\\
b_3 &= -\frac{a_3}{a_1^4}+\frac{2a_2^2}{a_1^5},\\
b_4 &= -\frac{a_4}{a_1^5}+\frac{2a_2a_3}{a_1^7}+\frac{a_2^3}{a_1^7}+\frac{3a_3a_2}{a_1^6}+\frac{6a_2^3}{a_1^7}.
\end{aligned}$$

Let us now consider the case of two variables. We start with the implicit function theorem.

THEOREM 2.2. *Let $f(x,y)$ be a real analytic function with $f(0,0)=0$, represented by the power series*

$$f(x,y)=\sum_{k+l=1}^{\infty} a_{k,l}x^k y^l \tag{2.4}$$

with $a_{0,1}\neq 0$ and $|a_{k,l}|\le M(1/R)^{k+l}$, $k,l=1,2,\dots$. Then there is a unique analytic function $y=h(x)$ such that $h(0)=0$ and $f(x,h(x))\equiv 0$ for any x with $|x|\le R_1$. The function $h(x)$ is given by a convergent power series

$$h(x)=\sum_{k=1}^{\infty} b_k x^k \tag{2.5}$$

with $|b_k|\le M_1(1/R_1)^k$, $k=1,2,\dots$, where

$$M_1=\min\biggl(\frac{|a_{0,1}|R^2}{128M},\frac{R}{2}\biggr),\qquad R_1=\min\biggl(\frac{3}{4}\,\frac{M_1}{|\frac{a_{1,0}}{a_{0,1}}|+\frac14},M_1\biggr).$$

Proof. 1. We bound the second (partial) derivatives of f on the polydisk $\{|x|,|y|\le R/2\}$ by a constant $M_2=16M/R^2$, as in the proof of Theorem 2.1:

$$\begin{aligned}
|f''_{xy}| &= \Bigl|\sum_{k,l=1}^{\infty} a_{kl}klx^{k-1}y^{l-1}\Bigr|\\
&\le \Bigl|\sum_{k,l=1}^{\infty}\Bigl(\frac{R}{2}\Bigr)^{k+l-2} klMR^{-(k+l)}\Bigr| \le \frac{M}{R^2}\Bigl|\sum_{k,l=1}^{\infty} kl\Bigl(\frac12\Bigr)^{k+l-2}\Bigr|\\
&= \frac{M}{R^2}\Bigl|\sum_{k=1}^{\infty} k\Bigl(\frac12\Bigr)^{k-1}\Bigr|\Bigl|\sum_{l=1}^{\infty} l\Bigl(\frac12\Bigr)^{l-1}\Bigr| = \frac{16M}{R^2}.
\end{aligned}$$

In the same fashion f''_{xx} and f''_{yy} are also bounded by $M_2=16M/R^2$.

2. Fix $R' = |a_{01}|/8M_2$. When $|x|, |y| \le R'$ we have $|f'_y(x,y) - a_{0,1}| \le \frac{1}{4}|a_{0,1}|$. Indeed,

$$|f'_y(x,y) - f'_y(0,0)| = \Big| \int_0^y f''_{yy}(0,t)\,dt + \int_0^x f''_{xy}(t,y)\,dt \Big|$$
$$\le 2M_2R' = \frac{1}{4}\,|a_{01}|.$$

In particular, for $|x|, |y| \le R'$ the partial derivative $f'_y(x,y)$ does not vanish.

3. In the polydisk $\Delta = \{|x|, |y| \le R'\}$ the partial derivative $f'_x(x,y)$ is also bounded from above. Indeed, $|f'_x(x,y) - a_{1,0}| \le \frac{1}{4}|a_{0,1}|$, and hence $|f'_x(x,y)| \le |a_{1,0}| + \frac{1}{4}|a_{0,1}|$ for $(x,y) \in \Delta$.

4. As a result, as long as the solution $y = h(x)$ of $f(x,y) = 0$ remains in the polydisk Δ, the derivative $h'(x) = -f'_x(x,y)/f'_y(x,y)$ satisfies

$$|h'(x)| \le A = \frac{4}{3}\left(\left|\frac{a_{1,0}}{a_{0,1}}\right| + \frac{1}{4}\right).$$

5. Now, for $|x| \le R_1 = \min(R'/A, R')$ we get

$$|h(x)| \le |x| \max_{t\in[0,x]}(|h'(t)|) \le R_1A \le R'.$$

Hence, for $|x| \le R_1$, the point $(x,y) = (x,h(x))$ remains in Δ. This justifies a posteriori the above inequalities. In particular, $|h(x)|$ is bounded by R'.

6. By the Cauchy formula we now get $|b_k| \le M_1(1/R_1)^k$, $k = 1, 2, \ldots,$ where $M_1 = R'$ and R_1 were defined above. This completes the proof.

To get the explicit expressions for the first Taylor coefficients b_k of the implicit function $h(y)$ through the Taylor coefficients $a_{k,l}$ of the input function $f(x,y)$, we use a recurrence relation given in Proposition 2.3 below. It can be produced by substituting (2.5) into (2.4), equating the result to zero and then equating to zero the coefficients of all powers of x.

PROPOSITION 2.3. *The coefficients b_k satisfy*

$$\sum_{i+j\le l} a_{ij}\Big(\sum_{\sum_{k=1}^j i_k = l-i} \prod_{m=1}^j b_{i_m}\Big) = 0$$

and so

$$b_m = -\frac{1}{a_{01}} \sum_{2\le i+j\le l} a_{ij}\Big(\sum_{\sum_{k=1}^j i_k = l-i} \prod_{m=1}^j b_{i_m}\Big) = 0.$$

Explicitly,

$$b_1 = -\frac{a_{10}}{a_{01}},$$

$$b_2 = -\frac{a_{20}}{a_{01}} - \frac{a_{02}a_{10}^2}{a_{01}^3} - \frac{a_{11}a_{10}}{a_{01}^2},$$

$$b_3 = -\frac{a_{30}}{a_{01}} - \frac{2a_{20}a_{10}a_{02}}{a_{01}^3} - \frac{2a_{10}^3a_{02}}{a_{01}^5} - \frac{2a_{10}^2a_{11}}{a_{01}^4}$$
$$+\frac{a_{03}a_{10}^3}{a_{01}^4} + \frac{a_{11}a_{20}}{a_{01}^2} + \frac{a_{11}a_{02}a_{10}^2}{a_{01}^4} + \frac{a_{11}^2a_{10}}{a_{01}^3} - \frac{a_{12}a_{10}^2}{a_{01}^3} + \frac{a_{21}a_{10}}{a_{01}^2}.$$

Finally, we provide a quantitative version of the inverse function theorem in two variables.

THEOREM 2.3. *Let $f:\mathbb{R}^2\to\mathbb{R}^2$ be a real analytic mapping with $f(0,0) = (0,0)$, given in coordinate form by $u_1 = f_1(x,y)$, $u_2 = f_2(x,y)$. Here $f_i(x,y)$, $i=1,2$, are real analytic functions with $f_i(0,0)=0$, represented by the convergent power series*

$$f(x,y) = \sum_{k,l=1}^{\infty} \begin{pmatrix} c_{kl}x^ky^l \\ d_{kl}x^ky^l \end{pmatrix} \tag{2.6}$$

with $\left\|\binom{c_{kl}}{d_{kl}}\right\|_\infty \le M(1/R)^{k+l}$, $k,l=0,1,\ldots$. Assume in addition that the Jacobian of f at zero, i.e. the determinant of the matrix $Df(0)$, does not vanish. Define

$$\delta = \frac{1}{2}\,\frac{1}{\|Df(0)^{-1}\|}.$$

Then there is an inverse mapping $g=f^{-1}:\mathbb{R}^2\to\mathbb{R}^2$, given in coordinate form by $x=g_1(u_1,u_2)$, $y=g_2(u_1,u_2)$, which is defined and analytic in the polydisk $\Delta=\{|u_1|\le R_1, |u_2|\le R_1\}$. The functions $g_i(u_1,u_2)$ are represented by convergent power series

$$g(u_1,u_2) = \sum_{k,l=1}^{\infty} \begin{pmatrix} a_{kl}u_1^ku_2^l \\ b_{kl}u_1^ku_2^l \end{pmatrix} \tag{2.7}$$

with $\left\|\binom{a_{kl}}{b_{kl}}\right\|_\infty \le M_1(1/R_1)^{k+l}$, $k,l=0,1,\ldots$. Here

$$M_1 = \min\left(\frac{\delta}{M_2},\frac{R}{2}\right) = \min\left(\frac{\delta R^2}{64M},\frac{R}{2}\right), \qquad R_1 = \delta M_1.$$

Proof. The proof goes essentially along the lines of the proof of Theorem 2.1.

1. We bound the first and second partial derivatives of f_1, f_2 on the polydisk $|x|,|y|\le R/2$ by a constant $M_2 = 16M/R^2$, as in the proof of Theorem 2.3.

2. Consider a ball $\mathcal{B}_\delta$ in the space of 2×2 matrices, with respect to the Euclidean matrix norm, centered at the matrix $Df(0)$. Recall that δ has been defined as $1/2\|Df(0)^{-1}\|$. Therefore, for any $B\in\mathcal{B}_\delta$ and any vector v we have $\|Bv\|\ge\delta\|v\|$.

3. Fix $R'\le R/2$ in such a way that for x,y inside the polydisk $\Delta_{R'} = \{|x|,|y|\le R'\}\subset\mathbb{C}^2$ we have $Df(x,y)\in\mathcal{B}_\delta$. To do this we use the bound

on the second derivative of f obtained in step 1: $\|D^2 f\| \le M_2$. We have $Df(x,y) - Df(0,0) = \int_0^1 D^2 f(xt,yt)(v)\,dt$ where v is the vector (x,y) and the second differential $D^2 f(xt,yt)$ is applied to v. Hence

$$\|Df(x,y) - Df(0,0)\| \le M_2 \|(x,y)\|.$$

Now for x, y inside the polydisk $\Delta_{R'}$ with $R' = \delta/M_2$ we get $\|Df(x,y) - Df(0,0)\| \le \delta$ and hence $Df(x,y) \in \mathcal{B}_\delta$.

4. Now we conclude from steps 2 and 3 that f is one-to-one on the polydisk $\Delta_{R'}$ in $\mathbb{C}^2$, and $f(\Delta_{R'})$ contains a polydisk Δ_{R_1} of radius $R_1 = \delta R'$. To do this, for any two points (x_1,y_1) and (x_2,y_2) write the difference $f(x_2,y_2) - f(x_1,y_1)$ as the integral

$$f(x_2,y_2) - f(x_1,y_1) = \int_0^1 Df(x_1 + t(x_2 - x_1), y_1 + t(y_2 - y_1))(v)\,dt.$$

Here v is the vector $(x_2 - x_1, y_2 - y_1)$ and the differential

$$Df(x_1 + t(x_2 - x_1), y_1 + t(y_2 - y_1))$$

is applied to v. Notice that this integral can be rewritten as

$$\Big(\int_0^1 Df(x_1 + t(x_2 - x_1), y_1 + t(y_2 - y_1))\Big)(v) = D(v).$$

Now the integral is a *convex combination of the differentials*

$$Df(x_1 + t(x_2 - x_1), y_1 + t(y_2 - y_1)).$$

By step 3 each of these differentials belongs to $\mathcal{B}_\delta$, and since $\mathcal{B}_\delta$ is a convex set, we conclude that $D \in \mathcal{B}_\delta$. (See [5, 6], where similar convexity arguments are applied.)

By the property of $\mathcal{B}_\delta$ stated in step 3 we see that $\|D(v)\| \ge \delta\|v\|$. This inequality implies that f is one-to-one on $\Delta_{R'}$. Applying it to $(x_1,y_1) = (0,0)$ we deduce that $f(\Delta_{R'})$ contains a polydisk Δ_{R_1} of radius $R_1 = \delta R'$. To show this we notice that the topological degree of f is 1 on the boundary of $\Delta_{R'}$ with respect to any point of Δ_{R_1}.

5. Hence f^{-1} is analytic in Δ_{R_1} and $f^{-1}(\Delta_{R_1}) \subset \Delta_{R'}$. In particular, each component of $f^{-1}(y)$ is bounded in absolute value by R' for $y \in \Delta_{R_1}$.

6. By the Cauchy formula we now get the required bounds on the Taylor coefficients of f^{-1}:

$$|a_{kl}|, |b_{kl}| \le \frac{M_1}{R_1^{k+l}}.$$

Here

$$M_1 = R' = \min\Big(\frac{\delta}{M_2}, \frac{R}{2}\Big) = \min\Big(\frac{\delta R^2}{16M}, \frac{R}{2}\Big), \qquad R_1 = \delta R'.$$

Explicit expressions for the first Taylor coefficients b_{lk} of the inverse functions $g_i(u_1, u_2)$ through the Taylor coefficients a^i_{kl} of the input mapping $f(x, y)$ can be obtained as follows:

The relation $f \circ g = \mathrm{id}$ gives us the equations

$$\begin{aligned}
g_1(u_1,u_2) &= a_{10}f_1(x,y) + a_{01}f_2(x,y) + a_{20}f_1(x,y)^2 \\
&\quad + a_{11}f_1(x,y)f_2(x,y) + a_{02}f_2(x,y)^2, \\
g_1(u_1,u_2) &= a_{10}(c_{10}x + c_{01}y + c_{11}xy + c_{20}x^2 + c_{02}y^2) \\
&\quad + a_{01}(d_{10}x + d_{01}y + d_{11}xy + d_{20}x^2 + d_{02}y^2) \\
&\quad + a_{20}(c_{10}x + c_{01}y + c_{11}xy + c_{20}x^2 + c_{02}y^2)^2 \\
&\quad + a_{11}(c_{10}x + c_{01}y + c_{11}xy + c_{20}x^2 + c_{02}y^2) \\
&\quad \times (d_{10}x + d_{01}y + d_{11}xy + d_{20}x^2 + d_{02}y^2) \\
&\quad + a_{02}(d_{10}x + d_{01}y + d_{11}xy + d_{20}x^2 + d_{02}y^2)^2 = x, \\
g_2(u_1,u_2) &= b_{10}f_1(x,y) + b_{01}f_2(x,y) + b_{20}f_1(x,y)^2 \\
&\quad + b_{11}f_1(x,y)f_2(x,y) + b_{02}f_2(x,y)^2, \\
g_2(u_1,u_2) &= b_{10}(c_{10}x + c_{01}y + c_{11}xy + c_{20}x^2 + c_{02}y^2) \\
&\quad + b_{01}(d_{10}x + d_{01}y + d_{11}xy + d_{20}x^2 + d_{02}y^2) \\
&\quad + b_{20}(c_{10}x + c_{01}y + c_{11}xy + c_{20}x^2 + c_{02}y^2)^2 \\
&\quad + b_{11}(c_{10}x + c_{01}y + c_{11}xy + c_{20}x^2 + c_{02}y^2) \\
&\quad \times (d_{10}x + d_{01}y + d_{11}xy + d_{20}x^2 + d_{02}y^2) \\
&\quad + b_{02}(d_{10}x + d_{01}y + d_{11}xy + d_{20}x^2 + d_{02}y^2)^2 = y,
\end{aligned}$$

Comparing the coefficients of the monomials of these equations gives

$$\begin{aligned}
a_{10}c_{10} + a_{01}d_{10} &= 1, \\
a_{10}c_{01} + a_{01}d_{01} &= 0, \\
b_{10}c_{10} + b_{01}d_{10} &= 0, \\
b_{10}c_{01} + b_{01}d_{01} &= 1.
\end{aligned}$$

This can be rewritten as

$$\begin{pmatrix} a_{10} & b_{10} \\ a_{01} & b_{01} \end{pmatrix} \begin{pmatrix} c_{10} & d_{10} \\ c_{01} & d_{01} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and so

$$\begin{pmatrix} a_{10} & b_{10} \\ a_{01} & b_{01} \end{pmatrix} = \begin{pmatrix} d_{01} & -d_{10} \\ -c_{01} & c_{10} \end{pmatrix} / (c_{10}d_{01} - c_{01}d_{10}).$$

Monomials of higher terms give us

$$a_{10}c_{20} + a_{01}d_{20} + a_{20}c_{10}^2 + a_{11}c_{10}d_{10} + a_{02}d_{10}^2 = 0,$$
$$a_{10}c_{11} + a_{01}d_{11} + a_{20}2c_{10}c_{01} + a_{11}(c_{10}d_{01} + c_{01}d_{10}) + a_{02}(2d_{10}d_{01}) = 0,$$
$$a_{10}c_{02} + a_{01}d_{02} + a_{20}c_{01}^2 + a_{11}c_{01}d_{01} + a_{02}d_{01}^2 = 0,$$

$$b_{10}c_{20} + b_{01}d_{20} + b_{20}c_{10}^2 + b_{11}c_{10}d_{10} + b_{02}d_{10}^2 = 0,$$
$$b_{10}c_{11} + b_{01}d_{11} + b_{20}2c_{10}c_{01} + b_{11}(c_{10}d_{01} + c_{01}d_{10}) + b_{02}(2d_{10}d_{01}) = 0,$$
$$b_{10}c_{02} + b_{01}d_{02} + b_{20}c_{01}^2 + b_{11}c_{01}d_{01} + b_{02}d_{01}^2 = 0.$$

Moving the known terms to the right hand side and setting $(a_{20}, a_{11}, a_{02}, b_{20}, b_{11}, b_{02})$ to be the vector of unknowns leads us to the linear system with the following coefficient matrix and right hand side:

$$\left(\begin{array}{cccccc|c} c_{10}^2 & c_{10}d_{10} & d_{10}^2 & 0 & 0 & 0 & -(a_{10}c_{20}+a_{01}d_{20}) \\ 2c_{10}c_{01} & c_{10}d_{01}+c_{01}d_{10} & 2d_{10}d_{01} & 0 & 0 & 0 & -(a_{10}c_{11}+a_{01}d_{11}) \\ c_{01}^2 & c_{01}d_{01} & d_{01}^2 & 0 & 0 & 0 & -(a_{10}c_{02}+a_{01}d_{02}) \\ 0 & 0 & 0 & c_{10}^2 & c_{10}d_{10} & d_{10}^2 & -(b_{10}c_{20}+b_{01}d_{20}) \\ 0 & 0 & 0 & 2c_{10}c_{01} & c_{10}d_{01}+c_{01}d_{10} & 2d_{10}d_{01} & -(b_{10}c_{11}+b_{01}d_{11}) \\ 0 & 0 & 0 & c_{01}^2 & c_{01}d_{01} & d_{01}^2 & -(b_{10}c_{02}+b_{01}d_{02}) \end{array}\right).$$

This way we can *sequentially* calculate the unknown coefficients.

In fact, we can strongly simplify the computations of the Taylor coefficients of the inverse mapping. Let us first perform a *linear* change of variables (in the source or in the target) which brings the Jacobian matrix of f at the origin to the unit matrix I. Of course, we have to transform accordingly the rest of the Taylor coefficients of f. Now in this special case the formulae for the Taylor coefficients of the inverse mapping g are particularly simple:

PROPOSITION 2.4. *When $Df(0,0) = I$ the above matrix reduces to*

$$\left(\begin{array}{cccccc|c} 1 & 0 & 0 & 0 & 0 & 0 & -c_{20} \\ 0 & 1 & 0 & 0 & 0 & 0 & -c_{11} \\ 0 & 0 & 1 & 0 & 0 & 0 & -c_{02} \\ 0 & 0 & 0 & 1 & 0 & 0 & -d_{20} \\ 0 & 0 & 0 & 0 & 1 & 0 & -d_{11} \\ 0 & 0 & 0 & 0 & 0 & 1 & -d_{02} \end{array}\right).$$

3. Quantitative Sard theorem. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a C^k-mapping, $k \geq 2$. The point $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ is called a *regular point* of f if the differential $Df(x)$ has the maximal possible rank $r = \min(n, m)$. The points

x where the rank of $Df(x)$ is strictly less than r are called *critical points* of f, and the values of f at its critical points are called *critical values*.

The classical Sard theorem (see [53–55, 19, 59]) states that if the mapping f is smooth enough, then the Lebesgue measure of the set of critical values is zero. This theorem is one of the main technical tools of singularity theory (where it appears mostly in a different but essentially equivalent form: as the "transversality theorem").

A typical conclusion of the transversality theorem is that a generic submanifold M in $\mathbb{R}^n$ is transversal to any fixed submanifold N. In our quantitative approach we want to know an explicit lower bound on the "measure of transversality". Going back to the Sard theorem we see that what is needed is the bound on the measure (or, better, on the "size") of not only critical, but also *near-critical values*. These are the values of f at the *near-critical points* where the differential is not exactly degenerate, but close to degenerate.

However, the classical Sard theorem (see [53–55, 19]) does not provide such information. We believe that the absence of the quantitative Sard theorem was one of the most essential obstructions to making singularity theory really quantitative and applicable.

The required quantitative version of the Sard theorem has been obtained in [62] on the base of certain geometric results in real algebraic geometry. Below we present the simplest version of this theorem.

To simplify the presentation, we consider only the case of *functions.* The general statement and proof of the quantitative Sard theorem, as well as some examples and applications can be found in [62–66].

So let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^1-function. For $\gamma \geq 0$, let

$$\Sigma(f, \gamma) = \{x \mid \|\operatorname{grad} f(x)\| \leq \gamma\}.$$

Let $B^n_r \subset \mathbb{R}^n$ be some ball of radius r. We denote $\Sigma(f, \gamma) \cap B^n_r$ by $\Sigma(f, \gamma, r)$ and $f(\Sigma(f, \gamma, r)) \subseteq \mathbb{R}$ by $\Delta(f, \gamma, r)$.

$\Sigma(f, \gamma, r)$ and $\Delta(f, \gamma, r)$ are the sets of γ-*critical points* and γ-*critical values* of f on B^n_r, respectively. For $\gamma = 0$ we get the usual critical points and values.

For a C^k-function $f : \mathbb{R}^n \to \mathbb{R}$ we define $R_k(f)$ as $R_k(f) = (M_k/k!)r^k$ where M_k is the maximum of the kth order derivatives of f on the ball B^n_r. Then $R_k(f)$ is essentially the remainder term in the $k-1$-order Taylor formula for f on the ball B^n_r.

THEOREM 3.1. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^k-function. For each non-negative γ and for each $\varepsilon > R_k(f)$ the set $\Delta(f, \gamma, r)$ can be covered by*

$$C_1(n, k) + C_2(n, k)\gamma(r/\varepsilon)$$

intervals of length ε, and for $\varepsilon \le R_k(f)$ it can be covered by

$$C_3(n,k)\left(\frac{R_k(f)}{\varepsilon}\right)^{n/k} + C_4(n,k)\gamma\left(\frac{r}{\varepsilon}\right)\left(\frac{R_k(f)}{\varepsilon}\right)^{(n-1)/k}$$

such intervals. In particular, for $\gamma = 0$, the set $\Delta(f,0,r)$ of exactly critical values of f can be covered by

$$C_1(n,k) + C_3(n,k)\left(\frac{R_k(f)}{\varepsilon}\right)^{n/k}$$

intervals of length ε.

The expression provided by Theorem 3.1 is not very simple. However, counting the number of covering ε-intervals as $\varepsilon \to 0$, one can easily see that if $k > n$ then the measure of $\Delta(f,\gamma,r)$ tends to zero as $\gamma \to 0$.

The following corollary is, essentially, a special case of a more general result of Corollary III.2.3 in [52].

Corollary 3.1. *Assume that the smoothness k is greater than n. Then for γ sufficiently small the measure of the set of γ-critical values of f satisfies*

$$m(\Delta(f,\gamma,r)) \le c\gamma^{(k-n)/(k-1)}.$$

Here c is a constant depending on k, n, r and $R_k(f)$. In particular, the measure of $\Delta(f,\gamma,r)$ tends to zero as $\gamma \to 0$.

Proof. Since by assumption γ is sufficiently small, and ε will be chosen later to be of the order of $\gamma^{k/(k-1)}$, we use the second inequality of Theorem 3.1. (One can drop this assumption, allowing for somewhat more complicated expressions. See also [52, Corollary III.2.3].)

Taking a larger constant, we see that $\Delta(f,\gamma,r)$ can be covered by

$$C\left[\left(\frac{1}{\varepsilon}\right)^{n/k} + \gamma\left(\frac{1}{\varepsilon}\right)^{(n+k-1)/k}\right]$$

intervals of length ε. Put now $\varepsilon_0 = \gamma^{k/(k-1)}$. Then both terms above are equal, and we conclude that $\Delta(f,\gamma,r)$ can be covered by $2C(1/\gamma)^{n/(k-1)}$ intervals of length ε_0. Finally, for $c = 2C$, the measure of $\Delta(f,\gamma,r)$ does not exceed $c\varepsilon_0(1/\gamma)^{n/(k-1)} = c\gamma^{(k-n)/(k-1)}$.

Corollary 3.2. *For $k > n$ the measure of the set $\Delta(f,0,r)$ of exactly critical values of f is zero.*

This is the usual Sard theorem.

Remark. In the computations above we do not take into account the more accurate definition of the "degree of smoothness" of f, which is appropriate in the quantitative Sard theorem. See [62, 66].

We do not give here the proof of Theorem 3.1, addressing the reader to [62, 66]. Let us only notice that the starting point is the following result for polynomials:

THEOREM 3.2. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a polynomial of degree d. Then for any $\gamma \geq 0$ the set $\Delta(f, \gamma, r)$ can be covered by $N(n, d)$ intervals of length γr. The constant $N(n, d)$ here depends only on n and d.*

The proof of Theorem 3.2 is based on metric bounds on real semi-algebraic sets, much in the spirit of S. Łojasiewicz's results and approach. See also [8] where an accurate bound on $N(n, d)$ is obtained.

Theorem 3.1 follows from Theorem 3.2 combined with the approximation of f by its local Taylor polynomials on an appropriate grid.

4. Quantitative Morse theorem. We consider smooth functions $f : B^n \to \mathbb{R}$, where B^n is the closed unit ball in $\mathbb{R}^n$. Probably, the first and most basic result of "proper" singularity theory is the Morse theorem ([49, 50, 48]), describing typical singularities of f. It states that "generically" f has the following properties:

(i) All critical points x_i of f are non-degenerate (i.e. the Hessian $H(f)$ is non-degenerate at each x_i). Consequently, the number of these critical points is finite.
(ii) All the critical values are distinct, i.e. $f(x_i) \neq f(x_j)$ for $i \neq j$.
(iii) Near each point x_i there is a new coordinate system $y_1, \ldots, y_n$, centered at this point, such that
$$f(y_1, \ldots, y_n) = y_1^2 + \cdots + y_l^2 - y_{l+1}^2 - \cdots - y_n^2 + \text{const}.$$

In particular, we can perturb any given f_0 by an arbitrarily small (in C^∞-norm) addition h so that $f = f_0 + h$ has properties (i)–(iii) as above.

Now a parallel quantitative result is the following:

THEOREM 4.1. *Fix $k \geq 3$. Let f_0 be a C^k-function with all derivatives up to order k uniformly bounded by K. Then for any given $\varepsilon > 0$, we can find h with $\|h\|_{C^k} \leq \varepsilon$ such that for $f = f_0 + h$,*

(i) *At each critical point x_i of f, the smallest eigenvalue of the Hessian $H(f)$ at x_i is at least $\psi_1(K, \varepsilon) > 0$.*
(ii) *The distance between any two different critical points x_i and x_j of f is not smaller than $d(K, \varepsilon)$. Consequently, the number of critical points x_i does not exceed $N(K, \varepsilon)$.*
(iii) *For any $i \neq j$, the distance between the critical values $f(x_i)$ and $f(x_j)$ is not smaller than $\psi_2(K, \varepsilon)$.*
(iv) *For $\delta = \psi_3(K, \varepsilon) > 0$ and for each critical point x_i of f, in a δ-neighborhood U_δ of x_i there is a new coordinate system $y_1, \ldots, y_n$,*

centered at x_i, such that

$$f(y_1,\dots,y_n) = y_1^2 + \cdots + y_l^2 - y_{l+1}^2 - \cdots - y_n^2 + \text{const.}$$

The C^{k-1}-norm of the coordinate transformation from the original coordinates to $y_1,\dots,y_n$ (and of the inverse transformation) does not exceed $M(K,\varepsilon)$.

Here ψ_1,ψ_2,ψ_3,d (tending to zero as $\varepsilon \to 0$) and N, M (tending to infinity) are explicitly given functions, depending only on k, K and ε. The neighborhoods U_δ of the singular points x_i play an important role in what follows. Let us call them the *controlled neighborhoods* of the corresponding singular points x_i.

REMARK. In the paper of L. Niederman ([52]), which was already mentioned in Section 3 above, another version of the "quantitative Morse theorem" is proved (see [52, Theorems III.2.5 and III.2.6]). It implies, in particular, a stronger assertion than statement (i) of Theorem 4.1: the Hessian of f is large not only at the critical, but also at the near-critical points.

A sketch of the proof of the first three statements of Theorem 4.1 is given in [66]. We shall now prove a two-dimensional version of the last statement of this theorem.

THEOREM 4.2. *Let $f(x,y)$ be a real analytic function with $f(0,0)=0$ and $df(0,0)=0$, represented by a power series*

$$f(x,y) = \sum_{k+l=2}^{\infty} a_{k,l} x^k y^l \tag{4.1}$$

that satisfies $|a_{k,l}| \le M(1/R)^{k+l}$. Assume also that the Hessian $H(f)(0,0)$ is a non-degenerate matrix with eigenvalues $\lambda_1,\lambda_2 \ne 0$. Denote $\min\{|\lambda_1|,|\lambda_2|\}$ by $a>0$. Then there are new coordinates $u_1 = u_1(x,y)$, $u_2 = u_2(x,y)$ in a neighborhood U of the origin in $\mathbb{R}^2$ such that

$$f(x,y) = \lambda_1 u_1^2 + \lambda_2 u_2^2 \tag{4.2}$$

at each point of U. The neighborhood U of the origin in $\mathbb{R}^2$ contains a disk D_κ of radius

$$\kappa = \frac{aR}{tM}.$$

The Taylor coefficients $b^i_{k,l}$, $i=1,2$, of the coordinate functions $u_1 = u_1(x,y)$, $u_2 = u_2(x,y)$ satisfy $|b^i_{k,l}| \le 2(1/\kappa)^{k+l}$.

The Taylor coefficients $c^i_{k,l}$, $i=1,2$, of the inverse coordinate transformation $x = g_1(u_1,u_2)$, $y = g_2(u_1,u_2)$ satisfy $|c^i_{k,l}| \le M_1(1/R_1)^{k+l}$, where the constants M_1 and R_1 are given by the expressions in Theorem 2.3. (The in-

put constants M and R in Theorem 2.3 *are taken to be* 2 *and* κ, *respectively, while the constant* δ *in Theorem* 2.3 *is set to be equal to* 1/2.)

Proof. The proof is constructive, in the sense that it provides an algorithm (and explicit expressions) for computing the Taylor series of the new coordinate functions $u_1 = u_1(x,y), u_2 = u_2(x,y)$. It consists of several steps.

1. Since the function $f(x,y)$ vanishes at the origin together with its first derivatives, the Taylor series of f starts from the terms of order two. These order two terms represent a non-degenerate quadratic form $\tilde{H}(f)$. We perform a linear coordinate transformation $(x_1, y_1) = L(x,y)$ which brings the quadratic form $\tilde{H}(f)$ to its diagonal form $\tilde{H}(f)(x,y) = \lambda_1 x_1^2 + \lambda_2 y_1^2$.

2. So assume that the quadratic part of $f(x,y)$ already has this form:

$$f(x,y) = \lambda_1 x^2 + \lambda_2 y^2 + \sum_{k+l=3}^{\infty} a_{k,l} x^k y^l. \tag{4.3}$$

Let us rewrite this expression as follows:

$$f(x,y) = \lambda_1 x^2 + \sum_{k+l\ge 3,\, l=0,1} a_{k,l} x^k y^l + \lambda_2 y^2 + \sum_{k+l\ge 3,\, l\ge 2} a_{k,l} x^k y^l. \tag{4.4}$$

Each term of the first sum is divisible by x^2 while each term of the second sum is divisible by y^2. So we can write

$$\begin{aligned} f(x,y) = {} & \lambda_1 x^2 \Bigg(1 + \sum_{k+l\ge 3,\, l=0,1} \frac{a_{k,l}}{\lambda_1}\, x^{k-2} y^l\Bigg) \\ & + \lambda_2 y^2 \Bigg(1 + \sum_{k+l\ge 3,\, l\ge 2} \frac{a_{k,l}}{\lambda_2}\, x^k y^{l-2}\Bigg). \end{aligned} \tag{4.5}$$

Using the notations

$$\sum_{k+l\ge 3,\, l=0,1} \frac{a_{k,l}}{\lambda_1}\, x^{k-2} y^l = q_1(x,y), \qquad \sum_{k+l\ge 3,\, l\ge 2} \frac{a_{k,l}}{\lambda_2}\, x^k y^{l-2} = q_2(x,y),$$

respectively, we get

$$f(x,y) = \lambda_1 x^2 (1 + q_1(x,y)) + \lambda_2 y^2 (1 + q_2(x,y)). \tag{4.6}$$

Finally, let $s_1(x,y) = (1 + q_1(x,y))^{1/2}$, $s_2(x,y) = (1 + q_2(x,y))^{1/2}$, where we choose the branch of the square root taking the value 1 at 1, and let $u_1(x,y) = x s_1(x,y)$, $u_2(x,y) = y s_2(x,y)$. Clearly, the required identity

$$f(x,y) = \lambda_1 u_1^2 + \lambda_2 u_2^2 \tag{4.7}$$

is satisfied.

It remains to notice that $s_1(x,y)$ and $s_2(x,y)$ are analytic functions of x, y in a neighborhood V of the origin, and their Taylor expansion can be obtained by substituting the sums above into the binomial series $\sqrt{1+v} = 1 + \frac{1}{2} v + \cdots$.

To conclude that u_1, u_2 form a coordinate system and to estimate the size of the coordinate neighborhood V, as well as the Taylor coefficients of the inverse mapping, we first estimate the expressions $q_1(x,y)$ and $q_2(x,y)$. Indeed, for $|x|,|y| \le \kappa$, using the assumptions on $a_{k,l}$, we get

$$|q_1(x,y)| \le \frac{2M}{a\lambda_1}\kappa \le \frac{2M\kappa}{aR}, \quad |q_2(x,y)| \le \frac{2M}{a\lambda_2}\kappa \le \frac{2M\kappa}{aR}. \tag{4.8}$$

In particular, taking $\kappa = aR/4M$ and denoting the polydisk $\{|x|,|y| \le \kappa\}$ by V we conclude that for $x, y \in V$ the absolute value of $q_1(x,y)$ and $q_2(x,y)$ does not exceed $1/2$. Hence for such x, y the functions $s_1(x,y)$ and $s_2(x,y)$ are analytic and bounded by 2. Therefore, the new coordinate functions $u_1(x,y) = xs_1(x,y)$, $u_2(x,y) = ys_2(x,y)$ are defined in a neighborhood U which contains the polydisk V. By the Cauchy formula we also get the required bound for the Taylor coefficients $b^i_{k,l}$ of u_1, u_2: $|b^i_{k,l}| \le 2(1/\kappa)^{k+l}$.

Now we apply the inverse function theorem (Theorem 2.3 above). We have to replace R by κ and M by 2 in the bounds for the direct transformation. Notice also that the differential of our transformation $u_1(x,y) = xs_1(x,y)$, $u_2(x,y) = ys_2(x,y)$ is the unit 2×2 matrix, since $s_1(x,y)$ and $s_2(x,y)$ take the value 1 at the origin. Hence the parameter δ in Theorem 2.3 in our case is $1/2$. This completes the proof of Theorem 4.2.

Proposition 4.1. *The Taylor coefficients of the new coordinate system u_1, u_2 are explicitly given through the Taylor coefficients of $f(x,y)$ as follows*:

$$\begin{aligned} u_1(x,y) = x &+ \left(\frac{a_{21}}{2\lambda_1}\right)xy + \left(\frac{a_{30}}{2\lambda_1}\right)x^2 + \left(\frac{a_{40}}{2\lambda_1} - \frac{3}{2}\,\frac{a_{21}a_{30}}{\lambda_1^2}\right)x^2y \\ &+ \left(\frac{a_{40}}{2\lambda_1} - \frac{3}{4}\,\frac{a_{30}^2}{\lambda_1^2}\right)x^3 + \left(-\frac{3}{4}\,\frac{a_{21}^2}{\lambda_1^2}\right)xy^2 + \cdots, \\ u_2(x,y) = y &+ xy\left(\frac{1}{2}\,\frac{a_{30}}{\lambda_2}\right) + y^2\left(\frac{1}{2}\,\frac{a_{12}}{\lambda_2}\right) + x^2y\left(\frac{1}{2}\,\frac{a_{22}}{\lambda_2} - \frac{3}{4}\,\frac{a_{12}^2}{\lambda_2^2}\right) \\ &+ xy^2\left(\frac{1}{2}\,\frac{a_{13}}{\lambda_2} - \frac{3}{2}\,a_{30}a_{12}\,\frac{1}{\lambda_2^2}\right) + y^3\left(\frac{1}{2}\,\frac{a_{04}}{\lambda_2} - \frac{3}{4}\,\frac{a_{03}^2}{\lambda_2^2}\right) + \cdots. \end{aligned}$$

The Taylor coefficients of the inverse coordinate transformation $x = g_1(u_1,u_2)$, $y = g_2(u_1,u_2)$ can be explicitly obtained through the Taylor coefficients of $f(x,y)$ as follows: we substitute the expressions of Proposition 4.1 into the expressions for the coefficients of the inverse function, as given in Section 2, after Theorem 2.3.

5. Stability of Morse functions. Another typical result of the classical singularity theory is the "stability theorem", which in the case of Morse singularities takes the following form: if f satisfies conditions (i)–(iii) above, then any small perturbation f_1 of f is equivalent to f via the diffeomorphisms of the source and target.

(In this form the result is true for functions on compact manifolds without boundary. In the case of functions defined on the unit ball, or on any other manifold with boundary, one has to care about singularities of f restricted to the boundary.)

A parallel result of quantitative singularity theory is the following:

THEOREM 5.1. *Let f be a C^k-function with all derivatives up to order k uniformly bounded by K. Let f satisfy*:

(a) *At each critical point x_i of f, the smallest eigenvalue of the Hessian $H(f)$ at x_i is at least $\psi_1 > 0$.*

(b) *For any $i \neq j$, the distance between the critical values $f(x_i)$ and $f(x_j)$ is not smaller than $\psi_2 > 0$.*

Then there is $\varepsilon_0 > 0$ (*depending only on K, ψ_1, ψ_2*) *such that for any given ε with $\varepsilon_0 > \varepsilon > 0$, and for any f_1 which is closer than ε to f in C^k-norm*, *f_1 is equivalent to f via diffeomorphisms G and H of the source and target*, *respectively. G and H differ* (*in C^{k-1}-norm*) *from the identical diffeomorphisms by not more than $s(K, \psi_1, \psi_2, \varepsilon)$. Here $s(K, \psi_1, \psi_2, \varepsilon)$ is an explicitly given function of its arguments*, *which tends to zero as $\varepsilon \to 0$.*

We plan to give the proof of Theorem 5.1 separately.

6. Organizing center. The next “quantitative” result has no direct analogy in the classical singularity theory. It states that for a generic mapping each of its “near-singular” points belongs to a controlled neighborhood of one of exact singular points (its “organizing center”).

This result answers (for the Morse singularities) an important problem in applications of singularity theory: the problem of identification of the organizing center for near-singularities. The notion of an organizing center was introduced by R. Thom (see [57]). One of interpretations of this notion is as follows: *when we detect a “near-singularity”, find its organizing center, which is a nearby exact singularity whose “controlled neighborhood” contains the original near-singular point.*

Theorem 6.1 below shows that this is possible for Morse singularities. It shows that (at least in principle) we can relate to each near-singularity its organizing center. We believe that this fact (extended to a wider range of singularities and supplemented with effective and efficient estimates of the parameters involved) may be of basic importance for applications of singularity theory. One can hope that progress in this direction may transform some inspiring ideas and approaches of [57] into theorems and working algorithms.

Theorem 6.1. *Let $f_0 : B^n \to \mathbb{R}$ be a C^k-function with all derivatives up to order k uniformly bounded by K. Then for any given $\varepsilon > 0$, we can find h with $\|h\|_{C^k} \leq \varepsilon$ such that for $f = f_0 + h$ conditions* (i)–(iv) *of Theorem* 4.1 *are satisfied, as well as the following additional condition*:

(v) *There is an explicit function $\eta(K, \varepsilon) > 0$ such that any point x with the norm of* $\operatorname{grad} f(x)$ *smaller than $\eta(K, \varepsilon)$ belongs to one of the controlled neighborhoods of the singular points x_i of f.*

Sketch of proof. Consider the mapping $Df : B^n \to \mathbb{R}^n$, where Df is the differential (or gradient) of f. The critical points x_i of f are exactly the preimages of zero under Df. If zero is a regular value of Df then the Hessian $H(f)$ is non-degenerate at each x_i (being the Jacobian of Df).

Now consider linear functions $h : B^n \to \mathbb{R}$. Zero is a γ-near-singular value of Df for $f = f_0 + h$ if and only if the point $-Dh$ is a γ-near-singular value of Df_0. The bound on the geometry of the near-critical values of Df_0, provided by the appropriate version of the quantitative Sard theorem (see [62, 66]) implies the following: *For any $r > 0$ there are points v in $\mathbb{R}^n$, at a distance at most r from zero, such that the entire ball B in $\mathbb{R}^n$ of radius $\eta(K, r)$, centered at v, consists of $\gamma(K, r)$-regular values of Df_0.* Here $\gamma(K, r)$ and $\eta(K, r)$ are explicitly given functions, tending to zero as $r \to 0$.

Now for a given $\varepsilon > 0$ let us pick a certain $\gamma(K, \varepsilon)$-regular value v of Df_0, at a distance at most ε from zero, with the property that the entire ball B in $\mathbb{R}^n$ of radius $\eta(K, \varepsilon)$, centered at v, consists of $\gamma(K, \varepsilon)$-regular values of Df_0. Let h be a linear function with $Dh = -v$. Then any point x with the norm of $\operatorname{grad} f(x)$ smaller than $\eta(K, \varepsilon)$ satisfies $Df_0(x) \in B$. Hence it is a $\gamma(K, \varepsilon)$-regular point for Df_0, i.e. the minimal eigenvalue of the Hessian $H(f)$ at x is bounded from below by $\gamma(K, \varepsilon)$.

To complete the proof, we apply a quantitative inverse function theorem (one of its versions is Theorem 2.3 of Section 2). It shows that with our lower bound on the Hessian (and with the global bound on higher derivatives) a certain neighborhood of x is mapped by Df *onto* the ball of an explicitly given radius in $\mathbb{R}^n$. With a proper choice of the function $\eta(K, \varepsilon)$ this last ball contains the origin. This means that in a neighborhood of x there is a true singular point x_i of f. Once more, with a proper tuning of the inequalities, we conclude that x belongs to the "controlled neighborhood" of x_i (as defined in Section 4). This completes the proof.

Remark. The results of L. Niederman in [52] provide an important information on the position of near-singular points. In particular, we can replace the first part of the proof above by an application of Theorems III.2.4 and III.2.5 of [52].

References

[1] V. I. Arnol'd, S. M. Gusein-Zade and A. N. Varchenko, *Singularities of Differentiable Maps*, Volumes I and II, Monogr. Math. 82 and 83, Birkhäuser Boston, Boston, MA, 1985 and 1988.

[2] D. Auroux, *Asymptotically holomorphic families of symplectic submanifolds*, Geom. Funct. Anal. 7 (1997), 971–995.

[3] M. Briskin, Y. Elichai and Y. Yomdin, *How can singularity theory help in image processing?*, in: Pattern Formation in Biology, Vision and Dynamics, M. Gromov and A. Carbone (eds.), World Sci., 2000, 392–423.

[4] Th. Bröcker and L. Lander, *Differentiable germs and catastrophes*, London Math. Soc. Lecture Note Ser. 17, Cambridge Univ. Press, 1975.

[5] F. H. Clarke, *Generalized gradients and applications*, Trans. Amer. Math. Soc. 205 (1975), 247–262.

[6] —, *On the inverse function theorem*, Pacific J. Math. 64 (1976), 97–102.

[7] D. Cohen, *Some quantitative results in singularity theory*, MA Thesis, Weizmann Institute of Science, 2004.

[8] D. D'Acunto and K. Kurdyka, *Bounds for gradient trajectories of definable functions with applications to robotics and semialgebraic geometry*, preprint, 2003.

[9] J. Damon, *Local Morse theory for solutions to the heat equation and Gaussian blurring*, J. Differential Equations 115 (1995), 368–401.

[10] —, *Generic properties of solutions to partial differential equations*, Arch. Rat. Mech. Anal. 140 (1997), 353–403.

[11] —, *Generic structure of two-dimensional images under Gaussian blurring*, SIAM J. Appl. Math. 59 (1999), 97–138.

[12] —, *Properties of ridges and cores for two-dimensional images*, J. Math. Imaging Vision 10 (1999), 163–174.

[13] —, *Smoothness and geometry of boundaries associated to skeletal structures I: sufficient conditions for smoothness*, Ann. Inst. Fourier (Grenoble) 53 (2003), 1941–1985.

[14] S. K. Donaldson, *Symplectic submanifolds and almost-complex geometry*, J. Differential Geom. 44 (1996), 666–705.

[15] —, *Lefschetz pencils on symplectic manifolds*, ibid. 53 (1999), 205–236.

[16] —, *Lefschetz fibrations in symplectic geometry*, in: ICM 98 II (1998), 309–314.

[17] Y. Elihai and Y. Yomdin, *Global motion planning algorithm, based on a high order discretization and on hierarchies of singularities*, in: Proc. 28th CDC (Tampa, FL, 1989), 1173–1174.

[18] —, —, *Flexible high-order discretization of geometric data for global motion planning*, Theoret. Comput. Sci. 157 (1996), 53–77.

[19] H. Federer, *Geometric Measure Theory*, Grundlehren Math. Wiss. 153, Springer, 1969.

[20] M. Golubitsky and V. Guillemin, *Stable Mappings and Their Singularities*, Grad. Texts in Math. 14, Amer. Math. Soc., 1973.

[21] M. Gromov, *On the entropy of holomorphic maps*, Enseign. Math. 49 (2003), 217–235 (preprint: SUNY, 1977).

[22] —, *Three remarks on geodesic dynamics and fundamental group*, ibid. 46 (2000), 391–402 (preprint: SUNY, 1976).

[23] —, *Entropy, homology and semialgebraic geometry*, in: Séminaire Bourbaki, Vol. 1985/86, Astérisque 145–146 (1987), 225–240.

[24] M. Gromov, *Spectral geometry of semi-algebraic sets*, Ann. Inst. Fourier (Grenoble) 42 (1992), no. 1–2, 249–274.
[25] J. Guckenheimer, *Catastrophes and partial differential equations*, ibid. 23 (1973), no. 2, 31–59.
[26] J. Guckenheimer and Y. Xiang, *Defining equations for bifurcations and singularities*, Moscow Math. J. 3 (2003), 935–946.
[27] P. Henrici, *Applied and Computational Complex Analysis*, Vol. 1, Wiley, New York, 1988.
[28] A. Ibort, D. Martinez-Torres and F. Presas, *On the construction of contact submanifolds with prescribed topology*, J. Differential Geom. 56 (2000), 235–283.
[29] V. Kaloshin, *Prevalence in spaces of smooth mappings and a prevalent Kupka–Smale theorem*, Dokl. Akad. Nauk 355 (1997), 306–307.
[30] —, *An extension of the Artin–Mazur theorem*, Ann. of Math. (2) 150 (1999), 729–741.
[31] —, *Generic diffeomorphisms with superexponential growth of number of periodic points*, Comm. Math. Phys. 211 (2000), 253–271.
[32] K. Kurdyka, T. Mostowski and A. Parusiński, *Proof of the gradient conjecture of R. Thom*, Ann. of Math. (2) 152 (2000), 763–792.
[33] K. Kurdyka and P. Orro, *Distance géodésique sur un sous-analytique*, Rev. Mat. Univ. Compl. Madrid 10 (1997), Special Issue, suppl., 173–182.
[34] K. Kurdyka, P. Orro and S. Simon, *Semialgebraic Sard theorem for generalized critical values*, J. Differential Geom. 56 (2000), 67–92.
[35] S. Łojasiewicz, *Ensembles semi-analytiques*, preprint, I.H.E.S., 1965.
[36] —, *Sur les ensembles semi-analytiques*, in: Actes du Congrès International des Mathématiciens (Nice, 1970), Tome 2, Gauthier-Villars, Paris, 1971, 237–241.
[37] —, *Introduction to Complex Analytic Geometry*, Birkhäuser, Basel, 1991.
[38] S. Łojasiewicz and M.-A. Zurro, *On the gradient inequality*, Bull. Polish Acad. Sci. Math. 47 (1999), 143–145.
[39] J. Martinet, *Singularities of Smooth Functions and Maps*, Cambridge Univ. Press, 1982.
[40] J. Mather, *Notes on Topological Stability*, Harvard Univ. 1970.
[41] —, *Stability of C^∞ mappings, I: The division theorem*, Ann. of Math. 87 (1968), 89–104.
[42] —, *Stability of C^∞ mappings, II: Infinitesimal stability implies stability*, ibid. 89 (1969), 254–291.
[43] —, *Stability of C^∞ mappings, III: Finitely determined map-germs*, Publ. Math. IHES 35 (1968), 127–156.
[44] —, *Stability of C^∞ mappings, IV: Classification of stable germs by R-algebras*, ibid. 37 (1969), 223–248.
[45] —, *Stability of C^∞ mappings, V: Transversality*, Adv. Math. 4 (1970), 301–336.
[46] —, *Stability of C^∞ mappings, VI: The nice dimensions*, in: Liverpool Singularities I, Lecture Notes in Math. 192, Springer, 1971, 207–253.
[47] —, *On Nirenberg's proof of the Malgrange preparation theorem*, ibid., 116–120.
[48] J. Milnor, *Morse Theory*, Princeton Univ. Press, Princeton, NJ, 1963.
[49] M. Morse, *The critical points of a function of n variables*, Trans. Amer. Math. Soc. 33 (1931), 71–91.
[50] —, *The Calculus of Variations in the Large*, Amer. Math. Soc., 1934.
[51] L. Niederman, *Hamiltonian stability and subanalytic geometry*, preprint, Univ. Paris XI, 2004.

[52] L. Niederman, *Prevalence of exponential stability among nearly-integrable Hamiltonian systems*, preprint, Univ. Paris XI, 2004.

[53] A. Sard, *The measure of the critical values of differentiable maps*, Bull. Amer. Math. Soc. 48 (1942), 883–890.

[54] —, *Images of critical sets*, Ann. of Math. 68 (1958), 247–259.

[55] —, *Hausdorff measure of critical images on Banach manifolds*, Amer. J. Math. 87 (1965), 158–174.

[56] J.-C. Sikorav, *Construction de sous-variétés symplectiques (d'après S. K. Donaldson et D. Auroux)*, in: Séminaire Bourbaki, Vol. 1997/98, exp. no. 844, Astérisque 252 (1998), 231–253.

[57] R. Thom, *Ensembles et morphismes stratifiés*, Bull. Amer. Math. Soc. 75 (1969), 240–284.

[58] —, *Stabilité structurelle et morphogenèse*, W. A. Benjamin, 1972. English edition: R. Thom, *Structural Stability and Morphogenesis*, Benjamin-Cumming, London, 1975.

[59] H. Whitney, *A function not constant on a connected set of critical points*, Duke Math. J. 1 (1935), 514–517.

[60] —, *On singularities of mappings of Euclidean spaces I, Mappings of the plane into the plane*, Ann. of Math. 62 (1955), 374–410.

[61] Y. Xiang, *Computing Thom–Boardman singularities*, Ph.D. Thesis, Cornell Univ., 1998.

[62] Y. Yomdin, *The geometry of critical and near-critical values of differentiable mappings*, Math. Ann. 264 (1983), 495–515.

[63] —, *Global bounds for the Betti numbers of regular fibers of differentiable mappings*, Topology 24 (1985), 145–152.

[64] —, *A quantitative version of the Kupka–Smale theorem*. Ergodic Theory Dynam. Systems 5 (1985), 449–472.

[65] —, *Sard's theorem and its improved versions in numerical analysis*, in: Computational Solution of Nonlinear Systems of Equations (Fort Collins, CO, 1988), Lectures in Appl. Math. 29, Amer. Math. Soc., 1990, 701–706.

[66] Y. Yomdin and G. Comte, *Tame Geometry with Applications in Smooth Analysis*, Lecture Notes in Math. 1834, Springer, Berlin, 2004.

Department of Mathematics
The Weizmann Institute of Science
Rehovot 76100, Israel
E-mail: yosef.yomdin@weizmann.ac.il

Reçu par la Rédaction le 31.1.2005
Révisé le 6.10.2005 (1628)