

# Contents

Foreword	ix
Chapter 1. Finiteness Problems in Differential Equations and Diophantine Geometry D. SCHLOMIUK	1
References	8
Chapter 2. Linear Differential Equations, Fuchsian Inequalities and Multiplicities of Zeros A. A. BOLIBRUKH	11
1. Introduction	11
2. Systems with regular singular points	12
3. Irregular singular points	22
4. Fuchsian inequalities	28
5. Multiplicities of zeroes	34
References	38
Chapter 3. Quantitative Theory of Ordinary Differential Equations and the Tangential Hilbert 16th Problem SERGEI YAKOVENKO	41
1. Hilbert 16th problem: Limit cycles, cyclicity, Abelian integrals	41
2. Abelian integrals and differential equations	59
3. Quasialgebraicity of function fields	71
4. Quantitative theory of differential equations	82
5. Isomonodromic reduction principle and the Riemann–Hilbert problem	95
References	106
Chapter 4. Around the Hilbert–Arnol’d Problem V. KALOSHIN	111
1. Around the Hilbert’s 16th problem and an estimate for cyclicity of elementary polycycles	111
2. Normal forms and the Khovanskiĭ method	123
3. Stratifications and Bézout’s theorem for chain maps	134
4. Bifurcation of spatial polycycles and blow-up along the diagonal of the space of multijets	146
References	158
Chapter 5. Finiteness Results in Differential Algebraic Geometry and Diophantine Geometry ALEXANDRU BUIUM	163
1. Finiteness results in Diophantine geometry	163

2. Finiteness results in differential algebraic geometry	168
3. Discussion of the proofs	170
Conclusion	173
References	174
Appendix A. Abstract: o-Minimal Structures, Real Analytic Geometry, and Transseries L. VAN DEN DRIES	175
Appendix B. List of Lectures Delivered at the Workshop	177
Appendix C. Photographs of Some Workshop Participants	179
Appendix D. List of Participants	181

## Around the Hilbert – Arnol’d Problem

V. KALOSHIN

### 1. Around the Hilbert’s 16th problem and an estimate for cyclicity of elementary polycycles

**1.1. The Hilbert 16th problem and its offsprings.** Consider a polynomial vector field on the real  $(x, y)$ -plane

$$(1) \quad \begin{cases} \dot{y} = P_n(x, y) \\ \dot{x} = Q_n(x, y) \end{cases} \quad P_n, Q_n \text{ polynomials, } \deg P_n, \deg Q_n \leq n.$$

A limit cycle of a polynomial vector field (1) is an isolated periodic solution. Define

$$H(n) = \boxed{\text{uniform bound for the number of limit cycles of (1)}}.$$

One way to formulate the Hilbert 16th problem is the following:

**HILBERT 16TH PROBLEM (HP).** *Estimate  $H(n)$  for any  $n \in \mathbb{Z}_+$ .*

To prove that  $H(1) = 0$  is an exercise, but to find  $H(2)$  is already a difficult unsolved problem (see [DRR, DMR] for work in this direction). Below we discuss two of the most significant branches of research HP has generated: existential and tangential Hilbert 16th problems.

1.1.1. *The tangential Hilbert 16th problem.* Consider a polynomial perturbation of a Hamiltonian polynomial line field

$$(2) \quad \begin{cases} \dot{x} = -\frac{\partial H}{\partial x} - \varepsilon P(x, y) \\ \dot{y} = \frac{\partial H}{\partial y} - \varepsilon Q(x, y). \end{cases}$$

For  $\varepsilon = 0$  the line field (2) does not have any limit cycles at all (all cycles are nonisolated). An oval (topological circle)  $\gamma$  of the level curve  $H(x, y) = h$  generates a limit cycle for a small nonzero value of  $\varepsilon$  if the accumulated energy dissipation is zero in the first approximation, i.e., when

$$(3) \quad \oint_{\gamma} P(x, y) dx + Q(x, y) dy = 0, \quad \gamma \subseteq \{H(x, y) = h\}.$$

The left-hand side expression is called a *complete Abelian integral*. If the polynomials  $H$ ,  $P$ , and  $Q$  are fixed, the integral (3) defines a multivalued function  $I(h)$ . The multivalued possibility appears when the corresponding level curve  $\{H(x, y) = h\}$  has several disjoint ovals.

TANGENTIAL HILBERT 16TH PROBLEM (THP) ([AVI<sup>+</sup>]). *For any collection of polynomials  $P$  and  $Q \in \mathbb{R}[x, y]$  of degree  $\leq n$  and  $H$  of degree  $\leq n + 1$  give an upper bound  $\text{TH}(n)$  on the number of real ovals over which the integral (3) vanishes, but not identically.*

In the latter case the perturbation (2) is a Hamiltonian system for  $\varepsilon \neq 0$  so it does not have any limit cycles at all. Even though the tangential Hilbert problem is not yet solved, in contrast to the Hilbert 16th problem, there are several quite general results related to it. Khovanskiĭ [K1] and Varchenko [V] proved

FINITENESS THEOREM. *For any  $n \in \mathbb{Z}_+$  the number of isolated zeroes of Abelian integrals is uniformly bounded over all Hamiltonians and forms of degree  $\leq n$ .*

For various other results estimating  $H(n)$  in various particular cases see [G, I1, M1, NDY, P2], and the lecture course [Y1] in the present volume for more references.

If we consider (3) over the field of complex numbers we have that an Abelian integral satisfies a Fuchsian equation or Picard–Fuchs equation (see, e.g., [AA]), i.e., an equation of the form

$$(4) \quad \dot{z} = \sum_j \frac{A_j}{t - \alpha_j} z_j,$$

where the  $A_j$  are constant matrices and  $z = (z_1, \dots, z_p) \in \mathbb{C}^p$  is a complex vector for some  $p$ . Investigation of various properties of fuchsian equations is the main topic of lectures of Bolibrukh [B1] in the present volume.

1.1.2. *From existential Hilbert 16th problem to Hilbert–Arnol'd problem.* A qualitative form of Hilbert 16th problem is the following:

EXISTENTIAL HILBERT 16TH PROBLEM (EHP). *Prove that  $H(n) < \infty$  for any  $n \in \mathbb{Z}_+$ .*

The problem of the finiteness of the number of limit cycles for an individual polynomial line field (1) is called the *Dulac problem* after the pioneering work of Dulac [D1] who claimed in 1923 to have solved this problem, but an error was found by Ilyashenko [I2] 60 years later. The Dulac problem was solved by two independent, rather different and incredibly complicated proofs given almost simultaneously by Ilyashenko and Écalle.

INDIVIDUAL FINITENESS THEOREM (IFT) ([I3, É]). *Any polynomial line field (1) has only a finite number of limit cycles.*

However, neither proof allows any generalization to solve EHP. Consider the equation (1) for different polynomials  $(P_n(x, y), Q_n(x, y)) \in (\mathbb{R}[x, y])^2$  as the family of vector fields on  $\mathbb{R}^2$  depending on the parameters of the polynomials. Using a central projection  $\pi: \mathbb{S}^2 \rightarrow \mathbb{R}^2$  and homogeneity with respect to parameters of the equation (1) (vector fields  $(\lambda P_n(x, y), \lambda Q_n(x, y))$  and  $(P_n(x, y), Q_n(x, y))$  for any  $\lambda \neq 0$  have the same trajectories) one can construct a *finite parameter family of analytic line fields on the sphere  $\mathbb{S}^2$  with a compact parameter base  $B$*  (see, e.g., [IY2] for details). After this reduction the existential Hilbert problem becomes a particular case of the following:

GLOBAL FINITENESS CONJECTURE (GFC) ([R1]). *For any analytic family of line fields on  $\mathbb{S}^2$  with a compact parameter base  $B$  the number of limit cycles is uniformly bounded over all parameter values.*

We refer the reader to the volumes [IY2, S1], and a book [R2], where various developments of these and related problems are discussed. Families of analytic fields are difficult to analyze. In the middle of the 1980's Arnol'd [AA] proposed to consider generic families of smooth vector fields as the first step toward understanding families of analytic vector fields. A smooth analog of the global finiteness conjecture is the following:

HILBERT – ARNOL'D PROBLEM (HAP) ([I4]). *Prove that, in a generic finite parameter family of vector fields on the sphere  $S^2$  with a compact base  $B$ , the number of limit cycles is uniformly bounded.*

Assume for a moment that an analytic (or a generic smooth) vector field on the sphere  $\mathbb{S}^2$  has an infinite number of limit cycles. By the Poincaré–Bendixson theorem, any limit cycle should surround an equilibrium point and, since our vector field has at most finitely many equilibria, there should be an infinite “nested” sequence around one of the equilibria. Then this “nested” sequence of limit cycles has to accumulate (in the sense of the Hausdorff metric) to a certain contour (polygon) consisting of equilibria (as vertices) and separatrix curves (sides of that polygon) connecting them. Such objects are called *polycycles*. It turns out that a possible solution to the Hilbert–Arnol'd problem reduces to the investigation of bifurcation of polycycles. Let us give several definitions.

DEFINITION 1. A polycycle  $\gamma$  of a vector field on the sphere  $\mathbb{S}^2$  is a cyclically ordered collection of equilibrium points  $p_1, \dots, p_k$  (with possible repetitions) and arcs  $\gamma_1, \dots, \gamma_k$  (distinct integral curves consisting possibly of equilibrium points) connecting them in a specific order: the  $j$ th arc  $\gamma_j$  connects  $p_j$  with  $p_{j+1}$  for  $j = 1, \dots, k$ .

A polycycle  $\gamma$  is called *monodromic* if one can choose a segment  $\Sigma$  transversal to  $\gamma$  such that on one side  $U \subset \Sigma$  of  $p = \Sigma \cap \gamma$  a Poincaré return map  $\Delta_\gamma: U \rightarrow \Sigma$  is defined with  $\Delta_\gamma(p) = p$ .

NONACCUMULATION THEOREM ([I3, É]). *For any analytic monodromic polycycle  $\gamma$  there is a tube neighborhood free from limit cycles or a Poincaré return map  $\Delta_\gamma: \Sigma \supset U \rightarrow \Sigma$  can't have infinitely many fixed points accumulating to  $p = \Sigma \cap \gamma$ .*

This theorem along with the above compactness arguments implies IFT. Both proofs of Ilyashenko and Ecalle deal with the analysis of type of germs of maps arising as Poincaré return maps of an analytic monodromic polycycle. The lectures by van den Dries discussed the theory of o-minimality.<sup>1</sup> This theory deals with classes of functions which satisfy certain axioms. A basic example of o-minimal class of functions is polynomials and analytic functions. In particular, if a map of a compact interval  $\Delta: U \rightarrow \mathbb{R}$  belongs to an o-minimal class of functions, then the equation  $\Delta(x) = x$  has finitely many solutions. One hopes that deeper understanding of o-minimal structures would allow one to include Poincaré return maps of monodromic polycycles into an o-minimal class and give an independent proof of

<sup>1</sup>For information on o-minimality and on model theory and its applications cf. [vdD, vdDS] and Lou van Dries webpage <http://www.math.uiuc.edu/People/vddnest.html>.

the nonaccumulation theorem. Finiteness theorems for differentiable function fields are discussed in the lectures by Buium [B2] in the present volume.

**DEFINITION 2.** Let  $\{\dot{x} = v(x, \varepsilon)\}_{\varepsilon \in B^k}$ ,  $x \in \mathbb{S}^2$ , be a  $k$ -parameter family of vector fields on  $\mathbb{S}^2$  having a polycycle  $\gamma$  for some parameter value  $\varepsilon_* \in B^k$ , where  $B^k$  is the unit ball in  $\mathbb{R}^n$ . The polycycle  $\gamma$  has cyclicity  $\mu$  in the family  $\{v(x, \varepsilon)\}_{\varepsilon \in B^k}$  if there exist neighborhoods  $U$  and  $V$  such that  $\mathbb{S}^2 \supseteq U \supset \gamma$ ,  $B^k \supseteq V \ni \varepsilon_*$  and for any  $\varepsilon \in V$  the field  $v(\cdot, \varepsilon)$  has no more than  $\mu$  limit cycles inside  $U$  and  $\mu$  is the minimal number with this property.

**EXAMPLES.** (1) In a generic  $n$ -parameter family, the maximal multiplicity of a degenerate limit cycle does not exceed  $n + 1$ , e.g., in codimension 1 a semistable limit cycle has cyclicity 2. Thus, the cyclicity of a trivial polycycle (a polycycle without singular points) in a generic  $n$ -parameter family does not exceed  $n + 1$ .

(2) (Andronov–Leontovich, 1930s; Hopf, 1940s). A nontrivial polycycle of codimension 1 has cyclicity at most 1.

(3) (Takens, Bogdanov, Leontovich, Mourtada, Grozovskii, early 1970s–1993 (see [G2, KS], and references there)). A nontrivial polycycle of codimension 2 has cyclicity at most 2.

There is no standard definition of a polycycle. For example in [DRR] the authors use the following definition: A *graphic* is a collection of singularities connected by phases curves. A polycycle is a graphic for which the first Poincaré return map is defined.

**DEFINITION 3.** The bifurcation number  $B(k)$  is the maximal cyclicity of a nontrivial polycycle occurring in a generic  $k$ -parameter family of  $C^\infty$  vector fields.

The definition of  $B(k)$  does not depend on a choice of the base of the family, it depends only on the number  $k$  of parameters.

**LOCAL HILBERT – ARNOL'D PROBLEM (LHAP) ([I4]).** *Prove that for any finite  $k$ , the bifurcation number  $B(k)$  is finite and find an upper estimate for  $B(k)$ .*

It turns out that a solution to the local Hilbert–Arnol'd problem implies a solution to the Hilbert–Arnol'd problem.

Similarly to the generic smooth vector fields, in the case of analytic vector fields one can define a so-called *limit periodic set* [FP, R1], which is either a polycycle or has an arc of equilibrium points,<sup>2</sup> and formulate the

**LOCAL FINITENESS CONJECTURE (LFC) ([R1]).** *Prove that any limit periodic set occurring in an analytic family of vector fields on  $\mathbb{S}^2$  has finite cyclicity in this family.*

Smooth vector fields are more flexible than analytic vector fields and easier to analyze. A strategy to attack the existential Hilbert problem, proposed by Arnol'd [AA] (see also [IK]), is first to understand generic smooth vector fields and then try to apply the developed methods to analytic vector fields. We stress out that a solution to the (local) Hilbert–Arnol'd problem would not necessarily solve the (finite cyclicity) global finiteness conjecture, because the study of generic vector fields does not include the study of nongeneric vector fields and nongeneric limit sets, which are unavoidable in the analytic setting. However, it might give some insight. Let us summarize the discussion in the form of a part of Figure 1.

<sup>2</sup>Generic vector fields cannot have an arc of equilibrium points.

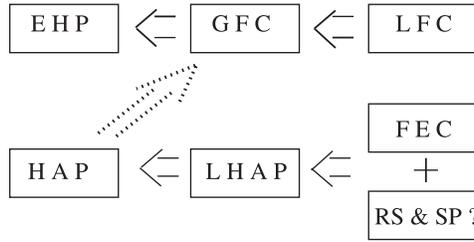


FIGURE 1. Existential Hilbert problem and its offsprings

1.1.3. *Cyclicity of elementary polycycles.* Now we shall formulate the main result of this course of lectures.

DEFINITION 4. An equilibrium point of a vector field on the two-sphere is called *elementary* if at least one eigenvalue of its linear part is nonzero. A polycycle is called an *elementary* polycycle if all its singularities are elementary.

The local Hilbert – Arnol’d problem was solved under the additional assumption that a polycycle have only elementary singularities.

DEFINITION 5. The elementary bifurcation number  $E(k)$  is the maximal cyclicity of a nontrivial elementary polycycle occurring in a generic  $k$ -parameter family of  $C^\infty$  vector fields.

From Examples (2) and (3) above it follows that

$$E(1) = 1, \quad E(2) = 2.$$

Information about the behavior of the function  $k \mapsto E(k)$  has been obtained recently. The first crucial step was done by Ilyashenko and Yakovenko:

FINITENESS OF ELEMENTARY CYCLICITY (FEC) ([IY2]). *For any  $n$  the elementary bifurcation number  $E(n)$  is finite.*

COROLLARY 1. *Under the assumption that generic  $C^\infty$  families of vector fields have only elementary singularities, the global Hilbert – Arnol’d problem is solved, i.e., any generic finite-parameter  $C^\infty$  family of vector fields on the sphere  $\mathbb{S}^2$  with a compact base and only elementary equilibria has a uniform upper bound for the number of limit cycles.*

MAIN THEOREM ([K6]). *For any  $k \in \mathbb{Z}_+$*

$$(5) \quad E(k) \leq 2^{25k^2}.$$

This is the first explicit general estimate for cyclicity of polycycles. The case of a polycycle consisting of only one singular point with no arcs at all is well known. An elementary equilibrium point can generate limit cycles in its small neighborhood if it is a slow focus, that is, the linearization matrix has a pair of two imaginary eigenvalues. This bifurcation was investigated by Takens [T1].

COROLLARY 2. *As in Corollary 1, under the assumption that all the polycycles are elementary the Main Theorem gives a solution to the local Hilbert – Arnol’d problem.*

1.1.4. *Resolution of singularities (RS) or blow-up of singularities of vector fields and singular perturbations (SP)*. In this section we discuss resolution of singularities and singular perturbations which might lead to a generalization of the main result and to a solution to the local Hilbert – Arnol'd problem (see the box with **RS & SP?** in Figure 1).

Let  $\dot{x} = v(x)$  be a  $C^\infty$  vector field on  $\mathbb{R}^2$  such that  $v(0) = 0$ . A vector field satisfies a *Lojasiewicz condition* if there exists  $k \in \mathbb{Z}_+$  and  $c > 0$  such that

$$(6) \quad \|v(x)\| \geq c\|x\|^k$$

for all  $x$  in some neighborhood of 0. It can be shown [**D2**] that any generic  $C^\infty$  finite-parameter family of vector fields on the sphere  $\mathbb{S}^2$  has only vector fields with equilibrium points satisfying a Lojasiewicz condition for some  $k \in \mathbb{Z}_+$  and  $c > 0$ .

To define a *blow-up* for a  $C^\infty$  vector field  $\dot{x} = v(x)$  on  $\mathbb{R}^2$  with an equilibrium at 0, i.e.,  $v(0) = 0$  consider the map

$$(7) \quad \phi: \mathbb{S}^1 \times \mathbb{R} \rightarrow \mathbb{R}^2; \quad \phi(\theta, r) \mapsto (r \cos \theta, r \sin \theta).$$

Then the pull-back  $\hat{v}$ , with  $\phi(\hat{v}) = v$ , is a  $C^\infty$  vector field on  $\mathbb{S}^1 \times \mathbb{R}$ , i.e.,  $d\phi_0(\hat{v}(0)) = X \circ \phi(0)$ , where  $\hat{v}$  is the blown-up vector field.

**DESINGULARIZATION THEOREM ([D2]).** *If a  $C^\infty$  vector field  $\dot{x} = v(x)$  on  $\mathbb{R}^2$  with  $v(0) = 0$  satisfies a Lojasiewicz condition, then there is a finite sequence of blow-ups leading to a vector field with only elementary equilibria.*

Sometimes this theorem is called the Bendixson – Seidenberg – Dumortier theorem [**B, S2, D2**]. Bendixson stated it without a proof it on the brink of the twentieth century. Seidenberg proved it in the analytic case and Dumortier did it for  $C^\infty$  vector fields with a Lojasiewicz condition. A quantitative version of the desingularization theorem, which estimates the number of necessary blow-ups, was obtained by Kleban [**K3**].

This theorem reduces consideration of an *individual* vector field, occurring in a generic finite-parameter family, with equilibria without restriction to an individual vector field with only *elementary* equilibria.

However, in order to extend an estimate on cyclicity of elementary polycycles (5) to an estimate on cyclicity of a generic nonelementary polycycle (LHAP) one needs the *desingularization theorem for families* of generic  $C^\infty$  vector fields. Different approaches to attack this problem were proposed by Denkowska and Roussarie [**DR**] and by Trifonov [**T4**].

An approach proposed by Trifonov leads to the dynamical phenomenon called *Singular Perturbation* (SP): in the simplest case one needs to analyze families of vector fields on the plane which for some values of the parameters have a curve of equilibria. Certainly a generic finite-parameter family of vector fields has no curve of equilibria; however, after even one step of blow-up such a curve can occur [**T4**]. Appearance of curves of equilibria after a desingularization in a family now seems to be the main obstacle standing between an estimate on cyclicity of elementary polycycles (5) and the local Hilbert – Arnol'd problem (see [**T4, IY2, R2**] for more).

**1.2. Bifurcation of spatial polycycles and multiplicity of generic germs.** In this part we present by-product results of the Main Theorem. The first result is an extension of the Main Theorem on the estimate of cyclicity of a planar elementary polycycle to an estimate on cyclicity of a spatial quasidelementary polycycle (see Section 1.2.1). The second result gives an estimate on cyclicity of

generic germs of smooth mappings which is a partial answer to Arnol'd's question [A] (see Section 1.2.4).

1.2.1. *Bifurcation of spatial polycycles.* The definition of a polycycle in the multidimensional case is the same as in the planar case (see Definition 1). When Ilyashenko and Yakovenko proved the Finiteness of Cyclicity for elementary polycycles, Arnol'd posed the question: *What can be said about bifurcations of spatial polycycles?*

Another sufficient reason to look at this problem is because the planar argument (the Poincaré–Bendixson theorem) implies that a collection of an infinite number of limit cycles of uniformly bounded length, located in a bounded domain, accumulates to a limit cycle. Indeed, consider a vector field  $\dot{x} = v(x)$  of finite codimension in  $\mathbb{R}^3$  (dimension 3 can be replaced by any  $N > 2$  anywhere in this section), i.e., a vector field which occurs in a generic  $C^\infty$  finite-parameter family. Then  $v(x)$  has only isolated singular points. Fix a positive number  $L$  and assume that in a compact region of the phase space there is an infinite number of phase curves of length less than  $L$  corresponding to limit cycles of  $v(x)$ . Then a subset of these limit cycles must accumulate to a separatrix polygon (polycycle).

Bifurcation properties of spatial polycycles are much richer than those of planar polycycles. The first important 3-dimensional feature is the existence of limit cycles that wind several times around a polycycle. This happens because a Poincaré return map is a 2-dimensional map and it might have not only fixed points, but also periodic points of higher periods as well. We call a periodic trajectory that “turns” around a whole polycycle exactly  $m$  times before closing up an  $m$ -cycle. Such a trajectory corresponds to a periodic point of a corresponding Poincaré return map of minimal period  $m$ . On the plane, for topological reasons, only 1-cycles exist. The definition of cyclicity requires some additional care.

Consider an  $n$ -parameter family of flows  $\{\dot{x} = v(x, \varepsilon)\}_{\varepsilon \in B^n}$  in  $\mathbb{R}^3$ . Let  $\gamma \subset \mathbb{R}^3$  be a polycycle of the field  $\dot{x} = v(x, \varepsilon^*)$  for some  $\varepsilon^* \in B^n$ . Then  $\gamma$  can be represented as a union of a finite number of equilibrium points  $\{p_j\}_{j \in J}$  and phase curves  $\{\gamma_j\}_{j \in J}$  connecting them. A *tube neighborhood*  $T_\gamma$  of the polycycle  $\gamma$  is a union of neighborhoods of equilibria  $\{p_j\}_{j \in J}$  and tube neighborhoods  $\{T_j\}_{j \in J}$  of phase curves  $\{\gamma_j\}_{j \in J}$ .

DEFINITION 6. Let  $m \in \mathbb{Z}_+$ . Then  $m$ -cyclicity of the polycycle  $\gamma$  in the family  $\{\dot{x} = v(x, \varepsilon)\}_{\varepsilon \in B^n}$ , denoted by  $\mu(m, \gamma)$ , is the minimal number  $\mu(m, \gamma)$  for which there is a tube neighborhood  $T_\gamma$  of the polycycle  $\gamma \subset T_\gamma \subset \mathbb{R}^3$ , a neighborhood  $V$  of the parameter  $\varepsilon^* \in V \subset B^n$  and for each  $j \in J$  a Poincaré section  $L_{\gamma, j}$  (a hyperplane) transversal to the corresponding  $\gamma_j$  such that the following condition holds:

- (1) for any parameter  $\varepsilon \in V$  the corresponding vector field  $\dot{x} = v(x, \varepsilon)$  has at most  $\mu(m, \gamma)$  limit cycles in  $T_\gamma$ ;
- (2) Each of these limit cycles  $l_i(\varepsilon)$  intersects each Poincaré section  $L_{\gamma, j}$  in exactly  $m$  different points;
- (3) In the sense of the Hausdorff metric for each  $j \in J$  the distance between each part of  $l_j(\varepsilon)$  which lies between two consecutive intersections of  $L_{\gamma, j}$  and the polycycle  $\gamma$  tends to 0 as  $\varepsilon$  tends to 0.

Now we discuss a classical example of a polycycle which has infinite  $m$ -cyclicity for any  $m \geq 1$ .

1.2.2. *The Shil'nikov polycycle.* Consider a flow  $\phi_t$  in  $\mathbb{R}^3$  with a hyperbolic equilibrium point  $O$  that has one positive eigenvalue  $\lambda$  and two complex conjugates  $\mu \pm \omega i$  with negative real part. Suppose that the sum of  $\lambda + \mu$  is positive, and the unstable one-dimensional manifold returns to the stable one, which is two-dimensional. Thus, the equilibrium  $O$  has a homoclinic orbit that tends back to  $O$  along the unstable manifold as  $t \rightarrow -\infty$ , and spirals around  $O$  on the stable manifold as  $t \rightarrow +\infty$ . In 1965 Shil'nikov [S3] discovered that the Poincaré map along this polycycle has a countable number of pairwise disjoint subdomains so that a restriction to any of them gives a Smale horseshoe. Any such horseshoe is structurally stable, therefore, the polycycle described above (the Shil'nikov polycycle) has infinite  $m$ -cyclicity for all  $m \in \mathbb{Z}_+$  (see, e.g., [GH, IL]). The codimension of this polycycle is 1.

However, it seems reasonable to state the following

CONJECTURE 1 (Arnol'd–Ilyashenko–Yakovenko). If a spatial polycycle  $\gamma \in \mathbb{R}^3$  has finite codimension  $k$  and all its equilibrium points are saddles with real eigenvalues or saddlenodes with at most one zero eigenvalue and the other eigenvalues are real, then the  $m$ -cyclicity of  $\gamma$ , denoted  $C(m, \gamma)$ , is finite for each  $m \in \mathbb{Z}_+$ .

Using the ideas and methods for the planar problem and a result of Grigoriev–Yakovenko [GY], the Arnol'd–Ilyashenko–Yakovenko's conjecture has been solved in arbitrary dimension  $N > 2$  with additional nondegeneracy assumptions on the polycycle's equilibria.

1.2.3. *An estimate of the cyclicity of a quasidelementary spatial polycycle.* In the planar case we considered polycycles with elementary equilibria only, now we define a class of points called *quasidelementary equilibria*. The author has shown that polycycles with only quasidelementary equilibria, have finite  $m$ -cyclicity for any  $m \geq 1$ . Moreover, there exists an explicit upper bound for  $m$ -cyclicity.

Recall some standard definitions from normal form theory.

DEFINITION 7. The set of complex numbers  $\lambda_1, \dots, \lambda_N \in \mathbb{C}$  is called:

- (a) nonresonant if there is no integral relation among the numbers  $\lambda_j$  of the form  $\lambda_j = \sum_{i=1}^N k_i \lambda_i$ , where  $k_i \in \mathbb{Z}_+$  for  $i = 1, \dots, n$  and  $\sum_{i=1}^N k_i \geq 2$ .

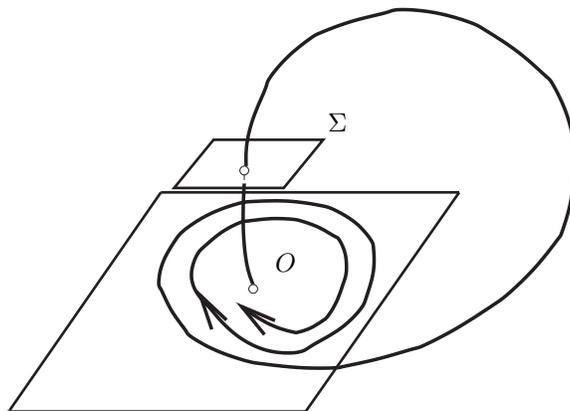


FIGURE 2. The Shil'nikov polycycle

- (b) *strongly simply resonant* if all the nontrivial resonance relations  $\lambda_j = \sum_{i=1}^N k_i \lambda_i$ , where  $k_i \in \mathbb{Z}_+$ ,  $i = 1, \dots, n$  and  $\sum_{i=1}^N k_i \geq 2$ , follow from a single one  $\sum_{i=1}^N k_i^* \lambda_i = 0$ , where  $k_i^* \in \mathbb{Z}_+$ .

DEFINITION 8. We shall call an equilibrium point of a differential equation *quasielementary* if the linearization matrix of the equation at this point has only *real* eigenvalues, at most one of them is zero, and they satisfy one of the following conditions:

- (1) they are nonresonant, and we call such an equilibrium a *nonresonant saddle*;
- (2) they form a strongly simply resonant set of numbers — a *strongly simply resonant saddle*;
- (3) one eigenvalue is zero with Lojasiewicz exponent 2 (see (6)) and the others form a nonresonant set.

A polycycle is called *quasielementary* if all its vertices are quasielementary.

Note that the class of quasielementary points in the case of the plane ( $N = 2$ ) coincides with the class of elementary points, except for the multiplicity two condition for saddlenodes. In a sense, Theorem 1 below is a generalization of Theorem 5 of Section 2.2 below.

DEFINITION 9. The *quasielementary bifurcation number*  $\text{QE}(N, n, m)$  is the maximal  $m$ -cyclicity of a quasielementary polycycle occurring in a generic  $n$ -parameter family of vector fields in  $\mathbb{R}^N$ .

THEOREM 1 ([K8]). *For any positive integer  $N$  (dimension of the phase space),  $n$  (number of parameters),  $m$  (number of turns around a polycycle), and  $T = 6Nm$  we have*

$$\text{QE}(N, n, m) \leq 2^{T^2}.$$

In the next section we describe another by-product of the Main Theorem.

1.2.4. *Geometric multiplicity of germs of generic maps.* Let  $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$  be a generic  $C^k$  smooth map,  $k \geq n + 1$ . Fix a point  $a \in \mathbb{R}^n$  and denote  $F(a)$  by  $b$ .

DEFINITION 10. The geometric multiplicity of a map germ  $F: (\mathbb{R}^n, a) \rightarrow (\mathbb{R}^n, b)$  at  $a$ , denoted by  $\mu_a^G = \mu_a^G(F)$ , is the maximal number of isolated preimages  $F^{-1}(\tilde{b})$  close to  $a$ :

$$(8) \quad \mu_a^G(F) = \limsup_{r \rightarrow 0} \sup_{\tilde{b} \in \mathbb{R}^n} \#\{x \in B_r(a) : F(x) = \tilde{b}\}.$$

For example, the geometric multiplicity of the function  $f: x \rightarrow x^2$  at 0 is two, but the geometric multiplicity of  $f: x \rightarrow x^3$  at 0 is one, even though 0 is a degenerate point of second order.

In the complex case the geometric multiplicity equals the usual multiplicity (see, e.g., [AGV]). In the real case the first is no greater than the second.

DEFINITION 11. Define the geometric multiplicity of  $n$ -dimensional germs,  $\mu^G(n)$ , as follows. Let  $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$  be a generic map. The geometric multiplicity of  $F$  equals the least upper bound of the geometric multiplicities of  $\mu_a^G(F)$  taken over all points  $a \in \mathbb{R}^n$ . Then the geometric multiplicity of  $n$ -dimensional

germs is the maximum of the geometric multiplicities of all generic maps  $F$  from  $\mathbb{R}^n$  to  $\mathbb{R}^n$

$$(9) \quad \mu^G(n) = \sup_{\substack{F\text{-generic} \\ a \in \mathbb{R}^n}} \mu_a^G(F).$$

It turns out that the geometric multiplicity of  $n$ -dimensional germs is finite for all positive integers  $n$  and depends only on the dimension  $n$ .

REMARK 1. For example, for  $n = 2$  the Whitney theorem about maps of surfaces states that a generic  $C^2$  map of two-dimensional manifolds  $F: M^2 \rightarrow N^2$  can have only three different types of germs: one-to-one, a fold, and a pleat (see, e.g., [AGV]). This implies that  $\mu^G(2) = 3$ .

A natural problem posed by Arnol'd [A] is to give estimates for the geometric multiplicity  $\mu^G(n)$  of  $n$ -dimensional germs.

In the case of complex analytic maps of  $\mathbb{C}^n$  into  $\mathbb{C}^n$  Gabriellov and Khovanskiĭ [GAK, Theorem 7] obtained an estimate on  $\mu^G(n)$  of the type:  $\mu^G(n) \leq n^n$ . The upper bound for the geometric multiplicity for  $n$ -dimensional smooth germs of generic maps is given by

THEOREM 2 ([K6]). *The geometric multiplicity of germs of a generic  $C^k$  smooth map  $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $k > n$  admits the following upper estimate:*

$$(10) \quad \mu_a^G(F) \leq 2^{n(n-1)/2+1} n^n, \quad \forall a \in \mathbb{R}^n.$$

Using the same method one can prove

THEOREM 3 ([K6]). *Let  $F: \mathbb{R}^n \rightarrow \mathbb{R}^N$  be a generic  $C^k$  smooth map with  $k > n$ ,  $N \geq n$  and  $P: \mathbb{R}^N \rightarrow \mathbb{R}^n$  be a polynomial of degree  $d$ . Then the geometric multiplicity of germs of a chain map  $P \circ F: \mathbb{R}^n \rightarrow \mathbb{R}^n$  admits the following upper estimate:*

$$(11) \quad \mu_a^G(P \circ F) \leq 2^{n(n-1)/2+1} (dn)^n, \quad \forall a \in \mathbb{R}^n.$$

An interesting feature of this theorem is that the geometric multiplicity does not depend on the dimension  $N$  of the intermediate space.

The problem of an upper estimate of geometric multiplicity of germs of generic smooth maps  $F: (\mathbb{R}^n, 0) \rightarrow (\mathbb{R}^n, 0)$  or chain maps  $P \circ F: (\mathbb{R}^n, 0) \rightarrow (\mathbb{R}^n, 0)$  is closely related to the problem of estimating the cyclicity of elementary polycycles as the reader will see below.

All the results (the Main Theorem, Theorem 1, Theorem 2, and Theorem 3) were announced in [K6].

**1.3. Three stages of the proof of the Main Theorem and outline of the content of the lectures.** The Main Theorem is a quantitative extension of the Ilyashenko–Yakovenko finiteness theorem. The paper of Ilyashenko–Yakovenko [IY2] was a cornerstone for the present paper. In [IY2] the authors made an important step: they found a route from *bifurcation theory* to *singularity theory* using the Khovanskiĭ reduction method [K1]. In [K6] we follow this route at the beginning and using some new ideas get an estimate for the cyclicity of elementary polycycles. Below we outline the main steps of the proof of the Main Theorem and describe the contents of the coming sections.

The proof of the Main Theorem consists of three steps. The relation to the proof of the finiteness theorem [IY2] is discussed in Section 1.3.2.

Step 1. *Normal forms for local families of vector fields and their integration.* We use normal forms to establish an explicit form for the Poincaré correspondence map near equilibrium points on the polycycle under consideration. In [MR] and later in [IY2] it is noticed that these maps satisfy Pfaffian (polynomial differential) equations with the coefficients of polynomials depending smoothly on the parameters of the family. As a result a *basic system* of equations, determining the number of limit cycles, is obtained.

Step 2. *The Khovanskii reduction method.* We discuss a variation of the Khovanskii method [K2]. This method allows us to investigate systems of equations that involve functions satisfying Pfaffian equations. It turns out that the number of solutions of the basic system can be estimated by the number of solutions of a *mixed functional-Pfaffian* system. After an application of the Khovanskii method to the mixed functional-Pfaffian system we obtain several *chain maps*: the maps of the form

$$(12) \quad x \mapsto (P_1, \dots, P_n) \circ (x, f(x), f'(x), \dots, f^{(n)}(x)),$$

where  $x$  is a point near  $0 \in \mathbb{R}^n$ ,  $f$  is a generic function,  $f^{(k)}(x)$  is the collection of all derivatives of  $f$  of order  $k$ , and  $P = (P_1, \dots, P_n)$  is a vector-polynomial given by its coordinate functions of known degree.

It turns out that the problem of estimating the number of limit cycles reduces to estimating the number of small regular preimages of some *special* points of the chain map. Special points form an open cone-like semialgebraic set  $K$  approaching  $0$  in the image, e.g., if  $K \subset \mathbb{R}^2$ , then  $K = \{(x_1, x_2) : 0 < x_2 < x_1^m\}$  for some  $m \in \mathbb{Z}_+$ .

Denote by  $F$  the map  $F: x \mapsto (x, f(x), f'(x), \dots, f^{(n)}(x))$  which is called the  $n$ th jet of  $f$ . Denote by  $L_F$  the linearization of  $F$  at the point  $x = 0$ .

Lecture 2 highlights Step 1 and in a simple nontrivial case.

Step 3. *Bézout's theorem for the Chain maps.* We shall construct an algebraic set  $\Sigma$  in the image of  $F$  (in the space of  $n$ -jets) so that if  $F$  is transversal to  $\Sigma$ , then the number of preimages of any point  $a$  from a set of special points  $K$  is *the same* for  $F$  and its linearization  $L_F$  at zero:

$$(13) \quad \#\{x : P \circ F(x) = a\} = \#\{x : P \circ L_F(x) = a\} \leq \prod_{j=1}^k \deg P_j.$$

Since  $L_F$  is a linear map, one can apply Bézout's theorem to estimate the right-hand part of the equality. This observation completes the proof of the Main Theorem.

In order to prove existence of such a set  $\Sigma$  we need to apply stratification theory originating in works of Whitney [W2], Thom [T2], and Mather [M2]. More exactly, we need to prove the existence of the so-called  $a_P$ -stratification introduced by Thom in some special case [K6]. Lecture 3 presents necessary notions from stratification theory and states the required result on existence of an  $a_P$ -stratification.

1.3.1. *Multichain maps and bifurcation of spatial polycycles.* In order to get an estimate on cyclicity of spatial polycycles in Theorem 1 we face the problem of estimating the geometric multiplicity of *multichain maps* of the form

$$(14) \quad P \circ (F, F): B^n \times B^n \rightarrow \mathbb{R}^{2n},$$

where  $B^n \subset \mathbb{R}^n$  is a unit ball,  $F: B^n \rightarrow \mathbb{R}^N$  is a generic map and  $P: \mathbb{R}^{2N} \rightarrow \mathbb{R}^{2n}$  is a vector-polynomial of known degree. The appearance of this problem is described

$$\begin{array}{ccc}
 & & \mathcal{DD}_2(B^n, \mathbb{R}^N) \\
 & \nearrow \mathcal{D}_2(F) & \downarrow \pi_2 \\
 B^n \times B^n & \xrightarrow{(F, F)} & \mathbb{R}^{2N}
 \end{array}$$

FIGURE 3. Polynomial blow-up of the multijet space

with many more details in Lecture 4. It is no longer possible to treat the 2-tuple map

$$(15) \quad (F, F): B^n \times B^n \rightarrow \mathbb{R}^N \times \mathbb{R}^N$$

as a generic map.

Step 4. *Blow-up along the diagonal in the multijet space.* Grigoriev and Yakovenko [GY] constructed a so-called *space of divided differences* or  $\mathcal{DD}_2$ -space and the commutative diagram of Figure 3 where  $\mathcal{D}_2(F): B^n \times B^n \rightarrow \mathcal{DD}_2(B^n, \mathbb{R}^N)$  is a smooth map,  $\pi_2: \mathcal{DD}_2(B^n, \mathbb{R}^N) \rightarrow \mathbb{R}^{2N}$  is an explicitly computable polynomial and

$$(16) \quad \pi_2 \circ \mathcal{D}_2(F) = (F, F): B^n \times B^n \rightarrow \mathbb{R}^{2N}.$$

It turns out that one can treat  $\mathcal{D}_2(F)$  as a generic map for a generic  $F$  and impose various transversality conditions. Therefore, we can represent the multichain map  $P \circ (F, F)$ , given by (14), in the form

$$(17) \quad P \circ (F, F) = (P \circ \pi_2) \circ \mathcal{D}_2(F),$$

where  $P \circ \pi_2$  is a polynomial since  $\pi_2$  is a polynomial, and  $\mathcal{D}_2(F)$  is a smooth map. Moreover, it turns out that  $\mathcal{D}_2(F)$  is generic for a generic  $F$ . Now we can apply Bézout's theorem to the chain map  $(P \circ \pi_2) \circ \mathcal{D}_2(F)$ . In Lecture 4 we describe this construction with details and in greater generality and exhibit application of this construction to an old problem on the rate of growth of the number of periodic points for generic diffeomorphisms in smooth dynamic systems (see, e.g., [AM, S4]).

1.3.2. *Relation of the proof of the Main Theorem and the Ilyashenko–Yakovenko finiteness theorem [IY2].* Steps 1 of both proofs [IY2] and [K6] are the same. We shall just present the table of required normal forms from [IY2], which were obtained in [IY1]. Step 2 in this proof is slightly different from the one in [IY2] and this is the first novel point. After application of the Khovanskiĭ method we obtain the same collection of chain maps of the form (12) as in [IY2]. However, in [IY2] the authors investigate the number of regular preimages of points in the image by the chain maps *without any restriction* on those points. In the present proof, using new additional arguments in the Khovanskiĭ method, we reduce consideration to only preimages of *special* points, i.e., points from a tiny cone-like set in the image. At this point our proof continues on independently. The proof from [K6] can be considered as an independent simplified proof of the Ilyashenko–Yakovenko's finiteness theorem modulo the derivation of a mixed functional-Pfaffian system.

*Acknowledgments.* I would like to thank Dana Schlomiuk for giving me a great opportunity to give a course of lectures in a workshop held in Montréal during June 2000 and for careful reading of the present manuscript. Many helpful remarks improved its overall quality. Special thanks go to my teacher Yulij Ilyashenko who patiently taught me bifurcation theory and whose long lasting support and encouragement have been a crucial source of energy. Discussions with William

Cowieson, Andrei Gabrielov, Askold Khovanskiĭ, Pavao Mardešić, John Mather, Robert Moussu, Oleg Shelkovernikov, and Sergei Yakovenko have been very helpful for me. I would like also to thank Christiane Rousseau and Robert Roussarie for inviting me to give series of lectures on the bifurcation of limit cycles in Montréal and Dijon and for useful discussions. These series of lectures have been extremely helpful in improving the presentation of the present lectures. I would like to acknowledge financial support and the very productive atmosphere of IPST, University of Maryland and Courant Institute of Mathematical Science, NYU, where various parts of the work have been done. While in the Institute for Physical Sciences and Technology I enjoyed fruitful discussions with James Yorke and Brian Hunt.

### 2. Normal forms and the Khovanskiĭ method

We explain the proof of the Main Theorem in the simplest nontrivial case  $n = 2$ . Consider a generic 2-parameter  $C^\infty$  family of vector fields  $\{\dot{x} = v(x, \varepsilon)\}_{\varepsilon \in B^2}$  and suppose that for  $\varepsilon = 0$  the vector field  $\dot{x} = v(x, 0)$  has a polycycle  $\gamma$  which consists of two saddles  $p_1, p_2$  and two separatrices  $\gamma_1, \gamma_2$  connecting them. Consider a segment  $\Sigma$  transversal to, say,  $\gamma_1$  and denote by  $\Delta: \Sigma \supset U \rightarrow \Sigma$  the Poincaré return map, which is defined on some open set  $U$  in  $\Sigma$ . In order to estimate the number of limit cycles bifurcating from the polycycle  $\gamma$  we need to estimate the number of isolated fixed points  $\#\{x \in U : \Delta(x) = x\}$ .

Using the standard approach we decompose the Poincaré map  $\Delta$  into a composition of four maps: two local  $\Delta_1$  and  $\Delta_2$  in neighborhoods of equilibria  $p_1$  and  $p_2$  respectively and two semilocal  $f_1$  and  $f_2$  along connecting separatrices  $\gamma_1$  and  $\gamma_2$  respectively to be defined precisely below (see Figure 4). After that we replace the equation  $\Delta(x) = x$  by the system of equations corresponding to  $\Delta \equiv \Delta_2 \circ f_2 \circ \Delta_1 \circ f_1$ . To understand properties of local maps  $\Delta_i, i = 1, 2$  we use normal form theory.

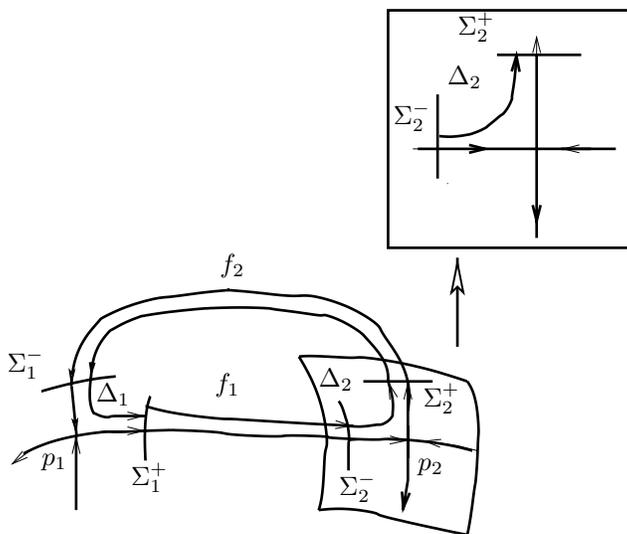


FIGURE 4. Construction of “entrance” and “exit” transversals

We say that a finite parameter family of vector fields is generic if generic finite parameter families form a dense set and pass through every point. It is essential that generic families pass through every point. This is somewhat similar to the notion of prevalence [K1].

### 2.1. Normal forms and a basic system determining the number of limit cycles.

2.1.1. *Polynomial normal forms of local families and Pfaffian Poincaré return maps.* It turns out that in a small neighborhood of an elementary equilibrium point there exist finitely differentiable normal coordinates (in the Cartesian product of the phase space and the parameter space), the so-called normal form of an equilibrium point. The list of finitely differentiable normal forms was obtained in [IY1]. The main feature of the list: all normal forms are *polynomial and integrable*. The smaller the neighborhood of a normal form, the higher is its smoothness. So smoothness can be chosen arbitrary large. All normal forms are summarized in Table 1 below.

In a small neighborhood of an elementary equilibrium point one can choose two small segments, say  $\Sigma^-$  and  $\Sigma^+$ , transversal to the vector field for the critical value of the parameter and explicitly calculate the Poincaré (correspondence) map which maps a point from one segment, say  $\Sigma^-$ , along the corresponding phase curve to a point from the other segment  $\Sigma^+$  (see Figure 4). For an appropriate choice of segments  $\Sigma^-$ ,  $\Sigma^+$  and coordinate functions  $x$ ,  $y$  in  $\Sigma^-$ ,  $\Sigma^+$  respectively, and a smooth function  $\lambda(\varepsilon)$  in the original parameter  $\varepsilon$  of the family, the Poincaré return map  $\Delta_\varepsilon: x \rightarrow y$ ,  $\varepsilon \in \mathbb{R}^k$ , can be explicitly computed. Moreover, there is a Pfaffian equation with 1-form  $\omega$  of the form

$$(18) \quad P(x, y, \lambda(\varepsilon)) dx + Q(x, y, \lambda(\varepsilon)) dy = 0$$

vanishing on the graph of  $y = \Delta_\varepsilon(x)$ , where  $P(x, y, \lambda(\varepsilon))$  and  $Q(x, y, \lambda(\varepsilon))$  are polynomials in  $x$  and  $y$ . This was first noticed by Moussu – Roche [MR].

EXAMPLE 1. Consider a nonresonant saddle on the plane in a family of  $C^\infty$  vector fields. There is a normal form which after an appropriate rescaling is given by the equation

$$(19) \quad \begin{cases} \dot{x} = \lambda_1(\varepsilon)x \\ \dot{y} = -\lambda_2(\varepsilon)y, \end{cases}$$

where  $\lambda_1(\varepsilon)$  and  $\lambda_2(\varepsilon)$  are smooth functions and two transversal “exit”-“entrance” sections are  $\Sigma^- = \{y = 1\}$ ,  $\Sigma^+ = \{x = 1\}$ .

Then for  $\lambda(\varepsilon) = \lambda_1(\varepsilon)/\lambda_2(\varepsilon)$  the function  $u(t) = x(t)y^{\lambda(\varepsilon)}(t)$  is a first integral. Therefore, if a trajectory starts from  $(x(0), y(0)) \in \Sigma^-$  and ends at  $(x(t^*), y(t^*)) \in \Sigma^+$  (see Figure 5,  $S_\mu$ -case), then  $y^{\lambda(\varepsilon)}(t^*) = x(0)$  or in the induced on  $\Sigma^-$  and  $\Sigma^+$  coordinates from  $\mathbb{R}^2$  we have  $y^{\lambda(\varepsilon)} = x$ . It is easy to see that the 1-form  $\omega = x dy + \lambda(\varepsilon)y dx$  vanishes on the graph  $y^{\lambda(\varepsilon)} = x$ . All necessary information about normal forms, Poincaré correspondence maps, and corresponding Pfaffian forms is in the Table 1 below. For completeness let's give the necessary definitions from the theory of normal forms.

2.1.2. *Definitions and a collection of normal forms.* A local family of planar vector fields is the germ of a map,

$$v: (\mathbb{R}^2, 0) \times (\mathbb{R}^k, 0) \rightarrow (\mathbb{R}^2, 0), \quad (x, y, \varepsilon) \mapsto v(x, y, \varepsilon).$$

TABLE 1. Unfolding of elementary equilibrium points on the plane

Type	Normal forms	Poincaré correspondence maps	Pfaffian equations
$S_0$	$\dot{x} = x, \dot{y} = -\lambda y,$ $\lambda = \lambda_0 \in \mathbb{R}^1$	$y = x^\lambda,$ $x > 0, y > 0$	$x dy - \lambda y dx = 0$
$S_\mu$	$\dot{x} = x(n/m + P_\mu(u, \lambda)),$ $\dot{y} = -y,$ $u = u(x, y) = x^m y^n,$ $P_\mu(u, \lambda) = \pm u^\mu(1 + \lambda_\mu u^\mu) +$ $W_{\mu-1}(u, \lambda),$	$0 = m \log y +$ $\int_{x^m}^{y^n} du/u P_\mu(u, \lambda),$ $x > 0, y > 0,$ $\lambda = (\lambda_1, \dots, \lambda_\mu)$	$y P_\mu(y^n, \lambda) dx -$ $(n/m + P_\mu(y^n, \lambda)) \times$ $x P_\mu(x^m, \lambda) dy = 0$
$D_\mu^c$	$\dot{x} = Q_\mu(x, \lambda),$ $\dot{y} = -y,$ $Q_\mu(x, \lambda) = \pm x^{\mu+1}(1 + \lambda_\mu x^\mu) +$ $W_{\mu-1}(x, \lambda),$	$y = C(\lambda)x,$ $C = \int_{-1}^1 du/Q_\mu(u, \lambda),$ $x, y \in \mathbb{R}^1$	$x dy - y dx = 0$
$D_\mu^h$	$\lambda = (\lambda_1, \dots, \lambda_\mu)$	$0 = \log y +$ $\int_x^1 du/Q_\mu(u, \lambda),$ $y > 0, x \in \mathbb{R}^1$	$Q_\mu(x, \lambda) dy -$ $y dx = 0$

In the first column we use the following notation for elementary equilibria (the subscript indicates the degree of degeneracy):

$S_0$  — Nonresonant saddle;

$S_\mu$  — Resonant saddle whose quotient equation (the differential equation for  $u = x^m y^n$  below) has the singular point of multiplicity  $\mu + 1$  at the origin,  $\mu \geq 1$ ; if we want to specify explicitly the resonance between the eigenvalues, we use the extended notation  $S_\mu^{(n:m)}$  assuming that the natural numbers  $m, n$  are mutually prime;

$D_\mu$  — Degenerate saddlenode of multiplicity  $\mu$ ;

$W_{\mu-1}(z, \lambda) = \lambda_0 + \lambda_1 z + \dots + \lambda_{\mu-1} z^{\mu-1}$  is a Weierstrass polynomial of degree  $\mu - 1$ .

For different technical remarks concerning this table see [IY2, §1.1]. We just briefly describe each column.

The second column has the corresponding normal forms. In the third column of the table the Poincaré correspondence maps  $y = \Delta(x, \lambda)$  for the polynomial normal forms are given. They are implicitly defined by the equations relating  $x$  to  $y$ , these equations depending explicitly on the parameters  $\lambda$  and thus implicitly on the original parameters  $\varepsilon$ . The choice of segments transversal to the phase curves of the family is described in Figure 4. The last column has Pfaffian equations vanishing on the graphs of the corresponding Poincaré maps.

A  $C^r$ -smooth conjugacy between two local families  $v$  and  $w$  of the above form is a map

$$H: (\mathbb{R}^2, 0) \times (\mathbb{R}^k, 0) \rightarrow (\mathbb{R}^2, 0), \quad (x, y, \varepsilon) \mapsto H(x, y, \varepsilon),$$

such that

$$H_* v(x, y, \varepsilon) = w(H(x, y, \varepsilon), \varepsilon),$$

where  $H_*$  stands for the Jacobian matrix with respect to the variables  $x, y$ . (this definition does not yet allow for reparametrization of a local family). Two families are finitely differentially equivalent if for any  $r < \infty$  there exists a  $C^r$ -conjugacy between them. The two families  $v, w$  are orbitally equivalent if there exists the

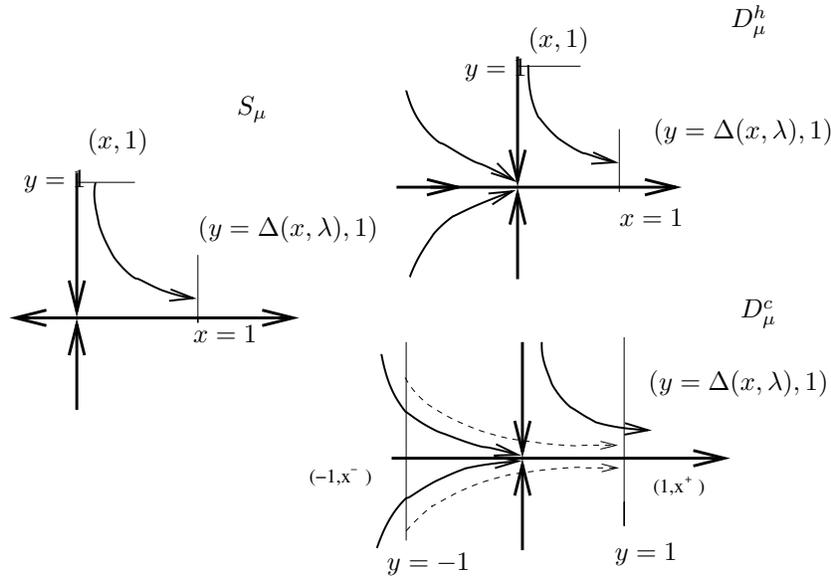


FIGURE 5. Poincaré correspondence maps

germ of a nonvanishing function  $\phi: (\mathbb{R}^2, 0) \times (\mathbb{R}^k, 0) \rightarrow \mathbb{R}^1$  such that  $v$  is equivalent to  $\phi \circ w$ .

To allow for a reparametrization of local families, we say that a family  $v(\cdot, \varepsilon)$  is induced from another family  $w(\cdot, \lambda)$ ,  $\lambda \in (\mathbb{R}^m, 0)$ , if  $v(\cdot, \varepsilon) = w(\cdot, \lambda(\varepsilon))$ , where  $\lambda(\varepsilon)$  is the germ of a smooth map  $(\mathbb{R}^k, 0) \rightarrow (\mathbb{R}^m, 0)$ . The number of new parameters  $m$  may be different from  $k$ .

Assume that the family  $w(\cdot, \lambda)$  is *global* (i.e., the expression  $w(x, y, \lambda)$  makes sense for all  $(x, y, \lambda) \in \mathbb{R}^{m+2}$ ); this happens in particular when  $w$  is *polynomial* in all its arguments. Restricting the parameters  $\lambda$  onto a small neighborhood of a certain point  $(0, 0, c) \in \mathbb{R}^2 \times \mathbb{R}^m$ , we obtain a *localization* of the global family  $w$ , which formally becomes a local family after the parallel translation  $\lambda \mapsto \lambda - c$ .

- DEFINITION 12. (1) A local family  $v = v(\cdot, \lambda)$  is *finitely smooth orbital versal unfolding* (in short, versal unfolding) of the germ  $v(\cdot, 0)$  if any other local family unfolding this germ is finitely differentiable orbitally equivalent to a family induced from  $v$ .
- (2) A polynomial family  $w(\cdot, \lambda)$ ,  $\lambda \in \mathbb{R}^m$ , is a *global finitely smooth orbital versal unfolding* (in short, global versal unfolding) for a certain class of local families of vector fields if any local family from this class is finitely differentiable orbitally equivalent to a local family induced from some localization of  $w$ .

To investigate a versal unfolding means to investigate at the same time all smooth local finite-parameter families which unfold the same germ  $v(\cdot, 0)$ . The main result describing versal unfoldings of germs of elementary singularities on the plane is given by the following

THEOREM 4 ([IY1]). *Suppose that a generic finite-parameter family of smooth vector fields on the plane possesses an elementary singular point for a certain value*

of the parameters. If this point has at least one hyperbolic sector, then the family is finitely differentiable orbitally equivalent to a family induced from some localization of one of the families given in the second column of Table 1.

2.1.3. *Singular-regular systems determining the number of limit cycles.* Recall that for simplicity we consider a 2-parameter family of vector fields  $\{\dot{x} = v(x, \varepsilon)\}_{\varepsilon \in B^2}$  and suppose that for  $\varepsilon = 0$  the vector field  $\dot{x} = v(x, 0)$  has a polycycle  $\gamma$  which consists of two saddles  $p_1, p_2$  and two separatrices connecting them,  $\gamma_1, \gamma_2$  (see Figure 4). For each saddle  $\{p_j\}_{j=1,2}$  there is a neighborhood  $\{U_j\}_{j=1,2}$  with a  $C^r$ -normal coordinate charts. Consider transversal segments “entrance”  $\Sigma_j^-$  and “exit”  $\Sigma_j^+$  which are parallel to coordinate axis of the normal chart such that the phase curve  $\gamma_{j-1}$  enters the neighborhood  $U_j$  through  $\Sigma_j^-$  and the phase curve  $\gamma_j$  exists  $U_j$  through  $\Sigma_j^+$ . The normal coordinates induce coordinates  $x_j$  and  $y_j$  on  $\Sigma_j^-$  and  $\Sigma_j^+$  respectively. For some parameter values the corresponding vector field defines the following collection of Poincaré correspondence maps:

$$(20) \quad \begin{aligned} \Delta_j(\cdot, \varepsilon): x_j &\rightarrow y_j = \Delta_j(x_j, \varepsilon), & j = 1, 2 \\ f_j(\cdot, \varepsilon): y_j &\rightarrow x_{j+1} = f_j(y_j, \varepsilon), & j = 1, 2 \pmod{2}, \end{aligned}$$

where  $\Delta_j(\cdot, \varepsilon)$  is a local Poincaré map from the “entrance” segment  $\Sigma_j^-$  to the “exit” segment  $\Sigma_j^+$  and  $f_j(\cdot, \varepsilon)$  is a semilocal Poincaré map along the phase curve  $\gamma_j$  from the “exit” segment  $\Sigma_j^+$  to the “entrance” segment  $\Sigma_{j+1}^-$ .

Now we decompose the monodromy map (the Poincaré first return map) along the polycycle  $\gamma$  into the chain of the two local singular maps  $\{\Delta_j(\cdot, \varepsilon)\}_{j=1,2}$  and the semilocal regular maps  $\{f_j(\cdot, \varepsilon)\}_{j=1,2}$  of total length 4. Limit cycles correspond to the fixed points of the monodromy. But instead of writing one equation for the fixed points of the monodromy we consider a system of 4 equations, which we call the *preliminary basic system*:

$$(21) \quad \begin{cases} y_1 = \Delta_1(x_1, \varepsilon), \\ x_2 - f_1(y_1, \varepsilon) = 0, \\ y_2 = \Delta_2(x_2, \varepsilon), \\ x_1 - f_2(y_2, \varepsilon) = 0. \end{cases}$$

Recall that  $x_j$ 's are  $C^r$ -normal coordinates on  $\Sigma_j^-$  and  $y_j$ 's are  $C^r$ -normal coordinates on  $\Sigma_j^+$ . Thus the system involves  $C^r$ -smooth regular functions  $f_j$ 's and the maps  $\Delta_j$  from the list (modulo reparametrization  $\varepsilon \rightarrow \lambda(\varepsilon)$ ), that are essentially singular. The problem now is to estimate the number of small isolated solutions uniformly over all sufficiently small parameter values.

Suppose for  $\varepsilon = \varepsilon^*$  the system (21) has the maximal number of isolated solutions. Since each isolated solution of this system corresponds to an isolated solution of the 1-dimensional Poincaré return map  $\Delta(x_1, \varepsilon) = x_1$ , one can choose a small  $\delta_1$  so that the number of regular (nondegenerate) solutions of  $\Delta(x_1, \varepsilon) = x_1 + \delta_1$  bounds the number of isolated solutions to (21) from above (see Figure 6, cf. [IY2, Figure 8].)

Recall that a point  $x \in \mathbb{R}$  is *nondegenerate* or *regular* for the map  $\Delta$  if the derivative  $\Delta'(x) \neq 0$  in the 1-dimensional case and  $x \in \mathbb{R}^n$  is a regular point of a smooth map  $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$  if the rank of the linearization  $dF(x)$  at  $x$  is maximal. Direct calculation shows that a regular solution to  $\Delta(x_1, \varepsilon) = x_1$  (resp.

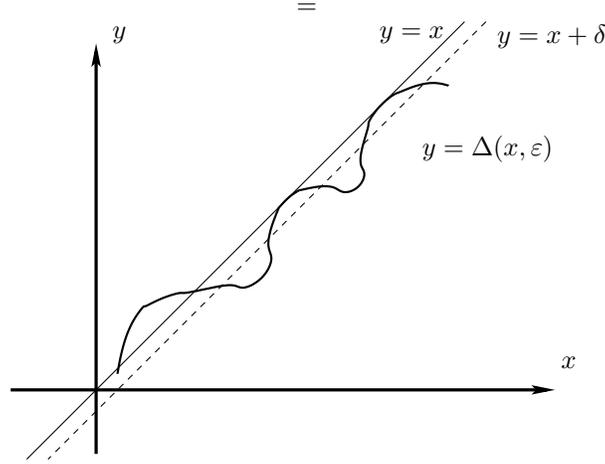


FIGURE 6. Isolated and regular solutions

$\Delta(x_1, \varepsilon) = x_1 + \delta_1$ ) corresponds to a regular solution to the system (21) (see [IY2, Lemma 3.3]).

Moreover, if  $\delta_2$  is nonzero and much smaller than  $\delta_1$ , then by the implicit function theorem the number of regular solutions of the system

$$(22) \quad \begin{cases} y_1 = \Delta_1(x_1, \varepsilon), \\ x_2 - f_1(y_1, \varepsilon) = \delta_1, \\ y_2 = \Delta_2(x_2, \varepsilon), \\ x_1 - f_2(y_2, \varepsilon) = \delta_2 \end{cases}$$

is the same as that for  $\Delta(x_1, \varepsilon) = x_1 + \delta_1$ . Therefore, it suffices to estimate the number of small regular solutions to the system (22) provided that  $1 \gg |\delta_1| \gg |\delta_2| \geq 0$ .

## 2.2. The reduction method of Khovanskiĭ.

2.2.1. *A mixed singular-regular functional system.* The system (22) is not easy to analyze, because it has the singular functions  $\Delta_j$ . The first key idea of the second step of the Main Result [MR, IY2] is to replace these singular equations in (22) by the singular functional-Pfaffian equations which have polynomial differentials of the form (18). As a result we can obtain the *mixed* functional singular-regular system of the following form

$$(23) \quad \begin{cases} \mathcal{F}_1(x_1, y_1, \varepsilon) = 0 \\ F_1(x, y, \varepsilon) = \delta_1, \\ \mathcal{F}_2(x_1, y_1, \varepsilon) = 0, \\ F_2(x, y, \varepsilon) = \delta_2, \end{cases}$$

$$d\mathcal{F}_j(x_j, y_j, \varepsilon) = P_j(x_j, y_j, \varepsilon) dx_j + Q_j(x_j, y_j, \varepsilon) dy_j,$$

$$F_j(x, y, \varepsilon) = x_{j+1} - f_j(y_j, \varepsilon), \quad X = (x_1, y_1, x_2, y_2) \in (\mathbb{R}^4, 0), \quad \varepsilon \in (\mathbb{R}^2, 0),$$

where  $\mathcal{F}_j$  are functions such that their differentials are polynomial 1-forms of the type from column 4 of Table 1 and  $1 \gg |\delta_1| \gg |\delta_2| \geq 0$ . In order to simplify

considerations we replaced the eigenfunctions  $\lambda_j(\varepsilon)$  (see (19)) by  $\varepsilon$ . What we are interested in is the upper estimate for the number of small regular solutions to (23), uniform over all parameters and all sufficiently small values of  $\delta$ 's with  $1 \gg |\delta_1| \gg |\delta_2| \geq 0$ .

2.2.2. *Reduction of the mixed functional system (23) to chain maps  $P \circ F$ .* Let  $F = (F_1, F_2): \mathbb{R}^4 \rightarrow \mathbb{R}^2$  be a smooth map formed by functions  $F_1$  and  $F_2$ . Denote by  $J^m(\mathbb{R}^4, \mathbb{R}^2)$  the space of  $k$ -jets of maps of  $\mathbb{R}^4$  to  $\mathbb{R}^2$ . Fix coordinates in the source  $X = (x_1, \dots, x_4)$  and the target  $(\delta_1, \delta_2)$ . Then the space  $J^m(\mathbb{R}^4, \mathbb{R}^2)$  consists of coordinates in the source, the target, and all partial derivatives of  $F$  of order at most  $k$

$$(24) \quad \left\{ (x_1, \dots, x_4); (F_1(X), F_2(X)); \left( \frac{\partial^\alpha F_i}{\partial^{\alpha_1} x_1 \dots \partial^{\alpha_4} x_4}, \forall i = 1, 2, \alpha_j \geq 0, \text{ such that } \sum_{j=1}^4 \alpha_j \leq m \right) \right\}.$$

We shall call these coordinates on the  $m$ -jet space  $J^m(\mathbb{R}^4, \mathbb{R}^2)$  the *natural coordinates*. With this coordinate system the space of  $m$ -jets has a natural linear structure. We also denote by  $j^m F$  the  $m$ th jet of the map  $F$ . Denote also by  $B_r(0) \subset \mathbb{R}^4$  the  $r$ -ball centered at the origin. We call a polynomial map  $P: \mathbb{R}^N \rightarrow \mathbb{R}^n$  nontrivial if the image  $P(\mathbb{R}^N)$  has nonempty interior in  $\mathbb{R}^n$ .

Our goal now is to realize Step 2 of our program outlined in Section 1.3, i.e., estimate the small number of solutions to (23) via geometric multiplicity of the chain maps (12) or prove the following

**THEOREM 5** (cf. [K6, Theorem 10]). *Suppose that the degrees of polynomial 1-forms from (23) are bounded by some  $d \in \mathbb{Z}_+$ . Then for a sufficiently small  $r > 0$  there exists a set of 3 (= the number of singular equations + 1) explicitly computable nontrivial polynomials  $P^k = (P_1^k, \dots, P_2^k): J^2(\mathbb{R}^4, \mathbb{R}^2) \rightarrow \mathbb{R}^2$ ,  $k = 0, 1, 2$  defined on the space of 2-jets  $J^2(\mathbb{R}^4, \mathbb{R}^2)$  such that for a generic  $C^3$  smooth<sup>3</sup> map  $F: \mathbb{R}^4 \rightarrow \mathbb{R}^2$  the number of regular solutions to the system (23) inside the ball  $B_r(0)$  is bounded by the number of small regular solutions*

$$(25) \quad \#\{X \in B_r(0) : (F_1, F_2)(X) = (\delta_1, \delta_2), (P_1^0, P_2^0) \circ j^2 F(X) = (\delta_3, \delta_4)\} \\ + \frac{1}{2} \sum_{k=1,2} \#\{X \in B_r(0) : (F_1, F_2)(X) = (\delta_1, \delta_2), (P_1^k, P_2^k) \circ j^2 F(X) = (\delta_3, \delta_4)\},$$

where  $1 \gg |\delta_1| \gg \dots \gg |\delta_4| \geq 0$  decrease to zero sufficiently fast. The degrees of the polynomials satisfy the inequalities  $\deg P_i^k \leq 2^i(d+1)$  for all  $k$  and  $i$ .

**REMARK 2.** We cannot find a direct reference in the book of Khovanskiĭ [K2], but this theorem is in the spirit of the results about perturbations discussed in Section 5.2 of this book. In fact this theorem is due to Khovanskiĭ.

2.2.3. *An application of Khovanskiĭ's method to the system (23) or a proof of Theorem 5.* The method is based on the following version of Rolle's lemma

**LEMMA 1.** *Consider  $C^2$  functions  $f: S^1 \rightarrow \mathbb{R}^1$  on the circle and  $g: [0, 1] \rightarrow \mathbb{R}^1$  on the segment, with a finite number of critical points. Then for any  $a \in \mathbb{R}$  and*

<sup>3</sup>Required smoothness 3 = the number of singular equations + 1

any sufficiently small  $\delta > 0$

$$(26) \quad \begin{aligned} \#\{x : f(x) = a\} &\leq \#\{x : f'(x) = \delta\} \\ \#\{x : g(x) = a\} &\leq \#\{x : g'(x) = \delta\} + 1. \end{aligned}$$

PROOF. One proves first the formula for  $\delta = 0$  using the fact that between any two consecutive preimages there is a point where the derivative is zero. Then one uses the nondegeneracy of critical points.  $\square$

Now using this lemma we shall prove Theorem 5.

PROOF OF THEOREM 5. Denote by  $\rho_r(x) = r - \sum_j x_j^2$  the function which measures distance to the boundary of the  $r$ -ball  $B_r(0)$  and vanishes on the boundary  $\partial B_r(0)$ . Recall that  $r$  is sufficiently small.

Denote by  $G_1 : \mathbb{R}^4 \rightarrow \mathbb{R}^3$  the map defined by the coordinate functions  $(\mathcal{F}_2, F_1, F_2)$ . Then the system (23) under investigation becomes the map  $(\mathcal{F}_1, G_1) : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ , given by its coordinate functions. In terms of this map we need to estimate the number of small preimages of points of the form

$$\#\{(\mathcal{F}_1, G_1)^{-1}(0, 0, \delta_1, \delta_2) \cap B_r(0)\},$$

where  $1 \gg |\delta_1| \gg |\delta_2| \geq 0$ .

Let's estimate the number of small preimages of a point

$$(27) \quad \#\{(\mathcal{F}_1, G_1)^{-1}(a_1, a_2, \delta_1, \delta_2) \cap B_r(0)\},$$

where  $1 \gg |\delta_1| \gg |\delta_2| \geq 0$  and  $a_1, a_2$  are arbitrary. Since there is no restriction on  $a_1, a_2$  the number of solutions may only increase.

Step 1. Eliminate one singular equation, say,  $\mathcal{F}_1 = 0$  and replace it by two chain-type equations  $\{P_1^i \circ j^1 F\}_{i=0,1}$  so that for a sufficiently small  $|\delta_2| \gg |\delta_3| \geq 0$  the number of small regular preimages

$$(28) \quad \#\{(G_1, P_1^i \circ j^1 F(X))^{-1}(a_2, \delta_1, \delta_2, \delta_3) \cap B_r(0)\}_{i=0,1}$$

is at least the number of small regular preimages of (27) for any  $a_1$ , i.e.,

$$(29) \quad \begin{aligned} \#\{X \in B_r(0) : (\mathcal{F}_1, G_1)(X) = (a_1, a_2, \delta_1, \delta_2)\} \\ \leq \#\{X \in B_r(0) : (G_1, P_1^0 \circ j^1 F)(X) = (a_2, \delta_1, \delta_2, \delta_3)\} \\ + \frac{1}{2} \#\{X \in B_r(0) : (G_1, \rho_r)(X) = (a_2, \delta_1, \delta_2, \delta_3)\} \end{aligned}$$

Consider a regular value  $(a_2, \delta_1, \delta_2) \in \mathbb{R}^3$  of the map  $G_1$ . By the rank theorem [GG] the level set  $L_{(a_2, \delta_1, \delta_2)} = G_1^{-1}(a_2, \delta_1, \delta_2) \cap B_r(0) \subset B_r(0)$  is a smooth 1-dimensional manifold in the  $r$ -ball. It consists of a finite number of connected parts either compact — topological circles, denoted by  $\{S_i\}_{i \in I(a_2, \delta_1, \delta_2)}$ , or noncompact — curves  $\{L_j\}_{j \in J(a_2, \delta_1, \delta_2)}$  reaching the boundary  $\partial B_r(0)$ . It is easy to see that

$$(30) \quad \begin{aligned} \#\{X \in B_r(0) : (\mathcal{F}_1, G_1)(X) = (a_1, a_2, \delta_1, \delta_2)\} \\ = \sum_{i \in I(a_2, \delta_1, \delta_2)} \#\{X \in S_i : \mathcal{F}_1(X) = a_1\} + \sum_{j \in J(a_2, \delta_1, \delta_2)} \#\{X \in L_j : \mathcal{F}_1(X) = a_1\}. \end{aligned}$$

Let us estimate the first sum on the right-hand side. Fix a circle, say,  $S_1$ . Restrict the function  $\mathcal{F}_1$  to  $S_1$  and denote the result by  $f_1 = \mathcal{F}_1|_{S_1} : S^1 \rightarrow \mathbb{R}$  (see Figure 7). We get a function  $f_1$  on the circle. Notice that the condition  $f_1'(X) = 0$

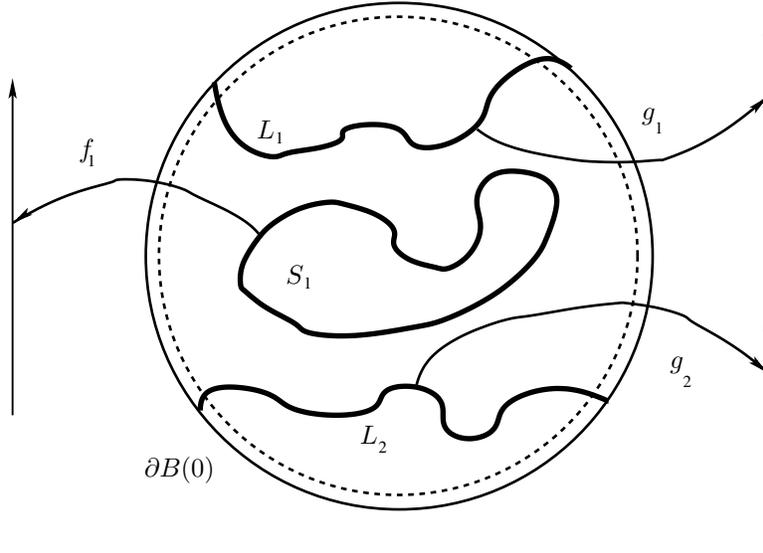


FIGURE 7. Application of Rolle's lemma

is equivalent to the condition that the Jacobian of the map  $(\mathcal{F}_1, G_1)$ , denoted by  $J_{\mathcal{F}_1, G_1}(X)$ , is zero.

$$(31) \quad \boxed{f_1'(X) = 0} \iff \boxed{J_{\mathcal{F}_1, G_1}(X) = 0}$$

Recall now that the differentials  $d\mathcal{F}_j(X) = P_j dx_j + Q_j dy_j$ ,  $j = 1, 2$  are polynomial, therefore, we have

$$(32) \quad J_{\mathcal{F}_1, G_1}(X) = *(d\mathcal{F}_1(X) \wedge d\mathcal{F}_2(X) \wedge dF_1(X) \wedge dF_2(X)) \\ = \det(\nabla\mathcal{F}_1(X), \nabla\mathcal{F}_2(X), \nabla F_1(X), \nabla F_2(X)) = P_1^0 \circ j^1 F(X).$$

where  $*$  is a natural isomorphism between the space of functions on  $\mathbb{R}^4$  and 4-forms on  $\mathbb{R}^4$  and  $\nabla F(X)$  is the gradient vector of a function  $F: \mathbb{R}^4 \rightarrow \mathbb{R}$ . Since  $\deg P_j, \deg Q_j \leq d$ , the degree of  $P_1^0$  is bounded by  $2(d+1)$ .

Now we can apply Rolle's Lemma 1 with  $f = f_1$  and get that for any  $a_1$  and a sufficiently small  $\delta_3 \neq 0$

$$(33) \quad \sum_i \#\{X \in S_i : \mathcal{F}_1(X) = a_1\} \leq \sum_i \#\{X \in S_i : J_{\mathcal{F}_1, G_1}(X) = \delta_3\}.$$

The second sum can be estimated in almost the same way. Instead of using Lemma 1 with  $f = f_i$  we need to use Lemma 1 with  $g = g_j = \mathcal{F}_1|_{L_j}: [0, 1] \rightarrow \mathbb{R}$  (see Figure 7). Denote the number of components reaching the boundary  $|J|$  by  $k$ . Then

$$(34) \quad \sum_{j=1}^k \#\{X \in L_j : \mathcal{F}_1(X) = a_1\} \leq \sum_{j=1}^k \#\{X \in L_j : J_{\mathcal{F}_1, G_1}(X) = \delta_3\} + k$$

In order to find the number of components reaching the boundary notice that each such component intersects the sphere  $\rho_r^{-1}(\delta_3)$  for  $\delta_3 > 0$  in at least two points. So  $P_1^1 \circ j^1 F(X) = \rho_r(X)$  and the second term in inequality (29) corresponds to

the number of noncompact components (the boundary term). This completes the proof of Step 1 or proves (29).

For  $i = 0, 1$  denote by  $G_2^i: \mathbb{R}^4 \rightarrow \mathbb{R}^3$  the maps defined by their coordinate functions  $(F_1, F_2, P_1^i \circ j^1 F)$ . Let's fix  $i = 0$  or 1.

Step 2. Eliminate the second singular equation  $\mathcal{F}_2 = 0$  and replace it by two chain-type equations  $\{P_2^i \circ j^2 F\}_{i=0,1,2}$  so that for a sufficiently small  $0 \leq |\delta_4| \ll |\delta_3|$  the number of small regular preimages  $\{(\mathcal{F}_2, G_2^i)^{-1}(a_2, \delta_1, \delta_2, \delta_3)\}_{i=0,1}$  is at least the number of small regular preimages  $(G_2^i, P_2^i \circ j^2 F)^{-1}(\delta_1, \delta_2, \delta_3, \delta_4)$  for any  $a_2$ , i.e.,

$$(35) \quad \#\{X \in B_r(0) : (\mathcal{F}_2, G_2^i)(X) = (a_2, \delta_1, \delta_2, \delta_3)\} \\ \leq \#\{X \in B_r(0) : (G_2^i, P_2^i \circ j^2 F)(X) = (\delta_1, \delta_2, \delta_3, \delta_4)\} \\ + \frac{1}{2}\#\{X \in B_r(0) : (G_1, \rho_r)(X) = (\delta_1, \delta_2, \delta_3, \delta_4)\}$$

The proof of this inequality is very similar to the proof of Step 1. We reproduce a shortened version of it in order to show why *the condition*  $|\delta_4| \ll |\delta_3|$  *is necessary* for the Khovanskiĭ method to work.

Let's choose a regular value  $(\delta_1, \delta_2, \delta_3)$  for the map  $G_2^i$  and consider the level set  $L_{(\delta_1, \delta_2, \delta_3)} = (G_2^i)^{-1}(\delta_1, \delta_2, \delta_3)$  which by the rank theorem is a smooth 1-dimensional manifold consisting of a finite number of connected components either compact — topological circles, denoted by  $\{S_i\}_{i \in I(\delta_1, \delta_2, \delta_3)}$ , or noncompact — curves  $\{L_j\}_{j \in J(\delta_1, \delta_2, \delta_3)}$  reaching the boundary  $\partial B_r(0)$ .

Then we restrict  $\mathcal{F}_2$  to  $L_{(\delta_1, \delta_2, \delta_3)}$  and get a finite collection of functions  $\{f_i = \mathcal{F}_1|_{S_i}: S^1 \rightarrow \mathbb{R}\}_{i \in I}$  on circles and  $\{g_j = \mathcal{F}_1|_{L_j}: [0, 1] \rightarrow \mathbb{R}\}_{j \in J}$  on the interval  $[0, 1]$ . In order to use Rolle's Lemma 1 we need to compute the condition  $f'_i(X) = 0$  (resp.  $g'_j(X) = 0$ ). This is equivalent to the Jacobian  $J_{\mathcal{F}_2, G_2^i}(X)$  of the map  $(\mathcal{F}_2, G_2^i)$  being equal to 0

$$(36) \quad J_{\mathcal{F}_2, G_2^i}(X) = *(d\mathcal{F}_2(X) \wedge dF_1(X) \wedge dF_2(X) \wedge d(P_1^i \circ j^1 F)(X)) \\ = \det(\nabla \mathcal{F}_2(X), \nabla F_1(X), \nabla F_2(X), \nabla(P_1^i \circ j^1 F)(X)) \\ = P_2^i \circ j^2 F(X).$$

Since  $d\mathcal{F}_2(X) = P_2 dx_2 + Q_2 dy_2$  and  $\deg P_2, \deg Q_2 \leq s$ , the degree of  $P_2^i$  is bounded by  $4(s+1)$ . An easy calculation shows that each time we take a Jacobian of a chain-map  $P \circ j^k F$  its degree at most doubles.

Now we would like to apply Rolle's Lemma 1 with  $f = f_i$  (resp.  $g = g_j$ ) and substitute a singular equation  $\mathcal{F}_2$  by the equation  $J_{\mathcal{F}_2, G_2^i}(X) = \delta_4$ . This equation has to be equivalent to the fact that the derivative  $f'(X)$  (resp.  $g'(X)$ ) is small or the covectors  $\nabla \mathcal{F}_2(X)$ ,  $\nabla F_1(X)$ ,  $\nabla F_2(X)$ , and  $\nabla(P_1^i \circ j^1 F)(X)$  have to be almost linearly dependent. However, the determinant (36) can be almost zero *not because gradient vectors are almost linearly dependent, but because one of the gradient vectors is small*. In order to avoid this problem let's make the following remark: for a fixed regular value  $(\delta_1, \delta_2, \delta_3)$  the level set  $L_{(\delta_1, \delta_2, \delta_3)}$  is a smooth compact 1-dimensional manifold possibly with boundary and the lengths of the gradient vectors  $\nabla F_1(X)$ ,  $\nabla F_2(X)$ , and  $\nabla(P_1^i \circ j^1 F)(X)$  have to be bounded away from zero. Knowing how far these lengths are from zero we can choose  $\delta_4$  of much smaller size to guarantee almost linear dependence of the gradient vectors. This proves inequality (35).

This argument allows one to apply Rolle's Lemma 1 in the described fashion inductively in any dimension and eliminate an arbitrary number of singular Pfaffian equations. This completes the proof of Theorem 5. See [IK, K6] for a more general treatment.  $\square$

2.2.4. *Geometric multiplicity of chain maps.* Let  $P: \mathbb{R}^N \rightarrow \mathbb{R}^n$  be a nontrivial vector-polynomial, i.e., the image  $P(\mathbb{R}^N) \subset \mathbb{R}^n$  has nonempty interior, let  $B^n \subset \mathbb{R}^n$  be a unit ball, and  $F: B^n \rightarrow \mathbb{R}^N$  be a generic sufficiently smooth map with  $N \geq n$ . We call the composition of a generic smooth map and a nontrivial polynomial

$$(37) \quad P \circ F: B^n \rightarrow \mathbb{R}^n$$

a *chain map*. More generally, let  $P: J^m(B^n, \mathbb{R}^N) \rightarrow \mathbb{R}^n$  be a nontrivial vector-polynomial, defined on the space of  $m$ -jets for some  $m \in \mathbb{Z}_+$ . Then a *chain map* is

$$(38) \quad P \circ j^m F: B^n \rightarrow \mathbb{R}^n$$

As a result of the application of Theorem 5 to the system (23) we need to estimate the number of small regular preimages of a special point of a *chain map* or its geometric multiplicity, defined in (8). Actually, application of Theorem 5 gives not a generic smooth map, but a jet of a generic smooth map. To simplify the discussion we consider the case of a smooth map (37). The general jet case can be treated using the same method.

The next two lectures are devoted to a proof of Bézout's theorem for chain maps. Recall that  $B_r(0) \subset \mathbb{R}^n$  denotes the  $r$ -ball centered at the origin.

**THEOREM 6** (cf. [K6, Theorem 3]). *Let  $P = (P_1, \dots, P_n)$  be a nontrivial polynomial defined on the space of  $m$ -jets  $P: J^m(B^n, \mathbb{R}^N) \rightarrow \mathbb{R}^n$  and let  $F: B^n \rightarrow \mathbb{R}^N$  be a  $C^k$  smooth mapping,  $k > m$ , and  $N > n$ . Suppose  $F$  satisfies a transversality condition depending only on  $P$ . Then for a sufficiently small  $r$ , to find a geometric multiplicity of the chain map (37) at 0 one can replace  $j^m F$  by  $L_{F,0,m}$ , its linear part at 0. Namely,*

$$(39) \quad \#\{X \in B_r(0) : P_1 \circ j^m F(X) = \delta_1, \dots, P_n \circ j^m F(X) = \delta_n\} \\ = \#\{X \in B_r(0) : P_1 \circ L_{F,0,m}(X) = \delta_1, \dots, P_n \circ L_{F,0,m}(X) = \delta_n\},$$

where  $1 \gg |\delta_1| \gg \dots \gg |\delta_n| \geq 0$ .

**REMARK 3.** By Bézout's theorem the number of isolated solutions to the equation in the right-hand side of (39) can be bounded by the product  $\prod_{i=1}^n \deg P_i$ .

The classical transversality theorem [AGV] says that for any given transversality condition, a generic mapping  $F: B^n \rightarrow \mathbb{R}^N$  satisfies this condition and generic mappings form an open dense set in the space of smooth mapping of  $B^n$  to  $\mathbb{R}^N$ . Moreover, a mapping  $F$  with "probability one" satisfies a previously given transversality condition. For definitions of "probability one" or "prevalence" see [HSY, K1].

*Acknowledgments.* I would like to thank Askold Khovanskiĭ whose deep insight helped me to make a significant simplification of the application of Khovanskiĭ's method.

### 3. Stratifications and Bézout's theorem for chain maps

In this lecture, first in Section 3.1 we describe a geometric picture behind Bézout's theorem for chain maps (Theorem 6) formulated at the end of the last lecture. It turns out that the proof of this theorem reduces to a question about the existence of a certain, so-called,  $a_P$ -stratification for the outer part  $P$  of the chain map (38). Then in Section 3.2 we define the necessary notions from stratification theory, including  $a_P$ -stratification and discuss the question of existence of an  $a_P$ -stratification. In general, it does not always exist, as examples from Section 3.3.1 of Thom and Grinberg show. At the end of this section we state Hironaka's theorem on the existence of  $a_P$ -stratification for polynomial functions, i.e., for polynomial maps with 1-dimensional image, and its extension, a theorem on the existence of an  $a_P$ -stratification for maps with a multidimensional image proven in [K6]. Such a theorem is required for the proof of the Main Result. Finally, in Section 3.4 we present a geometric proof of Hironaka's theorem based on the author's proof of the existence of Whitney's stratification [K7]. A proof of the existence of Whitney's stratification is also outlined.

**3.1. A heuristic description.** Consider a chain map  $P \circ F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , where  $F: \mathbb{R}^2 \rightarrow \mathbb{R}^N$  is a generic  $C^k$  smooth map,  $N, k > 2$  and  $P = (P_1, P_2): \mathbb{R}^N \rightarrow \mathbb{R}^2$  is a polynomial of degree  $d$ . Fix a small positive  $r$ . We would like to estimate the maximal number of small preimages

$$(40) \quad \#\{x \in B_r(0) : P_1 \circ F(x) = \varepsilon, P_2 \circ F(x) = 0\}$$

for a small enough  $\varepsilon$ .

To show the idea put  $N = 3$ ,  $P_1(x, y, z) = x^2 + y^2$ , and  $P_2(x, y, z) = xy$ . Assume also that  $F(0) = 0$ . Denote the level set by  $V_\varepsilon = \{P_1 = \varepsilon, P_2 = 0\}$ . The level set  $V_\varepsilon$  for  $\varepsilon > 0$  consists of 4 parallel lines (see Figure 8).

Notice that in our notation the number of intersections of  $F(B_r(0))$  with  $V_\varepsilon$  equals the number of preimages of the point  $(\varepsilon, 0)$  under  $P \circ F$ , see (40).

It is easy to see from Figure 8 that if  $F$  is transverse to  $V_0$  it is transverse to  $V_\varepsilon$  for any small  $\varepsilon > 0$ . Moreover, the number of intersections  $F(B_r(0))$  with  $V_\varepsilon$  equals 4 (see the points  $P_1, \dots, P_4$ ).

Another way to calculate the same number is as follows. Let us replace  $F$  by its linear part  $L_F$  at zero. Then

$$\begin{aligned} \#\{x \in B_r(0) : P_1 \circ F(x) = \varepsilon, P_2 \circ F(x) = 0\} \\ = \#\{x \in B_r(0) : P_1 \circ L_F(x) = \varepsilon, P_2 \circ L_F(x) = 0\} \end{aligned}$$

and solving this polynomial system also yields 4.

The idea behind this picture is the following: Consider an arbitrary  $N$  and a polynomial  $P = (P_1, P_2): \mathbb{R}^N \rightarrow \mathbb{R}^2$  of degree at most  $d$  and  $N > 2$ . Define the algebraic variety  $V_\varepsilon = (P_1, P_2)^{-1}(\varepsilon, 0)$  as the level set.

Assume for simplicity that for any small  $\varepsilon \neq 0$  the level set  $V_\varepsilon$  is a manifold of codimension 2. We shall get rid of this assumption later (see Theorem 9(b)). It turns out that there exists a special partition  $\mathcal{V}_0 = \{V_i\}_{i \in \mathcal{A}}$  of  $V_0 = \bigsqcup_{i \in \mathcal{A}} V_i$  into semialgebraic parts which are attached to their neighbors "regularly" see Definition 14) such that it depends only on  $P$  and satisfies the following condition. We say  $F$  is transverse to a stratified set  $(V_0, \mathcal{V}_0)$  if  $F$  is transverse to each stratum

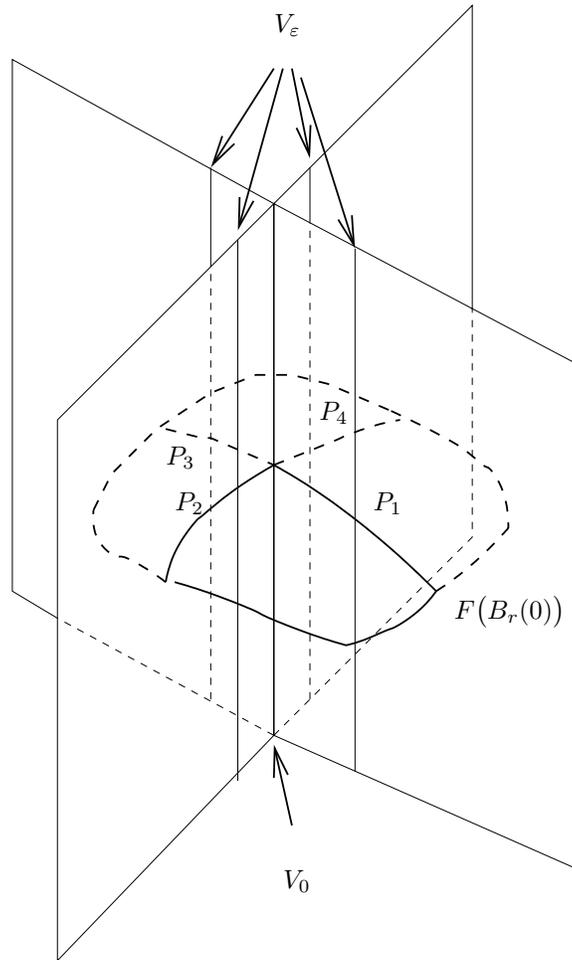


FIGURE 8. The Idealistic Example

$V_i \in \mathcal{V}_0$ ; then

$$(41) \quad \boxed{F \text{ is transverse to } (V_0, \mathcal{V}_0)} \implies \boxed{F \text{ is transverse to } V_\varepsilon}$$

LEMMA 2. Let  $B_r(0)$  be the  $r$ -ball centered at the point  $0 \in \mathbb{R}^2$  and let  $L_{F,0}$  denote the linearization of  $F$  at the point  $a$ . Under condition (41), the number of intersections of the image  $F(B_r(0))$  with  $V_\varepsilon$  coincides with the number of intersections of the image  $L_{F,0}(B_r(0))$  with  $V_\varepsilon$ , provided  $r$  is small enough. That is:

$$(42) \quad \#\{x \in B_r(0) : (P_1, P_2) \circ F(x) = (\varepsilon, 0)\} \\ = \#\{x \in B_r(0) : (P_1, P_2) \circ L_{F,0}(x) = (\varepsilon, 0)\}.$$

REMARK 4. The argument below is independent of the codimension of  $V_\varepsilon$ . We only need condition (41) and the fact that the codimension of  $V_\varepsilon$  coincides with the dimension of the preimage of a chain map  $P \circ F$ .

PROOF. Consider the 1-parameter family of maps  $F_t = tF + (1-t)L_F$  deforming the linear part of  $F$  into  $F$ . Clearly,  $F_1 \equiv F$  and  $F_0 \equiv L_F$ . Fix a small  $r > 0$ . Since  $F$  is transverse to  $V_0$  at 0 all  $F_t$  are transverse to  $V_0$  at 0. Condition (41) implies that for all small  $\varepsilon$  and all  $t \in [0, 1]$  we have  $F_t$  is transverse to  $V_\varepsilon$ .

Therefore, *the number of intersections of  $F_t(B_r(0))$  with  $V_\varepsilon$  is independent of  $t$* . Indeed, assume that  $\#\{F_{t_1}(B_r(0)) \cap V_\varepsilon\} \neq \#\{F_{t_2}(B_r(0)) \cap V_\varepsilon\}$  for some  $t_1 < t_2$ . Then as  $t_1$  increases to  $t_2$  there is a point  $t^*$  where the number of intersections either drops or jumps. At this point  $t^*$  the condition of transversality of  $F_{t^*}$  and  $V_\varepsilon$  must fail. This completes the proof of the lemma.  $\square$

### 3.2. Basic definitions of stratified sets, maps, etc.

3.2.1. *Stratified sets.* A stratification of a set, e.g., an analytic variety, is, roughly, a partition of it into manifolds so that these manifolds fit together “regularly.” Stratification theory was originated by Thom [T2] and Whitney [W2] for algebraic and analytic sets. It was one of the key ingredients in Mather’s proof of the topological stability theorem [M2]. For the history and further applications of stratification theory see [GM, dPW].

We consider here only the category of real (semi)algebraic sets for simplicity. Theorems on existence of stratifications proven here in the category of semialgebraic sets can be proven for the categories of complex or real (semi)analytic sets using similar methods. Call a subset  $V \subset \mathbb{R}^N$  a *semivariety* if locally at each point  $x \in \mathbb{R}^N$  it is a finite union of subsets defined by equations and inequalities

$$(43) \quad f_1 = \cdots = f_k = 0 \quad g_1 > 0, \dots, g_l > 0$$

where  $f_i$ ’s and  $g_j$ ’s are real algebraic depending on  $x$ . Semivarieties are closed under Boolean operations.

DEFINITION 13 (Whitney). Let  $V_i, V_j$  be disjoint manifolds in  $\mathbb{R}^N$ ,  $\dim V_j > \dim V_i$ , and let  $x \in V_i \cap \overline{V_j}$ . A triple  $(V_j, V_i, x)$  is called a (resp. b)-regular if

- (a) when a sequence  $\{y_n\} \subset V_j$  tends to  $x$  and  $T_{y_n} V_j$  tends in the Grassmannian bundle to a subspace  $\tau_x$  of  $\mathbb{R}^N$ , then  $T_x V_i \subset \tau_x$ ;
- (b) when sequences  $\{y_n\} \subset V_j$  and  $\{x_n\} \subset V_i$  each tends to  $x$ , the unit vector  $(x_n - y_n)/|x_n - y_n|$  tends to a vector  $v$ , and  $T_{y_n} V_j$  tends to  $\tau_x$ , then  $v \in \tau_x$ .<sup>4</sup>

$V_j$  is called a (resp. b)-regular over  $V_i$  if each triple  $(V_j, V_i, x)$  is a (resp. b)-regular.

REMARK 5. Since the Grassmannian manifold of  $\dim V_j$ -planes in  $m$ -dimensional space is compact, existence of limits in the definition above can be reached by choosing a subsequence  $\{x_{n_k}\}_{k \in \mathbb{Z}_+}$  or  $\{y_{n_k}\}_{k \in \mathbb{Z}_+}$  if necessary.

DEFINITION 14 (Whitney). Let  $V$  be a semivariety in  $\mathbb{R}^N$ . A disjoint decomposition

$$(44) \quad V = \bigsqcup_{i \in I} V_i, \quad V_i \cup V_j = \emptyset \text{ for } i \neq j$$

into smooth semivarieties  $\mathcal{V} = \{V_i\}_{i \in I}$ , called strata, is called an a (resp. b)-regular stratification if

- (1) each point has a neighborhood intersecting only finitely many strata;

<sup>4</sup>This way of defining b-regularity is due to Mather [M2]. Whitney’s definition [W2] is equivalent to this one provided a-regularity holds.

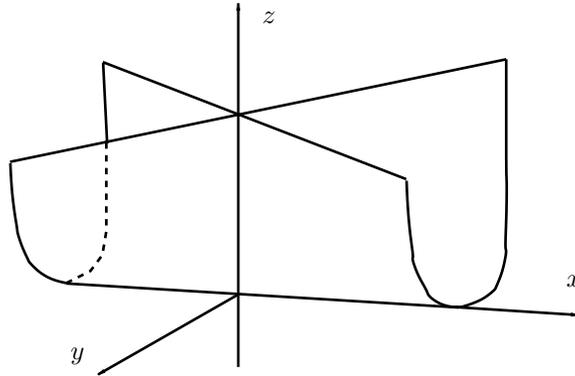


FIGURE 9. Whitney's umbrella

- (2) the frontier  $\overline{V_j} \setminus V_j$  of each stratum  $V_j$  is a union of other strata  $\bigsqcup_{i \in J(i)} V_i$ ;
- (3) any triple  $(V_j, V_i, x)$  such that  $x \in V_i \subset \overline{V_j}$  is a (resp. b)-regular.

The classical example of a stratified algebraic set in  $\mathbb{R}^3$  is the so-called *Whitney umbrella*. It is defined as follows

EXAMPLE 2. Consider the 2-dimensional algebraic variety in  $\mathbb{R}^3$ , defined by

$$(45) \quad V = \{(x, y, z) \in \mathbb{R}^3 : y^2 = zx^2\}.$$

The first natural partition of  $V$  into smooth parts (strata) is the vertical line  $V_1 = \{x = y = 0\}$  and the complement  $V_2 = V \setminus V_1$ . However,  $V_2$  does not fit regularly to  $V_1$  at the origin. To see that consider the sequence of the form  $(x_n, 0, 0) \in V_2$  with  $x_n \rightarrow 0$  as  $n \rightarrow \infty$ . It is easy to see that after we refine  $V_1$  into  $V_{0'} = \{0\}$  and  $V_{1'} = V_1 \setminus V_{0'}$  and put  $V_2 = V_{2'}$  the partition  $V = \bigsqcup_{i'=1,2,3} V_{i'}$  becomes the stratified manifold  $(V, \{V_{0'}, V_{1'}, V_{2'}\})$ .

THEOREM 7 ([W2, T3, L1]). *For any semivariety  $V$  in  $\mathbb{R}^N$  there is an a (resp. b)-regular stratification.*

REMARK 6. This theorem is not true for smooth sets. To see that one can construct a 2-surface in  $\mathbb{R}^3$  which looks like a corkscrew.

Existence of stratifications in the complex analytic case was proved by Whitney [W2]. Later Thom published a sketch of a proof [T3]. Then Łojasiewicz [L1] extended these results to the semianalytic case. Later other proofs were found. In [H] Hironaka found a nice proof using his resolution of singularities. J. Bochnak, M. Coste, and M.-F. Roy [BCR] and Z. Denkowska, K. Wachta [DW] follow the classical route of Whitney [W2] via the wing lemma. [BCR] uses a Nash wing lemma and [DW] apply the parameterized Puiseux theorem of W. Pawlucki [P1]. T. Wall [W1] and S. Łojasiewicz, J. Stasica, and K. Wachta [LSW] found proofs which use Milnor's curve selection lemma [M3]. The latter proof also uses the subanalyticity of the tangent map (for which an elementary proof was given by Z. Denkowska and K. Wachta [DW]). In [K7] the author gives a geometric proof based on the simple observation that regularity of stratifications is related to uniqueness of the limit of the tangent planes to a bigger stratum as they approach a smaller stratum. This proof is outlined in Section 3.4.1. For a nice exposition of the theory of semianalytic and subanalytic sets see [L2].

3.2.2. *Stratified maps and  $a_P$ -stratification.* First we define a smooth map of a stratified set  $(V, \mathcal{V})$ :

DEFINITION 15. Let  $(V, \mathcal{V})$  be a stratified set in  $\mathbb{R}^N$ ,  $V \subseteq \mathbb{R}^N$ ; then a map  $P: V \rightarrow \mathbb{R}^n$  is called  $C^2$ -smooth if it can be extended to a  $C^2$  smooth map of an open set  $U$ ,  $V \subset U \subset \mathbb{R}^N$ , denoted by  $\mathbf{P}: U \rightarrow \mathbb{R}^n$ , whose restriction to  $V$  coincides with  $P$ .

A stratification  $V = \bigcup_i V_i$  stratifies a smooth map  $P: V \rightarrow \mathbb{R}^n$  if the restriction of  $P$  to a stratum  $V_i$  has constant rank or rank  $dP|_{V_i}(x)$  is independent of  $x \in V_i$ .

A map  $G: \mathbb{R}^n \rightarrow \mathbb{R}^N$  is called transverse to a stratified set  $(V, \mathcal{V})$  if  $G$  is transverse to each stratum  $V_i \in \mathcal{V}$ .

EXAMPLE 3. With the notations of the Example 2 of the Whitney umbrella consider the Whitney umbrella  $V$  and the projection  $P = \pi|_V: V \rightarrow \mathbb{R}^2$  along the  $z$ -coordinate restricted to it. Then the stratification  $(V, \{V_0', V_1', V_2'\})$  of  $V$  is also an  $a_P$ -stratification of  $P: V \rightarrow \mathbb{R}^2$ , a term to be defined below.

Let  $V_i$  be a stratum of a stratification  $(V, \mathcal{V})$ ,  $\mathcal{V} = \{V_i\}_{i \in \mathcal{A}}$  and  $a \in N$ . Denote by  $L_{a,i} = (P^{-1}(a) \cap V_i)$  the level sets of  $P$  in  $V_i$ . By the rank theorem [GG], if a stratification  $(V, \mathcal{V})$  stratifies a smooth map  $P$ , then for each stratum  $V_i$  the number  $d_i(P) = \dim V_i - \text{rank } dP|_{V_i}$  is well defined and equals the dimension of any nonempty level set  $L_{a,i}$ .

Roughly speaking, an  $a_P$ -stratification is a stratification of a map  $P: V \rightarrow N$  such that it is also an  $a$ -stratification of its level sets, i.e., for any sequence of points  $\{b_k\} \subset P(V_j)$  converging to a point  $a \in P(V_i)$  the corresponding level sets  $L_{b_k,j} = (P^{-1}(b_k) \cap V_j) \subset V_j$  approach the limiting level set  $L_{a,i} \subset V_i$  “regularly.” A precise definition is the following

DEFINITION 16. Let  $P: \mathbb{R}^N \rightarrow \mathbb{R}^n$  be a  $C^2$  smooth map and let  $V_j$  and  $V_i$  be submanifolds of  $\mathbb{R}^N$  such that  $V_i \subset \overline{V_j} \setminus V_j$  and the restrictions  $P|_{V_j}$  to  $V_j$  and  $P|_{V_i}$  to  $V_i$  have constant ranks.  $V_j$  is called  $a_P$ -regular over  $V_i$  with respect to the map  $P$  at a point  $x \in V_i \cap \overline{V_j}$  if for any sequence  $\{x_n\} \subset V_j$  converging to  $x$  the sequence of tangent planes to the level sets  $T_n = \ker dP|_{V_j}(x_n)$  converges in the corresponding Grassmannian manifold of  $\dim \ker dP|_{V_j}$ -dimensional planes to a plane  $\tau$  and

$$(46) \quad \lim \ker dP|_{V_j}(x_n) = \tau \supseteq \ker dP|_{V_i}(x)$$

DEFINITION 17 (Thom). A  $C^2$  smooth map  $P: V \rightarrow \mathbb{R}^n$  of a stratifiable set  $V$  to  $\mathbb{R}^n$  is called  $a_P$ -stratifiable if there exists a stratification  $(V, \mathcal{V})$  such that the following conditions hold:

- (a)  $(V, \mathcal{V})$  stratifies the map  $P$  (see Definition 15);
- (b) for all pairs  $V_j$  and  $V_i$  from  $\mathcal{V}$  such that  $V_i \subseteq \overline{V_j} \setminus V_j$  the stratum  $V_j$  is  $a_P$ -regular over the stratum  $V_i$  with respect to  $P$  at the point  $x$  for all  $x \in V_i \cap \overline{V_j}$ .

REMARKS. (1) The original definition of  $a_P$ -stratification requires an appropriate stratification of the image too [M2], but for simplicity we do not require the stratification of the image and this weaker definition turns out to be sufficient for our purposes.

(2) With the notations above for an  $a_P$ -stratification to exist we must have  $d_i(P) \leq d_j(P)$  for each  $V_i \subseteq \overline{V_j} \setminus V_j$ , i.e., nonempty level sets  $L_{b,j}$  inside the bigger stratum  $V_j$  have dimension  $d_j(P)$  greater or equal to the dimension  $d_i(P)$  of

nonempty level sets  $L_{a,i}$  in the smaller stratum  $V_i$ . Otherwise  $\dim \ker dP|_{V_j}(x_n) < \dim \ker dP|_{V_i}(x)$  and condition (46) can't be satisfied.

3.2.3. *For  $a_P$ -stratifications condition (41) holds.* The heuristic description given above shows that the key to a proof of Bézout's theorem is condition (41) (Lemma 2). Now we prove that existence of an  $a_P$ -stratification of a polynomial map  $P$  is sufficient for condition (41) to hold.

Let  $P = (P_1, P_2): \mathbb{R}^N \rightarrow \mathbb{R}^2$  be a nontrivial vector-polynomial, i.e., the image  $P(\mathbb{R}^N)$  has nonempty interior. Denote by  $V = P_2^{-1}(0)$  and  $V_0 = (P_1, P_2)^{-1}(0)$  these level sets. Assume that there exists a stratification  $(V, \mathcal{V})$  which stratifies the map  $P|_V$  and such that the zero level set  $V_0$  is also stratified by a stratification  $(V_0, \mathcal{V}_0)$  with  $V_0 = \bigsqcup_{i \in \mathcal{A}_0} V_i$

LEMMA 3. *With the above notation, if each stratum  $V_j \in \mathcal{V} \setminus \mathcal{V}_0$  is  $a_P$ -regular over each stratum  $V_i \in \mathcal{V}_0$  with respect to the polynomial  $P$ , then any  $C^2$  smooth map  $F: \mathbb{R}^2 \rightarrow \mathbb{R}^N$  which is transverse to  $(V_0, \mathcal{V}_0)$  is also transverse to each level set  $\mathcal{V}_{b,j}$  for any small  $b$  and this is equivalent to condition (41).*

PROOF. Pick a point  $x$  in  $V_i \subset V_0$  and a point  $y \in V_j$ . Notice that  $\ker dP|_{V_i}(x)$  is the tangent plane to the level set  $\{P^{-1}(P(x)) \cap V_i\}$  at the point  $x$  and  $\ker dP|_{V_j}(y)$  is the tangent plane to the level set  $\{P^{-1}(P(y)) \cap V_j\}$ .

By condition (46), if a map  $F: X \rightarrow \mathbb{R}^N$  is transverse to  $\ker dP|_{V_i}(x)$  at a point  $x$ , then  $F$  is transverse to  $\ker dP|_{V_j}(y)$  for any  $y \in V_j$  nearby  $x$ . This completes the proof of the lemma.  $\square$

### 3.3. Existence of $a_P$ -stratifications for polynomial maps.

3.3.1. *Examples of nonexistence due to Thom and Grinberg.* Existence of  $a_P$ -stratifications is a nontrivial question. There are some obvious obstacles. For example, let  $V \subset \mathbb{R}^N$  be an algebraic variety and let  $P: \mathbb{R}^N \rightarrow \mathbb{R}^n$  be a polynomial map. Assume that  $(V, \mathcal{V})$  stratifies  $P$ . Take two strata  $V_i$  and  $V_j$  so that  $V_j$  lies "over"  $V_i$ , i.e.,  $V_i \subseteq \overline{V_j} \setminus V_j$ ; then the condition (46) can't be satisfied if the dimension of the level sets  $d_i(P)$  in the lower stratum  $V_i$  is strictly less than that of  $d_j(P)$  in the lower stratum  $V_j$ , i.e.,  $\dim \ker dP|_{V_i}(y) < \dim \ker dP|_{V_j}$ . In this case a plane  $\ker dP|_{V_j}(x)$  of the upper stratum  $V_j$  should belong to a plane  $\tau$  of smaller dimension by condition (46), which is impossible. Thom constructed the first example for which this happens [GWdPL].

EXAMPLE 4 (Thom). Consider the vector-polynomial  $P$  of the form

$$(47) \quad P: \begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} x \\ xy \end{pmatrix}.$$

The line  $\{x = 0\}$  is the line of critical points of  $P$ . Outside of the line  $\{x = 0\}$  the map  $P$  is a diffeomorphism; therefore, the preimage  $P^{-1}(a)$  of any point  $a \neq 0$  is 0-dimensional. On the other hand, the preimage of 0 is the 1-dimensional line  $\{x = 0\}$ . Thus,  $a_P$ -regularity fails to exist.

DEFINITION 18. Let us call an algebraic set  $V$  rank compatible with respect to a polynomial  $P$  if there exists a stratification  $(V, \mathcal{V})$  which stratifies  $P$  and for any pair  $V_i$  and  $V_j$  from  $\mathcal{V}$  such that  $V_i \subseteq \overline{V_j} \setminus V_j$ , the dimension of the level sets  $d_i(P)$  in the lower stratum  $V_i$  does not exceed the dimension of the level sets  $d_j(P)$  in the upper stratum  $V_j$ .

It turns out that even if an algebraic set  $V$  is rank compatible with respect to a polynomial  $P$ , then an  $a_P$ -stratification still does not always exist. Let us present an example with this property due to M. Grinberg. It seems that the existence of a counterexample was known before, but we did not find an appropriate reference.

Let  $V = \{(x, y, z, t) \in \mathbb{R}^4 : x^2 = t^2y + z\}$  be the three-dimensional algebraic variety and  $P : V \rightarrow \mathbb{R}^2$  be the natural projection associated to the last two coordinates, i.e.,  $P : (x, y, z, t) \rightarrow (z, t)$ .

LEMMA 4. *With the above notations the set  $V$  is rank compatible with the polynomial map  $P$  and does not have an  $a_P$ -stratification.*

PROOF. Consider a rank stratification of  $V$ . Such a stratification consists of three strata:  $V_1 = \{x = t = z = 0\}$ ,  $V_2 = \{t = 0, x^2 = z, x \neq 0\}$ , and  $V_3 = \{t \neq 0\}$ . On each stratum  $\text{rank } P|_{V_i} = i - 1$ . The level sets  $P^{-1}(t, z)$  are parabolas for  $t \neq 0$  and lines for  $t = 0$ .

Show that for each point  $\mathbf{a} = (0, a, 0, 0) \in V_1$  there exists a 1-parameter family of level sets such that at the point  $\mathbf{a}$  the property of  $a_P$ -regularity of  $V_3$  over  $V_1$  fails.

Consider the preimage of the curve  $\{z = -at^2\} \subset \mathbb{R}^2$ . This is an algebraic variety of the form  $W_a = \{x^2 = t^2(y - a)\}$ . One can see that  $W_a$  is the Whitney umbrella (see Figure 9). The level set  $x^2 = t_0^2(y - a)$  is a parabola. As  $t_0 \rightarrow 0$  this parabola tends to the semiline  $\{x = t = z = 0, y \geq a\}$ . At the point  $\mathbf{a} \in V_1$  the property of  $a_P$ -regularity of  $V_3$  over  $V_1$  clearly fails. This completes the proof of the lemma.  $\square$

### 3.3.2. Existence of $a_P$ -stratifications (Hironaka's theorem and its extension).

As we have seen above, sometimes  $a_P$ -stratifications exist, sometimes they do not. Let us state a positive result on their existence.

THEOREM 8 (Hironaka [H]). *If  $V \subset \mathbb{R}^N$  is a semialgebraic variety and  $P : \mathbb{R}^N \rightarrow \mathbb{R}$  is a polynomial function, then there exists an  $a_P$ -stratification  $(V, \mathcal{V})$  of  $V$  with respect to  $P$  by semialgebraic strata.*

In the next section we give a geometric proof of this result based on the proof of the existence of Whitney's stratifications due to the author [K7]. Below we describe an extension of Hironaka's theorem to maps with a multidimensional image proven in [K6]. This extension is sufficient to prove Bézout's theorem for chain maps (Theorem 6).

TARSKI – SEIDENBERG PRINCIPLE (see, e.g., [BCR, J]). *For any semialgebraic  $V$  in  $\mathbb{R}^N$  and a polynomial map  $P : \mathbb{R}^N \rightarrow \mathbb{R}^n$  the image  $P(V)$  is semialgebraic.*

Let  $\mathbb{R}^N$  and  $\mathbb{R}^n$  be Euclidean spaces with the fixed coordinate systems  $x = (x_1, \dots, x_N) \in \mathbb{R}^N$  and  $a = (a_1, \dots, a_n) \in \mathbb{R}^n$  with  $N \geq n$  and a nontrivial vector-polynomial  $P : \mathbb{R}^N \rightarrow \mathbb{R}^n$ . Recall that  $P$  is nontrivial if the image  $P(\mathbb{R}^N)$  has nonempty interior. For brevity in what follows we use *polynomial* for *vector-polynomial*.

DEFINITION 19. Let  $m \in \mathbb{Z}_+$  and  $\delta > 0$ . We call the  $(m, \delta)$ -cone  $K_{m, \delta}^n$  the following set of points:

$$(48) \quad K_{m, \delta}^n = \{a = (a_1, \dots, a_N) \in \mathbb{R}^N : 0 < |a_1| < \delta, 0 < |a_{j+1}| < |a_1 \cdots a_j|^m \text{ for } j = 1, \dots, N - 1\}.$$

Let  $m' \in \mathbb{Z}_+^N$ . If  $m' \geq m$  and  $\delta' \leq \delta$ , then we say that the  $(m', \delta')$ -cone  $K_{m', \delta'}^n$  is a refinement of the  $(m, \delta)$ -cone  $K_{m, \delta}^n$ .

Define the following sets:

$$(49) \quad V_{m, \delta, P} = \text{closure}\{P^{-1}(K_{m, \delta}^n)\}, \quad V_{0, m, P} = \bigcap_{\delta > 0} V_{m, \delta, P}$$

Then one has

**THEOREM 9 ([K6]).** *For any nontrivial polynomial  $P: \mathbb{R}^N \rightarrow \mathbb{R}^n$  there exist an integer  $m \in \mathbb{Z}_+$  and a positive  $\delta$  such that the following conditions hold*

- (a) *the set  $V_0 = V_{0, m, P}$  (see (49)) is a semialgebraic set of codimension at least  $n$ .*
- (b) *the set  $V_{m, \delta, P}$  consists of regular points of  $P$ , i.e., if  $b \in V_{m, \delta, P}$ , then the level set  $P^{-1}(P(b))$  is a manifold of codimension  $n$ .*
- (c) *there exists a stratification of  $V_0$  by semialgebraic strata  $(V_0, \mathcal{V}_0)$  satisfying the property:  $V_{m, \delta, P}$  is  $a_P$ -regular over any strata  $V_i \in \mathcal{V}_0$  with respect to  $P$ .*

**REMARK 7.** In order to have the compatibility condition for the limiting set  $V_{m, \delta, P}$  with the regular level sets  $P^{-1}(a)$  in the definition of the  $(m, \delta)$ -cone  $K_{m, \delta}^n$  (48) it is necessary that the range of values (“smallness”) of  $a_{j+1}$  depends on all  $a_i$ 's with  $i = 1, \dots, j$ . Indeed, consider the following

**EXAMPLE 5.** Let  $x = (x_1, x_2, x_3)$  denote a point in  $\mathbb{R}^3$  and  $P = (P_1, P_2, P_3): \mathbb{R}^3 \rightarrow \mathbb{R}^3$  be a polynomial map, given by

$$(50) \quad P_1(x) = x_1, \quad P_2(x) = x_1x_2, \quad P_3(x) = x_1x_2x_3.$$

If the definition of the  $(m, \delta)$ -cone is

$$(51) \quad K_{m, \delta}^3 = \{a = (a_1, a_2, a_3) \in \mathbb{R}^3 : 0 < |a_1| < \delta, 0 < |a_{2(\text{resp. } 3)}| < |a_1|^m\},$$

then the limiting set  $V_{m, \delta, P}$ , defined by (49), is 1-dimensional for any positive  $m$ . However, all level sets  $P^{-1}(a)$  with  $a_1a_2 \neq 0$  are 0-dimensional. In this case the compatibility condition in Definition 18 fails.

**3.4. A proof of Hironaka's theorem on existence of  $a_P$ -stratifications for polynomial functions.** In this section we present a geometric proof of Hironaka's theorem based on a proof of Whitney's Theorem 7 on existence of  $a$ -stratifications due to the author [K7]. First, we briefly outline the latter proof and then prove Hironaka's theorem following the same path.

**3.4.1. An outline of a proof of Whitney's Theorem 7 on existence of  $a$ -stratifications.** The outline given below works to prove  $b$ -stratifications too after a slight modification [K7].

A semivariety  $V$  has well-defined dimension, say  $d \leq N$ . Denote by  $V_{\text{reg}}$  the set of points where  $V$  is locally a real algebraic submanifold of  $\mathbb{R}^N$  of dimension  $d$ .  $V_{\text{reg}}$  is a semivariety, moreover,  $V_{\text{sing}} = V \setminus V_{\text{reg}}$  is a semivariety of positive codimension in  $V$ , i.e.,  $\dim V_{\text{sing}} < \dim V$ . (For the algebraic case see, e.g., [M3].)

Step 1. There is a filtration of  $V$  by semivarieties

$$(52) \quad V^0 \subset V^1 \subset \dots \subset V^d = V,$$

where for each  $k = 1, \dots, d$  the set  $V^k \setminus V^{k-1}$  is a manifold of dimension  $k$ . This is not difficult to see, e.g., [M3]. Indeed, consider  $V_{\text{sing}} \subset V$ , then  $V \setminus V_{\text{sing}}$  is a

manifold of dimension  $d$  and  $\dim V_{\text{sing}} < d$ . Inductive application of these arguments completes the proof.

A *refinement* of a decomposition  $V = \bigsqcup_{i \in I} V_i$  is a decomposition  $V = \bigsqcup_{i' \in I'} V_{i'}$  such that any stratum  $V_j$  of the first decomposition is a union of some strata of the second one, i.e., there is a set  $I'(j) \subset I'$  such that  $V_j = \bigsqcup_{i' \in I'(j)} V_{i'}$ .

Step 2. Let  $V \subset \mathbb{R}^N$  be a manifold and  $W \subset V$  be a semivariety. Denote by  $\text{Int}_V(W)$  the set of interior points of  $W$  in  $V$  w.r.t. the topology induced from  $\mathbb{R}^N$ . Let  $V_i$  and  $V_j$  be a pair of distinct strata. For each point  $x \in V_i \cap \overline{V_j}$  denote by  $V_j^{\text{con},x}$  the local connected component of  $V_j$  at  $x$ , i.e., the connected component of the intersection of  $V_j$  with a small ball centered at  $x$  and call it *essential* if the closure of  $V_j^{\text{con},x}$  has  $x$  in the interior,  $x \in \text{Int}_{V_i}(V_i \cap \overline{V_j^{\text{con},x}})$ . Denote by  $V_j^{\text{ess},x}$  the union of all local essential components of  $V_j$ . A semialgebraic set  $V_j$  can have only finitely many local connected components (see, e.g., [M3]).

**THEOREM 10.** *For any two disjoint strata  $V_j$  and  $V_i$  the set of points*

$$\text{Sing}_a(V_j, V_i) = \{x \in V_i \cap \overline{V_j} : (V_j^{\text{ess},x}, V_i, x) \text{ is not a-regular}\},$$

*is a semivariety in  $V_i$  and  $\dim \text{Sing}_a(V_j, V_i) < \dim V_i$ .*

Let us show that this theorem is sufficient to prove the a-regular case of Theorem 7. Consider a decomposition  $V = \bigsqcup_{i \in I} V_i$  and split the strata into two groups: the first group consists of strata of dimension at least  $k$  and the second group consists of the rest. Suppose that each stratum from the first group is a-regular over each stratum from the second group. Then by definition of a-regularity any refinement of a stratum from the second group preserves this a-regularity.

Now apply this refinement inductively. Consider strata in  $V^d \setminus V^{d-1}$  of dimension  $d$ . Using Theorem 10 and the result of Łojasiewicz [Ł1] that the frontier of a semivariety has dimension less than the semivariety itself, refine  $V^{d-1}$  so that each  $d$ -dimensional stratum is a-regular over each stratum in  $V^{d-1}$ . The above remark shows that any further refinement of the strata in  $V^{d-1}$  preserves the a-regularity of strata from  $V^d \setminus V^{d-1}$  over it. This reduces the problem of existence of stratification for  $d$ -dimensional semivarieties to the same problem for  $(d-1)$ -dimensional semivarieties. Induction on dimension completes the proof of Theorem 7.

Our proof is based on the observation that if  $V_i \subset \overline{V_j}$  are a pair of strata, a-regularity of  $V_j$  over  $V_i$  at  $x$  in  $V_i$  is closely related to whether the limit of tangent planes  $T_y V_j$  is unique or not as  $y$  from  $V_j$  tends to  $x$ . The rest of the paper is devoted to the proof of Theorem 10, which consists of two steps. In Lemma 5 we relate a-regularity with (non)uniqueness of limits of tangent planes  $T_y V_j$ ; then, based on it and Rolle's lemma in Lemma 6, one can prove Theorem 10.

Let  $V_i$  and  $V_j$  be a pair of distinct strata in  $\mathbb{R}^N$ . Define

$$(53) \quad \text{Un}(V_j, V_i) = \{x \in V_i \cap \overline{V_j} : \text{for any } V_j^{\text{con},x}, \text{ there exists } \tau_x \subset T_x \mathbb{R}^N \\ \text{such that for any } \{y_n\} \subset V_j^{\text{con},x} \text{ tending to } x, T_{y_n} V_j \rightarrow \tau_x\},$$

The proof consists of two lemmas.

**LEMMA 5.** *With the above notations we have*

$$(54) \quad \text{Sing}_a(V_j, V_i) \subset V_i \setminus \text{Un}(V_j, V_i),$$

*where  $\text{Sing}_a(V_j, V_i)$  is the set of a-singular points in  $V_i$  defined in Theorem 10.*

LEMMA 6. *With the above notations there is a set of strata  $\{V_j^p\}_{p \in \mathbb{Z}}$  (resp.  $\{V_i^p\}_{p \in \mathbb{Z}}$ ) in  $V_j$  (resp. in  $V_i$ ), each of positive codimension in  $V_j$  (resp. in  $V_i$ ), such that*

$$(55) \quad \begin{aligned} \text{Sing}_a(V_j, V_i) &\subset \bigcup_{p \in \mathbb{Z}} \text{Sing}_a(V_j^p, V_i) \bigcup_{p \in \mathbb{Z}} V_i^p. \\ V_i \setminus \text{Un}(V_j, V_i) &\subset \bigcup_{p \in \mathbb{Z}} V_i \setminus \text{Un}(V_j^p, V_i) \end{aligned}$$

REMARKS. (1) Inductive application of this lemma to the right-hand side of the first line of (55) reduces dimensions of  $V_j^p$ 's up to  $\dim V_i$ .

(2) The dimension of the frontier of a semivariety  $(\text{Sing}_a(V_j^p, V_i) \subset V_i \cap \overline{V_j^p})$  has dimension strictly smaller than that of the semivariety  $(V_j^p)$  itself.

(3) By Lemma 5 the set  $\text{Sing}_a(V_j, V_i)$  is a semivariety. Since a countable union of semivarieties of positive codimension in  $V_i$  contains  $\text{Sing}_a(V_j, V_i)$  we have  $\text{Sing}_a(V_j, V_i)$  has positive codimension in  $V_i$ , which proves Theorem 10.

Since this proves Theorem 10, as a consequence this proves Theorem 7 too. We are not going to prove these lemmas; however, we would like to exhibit the geometric idea behind the proof of Lemma 6. The section below is devoted to the idea of construction of proper subvarieties in the bigger stratum  $V_j$  approaching the non-a-regular points  $\text{Sing}_a(V_j, V_i) \subset V_i$ .

3.4.2. *Separation of Planes and dimension reduction in Lemma 6.* Let  $\tau_0$  and  $\tau_1$  be two distinct oriented  $k$ -dimensional planes in  $\mathbb{R}^N$ . An oriented  $(m - k)$ -dimensional plane  $l$  in  $\mathbb{R}^N$  separates  $\tau_0$  and  $\tau_1$  if  $l$  is transverse to  $\tau_0$  and  $\tau_1$  and the orientations induced by  $\tau_0 + l$  and  $\tau_1 + l$  in  $\mathbb{R}^N$  are different. Notice that there always exists an open set of orientable  $(m - k)$ -planes separating any two distinct orientable  $k$ -planes.

ROLLE'S LEMMA. *If a continuous family of orientable  $k$ -planes  $\{\tau_t\}_{t \in [0,1]}$  connects  $\tau_0$  and  $\tau_1$  and an orientable  $(m - k)$ -plane  $l$  separates  $\tau_0$  and  $\tau_1$ , then for some  $t^* \in (0, 1)$  transversality of  $\tau_{t^*}$  and  $l$  fails.*

In what follows we use the transversality theorem [GG] which says: *if  $V \subset \mathbb{R}^N$  is a manifold, then almost every plane of dimension  $k$  is transverse to  $V$ .*

AN OUTLINE OF THE PROOF OF LEMMA 6. Let  $x \in \text{Sing}_a(V_j, V_i)$ , then by Lemma 5 there are sequences  $\{y'_n\}, \{y_n\} \subset V_j^{\text{con}, x}$  with different limiting tangent planes  $\tau = \lim T_{y_n} V_j$  and  $\tau' = \lim T_{y'_n} V_j$ . Choose an orientation of  $T_{y_0} V_j$ . By connecting  $y_0$  locally with all other points  $\{y'_n\}$  one can induce an orientation on all other  $T_{y'_n} V_j$  so that the orientations of  $\tau_0$  and  $\tau_1$  coincide with the orientations of the limits.

Denote  $\dim V_j$  by  $k$ . There is an oriented  $(N - k)$ -plane  $l_j$  separating  $\tau_0$  and  $\tau_1$  and transverse to  $V_j$  (by the transversality theorem). Consider the orthogonal projection  $\pi_{l_j}$  along  $l_j$  onto its orthogonal complement  $l_j^\perp$ . Denote by  $p_{l_j, j}$  its restriction to  $V_j$ ,  $p_{l_j, j} = \pi_{l_j}|_{V_j} : V_j \rightarrow l_j^\perp$ . Denote by  $\text{Crit}(l_j, V_j)$  the set of critical points of  $p_{l_j, j}$  in  $V_j$  where the rank of  $p_{l_j, j}$  is not maximal. Then  $\text{Crit}(l_j, V_j)$  is a semivariety in  $V_j$  and  $\dim \text{Crit}(l_j, V_j) < \dim V_j$ . Connect two points  $y_n^0$  and  $y_n^1$  by a curve in  $V_j$ ; then  $T_{y_n^0} V_j$  deforms continuously to  $T_{y_n^1} V_j$ . Then by Rolle's lemma there is a critical point of  $p_{l_j, j}$  in  $V_j^{\text{con}, x}$  arbitrarily close to  $x$ . Thus  $x \in \overline{\text{Crit}(l_j, V_j)}$ .

By the transversality theorem there is a countable dense set of orientable  $(N - k)$ -planes  $\{l_j^p\}_{p \in \mathbb{Z}_+}$  transverse to  $V_j$  and separating any two distinct orientable  $k$ -planes  $\tau_0$  and  $\tau_1$ . Therefore, we have that

$$(56) \quad \text{Sing}_a(V_j, V_i) \subset V_i \setminus \text{Un}(V_j, V_i) \subset \bigcup_{p \in \mathbb{Z}_+} \{\overline{\text{Crit}(l_j^p, V_j)} \setminus \text{Crit}(l_j^p, V_j)\}.$$

These sets  $\{\text{Crit}(l_j^p, V_j) = V_j^p\}_{p \in \mathbb{Z}_+}$  are the proper subsets in  $V_j$  we are looking for. Using some additional simple argument given in [K7] one can complete the proof of Lemma 6.  $\square$

3.4.3. *A proof of  $a_P$ -stratifications for polynomial functions.* The proof below also consists of two steps.

Step 1. *Construct a rank stratification of  $P$ .* Consider an  $a$ -regular stratification  $(V, \mathcal{V}^0)$  of  $V$  by semialgebraic strata. Such a stratification always exists by Whitney's Theorem 7 proved above. Now we refine the stratification  $\mathcal{V}^0$  to a stratification  $\mathcal{V}^1$  so that  $\mathcal{V}^1$  is a rank stratification of  $P$  or the restriction of  $P$  to any stratum  $V_i \in \mathcal{V}^1$  is a map of constant rank. Notice that it is sufficient to refine each stratum  $V_i \subset \mathcal{V}^0$  so that  $P$  restricted to each stratum  $V_i^j \subset V_i$  has constant rank.

There are two cases: if  $P(V_i)$  is a point, then the rank of  $P|_{V_i}$  is identically zero and  $V_i$  stays unchanged and if  $P(V_i)$  contains an open set, then denote by  $\Sigma_{i,P} \subset V_i$  the set of critical points of  $P|_{V_i}$ . By Sard's lemma for algebraic sets [M] the set  $\Sigma_{i,P}$  is a semialgebraic set of positive codimension in  $V_i$ . Now refine each  $\Sigma_{i,P}$  to be an  $a$ -regular stratification of  $\Sigma_{i,P}$ . This is possible by Whitney's Theorem 7. Denote such a stratification by  $(V, \mathcal{V}^1)$ . By our construction  $\mathcal{V}^1$  is an  $a$ -regular rank stratification of  $P|_V$ , i.e.,  $P$  has constant rank on each stratum and the strata "fit"  $a$ -regularly.

Step 2. It is sufficient to prove the following

**THEOREM 11.** *Let  $V_i, V_j \subset \mathcal{V}^1$  be two strata in  $\mathbb{R}^N$  and  $P: \mathbb{R}^N \rightarrow \mathbb{R}$  be a constant on each strata  $V_i$  and  $V_j$ . Then the set of singular points*

$$\text{Sing}_{a,P}(V_j, V_i) = \{x \in V_i \cap \overline{V_j} : V_j \text{ is not } a_P\text{-regular over } V_i \text{ at } x \text{ w.r.t. } P\}$$

*is semialgebraic and has positive codimension in  $V_i$ .*

Inductive refinement arguments from Section 3.4.1 along with Theorem 11 complete the proof of Hironaka's Theorem 8. The rest of the section is devoted to a proof of Theorem 11.

**PROOF OF THEOREM 11.** Similarly to the existence proof of Whitney's Theorem 7 above we define the set with a unique limit of tangent planes to level sets of  $P$

$$(57) \quad \text{Un}_P(V_j, V_i) = \{x \in V_i : \text{for any } V_j^{\text{con},x} \text{ there is } \tau_{x,P} \text{ such that} \\ \lim_{y_n \rightarrow x} \ker dP|_{V_j^{\text{con},x}}(y_n) = \tau_{x,P} \text{ and is unique}\}.$$

**LEMMA 7.** *With the above notations  $\text{Un}_P(V_j, V_i)$  and  $\text{Sing}_{a,P}(V_j, V_i)$  are semi-varieties and*

$$(58) \quad \text{Sing}_{a,P}(V_j, V_i) \subset V_i \setminus \text{Un}_P(V_j, V_i).$$

LEMMA 8. *With the above notations there is a set of strata  $\{V_j^p\}_{p \in \mathbb{Z}}$  (resp.  $\{V_i^p\}_{p \in \mathbb{Z}}$ ) in  $V_j$  (resp. in  $V_i$ ) each of positive codimension in  $V_j$  (resp. in  $V_i$ ) such that*

$$(59) \quad \text{Sing}_{a,P}(V_j, V_i) \subset \bigcup_{p \in \mathbb{Z}} \text{Sing}_{a,P}(V_j^p, V_i) \bigcup_{p \in \mathbb{Z}} V_i^p.$$

REMARK 8. Similarly to the remarks after Lemma 6 this lemma allows one to reduce the dimension of  $V_j$ 's and prove that  $\text{Sing}_{a,P}(V_j, V_i)$  has positive codimension in  $V_i$ . This would prove Theorem 11 and as a consequence it would prove Theorem 8. So what is left to prove is Lemmas 7 and 8.

PROOF OF LEMMA 7. The proof goes by contradiction. Let  $x \in \text{Un}_P(V_j, V_i) \cap \text{Sing}_{a,P}(V_j, V_i)$ . Then for any local connected component  $V^{\text{con},x}$  and any sequence  $\{y_n\} \subset V^{\text{con},x}$  there is a limiting plane  $\lim \ker dP|_{V_j}(x) = \tau_{x,P}$ . Moreover, we have  $\ker dP|_{V_i}(x) \not\subset \tau_{x,P}$ . Thus, there is a unit vector  $v \in \ker dP|_{V_i}$  and  $v \notin \tau_{x,P}$ . The contradiction we are going to get is to find a sequence of points  $\{y_n\} \subset V^{\text{con},x}$  such that  $\lim \ker dP|_{V_j}(y_n) = \tau' \supset v$ . The rest of the proof is devoted to the construction of such a sequence.

By the theorem on implicit functions one can straighten  $V_i$  along with the nonempty level sets  $P^{-1}(a) \cap V_i$ . Then the ray  $l_v(x) = \{y \in \mathbb{R}^N : (y-x)/|y-x| = v\} \subset P^{-1}(P(x)) \subset V_i$  belongs to the level set  $P^{-1}(P(x))$ . By an extension of Wall [W1] of Milnor's curve selection lemma there is a 2-dimensional "wing"  $V_{j,v} \subset V_j$  such that  $l_v(x) \subset \overline{V_{j,v}}$ .

By Lemma 6 the set of points with nonunique limit  $\text{Un}(V_{j,v}, l_v(x))$  is 0-dimensional. Therefore, by Lemma 5 there is a neighborhood  $U_x$  of  $x$  such that any  $y \in U_x \cap l_v(x)$  may be distinct from  $x$ ,  $V_{j,v}$  is a-regular over  $l_v(x)$  at  $y$  and the limit  $\tau_y$  of tangent planes  $T_{y_n} V_{j,v}$  as  $y \rightarrow x$  is unique. The last two properties imply that  $\tau_y$  depends continuously on  $y$  as long as the limit  $\tau_y$  is unique. Therefore, there is a neighborhood  $U_y \subset U_x \setminus x$  of  $y$  such that  $V_{j,v} \cap U_y$  is a  $C^1$ -manifold with boundary.

Consider a  $C^1$ -smooth one-sided chart in  $U_y \cap V_{j,v}$  and let  $\pi_L$  be the map from  $U_y \cap V_{j,v}$  into the 2-dimensional plane  $\mathbb{R}^2$ . The image  $\pi_L(V_i)$  is a line in  $\mathbb{R}^2$  and  $\pi_L(V_{j,v})$  is a one-sided neighborhood of this line. Using Rolle's type of argument it is easy to show this for a sequence  $\{y_n\} \subset \pi_L(V_{j,v})$  from a semineighborhood of  $\pi_L(x') \in \mathbb{R}^2$  such that  $y_n \rightarrow \pi_L(y)$  and  $T_{y'_n} \{\pi_L \circ P^{-1}(P(\pi_L^{-1}(y'_n)))\} \rightarrow \pi_L(v)$ . This implies that  $\lim \ker dP|_{V_{j,v}}(y'_n)$  tends to  $v$ , however,  $V_{j,v} \subset V_j^{\text{con},x}$ . So  $v = \lim_{n \rightarrow \infty} \ker dP|_{V_{j,v}}(\pi_L^{-1}(y_n)) \subset \lim_{n \rightarrow \infty} \ker dP|_{V_j^{\text{con},x}}(\pi_L^{-1}(y_n)) = \tau$ . This is a contradiction with  $v \notin \tau$ .  $\square$

PROOF OF LEMMA 8. The proof is almost the same as the proof of Lemma 6 outlined above [K7].  $\square$

This completes the proof of Theorem 11 of Hironaka.  $\square$

*Acknowledgments.* I would like to thank my thesis advisor John Mather and David Nadler for stimulating discussions and numerous remarks on the mathematics of this lecture.

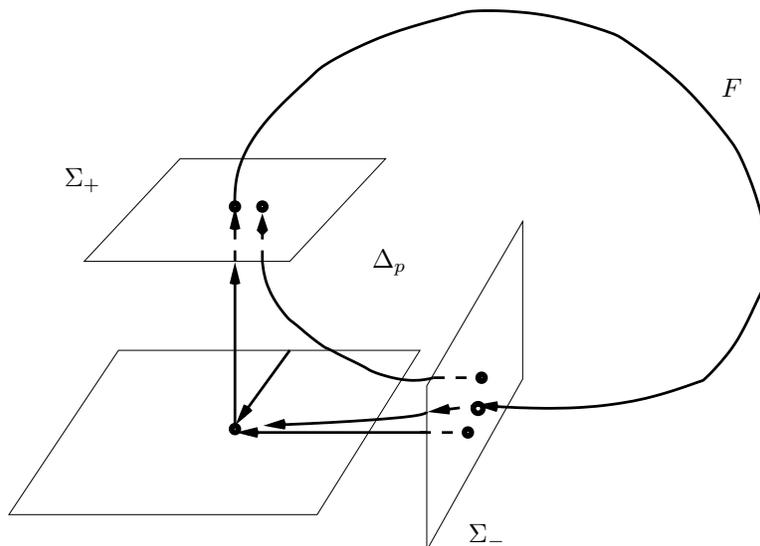


FIGURE 10. A saddle loop polycycle

#### 4. Bifurcation of spatial polycycles and blow-up along the diagonal of the space of multijets

In this lecture we discuss an essential ingredient of the proof of Theorem 1 [K8] about an estimate on cyclicity of spatial quasidelementary polycycles. First, in Section 4.1 we motivate the appearance of multichain maps (38) to get an estimate on cyclicity of spatial polycycles. Similarly to the planar case the question of estimating cyclicity of a quasidelementary polycycle reduces to estimating the geometric multiplicity of a multichain map of the form (38). To get an estimate on geometric multiplicity of a multichain map (38) one needs to prove a Bézout type theorem for multichain maps. However, a straightforward way to prove a Bézout type theorem for multichain maps faces a typical problem in singularity theory, namely, the problem that transversality fails on the diagonal in the space of multijets (see, e.g., [GG, M2]). We shall overcome this problem using a construction of Grigoriev–Yakovenko [GY] of blow-up along the diagonal in the space of multijets and a special multijet transversality theorem [GY]. This construction and its relation to Newton Interpolation Polynomials are described in Section 4.2. In Section 4.3 we describe the problem of rate of growth of the number of periodic points from Smooth Dynamical Systems (see, e.g., [AM, S4]) and outline the main result of the author along with Brian Hunt [KH, K5] in this direction. Finally, in Section 4.4 we outline how Newton Interpolation Polynomials can be applied to perturb trajectories and control the number of periodic points of diffeomorphisms.

**4.1. Multichain maps and spatial polycycles.** Consider the simplest example of a polycycle  $\gamma$  in  $\mathbb{R}^3$  consisting of a saddle equilibrium  $p$  and a connecting separatrix  $\gamma_p$  (see Figure 10).

Let  $\Sigma^-$  and  $\Sigma^+$  be “entrance” and “exit” transverse sections to  $\gamma_p$  chosen so that in  $C^r$ -normal coordinates the Poincaré return map  $\Delta_p$  along the polycycle  $\gamma$  has a “nice” form. We decompose the Poincaré return map  $\Delta$  along the polycycle

$\gamma$  into the composition of a local Poincaré map  $\Delta_p$  in a neighborhood of  $p$  and a semilocal map  $f$  along  $\gamma_p$ . Consider 2-cycles bifurcating from  $\gamma$ . Denote by  $x_1, x_2$  and  $y_1, y_2$  the first and the second intersection of each one of the 2-cycles with  $\Sigma^+$  and  $\Sigma^-$  respectively. Then the equation determining the number of 2-cycles has the form

$$(60) \quad \begin{cases} y_1 = f(x_1, \varepsilon) \\ x_2 = \Delta(y_1, \varepsilon) \\ y_2 = f(x_2, \varepsilon) \\ x_1 = \Delta(y_2, \varepsilon). \end{cases}$$

It is important that the first and the third equation have the same functional parts. Notice also that each of the equations from (60) is an equality in  $\mathbb{R}^2$  so it itself consists of two 1-dimensional equalities. So the total number of equations in (60) is 8. Following the strategy of the planar case from Section 2.2 we apply the Khovanskiĭ method to the system (60). Compare this system with the system (21) or (23). It is not difficult to see that the result of application of the Khovanskiĭ method gives a map of the form

$$(61) \quad P \circ (j^7 f, j^7 f): \mathbb{R}^8 \rightarrow \mathbb{R}^8.$$

To simplify our considerations we redenote the seventh jet  $j^7 f$  of  $f$  by a map  $F$  and consider the multichain map

$$(62) \quad P \circ (F, F): \mathbb{R}^8 \rightarrow \mathbb{R}^8,$$

where  $F$  is a *generic map* in a sense that it satisfies *any pre-given transversality condition*. It is clear that even if  $F$  is generic we cannot assume that the 2-tuple  $(F, F)$  is a generic map, simply because the first and the second components are the same. Let's explain why genericity fails for a 2-tuple mapping by an example.

#### 4.1.1. *Genericity (resp. transversality) fails for 2-tuple mappings!*

CLASSICAL TRANSVERSALITY THEOREM (e.g., [AGV, GG]). *Let  $N$  and  $m$  be positive integers and  $M$  be a smooth compact manifold in  $\mathbb{R}^m$ . Then for an open dense set of smooth mappings  $F: \mathbb{R}^N \rightarrow \mathbb{R}^m$  we have that  $F$  is transverse to  $M$ . In particular, this means that  $F^{-1}(M)$  is a smooth manifold.*

REMARKS. (1) It is an exercise from calculus to construct a set on the unit interval  $[0, 1]$  which is open dense and has arbitrary small positive measure. To justify that the transversality property is indeed generic, there is the so-called prevalent extension of the classical transversality theorem which says that for an a.e. mapping  $F: B^n \rightarrow \mathbb{R}^N$  we have that  $F$  is transverse to  $M$ . More exactly, for a generic finite-parameter family of mappings  $\{F_\varepsilon: B^n \rightarrow \mathbb{R}^N\}_{\varepsilon \in B^k}$  for a.e. parameter value,  $F_\varepsilon$  is transverse to  $M$ . See [HSY, K1] for more details.

(2) The fact that transversality of  $F$  to  $M$  implies that  $F^{-1}(M)$  is a smooth manifold follows from the theorem on implicit functions (see [AGV, GG]).

Since for a genericity condition on  $F$  what we need is that  $F$  has to satisfy a transversality condition, to show the impossibility of application of the classical transversality theorem we give a trivial example when transversality fails for an open set of 2-tuples  $(F, F)$ .

EXAMPLE 6. In the classical transversality theorem put  $n = m = 1$ . Consider the function  $f: x \rightarrow x^2$  for  $x \in I = [-1, 1]$  and the corresponding 2-tuple  $f \times f: I \times I \rightarrow \mathbb{R} \times \mathbb{R}$ , given by  $f \times f: (x_1, x_2) \rightarrow (x_1^2, x_2^2) = (y_1, y_2)$ . Let  $M = \{y_1 = y_2\} \subset \mathbb{R} \times \mathbb{R}$  be the diagonal. Then for each  $\tilde{f}$  which is  $C^1$ -close to  $f$  the preimage  $(\tilde{f}, \tilde{f})^{-1}(M)$  is a topological cross (not a manifold). This, in particular, implies that  $(\tilde{f}, \tilde{f})$  is not transverse to  $M$ , otherwise the preimage of a manifold would be a manifold.

To see this, notice that a function  $\tilde{f}$  close to  $f$  has to have a local minimum  $\tilde{x}$  close to 0 and  $\tilde{x}$  is a nondegenerate local minimum, i.e.,  $\tilde{f}: x \rightarrow \tilde{\varepsilon} + \tilde{a}(x - \tilde{x})^2 + o((x - \tilde{x})^2)$  with  $\tilde{a} \neq 0$ . Then  $\tilde{f}(x_1) - \tilde{f}(x_2) = 0$  has two intersecting curves of solutions  $x_1 = x_2$  and  $x_1 - \tilde{x} \approx -(x_2 - \tilde{x})$  which form a cross. This completes the proof of the claim in the example.

To explain what happens in this example we derive a general framework due to Grigoriev-Yakovenko [GY].

4.1.2. *Blow-up along the diagonal for 2-tuples in the 1-dimensional case.* For a smooth function  $f: \mathbb{R} \rightarrow \mathbb{R}$  consider the maps

$$(63) \quad (x_1, x_2) \xrightarrow{\mathcal{D}_2 \tilde{f}} \left( x_1, x_2, \tilde{f}(x_1), \frac{\tilde{f}(x_2) - \tilde{f}(x_1)}{x_2 - x_1} \right) = (x_1, x_1, u_1, u_2) \subset \mathbb{R}^4, \\ (x_1, x_2, u_1, u_2) \xrightarrow{\pi_2} (x_1, x_2, u_1, u_1 + u_2(x_2 - x_1)).$$

Direct calculation shows that  $\pi_2 \circ \mathcal{D}_2 \tilde{f} \equiv (\tilde{f}, \tilde{f})$ . Therefore,

$$(64) \quad (\tilde{f}, \tilde{f})^{-1}(M) = (\mathcal{D}_2 \tilde{f})^{-1} \circ \pi_2^{-1}(M).$$

This is incorporated into Figure 3 of with  $n = N = 1$ . By definition  $\mathcal{D}_2(f): I \times I \rightarrow \mathcal{DD}_2(I, \mathbb{R})$  is a smooth map,  $\pi_2: \mathcal{DD}_2(I, \mathbb{R}) \rightarrow \mathbb{R}$  is an explicitly computable polynomial map, and  $\pi_2 \circ \mathcal{D}_2(f) = (f, f): I \times I \rightarrow \mathbb{R}^2$ . Notice that outside of the diagonal  $\{x_1 = x_2\}$  the map  $\pi_2$  is one-to-one. However, the preimage of the set  $\pi_2^{-1}\{x_1 = x_2, f(x_1) = f(x_2)\}$  is of dimension 3 while the set  $\{x_1 = x_2, f(x_1) = f(x_2)\}$  itself is of dimension 2. This, in particular, means that  $\pi_2$  is a blow-up along the diagonal.

Consider  $\pi_2^{-1}(M) = \{u_2(x_2 - x_1) = 0\} \subset \mathbb{R}^4$ . This is the union of two intersecting hyperplanes. If the map  $\mathcal{D}_2 \tilde{f}$  is transverse to  $\{u_2(x_2 - x_1) = 0\}$ , then the preimage  $(\mathcal{D}_2 \tilde{f})^{-1}(\{u_2(x_2 - x_1) = 0\})$  has to be the union of two intersecting curves. It turns out that the assumption that  $\mathcal{D}_2 \tilde{f}$  is *generic for a generic  $\tilde{f}$*  is satisfied or for a generic  $\tilde{F}$  the map  $\mathcal{D}_2 \tilde{f}$  is transverse to both hyperplanes of  $\{u_2(x_2 - x_1) = 0\}$ . Let's justify that.

A "PROOF" OF THE CLASSICAL TRANSVERSALITY THEOREM OF THOM (see, e.g., [AGV, GG]). Transversality is an open property, i.e., if  $F$  is transverse to  $M$ , then for all  $\tilde{F}$  sufficiently close to  $F$  we have  $\tilde{F}$  is transverse to  $M$  too. So it is sufficient to show that by an arbitrary small perturbation of any mapping  $F: \mathbb{R}^N \rightarrow \mathbb{R}^m$  one can reach transversality to a compact manifold  $M$ . Let's prove it now.

Consider a smooth mapping  $F: B^n \rightarrow \mathbb{R}^N$ . Include this mapping into the  $m$ -parameter family  $\mathbf{F}: \mathbb{R}^n \times \mathbb{R}^N \rightarrow \mathbb{R}^n \times \mathbb{R}^N$ , given by  $\mathbf{F}(x, \varepsilon) = (x, F(x) + \varepsilon)$ . The determinant of the linearization of the mapping (the Jacobian)  $J_{\mathbf{F}}(x, \varepsilon)$  is constant and identically equals 1. Therefore,  $\mathbf{F}$  is a diffeomorphism and  $M_{\mathbf{F}} = \mathbf{F}^{-1}(\mathbb{R}^n \times M)$  is a manifold in the preimage  $\mathbb{R}^n \times \mathbb{R}^N$ .

FACT. *If  $\varepsilon$  is a regular point of the projection  $\pi_{M,F} = \pi|_{M_F}: \mathbb{R}^n \times \mathbb{R}^N \rightarrow \mathbb{R}^N$  along the  $x$ -coordinate, restricted to  $M_F$ , then  $F_\varepsilon = F(x) + \varepsilon$  is transverse to  $M$ .*

This follows from the implicit function theorem.

SARD'S LEMMA (e.g., [M4]). *A.e.  $\varepsilon \in \mathbb{R}^N$  is a regular value for the projection map  $\pi_{M,F}$ .*

Thus, one can choose a regular value  $\varepsilon$  arbitrarily close to 0. For such an  $\varepsilon$  the mapping  $F_\varepsilon$  is transverse to  $M$ . This completes the proof of the classical transversality theorem.  $\square$

Now we are ready to state the main result of this section.

4.1.3. *Multijet transversality theorem due to Grigoriev–Yakovenko.*

THEOREM 12 ([GY]). *Let  $M \subset \mathbb{R}^n \times \mathbb{R}^N$  be a smooth variety and  $B^n \subset \mathbb{R}^n$  be a unit ball. Then for an open dense set of smooth mappings  $F: B^n \rightarrow \mathbb{R}^N$  the set  $(F \times F)^{-1}(M)$  is stratified.*

*Moreover, let  $k \in \mathbb{Z}_+$  and  $M \subset \mathbb{R}^n \times \cdots \times \mathbb{R}^N$  ( $k$  times) be a smooth algebraic variety. Then for an open dense set of smooth mappings  $F: B^n \rightarrow \mathbb{R}^N$  the set  $(F \times \cdots \times F)^{-1}(M)$ , with  $k$  repetitions, is stratified.*

*Moreover, let  $n, k \in \mathbb{Z}_+$  and  $M \subset J^m(B^n, \mathbb{R}^N) \times \cdots \times J^m(B^n, \mathbb{R}^N)$  ( $k$  times) be a smooth algebraic variety. Then for an open dense set of smooth mappings  $F: B^n \rightarrow \mathbb{R}^N$  the set  $(F \times \cdots \times F)^{-1}(M)$ , with  $k$  repetitions, is stratified.*

PROOF OF THE THEOREM FOR THE MODEL EXAMPLE  $n = m = 1$  AND  $k = 2$ . Consider the map  $\mathcal{D}_2(F): (x_1, x_2, \varepsilon_1, \varepsilon_2) \mapsto (x_1, x_2, u_1, u_2)$ , defined by the formula (63). Direct calculations show that the determinant of the linearization (the Jacobian)  $J_{\mathcal{D}_2(F)}(x_1, x_2, \varepsilon_1, \varepsilon_2) \equiv 1$  and is formed by an upper triangular matrix with units on the diagonal. Since,  $\mathcal{D}_2(F)$  is a diffeomorphism one can apply arguments of the proof of the classical transversality theorem given above.  $\square$

A proof of the theorem in the general case follows along the same lines. The main difficulty is to construct Figure 3 in the general case. This is the subject of the next subsection.

COROLLARY 3. *For an open dense set of smooth functions  $F: I \rightarrow \mathbb{R}$  the preimage  $(F, F)^{-1}(M)$  is a 1-dimensional stratified manifold, i.e., a locally finite union of points and curves.*

**4.2. Newton interpolation polynomials and blow-up along the diagonal in the space of multijets.** This section is devoted to the description of the Grigoriev–Yakovenko construction of blow-up along the diagonal in the space of multijets in the general case. Let  $F: B^n \rightarrow \mathbb{R}^N$  be a smooth map of a unit ball  $B^n \subset \mathbb{R}^n$ ,  $j^m F: B^n \rightarrow J^m(B^n, \mathbb{R}^N)$  be an  $m$ th-order jet of  $F$ , and  $\mathcal{J}^{m,k}(B^n, \mathbb{R}^N) = J^m(B^n, \mathbb{R}^N) \times \cdots \times J^m(B^n, \mathbb{R}^N)$  ( $k$  repetitions) be the space of  $k$ -tuple  $m$ -jets. Denote the  $k$ -tuple of the  $m$ -jet of a map  $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$  by  $\mathcal{J}^{m,k} F(x_1, \dots, x_k) = (j^m F(x_1), \dots, j^m F(x_k))$ . The goal of this section is to define entries of an extension of Figure 3: the so-called space of divided differences  $\mathcal{DD}_k^m(B^n, \mathbb{R}^N)$ , the Newton map  $\pi_k^m: \mathcal{DD}_k^m(B^n, \mathbb{R}^N) \rightarrow \mathcal{J}^{m,k}(B^n, \mathbb{R}^N)$ ,  $\mathcal{D}_k^m(F): \underbrace{B^n \times \cdots \times B^n}_{k \text{ repetitions}} \rightarrow \mathcal{DD}_k^m(B^n, \mathbb{R}^N)$ . We use the exposition from [GY].

$$\begin{array}{ccc}
& & \mathcal{DD}_k^m(B^n, \mathbb{R}^N) \\
& \nearrow \mathcal{D}_k^m(F) & \downarrow \pi_k^m \\
\underbrace{B^n \times \cdots \times B^n}_{k \text{ repetitions}} & \xrightarrow{\mathcal{J}^{k,m}F} & \mathcal{J}^{m,k}(B^n, \mathbb{R}^N)
\end{array}$$

FIGURE 11. Polynomial blow-up of the multijet space

4.2.1. *Divided differences.* In order to extend the above construction we need to define so-called *divided differences*. Let  $g: \mathbb{R} \rightarrow \mathbb{R}$  be a sufficiently smooth function of one real variable.

DEFINITION 20. The *first order divided difference* of  $g$  is defined as

$$(65) \quad \Delta g(x_1, x_2) = \frac{g(x_2) - g(x_1)}{x_2 - x_1}$$

for  $x_2 \neq x_1$  and extended by its limit value as  $g'(x)$  for  $x = x_2 = x_1$ . Clearly, if  $g$  is a  $C^r$ -smooth function, then  $\Delta g$  is at least a  $C^{r-1}$ -smooth function of its arguments.

Iterating this construction we define *divided differences of the  $s$ th order* for any  $s \in \mathbb{Z}_+$  as

$$\Delta^s g(x_1, \dots, x_{s+1}) = \frac{\Delta^{s-1} g(x_1, \dots, x_{s-1}, x_{s+1}) - \Delta^{s-1} g(x_1, \dots, x_{s-1}, x_s)}{x_{s+1} - x_s}$$

for  $x_{s+1} \neq x_s$  and extended by its limit value as  $\partial \Delta^{s-1} g(x) / \partial x_s$  for  $x = x_{s+1} = x_s$ . Clearly, if  $g$  is a  $C^r$ -smooth function, then  $\Delta g$  is at least a  $C^{r-s}$ -smooth function of its arguments.

Notice that  $\Delta^s$  is linear as a function of  $g$ , and one can show that it is a symmetric function of  $x_1, \dots, x_{s+1}$ ; in fact, by induction it follows that

$$(66) \quad \Delta^s g(x_1, \dots, x_s) = \sum_{i=1}^{s+1} \frac{g(x_i)}{\prod_{j \neq i} (x_i - x_j)}$$

Another identity that is proved by induction will be more important for us, namely

$$(67) \quad \Delta^s x^l(x_1, \dots, x_{s+1}) = p_{l,s}(x_1, \dots, x_{s+1}),$$

where  $p_{l,s}(x_1, \dots, x_{s+1})$  is 0 for  $s > l$  and for  $s \leq l$  is the sum of all degree  $l - s$  monomials in  $x_0, \dots, x_s$  with unit coefficients,

$$(68) \quad p_{l,s}(x_1, \dots, x_{s+1}) = \sum_{r_0 + \cdots + r_s = l - s} \prod_{j=1}^{s+1} x_j^{r_j}.$$

The divided differences form coefficients for the Newton interpolation formula. For all  $C^\infty$  functions  $g: \mathbb{R} \rightarrow \mathbb{R}$  we have

$$(69) \quad g(x) = \Delta^0 g(x_1) + \Delta^1 g(x_1, x_2)(x - x_1) + \cdots \\ + \Delta^{k-1} g(x_1, \dots, x_{k-2})(x - x_1) \cdots (x - x_{k-3}) \\ + \Delta^k g(x_1, \dots, x_{k-1}, x)(x - x_1) \cdots (x - x_{k-2})$$

identically for all values of  $x, x_1, \dots, x_k$ . All terms of this representation are polynomial in  $x$  except for the last one which we view as a remainder term. The sum of the polynomial terms is the degree  $(k - 1)$  *Newton interpolation polynomial* for  $g$  at

$\{x_s\}_{s=1}^k$ . To obtain a degree  $2k-1$  interpolation polynomial for  $g$  and its derivative at  $\{x_s\}_{s=1}^k$ , we simply use (69) with  $k$  replaced by  $2k$  and the  $2k$ -tuple of points  $\{x_{s \pmod k}\}_{s=1}^{2k}$ . Similarly one can construct an interpolation polynomial for  $g$  and its derivatives up to any finite order.

All terms of this representation, except for the last one, are polynomial in  $x$  and their sum is the  $k$ th order *Newton Interpolation Polynomial* denoted by  $\mathcal{P}_{k-1}(x, \mathbf{X}_k)$ , where  $\mathbf{X}_k = (x_1, \dots, x_k)$ .

Now we can define entries of Figure 11 in the case  $m = 0$ . Let  $\mathcal{DD}_k(I, \mathbb{R}) = \underbrace{I \times \dots \times I}_{k \text{ times}} \times \mathbb{R}^k = (x_1, \dots, x_k; u_0, u_1, \dots, u_{k-1})$ . It is called the *space of divided differences*. Then

$$(70) \quad \begin{aligned} \mathcal{D}_k(f) &: \underbrace{I \times \dots \times I}_{k \text{ times}} \rightarrow \mathcal{DD}_k(I, \mathbb{R}), \\ \mathcal{D}_k(f) &: (x_1, \dots, x_k) \mapsto (x_1, \dots, x_k; u_0, \dots, u_{k-1}), \quad u_\alpha = \Delta^\alpha f(x_1, \dots, x_{\alpha+1}) \\ \pi_k &: \mathcal{DD}_k(I, \mathbb{R}) \rightarrow \mathbb{R}^{2k}, \end{aligned}$$

where  $\mathcal{D}_k(f)$  is smooth, provided that  $F$  is smooth,  $\pi_n: \mathcal{DD}_k(I, \mathbb{R}) \rightarrow \mathbb{R}^{2k}$  is a Newton Interpolation Polynomial of the form (69), and

$$(71) \quad \pi_k \circ \mathcal{D}_k(f) = (f, \dots, f): I \times \dots \times I \rightarrow \mathbb{R}^{2k},$$

where  $f$  and  $\mathbb{R}$  are repeated  $k$  times.

4.2.2. *Language of divided differences and the Newton interpolation formula.* In this section we introduce the construction of the divided differences space  $\mathcal{DD}_k(B^n, \mathbb{R}^N)$  and the corresponding map  $\mathcal{D}_k(F)$  and the polynomial  $\pi_k$  presented on Figure 3.

Let  $F: \mathbb{R}^n \rightarrow \mathbb{R}$  be a smooth function in  $n$  real variables  $x_1, \dots, x_n$ .

DEFINITION 21. The *first order divided difference* of  $F$  in the variable  $x_k$  is the function of  $n+1$  variables  $x_1, \dots, x_{k-1}, x'_k, x''_k, \dots, x_n$  defined as

$$(72) \quad \begin{aligned} \Delta_{x_k} F(x_1, \dots, x_{k-1}, x'_k, x''_k, \dots, x_n) \\ = \frac{F(x_1, \dots, x_{k-1}, x'_k, \dots, x_n) - F(x_1, \dots, x_{k-1}, x''_k, \dots, x_n)}{x'_k - x''_k} \end{aligned}$$

for  $x'_k \neq x''_k$  and extended by its limit value as  $\partial F(x_1, \dots, x_{k-1}, x'_k, \dots, x_n) / \partial x_k$  for  $x'_k = x''_k = x_k$ . Clearly, if  $F$  is a  $C^r$  function, then (e.g., by Hadamard lemma),  $\Delta_{x_k} F$  is at least a  $C^{r-1}$ -smooth function of its arguments.

It turns out that iterating this construction is possible [GY], which leads to

DEFINITION 22. Let  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}_+^n$  be a multiindex, let  $F$  be as above. Then  $\Delta_x^\alpha F = \Delta_{x_1}^{\alpha_1} \dots \Delta_{x_n}^{\alpha_n} F$  is called the mixed divided difference of order  $|\alpha| = \alpha_1 + \dots + \alpha_n$ . This divided difference is a smooth function of  $n+|\alpha|$  arguments subdivided into  $n$  groups of  $\alpha_1 + 1, \dots, \alpha_n + 1$  variables, symmetric with respect to permutations of variables within the same groups.

As direct calculations show the operators  $\Delta_{x_j}$  and  $\Delta_{x_i}$  commute for  $i \neq j$ , and, therefore, we can use the multiindex notation for divided differences.

4.2.3. *The Newton interpolation formula (in multivariables).* Let

$$X^1 = (x_1^1, \dots, x_N^1) \subset \mathbb{R}, \dots, X^n = (x_1^n, \dots, x_N^n) \subset \mathbb{R}$$

be subsets consisting of the same number of points, each  $X^j$  belonging to the corresponding  $j$ th coordinate axis of points in  $\mathbb{R}^N$ . Then, given a multiindex  $\alpha \in \mathbb{Z}_+^N$  and a smooth function  $F(x) = F(x^1, \dots, x^N)$  in  $N$  variables we can form the divided difference  $\Delta_x^\alpha F(X^1, \dots, X^N)$ .

In terms of the divided differences one can write the Newton interpolation polynomial as follows:

$$(73) \quad \mathcal{P}(t^1, \dots, t^N) = \sum_{0 \leq \alpha_i \leq n} \Delta_x^\alpha F(X^1, \dots, X^N) \prod_{i=1}^{\alpha_1} \dots \prod_{i=n}^{\alpha_n} (t^1 - x_{i_1}^1) \dots (t^n - x_{i_n}^n).$$

The polynomial  $\mathcal{P}(t^1, \dots, t^N)$  has degree  $\leq kn$  in the variables  $t = (t^1, \dots, t^N)$ . The Newton interpolation formula implies that the difference  $F(t) - \mathcal{P}(t^1, \dots, t^N)$  vanishes at all points of the Cartesian product grid  $\mathbf{X} = X^1 \times \dots \times X^N \subset \mathbb{R}^N$ . Moreover, if for each  $X^j = (x_1^j, \dots, x_n^j)$  we denote by  $\text{diag}^k(X^j)$  the set  $(x_1^j, \dots, x_n^j)$  repeated  $(m+1)$  times

$$(74) \quad \underbrace{(x_1^j, \dots, x_n^j)}_{(m+1) \text{ times}}, \dots, \underbrace{(x_1^j, \dots, x_n^j)}_{(m+1) \text{ times}}$$

then, to obtain interpolation of the  $m$ th jet of  $F$ , we replace each  $X^j = (x_1^j, \dots, x_n^j)$  by  $\text{diag}^m(X^j)$ . The degree of the interpolating polynomial will be  $\leq nk(m+1)$ .

In the case of a multivariate function  $F: B^n \rightarrow \mathbb{R}^N$  the interpolating polynomial  $\mathcal{P}(t^1, \dots, t^N)$  becomes an  $N$ -dimensional vector and is interpolating by coordinate functions of  $F = (F^1, \dots, F^N)$ .

**DEFINITION 23.** Let  $\mathcal{DD}_k^m(B^n, \mathbb{R}^N)$  be the collection of all divided differences with  $m$  repetitions,  $\{\Delta_x^\alpha F(\text{diag}^m(X^1), \dots, \text{diag}^m(X^n))\}_\alpha$ ,  $\alpha_i \leq (m+1)k$ ,  $i = 1, \dots, n$ . This is a linear space naturally equipped with the coordinates  $\{x_i, u_\alpha : 0 \leq i \leq N, \alpha_i \leq (m+1)\}$ , where  $x_i$  (resp.  $u_\alpha$ ) are vectors from  $\mathbb{R}^n$  (resp.  $\mathbb{R}^N$ ). The dimension of this space is equal to  $kn + N((m+1)k)^n$ .

The map  $\mathcal{D}_k^m F$  is defined by

$$(75) \quad \mathcal{D}_k^m F: (x_1, \dots, x_k) \rightarrow (x_1, \dots, x_k, \{u_\alpha\}_\alpha),$$

where  $u_\alpha = \Delta_x^\alpha F$ ,  $\forall i \alpha_i \leq (m+1)n$ .

The multivariate interpolation formula together with its derivatives in  $t_j$  evaluated at the points of the grid can be interpreted as a polynomial map restoring multijets from divided differences.

**NEWTON INTERPOLATION ON  $\mathbb{R}^m$  (abstract version).** *The multivariate Newton interpolation formula (73) defines a polynomial interpolation map*

$$\pi_k^m: \mathcal{DD}_k^m(B^n, \mathbb{R}^N) \rightarrow \mathcal{J}^{m,k}(B^n, \mathbb{R}^N)$$

such that  $\mathcal{J}^{m,k} f = \pi_k^m \circ \mathcal{D}_k^m F$ . The degrees of the components of  $\pi_k^m$  do not exceed  $(k+1)nN$ .

In the next section we present an application of Newton interpolation polynomials and Figure 11 to an old problem in dynamical systems: the problem of rate of growth of the number of periodic points for generic diffeomorphisms (see, e.g., [AM, S4]).

### 4.3. Rate of growth of the number of periodic points for generic diffeomorphisms and Newton interpolation polynomials.

4.3.1. *Statement of the problem.* Let  $\text{Diff}^r(M)$  be the space of  $C^r$  diffeomorphisms of a finite-dimensional smooth compact manifold  $M$  with the uniform  $C^r$ -topology,  $\dim M \geq 2$ , and let  $f \in \text{Diff}^r(M)$ . Consider the number of *isolated* periodic points of period  $n$

$$(76) \quad P_n(f) = \#\{\text{isolated } x \in M : x = f^n(x)\}.$$

The main question of this paper is:

How quickly can  $P_n(f)$  grow with  $n$  for a “generic” diffeomorphism  $f$ ?

We put the word “generic” in brackets because as the reader will see the answer depends on the notion of genericity.

We call a diffeomorphism  $f \in \text{Diff}^r(M)$  an *Artin–Mazur diffeomorphism* (or simply *A–M diffeomorphism*) if the number of isolated periodic orbits of  $f$  grows at most exponentially fast, i.e., for some number  $C > 0$

$$(77) \quad P_n(f) \leq \exp(Cn) \quad \text{for all } n \in \mathbb{Z}_+.$$

Artin & Mazur [AM] proved the following

**THEOREM 13.** *For any  $0 \leq r \leq \infty$ , A–M diffeomorphisms form a dense set of diffeomorphisms in  $\text{Diff}^r(M)$  with the uniform  $C^r$ -topology.*

In [K3] an elementary proof of the following extension of the Artin–Mazur result is given.

**THEOREM 14.** *For any  $0 \leq r < \infty$  A–M diffeomorphisms with all periodic points hyperbolic are dense in  $\text{Diff}^r(M)$  with the uniform  $C^r$ -topology.*

According to the standard terminology a set in  $\text{Diff}^r(M)$  is called generic (or residual) if it contains a countable intersection of open dense sets and a property is called (Baire) generic if diffeomorphisms with that property form a residual set. It turns out the A–M property is not generic, as is shown in [K4]. Moreover:

**THEOREM 15 ([K4]).** *For any  $2 \leq r < \infty$  there is an open set  $\mathcal{N} \subset \text{Diff}^r(M)$  such that for any given sequence  $a = \{a_n\}_{n \in \mathbb{Z}_+}$  there is a Baire generic set  $\mathcal{R}_a$  in  $\mathcal{N}$  depending on the sequence  $a_n$  with the property that if  $f \in \mathcal{R}_a$ , then for infinitely many  $n_i \in \mathbb{Z}_+$  we have  $P_{n_i}(f) > a_{n_i}$ .*

Since any two residual sets have nonempty intersection Theorem 15 implies that A–M diffeomorphisms are not generic. The proof of this theorem is based on a result of Gonchenko–Shil’nikov–Turaev [GST]. Two slightly different detailed proofs of their result are given in [K4, GTS]. The proof in [K4] relies on a strategy outlined in [GST].

However, it seems unnatural that if you pick a diffeomorphism at random then it may have an arbitrarily fast growth of number of periodic points. Moreover, Baire generic sets in Euclidean spaces can have zero Lebesgue measure. Phenomena which are Baire generic but have a small probability are well known in dynamical systems, KAM theory, number theory, etc. (see [O, HSY, K1] for various examples).

This partially motivates the problem posed by Arnol’d [A]: *Prove that “with probability one”  $f$  is an A–M diffeomorphism.* Arnol’d suggested the following interpretation of “with probability one”: *for a (Baire) generic finite parameter*

family of diffeomorphisms  $\{f_\varepsilon\}$ , for Lebesgue almost every  $\varepsilon$  we have that  $f_\varepsilon$  is A – M. (cf. [K1]). As Theorem 15 shows, a result on the genericity of the set of A – M diffeomorphisms based on topology (Baire) is likely to be extremely subtle, if possible at all.<sup>5</sup> We use instead a notion of “probability one” based on prevalence [HSY, K1], which is independent of Baire genericity. We also are able to state the result in the form Arnol'd suggested for generic families using this measure-theoretic notion of genericity. The main result in this direction is a partial solution to Arnol'd's problem. It says that for a prevalent diffeomorphism  $f \in \text{Diff}^r(M)$ ,  $r > 1$ , and any  $\delta > 0$  there exists  $C = C(\delta) > 0$  such that for all  $n \in \mathbb{Z}_+$

$$(78) \quad P_n(f) \leq \exp(Cn^{1+\delta})$$

This theorem is announced in [KH]. A major part of the proof is worked out in [K5]. We omit the precise statement, which requires an additional discussion.

#### 4.4. Dynamical usage of Newton interpolation polynomials.

4.4.1. *Perturbation of recurrent trajectories by Newton interpolation polynomials.* Let us start with several remarks which were the starting point of this paper. In order to keep notations and formulas simple we consider the 1-dimensional maps, but the reader should always have in mind that our consideration is designed for multidimensional diffeomorphisms.

Consider a map  $f: I \hookrightarrow I$  of the interval  $I = [-1, 1]$ . Recall that a trajectory  $\{x_k\}_{k \in \mathbb{Z}}$  of  $f$  is called *recurrent* if it returns arbitrarily close to its initial position — that is, for all  $\delta > 0$  we have  $|x_0 - x_n| < \delta$  for some  $n > 0$ . A very basic question of closing lemma type is how much one should perturb  $f$  to create a periodic point  $x_0$ . Let us give a “baby” answer

**BABY CLOSING LEMMA.** *Let  $\{x_k = f^k(x_0)\}_{k=0}^n$  be a trajectory of length  $n + 1$  of a map  $f: I \hookrightarrow I$ . Let  $u = (x_n - x_0) / \prod_{k=0}^{n-2} (x_{n-1} - x_k)$ . Then  $x_0$  is a periodic point of period  $n$  of the map*

$$(79) \quad \tilde{f}(x) = f(x) + u \prod_{k=0}^{n-2} (x - x_k)$$

Of course  $\tilde{f}$  is close to  $f$  only if  $u$  is sufficiently small, meaning that  $|x_0 - x_n|$  is small compared to  $\prod_{k=0}^{n-2} (x_{n-1} - x_k)$ . However, this product is likely to contain small factors for a recurrent trajectory. In general, it is difficult to control the effect of perturbations for recurrent trajectories. The simple reason is because *one cannot perturb  $f$  at two nearby points independently.*

It is important for the proof in [K5] to control the derivative of  $f$  along periodic orbits. If for some  $x \in I$ ,  $\gamma > 0$  and some positive integer  $n$  we have  $f^n(x) = x$  and  $|(f^n)'(x) - 1| > \gamma$ , then this implies that the interval around  $x$  of size  $\|f\|_{C^1}^{-n} \gamma$  is free from periodic points of the same period (see [KH, Proposition 1.1]). The quantity  $\gamma$  is called *hyperbolicity* and  $x$  is called  $(n, \gamma)$ -hyperbolic. This quantity was introduced by Gromov [G1] and Yomdin [Y2]. If one can estimate hyperbolicity for all points of period  $n$  from below, then one can estimate the number of periodic points of period  $n$ . The upper bound (78) is obtained by proving a lower bound on

<sup>5</sup>For example, using techniques from [GTS, K4] one can prove that for a Baire generic finite-parameter family  $\{f_\varepsilon\}$  and a Baire generic parameter value  $\varepsilon$  the corresponding diffeomorphism  $f_\varepsilon$  is not A – M. Unfortunately, how to estimate from below the measure of non-A – M diffeomorphisms in a Baire generic finite-parameter family is so far an unreachable question.

the rate of decay of hyperbolicity with period for prevalent diffeomorphisms. This is the reason the proof needs to control the derivative along trajectories.

The closing lemma above also gives an idea of how much we must change the parameter  $u$  to make a point  $x_0$  that is  $(n, \gamma)$ -periodic not be  $(n, \gamma)$ -periodic for a given  $\gamma > 0$ , which as we described above is one way to make a map that is “bad” for the initial condition  $x_0$  become “good.” To make use of our other alternative we must determine how much we need to perturb a map  $f$  to make a given  $x_0$  be  $(n, \gamma)$ -hyperbolic for some  $\gamma > 0$ .

**PERTURBATION OF HYPERBOLICITY.** Let  $\{x_k = f^k(x_0)\}_{k=0}^{n-1}$  be a trajectory of length  $n$  of a  $C^1$  map  $f : I \rightarrow I$ . Then for the map

$$(80) \quad f_v(x) = f(x) + v(x - x_{n-1}) \prod_{k=0}^{n-2} (x - x_k)^2$$

such that  $v \in \mathbb{R}$  and

$$(81) \quad \left| |(f_v^n)'(x_0)| - 1 \right| = \left| \left| \prod_{k=0}^{n-1} f'(x_k) + v \prod_{k=0}^{n-2} (x_{n-1} - x_k)^2 \prod_{k=0}^{n-2} f'(x_k) \right| - 1 \right| > \gamma$$

we have that  $x_0$  is an  $(n, \gamma)$ -hyperbolic point of  $f_v$ .

Once again we can see that the product of distances  $\prod_{k=0}^{n-2} |x_{n-1} - x_k|$  along the trajectory is an important quantitative characteristic of how much freedom we have to perturb.

The perturbations (79) and (80) are reminiscent of Newton interpolation polynomials. Let us put these formulas into a general setting using singularity theory.

**4.4.2. Distance to the diagonal in the multijet space.** Consider the  $2n$ -parameter family of a perturbation of a map  $f : I \hookrightarrow I$  by polynomials of degree  $2n - 1$

$$(82) \quad f_\varepsilon(x) = f(x) + \sum_{k=0}^{2n-1} \varepsilon_k x^k.$$

Define a map

$$(83) \quad \mathcal{J}_n^1 f : \underbrace{I \times \cdots \times I}_{n \text{ times}} \times \mathbb{R}^{2n} \rightarrow \underbrace{I \times \cdots \times I}_{n \text{ times}} \times \underbrace{(I \times \mathbb{R}) \times \cdots \times (I \times \mathbb{R})}_{n \text{ times}}$$

$$\mathcal{J}^{1,n} f(x_0, \dots, x_{n-1}, \varepsilon) = (x_0, \dots, x_{n-1}, f_\varepsilon(x_0), f'_\varepsilon(x_0), \dots, f_\varepsilon(x_{n-1}), f'_\varepsilon(x_{n-1})).$$

This map is called the  $n$ -tuple 1-jet map. The 1-jet of a function means that we take into account not only the image of a point, but also its derivative. The 1-jet of a function is usually denoted by  $j^1 f_\varepsilon(x) = (x, f_\varepsilon(x), f'_\varepsilon(x))$ . The space of 1-jets of functions on the interval  $I$  is denoted by  $\mathcal{J}^1(I, \mathbb{R})$ . The product of  $n$  copies of  $\mathcal{J}^1(I, \mathbb{R})$  is *multijet space* and is denoted by

$$(84) \quad \mathcal{J}^{1,n}(I, \mathbb{R}) = \underbrace{\mathcal{J}^1(I, \mathbb{R}) \times \cdots \times \mathcal{J}^1(I, \mathbb{R})}_{n \text{ times}}.$$

We need to include derivatives into our consideration, because we are interested in hyperbolicity (property of derivatives) of periodic points. The set of points

$$(85) \quad \Delta_n(I) = \{ \{x_0, \dots, x_{n-1}\} \times \mathbb{R}^{2n} \subset \underbrace{I \times \cdots \times I}_{n \text{ times}} \times \mathbb{R}^{2n} : \exists (i, j) \text{ s.t. } i \neq j, x_i = x_j \}$$

is called the *diagonal* in the space of multijets. In singularity theory the space of multijets is defined outside of the diagonal  $\Delta_n(I)$  and is usually denoted by  $\mathcal{J}^{1,n}(I, \mathbb{R}) = \mathcal{J}^{1,n}(I, \mathbb{R}) \setminus \Delta_n(I)$  (see [GG]).

It is easy to see that a *recurrent trajectory*  $\{x_k\}_{k \in \mathbb{Z}_+}$  is located in a neighborhood of the diagonal  $\Delta_n(I)$  in the space of multijets for a sufficiently large  $n$ . If  $\{x_k\}_{k=0}^{n-1}$  is a part of a recurrent trajectory of length  $n$ , then the product of distances along the trajectory

$$(86) \quad \prod_{k=0}^{n-2} |x_{n-1} - x_k|$$

measures how close  $\{x_k\}_{k=0}^{n-1}$  is to the diagonal  $\Delta_n(I)$ , or how independently one can perturb points of a trajectory. One can also say that (86) is a quantitative characteristic of how recurrent a trajectory of length  $n$  is. Introduction of this *product of distances along a trajectory* is a new central point of the method.

4.4.3. *Newton interpolation and blow-up along the diagonal in multijet space.* Now look at Grigoriev–Yakovenko’s construction [GY] in the 1-dimensional case with more details. This construction puts the “closing lemma” and “perturbation of hyperbolicity” statements above into a general framework.

Again consider the  $2n$ -parameter family (82) of perturbations of a  $C^1$  map  $f: I \rightarrow I$  by polynomials of degree  $2n - 1$ . Our goal now is to describe how such perturbations affect the  $n$ -tuple 1-jet of  $f$ , and since the operator  $j^{1,n}$  is linear in  $f$ , for the time being we consider only the perturbations  $\phi_\varepsilon$  and their  $n$ -tuple 1-jets. For each  $n$ -tuple  $\{x_k\}_{k=0}^{n-1}$  there is a natural transformation  $\mathcal{J}^{1,n}: I^n \times \mathbb{R}^{2n} \rightarrow \mathcal{J}^{1,n}(I, \mathbb{R})$  from  $\varepsilon$ -coordinates to jet-coordinates, given by

$$(87) \quad \mathcal{J}^{1,n}(x_0, \dots, x_{n-1}, \varepsilon) = j^{1,n}\phi_\varepsilon(x_0, \dots, x_{n-1}).$$

Instead of working directly with the transformation  $\mathcal{J}^{1,n}$ , we introduce intermediate  $u$ -coordinates based on Newton interpolation polynomials. The relation between  $\varepsilon$ -coordinates and  $u$ -coordinates is given implicitly by

$$(88) \quad \phi_\varepsilon(x) = \sum_{k=0}^{2n-1} \varepsilon_k x^k = \sum_{k=0}^{2n-1} u_k \prod_{j=0}^{k-1} (x - x_{j(\bmod n)}).$$

Based on this identity, we can define functions  $\mathcal{D}_n^1: I^n \times \mathbb{R}^{2n} \rightarrow I^n \times \mathbb{R}^{2n}$  and  $\pi_n^1: I^n \times \mathbb{R}^{2n} \rightarrow \mathcal{J}^{1,n}(I, \mathbb{R})$  so that  $\mathcal{J}^{1,n} = \pi_n^1 \circ \mathcal{D}_n^1$ , or in other words the diagram in Figure 11 commutes. This definition coincides with the one we gave before. We will show later that  $\mathcal{D}_n^1$  is invertible, while  $\pi_n^1$  is invertible away from the diagonal  $\Delta_n(I)$  and defines a blow-up along it in the space of multijets  $\mathcal{J}^{1,n}(I, \mathbb{R})$ . Consider Figure 11 for  $m = n = N = 1$ .

Recall that the intermediate space, denoted by  $\mathcal{DD}_n^1(I, \mathbb{R})$ , is called the *space of divided differences* and consists of  $n$ -tuples of points  $\{x_k\}_{k=0}^{n-1}$  and  $2n$  real coefficients  $\{u_k\}_{k=0}^{2n-1}$ . Here are explicit coordinate-by-coordinate formulas defining

$$\pi_n^1: \mathcal{DD}_n^1(I, \mathbb{R}) \rightarrow \mathcal{J}^{1,n}(I, \mathbb{R}).$$

$$\begin{aligned}
& \phi_\varepsilon(x_0) = u_0, \\
& \phi_\varepsilon(x_1) = u_0 + u_1(x_1 - x_0), \\
(89) \quad & \phi_\varepsilon(x_2) = u_0 + u_1(x_2 - x_0) + u_2(x_2 - x_0)(x_2 - x_1), \\
& \quad \vdots \\
& \phi_\varepsilon(x_{n-1}) = u_0 + u_1(x_{n-1} - x_0) + \cdots + u_{n-1}(x_{n-1} - x_0) \cdots (x_{n-1} - x_{n-2}), \\
& \phi'_\varepsilon(x_0) = \frac{\partial}{\partial x} \left( \sum_{k=0}^{2n-1} u_k \prod_{j=0}^{k-1} (x - x_{j(\text{mod } n)}) \right) \Big|_{x=x_0}, \\
(90) \quad & \quad \vdots \\
& \phi'_\varepsilon(x_{n-1}) = \frac{\partial}{\partial x} \left( \sum_{k=0}^{2n-1} u_k \prod_{j=0}^{k-1} (x - x_{j(\text{mod } n)}) \right) \Big|_{x=x_{n-1}},
\end{aligned}$$

These formulas are very useful for dynamics. For a given base map  $f$  and initial point  $x_0$ , the image  $f_\varepsilon(x_0) = f(x_0) + \phi_\varepsilon(x_0)$  of  $x_0$  depends only on  $u_0$ . Furthermore the image can be set to any desired point by choosing  $u_0$  appropriately — we say then that it depends nontrivially on  $u_0$ . If  $x_0, x_1$ , and  $u_0$  are fixed, the image  $f_\varepsilon(x_1)$  of  $x_1$  depends only on  $u_1$ , and as long as  $x_0 \neq x_1$  it depends nontrivially on  $u_1$ . More generally for  $0 \leq k \leq n-1$ , if pairwise distinct points  $\{x_j\}_{j=0}^k$  and coefficients  $\{u_j\}_{j=0}^{k-1}$  are fixed, then the image  $f_\varepsilon(x_k)$  of  $x_k$  depends only and nontrivially on  $u_k$ .

Suppose now that an  $n$ -tuple of points  $\{x_j\}_{j=0}^n$  not on the diagonal  $\Delta_n(I)$  and Newton coefficients  $\{u_j\}_{j=0}^{n-1}$  are fixed. Then derivative  $f'_\varepsilon(x_0)$  at  $x_0$  depends only and nontrivially on  $u_n$ . Likewise for  $0 \leq k \leq n-1$ , if distinct points  $\{x_j\}_{j=0}^n$  and Newton coefficients  $\{u_j\}_{j=0}^{n+k-1}$  are fixed, then the derivative  $f'_\varepsilon(x_k)$  at  $x_k$  depends only and nontrivially on  $u_{n+k}$ .

As Figure 12 illustrates, these considerations show that for any map  $f$  and any desired trajectory of distinct points with any given derivatives along it, one can choose Newton coefficients  $\{u_k\}_{k=0}^{2n-1}$  and explicitly construct a map  $f_\varepsilon = f + \phi_\varepsilon$  with such a trajectory. Thus we have shown that  $\pi_n^1$  is invertible away from the diagonal  $\Delta_n(I)$  and defines a blow-up along it in the space of multijets  $\mathcal{J}^{1,n}(I, \mathbb{R})$ .

The function  $\mathcal{D}^{1,n}: I^n \times \mathbb{R}^{2n} \rightarrow \mathcal{DD}^{1,n}(I, \mathbb{R})$  was explicitly defined using so-called divided differences above.

Recall that  $\mathcal{D}_n^1$  was defined implicitly by (88). We have described how to use divided differences to construct a degree  $2n-1$  interpolating polynomial of the form on the right-hand side of (88) for an arbitrary  $C^\infty$  function  $g$ . Our interest then is in the case  $g = \phi_\varepsilon$ , which as a degree  $2n-1$  polynomial itself will have no remainder term and coincide exactly with the interpolating polynomial. Thus  $\mathcal{D}^{1,n}$  is given coordinate-by-coordinate by

$$\begin{aligned}
(91) \quad u_m &= \Delta^m \left( \sum_{k=0}^{2n-1} \varepsilon_k x^k \right) (x_0, \dots, x_{m(\text{mod } n)}) \\
&= \varepsilon_m + \sum_{k=m+1}^{2n-1} \varepsilon_k p_{k,m}(x_0, \dots, x_{m(\text{mod } n)})
\end{aligned}$$

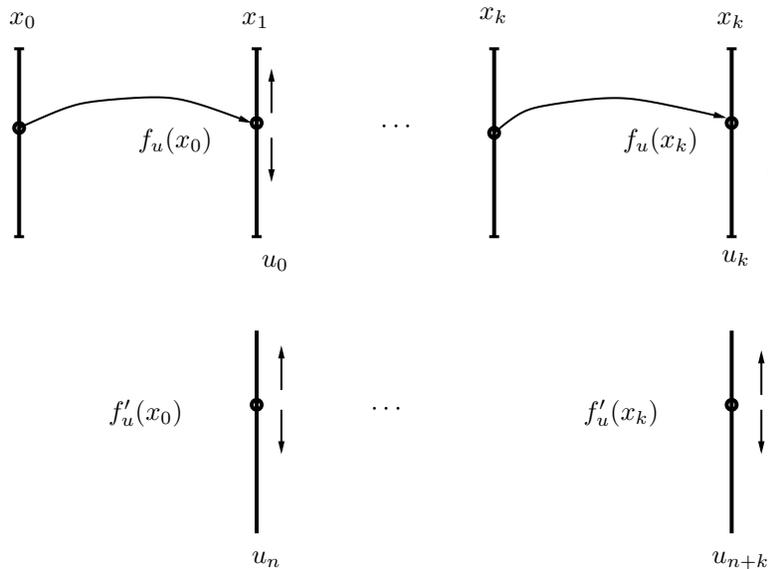


FIGURE 12. Newton coefficients and their action

for  $m = 0, \dots, 2n - 1$ . We call the transformation given by (91) the *Newton map*. Notice that for fixed  $\{x_k\}_{k=0}^{2n-1}$ , the Newton map is linear and given by an upper triangular matrix with units on the diagonal. Hence it is Lebesgue volume-preserving and invertible, whether or not  $\{x_k\}_{k=0}^{2n-1}$  lies on the diagonal  $\Delta_n(I)$ .

We call the basis of monomials

$$(92) \quad \prod_{j=0}^k (x - x_{j(\text{mod } n)}) \quad \text{for } k = 0, \dots, 2n - 1$$

in the space of polynomials of degree  $2n - 1$  the *Newton basis* defined by the  $n$ -tuple  $\{x_k\}_{k=0}^{n-1}$ . The Newton map and the Newton basis, and their analogues in dimension  $N$ , are useful tools for perturbing trajectories and proving (78).

*Acknowledgments.* In this lecture I have used fragments of the announcement [KH]. The good presentation of this announcement is entirely due to my coauthor Brian Hunt. Needless to say that numerous communications with him and also John Mather were very important for me.

### References

- [AA] D. V. Anosov and V. I. Arnol'd (eds.), *Dynamical systems*. I, Encyclopaedia Math. Sci., vol. 1, Springer-Verlag, Berlin, 1988.
- [A] V. Arnol'd, *Arnold problems*, with preface by M. B. Sevryuk and V. B. Filippov, Izdatel'stvo FAZIS, Moscow, 2000 (Russian).
- [AGV] V. I. Arnol'd, S. M. Guseĭn-Zade, and A. N. Varchenko, *Singularities of differentiable maps*. Vol. I: *The classification of critical points, caustics and wave fronts*, Monogr. Math., vol. 82, Birkhäuser, Boston, MA, 1985.
- [AM] M. Artin and B. Mazur, *On periodic points*, Ann. of Math. (2) **81** (1965), 82–99.
- [AVI<sup>+</sup>] V. I. Arnol'd, M. I. Vishik, Yu. S. Ilyashenko, A. S. Kalashnikov, V. A. Kondrat'ev, S. N. Kruzhkov, E. M. Landis, V. M. Millionshchikov, O. A. Oleĭnik, A. F. Filippov, and M. A. Shubin, *Some unsolved problems in the theory of differential equations and*

- mathematical physics*, Uspekhi Mat. Nauk **44** (1989), no. 4(268), 191–202 (Russian); English transl., Russian Math. Surveys **44** (1989), no. 4, 157–171.
- [B] I. Bendixson, *Sur les courbes définies par des équations différentielles*, Acta Math. **24** (1901), 1–88.
- [BS] I. S. Berezin and N. P. Shidkov, *Computing methods*. Vol. I, Pergamon, Oxford, 1965.
- [BCR] J. Bochnak, M. Coste, and M.-F. Roy, *Real algebraic geometry*, Ergeb. Math. Grenzgeb. (3), vol. 36, Springer-Verlag, Berlin, 1998.
- [BI] A. A. Bolibrukh, *Linear differential equations, Fuchsian inequalities and multiplicities of zeros*, in this volume.
- [B2] A. Buium, *Finiteness results in differential algebraic geometry and Diophantine geometry*, in this volume.
- [DR] Z. Denkowska and R. Roussarie, *A method of desingularization for analytic two-dimensional vector field families*, Bol. Soc. Brasil. Mat. (N.S.) **22** (1991), no. 1, 93–126.
- [DW] Z. Denkowska and K. Wachta, *Une construction de la stratification sous-analytique avec la condition (w)*, Bull. Polish Acad. Sci. Math. **35** (1987), no. 7-8, 401–405.
- [vdD] L. van den Dries, *o-minimal structures and real analytic geometry*, Current Developments in Mathematics (Cambridge, MA, 1998), Int. Press, Somerville, MA, 1999, pp. 105–152.
- [vdDS] L. van den Dries and P. Speissegger, *o-minimal preparation theorems*, Quad. Mat., (to appear).
- [D1] H. Dulac, *Sur les cycles limites*, Bull. Soc. Math. France **51** (1923), 45–188.
- [D2] F. Dumortier, *Singularities of vector fields on the plane*, J. Differential Equations **23** (1977), no. 1, 53–106.
- [DMR] F. Dumortier, M. El Morsalani, and C. Rousseau, *Hilbert’s 16th problem for quadratic systems and cyclicity of elementary graphics*, Nonlinearity **9** (1996), no. 5, 1209–1261.
- [DRR] F. Dumortier, R. Roussarie, and C. Rousseau, *Hilbert’s 16th problem for quadratic vector fields*, J. Differential Equations **110** (1994), no. 1, 86–133.
- [É] J. Écalle, *Introduction aux fonctions analysables et preuve constructive de la conjecture de Dulac*, Actualités Math., Herman, Paris, 1992.
- [FP] J.-P. Francoise and C. C. Pugh, *Keeping track of limit cycles*, J. Differential Equations **65** (1986), no. 2, 139–157.
- [GAK] Gabrielov A. and A. Khovanskiĭ, *Multiplicity of a Noetherian intersection*, Geometry of Differential Equations, Amer. Math. Soc. Transl. Ser. 2, vol. 186, Amer. Math. Soc., Providence, RI, 1998, pp. 119–131.
- [G] L. Gavrilov, *The infinitesimal 16th Hilbert problem in the quadratic case*, Invent. Math. **143** (2001), no. 3, 449–497.
- [GWDPL] C. G. Gibson, K. Wirthmüller, A. A du Plessis, and E. J. N. Loojenga, *Topological stability of smooth mappings*, Lecture Notes in Math., vol. 552, Springer-Verlag, Berlin–New York, 1976.
- [GG] M. Golubitsky and V. Guillemin, *Stable mappings and their singularities*, Grad. Texts in Math., vol. 14, Springer-Verlag, New York–Heidelberg, 1973.
- [GST] S. V. Gonchenko, L. P. Shil’nikov, and D. V. Tuvaev, *On models with nonrough Poincaré homoclinic curves*, Phys. D **62** (1993), no. 1-4, 1–14.
- [GTS] S. V. Gonchenko, D. V. Tuvaev, and L. P. Shil’nikov, *Homoclinic tangencies of an arbitrary order in Newhouse domains*, Dynamical Systems (Moscow, 1998), Itogi Nauki Tekh. Ser. Sovrem. Mat. Prilozh. Temat. Obz., vol. 67, Vseross. Inst. Nauchn. i Tekhn. Inform. (VINITI), Moscow, 1999, pp. 69–128 (Russian).
- [GM] M. Goresky and R. MacPherson, *Stratified Morse theory*, Ergeb. Math. Grenzgeb. (3), vol. 14, Springer-Verlag, Berlin, 1988.
- [GY] A. Grigoriev and S. Yakovenko, *Topology of generic multijet preimages and blow-up via Newton interpolation*, J. Differential Equations **150** (1998), no. 2, 349–362.
- [G1] M. Gromov, *On entropy of holomorphic maps* (1985), (preprint).
- [G2] T. Grozovskĭ, *Bifurcations of “apple” and “half-apple” polycycles in generic two-parameter families*, Differ. Uravn. **32** (1996), no. 4, 458–469; English transl., Differential Equations **32** (1996), no. 4, 459–470.
- [GH] J. Guckenheimer and P. Holmes, *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, Appl. Math. Sci., vol. 42, Springer-Verlag, New York, 1983.

- [H] H. Hironaka, *Subanalytic sets*, Number Theory, Algebraic Geometry and Commutative Algebra, in Honor of Yasuo Akizuki, Kinokuniya, Tokyo, 1973, pp. 453–493.
- [HSY] B. R. Hunt, T. Sauer, and J. A. Yorke, *Prevalence: a translation-invariant ‘almost every’ on infinite-dimensional spaces.*, Bull. Amer. Math. Soc. (N.S.) **27** (1992), no. 2, 217–238; addendum, Bull. Amer. Math. Soc. (N.S.) **28** (1993), no. 2, 306–307.
- [I1] Yu. S. Ilyashenko, *The multiplicity of limit cycles that arise in the perturbation of a Hamiltonian equation of the class  $\omega' = P_2/Q_1$  in a real and complex domain*, Trudy Sem. Petrovsk. **3** (1978), 49–60 (Russian); English transl., 191–202.
- [I2] ———, *Dulac’s memoir “On limit cycles” and related questions of the local theory of differential equations*, Uspekhi Mat. Nauk **40** (1985), no. 6(246), 41–78 (Russian); English transl., Russian Math. Surveys **40** (1985), no. 6, 1–50.
- [I3] ———, *Finiteness theorem for limit cycles*, Transl. Math. Monogr., vol. 94, Amer. Math. Soc., Providence, RI, 1991.
- [I4] ———, *Normal forms for local families and nonlocal bifurcations*, Complex Analytic Methods in Dynamical Systems (Rio de Janeiro, 1992), Astérisque, vol. 222, Soc. Math. France, Paris, 1994, pp. 233–258.
- [IK] Yu. S. Ilyashenko and V. Yu. Kaloshin, *Bifurcation of planar and spatial polycycles: Arnold’s program and its development*, The Arnoldfest (Toronto, 1997), Fields Inst. Commun., vol. 24, Amer. Math. Soc., Providence, RI, 1999, pp. 241–271.
- [IL] Yu. S. Ilyashenko and W. Li, *Nonlocal bifurcations*, Math. Surveys Monogr., vol. 66, Amer. Math. Soc., Providence, RI, 1999.
- [IY1] Yu. S. Ilyashenko and S. Yakovenko, *Finitely smooth normal forms of local families of diffeomorphisms and vector fields*, Uspekhi Mat. Nauk **46** (1991), no. 1(277), 3–39 (Russian); English transl., Russian Math. Surveys **46** (1991), no. 1, 1–43.
- [IY2] ———, *Finite cyclicity of elementary polycycles in generic families*, Concerning the Hilbert 16th Problem, Amer. Math. Soc. Transl. Ser. 2, vol. 165, Amer. Math. Soc., Providence, RI, 1995.
- [J] N. Jacobson, *Basic algebra*. I, W. H. Freeman and Co., San Francisco, CA, 1974.
- [K1] V. Yu. Kaloshin, *Some prevalent properties of smooth dynamical systems*, Tr. Mat. Inst. Steklova **213** (1997), 123–151 (Russian); English transl., Proc. Steklov Inst. Math. **1996**, no. 2(213), 115–140.
- [K2] ———, *Prevalence in spaces of finitely smooth mappings*, Funktsional. Anal. i Prilozhen. **31** (1997), no. 2, 27–33 (Russian); English transl., Funct. Anal. Appl. **31** (1997), no. 2, 95–99.
- [K3] ———, *An extension of the Artin–Mazur theorem*, Ann of Math. (2) **50** (1999), no. 2, 729–741.
- [K4] ———, *Generic diffeomorphisms with superexponential growth of number of periodic points*, Comm. Math. Physics **211** (2000), no. 1, 253–271.
- [K5] ———, *Stretched exponential estimate on the rate of growth of the number of periodic points for prevalent diffeomorphisms*, Princeton University, 2001.
- [K6] ———, *The Hilbert–Arnol’d problem and estimates on cyclicity of elementary polycycles in the plane and in space*, Funktsional. Anal. i Prilozhen. **35** (2001), no. 2, 78–81 (Russian); English transl., Functional Anal. Appl. **35** (2001), no. 2, 146–147.
- [K7] ———, *A geometric proof of existence of Whitney’s stratifications*, Moscow Math. J., (to appear).
- [K8] ———, *Bifurcations of spatial polycycles*, (in preparation) (Russian).
- [KH] V. Yu. Kaloshin and B. R. Hunt, *A stretched exponential estimate on the rate of growth of the number of periodic points for prevalent diffeomorphisms*. I, Electron. Res. Announc. Amer. Math. Soc. **7** (2001), 17–27; II, 28–36.
- [K1] A. G. Khovanskii, *Real analytic manifolds with the property of finiteness and complex abelian integrals*, Funktsional. Anal. i Prilozhen. **18** (1984), no. 2, 40–50 (Russian); English transl., Functional Anal. Appl. **18** (1984), no. 2, 119–127.
- [K2] ———, *Fewnomials*, Transl. Math. Monogr., vol. 88, Amer. Math. Soc., Providence, RI, 1991.
- [K3] O. Kleban, *Order of the topologically sufficient jet of a smooth vector field on the real plane at a singular point of finite multiplicity*, Concerning the Hilbert 16th Problem, Amer. Math. Soc. Transl. Ser. 2, vol. 165, Amer. Math. Soc., Providence, RI, 1995.

- [KS] A. Kotova and V. Stanzo, *On few-parameter generic families of vector fields on the two-dimensional sphere*, Concerning the Hilbert 16th Problem, Amer. Math. Soc. Transl. Ser. 2, vol. 165, Amer. Math. Soc., Providence, RI, 1995, pp. 155–201.
- [K] T.-C. Kuo, *The ratio test for analytic Whitney stratifications*, Proceedings of Liverpool Singularities-Symposium. I (1969/70), Lecture Notes in Math., vol. 192, Springer, Berlin, 1971, pp. 141–149.
- [L1] S. Lojasiewicz, *Ensembles semi-analytiques*, Institut de Hautes Études Scientifiques, 1965, (lecture notes).
- [L2] ———, *Sur la géométrie semi- et sous-analytique*, Ann. Inst. Fourier (Grenoble) **43** (1993), no. 5, 1575–1595.
- [LSW] S. Lojasiewicz, J. Stasica, and K. Wachta, *Stratifications sous-analytiques. Condition de Verdier*, Bull. Polish Acad. Sci. Math. **34** (1986), no. 9-10, 531–539.
- [M1] P. Mardešić, *An explicit bound for the multiplicity of zeros of generic Abelian integrals*, Nonlinearity **4** (1991), no. 3, 845–852.
- [M2] J. Mather, *Notes on topological stability*, Harvard University, 1970.
- [M3] J. W. Milnor, *Singular points of complex hypersurfaces*, Ann. of Math. Stud., vol. 61, Princeton Univ. Press, Princeton, NJ, 1968.
- [M4] ———, *Topology from the differentiable viewpoint*, Princeton Landmarks Math., Princeton Univ. Press, Princeton, NJ, 1997.
- [MR] R. Moussu and C. Roche, *Théorie de Hovanskiĭ et problème de Dulac*, Invent. Math. **105** (1991), no. 2, 431–441.
- [M] D. Mumford, *Algebraic geometry. I: Complex projective varieties*, Grundlehren Math. Wiss., vol. 221, Springer-Verlag, Berlin–New York, 1976.
- [NDY] Novikov D. and S. Yakovenko, *Simple exponential estimate for the number of zeroes of complete Abelian integrals*, Ann. Inst. Fourier (Grenoble) **45** (1995), no. 4, 897–927.
- [O] J. Oxtoby, *Measure and category: A survey of the analogies between topological and measure spaces*, Grad. Texts in Math., vol. 2, Springer-Verlag, New York–Berlin, 1980.
- [P1] W. Pawłucki, *Le théorème de Puiseux pour une application sous-analytique*, Polish Acad. Sci. Math. **32** (1984), no. 9-10, 555–560.
- [P2] G. S. Petrov, *Nonoscillation of elliptic integrals*, Funktsional. Anal. i Prilozhen. **24** (1990), no. 3, 45–50 (Russian); English transl., Funct. Anal. Appl. **24** (1990), no. 3, 205–210.
- [dPW] A. du Plessis and T. Wall, *The geometry of topological stability*, London Math. Soc. Monogr. (N.S.), vol. 9, Oxford Univ. Press, New York, 1995.
- [R1] R. Roussarie, *A note on finite cyclicity property and Hilbert's 16th problem*, Dynamical Systems (Valparaíso, 1986) (R. Bamón, R. Labarca, and J. Palis Jr., eds.), Lecture Notes in Math., vol. 1331, Springer-Verlag, Berlin, 1988, pp. 161–168.
- [R2] ———, *Bifurcations of planar vector fields and Hilbert's sixteenth problem*, Progr. Math., vol. 164, Birkhäuser, Basel, 1998.
- [S1] D. Schlomiuk (ed.), NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., vol. 408, Kluwer, Dordrecht, 1993.
- [S2] A. Seidenberg, *Reduction of singularities of the differential equations  $A dy = B dx$* , Amer. J. Math. **90** (1968), 248–269.
- [S3] L. P. Shil'nikov, *A case of the existence of a denumerable set of periodic motions*, Dokl. Akad. Nauk SSSR **160** (1965), 558–561 (Russian).
- [S4] S. Smale, *Differentiable dynamical systems*, Bull. Amer. Math. Soc. **73** (1967), 747–817.
- [T1] F. Takens, *Unfoldings of certain singularities of vector fields: generalized Hopf bifurcations*, J. Differential Equations **14** (1973), 476–493.
- [T2] R. Thom, *Ensembles et morphismes stratifiés*, Bull. Amer. Math. Soc. **75** (1969), 240–284.
- [T3] ———, *Propriétés différentielles locales des ensembles analytiques (d'après H. Whitney)*, Exp. No. 281, Séminaire Bourbaki, Vol. 9, Soc. Math. France, Paris, 1995, pp. 69–80.
- [T4] S. Trifonov, *Desingularization in families of analytic differential equations*, Concerning the Hilbert 16th Problem, Amer. Math. Soc. Transl. Ser. 2, vol. 165, Amer. Math. Soc., Providence, RI, 1995, pp. 97–129.

- [V] A. Varchenko, *Estimation of the number of zeros of an abelian integral depending on a parameter, and limit cycles*, Funktsional. Anal. i Prilozhen. **18** (1984), no. 2, 14–25 (Russian); English transl., Functional Anal. Appl. **18** (1984), no. 2, 98–108.
- [W1] T. C. Wall, *Regular stratifications*, Dynamical Systems (Warwick, 1974), Lecture Notes in Math., vol. 468, Springer-Verlag, Berlin–New York, 1975, pp. 332–344.
- [W2] H. Whitney, *Tangents to an analytic variety*, Ann. of Math. (2) **81** (1965), 496–549.
- [Y1] S. Yakovenko, *Quantitative theory of ordinary differential equations and tangential Hilbert 16th problem*, in this volume.
- [Y2] Y. Yomdin, *A quantitative version of the Kupka–Smale theorem*, Ergodic Theory Dynam. Systems **5** (1985), no. 3, 449–472.