Lecture Notes in Mathematics 1834

Editors:
J.–M. Morel, Cachan
F. Takens, Groningen
B. Teissier, Paris

Springer
Berlin
Heidelberg
New York
Hong Kong
London
Milan
Paris
Tokyo

Yosef Yomdin
Georges Comte

Tame Geometry with Application in Smooth Analysis

Springer

Authors

Yosef Yomdin

Department of Mathematics
Weizmann Institute of Science
Rehovot 76100, Israel

e-mail: yomdin@wisdom.weizmann.ac.il

Georges Comte

Laboratoire J. A. Dieudonné
UMR CNRS 6621
Université de Nice Sophia-Antipolis
28, avenue de Valrose
06108 Nice Cedex 2, France

e-mail: comte@math.unice.fr

Cataloging-in-Publication Data applied for

Bibliographic information published by Die Deutsche Bibliothek

Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at http://dnb.ddb.de

Mathematics Subject Classification (2000): 28A75, 14Q20, 14P10, 26B5, 26B15, 32S15

ISSN 0075-8434
ISBN 3-540-20612-4 Springer-Verlag Berlin Heidelberg New York

Springer-Verlag is a part of Springer Science + Business Media

http://www.springeronline.com

Typesetting: Camera-ready TeX output by the author

SPIN: 10973455 41/3142/ du - 543210 - Printed on acid-free paper

Preface

This book presents results and methods developed during quite a long period of time and many people helped in this work. We would like to thank Y. Kannai for pointing out, in the very beginning of the work on quantitative transversality, the relevance of metric entropy. Since 1983 M. Gromov encouraged this research and helped us in many fruitful discussions of qualitative transversality and Semialgebraic Geometry in Dynamics and Analysis. We would like to thank him especially for suggesting a problem of quantitative Kupka-Smale, for his contribution to C^k-reparametrization of semialgebraic sets and applications to dynamics, for providing a central (for this book) reference to Multidimensional Variations and to books of Vitushkin and Ivanov and for encouraging writing preliminary texts, which were used in this book. This book would not have be written without the help and encouragement of J. -J. Risler and B. Teissier and numerous fruitful discussions with them during all the long period of the book's preparation. It is a pleasure to thank M. Giusti, J. -P. Henry and M. Merle for their invitation to give a course on Metric Semialgebraic Geometry at École Polytechnique in 1985-86, and again M. Merle for his remarks during lectures given at the University of Nice - Sophia Antipolis in 1999-2000, and P. Milman for fruitful discussions and for an invitation to the University of Toronto, where the preliminary text has been written and typed. We would like to thank D. Trotman, who read and corrected the text, and also indicated precious references. We would like to thank M. Briskin, Y. Elichai, J.-P. Françoise, G. Loeper and N. Roytvarf for their help and contribution. Many of the results and methods presented in this book have been obtained in a long collaboration with them.

Last not least, many thanks belong to the University of Beer-Sheva, to the Max Planck Institut für Mathematik, to IHES, to the University of Toronto and to our home Universities - the University of Nice-Sophia Antipolis and the Weizmann Institute of Science, where many of the results and methods in this book have been developed, and to the Israel Science Foundation and the Minerva Foundation, which have supported the research of one of the authors for many years.

Table of Contents

1 Introduction and Content

1.1 Motivations

This book deals with several related topics:

– Geometry of real semialgebraic and tame sets (i.e. sets definable in some o-minimal structure on $(\mathbb{R}, +, .)$), with the stress on "metric" characteristics: lengths, volumes in different dimensions, curvatures etc...
– Behaviour of these characteristics under polynomial mappings.
– Integral geometry, especially the so-called Vitushkin variations, with the stress on applications to semialgebraic and tame sets.
– Geometry of critical and near critical values of differentiable mappings. Some fractal geometry naturally arising in this context.

Below we give a short description of each of these topics, their mutual dependance and logical order. Motivation for the type of question asked in this book, comes from several different sources: the main ones are Differential Topology, Singularity Theory, Smooth Dynamics, Control, Robotics and Numerical Analysis.

One of the main analytic results, underlying most basic constructions of Differential Analysis, Differential Topology, Differential Geometry, Differential Dynamics, Singularity Theory, as well as nonlinear Numerical Analysis, is the so-called Sard (or Morse-Sard, or Morse-Morse-Sard, see [Morse 1,2], [Mors], [Sar 1-3]) theorem. It asserts that the set of critical values of a sufficiently smooth mapping has measure zero. Mostly this theorem appears as an assumption, that the set $Y(c)$ of solutions of an equation $f(x) = c$ is a smooth submanifold of the domain of x, for almost any value c of the right hand side (in the semialgebraic or tame case, an asymptotic version of the Morse-Sard theorem (see [Rab], [Kur-Orr-Sim]) shows that f induces a fibration over the connected components of the "good values" c).

Another typical appearence of the Morse-Sard theorem is in the form of various transversality statements: by a small perturbation of the data we can always achieve a situation where all the submanifolds of interest intersect one another in a transversal way. Fig. 1.1 shows a non-transversal (a) and transversal (b, c) intersections of two plane curves.

Y. Yomdin and G. Comte: LNM 1834, pp. 1–22, 2004.

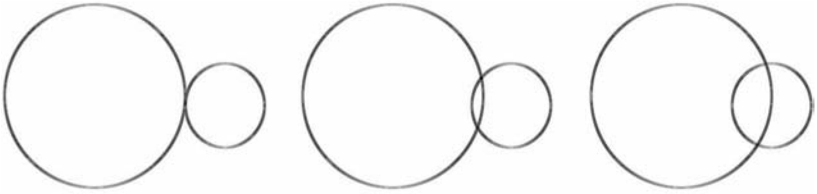

Fig. 1.1a Fig. 1.1b Fig. 1.1c

Fig. 1.1.

Technically, the Morse-Sard theorem is a rather subtle fact. It is well known since the classical examples of Whitney [Whi 1], that the geometry of the critical points and of the critical values of a differentiable function can be very complicated. In particular, the requirement that the function must have at least the same number of derivatives as the dimension of the domain, cannot be relaxed. And usually analytic facts that incorporate in an essential way the existence and the properties of the high order derivatives, are deep, both conceptually and technically.

However, in classical applications the Morse-Sard theorem appears as a background fact, as a default assumption, that all the transversality and regularity properties required can be achieved by small perturbations of the data. The interplay between the high order analytic structure of the mappings involved and their geometry rarely becomes apparent. The main reason is that the classical Morse-Sard theorem is basically *qualitative*. Its conclusion appears as an "existential" fact and it provides no quantitative information on the solutions, submanifolds etc... in question.

A very natural *quantitative* setting of the problems, covered by the Morse-Sard theorem, is possible. In the case of transversality, the question is: *given a maximal size of perturbations allowed, how strong a transversality of the perturbed submanifolds can be achieved?* (For plane curves the transversality can be measured just by the angle between the curves at their intersection points. Fig. 1.1c presents a strong transversality, in contrast to a weak one in Fig. 1.1b).

Concerning the set $Y(c)$ of solutions of the equation $f(x) = c$, a natural quantitative question is: *How much of the complexity of* $Y(c)$ *(geometric or topological) can be eliminated by a perturbation of* c *within the allowed range? What is the average complexity (with respect to* c*) of* $Y(c)$*?*

In nonlinear Numerical Analysis, a typical conclusion, provided by the classical Morse-Sard theorem, is that with probability one certain determinants do not vanish. However, to organize computations in a stable way it is necessary to get a quantitative information: *how big are these determinants usually ?*

The importance of this sort of quantitative information has been realized during the last decades in many fields. In the study of the complexity of algorithms (especially, in the work of Shub and Smale on complexity of the Newton type algorithms [Shu-Sma 1-7], [Sma 3,4]), quantitative considerations of the above type appear as one of the main tools (although mainly in situations where a direct treatment without Morse-Sard's theorem is possible).

In a recent study of high order numerical algorithms ([Eli-Yom 1-5], [Bri-Yom 6], [Bri-Eli-Yom], [Bic-Yom], [Wie-Yom], [Yom 24], [Y-E-B-S]) it became apparent that Quantitative Morse-Sard theorem and quantitative transversality may be crucially important in efficient organization of a high order data and in its efficient processing. We discuss this issue in some detail in Section 1.1.3, Chapter 1, and in Section 10.1.4, Chapter 10 below.

In Differential Dynamics, a number of "quantitative" problems have been posed by M. Gromov in the early eighties.

These concerned a quantitative behavior of periodic points, estimates for the volume growth and entropy etc... (See [Gro 1-4]). A "Quantitative Kupka-Smale theorem", bounding a typical quantitative behavior of periodic points and conjectured by M. Gromov, has been obtained in [Yom 4]. Very recently striking results in this direction have been obtained by Kaloshin [Kal 1-4] (some of these dynamical results are briefly discussed in Section 10.1.3, Chapter 10 below).

Important applications of quantitative transversality in symplectic geometry appeared recently in S. K. Donaldson's papers ([Don 1-3], see also [Sik]). These results have been further extended in [Aur], [Ibo] and other publications.

As for the Morse-Sard theorem itself, its sharpest quantitative version, concerning entropy dimension, has been obtained in [Yom 1]. Further applications, answering, in particular, a part of the quantitative questions above, appeared in [Yom 3,4,7,10,17,18,20]. More recently additional geometric and analytic information, related to different versions of the Morse-Sard theorem, concerning Hausdorff measure and dimension, has been obtained in [Bat 1-6], [Bat-Mor], [Bat-Nor], [Com 1], [Nor 1-4], [Nor-Pug], [Roh 1-3], [Yom 13-15,19], culminating in [Mor], in which the sharpest possible statement is given. Concerning singular values at infinity and the so-called Malgrange condition, one can see [Kur-Orr-Sim] for the semialgebraic case and [D'Ac] for the o-minimal case.

One of the main goals of this book is to give a proof and an "explanation" of the quantitative Morse-Sard theorem and related results. This is done via the study of the same questions first for polynomial (or tame) mappings. Indeed, while the classical Morse-Sard theorem is trivial for polynomials (critical values always form a semialgebraic set of a dimension smaller than that of the ambient space, and thus have Lebesgue measure zero), the quantitative questions above turn out to be nontrivial and highly productive. They are answered in this book by a combination of the methods of Real Semialgebraic and Tame Geometry and Integral Geometry.

One of the important advantages of this approach is that it allows one to separate the role of high differentiability and that of algebraic geometry in a smooth setting: all the geometrically relevant phenomena appear already for polynomial mappings. The geometric properties obtained are "stable with respect to approximation", and so can be imposed on smooth functions via polynomial approximation. The only role of high differentiability is to control the rate of this approximation. (In fact, the high order differentiability turns out to be not relevant at all in this circle of problems ! It is the rate of approximation by semialgebraic functions, that really counts. See Section 10.2 below).

Now the study of metric Semialgebraic Geometry with the above applications in view, essentially forces us to extend the tools beyond the usual lengths, areas etc... It is explained in detail below why using metric entropy (the minimal number of balls of a prescribed radius, covering a given set) and multidimensional variations (the average number of connected components in plane crossections of different dimensions) is most natural and rewarding in our setting.

In conclusion, let us express our hope that the results and methods presented in this book form only a beginning of the future "Quantitative Singularity Theory". The ultimate need for this theory is by now realized in many fields of mathematics. Quantitative Sard theorem, Quantitative Transversality and "Near Thom-Boardman Singularities" treated in this book definitely belong to this future theory, whose possible contours are discussed in some detail in Section 10.3.7 below.

1.2 Content and Organization of the Book

In the next section of this introduction we explain the main ideas of the semialgebraic part of the book, using a rather instructive example of the motion control problem in robotics. In the last section of the introduction we give an accurate proof of the simplest version of the generalized Sard theorem. This proof illustrates in a simple and transparent form (and without technicalities, unavoidable in a general setting) a good part of the ideas and methods developed below.

Chapter 2 is devoted to a precise introduction and a rather detailed study of the metric entropy of subsets of Euclidean spaces. We believe that the "transversality" results of Proposition 2.2 and Corollary 2.3, as well as a geometric interpretation of the entropy dimension, given by Theorem 2.9, are new.

In Chapter 3 we recall the theory of multidimensional variations, developed by A. G. Vitushkin ([Vit 1,2]), L.D. Ivanov ([Iva 1,2]), and others.

In general, to handle multidimensional variations is not an easy task. Most of the results for which this theory was initially developed (in particular, restrictions on composition representability of smooth functions [Vit 3]), had been later obtained by different (easier) methods. As a result, today it is not easy to find a presentation of this theory, especially in English. We believe that multi-dimensional variations, as applied to semialgebraic sets, give a very convenient and adequate geometric tool. Indeed, by definition, the i-th variation of $A \subseteq \mathbb{R}^n$, $V_i(A)$, is the average of the number of connected components of the section $A \cap P$ over all the $(n-i)$-dimensional affine planes P in $\mathbb{R}^n$. For A-semialgebraic, the number of connected components of $A \cap P$ is always bounded in terms of the diagram of A (i.e. of the degrees of the defining polynomials and of their set-theoretic formula), and hence to bound variations we need just to estimate the size of various projections of A.

On the other hand, the following basic inequality relates multidimensional variations with metric entropy: For any $A \subseteq \mathbb{R}^n$,

$$M(\epsilon, A) \leqslant C(n) \sum_{i=1}^{n} V_i(A) (\frac{1}{\epsilon})^i \ .$$

We give in Section 1.3 a rather detailed introduction to the theory of variations, in particular providing a complete proof of the above inequality for a general subset $A \subseteq \mathbb{R}^n$ (following [Zer]). We hope that this section, together with Section 5, where variations of semialgebraic sets are studied, can fill to some extent the gap in the literature on this subject.

In Chapter 4 we give some generalities on semialgebraic and tame sets, and prove explicitly (and with explicit bounds) the properties required in the rest of the book: bounds on the number of connected components, "covering theorems" (such as Theorem 1.3 stated below), etc...

Chapter 5 is devoted to variations of semialgebraic and tame sets. We stress the properties which are not true in general: comparison of variations of two tame sets, close to one another in the Hausdorff metric, in particular, of a set and its δ-neighborhood, correlations between variations of the same set, in different dimensions (in general, $V_i(A)$ for different i are "independent"), bounds on the radius of a maximal ball, contained in a δ-neighborhood of a set, etc...

Chapter 6 has a somewhat technical character. To study the behavior of tame and semialgebraic sets under mappings (in the same category), we have to measure properly the size of the first differential of the mappings. Roughly, we use as the "sizes" of a linear mapping (in different dimensions) the semiaxes of the ellipsoid, which is the image of the unit ball under this mapping. This leads to some exterior algebra (sometimes not completely trivial, especially as we want to deal with plane sections and integration).

Chapter 7 contains the main results of this book, as far as the tame (semialgebraic) sets and mappings are concerned. Basically they have the following form: assuming that the size of the differential Df of f is bounded (in one sense or another) on a set A, we estimate variations (and hence metric entropy) of the image $f(A)$ (Theorems 7.1 and 7.2). We deduce from the result the quantitative Morse-Sard theorem in the polynomial case (Theorem 7.5). In particular, we obtain, as a special case, Theorem 1.6 below.

Chapter 8 continues the line of Chapter 7, with somewhat more special results, related to "quantitative transversality" on one side, and to the behavior of mappings on more complicated singularities.

Finally, in Chapter 9 we apply the results of Chapters 7 and 8 to mappings of finite smoothness. The main tool is a Taylor approximation of the mappings; then we use appropriate "semialgebraic" results. Since these results "survive under approximation", it remains to count the total number of Taylor polynomials in the approximation. Consequently, the results have the form of the corresponding "semialgebraic" estimate with a "remainder term", taking into account a finite smoothness (Quantitative Morse-Sard theorem, Theorem 9.2). Considered from the point of view of Differential Analysis and Topology, the results of Chapter 9 give far-reaching improvements and generalizations of the usual Morse-Sard theorem.

In Chapter 10 we give a short overview of some additional applications of the results and methods presented in this book, and of some directions of their further development. The applications include:

– Maxima of smooth families
– Further applications in differential topology
– Smooth Dynamics
– Numerical Analysis

In some details the Semialgebraic Complexity of functions is defined and discussed.

We discuss briefly the following directions of further development:

– Asymptotic critical values
– Morse-Sard theorem in Sobolev spaces
– Real equisingularity
– "C^k-resolution" of semialgebraic sets and mappings
– Bernstein type inequalities for algebraic functions
– Polynomial Control problems
– Quantitative Singularity Theory

1.3 The Motion Planing Problem in Robotics as an Example

Probably the most natural example, where many of the results of this book have immediate and direct interpretation, is provided by various aspects of the so-called "Motion Planning Problem" in Robotics. This example allows one to understand the power of the methods discussed in this book, as well as their limitations. Moreover, we shall try to show in this example what should be done in general in order to transform the enormous analytic power of high order analytic and geometric methods into efficient computational tools.

The problem of motion planning is real, important and difficult, and it may be analyzed and (in principle) solved completely in the framework of Semialgebraic Geometry (although, as we explain below, to deliver its full power, Semialgebraic geometry must be combined with Singularity Theory and with a clever data representation). The main objects of this book, like "effective curves selection" inside semialgebraic sets, covering of semialgebraic sets via polynomial mappings, critical and near-critical points and values of polynomial mappings, – become directly visible in motion planning. The equations arising in the simplest examples are of reasonable degrees, and they can be explicitely solved and analyzed on popular symbolic algebra packages. On the other hand, such practical experiments show immediately the (very narrow) limits of a direct applicability of algebro-geometric methods.

All this justifies, in our view, a rather detailed presentation of the motion planning problem, given below. This presentation follows mostly [Sch-Sha], [Eli-Yom 3], [Tan-Yom] and [Sham-Yom].

Let B be a system comprising a collection of rigid subparts, some of which might be attached to each other at certain joints, while others might move independently. Suppose B has a total of ℓ degrees of freedom, that is, each placement of B can be specified by ℓ real parameters, each representing some relationship (orientation, displacement, etc...) between certain subparts of B. Suppose further that B is free to move in a two- or three-dimensional space amidst a collection of obstacles O whose geometry is known. Typical values of ℓ range from 2 (for a rigid object translating on a planar floor without rotating) to 6 (the typical number of joints for a manipulator arm). The values can also be much larger – for example, when we need to coordinate the motion of several independent systems in the same workspace.

Let $P \subseteq \mathbb{R}^\ell$ denote the space of the parameters of our problem.

The motion-planning problem for B is: given an initial placement Z_1 and a desired target placement Z_2 of B, determine whether there exists a continuous obstacle-avoiding motion of B from Z_1 to Z_2, and, if so, plan such a motion.

Let us consider two examples. The first one is shown in Fig. 1.2. This is a plane "robotic manipulator", consisting of two bars b_1 and b_2. The bar b_1

has its endpoint e_1 fixed at the origin, and the endpoint of b_2 is fixed at the second endpoint e_2 of b_1. Both b_1 and b_2 can rotate freely at e_1 and e_2.

O_1, O_2 and O_3 denote the obstacles, and the initial placement Z_1 and the desired target placement Z_2 are shown on the picture. Taking as free parameters the angles φ_1 and φ_2 shown in Fig. 1.2, we get the space P of parameters as the square $[0, 2\pi] \times [0, 2\pi]$ in $\mathbb{R}^2$ (or rather a torus T^2 – this more accurate topological representation sometimes helps).

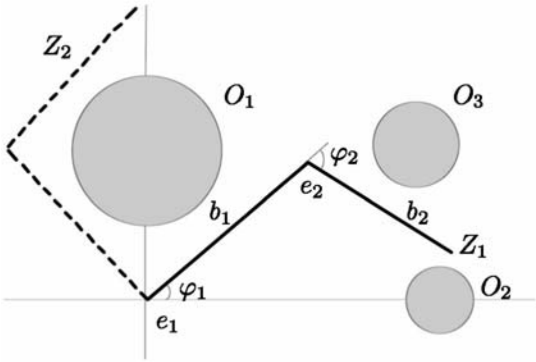

Fig. 1.2.

Another example of a motion-planning problem is represented in Fig. 1.3.

We have to move the plane rectangle B from the initial position Z_1 into the target position Z_2 avoiding the obstacles $O_1, \ldots, O_z$. (One can consider this task as a version of a well-known geometric problem: what is the minimal possible area of a plane domain, inside which we can turn a length 1 needle 180 degrees (see [Tao]))?

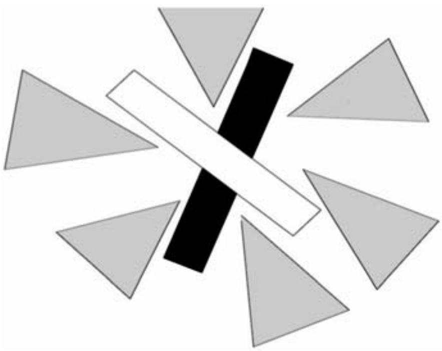

Fig. 1.3.

Here we have 3 degrees of freedom; as the parameters can be taken to be the coordinates (x, y) of the barycenter b of B and the rotation angle φ. Probably a direct examination of these problems will not provide a definite answer (at least for most readers). However, the solution will be greatly simplified if we pass to the so-called "free configuration space" of the problem. Generally the free configuration space of the moving system B denoted FP is the ℓ-dimensional parametric space of all free placements of B (the set of placements of B in which B does not intersect any obstacle). Each point z in FP is a ℓ-tuple giving the values of the parameters controlling the ℓ degrees of freedom of B at the corresponding placement. Clearly, finding a motion from a placement Z_1 represented by $Z_1 \in P$, to Z_2 represented by Z_2, is equivalent to joining Z_1 and Z_2 by a continuous path in FP.

The free configuration space FP of the first problem is shown in Fig. 1.4, together with the initial and target configurations Z_1, Z_2.

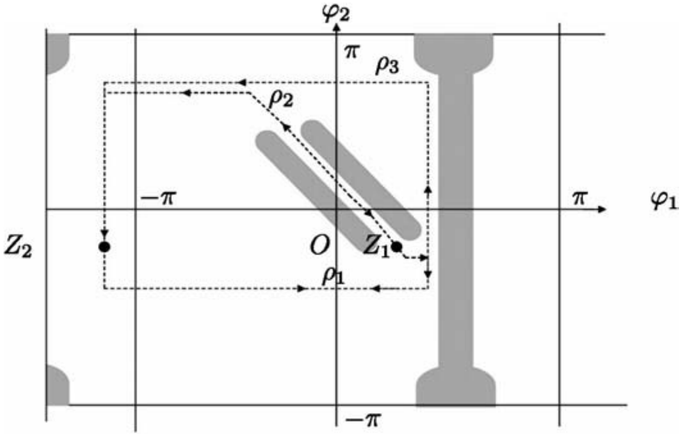

Fig. 1.4.

Now one sees immediately that the solution exists, since Z_1 and Z_2 belong to the same connected component of FP. Three of the "control (or configuration) trajectories" joining Z_1 and Z_2 are shown in Fig. 1.4, and the corresponding evolution of the manipulator is given in Fig. 1.5.

This figure shows one of the three solutions, represented on Fig. 1.4, namely ρ_1. It consists of 4 rotations (3 of them are consecutive, illustrated by arcs 1, $1'$, 2 and 3).

Thus the main difficulty in solving the motion planning problem consists in the construction of the free configuration space. This construction is nontrivial already in the first example considered. In the second example the free configuration space FP is fairly complicated: it looks like a spiralled worm-

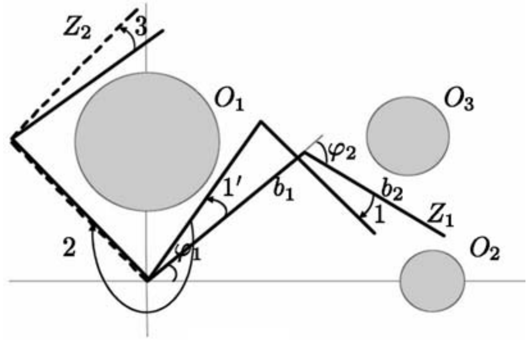

Fig. 1.5.

hole in the three-dimensional cube, and we do not show it here. However, the solution turns out to exist, and is shown in Fig. 1.6.

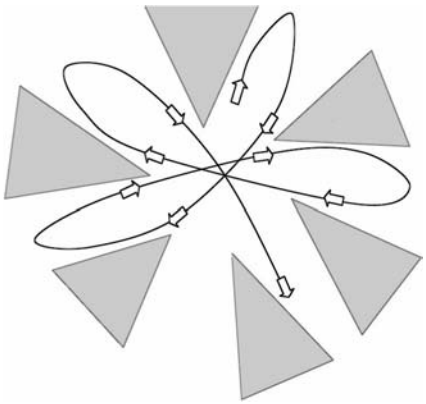

Fig. 1.6.

Now the basic fact is that if each part of the system B and each obstacle O are *semialgebraic* (i.e., representable by a finite number of polynomial equations, inequalities and set-theoretic operations), then the free configuration space FP is *semialgebraic*, and can be computed effectively from B and O.

There exists also an effective procedure to decide whether two given points belong to the same connected component of a given semialgebraic set. Consequently, for semialgebraic data (which is a very natural assumption) the motion planning problem can be effectively solved. See [Sch-Sha] for details.

Important remark. "Effectively" does not mean "efficiently"! The complexity of the algorithms, based on the direct approach as above and using symbolic computations with semialgebraic sets, is known to be extremely high. It becomes prohibitive in practical applications even for rather simple motion planning tasks.

The reason is that the *maximal possible* complexity of semialgebraic sets of a given degree is indeed very high – it grows at least as the degree to the power of the dimension. For example, let us take as the complexity measure the number of connected components of a semialgebraic set (this characteristic is intensively used below). Easy examples (also given below) show that this number can be as high as prescribed for very simple defining equations and inequalities.

A straightforward symbolic computation must take into account the "worst case", so it has to process each connected component separately. Inside this processing further ramifications appear, with the same number of choices as above, and so on. For the degrees of order 10 and the dimension 6, like in simplest practical applications, all this together is too much.

However an adaptive approach, which follows a natural "hierarchy of singularities" in the problem, reduces dramatically the complexity of computations. Indeed, in most cases we can expect our equations to be *non-degenerate*, in an appropriate sense (this is a virtue of the Morse-Sard theorem !). But a zero set of a non-degenerate system of equations is a regular manifold, and *locally* it has exactly one component.

Next after the non-degenerate case, we have to consider degenerations of "codimension one" in the sense of Singularity Theory (see [Arn-Var-Gus], [Boa], [Gol-Gui], ...). These degenerations are much less probable than the regular situation, but their explicit consideration is important, especially taking into account, that we have to treat not exactly singular, but rather "near-singular" cases. The *local* complexity of the solutions for systems of equations and inequalities with a degeneration of codimension one is still rather small.

Next we continue to the codimension two singularities, and so on. In general, we follow the "hierarchy of singularities" in our specific problem (as it is explained above, essentially this problem is to describe the free configuration space FP of the motion). Some initial steps in this hierarchy are described in [Eli-Yom 1-3].

There are several basic problems in this approach: first, what is the "locality size" which guarantees the expected low complexity of the small codimension singularities ? Second, where to stop in the hierarchy of singularities ? Third, how to treat "near-singular" situations ?

We hope that the answer to these questions can be provided by the future "Quantitative Singularity Theory". The "Quantitative Sard Theorem" and the "Quantitative Transversality" considered in this book form the first steps in this direction. See Section 10.3.7.

However, one can develop practical algorithms, based on a high order approximation and on hierarchy of singularities, before the theoretical foundations have been completed. Simple empirical procedures in most cases provide a reasonable answer to the problems above. As far as the motion planning is concerned, such an algorithm has been developed and initially tested (see [Eli-Yom 1-3]). It is based on an approximation of the free space FP on a certain grid, while at each gridpoint a semialgebraic representation as above is used. However, the hierarchy of the allowed degenerations at each gridpoint is restricted in such a way that the overall complexity of computations remains strictly bounded. More degenerate situations are treated (within the prescribed accuracy) simply by an appropriate subdivision of the grid. The efficiency of this algorithm confirms (in very limited cases, as of today) our theoretical expectations.

Of course, the discussion above is applicable not only to the Motion Planning problem. A combination of a high order approximation of the data, its further structuring and organization along the hierarchy of singularities in the problem, and its analytic processing, present a powerful computational approach in many important problems. The "Quantitative Singularity Theory" will form a theoretical basis of this approach. We discuss it in more detail (including, in particular, some specific implementations) in Section 10.1.4 of Chapter 10 below.

This is the place to say that we *do claim* that various theorems in semialgebraic geometry given below are (or may be) useful in motion planning and other applications, *but only when combined with a clever data approximation, with an analysis of the hierarchy of singularities of the problem, and with an appropriate scheme of numerical computations.* It is not the purpose of this book to develop these methods (see however [Eli-Yom 1-5], [Bri-Yom 6], [Bri-Eli-Yom], [Bic-Yom], [Wie-Yom], [Yom 24], [Y-E-B-S] and Section 10.1.4 of Chapter 10 below). Consequently, all the examples of "applications" given in this book are pure illustrations of mathematical results, and any attempt of their *straightforward* application in computations is in our opinion completely hopeless.

After this warning we return to the description of our approach to the example of a motion planning problem.

There are two general and well-known principles in real semialgebraic geometry (although their specific implementation can be rather nontrivial or impossible).

The first says that any reasonable operation with semialgebraic data leads to a semialgebraic "output", with a combinatorial complexity (i.e. the degrees of the polynomials and the set theoretic formula in a representation – below we call these data *the diagram* of the set) depending only on that of the input.

The second principle claims that any reasonable metric characteristic of a semialgebraic set of a given combinatorial complexity inside a ball of a given radius, can be bounded in terms of the complexity and the radius of the ball. A good part of this book is devoted to various specific manifestations of these general principles.

The simplest, but rather useful, example where both these principles work, is the following result (see Theorem 4.12, Chapter 4 below; see also [Den-Kur], [Har], [Kur], [Tei 1], [Yom 1,5], and [D'Ac-Kur] for an explicit value of the bound $K(D)$ of Theorem 1.1).

Theorem 1.1. *Let $A \subseteq \mathbb{R}^n$ be a semialgebraic set with a given diagram D. Then for the ball B_R of radius R, centered at the origin of $\mathbb{R}^n$, any two points z_1 and z_2, belonging to the same connected component of $A \cap B_R$, can be joined inside $A \cap B_R$ by a semialgebraic curve ℓ, such that the diagram of ℓ depends only on D, and the length of ℓ does not exceed $K(D) \cdot R$, with the constant $K(D)$ depending only on D.*

Corollary 1.2. *If a solution to a motion planning problem with semialgebraic data exists, it can be given by a semialgebraic path in the parameter space, whose complexity and length depend only on the combinatorial complexity of the data.*

In this book we mostly study not just semialgebraic sets, but rather their behavior under polynomial (or, more generally, semialgebraic – i.e. those with a semialgebraic graph) mappings. In the context of a motion planning problem, an important and highly nontrivial such mapping appears very naturally. This is the so-called *"kinematic mapping"* φ of the manipulator; it associates to any given values of the control parameters the position of the "tooling device" (or of a prescribed point, or of any prescribed part of the manipulator).

The kinematic mapping of a manipulator (as described above) is always semialgebraic, assuming that the controls are properly parametrized. Now, the main practical problem that appears in the programming of industrial robots, is the so-called "inverse kinematic problem":

For a given trajectory s of the tooling device in the workspace, find a corresponding trajectory σ in the space of controls (i.e. such that $s = \varphi(\sigma)$, where φ, as above, is the kinematic mapping). The initial motion-planning is a part of the inverse kinematic problem, since the manipulator in the process of motion is naturally assumed to avoid collisions. The inverse kinematic problem is usually redundant, since the number ℓ of the degrees of freedom of the manipulator is normally chosen to be bigger than the dimension of the configuration space of the tooling device (to provide a flexibility in programming).

There are various approaches to the inverse kinematic problem, mostly dealing with one or another way to eliminate the above-mentionned redundancy. One of these approaches is given in [Sha-Yom], together with some literature on the subject. Usually the redundancy is eliminated by introducing a certain distribution in the parameter space, transversal to the fibers of the kinematic mapping (and of a complementary dimension). Any motion of the tooling device can be now lifted to the parameter space, in a locally unique way.

Without additional restrictions this lifting is not semialgebraic. Moreover, it is not easy to combine this lifting with the requirement of the collision avoidance. We consider a combination of the redundancy elimination with the semialgebraic motion planning (as shortly presented below) a very important problem.

The following results provide a semialgebraic solution of bounded complexity to the inverse kinematic problem:

Theorem 1.3. (see Theorem 4.10 below) *Let $f : A \to B$ be a semialgebraic mapping between two semialgebraic sets. Then for any semialgebraic curve s in $f(A) \subseteq B$ there exists a semialgebraic curve σ in A, such that $f(\sigma) = s$. The diagram of σ depends only on the diagrams of f, A, B and s.*

Corollary 1.4. *For any semialgebraic trajectory s of the tooling device in its workspace, there exists a semialgebraic control trajectory σ, such that $s = \varphi(\sigma)$. If a solution without collisions exists, σ can be chosen to be noncolliding. The combinatorial complexity and the length of σ are bounded in terms of the complexity of the data.*

In fact, Theorem 1.3 is a special case of the following general "covering theorem" (see Theorem 4.10, Chapter 4 below):

Theorem 1.5. *For $f : A \to B$ as above, and for S a semialgebraic set in $f(A) \subseteq B$, there exists a semialgebraic set $\Sigma \subseteq A$, with $\dim\Sigma = \dim S$, such that $f(\Sigma) = S$. The diagram of Σ (and hence the bounds on its geometry) depends only on the diagrams of f, A, B and S.*

Obviously, this result has a natural interpretation in terms of control of a manipulator, whose tooling device has to cover a surface S in the workspace.

The next topic, which is central for this book, is the geometry of critical and near-critical values of semialgebraic (and later smooth) mappings. The near-critical points of f are those where the differential Df is "almost singular" (in an appropriate sense – see below). The near critical values are values of f at the near-critical points.

As applied to the kinematic mappings φ of a manipulator, these notions become quite relevant: near-critical points of φ are those, where some of the controls do not affect the position of the tooling device. Near critical values are the positions which we can get with such "bad" control. There are obvious reasons (especially as the dynamics of the motion is incorporated) to avoid such positions.

First of all, let us see what the critical set and the critical image look like in the first example above (with the "tooling device" just the endpoint of the manipulator). One can easily see that critical positions of the manipulator correspond exactly to $\varphi_2 = 0$ and $\varphi_2 = \pi$ (see Fig. 1.7).

In both these configurations the controls φ_1 and φ_2 do not affect the distance of the endpoint from the origin. For φ_2 near 0 or near π an easy

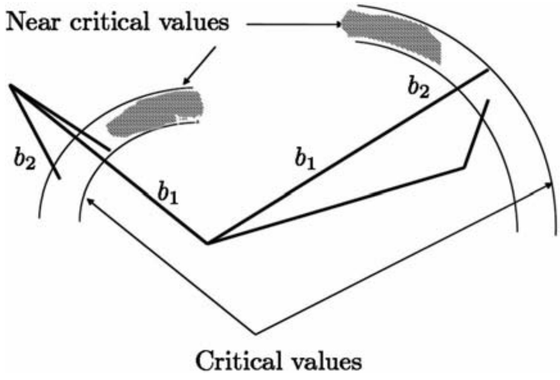

Fig. 1.7.

computation gives for the distance r of the endpoint from the origin (assuming $|b_1| > |b_2|$):

$$r \sim |b_1| + |b_2| - c\varphi_2^2 \ ,$$

or

$$r \sim |b_1| - |b_2| - c'(\pi - \varphi_2)^2 \ .$$

Hence the set of near-critical points of the kinematic mappings, where $|\frac{\partial r}{\partial \varphi_2}| \leqslant \gamma \ (\frac{\partial r}{\partial \varphi_1} \equiv 0)$, consists of two strips, $|\varphi_2| \leqslant \frac{\gamma}{2c}$ and $|\pi - \varphi_2| \leqslant \frac{\gamma}{2c'}$. The corresponding set of critical values consists of two rings,

$$|b_1| + |b_2| - \frac{\gamma^2}{4c} \leqslant r \leqslant |b_1| + |b_2|$$

and

$$|b_1| - |b_2| \leqslant r \leqslant |b_1| - |b_2| + \frac{\gamma^2}{4c'}$$

(see Fig. 1.7). Notice that the singularities of the kinematic mappings ϕ in this example are of the "fold" type, according to the Whitney classification (see [Whi 3], [Boa], [Gol-Gui]). After an appropriate coordinate change it can be locally written in the form

$$\begin{cases} y_1 = x_1 \\ y_2 = x_2^2 \ . \end{cases}$$

(Up to a distorsion of the change of coordinates, our definitions of near-critical points and values are invariant, so we can perform computations using Whitney normal forms).

Let us consider briefly one additional example of a manipulator. It consists of a bar b, which can slide in a frame F, which in turn slides along an ellipse E (see Fig. 1.8). Thus b always remains orthogonal to E, but can slide freely in this direction.

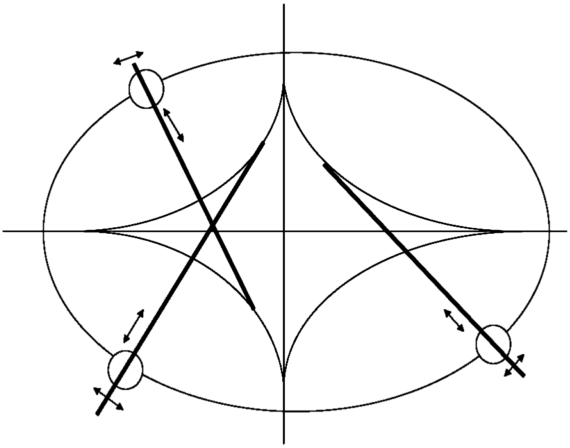

Fig. 1.8.

The "tooling device" once more is the endpoint e of b, and the kinematic mapping associates to the control parameters (position of the frame F on the ellipse E and the position of the bar b inside F) the endpoint e on the plane. Direct computations here are somewhat more involved. However, this example is well studied in Singularity Theory (see [Gol-Gui]). The set of critical values here is the curve Γ, shown in Fig. 1.8. It consists of the endpoint positions, as the distance of e from F is equal to the curvature radius of the ellipse E at F.

All the points on this curve Γ are folds, except the four vertices, at which the kinematic mapping has a "cusp" singularity, according to Whitney's classification ([Whi 3]). In a properly chosen system of local coordinates it can be written as

$$\begin{cases} y_1 = x_1 \\ y_2 = x_2^3 - x_1 x_2 \ . \end{cases}$$

As the γ-near-critical values are concerned, so here they form a strip around Γ of width of order γ^2 near the fold-points. At cusps the situation is more complicated.

In all these examples we see that the γ-critical values of ϕ form a "small" set: as γ tends to zero, the area of $\Delta(\phi)$ tends to zero. This is a general fact, and in this book we study the "size" of near-critical values in detail. However, the "area" is not convenient to measure this size. Instead, we bound the *metric entropy* of near critical values. For a compact X, the ϵ-entropy $M(\epsilon, X)$ is the minimal number of ϵ-balls that cover X. Let us give a couple of reasons why metric entropy is better for our purposes. Other reasons are given in Section 1.3 and scattered all over the book.

To keep this introduction to a reasonable size, we do not give here formal definitions of the metric entropy, near-critical values etc., but rather short explanations of these notions. Accurate definitions of all the notions related to metric entropy are given in Chapter 2. For near critical points and values this is done in Chapters 6 and 7.

First of all, in the motion planning context we would like to avoid not only the set of critical values, but a certain neighborhood. The fact that "area" of a set is small does not imply restrictions on its δ-neighborhood (it can consist of a dense collection of curves of very small length, etc...). Ultimately, the set of rational points in $\mathbb{R}^n$ has Lebesgue measure 0, while its δ-neighborhood for any $\delta > 0$ is all the space $\mathbb{R}^n$.

On the contrary, metric entropy is stable with respect to taking neighborhoods: if certain balls of radius ϵ cover X, the balls of radius $\epsilon + \delta$, centred at the same points, cover the δ-neighborhood X_δ.

Another reason is that in many cases we would like to restrict ourselves to a certain grid and to find "good points" in this grid. Once more, a small Lebesgue measure of a "bad" set does not guarantee that we can find good points in any prescribed grid (take once more all the rational points).

It is easy to show (see Chapter 2 below) that if metric entropy of a bad set is small, then in any sufficiently dense grid most of the points are good. Thus metric entropy is a stronger and more convenient geometric invariant for our purposes.

The third reason to work with it is that it is also much more natural for the problems treated in this book. For semialgebraic sets and mappings the behavior of the metric entropy reflects in a very transparent way their geometry in different dimensions. For smooth mappings, robustness of the entropy allows for a direct application of semialgebraic results via polynomial approximation. Moreover, for mappings of finite smoothness, metric entropy (and a related notion of block or entropy-dimension) turns out to be *the correct geometric invariant:* sets of critical values of C^k mappings can be characterized in such terms.

Let us make a more accurate definition of γ-critical sets and values: *x is a γ-critical point of f if the differential $Df_{(x)}$ maps the unit ball into an ellipsoid with the smallest semiaxis $\leqslant \gamma$. The set of γ-critical points of f is denoted by $\Sigma(\gamma, f)$, and the set of γ-critical values of f is $\Delta(\gamma, f) = f(\Sigma(\gamma, f))$.*

Now we are ready to state one of our main results:

Theorem 1.6. (Theorem 7.5, Chapter 7) *Let $f : A \to \mathbb{R}^m$ be a semialgebraic mapping of two semialgebraic sets, with $\operatorname{diam}(A) = R$ and $rank(Df_{|A}) \leqslant q$ and the norm of $Df(x)$ bounded by 1 for any x in A. Then for each $\gamma \geq 0$, $\epsilon \geq 0$. the ϵ-entropy of the set of γ-critical values $\Delta(\gamma, f)$ satisfies*

$$M(\epsilon, \Delta(\gamma, f)) \leqslant C_0 + C_1 (\frac{R}{\epsilon})^{q-1} + C_2 \gamma (\frac{R}{\epsilon})^q ,$$

where the constants C_0, C_1, C_2 depend only on the diagrams of f and A. In particular, the q-dimensional Lebesgue measure of $\Delta(\gamma, f)$ does not exceed $C_3 R^q \gamma$.

This result can be interpreted in terms of the kinematic mapping as follows:

Corollary 1.7. *For a kinematic mapping φ of a manipulator, the set of γ-bad positions $\Delta(\gamma, \varphi)$ has a volume at most $C\gamma$, with the constant C depending only on the diagrams and the size of the manipulator. Moreover, for γ small, most of the points in any sufficiently dense grid in the workspace of the manipulator are "γ-good".*

Now we can combine Corollary 1.4 and Corollary 1.7 and produce a semialgebraic solution to motion planning and inverse kinematic problems, which in addition avoids γ-bad positions for a prescribed sufficiently small γ. Moreover, in principle, this solution can be constructed explicitly, using the methods of this book. A part of the way from here to the efficient motion planning algorithm has been discribed above.

1.4 A Proof of the Morse-Sard Theorem in the Simplest Case

In this last section of the introduction we give an accurate proof of the simplest version of the generalized Sard theorem. It deals with a C^k-function from a certain ball in R^n to R (and not with a mapping into a higher-dimensional space).

This proof illustrates in a simple and transparent form (and without technicalities, unavoidable in a general setting) a good part of the ideas and methods developed below.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^1-function. For $\gamma \geq 0$, let $\Sigma(f, \gamma) = \{x / \|\text{grad} f(x)\| \leqslant \gamma\}$. Let $B^n_r \subseteq \mathbb{R}^n$ be some ball of radius r. We denote $\Sigma(f, \gamma) \cap B^n_r$ by $\Sigma(f, \gamma, r)$ and $f(\Sigma(f, \gamma, r)) \subseteq \mathbb{R}$ by $\Delta(f, \gamma, r)$. $\Sigma(f, \gamma, r)$ and $\Delta(f, \gamma, r)$ are the set of γ-critical points and γ-critical values of f on B^n_r, respectively. For $\gamma = 0$ we get the usual critical points and values.

First of all, we consider the case of f a polynomial.

Theorem 1.8. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a polynomial of degree d. Then for any $\gamma \geq 0$ the set $\Delta(f, \gamma, r)$ can be covered by $N(n, d)$ intervals of length γr. The constant $N(n, d)$ here depends only on n and d.*

Proof. The set $\Sigma = \Sigma(f, \gamma, r)$ of γ-critical points of f inside the closed ball B^n_r is a semialgebraic set (defined by $\|\text{grad} f\|^2 \leqslant \gamma^2$). Hence the number of connected components of Σ does not exceed $N_1(n, d)$. For each connected component Σ_i of Σ its image $f(\Sigma_i)$ is an interval Δ_i.

Let us take two points x^1_i and x^2_i in Σ_i, such that $y^1_i = f(x^1_i)$ and $y_i = f(x^2_i)$ are the end points of the interval Δ_i. By Theorem 1.1 above, x^1_i and x^2_i can be joined inside Σ_i by a (semialgebraic) curve s of length at most $K(n, d) \cdot r$. Now $|\Delta_i| = |y^2_i - y^1_i| = |\int_s \text{grad} f \cdot ds|$. But the norm of $\text{grad} f(x)$

for x in Σ does not exceed γ, and hence the last integral is bounded by γ.length$(s) \leqslant K(n,d) \cdot \gamma \cdot r$. Therefore, each Δ_i can be covered by at most $K(n,d)$ intervals of length γr, and $\Delta = \cup \Delta_i$ can be covered by $N_1(n,d) \cdot K(n,d) = N(n,d)$ such intervals. □

In the proof proposed above, a possible bound for $N_1(n,d)$ can be produced, via Bézout's Theorem. Following Theorem 4.9, Chapter 4, we have: $N_1(n,d) \leqslant d(2d-1)^{n-1}$ (because the polynomial $||grad(f)||^2 - \gamma^2$ is of degree $2(d-1)$). The difficult point concerns in fact the obtaining of a bound for $K(n,d)$.

We can modify a little bit the proof of Theorem 1.8 in the following way. We consider s a semialgebraic set of Σ of dimension $\leqslant 1$, not necessarily connected, (s being 0-dimensional in the case Δ is itself 0-dimensional), such that $f(s) = \Delta$. Such a set s is given by Theorem 4.10, and in this construction the diagram of s depends only on n and d. Let us consider now a connected component Δ_i of Δ which is not a point. We can find a semialgebraic set $s_i \subset s$ such that $f(s_i) = \Delta_i$ and such that the s_j are disjoint. The same arguments as in the proof of Theorem 1.8 show that the total length of $\bigcup_i \Delta_i$ is less than $\gamma \cdot \sum_i length(s_i) \leqslant \gamma \cdot length(s)$. But we have by Lemma 4.13: $length(s) \leqslant K'(n,d) \cdot r$. Considering that the number of connected components of Δ which are points is less than the number of connected components of Σ, which, in turn, is less than $d(2d-1)^{n-1}$, we obtain that we can cover Δ by $d(2d-1)^{n-1} + K'(n,d)$ intervals of length $\gamma . r$.

This bound is better than the one given in the proof of Theorem 1.8, because the product has been replaced by a sum. Nevertheless, the point is again to evaluate $K'(n,d)$.

Remark. In [D'Ac-Kur], it is shown that a possible bound for $K(n,d)$ is:

$$2 \cdot c(n,1) \cdot ((6d-4)^{n-1} + 2(6d-3)^{n-2}),$$

where the constant $c(n,1) = \Gamma(1/2)\Gamma((n+1)/2)/\Gamma(n/2)$ is introduced in Chapter 3 (Γ being the Euler function). In addition in [D'Ac-Kur] ([D'Ac-Kur], Theorem 10.1), it is shown that a possible bound for the constant $N(n,d)$ is:

$$d(2d-1)^{n-1} + 2 \cdot c(n,1) \cdot ((6d-4)^{n-1} + 2(6d-3)^{n-2}.$$

Now the property given by Theorem 1.8 is compatible with approximations. Let $g : B_r^n \to \mathbb{R}$ be a k times differentiable function, and let P be the Taylor polynomial of degree $k-1$ of g at the center of B_r^n. We have

$$\max_{x \in B_r^n} |g(x) - P(x)| \leqslant R_k(g)$$

$$\max_{x \in B_r^n} \|dg(x) - dP(x)\| \leqslant \frac{k}{r} R_k(g) \ ,$$

where $R_k(g) = \frac{1}{k!} \max\|d^k g\| \cdot r^k$ is the remainder term in the Taylor formula.

Hence, the critical points of g are at most γ_0-critical for P, where $\gamma_0 = \frac{k}{r} R_k(g)$, i.e. $\Sigma(g,0) \subseteq \Sigma(P, \gamma_0, r)$. Hence $\Delta(g,0) = g(\Sigma(g,0,r)) \subseteq g(\Sigma(P,\gamma_0,r))$. Finally, since $|g - P| \leqslant R_k(g)$, $g(\Sigma(P,\gamma_0,r))$ is contained in a $R_k(g)$-neighborhood of $P(\Sigma, \gamma_0, r)) = \Delta(P, \gamma_0, r)$.

Now by Theorem 1.8, $\Delta(P, \gamma_0, r)$ can be covered by at most $N(n, k-1)$ intervals of length $\gamma_0 r = k \cdot R_k(g)$, and hence by $k \cdot N(n, k-1)$ intervals of length $R_k(g)$. Thus the $R_k(g)$-neighborhood of $\Delta(P, \gamma_0, r)$, and hence $\Delta(g, 0, r)$ can be covered by the same number of intervals of length $3R_k(g)$, or by triple the number of $R_k(g)$-intervals. We proved the following result:

Theorem 1.9. *Let $g : B_r^n \to \mathbb{R}$ be a k times differentiable function. Then the set $\Delta(g, 0, r)$ of the critical values of g on the ball B_r^n can be covered by at most $N_2(n,k)$ intervals of length $R_k(g)$, where*

$$N_2(n,k) = 3 \cdot k \cdot N(n, k-1) \ \text{ depends only on } n \text{ and } k \ .$$

□

This result can be considered as a "Taylor formula" for the property of polynomials, given by Theorem 1.8.

Let us continue a little bit in this direction, considering the following question: for an arbitrary $\epsilon > 0$, how many intervals of length ϵ do we need to cover $\Delta(g, 0, r)$? To answer this question we first find r' such that the remainder term of g on any ball of radius r' is at most ϵ:

$$\frac{1}{k!} \max\|d^k g\| \cdot (r')^k = R_k(g) \cdot (\frac{r'}{r})^k = \epsilon \ \text{, i.e. } r' = r \cdot [\frac{\epsilon}{R_k(g)}]^{1/k} \ .$$

Now we cover B_r^n by subballs of radius r'. We need at most $C(n)(\frac{r}{r'})^n = C(n)[\frac{R_k(g)}{\epsilon}]^{n/k}$ such balls.

Theorem 1.9 guarantees that critical values of g on each small ball can be covered by at most $N_2(n,k)$ ϵ-intervals, and to cover all the critical values of g we need therefore at most $N_2(n,k) \cdot C(n) \cdot [\frac{R_k(g)}{\epsilon}]^{n/k}$ such intervals. We proved:

Theorem 1.10. *For g as above and for any ϵ, $0 < \epsilon \leqslant R_k(g)$, the set of critical values of g can be covered by $N_3(n,k)[\frac{R_k(g)}{\epsilon}]^{n/k}$ intervals of length ϵ.* □

Corollary 1.11. (Morse-Sard Theorem) If $g \in C^k$ with $k > n$, then the measure of the critical values of g is zero.

Proof. For any ϵ the measure of the critical values of g is bounded by ϵ times the number of ϵ-intervals covering our set. Hence

$$m(\Delta) \leqslant \lim_{\epsilon \to 0} \epsilon \cdot C \cdot (\frac{1}{\epsilon})^{n/k} = \lim_{\epsilon \to 0} C \epsilon^{1-n/k} = 0 \text{ for } \frac{n}{k} < 1 \ .$$

□

Concluding this section, let us discuss again the problem of finding the explicit (and optimal) constants in the inequalities above, and throughout the book. In principle, an application of the methods of Chapter 4 below allows one to get an explicit estimate for each of the "algebraic" constants. Let us illustrate our principle with the constant $N(n,d)$ of Theorem 1.8, although the following lines do not give a general proof, but rather a description of a general philosophy (see [D'Ac-Kur] for a rigourous proof of the obtaining of this bound):
we approximate the boundary of $\Sigma(f,\gamma,r)$ by a smooth semialgebraic set $Z = \{p = ||grad(f)||^2 - \gamma^2 = \delta\}$. Then we fix a generic linear form ℓ on $\mathbb{R}^n$ and define the curve S of all the critical points of ℓ on $Z \cap \{f = t\}$. Assuming that $Z \cap \{f = t\}$ is smooth and compact, we obtain explicit equations for S. Performing a linear change of variables in $\mathbb{R}^n$, we can assume that $\ell(x) = x_n$. Hence S consists of the points x in Z where the vector $(0, \cdots, 0, 1)$ is a linear combination of $grad(p)_{(x)}$ and $grad(f)_{(x)}$. This condition is given by the equations:

$$\frac{\partial f}{\partial x_1}\frac{\partial p}{\partial x_i} - \frac{\partial p}{\partial x_1}\frac{\partial f}{\partial x_i} = 0, \qquad i = 2, \cdots, n-1,$$

each of degree $(2d-3)(d-1)$, the equation defining Z being of degree $2d-2$.

We find, using Corollary 4.9, that the number of connected components $k(n,d)$ of S in generic hyperplane section is less than:

$$1/2(2d-2+(2d-3)(d-1)(n-2)+2)(2d-2+(2d-3)(d-1)(n-2)+1)^{n-2}$$

$$= 1/2(2d+(2d-3)(d-1)(n-2))(2d+(2d-3)(d-1)(n-2)-1)^{n-2}$$

Finally, following the proof of Theorem 1.8, and considering that $Z = Z_\delta$ is a uniform approximation of $\Sigma(f,\gamma,r)$ we obtain:

$$N(n,d) \leqslant d(2d-1)^{n-1} + c(n,1) \cdot k(n,d)$$

However, in most of the results in this book we do not give such explicit estimates. The reason on one side is that their producing is rather lengthy. On the other side, in all the applications in Smooth Analysis, considered in this book, the explicit estimates of the constants are not crucial. So just presenting explicit estimates of numerous constants below would not be especially instructive.

On the other hand, in applications in Numerical Analysis, discussed in Section 10.1.4, Chapter 10, accurate estimates of the algebraic constants are crucial. Producing such *accurate* estimates is not an easy task. In some cases even the asymptotics is not known. We consider this as an important open problem. On these questions, we can at least refer to [And-Brö-Rui], [Hei-Rec-Roy], [Hei-Roy-Sol 1,2,3,4], [Ren 1, 2, 3] (and of course to references contained in these papers), where complexity and effectiveness for classical problems (quantifier elimination, piano's mover problem, constructing Whithney stratifications...) are given and discussed.

2 Entropy

Abstract. We define in this chapter the entropy dimension of a set. We also recall the definition of Hausdorff measures and we compare the entropy and the Hausdorff dimensions, showing that the first one is bigger than the second one.

In Chapter 1 we have already considered the number of ϵ-intervals one needs to cover a given set. This metric invariant is very convenient in our approximation approach, in particular, because of its "stability": knowing this number for a set, we easily compute it for the ϵ-neighborhood of this set. Of course, the usual measure does not share this property. On the other hand, it turns out that in terms of this number we can formulate rather delicate properties, relevant in the study of smooth functions.

Definition 2.1. Let X be a metric space, $A \subset X$ a relatively compact subset. For any $\epsilon > 0$, denote by $M(\epsilon, A)$ the minimal number of closed balls of radius ϵ in X, covering A (note that this number does exist because A is relatively compact). The real number $H_\epsilon(A) = \log_2 M(\epsilon, A)$ is called the *ϵ-entropy* of the set A.

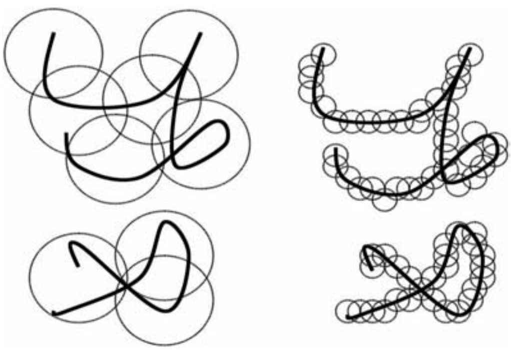

Fig. 2.1.

This terminology, introduced in [Kol-Tih], reflects the fact that $H_\epsilon(A)$ is the amount of information we need to describe a point in A with the accuracy

Y. Yomdin and G. Comte: LNM 1834, pp. 23–32, 2004.

ϵ (or digitally memorize A with accuracy ϵ). Thus, the behavior of $M(\epsilon, A)$ for various ϵ reflects not only "massiveness" of the set A, but also its geometry in X. A detailed study of metric entropy can be found in [Kol-Tih]. See also [Lor 2] and [Tri].

A subset Z of X is called an *ϵ-net* for A, if for any $y \in A$ there is $z \in Z$ with $\mathrm{d}(z, y) \leqslant \epsilon$. Consequently, $M(\epsilon, A)$ coincides with the minimal number of elements in ϵ-nets for A (these elements being the centres of covering balls).

On the other hand, call the set $W \subset A$ *ϵ-separated*, if for any distinct $w_1,\ w_2 \in W$, $\mathrm{d}(w_1, w_2) > \epsilon$. Denoting by $M'(\epsilon, A)$ the maximal number of elements in ϵ-separated subsets of A, we have easily: $M'(2\epsilon, A) \leqslant M(\epsilon, A) \leqslant M'(\epsilon, A)$. Indeed, if $x_1, \dots x_{M'(2\epsilon,A)}$ are in a 2ϵ-separated subset of A, every ϵ-ball containing x_j does not contain x_k, for $k \neq j$ and thus one needs at least $M'(2\epsilon, A)$ ϵ-balls to cover A, and if $x_1, \dots, x_{M'(\epsilon,A)}$ are ϵ-separated points in A, by definition of $M'(\epsilon, A)$, there exists no $y \in A$ such that $\mathrm{d}(y, x_j) > \epsilon$ for all $j \in \{1, \dots, M'(\epsilon, A)\}$, thus the balls $B(x_j, \epsilon)$ cover A, showing that $M(\epsilon, A)$ is less than $M'(\epsilon, A)$.

The proof of the following properties of $M(\epsilon, A)$ is immediate.

(1) $A \subset B \Rightarrow M(\epsilon, A) \leqslant M(\epsilon, B)$.

(2) $M(\epsilon, \bar{A}) = M(\epsilon, A)$, $\bar{A}$ the closure of A.

(3) $M(\epsilon_1, A) \geq M(\epsilon_2, A)$, for $\epsilon_1 \leqslant \epsilon_2$.

(4) $M(\epsilon, A \cup B) \leqslant M(\epsilon, A) + M(\epsilon, B)$, and if $\inf_{x \in A, y \in B} d(x, y) = \delta > 0$, then for $\epsilon < \delta/2$, $M(\epsilon, A \cup B) = M(\epsilon, A) + M(\epsilon, B)$.

(5) Let A_η denote the η-neighborhood of A. If for a given $\epsilon \leqslant \eta$, $\mu(\epsilon, \eta) \in \mathbb{N} \cup \{\infty\}$ denotes the supremun of $M(\epsilon, B)$ over all the η-balls B in X, then:

$$M(\epsilon, A_\eta) \leqslant \mu(\epsilon, 2\eta).M(\eta, A).$$

Indeed if some η-balls cover A, the 2η-balls centred at the same points cover A_η. Of course we have $\mu(2\epsilon, 2\epsilon) = 1$, and thus in particular:

$$M(2\epsilon, A_\epsilon) \leqslant M(\epsilon, A).$$

(6) Define the *Hausdorff distance* between $A_1, A_2 \subset X$ as follows:

$$d_{\mathcal{H}}(A_1, A_2) = \max\big(\sup_{x \in A_1} \mathrm{d}(x, A_2); \sup_{y \in A_2} \mathrm{d}(y, A_1)\big).$$

Then, if $\mathrm{d}(A_1, A_2) \leqslant \epsilon$, we have $A_1 \subset A_{2,\epsilon}$, $A_2 \subset A_{1,\epsilon}$, and by the last inequality we obtain:

$$M(2\epsilon, A_1) \leqslant M(\epsilon, A_2) \quad \text{and}$$

$$M(2\epsilon, A_2) \leqslant M(\epsilon, A_1).$$

Another "effectivity" property of ϵ-entropy is the following:

(7) For any 2ϵ-separated set W in X, the intersection $W \cap A$ contains at most $M(\epsilon, A)$ points. Indeed if we consider a covering of A by $M(\epsilon, A)$ ϵ-balls, and N points in A with $N > M(\epsilon, A)$, necessarily two of these points are in the same ϵ-ball, thus these points are not in a 2ϵ-separated set of A.
This means that knowing $M(\epsilon, A)$ to be small, we can find effectively points not in A. In fact most of the points in any regular net in $\mathbb{R}^n$ will be out of A, for $A \subset \mathbb{R}^n$ having small ϵ-entropy. Also this property is not shared by the usual measure, e.g. the measure of the rational points is zero, but in numerical analysis we can work only with rational points.

The behavior of metric entropy under mappings with known metric properties can be easily described:

(8) Let $f : X \to Y$ be an Hölderian mapping, i.e. such that there exist two reals $K > 0$ and α such that, for all $x, y \in X$,

$$\mathrm{d}_Y(f(x); f(y)) \leqslant K\big(\mathrm{d}_X(x; y)\big)^\alpha.$$

Then the image by f of a covering of $A \subset X$ by sets with diameter less than ϵ is a covering of $f(A) \subset Y$ by sets with diameter less than $K\epsilon^\alpha$. Consequently, for $A \subset X$ and any $\epsilon > 0$,

$$M(K\epsilon^\alpha, f(A)) \leqslant M(\epsilon, A),$$

and of course if f is a Lipschitzian mapping with Lipschitz constant K, then:

$$M(K\epsilon, f(A)) \leqslant M(\epsilon, A).$$

The following properties give the simplest version of the quantitative transversality theorem. We state them in a rather abstract form.

(9) Let X, Y be two metric spaces and let for each $t \in Y$, $f_t : X \to X$ be a homeomorphism. Let for any $A_1, A_2 \in X$, $\Sigma(A_1, A_2) \subset Y$ denote the set of $t \in Y$, for which $f_t(A_1) \cap A_2 \neq \emptyset$.

Notation. For a given $\epsilon > 0$, define $\eta(\epsilon)$ as follows: $2\eta(\epsilon)$ is the supremum, over all pairs of balls B_1, B_2 of radius ϵ in X, of the diameter of $\Sigma(B_1, B_2)$.

In general, $\eta(\epsilon)$ measures the "nondegeneracy" of the action of the parameter t on X. Our main example is the following: $X = Y = \mathbb{R}^n$, $f_t(x) = t + x$. Then clearly $\eta(\epsilon) = 2\epsilon$, since the set of t for which $t + B_1 \cap B_2 \neq \emptyset$ is a ball of radius 2ϵ in $\mathbb{R}^n$.

Proposition 2.2. *Let $A_1, A_2 \subset X$. Then for any $\epsilon > 0$ and $\xi = \eta(2\epsilon)$, we have:*

$$M(\xi, \Sigma(A_{1,\epsilon}, A_{2,\epsilon}))_Y \leqslant M(\epsilon, A_1).M(\epsilon, A_2).$$

Proof. We cover the ϵ-neighborhoods $A_{1,\epsilon}$ and $A_{2,\epsilon}$ by $M(\epsilon, A_1)$ and $M(\epsilon, A_2)$ 2ϵ-balls B_i and B'_j, respectively. Then the set of $t \in Y$ for which $f_t(A_{1,\epsilon})$ intersects $A_{2,\epsilon}$ is contained in the union $\bigcup_{i,j} \Sigma(B_i, B'_j)$. But each of these sets is contained in some ball of radius $\xi = \eta(2\epsilon)$ in Y, by definition of $\eta(2\epsilon)$. Thus one needs less than $M(\epsilon, A_1).M(\epsilon, A_2)$ ξ-balls to cover $\Sigma(A_{1,\epsilon}, A_{2,\epsilon})$. □

Corollary 2.3. *Let $A_1, A_2 \subset \mathbb{R}^n$ be bounded subsets. Assume that*

$$M(\epsilon, A_1) \leqslant K_1(\frac{1}{\epsilon})^{\alpha} \text{ and } M(\epsilon, A_2) \leqslant K_2(\frac{1}{\epsilon})^{\beta},$$

with $\alpha + \beta < n$. Then for any $\epsilon > 0$ there is a point t in any ball of radius $r > C(n)(K_1.K_2)^{\frac{1}{n}} \epsilon^{1-\frac{\alpha+\beta}{n}}$, such that $t + A_{1,\epsilon}$ does not intersect $A_{2,\epsilon}$.

Proof. By proposition 2.2, the set Σ of all t for which $t + A_{1,\epsilon}$ intersects $A_{2,\epsilon}$ satisfies:

$$M(4\epsilon, \Sigma) \leqslant M(\epsilon, A_1).M(\epsilon, A_2) \leqslant K_1.K_2(\frac{1}{\epsilon})^{\alpha+\beta}.$$

On the other hand, for a ball B_r of radius r in $\mathbb{R}^n$, $M(4\epsilon, B_r)$ is not smaller than $C'(n)(\frac{r}{\epsilon})^n$ (for $\epsilon \leqslant r$, and where $C'(n)$ is a constant which only depends on n). Hence if $K_1.K_2(\frac{1}{\epsilon})^{\alpha+\beta} < C'(n)(\frac{r}{\epsilon})^n$, i.e. for $r > (\frac{K_1.K_2}{C'(n)})^{\frac{1}{n}} \epsilon^{1-\frac{\alpha+\beta}{n}}$, we have:

$$M(4\epsilon, \Sigma) < M(4\epsilon, B_r).$$

Thus by property (1), we cannot have $B_r \subset \Sigma$. We conclude that any r-ball contains points which are not in Σ. □

Using (7) above we can give a more "effective" version of this corollary:

Corollary 2.4. *Let $A_1, A_2, K_1, K_2, \alpha, \beta$ be as above. Then for any $\epsilon > 0$, in any 2ϵ-separated set in $\mathbb{R}^n$ containing more than $K_1.K_2(\frac{2}{\epsilon})^{\alpha+\beta}$ elements, there is a t such that $(t + A_{1,\epsilon}) \cap A_{2,\epsilon} = \emptyset$.*

Proof. By (7), any 2ϵ-separated set of $\Sigma = \Sigma(A_{1,\epsilon}, A_{2,\epsilon})$ does not contain more than $M(\epsilon, \Sigma)$ points. The proof of corollary 2.3 shows that we have

$$M(\epsilon, \Sigma) \leqslant K_1.K_2(\frac{2}{\epsilon})^{\alpha+\beta}.$$

Hence if a 2ϵ-separated set of Σ contains more than $K_1.K_2(\frac{2}{\epsilon})^{\alpha+\beta}$ elements, we have a contradiction. □

Notice that the number of elements of a regular 2ϵ-net, say in a δ-cube in $\mathbb{R}^n$, which is $(\frac{\delta}{2\epsilon})^n$, becomes greater than $K_1.K_2(\frac{2}{\epsilon})^{\alpha+\beta}$, if $\epsilon < [\frac{\delta^n}{K_1.K_2 2^{n-\alpha-\beta}}]^{\frac{1}{n-\alpha-\beta}}$ (since $n > \alpha + \beta$), and hence we can find the required t in this specific net.

Notice also that in principle the statement of Corollary 2.4 allows definite verification by computations with bounded accuracy and time. Indeed, we must check only a finite number of points t in the net, and for each verify the fact of nonintersecting, say, the $\frac{\epsilon}{2}$-neighborhood of A_1 and A_2. But it is enough to make computations with accuracy $\frac{\epsilon}{3}$ to establish this fact definitively.

(10) For subsets in $\mathbb{R}^n$, we can compare $M(\epsilon,.)$ with the Hausdorff measure.

Definition 2.5. For a (bounded) set $A \subset \mathbb{R}^n$ and $\beta \geq 0$, the *β-dimensional spherical Hausdorff measure* $\mathcal{S}^\beta$ is defined as $\mathcal{S}^\beta(A) = \lim_{\epsilon\to 0} \mathcal{S}^\beta_\epsilon$, where $\mathcal{S}^\beta_\epsilon$ is the lower bound of all the sums of the form $\sum_{i=0}^{\infty} r_i^\beta$, where $r_i \leqslant \epsilon$ are the radii of the balls B_i, $i = 1, \ldots,$ and $A \subset \bigcup_{i=0}^{\infty} B_i$.

For more details about Hausdorff measures, see [Fed 2] or [Fal]. Remark that the usual β-dimensional Hausdorff measure $\mathcal{H}^\beta$ is defined in the same way, but with the B_i's being arbitrary sets with diameter less than ϵ, in the above definition. When A is an $(\mathcal{H}^m, m)$-rectifiable subset of $\mathbb{R}^n$, we have $\mathcal{H}^m = \mathcal{S}^m$ ([Fed 2], 3.2.26).

Notice that $\mathcal{S}^0_\epsilon = M(\epsilon, A)$, but $\mathcal{S}^0$ is equal to the number of points in A. For any set $A \in \mathbb{R}^n$, and for a suitable constant $c(n)$, $c(n)\mathcal{S}^n(A)$ is equal to the Lebesgue measure of A.

Proposition 2.6. *For any bounded $A \subset \mathbb{R}^n$, and $\beta \geq 0$, we have: $\mathcal{S}^\beta(A) \leqslant \liminf_{\epsilon\to 0} \epsilon^\beta M(\epsilon, A)$.*

Proof. The proof follows easily from the definitions. We have $\mathcal{S}^\beta_\epsilon = \inf\{\sum_{i=0}^{\infty} r_i^\beta ; A \subset \bigcup_{i=0}^{\infty} B_i \} \leqslant M(\epsilon, A).\epsilon^\beta$, hence we obtain the desired inequality: $\mathcal{S}^\beta(A) = \lim_{\epsilon\to 0} \mathcal{S}^\beta_\epsilon \leqslant \liminf_{\epsilon\to 0} \epsilon^\beta M(\epsilon, A)$. □

For the usual Lebesgue measure m on $\mathbb{R}^n$, we have a similar inequality:

Proposition 2.7. *For A a bounded subset in $\mathbb{R}^n$, we have the following inequality: $m(A) \leqslant V_n \inf_{\epsilon>0} \epsilon^n M(\epsilon, A)$, where V_n is the volume of the unit ball in $\mathbb{R}^n$.*

Proof. Let $A \subset \bigcup_{i=0}^{M(\epsilon,A)} B_i$, where B_i is an ϵ-ball. Hence we have: $m(A) \leqslant \sum_{i=0}^{M(\epsilon,A)} m(B_i) = \sum_{i=0}^{M(\epsilon,A)} V_n \epsilon^n = V_n \epsilon^n M(\epsilon, A)$. □

(11) Hausdorff and entropy dimensions

Let $0 \leqslant \alpha < \beta < \gamma \leqslant n$ be three real numbers. For every $A \in \mathbb{R}^n$, every $\epsilon > 0$, and every covering $\bigcup_{i=0}^{\infty} B_i$ of A by balls of radius $\leqslant \epsilon$, we have:

$$\epsilon^{\beta-\gamma} \sum_{i=0}^{\infty} r_i^{\gamma} \leqslant \sum_{i=0}^{\infty} r_i^{\beta} \leqslant \epsilon^{\beta-\alpha} \sum_{i=0}^{\infty} r_i^{\alpha},$$

It follows easily that for any bounded $A \subset \mathbb{R}^n$, there exists $0 \leqslant \beta \leqslant n$ such that $\mathcal{S}^{\alpha}(A) = \infty$ and $\mathcal{S}^{\gamma}(A) = 0$, for all α and γ such that $0 \leqslant \alpha < \beta < \gamma \leqslant n$. This $\beta = \inf\{\gamma; \mathcal{S}^{\gamma}(A) = 0\} = \sup\{\alpha; \mathcal{S}^{\alpha}(A) = \infty\}$ is called the *Hausdorff dimension* of A, and is denoted $\dim_{\mathcal{H}}(A)$.

Although $\mathcal{H}^{\beta}(A)$ and $\mathcal{S}^{\beta}(A)$ may differ, it is a simple exercise to check that the Hausdorff dimensions defined by $\mathcal{S}$ and $\mathcal{H}$ are the same: for instance, if β is a real number bigger than the Hausdorff dimension of A defined by the measures $\mathcal{H}^{\eta}$, we have $\mathcal{H}^{\beta}(A) = 0$, and for sufficiently small δ, $\mathcal{H}^{\beta}_{\delta}(A) < 1$. So let (E_i) be a covering of A, such that $\operatorname{diam}(E_i) \leqslant \delta$ and $\sum_{i=0}^{\infty} \operatorname{diam}(E_i)^{\beta} < 2$. Each E_i is contained in a ball B_i of radius $\operatorname{diam}(E_i)$, which is less than δ, thus we have $\mathcal{S}^{\beta}_{\delta} \leqslant \sum_{i=0}^{\infty} \operatorname{diam}(E_i)^{\beta} < 2$. It follows that $\mathcal{S}^{\beta}(A) \leqslant 2$, proving that the Hausdorff dimension of A defined by the spherical Hausdorff measures $\mathcal{S}^{\eta}$ is less than β, and finally less than the Hausdorff dimension of A defined by the Hausdorff measures $\mathcal{H}^{\eta}$. The opposite inequality can be proved in the same way.

Among the properties of the Hausdorff dimension, one has the following:

(a) For A a smooth m-dimensional submanifold in $\mathbb{R}^n$, $\dim_{\mathcal{H}}(A) = m$.

(b) $\dim_{\mathcal{H}}(\bigcup_{i=1}^{\infty}) = \sup_i \dim_{\mathcal{H}}(A_i)$.

The entropy dimension of A reflects the asymptotic behavior of $M(\epsilon, A)$ as $\epsilon \to 0$. In many cases it characterizes the set A much more precisely than the Hausdorff dimension.

Definition 2.8. The *entropy dimension* $\dim_e(A)$ is defined as:

$$\dim_e(A) = \limsup_{\epsilon \to 0} \frac{\log M(\epsilon, A)}{\log(\frac{1}{\epsilon})}.$$

Thus $\dim_e(A)$ is the infinum of β for which $M(\epsilon, A) \leqslant (\frac{1}{\epsilon})^{\beta}$, for sufficiently small ϵ.

Clearly, for A a compact smooth m-dimensional manifold in $\mathbb{R}^n$ and for $\epsilon \to 0$, $M(\epsilon, A) \sim c(m).Vol_m(A).(\dfrac{1}{\epsilon})^m$, hence $\dim_e(A) = m = \dim_{\mathcal{H}}(A)$.

On the other hand in many cases these dimensions are quite different. For instance, by the property (2) of $M(\epsilon, A)$ above, we have: $\dim_e(A) = \dim_e(\bar{A})$. But for $A = \mathbb{Q} \cap [0;1]$, $\dim_{\mathcal{H}}(A) = 0$ (A being countable) and $\dim_{\mathcal{H}}(\bar{A}) = 1$. In fact we can easily see that we always have:

$$\dim_{\mathcal{H}}(A) \leqslant \dim_e(A).$$

This inequality follows from Proposition 2.6: let β be a real such that $M(\epsilon, A) \leqslant (\dfrac{1}{\epsilon})^\beta$ (for small ϵ). Proposition 2.6 allows us to write:

$$\mathcal{S}^\beta \leqslant \liminf_{\epsilon \to 0} \epsilon^\beta M(\epsilon, A) \leqslant 1.$$

Thus $\dim_{\mathcal{H}}(A) \leqslant \beta$, and finally $\dim_{\mathcal{H}}(A) \leqslant \dim_e(A)$.

As an exercise, let us compute these two dimensions for the classical Cantor set $C_{\frac{1}{3}} \subset [0;1]$. This set is obtained by removing from $[0;1]$ the interval $[\frac{1}{3}; \frac{2}{3}]$: one obtains two intervals C^1 and C^2. We then proceed in the same way with these two intervals, and we construct a sequence $C^{i_1, i_2, \ldots, i_n}$, $i_k \in \{1, 2\}, k \in \{1, \ldots, n\}$, of intervals such that $\mathrm{diam}(C^{i_1, i_2, \ldots, i_n}) = (\frac{1}{3})^n$, $C^{i_1, i_2, \ldots, i_n, i_{n+1}} \subset C^{i_1, i_2, \ldots, i_n}$, $\mathrm{d}_{\mathcal{H}}(C^{i_1, i_2, \ldots, i_n, 1}; C^{i_1, i_2, \ldots, i_n, 2}) = (\frac{1}{3})^{n+1}$. We define $C_{\frac{1}{3}}$ as the set consisting of all the points $\bigcap_{n \in \mathbb{N}} C^{i_1, i_2, \ldots, i_n}$ (see Fig. 2.2).

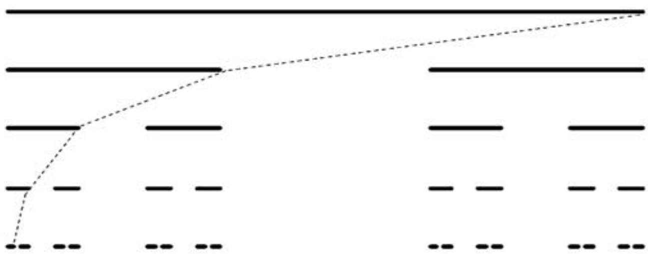

Fig. 2.2.

We can construct $C_{\frac{1}{k}}$, with $k > 2$, by considering intervals of length $\frac{1}{k^n}$ instead of $\frac{1}{3^n}$ at the step n. Notice that $C_{\frac{1}{3}}$ is obtained from $\widetilde{C}^1 = C_{\frac{1}{3}} \cap C^1$ by a homothety with centre the origin and ratio 3. Thus if $d = \dim_{\mathcal{H}}(C_{\frac{1}{3}})$, we have $\mathcal{S}^d(C_{\frac{1}{3}}) = 3^d \mathcal{S}^d(\widetilde{C}^1)$. Furthermore, $\widetilde{C}^1 \cup \widetilde{C}^2 = C_{\frac{1}{3}}$ and by symmetry $\mathcal{S}^d(C_{\frac{1}{3}}) = 2\mathcal{S}^d(\widetilde{C}^1)$. Now, as $3^d \mathcal{S}^d(\widetilde{C}^1) = 2\mathcal{S}^d(\widetilde{C}^1)$, assuming that $0 < \mathcal{S}^d(C_{\frac{1}{3}}) < \infty$, we have:

$$\dim_{\mathcal{H}}(C_{\frac{1}{3}}) = \frac{\log(2)}{\log(3)}$$

$$(\text{ and similarly } \dim_{\mathcal{H}}(C_{\frac{1}{k}}) = \frac{\log(2)}{\log(k)})$$

(for a rigorous proof see [Fal]).

Let us now compute $\dim_e(C_{\frac{1}{3}})$. For $\frac{1}{2.3^{n+1}} \leqslant \epsilon < \frac{1}{2.3^n}$, we have $M(\epsilon, C_{\frac{1}{3}}) = 2^{n+1}$, hence:

$$\frac{(n+1)\log(2)}{\log(2)+(n+1)\log(3)} \leqslant \frac{\log(M(\epsilon, C_{\frac{1}{3}}))}{\log(\frac{1}{\epsilon})} \leqslant \frac{(n+1)\log(2)}{\log(2)+n\log(3)}.$$

This proves that:

$$\dim_e(C_{\frac{1}{3}}) = \frac{\log(2)}{\log(3)} = \dim_{\mathcal{H}}(C_{\frac{1}{3}})$$

$$\text{and similarly } \dim_e(C_{\frac{1}{k}}) = \frac{\log(2)}{\log(k)} = \dim_{\mathcal{H}}(C_{\frac{1}{k}}).$$

For $A = \{1, \frac{1}{2}, \frac{1}{3}, \ldots\}$, $\dim_e(A) = 1/2$ (see below), while the Hausdorff dimension of this countable set is of course 0.

The entropy dimension of sets in $\mathbb{R}$ may be defined in many different ways (see [Tri]).

Theorem 2.9. (A. S. Besicovitch, S. J. Taylor) *Let $A \subset [a;b]$ be a closed subset.*

(i) If $m(A) > 0$, then $\dim_e(A) = 1$

(ii) Let $m(A)$ be zero, i.e $A = [a;b] \setminus \bigcup_{i=1}^{\infty} V_i$, with V_i open disjoint intervals and $\sum_{i=1}^{\infty} \alpha_i = b - a$, where α_i is the length of V_i. Then:

$$\dim_e(A) = \inf\{\beta; \sum_{i=1}^{\infty} \alpha_i^{\beta} < \infty\}.$$

For example, for $A = \{1, \frac{1}{2^a}, \frac{1}{3^a}, \ldots\}$, $\alpha_i \sim \frac{1}{i^{a+1}}$, and hence $\dim_e(A) = \frac{1}{a+1}$.

This theorem allows us to compute again $\dim_e(C_{\frac{1}{3}})$: for $2^{n-1} \leqslant i < 2^n$, $\alpha_i = (\frac{1}{3})^n$, thus $\sum_{i=1}^{\infty} \alpha_i^{\beta} = \frac{1}{3}\sum_{i=1}^{\infty}(\frac{2}{3^{\beta}})^{i-1}$, and this sum is convergent if and only if $\beta > \dim_e(A) = \frac{\log(2)}{\log(3)}$.

The following construction ([Yom 13]) generalizes Theorem 2.9 to higher dimensions:

Definition 2.10. Let $A \subset \mathbb{R}^n$ be a bounded subset. For a given $\beta > 0$, points $x_1, \dots, x_p \in A$, and a connected tree T with vertices x_i, $\rho_\beta(x_1, \dots, x_p, T)$ is the sum $\sum_{e \in T} |e|^\beta$, where for an edge e in T connecting x_i and x_j, $|e| = \mathrm{d}(x_i; x_j)$. Let $\rho_\beta(x_1, \dots, x_p)$ be the inf $\rho_\beta(x_1, \dots, x_p, T)$ over all the trees T connecting $x_1, \dots, x_p$. Finally, let $V_\beta(A) = \sup_{p, x_1, \dots, x_p \in A} \rho_\beta(x_1, \dots, x_p)$.

One can easily see that in the situation of Theorem 2.9, $V_\beta(A) = \sum_{i=1}^{\infty} \alpha_i^\beta$.

Theorem 2.11. *For any bounded $A \subset \mathbb{R}^n$, we have the following characterization of* $\dim_e$*:* $\dim_e(A) = \inf\{\beta;\ V_\beta(A) < \infty\}$.

The invariant V_β presents some interesting features. For instance the proof we have for the following property: for any $A \subset B^n$, $V_n(A) < \infty$, is nontrivial, although of course elementary. V_β turns to be intimately related to the properties of critical values of differentiable functions. In fact the Morse-Sard theorem (see [Fed 2], [Com 1]) claims that for a C^k-smooth function $f : B^n \to \mathbb{R}$, $\mathcal{H}^{\frac{n}{k}}(\Delta(f)) = 0$ (furthermore, this bound is the sharpest one, as proved in [Com 1]) and thus $\dim_{\mathcal{H}}(\Delta(f)) \leqslant \frac{n}{k}$. Our Theorem 1.10 above implies immediately a much stronger result: $\dim_e(\Delta(f)) \leqslant \frac{n}{k}$ (again this is the sharpest bound by [Com 1]). In particular the set $\{1, \frac{1}{2^a}, \frac{1}{3^a}, \dots, 0\}$ cannot be the set of critical values of f, if $k > n(a+1)$, while the Morse-Sard theorem gives no restrictions for countable sets to be sets of critical values.

However, the necessary and sufficient conditions for a given set to be the set of critical values, are given just in terms of V_β:

Theorem 2.12. ([Yom 13], [Bat-Nor]) *The compact set $A \subset \mathbb{R}$ is contained in $\Delta(f)$ for some C^k-smooth function $f : B^n \to \mathbb{R}$ if and only if $V_{\frac{n}{k}}(A) < \infty$.*

The rest of this book presents many examples of computation (or estimation) of $M(\epsilon, A)$. Usually we are interested not only in the asymptotic behavior of $M(\epsilon, A)$ as $\epsilon \to 0$, but in estimating $M(\epsilon, A)$ for any ϵ. The following example illustrates the problems which can arise here.

Let A be a compact surface in $\mathbb{R}^3$. For $\epsilon \to 0$, $M(\epsilon, A) \sim c.\mathcal{H}^2(A).(\frac{1}{\epsilon})^2$, with c some absolute constant. However, this expression does not bound $M(\epsilon, A)$ for ϵ relatively big: indeed, taking our surface to be very "thin" and "long", we can get $\mathcal{H}^2(A) \to 0$, but $M(\epsilon, A) \sim \dfrac{1}{\epsilon}\mathrm{length}(A)$.

On the other side, taking the surface with a fixed number of connected components and with the area and the "length" tending to zero, we see that the number of connected components of A should also enter the upper bound expression (see Fig. 2.3).

Indeed the correct bound has the following form (see [Iva 1], [Leo-Mel] and theorems 3.5 and 3.6 below):

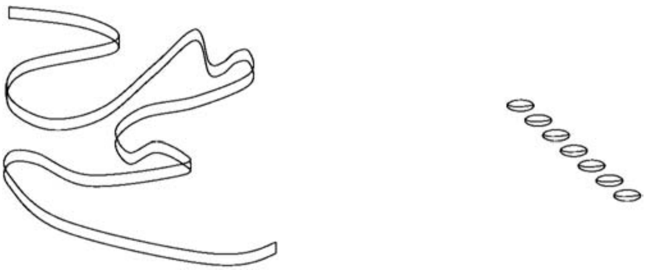

Fig. 2.3.

$$M(\epsilon, A) \leqslant \tilde{V}_0(A) + C_1\tilde{V}_1(A)\frac{1}{\epsilon} + C_2\tilde{V}_2(A)\frac{1}{\epsilon^2},$$

where $\tilde{V}_0(A)$ is the number of connected components of A, $\tilde{V}_2(A) = \mathcal{H}^2(A)$, and $\tilde{V}_1(A) = \int_A (k_1 + k_2)\ \mathrm{d}\mathcal{H}^2$, where k_1 and k_2 are the absolute values of the mean curvatures of A.

3 Multidimensional Variations

Abstract. We define in this chapter the multidimensional variations, study their properties and show how the ϵ-entropy of a subset A of $\mathbb{R}^n$ can be bounded in terms of variations of A. This form one of the main technical tools used in this book.

In this chapter we present part of the theory of multidimensional variations, developed by A. G. Vitushkin ([Vit 1], [Vit 2]), L. D. Ivanov ([Iva 1]) and others ([Leo-Mel], [Zer]...). Although in general to handle multidimensional variations is not an easy task, we will use them only for "tame sets" (algebraic, semialgebraic, analytic, semianalytic, subanalytic: see [Łoj] or [Den-Sta], definable in o-minimal structures: see Chapter 4 for a brief introduction or [Dri-Mil], [Shi], or [Dri]). In this case variations present a very convenient tool.

Our main goal is to bound $M(\epsilon, A)$ for a given A. As the last example of Chapter 2 suggests, the likely form of the required upper bound is (for $A \subset \mathbb{R}^n$):

$$M(\epsilon, A) \leqslant C(n) \sum_{i=0}^{n} V_i(A) (\frac{1}{\epsilon})^i,$$

where $V_0(A)$ should be the number of connected components of A, $V_n(A)$ its volume, and $V_i(A)$ should reflect the i-dimensional "size" of A. It turns out that the Vitushkin variations $V_i(A)$ provide the required inequality.

Let G_n^k denote the space of all the k-dimensional linear subspaces in $\mathbb{R}^n$. We have on the orthogonal group $\mathcal{O}_n(\mathbb{R})$ of $\mathbb{R}^n$ a unique invariant probability measure. Taking the image of this Haar measure under the action of $\mathcal{O}_n(\mathbb{R})$ on G_n^k, we obtain the standard probability measure $\gamma_{k,n}$ (denoted dP for simplicity) on G_n^k. This measure is of course invariant under the action of $\mathcal{O}_n(\mathbb{R})$ on G_n^k.

Let now $\bar{G}_n^k$ denote the space of all the k-dimensional affine subspaces in $\mathbb{R}^n$. Representing elements $\bar{P}$ of $\bar{G}_n^{n-k}$ by pairs $(x, P) \in \mathbb{R}^n \times G_n^{n-k}$, where $x \in P$, and $\bar{P} = \bar{P}_x$ is the k-dimensional affine subspace of $\mathbb{R}^n$, orthogonal to P at x, we have the standard measure on $\bar{G}_n^{n-k}$: $\bar{\gamma}_{n-k,n} = m \otimes \gamma_{k,n}$ (m being

Y. Yomdin and G. Comte: LNM 1834, pp. 33–45, 2004.

the Lebesgue measure on P, identified with $\mathbb{R}^{n-k}$). We will denote m by dx and $\bar{\gamma}_{n-k,n}$ by $d\bar{P}$, for simplicity; thus we have: $d\bar{P} = dx \otimes dP$.

Definition 3.1. Let A be a bounded subset of $\mathbb{R}^n$. Define $V_0(A)$ as the number of connected components of A. For $i = 1, 2, \ldots, n$, define the *i-th variation of* A, $V_i(A)$, as:

$$V_i(A) = c(n,i) \int_{\bar{P} \in \bar{G}_n^{n-i}} V_0(A \cap \bar{P}) \ d\bar{P}.$$

Here the coefficient $c(n,i)$ is choosen in such a way that $V_i(Q^i) = 1$, where $Q^i = [0,1]^i$ is the unit i-dimensional cube in $\mathbb{R}^n$. In the above notations we can also represent $V_i(A)$ as follows:

$$V_i(A) = c(n,i) \int_{P \in G_n^i} \Big(\int_{x \in P} V_0(A \cap \bar{P}_x) \ dx \Big) \, dP \qquad \text{or as:}$$

$$V_i(A) = c(n,i) \int_{P \in G_n^i} \Big(\int_{x \in P} V_0(A \cap \pi_P^{-1}(x)) \ dx \Big) \, dP,$$

where π_P is the orthogonal projection of $\mathbb{R}^n$ onto the i-dimensional linear subspace P.

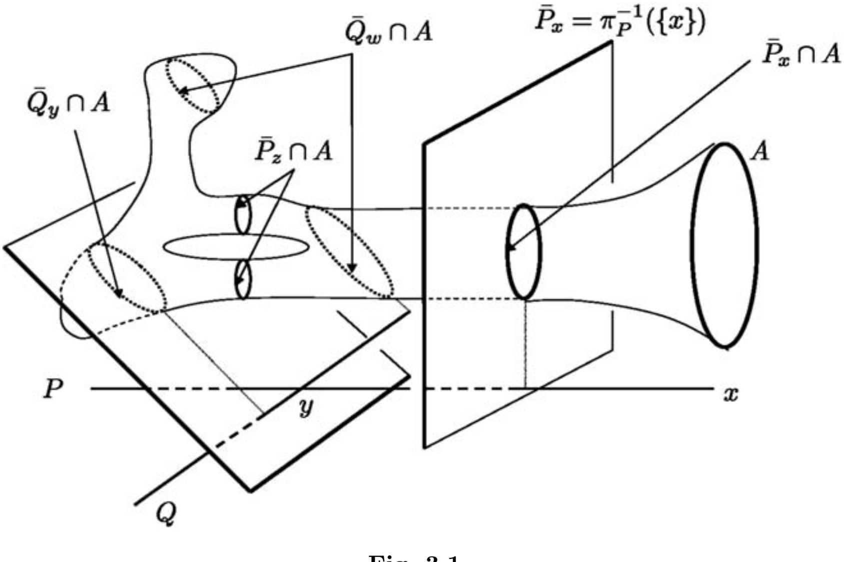

Fig. 3.1.

Remark. Of course when A is a tame set, it is easy to see that the function $\bar{P} \mapsto V_0(A \cap \bar{P})$ is measurable. More generally, if A is a closed set of $\mathbb{R}^n$, this function is still measurable (see [Vit 1,2], [Zer]).

The following properties of V_i can be proved more or less directly (see [Iva 1]) :

(1) By definition, $V_0(A)$ is the number of connected components of A, and $V_n(A) = m(A)$ (because $V_0(A \cap \pi^{-1}_{\mathbb{R}^n}(x))$ is $\mathbf{1}_A(x)$, the characteristic function of the set A).

(2) For A a smooth ℓ-dimensional submanifold of $\mathbb{R}^n$, $V_i(A) = 0$, $i > \ell$, because a generic $(n-\ell)$-plane does not encounter A, by the classical Sard Theorem. Furthermore, by the classical so-called Cauchy-Crofton formula in integral geometry (see [Buf]; [Cau]; [Cro]; [Fav]; [Leb]; [Fed 1], 5.11; [Fed 2], 2.10.15; [San]; [Lan 1,2]), we have: $V_\ell(A) = Cte \cdot \mathcal{H}^\ell(A) = Cte \cdot Vol_\ell(A)$. But, because $c(n,\ell)$ is chosen in such a way that $V_\ell([0,1]^\ell) = 1$, we obtain that $c(n,\ell)$ is the constant figuring in the Cauchy-Crofton fomula, i.e. (see [Fed 2]) $c(n,\ell) = \Gamma(\frac{1}{2})\Gamma(\frac{n+1}{2})/\Gamma(\frac{\ell+1}{2})\Gamma(\frac{n-\ell+1}{2})$, where $\Gamma(x) = \displaystyle\int_{s\in[0,+\infty]} e^{-s}s^{x-1}\ ds$ is the classical Euler function, satisfaying $\Gamma(\dfrac{1}{2}) = \sqrt{\pi}$ and $\Gamma(x+1) = x\Gamma(x)$. An easy computation allows us to find again this expression of $c(n,\ell)$. For instance, for $\ell = n-1$, by the Cauchy-Crofton formula:
$Vol_\ell(S^\ell) = c(\ell+1,\ell)\displaystyle\int_{P\in G^\ell_{\ell+1}} \Big(\int_{x\in P} V_0(S^\ell \cap \pi_P^{-1}(x))\ dx\Big)\ dP$
$= c(\ell+1,\ell)2Vol_\ell(B^\ell)$, hence we obtain: $c(\ell+1,\ell) = \dfrac{Vol_\ell(S^\ell)}{2Vol_\ell(B^\ell)} = \dfrac{\ell}{2}\dfrac{Vol_\ell(S^\ell)}{Vol_{\ell-1}(S^{\ell-1})} = \dfrac{\ell}{2}\sqrt{\pi}\dfrac{\Gamma(\ell/2)}{\Gamma((\ell+1)/2)}$.

(3) If $V_i(A) = 0$, then $V_j(A) = 0$ for $j \geq i$.

(4) For a convex subset A in $\mathbb{R}^n$, $V_i(A) = W_i(A, B^n)$, where W_i denotes the Minkowski mixed volume, and $B^n \subset \mathbb{R}^n$ is the unit ball. (see (**10**) for more details on this point.)

(5) $V_i(A)$ are invariants of the isometries of $\mathbb{R}^n$.

(6) Homogeneity property: For $\lambda \in \mathbb{R}$, $V_i(\lambda A) = \lambda^i V_i(A)$.

(7) $V_i(A \cup B) \leqslant V_i(A) + V_i(B)$. If $\bar{A} \cap \bar{B} = \emptyset$ we have the equality.

(8) Inductive formula for variations:

$$V_i(A) = c(n,i,j)\int_{\bar{P}\in\bar{G}^{n-j}_n} V_{i-j}(A \cap \bar{P})\ d\bar{P}.$$

In this formula, $A \cap \bar{P}$ is a subset of $\bar{P} = \mathbb{R}^{n-j}$, thus we have in this formula: $V_{i-j}(A \cap \bar{P}) = c(n-j, i-j)\displaystyle\int_{\bar{P}\supset\bar{Q}\in\bar{G}^{i-j}_{n-j}} V_0(A \cap \bar{P} \cap \bar{Q})\ d\bar{Q}.$

It is an exercise to compute the constant $c(n,i,j)$; for instance when $i=n-1$ and $j=1$ we have:

$$V_{n-1}(A)=c(n,n-1,1)\int_{\bar{P}\in\bar{G}_n^{n-1}} V_{n-2}(A\cap\bar{P})\ d\bar{P}.$$

Taking $A=B^{n-1}$, the unit $(n-1)$-ball of $\mathbb{R}^n$, we obtain: $V_{n-1}(B^{n-1}) = Vol_{n-1}(B^{n-1})=c(n,n-1,1)\int_{\bar{P}\in\bar{G}_n^{n-1}} Vol_{n-2}(B^{n-1}\cap\bar{P})\ d\bar{P}$. Now the hyperplane $\bar{P}$ is given by $(x=\sin(\theta),\xi)$, where ξ is a unit vector of $\mathbb{R}^n$, and $\theta\in[0,\pi/2]$. If the angle of ξ with $(B^{n-1})^{\perp}$ is $\alpha\in[0,\pi/2]$, the $(n-2)$-ball $B^{n-1}\cap\bar{P}$ has radius $\sqrt{1-\frac{\sin^2(\theta)}{\sin^2(\alpha)}}=\sqrt{1-\frac{x^2}{\sin^2(\alpha)}}$, hence:

$$\frac{Vol_{n-1}(B^{n-1})}{Vol_{n-2}(B^{n-2})}=\frac{2c(n,n-1,1)}{Vol_{n-1}(S^{n-1})}\int_{\xi\in\frac{1}{2}S^{n-1}}\int_{x=0}^{\sin(\alpha)}\left(1-\frac{x^2}{\sin^2(\alpha)}\right)^{\frac{n-2}{2}}dxd\xi$$

$$=\frac{2c(n,n-1,1)}{Vol_{n-1}(S^{n-1})}\int_{\xi\in\frac{1}{2}S^{n-1}}\sin(\alpha)\int_{u=0}^{\frac{\pi}{2}}\cos^{n-1}(u)dud\xi.$$

Let us denote $\mathcal{I}_n=\int_{u=0}^{\frac{\pi}{2}}\cos^n(u)du=\int_{u=0}^{\frac{\pi}{2}}\sin^n(u)du$, and let us recall that $\mathcal{I}_n=\frac{\pi}{2}\cdot\frac{n!}{2^n(\frac{n}{2}!)^2}$, when n is even, and $\mathcal{I}_n=\frac{2^{(n-1)}[(\frac{n-1}{2})!]^2}{n!}$, when n is odd. We obtain[1]:

$$\frac{O_{n-2}/(n-1)}{O_{n-3}/(n-2)}=\frac{2c(n,n-1,1)}{O_{n-1}}O_{n-2}\int_{\alpha=0}^{\frac{\pi}{2}}\sin^{n-1}(\alpha)I_{n-1}d\alpha.$$

And finally $c(n,n-1,1)=\dfrac{(n-2)\cdot O_{n-1}}{2.I_{n-1}^2.(n-1)\cdot O_{n-3}}$.

(9) The variations of different orders of a set are independent ([Vit 1,2], 22, Theorem 1): given any numbers $\rho>0$ and $0\leqslant A_i\leqslant+\infty$, $i=0,\dots,n$, with A_0 an integer and $A_n<\rho^n$, one can construct a closed set A lying in the cube $[0,\rho]^n$ such that:

$$V_i(A)=A_i\quad i=1,\dots,n.$$

[1] We denote O_n the n-volume of the n-unit sphere S^n. Let us recall that $\int_{\xi\in S^n}\Psi(\xi)\ d\xi=\int_{\alpha_1}\cdots\int_{\alpha_n}\Psi(\varphi(\alpha_1,\dots,\alpha_n))\prod_{j=1}^{n}\sin^{j-1}(\alpha_j)\ d\alpha_1\dots d\alpha_n$, for $\Psi: S^n\to\mathbb{R}$, where $\xi=\varphi(\alpha_1,\dots,\alpha_n)$ are the spherical coordinates.
In particular, we get: $O_n=O_{n-1}\int_{\alpha=0}^{\pi}\sin^{n-1}(\alpha)\ d\alpha=\frac{2\pi^{(n+1)/2}}{\Gamma((n+1)/2)}$.

(10) The variations $V_i(A)$, defined as the mean value of the number of connected components of the slices $\bar{P} \cap A$ are constructed in the same way as a lot of invariants in integral geometry. For instance, instead of the number $V_0(A \cap \bar{P})$ of connected components of $\bar{P} \cap A$, one may consider the Euler-Poincaré characteristic $\chi(\bar{P} \cap A)$, and the mean value:

$$\Lambda_i(A) = c(n,i) \int_{\bar{P} \in \bar{G}_n^{n-i}} \chi(A \cap \bar{P}) \, d\bar{P}$$

over all $(n-i)$-dimensional affine planes of $\mathbb{R}^n$. One obtains the so-called Lipschitz-Killing curvature $\Lambda_i(A)$ of the set A (see [Bla], [Ste], [Wey] for the emergence of these invariants and for instance [Brö], [Brö-Kup], [Kla], [Kup], [Sch], [Sch-McM], for a complete overview on this subject. See also [Fu] for a generalization of these curvatures via the so called "normal cycle"). These curvatures are characterized by the following property: assume that A is smooth, then the n-volume of the ϵ-neighborhood A_ϵ of A is a polynomial function in ϵ: $Vol_n(A_\epsilon) = \sum_{j=0}^{n} \Lambda_{n-j}(A) \cdot \mu_j \cdot \epsilon^j$ (here μ_j is the j-volume of the j-dimensional unit ball). In the general case (i.e. A may have singularities) we have to consider the generalized volume $\mathcal{V}_A(\epsilon) = \int_{x \in \mathbb{R}^n} \chi(A \cap B_\epsilon^n) \, dx$ instead of Vol_n, in order to get $\mathcal{V}_A(\epsilon) = \sum_{j=0}^{n} \Lambda_{n-j}(A) \cdot \mu_j \cdot \epsilon^j$.
Furthermore, note that $V_i(A) = \Lambda_i(A) (= W_i(A, B^n)$, see **(4)**) in the case A is convex, since in this situation $V_0(A \cap \bar{P}) = \chi(A \cap \bar{P}) = 1$, and $V_i(A) = \Lambda_i(A)$ in the case $\dim(A) = i$, since in this situation $V_0(A \cap \bar{P}) = \chi(A \cap \bar{P})$. The Lipschitz-Killing curvatures Λ_i are deeply connected with the (generalized) volume growth of the ϵ-neighborhood of A; the main result of this chapter (Theorem 3.5) is to show how the multidimensional variations V_i are connected with the ϵ-entropy of A.

The philosophy is the following: contrary to the Euler-Poincaré characteristic, the number of connected components is not an additive function, we thus cannot expect equality in formulas, but only inequalities.

To relate variations with the ϵ-entropy, we need the following modification:

Definition 3.2. Let B be a subset of $\mathbb{R}^n$. We denote by $V_0(A, B)$ the number of connected components of A lying strictly in B. The *higher variations $V_i(A, B)$ of A in B* are defined respectively by:

$$V_i(A, B) = c(n,i) \int_{\bar{P} \in \bar{G}_n^{n-i}} V_0(A \cap \bar{P}, B) \;\; d\bar{P}.$$

Remark. We also have an inductive formula for the relative variations, as in property (8):

$$V_i(A,B) = c(n,i,j) \int_{\bar{P}\in\bar{G}_n^{n-j}} V_{i-j}(A\cap\bar{P}, B\cap\bar{P})\ d\bar{P}$$

$$c(n,i,j)c(n-j,i-j) \int_{\bar{P}\in\bar{G}_n^{n-j}} \int_{\bar{P}\supset\bar{Q}\in\bar{G}_{n-j}^{i-j}} V_0(A\cap\bar{Q}, B\cap\bar{Q})\ d\bar{Q}d\bar{P}.$$

Proposition 3.3. *If $B_1,\dots,B_k$ are disjoint sets, then for any set $A \subset \mathbb{R}^n$ and for all $i = 0,\dots,n$, $V_i(A) \geq \sum_{j=0}^{n} V_i(A, B_j)$.*

Proof. The inequality is immediate for $i = 0$, and since the V_i's are defined by integration of V_0, it follows for V_i, $i > 0$. □

Now the main property of variations, distinguishing them among the usual metric invariants, is the following ([Iva 1], Theorem II.5.1; [Vit 1,2], 21, Lemma 1; [Zer], Theorem 2):

Theorem 3.4. *There exists a constant $c(n)$, depending only on n, such that, for any nonempty $A \subset \mathbb{R}^n$, and B_r a ball of radius r centered at $x \in A$, we have:*

$$\sum_{i=0}^{n} \frac{1}{r^i} V_i(A, B_r) \geq c(n). \qquad (*)$$

Remarks. Notice that the expected lower bound $c(n)$ for the sum of all the variations of sets A in B_1 is necessarily less than 1, because on one hand, by **(9)** the reals $V_i(A)$, $i \neq 0$, may be as small as we want, and $V_i(A, B_1) \leqslant V_i(A)$ by Proposition 3.3, and because on the other hand, 1 is the minimal number of connected components of a nonempty set (i.e. $V_0(A) \geq 1$)! Of course, the important point of the theorem is that the centre of the ball B_r lies in the set A: as mentioned in property **(9)** above, one can construct a set of arbitrary small variations in B_1, with no connected component inside B_1 ($V_0(A, B_1) = 0$). This set will certainly not contain the center of B_1 but will necessarily be very close to the boundary of B_1.

The proof of this theorem is rather tricky and complicated, so let us give the main idea of the proof.

First of all, by the homogeneity property (property **(6)** above) of variations it is enough to prove $(*)$ for $r = 1$. Assume now for simplicity that $n = 2$ and A is a tame set (see Definiion 4.17).

Let A_0 be the connected component of A containing x, the center of the ball B with radius 1.

If $A_0 \subset B$, then already $V_0(A, B) \geq 1$ (Fig. 3.2a).

If $A_0 \cap (\mathbb{R}^2 \setminus B) \neq \emptyset$, we can find a tame (and thus rectifiable) curve Γ in A, connecting x with the boundary of B. The length of $\Gamma \cap B_r$ is at least r, for any $r \in [0, 1]$. Hence by a classical integral geometry argument, the affine

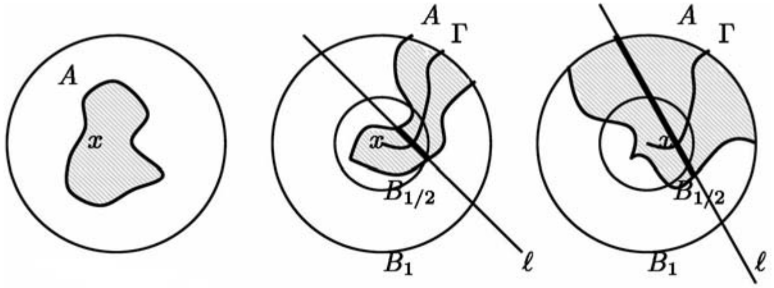

Fig. 3.2a Fig. 3.2b Fig. 3.2c

Fig. 3.2.

lines ℓ intersecting $\Gamma \cap B_{1/2}$ are in a set $E \subset \bar{G}_2^1$ of measure M (M does not depend on A). (It is a direct consequence of the Cauchy-Crofton formula: if ℓ intersects $[x;y]$, where $y = \Gamma \cap B_{1/2}$, ℓ also intersects $\Gamma \cap B_{1/2}$, and by the Cauchy-Crofton formula, we have: $c(2,1)\displaystyle\int_{\ell \in E} d\ell = c(2,1).M \geq V_1([x;y]) = \frac{1}{2}$. Hence $M \geq \dfrac{1}{2c(2,1)}$.)

If for half of these lines the component of $A \cap \ell$ containing $\Gamma \cap \ell$ lies in B, then $V_1(A,B)$ is bigger than $\frac{1}{2}M$ (Fig. 3.2b).

Finally if for more than half of the lines ℓ intersecting $\Gamma \cap B_{1/2}$ the component of $A \cap \ell$ containing $\Gamma \cap \ell$ hits also the boundary of B, than the length of this component is bigger than $\frac{1}{2}$, and an by integral geometry argument (Fubini's Theorem) we see that the area of $A \cap B$, i.e. $V_2(A,B)$, is bigger than a constant depending only on M (Fig. 3.2c).

Considering this draft of proof of Theorem 3.4, we see that the sum of the variations of A in B is in fact related to the *depth of embedding of A in B*, i.e. the real $\rho = \sup_{x \in A} \mathrm{d}(x, \mathbb{R}^2 \setminus B)$. One can actually prove (see [Vit 1,2], [Zer]) that, for any ball B of $\mathbb{R}^n$ (not necessarily centered at a point lying in A) and of radius at most 1, we have:

$$\sum_{i=0}^{n} V_i(A,B) \geq c(n)\rho^n. \qquad (**)$$

In fact to prove $(**)$, it is enough to suppose that B is a ball centered at a point of A and of radius ρ (by proposition 3.3), and by the homogeneity property, it is enough to assume that $\rho = 1$. Finally, to prove both $(*)$ and $(**)$, it is enough to prove that:

$$\sum_{i=0}^{n} V_i(A,B_1) \geq c(n), \qquad (***)$$

where B_1 is a ball centered at a point of A and of radius 1.

Proof of Theorem 3.4. As noticed above, it is enough to prove $(***)$ to obtain $(*)$ and $(**)$. Of course, if the connected component of A containing 0 lies in $B = B_1$, we have $V_0(A, B) \geq 1$ and the theorem is proved. In what follows we thus assume that this connected component hits the boundary of B. We will follow [Zer], Theorem 2, and prove $(***)$ by induction on n, the dimension of the ambient space $\mathbb{R}^n$. If $n = 1$, one of the intervals $[0, 1]$, $[-1, 0]$ is contained in A, thus $V_1(A, B) \geq 1$.

We now suppose that $n > 1$ and that the theorem (in fact $(**)$) is proved for $m < n$, i.e. that we have: $\sum_{i=0}^{m} V_i(A', B') \geq c(m)\rho^m$ for any set $A' \subset \mathbb{R}^m$ with depth of embedding ρ in B', a ball of radius less than one.

An affine hyperplane $\bar{P}$ of $\mathbb{R}^n$ is given by (x, P) (see the notations at the beginning of this chapter) or by $(\sin(\theta), \xi)$, with $\theta \in [0, \frac{\pi}{2}]$ and $\xi \in S$, the unit $(n-1)$-sphere. Let us denote by $T_\rho(\theta)$ the set of $\xi \in S$ such that the depth embedding of $\bar{P}_{(\sin(\theta),\xi)} \cap A$ in $\bar{P} \cap B$ is at least ρ. We remark that the radius of the ball $\bar{P} \cap B$ is $\cos(\theta)$. Let θ be such that $\cos(\theta) > \rho$, then there exist $\xi \in S$ and $y \in A \cap \bar{P}_{(\sin(\theta),\xi)}$ such that $\mathrm{d}(y, \bar{P} \cap S) = \rho$

(if such ξ and y do not exist, then for all ξ in S, the sphere in $\bar{P}(\sin(\theta), \xi)$ of centre 0 and radius $\cos(\theta) - \rho$ does not have common points with A, thus A does not have points in some sphere $S_{r=r(\theta,\rho)}$, which is impossible because A connects 0 to S).

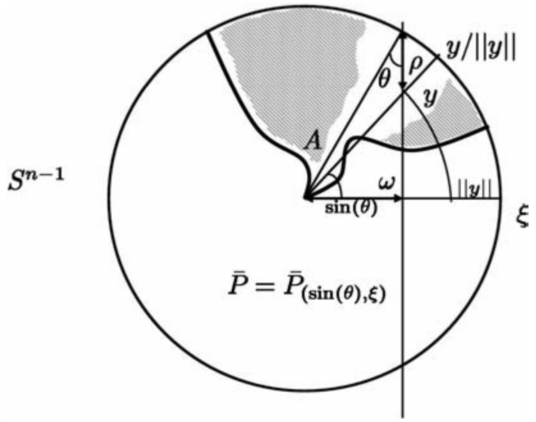

Fig. 3.3.

This implies that $\xi \in T_\rho(\theta)$. If ω is the angle between ξ and $\dfrac{y}{||y||}$ we have (see Fig. 3.3) $\sin(\theta) = ||y|| \cos(\omega)$ ($\Longleftrightarrow y \in \bar{P}$), and $\mathrm{d}(y, \bar{P} \cap S) = \cos(\theta) - ||y|| \sin(\omega)$, thus $\omega = \omega(\theta, \rho)$ is given by:

$$\mathrm{d}(y, \bar{P} \cap S) = \frac{\cos(\theta + \omega)}{\cos(\omega)} = \rho.$$

Now by symmetry, $T_\rho(\theta)$ contains the sphere of center $\dfrac{y}{||y||}$ and radius $\omega = \omega(\theta,\rho)$.

Let $K_{(s,||y||)}$ be the connected component of A containing y in the closure of $B_{||y||} \setminus B_s$ for $(s < ||y||)$ and $\tilde{K}_{(\theta,\rho)}$ the image of $K_{(\sin(\theta),\frac{\sin(\theta)}{\cos(\omega(\theta,\rho))})}$ by the map $z \mapsto \left(\dfrac{z}{||z||}, \arccos\dfrac{\sin(\theta)}{||z||}\right)$; $\tilde{K}_{(\theta,\rho)}$ is connected in $S \times [0,\omega(\theta,\rho)]$.

Furthermore, if $(\xi',\omega') \in \tilde{K}_{(\theta,\rho)}$ then $T_\rho(\theta)$ contains the sphere of center ξ' and radius ω'; indeed $(\xi',\omega') \in \tilde{K}_{(\theta,\rho)} \iff \dfrac{\sin(\theta)}{\cos(\omega')}\xi' \in A$, and then the sphere of center ξ' and radius ω' is contained in $T_{\rho'}(\theta)$, with $\rho' = \dfrac{\cos(\theta+\omega')}{\cos(\omega')}$, but now $\rho' \geq \rho$ implies $T_{\rho'}(\theta) \subset T_\rho(\theta)$.

Notice that y may be choosen such that $K_{(s,||y||)}$, the connected component of A containing y in the closure of $B_{||y||} \setminus B_s$, hits the boundary of B_s, for all $s \in [0,||y||]$. Consequently we easily check that $K_{(\sin(\theta),\frac{\sin(\theta)}{\cos(\omega(\theta,\rho))})}$ contains points of norm s, for all $s \in [\sin(\theta), ||y|| = \dfrac{\sin(\theta)}{\cos(\omega(\theta,\rho))}]$.

Finally we have proved that there exists a connected set $\widetilde{K}_{(\theta,\rho)}$ in $S \times [0,\omega(\theta,\rho)]$, such that for all (ξ',ω') in $\tilde{K}_{(\theta,\rho)}$, $T_\rho(\theta)$ contains the sphere of center ξ' and radius ω', and that there exist ξ_0 and $\xi_{\omega(\theta,\rho)} = \dfrac{y}{||y||}$ with $(\xi_{\omega(\theta,\rho)},\omega(\theta,\rho)),(\xi_0,0) \in \tilde{K}_{(\theta,\rho)}$. We conclude that $T_\rho(\theta)$ necessarily contains a ball of radius $\omega(\theta,\rho)$, and thus that the measure $\mathcal{A}(\rho,\theta)$ of the set of hyperplanes $P_{(\sin(\theta),\xi)}$ of $\mathbb{R}^n$, with $\xi \in T_\rho(\theta)$, is such that (see Footnote 1 on Page 36):

$$\mathcal{A}(\rho,\theta) \geq \frac{1}{\mathcal{I}_{n-1}} \int_{\nu=0}^{\omega(\theta,\rho)} \sin^{n-2}(\nu)\ \ d\nu, \tag{$\sharp$}$$

with $\mathcal{I}_{n-1} = \displaystyle\int_{\nu=0}^{\pi} \sin^{n-1}(\nu)\ \ d\nu.$

Now we have by the inductive formula for relative variations:

$$\sum_{i=0}^{n} V_i(A,B) = \sum_{i=1}^{n} V_i(A,B)$$

$$= \sum_{i=1}^{n} c(n,i,1) \int_{\bar{P}\in\bar{G}_n^{n-1}} V_{i-1}(A\cap\bar{P}, B\cap\bar{P})\ \ d\bar{P}.$$

Let us denote $\min_{i\in\{1,\ldots,n\}}(c(n,i,1))$ by $\mu(n)$. It follows that:

$$\sum_{i=0}^{n} V_i(A,B) \geq \mu(n) \int_{\bar{P}\in\bar{G}_n^{n-1}} \sum_{i=0}^{n-1} V_{i-1}(A\cap\bar{P}, B\cap\bar{P})\ \ d\bar{P}$$

and thus by the induction hypothesis:

$$\sum_{i=0}^{n} V_i(A,B) \geq \mu(n)c(n-1) \int_{\bar{P}\in\bar{G}_n^{n-1}} \delta_A^{n-1}(\bar{P}) \ \ d\bar{P},$$

where $\delta_A^{n-1}(\bar{P})$ is the depth of embedding of $A\cap\bar{P}$ in $B\cap\bar{P}$.
We have: $\displaystyle\int_{\bar{P}\in\bar{G}_n^{n-1}} \delta_A^{n-1}(\bar{P}) \ \ d\bar{P} = \int_{\theta=0}^{\frac{\pi}{2}} \int_{\rho=0}^{\cos(\theta)} \rho^{n-1} \ \ d[-\mathcal{A}(\rho,\theta)]d[\sin(\theta)]$,
and after an integration by parts:

$$\int_{\bar{P}\in\bar{G}_n^{n-1}} \delta_A^{n-1}(\bar{P}) \ \ d\bar{P} = \int_{\theta=0}^{\frac{\pi}{2}} \int_{\rho=0}^{\cos(\theta)} \cos(\theta)\mathcal{A}(\rho,\theta) \ \ d[\rho^{n-1}]d\theta.$$

Using (♯), we find:

$$\int_{\bar{P}\in\bar{G}_n^{n-1}} \delta_A^{n-1}(\bar{P}) \ \ d\bar{P} \geq$$

$$\frac{1}{\mathcal{I}_{n-1}} \int_{\theta=0}^{\frac{\pi}{2}} \int_{\rho=0}^{\cos(\theta)} \cos(\theta) \int_{\nu=0}^{\omega(\theta,\rho)} \sin^{n-2}(\nu) \ \ d\nu d[\rho^{n-1}]d\theta.$$

By Fubini's theorem, we obtain:

$$\int_{\bar{P}\in\bar{G}_n^{n-1}} \delta_A^{n-1}(\bar{P}) \ \ d\bar{P} \geq$$

$$\frac{1}{\mathcal{I}_{n-1}} \int_{\theta=0}^{\frac{\pi}{2}} \cos(\theta) \int_{\nu=0}^{\frac{\pi}{2}-\theta} \frac{\cos^{n-1}(\theta+\nu)\sin^{n-2}(\nu)}{\cos^{n-1}(\nu)} \ \ d\nu d\theta$$

$$= \frac{1}{\mathcal{I}_{n-1}} \int_{\nu=0}^{\frac{\pi}{2}} \frac{\sin^{n-2}(\nu)}{\cos^{n-1}(\nu)} \int_{\theta=0}^{\frac{\pi}{2}-\nu} \cos^{n-1}(\theta+\nu)\cos(\theta) \ \ d\theta d\nu$$

$$= \frac{1}{\mathcal{I}_{n-1}} \int_{\nu=0}^{\frac{\pi}{2}} \frac{\sin^{n-2}(\nu)}{\cos^{n-1}(\nu)} \int_{\zeta=\nu}^{\frac{\pi}{2}} \cos^{n-1}(\zeta)\cos(\zeta-\nu) \ \ d\zeta d\nu$$

$$= \frac{1}{\mathcal{I}_{n-1}} \int_{\nu=0}^{\frac{\pi}{2}} \frac{\sin^{n-2}(\nu)}{\cos^{n-1}(\nu)} \int_{\zeta=\nu}^{\frac{\pi}{2}} \big(\cos^{n}(\zeta)\cos(\nu)+\cos^{n-1}(\zeta)\sin(\zeta)\sin(\nu)\big) \ \ d\zeta d\nu$$

$$\geq \frac{1}{\mathcal{I}_{n-1}} \int_{\nu=0}^{\frac{\pi}{2}} \frac{\sin^{n-1}(\nu)}{\cos^{n-1}(\nu)} \int_{\zeta=\nu}^{\frac{\pi}{2}} \cos^{n-1}(\zeta)\sin(\zeta) \ \ d\zeta d\nu = \frac{1}{n^2\mathcal{I}_{n-1}}.$$

Finally, we have proved: $\displaystyle\sum_{i=0}^{n} V_i(A,B) \geq \frac{\mu(n)c(n-1)}{n^2\mathcal{I}_{n-1}} = c(n).$ □

Now we are ready to prove the inequality bounding ϵ-entropy in terms of variations.

Theorem 3.5. ([Iva 1], p. 246) *Let A be a bounded subset of $\mathbb{R}^n$. Then for any $\epsilon > 0$,*

$$M(\epsilon, A) \leqslant C(n) \sum_{i=0}^{n} \frac{1}{\epsilon^i} V_i(A),$$

where $C(n) = \dfrac{2^n}{c(n)}$ and $c(n)$ is the constant of Theorem 3.4

Proof. We recall that if $M'(\epsilon, A)$ is the maximal number of points x_j in A such that $\mathrm{d}(x_i, x_j) > \epsilon$, for $i \neq j$, then $A \subset \bigcup_{j=1}^{M'(\epsilon,A)} B_{(x_j,\epsilon)}$ (see the beginning of Chapter 2). Thus $M(\epsilon, A) \leqslant M'(\epsilon, A)$, and the balls $B_{(x_j,\epsilon/2)}$ are disjoint.

Now let $x_1, \dots, x_q$, be some ϵ-separated set in A, with $q = M(\epsilon, A)$. Consider the balls B_j of radius $\epsilon/2$, centered at x_j, for $j \in \{1, \dots, q\}$. These balls are disjoint, and hence, by proposition 3.3, $V_i(A) \geq \sum_{j=1}^{q} V_i(A, B_j)$. Multiplying these inequalities by $(\frac{2}{\epsilon})^i$ and adding them for $i = 0, 1. \dots, n$, we get by theorem 3.4:

$$\sum_{i=0}^{n} V_i(A)(\frac{2}{\epsilon})^i \geq \sum_{j=1}^{q} \sum_{i=1}^{n} (\frac{2}{\epsilon})^i V_i(A, B_j) \geq 2^n \sum_{j=1}^{q} c(n) = 2^n q c(n).$$

Thus:

$$M(\epsilon, A) = q \leqslant \frac{2^n}{c(n)} \sum_{i=1}^{n} \frac{1}{\epsilon^i} V_i(A)$$

$$\leqslant C(n) \sum_{i=1}^{n} \frac{1}{\epsilon^i} V_i(A), \text{ where } C(n) = \frac{2^n}{c(n)}.$$

□

In general, computation of variations is probably not easier than the direct computation of the ϵ-entropy. Many properties of V_i's are not yet understood enough, e.g. the behavior V_i under nonlinear transformations of $\mathbb{R}^n$. On the other hand, in many cases V_i can be estimated or computed quite effectively.

For a bounded $\mathcal{C}^k$-smooth submanifold A of dimension ℓ in $\mathbb{R}^n$, the V_i's are known to be finite for $k > k_0 = 2 - \dfrac{2}{\ell - i + 2}$ ([Leo-Mel]).

For a $\mathcal{C}^2$-smooth submanifold A in $\mathbb{R}^2$, $V_i(A)$ can be bounded (and for A convex, exactly computed) in terms of curvature integrals ([Leo-Mel]). We give here the computation of [Leo-Mel] in the case when A is a surface in $\mathbb{R}^3$. Then $V_0(A) = B_0(A)$, $V_2(A) = Vol_2(A)$ and it remains to bound $V_1(A)$. We have:

$$V_1(A) = c(3,1) \int_{l \in \mathbb{RP}^2} \int_{x \in l} V_0(A \cap \bar{P}_x)\, dx dl.$$

Now, by Sard's theorem, for all $l \in \mathbb{RP}^2$ there exists Ω_l in $\pi_l(A)$ with $m(A \setminus \Omega_l) = 0$, such that for all $x \in \Omega_l$, $S = A \cap P_x$ is a smooth curve (of P_x). We obtain $V_0(S) \leqslant \frac{1}{2\pi} \int_{s \in S} k(s)\, ds$, where k is the absolute value of the curvature. (Notice that for S convex we can integrate the curvature itself, and we have an equality). Thus $V_1(A) \leqslant \frac{1}{2\pi} c(3,1) \int_{l \in \mathbb{RP}^2} \int_{x \in l} \int_{s \in A \cap P_x} k(s)\; ds\, dx\, dl$. Now $k(s) = \frac{1}{\sin(\theta)} k(y,l)$, where $k(y,l)$ is the mean curvature of A at $y \in A$ in the direction of the plane passing through y and orthogonal to l (Fig. 3.4).

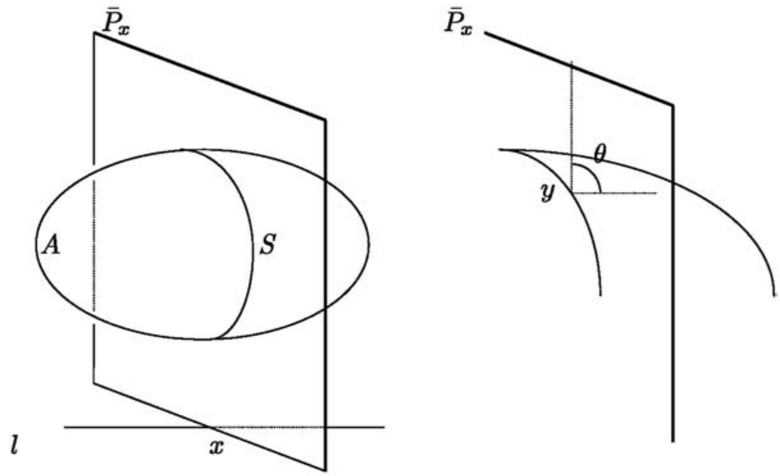

Fig. 3.4.

But $\frac{dx}{\sin(\theta)} ds = dy$. Therefore:

$$V_1(A) \leqslant \frac{1}{2\pi} c(3,1) \int_{l \in \mathbb{RP}^2} \int_{y \in A} k(y,l)\; dy\, dl$$
$$= \frac{1}{2\pi} c(3,1) \int_{y \in A} \int_{l \in \mathbb{RP}^2} k(y,l)\; dl\, dy.$$

Finally, $\int_{l \in \mathbb{RP}^2} k(y,l)\; dl \leqslant c'(k_1(y) + k_2(y))$, by the Euler formula, where $k_1(y)$ and $k_2(y)$ are the absolute values of the main curvatures of A at $y \in A$. Thus we've proved the following.

Theorem 3.6. ([Leo-Mel]) *For A a compact $\mathcal{C}^2$ surface in $\mathbb{R}^3$,*

$$V_1(A) \leqslant C \int_{y \in A} (k_1(y) + k_2(y))\, dy.$$

Via theorem 3.4 this proves also the formula for $M(\epsilon, A)$, given in the end of Chapter 2. We mention also the following result: if $A_k \xrightarrow[n \to \infty]{} A$ in the Hausdorff metric, and $V_i(A_k)$ are uniformly bounded, $i = 0, 1, \ldots, n$, then $\lim\limits_{k \to \infty} V_i(A_k) \geq V_i(A)$ ([Iva 1], theorem II.6.1).

For semialgebraic sets see results of Chapter 5 below.

4 Semialgebraic and Tame Sets

Abstract. We prove in this chapter a classical result: the number of connected components of a plane section $P \cap A$ of a semialgebraic set A is uniformly bounded with respect to P. An explicit bound is given in terms of the diagram of A and the dimension of P. We give a construction which provides a semialgebraic section of bounded complexity for any polynomial mapping of semialgebraic sets. In particular, any two points in a connected semialgebraic set can be joined by a semialgebraic curve of bounded complexity. We also give the definition of an o-minimal structure on the real field and show that in such a category the uniform bound for the number of connected components of plane sections holds.

Definition 4.1. A set $A \subset \mathbb{R}^n$ is called *semialgebraic*, if it can represented in a form $A = \bigcup_{i=1}^{p} A_i$, with $A_i = \bigcap_{j=1}^{j_i} A_{ij}$, where each A_{ij} has the form

$$\{(x_1, \ldots, x_n) \in \mathbb{R}^n;\ p_{ij}(x_1, \ldots, x_n) > 0\},$$
$$\{(x_1, \ldots, x_n) \in \mathbb{R}^n;\ p_{ij}(x_1, \ldots, x_n) \geq 0\},$$

p_{ij} being a polynomial (of degree d_{ij}).

Of course a representation of A in the above form is not unique.

As an exercise one can prove that the semialgebraic sets of $\mathbb{R}$ are the finite unions of points and intervals.

Definition 4.2. The set of data: $\left(n, p, j_1, \ldots, j_p, (d_{ij})_{\substack{i=1,\ldots,p \\ j=1,\ldots,j_i}}\right)$ is called the *diagram D of (the representation of) the set A.*

The properties of semialgebraic sets are studied in detail, e.g. in [Arn], [Ben-Ris], [Boc-Cos-Roy], [Cos], [Har 3], [Łoj], [Mil], [Pet-Ole], [Tho 1]. We recall briefly the most important of these properties:

Proposition 4.3. *Let $A \subset \mathbb{R}^n$ be a semialgebraic set. Then the set $\pi(A)$, where $\pi : \mathbb{R}^n \to \mathbb{R}^m$ is the canonical projection, is also semialgebraic, and the diagram of $\pi(A)$ depends only on the diagram of A (Tarski-Seidenberg theorem). As a consequence, the sets $\bar{A}$, ∂A, and each connected component*

Y. Yomdin and G. Comte: LNM 1834, pp. 47–58, 2004.

of A are semialgebraic, and the diagram of these sets depends only on the diagram of A (see [Ben-Ris], [Cos], [Har 3], [Łoj]*).*

Proposition 4.4. *Any semialgebraic set admits semialgebraic stratification (i.e. a partition into smooth submanifolds $(A_i)_{i\in I}$ of $\mathbb{R}^n$- the strata, that are semialgebraic sets of $\mathbb{R}^n$, such that the family $(A_i)_{i\in I}$ is locally finite and verifies the following property, called the frontier property: if $\bar{A}_i \bigcap A_j \neq \emptyset$ then $A_j \subset \bar{A}_i$. See Fig. 4.1 for an example of strafication), with the number of strata and their diagrams depending only on the diagram of the initial set (see* [Łoj]*). The same is true for triangulations (see* [Cos], [Har 3], [Łoj]*).*

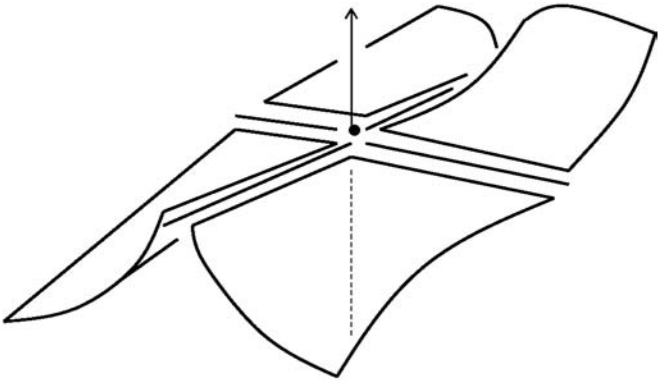

Fig. 4.1.

We define $\dim(A)$ as the maximal dimension of the strata in some stratification of A.

Proposition 4.5. *Let $A \subset \mathbb{R}^n$ be a semialgebraic set. Then all the Betti numbers $b_i(A)$, $i = 1, \ldots, n$, are bounded by constants $B_i(D)$ depending only on the diagram $D(A)$. In particular, the number of connected components of A is bounded by $B_0(D)$.*

This follows from Proposition 4.4 or can be obtained directly (see [Arn], [Ben-Ris], [Boc-Cos-Roy], [Mil], [Pet-Ole 1,2], [Tho]).

Since we need an explicit bound for $B_0(D)$, this last result we prove below.

In most constructions in this chapter we try to use only "direct" methods, avoiding e.g. the use of projections and hence of the theorem of Tarski-Seidenberg. The bounds are obtained by reduction to the case, where Bezout's theorem is applicable. We recall this theorem, in the real case, beelow. However we mention again, as in Chapter 1, references concerning effectiveness of algorithm involving Tarski-Seidenberg's principle: [And-Brö-Rui], [Hei-Rec-Roy], [Hei-Roy-Sol 1,2,3,4], [Ren 1, 2, 3] etc...

Proposition 4.6. (Bezout's theorem, see e.g. [Ben-Ris]) *Let*

$$p_1 = p_2 = \ldots = p_n = 0 \qquad (*)$$

be a system of polynomial equations in $\mathbb{R}^n$*, of degrees* $d_1, d_2, \ldots, d_n$*, respectively. Then the number of nondegenerate real solutions of* $(*)$ *(*x *is such a solution if the rank of* $\left(\frac{\partial p_i}{\partial x_j}(x)\right)_{\substack{j=1,\ldots,n \\ i=1,\ldots,n}}$ *equals* n*) is bounded by* $\prod_{i=1}^{n} d_i$*.*

Proof. Consider the complexification of equations $(*)$ on $\mathbb{C}^n$ and notice that a nondegenerate real zero x of $(*)$ is also a nondegenerate solution of the complexified system. Hence the inequality follows from the complex Bezout theorem (see [Ben-Ris] for instance). $\square$

Remark. The above inequality is not true in general for isolated, but possibly degenerate, solutions of $(*)$. For instance, let us consider the following system:

$$\begin{array}{ll} f_1(x_1, \ldots, x_{n-1}) = 0 & \deg d \\ \vdots & \vdots \\ f_{n-1}(x_1, \ldots, x_{n-1}) = 0 & \deg d \end{array}$$

and let this system have d^{n-1} nondegenerate solutions. Now consider the system in $\mathbb{R}^n$:

$$\begin{array}{ll} g = f_1^2 + \ldots + f_{n-1}^2 = 0 & \deg 2d \\ x_n = 0 & \deg 1 \\ \vdots & \vdots \\ x_n = 0 & \deg 1 \end{array}$$

Clearly this system has d^{n-1} isolated (of course degenerate) solutions. But the product of degrees is $2d < d^{n-1}$, for d sufficiently big and $n \geq 3$.

The reason is that a degenerate isolated zero of a real system $(*)$ can lie on a higher-dimensional component of zeros of the complexification of $(*)$.

Definition 4.7. For any finite sequence D of integers of diagram type $D = \left(n, p, j_1, \ldots, j_p, (d_{ij})_{\substack{i=1,\ldots,p \\ j=1,\ldots,j_i}}\right)$ define $B_0(D)$ as $\frac{1}{2}\sum_{i=1}^{p} d_i(d_i - 1)^{n-1}$, where $d_i = \sum_{j=1}^{j_i} d_{ij}$. For a semialgebraic set A, define $\widehat{B}_0(A)$ as the infimum of $B_0(D)$ over the diagrams D of all representations of A.

Theorem 4.8. *For any semialgebraic set* $A \subset \mathbb{R}^n$*, the number of bounded connected components of* A*,* $\widetilde{B}_0(A)$*, is bounded by* $\widehat{B}_0(A)$*.*

Proof. We follow the proof of [Mil], [War].

Obviously, it is enough to prove that the number of connected components of $A = \bigcap_{j=1}^{q} \{p_j \underset{\geq}{>} 0\}$, $\deg p_j = d_j$, is at most $\frac{1}{2}d(d-1)^{n-1}$, $d = \sum_{j=1}^{q} d_j$.

We may assume that only the inequalities $\geq$ define A, and hence that A is closed. Indeed, let us choose a point x_α in each (not necessarily bounded) connected components A_α of A. The number of connected component A_α, and hence of x_α is finite (see e.g. [Mil], [Whi 2]).

If one of the inequalities defining A has the form $\{p_j > 0\}$, let us denote $\min_\alpha(p_j(x_\alpha)) = \delta > 0$. Hence if we replace in the definition of A this inequality by $\{p_j - \frac{\delta}{2} \geq 0\}$, we obtain a new set $A' \subset A$. Of course any connected components of A' lies in exactly one connected component of A, and all the points x_α are still in A', therefore $\widetilde{B}_0(A') \geq \widetilde{B}_0(A)$.

We may assume that each component of A has a nonempty interior. Indeed, $A = \bigcap_{j=1}^{q} \{p_j \geq 0\}$ is closed, hence the minimal distance between the components A_α of A (inside a ball B containing all the bounded components of A) is $\rho > 0$. Let U be the open $\frac{\rho}{3}$-neighborhood of A, and let $\xi = \max_{x \in B \setminus U} \min_{1 \leqslant j \leqslant q} p_j(x)$. We have $\xi < 0$, because the continuous function $\min_{1 \leqslant j \leqslant q} p_j$ reaches its maximun on the compact set $B \setminus U$. Defining $A' = \bigcap_{j=1}^{q} \{p_j - \frac{1}{2}\xi \geq 0\}$, we have $A \subset A' \subset U$, because $\xi < \frac{\xi}{2} < 0$. By the choice of ρ, we also have $\widetilde{B}_0(A) = \widetilde{B}_0(U)(\leqslant \widetilde{B}_0(A'))$. Now any connected component of A' containing a component of A has a nonempty interior, indeed if it is not the case one can find a sequence of points x_n with limit $x \in A$ such that $p_j(x_n) < \frac{1}{2}\xi$ (for some $j \in \{1, \ldots, q\}$); which is a contradiction. We conclude that the number of bounded connected components of A' having a nonempty interior is greater than $B_0(A)$. The diagram of A and A' being the same, it suffices to prove the theorem for connected components with non empty interior.

Thus we may assume $A = \bigcap_{j=1}^{q} \{p_j \geq 0\}$, with each bounded component of A having a nonempty interior (we assume that A has bounded components of course). Let $p = \prod_{j=1}^{q} p_j$, $\deg(p) = \sum_{j=1}^{q} \deg(p_j) = d$.

Any bounded connected component of A contains at least a component of $\bigcap_{j=1}^{q} \{p_j > 0\}$, because a component of $\bigcap_{j=1}^{q} \{p_j = 0\}$ cannot have a nonempty interior. Hence any bounded connected component of A contains at least a component of $\{p > 0\}$ and finally: $\widetilde{B}_0(A) \leqslant \widetilde{B}_0(\{p > 0\})$.

The components of $\{p > 0\}$ are open, hence the image by p of such a component is a non trivial interval of the type $]0, C[$ or $]0, C]$ $(C \in \mathbb{R}_+ \cup \{\infty\})$. By Sard's theorem, let $\eta > 0$ be a sufficiently small regular value of p such that in each bounded component of $\{p > 0\}$ there is at least one component Z_α of the regular hypersurface $\{p = \eta\} = Z$.

Now take a generic linear form l on $\mathbb{R}^n$. We may assume all the critical points of l on Z to be nondegenerate, and on each Z_α there are at least two critical points of l – the minimum and the maximum. But the critical points of l on Z are defined by the following system of equations (assuming $l = x_1$):

$$\begin{array}{rclc} p - \eta & = & 0 & \deg\ d \\ \dfrac{\partial p}{\partial x_2} & = & 0 & \deg\ d-1 \\ \dfrac{\partial p}{\partial x_3} & = & 0 & \deg\ d-1 \\ \vdots & & & \vdots \\ \dfrac{\partial p}{\partial x_n} & = & 0 & \deg\ d-1 \end{array}$$

By proposition 4.6, the number of critical points is at most $d(d-1)^{n-1}$, and therefore:

$$\widetilde{B}_0(A) \leqslant \widetilde{B}_0(Z) \leqslant \frac{1}{2} d(d-1)^{n-1}. \qquad \square$$

Corollary 4.9. *Let $A \subset \mathbb{R}^n$ be a semialgebraic set with diagram $\big(n, p, j_1, \ldots, j_p, (d_{ij})_{\substack{i=1,\ldots,p \\ j=1,\ldots,j_i}}\big)$. Then:*

- *the number of connected components of the intersection of A with any ball B_r in $\mathbb{R}^n$ is bounded by $\frac{1}{2}\sum_{i=1}^{p}(d_i+2)(d_i+1)^{n-1}$, where $d_i = \sum_{j=1}^{j_i} d_{ij}$,*
- *the number of bounded connected components of $A \cap P$, where P is a ℓ-plane of $\mathbb{R}^n$, is bounded by $\frac{1}{2}\sum_{i=1}^{p}(d_i+2)(d_i+1)^{\ell-1}$.*
- *In particular the number of connected components of $A \cap P$ itself is also bounded by $\frac{1}{2}\sum_{i=1}^{p}(d_i+2)(d_i+1)^{\ell-1}$.*

Proof. For the first bound, we add to the inequalities defining A the inequality $r^2 - \sum_{i=1}^{n} x_i^2 \geq 0$ of degree 2, and for the second bound we substitute $n - \ell$ variables in the equations by the others. The bound $\frac{1}{2}\sum_{i=1}^{p} d_i(d_i-1)^{\ell-1}$ does not depend on the radius of the ball, but only the degrees, thus it also bounds the number of connected components of $A \cap P$ itself. □

We need also the following construction, concerning polynomial mappings of semialgebraic sets (see the proof of Theorem 7.1). In the o-minimal case, this result is called "definable choice".

Theorem 4.10. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a polynomial mapping of degree d, and let A be a compact semialgebraic set in $\mathbb{R}^n$. Let B be a semialgebraic subset in $f(A) \subset \mathbb{R}^m$. Then there exists a semialgebraic subset $C \subset A$ with* $\dim(C) = \dim(B)$, *such that $f(C) = B$, and the diagram $D(C)$ depends only on $D(A)$, $D(B)$, n, m and d.*

Proof. It follows immediately by using, say, the stratification of the mapping $f_{|A}$ on B (see [Har 3]). We give here a more direct proof, assuming A to be compact.

Indeed, in this case for each $y \in B$, $f^{-1}(y) \cap A$ is a compact semialgebraic subset in $\mathbb{R}^n$. Take $x(y)$ to be the maximal point in $f^{-1}(\{y\}) \cap A$, according to the lexicographic order in $\mathbb{R}^n$. Then clearly the set $C \subset A$, formed by all the points $x(y)$, $y \in B$ is semialgebraic, with the diagram depending only on the required data, and $f_{|C} : C \to B$ is one-to-one. □

This proof was suggested by A. Tannenbaum. But of course the definition of the semialgebraic set C is given by means of projections (because we define maximal points), and thus involve the theorem of Tarski-Seidenberg. Consequently on one hand the complexity (of a diagram) of C, which of course depends only on the complexity of the diagrams of A and B, may be big, and on the other hand in both proofs above it is difficult to write down some specific representation of C. Therefore we state below a weaker result, which we can prove, however, in a more constructive way.

Exercise 4.11. *Let $f = (f_1, \dots, f_m) : \mathbb{R}^n \to \mathbb{R}^m$, $n \geq m$ be a polynomial mapping,* $\deg(f_j) = d_j$. *Let A be a semialgebraic set inside some ball $B^n_r \subset \mathbb{R}^n$. Then for any $\xi > 0$ there exists a semialgebraic set C (depending on ξ) with the following properties:*

1. *$C \subset A_\xi$, the ξ-neighborhood of A.*
2. *The Hausdorff distance between $f(C)$ and $f(A)$ is at most $K\xi$, with K a Lipschitz constant of f on B^n_r.*
3. $\dim(C) \leqslant m$.
4. *Any ℓ-dimensional plane in $\mathbb{R}^n$ intersects C in at most ν connected components, where ν is explicit in terms of m, n, ℓ , d_k and $\bar{d}_{ij}$,, assuming that A is given as $\displaystyle\bigcup_{i=1}^{p} \bigcap_{j=1}^{q} \{p_{ij} \geq 0\}$ and* $\deg(p_j) = \bar{d}_{ij}$.

Hint. Find explicit equations for the α-neighborhood of the boundary of A, for the boundary of the α-neighborhood of A, and use the polar variety of these neighboroods relative to a convenient projection, to obtain the set C (see [Lê-Tei], [Hen-Mer] or [Hen-Mer-Sab] for the general theory of polar varieties).

Another result of the same type is the following. We have already stated this result in Chapter 1 (Theorem 1.1).

Theorem 4.12. *Let $A \subset \mathbb{R}^n$ be a semialgebraic set, B_r^n a ball of radius r. Then any two points x, y, in the same connected component of $A \cap B_r^n$, can be joined in $A \cap B_r^n$ by a by a piecewise smooth connected curve of length $\leqslant K \cdot r$, where K depends only on $D(A)$.*

Proof. (see [D'ac-Kur], [Den-Kur], [Har 3], [Kur], [Kur-Orr-Sim], [Tei 1], [Yom 1], [Yom 5]). We can assume that A is compact because it suffices to show the theorem for semialgebraic sets of the type: $\bigcup_{i=1}^{p} \bigcap_{j=1}^{j_i} A_{ij}$, where each A_{ij} has the form $\{(x_1, \ldots, x_n) \in \mathbb{R}^n;\ p_{ij}(x_1, \ldots, x_n) \geq 0\}$. Indeed, two points x and y in the same connected component of A may be joined by a curve s (this is a direct consequence of the triangulability of semialgebraic sets [Har 1], [Hir 2]), and if $p_{ij} > 0$ is an inequality defining A, since s is compact, we have $p_{ij|s} \geq \delta > 0$. Thus it suffices to prove the theorem for each set $A^\delta \subset A$, where the inequalities $\{p_{ij} > 0\}$ are replaced $\{p_{ij} \geq \delta\}$ in the definition of A^δ (of course $D(A^\delta) = D(A)$).

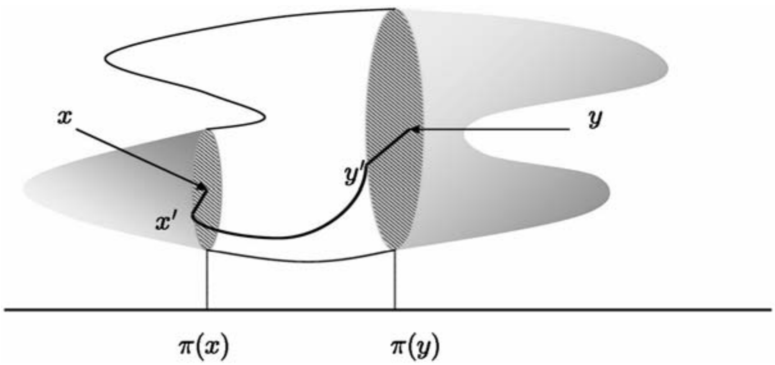

Fig. 4.2.

We propose proof by induction on the dimension of A (see Fig. 4.2) For this proof we do not need Theorem 4.10.

Lemma 4.13 is the first step of the induction. We assume $A \subset B_r^n$. Now take x and y in the same connected component of A, and let $\pi : A \to \mathbb{R}$ be the restriction of the standard projection $\mathbb{R}^n \to \mathbb{R}$. In the connected component of x in the fiber $\pi^{-1}(\pi(x))$ there exists a point x' of the boundary ∂A of A and in the same way in the connected component of y in the fiber $\pi^{-1}(\pi(y))$

there exists a point y' of the boundary of A. Now by induction hypothesis x and x' may be joined by an algebraic curve of length less than $K_1 \cdot r$, x' and y' by an algebraic curve of length less than $K_2 \cdot r$ (because $\dim(\partial A) < \dim(A)$), and y and y' by an algebraic curve of length less than $K_3 \cdot r$. Now the Theorem is a consequence of the following Lemma. □

Remark. Of course finding an explicit diagram of the boundary of a connected component of a given semialgebraic set is not an easy task (because Tarski-Seidenberg's principle is involved)! Consequently producing an explicit bound for the length of the curve provided by the proposed proof of Theorem 4.12 is, in general, a difficult problem, and the bound obtained, for sure, will not be sharp. We can find such an explicit bound in [D'Ac-Kur], for A the γ-critical set of f. We also propose an alternative method to produce such a bound in Chapter 1.

Lemma 4.13. Let Γ be a $(\mathcal{H}^\ell, \ell)$-rectifiable set in $B_r^n \subset \mathbb{R}^n$ (ie $\mathcal{H}^\ell$ almost of Γ is contained in the union of some countable family of ℓ-dimensional submanifolds of class 1 of $\mathbb{R}^n$) such that the number of points of $\Gamma \cap P$, where P is a generic $(n-\ell)$-plane in $\mathbb{R}^n$, is bounded by $B \in \mathbb{N}$. The ℓ-volume of Γ is $\leqslant C \cdot B \cdot r^\ell$, with C a constant depending only on n and ℓ.
In particular, the length of a semialgebraic curve of bounded complexity in B_r^n is bounded by $K \cdot r$.

Proof. The Lemma is a direct consequence of the Cauchy-Crofton formula for the volume: $Vol_\ell(\Gamma) = c \int_{P \in G_n^\ell} \int_{y \in \pi_P(\Gamma)} \mathrm{card}(\pi_P^{-1}(y) \cap \Gamma)\ dP$, where π_P is the orthogonal projection of $\mathbb{R}^n$ onto the ℓ-dimensional linear plane P of $\mathbb{R}^n$, and c depends only on n and ℓ (see Chapter 2). Hence we have: $Vol_\ell(\Gamma) \leqslant c.B \int_{P \in G_n^\ell} Vol_\ell(\pi_P(\Gamma))\ dP$. But $\Gamma \subset B_r^n$, thus $\pi_P(\Gamma) \subset B_r^\ell$, and $Vol_\ell(\pi_P(\Gamma)) \leqslant C'.r^\ell$, with $C' = Vol_\ell(B_1^\ell)$. □

Let us compare the statement of theorem 4.12 with another metric result on semialgebraic sets.

Theorem 4.14. ([Har 3], [Łoj]) *For any compact semialgebraic set, there are constants K and $\alpha > 0$, such that any x and y belonging to the same component of A, can be joined in A by a curve of length $\leqslant K||x-y||^\alpha$.*

Consider, for example, curves of degree 2 in the plane. For any such curve,the biggest exponent α in this theorem is 1: take x tending to y in A; the ratio $\mathrm{d}_A(x,y)/||x-y||^\alpha$, where $\mathrm{d}_A(x,y)$ is the distance from x to y in A, is bounded only for $\alpha \leqslant 1$. Furthermore, for any curve of degree 2 in the plane, there obviously exists a constant K such that any x and y in A can be joined by a curve in A of length $\leqslant K||x-y||$. But the constant K strongly depends on the concrete curve, as it is illustrated in Fig. 4.3. This is in contrast with the conclusion of Theorem 4.12: any two points on any parabola in B_r^2 can be joined by a curve of length $\leqslant C \cdot r$.

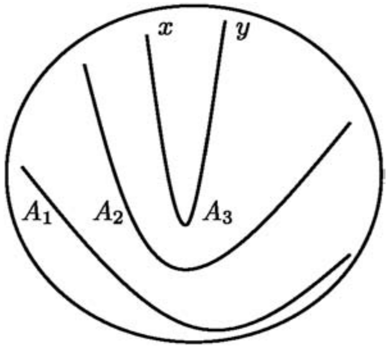

Fig. 4.3.

The fact that all the coefficients in our inequalities depend only on degrees, and not on a specific choice of the polynomials, is essential in our approach (see for instance the proof of Corollary 4.9).

More generally we say that a closed set A has the *Whitney property (with exponent α)*, if for each $a \in A$ there exists a neighbourhood U of a and two positive constants K and α such that any points x and y in U can be joined by a curve of length $\leqslant K.||x-y||^{\alpha}$.

It has been proved that each closed subanalytic set has the Whitney property in [Sta] (see also [Kur], in which the author proves that any subanalytic set A admits an analytic stratification such that each stratum has the Whitney property with $\alpha = 1$). For questions concerning the Whitney property of the geodesic distance see [Kur-Orr], and for questions concerning the subanalyticity or the regularity of the sub-Riemannian dictance see [Agr1], [Agr2] for instance.

The property mentioned in corollary 4.9 is classically called the Gabrielov property (see [Gab 4]). More precisely, we will say that a set A has the *local Gabrielov property* if for any $a \in A \subset \mathbb{R}^n$ there exists a neighbourhood U of a and an integer $B(= B(a,U))$ such that for any ℓ-dimensional affine plane P of $\bar{G}_n^{\ell}$, the number of connected components of $U \cap A \cap P$ is bounded by B.

If we can take $U = \mathbb{R}^n$ in the above definition, we will say that A has the *global Gabrielov property*

The Corollary 4.9 says that any semialgebraic set has the Gabrielov property (with explicit bound B depending only on ℓ and the degrees of the polynomials in the definitions of A).

However not only semialgebraic sets, but a very large class of sets has the Gabrielov property. Let us give two definitions:

Definition 4.15. ([Dri-Mil]) An *analytic-geometric category* $\mathcal{C}$ is the datum for each real analytic manifold M of the collection $\mathcal{C}(M)$ of sets of M such that the five following conditions are satisfied (for each real analytic manifold N):

AG1 $\mathcal{C}(M)$ is a boolean algebra (for $\cup, \cap$) of subsets of M and $M \in \mathcal{C}(M)$.

AG2 If $A \in \mathcal{C}(M)$, then $A \times \mathbb{R} \in \mathcal{C}(M \times \mathbb{R})$.

AG3 If $f : M \to N$ is a proper real analytic map and $A \in \mathcal{C}(M)$, then $f(A) \in \mathcal{C}(N)$.

AG4 If $A \subset M$ and $(U_i)_{i \in I}$ is an open covering of M, then $A \in \mathcal{C}(M)$ if and only if $A \cap U_i \in \mathcal{C}(M)$ for all $i \in I$.

AG5 Every bounded set in $\mathcal{C}(\mathbb{R})$ has finite boundary.

Let us note that the subanalytic sets of any real analytic manifold M (ie sets which locally are projections of relatively compact semianalytic sets. The term subanalytic was introduced by Hironaka in [Hir 1], but the notion has been first considered by Thom in [Tho 2] and Łojasiewicz [Łoj]. See for instance [Bie-Mil 5] or [Den-Sta]) form an analytic-geometric category, which is the smallest analytic-geometric category.

Because in AG3 we allow analytic functions, and AG5 concerns only bounded sets of $\mathcal{C}(\mathbb{R})$, the behaviour at infinity of sets in analytic-geometric categories is not controlled, and these sets are not as globally nice as semialgebraic sets. For globally nice sets we have the following categories:

Definition 4.16. ([Dri], [Dri-Mil]) A *structure on the real field* $(\mathbb{R}, +, .)$ is a sequence $(S_n)_{n \in \mathbb{N}}$ such that for each $n \in \mathbb{N}$:

S1 S_n is a boolean algebra of subsets of $\mathbb{R}^n$, with $\mathbb{R}^n \in S_n$.

S2 S_n contains the diagonal $\{(x_1, \dots, x_n) \in \mathbb{R}^n ; x_i = x_j\}$ for $1 \leqslant i < j \leqslant n$.

S3 If $A \in S_n$, then $A \times \mathbb{R}$ and $\mathbb{R} \times A$ belong to S_{n+1}.

S4 If $A \in S_{n+1}$ then $\pi(A) \in S_n$, where $\pi : \mathbb{R}^{n+1} \to \mathbb{R}^n$ is the standard projection.

S5 S_3 contains the graphs of addition and multiplication.

The structure is said to be *o-minimal* if it satisfies the following additional axiom:

S6 (o-minimal axiom) S_1 consists of the finite unions of intervals of all kinds.

Roughly speaking, each analytic-geometric category gives rise to an o-minimal structure, after compactification (see [Dri-Mi]).

The smallest structure $\mathcal{S}(\mathbb{R}, +, .)$ on $(\mathbb{R}, +, .)$ consists of the semialgebraic sets defined over $\mathbb{Q}$ (the p_{ij}'s are in $\mathbb{Q}[X_1, \dots, X_n]$ in the Definition 4.1).

We can construct structures satisfying axioms S1 to S5 in the following way: we consider a family of functions $(f_j)_{j \in J}$ and the smallest structure on $(\mathbb{R}, +, .)$ containing the graphs of the f_j's. When $f_j = j \in \mathbb{R}$, the constant function which equals j, we obtain the family of semialgebraic sets (this is nothing else than the Tarski-Seidenberg theorem).

When f_j ranges over all restrictions of analytic functions on closed balls of $\mathbb{R}^n$, we obtain the family of globally subanalytic sets, denoted $\mathcal{S}(\mathbb{R}_{an})$.

If we consider in addition the function $x \mapsto e^x$, we obtain the so-called Log-Analytic structure, denoted $\mathcal{S}(\mathbb{R}_{an,exp})$.

Of course we have: $\mathcal{S}(\mathbb{R},+,.) \subset \mathcal{S}(\mathbb{R}_{an}) \subset \mathcal{S}(\mathbb{R}_{an,exp})$. It has been proved in [Wil] that the structure $\mathcal{S}(\mathbb{R}_{an,exp})$ is o-minimal (so are the structures $\mathcal{S}(\mathbb{R}_{an})$, $\mathcal{S}(\mathbb{R},+,.)$ and the structure consisting of all semialgebraic sets).

See also [Shi], for an interesting and slightly different (actually, a more general) viewpoint.

Definition 4.17. We will say that a set belonging to an analytic-geometric category or to an o-minimal structure is a *tame set* (see [Tei 2]).

We have the following general result:

Theorem 4.18. Every tame set has the local Gabrielov property (the global Gabrielov property if the set is in an o-minimal structure).

Proof. We first prove the local property, hence we suppose that the tame set $A \subset \mathbb{R}^n$ lies in a closed ball B_r^n of radius $r > 0$. We denote by $\bar{G}_n^\ell(r)$ the subset of affine ℓ-dimensional planes of $\mathbb{R}^n$ which encounter the ball B_r, this set is compact, and we denote by B_r^ℓ the closed ball of radius r centered at the origin of $\mathbb{R}^\ell$. We suppose that $\ell < n$, because for $\ell = n$ the Gabrielov property just says that A has a finite number of connected components, which is true as an easy consequence of Axiom $S6$.

Let us consider the partition $E \cup F$ of the compact set $\bar{G}_n^\ell(r) \times B_r^\ell$, where:

$$E = \{(P,x) \in \bar{G}_n^\ell(r) \times B_r^\ell;\quad x \in A \cap P\}$$

$$F = \{(P,x) \in \bar{G}_n^\ell(r) \times B_r^\ell;\quad x \notin A \cap P\},$$

and finally let us denote $\pi : \bar{G}_n^\ell(r) \times B_r^\ell :\to \bar{G}_n^\ell(r)$ the standard projection.

Now, A being in an analytic-geometric category or in an o-minimal category, the proper map π (its graph) lies in the same category, and thus admits a Whitney stratification (see [Dri-Mil], [Loi]): there exists a tame stratification Σ of $\bar{G}_n^\ell(r) \times B_r^\ell$ compatible with the family (E,F) (E and F are unions of strata) and a tame stratification Σ' of $\bar{G}_n^\ell(r)$ such that for each stratum $\sigma' \in \Sigma'$, $\pi^{-1}(\sigma')$ is a union of strata and the restriction of π to each of these strata is a submersion over σ' (of course this property does not require the Gabrielov property!).

The first isotopy lemma of Thom-Mather ([Tho 2], theorem 1G1; [Ma], proposition 3.11) gives us a local topological trivialisation of π over each stratum $\sigma' \in \Sigma'$, furthermore this trivialisation is compatible with Σ. The stratification Σ' being locally finite and $\bar{G}_n^\ell(r)$ being compact, Σ' is finite. Let us write $\Sigma' = \{\sigma'_1, \ldots, \sigma'_s\}$. For any $q \in \{1, \ldots, s\}$ and any $P, P' \in \sigma_q$, the fibers $\pi^{-1}(P)$ and $\pi^{-1}(P')$ are homeomorphic, and the homeomorphism preserves the sets E and F. In particular the number of connected components of $A \cap P$ and $A \cap P'$ is the same, and of course this number is finite for tame sets by Axiom $S6$, as mentionned above. Consequently the number of connected components of $A \cap P$ is uniformly bounded with respect to P.

If A lies in an o-minimal structure, the standard (semialgebraic) compactification of A in S^n (for instance) is also o-minimal, thus we can take $r = \infty$ in the proof above. $\square$

This proof is short, but it uses two deep results: the existence of Whitney stratifications of a proper tame map and the first isotopy lemma of Thom-Mather. Of course, as mentionned in the proof, these results do not require the Gabrielov property (at least for planes $\neq \mathbb{R}^n$). For the semialgebraic case, we may refer to [Har 2], that do not requires the firts isotopy lemma and provides a method without integration of vector fields.

In what follows we will use the global Gabrielov property.

5 Variations of Semialgebraic and Tame Sets

Abstract. We study here multidimensional variations of semialgebraic and tame sets, partly following [Vit 1], [Iva 1]. The stress is lain on properties which distinguish tame sets, in particular, correlations between variations of ϵ-neighborhood, comparison of variations of two sets with a small Hausdorff distance etc...

The property which makes computation of variations of a semialgebraic or an o-minimal set A tractable, is that the number of connected components of A, as well as of its sections, is uniformly bounded (in terms of the diagram of A, if A is semialgebraic). We call this property the global Gabrielov property (see Chapter 4, Theorem 4.18). We recall the bound given for semialgebraic sets by Corollary 4.9:

- *For a semialgebraic set $A \subset \mathbb{R}^n$ with diagram:*
$$\big(n, p, j_1, \dots, j_p, (d_{ij})_{\substack{i=1,\dots,p \\ j=1,\dots,j_i}}\big),$$
the number of connected components of $A \cap P$, where P is a ℓ-plane of $\mathbb{R}^n$ ($1 \leqslant \ell \leqslant n$), is bounded by $B_{0,\ell}(A)$, the infimum over all representations of A of the numbers $\frac{1}{2}\sum_{i=1}^{p}(d_i+2)(d_i+1)^{\ell-1}$, with $d_i = \sum_{j=1}^{j_i} d_{ij}$ (Corollary 4.9).
- *For a tame set A in an o-minimal structure, we denote also by $B_{0,\ell}(A)$ the uniform bound of the number of connected components of $A \cap P$, for P an ℓ-plane of $\mathbb{R}^n$ (Theorem 4.18).*

Now we can prove the first among the specific properties of variations for global tame sets.

Theorem 5.1. *Let $A \subset \mathbb{R}^n$ be a tame set, $B \subset \mathbb{R}^n$ a bounded and closed set. If $A \subset B$, then $V_i(A) \leqslant B_{0,n-i}(A).V_i(B)$ (for all $i \in \{0, \dots, n\}$).*

Proof. The assumption that B is closed implies that $V_0(B)$ is a measurable function (see [Zer]); since $A \subset B$ is tame and B is bounded, A is in fact globally tame. We have:

$$V_i(A) = c(n,i) \int_{P \in G_n^i} \int_{x \in P} V_0(A \cap P_x)\ dx\ dP$$

Y. Yomdin and G. Comte: LNM 1834, pp. 59–73, 2004.

$$\leqslant c(n,i)B_{0,n-i}(A)\int_{P\in G_n^i}\int_{x\in P}\mathbf{1}_{\pi_{P(A)}}(x)\ \ dx\ dP,$$

where $\mathbf{1}_{\pi_{P(A)}}$ is the indicator function of the projection $\pi_{P(A)}$ of A onto P. But since $A\subset B$, for any $P\in G_n^i$, $\mathbf{1}_{\pi_{P(A)}}\leqslant\mathbf{1}_{\pi_{P(B)}}$, and we can continue our inequalities:

$$V_i(A)\leqslant c(n,i)B_{0,n-i}(A)\int_{P\in G_n^i}\int_{x\in P}\mathbf{1}_{\pi_{P(B)}}(x)\ \ dx\ dP$$

$$\leqslant c(n,i)B_{0,n-i}(A)\int_{P\in G_n^i}\int_{x\in P}V_0(B\cap P_x)\ \ dx\ dP=B_{0,n-i}(A)V_i(B),$$

□

Remark. Of course no inequality of this type holds for nontame sets: e.g., we can take as A a curve of infinite length in the unit square B in $\mathbb{R}^2$ and then $V_1(A)=\infty$, $V_1(B)\leqslant c(2,1)\sqrt{2}=\dfrac{\pi}{\sqrt{2}}$.

We find again 4.13, as a corollary of 5.1:

Corollary 5.2. *Let $A\subset\mathbb{R}^n$ be a tame set of dimension ℓ. Then for any ball B_r^n of radius r in $\mathbb{R}^n$,*

$$Vol_\ell(A\cap B_r^n)=V_\ell(A\cap B_r^n)\leqslant Vol_\ell(B_1^\ell)\,.\,c(n,k)\cdot B_{0,n-\ell}(A)\,.\,r^\ell.$$

□

Remark. The Corollary shows that a semialgebraic subset $A\subset B_1^n$ of given complexity, does not have too big volume, or conversely that a semialgebraic subset $A\subset B_1^n$ of big volume has big complexity.

When A is a bounded tame set of dimension ℓ, the projection $\pi_P(A)$ of A onto the ℓ-dimensional vector plane P splits into a finite number of tame domains, say $K_1^P,\dots,K_{n_P}^P$, such that A covers each K_j^P with the multiplicity $E_j^P\in\mathbb{N}$ and the numbers n_P and E_j^P are uniformly bounded as P ranges over all elements of G_n^ℓ, by the same arguments as in 4.18. The function: $s_P=\displaystyle\sum_{j=1}^{n_P}E_j^P\mathbf{1}_{K_j^P}$ is a tamely constructible function (there exists a (finite) tame stratification of P such that s_P is constant on each stratum).

The representation of this function, of course, is not unique, but by additivity, its volume defined as $Vol(s_P)=\sum_{j=1}^{n_P}E_j^P Vol_\ell(K_j^P)$ is well defined. Now by definition of the ℓ-variation of A, we have

$$V_\ell(A)=c(n,\ell)\int_{P\in G_n^\ell}Vol(s_P)\ \ dP,$$

and the classical Cauchy-Crofton formula ([Fed 1], 5.22 or [Fed 2], 2.10.15) gives:

$$Vol_\ell(A)=V_\ell(A)=c(n,\ell)\int_{P\in G_n^\ell}Vol(s_P)\ \ dP.$$

Instead of considering bounded tame sets, one can consider germs of tame sets (at the origin, for simplicity). Let us denote the germ of A by A_0.

Note that the projection of such a germ onto a given ℓ-dimensional vector plane P is not well defined (for instance the projection (in some chart) onto $\mathbb{R}^2$ of the germ at a point $(a,0) \in \mathbb{P}^1 \times \mathbb{R}^2$ of the blowing-up of the plane $\mathbb{R}^2$ at the origin depends on the chosen representation of the germ. Following the terminology of Thom, this mapping is not "sans éclatement"). However, the projection of A_0 is well defined for generic directions, the transverse directions, namely projections onto planes P such that $P^\perp \cap \mathcal{C}_0 A_0 = \{0\}$, where $\mathcal{C}_0 A_0$ is the tangent cone of A_0 at the origin (the tangent cone of the germ A_0 is the semi-cone of $\mathbb{R}^n$ consisting of all limit secants to A_0 at the origin).

It follows that a generic projection $\pi_P(A_0)$ of the germ A_0 splits into a finite number of germs of tame domains $k_1^P, \dots, k_{\nu_P}^P$, such that A_0 covers the germ k_j^P with multiplicity $e_j^P \in \mathbb{N}$. With the same arguments as in 4.18 (the generic local topological triviality provided by the first isotopy lemma of Thom) one can show that the numbers ν_P and e_j^P are uniformly bounded as P ranges over all elements of G_n^ℓ (in fact one can easily see that $\nu_P \leqslant n_P$ and $e_j^P \leqslant E_j^P \leqslant B_{0,n-\ell}(A)$, for any bounded set A that represents A_0).

Now we can define, as above, a tamely constructible function-germ $\sigma_P = \sum_{j=1}^{\nu_P} e_j^P \mathbf{1}_{k_j^P}$, which is the "localization" of s_P. The volume of σ_P makes no sense, but one can define the "localisation of the volume of σ_P": the density $\Theta_\ell(\sigma_P)$ (or the Lelong number) of σ_P by the following formulas:

$$\Theta_\ell(\sigma_P) = \sum_{j=1}^{\nu_P} e_j^P \Theta_\ell(k_j^P),$$

where the density at the origin $\Theta_\ell(B_0)$ of a tame germ B_0 is by definition the limit when r tends to 0, of the ratio: $\dfrac{Vol_\ell(B \cap B_r^n)}{Vol_\ell(B_r^\ell)}$, (this limit exists for tame sets, by [Kur-Rab]), for any set B which represents the germ B_0.

One can prove the following "localized" version of the Cauchy-Crofton formula for the volume (see [Com 2], [Com 3]):

$$\Theta_\ell(A_0) = \int_{P \in G_n^\ell} \Theta_\ell(\sigma_P) \; dP.$$

Notice that the measure that occurs in this formula is the probability measure dP on G_n^ℓ, and not $c(n,\ell)\, dP$ as in the Cauchy-Crofton formula for the volume.

Actually, one has a more general result:

Theorem 5.3. **([Com 2], [Com 3])** *Let A_0 be a germ at the origin of a tame set of dimension ℓ in $\mathbb{R}^n$, $\mathcal{G} \subset G_n^\ell$ a tame set, $G < \mathcal{O}_n(\mathbb{R})$ a group acting transitively on $\mathcal{G}$ and μ a G-invariant measure on $\mathcal{G}$. Suppose that* *the tangent spaces of $\mathcal{C}_0 A_0$ are in $\mathcal{G}$, that there exists $P^0 \in \mathcal{G}$ such that $\{g \in G;\ g.P^0 = P^0\}$ acts transitively on P^0, and that $\mu(\mathcal{G}) = \mu(\mathcal{G} \cap \mathcal{T}_{A_0}) = 1$ (where $\mathcal{T}_{A_0}$ is the set of all transverse projections to A_0). Then we have:*

$$\Theta_\ell(A_0) = \int_{P \in \mathcal{G}} \Theta_\ell(\sigma_P)\ \ d\mu(P).$$

In particular, when A_0 is the germ of a complex analytic set A in $\mathbb{C}^n \sim \mathbb{R}^{2n}$, the function σ_P is a function that almost everywhere equals $e(A, 0)$, the local multiplicity of A at the origin (see [Whi 4]), for generic P in the complex Grassmann manifold $\widetilde{G}_n^\ell \subset G_{2n}^{2\ell}$. Taking $\mu = d\tilde{P}$ the invariant probability measure on $\widetilde{G}_n^\ell$ in Theorem 5.3, we obtain a theorem of Draper, which states that the Lelong number of a complex analytic set is its multiplicity:

Theorem 5.4. ([Dra], [Dem]) With the above notations, we have:

$$\Theta_{2\ell}(A_0) = e(A, 0).$$

□

Remarks.

- These results illustrate a general principle: global data on algebraic sets, such as the number of connected components in a plane subspace of the ambient space, uniform bounds for these numbers, Betti numbers, etc... are related to the degree, whereas the localization of these data, or local invariants, are related to the multiplicity.
- It is a direct consequence of Corollary 5.2 and of the definition of the density, that a semialgebraic subset $A \subset \mathbb{R}^n$ has its density bounded by its complexity, namely:
$$\Theta_\ell(A_0) \leqslant c(n,k) B_{0,n-\ell}.$$
However, one can obtain a better bound for the density of A in terms of complexity, by using the Cauchy-Crofton formula for the density (Theorem 5.3):

Theorem 5.5. *Let A_0 be a germ of a tame set of dimension ℓ at the origin of $\mathbb{R}^n$. The following upper bound for the density $\Theta_\ell(A_0)$ of A_0 holds:*

$$\Theta_\ell(A_0) \leqslant B_{0,n-\ell}.$$

Proof. We just notice that $\Theta_\ell(\sigma_P) \leqslant B_{0,n-\ell}$. □

Remark. We can find a better bound for $\Theta_\ell(A_0)$, say $\Theta_\ell(A_0) \leqslant \beta_{0,n-\ell}$, where $\beta_{0,n-\ell}$ has the same nature as $B_{0,n-\ell}$, the multiplicity m_{ij} of the polynomials p_{ij} replacing the degrees d_{ij}. We just have to prove the local version of Bezout's theorem, which only involves the multiplicity m_{ij}.

Now let us go back to the study of the entropy of tame sets.

In general, the fact that the measure $m(A)$ is small says nothing about $M(\epsilon, A)$ – take A to be the set of rational points in $[0;1]$. But for semialgebraic and tame sets, if the volume is small, the contribution to the ϵ-entropy of its "geometric complexity" is bounded by the algebraic complexity of this set (by its diagram in the semialgebraic case).

Proposition 5.6. *Let $A \subset B_r^n \subset \mathbb{R}^n$ be a tame set. Then:*

$$\begin{aligned} M(\epsilon, A) &\leqslant C(n) Vol_n(A).(\frac{1}{\epsilon})^n \\ &\quad + C(n).(\frac{1}{\epsilon})^{n-1} \sum_{j=0}^{n-1} B_{0,n-j}(A).Vol_j(B_1^j).r^j.\epsilon^{n-j-1} \\ &\leqslant C(n) Vol_n(A).(\frac{1}{\epsilon})^n + C[1 + (\frac{r}{\epsilon})^{n-1}], \end{aligned}$$

where $C(n)$ is given by 3.4 *and* 3.5, *and C depends only on $D(A)$ in the semialgebraic case .*

Proof. By Theorem 3.5 we have:

$$\begin{aligned} M(\epsilon, A) &\leqslant C(n) \sum_{i=0}^{n} \frac{1}{\epsilon^i} \cdot V_i(A) \leqslant C(n) \cdot Vol_n(A) \cdot (\frac{1}{\epsilon})^n \\ &\quad + C(n) \sum_{j=0}^{n-1} B_{0,n-j}(A) \cdot Vol_j(B_1^j) \cdot (\frac{r}{\epsilon})^j, \end{aligned}$$

the last inequality being a consequence of 5.2. But the last term of this sum is bounded by:

$$C(n) \cdot B(A) \sum_{j=0}^{n-1} (\frac{r}{\epsilon})^j \leqslant n \cdot C(n) \cdot B(A)(1 + (\frac{r}{\epsilon})^{n-1}),$$

where $B(A) = \max_{j \in \{1,\dots,n-1\}} \big(B_{0,n-j}(A) \cdot Vol_j(B_1^j)\big)$, which proves the last inequality with $C = n \cdot C(n) \cdot B(A)$. □

If we know *a priori* that our semialgebraic set has dimension less than n, we get nontrivial bounds on its ϵ-entropy without any specific information, except its algebraic complexity:

Corollary 5.7. *Let $A \subset \mathbb{R}^n$ be a tame set of dimension $\ell < n$. Then for any ball $B_r^n \subset \mathbb{R}^n$:*

$$M(\epsilon, A \cap B_r^n) \leqslant C(n) \sum_{j=0}^{\ell} B_{0,n-j}(A) Vol_j(B_1^j)(\frac{r}{\epsilon})^j \leqslant C[(\frac{r}{\epsilon})^\ell + 1],$$

where C depends only on the diagram $D(A)$ of A in the semialgebraic case.

Proof. It follows immediately from Proposition 5.6, taking into account that $V_j(A) = 0$, for $j > \ell$. □

Thus semialgebraic sets of fixed algebraic complexity and of dimension less than n, cannot be "too dense" inside the ball B_r^n.

In the same way, when a semialgebraic set $A \subset \mathbb{R}^n$ approaches a given set $B \subset \mathbb{R}^n$, the dimension of A and the complexity of A cannot be too small. More precisely, we have:

Proposition 5.8. (see also [Zer]) *Let $A \subset B_1^n \subset \mathbb{R}^n$ be a semialgebraic set, $B \subset \mathbb{R}^n$ and $1 > \eta > 0$ such that $B \subset A_\eta$ (A_η being the η-neighborhood of A). Then we have:*

$$M(\eta, B) \leqslant (\frac{4}{\eta})^k C(n).\nu(k).\alpha(n),$$

where k is the dimension of A, $C(n)$ is given by Theorem 3.5, $\nu(k) = \sum_{i=0}^{k} c(n,i).Vol_i(B_1^i)$, and $\alpha(n) = \frac{1}{2}\sum_{i=1}^{p}(d_i+2)(d_i+1)^{n-1}$, assuming that A has a diagram $\big(n, p, j_1, \dots, j_p, (d_{ij})_{\substack{i=1,\dots,p \\ j=1,\dots,j_i}}\big)$.

Proof. Let us write $M = M(4\eta, B)$, and let $x_1, \dots, x_M$, let be M points in B such that $\mathrm{d}(x_i, x_j) > 4\eta$, for $i \neq j$. By assumption, there exist $y_1, \dots, y_M$ in A with $\mathrm{d}(x_j, y_j) < \eta$. The balls B_j of center y_j and radius η are disjoint, thus we have, by Proposition 3.3:

$$V_i(A) \geq \sum_{j=1}^{M} V_i(A, B_j), \quad \text{for all } \ i = 0, \dots, k.$$

It follows that:

$$\sum_{i=0}^{k} \frac{V_i(A)}{\eta^i} \geq \sum_{j=1}^{M}\sum_{i=0}^{k} \frac{V_i(A, B_j)}{\eta^i} \geq M.c(n), \quad \text{(by Theorem 3.4)}.$$

On the other hand, by Theorem 5.2,

$$V_i(A) \leqslant c(n,i).\alpha(n-i).Vol_i(B_1^i),$$

hence:

$$M(4\eta, B) \leqslant \frac{1}{c(n)}.\sum_{i=0}^{k} c(n,i).\alpha(n-i)\frac{Vol_i(B_1^i)}{\eta^i}$$

$$\leqslant \frac{1}{c(n)}.\alpha(n)\sum_{i=0}^{k}\frac{c(n,i).Vol_i(B_1^i)}{\eta^k}.$$ □

Remarks.

- Proposition 5.8 implies that when $M(\epsilon, B)$ is big, k and the degrees d_{ij} cannot be too small.
- If for all $\epsilon > 0$, there exists a semialgebraic set A^ϵ of dimension k such that $B \subset (A^\epsilon)_\epsilon$, and if the complexity of A^ϵ is uniformly bounded, then $\dim_e(B) \leqslant k$.

The properties of the ϵ-entropy allow us easily to get bounds also for the volume of η-neighborhoods of tame sets.

Theorem 5.9. *Let A be a tame set of dimension $\ell < n$. Then for any ball $B_r^n \subset \mathbb{R}^n$ and for any $\eta > 0$, the volume of the η-neighborhood of $A \cap B_r^n$ is bounded as follows:*

$$Vol_n\big((A \cap B_r^n)_\eta\big) \leqslant C'(n) \sum_{j=0}^{\ell} B_{0,n-j}(A) \cdot r^j \cdot \eta^{n-j} \leqslant C' \cdot [r^\ell \eta^{n-\ell} + \eta^n],$$

with C' depending only on $D(A)$ in the semialgebraic case.

Proof. We have: $Vol_n((A \cap B_r^n)_\eta) \leqslant Vol_n(B_{2\eta}^n) \cdot M(\eta, A \cap B_r^n)$, since $M(\eta, A \cap B_r^n)$ balls of radius 2η cover the η-neighborhood of $A \cap B_r^n$. Applying the result of Corollary 5.7, we obtain the required inequality. □

Naturally, this expression is quite similar to the Weyl or Steiner formula for volume of tubes (see [Wey], [Ste]).

Turning back to variations, we notice that in general if two sets are closed in the Hausdorff metric, their variations may be quite different: take once more $A = [0;1]$, and $B = \mathbb{Q} \cap A$. Then $d_{\mathcal{H}}(A, B) = 0$, but $V_1(A) = 1$, $V_1(B) = 1$ and $V_0(A) = 1$, $V_0(B) = \infty$. Another example is the following: take $A = [0;1]$ and B a spiral in $A \times [0;\epsilon]$ of infinite length. Then $V_0(A) = V_0(B) = 1$, but $V_1(A) = 1$ and $V_1(B) = \infty$.

However, for semialgebraic sets of a fixed algebraic complexity, or for tame sets, the situation is different:

Theorem 5.10. *Let $A, B \subset B_r^n$ be two tame sets. If $d_{\mathcal{H}}(A, B) = \delta$, then $|Vol_n(A) - Vol_n(B)| \leqslant C(\delta r^{n-1} + \delta^n)$, with C depending only on the bounds $B_{0,j}(\partial A)$ and $B_{0,j}(\partial B)$, and thus on $D(A)$ and $D(B)$, when A and B are semialgebraic. In particular, $|Vol_n(A) - Vol_n(B)| \xrightarrow[\delta \to 0]{} 0$.*

Proof. The hypothesis $d_{\mathcal{H}}(A, B) = \delta$ implies that $(A \setminus B) \cup (B \setminus A) \subset (\partial A \cup \partial B)_\delta$. Hence

$$|Vol_n(A) - Vol_n(B)| \leqslant Vol_n(A \setminus B) + Vol_n(B \setminus A) \leqslant Vol_n\big((\partial A \cup \partial B)_\delta\big).$$

But ∂A and ∂B are tame sets of dimension less than $n - 1$ (with diagrams depending only on $D(A)$ and $D(B)$, in the semialgebraic case), hence by Theorem 5.9, we have:

$$|Vol_n(A) - Vol_n(B)| \leqslant C[r^{n-1}\delta + \delta^n].$$

$\square$

Also taking the δ-neighborhoods for tame sets does not change variations too much:

Theorem 5.11. *Let $A \subset B_r^n$ be a tame set. Then, for all $i \in \{1, \ldots, n\}$:*

$$V_i(A_\delta) \leqslant V_i(A) + \widehat{C}(\delta r^{i-1} + \delta^i), \qquad V_0(A_\delta) \leqslant V_0(A),$$

where the constant $\widehat{C}$ depends only on constants of the type $B_{0,n-j}$, and thus in the semialgebraic case, only on $D(A)$. In particular $\lim_{\delta \to 0} V_i(A_\delta) \leqslant V_i(A)$. (For A closed, in fact $\lim_{\delta \to 0} V_i(A_\delta) = V_i(A)$)

Proof. We denote below C, C' etc..., the constants of type $B_{0,n-j}$ (depending only on $D(A)$ in the semialgebraic case). We have:

$$V_i(A_\delta) = c(n,i) \int_{P \in G_n^i} \int_{x \in P} V_0(A_\delta \cap P_x)\, dx\ dP.$$

Let us estimate the inner integral. Let π_P denote the orthogonal projection onto P. Of course, we have $\pi(A) \subset \pi(A_\delta) \subset \pi(A)_\delta$. By theorem 5.9:

$$Vol_i(\pi_P(A_\delta) \setminus \pi_P(A)) \leqslant C(\delta r^{i-1} + \delta^i).$$

We have:

$$\int_{x \in P} V_0(A_\delta \cap P_x)\, dx =$$

$$\int_{x \in \pi_P(A_\delta) \setminus \pi_P(A)} V_0(A_\delta \cap P_x)\, dx + \int_{x \in \pi_P(A)} V_0(A_\delta \cap P_x)\, dx.$$

Since A_δ is tame, we have the global Gabrielov property: there exists C' such that $V_0(A_\delta \cap P_x) \leqslant C'$, and consequently the first integral is bounded by:

$$C'C(\delta r^{i-1} + \delta^i).$$

Now let $\Sigma \subset \pi_P(A)$ be the set of $x \in P$, for which $V_0(A_\delta \cap P_x) > V_0(A \cap P_x)$. Let us prove the following lemma:

Lemma 5.12. Let $\Sigma \subset \Delta_\delta$, where Δ is the set of critical values of $\pi_P : A \to P$. (Of course, the singular points of A are included in the set of critical points of $\pi_P : A \to P$).

Proof of Lemma 5.12. If for some $x \in \Sigma$ there is a component of $A_\delta \cap P_x$, not containing points of $A \cap P_x$, then at a distance at most δ from this component, there is a point $y \in A$, at which the distance to P_x achieves its local minimum. Clearly y is a critical point of $\pi_{P|A}$. Thus $\pi_P(y) \in \Delta$, and $\mathrm{d}(x, \pi_P(y)) \leqslant \delta$. $\square$

Now:

$$\int_{x\in\pi_P(A)} V_0(A_\delta\cap P_x)\,dx = \int_{x\in\Delta_\delta} V_0(A_\delta\cap P_x)\,dx + \int_{x\in\pi_P(A)\setminus\Delta_\delta} V_0(A_\delta\cap P_x)\,dx.$$

But $\dim(\Delta) < i$ (by Sard's theorem), hence: $Vol_i(\Delta_\delta) \leqslant C''[\delta r^{i-1} + \delta^i]$, by theorem 5.8, and once more:

$$\int_{x\in\Delta_\delta} V_0(A_\delta \cap P_x)\,dx \leqslant C'C''(\delta r^{i-1} + \delta^i).$$

Furthermore:

$$\int_{x\in\pi_P(A)\setminus\Delta_\delta} V_0(A_\delta \cap P_x)\,dx \leqslant \int_{x\in\pi_P(A)\setminus\Delta_\delta} V_0(A \cap P_x)\,dx$$

$$\leqslant \int_{x\in\pi_P(A)} V_0(A \cap P_x)\,dx.$$

Combining these estimates, we get:

$$\int_{x\in P} V_0(A_\delta \cap P_x)\,dx \leqslant \int_{x\in P} V_0(A \cap P_x)\,dx + \bar{C}(\delta r^{i-1} + \delta^i),$$

which substituting into the integral over G_n^i gives:

$$V_i(A_\delta) \leqslant V_i(A) + \widehat{C}(\delta r^{i-1} + \delta^i).$$

□

Of course, nothing of this type holds for non-tame sets: for $A = \mathbb{Q}\cap[0;1]$, $A_\delta = [-\delta; 1+\delta]$, and $V_1(A_\delta) = 1 + 2\delta$, $V_1(A) = 0$.

As an easy consequence of Theorem 5.11 we obtain the following extension of Theorem 5.10 to lower variations:

Theorem 5.13. *Let $A, B \subset B_r^n$ be tame sets. Then if $d_{\mathcal{H}}(A,B) = \delta$,*

$$V_i(A) \leqslant \widetilde{C}\big(V_i(B) + \delta r^{i-1} + \delta^i\big)$$

$$V_i(B) \leqslant \widetilde{C}\big(V_i(A) + \delta r^{i-1} + \delta^i\big),$$

where $\widetilde{C}$ depends only on constants of the type $B_{0,j}$, and thus only on $D(A)$ and $D(B)$ when A and B are semialgebraic. In particular, if $d_{\mathcal{H}}(A,B) \to 0$, then $\limsup V_i(A) \leqslant \widetilde{C} V_i(B)$ and $\limsup V_i(B) \leqslant \widetilde{C} V_i(A)$.

Proof. We have $A \subset B_\delta$ and $B \subset A_\delta$ and applying Theorem 5.11 and Theorem 5.1, we obtain the required inequality.

The following easy example shows that for lower variations V_i, $i < n$, the constant $\widetilde{C}$ in the bounds of Theorem 5.13 cannot be eliminated.

$$\text{Take } A, B \subset \mathbb{R}^2, \quad A = B_1^2 \cap \{y^2 = 0\}, \quad B = B_1^2 \cap \{y^2 = \delta^2\}.$$

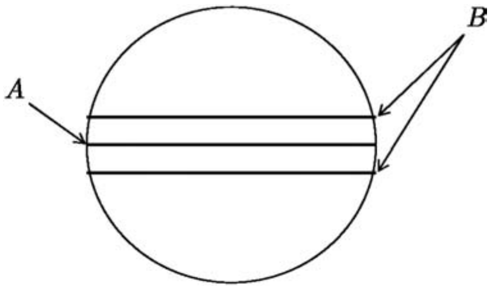

Fig. 5.1.

Then:

$$d_{\mathcal{H}}(A,B)=\delta, \quad V_1(A)=1, \quad V_1(B) \underset{\delta\to 0}{\to} 2. \qquad \text{(see Fig. 5.1)}$$

In fact, for tame sets, ϵ-entropy and variations are equivalent in the following sense:

Theorem 5.14. *For A a bounded tame set of dimension ℓ, we have:*

$$C_1 \sum_{i=0}^{\ell} V_i(A).(\frac{1}{\epsilon})^i \leqslant M(\epsilon,A) \leqslant C_2 \sum_{i=0}^{\ell} V_i(A).(\frac{1}{\epsilon})^i,$$

where C_1 and C_2 depend only on constants of the type $B_{0,j}$, and thus only on $D(A)$ in the semialgebraic case.

Proof. The upper bound is given in 3.5; it remains to prove the lower bound.

We have: $A \subset \bigcup_{j=1}^{q} B_j$, with $q=M(\epsilon,A)$ and B_j some balls of radius ϵ. Hence, by the property (7) of variations,

$$V_i(A) \leqslant \sum_{j=1}^{q} V_i(B_j \cap A).$$

But by Theorem 5.1: $V_i(B_j\cap A) \leqslant c(n,i).B_{0,n-i}(A\cap B_j).Vol_i(B_1^i).\epsilon^i = C.\epsilon^i$, hence:

$$V_i(A) \leqslant qC.\epsilon^i \quad \text{or} \quad M(\epsilon,A)=q \geq \frac{1}{C}V_i(A).\frac{1}{\epsilon^i}.$$

Adding all these inequalities for $i=0,\dots,\ell$, we obtain:

$$M(\epsilon,A) \geq C_1 \sum_{i=1}^{\ell} V_i(A).\frac{1}{\epsilon^i}.$$

□

Remark. While the upper bound for ϵ-entropy in terms of variations is valid for any set, the lower holds only for “simple” ones. For instance for

$A = [0;1] \cap \mathbb{Q}$, the set of rational points in $[0;1]$, $V_0(A) = \infty$, but $M(\epsilon, A) = M(\epsilon, [0;1]) = [\frac{1}{\epsilon}] + 1$.

Another property of variations, which distinguish tame sets among more complicated ones, is the following:

Theorem 5.15. *Let A be a tame set. Then for each $i \geq 0$ and $0 \leqslant j \leqslant i$:*

$$V_i(A) \leqslant C.V_j(A).V_{i-j}(A),$$

where C depends only on constants of the type $B_{0,k}$, and thus only on $D(A)$ in the semialgebraic case.

Proof. By the inductive formula for variations (property 8 of Chapter 3), we have:

$$V_i(A) = c(n,i,j) \int_{\bar{P} \in \bar{G}_n^{n-j}} V_{i-j}(A \cap \bar{P}) \, d\bar{P}.$$

Now by Theorem 5.1, $V_{i-j}(A \cap \bar{P}) \leqslant B_{0,i-j}(A \cap \bar{P}) V_{i-j}(A)$. Hence:

$$V_i(A) \leqslant c(n,i,j).B_{0,i-j}(A \cap \bar{P}).V_{i-j}(A) \int_{\bar{P} \in \bar{G}_n^{n-j}} \mathbf{1}_{\bar{P} \cap A} \, d\bar{P}$$

$$\leqslant c(n,i,j).B_{0,i-j}(A \cap \bar{P}).V_{i-j}(A) \int_{\bar{P} \in \bar{G}_n^{n-j}} V_0(\bar{P} \cap A) \, d\bar{P}$$

$$= \frac{c(n,i,j)}{c(n,j)}.B_{0,i-j}(A \cap \bar{P}).V_{i-j}(A).V_j(A) = C.V_{i-j}(A).V_j(A).$$

□

Remark. Theorem 7.6 below generalizes this property for mappings of semialgebraic sets.

Corollary 5.16. *Let $A \subset B_r^n$ be a tame set. Then for any $i = 0, 1, \dots, n$ and $0 \leqslant j \leqslant i$,*

$$V_j(A) \geq \frac{C}{r^{i-j}} \cdot V_i(A).$$

Proof. By Theorem 5.1, $V_{i-j}(A) \leqslant \widetilde{C}.r^{i-j}$. □

Remark. Theorem 5.15 and Corollary 5.16 show that there exist strong correlations among the variations of a tame set, or semialgebraic sets of fixed complexity. In general, variations are independent (property (**9**) of Chapter 3). Let us give here an example, contradicting Corollary 5.16, for sets of unbounded complexity. Let S be a curve in $B_1^2 \subset \mathbb{R}^2$ of infinite length. Let à be the set in $\mathbb{R}^3$ defined by $A = (S \times [0;1]) \cup (B_1^2 \times (\{0\} \cup \{1\}))$. Then one can easily check that any plane cuts A in a connected set. Hence $V_1(A) \leqslant V_1([0;1]^3) = 1$, while $V_2(A) \leqslant Vol_2([0;1]^3) = \infty$.

As a consequence we obtain the following important property of semialgebraic and tame sets:

Theorem 5.17. *Let $A \subset B_r^n$ be a tame set. There are constants $C_0, \ldots, C_n$ depending only on constants of the type $B_{0,j}$, and thus only on $D(A)$ in the semialgebraic case, such that if for some ϵ, $M(\epsilon, A)$ is strictly greater than $\sum_{i=0}^{q} C_i(\frac{r}{\epsilon})^i$, then:*

$$V_{q+1}(A) \geq C'\epsilon^{q+1}.(\sum_{i=0}^{n-q}(\frac{r}{\epsilon})^i)^{-1}.[M(\epsilon, A) - \sum_{i=0}^{q} C_i(\frac{r}{\epsilon})^i].$$

Proof. By Theorem 5.1:

$$M(\epsilon, A) \leqslant \sum_{i=0}^{n} V_i(A)(\frac{1}{\epsilon})^i \leqslant \sum_{i=0}^{q} C_i(\frac{r}{\epsilon})^i + \sum_{i=q+1}^{n} V_i(A)(\frac{1}{\epsilon})^i.$$

Hence

$$\sum_{i=q+1}^{n} V_i(A)(\frac{1}{\epsilon})^i \geq M(\epsilon, A) - \sum_{i=0}^{q} C_i(\frac{r}{\epsilon})^i.$$

By Corollary 5.16, $V_i(A) \leqslant C r^{i-q-1} V_{q+1}(A)$, $i \geq q+1$, i.e.:

$$\begin{aligned} \sum_{i=q+1}^{n} V_i(A)(\frac{1}{\epsilon})^i &\leqslant C V_{q+1}(A) \sum_{i=q+1}^{n} r^{i-q-1}.(\frac{1}{\epsilon})^i \\ &= C V_{q+1}(A).\frac{1}{r^{q+1}} \sum_{i=q+1}^{n} (\frac{r}{\epsilon})^i \\ &= C V_{q+1}(A).(\frac{r}{\epsilon})^{q+1}.\frac{1}{r^{q+1}} \sum_{i=0}^{n-q}(\frac{r}{\epsilon})^i. \end{aligned}$$

□

This theorem shows, in particular, that if A contains a “too dense” grid, then it must have components of high dimension.

Corollary 5.18. *If for some $\epsilon > 0$ and for some ϵ-separated set $Z \subset B_r^n$, $A \cap Z$ contains $p \geq \sum_{i=0}^{q} C_i(\frac{r}{\epsilon})^i$ points, then:*

$$\begin{aligned} V_{q+1}(A) &\geq C'\epsilon^{q+1}(\sum_{i=0}^{n-q}(\frac{r}{\epsilon})^i)^{-1}(p - \sum_{i=0}^{q} C_i(\frac{r}{\epsilon})^i) \\ &\geq C''\frac{\epsilon^n}{r^{n-q-1}}.(p - \sum_{i=0}^{q} C_i(\frac{r}{\epsilon})^i) > 0. \end{aligned}$$

In particular, $\dim(A) \geq q+1$. $\square$

The following example shows that in general the degree ϵ^n in the expression cannot be improved.

Let A be the ball B_r^n itself and $q < n$. Then for ϵ small, $M(\epsilon, A) \sim (\frac{r}{\epsilon})^n$, and the expression in Theorem 5.17 gives $V_{q+1}(B_r^n) \geq C\dfrac{\epsilon^n}{r^{n-q-1}}M(\epsilon, A) = Cr^{q+1}$, the sharp bound, up to coefficients.

However, if we know a priori that $\dim(A) \leqslant n$, we get the following:

Corollary 5.19. *If under the assumptions of Corollary* 5.17 *we have in addition* $\ell = \dim(A)$, $q+1 \leqslant \ell \leqslant n$, *then:*

$$V_{q+1}(A) \geq C'' \frac{\epsilon^\ell}{r^{\ell-q-1}}(p - \sum_{i=0}^{q} C_i(\frac{r}{\epsilon})^i) > 0.$$

In particular, for $\ell = q+1$,

$$V_{q+1}(A) \geq C''\epsilon^{q+1}(p - \sum_{i=0}^{q} C_i(\frac{r}{\epsilon})^i) > 0.$$

Proof. We replace, in the proof of Theorem 5.17 the inequality:

$$M(\epsilon, A) \leqslant \sum_{i=0}^{n} V_i(A)(\frac{1}{\epsilon})^i$$

by the inequality:

$$M(\epsilon, A) \leqslant \sum_{i=0}^{\ell} V_i(A)(\frac{1}{\epsilon})^i.$$

$\square$

Remark. Results of Theorem 5.17 and Corollaries 5.18 and 5.19 can be considered as the generalization of the following fact: if the polynomial vanishes at “too many” points, then it vanishes identically. Indeed, these results can be formulated as follows: if the semialgebraic set A contains more “separated” points than is prescribed for the sets of dimension $\leqslant q$ and of the same complexity, than its dimension is at least $q+1$ and the $(q+1)$-volume is not too small.

Theorem 5.20. *Let* $A \subset \mathbb{R}^n$ *be a tame set. Then for each* $i = 0, 1, \ldots, n$, *there exists a tame set* $C_i \subset A$, *with* $\dim(C_i) = i$ *and* $V_i(C_i) \geq \lambda V_i(A)$. *When* A *is semialgebraic,* C_i *is also semialgebraic and the diagram of* C_i *and* λ *depend only on the diagram of* A.

Proof. We have:

$$V_i(A) = c(n,i) \int_{P \in G_n^i} \int_{x \in P} V_0(A \cap P_x) \; dx \; dP$$

$$\leqslant c(n,i)B_{0,n-i}\int_{P\in G_n^i} Vol_i(\pi_P(A))\ dP$$

In particular, for some $P \in G_n^i$, $Vol_i(\pi_P(A)) \geq \lambda V_i(A)$. Now by Theorem 4.10 (this result being valid for tame sets, because one can stratify a tame map), there is an i-dimensional set $C \subset A$ of large enough i-dimensional measure, such that $\dim(C) = i$, and $\pi_P(C) = \pi_P(A)$. Thus we get:

$$V_i(C) = Vol_i(C) \geq Vol_i(\pi_P(C)) = Vol_i(\pi_P(A)) \geq \lambda V_i(A).$$

□

Another simple but important property of semialgebraic sets of fixed algebraic complexity or of bounded tame sets is the following:

Theorem 5.21. *Let A be a set in an o-minimal structure (or a bounded tame set),* $\dim(A) \leqslant n-1$. *Then for each $\eta > 0$, the maximal radius for a ball contained in the η-neighborhood A_η of A, is at most $C\eta$, where C depends only on constants of the type $B_{0,j}$, and consequently, only on the diagram of A in the semialgebraic case.*

Proof. Assume that the ball $B_r^n \subset A_\eta$. Let B_{2r}^n be the ball of radius $2r$, centred at the same point. Then for $\eta \leqslant r$, $A_\eta \cap B_r^n \subset (A \cap B_{2r}^n)_\eta$. By theorem 5.8,

$$Vol_n\big((A\cap B_{2r}^n)_\eta\big) \leqslant C'[r^\ell \eta^{n-\ell} + \eta^n] \leqslant C'' r^\ell \eta^{n-\ell},$$

where $\ell = \dim(A) < n$. Thus $Vol_n(A_\eta \cap B_r^n) \leqslant C'' r^\ell \eta^{n-\ell}$. But by the assumption, $A_\eta \cap B_r^n = B_r^n$, i.e. $Vol_n(B_1^n) r^n \leqslant C'' r^\ell \eta^{n-\ell}$. Hence:

$$Vol_n(B_1^n).r^{n-\ell} \leqslant C''\eta^{n-\ell}, \quad \text{or} \quad r \leqslant C\eta, \quad \text{where} \quad C = \Big(\frac{C''}{Vol_n(B_1^n)}\Big)^{\frac{1}{n-\ell}}.$$

□

Of course, nothing of this sort is true for non-tame sets: take A to be the set of rational points in $[0;1]$.

Remark. The property given by Theorem 5.21 is much more precise than the similar conclusion one can obtain by comparison of "global volumes". Indeed, the volume of an η-neighborhood of $A \subset B_1^n$, $\dim(A) = \ell$, is of order $\eta^{n-\ell}$. Comparing this with the volume of the ball of radius r, $Vol_n(B_1^n) r^n$, we obtain only $Vol_n(B_1^n) r^n \leqslant \eta^{n-\ell}$, or $r \leqslant C\eta^{1-\frac{1}{n}}$, instead of $r \leqslant C\eta$

Theorem 5.20 may be generalized in the following way: if a tame set B lies in the η-neighborhood of an ℓ-dimensional tame set, then the variations $V_j(B)$, $j > \ell$, are bounded in terms of lower variations of B and η.

Theorem 5.22. *Let $A \subset B_r^n$ be a tame set,* $\dim(A) = \ell < n$. *Let $\eta > 0$ be fixed. Then if another tame set $B \subset B_r^n$ is contained in the η-neighborhood A_η of A, its variations satisfy the following inequality:*

$$\textit{for any} \quad j > \ell \qquad V_j(B) \leqslant \bar{C}\cdot V_0(B)\cdot \eta^j + \bar{C}\cdot \sum_{i=1}^{\ell} [V_i(B) + \eta r^{i-1} + \eta^i]\eta^{j-i},$$

where $\bar{C}$ depends only on constants of the type $B_{0,j}$, and consequently, only on the diagram of A and B, in the semialgebraic case.

Proof. If $B \subset A_\eta$, then also $B \subset (A \cap B_\eta)_\eta$. Now by Theorem 5.11,

$$V_i(B_\eta) \leqslant C \cdot V_i(B) + \widehat{C}(\eta r^{i-1} + \eta^i).$$

Hence by Theorem 5.1:

$$V_i(A \cap B_\eta) \leqslant C \cdot V_i(B_\eta) \leqslant C[V_i(B) + \widehat{C}(\eta r^{i-1} + \eta^i)],$$

for $i = 1, \ldots, \ell$, $V_0(A \cap B_\eta) \leqslant C \cdot V_0(B)$ and $V_i(A \cap B_\eta) = 0$, for $i > \ell$. Let us prove the following proposition, in order to finish the proof of Theorem 5.22.

Proposition 5.23. *Let $A \subset \mathbb{R}^n$ be a bounded tame set of dimension $\ell < n$. Then for $\eta \geq 0$ and for $j > \ell$,*

$$V_j(A_\eta) \leqslant C_1 \sum_{i=0}^{\ell} V_i(A)\eta^{j-i}.$$

Proof. As usual, it is enough to bound the j-volume of the projections $\pi_P(A_\eta)$ on j-dimensional subspaces P of $\mathbb{R}^n$. But $\pi_P(A_\eta) \subset \big(\pi_P(A)\big)_\eta$, and $V_i(\pi_P(A)) \leqslant V_i(A)$. Hence

$$Vol_j\big(\pi_P(A)\big)_\eta \leqslant c\eta^j M(\eta, \pi_P(A)) \leqslant C_1\eta^j \sum_{i=0}^{\ell} V_i(A)(\frac{1}{\eta})^i = C_1 \sum_{i=0}^{\ell} V_i(A)\eta^{j-i}.$$

□

Applying Proposition 5.21 in our situation, it follows that for $j > \ell$:

$$\begin{aligned} V_j(B) &\leqslant C'V_j\big((A \cap B_\eta)_\eta\big) \leqslant C'C_1 \sum_{i=0}^{\ell} V_i(A \cap B_\eta)\eta^{j-i} \\ &\leqslant C'C_1\widehat{C} \sum_{i=1}^{\ell} [V_i(B) + \eta r^{i-1} + \eta^i]\eta^{j-i} + C'C_1\widehat{C}V_0(B)\eta^j \\ &= C_2 \sum_{i=0}^{\ell} V_i(B)\eta^{j-i} + C_3 \sum_{i=0}^{\ell} r^{i-1}\eta^{j-i+1} + C_4\eta^j. \end{aligned}$$

□

6 Some Exterior Algebra

Abstract. We give in this chapter some basic definitions and well-known results in exterior algebra, in order to get a convenient definition of a size for differentials of mappings. The behaviour of this size under projections and restrictions to subspaces (as required by the variations approach) is studied.

In the next chapter we describe variations of the images and the preimages of polynomial mappings (generally speaking, on semialgebraic sets). This description is given in terms of (as usual) "algebraic complexity" of the sets and mapping involved, and it requires some metric information on the mapping. Usually this information concerns upper or lower bounds for the first differential (but our approach allows one to encounter also more delicate properties: e.g., the presence of higher Thom-Boardman singularities, as in Theorem 8.10 below).

The bounds on variations, obtained in this chapter, being restated in term of ϵ-entropy, form a part of our main results in the algebraic category: the quantitative Morse-Sard theorem and the quantitative transversality theorem.

First of all, to describe quantitatively the behavior of the differential of a considered mapping, we need some results from exterior algebra (for complements see, for instance, [Bou], Algèbre, Chap. 3), and for the sake of non-familiar readers we develop below some very basic results in this area.

Of course the well-informed reader can immediately go to Lemma 6.2.

Let $\bigotimes_i \mathbb{R}^n = \mathbb{R}^n \bigotimes \ldots \bigotimes \mathbb{R}^n$ denote, as usual, the tensor product of $\mathbb{R}^n, \ldots, \mathbb{R}^n$ (i times). We recall that this vector space is formally defined as the quotient E/F, where E is the vector space of real valued functions on $(\mathbb{R}^n)^i$ that vanish outside finite sets ($\mathbf{1}_{[\{(v_1,\ldots,v_i)\}]}$ being denoted by $v_1 \otimes \ldots \otimes v_i$) and F is the subspace of E generated by elements of the type:

$$\begin{aligned}&(v_1 \otimes \ldots \otimes v_{j-1} \otimes x \otimes v_{j+1} \otimes \ldots \otimes v_i)\\ &\quad +(v_1 \otimes \ldots \otimes v_{j-1} \otimes y \otimes v_{j+1} \otimes \ldots \otimes v_i)\\ &\quad -(v_1 \otimes \ldots \otimes v_{j-1} \otimes (x+y) \otimes v_{j+1} \otimes \ldots \otimes v_i)\end{aligned}$$

and

Y. Yomdin and G. Comte: LNM 1834, pp. 75–82, 2004.

$$v_1 \otimes \ldots \otimes v_{j-1} \otimes \lambda x \otimes v_{j+1} \otimes \ldots \otimes v_i$$
$$-\lambda(v_1 \otimes \ldots \otimes v_{j-1} \otimes x \otimes v_{j+1} \otimes \ldots \otimes v_i),$$

for all $x, y \in \mathbb{R}^n$, all $\lambda \in \mathbb{R}$ and all $j \in \{1, \ldots, i\}$.

We also denote by $v_1 \otimes \ldots \otimes v_i$ the class of $v_1 \otimes \ldots \otimes v_i$ in $\bigotimes_i \mathbb{R}^n$.

Now let us consider the subspace G of $\bigotimes_i \mathbb{R}^n$ generated by all elements of the type:

$$v_1 \otimes \ldots \otimes v_{j-1} \otimes x \otimes x \otimes v_{j+2} \otimes \ldots \otimes v_i.$$

The vector space $\bigotimes_i \mathbb{R}^n / G$, denoted by $\bigwedge_i \mathbb{R}^n$, is by definition the i-th exterior product of $\mathbb{R}^n$. The class of $v_1 \otimes \ldots \otimes v_i$ in $\bigwedge_i \mathbb{R}^n$ is denoted by $v_1 \wedge \ldots \wedge v_i$.

Of course we have:

$$\begin{aligned}
&v_1 \otimes \ldots \otimes v_{j-1} \otimes x \otimes y \otimes v_{j+2} \otimes \ldots \otimes v_i \\
&\quad + v_1 \otimes \ldots \otimes v_{j-1} \otimes y \otimes x \otimes v_{j+2} \otimes \ldots \otimes v_i \\
&= v_1 \otimes \ldots \otimes v_{j-1} \otimes (x+y) \otimes (y+x) \otimes v_{j+2} \otimes \ldots \otimes v_i \\
&\quad -(v_1 \otimes \ldots \otimes v_{j-1} \otimes x \otimes x \otimes v_{j+2} \otimes \ldots \otimes v_i) \\
&\quad -(v_1 \otimes \ldots \otimes v_{j-1} \otimes y \otimes y \otimes v_{j+2} \otimes \ldots \otimes v_i) \in G,
\end{aligned}$$

and in particular:

$$v_1 \wedge \ldots \wedge v_j \wedge v_{j+1} \ldots \wedge v_i = -(v_1 \wedge \ldots \wedge v_{j+1} \wedge v_j \ldots \wedge v_i).$$

If $(e_1, \ldots, e_n)$ denotes a basis of $\mathbb{R}^n$, elements of the type $e_J = e_{j_1} \wedge \ldots \wedge e_{j_i}$ give a basis of $\bigwedge_i \mathbb{R}^n$, where $J = \{j_1 < \ldots < j_i\} \subset \{1, \ldots, n\}$, and consequently the dimension of $\bigwedge_i \mathbb{R}^n$ is C_n^i.

We define on $\bigwedge_i \mathbb{R}^n$ the following scalar product, by defining it for homogeneous elements of the type $w = v_1 \wedge \ldots \wedge v_i$:

$$(w, w') = \det\Big((v_k, v'_j)\Big), \quad k, j = 1, \ldots, i.$$

If $(e_1, \ldots, e_n)$ is an orthonormal basis of $\mathbb{R}^n$, then $(e_I)_{I \in \Lambda(n,i)}$, where $\Lambda(n,i)$ is the set of all subsets $I = \{j_1 < \ldots < j_i\} \subset \{1, \ldots, n\}$, is an orthonormal basis of $\bigwedge_i \mathbb{R}^n$, and in particular if we write $w = \sum_{I \in \Lambda(n,i)} \xi_I e_I$, $w' = \sum_{I \in \Lambda(n,i)} \xi'_I e_I$, with $\xi_I, \xi'_I \in \mathbb{R}$, we obtain:

$$(w, w') = \sum_{I \in \Lambda(n,i)} \xi_I . \xi'_I.$$

We can naturally define a product $\wedge : \bigwedge_i \mathbb{R}^n \times \bigwedge_j \mathbb{R}^n \to \bigwedge_{i+j} \mathbb{R}^n$, by the equality: $(v_1 \wedge \ldots \wedge v_i) \wedge (v'_1 \wedge \ldots \wedge v'_j) = v_1 \wedge \ldots \wedge v_i \wedge v'_1 \wedge \ldots \wedge v'_j$.

It is then immediate to check that

$$||v_1 \wedge \ldots \wedge v_i \wedge v_1' \wedge \ldots \wedge v_j'||_{i+j} \leqslant ||v_1 \wedge \ldots \wedge v_i||_i . ||v_1' \wedge \ldots \wedge v_j'||_j .$$

It is an easy exercice to check that the volume of a parallelepiped $\Pi_k = \Pi_k(v_1, \ldots, v_k)$ spanned by k vectors $v_1, \ldots, v_k$ in $\mathbb{R}^n$ is given by the following induction on k:

- $Vol_k(\Pi_k) = 0$ if the v_i's are dependant ,
- $Vol_1(\Pi_1) = ||v_1||$, and
  $Vol_{j+1}(\Pi_{j+1}(v_1, \ldots, v_{j+1})) = |(v_{j+1}, \nu_{j+1})|.Vol_j(\Pi_j(v_1, \ldots, v_j))$,

for all $j \in \{1, \ldots, k-1\}$, where ν_j is a unit normal vector, in the $(j+1)$-dimensional vector space spanned by $v_1, \ldots, v_{j+1}$, to the j-dimensional vector space spanned by $v_1, \ldots, v_j$.

From these formulas, one can prove easily by induction on k that, geometrically, the norm of a homogeneous element $w = v_1 \wedge \ldots \wedge v_n$, which is $||w|| = \sqrt{\det\Big((v_k, v_j)\Big)}$, is nothing else than the volume of the parallelepiped $\Pi_k = \Pi_k(v_1, \ldots, v_k)$ spanned by the k vectors $v_1, \ldots, v_k$ in $\mathbb{R}^n$.

Let now $\mathrm{L} : \mathbb{R}^n \to \mathbb{R}^m$ be a linear mapping. This mapping induces, for each i, the linear mapping $\mathrm{L}_i : \bigwedge_i \mathbb{R}^n \to \bigwedge_i \mathbb{R}^m$, defined on homogeneous elements by:

$$\mathrm{L}_i(v_1 \wedge \ldots \wedge v_i) = \mathrm{L}(v_1) \wedge \ldots \wedge \mathrm{L}_i(v_i).$$

This linear mapping has a norm, induced by the scalar products on $\bigwedge_i \mathbb{R}^n$ and $\bigwedge_i \mathbb{R}^m$ respectively.

Notations. For $i = 1, \ldots, \min(n, m)$, let us denote

- $||\mathrm{L}_i|| = \max\limits_{||w||=1, w \in \wedge_i \mathbb{R}^n} ||\mathrm{L}_i(w)||$ by $w_i(\mathrm{L})$, and $w_0(\mathrm{L}) = 1$.
- $\lambda_i(\mathrm{L})$, $i = 1, \ldots, \min(n, m)$, the semiaxes of the ellipsoid $\mathrm{L}(B_1^n) \subset \mathbb{R}^m$, in a decreasing order: $\lambda_1(\mathrm{L}) \geq \lambda_2(\mathrm{L}) \geq \ldots \geq \lambda_q(\mathrm{L})$, with $q = \min(n, m)$. $\lambda_0(\mathrm{L}) = 1$.

The proposition says that $w_i(\mathrm{L})$ is the maximun i-dimensional volume of the image under L of a unit i-dimensional cube of $\mathbb{R}^n$, when $i = \min(n, m)$, $w_i(\mathrm{L})$ equals the square root of the determinant of ${}^t\mathrm{LL}$, and thus $w_i(\mathrm{L})$ equals $\det(\mathrm{L})$, for $i = m = n$.

Proposition 6.1. *Let* $\mathrm{L} : \mathbb{R}^n \to \mathbb{R}^m$ *be a linear map, and* $i \in \{1, \ldots, q = min(n, m)\}$. *then:*

$$w_i(\mathrm{L}) = \lambda_0(\mathrm{L})\lambda_1(\mathrm{L}) \ldots \lambda_i(\mathrm{L}),$$

and this number equals the maximun i-dimensional volume of the image under L *of a unit i-dimensional cube of* $\mathbb{R}^n$.

More precisely, there exist orthonormal bases $e_1, \ldots, e_n$ *and* $e_1', \ldots, e_m'$ *of* $\mathbb{R}^n$ *and* $\mathbb{R}^m$, *respectively, such that:*

- $\mathrm{L}(e_j) = \lambda_j(\mathrm{L})e'_j$, *for* $j \in \{1, \dots, q\}$, *and* $\mathrm{L}(e_j) = 0$, *for* $j > q$ *(if* $n > m$*).*
- *The norm* $w_i(\mathrm{L}) = ||\mathrm{L}_i|| = \max\limits_{||w||=1, w \in \wedge_i \mathbb{R}^n} ||\mathrm{L}_i(w)||$ *is attained on the homogeneous* $w = e_1 \wedge \dots \wedge e_i$, *and hence is equal to* $\lambda_0(\mathrm{L}) \dots \lambda_i(\mathrm{L})$. *In particular, this norm is equal to the maximal expansion of the* i*-dimensional volume by* L.

Proof. The statements of this proposition are well known and their proof is an exercise in linear algebra. For sake of completeness we give it here.

Consider the symetric bilinear nonnegative form ω on $\mathbb{R}^n$ given by

$$\omega(u, v) = (\mathrm{L}u, \mathrm{L}v) = {}^t v({}^t\mathrm{L}\mathrm{L})u.$$

Let $e_1, \dots, e_n$ be an orthonormal basis of $\mathbb{R}^n$ in which ω has a diagonal form: if $u = \sum\limits_{j=1}^{n} \alpha_j e_j$, $v = \sum\limits_{j=1}^{n} \beta_j e_j$, we have $\omega(u, v) = \sum\limits_{j=1}^{n} \lambda_j^2 \beta_j \alpha_j$, with $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k > 0$, $\lambda_{k+1} = \dots = \lambda_n = 0$ and $k =$rank(L) $=$rank(ω).

Let $e'_j = \dfrac{1}{\lambda_j}\mathrm{L}(e_j)$, for $j = 1, \dots, k$. Of course we have $(e'_j, e'_l) = \delta_{j,l}$, and hence e'_j form a part of an orthonormal basis e'_j, $j = 1, \dots, m$ in $\mathbb{R}^m$ (we can suppose that $\mathrm{L}(e_j) = 0$, for $i \geq k$). Now if $u = \sum\limits_{j=1}^{n} \alpha_j e_j$, is such that $\sum\limits_{j=1}^{n} \alpha_j^2 \leqslant 1$, we have $\mathrm{L}(u) = \sum\limits_{j=1}^{k} \alpha_j \lambda_j \mathrm{L}(e_j) = \sum\limits_{j=1}^{k} x_j e'_j$, and

$$\frac{x_1^2}{\lambda_1^2} + \dots + \frac{x_k^2}{\lambda_k^2} \leqslant 1.$$

Hence, the numbers $\lambda_1, \dots, \lambda_k$ are the semiaxes of the ellipsoid $\mathrm{L}(B_1^n)$, which is given by

$$\frac{x_1^2}{\lambda_1^2} + \dots + \frac{x_k^2}{\lambda_k^2} \leqslant 1, \quad x_{k+1} = \dots = x_m = 0.$$

This proves the first part of the proposition.

Now let us define the following symmetric bilinear form ω_i on $\bigwedge_i \mathbb{R}^n$:

$$\omega_i(w_1, w_2) = \big(\mathrm{L}_i(w_1), \mathrm{L}_i(w_2)\big).$$

This form is diagonal on the basis $(e_I)_{I \in \Lambda(n,i)}$. Indeed:

$$\mathrm{L}_i(e_{j_1} \wedge \dots \wedge e_{j_i}) = \lambda_{j_1} \dots \lambda_{j_i}(e'_{j_1} \wedge \dots \wedge e'_{j_i}), \text{ and } (e'_I, e'_J) = \delta_{I,J}.$$

Hence $w_i(\mathrm{L}) = ||\mathrm{L}_i||$ is the maximal eigenvalue of ω_i, which is $\lambda_1 \dots \lambda_i$, since $\lambda_1 \geq \dots \geq \lambda_q > 0$. Thus this norm is attained on the homogeneous elements $e_1 \wedge \dots \wedge e_i$ and is consequently the i-volume of the parallelepiped (in fact the cube) spanned by $\mathrm{L}(e_1), \dots, \mathrm{L}(e_i)$ in $\mathbb{R}^m$. But any unit i-cube in $\mathbb{R}^n$ is spanned by some $e_{j_1}, \dots, e_{j_i}$ (up to an isometry) and the i-volume of its image under L is less than $Vol_i\big(\Pi_i(\mathrm{L}(e_1), \dots \mathrm{L}(e_i))\big) = \lambda_1 \dots \lambda_i$. $\square$

Remark. The statement of Proposition 6.1 is in fact equivalent to the representation of L as the composition of an orthogonal and a symmetric mapping.

Below we express our assumptions on the differentials of the mappings considered in terms of $\lambda_i(df)$ and $w_i(df)$.

Lemma 6.2. *For* $\mathrm{L}_1, \mathrm{L}_2 : \mathbb{R}^n \to \mathbb{R}^m$, *two linear mappings, we have:*

$$w_i(\mathrm{L}_1 + \mathrm{L}_2) \leqslant k(n, m, i) \sum_{j=0}^{i} w_j(\mathrm{L}_1) w_{i-j}(\mathrm{L}_2), \quad i = 0, \dots, \min(n, m).$$

Proof. Let $w = v_1 \wedge \dots \wedge v_i \in \bigwedge_i \mathbb{R}^n$ be the element on which the norm $w_i(\mathrm{L}_1 + \mathrm{L}_2) = ||\mathrm{L}_1 + \mathrm{L}_2||$ is attained. Then:

$$(\mathrm{L}_1 + \mathrm{L}_2)_i(w) = (\mathrm{L}_1(v_1) + \mathrm{L}_2(v_2)) \wedge \dots \wedge (\mathrm{L}_1(v_i) + \mathrm{L}_2(v_i))$$
$$= \sum_{I \subset \{1, \dots, i\}} \pm \Big(\bigwedge_{j \in I} \mathrm{L}_1(v_j) \Big) \wedge \Big(\bigwedge_{k \in \bar{I}} \mathrm{L}_2(v_k) \Big).$$

Hence:

$$w_i(\mathrm{L}_1 + \mathrm{L}_2) = ||(\mathrm{L}_1 + \mathrm{L}_2)_i(w)|| \leqslant \sum_{I \subset \{1, \dots, i\}} || \bigwedge_{j \in I} \mathrm{L}_1(v_j) ||.|| \bigwedge_{k \in \bar{I}} \mathrm{L}_2(v_k) ||$$

$$\leqslant \sum_{I \subset \{1, \dots, i\}} ||(\mathrm{L}_1)_{|I|}||.||(\mathrm{L}_2)_{i-|I|}|| \leqslant k(n, m, i) \sum_{j=0}^{i} w_j(\mathrm{L}_1) w_{i-j}(\mathrm{L}_2).$$

$\square$

The next lemma will be used in the proof of the "quantitative transversality" theorem below. A general structure of transversality results is the following: we have a mapping $F : N \times T \to M$, where T is the space of parameters. We assume that the parameters act non-degenerately, i.e. that $F_{|\{*\} \times T} : \{*\} \times T \to M$ is non degenerate. The desirable conclusion is that for a generic value of parameters $t_0 \in T$, $f_t = F_{N \times \{t_0\}} : N \times \{t_0\} \to M$ is in some sense nondegenerate. The following lemma gives a linear version of this statement; $\mathbb{R}^q$ in this lemma serving both as T and M.

Consider the product $\mathbb{R}^p \times \mathbb{R}^q$, and let

$$\pi_1 : \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}^p, \qquad \pi_2 : \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}^q$$

be the natural projections.

For a linear mapping $\mathrm{L} : \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}^q$ denote by L' and L'' the restrictions

$$\mathrm{L}' = \mathrm{L}_{|\mathbb{R}^p \times \{0\}} : \mathbb{R}^p \times \{0\} \to \mathbb{R}^q$$

$$\mathrm{L}'' = \mathrm{L}_{|\{0\}\times\mathbb{R}^q} : \{0\} \times \mathbb{R}^q \to \mathbb{R}^q.$$

Then $\mathrm{L} = \mathrm{L}' \circ \pi_1 + \mathrm{L}'' \circ \pi_2$. Let $\mathrm{T} \subset \mathbb{R}^p \times \mathbb{R}^q$ be a linear subspace, $\widetilde{\pi}_i = \pi_{i|\mathrm{T}}$ $(i = 1, 2)$, $\widetilde{\mathrm{L}} = \mathrm{L}_{|\mathrm{T}}$. Then $\widetilde{\mathrm{L}} = \mathrm{L}' \circ \widetilde{\pi}_1 + \mathrm{L}'' \circ \widetilde{\pi}_2$.

We assume also that L'' is regular, i.e. that $(\mathrm{L}'')^{-1} : \mathbb{R}^q \to \mathbb{R}^q$ exists.

Lemma 6.3. *Under the above assumptions,*

$$w_i(\widetilde{\pi}_2) \leqslant k.w_i(\mathrm{L}''^{-1}) \sum_{j=0}^{i} w_j(\widetilde{\mathrm{L}}).w_{i-j}(\mathrm{L}'),$$

$i = 1, \ldots, D = min\big(q, \dim(P)\big)$. *Here the constant k depends only on p, q,* $\dim(T)$.

Remark. In the special case, where $\mathrm{T} = \ker(\mathrm{L})$, we obtain $\dim(\mathrm{T}) = p$, $w_D(\widetilde{\pi}_2) \leqslant k.w_D(\mathrm{L}''^{-1})w_D(\mathrm{L}')$ (since only $w_0(\widetilde{\mathrm{L}})$ is not zero). Thus if L' is degenerate, also $\pi_2 : \mathrm{T} \to \mathbb{R}^q$ is degenerate. This last statement is used in standard reduction of the transversality theorem to the Sard theorem. So lemma 6.3 can be considered as a quantitative version of this simple statement.

Proof. We have:

$$\widetilde{\pi}_2 = \mathrm{L}''^{-1} \circ (\widetilde{\mathrm{L}} - \mathrm{L}' \circ \widetilde{\pi}_1).$$

Hence for all $i \in \{1, \ldots, D\}$:

$$w_i(\widetilde{\pi}_2) \leqslant w_i(\mathrm{L}''^{-1}) w_i(\widetilde{\mathrm{L}} - \mathrm{L}' \circ \widetilde{\pi}_1) \leqslant k.w_i(\mathrm{L}''^{-1}) \sum_{j=0}^{i} w_j(\widetilde{\mathrm{L}}) w_{i-j}(\mathrm{L}'),$$

by lemma 6.2, since of course, $w_{i-j}(\mathrm{L}' \circ \widetilde{\pi}_1) \leqslant w_{i-j}(\mathrm{L}')$. □

The last algebraic lemma we need has a technical character.

Let $\mathrm{L} : \mathbb{R}^n \to \mathbb{R}^m$ be a linear mapping. Let $i \leqslant m$ and let G_m^i denote as usual the Grassmann manifold of all the i-dimensional linear subspaces P of $\mathbb{R}^m$. For $P \in G_m^i$, let $\pi_P : \mathbb{R}^m \to P$ denote the orthogonal projection onto P and $\pi_{\ker\,(\pi_P \circ \mathrm{L}\,)} : \mathbb{R}^n \to \ker(\pi_P \circ \mathrm{L})$ the orthogonal projection onto $\ker(\pi_P \circ \mathrm{L})$.

Let, for a Euclidean space V, ω_V denote the volume form on V. For any $P \in G_m^i$ define an n-form $\omega(P)$ on $\mathbb{R}^n$ as follows:

$$\omega(P) = [(\pi_{\ker(\pi_P \circ \mathrm{L})})^* \omega_{\ker(\pi_P \circ \mathrm{L})}] \wedge [(\pi_P \circ \mathrm{L})^* \omega_P],$$

and the n-form $\displaystyle\int_{P \in G_m^i} \omega(P)\, dP$ by the following formula:

$$[\int_{P \in G_m^i} \omega(P)\, dP](u_1, \ldots, u_n) = \int_{P \in G_m^i} \omega(P)(u_1, \ldots, u_n)\, dP,$$

for all $(u_1, \dots, u_n) \in (\mathbb{R}^n)^n$.

Of course, if $\mathrm{rank(L)} < i$, we have $\mathrm{rank}(\pi_P \circ \mathrm{L}) < i$ and the i-form $[(\pi_P \circ \mathrm{L})^*(\omega_P)]$ is identically equal to 0. Thus we suppose that $k = \mathrm{rank(L)} \geq i$ and $\mathrm{rank}(\pi_P \circ \mathrm{L}) = i$.

Lemma 6.4. *With the notations above, we have:*

$$\int_{P \in G_m^i} \omega(P)\ dP \leqslant w_i(\mathrm{L}) \cdot \omega_{\mathbb{R}^n} \leqslant C \cdot \int_{P \in G_m^i} \omega(P)\ dP,$$

where C_2 depends only on n, m and i.

Proof. Clearly:

$$(\pi_P \circ \mathrm{L})^*(\omega_P) = (\pi_P \circ \mathrm{L} \circ \pi_{\ker^\perp(\pi_P \circ \mathrm{L})})^*(\omega_P)$$

$$\leqslant w_i(\pi_P \circ \mathrm{L}).(\pi_{\ker^\perp(\pi_P \circ \mathrm{L})})^* \omega_{\ker^\perp(\pi_P \circ \mathrm{L})} \leqslant w_i(\mathrm{L}).(\pi_{\ker^\perp(\pi_P \circ \mathrm{L})})^* \omega_{\ker^\perp(\pi_P \circ \mathrm{L})}.$$

Hence:

$$\omega(P) \leqslant w_i(\mathrm{L}).[(\pi_{\ker(\pi_P \circ \mathrm{L})})^* \omega_{\ker(\pi_P \circ \mathrm{L})}] \wedge [(\pi_{\ker^\perp(\pi_P \circ \mathrm{L})})^* \omega_{\ker^\perp(\pi_P \circ \mathrm{L})}]$$

$$= w_i(\mathrm{L}).\omega_{\mathbb{R}^n}.$$

Integrating over G_m^i, we obtain the left-hand side inequality.

We recall (see Proposition 6.1) that we denote by $e_1, \dots, e_k, e_{k+1}, \dots, e_n$ an orthonormal basis of $\mathbb{R}^n$ such that $\ker(\mathrm{L}) = < e_{k+1}, \dots, e_n >$, $e'_1 = \dfrac{\mathrm{L}(e_1)}{\lambda_1(\mathrm{L})}, \dots, e'_k = \dfrac{\mathrm{L}(e_k)}{\lambda_k(\mathrm{L})}$ is a part of an orthonormal basis of $\mathbb{R}^m$, and $\lambda_1(\mathrm{L}) \geq \dots \geq \lambda_k(\mathrm{L}) > 0$ are the positive eigenvalues of the nonnegative symmetric form ${}^t\mathrm{LL}$.

Now on the other hand, we have:

$$(\pi_P \circ \mathrm{L})^* \omega_P \geq w_i(\mathrm{L}).\mathrm{J_L}(P).(\pi_{\ker^\perp(\pi_P \circ \mathrm{L})})^* \omega_{\ker^\perp(\pi_P \circ \mathrm{L})}, \qquad (*)$$

where $\mathrm{J_L}(P) = \mathrm{J}_{(\pi: < e'_1, \dots, e'_i > \to P)}$ and where $\mathrm{J}_{(\pi: < e'_{\ell_1}, \dots, e'_{\ell_i} > \to P)}$ is the Jacobian of the projection onto P of the i-dimensional subspace of $\mathbb{R}^n$ spanned by $e'_{\ell_1}, \dots, e'_{\ell_i}$.

For linearity reasons, to prove $(*)$ it suffices to prove that the inequality $(*)$ holds for every $(e_{\ell_1}, \dots, e_{\ell_i})$, $\ell_1, \dots, \ell_i \in \{1, \dots, k\}$, because of course we have $\ker^\perp(\pi_P \circ \mathrm{L}) \subset < e_1, \dots, e_k >$.

The following equality holds:

$$[(\pi_P \circ \mathrm{L})^* \omega_P](e_{\ell_1}, \dots, e_{\ell_i}) = \omega_P(\lambda_{\ell_1}(\mathrm{L}).\pi_P(e'_{\ell_1}), \dots, \lambda_{\ell_i}(\mathrm{L}).\pi_P(e'_{\ell_i}))$$

$$= \epsilon.\lambda_{\ell_1}(\mathrm{L}) \dots \lambda_{\ell_i}(\mathrm{L}).\mathrm{J}_{(\pi: < e'_{\ell_1}, \dots, e'_{\ell_i} > \to P)}, \qquad (1)$$

where $\epsilon = +1$ or -1.

But we have also:

$$[(\pi_{\ker^\perp(\pi_P\circ \mathrm{L})})^* \omega_{\ker^\perp(\pi_P\circ \mathrm{L})}](e_{\ell_1},\dots,e_{\ell_i}) = \epsilon.|\det(\pi_{\ker^\perp(\pi_P\circ \mathrm{L})})|. \tag{2}$$

It follows from (1) and (2) that:

$$[(\pi_P\circ \mathrm{L})^*\omega_P](e_{\ell_1},\dots,e_{\ell_i}) \geq$$

$$\lambda_{\ell_1}(\mathrm{L})\dots\lambda_{\ell_i}(\mathrm{L})\cdot \mathrm{J}_{(\pi:<e'_{\ell_1},\dots,e'_{\ell_i}>\to P)}\cdot[(\pi_{\ker^\perp(\pi_P\circ \mathrm{L})})^* \omega_{\ker^\perp(\pi_P\circ \mathrm{L})}](e_{\ell_1},\dots,e_{\ell_i}),$$

for all $\ell_1,\dots,\ell_i\in\{1,\dots,k\}$. Now considering that there exists a constant k such that:

$$[(\pi_P\circ \mathrm{L})^*\omega_P] = k\cdot[(\pi_{\ker^\perp(\pi_P\circ \mathrm{L})})^* \omega_{\ker^\perp(\pi_P\circ \mathrm{L})}],$$

we obtain that:

$$k\cdot[(\pi_{\ker^\perp(\pi_P\circ \mathrm{L})})^* \omega_{\ker^\perp(\pi_P\circ \mathrm{L})}] \geq$$

$$\lambda_{\ell_1}(\mathrm{L})\dots\lambda_{\ell_i}(\mathrm{L})\cdot \mathrm{J}_{(\pi:<e'_{\ell_1},\dots,e'_{\ell_i}>\to P)}.[(\pi_{\ker^\perp(\pi_P\circ \mathrm{L})})^* \omega_{\ker^\perp(\pi_P\circ \mathrm{L})}],$$

for all $\ell_1,\dots,\ell_i\in\{1,\dots,k\}$. Taking $\ell_1=1,\dots,\ell_i=i$, it follows in particular that:

$$k\geq w_i(\mathrm{L})\cdot \mathrm{J}_{(\pi:<e'_1,\dots,e'_1>\to P)},$$

and this inequality proves $(*)$.

Now from $(*)$ we obtain:

$$\omega(P)\geq w_i(\mathrm{L})\cdot \mathrm{J_L}(P)\cdot[(\pi_{\ker(\pi_P\circ \mathrm{L})})^*\omega_{\ker(\pi_P\circ \mathrm{L})}]\wedge[(\pi_{\ker^\perp(\pi_P\circ \mathrm{L})})^* \omega_{\ker^\perp(\pi_P\circ \mathrm{L})}]$$

$$= w_i(\mathrm{L})\cdot \mathrm{J_L}(P)\cdot\omega_{\mathbb{R}^n},$$

and integrating over G^i_m, we deduce that:

$$\int_{P\in G^i_m}\omega(P)\ dP\geq w_i(\mathrm{L})\cdot\int_{P\in G^i_m}\mathrm{J_L}(P)\ dP\ \cdot\omega_{\mathbb{R}^n},$$

which proves the right-hand side inequality of lemma 6.4, with the constant C equal to $1/\int_{P\in G^i_m}\mathrm{J_L}(P)\ dP$, and of course by the homogeneity of G^i_m, this constant does not depend on L, but only on m and i. □

7 Behaviour of Variations under Polynomial Mappings

Abstract. We study here the multidimensional variations of the image under a polynomial mapping of a semialgebraic set. We bound from above the i-th variation of the image by the i-th variation of the set and by the i-th Jacobian. This allows us to prove the quantitative Sard theorem for polynomial functions. We also define and study the "variations" of a polynomial mapping, and we finally bound from below the variation of the image.

Let $f : N \to M$ be a differentiable mapping of the manifolds N and M of dimensions respectively n and m, with $n \leqslant m$. Then the behavior of the n-volume under f is quite regular. It is described by the so-called *area formula* in geometric measure theory (see [Fed 2]). For any $\mathcal{H}^n$-measurable set $A \in N$, we have:

$$\int_{y \in M} N(f_{|A}, y)\ d\mathcal{H}^n(y) = \int_{x \in A} w_n(Df_{(x)})\ d\mathcal{H}^n(x),$$

where $N(f_{|A}, y)$ is $\sharp(f^{-1}(\{y\}) \cap A)$, the number of points of the set $f^{-1}(\{y\}) \cap A$ and $w_n(Df_{(x)})$ is defined in the preceding chapter. In particular, the characteristic function of $f(A)$, $\mathbf{1}_{f(A)}(y) \leqslant N(f_{|A}, y)$, and thus we obtain the following inequality, showing that the n-dimensional volume under f is bounded by the integration of the Jacobian:

$$\mathcal{H}^n(f(A)) \leqslant \int_{x \in A} w_n(Df_{(x)})\ d\mathcal{H}^n(x).$$

It follows that when A is a critical set for f (for all $x \in A$, $Df_{(x)} = 0$), then the n-volume of $f(A)$ is null.

However, when $n > m$ the situation is much more complicated; e.g. the function f of Whitney (see [Whi 1]) is C^{n-1} smooth on B_1^n, but $f(A) = [0; 1]$, although A is a connected set of critical points of f (see also the references of [Bat] and [Nor] for examples of functions with dense critical values set).

In the case $n > m$, we have the so-called *coarea formula*:

$$\int_{y \in M} \mathcal{H}^{n-m}(A \cap f^{-1}(\{y\}))\ d\mathcal{H}^m(y) = \int_{x \in A} w_m(Df_{(x)})\ d\mathcal{H}^n(x),$$

Y. Yomdin and G. Comte: LNM 1834, pp. 83–98, 2004.

which says that, for A a subset of critical points, and $\mathcal{H}^m$-almost all critical values y of f, the fiber $A \cap f^{-1}(\{y\})$ is $\mathcal{H}^{n-m}$-null.

Our first result in this chapter shows that for polynomial mappings of semialgebraic sets of different dimensions, but of fixed algebraic complexity, the formula of integration with the Jacobian is still valid in some sense, for variations.

Theorem 7.1. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a polynomial mapping, $f = (f_1, \ldots, f_m)$, $\deg(f_j) = \bar{d}_j$, for $j \in \{1, \ldots, m\}$. Let $A \subset \mathbb{R}^n$ be a semialgebraic subset. Assume that at each $x \in A$ and for any i-dimensional affine space $\bar{P} \subset \mathbb{R}^n$ tangent[1] to A at x, $w_i(Df_{(x)|\bar{P}}) \leqslant \gamma$. Then:*

$$V_i(f(A)) \leqslant C.\gamma.V_i(A),$$

with the constant C depending only on $\bar{d}_1, \ldots, \bar{d}_m$, m, n and $D(A)$, the diagram of A.

Remark. We do not know whether or not the constant C in the above inequality really depends on the degrees $\bar{d}_1, \ldots, \bar{d}_m$.

Proof. By definition,

$$V_i(f(A)) = c(m,i) \int_{P \in G_m^i} \int_{x \in P} V_0(f(A) \cap \bar{P}_x)\, dx\, dP.$$

Obviously we have: $V_0(f(A) \cap \bar{P}_x) \leqslant V_0(A \cap f^{-1}(\bar{P}_x))$. Let us estimate the integer $V_0(A \cap f^{-1}(\bar{P}_x))$ with Corollary 4.9.

There exist $L_1, \ldots, L_i : \mathbb{R}^m \to \mathbb{R}$, i affine linear forms in general position such that $\bar{P}_x = L_1^{-1}(\{0\}) \cap \ldots \cap L_i^{-1}(\{0\})$. The algebraic set $f^{-1}(P_x))$ is thus given by $\bigcap_{j=1}^{i} B_j$, where $B_j = \{x \in \mathbb{R}^n; (L_j \circ f)(x) = 0\}$. The degree of $L_j \circ f$ is less than $\mu = \max_{j \in \{1,\ldots,i\}} \bar{d}_j$. Now $A \cap f^{-1}(\bar{P}_x))$ is given by $\bigcup_{k=1}^{p} (\bigcap_{j=1}^{j_k} A_{kj} \bigcap_{j=1}^{i} B_j)$, where $A_{kj} = \{x \in \mathbb{R}^n; p_{kj}(x)\ ?\ 0\}$ and $?$ is one of the symbols: $>, \geq$, and $\deg(p_{kj}) = \bar{d}_{kj}$. By Corollary 4.9, we obtain:

$$V_0(A \cap f^{-1}(\bar{P}_x)) \leqslant C' = \frac{1}{2} \sum_{k=1}^{p} (\bar{d}_k + i\mu + 2)(\bar{d}_k + i\mu + 1)^{n-1}, \text{ with } \bar{d}_k = \sum_{j=1}^{j_k} \bar{d}_{kj}.$$

We thus have:

$$V_i(f(A)) \leqslant c(m,i).C' \int_{P \in G_m^i} Vol_i([\pi_P \circ f](A))\, dP, \qquad (*)$$

[1] The affine space $\bar{P} \subset \mathbb{R}^n$ is tangent to A at $x \in A$ if every direction ν of $\bar{P}$ is a tangential direction of A, that is to say there exists a smooth curve α in A with $\alpha(0) = x$ and $\alpha'(0) = \nu \in \bar{P}$.

where $\pi_P : \mathbb{R}^m \to P$ is the orthogonal projection.

It remains to estimate the i-size of $\pi_P(f(A))$. For this we cannot immediately use the area formula for $\pi_P \circ f\ : A^0 \to P$ (A^0 being the nonsingular locus of A), because the interesting case is $\dim(A) \geq i$. But Theorem 4.10 will allow us to use this formula: there exists a semialgebraic subset $C \subset A$, such that $\dim(C) = \dim(\pi_P(f(A)))$, $\pi_P(f(A)) = \pi_P(f(C))$, and the diagram $D(C)$ only depends on $D(A)$.

Of course we want to estimate $V_i(f(A))$ for $i \leqslant \dim(f(A))$, and for generic $P \in G^i_m$ we have $\dim(C) = \dim(\pi_P(f(A))) = \dim(f(A))$; it implies that generically $\dim(C) = i$.

Now we can use the area formula for $\pi_P \circ f\ : C^0 \to P$:

$$Vol_i([\pi_P \circ f](A)) \leqslant \int_{x \in C^0} Jac((\pi_P \circ f)_{|C^0})(x)\ d\mathcal{H}^i(x).$$

But by our assumptions

$$Jac((\pi_P \circ f_{|C^0}))(x) = w_i(D(\pi_P \circ f_{|C^0})_{(x)}) \leqslant w_i(Df_{(x)|\ \mathrm{T}_x C^0}) \leqslant \gamma,$$

where $\mathrm{T}_x C^0$ is the tangent space of C^0 at x. Hence, by Theorem 5.1:

$$Vol_i([\pi_P \circ f](A)) \leqslant \gamma Vol_i(C) = \gamma V_i(C) \leqslant \gamma.C''.V_i(A),$$

and finally

$$V_i(f(A)) \leqslant c(m,i).C'.C''.\gamma.V_i(A) \int_{P \in G^i_m} dP = C.\gamma.V_i(A).$$

□

The proof above is based on Theorem 4.10, which gives no explicit bound on the complexity of the covering set C. Below we prove a little bit weaker result, which however uses Exercise 4.11 instead of Theorem 4.10, and hence gives an explicit bound for the coefficient in the inequality. We also use a stronger, but easier to verify, condition $w_i(Df_{(x)}) \leqslant \gamma$, instead of requiring this inequality only on i-dimensional subspaces tangent to A.

Theorem 7.2. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a polynomial mapping, $f = (f_1, \ldots, f_m)$, $\deg(f_j) = \bar{d}_j$, for $j \in \{1, \ldots, m\}$. Let $A \subset \mathbb{R}^n$ be a bounded semialgebraic subset. Assume that at each $x \in A$, $w_i(Df_{(x)}) \leqslant \gamma$. Then:*

$$V_i(f(A)) \leqslant \widetilde{C}.\gamma.V_i(A),$$

where

$$\widetilde{C} = \frac{1}{4} c(m,i).\Big(\sum_{k=1}^{p} (d_k + i\mu + 2)(d_k + i\mu + 1)^{n-1}\Big)\kappa,$$

$$\mu = \max_{j \in \{1,\ldots,i\}} \bar{d}_j,\ d_k = \sum_{j=1}^{j_k} d_{kj},$$

and κ only depends on $d_k, \bar{d}_j, n, i$, assuming that A is given by $\bigcup_{k=1}^{p}(\bigcap_{j=1}^{j_k} A_{kj})$, *where* $A_{kj} = \{x \in \mathbb{R}^n; p_{kj}(x) \ ? \ 0\}$ *and ? is one of the symbols: $>, \geq$, and* $\deg(p_k j) = d_{kj}$.

Proof. We proceed exactly as in the proof of theorem 7.1. We obtain the inequality $(*)$:

$$V_i(f(A)) \leqslant c(m,i).C' \int_{P\in G^i_m} Vol_i([\pi_{\bar{P}} \circ f](A))\, dP, \qquad (*)$$

where $C' = \dfrac{1}{2}\sum_{k=1}^{p}(d_k+i\mu+2)(d_k+i\mu+1)^{n-1}$, $\mu = \max_{j\in\{1,\dots,i\}}\bar{d}_j$ and $d_k = \sum_{j=1}^{j_k} d_{kj}$. To estimate $Vol_i(\pi_P(f(A)))$ we use Exercise 4.11, applied to the polynomial $\pi_P \circ f : \mathbb{R}^n \to P \in G^i_m$, instead of Theorem 4.10. It provides for any $\xi > 0$ a semialgebraic set C, such that C is contained in the ξ-neighborhood of A, $\mathrm{d}_{\mathcal{H}}(\pi_P(f(C); \pi_P(f(A)))) \leqslant K\xi$, for all $P \in G^i_m$, where K is a Lipschitz constant of f on a bounded subset of $\mathbb{R}^n$ containing A, and $\dim(C) \leqslant i$. We obtain by Theorem 5.9:

$$Vol_i(\pi_P(f(A))) = \lim_{\xi\to 0} Vol_i(\pi_P(f(C))).$$

And since we can assume that $w_i(\pi_P \circ f)_{|C^0} \leqslant \gamma + \eta(\xi)$, with $\eta(\xi)$ arbitrarily small as $\xi \to 0$, we obtain by the area formula:

$$Vol_i(\pi_P(f(A))) = \lim_{\xi\to 0} Vol_i(\pi_P(f(C))) \leqslant \gamma.Vol_i(C) = \gamma.V_i(C).$$

Furthermore by Theorem 5.1, we have $V_i(C) \leqslant B_{0,n-i}(C).V_i(A_\xi)$, where $B_{0,n-i}(C)$ is a bound for the number of connected components of $C \cap Q$, for any $Q \in \bar{G}^{n-i}_n$. Now, by Theorem 5.11, $\lim_{\xi\to 0} V_i(A_\xi) \leqslant V_i(A)$. But again by Exercise 4.11 we can take $B_{0,n-i}(C)$ depending only on $d_k, \bar{d}_j, n, i$, and we get from $(*)$ the desired inequality:

$$Vol_i(f(A)) \leqslant c(m,i).C'.B_{0,n-i}(C).\gamma.V_i(A).$$

□

Remark. In the rest of this chapter we prove the results requiring "covering" subsets C, using Theorem 4.10 to simplify the arguments. However, in any of these results, Theorem 4.11 could be used exactly as in the proof above. This allows us to obtain the explicit expressions for the coefficients below.

Definition 7.3. Let $\Lambda = (\lambda_1, \dots, \lambda_q)$, with $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_q \geq 0$, be given.

For any differentiable $f : \mathbb{R}^n \to \mathbb{R}^m$, with $q = \min(n, m)$, define $\Sigma(f, \Lambda)$ as the set $\{x \in \mathbb{R}^n;\ \lambda_i(Df_{(x)}) \leqslant \lambda_i, i = 1, 2, \dots q\}$, where $\lambda_i(Df_{(x)})$ is defined in Chapter 6. The set $\Sigma(f, \Lambda)$ is the set of points x such that $Df_{(x)} : \mathbb{R}^n \to \mathbb{R}^m$ transforms B_1^n into an ellipsoid in $\mathbb{R}^m$ with semiaxes (in decreasing order) smaller than $\lambda_1, \dots, \lambda_q$, respectively.

For a semialgebraic set A we denote by $\Sigma(f, \Lambda, A)$ the set $\Sigma(f, \Lambda) \cap A$, and by $\Delta(f, \Lambda, A)$ the set $f(\Sigma(f, \Lambda, A)) \subset \mathbb{R}^m$.

Corollary 7.4. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a polynomial mapping, $f = (f_1, \dots, f_m)$, $\deg(f_j) = d_j$, and A a bounded semialgebraic set of $\mathbb{R}^n$. Then:*

$$V_i(\Delta(f, \Lambda, A)) \leqslant C.\lambda_0\lambda_1 \dots \lambda_i.V_i(A),$$

where $\lambda_0 = 1$ by definition, and C depends only on $d_1, \dots, d_m$, n, m and $D(A)$.

In particular, when $A = B^n_{(a,r)}$, we get:

$$V_i(\Delta(f, \Lambda, B^n_{(a,r)})) \leqslant C.\lambda_0\lambda_1 \dots \lambda_i.r^i,$$

where C depends only on $d_1, \dots, d_m$, n and m.

Proof. We just have to apply Theorem 7.2 to the set $\Sigma(f, \Lambda, A)$, which is semialgebraic with diagram depending only on $d_1, \dots, d_m$, n, m and $D(A)$. We have $w_i(Df_{(x)}) \leqslant \gamma = \lambda_0\lambda_1 \dots \lambda_i$ for all $x \in \Sigma(f, \Lambda, A)$. Hence we obtain:

$$V_i(\Delta(f, \Lambda, A)) \leqslant C'.\lambda_0\lambda_1 \dots \lambda_i.V_i(\Sigma(f, \Lambda, A)).$$

But by Theorem 5.1, $V_i(\Sigma(f, \Lambda, A)) \leqslant C'.V_i(A)$.
In particular, when $A = B^n_{(a,r)}$, we get:

$$V_i(\Delta(f, \Lambda, B^n_{(a,r)})) \leqslant C''.\lambda_0\lambda_1 \dots \lambda_i.r^i.$$

□

As an immediate consequence we obtain one of the main results of this section - the quantitative Morse-Sard theorem for polynomial mappings:

Theorem 7.5. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a polynomial mapping, $f = (f_1, \dots, f_m)$, $\deg(f_j) = d_j$ and A a bounded semialgebraic set in $\mathbb{R}^n$. Then for any $\Lambda = (\lambda_1, \dots, \lambda_q)$, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_q \geq 0$, $q = \min(n, m)$,*

$$M(\epsilon, \Delta(f, \Lambda, A)) \leqslant C \sum_{i=0}^{q} \lambda_0 \dots \lambda_i \cdot V_i(A) \cdot (\frac{1}{\epsilon})^i,$$

where C depends only on $d_1, \dots, d_m$, n, m and $D(A)$. In particular, for $A = B^n_{(a,r)}$, we obtain:

$$M(\epsilon, \Delta(f, \Lambda, B^n_{(a,r)})) \leqslant C \sum_{i=0}^{q} \lambda_0 \dots \lambda_i \cdot (\frac{r}{\epsilon})^i.$$

Proof. It is a direct consequence of Theorem 3.5 and Corollary 7.4. □

Now let us consider the set $\Delta_f^0 = \{x \in \mathbb{R}^n;\ Df_{(x)} = 0\}$. Clearly $\Delta_f^0 = \Delta(f, \Lambda, B^n_{(a,r)})$, for $\Lambda = (0, \ldots, 0)$. Hence we get:

Corollary 7.6. *For $f\ : \mathbb{R}^n \to \mathbb{R}^m$ a polynomial mapping, with $f = (f_1, \ldots, f_m)$, $\deg(f_j) = d_j$.*

$$M(\epsilon, \Delta_f^0) \leqslant C,$$

where $C = C(n, d_1, \ldots, d_m)$ depends only on $d_1, \ldots, d_m, n$ and m. □

Clearly this result recovers the simple fact that the critical values of rank 0 of a polynomial mapping form a finite set with a number of points bounded by the degrees (once we know that one can stratify a semialgebraic set with a number of connected strata depending only on the degrees, this fact is obvious).

Let us consider now $\Delta_f^\epsilon = \Delta(f, \Lambda, B^n_{(a,r)})$, for $\Lambda = (\frac{\epsilon}{r}, \ldots, \frac{\epsilon}{r})$. We can now prove an extended version of Theorem 1.8:

Theorem 1.8. (extended) *Let $f\ : \mathbb{R}^n \to \mathbb{R}^m$ be a polynomial mapping, with $f = (f_1, \ldots, f_m)$, $\deg(f_j) = d_j$. Then for any $\epsilon > 0$ the set Δ_f^0 can be covered by $N(n, m, d_1, \ldots, d_m)$ balls of radius ϵ. The constant $N(n, m, d_1, \ldots, d_m)$ depends only on $d_1, \ldots, d_m, n$ and m.*

In particular the set Δ_f^0 of rank-0 critical values of $f_{|B_r^n}$ contains at most $N(n, m, d_1, \ldots, d_m)$ points, for any $r > 0$, hence the number of critical values of f is bounded by $N(n, m, d_1, \ldots, d_m)$.

Proof. The proof follows directly from Theorem 7.5. □

Remarks.

- We notice that, as a consequence of Theorem 7.5 and its smooth counterpart, some "rigidity" results for smooth functions can be obtained: if the function has inside a given ball "too many" critical points, with the values at these points "too separated", its high order derivatives must be "big". We do not touch these questions here (see [Yom 1], Theorem 3.9, Corollary 3.10 and Theorem 3.11.)
- Obviously Theorem 7.1, Theorem 7.2, Corollary 7.4, Theorem 7.5, Corollary 7.6 and Theorem 1.8 are true for $f : \mathbb{R}^n \to \mathbb{R}^m$ a mapping definable in some o-minimal structure (i.e. a mapping whose graph is a definable set in some o-minimal structure). The constants that appear in these statements depending only on the set A and the mapping f (the notions of degree and diagram making no more sense in o-minimal structures).

Theorem 7.2 gives an upper bound for the variations of the images of the polynomial mapping, in terms of the uniform upper bounds for the first derivatives of the mapping.

The following example shows that the uniform lower bounds for the derivatives do not allow one to bound variations of the image from below.

Let $A \subset \mathbb{R}^2$ be the rectangle $[0;\delta] \times [0;1]$ and $f = \pi_1 : (x,y) \mapsto x$ (see Fig. 7.1).

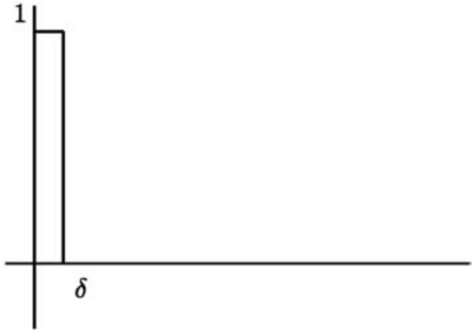

Fig. 7.1.

Then $w_1(f) = 1$, $V_1(A) \geq 1$ (because for instance an affine line that encounters $I = \{0\} \times [0;1]$ also encounters A, and by choice of the coefficient $c(2,1)$, the measure of the set of affine lines that encounter I is $\dfrac{1}{c(2,1)}$), although $f(A) = [0;\delta]$, and $V_1(f(A)) = \delta$ can be arbitrarily small.

If we use the integral norms of the derivatives, the situation is the opposite: in general we can bound variations of the image from below, but not from above.

In the results below, we assume that $\dim(A) = s$ for a semialgebraic set $A \subset \mathbb{R}^n$, and then we integrate on A with respect to the s-dimensional Hausdorff measure $\mathcal{H}^s$ of A. The regular part A^0 of A having the same measure as A, it is the same to integrate on A^0, and thus one can integrate the volume form on A^0 (there exists $Z \subset A^0$ such that $\dim(Z) < \dim(A^0)$ and $A^0 \setminus Z$ is orientable.)

Theorem 7.7. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a polynomial mapping, $f = (f_1, \ldots, f_m)$, $\deg(f_j) = d_j$, and let $A \subset \mathbb{R}^n$ be a semialgebraic set with $\dim(A) = s$, $i \leqslant s$. Then*

$$\int_{x \in A} w_i(D[f_{|A^0}]_{(x)}) \ \ d\mathcal{H}^s(x) \leqslant C \cdot V_i(f(A)) \cdot V_{s-i}(A),$$

where the constant C depends only on $d_1, \ldots, d_m, n, m$ and $D(A)$.

Proof. Let G_m^i denote, as usual, the grassmannian of all the i-dimensional linear spaces of $\mathbb{R}^m$. Let $x \in A^0$ and $\mathrm{T}_x A^0$ be the tangent space of A^0 at x, $\mathrm{T}_x A^0 \sim \mathbb{R}^s$. We consider the linear mapping $Df_{|A^0}(x) : \mathrm{T}_x A^0 \to \mathbb{R}^m$ as

$Df_{|A^0}(x) \ : \mathbb{R}^s \to \mathbb{R}^m$, and for each $P \in G^i_m$ we define an s-form on $\mathrm{T}_x A^0$, $\omega(P)$, as in lemma 6.4 above, by:

$$\omega(P) = [(\pi_{\ker(\pi_P \circ Df_{|A^0(x)})})^* \omega_{\ker(\pi_P \circ Df_{|A^0(x)})}] \wedge [(\pi_P \circ f_{|A^0})^* \omega_P],$$

where ω_V denotes the volume form on the Euclidean space V.

By Lemma 6.4:

$$w_i(Df_{|A^0(x)}) \cdot \omega_A \leqslant C \cdot \int_{P \in G^i_m} \omega(P)\ dP \leqslant C \int_{P \in G^i_m} \omega(P)\ dP,$$

where ω_A is the volume form on A^0.

Hence:

$$\begin{aligned}\int_{x \in A^0} w_i(Df_{|A^0(x)})\ d\mathcal{H}^s(x) &= \int_{x \in A^0} w_i(Df_{|A^0(x)})\ \ \omega_A \\ &\leqslant C \cdot \int_{x \in A^0} [\int_{P \in G^i_m} \omega(P)\ dP] \\ &= C \cdot \int_{P \in G^i_m} \Big[\int_{x \in A^0} [(\pi_{\ker(\pi_P \circ Df_{|A^0(x)})})^* \omega_{\ker(\pi_P \circ Df_{|A^0(x)})}] \\ &\qquad \wedge [(\pi_P \circ f_{|A^0})^* \omega_P] \ \Big]\ dP\end{aligned}$$

$$= C \cdot \int_{P \in G^i_m} \Big[\int_{\xi \in (\pi_P \circ f)(A)} \Big[\int_{x \in (\pi_P \circ f)^{-1}(\xi)} \omega_{(\pi_P \circ f)^{-1}(\xi)}(x) \ \Big]\ \omega_P(\xi) \ \ \Big]\ dP.$$

The integral $\displaystyle\int_{x \in (\pi_P \circ f)^{-1}(\xi)} \omega_{(\pi_P \circ f)^{-1}(\xi)}(x)$, for ξ a regular value of $\pi_P \circ f$, is simply the integral of the volume form on the fiber $(\pi_P \circ f)^{-1}(\xi)$, and hence is equal to $Vol_{s-i}((\pi_P \circ f)^{-1}(\xi))$. In turn, this volume is equal to the $(s-i)$-variation of $(\pi_P \circ f)^{-1}(\xi)$, and it is bounded by $\widetilde{C}.V_{s-i}(A)$ (by Theorem 5.1), since $(\pi_P \circ f)^{-1}(\xi)$ is a semialgebraic subset of A, with diagram depending only on $D(A)$ and $d_1, \dots, d_m$.

We can assume that for almost all $P \in G^i_m$, almost all the values $\xi \in (\pi_P \circ f)(A) \subset P$ are regular for $\pi_P \circ f$, because if it is not the case, $w_i(Df_{|A^0(x)}) = 0$ for all $x \in A^0$. Hence:

$$\int_{x \in A^0} w_i(Df_{|A^0(x)})\ d\mathcal{H}^s(x) \leqslant C \cdot \widetilde{C}.V_{s-i}(A) \int_{P \in G^i_m} [\int_{\xi \in (\pi_P \circ f)(A)} \omega_P(\xi)]\ dP$$

$$\leqslant C \cdot \widetilde{C} \cdot V_{s-i}(A) \int_{P \in G^i_m} [\int_{x \in P} V_0(f(A) \cap P_x)\ dx]\ dP$$

$$= C \cdot \widetilde{C} \cdot V_{s-i}(A) \cdot \frac{1}{c(m,i)} V_i(f(A)) = C' \cdot V_i(f(A)) \cdot V_{s-i}(A)$$

□

Remark. As an immediate corollary we obtain again the result of Theorem 5.15: $V_j(A) \leqslant c \cdots V_i(A) V_{j-i}(A)$, for all $j \in \{0, \ldots, n\}$ and all $i \leqslant j$.

Indeed, for $j = s = \dim(A)$, applying Theorem 7.7 to the inclusion mapping of A into $\mathbb{R}^n$, we have for all $i \leqslant j$, $w_i(Df_{|A^0(x)}) = 1$, thus:

$$\int_{x \in A^0} w_i(Df_{|A^0(x)})\, d\mathcal{H}^j(x) = Vol_j(A) = V_j(A) \leqslant C V_i(A) V_{j-i}(A).$$

Now for $j \leqslant s = \dim(A)$, by Theorem 5.20 there exists a semialgebraic set $C_j \subset A$, with $\dim(C_j) = j$, $D(C_j)$ depending only on $D(A)$, such that $V_j(C_j) \geq \lambda V_j(A)$, λ depending only on $D(A)$. Applying Theorem 7.7 and the same remarks as above to C_j, we have:

$$V_j(A) \leqslant \frac{1}{\lambda} V_j(C_j) \leqslant \frac{C}{\lambda} V_i(C_j) V_{j-i}(C_j) \leqslant \frac{C}{\lambda} V_i(A) V_{j-i}(A).$$

(The last inequality being a consequence of Theorem 5.1.)

In fact, Theorem 7.7 can be formulated in a stronger form, if we introduce the notion of variations of functions and mappings, following more or less [Iva 1].

Definition 7.8. Let $A \subset \mathbb{R}^n$ be a compact subset, and let $f : A \to \mathbb{R}^m$ be a continuous mapping. For $i = 0, 1, \ldots, n$ and $\ell = 0, 1, \ldots, m$, the variation $V_{i,\ell}(f)$ is defined as follows:

$$V_{i,\ell}(f) = c(i, \ell, n, m) \int_{\bar{P} \in \bar{G}_m^{m-\ell}} V_i(f^{-1}(\bar{P}))\, d\bar{P},$$

with $c(i, \ell, n, m)$ a well-chosen constant depending only on i, ℓ, n and m.

The following properties of $V_{i,\ell}(f)$ are immediate:

(1) $V_{i,0}(f) = V_i(A)$, if we choose $c(i, 0, n, m) = 1$.

(2) $V_{i,m}(f) = c(i, m, n, m) \int_{\xi \in \mathbb{R}^m} V_i(f^{-1}(\xi))\, d\xi$.

(3) $V_{0,\ell}(f) \geq V_\ell(f(A))$, if we choose $c(0, \ell, n, m) = c(\ell, m)$. Indeed $V_{0,\ell}(f) = c(0, \ell, n, m) \int_{\bar{P} \in \bar{G}_m^{m-\ell}} V_0(f^{-1}(\bar{P}))\, d\bar{P}$, but clearly the number of connected components of $f(A) \cap P = f(f^{-1}(P))$ is less than the number of connected components of $f^{-1}(P)$, and hence:

$$V_{0,\ell}(f) \geq c(0, \ell, n, m) \int_{\bar{P} \in \bar{G}_m^{m-\ell}} V_0(f(A) \cap \bar{P})\, d\bar{P} = V_\ell(f(A)).$$

(4) Inductive formula for variations of mappings. For $j \leqslant i$, we have:

$$V_{i,\ell}(f) = \tilde{c}(j, i, \ell, n, m) \int_{\bar{P} \in \bar{G}_n^{n-j}} V_{i-j,\ell}(f_{|A \cap \bar{P}})\, d\bar{P}.$$

In particular, for $j = i$ we obtain:

$$V_{i,\ell}(f) = \widetilde{c}(i,i,\ell,n,m) \int_{\bar{P}\in\bar{G}_n^{n-i}} V_{0,\ell}(f_{|A\cap\bar{P}})\ d\bar{P}.$$

Combining **(3)** and **(4)**, we obtain:

(5) $V_{i,\ell}(f) \geq \widetilde{c}(i,i,\ell,n,m) \int_{\bar{P}\in\bar{G}_n^{n-i}} V_\ell(f(A\cap\bar{P}))\ d\bar{P}.$

There is an important special case, where the equality holds in **(3)** and **(5)**:

(6) If $f : A \to \mathbb{R}^m$ is an embedding, then:

$$V_{0,\ell}(f) = V_\ell(f(A)),$$

$$V_{i,\ell}(f) = \widetilde{c}(i,i,\ell,n,m) \int_{\bar{P}\in\bar{G}_n^{n-i}} V_\ell(f(A\cap\bar{P}))\ d\bar{P}.$$

Indeed in this case, $V_0(f^{-1}(\bar{P})) = V_0(f(A)\cap\bar{P})$.

More generally, we have the following:

(7) If $f : A \to f(A)$ is a tame finite mapping of multiplicity at most q, then

$$V_{0,\ell}(f) \leqslant q.V_\ell(f(A)),$$

$$V_{i,\ell}(f) \leqslant q.\widetilde{c}(i,i,\ell,n,m) \int_{\bar{P}\in\bar{G}_n^{n-i}} V_\ell(f(A\cap\bar{P}))\ d\bar{P}.$$

Indeed, under our assumption, over the complement of a set of dimension $< \dim(A)$ in $f(A)$, f is a q-covering. Hence over each connected component of $f(A)\cap\bar{P}$, we have at most q components of $f^{-1}(\bar{P})$. Thus:

$$V_{0,\ell}(f) = c(0,\ell,n,m) \int_{\bar{P}\in\bar{G}_m^{m-\ell}} V_0(f^{-1}(\bar{P}))\ d\bar{P}$$

$$\leqslant c(0,\ell,n,m) \int_{\bar{P}\in\bar{G}_m^{m-\ell}} q.V_0(f(A)\cap\bar{P})\ d\bar{P} = qV_\ell(f(A)).$$

There is one special situation, where the variations are reduced to the usual integral-geometric invariant and hence can be computed in "closed form". Namely, for $i + \ell = n$ and $f : \mathbb{R}^n \to \mathbb{R}^m$ smooth and semialgebraic, we have:

$$V_{i,\ell}(f) = c(i,\ell,n,m) \int_{\bar{P}\in\bar{G}_m^{m-\ell}} V_i(f^{-1}(\bar{P}))\ d\bar{P}$$

$$= c(i,\ell,n,m) \int_{\bar{P}\in\bar{G}_m^{m-\ell}} Vol_i(f^{-1}(\bar{P}))\ d\bar{P},$$

since by Sard's theorem there exists a smooth semialgebraic set $Y \subset f(\mathbb{R}^n)$, such that $\mathrm{codim}(f(\mathbb{R}^n) \setminus Y) \geq 1$ and $f : \mathbb{R}^n \to Y$ is a submersion. Thus for a generic $\bar{P} \in \bar{G}_m^{m-\ell}$, $\dim(Y \cap \bar{P}) = \dim(Y) - \ell$, and $\dim(f^{-1}(\bar{P})) = n - \mathrm{codim}_Y(Y \cap \bar{P}) = n - \ell = i$.

To compute these variations, we need some additional constructions in linear algebra.

Definition 7.9. Let $\mathrm{L} : \mathbb{R}^n \to \mathbb{R}^m$ be a linear mapping. For $i \leqslant q = \min(n, m)$, we define $\omega_i(\mathrm{L})$ as follows:

$$\omega_i(\mathrm{L}) = \frac{1}{V_i} \int_{P \in G_m^i} Vol_i\big([\pi_P \circ \mathrm{L}](B_1^n)\big)\ dP,$$

where B_1^n is as usual the unit ball centered at the origin of $\mathbb{R}^n$, and V_i is the volume of the unit i-dimensional ball.

Remark. If $\mathrm{rank}(\mathrm{L}) = i$, $\dim(\mathrm{L}(B_1^n)) = i$, and the Cauchy-Crofton formula for the volume gives:

$$\omega_i(\mathrm{L}) = \frac{1}{V_i.c(m,i)} c(m,i) \int_{P \in G_m^i} Vol_i\big(\pi_P(\mathrm{L}(B_1^n))\big)\ dP$$

$$= \frac{Vol_i(\mathrm{L}(B_1^n))}{V_i.c(m,i)} = \frac{\lambda_1(\mathrm{L}) \ldots \lambda_i(\mathrm{L})}{c(m,i)} = \frac{w_i(\mathrm{L})}{c(m,i)}.$$

In the notations of lemma 6.4 above, we have now the following:

Lemma 7.10. *For $P \in G_m^i$ and for*

$$\omega(P) = [(\pi_{\ker(\pi_P \circ \mathrm{L})})^* \omega_{\ker(\pi_P \circ \mathrm{L})}] \wedge [(\pi_P \circ \mathrm{L})^* \omega_P],$$

we have:

$$\int_{P \in G_m^i} \omega(P)\ dP = \omega_i(\mathrm{L}).\omega_{\mathbb{R}^n}.$$

Proof. We clearly have:

$$[(\pi_P \circ \mathrm{L})^* \omega_P] = \lambda_1(\pi_P \circ \mathrm{L}) \ldots \lambda_i(\pi_P \circ \mathrm{L})(\pi_{\ker^\perp(\pi_P \circ \mathrm{L})})^* \omega_{\ker^\perp(\pi_P \circ \mathrm{L})}.$$

Of course we suppose that $\mathrm{rank}(\mathrm{L}) \geq i$, and thus that for generic $P \in G_m^i$, $\mathrm{rank}(\pi_P \circ \mathrm{L}) = i$ (if it is not the case, each side of the equality we want to prove is 0).

Now $\lambda_1(\pi_P \circ \mathrm{L}) \ldots \lambda_i(\pi_P \circ \mathrm{L}).Vol_i(B_1^i) = Vol_i([\pi_P \circ \mathrm{L}](B_1^n))$, and

$$\int_{P \in G_m^i} \omega(P) dP = \frac{1}{V_i} [(\pi_{\ker(\pi_P \circ \mathrm{L})})^* \omega_{\ker(\pi_P \circ \mathrm{L})}] \wedge [(\pi_{\ker^\perp(\pi_P \circ \mathrm{L})})^* \omega_{\ker^\perp(\pi_P \circ \mathrm{L})}] \times$$

$$\int_{P\in G^i_m} Vol_i([\pi_P \circ \mathrm{L}](B^n_1))dP$$

$$= \omega_i(\mathrm{L}).\omega_{\mathbb{R}^n}. \qquad \square$$

Now Lemma 6.4 can be rewritten as follows:

Lemma 7.11. *Let* $\mathrm{L} : \mathbb{R}^n \to \mathbb{R}^m$ *be a linear mapping. For* $i \leqslant q = min(n, m)$, *we have:*

$$\omega_i(\mathrm{L}) \leqslant w_i(\mathrm{L}) \leqslant C \cdot \omega_i(\mathrm{L}),$$

where C *depends only on* n, m *and* i. $\square$

Remark. By the remark following Definition 7.9, the equality holds in Lemma 7.11, when $\mathrm{rank}(\mathrm{L}) = i$. In this case we have: $\omega_i(\mathrm{L}) = \dfrac{w_i(\mathrm{L})}{c(m,i)}$.

Theorem 7.12. *Let* A *be a compact smooth* s*-dimensional submanifold of* $\mathbb{R}^n$ *and* $f : A \to \mathbb{R}^m$ *be a smooth mapping. We have:*

$$V_{s-i,i}(f) = c(s-i, i, n, m) \int_{x\in A} \omega_i(Df_{|A^0(x)})\ d\mathcal{H}^s(x).$$

Proof. By definition

$$V_{s-i,i}(f) = c(s-i, i, n, m) \int_{\bar{P}\in \bar{G}^{m-i}_m} V_{s-i}(f^{-1}(\bar{P}))\ d\bar{P}$$

$$= c(s-i, i, n, m) \int_{P\in G^i_m} \int_{\xi\in P} V_{s-i}([\pi_P \circ f]^{-1}(\xi))\ d\xi\ dP.$$

For all $P \in G^i_m$, almost all the points $\xi \in P$ are regular values of $(\pi_P \circ f)_{|A^0}$, by the classical Sard theorem, hence $[\pi_P \circ f]^{-1}(\xi) = Y_{P,\xi}$ is empty or is a smooth submanifold in A^0 of dimension $s-i$. We obtain:

$$V_{s-i,i}(f) = c(s-i, i, n, m) \int_{P\in G^i_m} \int_{\xi\in P} Vol_{s-i}(Y_{P,\xi})\ d\xi\ dP$$

$$= c(s-i, i, n, m) \int_{P\in G^i_m} \int_{\xi\in P} \int_{Y_{P,\xi}} \omega_{Y_{P,\xi}}\ d\xi\ dP = c(s-i, i, n, m)$$

$$\int_{P\in G^i_m} \Big[\int_{x\in A^0} [(\pi_{\ker(\pi_P \circ Df_{|A^0(x)})})^* \omega_{\ker(\pi_P \circ Df_{|A^0(x)})}] \wedge [(\pi_P \circ f_{|A^0})^* \omega_P] \Big]\ dP$$

$$= c(s-i, i, n, m) \int_{P\in G^i_m} [\int_{x\in A^0} \omega_P(x)]\ dP$$

$$= c(s-i, i, n, m) \int_{x\in A^0} [\int_{P\in G^i_m} \omega_P(x)\ dP]$$

$$= c(s-i,i,n,m) \int_{x\in A^0} \omega_i(Df_{|A^0(x)})\ \omega_{A^0}$$

$$= c(s-i,i,n,m) \int_{x\in A} \omega_i(Df_{|A^0(x)})\ d\mathcal{H}^s(x). \qquad \square$$

Now we turn back to semialgebraic sets and their polynomial mappings. The following theorem shows that, as above, variations of semialgebraic sets and mappings, are equivalent, up to coefficients, to their "integral-geometric" invariants - average volume of projections.

Theorem 7.13. *Let $A \subset \mathbb{R}^n$ be a bounded semialgebraic set, $f = (f_1, \ldots, f_m) : A \subset \mathbb{R}^n \to \mathbb{R}^m$ be a polynomial mapping, with $\deg(f_j) = d_j$. Then:*

(1) *For $i = 0, \ldots, n$; $\ell = 0, \ldots, m$: $V_{i,\ell}(f) \leqslant C_1.V_i(A).V_\ell(f(A))$.*

(2) *For $A \subset B^n_{r_1}$, $f(A) \subset B^m_{r_2}$: $V_{i,\ell}(f) \leqslant C_2.r_1^i.r_2^\ell$.*

(3) *$V_{i,m}(f) \leqslant C_3.V_m(f(A)).V_i(A)$.*

(4) *$V_{0,\ell}(f) \leqslant C_4.V_\ell(f(A))$.*

(5) *$V_{i,\ell}(f) \leqslant C_5. \int_{P\in G^{n-\ell}_n} V_\ell(f(A\cap P))\ dP$.*

All the constants here depend only on $D(A)$, m, n, $d_1, \ldots, d_m$.

Proof. (1). We have:

$$V_{i,\ell}(f) = c(i,\ell,n,m) \int_{\bar{P}\in \bar{G}^{m-\ell}_m} V_i(f^{-1}(\bar{P}))\ d\bar{P},$$

and by Theorem 5.1,

$$V_i(f^{-1}(\bar{P})) \leqslant B_{0,n-i}(f^{-1}(\bar{P})).V_i(A) \leqslant B_{0,n-i}(f^{-1}(\bar{P})).V_i(A).V_0(f(A)\cap \bar{P}).$$

Now, we have shown in the proof of Theorem 7.1 that there exists C', a constant depending only on $D(A)$, such that: $V_0(f^{-1}(\bar{P})) \leqslant C'$. It follows that:

$$V_{i,\ell}(f) \leqslant c(i,\ell,n,m).C'.V_i(A) \int_{\bar{P}\in \bar{G}^{m-\ell}_m} V_0(f(A)\cap \bar{P})\ d\bar{P}$$
$$= c(i,\ell,n,m).C'.V_i(A).V_\ell(f(A)).$$

Inequalities **(2)**, **(3)** and **(4)** follow immediately from **(1)**. Inequality **(5)** is a consequence of inequality **(4)** and of property **(4)** of $V_{i,\ell}(f)$ listed above. $\square$

Combining Lemma 7.10, Theorem 7.11, and Theorem 7.12.(**1**), we obtain:

Corollary 7.14. *Let $A \subset \mathbb{R}^n$ be a semialgebraic set of dimension s, $f = (f_1, \ldots, f_m) : A \subset \mathbb{R}^n \to \mathbb{R}^m$ be a polynomial mapping, with $\deg(f_j) = d_j$. Then:*

$$C_1 \cdot V_{s-i,i}(f) \leqslant \int_{x\in A} w_i(Df_{|A^0(x)})\ d\mathcal{H}^s(x)$$
$$\leqslant C_2 \cdot V_{s-i,i}(f) \leqslant C_3 \cdot V_{s-i}(A) \cdot V_i(f(A)),$$

where C_1, C_2 and C_3 depend only on $D(A)$, m, n, $d_1, \ldots, d_m$. □

Remark. Theorem 7.12.(**1**), Theorem 7.11 and Lemma 7.8, or directly Corollary 7.14 give again Theorem 7.7, for a semialgebraic set of dimension s:

$$V_i(f(A)) \geq \frac{C}{V_{s-i}(A)} \int_{x\in A} w_i(Df_{|A^0(x)})\ d\mathcal{H}^s(x).$$

From Theorem 7.2 and corollary 7.14 (or Theorem 7.7), we have, for A a semialgebraic set of $\mathbb{R}^n$ of dimension s and $f : \mathbb{R}^n \to \mathbb{R}^m$ a polynomial mapping such that $w_i(Df_{(x)}) \leqslant\gamma$ on A:

$$\frac{C}{\sigma_i.V_{s-i}(A)}.\int_{x\in A} w_i(Df_{|A^0(x)})\ d\mathcal{H}^s(x) \leqslant V_i(f(A)) \leqslant C'.\gamma.V_i(A). \qquad (**)$$

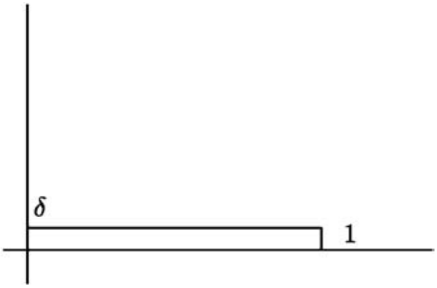

Fig. 7.2.

We have shown (example of Fig. 7.2) that it is not possible in general to obtain uniform lower bounds for $V_i(f(A))$ in terms of lower bounds for Df (one needs to integrate $w_i(Df)$ on A to obtain such a bound). Clearly, it is also impossible, in general, to give an upper bound for $V_i(f(A))$ in terms of $\displaystyle\int_{x\in A} \omega_i(Df_{(x)})\ d\mathcal{H}^s(x)$.

For instance, let A be $[0;1] \times [0;\delta] \subset \mathbb{R}^2$, $f(x,y) = x$, $w_1(f) = 1$, $V_1(f(A)) = 1$, but $\displaystyle\int_{x\in A} w_i(Df_{(x)})\ d\mathcal{H}^s(x) = \delta \to 0$ (see Fig. 7.2).

To have upper bounds for the variations of the images in integral terms we need some assumptions, providing the fibers of f are big:

Theorem 7.15. *Let $A \subset \mathbb{R}^n$ be a semialgebraic set of dimension s, $f = (f_1, \ldots, f_m) : \mathbb{R}^n \to \mathbb{R}^m$ be a polynomial mapping, with* $\deg(f_j) = d_j$, *and* *assume that for each $P \in G_m^{m-i}$ with $P \cap f(A) \neq \emptyset$, $V_{s-i}(f^{-1}(P)) \geq K$. Then:*

$$V_i(f(A)) \leqslant \frac{\widetilde{C}}{K} \int_{x \in A} w_i(Df_{|A^0(x)})\ d\mathcal{H}^s(x),$$

where the constant $\widetilde{C}$ only depends on $D(A)$, m, n, $d_1, \dots, d_m$.

Proof. We have, from Corollary 7.14:

$$\int_{x \in A} w_i(Df_{|A^0(x)})\ d\mathcal{H}^s(x) \geq C_1.V_{s-i,i}(f)$$

$$= C_1.c(s-i,i,n,m) \int_{\bar{P} \in \bar{G}_m^{m-i}} V_{s-i}(f^{-1}(\bar{P}))\ d\bar{P}.$$

But by assumption:

$$\int_{\bar{P} \in \bar{G}_m^{m-i}} V_{s-i}(f^{-1}(\bar{P}))\ d\bar{P} \geq K \int_{\bar{P} \in \bar{G}_m^{m-i}} \bar{\chi}_{[f(A)]}(P)\ d\bar{P},$$

where $\bar{\chi}_{[f(A)]}(P) = 1$ when $P \cap f(A) \neq \emptyset$ and $\bar{\chi}_{[f(A)]}(P) = 0$ when $P \cap f(A) = \emptyset$. We also have:

$$V_i(f(A)) = c(m,i) \int_{\bar{P} \in \bar{G}_m^{m-i}} V_0(f(A) \cap \bar{P})\ d\bar{P}$$

$$\leqslant c(m,i).B_{0,m-i}(f(A)) \int_{\bar{P} \in \bar{G}_m^{m-i}} \bar{\chi}_{[f(A)]}(P)\ d\bar{P}.$$

Finally, we get:

$$V_i(f(A)) \leqslant \frac{c(m,i).B_{0,m-i}(f(A))}{C_1.c(s-i,i,n,m).K} \int_{x \in A} w_i(Df_{|A^0(x)})\ d\mathcal{H}^s(x).$$

□

Under similar assumptions we can give also the lower bounds for the variations of the image in terms of the uniform lower bound of $w_i(Df_{|A^0(x)})$, as a consequence of the inequality in the remark (**) that follows Corollary 7.14.

Theorem 7.16. *Let $A \subset \mathbb{R}^n$ be a semialgebraic set of dimension s, $f = (f_1, \dots, f_m) : \mathbb{R}^n \to \mathbb{R}^m$ be a polynomial mapping, with* $\deg(f_j) = d_j$, *satisfying $w_i(Df_{|A^0(x)}) \geq \Gamma$ ($i \leqslant s$), then:*

$$V_i(f(A)) \geq \frac{C.\Gamma}{V_{s-i}(A)}\ Vol_s(A) = \frac{C.\Gamma}{V_{s-i}(A)}\ \frac{Vol_s(A)}{V_i(A)}\ V_i(A),$$

where C is a constant depending only on $D(A)$, m, n, $d_1, \dots, d_m$. □

When $\Gamma \leqslant w_i(Df_{|A^0(x)}) \leqslant \gamma$, we thus obtain, from (**):

$$\frac{C.\Gamma}{V_{s-i}(A)}\ Vol_s(A) \leqslant V_i(f(A)) \leqslant C'.\gamma.V_i(A).$$

Remark. The assumption $\dfrac{Vol_s(A)}{V_i(A)} \geq K$ means that the sections of A in the directions orthogonal to its maximal i-dimensional section, are at least K-big.

K has the physical dimension of $V_{s-i}(A)$, but of course, if $V_{s-i}(A)$ is big, K may be small: $A = [0;\delta] \times [0;1]$, $s = 2$, $i = s - i = 1$, $V_1(A) \geq 1$, $Vol_2(A) = \delta$.

From Theorem 7.16 we obtain:

Corollary 7.17. *Let $A \subset B_r^n \subset \mathbb{R}^n$ be a semialgebraic set of dimension s, $f = (f_1, \dots, f_m) : \mathbb{R}^n \to \mathbb{R}^m$ be a polynomial mapping, with $\deg(f_j) = d_j$, satisfying $w_i(Df_{|A^0(x)}) \geq \Gamma > 0$ ($i \leqslant s$), then for any ball $B_\delta^m \subset \mathbb{R}^m$:*

$$Vol_s(f^{-1}(B_\delta^m) \cap A) \leqslant \frac{c}{\Gamma}\, \delta^i . r^{s-i},$$

where c is a constant depending only on $D(A)$, m, n, $d_1, \dots, d_m$.

Proof. From Theorem 7.16 we have

$$Vol_s(\widetilde{A}) \leqslant \frac{\sigma_i}{C.\Gamma}\, V_{s-i}(\widetilde{A}).V_i(\widetilde{f}(\widetilde{A})),$$

where $\widetilde{f} = f_{|A \cap f^{-1}(B_\delta^m)} : \ \widetilde{A} = A \cap f^{-1}(B_\delta^m) \to B_\delta^m$. But $V_{s-i}(\widetilde{A}) \leqslant c_1.r^{s-i}$, since $\widetilde{A} \subset B_r^n$, and $V_i(f(A)) \leqslant c_2.\delta^i$ (by Theorem 5.1). $\square$

8 Quantitative Transversality and Cuspidal Values

Abstract. We consider families of polynomial mappings f_t, and we study the set of parameters t for which f_t has a near-critical point with value near the origin (it is well-known that general transversality results can be reduced to this situation). The variations of this critical set are bounded by its level of degeneracy. We also apply similar methods to Thom-Boardman singularities.

A general transversality result asserts that by a small perturbation any two given submanifolds can be brought into a "general position": at each intersection point their tangent spaces span the ambient space. In the introduction a quantitative question of this sort has been posed: how big a transversality can be achieved by a perturbation of a prescribed size? In this chapter we answer this question in a polynomial case, in theorem 8.1 and corollary 8.3 below. Technically, the transversality question is usually reduced to the case where one of the manifolds is just the origin. This is done by composing the imbedding mapping of the first manifold with the projection along the second one inside its tubular neighborhood. To reduce technicalities we start below with such a situation: we consider a mapping of the product of two euclidean spaces (the second representing the parameters) into the third one, and define the set Δ of those parameters, for which the mapping on the first factor is not transversal to the origin. Being quantitative, this definition in fact involves situations, close to non-transversal, and not only to the origin, but to any point, close to the origin.

A different result, based on a similar technique, is given by theorem 8.10 and corollary 8.11. It shows that the set of special critical values, attained by the mapping at "higher order" singular points, is smaller than the set of all the critical values.

Let $f = (f_1, \dots, f_m) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ be a polynomial mapping, with $\deg(f_j) = d_j$, for $j = 1, \dots m$. For some fixed $r > 0$, we consider the restriction $f : B_r^n \times B_r^m \to \mathbb{R}^m$, that we will also denote f. We will consider $t \in B_r^m$ as a parameter, and as in the usual transversality theorem, our aim is to show, that for a typical value of t, the mapping $f_t = f(.,t) : B_r^n \to \mathbb{R}^m$ is

Y. Yomdin and G. Comte: LNM 1834, pp. 99–107, 2004.

nondegenerate, in some sense[1]. To achieve this, a necessary assumption is, at least, that the parameters act nondegenerately. Hence we assume that for any $(x,t) \in B_r^n \times B_r^m$, the linear mapping $D_t f_{(x,t)} : \mathbb{R}^m \to \mathbb{R}^m$ is onto, where $D_t f_{(x,t)}$ (resp. $D_x f_{(x,t)}$) denotes the restriction of $Df_{(x,t)} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ to $\{0\} \times \mathbb{R}^m$ (resp. to $\mathbb{R}^n \times \{0\}$). We thus have, for any $(x,t) \in B_r^n \times B_r^m$ and any $i = 1, \dots, m$, $\lambda_i(D_t f_{(x,t)}) \neq 0$. But by compactness of $\bar{B}_r^n \times \bar{B}_r^m$, we can assume the following:

$(*)$ *There exist $\rho_1 \geq \rho_2 \geq \dots \rho_m > 0$, such that for any $(x,t) \in B_r^n \times B_r^m$, and for $i = 1, \dots, m$: $\lambda_i(D_t f_{(x,t)}) \geq \rho_i > 0$.*

Now for $\Lambda = (\lambda_1, \dots, \lambda_m)$, with $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq 0$, and $\delta > 0$, let $\Sigma(f, \Lambda, \delta)$ be the set defined as follows:

$$\Sigma(f, \Lambda, \delta) =$$

$$\{(x,t) \in B_r^n \times B_r^m \ ; \ \lambda_i(D_x f_{(x,t)}) \leqslant \lambda_i, \forall i = 1, \dots, m \text{ and } \|f(x,t)\| \leqslant \delta\}.$$

We denote also by $\Delta(f, \Lambda, \delta) \subset B_r^n$ the set $\pi_2(\Sigma(f, \Lambda, \delta))$, where the projection $\pi_2 : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ is the standard projection.

Thus $\Delta(f, \Lambda, \delta)$ consists of those parameters $t \in B_r^m$ for which there exists $x \in B_r^n$ with $\lambda_i(D_x f_{(x,t)}) \leqslant \lambda_i$, $i = 1, \dots, m$, and $f_t(x) \in B_\delta^m \subset \mathbb{R}^m$.

Theorem 8.1. *With the notations above, we have, for $s = 0, 1, \dots, m$:*

$$V_s(\Delta(f, \Lambda, \delta)) \leqslant \frac{c.M}{\rho_m \cdots \rho_{m-s+1}} \sum_{j=0}^{s} \lambda_0 \dots \lambda_j r^j \delta^{s-j},$$

with c depending only on $n, m, d_1, \dots, d_m$, and M bounding above $w_j(Df)$, on $B_r^n \times B_r^m$, for $j = 0, \dots, s$.

Proof. As usual, it is enough to bound the s-volume of the projections of $\Delta(f, \Lambda, \delta)$ on various s-dimensional subspaces $P \in G_m^s$, because

$$V_s(\Delta(f, \Lambda, \delta)) = c(m,s) \int_{P \in G_m^s} \int_{x \in P} V_0(\Delta(f, \Lambda, \delta) \cap P) \ dx \ dP$$

$$\leqslant c(m,s).B_{0,m-s}(\Delta(f, \Lambda, \delta)) \int_{P \in G_m^s} Vol_s(\pi_P(\Delta(f, \Lambda, \delta))) \ dP.$$

Let $P \in G_m^s$ be fixed. By Theorem 4.10 there exists a semialgebraic set $C \subset \Sigma(f, \Lambda, \delta)$, such that $D(C)$ depends only on $n, m, d_1, \dots, d_m$, $\pi_P(\Delta(f, \Lambda, \delta)) = (\pi_P \circ \pi_2)(\Sigma(f, \Lambda, \delta)) = (\pi_P \circ \pi_2)(C)$, $\dim(C) = \dim(\pi_P(\Delta(f, \Lambda, \delta))) = s$ (of course we consider only variations $V_s(\Delta(f, \Lambda, \delta))$ for $s \leqslant \dim(\Delta(f, \Lambda, \delta))$).

[1] Here "nondegenerate" means "volume-nondegenerate", i.e. at least the maximal m-volume of the image by $Df_{t\ (x,t)}$ of a unit cube of $\mathbb{R}^n$ is not too small, say bigger than λ, and "typical" means that the bad parameters t lie within a set of small variations depending on λ.

Now by the area formula (see Chapter 7) we have:

$$\begin{aligned} Vol_s((\pi_P \circ \pi_2)(\Sigma(f,\Lambda,\delta))) &\leqslant \int_{x\in C^0} w_s((\pi_P \circ \pi_2)_{|\mathrm{T}_x C^0})\ d\mathcal{H}^s(x) \\ &\leqslant \int_{x\in C^0} w_s(\pi_{2|\mathrm{T}_x C^0})\ d\mathcal{H}^s(x). \end{aligned}$$

Thus let us bound $w_s(\pi_{2|\mathrm{T}_x C^0})$.

For this, we apply Lemma 6.3, with $\mathrm{T} = \mathrm{T}_x C^0$, $\mathrm{L}' = D_x f_{(x,t)}$, $\mathrm{L}'' = D_t f_{(x,t)}$ and $\widetilde{\mathrm{L}} = D(f_{|C^0})_{(x,t)} = (Df_{(x,t)})_{|\mathrm{T}_x C^0}$. We have:

$$\begin{aligned} w_s(\pi_{2|\mathrm{T}_x C^0}) &\leqslant k.w_s([D_t f_{(x,t)}]^{-1}) \sum_{j=0}^{s} w_j(D(f_{|C^0})_{(x,t)}).w_{s-j}(D_x f_{(x,t)}) \\ &\leqslant \frac{k}{\rho_m \dots \rho_{m-s+1}} \sum_{j=0}^{s} w_j(D(f_{|C^0})_{(x,t)}).\lambda_0 \dots \lambda_{s-j}, \end{aligned}$$

by assumption on $D_t f$ and since $(x,t) \in \Sigma(f,\Lambda,\delta)$. The constant k, defined in lemma 6.3, depends only on combinatorial data.

It follows that:

$$Vol_s((\pi_P \circ \pi_2)(\Sigma(f,\Lambda,\delta))) \leqslant$$

$$\frac{k}{\rho_m \dots \rho_{m-s+1}} \sum_{j=0}^{s} \lambda_0 \dots \lambda_{s-j} \int_{x\in C^0} w_j(D(f_{|C^0})_{(x,t)})\ d\mathcal{H}^s(x)$$

$$\leqslant \frac{k}{\rho_m \dots \rho_{m-s+1}} \sum_{j=0}^{s} \lambda_0 \dots \lambda_{s-j}.C_3.M.V_{s-j}(C).V_j(f(C)),$$

the last inequality being a consequence of Corollary 7.14, and M being a constant such that $w_j(D(f_{|C^0})_{(x,t)}) \leqslant M$, for all $j \in \{0,\dots,s\}$.

Finally we have: $C \subset B_r^n \times B_r^m$ and $f(C) \subset B_\Delta^m$, thus by Theorem 5.1 we have $V_{s-j}(C) \leqslant c.r^{s-j}$ and $V_j(f(C)) \leqslant c'.\delta^j$, and thus:

$$Vol_s((\pi_P \circ \pi_2)(\Sigma(f,\Lambda,\delta))) \leqslant \frac{k.C_3.M.c.c'}{\rho_m \dots \rho_{m-s+1}} \sum_{j=0}^{s} \lambda_0 \dots \lambda_{s-j} r^{s-j} \delta^j.$$

□

As we can expect, the bound of Theorem 8.1 shows that for λ_i and δ small, variations of the set of "bad" parameters are small.

We can consider now some special cases of this theorem.

Corollary 8.2. *With the same notations as above,*

$$V_s(\Delta(f,\Lambda,0)) \leqslant \frac{c.M}{\rho_m \dots \rho_{m-s+1}} \lambda_0 \dots \lambda_s r^s.$$

□

This formula is very similar to the bound given in Corollary 7.4. In fact this Corollary is a special case of Theorem 8.1. Indeed, for a given polynomial mapping $f : \mathbb{R}^n \to \mathbb{R}^m$, let us consider the mapping $F : B_r^n \times B_r^m \to \mathbb{R}^m$ defined by $F(x,t) = f(x)+t$. We have $\rho_1 = \ldots = \rho_m = 1$, and from Theorem 8.1 we obtain:

$$V_s(\Delta(F,\Lambda,0)) \leqslant c.M.\lambda_0 \ldots \lambda_s r^s.$$

But $\Delta(F,\Lambda,0)$ is the set of values $f(x) = t$ such that $x \in B_r^n$, $\lambda_i(Df_{(x)}) \leqslant \lambda_i$, for $i = 0,\ldots,m$, and thus is $\Delta(f,\Lambda,B_r^n)$, in the notation of Corollary 7.4.

Let us return to the general case. If $t \notin \Delta(f,\Lambda,\delta)$, then for any $x \in B_r^n$ such that $||f(x,t)|| \leqslant \delta$, at least one of the quantities $\lambda_i(D_x f_{(x,t)})$ is greater than λ_i.

If we want to guarantee that for each i, $\lambda_i(D_x f_{(x,t)}) \geq \lambda_i$, we can proceed as follows:

Assume that everywhere on $B_r^n \times B_r^m$, we have:

$$\lambda_m(D_x f) \leqslant \ldots \leqslant \lambda_2(D_x f) \leqslant \lambda_1(D_x f) \leqslant K.$$

Then for $\Lambda_i = (K,\ldots,K,\lambda_i,\lambda_i,\ldots,\lambda_i)$ (λ_i being located at the i^{th} place), we have $t \notin \Delta(f,\Lambda_i,\delta)$ and $||f(x,t)|| \leqslant \delta$ implies for all $x \in B_r^n$: $\lambda_i(D_x f_{(x,t)}) \geq \lambda_i$.

Hence if we denote by $\bar{\Delta}(f,\Lambda,\delta)$ the union $\bigcup_{i=1}^{m} \Delta(f,\Lambda_i,\delta)$, we have: $t \notin \bar{\Delta}(f,\Lambda,\delta)$ and $||f(x,t)|| \leqslant \delta$ imply for all $x \in B_r^n$, and for all $i \in \{1,\ldots,m\}$: $\lambda_i(D_x f_{(x,t)}) \geq \lambda_i$.

But by Theorem 8.1, we have the following:

$$V_s(\Delta(f,\Lambda_i,\delta)) \leqslant \frac{c.M}{\rho_m \cdots \rho_{m-s+1}} [\sum_{j=0}^{i-1} K^j r^j \delta^{s-j} + K^{i-1} \sum_{j=i}^{s} \lambda_i^{j-i+1} r^j \delta^{s-j}].$$

Adding all these expressions, for i from 1 to m, we get a bound for $\bar{\Delta}(f,\Lambda,\delta)$, as follows:

$$V_s(\bar{\Delta}(f,\Lambda,\delta)) \leqslant C_s^1 \cdot (\lambda_1 + \ldots + \lambda_s) + C_s^2 \cdot \delta.$$

Thus for $\lambda_1,\ldots,\lambda_m$ and δ tending to 0 all the variations of the set $\bar{\Delta}(f,\Lambda,\delta)$ tend to 0. Consequently, for most of values of the parameters t, if $||f(x,t)|| \leqslant \delta$, then for each i, $\lambda_i(D_x f_{(x,t)}) \geq \lambda_i$.

Now we consider another generalization of Theorem 8.1. Let $A \subset \mathbb{R}^m$ and define $\Sigma(f,\Lambda,A,\delta)$ as the set:

$$\Sigma(f,\Lambda,A,\delta) =$$

$$\{(x,t) \in B_r^n \times B_r^m;\ \lambda_i(Df_{t\ (x)}) \leqslant \lambda_i, i = 1,\ldots,m, \text{ and } f_t(x) \in A_\delta\},$$

and

$$\Delta(f, \Lambda, A, \delta) = \pi_2\big(\Sigma(f, \Lambda, A, \delta)\big).$$

The set $\Delta(f, \Lambda, A, \delta)$ is the set of those parameters $t \in B_r^m$ for which there exists $x \in B_r^n$ with $\lambda_i(Df_{t(x)}) \leqslant \lambda_i$ for all $i \in \{1, \dots, m\}$, and there exists $a \in A$ with $||f(x,t) - a|| \leqslant \delta$.

Corollary 8.3. *With the notations above, we have, for all $s \in \{1, \dots, m\}$:*

$$V_s\big(\Delta(f, \Lambda, A, \delta)\big) \leqslant \frac{\widetilde{c}.M.M(\delta, A)}{\rho_m \cdots \rho_{m-s+1}} \sum_{j=0}^{s} \lambda_0 \dots \lambda_j r^j.\delta^{s-j},$$

where $\widetilde{c}$ depends only on $n, m, d_1, \dots, d_m$, and M bounds $w_j(Df)$ on $B_r^n \times B_r^m$ for $j = 0, \dots, s$.

Proof. We cover the δ-neighbourhood A_δ of A by $M(\delta, A)$ balls of radius 2δ, and to each of them we apply Theorem 8.1. □

By Theorem 3.5, we obtain:

Corollary 8.4. *With the notations above, we have:*

$$M\big(\epsilon, \Delta(f, \Lambda, A, \delta)\big) \leqslant$$

$$C.\widetilde{M}.M(\delta, A) \sum_{s=0}^{m} \frac{1}{\rho_m \cdots \rho_{m-s+1}} (\frac{1}{\epsilon})^s \sum_{j=0}^{s} \lambda_0 \dots \lambda_j r^j.\delta^{s-j},$$

where C depends only on $n, m, d_1, \dots, d_m$, and $\widetilde{M}$ bounds $w_j(Df)$ on $B_r^n \times B_r^m$ for $j = 0, \dots, m$. □

Substituting here $\epsilon = \delta$, we get:

Corollary 8.5. *With the notations above:*

$$M\big(\epsilon, \Delta(f, \Lambda, A, \epsilon)\big) \leqslant C \cdot \widetilde{C} \cdot \widetilde{M} \cdot M(\epsilon, A) \sum_{j=0}^{m} \lambda_0 \dots \lambda_j (\frac{r}{\epsilon})^j,$$

where C depends only on $n, m, d_1, \dots, d_m$, $\widetilde{M}$ bounds from above $w_j(Df)$ on $B_r^n \times B_r^m$ for $j = 0, \dots, m$, and $\widetilde{C}$ depends only on $\rho_1, \dots, \rho_m$. □

Assume now that $\lambda_j = 0$, for $j \geq p$, then:

$$M(\epsilon, \Delta(f, \Lambda, A, \epsilon)) \sim \widetilde{c}.\widetilde{M}.M(\epsilon, A)(\frac{1}{\epsilon})^p.$$

Hence we obtain the following:

Corollary 8.6. *With the notations above, and for $\Lambda = (\lambda_1, \dots, \lambda_p, 0, \dots, 0)$,*

$$\dim_e(\Delta(f, \Lambda, A, 0)) \leqslant \dim_e(A) + p.$$

Proof. We have $\Delta(f,\Lambda,A,0) \subset \Delta(f,\Lambda,A,\epsilon)$, thus there exists a constant $\hat{c}$ such that for all $\epsilon > 0$, $M(\epsilon,\Delta(f,\Lambda,A,0)) \leqslant \hat{c}.M(\epsilon,A)(\frac{1}{\epsilon})^p$. It follows that $M(\epsilon,\Delta(f,\Lambda,A,0)).\epsilon^{\beta+p}$ is bounded for all $\epsilon > 0$, and for all $\beta > \dim_e(A)$. Finally $\dim_e(\Delta(f,\Lambda,A,0)) \leqslant \dim_e(A)+p$. □

In particular, for $A \subset \mathbb{R}^m$such that $\dim_e(A) < m-p$, we have: $\dim_e(\Delta(f,\Lambda,A,0)) < m$, hence for a generic $t \in B_r^m$, $f_t(\Sigma(f,\Lambda,r)) \cap A = \emptyset$. In other words, for a generic t, if $f(x,t) \in A$, then the rank of $D_x f_{(x,t)}$ is greater than p.

Our next goal is to give a relative version of Corollary 7.4.

Let $f : B_r^n \to \mathbb{R}^m$ and $g : B_r^n \to \mathbb{R}^q$, $q \leqslant n$, be two polynomial mappings.

Let $\delta \geq 0$ be fixed. We assume that at each point $x \in g^{-1}(B_\delta^q)$, the mapping g is nondegenerate, for instance, we assume that:

$$\lambda_i(Dg_{(x)}) \geq \rho_i > 0, \ i = 0,1,\ldots,q, \ (\rho_0 = 1).$$

For $\Lambda\Lambda' = (\lambda_1,\ldots,\lambda_q,\lambda'_1,\ldots,\lambda'_{\min(n-q,m)})$, $\lambda_1 \geq \ldots \geq \lambda_q \geq 0$, $\lambda'_1 \geq \ldots \geq \lambda'_{\min(n-q,m)} \geq 0$, $(\lambda_0 = \lambda'_0 = 1)$, we denote by $\Sigma(f,g,\Lambda\Lambda',\delta)$ the set:

$$\Sigma(f,g,\Lambda\Lambda',\delta) = \{x \in B_r^n;\ \lambda_i(Df_{(x)|(\ker(Dg_{(x)})^\perp}) \leqslant \lambda_i,\ i = 1,\ldots,\lambda_q,$$

$$\lambda_j(Df_{(x)|\ker Dg_{(x)}}) \leqslant \lambda'_j,\ j = 1,\ldots,\min(n-q,m),\ \text{and}\ ||g(x)|| \leqslant \delta\}.$$

As usual, $\Delta(f,g,\Lambda\Lambda',\delta)$ denotes the set $f(\Sigma(f,g,\Lambda\Lambda',\delta))$.

Theorem 8.7. *With the notations above, we have, for $s = 0,1,\ldots,m$.*[2]*:*

$$V_s(\Delta(f,g,\Lambda\Lambda',\delta)) \leqslant K.\hat{C}\sum_{j=0}^{s} \frac{\lambda_0\lambda_1\ldots\lambda_{s-j}}{\rho_0\rho_1\ldots\rho_{s-j}}\lambda'_0\lambda'_1\ldots\lambda'_j.r^j.\delta^{s-j},$$

where K bounds from above $\sigma_{s-j}(Dg_{(x)|\mathrm{T}})$, for all $x \in \Sigma(f,g,\Lambda\Lambda',\delta)$, all $T \in G_n^s$ and all $j \in \{0,\ldots,s\}$ (σ_{s-j} being defined in Theorem 7.7.)

Proof. As in the proof of Theorem 8.1, it suffices to bound the s-volume of the projections $\pi_P(\Delta(f,g,\Lambda\Lambda',\delta)) \subset P$, where P is a s-dimensional linear subspace of $\mathbb{R}^m$. Let P be a fixed plane in G_m^s. By Theorem 4.10 there exists a semialgebraic set $C \subset \Sigma(f,g,\Lambda\Lambda',\delta)$, such that $D(C)$ depends only on $D(\Sigma(f,g,\Lambda\Lambda',\delta))$,

$$[\pi_P \circ f](C) = [\pi_P \circ f](\Sigma(f,g,\Lambda\Lambda',\delta)) = \pi_P(\Delta(f,g,\Lambda\Lambda',\delta)),$$

and

$$\dim(C) = \dim(\pi_P(\Delta(f,g,\Lambda\Lambda',\delta))) = s$$

[2] The convention here is that for $j>q$, $\lambda_j = 0$, for $j>min(n\text{-}q,m)$, $\lambda'_j = 0$, and for $j> q$, $\rho_j = 1$.

(we consider only variations $V_s(\Delta(f,g,\Lambda\Lambda',\delta))$ for $s \leqslant \dim(\Delta(f,g,\Lambda\Lambda',\delta))$.)

By the area formula, we have:

$$Vol_s(\pi_P(\Delta(f,g,\Lambda\Lambda',\delta))) \leqslant \int_{x\in C^0} w_s([\pi_P \circ Df_{(x)}]_{|\mathrm{T}_x C^0})\ d\mathcal{H}^s(x)$$

$$\leqslant \int_{x\in C^0} w_s(Df_{(x)|\mathrm{T}_x C^0})\ d\mathcal{H}^s(x).$$

It remains to estimate $w_s(Df_{(x)|\mathrm{T}_x C^0})$. By Lemma 6.2, with $\mathrm{T} = \mathrm{T}_x C^0$, $\pi_1 : \mathrm{T} \to \ker(Dg_{(x)})$, $\pi_2 : \mathrm{T} \to (\ker(Dg_{(x)}))^\perp$, $\mathrm{L}_1 = Df_{(x)|\ker(Dg_{(x)})} \circ \pi_1$, $\mathrm{L}_2 = Df_{(x)|(\ker(Dg_{(x)}))^\perp} \circ \pi_2$, $\mathrm{L}_1 + \mathrm{L}_2 = Df_{(x)|\mathrm{T}}$, we get:

$$w_s(Df_{(x)|\mathrm{T}_x C^0}) \leqslant k(s,m,s) \sum_{j=0}^{s} w_j(\mathrm{L}_1).w_{s-j}(\mathrm{L}_2).$$

Now by assumption, $w_j(\mathrm{L}_1) \leqslant \lambda'_0\lambda'_1 \ldots \lambda'_j$, for all $j \in \{0,\ldots,\min(n-q,m)\}$, and $w_{s-j}(\mathrm{L}_2) \leqslant \lambda_0\lambda_1\ldots\lambda_{s-j} \leqslant \dfrac{w_{s-j}(Dg_{(x)})}{\rho_0\rho_1\ldots\rho_{s-j}}\lambda_0\lambda_1\ldots\lambda_{s-j}$, for all $j \in \{0,\ldots,q\}$. We thus have:

$$w_s(Df_{(x)|\mathrm{T}_x C^0}) \leqslant k(s,m,s) \sum_{j=0}^{s} \frac{\lambda_0\lambda_1\ldots\lambda_{s-j}}{\rho_0\rho_1\ldots\rho_{s-j}} \lambda'_0\lambda'_1\ldots\lambda'_j w_{s-j}(Dg_{(x)})$$

$$Vol_s(\pi_P(\Delta(f,g,\Lambda\Lambda',\delta))) \leqslant \int_{x\in C^0} w_s(Df_{(x)|\mathrm{T}_x C^0})\ d\mathcal{H}^s(x)$$

$$\leqslant k(s,m,s) \sum_{j=0}^{s} \frac{\lambda_0\lambda_1\ldots\lambda_{s-j}}{\rho_0\rho_1\ldots\rho_{s-j}} \lambda'_0\lambda'_1\ldots\lambda'_j \int_{x\in C^0} w_{s-j}(Dg_{(x)})\ d\mathcal{H}^s(x)$$

$$\leqslant \sigma_{s-j}(g_{|C^0}).C.k(s,m,s) \sum_{j=0}^{s} \frac{\lambda_0\lambda_1\ldots\lambda_{s-j}}{\rho_0\rho_1\ldots\rho_{s-j}} \times$$

$$\lambda'_0\lambda'_1\ldots\lambda'_j V_{s-j}(g(C^0).V_{s-(s-j)}(C^0)$$

The last inequality being a consequence of Theorem 7.7. By assumption, $\sigma_{s-j}(g_{|C^0}) \leqslant K$, $C \subset B_r^n$, $g(C) \subset B_\delta^q$, hence by Theorem 5.1:

$$Vol_s(\pi_P(\Delta(f,g,\Lambda\Lambda',\delta))) \leqslant$$

$$K.C.k(s,m,s) \sum_{j=0}^{s} \frac{\lambda_0\lambda_1\ldots\lambda_{s-j}}{\rho_0\rho_1\ldots\rho_{s-j}} \lambda'_0\lambda'_1\ldots\lambda'_j.r^j.\delta^{s-j}.$$

□

First of all, we see that Corollary 7.4 is a special case of Theorem 8.7. Indeed, for $g : B_r^n \to B_r^n$ being the identity mapping, $\delta = r$, we have $\ker(Dg_{(x)}) = 0$, $\lambda'_0 = \lambda'_{\min(n-n,m)} = 1$, $\rho_0 = \ldots = \rho_n = 1$,$K = 1$, and $\Delta(f,g,\Lambda\Lambda',r) = \Delta(f,\Lambda,B_r^n)$, hence the only term remains:

$$V_s(\Delta(f, \Lambda, B^n_r)) \leqslant \hat{C}.\lambda_0\lambda_1 \dots \lambda_s.r^s.$$

In the next corollary we assume f to be the identity, $\lambda_0 = \dots = \lambda_q = 1$, $\lambda'_0 = \dots = \lambda'_{\min(n-q,m)} = 1$. Hence $\Delta(f, g, \Lambda\Lambda', \delta)$ is just the tube $\mathcal{T}_\delta = g^{-1}(B^q_\delta)$ around the fiber $\mathrm{Y}_\delta = g^{-1}(\{0\})$. In this case, considering the convention (See Footnote 2 on Page 104) of Theorem 8.7, we see that to have the term $\dfrac{\lambda_0\lambda_1 \dots \lambda_{s-j}}{\rho_0\rho_1 \dots \rho_{s-j}}\lambda'_0\lambda'_1 \dots \lambda'_j.r^j.\delta^{s-j} \neq 0$, we must have $s - j \leqslant q$ and $j \leqslant \min(n - q, n) = n - q$. Thus we obtain:

Corollary 8.8. *With the above notations, for all $s \in \{0, \dots, m\}$:*

$$V_s(\mathcal{T}_\delta) \leqslant K.\hat{C} \sum_{j=s-q}^{n-q} \frac{1}{\rho_0\rho_1 \dots \rho_{s-j}} r^j.\delta^{s-j}.$$

□

In particular, for the volume of $\mathcal{T}_\delta$ we have the following bound:

Corollary 8.9. *With the above notations:*

$$Vol_n(\mathcal{T}_\delta) \leqslant K.\hat{C}.\frac{1}{\rho_0\rho_1 \dots \rho_q} r^{n-q}.\delta^q.$$

□

Theorem 8.1 also can be obtained as a special case of Theorem 8.7.

The main application of Theorem 8.7 we give here concerns the "high order critical values", i.e. the values of the mapping on the "near-Thom-Boardman"singularities (see [Boa] or [Gol-Gui] for basic definitions). Although by our methods the general Thom-Boardman singularities can be treated, the expressions are very complicated. Thus we consider here only the simplest situation: let $f : B^n_r \to \mathbb{R}^n$ be a polynomial mapping, let J denote the Jacobian of f, and let $\delta \geq 0$ be given.

If we assume that for any x with $|J_{(x)}| \leqslant \delta$, $||\nabla J_{(x)}|| \geq \gamma > 0$, the set $J^{-1}(\{x\})$, $x \in B^1_\delta$, is a submanifold of dimension $n-1$ of $\mathbb{R}^n$. In particular we can restrict f to $\Sigma^1(f) = J^{-1}(\{0\})$, and consider the singularities of $f_{|\Sigma^1(f)} : \Sigma^1(f) \to \mathbb{R}^n$. The set of points x of $\Sigma^1(f)$ such that the map $f_{|\Sigma^1(f)}$ drops rank $s > 0$ (i.e. $\mathrm{rank}(Df_{(x)|\Sigma^1(f)}) \leqslant n - 1 - s$) is denoted $\Sigma^{1,s}(f)$ (one can of course define in the same way the sets $\Sigma^{r,s}(f)$, and so $\Sigma^{r,s,t,\dots}(f)$). The points of the $\Sigma^{r,s,t,\dots}(f)$'s are called Thom-Boardman singularities of f (see [Boa], or [Gol-Gui]). Now for $\Lambda = (\lambda_1, \dots, \lambda_{n-1})$, we denote by $\Sigma^1(f, \Lambda, \delta)$ the following set:

$$\Sigma^1(f, \Lambda, \delta) =$$

$$\{x \in B^n_r;\ |J_{(x)}| \leqslant \delta \text{ and } \lambda_i(Df_{(x)|\ker(DJ_{(x)})}) \leqslant \lambda_i, i = 1, \dots, \lambda_{n-1}\}.$$

Clearly, for λ_i and δ small (which is the case we usually have in mind), $\Sigma^1(f, \Lambda, \delta)$ is the set of points which are near Thom-Boardman singularities of type $\Sigma^{1,s}(f)$, for all $s \in \{1, \ldots, n-1\}$. For instance, for $\delta = 0$ and $\lambda_{n-s} = \ldots = \lambda_{n-1} = 0$, we obtain $\Sigma^{1,s}$.

As usual we denote by $\Delta^1(f, \Lambda, \delta)$ the image $f(\Sigma^1(f, \Lambda, \delta))$.

Theorem 8.10. *For $s \leqslant n-1$, we have (with the convention (See Footnote 2 on Page 104):*

$$V_s(\Delta^1(f,\Lambda,\delta)) \leqslant \hat{C}.K[\frac{M}{\gamma}.\lambda_0 \ldots \lambda_{s-1}.r^{s-1}.\delta + \lambda_0 \ldots \lambda_s.r^s],$$

$$\text{and for } s = n: \quad V_n(\Delta^1(f,\Lambda,\delta)) \leqslant \frac{\hat{C}.K}{\gamma}.\lambda_0 \ldots \lambda_{n-1}.r^{n-1}.\delta,$$

where K bounds above $\sigma_{s-j}(DJ_{(x)|\ \mathrm{T}})$, for all $x \in \Sigma^1(f, \Lambda, \delta)$, all $T \in G_n^s$, and all $j \in \{s-1, s\}$, and where M is an upper bound for

$$\lambda_1(Df_{(x)|(\ker(DJ_{(x)}))^{\perp}}),$$

for all $x \in \Sigma^1(f, \Lambda, \delta)$.

Proof. We remark that $\Sigma^1(f, \Lambda, \delta) = \Sigma(f, J, \Lambda\Lambda', \delta)$ (with the notations of Theorem 8.7), when $\Lambda\Lambda' = (M, \lambda_1, \ldots, \lambda_{n-1})$. We thus can apply Theorem 8.7 with $g = J$, $\lambda_1 = M$, $\lambda'_j = \lambda_j$ for $j = 1, \ldots, n-1$, and $\rho_1 = \gamma$. □

In particular, for $n = 2$, we have $\Lambda = \lambda$, and thus:

Corollary 8.11. *With the notations above:*

$$\begin{aligned} V_0(\Delta^1(f,\Lambda,\delta)) &\leqslant C_0 \\ V_1(\Delta^1(f,\Lambda,\delta)) &\leqslant C_1 \cdot K(\lambda \cdot r + \frac{M}{\gamma}\delta) \\ V_2(\Delta^1(f,\Lambda,\delta)) &\leqslant C_2 \cdot K\frac{M}{\gamma}\lambda \cdot r \cdot \delta. \end{aligned}$$

□

Thus for λ and δ small, the area V_2 of the "near-cuspidal" values $\Delta^1(f, \Lambda, \delta)$ is of order $\lambda \cdot \delta$, and not of order λ, as one expects for generic rank 1 near-critical values of f.

9 Mappings of Finite Smoothness

Abstract. We prove the quantitative Morse-Sard theorem for $\mathcal{C}^k$ mappings with n variables, i.e. we bound the ϵ-entropy of near-critical values. In particular, we give, for the entropy dimension of the rank-ν set of critical values, a bound depending only on n, ν and k. We then give examples showing that our statement is the best possible. We also give the $\mathcal{C}^k$ version of the polynomial quantitative transversality of Chapter 8.

In this chapter we investigate the properties of the set of Λ-critical values of a given $\mathcal{C}^k$-mapping, $k \geq 1$, $f : B_r^n \to \mathbb{R}^m$ [1].

In what follows, k is not necessarily an integer. Let us write, for $k > 1$, $k = p + \alpha$, with $p \in \mathbb{N} \setminus \{0\}$ and $\alpha \in]0;1]$; we say that f is a $\mathcal{C}^k = \mathcal{C}^{p+\alpha}$-mapping (of Hölder smoothness class $\mathcal{C}^k$) when f is p-times differentiable and there exists a constant $K > 0$ such that for every $x, y \in B_r^n$, $||D^p f_{(x)} - D^p f_{(y)}|| \leqslant K||x - y||^\alpha$ [2]. A $\mathcal{C}^1$-mapping is just a differentiable mapping with continuous derivative $x \to Df_{(x)}$.

Let us notice that a $\mathcal{C}^k$-mapping, with $k \in \mathbb{N} \setminus \{0, 1\}$, is a $(k-1)$-times differentiable mapping such that: $x \mapsto D^{k-1} f_{(x)}$ is Lipschitz.

We denote by $R_k(f)$ the following quantity:

$$R_k(f) = \frac{K}{p!} . r^{k=p+\alpha}.$$

For $\Lambda = (\lambda_1, \ldots, \lambda_q)$, $q = \min(n, m)$, the set $\Sigma(f, \Lambda, B_r^n)$ of Λ-critical points and the set $\Delta(f, \Lambda, B_r^n)$ of Λ-critical values are defined as above:

$$\Sigma(f, \Lambda, B_r^n) = \{x \in B_r^n;\ \lambda_i(Df_{(x)}) \leqslant \lambda_i,\ i = 1, \ldots, q\},$$

$$\Delta(f, \Lambda, B_r^n) = f(\Sigma(f, \Lambda, B_r^n)).$$

Proposition 9.1. *With the notations above, for any $\epsilon \geq R_k(f)$, we have:*

[1] As usual, a $\mathcal{C}^k$-mapping on B_r^n is a mapping which can be extended to a $\mathcal{C}^k$-mapping on some open neighborhood of $B_r^n \subset \mathbb{R}^n$.

[2] The norm of the p-th differential is of course the Euclidean norm on the corresponding space of multilinear mappings. As an easy exercice, one can check for instance that $x \to x^k$, for k> 1 is a $\mathcal{C}^k$-smooth mapping and that $x \to x^k \sin(1/x)$, for k> 2 is a $\mathcal{C}^{k/2}$-smooth mapping.

Y. Yomdin and G. Comte: LNM 1834, pp. 109–130, 2004.

$$M(\epsilon, \Delta(f, \Lambda, B_r^n)) \leqslant c \sum_{i=0}^{i=q} \lambda_0 \lambda_1 \dots \lambda_i (\frac{r}{\epsilon})^i,$$

where the constant c only depends on n, m and k.

Proof. Let P_p be the Taylor polynomial of f of degree p at $0 \in B_r^n$. We write $\varphi(x) = f(x) - P_p(x) = f(x) - \sum_{j=0}^{p} \frac{1}{j!} D^j f_{(0)}.x^j$. For all $j \in \{1, \dots, p\}$, we have $D^j \varphi_{(0)} = 0$, and $D^p \varphi_{(x)} = D^p f_{(x)} - D^p f_{(0)}$.

We have by the Taylor formula, for any $x \in B_r^n$:

$$\varphi(x) = f(x) - P_p(x) = \sum_{j=0}^{p-1} \frac{1}{j!} D^j \varphi_{(0)} x^j + \int_{t=0}^{t=1} \frac{(1-t)^{p-1}}{(p-1)!} D^p \varphi_{(tx)}.x^p \, dt.$$

Thus:

$$\begin{aligned} ||f(x) - P_p(x)|| &= || \int_{t=0}^{t=1} \frac{(1-t)^{p-1}}{(p-1)!} D^p \varphi_{(tx)}.x^p \, dt|| \\ &\leqslant \int_{t=0}^{t=1} \frac{(1-t)^{p-1}}{(p-1)!} ||D^p \varphi_{(tx)}||.||x||^p \, dt \\ &= \int_{t=0}^{t=1} \frac{(1-t)^{p-1}}{(p-1)!} ||D^p f_{(tx)} - D^p f_{(0)}||.||x||^p \, dt \\ &\leqslant \int_{t=0}^{t=1} \frac{(1-t)^{p-1}}{(p-1)!} .K.||tx - 0||^\alpha.||x||^p \, dt, \end{aligned}$$

and finally, we obtain:

$$||f(x) - P_p(x)|| \leqslant \frac{K}{p!} r^k = R_k(f) \leqslant \epsilon. \qquad \mathbf{(1)}$$

The same argument for $D\varphi$ gives us:

$$||Df_{(x)} - DP_{p(x)}|| \leqslant \frac{K}{(p-1)!} r^{k-1} = \frac{p}{r} R_k(f) \leqslant \frac{p}{r} \epsilon. \qquad \mathbf{(2)}$$

Hence if we denote by λ'_i the numbers $\lambda_i + \frac{p\epsilon}{r}$, for $i \geq 1$ and $\lambda'_0 = 1$, and put Λ' be $(\lambda'_1, \dots, \lambda'_q)$, by inequality **(2)**, we have $\Sigma(f, \Lambda, B_r^n) \subset \Sigma(P_p, \Lambda', B_r^n)$. Now $\Delta(f, \Lambda, B_r^n) \subset f\big(\Sigma(P_p, \Lambda', B_r^n)\big) \subset \big[\Delta(P_p, \Lambda', B_r^n)\big]_\epsilon$, by **(1)**.

Therefore we can write:

$$M(2\epsilon, \Delta(f, \Lambda, B_r^n)) \leqslant M(2\epsilon, \big[\Delta(P_p, \Lambda', B_r^n)\big]_\epsilon) \leqslant M(\epsilon, \Delta(P_p, \Lambda', B_r^n)),$$

and by Theorem 7.5:

$$M(2\epsilon, \Delta(f, \Lambda, B_r^n)) \leqslant C \sum_{i=0}^{q} \lambda'_0 \dots \lambda'_i (\frac{r}{\epsilon})^i,$$

or equivalently:

$$M(\epsilon, \Delta(f, \Lambda, B_r^n)) \leqslant \widetilde{C} \sum_{i=0}^{q} \lambda_0' \dots \lambda_i' (\frac{r}{\epsilon})^i,$$

with $\widetilde{C}$ depending only on p, n and m.

But for any $i \in \{1, \dots, q\}$,

$$\lambda_0' \dots \lambda_i' = \prod_{j=1}^{i} (\lambda_j + \frac{p\epsilon}{r}) = c' \sum_{\ell=0}^{i} \sum_{i_1 > \dots > i_\ell \in \{1,\dots,i\}} \lambda_{i_1} \dots \lambda_{i_\ell} (\frac{\epsilon}{r})^{i-\ell}.$$

We have $\lambda_{i_1} \dots \lambda_{i_\ell} \leqslant \lambda_1 \dots \lambda_\ell$, because $\lambda_1 \geq \dots \geq l_\ell$, and thus:

$$\lambda_0' \dots \lambda_i' \leqslant \widetilde{c} \sum_{\ell=0}^{i} \lambda_1 \dots \lambda_\ell (\frac{\epsilon}{r})^{i-\ell}.$$

We conclude that:

$$\begin{aligned} M(\epsilon, \Delta(f, \Lambda, B_r^n)) &\leqslant \widetilde{C}.\widetilde{c} \sum_{i=0}^{q} \sum_{\ell=0}^{i} \lambda_0 \lambda_1 \dots \lambda_\ell (\frac{r}{\epsilon})^{\ell-i} (\frac{r}{\epsilon})^i \\ &\leqslant c \sum_{i=0}^{q} \lambda_0 \dots \lambda_i (\frac{r}{\epsilon})^i. \end{aligned}$$

□

Thus for $\epsilon \geq R_k(f)$, we have for $\mathcal{C}^k$-smooth f exactly the same expression as for polynomials. This fact is one of many effects of "near-polynomiality", which can be described roughly as follows: if we consider a $\mathcal{C}^{(k=p+\alpha)}$-smooth function with accuracy $\epsilon \geq R_k(f)$, that is to say with accuracy "not too small", we cannot distinguish it from its Taylor polynomial of degree p, not only in the $\mathcal{C}^0$-norm, but also in the structure of critical points and values.

The next result is the main result of this chapter, and one of the main results of the book: the Quantitative Morse-Sard Theorem. It bounds for any $\epsilon > 0$ the entropy of the set of near-critical values of f (parametrised by $\lambda_1, \dots, \lambda_q$), in terms of ϵ, λ_i, r and the only data on f - its remainder term $R_k(f)$.

Theorem 9.2. *Let $f : B_r^n \to \mathbb{R}^m$ be a $\mathcal{C}^{k=p+\alpha}$-smooth mapping on B_r^n. Then for $\Lambda = (\lambda_1, \dots, \lambda_{q=\min(n,m)})$ and for $\epsilon > 0$, we have:*

$$\begin{aligned} M(\epsilon, \Delta(f, \Lambda, B_r^n)) &\leqslant c \sum_{i=0}^{q} \lambda_0 \lambda_1 \dots \lambda_i (\frac{r}{\epsilon})^i, \quad \text{for } \epsilon \geq R_k(f), \\ M(\epsilon, \Delta(f, \Lambda, B_r^n)) &\leqslant \widetilde{c} \sum_{i=0}^{q} \lambda_0 \lambda_1 \dots \lambda_i (\frac{r}{\epsilon})^i \left(\frac{R_k(f)}{\epsilon} \right)^{\frac{n-i}{k}}, \quad \text{for } \epsilon \leqslant R_k(f), \end{aligned}$$

where c and $\widetilde{c}$ depend only on n, m and k.

Proof. For $\epsilon \geq R_k(f)$ the result follows from Proposition 9.1.

Let $\epsilon \leqslant R_k(f)$. We cover B_r^n by balls of radius $r' < r$, where r' is chosen in such a way that $R_k(f_{|B_{r'}^n}) \leqslant \epsilon$. We can take for r', $\frac{K}{p!}(r')^k = \epsilon$, or:

$$r' = r.\left(\frac{\epsilon}{R_k(f)}\right)^{\frac{1}{k}}.$$

The number of such balls we need to cover B_r^n is at most:

$$N = C(\frac{r}{r'})^n = C.\left(\frac{R_k(f)}{\epsilon}\right)^{\frac{n}{k}}.$$

We apply Proposition 9.1 to the restriction of f to each of the N balls $B_{(x_j,r')}^n$, hence we obtain:

$$\begin{aligned} M(\epsilon, \Delta(f, \Lambda, B_r^n)) &\leqslant \sum_{j=1}^{N} M(\epsilon, \Delta(f_{|B_{(x_j,r')}^n}, \Lambda, B_{(x_j,r')}^n)) \\ M(\epsilon, \Delta(f, \Lambda, B_r^n)) &\leqslant C.\left(\frac{R_k(f)}{\epsilon}\right)^{\frac{n}{k}}.c\sum_{i=0}^{q} \lambda_0 \ldots \lambda_i(\frac{r'}{\epsilon})^i \\ &= \widetilde{c}\left(\frac{R_k(f)}{\epsilon}\right)^{\frac{n}{k}} \sum_{i=0}^{q} \lambda_0 \ldots \lambda_i(\frac{r}{\epsilon})^i(\frac{r'}{r})^i. \end{aligned}$$

Thus:

$$M(\epsilon, \Delta(f, \Lambda, B_r^n)) \leqslant \widetilde{c}\sum_{i=0}^{q} \lambda_0 \ldots \lambda_i(\frac{r}{\epsilon})^i\left(\frac{R_k(f)}{\epsilon}\right)^{\frac{n-i}{k}}.$$

□

Let us stress the fact that for $\epsilon \geq R_k(f)$ the expression of Theorem 9.2 is the same as in the polynomial case. So in a resolution coarser than the Taylor remainder term one cannot distinguish between the geometry of the critical values of f and of its Taylor polynomial approximation. In finer resolution a correction appears, expressed in terms of the remainder $R_k(f)$. In fact, Theorem 9.2, as well as most of the results of Chapter 9, can be considered as a "generalized Taylor formula" for the property in question: they contain a "polynomial term" and a correction, expressed through R_k.

Let us denote now:

$$\Sigma_f^\nu = \{x \in B_r^n; \text{rank}(Df_{(x)}) \leqslant \nu\}$$

and

$$\Delta_f^\nu = f(\Sigma_f^\nu),$$

respectively the rank-ν set of critical points and of critical values. The closed ball B_r^n being compact, there exist $\lambda_1 \geq \ldots\ldots \lambda_q \geq 0$ such that for all

$x \in B_r^n$ and for all $i \in \{1, \ldots, q\}$, $\lambda_i(Df_{(x)}) \leqslant \lambda_i$. Therefore we have $\Delta_f^\nu \subset \Delta(f, \Lambda, B_r^n))$, for $\Lambda = (\lambda_1, \ldots, \lambda_\nu, 0, \ldots, 0)$, and thus by Theorem 9.2:

$$M(\epsilon, \Delta_f^\nu) \leqslant M(\epsilon, \Delta_f^\nu \subset \Delta(f, \Lambda, B_r^n))) \leqslant \widetilde{c} \sum_{i=0}^{\nu} \lambda_0 \ldots \lambda_i \left(\frac{r}{\epsilon}\right)^i \left(\frac{R_k(f)}{\epsilon}\right)^{\frac{n-i}{k}},$$

so

$$\begin{aligned} M(\epsilon, \Delta_f^\nu) &\leqslant \widetilde{c} \sum_{i=0}^{\nu} \lambda_0 \ldots \lambda_i \left(\frac{r}{\epsilon}\right)^i \left(\frac{K.r^k}{p!\epsilon}\right)^{\frac{n-i}{k}} \\ &= \widetilde{c} \sum_{i=0}^{\nu} \lambda_0 \ldots \lambda_i \left(\frac{K}{p!}\right)^{\frac{n-i}{k}} .r^n. \left(\frac{1}{\epsilon}\right)^{i+\frac{n-i}{k}}. \end{aligned}$$

Finally we have bounded $M(\epsilon, \Delta_f^\nu)$ by a polynomial of degree $\nu + \dfrac{n-\nu}{k}$ in $\left(\dfrac{1}{\epsilon}\right)$, hence by definition the entropy dimension of Δ_f^ν is less than $\nu + \dfrac{n-\nu}{k}$, and so is the Hausdorff dimension of Δ_f^ν (by the inequality $\dim_{\mathcal{H}} \leqslant \dim_e$ of Chapter 2). We have proved:

Theorem 9.3. (Entropy Morse-Sard Theorem [Yom 1]) *Let $f : B_r^n \to \mathbb{R}^m$ be a $\mathcal{C}^{k=p+\alpha}$-mapping on B_r^n. Then:*

$$\dim_{\mathcal{H}}(\Delta_f^\nu) \leqslant \dim_e(\Delta_f^\nu) \leqslant \nu + \frac{n-\nu}{k}.$$

□

Comments.

• First of all, let us stress that the bound on the entropy dimension may be much more restrictive than that for the Hausdorff dimension. Many examples of this sort are given in Chapter 2. In particular, for $\nu = 0$, i.e. for critical values of rank 0, by Theorem 9.3 their entropy dimension is at most n/k. This implies for instance that the sequences $\{1, 1/2^\beta, \ldots, 1/n^\beta, \ldots\}$ cannot be contained among these critical values for $\beta < \dfrac{k}{n} + 1$. Of course, the Hausdorff dimension of any countable set is zero.

• Theorem 3.4.3 of [Fed 2] gives, for f a k-times continuously differentiable mapping ($k \in \mathbb{N} \setminus \{0\}$), $\mathcal{H}^{\nu + \frac{n-\nu}{k}}(\Delta_f^\nu) = 0$. In particular, it implies $\dim_{\mathcal{H}}(\Delta_f^\nu) \leqslant \nu + \dfrac{n-\nu}{k}$. Theorem 9.3 shows that the assumption k-times continuously differentiable can be weakened to $\mathcal{C}^{k-1+1}$, that is to say, f is $(k-1)$-times differentable and $D^{k-1}f$ is Lipschitz.

• In terms of entropy (as well as in terms of Hausdorff) dimension, Theorem 9.3 cannot be sharpened: we can construct a family of functions $f_\beta : [0;1]^2 \to [0;1]$, with $\beta \in]2; \infty[$, such that f_β is $\mathcal{C}^{\frac{2\ln 3}{\ln \beta}}$-smooth, $\Delta_{f_\beta}^0 = I_{\frac{1}{3}}$, the classical Cantor set of $I = [0;1]$ of entropy (and Hausdorff) dimension $\ln 2/\ln 3$ ([Com 1]).

In this case, Theorem 9.3 gives us: $(\ln 2/\ln 3 = \dim_{\mathcal{H}}(\Delta^0_{f_\beta}) =) \dim_e(\Delta^0_{f_\beta}) \leqslant \ln\beta/\ln 3$, and β is as close as we want to 2.

The function f_β is constructed in the following way.

First of all we consider the classical Cantor set C_β on the square $I^2 \subset \mathbb{R}^2$: C^1, C^2, C^3, C^4 are the squares of I^2 with side of length $1/\beta$ and centred respectively at $(1/4;1/4), (3/4;1/4), (3/4;3/4), (1/4;3/4)$. As for the classical Cantor set of I, constructed in Chapter 2, we iterate this construction of squares in order to obtain a sequence of squares: $(C^{i_1,i_2,\dots,i_n})$, with $i_k \in \{1,2,3,4\}$, $k \in \{1,\dots,n\}$. The set C_β is defined as the set consisting of all the points of I^2 of the type $\bigcap\limits_{n\in\mathbb{N}} C^{i_1,i_2,\dots,i_n}$.

Now let us consider $f_0 : I^2 \to \mathbb{R}$ a $\mathcal{C}^\infty$-smooth mapping, such that $f_{0|\partial I^2} \equiv 0$, $D^k f_0(\xi) \longrightarrow 0$ for ξ converging to a point of ∂I^2, $f_{0|C^i} \equiv \dfrac{9}{8} + b_i$, for $i \in \{1,2,3,4\}$ and $b_1 = 0, b_2 = 2/9, b_3 = 2/3, b_4 = 8/9$. The function f_0 allows us to construct a function $f_1 : I^2 \to \mathbb{R}$ such that $f_{0|I^2\setminus\cup_{i=1}^4 C^i} = f_{1|I^2\setminus\cup_{i=1}^4 C^i}$, as follows: if we denote by $\tau_i : C^i \to \mathbb{R}^2$ the translation which centres C^i at $(1/2;1/2)$, $\pi_\beta : I^2 \to \mathbb{R}^2$ defined by $\pi_\beta(\xi) = \beta.\xi$, and $\pi_{1/9} : \mathbb{R} \to \mathbb{R}$ defined by $\pi_{1/9}(t) = t/9$, we set $f_1 = f_0 + \sum\limits_{i=1}^{4} \pi_{1/9} \circ f_0 \circ \pi_\beta \circ \tau_i$.

By induction we obtain in the same way functions f_k, and finally, as a limit of f_k, a function $f_\beta : I^2 \to \mathbb{R}$, such that: for all $x \in I_{\frac{1}{3}}$ (the classical Cantor set of I constructed with the ratio 3), $f_\beta(x) = \sum\limits_{k\geq 0} \dfrac{1}{9^k}(8/9+\xi_k) = 1 + \sum\limits_{k\geq 0} \dfrac{\xi_k}{9^k}$, where $\xi_k = b_j$ if $x = \bigcap\limits_{n\in\mathbb{N}} C^{i_1,i_2,\dots,i_n}$ and $i_k = j$.

In particular f_β gives a bijection between $I_{\frac{1}{3}}$ and C_β. It is not difficult to prove that f_β is a $\mathcal{C}^{\frac{2\ln 3}{\ln\beta}}$-smooth mapping, with critical locus $I_{\frac{1}{3}}$, for $2 < \beta < 9^{\frac{1}{3}}$ (see [Com 1]).

Let us notice that if we had considered, in the above construction, the classical Cantor set $I_{\frac{1}{\alpha}}$ of I, constructed with the ratio $1/\alpha$ instead of $1/3$, one would have obtained a $\mathcal{C}^{\frac{2\ln\alpha}{\ln\beta}}$-smooth function $f_{\beta,\alpha}$, for $2 < \beta < (\alpha^2)^{1/\mathrm{E}(2\frac{\ln\alpha}{\ln\beta})}$, hence as $\alpha \longrightarrow \infty$, $f_{\beta,\alpha}$ can be as regular as desired.

Let us notice that by [Yom 13], Theorem 5.6, for functions $f : B^n_r \to \mathbb{R}$, the condition $\dim_e(A) \leqslant \frac{n}{k}$ is almost a sufficient condition for the set A to be a set of critical values of a $\mathcal{C}^k$-smooth function f: given a bounded set $A \subset \mathbb{R}$ such that $\dim_e(A) < \frac{n}{k}$, there exists a $\mathcal{C}^k$-smooth function $f : B^n_r \to \mathbb{R}$, with $A \subset \Delta^0_f$. Consequently, given $I_{\frac{1}{3}}$ as in the example above, $\dim_e(I_{\frac{1}{3}}) = \frac{\ln 2}{\ln 3} <$

$\frac{\ln\beta}{\ln 3}$, for $\beta \in]2,\infty[$, formally implies that there exists f_β, $\mathcal{C}^{2\frac{\ln 3}{\ln\beta}}$-smooth, such that $I_{\frac{1}{3}} \subset \Delta^0_{f_\beta}$.

• Unlike the Federer theorem, Theorem 9.3, tells nothing about $\mathcal{H}^{\nu+\frac{n-\nu}{k}}(\Delta^\nu_f)$: we do not know from 9.3 for instance, whether or not

$$\mathcal{H}^{\nu+\frac{n-\nu}{k}}(\Delta^\nu_f) = 0. \tag{$*$}$$

Of course, by [Com 1], equality ($*$) would be the best quantitative result in terms of Hausdorff measure.

• Equality ($*$) has been first proved in two particular cases:

(i) For $k = (n-\nu)/(m-\nu)$, that is to say: if $f : \mathbb{R}^n \to \mathbb{R}^m$ is a $\mathcal{C}^{\frac{n-\nu}{m-\nu}}$-smooth mapping, then $\mathcal{H}^m(\Delta^\nu_f) = 0$ ([Bat 1]).

(ii) For $\nu = 0$, that is to say: if $f : \mathbb{R}^n \to \mathbb{R}^m$ is a $\mathcal{C}^k$-smooth mapping, than $\mathcal{H}^{\frac{n}{k}}(\Delta^0_f) = 0$ ([Com 1]).

One can prove the classical Morse-Sard theorem (concerning the $\mathcal{H}^m$-nullity of Δ^ν_f, see [Sar 1,2], [Bat 1], [Nor 1]) by induction on ν, starting from $\nu = 0$, and using Fubini's theorem. Unfortunately Fubini's theorem does not hold anymore for non-integral Hausdorff measures (see [Fal]), thus one cannot prove ($*$) from *(ii)* as it is done in the classical case. One can only prove in this way that for any point $a \in \Delta^\nu_f$ there exists an open subset U of $\mathbb{R}^m$ containing a, such that $U \cap \Delta^\nu_f$ is $\mathcal{H}^\nu \otimes \mathcal{H}^{\frac{n-\nu}{k}}$-null, where $\mathcal{H}^\nu$ measures $\mathbb{R}^\nu$ and $\mathcal{H}^{\frac{n-\nu}{k}}$ measures $\mathbb{R}^{m-\nu}$ (in some suitable $\mathcal{C}^k$-coordinate system for U), but in general, as mentioned above, $\mathcal{H}^\nu \otimes \mathcal{H}^{\frac{n-\nu}{k}} \neq \mathcal{H}^{\nu+\frac{n-\nu}{k}}$ (on this question see [Fal], [Fed 2], [Nor 2] or for a counterexample see [Sar 3]).

• However if $f : \mathbb{R}^n \to \mathbb{R}^m$ is a $\mathcal{C}^{(p+\alpha)+}$-smooth mapping (it means that there exists for every $a \in \mathbb{R}^n$, a ball $B^n_{(a,r)}$, and a function $\epsilon_a : \mathbb{R} \to \mathbb{R}$ such that: $\epsilon_a(t) \underset{t\to 0}{\longrightarrow} 0$ and $||D^p f_{(x)} - D^p f_{(y)}|| \leqslant \epsilon_a(||x-y||)||x-y||^\alpha$, for all $x, y \in B^n_{(a,r)}$), the proof of theorem 3.4.3 of [Fed 2] gives ($*$): $\mathcal{H}^{\nu+\frac{n-\nu}{p+\alpha}}(\Delta^\nu_f) = 0$. Finally, let us note that the proof of theorem 3.4.3 of [Fed 2] gives also: $\dim_{\mathcal{H}}(\Delta^\nu_f) \leqslant \nu + \dfrac{n-\nu}{k}$, because if $f : \mathbb{R}^n \to \mathbb{R}^m$ is a $\mathcal{C}^k$-smooth mapping, it is also a $\mathcal{C}^{\ell+}$-smooth mapping, for any $1 \leqslant \ell < k$, and thus one has by [Fed 2], Theorem 3.4.3: $\mathcal{H}^{\nu+\frac{n-\nu}{\ell}}(\Delta^\nu_f) = 0$, for every $1 \leqslant \ell < k$, showing that $\dim_{\mathcal{H}}(\Delta^\nu_f) \leqslant \nu + \dfrac{n-\nu}{k}$.

• Finally it has been proved in [Mor] by Carlos Gustavo T. de A. Moreira that $\mathcal{C}^k$-regularity for f implies equality ($*$). The difficulty mentioned above concerning the use of Fubini's theorem has been overcome by a "careful decomposition of the critical set, combined with a parametrized strong version of the so-called A. P. Morse lemma.

Let us give Moreira's statement:

Theorem ([Mor]). *Let $f : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}^m$ be a $\mathcal{C}^{k=p+\alpha}$-smooth mapping, where $\mathcal{U}$ is an open subset of $\mathbb{R}^n$. We have:*

$$\mathcal{H}^{\nu+\frac{n-\nu}{k}}(\Delta_f^\nu) = 0.$$

Remark. This result, which concerns the Hausdorff dimension of Δ_f^ν, as well as Theorem 9.3, which concerns the entropy dimension of Δ_f^ν, cannot be sharpened, as it is indicated in comments above. Examples may be also find in [Mor].

The next example of applications of semialgebraic results in a smooth category concerns the following question: what additional information do the high smoothness assumptions give in situations where the Morse-Sard theorem is true a priori for $\mathcal{C}^1$-mappings, say if $n \leqslant m$.

We start with the case $n < m$. Here assuming $f : B_r^n \to \mathbb{R}^m$ is $\mathcal{C}^1$, we get immediately $M(\epsilon, f(B_r^n)) \sim c.(\frac{r}{\epsilon})^n$, since f, being Lipschitzian, preserves the ϵ-entropy.

Applying this situation in Theorem 9.2, we can assume that $\lambda_1 = \ldots = \lambda_q$ bound $\lambda_i(Df_{(x)})$, for each $x \in B_r^n$ and each $i \in \{1, \ldots, q\}$. We obtain the bound:

$$M(\epsilon, f(B_r^n)) \leqslant c' \sum_{i=0}^{i=q=n} (\frac{r}{\epsilon})^i \left(\frac{R_k(f)}{\epsilon}\right)^{\frac{n-i}{k}},$$

and for $k \geq 1$, $i \leqslant n$, the degree of $\frac{1}{\epsilon}$ in each term is $i + \frac{n-i}{k} = n - (n-i)(1 - \frac{1}{k}) \leqslant n$, and for $i = n$ this degree is exactly n, for any k. Thus Theorem 9.2 gives no additional informations for $k > 1$. Of course, in the global setting of Theorem 9.2, the above bound is the best possible even for f a linear mapping, since the ϵ-entropy of the n-dimensional subspace in $\mathbb{R}^m$ is of order $(\frac{1}{\epsilon})^n$.

However, if we ask another question of a more local nature, we find once more that the high smoothness of f strongly influences its behavior. The question is: how far should we move a given point in $\mathbb{R}^m$ to put it out of the ϵ-neighborhood of $f(B_r^n)$? In other words, what is the maximal radius of the ball, entirely contained in the ϵ-neighborhood of $f(B_r^n)$?

We recall that we have defined, for $k > 1$, $R_k(f) = \frac{K}{p!}.r^{k=p+\alpha}$, where K is defined by: $||D^p f_{(x)} - D^p f_{(y)}|| \leqslant K||x-y||^\alpha$, for all $x, y \in B_r^n$.

Now we define $R_1(f) = r. \sup_{\zeta \in B_r^n} ||Df_{(\zeta)}||$, Df being continuous.

Theorem 9.4. *Let $f : B_r^n \to \mathbb{R}^m$, $n < m$, be a $\mathcal{C}^k$-smooth mapping. Then for $\epsilon > 0$, the maximal radius of a ball contained in the ϵ-neighborhood of $f(B_r^n)$ is:*

$$c.\min\left(\epsilon\left(\frac{R_k(f)}{\epsilon}\right)^{\frac{n}{k(m-n)}},\epsilon\left(\frac{R_1(f)}{\epsilon}\right)^{\frac{n}{m}}\right),$$

where c depends only on n, m and k.

Actually we prove the following more precise result:

Theorem 9.5. *Let $f : B_r^n \to \mathbb{R}^m$, $n < m$, be a $\mathcal{C}^k$-smooth mapping, and B_δ^m be some ball of radius δ in $\mathbb{R}^m$. Then for $\delta \geq \epsilon > 0$:*

$$M(\epsilon, f(B_r^n) \cap B_\delta^m) \leqslant c'.\min\left(\left(\frac{R_1(f)}{\epsilon}\right)^n, \left(\frac{\delta}{\epsilon}\right)^n \left(\frac{R_k(f)}{\epsilon}\right)^{\frac{n}{k}}\right),$$

where c' depends only on n, m and k.

Of course Theorem 9.4 is a trivial consequence of Theorem 9.5: if B_δ^m is a ball contained in $f(B_r^n)$, we have by 9.5:

$$M(\epsilon, f(B_r^n) \cap B_\delta^m) = M(\epsilon, B_\delta^m) = C.(\frac{\delta}{\epsilon})^m \leqslant$$

$$\leqslant \min\left(\left(\frac{R_1(f)}{\epsilon}\right)^n, \left(\frac{\delta}{\epsilon}\right)^n \left(\frac{R_k(f)}{\epsilon}\right)^{\frac{n}{k}}\right),$$

Thus:

$$\delta \leqslant c.\min\left(\epsilon\left(\frac{R_k(f)}{\epsilon}\right)^{\frac{n}{k(m-n)}},\epsilon\left(\frac{R_1(f)}{\epsilon}\right)^{\frac{n}{m}}\right).$$

Proof of Theorem 9.5. The inequality:

$$M(\epsilon, f(B_r^n) \cap B_\delta^m) \leqslant M(\epsilon, f(B_r^n)) \leqslant c'.\left(\frac{R_1(f)}{\epsilon}\right)^n$$

follows from Property 2.8 of ϵ-entropy, since the Lipschitz constant of f is $R_1(f)/r$ on B_r^n. To prove that $M(\epsilon, f(B_r^n) \cap B_\delta^m) \leqslant c'.\left(\frac{\delta}{\epsilon}\right)^n \left(\frac{R_k(f)}{\epsilon}\right)^{\frac{n}{k}}$, we cover B_r^n by balls of radius $r' = r\left(\frac{\epsilon}{R_k(f)}\right)^{\frac{1}{k}}$. For any such ball $B_{r'}^n$, and P_p the Taylor polynomial for f of degree p at the centre of $B_{r'}^n$, we have:

$$||f(x) - P_p(x)|| \leqslant \frac{K}{p!} r'^k = \frac{K}{p!} r^k . \frac{\epsilon}{R_k(f)} = \epsilon, \text{ for any } \ x \in B_{r'}^n.$$

Hence:

$$f(B_{r'}^n) \cap B_\delta^m \subset \left[P_p(B_{r'}^n)\right]_\epsilon \cap B_\delta^n.$$

We thus get:

$$M(2\epsilon, f(B_{r'}^n) \cap B_\delta^m) \leqslant M(2\epsilon, \left[P_p(B_{r'}^n)\right]_\epsilon \cap B_\delta^n) \leqslant M(2\epsilon, \left[P_p(\mathbb{R}^n)\right]_\epsilon \cap B_\delta^n)$$

$$\leqslant M(2\epsilon, \left[P_p(\mathbb{R}^n)\cap B^n_{\delta+\epsilon}\right]_\epsilon) \leqslant M(\epsilon, P_p(\mathbb{R}^n)\cap B^n_{\delta+\epsilon}) \leqslant \widetilde{c}.(\frac{\epsilon+\delta}{\epsilon})^n \leqslant 2^n\widetilde{c}.(\frac{\delta}{\epsilon})^n,$$

since $P_p(\mathbb{R}^n)$ is a semialgebraic set of fixed complexity and of dimension $\leqslant n(< m)$ in $B^m_{\epsilon+\delta}$, and since we assume here that $\epsilon \leqslant \delta$. (see Corollary 5.7).

Finally, the number of balls $B^n_{r'}$ one needs to cover B^n_r is $C.(\frac{r}{r'})^n$ that is to say $C'(\frac{R_k(f)}{\epsilon})^{\frac{n}{k}}$, completing the proof. □

Remark. Notice that the latter quantity in Theorem 9.5 is smaller than the first one only if $\epsilon \geq R_k(f)(\frac{\delta}{R_1(f)})^k$. Since the bounds for the ϵ-entropy of the subsets of B^m_δ are meaningful only if $\epsilon < \delta$, we see that, at least as $\delta \longrightarrow 0$, for any $k > 1$, we have the range of values of ϵ, namely $\delta \geq \epsilon \geq R_k(f)(\frac{\delta}{R_1(f)})^k$, where our bounds for the ϵ-entropy of $f(B^n_{r'}) \cap B^m_\delta$ are strictly better than for $\mathcal{C}^1$-mappings.

The bound of Theorem 9.5 is "almost sharp". Below we give corresponding examples. Let us start with a mapping, given by an explicit analytic formula (but providing roughly twice lower smoothness than in the optimal construction.) Consider the following $\mathcal{C}^{\frac{k}{2}}$-smooth mapping $g : [0;1]^n \to \mathbb{R}^{2n}$, where $k > 2$:

$$g(t_1,\ldots,t_n) = (t_1^k,\ldots,t_n^k, t_1^k\sin(\frac{1}{t_1}),\ldots,t_n^k\sin(\frac{1}{t_n})).$$

We have:

$$Vol_n(g([0;1]^n)\cap B^{2n}_\delta) \geq \left(\frac{2(\sqrt{n})^{\frac{1}{k}-1}}{2^k\pi(k-1)}\right)^n \delta^n.(\frac{1}{\delta})^{\frac{n}{k}} + (\frac{\delta}{\sqrt{n}})^n. \qquad (**)$$

Indeed, on one hand for $(t_1,\ldots,t_n) \in [0;\delta^{\frac{1}{k}}]^n$, $g(t_1,\ldots,t_n) \subset [0;\delta]^{2n} \subset B^{2n}_{\delta\sqrt{n}}$, thus $Vol_n(g([0;1]^n)\cap B^{2n}_{\delta\sqrt{n}}) \geq Vol_n(g([0;\delta^{\frac{1}{k}}]^n)$, and on the other hand $Vol_n(g([0;\delta^{\frac{1}{k}}]^n) = \left(Vol_1(\gamma)\right)^n$, because $g([0;\delta^{\frac{1}{k}}]^n) = \gamma^n \subset (\mathbb{R}^2)^n$, where $\gamma = \{(t^k, t^k\sin(\frac{1}{t})) \in \mathbb{R}^2;\ t \in [0;\delta^{\frac{1}{k}}]\}$. Now let us show that $Vol_1(\gamma) \geq \frac{2}{2^k\pi(k-1)}\delta.(\frac{1}{\delta})^{\frac{1}{k}} + \delta^n$. We have:

$$Vol_1(\{(t^k, t^k\sin(\frac{1}{t}));\ t \in [\frac{1}{\pi(p+1)};\frac{1}{\pi p}]\}) \geq \left(\frac{1}{\pi(p+\frac{1}{2})}\right)^k.$$

Thus one gets:

$$Vol_1(\gamma) \geq \frac{1}{\pi^k}\sum_{p=p_0}^{+\infty}\frac{1}{(p+\frac{1}{2})^k} + \delta,$$

where p_0 is the smallest integer satisfying: $p_0 \geq \frac{1}{\pi \delta^{\frac{1}{k}}}$. We obtain:

$$Vol_1(\gamma) \geq \frac{1}{\pi^k} \int_{\frac{1}{\pi\delta^{1/k}}+1}^{+\infty} \frac{dt}{(t+\frac{1}{2})^k} + \delta \geq \frac{2}{2^k(k-1)\pi} \left(\frac{1}{\delta}\right)^{\frac{1}{k}-1} + \delta$$

which proves $(**)$. Consequently we have the following lower bound, for $\epsilon \to 0$:

$$M(\epsilon, g([0;1]^n) \cap B_\delta^{2n}) \geq c.(\frac{1}{\epsilon})^n Vol_n(g([0;1]^n) \cap B_\delta^{2n})$$

$$\geq c.\left(\frac{2(\sqrt{n})^{\frac{1}{k}-1}}{2^k\pi(k-1)}\right)^n (\frac{\delta}{\epsilon})^n.(\frac{1}{\delta})^{\frac{n}{k}} + c.(\frac{\delta}{\sqrt{n}\epsilon})^n,$$

where c depends only on n. More accurate analysis of the geometry of $g([0;1]^n) \cap B_\delta^{2n}$ allows one to show that this lower bound becomes valid starting with $\epsilon \sim \delta^{1+1/k}$ and smaller. The upper bound given by Theorem 9.5 being $c'.(\frac{\delta}{\epsilon})^n(\frac{1}{\epsilon^2})^{\frac{n}{k}}$ (g is a $\mathcal{C}^{k/2}$-smooth mapping), we obtain:

$$c.\left(\frac{2(\sqrt{n})^{\frac{1}{k}-1}}{2^k\pi(k-1)}\right)^n (\frac{\delta}{\epsilon})^n.(\frac{1}{\delta})^{\frac{n}{k}} + c.(\frac{\delta}{\sqrt{n}\epsilon})^n \leqslant M(\epsilon, g([0;1]^n) \cap B_\delta^{2n})$$

$$\leqslant c'.(\frac{\delta}{\epsilon})^n(\frac{1}{\epsilon^2})^{\frac{n}{k}}.$$

Hence the ratio of the lower and the upper bound is:

$$1 \geq \frac{c}{c'}.\left(\frac{2(\sqrt{n})^{\frac{1}{k}-1}}{2^k\pi(k-1)}\right)^n (\frac{\epsilon^2}{\delta})^{\frac{n}{k}} + \frac{c}{c'(\sqrt{n})^n}(\epsilon^2)^{\frac{n}{k}}$$

$$\geq \frac{c}{c'}.\left(\frac{2(\sqrt{n})^{\frac{1}{k}-1}}{2^k\pi(k-1)}\right)^n (\frac{1}{\delta})^{\frac{n}{k}+\frac{2n}{k^2}},$$

for $\epsilon = \delta^{1+\frac{1}{k}}$.

Additional results in this direction are given below.

Comparing the two expressions in Theorem 9.4, we see that the first one (the one which gives the best bound for highly differentiable functions) is better than the second (asymptotically, as $\epsilon \longrightarrow 0$), if:

$$\frac{n}{k(m-n)} < \frac{n}{m}, \quad \text{or } k > \frac{m}{m-n}.$$

Once more, this bound is virtually sharp.

Consider the mapping $g : [0;1]^n \to \mathbb{R}^{2n}$, built above: $g(t_1, \ldots, t_n) = (t_1^k, \ldots, t_n^k, t_1^k \sin(\frac{1}{t_1}), \ldots, t_n^k \sin(\frac{1}{t_n}))$. Its image is contained in the part of $\mathbb{R}^{2n}$, defined by $|x_{n+i}| \leqslant |x_i|$, $i = 1, \ldots, n$. But taking 2^n such mappings, we "cover" all the space. Thus it is enough to assume that the coordinates of the considered points satisfy: $|x_{n+i}| \leqslant |x_i|$, for all $i \in \{1, \dots, n\}$. So let a point $x = (x_1, \dots, x_n, x_{n+1}, \dots, x_{2n}) \in \mathbb{R}^{2n}$ be given, with $|x_{n+i}| \leqslant |x_i|$. We claim that the distance $\mathrm{d}(x, g([0;1]^n))$ is at most $||x||^{1+\frac{1}{k}}$. Indeed, we can find ξ_i, such that:

$$\frac{1}{x_i^{\frac{1}{k}}} \leqslant \frac{1}{\xi_i} \leqslant \frac{1}{x_i^{\frac{1}{k}}} + 2\pi \quad \text{and} \quad \sin(\frac{1}{\xi_i}) = \frac{x_{n+i}}{x_i} \in [-1; 1].$$

We thus have:

$$0 \leqslant x_i - \xi_i^k \leqslant x_i - \frac{1}{(\frac{1}{x_i^{\frac{1}{k}}} + 2\pi)^k} \leqslant x_i\left[(1+2\pi x_i^{\frac{1}{k}})^k - 1\right] \leqslant x_i(2k\pi x_i^{\frac{1}{k}}(1+2\pi x_i^{\frac{1}{k}})^{k-1})$$

and finally:

$$|x_i - \xi_i^k| \leqslant c.x_i^{1+\frac{1}{k}}, \tag{3}$$

where c depends only on k. Now, of course, we also have:

$$0 \leqslant |x_{n+i} - \xi^k \sin(\frac{1}{\xi_i})| = |x_i \sin(\frac{1}{\xi_i}) - \xi^k \sin(\frac{1}{\xi_i})| \leqslant |x_i - \xi_i^k| \leqslant c.x_i^{1+\frac{1}{k}}. \tag{4}$$

Inequalities (3) and (4) show that $\mathrm{d}(x, g([0;1]^n) \leqslant \widetilde{c}.||x||^{1+\frac{1}{k}}$, where $\widetilde{c}$ depends only on k and n. Therefore, if we take $\epsilon = \widetilde{c}.\delta^{1+\frac{1}{k}}$, the ball of radius δ centred at the origin is contained in the ϵ-neighborhood of $g([0;1]^n)$. Thus the lower bound for the maximal radius of a ball, contained in $\big(g([0;1]^n)\big)_\epsilon$, is $c'.\epsilon^{1-\frac{1}{k+1}}$. The upper bound, given by theorem 9.4 is $c.\epsilon^{1-\frac{n}{(k/2)(2n-n)}} = c.\epsilon^{1-\frac{2}{k}}$.

Examples in the same spirit can be given for any m and n, $n < m$, and with an essentially maximal possible differentiability:

Let $\Psi : B_1^n \to \mathbb{R}^n$ be a $\mathcal{C}^\infty$-mapping with the following properties:

- $\Psi_{|B_1^n \setminus B_{1/2}^n} \equiv 0$,
- $B_1^n \subset \Psi(B_{1/4}^n)$.

We assume that $m > n$ and consider, for any $s > 0$ in $\mathbb{R}^{m-n}$, the following net $\mathcal{Z}_s = \{Z_\alpha^s\}$: on each sphere S_{1/N^s}^{m-n-1}, the points Z_α^s form an $\dfrac{1}{N^{s+1}}$-net, $N \in \mathbb{N} \setminus \{0\}$, and the number of Z_α^s on S_{1/N^s}^{m-n-1} is $c.\left[\dfrac{1/N^s}{1/N^{s+1}}\right]^{m-n-1} = c.N^{m-n-1}$, where c only depends on n and m.

Now for each α and N, such that $Z_\alpha^s \in S_{1/N^s}^{m-n-1}$, let $r_\alpha = (\dfrac{1}{N})^{\frac{s}{k}}$, where $k(= p + \beta) \in \mathbb{R}$. For any $r'_\alpha \leqslant r_\alpha$, the mapping: $\dfrac{1}{N^s}\Psi(\dfrac{1}{r'_\alpha}x)$ is a $\mathcal{C}^k$-smooth mapping, such that all its derivatives up to order p, as well as the Hölder constant K of order β of $D^p\Psi$ (see the very beginning of this chapter for the definition of K) are bounded by a constant not depending on N, s, etc...

(but depending on the ratio r'_α/r_α.), and such a mapping is zero off $B^n_{r'_\alpha}$. It follows that when the balls $B^n_{r'_\alpha}$ centred at certain points x_α are contained in B^n_1 and disjoint, and all the ratios r'_α/r_α are uniformly bounded with respect to α, the mapping $\sum_\alpha \frac{1}{N^s}\Psi(\frac{1}{r'_\alpha}(x - x_\alpha))$ is a $\mathcal{C}^k$-mapping.

Now let $\Phi : B^n_r \to \mathbb{R}$ be a $\mathcal{C}^\infty$-function with the following properties:

- $\Phi_{|B^n_1\setminus B^n_{1/2}} \equiv 0$,
- $\Phi_{|B^n_{1/4}} \equiv 1$.

If the sum $\sum_\alpha r^n_\alpha$ converges, we can find balls $B^n_{r'_\alpha}$, contained in B^n_1, disjoint, and with $\frac{r'_\alpha}{r_\alpha} \geq c$, c not depending on α (see [Iva 1]). Let us take as x_α the centers of $B^n_{r'_\alpha}$.

Thus if we define $g : B^n_1 \to \mathbb{R}^m$ as follows:

$$g(x) = \sum_\alpha \left(\frac{1}{N^s}\Psi(\frac{1}{r'_\alpha}(x - x_\alpha)), Z^s_\alpha.\Phi(\frac{1}{r'_\alpha}(x - x_\alpha))\right) \in \mathbb{R}^n \times \mathbb{R}^{m-n} = \mathbb{R}^m,$$

g is a $\mathcal{C}^k$-smooth mapping of B^n_1.

By construction the sets $g(B^n_1)$ form an ϵ-net of the cone $||y_1|| \leqslant ||y_2||$ in $\mathbb{R}^m = \mathbb{R}^n \times \mathbb{R}^{m-n} = \{(y_1, y_2) \in \mathbb{R}^n \times \mathbb{R}^{m-n}\}$, intersected with the ball B_δ, where $\epsilon = \delta^{\frac{s+1}{s}}$. Indeed, for $\delta = \frac{1}{N^s}$, $\epsilon = \frac{1}{N^{s+1}} = \delta^{\frac{s+1}{s}}$. Now by construction, the image under g of the ball of radius $\frac{r'_\alpha}{4}$, centred at x_α, is the ball $(B^n_{\frac{1}{N^s}}, Z^s_\alpha) \subset \mathbb{R}^n \times \mathbb{R}^{m-n} = \mathbb{R}^n$. Thus B^m_δ is contained in the ϵ-neighborhood of $g(B^n_1)$, for $\delta = \epsilon^{1-\frac{1}{s+1}}$. It remains to determine the allowed values for s: the condition $\sum_\alpha r^n_\alpha < \infty$ means that $\sum_{N=1}^\infty N^{m-n-1}(\frac{1}{N})^{\frac{sn}{k}} = \sum_{N=1}^\infty N^{m-n-1-\frac{sn}{k}} < \infty$, ie $s > \frac{k(m-n)}{n}$.

Finally the lower bound for $\delta = \delta(\epsilon)$ in our example, is $\epsilon^{1-\frac{n}{n+k(m-n)}}$, while Theorem 9.4 gives an upper bound which is $\epsilon^{1-\frac{n}{k(m-n)}}$. In particular, for $k \longrightarrow \infty$, this bound is asymptotically sharp.

Estimation of the ϵ-entropy in this last example allows one to show essential sharpness (as $k \to \infty$) also of the bounds in Theorem 9.5.

Exactly the same arguments as in Theorem 9.5 can be applied in more general situations. For instance:

Theorem 9.6. *Let $f : B^n_r \to \mathbb{R}^m$ be a $\mathcal{C}^k$-smooth mapping, $q = min(n, m)$, $\Lambda = (\lambda_1, \ldots, \lambda_q)$, B^m_δ be some ball of radius δ in $\mathbb{R}^m$, and $\epsilon \leqslant \delta$. Then we have:*

$$M(\epsilon, \Delta(f, \Lambda, B^n_r) \cap B^m_\delta) \leqslant$$

$$\leqslant c.\left(\frac{R_k(f)}{\epsilon}\right)^{\frac{n}{k}} \sum_{i=0}^{q} min\left(\lambda_0 \dots \lambda_i (\frac{r}{\epsilon})^i (\frac{\epsilon}{R_k(f)})^{\frac{i}{k}}, (\frac{\delta}{\epsilon})^i\right).$$

Proof. The main argument in the proof of Theorem 9.5, consists in the approximation of f by its Taylor polynomial on balls of radius $r' = r\left(\frac{\epsilon}{R_k(f)}\right)^{\frac{1}{k}}$. As applied to $\Delta(f, \Lambda, B_r^n)$, it becomes the same as the argument in the proof of Theorem 9.2, thus we obtain the same bound: $M(\epsilon, \Delta(f, \Lambda, B_r^n) \cap B_\delta^m) \leqslant \widetilde{c}.\left(\frac{R_k(f)}{\epsilon}\right)^{\frac{n}{k}} \sum_{i=0}^{q} \lambda_0 \dots \lambda_i (\frac{r}{\epsilon})^i (\frac{\epsilon}{R_k(f)})^{\frac{i}{k}}$.

Let us now prove that: $M(\epsilon, \Delta(f, \Lambda, B_r^n) \cap B_\delta^m) \leqslant c'.\left(\frac{R_k(f)}{\epsilon}\right)^{\frac{n}{k}} \sum_{i=0}^{q} (\frac{\delta}{\epsilon})^i$. As in the proof of Theorem 9.5, we cover the ball B_r^n by $C'(\frac{R_k(f)}{\epsilon})^{\frac{n}{k}}$ balls of radius $r' = r\left(\frac{\epsilon}{R_k(f)}\right)^{\frac{1}{k}}$. Then approximating f on each of these small balls, by P_p, its Taylor polynomial of degree p, we obtain (cf. the proof of Theorem 9.5):

$$M(2\epsilon, \Delta(f, \Lambda, B_r^n) \cap B_\delta^m) \leqslant M(\epsilon, P_p(\mathbb{R}^n) \cap B_{\delta+\epsilon}^m) \leqslant M(\epsilon, P_p(\mathbb{R}^n) \cap B_{2\delta}^m).$$

Now by Theorem 3.5: $M(\epsilon, P_p(\mathbb{R}^n) \cap B_{2\delta}^m) \leqslant C(m) \sum_{i=0}^{m} \frac{V_i(P_p(\mathbb{R}^n) \cap B_{2\delta}^m)}{\epsilon^i}$. But $P_p(\mathbb{R}^n) \cap B_{2\delta}^m$ is a semialgebraic set of $B_{2\delta}^m$ of dimension less than q, and of fixed complexity, hence by Corollary 5.2, we obtain: $V_i(P_p(\mathbb{R}^n) \cap B_{2\delta}^m) \leqslant C.\delta^i$, for $i \leqslant q$, and $V_i(P_p(\mathbb{R}^n) \cap B_{2\delta}^m) = 0$, for $i > q$, completing the proof. □

We shall consider in detail only the following special case of Theorem 9.6.: let $\nu < m$ be given. As usual, we denote by Δ_f^ν the image by f of $\Sigma_f^\nu = \{x \in B_r^n;\ \text{rank}(Df_{(x)}) \leqslant \nu\}$. Considering $\Lambda = (\lambda_1, \dots \lambda_\nu, 0, \dots, 0)$ such that $\lambda_i(Df_{(x)}) \leqslant \lambda_i$, for all $x \in B_r^n$ and all $i \in \{1, \dots, \nu\}$, we have $\Delta_f^\nu = \Delta(f, \Lambda, B_r^n)$. We get by Theorem 9.6:

$$M(\epsilon, \Delta_f^\nu \cap B_\delta^m) \leqslant c.\left(\frac{R_k(f)}{\epsilon}\right)^{\frac{n}{k}} (\frac{\delta}{\epsilon})^\nu.$$

But of course we also have:

$$M(2\epsilon, [\Delta_f^\nu]_\epsilon \cap B_\delta^m) \leqslant M(2\epsilon, [\Delta_f^\nu \cap B_{\delta+\epsilon}^m]_\epsilon) \leqslant M(\epsilon, \Delta_f^\nu \cap B_{\delta+\epsilon}^m).$$

We conclude that:

$$M(\epsilon, [\Delta_f^\nu]_\epsilon \cap B_\delta^m) \leqslant c'.M(\epsilon, \Delta_f^\nu \cap B_{\delta+\epsilon}^m) \leqslant \widetilde{c}.\left(\frac{R_k(f)}{\epsilon}\right)^{\frac{n}{k}} (\frac{\delta+\epsilon}{\epsilon})^\nu \leqslant$$

$$\leqslant \widetilde{C}.\left(\frac{R_k(f)}{\epsilon}\right)^{\frac{n}{k}}(\frac{\delta}{\epsilon})^{\nu}.$$

Now if a ball B_δ^m is contained in the ϵ-neighborhood of Δ_f^ν, we get:

$$M(\epsilon,[\Delta_f^\nu]_\epsilon \cap B_\delta^m) = M(\epsilon, B_\delta^m) \leqslant \widetilde{C}.\left(\frac{R_k(f)}{\epsilon}\right)^{\frac{n}{k}}(\frac{\delta}{\epsilon})^{\nu}$$

Comparing this with $M(\epsilon, B_\delta^m) = C.(\frac{\delta}{\epsilon})^m$, we obtain the following corollary:

Corollary 9.7. *Let $f : B_r^n \to \mathbb{R}^m$ be a $\mathcal{C}^k$-smooth mapping. The maximal radius of a ball contained in $[\Delta_f^\nu]_\epsilon$ is $c.\epsilon\left(\frac{R_k(f)}{\epsilon}\right)^{\frac{n}{k(m-\nu)}}$.* □

Examples of the type considered above show that this bound is virtually sharp:

For instance, let us consider $\varphi : \mathbb{R} \to [0;1]$, a $\mathcal{C}^\infty$-smooth function, such that $\varphi(\frac{1}{2}) = 1$, $\varphi(x) = 0$, for all $x \in \mathbb{R} \setminus [0;1]$ and $|\varphi'(x)| \leqslant 2$, for all $x \in \mathbb{R}$. If we denote by f the function defined as follows:

$$f(x) = \sum_{p=1}^{\infty} \frac{1}{p^k}.\varphi(\frac{(2p-1)(2p+1)}{2}x - \frac{2p-1}{2}),$$

we obtain a $\mathcal{C}^{\frac{k}{2}}$-smooth function with $\Delta_f^0 = \{0, 1, \frac{1}{2^k}, \ldots, \frac{1}{p^k}, \ldots\}$. Now $[\Delta_f^0]_\epsilon$ contains an interval of length $\delta(\epsilon) = \epsilon + \frac{1}{p^k}$, for $\epsilon = \frac{1}{p^k} - \frac{1}{(p+1)^k}$. It follows, after an easy computation, that $\delta(\epsilon) = \epsilon^{1-\frac{1}{k+1}}(\frac{1}{(k+\eta(\frac{1}{p}))^{\frac{k}{k+1}}} + \epsilon^{\frac{1}{k+1}})$, where $\eta(\frac{1}{p})$ is a function tending to 0 as p tends to ∞. Finally, $\delta(\epsilon) \geq c.\epsilon^{1-\frac{1}{k+1}}$, with c depending only on k.

Under these asumptions, Corollary 9.7 gives $\delta(\epsilon) \leqslant c.\epsilon\left(\frac{R_k(f)}{\epsilon}\right)^{\frac{2}{k}} = c'\epsilon^{1-\frac{2}{k}}$.

Exercise. Find sequences $\alpha_p \to 0$, $\beta_p \to \infty$, for which the function $f(x) = \sum_{p=1}^{\infty} \frac{1}{p^k}\varphi(\beta_p(x-\alpha_p))$ is C^γ, with γ arbitrarily close to k.

A more general example is the following: assume $n \leqslant m$ and represent $\mathbb{R}^m$ as $\mathbb{R}^\nu \times \mathbb{R}^{m-\nu}$. We build $g : B_1^n \to \mathbb{R}^m$ exactly as the mapping g above, taking $\Psi : B_1^n \to \mathbb{R}^\nu$ to be the mapping covering the unit ball in $\mathbb{R}^\nu$. Thus the rank of Dg is ν on each ball of radius $\dfrac{1}{4}r'_\alpha$, centred at x_α. Therefore the Δ_g^ν form an ϵ-net in B_δ^m , $\epsilon = \delta^{\frac{s}{s+1}}$ and B_δ^m is contained in the ϵ-neighborhood of Δ_g^ν, for $\delta = \epsilon^{1-\frac{1}{s+1}}$ and $s > \dfrac{k(m-\nu)}{n}$. Thus the lower bound for $\delta = \delta(\epsilon)$ in this example is $\epsilon^{1-\frac{n}{n+k(m-\nu)}}$, while Corollary 9.7 gives: $\delta(\epsilon) \leqslant c.\epsilon^{1-\frac{n}{k(m-\nu)}}$, showing that this bound is asymptotically sharp, as $k \longrightarrow \infty$.

Our next result concerns quantitative transversality. As usual, we assume that the domain and the image of our mappings are the Euclidean balls. Moreover, we assume actually (although the statement of the Theorem 9.8 below is more general), that the submanifold in the image, to which our mapping should be transversal, is the origin. (The general situation is reduced to this special case by composing our mapping with the projection along the submanifold to its normal plane).

So let $f = (f_1, \ldots, f_m) : B_r^n \times B_r^m \to \mathbb{R}^m$ be a $\mathcal{C}^k$mapping. We assume the following:

– For any $(x,t) \in B_r^n \times B_r^m$, $\lambda_m(D_t f_{(x,t)}) \geq \rho > 0$. By compactness of $B_r^n \times B_r^m$, this assumption is of course equivalent to the following:

$$D_t f_{(x,t)} : \mathbb{R}^m \to \mathbb{R}^m$$

is onto.

From this assumption, it follows that, for any $x \in B_r^n$, the $\mathcal{C}^k$-smooth mapping $f(x,.) : B_r^m \to \mathbb{R}^m$ is locally invertible. If furthermore we assume that:

– For any $x \in B_r^n$, $f(x,.) : B_r^m \to \mathbb{R}^m$ is injective,

then $f(x,.) : B_r^m \to \mathbb{R}^m$ is globally invertible and $f^{-1}(x,.)$ satisfies the Lipschitz condition with some constant $\mathcal{L}$.

(This second assumption is technically convenient, but it can be easily avoided.)

Now let, $A_1 \subset B_r^n$, $A_2 \subset \mathbb{R}^m$, $\delta > 0$ and $\Lambda = (\lambda_1, \ldots, \lambda_m)$, with $\lambda_1 \geq \ldots \geq \lambda_m \geq 0$ be given. We define $\Sigma(f, \Lambda, A_1, A_2, \delta)$ as the following set:

$$\begin{aligned}&\Sigma(f, \Lambda, A_1, A_2, \delta)\\ &\quad = \{(x,t) \in A_1 \times B_r^m ; \lambda_i(D_x f_{(x,t)}) \leqslant \lambda_i,\ 1 \leqslant i \leqslant m, f(x,t) \in [A_2]_\delta\}.\end{aligned}$$

The set $\Delta(f, \Lambda, A_1, A_2, \delta)$ is finally defined as $\pi_2(\Sigma(f, \Lambda, A_1, A_2, \delta))$, where $\pi_2 : B_r^n \times B_r^m \to B_r^m$ is the restriction of the standard projection. Thus $\Delta(f, \Lambda, A_1, A_2, \delta)$ consists of those parameters $t \in B_r^m$, for which there exists $x \in A_1$, with $\lambda_i(D_x f_{(x,t)}) \leqslant \lambda_i$ for all $i \in \{1, \ldots, m\}$ and $f_t(x) \in [A_2]_\delta$. We have the following Theorem:

Theorem 9.8 . *With the notations above, for any ϵ, $0 < \epsilon < \dfrac{\rho r}{2k}$, we have:*

$$M(\epsilon, \Lambda(f, \Lambda, A_1, A_2, \delta)) \leqslant c_1 \frac{M(\delta, A_2)}{\rho^m} \sum_{j=0}^{m} \lambda_0 \ldots \lambda_j (\frac{r}{\epsilon})^j (1 + \frac{\delta}{\epsilon} + \ldots + (\frac{\delta}{\epsilon})^{m-j}),$$

if $\epsilon \geq R_k(f)$, and we have

$$M(\epsilon, \Delta(f, \Lambda, A_1, A_2, \delta)) \leqslant$$

$$\frac{c_1 . c_2}{\rho^m} M(r', A_1) M(\delta, A_2)(1 + \frac{\delta}{r'})^m \times$$

$$\sum_{j=0}^{m} \lambda_0 \ldots \lambda_j (\frac{r}{\epsilon})^j (\frac{\epsilon}{R_k(f)})^{\frac{j}{k}} (1 + \frac{\delta}{\epsilon} + \ldots + (\frac{\delta}{\epsilon})^{m-j}),$$

if $\epsilon \leqslant R_k(f)$, where $c_2 = c_2\big(\max_{(x,t) \in B_r^n \times B_r^m} (||f(x,t)||, ||Df_{(x,t)}||, 1), \mathcal{L}\big)$ and $r' = r(\frac{\epsilon}{R_k(f)})^{\frac{1}{k}}$.

Proof. We can assume that A_2 is a point (the origin in $\mathbb{R}^m$). Indeed, we can cover $[A_2]_\delta$ by $M(\delta, A_2)$ balls of radius 2δ, thus $\Delta(f, \Lambda, A_1, A_2, \delta)$ is the union of $M(\epsilon, A_2)$ sets of type $\Delta(f, \Lambda, A_1, \{0\}, 2\delta)$, and its entropy is at most $M(\epsilon, A_2)$ times the entropy of $\Delta(f, \Lambda, A_1, \{0\}, 2\delta)$.

Let us consider first the case $R_k(f) \leqslant \epsilon \leqslant \frac{\rho r}{2k}$. We denote by P_p the Taylor polynomial of degree p of f at the origin of $\mathbb{R}^n \times \mathbb{R}^m$, where $k = p + \alpha$, $\alpha \in]0;1]$. We have established in the proof of Theorem 9.1 the two following inequalities:

$$||f(x,t) - P_p(x,t)|| \leqslant R_k(f) \leqslant \epsilon, \quad \text{for any } (x,t) \in B_r^n \times B_r^m \qquad \textbf{(5)}$$

$$||Df_{(x,t)} - DP_{p(x,t)}|| \leqslant \frac{pR_k(f)}{r} \leqslant \frac{p}{r}\epsilon, \quad \text{for any } (x,t) \in B_r^n \times B_r^m \qquad \textbf{(6)}$$

Thus, by (6), for any $(x,t) \in B_r^n \times B_r^m$, $\lambda_m(D_t P_{p(x,t)}) \geq \rho - \frac{p\epsilon}{r} \geq \frac{\rho}{2}$, since by assumptions, $\epsilon \leqslant \frac{\rho r}{2k} \leqslant \frac{\rho r}{2p}$.

On the other hand, for any $(x,t) \in \Sigma(f, \Lambda, A_1, \{0\}, \delta)$, $||P_p(x,t)|| \leqslant \delta + \epsilon$, by (5), and $\lambda_i(D_x P_{p(x,t)}) \leqslant \lambda_i' = \lambda_i + \frac{p}{r}\epsilon$, $i \in \{1, \ldots, m\}$,by (6).

Therefore, if $\Lambda' = (\lambda_1', \ldots, \lambda_m')$, then:

$$\Sigma(f, \Lambda, A_1, \{0\}, \delta) \subset \Sigma(P_p, \Lambda', A_1, \{0\}, \delta + \epsilon),$$

and

$$\Delta(f, \Lambda, A_1, \{0\}, \delta) = \pi_2(\Sigma(f, \Lambda, A_1, \{0\}, \delta))$$

$$\subset \pi_2(\Sigma(P_p, \Lambda', A_1, \{0\}, \delta + \epsilon)) = \Delta(P_p, \Lambda', A_1, \{0\}, \delta + \epsilon).$$

Now we use the following simplified version of Corollary 8.4: if in the assumptions of Corollary 8.4, $\rho_1 \geq \ldots \geq \rho_m \geq \rho > 0$, then assuming $A = \{0\}$, we have:

$$M(\epsilon, \Delta(P_p, \Lambda', A_1, \{0\}, \delta+\epsilon))$$
$$\leqslant \frac{C}{\rho^m}\sum_{j=0}^{m} \lambda'_0 \ldots \lambda'_j (\frac{r}{\epsilon})^j (1 + \frac{\delta+\epsilon}{\epsilon} + \ldots + (\frac{\delta+\epsilon}{\epsilon})^{m-j})$$
$$\leqslant \frac{\hat{C}}{\rho^m}\sum_{j=0}^{m} \lambda'_0 \ldots \lambda'_j (\frac{r}{\epsilon})^j \big(1 + \frac{\delta}{\epsilon} + \ldots + (\frac{\delta}{\epsilon})^{m-j}\big).$$

It follows that:

$$M(\epsilon, \Delta(f, \Lambda, A_1, \{0\}, \delta)) \leqslant M(\epsilon, \Delta(P_p, \Lambda', A_1, \{0\}, \delta+\epsilon))$$
$$\leqslant \frac{C_1}{\rho^m}\sum_{j=0}^{m} (\lambda_0 + \frac{k\epsilon}{r}) \ldots (\lambda_j + \frac{k\epsilon}{r})(\frac{r}{\epsilon})^j (1 + \frac{\delta}{\epsilon} + \ldots + (\frac{\delta}{\epsilon})^{m-j})$$
$$= \frac{C_1}{\rho^m}\sum_{j=0}^{m} (1 + \frac{\delta}{\epsilon} + \ldots + (\frac{\delta}{\epsilon})^{m-j}) \sum_{0 \leqslant j_1 \leqslant \ldots \leqslant j_\ell \leqslant j} \lambda_{j_1} \ldots \lambda_{j_\ell} (\frac{k\epsilon}{r})^{j-\ell} (\frac{r}{\epsilon})^j$$
$$\leqslant \frac{C_2}{\rho^m}\sum_{j=0}^{m} (1 + \frac{\delta}{\epsilon} + \ldots + (\frac{\delta}{\epsilon})^{m-j}) \sum_{\ell=0}^{j} \lambda_0 \ldots \lambda_\ell (\frac{r}{\epsilon})^\ell \quad (\text{ since } \lambda_1 \geq \ldots \geq \lambda_m)$$
$$= \frac{C_2}{\rho^m}\sum_{\ell=0}^{m} \lambda_0 \ldots \lambda_\ell (\frac{r}{\epsilon})^\ell \sum_{j=\ell}^{m} (1 + \frac{\delta}{\epsilon} + \ldots + (\frac{\delta}{\epsilon})^{m-j})$$
$$\leqslant \frac{c_1}{\rho^m}\sum_{\ell=0}^{m} \lambda_0 \ldots \lambda_\ell (\frac{r}{\epsilon})^\ell (1 + \frac{\delta}{\epsilon} + \ldots + (\frac{\delta}{\epsilon})^{m-\ell}),$$

This completes the proof of Theorem 9.8 in the case $\epsilon \geq R_k(f)$.

If $\epsilon \leqslant R_k(f)$, we define, as above, r' by $r' = r(\frac{\epsilon}{R_k(f)})^{\frac{1}{k}}$. Hence on each subball $B_{r'}$ of radius r', we have: $R_k(f_{|B_{r'}}) = \epsilon$. It remains to cover the set $\Sigma(f, \Lambda, A_1, \{0\}, \delta)$ by some balls of radius r' and to apply the first part of the Theorem to the retriction of f on these balls.

To count the number of r'-balls we need, first we cover $A_1 \subset B^n_r$ by $M(r', A_1)$ r'-balls $B^n_{r'}$ in $\mathbb{R}^n$. Then we count the number of r'-balls we need to cover the set $\{(x,t) \in B^n_{r'} \times B^m_r;\ f(x,t) \in B^m_\delta\}$, which contains of course the set $\Sigma(f, \Lambda, A_1, \{0\}, \delta) \cap B^n_{r'} \times B^m_r$.

Lemma 9.9. *The set* $\{(x,t) \in B^n_{r'} \times B^m_r;\ f(x,t) \in B^m_\delta\}$ *can be covered by at most* $C_3\big(M_1(f).\mathcal{L}\big)^m(1 + \frac{\delta}{r'})^m$ *r'-balls in* $B^n_r \times B^m_r$, *where* $M_1(f) = \max_{(x,t)\in B^n_r \times B^m_r} (||f(x,t)||, ||Df_{(x,t)}||, 1)$.

Proof of Lemma 9.9. It is enough to show that this set is contained in $B^n_{r'} \times B^m_{\delta'}$, with $\delta' = \mathcal{L}(2\delta + M_1(f)r')$.

If $(x_0, t_0) \in \{(x,t) \in B^n_{r'} \times B^m_r;\ f(x,t) \in B^m_\delta\}$, we have for any $x_1 \in B^n_{r'}$:

$$||f(x_0,t_0) - f(x_1,t_0)|| \leqslant M_1(f).r',$$

hence, if (x_1, t_1) is also in $\{(x,t) \in B^n_{r'} \times B^m_r;\ f(x,t) \in B^m_\delta\}$:

$$\begin{aligned}||f(x_1,t_1) - f(x_1,t_0)|| &\leqslant ||f(x_1,t_1) - f(x_0,t_0)|| + ||f(x_0,t_0) - f(x_1,t_0)||\\ &\leqslant 2\delta + M_1(f).r'\end{aligned}$$

and by the Lipschitz condition on $f^{-1}(x_1,.)$, we obtain:

$$\begin{aligned}||t_1 - t_0|| &= ||f^{-1}(x_1,.)[f(x_1,t_1)] - f^{-1}(x_1,.)[f(x_1,t_0)]||\\ &\leqslant \mathcal{L}(2\delta + M_1(f).r').\end{aligned}$$

□

Finally, the total number of r'-balls we need to cover $\Sigma(f, \Lambda, A_1, \{0\}, \delta)$ is at most:

$$M(r', A_1).C_3(M_1(f) \cdot \mathcal{L})^m.(1 + \frac{\delta}{r'})^m,$$

and thus:

$$\begin{aligned}&M(\epsilon, \Delta(f, \Lambda, A_1, \{0\}, \delta))\\ &\leqslant \frac{c_1 \cdot c_2}{\rho^m} M(r', A_1)(1 + \frac{\delta}{r'})^m \sum_{j=0}^{m} \lambda_0 \ldots \lambda_j (\frac{r}{\epsilon})^j (\frac{\epsilon}{R_k(f)})^{\frac{j}{k}} (1 + \frac{\delta}{\epsilon} + \ldots + (\frac{\delta}{\epsilon})^{m-j}).\end{aligned}$$

Theorem 9.8 is proved. □

The expression of Theorem 9.8 is rather complicated, so we will give below some simplified versions of this inequality.

First of all, substituting $\delta = 0$, we obtain the following result (assuming of course $A_2 = \{0\} \in \mathbb{R}^m$):

Corollary 9.10. *The set $\Delta(f, \Lambda, A_1, \{0\}, 0)$ of parameters $t \in B^m_r$, for which there exists $x \in A_1$, such that f_t at x is Λ-not-transversal to $0 \in \mathbb{R}^m$, satisfies:*

$$M(\epsilon, \Delta(f, \Lambda, A_1, \{0\}, 0)) \leqslant c.M(r', A_1) \sum_{j=0}^{m} \lambda_0 \ldots \lambda_j (\frac{r}{\epsilon})^j (\frac{\epsilon}{R_k(f)})^{\frac{j}{k}},$$

where c depends on $M_1(f)$, $\mathcal{L}$ and ρ. □

In particular, for $A_1 = B^n_r$, we get:

Corollary 9.11. *The set $\Delta(f, \Lambda, A_1, \{0\}, 0)$ of parameters $t \in B^m_r$, for which there exists $x \in B^n_r$, such that f_t at x is Λ-not-transversal to $0 \in \mathbb{R}^m$, satisfies:*

$$M(\epsilon, \Delta(f, \Lambda, B^n_r, \{0\}, 0)) \leqslant \widetilde{c}. \sum_{j=0}^{m} \lambda_0 \ldots \lambda_j (\frac{r}{\epsilon})^j (\frac{R_k(f)}{\epsilon})^{\frac{n-j}{k}},$$

where $\widetilde{c}$ depends on $M_1(f)$, $\mathcal{L}$ and ρ. □

This last expression is identical to the expression of Theorem 9.2. Thus if the "quantitativeness" of our transversality theorem is restricted to the measure of transversality of a given submanifold (but not of its δ-neighborhood), the conclusion is the following: the set of Λ-bad parameters t behaves exactly as the set of Λ-critical values in the quantitative Sard theorem. Of course, this is a "quantitavization" of a well-known relation, used in the standard proofs of the transversality theorem.

We give only one more result in this direction.

Corollary 9.12. *With the notations above, and for $\lambda_m = 0$:*

$$\dim_{\mathcal{H}}(\Delta(f,\Lambda,B_r^n,\{0\},0)) \leqslant \dim_e(\Delta(f,\Lambda,B_r^n,\{0\},0)) \leqslant m-1+\frac{n-m+1}{k}.$$

That is to say, the set of parameters $t \in B_r^m$, for which there exists $x \in B_r^n$ such that $f_t(x) = 0$ and $rank(Df_{t(x)}) \leqslant m-1$, has Hausdorff and entropy dimension $< m$, provided $k > n-m+1$. □

Now we consider some cases where $\delta \neq 0$. First of all, assuming $k = 1$ (in this special case f is a differentiable mapping with continuous derivative and $R_1(f) = K.r$, where K is such that for all (x_0,t_0) and (x_1,t_1) in $B_r^n \times B_r^m$, $||Df_{(x_0,t_0)} - Df_{(x_1,t_1)}|| \leqslant K$), and substituting $\epsilon = \delta$, we obtain the following result, generalizing Proposition 2.2. The difference is that now we do not assume that parameters act independently of x, as translations of $\mathbb{R}^m$.

Corollary 9.13. *With the notations above, and for λ_j such that $\lambda_j(D_x f_{(x,t)}) \leqslant \lambda_j$, for all $(x,t) \in B_r^n \times B_r^m$:*

$$M\big(\epsilon,\Delta(f,\Lambda,A_1,A_2,\epsilon)\big) \leqslant c'.M(\epsilon,A_1)M(\epsilon,A_2).$$

That is to say: the set of parameters $t \in B_r^m$, for which there exists x in A_1 such that $f_t(x) \in [A_2]_\epsilon$, has its ϵ-entropy bounded by $c'.M(\epsilon,A_1)M(\epsilon,A_2)$, where c' only depends on $M_1(f)$, $\mathcal{L}$, ρ.

Proof. It suffices to put $\lambda_j = \max_{(x,t)\in B_r^n\times B_r^m} ||Df_{(x,t)}||$, $\epsilon = \delta$, and to apply Theorem 9.8. □

In general the substitution $\epsilon = \delta$ in Theorem 9.8 gives a quite precise result, when the λ_j's are not bigger than $\max_{(x,t)\in B_r^n\times B_r^m} ||Df_{(x,t)}||$:

Corollary 9.14. *With the notations above:*

$$M\big(\delta,\Delta(f,\Lambda,A_1,A_2,\delta)\big) \leqslant c'.M(r',A_1)M(\delta,A_2)\sum_{j=0}^{m}\lambda_0\dots\lambda_j(\frac{1}{\delta})^{j-\frac{j}{k}}.$$

□

Let us assume now that $A_1 = B_r^n$ and $A_2 = \{0\}$. Then $M(r',A_1)$, for $\epsilon = \delta$ is of order $(\frac{1}{\delta})^{\frac{n}{k}}$, and we get:

$$M\big(\delta, \Delta(f, \Lambda, B_r^n, \{0\}, \delta)\big) \leqslant c' \sum_{j=0}^{m} \lambda_0 \dots \lambda_j (\frac{1}{\delta})^{j+\frac{n-j}{k}}.$$

Furthermore, we have $m - j - \dfrac{n-j}{k} \geq 1 - \dfrac{n-m+1}{k}$, for $j \leqslant m-1$ and $k \geq 1$, thus for $j \leqslant m-1$ and $k \geq n-m+1$ we obtain:

$$m - j - \frac{n-j}{k} \geq 1 - \frac{n-m+1}{k} \geq 0.$$

It follows from Corollary 9.14 and Proposition 2.6 that for any $\delta > 0$:

$$\begin{aligned}
\mathcal{H}^m\big(\Delta(f, \Lambda, B_r^n, \{0\}, \delta)\big) &\leqslant \mathcal{S}^m\big(\Delta(f, \Lambda, B_r^n, \{0\}, \delta)\big) \\
&\leqslant \delta^m M\big(\delta, \Delta(f, \Lambda, B_r^n, \{0\}, \delta)\big) \\
&\leqslant c''.(\delta^{1-\frac{n-m+1}{k}} + \gamma\delta^{-\frac{n-m}{k}}) = \mu(\delta, \gamma), \qquad (\mathbf{7})
\end{aligned}$$

where c'' depends only on $M_1(f), \rho, \mathcal{L}, m,$, assuming $\lambda_1 = \ldots = \lambda_{m-1} = M_1(f)$ and $\lambda_m = \gamma$.

In particular, for any $k > n-m+1$, $\mu(\delta, \gamma) \longrightarrow 0$ as δ and γ tend to 0, and $\gamma \leqslant \delta^{\frac{n-m}{k}+\xi}$, $\xi > 0$.

Hence we have the following result:

Theorem 9.15. *In any set $E \subset B_r^m$, with $\mathcal{H}^m(E) > \eta > 0$, and for any δ, γ, with $\mu(\delta, \gamma) < \eta$, there is a value $t_0 \in E$ of the parameter t, such that for any $x \imath B_r^n$, if $||f_{t_0}(x) \leqslant \delta||$, then $\lambda_m(Df_{t_0(x)}) \geq \gamma$.* □

Inequality (7) can be applied in various situations. We give here only the following corollary:

Corollary 9.16. *In any ball $B_\sigma^m \subset B_r^m$, there is t_0 such that f_{t_0} is δ-transversal to $0 \in \mathbb{R}^m$ (i.e. at each x where $||f_{t_0}(x)|| \leqslant \delta$, $\lambda_j(Df_{t_0(x)}) \geq \delta$, for $j \in \{1, \ldots, m\}$), for $\delta = \left(\dfrac{\mathcal{H}^m(B_1^m)}{2c''}\right)^{\frac{k}{k-n-m+1}} \sigma^{\frac{mk}{k-n-m+1}}$.*

Proof. Substituting $\gamma = \delta$ in inequality (7), we obtain:

$$\mu(\delta, \delta) = c''(\delta^{1-\frac{n-m+1}{k}} + \delta^{1-\frac{n-m}{k}}) < 2c''\delta^{1-\frac{n-m+1}{k}},$$

hence for $\delta = \left(\dfrac{\mathcal{H}^m(B_1^m)\sigma^m}{2c''}\right)^{1/(1-\frac{n-m+1}{k})}$, $\mu(\delta, \delta) < \mathcal{H}^m(B_1^m)\sigma^m = \mathcal{H}^m(B_\sigma^m)$.

□

The general expression of Corollary 9.14 gives an interesting information for "small" sets A_1. Roughly, the point here is that the entropy of A_1 appears in this expression for the radius r' of covering balls ($r' \sim \epsilon^{\frac{1}{k}}$). We state here the corresponding results only in the case of the Sard theorem, i.e. for

$f(x,t) = g(x)+t$. Then $\Delta(f, \Lambda, A, \{0\}, 0)$ is exactly the set of Λ-critical values of g on the set A, and we have:

Theorem 9.17. *Let $f : B_r^n \to \mathbb{R}^m$ be a $\mathcal{C}^k$-smooth mapping, and $\Delta(f, \Lambda, A)$ be the set of Λ-critical values of f on a subset $A \subset B_r^n$. Then for $0 < \epsilon \leqslant R_k(f)$:*

$$M(\epsilon, \Delta(f, \Lambda, A)) \leqslant c.M(\epsilon^{\frac{1}{k}}, A) \sum_{j=0}^{m} \lambda_0 \dots \lambda_j (\frac{r}{\epsilon})^j (\frac{\epsilon}{R_k(f)})^{\frac{j}{k}}.$$

□

As a special case (for $A = B_r^n$) we obtain once more the result of Theorem 9.2.

In order to stress the difference, consider only the consequence, concerning the entropy dimension. Assuming that $A \subset \Sigma_f^\nu$ (the rank-ν set of critical points of f, i.e. the set of x such that $\operatorname{rank}(Df(x)) \leqslant \nu$)), we get:

$$M(\epsilon, \Delta_f^\nu) \leqslant c.M(\epsilon^{\frac{1}{k}}, A)(\frac{1}{\epsilon})^{\nu(1-\frac{1}{k})},$$

generalizing to $\mathcal{C}^k$-smooth mappings Corollary 8.5.

In particular, when $M(\epsilon, A) \leqslant (\frac{1}{\epsilon})^\alpha$, $M(\epsilon, \Delta_f^\nu) \leqslant c(\frac{1}{\epsilon})^{\frac{\alpha}{k}+\nu(1-\frac{1}{k})}$, and we have proved the following result, which extends Corollary 8.6 to $\mathcal{C}^k$-mappings:

Theorem 9.18. *Let $f : B_r^n \to \mathbb{R}^m$ be a $\mathcal{C}^k$-smooth mapping. If $A \subset B_r^n$ is contained in Σ_f^ν, then*

$$\dim_e(f(A)) \leqslant \frac{1}{k} \dim_e(A) + \nu(1 - \frac{1}{k}).$$

In particular, for $\nu = 0$:

$$\dim_e(f(A)) \leqslant \frac{1}{k} \dim_e(A).$$

□

Remark. It has been proved, in [Com 1], that, if $A \subset \Sigma_f^0$, $\mathcal{H}^s(A) = 0$ implies $\mathcal{H}^{\frac{s}{k}}(f(A)) = 0$, and thus that $\dim_{\mathcal{H}}(f(A)) \leqslant \frac{1}{k} \dim_{\mathcal{H}}(A)$. Consequently, Theorem 9.18 has an analogue for Hausdorff dimension.

Theorem 9.19. *Let $f : B_r^n \to \mathbb{R}$ be a $\mathcal{C}^k$-smooth function, if $A \subset B_r^n$ is a set of critical points, then:*

$$\dim_e(f(A)) \leqslant \frac{1}{k} \dim_e(A) \quad \textit{and} \quad \dim_{\mathcal{H}}(f(A)) \leqslant \frac{1}{k} \dim_{\mathcal{H}}(A).$$

□

10 Some Applications and Related Topics

In this chapter we briefly describe some further applications of the results and methods developed in the preceeding chapters.

Y. Yomdin and G. Comte: LNM 1834, pp. 131–169, 2004.

10.1 Applications of Quantitative Sard and Transversality Theorems

10.1.1 Maxima of smooth families

Let $h : B_r^n \times B_s^m \to \mathbb{R}$ be a continuous function. The function h can be considered as a family of functions $h_y : B_r^n \to \mathbb{R}$, $y \in B_s^m$. The (pointwise) maximum function m_h of this family is defined as

$$m_h(x) = \max_{y \in B_s^m} h(x, y)\ ,\ x \in B_r^n .$$

Functions of this form arise naturally in many problems of calculus of variation, optimization, control, etc. If we assume h to be $\mathcal{C}^k$ or analytic, it does not imply m_h to be even once differentiable. (Notice, however, that for h a polynomial or a semialgebraic function, m_h is semialgebraic, for h analytic m_h is subanalytic and for h a tame function, m_h is tame). Lipschitz constant is preserved by taking maxima.

Understanding the analytic nature of maxima functions is an important and mostly open problem. Besides its theoretical aspects, it presents a challenge in numerical applications: lack of smoothness of maxima functions prevents using many standard algorithms and packages. See, for example, [Roc], [Shap-Yom], where some partial results, questions and references can be found.

Not surprisingly, critical and near critical points and values of the family h can be traced in the structure of m_h. We give here a couple of examples, where nontrivial geometric restrictions on m_h are imposed by the Sard theorem. (Maxima of smooth families can be naturally studied also in the framework of Singularity Theory. Classification of their typical singularities can be found in [Bry], [Dav-Zak], [Mat], [Yom 11,12,25]).

Let f be a convex function of n variables. It can be always represented as a maximum function of a family of (supporting) linear functions. The question is what restrictions on f are imposed by the assumption that these linear functions smoothly depend on the parameter.

By the Alexandrov-Fenchel theorem, the graph of f has a well defined second fundamental form $W(x)$ at almost each of its points x. Denote by S_i the set of x for which the rank of $W(x)$ is at most i, and by D_i the image of S_i under the Gauss mapping (i.e. the set of the endpoints of the unit normal vectors to the graph of f at the points of S_i).

Theorem 10.1. *For f representable as a maximum function of a k-smooth m-parametric family of linear functions, the entropy dimension of D_i does not exceed $i + (n + m - i)/k$.*

In particular, for $n = 1$ and f having as a graph a convex polygon with an infinite number of faces, the representation as above implies that the box dimension of the normals to the faces in the unit circle does not exceed $(m + 1)/k$.

These results can be found in ([Yom 8,9,14,15]).

Another approach is based on the study of critical points and values of the maxima functions themselves. An important fact is that critical (and near-critical) points and values can be defined for any Lipschitz function f, using the notion of Clarke's generalized differential (see [Fed 2], [Cla 1,2] and [Roc]). Assuming that f is representable as a maximum of a smooth family, a version of the quantitative Sard theorem can be proved for f (see [Yom 15,24] and [Roh 1-3]). We do not state here these results, mentionning only that Theorem 4.2 of [Yom 24] implies that the entropy dimension of the critical values of a maximum function f of an m-parametric k-smooth family does not exceed $(n+m)/k$.

This provides another restriction on representability. For example, the function $x^q cosine(1/x)$ cannot be represented as a maximum of any k-smooth m-parametric family of functions for k greater than $(q+1)(m+1)$.

Undoubtedly, the examples above represent only the simplest manifestations of the interplay between the singular geometry of families and the analytic structure of their maxima functions. Much more can be conjectured with various strength of evidence. Some of these conjectures and observations can be found in ([Shap-Yom], [Yom 8,9,11,12,14-17,20,24]) and in Section 10.2 below.

10.1.2 Average topological complexity of fibers

Let $f : M^m \to N^n$ be a C^k mapping of compact manifolds. The following implication of the usual Sard theorem is by far the most frequently used in differential topology: for almost any y in N the fiber $f^{-1}(y)$ is a compact smooth submanifold of M.

A natural question is then what is a typical topological complexity of such a fiber (in particular, how many connected components may it have)? The usual Sard theorem gives no information of this type. On the other hand, the Quantitative Sard theorem, proved above, allows us to give explicit upper bounds for an average of Betti numbers of the fibers.

Let for y in N, $B_i(y)$ denote the i-th Betti number of the fiber $f^{-1}(y)$. We assume that the smoothness k of f is greater than $s = n - m + 1$. Then by Sard's theorem $B_i(y)$ is finite almost everywhere in N.

Theorem 10.2. *For any q between zero and $k - s/n$ the average of $(B_i(y))^q$ over N is finite.*

The proof of this theorem is given in [Yom 3]. It is based on an estimate of the average distance from a point in N to the set of critical values of f. This estimate is provided by the Quantitative Sard theorem.

Many additional results of the same spirit are obtained in [Yom 3]: existence of "simple" fibers in any subset of N of a given positive measure, average bounds on volume of fibers etc. For M and N euclidean balls, explicit bounds are given in [Yom 3] for all the above quantities, in terms of the bounds on the derivatives of f.

Also here, much more information can be presumably extracted by similar methods: estimates for "bounded triangulations" of the fibers, more delicate estimates of the geometry, including upper bounds for curvatures, estimates for the spectrum of certain differential operators on these fibers (compare Gromov's paper [Gro 4]).

Till this point all the estimates in this section concerned a "typical" level (fiber) of a mapping, or an "average" behavior of the fibers. There is another important effect, related to all the level sets of "near-polynomial" differentiable functions. It turns out that if a differentiable function f on the unit ball B^n is sufficiently close to a polynomial of degree $k-1$ (i.e. if its partial derivatives of the order k are sufficiently small with respect to its C^0-norm), then *all its level sets* resemble algebraic varieties of degree $k-1$.

More accurately, it was shown in [Yom 2] that *for f on B^n k times continuously differentiable, and with the C^0-norm equal to 1, if the norm of the k-th derivative of f in B^n is bounded by 2^{-k-1}, then the set of zeroes $Y(f)$ of f is contained in a countable union of smooth hypersurfaces. "Many" straight lines intersect $Y(f)$ in at most $k-1$ points, and the $n-1$-dimensional Hausdorff measure of $Y(f)$ is bounded by a constant $C(n,k)$, depending only on n and k.*

This is in a strict contrast with the fact that any closed set in B^n may be the set of zeroes $Y(f)$ of an infinitely differentiable f, if we do not assume restrictions on the derivatives.

We hope that many of the results of this book, concerning the ϵ-entropy of semialgebraic sets and its behavior under polynomial mappings, can be extended to "near-polynomial" functions. This approach presumably can be applied in one important problem in Analysis, namely, description of the "Nodal Sets" of the eigenfunctions of elliptic operators (see [CBar 1,2], [Donn-Fef 1,2], [Han-Har-Lin], [Har-Sim], [Hof-Hof-Nad]). Recently, the results of [Yom 2] have been applied in this context in [CBar 1].

10.1.3 Quantitative Kupka-Smale Theorem

A theorem of Kupka and Smale ([Kup 1], [Sma 1]) asserts that all the periodic points of a generic diffeomorphism, or closed orbits of a generic flow, are hyperbolic (i.e. no eigenvalue of the linearization of the mapping along the orbit sits on the unit circle). The proof consists in a (rather delicate, as many things in dynamics) application of a transversality theorem.

Quantitative transversality, obtained above, was used in [Yom 4] to get a much more precise result: estimates of hyperbolicity, and not only for closed, but for "almost closed" orbits.

More accurately, let $f : M^m \to M^m$ be a C^k diffeomorphism of a compact smooth m-dimensional manifold M. Let $d(x, y)$ be a certain distance on M. The point x in M is called (n, δ)-periodic, for n a positive integer and a certain positive real number δ, if $d(x, f^n(x)) \leqslant \delta$. The point x is called (n, γ)-hyperbolic, if all the eigenvalues of the differential $Df^n(x)$ are at a distance at least γ from the unit circle.

Theorem 10.3. *Let $k \geq 3$. There is a dense subset W in the space of all $\mathcal{C}^k$ diffeomorphisms of M with the following property:*
For any g in W there are positive constants a and b, depending on g, such that for any positive integer n any (n, a^{n^α})-periodic point of g is (n, b^{n^α})-hyperbolic.
Here $\alpha = \alpha(m, k) = \log_2(m^2 + mk + k + 1)$.

Hyperbolicity implies local uniqueness of solutions of the equation $f^n(x) = x$. In turn, quantitative hyperbolicity provides a bound on the distance between any two solutions. We get:

Corollary 10.4. *For any g in W there are positive constants c and C, such that:*

$i-$ The distance $d(x_1, x_2)$ between any two exactly periodic points x_1, x_2 of period n is not smaller than c^{n^α}.

$ii-$ The number of n-periodic ponts of g does not exceed C^{n^α}.

In fact, one can prove that for any given $\mathcal{C}^k$ diffeomorphism f of M and for any positive ϵ, there exists g ϵ-close to f in the C^k-norm, with properties as above. The constants a, b, c and C are given explicitly in terms of the bounds on the derivatives of f and ϵ.

These results give only the first step in understanding the quantitative behaviour of periodic points under perturbations of a diffeomorphism. The main open problem is whether one can replace the overexponential bounds above by exponential ones. (Indeed, our $\alpha(m, k)$ above is always greater than one. Its first values are $\alpha(1, 3) = 3.585...$, $\alpha(2, 3) = 4.585...$ etc.).

On the other hand, the theorem of Artin and Mazur [Art-Maz] guarantees the exponential growth of the number of periodic points with the period for a dense set of diffeomorphisms (in fact, for polynomial ones, with respect to a certain realization of M as a real algebraic manifold).

The main difficulty, which prevents obtaining exponential growth in the proofs in [Yom 4] is of a dynamical nature: iterations of a periodic point can come close to one another. This makes the control of the influence of perturbations very difficult. (The same difficulty manifests itself in many other dynamical questions, like the "Closing lemma").

There are examples, showing that while the first order control in situations as above is hopeless, it still may be regained in high order derivatives. A recent result of Grigor'ev and Yakovenko [Gri-Yak] which provides a transversality statement in a multijet situation, may also be relevant (as it indeed seems to be in Kaloshin's results).

The understanding of all the situation around the behavior of periodic and almost periodic points and trajectories was recently changed dramatically by striking results of Kaloshin ([Ka 1-3], [Ka-Hun]), who showed that any prescribed growth rate can occure generically on one hand, and that the exponential bound holds for "prevalent" diffeomorphisms, on the other hand!

10.1.4 Possible Applications in Numerical Analysis

10.1.4.1 Summary. Virtually any numerical algorithm in Nonlinear Analysis requires for its proper work that some Jacobians be nonzero. The usual Sard theorem essentially claims that this is the case for randomly chosen data. Some popular numerical algorithms, such as continuation methods for solving nonlinear systems of equations, explicitly involve a random choice of parameters or of a starting point, appealing to the Sard theorem as a justification (see, for example, [All], [Shu-Sma 1-7], [Sma 3,4]).

However, the role of this theorem is restricted to an a posteriori confirmation of the efficiency (experimentally well known) of the algorithms. It gives no specific recommendations on how to organize such algorithms or how to optimize the choice of the parameters involved.

In an important special case of solving systems of nonlinear algebraic equations, a thorough theoretical study of the efficience of the global Newton method was given by M. Shub and S. Smale ([Shu-Sma 1-7]). It involved a "quantitative" investigation of the degree of non-degeneracy of the data on each step of the algorithm. Although this investigation has been performed mostly without an explicit application of the Morse-Sard theorem, it stressed the importance of the "quantitative" information on the distribution of singularities.

In most of non-linear algorithms Quantitative Sard theorem and Quantitative Transversality results proved in Chapters 7–9 above allow us, at least in principle, to give explicit recommendations on how to organize these algorithms or how to optimize the choice of the parameters. In general, we believe that the importance of "Quantitative Singularity Theory" in practical computations, from Motion Planning and solving PDE's to Image Processing and Computational Biology, will grow. We hope that this field will become one of important applications of the methods presented in the book. This justifies in our view, a somewhat lengthy discussion of these topics in the Introduction, section 1-3, in this section, and in Section 10.3.7 below.

10.1.4.2 How to Treat Singularities? The stability, accuracy and running time of any numerical algorithm is determined by its "well-posedness", i.e. by the separation from zero of the involved determinants. Consequently, the overall running time, as well as the accuracy, of a non-linear numerical algorithm involving "near-singularities" can be estimated in terms of the lower bound on the near-degenerated Jacobians at these points. In turn, the Quantitative Sard theorem provides an explicit probability distribution for these

Jacobians. Thus, if the numerical scheme involves an adaptation of the parameters, according to the degree of degeneracy of the data (this is a typical case), the switching thresholds can be optimized to provide the best average performance. A purely illustrative example of this sort was given in [Yom 18].

On the other hand, there are many examples of non-linear numerical algorithms, where not only an adaptation of the parameters, according to the degree of degeneracy of the data is performed, but rather a switching from one algorithm to another, when singularities are encountered. Indeed, in most of practical problems singularities of different degrees of degeneracy cannot be avoided. Of course, these singularities can be treated just by increasing resolution and accuracy of the processing (adaptation of the parameters). However, a much more rewarding approach is to acknowledge the presence of a singularity and to treat it explicitly, via the methods and tools provided by an analytic study of the background problem. One of the most powerful such tools is a "Normal Form" of a singularity.

10.1.4.3 Normal Forms. In each specific problem, the mathematical objects involved (functions, surfaces, differential equations) are usually considered up to a certain equivalence relation, provided by a certain group of allowed transformations.

Normal Form is "the simplest" representative in each equivalence class of objects (for example, containing the minimal possible number of free parameters).

In no way were normal forms invented in Singularity Theory. The idea to bring a mathematical problem to its simplest possible form by changes of variables and other permitted transformations is certainly one of the most powerful and oldest discoveries in Analysis and in Mathematics in general.

Singularity Theory provides a general approach and techniques for finding normal forms of singularities of smooth mappings and other related objects. It stresses the importance of this notion in many theoretic and applied fields, extending far beyond the classical examples. Moreover, Singularity Theory provides a unified way to find in each specific problem its "hierarchy of singularities", according to their degree of degeneracy, and the corresponding hierarchy of normal forms. See [Tho 3], [Boa], [Arn-Gus-Var], [Gol-Gui], [Ma 1-8] for general methods of Singularity Theory and many specific results, [Zh] for a description of the normal forms of differential 1-forms and Pfaffian equations, [Bry], [Dav-Zak], [Mat], [Yom 26] for a description of the hierarchy of singularities of maximum function, and [Guck], [Dam 1-4], [Rie 1,2] for a treatment of singularities in solutions of PDE's and in Imaging. We give only a small fraction of the relevant literature.

One of the most important features of the normal form approach, especially relevant for numerical applications, is the following: *while the processing of singular data is usually ill-posed, bringing singularity to its normal form and finding the corresponding normalizing transformations is usually well-posed.*

Normal forms have been successfully applied *in the process of computations* in many numerical algorithms in various fields of applications. Much less was done with normal forms of singularities, considered as the *basic element of data representation.*

We believe that really efficient computational methods in many nonlinear problems must be based on a coordinated use of the hierarchy of singularities and of the corresponding hierarchy of their normal forms *both in data representation and in data processing.*

10.1.4.4 Flexible High Order Discretization. A unified approach to high-order data representation and processing, based on this principle, has been proposed and initially tested in [Eli-Yom 1-5], [Bri-Eli-Yom], [Bic-Yom], [Wie-Yom], [Koch-Seg-Yom 1-2], [Y-E-B-S], [Bri-Yom 6], [Yom 27,28]. We have started discussing this approach (as related to motion planning in robotics) in Section 1.1.4 of the Introduction. Let us give here a little bit more general description of this framework, since we believe it presents an important ground for application of the results and methods of the book.

Kolmogorov's Representation of Smooth Functions

First of all, if we want to use analytic methods of Singularity Theory *in the process of computations*, we have to provide a direct access to the high order derivatives of the data. Indeed, all the computations in Singularity Theory, related to the classification of singularities and to the procedure of bringing them to their normal forms, use relatively high order Taylor polynomials ("Jets") of the functions involved.

Computing high order derivatives from the conventional grid representation (where only the value of the function is stored at each grid point) is very unstable and unreliable. Consequently, we are forced to *keep explicitly high order derivatives at each grid point.* In an equivalent way, we say that a Taylor polynomial of a prescribed degree k (or a k-jet) is stored at each grid-point. This is a highly non-orthodox decision from the point ov view of the traditional numerical analysis. Fortunately, it is supported by the Kolmogorov theory of optimal representation of smooth functions (see [Kol], [Kol-Tih], [Tih 2] and many other publications).

The question of an optimal representation of smooth functions has been investigated by Kolmogorov in his work on ϵ-entropy of functional classes. The problem can be shortly described as follows: *How many bits do we need to memorize a C^k-function of n variables up to a prescribed accuracy $\epsilon > 0$?* Mathematically, this is exactly the question of computing (the logarithm of) the ϵ-entropy of the class of C^k-functions, with respect to the C^0-norm (see Chapter 2 above).

It was shown in [Kol], [Kol-Tih] that asymptotically, the best way to memorize a C^k-function up to the accuracy $\epsilon > 0$ is to store the coefficients of its k-th order Taylor polynomials at each point of some grid with the step

$h = O(\epsilon^{1/(k+1)})$. The corresponding estimate for the ϵ-entropy of the class of C^k-functions in the space C^0 is a double exponent with the upper term of order $(\frac{1}{\epsilon})^{\frac{n}{k}}$.

One of the main trade-offs in any numerical approach, based on a grid representation of the data, is between the density of the grid versus the processing complexity at each grid-point. Kolmogorov's representation tends to increase as far as possible the analytic power and flexibility of the local data representation at each grid-point, strongly expanding in this way this grid-point's "covering area".

As a result, a density of the grid can be strongly reduced, while preserving the required approximation accuracy. This reduction may lead to a major efficiency improvement, especially in the problems with the large number of unknowns and parameters.

Let us give a simple (and purely illustrative) example. Assume we have to approximate a function f of 10 variables on the unit cube Q, with the accuracy of $\epsilon = 0.01$. We use a uniform grid in Q with the step-size h and a Tailor polynomial approximation at each grid-point. Assuming that the derivatives of f up to the third order, are bounded by 100, we get according to the Tailor remainder formula, that the accuracy of the first order Tailor polynomial approximation within the distance h from the grid-point is $100h^2$, while the accuracy of the third order approximation is $10h^4$. Hence to get a required overall approximation accuracy of $\epsilon = 0.01$, we must take $h = 0.01$ in the first case and $h = 0.16$ in the second case. The size of the "covering area" of each point increases more than ten times. Hence the number of the required grid-points in the 10-dimensional cube Q drops 10^{10} times. On the other side, the complexity of the local representation and processing at each grid-point is roughly the number of the stored coefficients of the Taylor polynomial. For the third degree Taylor polynomial it is of order 200, while for the first degree Taylor polynomial it is 11. The 1:20 jump in local complexity is more than compensated by the 10^{10} reduction in the number of grid-points.

Of course, one will try not to work neither with 100^{10} nor with 10^{10} grid-points. It would be a clever decision also to avoid, if possible, a straightforward computation at each grid-point of a polynomial with 200 coefficients. Certainly, in a practical treatment of any problem with a 10-dimensional phase space all the attempts will be made to use methods not involving a scanning over its total definition domain.

Still, scanning over *some important areas in the phase-space* may be unavoidable. This is especially true in the practically important problems of the motion planning in robotics and in the qualitative-quantitative study of large systems of ODE's appearing in biology, with the goal to describe the "dynamical" effects like closed (periodic) trajectories, attractors, influence of the main parameters, etc. In each of these problems a clever use of the high order representation may give potentially a major improvement in efficiency.

So we suggest to use in practical computations the Kolmogorov representation of smooth functions. It is very compact, it allows for application *at each grid-point of analytic processing methods*, and it is compatible with the "Normal Form Representation".

The main drawback of this representation is its very high sensitivity to the noise in the numerical data. This is a major problem, but it can be solved by the "multi-order" strategy, shortly discussed below in this section.

"Non Conforming" Representation

The following feature of the Kolmogorov representation is very important in our approach. We assume that a required accuracy or tolerance $\epsilon > 0$ is given from the very beginning. We require the discretized data to *represent* the "true" function up to this accuracy in a C^k-norm. But we do not require the discretized data *to be itself* a C^k-function. In particular, the Taylor polynomials at the neighboring grid points *may disagree up to* ϵ.

This makes our representation very flexible. In particular, *we make no effort to subdivide the domain into pieces where each of the polynomials is used or to adjust their values (and/or the values of their derivatives) at the boundaries of such pieces.*

(Notice that this subdivision and adjustment is by far the most complicated part in many high order algorithms. Moreover, it introduces into the solution process a rigid combinatorial-geometric structure, which has nothing to do with the original problem to be solved).

In some applications, such as Computer Graphics and Engineering, it is important to be able to produce ultimately a truly continuous or smooth function from the discretized data. Mathematically, this is an interpolation problem (or a problem of extending a C^k function from a given subset (grid) to the whole space). As many other problems in high order processing, this extension problem is not easy. It was studied by Whitney ([Whi 5,6]) and others. Recently an important progress has been obtained in [Bie-Paw-Mi 2]. Also certain multivariate interpolation methods can be used (see, in particular, [Mic-Mi]). These ideas can be used to produce an efficient "numerical smoothing" of the high order data in the Kolmogorov representation (see [Bic-Yom], [Wie-Yom], [Yom 27]).

Absence of the rigid adjustment of the neighboring polynomials makes the Kolmogorov representation relatively insensitive to the precise geometry of the grid. The grid-points do not need to be the vertices of a regular partition of the domain. If necessary, they can "move" freely, just approximately preserving the mesh-step. The precise geometry of the grid is involved only in the computation of the optimal relaxation (or interpolation) coefficients, which is performed once forever, before the actual data processing.

Also near the boundary no adjustment of the grid geometry is necessary. Using the calculus of Taylor polynomials ("Jet calculus" – see below), we incorporate the boundary data into the structure of the Taylor polynomials at the grid points near the boundary.

Processing of Functions in Kolmogorov's Representation

An important requirement for a discretization scheme used in computations is that any analytic operation or functional, as applied to a discretized function, can be expressed (up to the prescribed accuracy ϵ) in terms of the discretization data. Kolmogorov representation satisfies this requirement. Moreover, many important analytic operations can be implemented much more efficiently with the explicit high order data at each grid-point. This is because having an access to the high-order derivatives allows for application of powerful analytic procedures, like the implicit function theorem, inversion of mappings, recurrent computation of Taylor coefficients, Runge-Kutta calculations for solving ODE's, and similar formulae in other problems. These operations form a "Jet-calculus", widely used in many areas of mathematics, but probably represented in the most coherent form in Singularity Theory (see, for example, [Arn-Gus-Var], [Ma 1-8], [Gol-Gui], [Zh]). See also [Eli-Yom 1-5], [Wie-Yom], [Bri-Eli-Yom], Bri-Yom 6] and [Yom 27] for some specific examples of applications of the "Jet-calculus".

It is important to stress, that the operations and formulae in the Jet-calculus play in our approach the same role as the standard arithmetic operations in the usual computations. They form a "Jet-calculus library" which is assumed to ultimately include any important analytic operation, expressed in a jet language. In any practical realization all the relevant Jet-calculus operations and formulae should be implemented in the most efficient way, exactly as the the standard arithmetic operations are implemented in the computer processors.

Multi-Order Strategy

As it was mentioned above, the main drawback of the Kolmogorov representation is a very high sensitivity of the high order derivatives to the noise in the numerical data. As a result, it is not easy to produce in a reliable way such a representation in most of practical problems. Also computations with the explicit high-order data are far from being robust. They involve a mixing of several scales and as a result a division by high degrees of ϵ. This is a major problem, and it is this problem that prevents a use of Kolmogorov-like representations in the conventional numerical algorithms. However, the advantages of this type of representations are so important, that in our opinion they justify a development of the new methods which will allow one to overcome the instability of the explicit high-order data processing.

The "multi-order" strategy in many cases solves the problem. It consists in a successive processing of the data: from the lowest order to higher and higher ones. At each order the maximal possible for this order accuracy is achieved, and then the next order data is included as a "correction" to the previous step. In many important situations this approach separates the scales and strongly reduces the stiffness of the equations to be solved. In particular, it excludes a necessity to divide the data by high degrees of ϵ. See [Wie-Yom] for one specific implementation of the multi-order approach in solving elliptic PDE's and [Yom 27] for its more general mathematical treatment.

As the processing of the noisy empirical data is concerned (like digital images) it is difficult to expect an overall smoothness. Here the multi-order strategy is used also in order to discover a "hidden regularity" in the data (which is closely related to the presence of singularities and to their Normal Form representation). Consequently, in the implementations related to noisy empirical data (like in Image Processing), the multi-order approach includes an additional ingredient: the higher order data is used not everywhere, but only in cases we the lower order analysis predicts a reliability of the higher order information. See [Bri-Eli-Yom] and [Bri-Yom 6].

10.1.4.5 Flexible High Order Discretization: a General Structure. In general, we combine the Kolmogorov representation of the smooth components of the data with the Normal Form representation of the singularities. The "geometric support" Z of the singularities (or the critical set) is explicitly memorized (Z can be called also the singular set of the data). Next, at each point of a certain grid in a vicinity of Z the following data is stored: the normal form of the local singularity and the coordinate transformations, (or other allowed transformations) which bring this singularity to its normal form (the "normalizing transformations"). Notice that in most cases the list of the normal forms used is simple: there are several discrete types of normal forms, and for each type, a small number of free parameters. Let us remind once more that *while the processing of singular data is usually ill-posed, bringing singularity to its normal form and finding the normalizing transformations is well-posed.* Besides the improved stability and accuracy in producing our representation from a noisy data, this fact has another important consequence: our data size is usually strongly reduced in comparison with the original data size.

10.1.4.6 Implementation Examples. The proposed approach has been experimentally tested in several directions. In addition to the specific publications mentioned below, we plan to present a general description of the underlying theory, of implemented algorithms as well as the experimental results in [Yom-Bri 6], [Yom 27]. Here we just mention shortly the main investigated problems.

Solving PDE's. Gromov's h-homotopy as a Numerical Method

Our implementation of the general method described above to solving elliptic PDE's is based on Kolmogorov's representation of smooth functions on the one hand, and on Gromov's approach to solving PDE's (differential relations, h-homotopy – see [Gro 5]) on the other. This method has been proposed in a general form in [Eli-Yom 1] and partially implemented in [Koch-Seg-Yom 1-2] and [Wie-Yom]. It consists of several steps.

1. All the functions involved (known and unknown) are represented by their Taylor polynomials of a fixed degree k at all the nodes of a certain fixed grid.

2. The Taylor polynomials for the unknown functions (whose coefficients form the basic set of the unknowns) are a priori parameterized to satisfy locally the PDE to be solved. For example, for the equation $\Delta u = 0$, harmonic polynomials are used at each grid point. The boundary conditions are incorporated into the parameterization of the Taylor polynomials at the grid points near the boundary.

However, for the parameter values, picked at random, the above Taylor polynomials satisfying the equation do not agree with one another. In Gromov's terminology [Gro 5], they form a nonholonomic section of a differential relation and do not represent a true solution of the equation. Thus the solution process consists of finding the values of the unknown parameters which minimize the discrepancy between the neighboring Taylor polynomials. This approach can be considered as a discretized realization of Gromov's h-homotopy. In such a terminology the standard methods use a true function which approximately satisfies the differential relation. Our method uses objects which are not true functions but "satisfy" the differential relation exactly at each grid-point.

3. We implement the last stage of the solution as a certain relaxation procedure where the Taylor polynomial at each node is corrected according to the neighboring ones. The mere presence of several Taylor coefficients at each node (instead of the only one in standard schemes) allows one to find relaxation coefficients which "cancel" the discretization error of the solution up to an order m which is much higher than k. For example, for $\Delta u = 0$ for the second degree Taylor polynomials, we get for an internal node the discretization error of order h^{10} where h is the step of the grid.

4. At the previous stage we got, at each grid point, a Taylor polynomial of degree k which agrees with the Taylor polynomial of the true solution (at this point) up to order $m > k$. It turns out that from this data one can usually reconstruct at each grid point the m-th degree Taylor polynomial of the true solution, with the same accuracy. For this reconstruction the same neighboring nodes as those in the relaxation steps can be used.

We've achieved a very high order of the discretization error (10, for the degree two Taylor polynomials and the Laplace equation), as compared to the size of the processed data, while preserving a stability of the classical methods. This is a general feature of the suggested algorithms. It is explained by the following two facts:

1. We use at each grid point jets which satisfy the initial differential equation. This reduces strongly the number of free parameters. For example, for two independent variables, general jets of degree k contain $\frac{(k+1)(k+2)}{2}$ parameters. Jets of degree k, satisfying a linear PDE with constant coefficients of order 2, have $2k+1$ free parameters.

(Notice that the requirement for local representing elements to satisfy the initial differential equation is usually not compatible with the precise adjustment of the values and the derivatives for the neighboring elements (since the last requirement leads to elements with compact support)).

2. We find a relaxation scheme (i.e., equations relating a Taylor polynomial at each grid point with its neighboring polynomials) which provides a highest order discretization error. Since we spend no degrees of freedom to provide boundary adjustment of neighboring polynomials, enough degrees of freedom remain to "cancel" the Taylor coefficients in a discretization error up to a high order. The following rough computation shows what order of accuracy can be expected

For a second order linear PDE and jets of degree k we have $2k+1$ parameters at each grid point. Thus at four nearest neighboring points, we have $8k+4$ parameters. Assuming that the systems that arise are solvable, we can reconstruct (at the central point) a jet of degree $4k$ satisfying the equation, or we can cancel the Taylor terms in the discretization error up to the same degree. In particular, for $\Delta u = 0$, for a regular grid and $k = 2$, the corresponding equations are solvable up to degree 10 (and not only 8) because of a symmetry in the problem.

Notice that wider neighborhood stencils can be involved in the relaxation procedure. However, geometrically compact schemes have important computational advantages.

Initial investigation of the parabolic equations has been started in [Bic-Yom]. It confirmed the results obtained in the elliptic case. Investigation of equations, developing singularities (like Burgers equation) has also been started.

Motion Planning; Inverse Kinematic problem

Motion planning in Robotics has been investigated in [Eli-Yom 1-3], [Sham-Yom] and in [Tan-Yom]. In section 1.1.4 of the Introduction our approach to the Motion planning has been described in some detail, so we do not repeat this discussion here.

In the implemented algorithm for a plane Motion planning the piecewise-linear approximation has been used at each grid-point, while the main effort concentrated on the efficient treatment of singularities of the data (on the boundary of the "configuration space"). A couple of additional important features, inherently related to our approach, have been investigated. It is a parallel realization of the algorithms, and the organization of the information flow. Both issues are governed (in our discretization scheme) by the mathematical nature of the problem. We do not discuss these topics here, referring the reader to [Eli-Yom 1-3] and [Yom 27].

Inversion of Mappings

For a given mapping of two Euclidean spaces its inversion is equivalent to solving a system of non-linear equations for each given right-hand side: for any y in the target we find all the preimages x under f of y.

The following specific problem has been considered: starting with an explicitly given direct mapping f, find a convenient representation of the inverse mapping f^{-1}. The construction of this representation may be mathematically involved. However, once constructed, it should allow for a fast computing of all the values of $f^{-1}(y)$ for each given y.

Such a statement of the problem has some practical justifications. In particular, in the "Ray tracing" procedure in the 3D Computer Graphics, certain nonlinear mappings (projections of the surfaces of the 3D objects onto the screen) must be inverted, in order to find a new color of each pixel on the screen, as the illumination conditions change. This must be done dynamically and in real time. This problem still presents a computational challenge even for home computers, not speaking about portable devices.

An algorithm for a numerical inversion of nonlinear mappings, described in [Eli-Yom 4], has been practically implemented only for mappings of the plane to itself, so we restrict our discussion to this case.

We use as an input the Kolmogorov representation of the direct mapping f (i.e. its approximation by the Taylor polynomials of degree k on a certain grid in the source). The inverse mapping f^{-1} is also represented on a certain grid in the target space. However, f^{-1} may have several branches over some regions. Accordingly, at each grid-point f^{-1} is represented by one ore several "models". The models may be of the following three types:

1. A regular point. A model of this type is just a Taylor polynomial of degree k, approximating the regular branch of the inverse mapping.

2. A fold. This model consists of the normal form "fold" $(y_1 = x_1, y_2 = x_2^2)$, and of the direct and the inverse coordinate transformations, bringing f to its normal form. These coordinate transformations are given by their Taylor polynomials of degree $k-1$.

3. A cusp. This model consists of the normal form "cusp" $(y_1 = x_1, y_2 = x_2^3 - 3x_1x_2)$, and of the direct and the inverse coordinate transformations, bringing f to its normal form. These coordinate transformations are given by their Taylor polynomials of degree $k-2$. (The reduction in the degrees of the Taylor polynomials for the normalizing coordinate transformations is caused by the mathematical structure of the Whitney procedure ([Whi 2]), which in turn reflects the natural "balance of the smoothness" in the problem).

The inversion of the normal forms (the standard fold and cusp) is assumed to be easily available. In practical realizations of the algorithm this inversion is implemented in a special fast subroutine. Now to compute the inverse function $f^{-1}(y)$ for any given y, we just substitute y into the corresponding model. For a regular point the k-jet of the inverse mapping at the point y is computed. For the fold (cusp) the point y is substituted into the k-jet of the the normalizing coordinate transformation in the target. Then the inverse of the standard fold (cusp) is applied. Finally, the result is substituted into the k-jet of the the inverse coordinate transformation in the source. In any case, the overall complexity of computations is comparable with that of computing a polynomial of degree k at one point.

The logical structure of the algorithm is as follows:

The algorithm analyses the input Taylor polynomials of the direct mapping. According to the Whitney description of the singularities of the plane to plane mappings, a decision is taken, whether the considered point is classified as a regular point, near-fold or near-cusp. In the first case the inverse jet is obtained by the standard jet-inversion formulae from the "Jet-calculus". If the point has been classified as a near-fold (near-cusp), the normalizing coordinate transformations are computed via a "jet implementation" of the Whitney procedure ([Whi 2]).

(Notice that the construction of a numerically stable jet implementation for the procedure of bringing a singularity to its normal form is not trivial. It requires, in particular, a careful study of the "balance of the smoothness" for this specific singularity. Sometimes it may be necessary to introduce additional parameters to the normal form, in order to increase stability of computations (see [Eli-Yom 4], [Yom 27]). However, as a stable and efficient "jet normalization" has been constructed, it becomes one of the standard formulae of the Jet calculus library).

The main problem in the implementation of the algorithm is in the classification of the points. Indeed, if the point has been classified as a near-fold, a real fold must be somewhere nearby, and the normalizing coordinate transformations to the normal form of this fold must be effectively bounded together with their derivatives (otherwise their computing will not be reliable enough). To guarantee this we must to know that this nearby fold (cusp) is sufficiently non-degenerate.

So we put a number of thresholds, determining the decision at each branching point, and the algorithm operates according to the basic pattern described above and to the chosen thresholds. Of course, as the step-size of the greed decreases, the non-degeneracy assumptions can be relaxed.

If the non-degeneracy assumptions are not satisfied in a neighborhood of a certain point, we either subdivide the grid or restrict the processing to a rough low-order approximation of the inverse mapping.

So one of the most important tasks in the practical implementation of the algorithm is a proper tuning of the thresholds. It is exactly the point where we expect a serious improvement via the “Quantitative Whitney theory” of singularities of mappings of the plane to itself. In Section 10.3.7 below we give a sketch of a possible form of this theory and its possible applications.

In the present realization of the algorithm the thresholds have been tuned empirically, separately for each example ([Eli-Yom 4]).

The algorithm has been practically implemented for $k = 2$ ([Eli-Yom 4]) and preliminary tested.

Images Representation, Compression and Processing

The most advanced implementation of our approach till now has been achieved in Image Processing. A nonlinear high-order approximation, together with the appropriate “Normal forms” capturing image singularities, have been used to achieve a very compact representation of images and “synthetic video” sequences ([Bri-Eli-Yom], [Eli-Yom 5], [Y-E-B-S], [Bri-Yom 6]). The results have been intensively tested against various known methods, and have been used in working practical applications.

We do not discuss here the details of this implementation. Such a discussion would go too far away from the mainstream of this book. Let us notice only that we indeed use in image analysis a third and fourth order local polynomial approximation, which would be impossible without a proper application of the multi-order strategy. We use three types of normal forms for singularities (edges, ridges and “patches”) and a low order Kolmogorov representation for smooth parts of the image. Since most of the edges and ridges in a typical image of the real world appear as the visible contours of the 3D objects, we do believe, that the methods of this book will ultimately help us to optimize our representation.

10.1.4.7 The Role of “Quantitative Singularity Theory”. As it was explained above, our method applies normal forms of singularities already on the level of the initial data structure, and then along all the route of the processing. One of the main problems in this approach is a treatment of near-singular situations, and switching from a regular to a singular representation. The cases, investigated in detail till now used an empiric approach to this problem. Undoubtedly, a proper use of the Quantitative Sard and Transversality theorems can provide a firm basis for an efficient realization of this type of algorithms. Moreover, each of the algorithms described above poses virtually the same mathematical problems, whose natural place is in the framework of “Quantitative Singularity Theory”. Let us describe here two of these problems. We shall return to them in Section 10.3.7 below, devoted to this theory.

1. Identification of the “Organizing Center”

This notion was introduced by R. Thom in [Tho 3]. In our interpretation the problem is *for a near-singular point to find a nearby "exact" singularity and the normalizing transformations for this singularity, via efficient and robust procedures.*

This is one of the most tricky problems in construction of virtually any algorithm involving an explicit treatment of singularities.

2. Finding probability distributions for the "degree of non-degeneracy" of singularities.

This problem is closely related to the first one. The estimates which appear in finding the "Organizing Center" involve as the main parameter the degree of non-degeneracy of the singularity to be found. Knowing a lower bound for this degree of non-degeneracy implies the tuning of the thresholds at the branching points of the algorithm.

The problem is *to find the probability distributions of various parameters of the singularities, responsible for the degree of their non-degeneracy. This should be done with respect to the natural probability distribution on the input data.*

Having these distributions, we can optimize our tuning of the thresholds in order to get the best average performance of the algorithm.

In Section 10.3.7 below we discuss these and other problems of Quantitative Singularity Theory in more detail, describing, in particular, the solution of the above problems in some special cases.

10.2 Semialgebraic Complexity of Functions

In Chapter 9 above it was shown how "stable" metric properties of semialgebraic sets and mappings can be applied in smooth analysis, via approximation of $\mathcal{C}^k$-functions by their Taylor polynomials.

It turns out that exactly the same method can be applied to a much wider class of functions than $\mathcal{C}^k$-ones. In fact, what we need for the results of Chapter 9 to be true, is not a regularity ($\mathcal{C}^k, \mathcal{C}^\omega$, etc.) of the functions considered, but rather their "complexity", measured as the rate of their best approximation by semialgebraic functions of given "combinatorial complexity".

Investigation of semialgebraic complexity presents the most direct continuation of the main lines of this book. Consequently, in this section we present with somewhat more detail (but without proofs) some definitions and results in this direction. Mostly we follow [Yom 16,17,20,24] and try to stress open problems and promising investigation directions.

To simplify presentation, we always assume our functions to be continuously differentiable, and as the main property under investigation we take the one given by the Quantitative Sard Theorem. This restriction is not essential. In fact, on one side, the approach can be generalized to Lipschitz functions (see [Cla 1,2], [Yom 11,12,15]), and, on the other side, much more general "geometric complexity bounds" (in the spirit of the above sections of this chapter) can be obtained for functions of bounded semialgebraic complexity.

10.2.1 Semialgebraic Complexity

Let $f : B_r^n \to \mathbb{R}$ be a $\mathcal{C}^1$-function. Let $g : B_r^n \to \mathbb{R}$ be a semialgebraic function. We do not assume g to be differentiable or even continuous. However, g is analytic on a complement of a semialgebraic subset $S(g)$, dim $S(g) < n$.

Definition 10.5. For f, g as above, the deviation $||f - g||_{\mathcal{C}^1}$ is defined as:

$$||f - g||_{\mathcal{C}^1} = \sup_{x \in B_r^n \setminus S(g)} (|f(x) - g(x)| + r\|\nabla f(x) - \nabla g(x)\|),$$

where $\|\ \|$ is the usual Euclidean norm of the gradients.

Now for any semialgebraic function $g : B_r^n \to \mathbb{R}$, let $C(g) = C(D(g))$, where $D(g)$ is the diagram of g, be the constant, defined in Theorem 7.5 above. This constant, essentially, bounds the entropy of near-critical values of g. (Strictly speaking, theorem 7.5 must be extended from polynomial to semialgebraic functions. However, such an extension is rather straightforward, if we do not ask for the best constants).

Definition 10.6. Let $f : B_r^n \to \mathbb{R}$ be a $\mathcal{C}^1$-function. A semialgebraic complexity $\sigma_s(f, \epsilon)$, for any $\epsilon > 0$, is defined as follows:

$$\sigma_s(f, \epsilon) = \inf\ C(g),$$

where the infinum is taken over all the semialgebraic functions g, such that:

$$||f - g||_{\mathcal{C}^1} \leqslant \epsilon.$$

In other words, $\sigma_s(f, \epsilon)$ is the minimal "$C(g)$-complexity" of semialgebraic functions g, ϵ-approximating f in $\mathcal{C}^1$-norm. Alternatively, we can define a "semialgebraic approximation rate" $E_s(f, d)$ as the inf $||f - g||_{\mathcal{C}^1}$ over all the semialgebraic g with $C(g) \leqslant d^n$. These definitions are motivated by classical Approximation Theory, where g are mostly taken to be polynomials (trigonometric polynomials, other orthogonal systems ...), and $C(g)$ is the degree (see [Ahi], [Lor 1-4], [War]). One of the most basic facts here is that the rate of a polynomial approximation of a given function is completely determined by its "regularity" in the usual sense: the number of continuous derivatives, in the finite smoothness case, or the size of the complex domain, to which the function can be extended, in the real analytic case.

More accurately, let us define the polynomial "complexity" and "approximation rate" as:

$$\sigma_p(f, \epsilon) = \inf_p C(p),$$

over all polynomials p with $||f - p||_{\mathcal{C}^1} \leqslant \epsilon$,

$$E_p(f,d) = \inf_p ||f-p||_{\mathcal{C}^1},$$

over all polynomials p with $C(p) \leqslant d^n$. (Notice that above, $C(p) = d^n$.) Written as above, the definition shows that σ_p and E_p are constructed exactly as σ_s and E_s, only with a subclass of all semialgebraic functions. This proves immediately that for any $\epsilon > 0$ and $d > 0$,

$$\sigma_s(f,\epsilon) \leqslant \sigma_p(f,\epsilon),$$

$$E_s(f,\epsilon) \leqslant E_p(f,\epsilon).$$

Now the classical Jackson's and Bernstein's theorems in Approximation Theory can be reformulated in our case as follows (see, for example, [Ahi], [Lor 1]):

Theorem 10.7. *If $f : B_r^n \to \mathbb{R}$ is $\mathcal{C}^k$, then*

$$\sigma_p(f,\epsilon) \leqslant C_1 \left(\frac{1}{\epsilon}\right)^{\frac{n}{k-1}},$$

$$E_p(f,d) \leqslant C_2 \left(\frac{1}{d}\right)^{k-1}.$$

Conversely, if $\sigma_p(f,\epsilon) \leqslant C\left(\frac{1}{\epsilon}\right)^{\frac{n}{k-1}}$ (or, equivalently, $E_p(f,d) \leqslant C'\left(\frac{1}{d}\right)^{k-1}$), then f is k times continuously differentiable on B_r^n.

For analytic functions the corresponding result is true, with:

$$\sigma_p(f,\epsilon) \sim |\log \epsilon|^n,$$

$$E_p(f,d) \sim q^d\ ,\ q < 1.$$

(We do not intend here to give accurate formulations of the results of Approximation Theory. Consequently, some details, sometimes important, are omitted).

Let us return to semialgebraic complexity. It is bounded by the polynomial one, and one can show that for generic $\mathcal{C}^k$ or analytic functions σ_s and σ_p are equivalent. On the other hand, we can see immediately, that semialgebraic complexity can be small for functions, not regular in the usual sense. Indeed, let $f(x)$ be defined as:

$$f(x) = \begin{cases} 0, & -1 \leqslant x \leqslant 0 \\ x^2, & 0 \leqslant x \leqslant 1\ . \end{cases}$$

f is $\mathcal{C}^1$, but not $\mathcal{C}^2$ on $[-1,1]$, and since f is itself semialgebraic, $\sigma_s(f,\epsilon) \leqslant$ *const*, and $E_s(f,d) = 0$ for d big enough. The same is true for any ($\mathcal{C}^1$) semialgebraic function f.

Below we give many examples of functions, whose semialgebraic complexity is better than their regularity prescribes. Then the following problem becomes a central one for understanding the relationship between "regularity" and "complexity properties": *Does low semialgebraic complexity imply existence of the high order derivatives in a certain generalized sense?*

There are some partial results in this direction. A very important class of nonsmooth functions with a low semialgebraic somplexity is given by maxima of smooth families, discussed above. For a function f, representable as a pointwise supremum of a bounded in C^2-norm family of C^2-smooth functions, its generalized Laplacian $\tilde{\Delta} f$ (in the distribution sense) is shown in [Yom 8,9] to be a measure with an explicitly bounded variation and singular part.

Question. Is a similar property true for f with $\sigma_s(f,\epsilon) \sim (1/\epsilon)^{n/2}$?

Another approach here is the following: A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to have a k-th Peano differential at $x_0 \in \mathbb{R}^n$, if there exists a polynomial $P : \mathbb{R}^n \to \mathbb{R}$ of degree k, such that $|f(x) - P(x)| = o(\|x - x_0\|^k)$. A classical result in convex geometry - the Alexandrov-Fenchel theorem - claims that any convex function has the second Peano differential almost everywhere. Suprema of C^2-families can be shown to admit a representation as a difference of two convex functions (see [Roc], [Shap-Yom]) and hence are almost everywhere twice Peano differentiable.

Conjecture. *If the semialgebraic complexity of a $\mathcal{C}^1$-function $f : B_r^n \to \mathbb{R}$ satisfies $\sigma(f,\epsilon) \leqslant C \left(\frac{1}{\epsilon}\right)^{\frac{n}{k}}$, then f has a k-th Peano differential almost everywhere.*

This conjecture is strongly supported by the following result of Dolgenko ([Dol 1,2], see also [Iva 2]): Define the *rational complexity* $\sigma_r(f,\epsilon)$ of f exactly as in definition 10.2, but restricting the approximating functions to the rational ones. Then, as shown in [Dol 1,2], the condition $\sigma_r(f,\epsilon) \leqslant C \left(\frac{1}{\epsilon}\right)^{\frac{n}{k}}$ does not imply even C^2-smoothness of f. However, one of the main results of [Dol 1,2] states that under this condition f has a k-th Peano differential almost everywhere.

The proof in [Dol 1,2] seems to apply to semialgebraic complexity with no essential modification.

Thus for a "$\mathcal{C}^k$-type" behavior of the complexity ($\sigma(f,\epsilon) \sim \left(\frac{1}{\epsilon}\right)^{n/k}$) we have some partial results and (hopefully) reasonable conjectures. The following interesting problem then naturally arises:

What kind of "regularity" can be expected for functions with an "analytic-type" behavior of complexity ($\sigma(f,\epsilon) \sim |\log \epsilon|^n$)?

In particular, we can expect for such functions existence, at almost every point of Peano differentials of any order.
Do the Taylor series, defined in this way, converge? What is their relation with the original function?

10.2.2 Semialgebraic Complexity and Sard Theorem

The result of this section is that the geometry of the set of critical values of the C^1-function is determined by its semialgebraic complexity (and not by its regularity, as it appears in standard settings of Sard-like results). So let $f : B_r^n \to \mathbb{R}$ be a C^1-function, $\Sigma(f)$ the set of its critical points, and $\Delta_f = f(\Sigma_f)$ the set of its critical values.

Theorem 10.8. *For any* $\epsilon > 0$,

$$M(\epsilon, \Delta_f) \leqslant \sigma_s(f, \epsilon) \ ,$$

where $\sigma_s(f, \epsilon)$ *is the semialgebraic complexity of* f.

Proof. It is completely identical with the proof of Proposition 9.1 (with a simplification, following from the fact that we consider only critical, and not near-critical, values). For a given $\epsilon > 0$, we find a semialgebraic function g, such that $||f - g||_{C^1} \leqslant \epsilon$. By definition of $||f - g||_{C^1}$ in Section 10.2 above, it follows that $|f(x) - g(x)| \leqslant \epsilon$ for any $x \in B_r^n$, and $\|\nabla f(x) - \nabla(g)\| \leqslant \epsilon/r$ for any $x \in B_r^n$, where g is smooth. Hence $\Sigma(f)$ is contained in the set of ϵ/r-critical points $\Sigma(g, \epsilon/r)$ of g. In turn, Δ_f is contained in an ϵ-neighborhood $\Delta_\epsilon(g, \epsilon/r)$. Therefore:

$$M(\epsilon, \Delta_f) \leqslant M(\epsilon, \Delta_\epsilon(g, \epsilon/r)) \leqslant C(g) \ ,$$

by Theorem 7.5 of Chapter 7 above. Now taking the infimum over all the semialgebraic g with $||f - g||_{C^1} \leqslant \epsilon$, we get

$$M(\epsilon, \Delta_f) \leqslant \inf_g C(g) = \sigma_s(f, \epsilon) \ .$$

This completes the proof. □

The semialgebraic complexity $\sigma_s(f, \epsilon)$ is a "correct" property of functions not only for the Sard Theorem itself, but for all the related properties, discussed in the above sections of this Chapter: transversality results, average number of connected components of the fiber, etc. Also "computational complexity" of most of the natural mathematical operations with f, is bounded in terms of $\sigma_s(f, \epsilon)$. In particular, this is true for solving equations $f = const$ with a prescribed accuracy (see [Yom 20]).

A very important exclusion, however, is given by the dynamical results around the $\mathcal{C}^k$ reparametrization theorem, shortly mentioned below in this Chapter. To bound complexity of the iterations of a mapping $f : M \to M$ (and, in particular, its entropy, volume growth, etc.), it is not enough to assume that $\sigma(f, \epsilon)$ is small. The problem is that a piecewise-smooth structure of f, to which $\sigma(f, \epsilon)$ is essentially insensible, in iterations can lead to an exponential growth of the number of smooth pieces, and thus to the blow-up of the complexity. Consequently, we consider the following problem as an important one for understanding the nature of various complexity notions:

Is it possible to replace the $\mathcal{C}^k$ or analyicity assumptions in "dynamical" complexity bounds (like those discussed below) by weaker "complexity"-type assumptions?

A natural candidate is provided by various types of lacunary series (in particular, Bernstein's quasianalytic classes), whose semialgebraic complexity may be low, despite the lack of the usual regularity, and whose structure excludes "pieces accumulation", mentioned above.

The following main examples of functions, whose semialgebraic complexity is better (sometimes much better) than their regularity prescribes, have been investigated in [Yom 16,17,20,24]: maxima of smooth families, functions, representable as compositions (see also [Vit 3]), lacunary series (in particular, Bernstein's quasianalytic classes) and certain functions on infinite-dimensional spaces. Below we give only one example of the last type.

We believe that richness of mathematical structures, involved in these examples, and their potential applicability to many important problems in Analysis justifies further investigation of semialgebraic complexity.

10.2.3 Complexity of Functions on Infinite-Dimensional Spaces

The notion of semialgebraic complexity applies equally well to functions on infinite-dimensional spaces. Moreover, it is in this setting that the difference between the notions of complexity and regularity becomes apparent. This is due to the fact that a "polynomial of an infinite number of variables" is not a "simple function" (unless it satisfies some additional restrictions). Just a regularity, which is also for functions of infinite number of variables roughly equivalent to the rate of their global polynomial approximation, provides no information on complexity (in contrast to the finite-dimensional case).

Let us start with an example of a polynomial on ℓ^2, which does not satisfy the usual Sard theorem (it belongs to I. Kupka [Kup2]).

Let $\ell^2 = \{x = \{x_1, \dots, x_i, \dots), \Sigma x_i^2 < \infty\}$ be the standard Hilbert space. We define a function f in the following way:

$$\begin{array}{rccl} f: & \ell^2 & \to & \mathbb{R} \\ & x & \to & f(x) = \displaystyle\sum_{i=1}^{\infty} \frac{1}{2^i}\varphi(ix_i), \end{array} \qquad (\mathbf{1})$$

where φ is the polynomial of degree 3, such that $\varphi(0) = \varphi'(0) = 0$, $\varphi(1) = 1$ and $\varphi'(1) = 0$. Thus φ has exactly two critical points 0 and 1 with the critical values 0 and 1 respectively. One can easily show that f is infinitely differentiable (in any reasonable definition of differentiability on ℓ^2). In fact f can be considered as a polynomial of degree 3 on ℓ^2. Now $x = (x_1, \dots, x_i, \dots)$ is a critical point of f if and only if for each i, ix_i is a critical point of φ. Thus

$$\Sigma(f) = \left\{ \left(a_1, \frac{a_2}{2}, \dots, \frac{a_i}{i}, \dots\right), a_i = 0, 1 \right\}$$

$$\text{and} \quad \Delta_f = \left\{ \sum_{i=1}^{\infty} \frac{1}{2^i} a_i, a_i = 0, 1 \right\} = [0, 1] \ .$$

Hence the critical values of a C^∞-function f cover the interval:

the Sard theorem is no more valid on ℓ^2. (See also [Hai] for an example of a function $f : \ell^2 \to \mathbb{R}$, having $[0,1]$ for critical values and a critical set of Hausdorff dimension 4) Notice that the function f can be approximated by "simple" ones, namely by the polynomials, depending only on a finite number of variables: $\sum_{i=1}^{N} \varphi(ix_i)$. Generalizing this remark, we show below that f violates the Sard theorem since the rate of its approximation by these "simple" polynomials is not high enough.

Kupka's example shows that in order to apply the approach of this chapter to an infinite-dimensional situation, we have first to find a good class of "simple" approximating functions. This suggests the following generalization of our main definition 10.2 :

Let V be a Banach space (of finite or infinite dimension) and let $B \subset V$ be the unit ball in V. We consider (Frechet) continuously differentiable functions on B and with the standard $\mathcal{C}^1$ norm $\| \ \|_{\mathcal{C}^1}$.

Now assume that some subclass Q of such functions is given, satisfying the following condition $(*)$:

Fix any $q \in Q$. Then for any $\epsilon > 0$, the set of ϵ-critical values of q on B, $\Delta(q, \epsilon)$, can be covered by $C(q)$ intervals of length ϵ, i.e. $M(\epsilon, \Delta(q, \epsilon)) \leqslant C(q)$, with $C(q)$ depending only on q and not on ϵ.

Definition 10.9. For any $\mathcal{C}^1$-function f on B, the Q-complexity $\sigma_Q(f, \epsilon)$ is defined as

$$\sigma_Q(f, \epsilon) = \inf_{q \in Q, ||f - q||_{\mathcal{C}^1} \leqslant \epsilon} C(q).$$

Theorem 10.10. — *For any $\epsilon > 0$,*

$$M(\epsilon, \Delta_f) \leqslant \sigma_Q(f, \epsilon).$$

The proof is identical to the proof of Theorem 10.8 above.

The main difficulty in the application of this result to specific functions on infinite-dimensional spaces consists of a choice of the approximating class Q. To understand Kupka's example, we shall take Q consisting of polynomials, depending on a finite number of variables, or more accurately, on a finite number of linear functionals on V.

Proposition 10.11. *Let V be a Banach space and let $\ell_1, \dots, \ell_n$ be linear functions on V. Let $p(x_1, \dots, x_n)$ be a polynomial of degree d. Then for a function $\tilde{p} : B \to \mathbb{R}$, $\tilde{p}(v) = p(\ell_1(v), \dots, \ell_n(v))$, $C(\tilde{p}) \leqslant n \cdot (2d)^n$.*

Here $B \subseteq V$ denotes the unit ball in V, and $C(\tilde{p})$, as above, is the minimal number of ϵ-intervals, covering the set of ϵ-critical values of $\tilde{p}$ on B. For $V = \ell^2$ and $\ell_1, \dots, \ell_n$ orthonormal, n can be omitted.

Applying this proposition to the partial sums of the infinite series, we get, after simple computations:

Theorem 10.12. *For* $f = \sum_{i=1}^{\infty} (\frac{1}{q})^i p_i(x_1, \dots, x_i)$, $\deg p_i = d$, $|p_i| \leqslant 1$ *on the unit ball and* $q > 1$, $\sigma_Q(f, \epsilon) \leqslant C(q, d)(\frac{1}{\epsilon})^{\log_q(2d)}$. *In particular, for* $q > 2d$, f *satisfies the Sard theorem (i.e.,* $m(\Delta_f) = 0$*).*

Returning to the Kupka example above, we see that it occurs exactly on the boundary: by Theorem 10.12, for any $q > 6$, functions $f = \sum_{i=1}^{\infty} \frac{1}{q^i} p_i(x, \dots, x_i)$ with $\deg p_i = 3$, satisfy the Sard theorem. In the specific form of the function f in Kupka's example above it is enough to take $q > 2$. By formal analogy we can say that the complexity of the function $f : \ell^2 \to \mathbb{R}$, $f = \sum_{i=1}^{\infty} \frac{1}{q^i} p_i(x_1, \dots, x_i)$, $\deg p_i = d$, is the same as the complexity of C^k-functions $g : B^n \to \mathbb{R}$, if $\frac{n}{k-1} = \log_q(2d) = \beta$. In particular, the sequences of the form $1, 1/2^s, 1/3^s, \dots, 1/k^s, \dots$, may appear among the critical values of both f and g only if $1/(s-1) \leqslant \beta$.

More results and discussions in this spirit can be found in [Yom 17,20,24]. See also [Sma 2], where an infinite dimensional version of Sard's Theorem is given. It is proved in [Sma 2] for nonlinear Fredholm mappings via reduction to a finite dimensional case; one can expect that the complexity approach as above will also work in this situation.

10.3 Additional Directions

10.3.1 Asymptotic Critical Values of Semialgebraic and Tame Mappings

We'd like to briefly mention here some results concerning the singular behaviour of semialgebraic or tame mappings at infinity. To do this, let us define the set $K(f)$ of generalized critical values of a semialgebraic or a tame mapping $f : \mathbb{R}^n \to \mathbb{R}^m$, $n > m$:

$$K(f) = \Delta_f \cup K_\infty(f),$$

Δ_f being the classical set of critical values of f (i.e. $\Delta_f = f(\Sigma_f)$, with Σ_f the set of points $x \in \mathbb{R}^n$ such that $Df_{(x)}$ is not onto) and $K_\infty(f)$ being the set of critical values at infinity or asymptotic critical values, defined by:

$$K_\infty(f) = \{y \in \mathbb{R}^m, \exists x_k \in \mathbb{R}^n, |x_k| \to \infty, f(x_k) \to y,\ |x_k| \cdot \lambda_m(Df_{(x_k)}) \to 0\}.$$

Of course, when x_k is a singular point, $\lambda_m(Df_{(x_k)}) = 0$, thus $K_\infty(f)$ is the union of $\mathrm{adh}(\Delta_f)$ (the closure of Δ_f) and of the set:

$\{y \in \mathbb{R}^m, \exists x_k \in \mathbb{R}^n \setminus \Sigma_f, |x_k| \to \infty, f(x_k) \to y,\ |x_k| \cdot \lambda_m(Df_{(x_k)}) \to 0\}$.

The set $K(f)$ has the following remarkable property (see [Rab] for a very general context, or [Kur-Orr-Sim] for the semialgebraic case):

Theorem. ([Rab]) *With the same notations as above, assuming that f is a $\mathcal{C}^2$-semialgebraic mapping, then $f : \mathbb{R}^n \setminus f^{-1}(K(f)) \to \mathbb{R}^m \setminus K(f)$ is a fibration over each connected component of $\mathbb{R}^m \setminus K(f)$.*

The question of the size of the set of generalized critical values $K(f)$ can be seen as a Sard problem at infinity. Of course when f is semialgebraic (resp. tame), $K(f)$ is a semialgebraic (resp. tame) subset of $\mathbb{R}^m$, and the question of the size of $K(f)$, is just the question of its dimension. This question has been solved in [Kur-Orr-Sim], essentially with techniques similar to those developped in [Yom 1] and in the present book.

Theorem. ([Kur-Orr-Sim]) *With the same notations as above, assuming that f is a $\mathcal{C}^1$-semialgebraic mapping, then the semialgebraic set $K(f)$ is closed and has dimension (strictly) less than m.*

The fact that $K(f)$ is a strict semialgebraic set of $\mathbb{R}$, i.e. is a finite set, for a polynomial function $f : \mathbb{R}^n \to \mathbb{R}$, plays a crucial role in the proof of the gradient conjecture of R. Thom (see [Kur-Mos], [Kur-Mos-Par]). We have an extension of this result in the more general setting of tame categories (the techniques of the proof are here quite different from the techniques of the proof in the semialgebraic case):

Theorem. ([D'Ac]) *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a tame $\mathcal{C}^1$-mapping, with the same notations as above, then the tame set $K(f)$ is finite.*

10.3.2 Morse-Sard Theorem in Sobolev Spaces

We essentially refer here to [Pas] (see [Eva-Gar] for the basic definifions). The main theorem of this paper is the following, generalizing the classical Morse-Sard Theorem for Sobolev functions.

Theorem. [Pas] *Let $n > m$ be two integers and let, for $p > n$, $f \in W^{n-m+1,p}_{\ell oc}(\mathbb{R}^n;\mathbb{R}^m)$. Then $\mathcal{H}^m(\Delta_f) = 0$.*

The continuous representative $\widetilde{f}$ of f is in $\mathcal{C}^{n-m,\alpha}$, for some $\alpha = \alpha(p) < 1$ (see [Eva-Gar]), and the quantitative Morse-Sard Theorem for $\widetilde{f}$ shows that the Hausdorff dimension of $\Delta_{\widetilde{f}}$ is at most $\min(m, m-1+\dfrac{n-m+1}{n-m+\alpha}) = m$. In other words, the classical Morse-Sard Theorem gives no information on the dimension of Δ_f, when $f \in W^{n-m+1,p}_{\ell oc}(\mathbb{R}^n;\mathbb{R}^m)$. As pointed out in [Pas], this is the effect of the existence of another weak derivative summable enough.

The idea of the proof is the following: we know that there exists, for every $\epsilon > 0$, a set $F_\epsilon \subset \mathbb{R}^n$ and a mapping $f_\epsilon \in \mathcal{C}^{n-m+1}(\mathbb{R}^n;\mathbb{R}^m)$, such that

$f_{|F_\epsilon} \equiv f_{\epsilon|F_\epsilon}$ and $\mathcal{H}^n(\mathbb{R}^n \setminus F_\epsilon) \leqslant \epsilon$. Let us notice that $\Sigma_f = \bigcup_{n\in\mathbb{N}} (\Sigma_f \cap F_{\frac{1}{n}}) \cup \mathcal{N}$, where $\mathcal{H}^n(\mathcal{N}) = 0$, and that, by the classical Morse-Sard Theorem:

$$\mathcal{H}^m(f(\Sigma_f \cap F_{\frac{1}{n}})) = \mathcal{H}^m(f_{\frac{1}{n}}(\Sigma_{f_{\frac{1}{n}}} \cap F_{\frac{1}{n}})) = 0.$$

The main point of [Pas] is then to prove that $\mathcal{H}^m(\mathcal{N}) = 0$, i.e. for several sets $A \subset \Sigma_f$, such that $\mathcal{H}^n(A) = 0$, we have $\mathcal{H}^m(f(A)) = 0$.

Classical examples show that this result cannot be sharpened.

10.3.3 From Global to Local: Real Equisingularity

We have indicated in Chapter 5, how to localize the ℓ-variation of sets $A \subset \mathbb{R}^n$ of dimension ℓ (i.e. how to localize the ℓ-volume of A), in order to obtain a local invariant for the germ A_0 (we assume that $0 \in A$), called the density of A_0 and denoted $\Theta_\ell(A, 0)$.

We can proceed in the same way for the variations V_i, with $i = \{1, \dots, n\}$ (see [Com-Gra-Mer]): let ϵ be a positive real number and let us denote A^ϵ the set $\frac{1}{\epsilon}.(A \cap B^n_\epsilon) \subset B^n_1$. Now if we fix $\bar{P} \in \bar{G}^{n-i}_n$, we remark that the sets $\bar{P} \cap A^\epsilon$ have the same topological type, for $\epsilon > 0$ small enough, and that the number of topological types of $\bar{P} \in \bar{G}^{n-i}_n$, with respect to $\bar{P}$ is finite. This fact is a direct consequence of the first isotopy lemma of Thom-Mather; we have to repeat the argument of Theorem 4.18:
let us denote $[\bar{G}^{n-i}_n]_1$ the compact set of $\bar{P} \in \bar{G}^{n-i}_n$ such that $\bar{P} \cap B^n_1 \neq \emptyset$, and

$$E = \{(\bar{P}, \epsilon, x); \bar{P} \in [\bar{G}^{n-i}_n]_1, \epsilon \in]0; 1], x \in A^\epsilon \cap \bar{P}\},$$

$$F = \{(\bar{P}, \epsilon, x); \bar{P} \in [\bar{G}^{n-i}_n]_1, , \epsilon \in]0; 1], x \in \mathrm{adh}(B^{n-i}_{(0,1)}) \setminus (A^\epsilon \cap \bar{P})\}.$$

We have: $\mathrm{adh}(E \cup F) = G = [\bar{G}^{n-i}_n]_1 \times [0; 1] \times \mathrm{adh}(B^n_{(0,1)})$. The projection $p : G \to [\bar{G}^{n-i}_n]_1 \times [0; 1]$ is a proper semialgebraic (or subanalytic in the subanalytic case) morphism. One can stratify this morphism in the following way: there exist a Whitney stratification Σ of G, compatible with E and F, and a Whitney stratification Σ' of $[\bar{G}^{n-i}_n]_1 \times [0; 1]$ (which is of course finite, since $[\bar{G}^{n-i}_n]_1 \times [0; 1]$ is compact), such that the fibers of $p^{-1}(\{(\bar{P}, \epsilon)\})$ and $p^{-1}(\{(\bar{Q}, \eta)\})$ are homeomorphic with respect to E and F. In particular the sets $A^\epsilon \cap \bar{P} = p^{-1}(\{(\bar{P}, \epsilon)\}) \cap E$ and $A^\epsilon \cap \bar{Q} = p^{-1}(\{(\bar{Q}, \eta)\}) \cap E$ are homeomorphic.

This shows that:

– the number of topological types of $A^\epsilon \cap \bar{P}$ is finite,

– for a given $\bar{P} \in [\bar{G}^{n-i}_n]_1$, the topological type of $A^\epsilon \cap \bar{P}$ does not depend on ϵ, for $\epsilon > 0$ small enough. In particular the integers $V_0(A^\epsilon \cap \bar{P})$ and $\chi(A^\epsilon \cap \bar{P})$ (where χ is the Euler-Poincaré characteristic) do not depend on ϵ, for $\epsilon > 0$ small enough.

We thus can define:

$$V_i^{\ell oc}(A_0) = \lim_{\epsilon \to 0} c(n,i) \int_{\bar{P} \in [\bar{G}_n^{n-i}]_1} V_0(A^\epsilon \cap \bar{P})\, d\bar{P}$$

$$= \lim_{\epsilon \to 0} \frac{c(n,i)}{\epsilon^i} \int_{\bar{P} \in \bar{G}_n^{n-i}} V_0(A \cap B_\epsilon^n \cap \bar{P})\, d\bar{P}$$

and

$$\Lambda_i^{\ell oc}(A_0) = \lim_{\epsilon \to 0} c(n,i) \int_{\bar{P} \in [\bar{G}_n^{n-i}]_1} \chi(A^\epsilon \cap \bar{P})\, d\bar{P}$$

$$= \lim_{\epsilon \to 0} \frac{c(n,i)}{\epsilon^i} \int_{\bar{P} \in \bar{G}_n^{n-i}} \chi(A \cap B_\epsilon^n \cap \bar{P})\, d\bar{P},$$

the local i-th variation of the germ A_0 and the local Lipschitz-Killing curvature of A_0 (the global Lipschitz-Killing curvatures have been introduced in [Wey], see also [Kla], [Had], [Sch], [Sch-McM]).

Of course we have a uniform bound on $\bar{P}$, in the semialgebraic case, for these invariants, in terms of the diagram of A.

One can study the relation between the geometry of a multigerm $(A_y)_{y \in Y}$ along a smooth set Y and the variation of the invariants $\Lambda_i^{\ell oc}(A_y)$ and $V_i^{\ell oc}(A_y)$ along Y. One can for instance prove that these invariants vary continuously along Y, when Y is a stratum of a Verdier stratification of $\mathrm{adh}(A)$ (see [Com-Gra-Mer]).

10.3.4 $\mathcal{C}^k$ Reparametrization of Semialgebraic Sets

Assume we are given an algebraic (or a semialgebraic) set $A \subset \mathbb{R}^n$. Its reparametrization is a subdivision of A into semialgebraic pieces A_j together with algebraic mappings $\psi_j : I_j \to A_j$, where I^{n_j} is a unit cube in $\mathbb{R}^{n_j}$. We assume additionally that ψ_j are onto and homeomorphic on the interiors of I^{n_j} and A_j.

A relatively easy fact, which can be proved completely in the framework of the methods presented above, is that for any compact semialgebraic set there exists a finite reparametrization (with the number of pieces bounded in terms of the diagram of A, i.e. in terms of the degrees and the number of the equations and inequalities, defining A). In a sense this result can be considered as a (strongly simplified) version of resolution of singularities of A.

Various "quantitative" questions can be asked in relation with reparametrizations of semialgebraic sets. Applications in dynamical systems motivate the following specific problem: for A inside I^n (the unite cube in $\mathbb{R}^n$), is it possible to add the requirement that the norm $\|\psi_j\|_{\mathcal{C}^k}$ be bounded by 1 (then the reparametrization is called a $\mathcal{C}^k$-one), and still to have for any such A the number of pieces, bounded in terms of the diagram of A?

The positive answer is rather straightforward for $k = 1$. However, for the derivatives of order two and higher new techniques have to be applied, in particular, Markov inequalities for polynomials, estimates of the derivatives etc. (Try to find such a reparametrization for $k = 2$ and the set A - a hyperbola $xy = \epsilon$, with the number of pieces not depending on ϵ).

The following result (in a weaker form) was obtained in [Yom 6,7] and in a final form in [Gro 3]:

Theorem 10.13. *For any natural k and for any semialgebraic A inside the unit cube in $\mathbb{R}^n$, there exists a $\mathcal{C}^k$-reparametrization of A, with the number of pieces, depending only on k and on the diagram of A.*

This theorem, combined with approximation by Taylor polynomials, proper rescalings and estimates of the derivatives of compositions, allows one to bound the local complexity of iterations of $\mathcal{C}^k$-smooth mappings. In particular, this provides an inequality between the topological entropy and the rate of the volume growth for such mappings ([Yom 6,7]).

As it was mentioned above, in these results $\mathcal{C}^k$-smooth mappings cannot be replaced by mappings of low semialgebraic complexity. Technically, the reason is that if we restrict a $\mathcal{C}^k$ function to smaller and smaller neighborhoods of the origin, and then rescale back to the unit ball, the derivatives tend to zero. Properly understood "complexity" of these rescaled functions also tends to zero, faster for larger k. This type of behavior is not shared by functions of low semialgebraic complexity: they may have a "conic singularity" near the origin, and then restriction to a smaller neighborhood and rescaling change nothing.

Dynamical problems (in particular, the study of the semicontinuity modulus of the topological entropy in analytic families) lead to the same question as above, where the $\mathcal{C}^k$-norm of the reparametrizing mappings is replaced by a certain analytic norm (see [Yom 21,23]).

Extension of the $\mathcal{C}^k$-reparametrization theorem to the analytic case is not straightforward. In dimension 1, it is roughly equivalent to the classical Bernstein inequality for polynomials (see [Ber]). However, in higher dimensions it requires a certain Bernstein-type inequality for algebraic functions (discussed below), which was proved only recently ([Roy-Yom]). We plan to present analytic reparametrization results in ([Yom 23]).

Recently $\mathcal{C}^k$-reparametrization theorem has been applied in the study of Anderson localization for Schrodinger operator on Z^2 with quasi-periodic potential ([Bou-Gol-Schl]). We hope that this rather unexpected application will allow one to better understand the analytic consequenses of this theorem, and possibly, of its sharpened version.

10.3.5 Bernstein-Type Inequalities for Algebraic Functions

Let D_R denote the closed disk of radius $R > 0$, centered at the origin in C.

Definition 10.14. *Let $R > 0$, and $K > 0$ be given and let f be holomorphic in a neighborhood of D_R. We say that f belongs to the Bernstein class $B_{R,K}$ if the maximum of the absolute value of f over D_R is at most K times the maximum over $D_{\frac{1}{2}R}$. The constant K is called the Bernstein constant of f.*

This definition is motivated by one of the classical Bernstein inequalities: *let $p(x)$ be a polynomial of degree d. Then*

$$max_{x \in E_R} ||p(x)|| \leqslant R^d max_{[-1,1]} ||p(x)||,$$

where E_R is the ellipse in C with the focuses at $-1, 1$ and the semiaxes R ([Ber]).

A problem of computing Bernstein constants of algebraic functions has recently appeared in several quite different situations.

In [Fef-Nar 1-3] this problem is investigated in relation with estimates of a symbol of some pseudodifferential operators. In [B-L-M-T 1,2] and [Bru 1], [Bar-Ple 1-3] (see also [Bos-Mi]) this problem is connected with some results in Potential Theory and with a characterization of algebraic subsets.

In [Roy 1,2], [Roy-Yom], [Bri-Yom 5], [Fra-Yom 1,2], [Yak 1,2], [Yom 22] and [Bru 2] Bernstein classes are used in counting zeroes in finite dimensional families of analytic functions (this problem is closely related to the classical problem of counting closed trajectories (limit cycles) of plane polynomial vector-fields).

In [Yom 6,7], [Yom 21] and [Yom 23] various forms of Bernstein inequality are used to prove results on a "$\mathcal{C}^k$-reparametrization" of semialgebraic sets, which, in turn allow one to control the complexity growth in iterations of smooth mappings (see Section 10.3.4 above). It was exactly the absence of the Bernstein inequality for algebraic functions, which restricted the results of [Yom 21] to one and two dimensional dynamics only.

By a *structural* Bernstein inequality for a certain class of functions, defined by algebraic data (algebraic functions, solutions of algebraic differential equations, etc.) we understand an inequality bounding the Bernstein constant of the function on a couple of concentric disks in terms of the degree and the relative position of these concentric disks in the maximal concentric disk of regularity only.

As the example of rational functions shows, in a sense, this is the best possible inequality one can expect for functions with singularities.

Let $y(x)$ be an algebraic function, given by an equation:

$$p_d(x)y^d + p_{d-1}(x)y^{d-1} + \cdots + p_1(x)y + p_0(x) = 0 \ ,$$

with $p_j(x)$ - polynomials in x of degree m. Let $\tilde{y}(x)$ be one of the branches of y and assume that $\tilde{y}$ is regular over D_R. (We can assume that D_R is a maximal disk of regularity of $\tilde{y}$, so its boundary contains poles or branching points of $\tilde{y}$).

Theorem 10.15. *For any $R_1 < R$, $\tilde{y} \in B_{R_1,K}$, with* $K = [\frac{4A(R+R_1)}{R-R_1}]^{2m+2}$. *Here A is an absolute constant.*

Theorem 10.15 provides a structural Bernstein inequality for algebraic functions of one variable. It can be easily extended to algebraic functions of several variables (see [Roy-Yom]).

One can restate this theorem (or, rather, its generalization to multivalued functions) in a more geometric way: if an algebraic curve Y of degree d in $\mathcal{C}^2$ is contained, over the disk D_R in the x-axis, in a tube of the size K, and it does not blow up to infinity over the disk D_{3R}, then it is contained in a tube of the size $C(d)K$ over the disk D_{2R}.

One can hope that this last result admits for a generalization to higher dimensions and more complicated (semi)algebraic sets. There is also an important problem of obtaining structural Bernstein inequalities for other classes of analytic functions, beyond algebraic ones (in particular, for solutions of algebraic differential equations and, hopefully, for their Poincare mappings). Some initial results in this direction are given in [Roy 1,2], [Roy-Yom], [Fra-Yom 1,2], [Bri-Yom 5], [Yom 22], [Bru 2].

In the next section we discuss shortly polynomial control problems, stressing situations, where semialgebraic geometry underlines the dynamics of the trajectories.

10.3.6 Polynomial Control Problems

A specific problem, related to the main topics of this book, and considered in [Bri-Yom 1-4] is the validity of Sard's theorem and its "quantitative" generalizations for an important class of nonlinear mappings, namely, input to state mappings of nonlinear finite dimensional control problems of the type:

$$\dot{x}(t) = f(x(t), u(t)), x(0) = x_0, \quad t \in [0, T], \tag{2}$$

where $x(t)$ is the state, and $u(t)$ is the control, at time t.

Input-to-state mapping J_f of (2) associates to each control $\tilde{u}$ the state $J_f(\tilde{u})$, to which $\tilde{u}$ steers the system from the initial state x_0 in time T.

Mappings J_f for nonlinear f are known to be complicated. However, the question of validity of Sard's theorem for these mappings is important from both the theoretical and computational points of view ([Sus], [Bri-Yom 1-4], [Zel-Zh]. See also [Mon], where the question of the validity of Sard's Theorem is presented as one of the important open problems in the field). Let us assume x and u to be one-dimensional. (Our methods work also in a multi-dimensional situation).

Let $v(t)$ be a perturbation of the control. Then the differential $DJ_f(u)(v)$ is given by the solution $z(T)$ of the linearized equation (2) along the trajectory $(x(t), u(t))$:

$$\dot{z}(t) = f_z(x(t), u(t))z(t) + f_u(x(t), u(t))v(t), z(0) = 0. \quad \textbf{(3)}$$

In particular, a control $u(t)$ is critical for J_f if and only if $f_u(x(t), u(t)) \equiv 0$.

As we assume f to be a polynomial, this last equation $f_u(x, u) = 0$ defines an algebraic curve Y in the plane (x, u). It allows one to express u as a (generally multivalued) function $u(x)$ of x.

Choosing a certain univalued branch $u(x)$ of this multivalued function and substituting this expression into the original equation (2) we get an ordinary differential equation:

$$\dot{x} = f(x, u(x)), x(0) = x_0, \quad t \in [0, T], \quad \textbf{(4)}$$

whose solution is uniquely defined on the all interval of existence.

Hence assuming that the control $u(t)$ is critical (and that it is continuous, i.e. does not jump from one branch of the algebraic curve Y to another) we get only a finite number of possibilities for the control u and for the solution x: at any double (multiple) point of the algebraic curve Y the control can switch from one branch to another. Clearly, the total number of such choices for u is bounded through the degree of the polynomial f.

This simple consideration shows that for the control problem (2) the number of critical values of the input-to-state mapping J_f (on the space of continuous controls) is finite, and bounded through the degree of the polynomial f. Assumption of continuity of the controls is not very natural in control (optimal controls may jump infinitely often) but in some cases it can be verified. In particular, the following result has been obtained in [Bri-Yom 1] using similar considerations: let the variables x and u in (2) be now two-dimensional. We assume that the possible values of the control u belong to a given compact polygon U in the plane.

Denote by $R(T, U)$ the time T reachable set of (2), i.e. the set of all the state positions $x(T)$, to which various controls $u(t)$ in U can drive the system in time T.

Theorem 10.16. *The number of the "outward" corners of the boundary of the time T reachable set $R(T, U)$ of the control problem* (2) *can be explicitly bounded through the degree of f and the number of vertices of U.*

Now let us return to one-dimensional x and u and consider near-critical controls. If the differential of the input-to-state mapping J_f is small, the differential equation (3) becomes a differential inequality, which leads to the requirement that the absolute value of $f_u(x, u)$ be small. This condition defines a semialgebraic set S in the plane. (All these objects of course depend on the parameter, measuring the size of the differential of J_f).

Therefore, near-critical trajectories $(x(t), u(t))$ lie in S. The complement to S consists of a finite (and bounded through the degree of f) number of "islands" O_i. Let us assume that $x(t)$ is monotone in t on $[0, T]$. (If a near critical trajectory $x(t)$ "turns back" at a certain moment t_0, one can show that it remains near the turning point $x(t_0)$for the rest of the time). Then for each island O_i in the plane (x, u) the trajectory $(x(t), u(t))$ can pass either above or below O_i. Now two trajectories, that pass on the same side of each of the islands O_i, are “visible” one from another. Using properties of semialgebraic sets, discussed in this book, one can join these trajectories inside S by paths of controllable lengths, and estimate the difference of the derivatives of $x(t)$. As a result, we get a differential inequality, which, in turn, implies that the endpoints of the two trajectories as above must be close to one another.

The following result is obtained in [Bri-Yom 2] by a detailed analysis on the above lines:

Denote by W_K the set of K-Lipschitz u on [0,1] with $|u(t)| \leqslant 1$, and fix the L_p-norm on the control space, $p \geq 1$.

Theorem 10.17. *Assume $x_0 = 0$ in (1.1). Let $f(x, u)$ be a polynomial of degree d, satisfying $|f(x, u)| \leqslant 1$ for $|x| \leqslant 1$, $|u| \leqslant 1$. Then for any $\gamma \geq 0$ the set of γ-critical values of J_f on W_K can be covered by $N(d) = d^\gamma 2^{3(d+1)^2}$ intervals of length $\delta = (qK)^{1/q} \gamma^{q/q+1}$. Here $1/p + 1/q = 1$.*

In particular, the measure of the γ-critical values of J_f does not exceed $N(d)\delta$.

Thus, a quantitave Sard theorem is valid for the control problems as above. On can apply the approach of “Semialgebraic Complexity” described in this chapter, and extend the result to right-hand sides more complicated than polynomials. However, since the growth of the estimate of theorem 10.17 in d is very fast, f above can be replaced only by analytic functions of a very restricted growth. Using the same considerations as above, but with an infinite number of the “islands” O_i, one can easily construct control problems of the above form with f infinitely smooth and with critical values of J covering the whole interval (see [Bri-Yom 3,4]).

The approach outlined above can be applied also in higher-dimensional control problems. In higher dimension the relation between the dynamics of near-critical trajectories and semi-algebraic geometry of the right hand side polynomials remains especially transparent for the trajectories of “rank zero”: those for which the norm of the differntial of the input to state mapping J is small. (For near-critical trajectories of higher rank their behavior is governed by a combination of semi-algebraic restrictions with the Pontryagin maximum principle, and all the considerations become more delicate).

However, also for near-critical trajectories of “rank zero” a new important dynamical problem enters: consider, for example, the case of two-dimensional x and one-dimensional u. Here as above critical trajectories of rank zero must lie in a semi-algebraic set S in a three-dimensional phase space (x, u), defined by the condition that $f_u(x, u)$ be small, where f is a (vector) right hand side of the equation (1.1). We can consider the complement of S as a set of islands O_i, but now these islands are three-dimensional bodies, and in addition to a possibility of passing above or under the island, the trajectory can rotate around it. See [Bri-Yom 3,4] for a detailed discussion of this direction.

The "combinatorial" bound on the behavior of near-critical trajectories can be saved only if the possible rotation around semialgebraic bodies can be bounded in algebraic terms. The bounds in this spirit exist, and they relate the near-critical controls with the well-known problem of bounding oscillation of polynomial vector fields. We conclude this section, presenting a recent result, obtained in [Gri-Yom]. (There exists a rich theory of nonoscillation of trajectories of algebraic vector fields (see [Yak 1,2]), which provides similar conclusions. However, geometric methods of [Gri-Yom] are well suited to extensions and applications in control).

For a Lipschitz vector field in $\mathbb{R}^n$, angular velocity of its trajectories with respect to any stationary point is bounded by the Lipschitz constant. The same is true for a rotation speed around any integral submanifold of the field. However, easy examples show that a trajectory of a $\mathcal{C}^\infty$-vector field in $\mathbb{R}^3$ can make in finite time an infinite number of turns around a straight line. We show that for a trajectory of a polynomial vector field in $\mathbb{R}^3$, its rotation rate around any algebraic curve is bounded in terms of the degree of the curve and the degree and size of the vector field. As a consequence, we obtain a linear in time bound on the number of intersections of the trajectory with any algebraic surface.

For an algebraic vector field v in $\mathbb{R}^3$ define its norm $\|v\|$ as the sum of the absolute values of the coefficients of the polynomials, defining this field.

Below we always assume that all the objects considered (trajectories of the vectorfields, algebraic curves) are contained in the unit ball B_1 in $\mathbb{R}^3$.

Theorem 10.18. *Rotation of any trajectory $w(t)$ of an algebraic vector field v in $\mathbb{R}^3$ around an algebraic curve V, between the time moments t_1 and t_2, is bounded by:*

$$C_1(d_1, d_2) + C_2(d_1, d_2)\|v\|(t_2 - t_1).$$

(Here the constants $C_1(d_1, d_2)$ and $C_2(d_1, d_2)$ depend only on the degrees d_1, d_2 of the field v and the curve V, respectively).

Theorem 10.19. *For any trajectory $w(t)$ of an algebraic vector field v, and for any algebraic surface W in $\mathbb{R}^3$, the number of intersection points of $w(t)$ with W between moments in time t_1 and t_2 is bounded by*

$$C_3(d_1, d_2) + C_4(d_1, d_2)\|v\|(t_2 - t_1).$$

We conclude this chapter with a discussion of some natural extensions of the Quantitative Sard theorem and Quantitative Transversality. These and other similar extensions may ultimately form what we call "Quantitative Singularity Theory".

10.3.7 Quantitative Singularity Theory

Quantitative Sard theorem, Quantitative Transversality and "Near Thom-Boardman Singularities" treated in Chapters 7–9 of this book definitely belong to a much wider domain which can be called "Quantitative Singularity Theory". In this section we give some examples, illustrating the possible contours of this future theory.

10.3.7.1 Quantitative Morse-Sard Theorem. Consider smooth functions $f : B^n \to \mathbb{R}$. Probably, the first and the most basic result of Singularity Theory is the Morse theorem ([Morse 1,2], [Mil 2]), describing typical singularities of f. It states that "generically" f has the following properties:

i. All the critical points x_i of f are nondegenerate (i.e. the Hessian $H(f)$ is non-degenerate at each x_i). Consequently, the number of these critical points is finite.
ii. All the critical values are distinct, i.e. $f(x_i) \neq f(x_j)$ for $i \neq j$.
iii. Near each point x_i there is a new coordinate system $y_1, ..., y_n$, centered at this point, such that

$$f(y_1, ..., y_n) = y_1^2 + y_2^2 + ... + y_l^2 - y_{l+1}^2 - y_{l+2}^2 - ... - y_n^2 + const.$$

In particular, for any given f_0 we can perturb it by an arbitrarily small (in C^∞-norm) addition h, in such a way that $f = f_0 + h$ has the properties i, ii, iii as above. Now a parallel result of Quantitative Singularity Theory in this situation is the following:

Statement 1. Fix $k \geq 3$. Let a C^k function f_0 be given with all the derivatives up to order k uniformly bounded by K. Then for any given $\epsilon > 0$, we can find h with $\|h\|_{C^k} \leqslant \epsilon$, such that for $f = f_0 + h$,

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

$$f(y_1, ..., y_n) = y_1^2 + y_2^2 + ... + y_l^2 - y_{l+1}^2 - y_{l+2}^2 - ... - y_n^2 + const.$$

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

Here $\psi_1, \psi_2, \psi_3, d$ (tending to zero as $\epsilon \to 0$) and N, M (tending to infinity) are explicitly given functions, depending only on K and ϵ. The neighborhoods U_δ of the singular points x_i play an important role in what follows. Let us call them the *controlled neighborhoods* of the corresponding singular points x_i.

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

Now consider linear functions $h : B^n \to \mathbb{R}$. Zero is a γ-near singular value of Df for $f = f_0 + h$ if and only if the point $-Dh$ is a γ-near singular value of Df_0.

At this step we apply the Quantitative Sard theorem of Chapter 9. It bounds the entropy of the near critical values of Df_0. Its result can be easily translated into the following form: *For any $r > 0$ there are $\gamma(K, r)$-regular values v of Df_0, at a distance at most r from zero.* Here $\gamma(K, r)$ is an explicitely given function, tending to zero as r tends to zero (see [Yom 3] and [Yom 18]).

For a given $\epsilon > 0$ let us pick a certain $\gamma(K, \epsilon)$-regular value v of Df_0, at a distance at most ϵ from zero, and let h be a linear function with $Dh = -v$. Then all the critical points of $f = f_0 + h$ have the Hessian with the minimal eigenvalue bounded from below by $\gamma(K, \epsilon)$. This proves the part i of the statement.

Having this lower bound for the Hessian (and the upper bound K for all the derivatives of f_0 up to order k), we can produce the bounds in ii and iii in a rather straightforward way. The part iv of the statement is obtained by a careful "quantification" of the conventional normalization procedure. See [Yom 28] for details.

Another typical result of the classical Singularity Theory is the "Stability theorem", which in the case of Morse singularities takes the following form: if f satisfies conditions i, ii, iii than any small perturbation f_1 of f is equivalent to f via the diffeomeophisms of the source and the target.

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

A parallel result of Quantitative Singularity Theory is the following:

Statement 2. Let a C^k function f be given, with all the derivatives up to order k uniformly bounded by K. Let f satisfy

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

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

Then there is $\epsilon_0 > 0$ (depending only on K, ψ_1, ψ_2) such that for any given ϵ, $\epsilon_0 > \epsilon > 0$, and for any f_1 which is closer than ϵ to f in C^k-norm, f_1 is equivalent to f via the diffeomeophisms G and H of the source and the target, respectively. G and H differ (in C^{k-1}-norm) from the identical diffeomeorphisms not more than by $s(K, \psi_1, \psi_2, \epsilon)$. Here $s(K, \psi_1, \psi_2, \epsilon)$ tends to zero as ϵ tends to zero. For the proof see [Yom 28]. The next "quantitative"

result has no direct analogy in the classical Singularity Theory. It claims that for a generic mapping each its "near-singular" point belongs to a controlled neighborhood of one of exact singular points (its "organizing center").

A more accurate statement of this result is as follows:

Statement 3. Let a C^k function f be given, with all the derivatives up to order k uniformly bounded by K. Then for any given $\epsilon > 0$, we can find h with $\|h\|_{C^k} \leqslant \epsilon$, such that for $f = f_0 + h$ the conditions i-iv of Statement 1 are satisfied, as well as the following additional condition.

v. There is $\eta(K, \epsilon) > 0$ such that for any point x if the norm of the grad$f(x)$ is smaller than $\eta(K, \epsilon)$ then x belongs to one of the controlled neighborhoods of the singular points x_i of f.

Sketch of the proof. As in the proof of Statement 1, we take h to be a linear function. The bound on the entropy of the near critical values of Df_0, provided by the Quantitative Sard theorem of Chapter 9, implies the following: *For any $r > 0$ there are the points v in $\mathbb{R}^n$, at a distance at most r from zero, such that the entire ball B in $\mathbb{R}^n$ of radius $\eta(K, r)$, centered at v, consists of $\gamma(K, r)$-regular values of Df_0.* Here $\gamma(K, r)$ and $\eta(K, r)$ are explicitely given functions, tending to zero as r tends to zero.

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

To complete the proof, we apply a "Quantitative Inverse Function Theorem" (its various forms are scattered over the literature). It shows that with our lower bound on the Hessian (and with the global bound on higher derivatives) a certain neighborhood of x is mapped by Df *onto* the ball of controlled radius in $\mathbb{R}^n$. With a proper choice of the function $\eta(K, \epsilon)$ this last ball contains the origin. This means that in a neighborhood of x there is a true singular point x_i of f. Once more, with a proper tuning of the inequalities, we get that x belongs to the controlled neighborhood of x_i. This completes the proof.

The result of Statement 3 answers the problem posed in Section 10.1.4 of this chapter (for the Morse singularities). It shows that (at least in principle) we can relate each "near-singularity" to its "organizing center". We believe that this fact (extended to a wider range of singularities and supplemented with effective and efficient estimates of the involved parameters) has a basic importance for applications of Singularity Theory. In some very special examples this was explained in Section 10.1.4. In general one can hope that a progress in this directions may transform some inspiring ideas and approaches of [Tho 3] into theorems and working algorithms.

10.3.7.2 Quantitative Singularity Theory: a General Setting. The next example, which presents some further main ideas and tools of Singularity Theory, would be the Whitney theory of singularities of mapping of the plane to itself. We believe that quantitative results in the spirit of Statements 1-3 above are still valid in this case, although the proofs become more tricky. Presumably, the same concerns the main body of the modern Singularity Theory, as it was formed in [Whi 3], [Tho 2,3], [Ma 1-8], [Boa], [Arn-Gus-Var] and many other publications. The main its tools include Sard and Transverasity theorems, Division and Preparations theorems and highly developed algebraic techniqies for classification and normalization of singularities. All these ingredients are basically "quantitative", so in principle one can expect each result of the classical Singularity Theory to exist in an explicit quantitative form. We believe that obtaining such results will not be a straightforward repetition of the existing proofs, but rather a discovery of a variety of unknown inportant phenomena.

10.3.7.3 Probabilistic Approach. Another problem posed in Section 10.1.4 of this chapter concerned a probabilistic distribution of singularities. Let us illustrate this problem by some examples. We restrict the discussion to the polynomial functions and mappings.

Let us remind Theorem 1.8, proved in the introduction to the book:

Theorem 1.8. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a polynomial of degree d. Then for any $\gamma \geq 0$ the set $\Delta(f, \gamma, r)$ can be covered by $N(n, d)$ intervals of length γr. The constant $N(n, d)$ here depends only on n and d.*

This result can be easily reinterpreted in probabilistic terms. Let c be picked randomly in the interval $[a, b]$.

Statement 4. With probability $p \geq 1 - \frac{N(n,d)\gamma}{b-a}$, c is a γ-regular value of f. Equivalently, for $f_c = f - c$, with probability $p \geq 1 - \frac{N(n,d)\gamma}{b-a}$ the following is true: at any point x satisfying $f_c(x) = 0$, the norm of the $\text{grad} f_c(x)$ is at least γ.

Following the results of Chapters 7 and 8 one can produce rather accurate probabilistic distributions of near critical values of polynomial mappings, with repect to their "degree of degeneracy". Indeed, most of the geometric estimates there are essentially sharp (up to constants) so as a result we get not only inequalities, but actual distributions.

One can also include in consideration additional geometric parameters, like distance of a certain point to the set of critical values (see [Yom 3] and Section 10.1.2, Chapter 10). On the base of the results of Chapter 9 the corresponding distributions for the case of finitely differentiable mappings can be obtained.

Having the distributions of "near-regularity" we cane produce estimates for the average complexity of the fibers of f, or for the expectation of this complexity (see [Yom 3]). Already these results (treating only the nonsingular objects, like fibers of a fixed mapping) may help in construction of certain algorithms, like "fiber tracing" (see [Yom 18]).

However, we believe that also in the genuine territory of Singularity Theory (like Morse theory, Whitney theory of plane mappings, etc.) the probabilistic distributions of the "degree of degeneracy" of various singularities can be obtained. In the case of the Morse theorem this can be done essentially in the lines of the application of the Quantitative Sard theorem, as described above (see [Yom 28]). For the Whitney theorem the situation is much more delicate, as one has to follow a geometry of the hierarchy of singularities. Theorem 8.10 of Chapter 8 above, bounding the size of the "near-cuspidal" values, provides one of the steps in this direction.

An important motivation for a probabilistic study of the distribution of singularities is provided by deep results on the asymptotics of the oscillating integrals (see, for example, [Arn-Gus-Var]). These results include, in particular, an important information on the asymptotics of the relative probabilities of different types of singularities.

In conclusion, let us mention that many other results and approaches extend the quantitative power of the analysis of singularities. In particular, see [Bie-Mi 5] as the resolution of singularities is concerned, [Bie-Mi 3,4,6] for the geometry of subanalytic sets, [Bie-Mi 1,2] and [Bie-Mi-Paw 1] for composite differentiable functions, [Mi] for the division theorem, [Guck] and [Dam 1-4] for singularities in PDE's.

Glossary

$\nabla f_{(x)}, grad(f)_{(x)}$	gradient of the function f at x	
B_r^n	n-dimensional ball of radius r	
A_δ	δ-neighbourhood of the set A	
Vol_i	i-volume in $\mathbb{R}^i$	
S^ℓ	ℓ-dimensional sphere	
$\Sigma(f,\gamma)$	γ-critical points of f	page 18
$\Delta(f,\gamma)$	γ-critical values of f	page 18
$\Sigma(f,\gamma,r)$	γ-critical points of f in B_r^n	page 18
$\Sigma(f,\gamma,r)$	γ-critical values of f in B_r^n	page 18
$M(\epsilon,A)$	minimal number of ϵ-balls to cover A	page 23
$H_\epsilon(A)$	ϵ-entropy of A	page 23
$\mu(\epsilon,\eta)$		page 24
$d_{\mathcal{H}}(A,B)$	Hausdorff distance between A and B	page 24
$\Sigma(A,B)$		page 25
$\eta(\epsilon)$		page 25
$\mathcal{S}^\beta$	β-dimensional spherical Hausdorff measure	page 27
$\mathcal{H}^\beta$	β-dimensional Hausdorff measure	page 27
$\dim_{\mathcal{H}}$	Hausdorff dimension	page 28
$\dim_e(A)$	dimension d'entropie	page 28
$C_{\frac{1}{3}}$	classical Cantor set in $[0,1]$ (with ratio $\frac{1}{3}$)	page 29
$V_\beta(A)$		page 31
G_n^k	Grassmann manifold of k-vector space of $\mathbb{R}^n$	page 33
$\gamma_{k,n}$	standard probability measure on G_n^k	page 33
$\bar{G}_n^k$	Grassmann manifold of k-affine space of $\mathbb{R}^n$	page 33
$\bar{\gamma}_{k,n}$	standard probability measure on $\bar{G}_n^k$	page 33
$V_0(A)$	number of connected components of A	page 34
$V_i(A)$	i-th variation of A	page 34
$c(n,\ell)$	universal constant	page 34
Γ	classical Euler function	page 35
$\mathbf{1}_A$	characteristic function of the set A	page 35
$\mathcal{I}_n$		page 36
$V_i(A,B)$	i-th variation of A relative to B	page 37
$D(A)$	diagramm of the semialgebraic set A	page 48
$B_0(D)$		page 49
$\widehat{B}_0(D)$		page 49
$\widetilde{B}_0(D)$		page 49

Y. Yomdin and G. Comte: LNM 1834, pp. 171–186, 2004.

	Gabrielov property	page 55
	structure on the real field	page 56
	o-minimality	page 57
	tame set	page 57
$B_{0,\ell}(A)$		page 59
Θ_ℓ	ℓ-density	page 61
$e(A,0)$	Local multiplicity at 0 of A	page 62
$\bigotimes_i \mathbb{R}^n$	tensor product of $\mathbb{R}^n$ (i times)	page 75
$\bigwedge_i \mathbb{R}^n$	i-th exterior product of $\mathbb{R}^n$	page 76
$w_i(\mathrm{L})$		page 77
$\lambda_i(\mathrm{L})$	i-th semiaxe of an ellipsoid	page 77
$\omega(P)$		page 83
$N(f_{\vert A}, y)$	number of points of $A \cap f^{-1}(\{y\})$	page 83
$\Sigma(f, \Lambda)$	Λ-near critical points of f	page 87
$\Sigma(f, \Lambda, A)$	Λ-near critical points of f in A	page 87
$\Delta(f, \Lambda, A)$	Λ-near critical values of f in A	page 87
$V_{i,\ell}(f)$	definition 7.8	page 91
$c(i, \ell, n, m)$	definition 7.8	page 91
$\omega_i(\mathrm{L})$	definition 7.9	page 93
$\Sigma(f, \Lambda, \delta)$		page 100
$\Delta(f, \Lambda, \delta)$		page 100
$\Sigma(f, \Lambda, A, \delta)$		page 102
$\Delta(f, \Lambda, A, \delta)$		page 103
$\Sigma(f, g, \Lambda, \Lambda', \delta)$		page 104
$\Delta(f, g, \Lambda, \Lambda', \delta)$		page 104
$\mathcal{T}_\delta$	tube $g^{-1}(B^q_\delta)$	page 106
$\Sigma^1(f)$		page 106
$\Sigma^1(f, A, \delta)$		page 106
$\Delta^1(f, A, \delta)$		page 107
$R_k(f)$		page 109
$\Sigma(f, \Lambda, B^n_r)$		page 109
$\Delta(f, \Lambda, B^n_r)$		page 109
Σ^ν_f	rank-ν set of critical points of f	page 112
Δ^ν_f	rank-ν set of critical values of f	page 112
$\Sigma(f, \Lambda, A_1, A_2, \delta)$		page 124
$\Delta(f, \Lambda, A_1, A_2, \delta)$		page 124
$m_h(x)$		page 132
$\sigma_s(f, \epsilon)$	definition 10.6	page 149
$\sigma_p(f, \epsilon)$		page 149
$E_p(f, d)$		page 149
$\sigma_Q(f, \epsilon)$	definition 10.9	page 154
$K(f)$	Set of generalized critical values	page 155
$K_\infty(f)$	Set of asymptotic critical values	page 155
Σ_f	Set of critical points	page 155
Δ_f	Set of critical values	page 155

References

[Agr 1] A. A. Agrachev; B. Bonnard; M. Chyba; I. Kupka, Sub-Riemannian sphere in Martinet flat case. *ESAIM Control Optim. Calc. Var.* **2**, (1997), 377-448.

[Agr 2] A. A. Agrachev, Compactness for sub-Riemannian length-minimizers and subanalyticity. Control theory and its applications *Rend. Sem. Mat. Univ. Politec. Torino* **56** (1998), no. 4, (2001), 1-12.

[Ahi] N. I. Achieser, Theory of Approximation. Dover Publications, New York, (1992).

[All] E. K. G. Allgower, Simplicial and continuation methods for approximating fixed points and solutions to systems of equations. *SIAM Review* **22**, (1980), 28-85.

[Alo] N. Alon, The number of polytopes, configurations and real matroids. *Mathematika* **33**, (1986), n. 1, 62-71.

[And-Brö-Rui] C. Andradas; L. Bröcker; J. M. Ruiz, Constructible sets in real geometry. *Ergebnisse der Mathematik und ihrer Grenzgebiete (3)*, **33**. Springer-Verlag, Berlin, (1996).

[Arn] V. I. Arnold, O. A. Oleinik, Topology of real algebraic varieties. (Russian) *Vestnik Moskov. Univ. Ser. I Mat. Mekh.* (1979), n. 6, 7-17.

[Arn-Gus-Var] V. I. Arnol'd; S. M. Gusein-Zade; A. N. Varchenko, Singularities of differentiable maps, Volumes I and II, *Monographs in Mathematics 82 (83)*, Birkhauser Boston, Inc., Boston, MA, 1985 (1988).

[Art-Mas] M. Artin; B. Mazur, On periodic points. *Ann. Math.* **81**, (1965), 82 - 99.

[Aur] D. Auroux, Asymptotically holomorphic families of symplectic submanifolds, *Geom. Funct. Anal.* **7** (1997), no. 6, 971-995.

[CBar 1] C. Bar, Zero sets of solutions to semilinear elliptic systems of first order, *Invent. Math.* **138** (1999), no. 1, 183-202.

[CBar 2] C. Bar, On nodal sets for Dirac and Laplace operators, *Commun. Math. Phys.* **188** (1997), 709-721.

[Bar-Ple 1] M. Baran; W. Pleśniak, Bernstein and van der Corput-Shaake type inequalities on semialgebraic curves. Studia Math. **125**, n. 1, 1997, 83-96.

[Bar-Ple 2] M. Baran, M.; W. Pleśniak, Polynomial inequalities on algebraic sets, *Stud. Math.* **141**, No. 3 (2000), 209-219.

[Bar-Ple 3] M. Baran; W. Pleśniak, Characterization of compact subsets of algebraic varieties in terms of Bernstein type inequalities, *Stud. Math.* **141**, No. 3 (2000), 221-234.

[Bat 1] S. M. Bates, Toward a precise smoothness hypothesis in Sard's theorem. *Proc. Amer. Math. Soc.* **117**, (1993), n. 1, 279-283.

[Bat 2] S. M. Bates, On smooth nonlinear surjections of Banach Spaces. *Isr. J. Math.* **100**, (1997), 209-220.

[Bat 3] S. M. Bates, On smooth rank 1 mappings of Banach Spaces onto the plane. *J. Diff. Geom.* **37**, n. 3, (1993), 729-733.

[Bat 4] S. M. Bates, On the image size of singular maps, I. *Proc. Amer. Math. Soc.* **114**, n. 3, (1992), 699-705.

[Bat 5] S. M. Bates, On the image size of singular maps, II. *Duke Math. J.* **68**, n. 3, (1992), 463-476.

[Bat 6] S. M. Bates, A smooth surjective rank 1 endomorphism of a Hilbert space. *Int. Math. Res. Not.* **6**, (1991), 71-73.

[Bat 7] S. M. Bates, On smooth, nonlinear surjections of Banach spaces. *Israel J. Math.*, **100**, (1997), 209-220.

[Bat-Mor] S. M. Bates, C. G. Moreira, De nouvelles perspectives sur le théorème de Morse-Sard. *C. R. Acad. Sci. Paris* **332**, (2001), Série I, 13-17.

[Bat-Nor] S. M. Bates; A. Norton, On sets of critical values in the real line. *Duke Math J.* **83**, No. 2, (1996), 399-413.

[Ben-Ris] R. Benedetti; J. J. Risler, Real algebraic and semi-algebraic sets. *Actualités Mathématiques*, Hermann, Paris, (1990).

[Ber] S. Bernstein, Sur une propriété des polynomes. *Proc. Kharkov Math. Society*, Serie 2, v.**14** (1913), 1-6.

[Bic-Yom] E. Bichuch, Y. Yomdin, High order Taylor discretization for numerical solution of parabolic PDE's, preprint, 1995.

[Bie-Mi 1] E. Bierstone; P. Milman, Composite differentiable functions, *Ann. of Math* (2) **116** (1982), no. 3, 541-558.

[Bie-Mi 2] E. Bierstone; P. Milman, Algebras of composite differentiable functions, *Singularities, Part 1 (Arcata, Calif., 1981)*, 127-136, Proc. Sympos. Pure Math., 40, *Amer. Math. Soc., Providence, RI* 1983.

[Bie-Mi 3] E. Bierstone; P. Milman, Relations between analytic functions I, *Ann. Inst. Fourier (Grenoble)* **37** (1987), no. 1, 187-239.

[Bie-Mi 4] E. Bierstone; P. Milman, Relations between analytic functions II, *Ann. Inst. Fourier (Grenoble)* **37** (1987), no. 2, 49-77.

[Bie-Mi 5] E. Bierstone; P. Milman, Semianalytic and subanalytic sets. *Inst. Hautes Études Sci. Publ. Math.* **67** (1988), 5-42.

[Bie-Mi 6] E. Bierstone; P. Milman, Canonical desingularization in characteristic zero by blowing up the maximum strata of a local invariant, *Invent. Math.* **128** (1997), no. 2, 207-302.

[Bie-Mi 7] E. Bierstone; P. Milman, Geometric and differential properties of subanalytic sets, *Ann. of Math* (2) **147** (1998), no. 3, 731-785.

[Bie-Mi-Paw 1] E. Bierstone; P. Milman, W. Pawlucki, Composite differentiable functions, *Duke Math. J.* **83** (1996), no. 3, 607-620.

[Bie-Mi-Paw 2] E. Bierstone; P. Milman, W. Pawlucki, Differentiable functions defined in closed sets. A problem of Whitney, *Invent. Math.* **151** (2003), no. 2, 329-352.

[Bla] W. Blaschke, Vorlesungen über Integralgeometrie. *3te Aufl. Deutscher Verlag der Wissenschaften,* Berlin, 1955.

[Boc-Cos-Roy] J. Bochnak; M. Coste; M. F. Roy, Real algebraic geometry. Translated from the 1987 French original. *Ergebnisse der Mathematik und ihrer Grenzgebiete (3)* **36**, Springer-Verlag, Berlin, (1998).

[Boa] J. M. Boardman, Singularities of differential maps. *Publ. Math. I.H.E.S.* **33**, (1967), 21-57.

[B-L-M-T] L. Bos; N. Levenberg; P. Milman; B.A. Taylor, Tangential Markov inequalities characterize algebraic submanifolds of C^n. *Indiana Univ. Math. J.* **44**, (1995), 115-137.

[Bos-Mi] L. Bos; P. Milman, Sobolev-Gagliardo-Nirenberg and Markov type inequalities on subanalytic domains, *Geom. Funct. Anal.* **5** (1995), no. 6, 853-923.

[B-L-M-T 2] L. Bos; N. Levenberg, P. Milman, B.A. Taylor, Tangential Markov inequalities on real algebraic varieties, *Indiana Univ. Math. J.*, **47**, (1998), no.4, 1257-1272.

[Bour-Gol-Schl] J. Bourgain; M. Goldstein; W. Schlag, Anderson localization for Schrodinger operators on Z^2 with quasi-periodic potential, *Acta Math.* **188** (2002), 41-86.

[Bou] N. Bourbaki, Algèbre. Paris, Hermann.

[Bri-Eli-Yom] M. Briskin; Y. Elichai; Y. Yomdin, How can Singularity theory help in Image processing, in *Pattern Formation in Biology, Vision and Dynamix*, M. Gromov, A. Carbone, Editors, World Scientific (2000), 392-423.

[Bri-Yom 1] M. Briskin; Y. Yomdin, Vertices of reachable sets in nonlinear control problems. Proc. 29th CDC, Honolulu, Hawaii, Dec. (1990), 2815-2816.

[Bri-Yom 2] M. Briskin; Y. Yomdin, Critical and near-critical values in polynomial control problems, I: One-dimensional case. Israel J. Math. **78**, (1992), 257-280.

[Bri-Yom 3] M. Briskin; Y. Yomdin, Extremal and near-extremal trajectories of rank zero in nonlinear control problems. in: "Optimization and Nonlinear Analysis", A. Ioffe, M. Marcus and S. Reich (eds.), Pitman Research Notes in Mathematics, Series **244**, 1992, 53-63.

[Bri-Yom 4] M. Briskin; Y. Yomdin, Semialgebraic geometry of polynomial control problems. in: Computational Algebraic Geometry, F. Eyssette; A. Galligo (eds), Birkhaüser, (1993), 21-28.

[Bri-Yom 5] M. Briskin; Y. Yomdin, Algebraic families of analytic functions, I. *J. of Diff. Equations*, Vol. **136**, n. 2, (1997), 248-267.

[Bri-Yom 6] M. Briskin; Y. Yomdin, Geometry of Images, in preparation.

[Brö] L. Bröcker, Lipschitz-Killing invariants (with an appendix by Andreas Bernig). *Preprint Münster University.*

[Brö-Kup] L. Bröcker; M. Kuppe, Integral geometry of tame sets. *Geom. Dedicata* **82**, (2000), n. 1-3, 285-323.

[Bru 1] A. Brudnyi, Bernstein-type inequality for algebraic functions. *Indiana Univ., Math. J.* **46**, 1, (1997), 95-117.

[Bru 2] A. Brudnyi, Cyclicity of Zeroes of Families of Analytic Functions, to appear.

[Bry] L. N. Bryzgalova, Singularities of a maximum of a function depending on parameters. *Funct. Anal. Appl.* **11**, (1977), 49-50.

[Buf] G. -L. Leclerc, comte de Buffon, Éssai d'arithmétique morale. *Volume des suppléments de l'édition in quarto de l'imprimerie royale* **11**, (1777).

[Cla 1] F. H. Clarke, Generalized gradients and applications. *Trans. Amer. Math. Soc.* **205**, (1975), 247-262.

[Cla 2] F. H. Clarke, On the inverse function theorem. *Pacific J. Math.* **64**, (1976), 97-102.

[Cau] A. L. Cauchy, Mémoire sur la rectification des courbes et la quadrature des surfaces courbes. *Mémoires de l'Académie des Sciences Paris* **22**, (1850), 3-15.

[Com 1] G. Comte, Sur les hypothèses du théorème de Sard pour le lieu critique de rang 0. *C. R. Acad. Sci. Paris, t.* **323**, *Série I*, (1996), 143-146.

[Com 2] G. Comte, Formule de Cauchy-Crofton pour la densité des ensembles sous-analytiques. *C. R. Acad. Sci. Paris, t.* **328**, *Série I*, (1999), 505-508.

[Com 3] G. Comte, Équisingularité réelle: nombre de Lelong et images polaires. *Ann. Scient. Éc. Norm. Sup.* **(33)**, (2000), 757-788.

[Com-Gra-Mer] G. Comte; Ph. Graftieaux; M. Merle, Équisingularité réelle II: Invariants locaux et conditions de régularité. In preparation.

[Com-Mi-Tro] G. Comte; P. Milman; D. Trotman, On Zariski's multiplicity problem, *Proc. Amer. Math. Soc.* **130** (2002), no. 7, 2045-2048.

[Cos] M. Coste, Ensembles semi-algébriques. Real algebraic geometry and quadratic forms (Rennes, (1981)), pp. 109-138, *Lecture Notes in Math.* **959**, Springer, Berlin-New York, (1982).

[Cro] W. Crofton, On the theory of local probability. *Phil. Trans. of the Royal Soc. London* **158**, (1868), 181-199.

[D'Ac] D. D'Acunto, Valeurs critiques asymptotiques d'une fonction définissable dans une structure o-minimale. *Annales Polonici Mathematici* **(LXXV.1)** (2000), 35-45.

[D'Ac-Kur] D. D'Acunto; K. Kurdyka, Bounds for gradients trajectories of definable functions with applications to robotics and semialgebraic geometry, *Preprint* (2003).

[Dam 1] J. Damon, Local Morse Theory for Solutions to the Heat Equation and Gaussian Blurring, *J. Diff. Eqns.* **115** (1995), 368-401.

[Dam 2] J. Damon, Generic properties of solutions to partial differential equations, *Arch. Rational Mech. Anal.* **140** (1997), no. 4, 353-403.

[Dam 3] J. Damon, Generic structure of two-dimensional images under Gaussian blurring, *SIAM J. Appl. Math.* **59** (1999), no. 1, 97-138.

[Dam 4] J. Damon, Properties of Ridges and Cores for Two-Dimensional Images, *J. Math. Imaging Vision* **10** (1999), no. 2, 163-174.

[Dav-Zak] A. A. Davydov; V. M. Zakalyukin, Point singularities of a conditional minimum on a three-dimensional manifold. *Tr. Mat. Inst. Steklova* **220** (1998), *Optim. Upr., Differ. Uravn. i Gladk. Optim.*, 113-129.

[Dem] J. P. Demailly, Nombres de Lelong généralisés, théorèmes d'intégralité et d'analyticité. *Acta Math.* **159**, (1987), n. 3-4, 153-169.

[Den-Kur] Z. Denkowska; K. Kurdyka, Une propriété métrique des fibres d'un sous-analytique. *C. R. Acad. Sci. Paris Série I* **299**, (1984), 799-801.

[Den-Sta] Z. Denkowska; J. Stasica, Ensembles sous-analytiques à la polonaise. Preprint, (1985).

[Dol 1] E. P. Dolgenko, The rate of a rational approximation and properties of functions. *Mat. Sb.* **56**, (1962), 403-432.

[Dol 2] E. P. Dolgenko, On the properties of functions of several variables, well approximable by rational functions. *Izv. Acad. Nauk SSSR* **26**, (1962), 641-652.

[Don 1] S. K. Donaldson, Symplectic submanifolds and almost-complex geometry. *J. Differential Geom.* **44**, (1996), n. 4, 666-705.

[Don 2] S. K. Donaldson, Lefschetz pencils on symplectic manifolds, *J. Differ. Geom* **53**, No. 2 (1999), 205-236.

[Don 3] S. K. Donaldson, Lefschetz fibrations in symplectic geometry, *ICM 98* **II** (1998), 309-314.

[Donn-Fef 1] H. Donnely; C. Fefferman, Nodal sets of eigenfunctons on Riemannian manifolds, *Invent. Math.* **93** (1988), 161-183.

[Donn-Fef 2] H. Donnely; C. Fefferman, Nodal sets of eigenfunctons of the Laplacian on surfaces, *J. Am. Math. Soc.* **3** (1990), 333-353.

[Dra] R. N. Draper, Intersection theory in analytic geometry. *Math. Ann.* **180**, (1969), 175-204.

[Dri] L. van den Dries. Tame topology and o-minimal structures. *London Mathematical Society Lecture Note Series* **248**. Cambridge University Press, (1998).

[Dri-Mac-Mar] L. van den Dries; A. Macintyre; D. Marker, The elementary theory of restricted analytic fields with exponentiation. *Ann. of Math.* **140**, (1994), 183-205.

[Dri-Mil 1] L. van den Dries; C. Miller, On the real exponential field with restricted analytic functions. *Israel J. Math* **85**, (1994), 19-56.

[Dri-Mil 2] L. van den Dries; C. Miller, Geometric categories and o-minimal structures. *Duke Math. J.* **84**, (1996), 497-540.

[Eli-Yom 1] Y. Elihai; Y. Yomdin, Flexible high-order discretization. Preprint, 1989.

[Eli-Yom 2] Y. Elihai; Y. Yomdin, Global motion planning algorithm, based on a high order discretization and on hierarchies of singularities. Proc. 28th CDC, Tampa, Florida, (1989), 1173-1174.

[Eli-Yom 3] Y. Elihai; Y. Yomdin, Flexible high order discretization of geometric data for global motion planning. Theoretical Computer Science A Vol. **157**, Elsevier Science B.V., (1996), 53-77.

[Eli-Yom 4] Y. Elihai; Y. Yomdin, Numerical inversion of plane to plane mappings, based on a high order discretization and on hierarchies of singularities. In preparation.

[Eli-Yom 5] Y. Elihai; Y. Yomdin, Normal forms representation: A technology for image compression, *SPIE Vol.1903, Image and Video Processing*, (1993), 204–214.

[Eva-Gar] L. C. Evans; R. F. Gariepy, Measure theory and fine properties of functions. *Studies in Advanced Mathematics*, CRC Press, Boca Raton, (1992), 268 p.

[Fal] K. J. Falconer, The geometry of fractal sets. *Cambridge Tracts in Mathematics* **85**, Cambridge University Press (1986).

[Fav] J. Favard, Une définition de la longueur et de l'aire. *C. R. Acad. Sci. Paris* **194**, (1932), 344-346.

[Fed 1] H. Federer, The (Φ, k) rectifiable subsets of n space. *Trans. Amer. Math. Soc.* **62**, (1947), 114-192.

[Fed 2] H. Federer, Geometric measure theory. *Grundlehren Math. Wiss.*, Vol. **153** *Springer-Verlag*, (1969).

[Fef-Nar 1] C. Fefferman; R. Narasimhan, Bernstein's inequality on algebraic curves. *Ann. Inst. Fourier, Grenoble* **43**, 5, (1993), 1319-1348.

[Fef-Nar 2] C. Fefferman; R. Narasimhan, On the polynomial-like behaviour of certain algebraic functions. *Ann. Inst. Fourier, Grenoble* **44**, 2, (1994), 1091-1179.

[Fef-Nar 3] C. Fefferman; R. Narasimhan, Bernstein's inequality and the resolution of spaces of analytic functions. *Duke Math. J.* **81**, 1, (1995), 77-98.

[Fen-Sch] S.-T. Feng; G. Schumacher, On the diameter of plane curves, *Compositio Math.* **119** (1999), no. 3, 331-334.

[Fra-Yom 1] J.-P. Françoise; Y. Yomdin, Bernstein inequality and applications to differential equations and analytic geometry. *J. of Functional Analysis*, (1997), Vol. **146**, n. 1, 1997, 185-205.

[Fra-Yom 2] J.-P. Françoise; Y. Yomdin, Projection of analytic sets and Bernstein inequalities. in *"Singularities Simposium - Lojasiewicz 70"*, B. Jakubczyk; W. Pawlucki and J. Stasica (eds.), Banach Center Publications, Vol. **44**, Warszawa (1998), 103-108.

[Fu] J. H. G. Fu, Curvatures measures of subanalytic sets, *Amer. J. Math.* **116** (1994), 819-880.

[Gab 1] A. Gabrielov, Multiplicities of zeros of polynomials on trajectories of polynomial vector fields and bounds on degree of nonholonomy. *Math. Res. Lett.* **2**, (1995), 1-15.

[Gab 2] A. Gabrielov, Multiplicities of Pfaffian intersections and the Lojasiewicz inequality. *Selecta Matematica*, New Series **1**, n. 1, (1995), 113-127.

[Gab 3] A. Gabrielov, Formal relations between analytic functions. *USSR Izv.* **7** (1973), 1056-1088.

[Gab 4] A. Gabrielov, Projections of semi-analytic sets. *Funct. Anal. Appl.* **2**, (4), (1968), 282-291.

[Gol-Gui] V. Golubitski; V. Guillemin, Stable mappings and their singularities. *Graduate Texts in Math.* **14**, (1973).

[Gri-Yak] A. Grigoriev; S. Yakovenko, Topology of generic multijet preimages and blow-up via Newton interpolation. *J. Differ. Equations* **150**, n. 2, (1998), 349-362.

[Gri-Yom] A. Grigoriev; Y. Yomdin, Rotation rate of a trajectory of an algebraic vector field around an algebraic curve. *Qual. Theory of Dyn. Systems* **2**, n. 1 (2001), 61-66.

[Gro 1] M. Gromov, On the entropy of holomorphic maps. Preprint.

[Gro 2] M. Gromov, Three remarks on geodesic dynamics and fundamental group. Preprint.

[Gro 3] M. Gromov, Entropy, homology and semialgebraic geometry. Séminaire Bourbaki, Vol. 1985/86. *Astérisque* **145-146**, (1987), 225-240.

[Gro 4] M. Gromov, Spectral geometry of semi-algebraic sets. Ann. Inst. Fourier (Grenoble) **42**, n. 1-2, (1992), 249-274.

[Gro 5] M. Gromov, Partial Differential Relations, Springer, 1986.

[Guck] J. Guckenheimer, Catastrophes and Partial Differential Equations, *Ann. Inst. Fourier* **23** (1973), 31-59.

[Had] H. Hadwiger, Vorlesungen über Inhalt, Oberfläche, und Isoperimetrie. Berlin, Springer Verlag, (1957).

[Hai] J. Hai, A counterexample to infinite-dimensional version of the Morse-Sard Theorem, *Appl. Math. J. Chinese Univ. Ser. B* **18** (2003), 64-68.

[Han-Har-Lin] Q. Han; R. Hardt; F. H. Lin, Geometric measure of singular sets of elliptic equations, *Comm. Pure Appl. Math.* **51** (1998), 1425-1443.

[Har] R. M. Hardt, Some analytic bounds for subanalytic sets. Differential geometric control theory (Houghton, Mich., (1982)), 259-267, *Progr. Math.* **27**, Birkhauser Boston, Boston, Mass., (1983).

[Har 1] R. Hardt, Triangulation of subanalytic sets and proper light subanalytic maps, *Invent. Math.* **38** (1977), 207-217.

[Har 2] R. Hardt, Semialgebraic local triviality in semialgebraic mappings, *Amer. Journal of Math.* **102** (1980), 291-302.

[Har-Sim] R. Hardt; L. Simon, Nodal sets for solutions of elliptic equations, *J. Diff. Geom.* **30** (1989), 505-522.

[Hei-Rec-Roy] J. Heintz; T. Recio; M. -F. Roy, Algorithms in real algebraic geometry and applications to computational geometry, *DIMACS Series in Discrete Mathematics and Theoretical Computer Science.* **6** (1991), 137-163.

[Hei-Roy-Sol 1] J. Heintz; M. -F. Roy; P. Solernó, On the complexity of semialgebraic sets, *Proc. Information Processing (IFIP 1989), San Francisco.* G. X. Ritter ed. North Holland, Amsterdam, (1989), 293-298.

[Hei-Roy-Sol 2] J. Heintz; M. -F. Roy; P. Solernó, Sur la complexité du principe de Tarski-Seidenberg, *Bull. Soc. Math. France* **118** (1990), 101-126.

[Hei-Roy-Sol 3] J. Heintz; M. -F. Roy; P. Solernó, Single exponential path finding in semialgebraic sets. Part II: The general case, *Algebraic geometry and its applications (West Lafayette, IN, 1990)*, Springer, New York, (1994), 449-465,.

[Hei-Roy-Sol 4] J. Heintz; M. -F. Roy; P. Solernó, Description of the connected components of a semialgebraic set in single exponential time. Discrete Comput. Geom. 11 (1994), no. 2, 121-140.

[Hen-Mer] J. -P. Henry; M. Merle, Conditions de régularité et éclatements. *Ann. Inst. Fourier (Grenoble)* **37**, (1987), 159-190.

[Hen-Mer-Sab] J. -P. Henry; M. Merle; C. Sabbah, Sur la condition de Thom stricte pour un morphisme analytique complexe. *Ann. Sci. École Norm. Sup. (4)* **17**, (1984), 227-268.

[Hir 1] H. Hironaka, Subanalytic sets, *Number Theory, algebraic geometry and commutative algebra, Kinokuniya, Tokyo* (1973), 453-493.

[Hir 2] H. Hironaka, Triangulations of algebraic sets, *Proc. of Symposia in Pure Math. Amer. Math. Soc.* **29** (1975), 165-185.

[Hof-Hof-Nad] M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof, N. Nadirashvili, Critical sets of smooth solutions to elliptic equations in dimension 3, *Indiana Univ. Math. J.* **45** (1996), 15-37.

[Ibo] A. Ibort, D. Martinez-Torres, F. Presas, On the construction of contact submanifolds with prescribed topology, *J. Differ. Geom* **56**, No. 2 (2000), 235-283.

[Iva 1] L. D. Ivanov, Variazii množhestv i funktsii. (Russian) [Variations of sets and functions] Edited by A. G. Vituškin. *Izdat. "Nauka"*, Moscow, (1975), 352 p.

[Iva 2] L. D. Ivanov, The approximation of l-smooth functions by rational ones in the integral metric. (Russian) *Izv. Akad. Nauk SSSR Ser. Mat.* **36**, (1972), 234-239.

[Izu] S. Izumiya, G.T. Kossioris, Semi-Local Classification of Geometric Singularities for Hamilton-Jacobi Equations, *J. Differ. Eqns.* **118** (1995), 166-193.

[Kal 1] V. Kaloshin, Prevalence in spaces of smooth mappings and a prevalent Kupka-Smale theorem. *Dokl. Akad. Nauk* **355**, (1997), 306-307.

[Kal 2] V. Kaloshin, An extension of the Artin-Mazur theorem. *Ann. Math. (2)* **150**, n. 2, (1999), 729-741.

[Kal 3] V. Kaloshin, Generic diffeomorphisms with superexponential growth of number of periodic points. *Comm. Math. Phys.* **211**, n. 1, (2000), 253-271.

[Kal-Hun] V. Kaloshin; B. Hunt, A stretched exponential bound on the rate of growth of the number of periodic points for prevalent diffeomorphisms, I and II. *Electron. Res. Announc. AMS* **7**, (2001), 17-27 and 28-36.

[Kho] A. G. Khovanskii, Fewnomials. AMS 1991, Providence, R.I. *Translations of mathematical monographs* **88**.

[Kla] D. Klain, A short proof of Hadwiger's characterization theorem. *Mathematika* **42**, (1995), 329-339.

[Koch-Seg-Yom1] E. Kochavi; R. Segev; Y. Yomdin, Numerical solution of field problems by nonconforming Taylor discretization, *Appl. Math. Modelling*, **15** (1991), 152–157.

[Koch-Seg-Yom2] E. Kochavi; R. Segev; Y. Yomdin, Modified algorithms for nonconforming Taylor discretization, *Computers & Structures* **49**, no. 6 (1993), 969–979.

[Kole-Bak] A. Kolesnikov; A. J. Baker, An efficient high-order Taylor weak statement formulation for the Navier-Stokes equations, *J. Comput. Phys.* **173** (2001), no. 2, 549-574.

[Kol] A. N. Kolmogorov, Asymptotic characteristics of some completely bounded metric spaces, *Dokl. Akad. Nauk SSSR* **108** (1956), 585-589.

[Kol-Tih] A. N. Kolmogorov; V. M. Tihomirov, ε-entropy and ε-capacity of sets in functional space. *Amer. Math. Soc. Transl.* **17**, (1961), 277-364.

[Kup] M. Kuppe, IntegralGeometrie Whitney-Stratifizierbarer Mengen, PhD, Münster-University, (1999).

[Kup 1] I. Kupka, Contribution à la théorie des champs génériques. *Contrib. Diff. Eqs.* **2**, (1963), 457-484. *Ibid* **3**, (1964), 411-420.

[Kup 2] I. Kupka, Counterexample to the Morse - Sard theorem in the class of infinite-dimensional manifolds. *Proc. AMS* **16**, (1965), 954-957.

[Kur] K. Kurdyka, On a subanalytic stratification satisfying a Whitney property with exponent 1. Real algebraic geometry (Rennes, 1991), *Lecture Notes in Math.* **1524**, Springer, Berlin, (1992), 316-322.

[Kur-Mos] K. Kurdyka; T. Mostowski, On the gradient conjecture of R. Thom. Preprint University of Savoie (France), (1998).

[Kur-Mos-Par] K. Kurdyka; T. Mostowski; A. Parusiǹski, Proof of the gradient conjecture of R. Thom. *Ann. of Math. (2)* **152**, (2000), n. 3, 763-792.

[Kur-Orr] K. Kurdyka; P. Orro, On geodesic distance in subanalytic set, *Revista Mat. Univ. Compl. Madrid* **10** (1997), Proceedings of Conf. Real Alg. Analyt. Geom., Segovia 1995, 173-182.

[Kur-Orr-Sim] K. Kurdyka; P. Orro; S. Simon, Semialgebraic Sard theorem for generalized critical values. *J. Differential Geom. (1)* **56**, (2000), 67-92.

[Kur-Rab] K. Kurdyka; G. Raby, Densité des ensembles sous-analytiques. *Ann. Inst. Fourier (Grenoble)* **39**, (1989), 753-771.

[Lan 1] R. Langevin, La petite musique de la géométrie intégrale. *La recherche de la vérité, Écrit. Math., ACL-Éd. Kangourou*, Paris, (1999), 117-143.

[Lan 2] R. Langevin, Introduction to integral geometry. Colóquio Brasileiro de Matemática. [21st Brazilian Mathematics Colloquium], Instituto de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, (1997), 160 p.

[Lan-Ros] R. Langevin; H. Rosenberg, Fenchel theorems for submanifolds of S^n, *Comm. Math. Helvetici* **71** (1996), 594-616.

[Lan-Shi] R. Langevin, Th. Shifrin, Polar varietes and integral geometry, *Amer. J. of Math.*, **104**, no 3 (1982), 553-605.

[Lê-Tei] D. T. Lê; B. Teissier, Variétés polaires locales et classes de Chern des variétés singulières. *Ann. of Math. (2)* **114**, (1981), 457-491.

[Leb] H. Lebesgue, Exposition d'un mémoire de M. W. Crofton. *Nouv. Ann. Math. (4)* **12**, (1912), 481-502.

[Leo-Mel] A. M. Leontovič; M. S. Melnikov, On the boundedness of variations of a manifold. (Russian) *Trudy Moskov. Mat. Obšč.* **14**, (1965), 306-337.

[Loi] T. L. Loi, Verdier and strict Thom stratifications in o-minimal structures. *Illinois J. Math.* **42**, (1998), n. 2, 347-356.

[Loj] S. Lojasiewicz, Ensembles semi-analytiques. Preprint, (1965).

[Lor 1] G. G. Lorentz, Approximation of functions. Second edition. *Chelsea Publishing Co.*, New York, (1986), 188 p.

[Lor 2] G. G. Lorentz. Approximation of functions. Holt, Rinehart and Winston, New York-Chicago, Toronto, Ont. (1966).

[Lor 3] G. G. Lorentz. Metric entropy and approximation. *Bull. Amer. Math. Soc.* **72**, (1966), 903-937.

[Lor 4] G. G. Lorentz. Lower bounds for the degree of approximation. *Trans. Amer. Math. Soc.* **97**, (1960), 25-34.

[Ma 1] J. Mather, Notes on topological stability. *Harvard University*, (1970).

[Ma 2] J. Mather, Stability of C^∞ Mappings, I: The division theorem, *Ann. of Math.*, **87**, no. 1 (1968), 89-104.

[Ma 3] J. Mather, Stability of C^∞ Mappings, II: Infinitesimal stability implies stability, *Ann. of Math.*, **89**, no. 2 (1969), 254-291.

[Ma 4] J. Mather, Stability of C^∞ Mappings, III: Finitely determined map-germs, *Publ. Math. IHES* **35** (1968), 127-156.

[Ma 5] J. Mather, Stability of C^∞ Mappings, IV: Classification of stable germs by R-algebras, *Publ. Math. IHES* **37** (1969), 223-248.

[Ma 6] J. Mather, Stability of C^∞ Mappings, V: Transversality, *Adv. Math* **4**, no. 3 (1970), 301-336.

[Ma 7] J. Mather, Stability of C^∞ Mappings, VI: The nice dimensions, *Liverpool Singularities I* (1971), 207-253.

[Ma 8] J. Mather, On Nirenberg's Proof of the Malgrange Preparation Theorem, *Liverpool Singularities I* (1971), 116-120.

[Mat] V. I. Matov, Topological classification of germs of maximum and minimax functions generic families. *Uspekhi Mat. Nauk* **37**, 4, (1982), 167-168.

[Mic-Mi] C. Micchelli; P. Milman, A formula for Kergin intrpolation in R^k, *J. Approx. Theory* **29** (1980), no. 4, 294-296.

[Mi] P. Milman, The Malgrange-Mather division theorem, *Topology* **16** (1977), no. 4, 395-401.

[Mil 1] J. Milnor. On the Betti numbers of real varieties. *Proc. Amer. Math. Soc.* **15**, (1964), 275-280.

[Mil 2] J. Milnor, Morse Theory, Princeton Univ. Press, Princeton, N.J. 1963.

[Min] H. Minkowski, Sur les surfaces convexes fermées, *C. R. Acad. Sci. Paris, t. 132,*, (1901), 21-24.

[Mon] R. Montgomery. A tour of subriemannian geometries, their geodesics and applications. *Math. Surveys and Monographs* **91**, AMS, (2002).

[Mor] C. G. T. de A. Moreira, Hausdorff measures and the Morse-Sard theorem. *Publ. Mat.* **45**, (2001), n. 1, 149-162.

[Mors] A. Morse, The behavior of a function on its critical set. *Ann. Math.* **40**, (1939), 62-70.

[Morse 1] M. Morse, The critical points of a function of n variables, *Trans. Amer. Math. Soc.*, **33**, (1931), 71-91.

[Morse 2] M. Morse, The calculus of variations in the large, New York, 1934.

[Noi-Tro] L. Noirel; D. Trotman, Subanalytic and semialgebraic realisations of abstract stratified sets, *Real algebraic geometry and ordered structures (Baton Rouge, LA, 1966)*, 203-207, Contemp. Math., 253, *Amer. Math. Soc., Providence, RI*, 2000.

[Nor 1] A. Norton, The Zygmund Morse-Sard theorem. *J. Geom. Anal.* **4**, n. 3, (1994), 403-424.

[Nor 2] A. Norton, A critical set with nonnull image has large Hausdorff dimension. *Trans. Amer. Math. Soc.* **296**, (1986), 367-376.

[Nor 3] A. Norton, Functions non constant on fractal quasi-arcs of critical points. *Proc. Amer. Math. Soc.* **106**, n. 2, (1989), 397-405.

[Nor 4] A. Norton, A C-infinity function with an interval of critical values. *Indiana Univ. Math. J.* **40**, n. 4, (1991), 1483-1488.

[Nor-Pug] A. Norton; Ch. Pugh, Critical sets in the plane. *Mich. Math. J.* **38**, n. 3, (1991), 441-459.

[Pas] L. de Pascale, The Morse-Sard Theorem in Sobolev Spaces. *Ind. Univ. Math. J.* **50**, n. 3, (2001).

[Pet-Ole 1] I. G. Petrovskii; O. A. Oleinik, On the topology of real algebraic surfaces. (Russian) *Izvestiya Akad. Nauk SSSR.* Ser. Mat. **13**, (1949), 389-402.

[Pet-Ole 2] I. G. Petrovskii; O. A. Oleinik, On the topology of real algebraic surfaces. *Amer. Math. Soc. Translation* **70** (1952), 20 p. Translated from Izvestiya Akad. Nauk SSSR. Ser. Mat. 13, 389-402 (1949);

[Pey] J. Peyrieres, Introduction aux mesures et dimension de packing, in *Fractals, dimensions non entieres et applications*, Scientific Editor G. Cherbit, Masson, Paris 1987.

[Rab] P. J. Rabier, Ehresmann fibrations and Palais-Smale conditions for morphisms of Finsler manifolds. *Ann. of Math. (2)*, **146**, n. 3, (1997), 647–691.

[Ren 1] J. Renegar, On the computational complexity and geometry of the first-order theory of the reals. I. Introduction. Preliminaries. The geometry of semi-algebraic sets. The decision problem for the existential theory of the reals. *J. Symbolic Comput.* **13**, no. 3, (1992), 255-299.

[Ren 2] J. Renegar, On the computational complexity and geometry of the first-order theory of the reals. II. The general decision problem. Preliminaries for quantifier elimination. *J. Symbolic Comput.* **13**, no. 3 (1992), 301-327.

[Ren 3] J. Renegar, On the computational complexity and geometry of the first-order theory of the reals. III. Quantifier elimination. *J. Symbolic Comput.* **13**, no. 3, (1992), 329-352.

[Rie 1] J. Rieger, Generic properties of edges and corners an smooth greyvalue surfaces, *Biol. Cybernet.* **66** (1992), 497-502.

[Rie 2] J. Rieger, Generic evolution of edges on families of diffused greyvalue surfaces, *J. Math. Imag. Vision* **5** (1995), 207-217.

[Ris 1] J. J. Risler. Complexité et géométrie réelle (d'après A. Khovansky). Séminaire Bourbaki, Vol. 1984/85. *Astérisque* **133-134**, (1986), 89-100.

[Ris 2] J. J. Risler. Additive complexity and zeros of real polynomials. *SIAM J. Comput.* **14**, (1985), 178-183.

[Roc] R. T. Rockafeller, Favorable classes of Lipschitz continuous functions in subgradient optimization. In *Progress in Nondifferentiable Optimization*, E. Nurminski,ed. Pergamon Press, New-York, (1981).

[Roh 1] A. Rohde, On Sard's theorem for nonsmooth functions. *Numer. Funct. Anal. Optim.* **18**, (1997), n. 9-10, 1023-1039.

[Roh 2] A. Rohde, On the density of near-critical values of continuous selections and maximum functions. *Approximation and optimization in the Caribbean, II (Havana, 1993)*, 569-581, Approx. Optim., 9, Lang, Frankfurt am Main, (1997).

[Roh 3] A. Rohde, On the ϵ - entropy of nearly critical values. *J. Approx. Theory* **76**, (1994), n. 2, 166-176.

[Roy 1] N. Roytvarf, Generalized moments, composition of polynomials and Bernstein classes. In "Entire functions in modern analysis. B. Ya. Levin memorial volume", *Isr. Math. Conf. Proc.* **15**, (2001), 339-355.

[Roy 2] N. Roytvarf, Markov type inequality for functional determinants. *Qual. Theory of Dyn. Systems* **2**, n. 1, (2001), 79-92.

[Roy-Yom] N. Roytvarf; Y. Yomdin, Bernstein Classes. *Annales De l'Institut Fourier, Grenoble* **47** (1997), 3, 825-858.

[San] L. A. Santalo. Integral geometry and geometric probability. With a foreword by Mark Kac. *Encyclopedia of Mathematics and its Applications* Vol. **1**. Addison-Wesley Publishing Co., Reading, Mass.-London-Amsterdam, (1976).

[Sar 1] A. Sard. The measure of the critical values of differentiable maps. *Bull. Amer. Math. Soc.* **48**, (1942), 883-890.

[Sar 2] A. Sard, Images of critical sets. *Ann. Math.* **68**, (1958), 247-259.

[Sar 3] A. Sard, Hausdorff measure of critical images on Banach manifolds. *Amer. J. Math.* **87**, (1965), 158-174.

[Sch] R. Schneider, Convex bodies: the Brunn-Minkowski theory. *Encyclopedia of Mathematics and its Applications* **44**. Cambridge University Press, Cambridge, (1993), 490 p.

[Sch-McM] R. Schneider; P. McMullen, Valuations on convex bodies. *Convexity and its applications*, Birkhäuser, Basel, (1983), 170-247.

[Sch-Sha] J. T. Schwartz; M. Sharir, On the "piano movers" problem II. General techniques for computing topological properties of real algebraic manifolds. *Adv. Appl. Math.* **4**, (1983), 298-351.

[Shap-Yom] A. Shapiro; Y. Yomdin, On Functions representable as a difference of two convex functions and necessary conditions in a constrained optimization. Preprint, Beer-Sheva, (1981).

[Sham-Yom] T. Shamir; Y. Yomdin, Repeatability of redundant manipulators: mathematical solution of the problem, *IEEE Trans. on Automatic Control*, **AC-33**, No. 11 (1988), 1004–1009.

[Shi] M. Shiota. Geometry of subanalytic and semialgebraic sets. *Progress in Mathematics* **150**, Birkhäuser Boston, (1997).

[Shu-Sma 1] M. Shub; S. Smale. Computational complexity. On the geometry of polynomials and a theory of cost. I. *Ann. Sci. éc. Norm. Sup.* $4^{\text{ème}}$ Série, t. **18** (1985), n. 1, 107–142.

[Shu-Sma 2] M. Shub; S. Smale, Computational complexity: on the geometry of polynomials and a theory of cost. II. *SIAM J. Comput.* **15** (1986), n. 1, 145-161.

[Shu-Sma 3] M. Shub; S. Smale, Complexity of Bezout's Theorem. I: Geometric aspects. *J. Am. Math. Soc.* **6**, n. 2, (1993), 459-501.

[Shu-Sma 4] M. Shub; S. Smale, Complexity of Bezout's Theorem. II: Volumes and probabilities. *Birkhauser Prog. Math.* **109**, (1993), 267-265.

[Shu-Sma 5] M. Shub; S. Smale, Complexity of Bezout's Theorem. III: Condition number and packing. *J. Complexity* **9**, n. 1, (1993), 4-14.

[Shu-Sma 6] M. Shub; S. Smale, Complexity of Bezout's Theorem. IV: Probability of success, extensions. *SIAM J. Number Anal.* **33**, n. 1, (1996), 128-148.

[Shu-Sma 7] M. Shub; S. Smale, Complexity of Bezout's Theorem. V: Polynomial time. *Theor. Comput. Sci.* **133**, n. 1, (1994), 141-164.

[Sik] J.-C. Sikorav, Construction de sous-variétés symplectiques (d'après S. K. Donaldson et D. Auroux), *Sḿinaire Bourbaki, Vol. 1997/98. Astérisque* No. 252 (1998), Exp. No. 844, 4, 231-253.

[Sla 1] V. V. Slavskii, Integral-Geometric relation with orthogonal projection for multidimensional surfaces, *Dokl. Akad. Nauk SSSR* **214** (1974), 48-51.

[Sla 2] V. V. Slavskii, Integral-Geometric relation with orthogonal projection for hypersurfaces, *Sibirsk. Mat. Z.* **16** (1975), 103-123.

[Sla 3] V. V. Slavskii, Integral-Geometric relation with orthogonal projection for surfaces, *Sibirsk. Mat. Z.* **16** (1975), 355-367.

[Sma 1] S. Smale, Stable manifolds for differential equations and diffeomorphisms. *Ann. Sc. Norm. Sup. Pisa* **18**, (1963), 97-116.

[Sma 2] S. Smale, An infinite dimensional version of Sard's theorem. *Amer. J. Math.* **87**, (1965), 861-866.

[Sma 3] S. Smale, On the efficiency of algorithms of analysis. *Bull. Amer. Math. Soc.* **13**, (1985), 87-121.

[Sma 4] S. Smale, Complexity theory and numerical analysis. *Acta umerica.* **6**, (1997), 523-551.

[Sta] J. Stasica, The Whitney condition for subanalytic sets. *Zeszyty Nauk. Uniw. Jagielloń. Prace Mat.* **23**, (1982), 211-221.

[Ste] J. Steiner, Über parallele Flächen. *Monatsber. preuss. Akad. Wiss.*, Berlin 114-118. *Ges. Werke, Vol. 2, Reimer,* Berlin (1882), 171-176.

[Sus] H. J. Sussman, Lie backets, real analyticity and geometric control. in *Differential Geometric Control Theory*, 1-116.

[Tan-Yom] A. Tannenbaum; Y. Yomdin, Robotic manipulators and the geometry of real semialgebraic sets. *IEEE Trans. on Robotics and Automation* **3,4**, (1987), 301-307.

[Tao] T. Tao, From rotating needles to stability of waves: emerging connections between Combinatorics, Analysis, and PDE. *Notices Amer. Math. Soc.* **18**, n. 3, (2001), 294-303.

[Tar] A. Tarski, A Decision Method for Elementary Algebra and Geometry. 2nd ed., *Univ. California Press*, Berkeley, 1951.

[Tei 1] B. Teissier, Sur trois questions de finitude en géométrie analytique réelle. (Appendix to: F. Trèves, On the local solvability and the local integrability of systems of vector fields.) *Acta Math.* **151**, (1983), n. 1-2, 39-48.

[Tei 2] B. Teissier, Tame and stratified objects. *Geometric Galois Actions, 1. Around Grothendieck's esquisse d'un programme. London Math. Soc. Lecture Note Series* **242**, (1997), 231-242.

[Tho 1] R. Thom. Sur l'homologie des variétés algébriques réelles. Differential and Combinatorial Topology (A Symposium in Honor of Marston Morse), Princeton Univ. Press, Princeton, N.J., (1965), 255-265.

[Tho 2] R. Thom, Ensembles et morphismes stratifiés. *Bull. Amer. Math. Soc.* **75**, (1969), 240-284.

[Tho 3] R. Thom, Stabilité structurelle et morphogénese, *W. A. Benjamin, Inc.*, 1972. English edition: R. Thom, Structural Stability and Morphogenesis, *Benjamin-Cumming, Inc.*, London-Amsterdam-Sydney-Tokyo, 1975

[Tih 1] V. M. Tihomirov, Diameters of sets in functional spaces and the theory of best approximations. *Russian Math. Surveys* **15**, (1960), n. 3, 75-111.

[Tih 2] V. M. Tihomirov, Kolmogorov's work on the ϵ-entropy of functional classes and superposition of functions. *Russian Math. Surveys* **18**, (1963), n. 5, 51-87.

[Tri] C. Tricot, Douze définitions de la densité logarithmique. *C. R. Acad. Sci. Paris Série. I Math.* **293**, (1981), n. 11, 549-552.

[Vit 1] A. G. Vitushkin, O mnogomernyh Variaziyah, Gostehisdat, Moskow, (1955).

[Vit 2] A. G. Vitushkin, Ozenka sloznosti zadachi tabulirovaniya, Fizmatgiz, Moskow, 1959. Translation: Theory of the transmission and processing of information. Pergamon Press, (1961).

[Vit 3] A. G. Vitushkin, On representation of functions by means of superposition and related topics, *Enseignement Math.* **23**, (1977), n. 3-4, 255-320.

[War] H. E. Warren, Lower bounds for approximation by nonlinear manifolds. *Trans. Amer. Math. Soc.* **133**, (1968), 167-178.

[Wey] H. Weyl, On the volume of tubes. *Amer. J. Math.* **61**, (1939), 461-472.

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

[Whi 2] H. Whitney, Elementary structure of real algebraic varieties. *Ann. Math.* **66**, (1957), 545-556.

[Whi 3] H. Whitney, On singularities of mappings of Euclidean spaces I, Mappings of the plane into the plane. *Ann. Math.* **62**, (1955), 374-410.

[Whi 4] H. Whitney, Complex analytic varieties. *Addison-Wesley*, (1972).

[Whi 5] H. Whitney, Differentiable functions defined in arbitrary subsets of Euclidean space, *Trans. Amer. Math. Soc.* **40** (1936), no. 2, 309-317.

[Whi 6] H. Whitney, On the extension of differentiable functions, *Bull. Amer. Math. Soc.* **50** (1944), 76-81.

[Wie-Yom] Z. Wiener; Y. Yomdin, From formal numerical solutions of elliptic PDE's to the true ones. *Mathematics of Computations* **69**, (2000), n. 229, 197-235.

[Wil] A. J. Wilkie, Model completeness results for expansions of the ordered field of real numbers by restricted Pfaffian functions and the exponential function. *J. Amer. Math. Soc.* **9**, (1996), 1051-1094.

[Yak 1] S. Yakovenko, On zeroes of functions from Bernstein classes. *Nonlinearity* **13**, n. 4, (2000), 1087-1094.

[Yak 2] S. Yakovenko, On functions and curves defined by ordinary differential equations. *Proceedings of the Arnoldfest, Toronto, 1997, Fields Inst. Commun., 24*, AMS, Providence, RI, (1999), 497-525.

[Yom 1] Y. Yomdin, The geometry of critical and near-critical values of differentiable mappings. *Math. Ann.* **264**, (1983), n. 4. 495-515.

[Yom 2] Y. Yomdin, The set of zeroes of an "almost polynomial" function. *Proc. Amer. Math. Soc.* **90** , (1984), n. 4, 538-542.

[Yom 3] Y. Yomdin, Global bounds for the Betti numbers of regular fibers of differentiable mappings. *Topology* **24** , (1985), n. 2, 145-152.

[Yom 4] Y. Yomdin, A quantitative version of the Kupka–Smale theorem. *Ergodic Theory Dynamical Systems* **5**, (1985), n. 3, 449-472.

[Yom 5] Y. Yomdin, Metric properties of semialgebraic sets and mappings and their applications in smooth analysis. Proceedings of the Second International Conference on Algebraic Geometry, La Rabida, Spain, 1984, J.M. Aroca, T. Sahcez-Geralda, J.L. Vicente, eds., Travaux en Course, Hermann, Paris (1987), 165-183.

[Yom 6] Y. Yomdin, Volume growth and entropy. *Israel J. Math.* **57**, (1987), n. 3, 285-300.

[Yom 7] Y. Yomdin, C^k-resolution of semialgebraic mappings. Addendum to: "Volume growth and entropy". *Israel J. Math.* **57**, n. 3, (1987), 301–317.

[Yom 8] Y. Yomdin, On functions representable as a supremum of a family of smooth functions. II. *SIAM J. Math. Anal.* **17**, n. 4, (1986), 961-969.

[Yom 9] Y. Yomdin, On functions representable as a supremum of a family of smooth functions. *SIAM J. Math. Anal.* **14** , n. 2, (1983), 239-246.

[Yom 10] Y. Yomdin, Critical values and representations of functions by means of compositions. *SIAM J. on Math. Anal.* **17**, n.1, (1986), 236-239.

[Yom 11] Y. Yomdin, Applications of a generalized differential of lipschitzian functions in some results of Singularity Theory, I and II. Preprints, Beer-Sheva, (1980).

[Yom 12] Y. Yomdin, Some results on finite determinacy and stability, not requiring the explicite use of smoothness. In: *Singularities, Part 2 (Arcata, Calif., 1981), Proc. Symp. Pure Math.* **40**, A.M.S. Providence, RI, (1983), 667-674.

[Yom 13] Y. Yomdin, Beta-spread of sets in metric spaces and critical values of smooth functions. Preprint, MPI Bonn, (1983).

[Yom 14] Y. Yomdin, On represntability of convex functions as maxima of linear families. Preprint, MPI Bonn, (1983).

[Yom 15] Y. Yomdin, Maxima of smooth families, III: Morse-Sard Theorem. Preprint, MPI Bonn, (1984).

[Yom 16] Y. Yomdin, Approximational complexity of functions and entropy. *Proceedings of the Israel Mathematical Union Conference* (Tel Aviv, 1987), 1-9, Tel Aviv Univ., Tel Aviv, (1987).

[Yom 17] Y. Yomdin, Approximational complexity of functions. *Geometric aspects of functional analysis* (1986/87), 21-43, *Lecture Notes in Math.*, **1317**, Springer, Berlin-New York, (1988).

[Yom 18] Y. Yomdin, Sard's theorem and its improved versions in numerical analysis. *Proc. A.M.S. Conf., Colorado, 1988. Lectures in Applied Math.* **29**, (1990), 701-706.

[Yom 19] Y. Yomdin, Surjective mappings whose differential is nowhere surjective. *Proc. A.M.S.*, Vol. **111**, n.1, (1991), 267-270.

[Yom 20] Y. Yomdin, Complexity of functions: some questions, conjectures and results. *J. of Complexity*, **7**, (1991), 70-96.

[Yom 21] Y. Yomdin, Local complexity growth for iterations of real analytic mappings and semi-continuity moduli of the entropy. *Erg. Theory and Dynamical Systems*, **11**, (1991), 583-602.

[Yom 22] Y. Yomdin, Global finiteness properties of analytic families and algebra of their Taylor coefficients. *Proceedings of the Arnoldfest, Toronto, 1997, Fields Inst. Commun., 24*, AMS, Providence, RI, (1999), 527-555.

[Yom 23] Y. Yomdin, Analytic reparametrization of semialgebraic sets and functions. In preparation.

[Yom 24] Y. Yomdin, Semialgebraic complexity of functions. *To appear in J. of Complexity.*

[Yom 25] Y. Yomdin, On the local structure of a generic central set, *Compositio Math.* **43** (1981), no. 2, 225-238.

[Yom 26] Y. Yomdin, Quantitative Singularity theory *A summary of a course given at the Eur. Math. Union Summer School, Cluj, Romania*, 1998, preprint.

[Yom 27] Y. Yomdin, Jet Representation and Processing of functions, in preparation.

[Yom 28] Y. Yomdin, Quantitative Morse Theorem, in preparation.

[Y-E-B-S] Y. Yomdin; Y. Elichai; S. Birman; E. Spiegel, Synthetic textures in MPEG, to appear.

[Zel-Zh] I. Zelenko; M. Zhitomirskii, Rigid curves of generic 2-distributions on 3-manifolds, *Duke Math. Journal* **79**, no. 2 (1995), 281-307.

[Zer] M. Zerner, Approximation par des variétés algébriques dans les espaces Hilbertiens. *Studia Mathematica* **t. 38**, (1970), 413-435.

[Zh] M. Zhitomirskii, Typical singularities of differential 1-forms and Pfaffian equations, *Translations of Math. Monographs*, Vol. 113, Amer. Math. Soc., Providence, 1992.