Gradient inequalities, generalizations and quantitative aspects

Krzysztof Kurdyka (Université Savoie Mont Blanc)

GKLW workshop S. Łojasiewicz in memory

IMPAN Warszawa December 2022

By Jarosława K.

Program

- **Lecture 1: Talweg and the length of gradient trajectories.**
- **Lecture 2: Effective estimates for the length of gradient trajectories in the polynomial case, also for discrete trajectories.**

Program

- **Lecture 1: Talweg and the length of gradient trajectories.**
- **Lecture 2: Effective estimates for the length of gradient trajectories in the polynomial case, also for discrete trajectories.**
- **Lecture 3: From Talweg to KL-inequality and the Łojasiewicz gradient inequality.**

Program

- **Lecture 1: Talweg and the length of gradient trajectories.**
- **Lecture 2: Effective estimates for the length of gradient trajectories in the polynomial case, also for discrete trajectories.**
- **Lecture 3: From Talweg to KL-inequality and the Łojasiewicz gradient inequality.**
- **Lecture 4: KL inequality for definable maps**

Program

- **Lecture 1: Talweg and the length of gradient trajectories.**
- **Lecture 2: Effective estimates for the length of gradient trajectories in the polynomial case, also for discrete trajectories.**
- **Lecture 3: From Talweg to KL-inequality and the Łojasiewicz gradient inequality.**
- **Lecture 4: KL inequality for definable maps**

Program

- **Lecture 1: Talweg and the length of gradient trajectories.**
- **Lecture 2: Effective estimates for the length of gradient trajectories in the polynomial case, also for discrete trajectories.**
- **Lecture 3: From Talweg to KL-inequality and the Łojasiewicz gradient inequality.**
- **Lecture 4: KL inequality for definable maps**

Let U be open in $\mathbb{R}^n$ or a Riemannian manifold. Let $f : U \to \mathbb{R}$ be a C^2 function. Set $\varphi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ defined by

$$\varphi(t) = \inf\{|\nabla f(x)| : x \in f^{-1}(t)\}$$

if $t \in f(U)$, and $\varphi(t) = +\infty$ otherwise. We say that $t_0 \in \mathbb{R}$ is a typical value of f if there exists $c_0 > 0$ such that $\varphi(t) \geqslant c_0$ in some neighborhood of t_0. The complement in $\mathbb{R}$ of the set of typical values of f is the set of generalized critical values of f denoted by $K(f)$. Clearly $K_0(f) \subset K(f)$, where $K_0(f)$ stands for the set of all critical values of f, but in general

$$K_0(f) \neq K(f).$$

For instance if f extends to a neighborhood of $\overline{U}$ in such a way that the extension has a critical point at $x \in \partial U$, but f has no critical points in U, then $f(x) \in K(f) \setminus K_0(f)$.

Exercise . Prove that

$$K(f) = \{t \in \mathbb{R} : \exists_{x_\nu \in U}\, f(x_\nu) \to t,\ |\nabla f(x_\nu)| \to 0,\ as\, \nu \to +\infty\}.$$

Let $\varepsilon \geq 0$, the set

$$V_\varepsilon(f) = \{x \in U : |\nabla f(x)| \leqslant (1+\varepsilon)\varphi(f(x))\}$$

is called the ε-ridge/valley set of f. Clearly $V_\varepsilon(f)$ contains the set of all critical points of f, of course $V_\varepsilon(f)$ depends on U.

Assume the set $K(f)$ of generalized critical values of f, is finite. Let

$$\Gamma^{\varepsilon}(f) \subset V_{\varepsilon}(f)$$

be a curve satisfying the following properties:

(i) $\Gamma^{\varepsilon}(f)$ is a finite union of smooth connected curves and points;

(ii) for any $t \in f(U) \setminus K(f)$ the set $f^{-1}(t) \cap \Gamma^{\varepsilon}(f)$ consists of exactly one point;

(iii) $\Gamma^{\varepsilon}(f)$ intersects transversally $f^{-1}(t)$ for all but finitely many $t \in f(U)$.

Such a curve $\Gamma^{\varepsilon}(f)$ will be called an ε-talweg of f. Note that the ε-talweg of f depends on U, the domain of f.

Let $0 < \tau \leq 1$. We say that a C^1 curve $x : I \to U$ is a τ-trajectory of ∇f if $x'(t) \neq 0$ for each $t \in I$ and

$$\langle \nabla f(x(t)), x'(t) \rangle \geq \tau |\nabla f(x(t))| \, |x'(t)|, \qquad t \in I. \tag{1}$$

In other words, the cosine of the angle between $\nabla f(x(t))$ and $x'(t)$ is greater than or equal to τ. Clearly a 1-trajectory of ∇f is a trajectory of ∇f in the standard sense.

Theorem (Generalized Comparison Principle)

Assume that $f : U \to \mathbb{R}$ is a C^2 function defined in an open subset U of $\mathbb{R}^n$ and $0 < \tau \leq 1$. Let $x : I \to U$ be a τ-trajectory of ∇f. Then the length of $x(t)$ is bounded by

$$\frac{1+\varepsilon}{\tau} \, Length \, \Gamma^{\varepsilon}(f).$$

Comparison Principle. *The length of any trajectory of ∇f is bounded by $(1+\varepsilon)$ $Length \, \Gamma^{\varepsilon}(f)$.*

Let $0 < \tau \leq 1$. We say that a C^1 curve $x : I \to U$ is a τ-trajectory of ∇f if $x'(t) \neq 0$ for each $t \in I$ and

$$\langle \nabla f(x(t)), x'(t) \rangle \geq \tau |\nabla f(x(t))| \, |x'(t)|, \qquad t \in I. \tag{1}$$

In other words, the cosine of the angle between $\nabla f(x(t))$ and $x'(t)$ is greater than or equal to τ. Clearly a 1-trajectory of ∇f is a trajectory of ∇f in the standard sense.

Theorem (Generalized Comparison Principle)

Assume that $f : U \to \mathbb{R}$ is a C^2 function defined in an open subset U of $\mathbb{R}^n$ and $0 < \tau \leq 1$. Let $x : I \to U$ be a τ-trajectory of ∇f. Then the length of $x(t)$ is bounded by

$$\frac{1+\varepsilon}{\tau} \, \mathit{Length}\, \Gamma^{\varepsilon}(f).$$

Comparison Principle. *The length of any trajectory of ∇f is bounded by $(1+\varepsilon)\, \mathit{Length}\, \Gamma^{\varepsilon}(f)$.*

Since $\Gamma^{\varepsilon}(f)$ meets transversally all but a finite number of fibers of f, we will assume that the curve $\Gamma^{\varepsilon}(f)$ is smooth, connected and transverse to every fiber of f. Moreover by deleting finitely many fibers $f^{-1}(t)$, $t \in K(f)$, we may assume that f has no generalized critical values in U.
Let $x : I \to U$ be a τ-trajectory of ∇f; note that it is an embedding of I into U. So, $X := x(I)$ is a smooth curve. Let $x(s)$ be the arc-length parametrization of X of ∇f and let $\gamma(u)$ be the arc-length parametrization of the curve $\Gamma^{\varepsilon}(f)$. We fix orientations so that both functions $s \mapsto (f \circ x)(s)$ and $u \mapsto (f \circ \gamma)(u)$ are strictly increasing.

Let $\eta : X \to \Gamma^\varepsilon(f)$ be given by $\eta = (f|_{\Gamma^\varepsilon(f)})^{-1} \circ (f|_X)$. We now compute η in our arc-length charts, that is, we consider $h(s) = \gamma^{-1} \circ \eta \circ x(s)$. Clearly, to prove Theorem 1 it is enough to show that $h'(s) \geqslant \frac{\tau}{1+\varepsilon}$. Taking the derivative with respect to s in the equality $(f \circ \eta \circ x)(s) = (f \circ x)(s)$ we obtain

$$\langle \nabla f((\eta \circ x)(s)), (\eta \circ x)'(s) \rangle = \langle \nabla f(x(s)), x'(s) \rangle.$$

Since $x(s)$ is a τ-trajectory of ∇f, by (1) we have

$$\langle \nabla f(x(s)), x'(s) \rangle \geq \tau |\nabla f(x(s))||x'(s)| = \tau |\nabla f(x(s))|$$

since $|x'(s)| = 1$. Therefore, by the Cauchy–Schwarz inequality

$$|\nabla f((\eta \circ x)(s))| \cdot |(\eta \circ x)'(s)| \geqslant \tau |\nabla f(x(s))|.$$

Since $\eta(x(s)) \in V_\varepsilon(f)$, we have

$$(1+\varepsilon)|\nabla f(x(s))| \geq |\nabla f((\eta \circ x)(s))|,$$

thus $|(\eta \circ x)'(s)| \geqslant \frac{\tau}{1+\varepsilon}$. But γ is an arc-length parametrisation, so

$$h'(s) = (\gamma^{-1} \circ \eta \circ x)'(s) = |(\eta \circ x)'(s)| \geq \frac{\tau}{1+\varepsilon},$$

and Theorem follows.

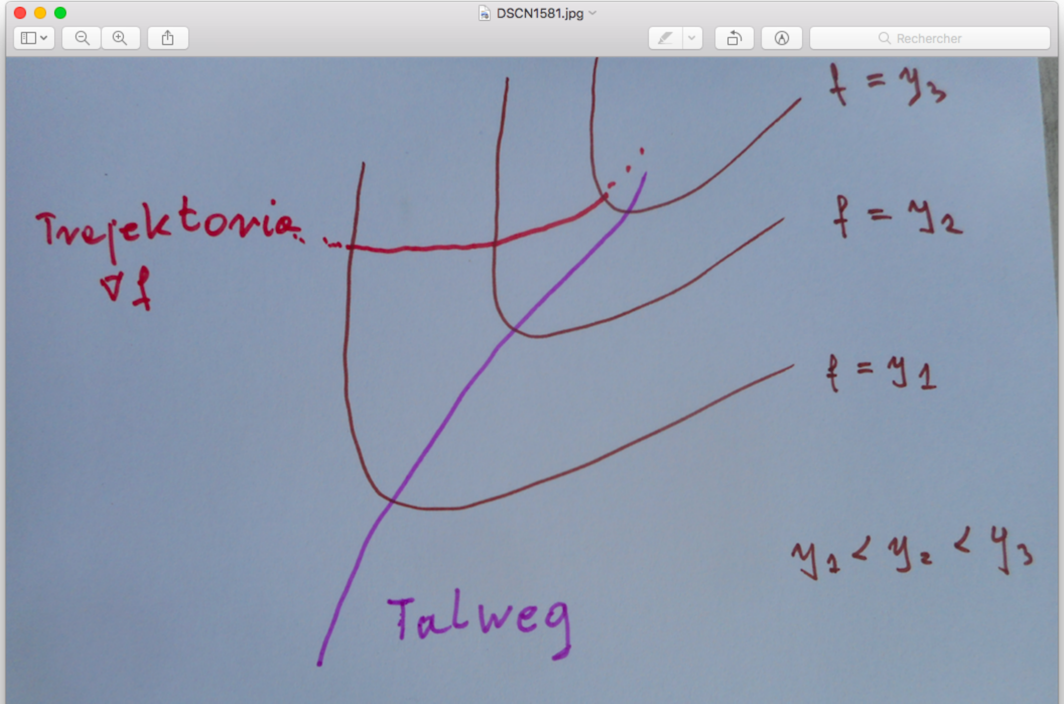
Comparison principle

Theorem (Comparison Principle, D'Acunto , K. 2005)

Let $f : U \to \mathbb{R}$ be a C^2 function defined in an open subset U of $\mathbb{R}^n$ or on a Riemannian manifold. Let $x : I \to U$ be a trajectory of ∇f. Then the length of the trajectory $x(t)$ is bounded by $(1+\varepsilon) Length\, \Gamma^{\varepsilon}(f)$.

For two values $y_1 < y_2$, consider a trajectory $x(t)$ of ∇f, starting at the level $f^{-1}(y_1)$ and ending at the level $f^{-1}(y_2)$. By the Comparison Principle Then the length of the trajectory $x(t)$

$$\text{Length of } x(t) \leq (1+\varepsilon)\text{Length } (f^{-1}(y_1, y_2) \cap \Gamma^{\varepsilon}(f)).$$


DSCN1581.jpg
Rechercher
f = y3
f = y2
f = y1
Trajektoria
∇f
Talweg
y1 < y2 < y3

Given $f : U \to \mathbb{R}$ of class C^2, assume that ε-talweg, denoted by $\Gamma^\varepsilon(f)$, is of finite length. Let $K \subset U$ be compact. Consider a sequence $a_\nu \in K$, $\nu \in \mathbb{N}$ such that the sequence $f(a_\nu) \in \mathbb{R}$, $\nu \in \mathbb{N}$ is decreasing, moreover $a_{\nu+1}$ is the orthogonal projection of a_ν on $f^{-1}(f(a_{\nu+1})) \cap K$.

Theorem (Convergence of proximal algorithm)

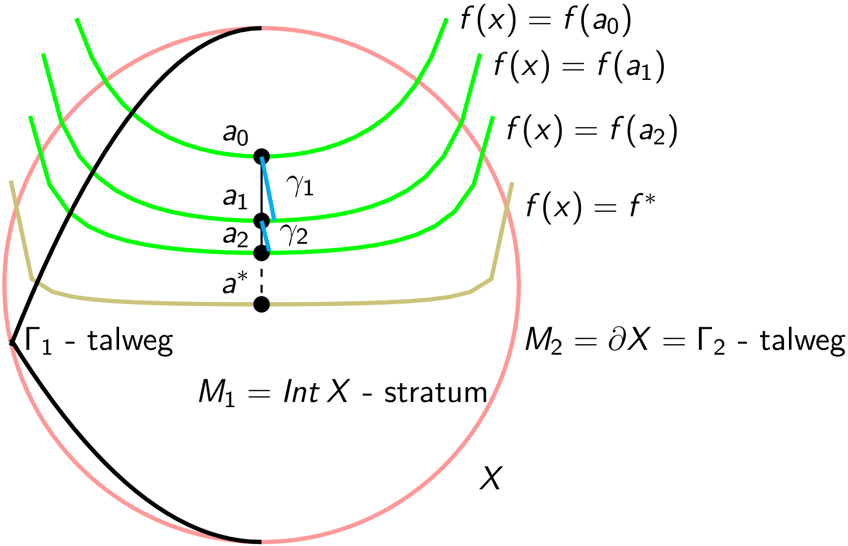
There exists $\lim_{\nu\to\infty} a_\nu = a_*$, *moreover*

$$|a_\nu - a_*| \le Length(\Gamma^\varepsilon(f) \cap f^{-1}([f(a_*), f(a_\nu)]).$$

Indeed, by Comparison Principle, the length of the trajectory γ_ν of ∇f, joining a_ν with the level $f^{-1}(f(a_{\nu+1}))$ is bounded by the length of $(\Gamma^\varepsilon(f) \cap f^{-1}([f(a_{\nu+1}), f(a_\nu)])$. On the other hand $|a_\nu - a_{\nu+1}|$ is bounded by the length of γ_ν. Clearly

$$|a_\nu - a_*| \le \sum_{k=\nu}^{\infty} |a_k - a_{k+1}|.$$

□



$f(x) = f(a_0)$
$f(x) = f(a_1)$
$f(x) = f(a_2)$
$f(x) = f^*$
a_0
a_1
a_2
a^*
γ_1
γ_2
Γ_1 - talweg
$M_2 = \partial X = \Gamma_2$ - talweg
$M_1 = Int\, X$ - stratum
X

Convexification and discrete gradient trajectories

Let $X \subset \mathbb{R}^n$ be a compact convex semialgebraic set and $f \geq m > 0$ a polynomial. Then for some $N \in \mathbb{N}$ (explicite) the function

$$\varphi_{N,\xi}(x) := (1 + |x - \xi|^2)^N f(x)$$

is μ-strongly convex on X for any $\xi \in X$.
Choose an arbitrary point $a_0 \in X$, and by induction set

$$a_\nu := \mathrm{argmin}_X \, \varphi_{N,a_{\nu-1}}.$$

Theorem (K.,Spodzieja 2015)

The limit

$$a^* = \lim_{\nu \to \infty} a_\nu$$

exists, and $f(a^*)$ is a lower critical value of f on X.

Talweg lines and gradient extremal lines

In the definition of ε-ridge/valley we allow $\varepsilon = 0$. In this case $V_0(f)$ is the set of points $x \in U$ at which $|\nabla f|$ has a global minimum on the fiber $f^{-1}(f(x))$. The set $V_0(f)$ depends on U, so it is more natural to consider also local minima of $|\nabla f|$ on fibers of f, that is the set

$$\Gamma_1(f) = \{x \in U : |\nabla f| \text{ has a loc. min. at } x \text{ on the fiber } f^{-1}(f(x))\}.$$

We will call this set the *talweg* (in ancient spelling *thalweg*) of f or the *ridge and valley lines* of f. Clearly $V_0(f) \subset \Gamma_1(f)$.

Actually the notion of ridge or bottom of a valley is not really well determined in the literature. There are several definitions which are used in applied sciences. These “natural” lines appear in classical geomorphology, specially in hydrology, oil recovery meteorology and recently in a very spectacular way in artificial vision.

In the late 19th century there were fascinating discussions on how to define mathematically the lines (“talwegs”) sketching the drainage pattern of a landscape of the graph of a function f in two variables: see, among many others, de Saint-Venant, Cayley, Maxwell, Jordan. Nowadays, the discussion is still alive and has been imported to image analysis. Some authors insisted that “talwegs” should be composed of trajectories of ∇f; they suggested that stable and unstable manifolds should be good candidates. However to define “talweg” in this way, f must have saddle points in U, but on the other hand we “see” ridges or valleys even when we do not “see” saddle points, for instance they are possibly outside of the domain U.

The set $\Gamma_1(f)$ is hence one of the possible candidates for a "talweg". It has several advantages, which we explain below. First let us introduce a larger set called the *gradient extremal set*

$$\Theta_1(f) = \{x \in U : \mathrm{d}(|\nabla f|^2) \wedge \mathrm{d}f = 0\}.$$

This is simply the set of critical points of the map $(f, |\nabla f|^2) : U \to \mathbb{R}^2$ and of course

$$\Gamma_1(f) \subset \Theta_1(f).$$

Clearly $\Theta_1(f)$ is closed in U.
Let $Hf(x)$ be the Hessian matrix of f at the point x. Note that $\nabla(|\nabla f|^2) = 2Hf \cdot \nabla f$. Hence we have the following

Lemma

The set $\Theta_1(f) = \{x \in \mathbb{R}^n : \mathrm{d}(|\nabla f|^2) \wedge \mathrm{d}f = 0\}$ is the set of points $x \in \mathbb{R}^n$ such that $\nabla f(x)$ is an eigenvector of $Hf(x)$.

The canonical stratification of the incidence variety Σ

Let $S_n(\mathbb{R})$ the space of $n \times n$ symmetric matrices with real entries. We define

$$\Sigma = \{(V, H) \in \mathbb{R}^n \times S_n(\mathbb{R}) : \exists \lambda \in \mathbb{R},\ H \cdot V = \lambda V\}.$$

Denote by $S_n^*(\mathbb{R}) \subset S_n(\mathbb{R})$ the space of $n \times n$ symmetric matrices with real entries and simple and nonzero eigenvalues.

Theorem (Stratification of Sigma)

Denote $\Sigma_0 = 0 \times S_n(\mathbb{R})$, $\Sigma_0^* = 0 \times S_n^*(\mathbb{R})$, $\Sigma_1 = \Sigma \setminus \Sigma_0$. *Then*

1. Σ *is algebraic of codimension* $n - 1$.
2. $\Sigma_0 \setminus \Sigma_0^*$ *is algebraic of codimension* $n + 1$,
3. $\Sigma_1 = \Sigma \setminus \Sigma_0$ *is nonsingular,*
4. (Σ_1, Σ_0^*) *satisfies Whitney's condition condition b (and even a stronger condition:* (Σ_1, Σ_0^*) *is locally analytically trivial along* Σ_0^**).*

Given a C^2 function $f : U \to \mathbb{R}$, where U is an open subset of $\mathbb{R}^n$, we shall consider $T_f : U \to \mathbb{R}^n \times S_n(\mathbb{R})$, the *Gauss–Hesse map* of f, given by

$$T_f(x) = (\nabla f(x), Hf(x)).$$

Recall that Σ consists of couples (eigenvector, symmetric matrix). Hence

$$\Theta_1(f) = (T_f)^{-1}(\Sigma). \tag{2}$$

Let U be an open subset of $\mathbb{R}^n$ and let $f : U \to \mathbb{R}$ be a C^∞ function. We denote by $C_f = \{x \in U : \nabla f(x) = 0\}$ the critical set of f. Recall that f is a *Morse function at* $x \in C_f$ if the Hessian matrix $Hf(x)$ has only nonzero eigenvalues. Suppose that the eigenvalues of $Hf(x)$ are simple; then $Hf(x)$ has n distinct eigenspaces $L_1, \ldots, L_n$ each of dimension 1.

Theorem (Local structure of the gradient extremal set of a generic C^∞ function)

Assume that T_f, the Gauss–Hesse map of f, is transverse to the strata of Σ. Then

1. *f is Morse function, hence the critical set C_f has only isolated points,*
2. *$\Theta_1(f) \setminus C_f$ is a C^∞ submanifold of dimension 1,*
3. *for any $x \in C_f$ there exists an open neighborhood U_x such that*

$$\Theta_1(f) \cap U_x = S_1 \cup \cdots \cup S_n, \qquad (3)$$

   - 3.1 *each S_i is a connected C^∞ submanifold of dimension 1, closed in U_x,*
   - 3.2 *$S_i \cap S_j = \{x\}$, for any $i \neq j$,*
   - 3.3 *L_i is tangent to S_i at x.*

Possibly $\Theta_1(f) = \emptyset$.

Bound for the length of gradient trajectories for generic C^∞ functions.

Let V be an open subset of U such that $\overline{V} \subset U$ is compact and the boundary B of V is smooth. Let $\Theta_2(f)$ denote the corresponding gradient extremal set of f on the boundary of V.

Theorem

There is an open and dense set $\mathcal{S} \subset C^\infty(U, \mathbb{R})$ such that, if $f \in \mathcal{S}$, then the length of any trajectory of ∇f in V is bounded by the length of $(\overline{V} \cap \Theta_1(f)) \cup \Theta_2(f)$, in particular it is finite.

Open Problems.

Describe global structure of the gradient extremal set of a generic C^∞ function f on a compact manifold M.

More precisely, assume that M is connected. To the best of my knowledge, the following questions are open:

- Is there a Morse function $f : M \to \mathbb{R}$ such that $\Theta_1(f)$ is connected ?
- If so, are the critical points of f restricted to $\Theta_1(f)$ exactly the critcal points of f on M ?
- If so, it should be possible to interpret Morse (Floer) Homology on $\Theta_1(f)$.

Lecture 2: Effective estimates for the length of gradient trajectories in the polynomial case.

o-minimal structures

A collection $\mathcal{M} = \bigcup_{n\in\mathbb{N}} \mathcal{M}_n$ is an o-minimal structure on $(\mathbb{R}, +, \cdot)$, where each $\mathcal{M}_n$ is a family of subsets of $\mathbb{R}^n$, if

1. each $\mathcal{M}_n$ is closed under finite set-theoretical operations;
2. if $A \in \mathcal{M}_n$ and $B \in \mathcal{M}_m$, than $A \times B \in \mathcal{M}_{n+m}$;
3. let $A \in \mathcal{M}_{n+m}$ and $\pi : \mathbb{R}^{n+m} \to \mathbb{R}^n$ be the projection on the first n coordinates, then $\pi(A) \in \mathcal{M}_n$;
4. let $f, g_1, ..., g_k \in \mathbb{Q}[X_1, ..., X_n]$, then $\{x \in \mathbb{R}^n : f(x) = 0, g_1(x) > 0, ..., g_k(x) > 0\} \in \mathcal{M}_n$;
5. $\mathcal{M}_1$ consists of all finite unions of open intervals and points.

For a fixed o-minimal structure $\mathcal{M}$, $A \in \mathcal{M}_n$ is called definable in $\mathcal{M}$. A map $f : A \to \mathbb{R}^m$, is $\mathcal{M}$ definable if its graph is $\mathcal{M}$- definable. Semialgebraic sets, globally subanalytic sets form o-minimal structures. An important new o-minimal structure is $\mathbb{R}_{exp}$ the expansion of semialgebraic sets by the global exponential function (suggested by Tarski).

Lemma (Infimum on fibers)

Let $f : A \to \mathbb{R}$ be a definable function such that $f(x) \geqslant 0$ for all $x \in A$. Let $G : A \to \mathbb{R}^m$ be a definable mapping and define a function $\varphi : G(A) \to \mathbb{R}$ by

$$\varphi(y) = \inf_{x \in G^{-1}(y)} f(x).$$

Then φ is definable.

Lemma (Definable choice)

Let $S \subseteq \mathbb{R}^{m+n}$ be a definable set and $\pi_m : \mathbb{R}^{m+n} \to \mathbb{R}^m$ be the projection on the first m coordinates. Then there exists a definable map $f : \pi_m(S) \to \mathbb{R}^n$ such that the graph of f is contained in S.

Lemma (Uniform finiteness)

Let $A \subset \mathbb{R}^{n+m}$ be a definable set and assume that for all $y \in \mathbb{R}^m$ the set $A_y = \{x \in \mathbb{R}^n : (x, y) \in A\}$ is finite. Then there exists an integer N such that $\mathrm{Card} A_y \leq N$ for all $y \in \mathbb{R}^m$.

Let Γ be a compact definable curve and let $\mathcal{H}$ denote the set of affine hyperplanes in $\mathbb{R}^n$. Then, for almost every $H \in \mathcal{H}$ (that is, except maybe for a definable subset $\mathcal{H}_1 \subset \mathcal{H}$ of codimension greater than or equal to 1), the set $\Gamma \cap H$ is finite. Let $i(\Gamma, H)$ denote the cardinality of $\Gamma \cap H$.

Theorem (Cauchy - Crofton formula)

There exists a normalization of the canonical measure μ on $\mathcal{H}$ such that the length of Γ can be expressed by the formula

$$Length(\Gamma) = \int_{\mathcal{H}} i(\Gamma, H) d\mu. \tag{4}$$

Corollary (Uniform bound for the length of definable curves)

Let $K \subset \mathbb{R}^n$ be a compact set and let $\mathcal{G} = \{\Gamma_p\}_{p\in\mathcal{P}}$ be a definable family of definable curves contained in K. Then there exists a constant $m_{\mathcal{G}} > 0$, depending only on the family $\mathcal{G}$, such that for any $p \in \mathcal{P}$,

$$Length(\Gamma_p) \leq m_{\mathcal{G}}.$$

The constant $m_{\mathcal{G}}$ in the Corollary is the product of some integer i_* by the normalized volume of the hyperplanes that intersect the compact set K. Let us denote by $\nu(n)$ this volume when $K = \overline{\mathbb{B}^n}$. Then we have the following

$$\nu(n) = \frac{n Vol_n(\mathbb{B}^n)}{Vol_{n-1}(\mathbb{B}^{n-1})}.$$

Using the Euler Γ function we obtain the following alternative formula:

$$\nu(n) = 2\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n+1}{2}\right)\Gamma\left(\frac{n}{2}\right)^{-1}$$

Note that in dimension 2, we have $\nu(2) = \pi$ and, for any $n \in \mathbb{N}$, $\nu(n) \leqslant 2n$. A simple computation shows that $\nu(n) \sim \sqrt{2\pi n}$ for sufficiently large n.

Lemma (Generalized critical values, definable case)

Let $f : U \to \mathbb{R}$ be a C^1 function defined on an open bounded subset U of $\mathbb{R}^n$. Assume that f is definable in an o-minimal structure. Then the set $K(f)$ is finite.

Proof Recall that

$$\varphi(t) = \inf\{|\nabla f(x)| : x \in f^{-1}(t)\}$$

is a definable function on $f(U) \subset \mathbb{R}$. Hence there are $t_0 < t_1 < \cdots < t_N$ such that f is monotone (and nonnegative) on each (t_i, t_{i+1}).
Recall that $t \in \mathbb{R}$ is a typical value of f if there exists $c > 0$ such that $\varphi(t) \geqslant c$ in some neighborhood of t. Otherwise $t \in K(f)$. For any $\varepsilon > 0$ set

$$\Sigma_\varepsilon := \{x \in U : \ |\nabla f(x)| \leq \varepsilon\}.$$

The collection Σ_ε, $\varepsilon > 0$ is a definable family of definable and bounded sets. For $\varepsilon > 0$ small enough $\Sigma_\varepsilon = \Sigma_\varepsilon^1 \cup \cdots \cup \Sigma_\varepsilon^k$, have the same number of connected components. Moreover, the geometric diameter of any Σ_ε^i is uniformly bounded by some constant $D > 0$. Thus, $f(\Sigma_\varepsilon^i)$ is a segment of the length at most εD.

Note that

$$K(f) = \bigcap_{\varepsilon>0} f(\Sigma_\varepsilon).$$

Therefore $K(f)$ consists of at most k points. $\square$

Let U be an open subset of $\mathbb{R}^n \times \mathbb{R}^k$ and $f : U \to \mathbb{R}$ a definable function. Then the set U is definable by definition. Let us denote by $\mathcal{P} \subset \mathbb{R}^k$ the projection of U on the last coordinates. For $p \in \mathcal{P}$ set

$$U_p = \{x \in \mathbb{R}^n : (x, p) \in U\}, \qquad f_p : U_p \to \mathbb{R},$$

where $f_p(x) = f(x, p)$, $x \in U_p$ is a definable function. Throughout this section, $\mathcal{F} = \{f_p\}_{p\in\mathcal{P}}$ denotes the definable family of such functions f_p.

Gradient trajectories of definable functions

Under these hypotheses, we state the main result of this section:

Theorem

Let $\mathcal{F}$ be a definable family of functions as above. Assume that for each $p \in \mathcal{P}$, the function f_p is of class C^2 on U_p and there exists a compact set $K \subset \mathbb{R}^n$ such that $U_p \subset K$. Then there exists a constant $M_{\mathcal{F}} > 0$ such that, for all $p \in \mathcal{P}$, the length of any trajectory of ∇f_p is bounded by $M_{\mathcal{F}}$.

Let $\mathcal{F}$ denote a definable family of functions satisfying the same assumptions as in Theorem except that the definable sets U_p are bounded but not necessarily contained in a fixed compact set K. Let d_p denote the diameter of U_p. Then we have the following corollary:

Corollary

There exists a constant $M_{\mathcal{F}} > 0$ such that for every $p \in \mathcal{P}$ the length of any trajectory of ∇f_p is bounded by $M_{\mathcal{F}} \cdot d_p$.

Bounds for gradient trajectories of polynomials

Throughout this section $f : \mathbb{R}^n \to \mathbb{R}$ denotes a polynomial function of degree d. We will give an explicit upper bound for the length of a trajectory of ∇f restricted to the unit ball $\mathbb{B}^n$. First observe that the following corollary follows easily from the previous Theorem.

Corollary

Let f be a polynomial in n variables of degree d. Then the length of any trajectory of ∇f in a ball of radius r is bounded by $rA(n,d)$, where $A(n,d)$ is a constant depending only on d and n.

In order to estimate $A(n,d)$ explicitly we will use the method described in Lecture 1. We shall construct explicitly a talweg of f, i.e. a semialgebraic curve Γ with the following property: if $y \in \overline{\mathbb{B}^n}$, then

$$|\nabla f(y)| \geq |\nabla f(x)| \quad \text{for some } x \in \Gamma \cap f^{-1}(f(y)). \tag{5}$$

In other words, we have to minimize $|\nabla f|^2$ on the fibers of f restricted to $\overline{\mathbb{B}^n}$. More precisely, we shall prove that, for a generic polynomial f of degree d, the set

$$\Gamma_1 = \{x \in \mathbb{B}^n : |\nabla f|^2 \text{ has a loc. min. at } x \text{ on } f^{-1}(f(x)) \cap \mathbb{B}^n\}$$

is of dimension at most 1. We shall prove that the set $\Gamma_2 \subset \mathbb{S}^{n-1}$ defined by

$$\Gamma_2 = \{x \in \mathbb{S}^{n-1} : |\nabla f|^2 \text{ has a loc. min. at } x \text{ on } f^{-1}(f(x)) \cap \mathbb{S}^{n-1}\}$$

is of dimension 1. Then we take $\Gamma = \Gamma_1 \cup \Gamma_2$ and for a generic polynomial we shall give explicit formulae for polynomials describing Γ_1 and Γ_2.
The following proposition shows that the locus of points at which the level sets of the polynomials f and $|\nabla f|^2$ are not transverse, generically defines an algebraic curve. Let $\mathbf{X} = (X_1, \ldots, X_n)$ and denote by $\mathbb{R}_d[\mathbf{X}]$ the space of polynomials in n variables of degree less than or equal to d.

Let $\mathbf{X} = (X_1, \ldots, X_n)$ and denote by $\mathbb{R}_d[\mathbf{X}]$ the space of polynomials in n variables of degree less than or equal to d.

Lemma (Generic transversality)

Let $n, d \geq 2$. Then there exists a semialgebraic set $E_d \subset \mathbb{R}_d[\mathbf{X}]$ of codimension greater than or equal to 1 such that, for any polynomial $f \in \mathbb{R}_d[\mathbf{X}] \setminus E_d$, the set

$$\Theta_1(f) = \{x \in \mathbb{R}^n : \mathrm{d}(|\nabla f|^2) \wedge \mathrm{d}f = 0\}$$

is either empty or a finite union of real algebraic curves.

Proof. Recall that by (2) we have $\Theta_1(f) = (T_f)^{-1}(\Sigma)$, where

$$\Sigma = \{(V, H) : \exists \lambda \in \mathbb{R} : H \cdot V = \lambda V\}$$

is an algebraic subset of $\mathbb{R}^n \times S_n(\mathbb{R})$ and $T_f = (\nabla f, Hf)$ is the Gauss–Hess map of f. Recall that codim $\Sigma = n - 1$. Let

$$E_d = \{f \in \mathbb{R}_d[\mathbf{X}] : T_f \text{ is not transverse to } \Sigma\}.$$

By routine arguments, E_d is semialgebraic. To show that the set $\mathbb{R}_d[\mathbf{X}] \setminus E_d$ is dense in $\mathbb{R}_d[\mathbf{X}]$ we make quadratic perturbation. $\square$

We now study Γ_2, the set of points of $\mathbb{S}^{n-1}$ where $|\nabla f|^2$ has a local minimum on the fibers of f restricted to the sphere. Let $r(x) = |x|^2$ and define

$$\Theta_2(f) = \{x \in \mathbb{S}^{n-1} : \mathrm{d}(|\nabla f|^2) \wedge \mathrm{d}f \wedge \mathrm{d}r = 0\}.$$

Note that $\Gamma_2 \subset \Theta_2(f)$ and $\Theta_2(f)$ is the set of critical points of $|\nabla f|^2$ on the fibers of f restricted to the sphere. For a generic polynomial f the set $\Theta_2(f)$ is a curve, namely we have:

Lemma (Generic transversality on the sphere)

There exists a semialgebraic set $F_d \subset \mathbb{R}_d[\mathbf{X}]$ of codimension greater than or equal to 1 such that, for any $f \in \mathbb{R}_d[\mathbf{X}] \setminus F_d$, the set $\Theta_2(f)$ is nonempty and is a finite union of real algebraic curves.

Proof. We proceed as before and define

$$\tilde{\Sigma} = \left\{(V, H, p) \in \mathbb{R}^n \times S_n(\mathbb{R}) \times \mathbb{S}^{n-1} : (H \cdot V) \wedge V \wedge p = 0\right\}.$$

and argue on transversality to $\tilde{\Sigma}$. □

Take $G_d = E_d \cup F_d$.

Theorem (Estimate in a generic case)

Let $n, d \geqslant 2$ be integers. Then for any $f \in \mathbb{R}_d[\mathbf{X}] \setminus G_d$, the length of any trajectory of ∇f in $\mathbb{B}^n$ is bounded by

$$A(n,d) = \nu(n)((3d-4)^{n-1} + 2(3d-3)^{n-2})$$

where $\nu(n)$ is a constant depending only on the dimension.

Proof. To apply Cauchy-Crofton formula we need

Lemma

For a generic affine hyperplane H the set $H \cap \Theta_1(f)$ has at most $(3d-4)^{n-1}$ points.

Recall that $\Theta_1(f) = \{x \in \mathbb{R}^n : \mathrm{d}(|\nabla f|^2) \wedge \mathrm{d}f = (2Hf \cdot \nabla f) \wedge \nabla f = 0\}$. So, it is contained in the zero set of $(n-1)$ polynomials of the degree at most $(d-2)+(d-1)-(d-1) = 3d-4$. We conclude with general version of Bézout's Theorem.

Similarly

Lemma

For a generic affine hyperplane H the set $H \cap \Theta_2(f)$ has at most $2(3d-3)^{n-2}$ points.

According to Comparison Principle the length of any ∇f in $\mathbb{B}^n$ is bounded by $A(n,d)$, for a generic polynomial $f \in \mathbb{R}_d[X_1,\ldots,X_n]$. □

Finally, by a perturbation argument (a bit involved) and Generalized Comparison Principle we obtain

Theorem

If $f : \mathbb{R}^n \to \mathbb{R}$, with $n \geqslant 2$, is a polynomial of degree $d \geq 2$, then the length of any trajectory of ∇f in a ball of radius r is bounded by

$$\nu(n) r A(n,d),$$

where $A(n,d) = (3d-4)^{n-1} + 2(3d-3)^{n-2}$.

Example for a lower bound

For integers d and n, we will denote by $D(n, d)$ the supremum of the lengths of gradient trajectories of polynomials of degree d in the unit ball in $\mathbb{R}^n$. Then

Theorem

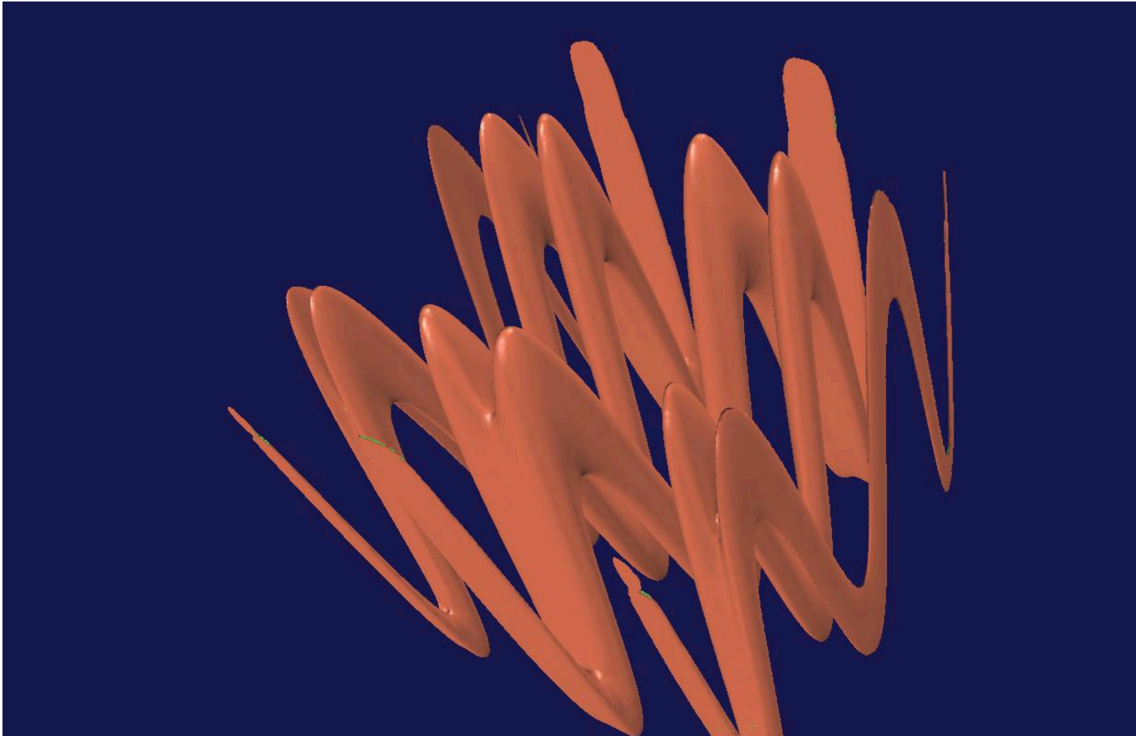
For any integers $n, d \geq 2$

$$D(n, 2d) \geqslant 2d^{n-1}, \qquad d \in \mathbb{N}.$$

Idea of example; the dth Chebyshev polynomial (of the first kind) $T_d(x)$ is determined by $T_d(\cos\theta) = \cos(d\theta)$. For $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, put $f_i(x) = x_{i+1} - T_d(x_i)$ and $p = \sum_{i=1}^{n-1} f_i^2$. Now we define our polynomial by

$$f(x) = \alpha(1 - x_1^2) - p(x),$$

where $\alpha > 0$ is small enough. For $n = 3$ and $d = 6$ the levels of f looks like



We state an estimate on the length of gradient trajectories of fewnomials. Let f_i, $i = 1, \ldots, n$, be polynomials in n variables with only K monomials. Then, Khovanskiĭ's classical result states that the number of non degenerate solutions of the system $f_1 = \cdots = f_n = 0$ is bounded by

$$2^n(1+n)^K 2^{\frac{K(K-1)}{2}}.$$

Theorem

Let f be a polynomial in n variables containing only k monomials. Then the length of the trajectories of ∇f inside a ball of radius r is bounded by

$$r\nu(n)N(n,k),$$

where $N(n,k) = 2^n[(1+n)^{K_1}2^{\frac{K_1(K_1-1)}{2}} + (1+n)^{K_2}2^{\frac{K_2(K_2-1)}{2}}]$,
$K_1 = 2n(k + \frac{n(n+3)}{2})^3$ *and* $K_2 = 6n(k + \frac{n(n+3)}{2})^3$.

Let $D \subset \mathbb{R}^n$ be path-connected, by $diam_g(D)$ we mean the infimum of $c >$ such that any pairs of points can be joined in D by path of length $\leq c$.

Theorem

Let $f : \mathbb{R}^n \to \mathbb{R}$, with $n \geqslant 2$, be a polynomial of degree $d \geq 2$ and let $B(r)$ be a ball of radius r in $\mathbb{R}^n$. Let D_i, $i \in I$, be all connected components of $\{f > 0\} \cap B(r)$. Then

$$\sum_{i \in I} diam_g(D_i) \ \leq \ Od^{n-1} = 2r\nu(n)(3d+2)^{n-1}. \tag{6}$$

In particular $diam_g(D_i) \leqslant 2r\nu(n)(3d+2)^{n-1}$ for any connected component D_i. More precisely, any two points in D_i can be joined in D_i by an arc which is a piecewise trajectory of ∇g or $-\nabla g$ of length not greater than $2r\nu(n)(3d+2)^{n-1}$, where g is a polynomial of degree at most $d+2$.

Lecture 3: From Talweg to KL-inequality and the Łojasiewicz gradient inequality

Let $f : U \to \mathbb{R}$ be a $C^2 1$ function defined on an open bounded subset U of $\mathbb{R}^n$. Assume that f is definable in an o-minimal structure. Then $\Gamma^1(f)$

$$\varphi(t) = \inf\{|\nabla f(x)| : x \in f^{-1}(t)\}$$

is a definable function on $f(U) \subset \mathbb{R}$. Let us take definable 1-talweg, denoted $\Gamma^1(f)$. That is, if $y \in \Gamma^1(f)$ then $|\nabla f(y)| \leq 2\varphi(f(y))$. Thus for any $x \in U$ such that $f(x) = f(y)$ we have

$$2|\nabla f(x)| \geq |\nabla f(y)|.$$

Let $\eta : (a, b) \to \Gamma^1(f)$ be an (almost) arc-length parametrization,$|\eta'| \leq 2$. Set $h = f \circ \eta : (a, b) \to \mathbb{R}$, recall that h is strictly increasing definable C^1. Put $c, d) = h((a, b))$. So $h : (a, b) \to (c, d)$ is a diffeomorphism. Set

$$\Psi = 4h^{-1} : (c, d) \to (a, b).$$

Let us take $x \in U$, there exists $y \in \Gamma^1(f)$ such that $f(x) = f(y) = t$. Let $s \in (a, b)$ such that $y = \eta(s)$. Then

$$|h'(s)| = |\langle \nabla(\eta(s)), \eta'(s) \rangle| \leq |\nabla f(\eta(s))||\eta'(s)\rangle| \leq 2|\nabla f(y)|.$$

So, $|h'(s)|^{-1}| \geq (2|\nabla f(y)|)^{-1}$. Now we compute

$$|\nabla(\Psi \circ f)(x)| = |\Psi'(f(x))||\nabla f(x)| \geq 4|h'(s)|^{-1}|(1/2)|\nabla f(y)|.$$

Therefore $|\nabla(\Psi \circ f)(x)| \geq 1$.

K-L inequality

Theorem (K 1998)

Let $f : U \to \mathbb{R}$ be a C^2 definable in an o-minimal structure $\mathcal{M}$, where $U \subset \mathbb{R}^n$ is open bounded set. Then there exists continuous strictly increasing function $\Psi : (\alpha, \beta) \to \mathbb{R}$ definable in the structure $\mathcal{M}$ such that

$$\|\nabla(\Psi \circ f)\| \geq 1,$$

holds in U.

The function Ψ is called a desingularizing function for f. Since Ψ is definable it is differentiable except finitely many points.
Frequently, a function $f : U \to \mathbb{R}$ (not necessarily definable, possibly even open in some infinite dimensional Hilbert space) satisfying the conclusion of this theorem, is said to has *KL* -property.

Applying K-L inequality

For two values $y_1 < y_2$, consider a trajectory $x(t)$ of ∇f, starting at the level $f^{-1}(y_1)$ and ending at the level $f^{-1}(y_2)$. Up to a reparametrization $x(t)$ is a trajectory of $\nabla(\Psi \circ f)$ starting at $(\Psi \circ f)^{-1}(\Psi(y_1))$ and ending at $(\Psi \circ f)^{-1}(\Psi(y_2))$. Take the arc length parametrization s of the trajectory $x(t)$, and set $g(s) = \Psi \circ f(x(s))$, then

$$g'(s) = \langle \nabla(\Psi \circ f)(x(s), x'(s) \rangle = \|\nabla(\Psi \circ f)\| \geq 1.$$

Hence the length of $x(t)$ is not greater then $\Psi(y_2) - \Psi(y_1)$.

Łojasiewicz's gradient inequality.

Theorem (Łojasiewicz 1962)

Let $f : U \to \mathbb{R}$ be a real analytic function, where $U \subset R^n$ is open and $K \subset U$ a compact set. Then for any $x^* \in K$ there exists constants $C > 0$, $\varepsilon > 0$ and $\rho < 1$ such that

$$\|\nabla f(x)\| \geq C|f(x) - f(x^*)|^\rho, \tag{7}$$

for $x \in K$ such that $|f(x) - f(x^*)| \leq \varepsilon$.

We deduce it from the KL-inequality. For simplicity, assume $f(x^*) = 0$. In the subanalytic case we can take $\Psi(y) = bsgn(y)|y|^{1-\rho}$, where $1 - \rho > 0$ and $b > 0$ is a constant. Since $\|\nabla(\Psi \circ f)\| \geq 1$, we get

$$|\Psi'(f(x)||\nabla f(x)| = b(1-\rho)|f(x)|^{-\rho}||\nabla f(x)| \geq 1.$$

So, $\|\nabla f(x)\| \geq C|f(x)|^\rho$ holds with $C = b(1-\rho)$.

By $F : (\mathbb{R}^n, a) \to (\mathbb{R}^m, 0)$, where $a \in \mathbb{R}^n$, we denote a mapping from a neighborhood $U \subset \mathbb{R}^n$ of the point a to $\mathbb{R}^m$ such that $F(a) = 0$. We put $V(F) = \{x \in U : F(x) = 0\}$.
If $F : (\mathbb{R}^n, a) \to (\mathbb{R}^m, 0)$ is a real analytic mapping, then there are positive constants C, η, ε such that the following *Łojasiewicz inequality* holds:

$$|F(x)| \geq C \operatorname{dist}(x, V(F))^{\eta} \quad \text{if} \quad |x - a| < \varepsilon, \tag{8}$$

where $|\cdot|$ is the Euclidean norm in $\mathbb{R}^n$ and $\operatorname{dist}(x, V)$ is the distance of $x \in \mathbb{R}^n$ to the set V ($\operatorname{dist}(x, V) = 1$ if $V = \emptyset$). The smallest exponent η in (8) is called the *Łojasiewicz exponent* of F at a and is denoted by $\mathcal{L}_a(F)$. It is known that $\mathcal{L}_a(F)$ is a rational number and (8) holds with any $\eta \geq \mathcal{L}_a(F)$ and some $C, \varepsilon > 0$.

Let $f : (\mathbb{R}^n, a) \to (\mathbb{R}, 0)$ be a real analytic function.
Then there are positive constants C, ε and a constant $\varrho \in [0, 1)$ such that the following *Łojasiewicz gradient inequality* holds

$$|\nabla f(x)| \geq C|f(x)|^{\varrho} \quad \text{if} \quad |x - a| < \varepsilon. \tag{Ł}$$

The smallest exponent ϱ in (Ł), denoted by $\varrho_a(f)$, is called the *Łojasiewicz exponent in the gradient inequality*. The number $\varrho_a(f)$ is rational and (Ł) holds with any exponent $\varrho \geq \varrho_a(f)$ and some positive constants C, ε.

In the case of an analytic function $f : (\mathbb{R}^n, a) \to (\mathbb{R}, 0)$ such that $V(f) = \{a\}$ (i.e. has a strict extremum at a), Gwoździewicz proved that

$$\mathcal{L}_a(f) = \frac{1}{1 - \varrho_a(f)} = \mathcal{L}_a(\nabla f) + 1. \tag{G1}$$

The above result is not true in the general case, even if we assume that f has an isolated singularity. For an arbitrary analytic function $f : (\mathbb{R}^n, a) \to (\mathbb{R}, 0)$ we have proved (with Spodzieja) that

$$\mathcal{L}_a(f) \leq \frac{1}{1 - \varrho_a(f)} \tag{9}$$

and there are positive constants C, ε such that

$$|\nabla f(x)| \geq C \operatorname{dist}(x, V(f))^{\varrho_a(f)/(1-\varrho_a(f))} \quad \text{if} \quad |x - a| < \varepsilon \tag{10}$$

If f has an isolated singularity at a, we proved that

$$\frac{1}{1 - \varrho_a(f)} \leq \mathcal{L}_a(\nabla f) + 1. \tag{11}$$

By Gwoździewicz's result (G1), the estimates (9) and (11) are exact in terms of $\varrho_a(f)$.

Bounds on $\varrho_a(f)$ for polynomials

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a polynomial of degree d. If f has a local strict extremum at a, Gwoździewicz proved that

$$\varrho_a(f) \leq 1 - \frac{1}{(d-1)^n + 1}. \tag{G2}$$

He used polar curves. In the general case, i.e. without the assumption that a is an isolated point of $V(f)$,

Theorem (D. D'Acunto and K. Kurdyka, A. Gabrielov)

$$\varrho_a(f) \leq 1 - \frac{1}{d(3d-4)^{n-1}} \quad \textit{for} \quad d \geq 2, \tag{\textit{D-K}}$$

Proof.

Let $a \in \Gamma^1(f) \subset \mathbb{R}^n$ and assume $\nabla f(a) = 0$.
Consider $\eta(s)$ arc-length parametrization of $\Gamma^1(f)$ and
$h(s) = f \circ \eta(s) = s^k a(s)$, $a(0) \neq 0$. How to control k ?
Recall that the talweg $\Gamma^1(f)$ is contained in

$$\Theta_1(f) = \{x \in \mathbb{R}^n : \mathrm{d}(|\nabla f|^2) \wedge \mathrm{d}f = (2Hf \cdot \nabla f) \wedge \nabla f = 0\}.$$

So, $\Theta_1(f)$ is contained in the zero set of $(n-1)$ polynomials of the degree at most $(d-2)+(d-1)-(d-1) = 3d-4$. The order k, of f on $\Theta_1(f)$ is bounded by

$$k \leq d(3d-4)^{n-1}$$

and we conclude as in the implication KL inequality to Łojasiewicz inequality.

□

S. Basu, A-M. Nezhad, *Improved effective Łojasiewicz inequality and applications*, preprint Purdue University, November 2022.

J. Bolte, A. Daniilidis, O. Ley, L. Mazet, *Characterizations of Łojasiewicz inequalities: subgradient flows, talweg, convexity*, Trans. Amer. Math. Soc. 362 (2010), no. 6, 3319–3363.

D. D'Acunto, *Sur les courbes intégrales du champs de gradients,* Thèse Université de Savoie (2001).

D. D'Acunto, K. Kurdyka *Bounds for gradient trajectories and geodesic diameter of real algebraic sets*, Bulletin of LMS 38 (6), 951-965, (2006).

D. D'Acunto, K. Kurdyka *Explicit bounds for the Lojasiewicz exponent in the gradient in- equality for polynomials*, Annales Polonici Mathematici, 87 (2005), 51–61.

D. D'Acunto, K. Kurdyka *Bounding the length of gradient trajectories* D. D'Acunto, K. Kurdyka Annales Polonici Mathematici 127(2022), 13–50.

A. Gabrielov, *Multiplicities of Pfaffian intersections, and the Łojasiewicz inequality.* Selecta Math. (N.S.) 1 (1995), no. 1, 113–127.

J. Gwoździewicz, *The Łojasiewicz exponent of an analytic function at an isolated zero*, Comment. Math. Helv. 74 (1999), no. 3, 364–375.

K. Kurdyka, *On gradients of functions definables in o-minimal structures,* Ann. Inst. Fourier (Grenoble) 48 (3), 769-783 (1998).

K. Kurdyka, *KL inequality for definable maps,* in preparation.

K. Kurdyka, S. Spodzieja, *Convexifying positive polynomials and sums of squares approximation* SIAM Journal on Optimization 25 (4), 2512–2536.

K. Kurdyka, S. Spodzieja, *Exponential convexifying of polynomials* Bulletin des Sciences Mathématiques 180, 103–197.

M. Lejeune-Jalabert, B. Teissier, *Clôture intégrale des idéaux et équisingularité. With an appendix by J-J. Risler.* , Ann. Fac. Sci. Toulouse Math. (6) 17 (2008), 781- 859.

- S. Łojasiewicz, *Une propriété topologique des sous-ensembles analytiques réels* Les Equations aux Dérivées Partielles (Paris, 1962) pp. 87–89 Editions du Centre National de la Recherche Scientifique, Paris.
- S. Łojasiewicz, *Ensembles semi-analytiques*, preprint IHES, (1965).