Uniwersytet Łódzki
Wydział Matematyki i Informatyki

Katedra Funkcji Analitycznych
i Równań Różniczkowych

Abdulljabar Naji Ahmed Abdullah

Logarytmiczne uwypuklenie wielomianów

Rozprawa Doktorska

Promotor

Prof. dr hab. Stanisław Spodzieja

Łódź 2022

"Mądra krytyka oświeca, głupia gasi"
A. Fredro
Panu prof. dr. hab. Stanisławowi Spodziei
za poświęcony czas, cierpliwość i wszechstronne wsparcie.

Spis treści

Summary of the doctoral dissertation

One of the fundamental problems of analysis, technology, economics and other branches of science is the search for minima and critical points of functions. One of the methods leading to this goal is the deformation of a given function to a convex function, searching for critical points of this deformation and iterating this process. Reducing a function to a convex or strongly convex function leads to easy determination of critical points and minima of this deformation. These are the exact points where the gradient is zero.The classic approach to convexifying of a function $f : \mathbb{R}^n \to \mathbb{R}$ on bounded and convex sets is to add a strongly convex function $b : \mathbb{R}^n \to \mathbb{R}$ such that $f + b$ is a strongly convex function on this set (see for instance [19], [11] and [20] for quadratic function $b(x) = \gamma|x|^2$, $\gamma > 0$). We describe this more precisely.

Recall that a function $g : X \to \mathbb{R}$ is *strongly convex* or *μ-strongly convex* if $X \subset \mathbb{R}^n$ is a convex set, g is of class $\mathscr{C}^1$ and

$$g(y) \geq g(x) + \langle y - x, \nabla g(x)\rangle + \frac{\mu}{2}|y - x|^2 \quad \text{for } x, y \in X,$$

where $\mu > 0$, $\langle \cdot, \cdot \rangle$ is the standard scalar product in $\mathbb{R}^n$ and ∇g is the gradient of g.

Let $b : \mathbb{R}^n \to \mathbb{R}$ be a $\mathscr{C}^k$ class μ-strongly convex function, $k \geq 2$, $\mu > 0$.

Let $X \subset \mathbb{R}^n$ be a compact and convex set, let $f : \mathbb{R}^n \to \mathbb{R}$ be a function of class $\mathscr{C}^k$ and let $D \in \mathbb{R}$ be a positive number such that

$$|\partial_\beta^2 f(x)| \leq D \quad \text{for } \ x \in X \text{ and } \beta \in S^{n-1},$$

where S^{n-1} the unit sphere in $\mathbb{R}^n$, i.e., $S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$, and $\partial^2_\beta f(x)$ is the second order derivative of f in the direction β at x. One can directly check that

Lemma 1 (see. uwaga 2.2.3 i fakt 3.1.1)**.** *For any $\xi \in \mathbb{R}^n$ and*

$$N > D/\mu,$$

the function $\phi_{N,\xi} : \mathbb{R}^n \to \mathbb{R}$ defined by

$$\phi_{N,\xi}(x) = Nb(x - \xi) + f(x), \quad x \in \mathbb{R}^n,$$

is strongly convex on X (more precisely $N\mu - D$-strongly convex).

In this paper, we will compare the above approach to convexifying of a function that takes only positive values with another approach of multiplying it by a power of a strongly convex function (see chapter 2 and chapter 3). The latter approach was proposed in [9] and continued in [8]. More precisely, in [9] a positive function f of class $\mathscr{C}^2$ is convex on a compact and convex set $X \subset \mathbb{R}^n$ by multiplying the function f by $(1 + |x|^2)^N$ for some N, and in [8] – by multiplying the function f by $\exp(N|x|^2)$. In the second chapter, we generalize these results and show that

Corollary 2 (see wniosek 2.2.2)**.** *If $X \subset \mathbb{R}^n$ is a compact and convex set and $f : X \to \mathbb{R}$ is a function of the class $\mathscr{C}^2$ that takes only positive values, then for any strongly convex function $b : \mathbb{R}^n \to \mathbb{R}$ there is $N_0 > 0$ such that for each $N \geq N_0$ and $\xi \in X$, the function*

$$\varphi_{N,\xi}(x) = b^N(x - \xi) f(x), \quad x \in \mathbb{R}^n, \tag{1}$$

is strongly convex on the set X.

In the case where the function f is a polynomial, the exponent N can be estimated efficiently in terms of the radius of the set X (i.e., $\sup\{|x| : x \in X\}$) of the modules of the polynomial coefficients and $m = \inf\{f(x) : x \in X\}$ (see wniosek 2.2.7). Therefore, in the case of positive functions on compact and convex sets, both adding a multiple of a strongly convex function to the function and multiplying it by the power of such a function have a similar effect, but the first method uses a smaller coefficient N. If we additionally assume that b is a logarithmically strongly convex function (i,e., $\ln b$ is a strongly convex function), then $\varphi_{N,\xi}$ is also a logarithmically strongly convex function (see wniosek 2.2.9). In the case when X is a semialgebraic, compact and convex set, the coefficients of the polynomials describing X and the coefficients of the polynomial f are integers (or rational numbers), the exponent N can be determined fully efficiently (see Twierdzenia 2.2.10 and 2.2.12). These theorems are obtained using the result of G. Jeronimo, D. Perrucci, E. Tsigaridas from [6].

In the third chapter, we will compare the classical approach to problem convexifying of a function with the above for any strongly convex function $b : \mathbb{R}^n \to \mathbb{R}$ and a positive function on a closed and convex (not necessarily bounded) set. In this chapter, we convexiying of a function f by multiplying it by $b(N(x-\xi))$ instead of $b^N(x-\xi)$. This approach simplifies some calculations.

Lemma 1 is difficult to apply to unbounded sets. Namely, we have

Fact 3 (see fakt 3.1.2)**.** *Let $b : \mathbb{R}^n \to (0, +\infty)$ be a convex function of class $\mathscr{C}^2$, let $X \subset \mathbb{R}^n$ be a convex and closed set, let $f : \mathbb{R}^n \to \mathbb{R}$ be a function of class $\mathscr{C}^2$ and let $N > 0$. If for any $\xi \in X$, the function $\phi_{N,\xi}$ defined by* (1) *is convex on X, then*

$$\partial_\beta^2 \phi_{N,\xi}(\xi) = N\partial_\beta^2 b(0) + \partial_\beta^2 f(\xi) \geq 0 \quad \textit{for any } \beta \in S^{n-1}.$$

In particular, $\partial_\beta^2 f$, $\beta \in S^{n-1}$, are bounded together from below on X.

Therefore, Lemma 1 can be extended to the case of unbounded sets only if the second-order directional derivatives $\partial_\beta^2 f$, $\beta \in S^{n-1}$, together are bounded from below on X. In the general case, instead of the constant N, we need to choose a function that depends on $|\xi|$. Namely, assuming that f has a second-order increase of a polynomial , i.e.,

$$|\partial_\beta^2 f(x)| \leq D(1 + |x|)^\alpha \quad \text{for } \ x \in X \text{ i } \beta \in S^{n-1},$$

for some $D > 0$, and $\alpha \in \mathbb{N}$ and $b : \mathbb{R}^n \to \mathbb{R}$ is a function of class $\mathscr{C}^k$, $k \geq 2$, and logarithmically μ-strongly convex, $\mu > 0$, such that

$$0 = \operatorname{argmin}_{\mathbb{R}^n} b \qquad \text{i} \qquad b(0) = 1,$$

where $\operatorname{argmin}_X b$ is a point in X where b takes the smallest value in X, we have

Lemma 4 (see Lemma 3.2.1). *Let*

$$N(|\xi|) = \frac{D}{\mu}\left(|\xi| + 1 + \sqrt{\frac{\alpha}{\mu}}\right)^{\alpha} + 1$$

Then, for any $\xi \in \mathbb{R}^n$ a function $\phi_\xi : \mathbb{R}^n \to \mathbb{R}$ defined by

$$\phi_\xi(x) = N(|\xi|)b(x - \xi) + f(x), \quad x \in \mathbb{R}^n,$$

is strongly convex on X (more precisely μ-strongly convex).

In particular, the assertion of the above lemma can be obtained for the function $\psi_\xi(x) = Nb(\xi)b(x - \xi) + f(x)$, for a sufficiently large constant N (see lemat 3.2.5).

In the case when we obtain the convexifying of a function by multiplying it by $x \mapsto b(N(x - \xi))$, where b is a strongly convex function or logarithmically strongly convex, we must of course assume that the function only takes positive values on X. Then we have

Fact 5 (see Fact 3.1.5). *Let $b : \mathbb{R}^n \to (0, +\infty)$ be a μ-strongly convex function of class $\mathscr{C}^2$ such that $0 = \operatorname{argmin}_{\mathbb{R}^n} b$, let $X \subset \mathbb{R}^n$ be a convex and closed set and let $f : \mathbb{R}^n \to (0, +\infty)$ be a function of class $\mathscr{C}^2$. Let*

$$\varphi_{N,\xi}(x) = b(N(x - \xi))f(x),$$

where $N > 0$.

(i) *If for any $\xi \in X$, $\varphi_{N,\xi}$ is a strictly convex function on X, then $C\partial_\beta^2 f(x) \geq -f(x)$ for any $x \in X$ and $\beta \in S^{n-1}$ and some constant $C > 0$.*

(ii) *If b is the logarithmically μ-strongly convex function and $C\partial_\beta^2 f(x) \geq -f(x)$ and $Cf(x) \geq |\partial_\beta f(x)|$ for any $x \in X$, $\beta \in S^{n-1}$ and some constant $C > 0$, then for $N > \frac{2C}{\sqrt{mu}} + \frac{1}{C\mu}$, the function $\varphi_{N,\xi}$ is strictly convex on X.*

The main difficulty in applying the above fact is the estimation of the constant C. This difficulty can be overcome when we convexifying of the polynomial. More specifically, let $f \in \mathbb{R}[x]$, where $x = (x_1, \ldots, x_n)$ is a system of a variable, be a polynomial of degree d, and let

$$f = f_0 + \cdots + f_d,$$

where f_j is a homogeneous polynomial of degree j or zero. Let

$$f_{d*} = \min_{|x|=1} f_d(x).$$

Obviously, $f_{d*} > 0$ if and only if the leading form f_d of the polynomial f takes only positive values $\mathbb{R}^n \setminus \{0\}$.

Let $b : \mathbb{R}^n \to \mathbb{R}$ be a logarithmically μ-strongly convex function of a class $\mathscr{C}^k$, $k \geq 2$. We assume that

$$0 = \operatorname{argmin}_{\mathbb{R}^n} b \quad \text{i} \quad b(0) = 1.$$

Then we can get a convexifying of polynomial f by multiplying it by the function $b(N(x-\xi))$, $\xi \in \mathbb{R}^n$. Namely, we have

Theorem 6 (see Theorem 3.3.1)**.** *Assume that $f_{d*} > 0$ and there exists $m > 0$ such that*

$$f(x) \geq m \quad \textit{for } x \in \mathbb{R}^m.$$

Then there is an effectively compatible N_0 such that for any $N > N_0$ and for any $\xi \in \mathbb{R}^n$ the function $\varphi_{N,\xi} : \mathbb{R}^n \to \mathbb{R}$ defined by

$$\varphi_{N,\xi}(x) = b(N(x-\xi))f(x)$$

is μ-strongly convex in $\mathbb{R}^n$.

Theorem 3.3.3 is a similar type. It can be shown that theories similar of theorems 6 and 3.3.3 hold for $\mathbb{R}^n \ni x \mapsto b^N(x-\xi)f(x) \in \mathbb{R}$, but the proofs of these versions are much more complicated in terms of computation.

Chapters 4, 5 and 6 we deal with iterations of a mapping that assigns to each point the only critical point of the convexifying of function f.

In Chapter 4, we assume that $X \subset \mathbb{R}^n$ is a convex and compact set, and that the function $b : \mathbb{R}^n \to \mathbb{R}$ is strongly convex of class $\mathscr{C}^k$, $k \geq 2$ such that $0 = \operatorname{argmin}_{\mathbb{R}^n}$. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function of the class $\mathscr{C}^k$. Then there is a number $N \geq 1$ such that for any $\xi \in \mathbb{R}^n$, the function

$$\phi_{N,\xi}(x) = Nb(x-\xi) + f(x)$$

is strongly convex on X. We define a mapping

$$\kappa_N : X \ni \xi \mapsto \operatorname{argmin}_X \phi_{N,\xi} \in \mathbb{R}^n.$$

If

$$X_{f\leq r} := \{x \in \mathbb{R}^n : f(x) \leq r\} \subset X,$$

we show that

Lemma 7 (see Lemma 4.1.3 i 4.1.4). *The following properties hold:*

(i) The mapping κ_N is a diffeomorphism of class $\mathscr{C}^{k-1}$ from $X_{f\leq r}$ to $Y = \kappa_N(X_{f\leq r}) \subset X_{f\leq r}$.

(ii)The set of fixed points of $\kappa_N|_{X_{f\leq r}}$ is equal to $\Sigma_f \cap X_{f\leq r}$, where Σ_f is the set of critical points of f.

In theorem 4.2.1 (c) we show that

Theorem 8. *If $f : \mathbb{R}^n \to \mathbb{R}$ is a semialgebraic function of class $\mathscr{C}^2$ then for any $\xi \in X_{f\leq r}$, the limit point $\lim_{\nu\to\infty} \kappa_N^\nu(\xi)$ exists and belongs to Σ_f.*

The proof of this theorem, is based on showing the monotonicity of the sequence $f(\xi_\nu)$ (see wniosek 4.1.5) and applying the comparison principle (see [9, Lemma 7.7]) to show that the series $\sum_{\nu=0}^{\infty} \mathrm{dist}(\kappa_N^\nu(\xi), f^{-1}(f(\kappa_N^{\nu+1}(\xi))))$ is convergent. The idea of this proof is based on the proof of [9, Theorem 7.5]). The proof of this theorem is not a direct transfer of [9, Theorem 7.5]), because in [9] considered convexification of f by multiplying it by $(1+|x|^2)^N$, and we consider convexification this function by adding Nb to it.

Assuming that the function f is semialgebraic, the above theorem allows us to define the mapping

$$\kappa_{N,*} : X_{f\leq r} \to \Sigma_f \cap X_{f\leq r}\,,$$

given by $\kappa_{N,*}(\xi) = \lim_{\nu\to\infty} \kappa_N^\nu(\xi)$.

Assuming that the function f has only one critical value on $X_{f\leq r}$, we will show that the mapping $\kappa_{N,*}$ is continuous. Namely, we have

Theorem 9 (see Theorem 4.3.1). *Let $0 \in \operatorname{Int} X_{f\leq r}$ and let $f(0)$ be the minimal value of f. Then there exists $f(0) < \delta < r$ such that the sequence κ_N^ν uniformly convergents to $\kappa_{N,*}$ in the set $U = X_{f\leq\delta}$. In particular the mapping*

$$\kappa_{N,*}|_U : U \to U \cap \Sigma_f$$

is continuous and $\kappa_{N,*}(\xi) = \xi$ for $\xi \in U \cap \Sigma_f$. Consequently $\kappa_{N,*}|_U$ is a deformation retraction and the set $U \cap \Sigma_f$ is a retract of U.

In Chapter 5, we transfer some properties of the mapping κ_N to the case of unbounded sets. Among other things, in the case where the convexification of the function f is of the form

$$\psi_\xi(x) = N(\xi)b(x-\xi) + f(x), \qquad (\xi, x) \in \mathbb{R}^n \times \mathbb{R}^n$$

and for a convex and closed set X,

$$\kappa(\xi) = \operatorname{argmin}_X \psi_\xi \in X,$$

we show

Fact 10 (see Fact 5.1.1)**.** *For any $\xi \in X$, the point $\kappa(\xi)$ is the anique lower critical point of ψ_ξ on X.*

A similar fact holds for the function

$$\kappa_N : X \ni \xi \mapsto \operatorname{argmin}_X b(N(x-\xi))f(x).$$

In Chapter 6, we trasver the results from chapter four for the mapping $X_{f\le r} \ni \xi \mapsto \operatorname{argmin}_{|x|\le R} b^N(x-\xi)f(x) \in X$, assuming that the set $X_{f\le r}$ is compact and convex. In this case, all of the above properties of κ_N are true. If we additionally assume that $b(x) = \exp(|x|^2)$ and f is a polynomial, then the mapping κ_N has some additional properties, among others it is an analytic and semialgebraic mapping, i.e. it is a Nash maping. Namely, we have

Fact 11 (see Fact 6.3.2)**.** *The mapping $\kappa_N : X_{f,r} \to \kappa_N(X_{f,r})$ is the inverse of*

$$\kappa_N(X_{f,r}) \ni x \mapsto x + \frac{1}{2Nf(x)}\nabla f(x) \in X_{f,r}\,,$$

so it is an analytic and semialgebraic mapping, i.e., it is a Nash mapping.

At the end of chapter six, we deal with the convergence problem of the sequence $\frac{\xi_\nu}{|\xi_\nu|}$, where $\xi_\nu = \kappa_N^\nu(\xi)$, $\nu \in \mathbb{N}$, and $\xi \in \mathbb{R}^n$, i.e. the problem of convergence of a sequence of spherical parts of the sequence ξ_ν. This is a transfer of Rene Thom's problem for the gradient field trajectory (solved in [7]) to the discrete case. We consider this problem assuming that $\xi_\nu \to 0$, when $\nu \to \infty$, and some additional assumptions. Namely, we show that the following fact holds.

Fact 12 (see Fact 6.3.19). *Let $f \in \mathbb{R}[x]$, $x = (x_1, \dots, x_n)$, be a polynomial of the form*

$$f = f_k + \cdots + f_d,$$

where f_j is a homogeneous polynomial of degree j or zero. Suppose $f_k(\theta) > 0$ for $\theta \in S^{n-1}$. Then there is the following limit

$$\lim_{\nu \to \infty} \frac{1}{|\xi_\nu|} \xi_\nu.$$

Moreover, the sequence $|\xi_\nu|$ is strictly decreasing from a certain point.

Wstęp

Jednym z podstawowych problemów analizy, techniki, ekonomii i innych gałęzi nauki jest poszukiwanie minimów i punktów krytycznych funkcji. Jedną z metod prowadzących do tego celu jest deformacja danej funkcji do funkcji wypukłej, poszukiwanie punktów krytycznych tej deformacji i iterowanie tego procesu. Sprowadzanie danej funkcji do funkcji wypukłej, czy silnie wypukłej prowadzi do łatwego wyznaczania punktów krytycznych i minimów tej deformacji. Są to dokładnie te punkty, w których gradient się zeruje. Klasycznym podejściem do uwypuklania funkcji $f : \mathbb{R}^n \to \mathbb{R}$ na zbiorach ograniczonych i wypukłych jest dodanie do tej funkcji takiej funkcji silnie wypukłej $b : \mathbb{R}^n \to \mathbb{R}$, że $f + b$ jest funkcją silnie wypukłą na tym zbiorze (patrz. np. [19], [11] i [20] dla funkcji kwadratowej $b(x) = \gamma|x|^2$, $\gamma > 0$). Omówimy to dokładniej.

Niech $X \subset \mathbb{R}^n$ będzie zbiorem wypukłym. Funkcję $g : X \to R$ nazywamy *silnie wypukłą*, gdy istnieje $\mu > 0$ takie, że

$$g(tx+(1-t)y) \leq tg(x)+(1-t)g(y)-t(1-t)\frac{\mu}{2}|x-y|^2 \text{ dla } x, y \in X \text{ i } 0 < t < 1.$$

Wtedy mówimy, że funkcja g jest μ-*silnie wypukła*. Jeśli dodatkowo funkcja g jest klasy $\mathscr{C}^1$, to powyższy warunek jest równoważny

$$g(y) \geq g(x) + \langle y - x, \nabla g(x)\rangle + \frac{\mu}{2}|y - x|^2 \quad \text{dla } x, y \in X,$$

gdzie $\langle \cdot, \cdot \rangle$ jest standardowym iloczynem skalarnym w $\mathbb{R}^n$, a ∇g jest gradientem funkcji g (patrz [16]). Jeśli g jest funkcją klasy $\mathscr{C}^2$, to równoznaczne z faktem, że pochodne kierunkowe drugiego rzędu $\partial_\beta^2 g(x)$ funkcji g we wszystkich kierunkach $\beta \in \mathbb{R}^n$ o długości 1 są ograniczone od dołu przez liczbę μ dla dowolnego $x \in X$. W pracy ten warunek będziemy często wykorzystywali.

Niech $b : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^k$, $k \geq 2$, μ-silnie wypukłą, $\mu > 0$.

Niech $X \subset \mathbb{R}^n$ będzie zbiorem zwartym i wypukłym, niech $f : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^k$ i niech $D \in \mathbb{R}$ będzie liczbą dodatnią taką, że

$$|\partial_\beta^2 f(x)| \leq D \quad dla\ x \in X \text{ i } \beta \in S^{n-1},$$

gdzie S^{n-1} jest sfera jenostkowa w $\mathbb{R}^n$. W klasycznym podejściu do uwypuklania funkcji na zbiorze zwartym, kluczową rolę odgrywa następujący łatwy do sprawdzenia lemat.

Lemat 1 (patrz uwaga 2.2.3 i fakt 3.1.1)**.** *Dla dowolnego $\xi \in \mathbb{R}^n$ oraz*

$$N > D/\mu,$$

funkcja $\phi_{N,\xi} : \mathbb{R}^n \to \mathbb{R}$ określona wzorem

$$\phi_{N,\xi}(x) = Nb(x - \xi) + f(x), \quad x \in \mathbb{R}^n,$$

jest silnie wypukła na zbiorze X (dokładniej $N\mu - D$-silnie wypukła).

W pracy porównamy powyższe podejście do uwypuklania funkcji, która przyjmuje tylko wartości dodatnie, z innym podejściem polegającym na pomnożeniu jej przez pewną potęgę funkcji silnie wypukłej (patrz rozdział 2 oraz rozdział 3). To drugie podejście zostało zaproponowane w pracy [9] i kontynuowane w pracy [8]. Dokładniej, w [9] uzyskano uwypuklenie funkcji dodatniej f klasy $\mathscr{C}^2$ na zbiorze zwartym i wypukłym $X \subset \mathbb{R}^n$ przez pomnożenie funkcji f przez $(1+|x|^2)^N$ dla pewnego N, a w [8] – przez pomnożenie funkcji f przez $\exp(N|x|^2)$. W rozdziale drugim uogólnimy te wyniki i pokażemy, że

Wniosek 2 (patrz wniosek 2.2.2)**.** *Jeśli $X \subset \mathbb{R}^n$ jest zbiorem zwartym i wypukłym oraz $f : X \to \mathbb{R}$ jest funkcją klasy $\mathscr{C}^2$ przyjmującą tylko wartości dodatnie, to dla dowolnej funkcji silnie wypukłej $b : \mathbb{R}^n \to \mathbb{R}$ istnieje $N_0 > 0$ takie, że dla każdego $N \geq N_0$ oraz $\xi \in X$, funkcja*

$$\varphi_{N,\xi}(x) = b^N(x - \xi) f(x), \quad x \in \mathbb{R}^n,$$

jest silnie wypukła na zbiorze X.

W przypadku, gdy funkcja f jest wielomianem, wykładnik N można oszacować efektywnie w terminach promienia zbioru X (to znaczy $\sup\{|x| : x \in X\}$)

wielkości modułów współczynników wielomianu oraz $m = \inf\{f(x) : x \in X\}$ (patrz wniosek 2.2.7). Wbec tego, w przypadku funkcji dodatnich na zbiorach zwartych i wypukłych, zarówno dodanie do funkcji pewnej wielokrotności funkcji silnie wypukłej jak i pomnożenie jej przez potęgę takiej funkcji dają podobny efekt, jednak pierwsza metoda prowadzi do użycia mniejszego współczynnika N. Jeśli dodatkowo założymy, że funkcja b jest logarytmicznie silnie wypukła (to znaczy $\ln b$ jest funkcją silnie wypukłą), to funkcja $\varphi_{N,\xi}$ również jest logarytmicznie silnie wypukła (patrz wniosek 2.2.9). W przypadku, gdy zbiór X jest semialgebraiczny, zwarty i wypukły, współczynniki wielomianów opisujących zbiór X oraz współczynniki wielomianu f są liczbami całkowitymi (lub wymiernymi), to wykładnik N można wyznaczyć w pełni efektywnie (patrz Twierdzenia 2.2.10 i 2.2.12). Twierdzenia te uzyskujemy stosując wynik G. Jeronimo, D. Perrucci, E. Tsigaridas z pracy [6].

W rozdziale trzecim porównamy klasyczne podejście do problemu uwypuklania funkcji z powyższym dla dowolnej funkcji silnie wypukłej $b : \mathbb{R}^n \to \mathbb{R}$ i funkcji dodatniej na zbiorze domkniętym i wypukłym (niekoniecznie ograniczonym). W tym rozdziale będziemy uwypuklać funkcje f przez pomnożenie jej przez $b(N(x-\xi))$, zamiast przez $b^N(x-\xi)$. Podejście takie upraszcza niektóre obliczenia.

Trudno jest zastosować Lemat 1 w przypadku zbiorów nieograniczonych. Mianowicie mamy

Fakt 3 (patrz fakt 3.1.2)**.** *Niech $b : \mathbb{R}^n \to (0,+\infty)$ będzie funkcją wypukłą klasy $\mathscr{C}^2$, Niech $X \subset \mathbb{R}^n$ będzie zbiorem wypukłym i domkniętym, niech $f : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^2$ i niech $N > 0$. Jeśli dla każdego $\xi \in X$ funkcja $\phi_{N,\xi}$ określona wzorem* (1) *jest wypukła na zbiorze X, to*

$$\partial^2_\beta \phi_{N,\xi}(\xi) = N\partial^2_\beta b(0) + \partial^2_\beta f(\xi) \geq 0 \quad \textit{dla każdego } \beta \in S^{n-1}.$$

W szczególności pochodne $\partial^2_\beta f$, $\beta \in S^{n-1}$, są wspólnie ograniczone od dołu na zbiorze X.

W związku z powyższym, lemat 1 można przenieść do przypadku zbiorów nieograniczonych tylko przy założeniu, że pochodne kierunkowe drugiego rzędu $\partial^2_\beta f$, $\beta \in S^{n-1}$, są ograniczone od dołu na zbiorze X. W ogólnym przypadku zamiast stałej N musimy wybrać funkcję zależną od $|\xi|$. Mianowicie, przy założeniu, że funkcja f posiada wzrost wielomianu drugiego rzędu, to znaczy

$$|\partial_\beta^2 f(x)| \leq D(1+|x|)^\alpha \quad \text{dla } \ x \in X \text{ i } \beta \in S^{n-1},$$

dla pewnych $D > 0$ i $\alpha \in \mathbb{N}$ oraz $b : \mathbb{R}^n \to \mathbb{R}$ jest funkcją klasy $\mathscr{C}^k$, $k \geq 2$, i logarytmicznie μ-silnie wypukłą, $\mu > 0$, taką że

$$0 = \operatorname{argmin}_{\mathbb{R}^n} b \qquad \text{i} \qquad b(0) = 1,$$

gdzie $\operatorname{argmin}_X b$ jest jedynym punktem zbioru X, w którym funkcja b przyjmuje najmniejszą wartość w zbiorze X, mamy

Lemat 4 (patrz lemat 3.2.1). *Niech*

$$N(|\xi|) = \frac{D}{\mu}\left(|\xi| + 1 + \sqrt{\frac{\alpha}{\mu}}\right)^\alpha + 1 \tag{1}$$

Wówczas dla dowolnego $\xi \in \mathbb{R}^n$ *funkcja* $\phi_\xi : \mathbb{R}^n \to \mathbb{R}$ *określona wzorem*

$$\phi_\xi(x) = N(|\xi|)b(x-\xi) + f(x), \quad x \in \mathbb{R}^n,$$

jest silnie wypukła na zbiorze X *(dokładniej* μ*-silnie wypukła).*

W szczególności można uzyskać tezę powyższego wniosku dla funkcji $\psi_\xi(x) = Nb(\xi)b(x-\xi) + f(x)$, dla dostatecznie dużej stałej N (patrz lemat 3.2.5).

W przypadku, gdy uwypuklenie funkcji otrzymujemy przez pomnożenie jej przez $x \mapsto b(N(x-\xi))$, gdzie b jest funkcją silnie wypukłą lub logarytmicznie silnie wypukłą, musimy oczywiście założyć, że funkcja przyjmuje tylko wartości dodatnie na zbiorze X. Wówczas mamy

Fakt 5 (patrz fakt 3.1.5). *Niech* $b : \mathbb{R}^n \to (0, +\infty)$ *będzie funkcją* μ*-silnie wypukłą klasy* $\mathscr{C}^2$ *taką, że* $0 = \operatorname{argmin}_{\mathbb{R}^n} b$*, niech* $X \subset \mathbb{R}^n$ *będzie zbiorem wypukłym i domkniętym oraz niech* $f : \mathbb{R}^n \to (0, +\infty)$ *będzie funkcją klasy* $\mathscr{C}^2$*. Niech*

$$\varphi_{N,\xi}(x) = b(N(x-\xi))f(x),$$

gdzie $N > 0$.

(i) *Jeśli dla dowolnego $\xi \in X$ funkcja $\varphi_{N,\xi}$ jest ściśle wypukła na zbiorze X, to $C\partial_\beta^2 f(x) \geq -f(x)$ dla dowolnych $x \in X$ i $\beta \in S^{n-1}$ i pewnej stałej $C > 0$.*

(ii) *Jeśli fankcja b jest logarytmicznie μ-silnie wypukła i $C\partial_\beta^2 f(x) \geq -f(x)$ oraz $Cf(x) \geq |\partial_\beta f(x)|$ dla dowolnych $x \in X$, $\beta \in S^{n-1}$ i pewnej stałej $C > 0$, to dla $N > \frac{2C}{\sqrt{\mu}} + \frac{1}{C\mu}$, funkcja $\varphi_{N,\xi}$ jest ściśle wypukła na zbiorze X.*

Główną trudnością w stosowaniu powyższego faktu jest oszacowanie stałej C. Trudność tę można pokonać w przypadku, gdy uwypuklamy wielomiany. Dokładniej, niech $f \in \mathbb{R}[x]$, gdzie $x = (x_1, \ldots, x_n)$ jest układem zmiennym, będzie wielomienem stopnia d i niech

$$f = f_0 + \cdots + f_d,$$

gdzie f_j jest wielomianem jednorodnym stopnia j lub zerem. Niech

$$f_{d*} = \min_{|x|=1} f_d(x).$$

Oczywiście $f_{d*} > 0$ wtedy i tylko wtedy, gdy forma wiodąca f_d wielomianu f przyjmuje tylko wartości dodatnie w $\mathbb{R}^n \setminus \{0\}$.

Niech $b : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją logarytmicznie μ-silnie wypukłą klasy $\mathscr{C}^k$, $k \geq 2$. Załóżmy, że

$$0 = \operatorname{argmin}_{\mathbb{R}^n} b \quad \text{i} \quad b(0) = 1.$$

Wówczas możemy uzyskać uwypuklenie wielomianu f przez pomnożenie go przez funkcję $b(N(x - \xi))$, $\xi \in \mathbb{R}^n$. Mianowicie, mamy

Twierdzenie 6 (patrz twierdzenie 3.3.1)**.** *Załóżmy, że $f_{d*} > 0$ oraz istnieje $m > 0$ takie, że*

$$f(x) \geq m \quad \text{dla } x \in \mathbb{R}^m.$$

Wówczas istnieje, efektywnie wyliczalne, N_0 takie, że dla dowolnego $N > N_0$ oraz dla dowolnego $\xi \in \mathbb{R}^n$, funkcja $\varphi_{N,\xi} : \mathbb{R}^n \to \mathbb{R}$ określona wzorem

$$\varphi_{N,\xi}(x) = b(N(x - \xi))f(x)$$

jest μ-silnie wypukła w $\mathbb{R}^n$.

Podobnego typu jest twierdzenie 3.3.3. Można pokazać, że zachodzą odpowiedniki twierdzeń 3.3.1 i 3.3.3 dla funkcji $\mathbb{R}^n \ni x \mapsto b^N(x-\xi)f(x) \in \mathbb{R}$, jednak dowody tych wersji są o wiele bardziej skomplikowane pod względem rachunkowym.

W rozdziałach czwartym, piątym i szóstym zajmujemy się iteracjami odwzorowania, które każdemu punktowi przypisuje jedyny punkt krytyczny uwypuklenia funkcji f.

W rozdziale czwartym zakładamy, że $X \subset \mathbb{R}^n$ jest zbiorem wypukłym i zwartym oraz, że funkcja $b : \mathbb{R}^n \to \mathbb{R}$ jest silnie wypukła klasy $\mathscr{C}^k$, $k \geq 2$, taka że $0 = \operatorname{argmin}_{\mathbb{R}^n} b$. Niech $f : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^k$. Wówczas istnieje liczba $N \geq 1$, taka, że dla dowolnego $\xi \in \mathbb{R}^n$, funkcja

$$\phi_{N,\xi}(x) = Nb(x-\xi) + f(x)$$

jest silnie wypukła na zbiorze X. Określamy odwzorowanie

$$\kappa_N : X \ni \xi \mapsto \operatorname{argmin}_X \phi_{N,\xi} \in \mathbb{R}^n.$$

Jeśli

$$X_{f\leq r} := \{x \in \mathbb{R}^n : f(x) \leq r\} \subset X,$$

to pokazujemy, że

Lemat 7 (patrz lematy 4.1.3 i 4.1.4)**.** *Zachodzą następujące własności:*

(i) Odwzorowanie κ_N jest dyfeomorfizmem klasy $\mathscr{C}^{k-1}$ ze zbioru $X_{f\leq r}$ na zbiór $Y = \kappa_N(X_{f\leq r}) \subset X_{f\leq r}$.

(ii) Zbiorem punktów stałych odwzorowania $\kappa_N|_{X_{f\leq r}}$ jest $\Sigma_f \cap X_{f\leq r}$, gdzie Σ_f jest zbiorem punktów krytycznych funkcji f.

W twierdzeniu 4.2.1 (c) pokazujemy, że

Twierdzenie 8. *Jeśli funkcja f jest semialgebraiczna, to dla dowolnego $\xi \in X_{f\leq r}$, punkt graniczny $\lim_{\nu\to\infty} \kappa_N^\nu(\xi)$ istnieje i należy do zbioru Σ_f.*

Dowód tego twierdzenia opiera sie na pokazaniu monotoniczności ciągu $f(\xi_\nu)$ (patrz wniosek 4.1.5) i zastosowaniu zasady porównania (patrz [9, Lemma 7.7]) do pokazania, że szereg $\sum_{\nu=0}^{\infty} \operatorname{dist}(\kappa_N^\nu(\xi), f^{-1}(f(\kappa_N^{\nu+1}(\xi))))$ jest zbieżny.

Idea tego dowodu jest wzorowana na dowodzie [9, Theorem 7.5]). Dowód tego twierdzenia nie jest bezpośrednim przeniesieniem dowodu [9, Theorem 7.5]), gdyż w [9] rozważano uwypuklanie funkcji f przez pomnożenie jej przez $(1 + |x|^2)^N$, a my rozważamy uwypuklanie tej funkcji przez dodanie do niej Nb.

Przy założeniu, że funkcja f jest semialgebraiczna, powyższe twierdzenie pozwala określić odwzorowanie

$$\kappa_* : X_{f\leq r} \to \Sigma_f \cap X_{f\leq r}\,,$$

wzorem $\kappa_{N,*}(\xi) = \lim_{\nu\to\infty} \kappa_N^\nu(\xi)$.

Przy założeniu, że funkcja f ma tylko jedną wartość krytyczną na zbiorze $X_{f\leq r}$, pokażemy, że odwzorowanie $\kappa_{N,*}$ jest ciągłe. Mianowicie, mamy

Twierdzenie 9 (patrz twierdzenie 4.3.1)**.** *Niech $0 \in \operatorname{Int} X_{f\leq r}$ oraz niech $f(0)$ będzie najmniejszą wartością funkcji f. Wtedy istnieje $f(0) < \delta < r$ taki, że ciąg κ_N^ν jest zbieżny jednostajnie do $\kappa_{N,*}$ w zbiorze $U = X_{f\leq\delta}$. W szczególności odwzorowanie*

$$\kappa_{N,*}|_U : U \to U \cap \Sigma_f$$

jest ciągłe oraz $\kappa_{N,*}(\xi) = \xi$ wtedy i tylko wtedy, gdy $\xi \in U\cap\Sigma_f$. W konsekwencji $\kappa_{N,*}|_U$ jest retrakcją, a zbiór $U \cap \Sigma_f$ jest retraktem zbioru U.

W rozdziale piątym przenosimy pewne własności odwzorowania κ_N na przypadek zbiorów nieograniczonych. Między innymi, w przypadku, gdy uwypuklenie funkcji f jest postaci

$$\psi_\xi(x) = N(\xi)b(x-\xi) + f(x), \qquad (\xi, x) \in \mathbb{R}^n \times \mathbb{R}^n$$

oraz dla zbioru wypukłego i domkniętego X,

$$\kappa(\xi) = \operatorname{argmin}_X \psi_\xi \in X,$$

pokazujemy

Fakt 10 (patrz fakt 5.1.1)**.** *Dla dowolnego $\xi \in X$ punkt $\kappa(\xi)$ jest jedynym dolnym punktem krytycznym funkcji ψ_ξ na zbiorze X.*

Przypomnijmy, że punkt $a \in X$ jest *dolnym punktem krytycznym funkcji* różniczkowalnej g na zbiorze X, jeśli

(2) $\langle \nabla g(a), x - a \rangle \geq 0 \quad \text{dla } x \in X \quad \text{w pewnym otoczeniu punktu } a.$

Podobny fakt zachodzi dla funkcji $\kappa_N : X \ni \xi \mapsto \operatorname{argmin}_X b(N(x-\xi))f(x)$.

W rozdziale szóstym przenosimy wyniki z rozdziału czwartego dla odwzorowania $X_{f\leq r} \ni \xi \mapsto \operatorname{argmin}_{|x|\leq R} b^N(x-\xi)f(x) \in X$, przy założeniu, że zbiór $\kappa_N : X_{f\leq r}$ jest zwarty i wypukły. W tym przypadku zachodzą wszystkie powyżej przedstawionme własności odwzorowania κ_N. Jeśli dodatkowo założymy, że $b(x) = \exp(|x|^2)$ oraz f jest wielomianem, to odwzorowanie κ_N ma pewne dodatkowe włsności, między innymi jest to odwzorowanie analityczne i semialgebraiczne, czyli jest odwzorowaniem Nasha. Mianowicie, mamy

Fakt 11 (patrz fakt 6.3.2)**.** *Jeśli f jest wielomianem dodatnim na zbiorze $X_{f\leq r}$ oraz $b(x) = \exp(|x|^2)$, to dla dostatecznie dużego N, odwzorowanie $\kappa_N : X_{f\leq r} \to \kappa_N(X_{f\leq r})$ jest odwrotnością odwzorowania*

$$\kappa_N(X_{f\leq r}) \ni x \mapsto x + \frac{1}{2Nf(x)}\nabla f(x) \in X_{f\leq r}\,,$$

więc jest analityczne i semialgebraiczne, t.j., jest odwzorowaniem Nasha.

Na końcu rozdziału szóstego zajmujemy się problemem zbieżności ciągu $\frac{\xi_\nu}{|\xi_\nu|}$, gdzie $\xi_\nu = \kappa_N^\nu(\xi)$, $\nu \in \mathbb{N}$, oraz $\xi \in \mathbb{R}^n$, czyli problem zbieżności ciągu części sferycznych ciągu ξ_ν. Jest to przeniesienie problemu René Thoma dla trajektorii pola gradientowego (rozwiązanego w [7]) do przypadku dyskretnego. Rozważamy ten problem przy założeniu, że $\xi_\nu \to 0$ gdy $\nu \to \infty$ oraz kilku dodatkowych założeniach. Mianowicie, pokazujemy, że zachodzi następujący fakt.

Fakt 12 (patrz fakt 6.3.19)**.** *Niech $f \in \mathbb{R}[x]$, $x = (x_1, \dots, x_n)$, będzie wielomienem stopnia postaci*

$$f = f_k + \cdots + f_d,$$

gdzie f_j jest wielomianem jednorodnym stopnia j lub zerem. Załóżmy, że $f_k(\theta) > 0$ dla $\theta \in S^{n-1}$. Wówczas istnieje następująca granica

$$\lim_{\nu\to\infty} \frac{1}{|\xi_\nu|}\xi_\nu.$$

Co więcej, ciąg $|\xi_\nu|$ jest ściśle malejący od pewnego miejsca.

Część wyników tej pracy zostało już opublikowane w pracy [1]. Dotyczy to punktu 2.2 i Rozdziału 6

Rozdział 1

Preliminaria

W tym rozdziale wprowadzimy podstawowe oznaczenia i przypomnimy elementarne fakty z zakresu funkcji wypukłych, funkcji semialgebraicznych, gradientu funkcji wypukłych, punktów krytycznych funkcji i logarytmicznego uwypuklenia funkcji, przydatne w dalszym ciągu pracy.

1.1 Oznaczenia

Symbolem $\mathbb{R}$ będziemy oznaczać ciało liczb rzeczywistych, a symbolem $\mathbb{Z}$ - pierścień liczb całkowitych. Przez $\mathbb{N}$ będziemy oznaczali zbiór liczb naturalnych.

Standardowym iloczynem skalarnym w przestrzeni $\mathbb{R}^n$ nazywamy funkcję $\langle\cdot,\cdot\rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ określoną wzorem:

$$\langle x, y\rangle = x_1y_1 + ... + x_ny_n \quad \text{dla } x = (x_1, ..., x_n),\ y = (y_1, ..., y_n).$$

Przez *normę euklidesową* w przestrzeni $\mathbb{R}^n$ rozumiemy funkcję $\mathbb{R}^n \ni x \mapsto |x| \in \mathbb{R}^0_+$ określoną wzorem $|x| = \sqrt{\langle x, x\rangle}$, to jest

$$|x| = \sqrt{x_1^2 + \cdots + x_n^2}, \quad \text{gdzie } x = (x_1, \ldots, x_n).$$

W dalszym ciągu pracy zakładamy, że przestrzeń $\mathbb{R}^n$ została wyposażona w metrykę wyznaczoną przez normę euklidesową. Odległość między punktami $x, y \in \mathbb{R}^n$ w tej metryce oznaczamy $|x - y|$.

Niech f będzie funkcją rzeczywistą określoną w zbiorze otwartym $D \subset \mathbb{R}^n$. Oznaczmy przez $\partial_v f(x)$ *pochodną kierunkową funkcji* f *w kierunku wektora* $v \in \mathbb{R}^n$ *w punkcie* $x \in D$. Przez $\partial_v f$ oznaczamy *pochodną kierunkową funkcji* f *w kierunku wektora* v, to jest funkcję $\partial_v f : D \ni x \mapsto \partial_v f(x) \in \mathbb{R}$. Pochodną kierunkową funkcji $\partial_v f$ w punkcie $x \in D$ w kierunku wektora $w \in \mathbb{R}^n$ oznaczamy $\partial^2_{v,w} f(x)$ i nazywamy *pochodna kierunkowa drugiego rzędu funkcji* f *w kierunku wektorów* v, w *w punkcie* x. W przypadku, gdy $w = v$, pochodną kierunkową $\partial^2_{v,v} f(x)$ nazywamy *pochodną kierunkową drugiego rzędu* funkcji f w kierunku wektora v w punkcie x i oznaczamy $\partial^2_v f(x)$.

Jeśli $v = (0, \dots, 0, 1, 0, \dots, 0)$, gdzie 1 jest na $i-$tym miejscu, to piszemy również $\partial_v f = \frac{\partial f}{\partial x_i}$ i nazywamy *pochodną cząstkową* funkcji f po i-tej zmiennej. Wtedy gradient $\nabla f : D \to \mathbb{R}^n$ funkcji f definiujemy wzorem

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right).$$

Przypomnijmy, że jeśli wszystkie pochodne cząstkowe funkcji f istnieją w zbiorze D i są one ciągłe, to mówimy, że funkcja f jest klasy $\mathscr{C}^1$. Jeśli wszystkie pochodne cząstkowe wszystkich pochodnych cząskowych (to jest pochodne drugiego rządu) funkcji f są ciągłe, to mówimy, że funkcja jest klasy $\mathscr{C}^2$. Jeśli funkcja f jest klasy $\mathscr{C}^k$ i wszystkie pochodne cząstkowe $k+1-$rzędu funkcji f istnieją w zbiorze D i są ciagłe, to mówimy, że funkcja jest klasy $\mathscr{C}^{k+1}$.

1.2 Funkcje wypukłe

Niech X będzie podzbiorem przestrzeni $\mathbb{R}^n$. Przypomnijmy, że zbiór X nazywamy *wypukłym*, jeśli

$$\forall_{x,y \in X} \ \ \forall_{0<t<1} \ \ (1-t)x + ty \in X.$$

Funkcję $f : X \to \mathbb{R}$ nazywamy *wypukłą*, gdy zbiór X jest wypukły i

$$\forall_{x,y \in X} \ \ \forall_{0<t<1} \ \ f\big(tx + (1-t)y\big) \leq tf(x) + (1-t)f(y). \tag{1.1}$$

Funkcję $f : X \to \mathbb{R}$ nazywamy *ściśle wypukłą*, gdy zbiór X jest wypukły i

(1.2) $\qquad \forall_{x,y\in X,\, x\neq y}\ \forall_{0<t<1}\ f\big(tx+(1-t)y\big) < tf(x)+(1-t)f(y).$

Niech $f : X \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^2$ oraz niech X będzie zbiorem wypukłym. Dla dowolnych $a \in X$ i $v \in \mathbb{R}^n$ kładziemy

$$I_{a,v} = \{t \in \mathbb{R} : a + tv \in X\}.$$

Oczywiście, $I_{a,v}$ jest przedziałem lub zbiorem jednoelementowym. Przypomnijmy pewne znane fakty (patrz [3], [4], Theorem 22.5 i [5], Theorem $A2.1$).

Fakt 1.2.1. *Następujące warunki są równoważne:*

(a) *Funkcja f jest wypukła.*

(b) *Dla dowolnego wektora $v \in \mathbb{R}^n$ i dowolnego punktu $a \in X$ funkcja $I_{a,v} \ni t \mapsto \partial_v f(a+tv) \in \mathbb{R}$ jest rosnąca (tzn. niemalejąsa).*

(c) *Dla dowolnego wektora $v \in \mathbb{R}^n$ i dowolnego punktu $a \in X$ zachodzi $\partial_v^2 f(a) \geq 0$.*

Fakt 1.2.2. *Następujące warunki są równoważne:*

(a) *Funkcja f jest ściśle wypukła.*

(b) *Dla dowolnego wektora $v \in \mathbb{R}^n$ o dodatniej długości i dowolnego punktu $a \in X$ funkcja $I_{a,v} \ni t \mapsto \partial_v f(a+tv) \in \mathbb{R}$ jest ściśle rosnąca.*

(c) *Funkcja f jest wypukła oraz dla dowolnego wektora $v \in \mathbb{R}^n$ o dodatniej długości i dowolnego punktu $a \in X$ zbiór $\{t \in I_{a,v} : \partial_v^2 f(a+tv) > 0\}$ jest zbiorem gęstym w $I_{a,v}$, pod warunkiem, że $I_{a,v}$ jest przedziałłem.*

Przypomnijmy jeden prosty, ale użyteczny fakt.

Fakt 1.2.3. *Jeśli funkcja $f : X \to \mathbb{R}$ jest ściśle wypukła klasy $\mathscr{C}^2$, to gradient $\nabla f : X \to \mathbb{R}^n$ jest odwzorowaniem różnowartościowym.*

Dowód. Weźmy dowolne $a, b \in X$ takie, że $a \neq b$. Wówczas $I_{a,b-a}$ jest przedziałem oraz $0, 1 \in I_{a,b-a}$. Ponieważ f jest funkcją ściśle wypukłą, więc z faktu 1.2.2 (b), funkcja

$$\varphi : I_{a,b-a} \ni t \mapsto \partial_{b-a} f(a + t(b-a)) \in \mathbb{R}$$

jest ściśle rosnąca. Zatem

$$\langle \nabla f(a), b-a\rangle = \varphi(0) < \varphi(1) = \langle \nabla f(b), b-a\rangle,$$

i w konsekwencji, $\nabla f(a) \neq \nabla f(b)$. $\square$

1.3 Funkcje silnie wypukłe

Przypomnijmy, że funkcja $f : X \to \mathbb{R}$ jest *silnie wypukła* lub *μ-silnie wypukła*, $\mu > 0$, jeśli zbiór $X \subset \mathbb{R}^n$ jest wypukły oraz

$$\forall_{x,y\in X}\ \forall_{0<t<1}\ \ f(tx+(1-t)y) \leq tf(x) + (1-t)f(y) - t(1-t)\frac{\mu}{2}|x-y|^2.$$

Jeśli dodatkowo funkcja f jest klasy $\mathscr{C}^1$ w pewnym zbiorze otwartym zawierającym zbiór X, to powyższy warunek jest równoważny następującemu warunkowi (patrz [16])

$$\forall_{x,y\in X}\ \ f(y) \geq f(x) + \langle y-x, \nabla f(x)\rangle + \frac{\mu}{2}|y-x|^2.$$

Oczywiscie każda funkcja silnie wypukła jest ściśle wypukła.

Niech $f : X \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^2$ w pewnym otoczeniu zbioru $X \subset \mathbb{R}^n$ i niech zbiór X bedzie wypukły.

Oznaczamy przez S^{n-1} sferę jednostkową w $\mathbb{R}^n$, to znaczy

$$S^{n-1} = \{v \in \mathbb{R}^n : |v| = 1\}.$$

Fakt 1.3.1 ([5], Theorem *A*2.4)**.** *Niech $\mu > 0$. Następujące warunki są równoważne:*

(a) *Funkcja f jest μ-silnie wypukła.*

(b) *Dla dowolnego wektora $v \in S^{n-1}$ zachodzi $\partial_v^2 f(x) \geq \mu$ w dowolnym punkcie $x \in X$.*

(c) *Dla dowolnego punktu $x \in X$, każda wartość własna macierzy Hessego*

$$H\big(f\big)(x) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right]_{1\leq i,j\leq n}$$

jest ograniczona od dołu przez μ.

Wprost z definicji dostajemy

Fakt 1.3.2. *Jeśli $f : \mathbb{R}^n \to \mathbb{R}$ jest funkcją silnie wypukłą klasy $\mathscr{C}^1$, to*

$$\lim_{|x|\to\infty} f(x) = +\infty$$

.

Wiadomo, że jeśli f jest funkcją ściśle (lub silnie) wypukłą na zbiorze $X \subset \mathbb{R}^n$, to istnieje jedyny punkt $x_0 \in X$ w którym funkcja f przyjmuje najmnejszą wartość. Punkt x_0 oznaczamy $\arg\min_X f$.

1.4 Funkcje logarytmicznie silnie wypukłe

Niech f będzie funkcją rzeczywistą klasy $\mathscr{C}^2$ określoną w otoczeniu zbioru wypukłego $X \subset \mathbb{R}^n$.

Jeśli $f(x) > 0$ dla $x \in X$, to funkcję f nazywamy *logarytmicznie wypukła*, *logarytmicznie ściśle wypukła* oraz *logarytmicznie μ-silnie wypukła* jeśli funkcja $\ln f$ jest odpowiednio wypukła, ściśle wypukła oraz $\mu-$silnie wypukła (por [16], [19]). Funkcje logarytmicznie μ-silnie wypukłe nazywamy również funkcjami *logarytmicznie silnie wypukłymi*. Oczywiście, $x \mapsto \exp(b(x))$, $x \mapsto \exp(\exp(b(x)), \ldots$, są funkcjami logarytmicznie silnie wypukłymi, dla silnie wypukłej funkcji b. Na przykład funkcja Gamma jest funkcją logarytmicznie wypukłą w zbiorze dodatnich liczb rzeczywistych.

Można bezpośrednio sprawdzić, że każda funkcja logarytmicznie wypukła, logarytmicznie ściśle wypukła i logarytmicznie silnie wypukła jest również odpowiednio wypukła, ściśle wypukła i silnie wypukła. Oczywiście każda logarytmicznie silna funkcja wypukła jest również logarytmicznie ściśle wypukła i w konsekwencji logarytmicznie wypukła.

Ponieważ suma funkcji silnie wypukłych jest silnie wypukła (patrz [4], Theorem 22.3), więc mamy

Fakt 1.4.1. *Iloczyn funkcji logarytmicznie wypukłych jest logarytmicznie wypukły. Co więcej, jeśli funkcje $b_1, b_2 : \mathbb{R}^n \to \mathbb{R}$ są odpowiednio logarytmicznie μ_1, μ_2-silnie wypukłe, to funkcja $b_1 b_2$ jest logarytmicznie $\mu_1 + \mu_2$- silnie wypukłą.*

Następujący fakt wynika z definicji.

Fakt 1.4.2. *Niech $b : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^2$ przyjmującą tylko wartości dodatnie i niech $\mu > 0$. Następujące warunki są równoważne:*

(a) *Funkcja b jest logarytmicznie μ-silnie wypukła.*

(b) *Dla dowolnych $x \in \mathbb{R}^n$ i $\beta \in S^{n-1}$, zachodzi nierówność*

$$b(x)\partial_\beta^2 b(x) \geq \mu b^2(x) + (\partial_\beta b(x))^2.$$

Dowód. Ponieważ b jest funkcją klasy $\mathscr{C}^2$ oraz $b(x) > 0$ dla $x \in \mathbb{R}^n$, to funkcja $\ln b$ jest klasy $\mathscr{C}^2$ w $\mathbb{R}^n$ oraz

$$\partial_\beta^2(\ln b)(x) = \frac{b(x)\partial_\beta^2 b(x) - \big(\partial_\beta b(x)\big)^2}{\big(b(x)\big)^2} \quad \text{dla } x \in \mathbb{R}^n, \quad \beta \in S^{n-1}. \tag{1.3}$$

Załóżmy, że funkcja b jest logarytmicznie μ-silnie wypukła w $\mathbb{R}^n$, czyli że funkcja $\ln b$ jest μ-silnie wypukła. Weźmy dowolny $x \in \mathbb{R}^n$ oraz $\beta \in S^{n-1}$. Wówczas z faktu 1.3.1 mamy, $\partial_\beta^2(\ln b)(x) \geq \mu$ i z (1.3),

$$\frac{b(x)\partial_\beta^2 b(x) - \big(\partial_\beta b(x)\big)^2}{\big(b(x)\big)^2} \geq \mu. \tag{1.4}$$

Stąd dostajemy (b).

Zakładając, że zachodzi warunek (b), widzimy że zachodzi również (1.4), a wobec (1.3) dostajemy, że $\partial_\beta^2(\ln b)(x) \geq \mu$ dla $x \in \mathbb{R}^n$ oraz $\beta \in S^{n-1}$. To wraz z faktem 1.3.1 daje, że $\ln b$ jest funkcja μ-silnie wypukłą w $\mathbb{R}^n$. To daje (a). $\square$

Z faktu 1.4.2 otrzymujemy

Wniosek 1.4.3. *Niech $b : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją logarytmicznie μ-silnie wypukłą klasy $\mathscr{C}^2$. Wówczas dla dowolnych $x \in \mathbb{R}^n$ i $\beta \in S^{n-1}$, mamy*

$$\mu b(x) \leq \partial_\beta^2 b(x) \quad \textit{oraz} \quad \sqrt{\mu}|\partial_\beta b(x)| \leq \sqrt{\big(\partial_\beta^2 b(x)\big)^2 - \mu^2 b^2(x)} < \partial_\beta^2 b(x).$$

Dowód. Z faktu 1.4.2 mamy,

$$b(x)\partial_\beta^2 b(x) \geq \mu b^2(x) + (\partial_\beta b(x))^2 \quad \text{dla każdych } x \in \mathbb{R}^n,\ \beta \in S^{n-1}. \tag{1.5}$$

Ponieważ $(\partial_\beta b(x))^2 \geq 0$, to $b(x)\partial_\beta^2 b(x) \geq \mu b^2(x)$, a ponieważ $b(x) > 0$ dla $x \in \mathbb{R}^n$, to dostajemy nierówność $\mu b(x) \leq \partial_\beta^2 b(x)$. To daje nierówność po lewej stronie tezy. W szczególności mamy $\left(\partial_\beta^2 b(x)\right)^2 - \mu^2 b^2(x) \geq 0$, a wiec $\sqrt{\left(\partial_\beta^2 b(x)\right)^2 - \mu^2 b^2(x)}$ istnieje. Stosując (1.5) i powyżej udowodnioną nierówność, mamy

$$\begin{aligned}(\partial_\beta b(x))^2 &\leq \left(\partial_\beta^2 b(x) - \mu b(x)\right) b(x) \leq \left(\partial_\beta^2 b(x) - \mu b(x)\right) \frac{1}{\mu} \partial_\beta^2 b(x) \\ &\leq \left(\left(\partial_\beta^2 b(x)\right)^2 - \mu b(x) \partial_\beta^2 b(x)\right) \frac{1}{\mu} \leq \left(\left(\partial_\beta^2 b(x)\right)^2 - \mu^2 b^2(x)\right) \frac{1}{\mu}\end{aligned}$$

Stąd dostajemy nierówność po prawej stronie tezy. $\square$

Weźmy funkcję $b : \mathbb{R}^n \to \mathbb{R}$ klasy $\mathscr{C}^k$ logarytmicznie μ-silnie wypukła, $k \geq 2$, $\mu > 0$. Załóżmy, że

$$0 = \operatorname{argmin}_{\mathbb{R}^n} b \quad \text{i} \quad b(0) = 1, \tag{1.6}$$

gdzie $\operatorname{argmin}_X g$ oznacza jedyny punkt $x \in X$, w którym funkcja $g : X \to \mathbb{R}$ ma globalne minimum. Z definicji otrzymujemy

Fakt 1.4.4. *Zgodnie z powyższymi założeniami, dla dowolnego $x \in \mathbb{R}^n$,*

$$\exp\left(\frac{\mu}{2}|x|^2\right) \leq b(x) \leq \exp\left(\langle x, \frac{1}{b(x)} \nabla b(x)\rangle - \frac{\mu}{2}|x|^2\right), \tag{1.7}$$

a więc $\mu|x|^2 b(x) \leq \langle x, \nabla b(x)\rangle$, i w konsekwencji

$$\mu|x| \leq \mu|x| \exp\left(\frac{\mu}{2}|x|^2\right) \leq |\nabla b(x)| \quad \textit{dla } x \in \mathbb{R}^n. \tag{1.8}$$

W szczególności,

$$\partial_\beta^2 b(x) \geq \mu \exp\left(\frac{\mu}{2}|x|^2\right) \quad \textit{dla dowolnego } \beta \in S^{n-1}. \tag{1.9}$$

Dowód. Ponieważ funkcja b jest logarytmicznie μ-silnie wypukła klasy $\mathscr{C}^2$, to $\ln b$ jest funkcją μ-silnie wypukłą klasy $\mathscr{C}^2$, a wiec zgodnie z definicją, dla dowolnych $x, y \in \mathbb{R}^n$ mamy

$$\ln b(x) \geq \ln b(y) + \langle x - y, \nabla(\ln b)(y)\rangle + \frac{\mu}{2}|x - y|^2 \tag{1.10}$$

Ponieważ $0 = \operatorname{argmin}_{\mathbb{R}^n} b$, to $0 = \operatorname{argmin}_{\mathbb{R}^n} \ln b$ i z warunku koniecznego istnienia ekstremum funkcji mamy $\nabla(\ln b)(0) = 0$. Ponieważ $b(0) = 1$, to dla $y = 0$ i dowolnego $x \in \mathbb{R}^n$, mamy $\ln b(x) \geq \frac{\mu}{2}|x|^2$, a więc otrzymujemy nierówność po lewej stronie (1.7). Podobnie z (1.10) dla $x = 0$ i $y \in \mathbb{R}^n$, mamy

$$\ln b(y) \leq \langle y, \nabla(\ln b)(y)\rangle - \frac{\mu}{2}|y|^2$$

Stąd otrzymujemy nierówność po prawej stronie (1.7).

Nierówność po lewej stronie (1.8) wynika z tego, że $\exp\left(\frac{\mu}{2}|x|^2\right) \geq 1$. Udowodnimy nierówność po prawej stronie (1.8). Z (1.7) i monotoniczności funkcji exp dla $x \in \mathbb{R}^n$ dostajemy

$$\mu|x|^2 b(x) \leq \langle x, \nabla b(x)\rangle - \frac{\mu}{2}|x| \leq \langle x, \nabla b(x)\rangle.$$

Ponieważ, z nierówności Schwarza, $\langle x, \nabla b(x)\rangle \leq |x||\nabla b(x)|$, to $\mu|x|^2 b(x) \leq |x||\nabla b(x)|$, a więc $\mu|x|b(x) \leq |\nabla b(x)|$. Stąd i ponownie z (1.7) mamy

$$\mu|x| \exp\left(\frac{\mu}{2}|x|^2\right) \leq |\nabla b(x)|,$$

czyli udowodnilismy nierówność po prawej stronie (1.8).

Nierówność (1.9) wynika z nierówności po lewej stronie (1.7) i nierównosci po lewej stronie tezy wniosku 1.4.3. $\square$

Uwaga 1.4.5. W [9] uzyskano uwypuklenie funkcji dodatniej f klasy $\mathscr{C}^2$ na zbiorze zwartym i wypukłym $X \subset \mathbb{R}^n$ poprzez pomnożenie f przez $(1+|x|^2)^N$ dla pewnej N, a w [8] – przez pomnożenie f przez $\exp(N|x|^2)$. W świetle faktu 1.4.4 widzimy podejście to było naturalne.

Z faktu 1.2.3 otrzymujemy

Wniosek 1.4.6. *Jeśli $f : \mathbb{R}^n \to \mathbb{R}$ jest funkcją logarytmicznie ściśle wypukłą, to odwzorowanie*

$$\frac{1}{f}\nabla f : X \ni x \mapsto \frac{1}{f(x)}\nabla f(x) \in \mathbb{R}^n$$

jest różnowartościowe.

Dowód. Istotnie, z definicji funkcja $\ln f$ jest ściśle wypukła oraz $\nabla(\ln f) = \frac{1}{f}\nabla f$. Zatem z faktu 1.2.3 dostajemy tezę. To kónczy dowód . $\square$

Bez założenia logarytmiczniej silnej wypukłości funkcji f, powyższy wniosek nie zachodzi. Pokazuje to następujący przykład.

Przykład 1.4.7. Niech $f(x) = 1 + x^2$. Wtedy

$$\frac{f'}{f}(x) = \frac{2x}{1+x^2}$$

i oczywiście ta funkcja nie jest różnowartościowa. Ponadto funkcja f jest silnie wypukła.

Lemat 1.4.8. *Niech $b : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją μ-silnie wypukłą klasy $\mathscr{C}^2$, niech $x_0 = \operatorname{argmin}_{\mathbb{R}^n} b$ i niech $X \subset \mathbb{R}^n$ będzie zbiorem wypukłym i zwartym. Jeśli $b(x) > 0$ dla $x \in X$ i x_0 jest punktem wewnętrznym zbioru X, to istnieje $\varepsilon > 0$ takie, że*

(i) *funkcja b jest logarytmicznie silnie wypukła w zbiorze $X_{x_0,\varepsilon} = \{x \in X : |x - x_0| \leq \varepsilon\}$,*

(ii) *funkcja $X_{x_0,\varepsilon} \ni x \mapsto \frac{1}{b(x)}\nabla b(x) \in \mathbb{R}^n$ jest różnowartościowa,*

(iii) *istnieje $\delta > 0$ takie, że dla każdego $x \in X$ takiego, że $\frac{|\nabla b(x)|}{b(x)} < \delta$, zachodzi $|x - x_0| < \varepsilon$.*

Dowód. Ponieważ $b(x) > 0$ dla $x \in X$ i b jest funkcja μ-silnie wypukła, to dla dowolnych $x \in X$ i $\beta \in S^{n-1}$, mamy

$$\partial_\beta(\ln b)(x) = \frac{\partial_\beta b(x)}{b(x)}, \qquad \partial_\beta^2(\ln b)(x) = \frac{b(x)\partial_\beta^2 b(x) - \big(\partial_\beta b(x)\big)^2}{\big(b(x)\big)^2}.$$

Z faktu 1.2.1 mamy

$$\partial_\beta^2(\ln b)(x) = \frac{\partial_\beta^2 b(x)}{b(x)} - \left(\frac{\partial_\beta b(x)}{b(x)}\right)^2 \geq \frac{\mu}{b(x)} - \left(\frac{\partial_\beta b(x)}{b(x)}\right)^2.$$

Ponieważ funkcja b jest klasy $\mathscr{C}^2$ spełniającą (1.6), więcz $\partial_\beta b(x_0) = 0$ i $b(0) = 1$. Zatem istnieje $\varepsilon > 0$ i $\eta_1 > 0$ takie, że $\partial_\beta b(x) \geq \eta_1$ dla $x \in X$ takich, że $|x - x_0| \leq \varepsilon$. To daje (i).

Teza (ii) bezpośrednio wynika z (i) i wniosku 1.4.6. Biorąc

$$\delta = \min\left\{\varepsilon,\ \inf\left\{\frac{|\nabla b(x)|}{b(x)} : x \in X,\ |x - x_0| \geq \varepsilon\right\}\right\},$$

gdzie $\inf\emptyset = +\infty$, zauważmy, że $\delta > 0$ i dostajemy tezę (iii). $\square$

1.5 Uwagi o funkcjach logarytmicznie silnie wypukłych ℓ-tego rzędu

Dla dowolnego $\ell = 0, 1, \dots$ określamy funkcję $\exp^{[\ell]}$ przez indukcję: $\exp^{[\ell]}(x) = x$ dla $\ell = 0$ i $\exp^{[\ell+1]} = \exp \circ \exp^{[\ell]}$ dla $\ell \geq 1$. Oczywiście funkcja $\exp^{[\ell]} : \mathbb{R} \to A_\ell$ jest ściśle rosnącym dyfeomorfizmem klasy $\mathscr{C}^\infty$, gdzie $A_\ell = \mathbb{R}$ dla $\ell = 0$ i $A_\ell = (\exp^{[\ell-1]}(0), +\infty)$ dla $\ell > 0$. Funkcję odwrotną do funkcji $\exp^{[\ell]}$ oznaczamy przez $\ln^{[\ell]}$. Dokładniej, $\ln^{[0]}(x) = x$, $\ln^{[1]} = \ln$ i $\ln^{[\ell+1]} = \ln \circ \ln^{[\ell]}$ dla $\ell \geq 2$. Oczywiście funkcja $\ln^{[\ell]}$ jest określona w zbiorze A_ℓ

Niech $g : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją. Mówimy, że funkcja g jest *logarytmicznie ℓ-tego rzędu μ-silnie wypukła* lub że jest *funkcją ℓ-logarytmicznie μ- silnie wypukłą*, jeśli $g(x) > \exp^{[\ell-1]}(0)$ dla $x \in \mathbb{R}^n$ i funkcja $\ln^{[\ell]} g$ jest μ-silnie wypukła.

Weźmy dowolną funkcję ℓ-logarytmicznie μ-silnie wypukłą $b : \mathbb{R}^n \to \mathbb{R}$ klasy $\mathscr{C}^k$, $k \geq 2$. Wtedy $b(x) > \exp^{[\ell-1]}(0)$, więc $\ln^{[j]} b(x) > \exp^{[\ell-j-1]}(0)$ dla $x \in \mathbb{R}^n$, $0 \leq j \leq \ell - 1$.

Oczywiście funkcje 1-logarytmicznie μ-silnie wypukłe są funkcjami logarytmicznie μ-silnie wypukłymi. Zatem z Faktu 1.4.2 otrzymujemy

Fakt 1.5.1. *Weźmy dowolne $\ell \geq 1$. Niech $b : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^2$ taką, że $b(x) > \exp^{[\ell-1]}(0)$ dla $x \in \mathbb{R}^n$ oraz niech $\mu > 0$. Wówczas następujące warunki są równoważne:*

(a) *Funkcja b jest ℓ-logarytmicznie μ-silnie wypukła.*

(b) *Dla dowolnego $x \in \mathbb{R}^n$ i dowolnego $\beta \in S^{n-1}$,*

$$\ln^{[\ell-1]} b(x) \partial_\beta^2 \ln^{[\ell-1]} b(x) \geq \mu \left(\ln^{[\ell-1]} b(x)\right)^2 + \left(\partial_\beta \ln^{[\ell-1]} b(x)\right)^2.$$

Wniosek 1.5.2. *Weźmy dowolne $\ell > 1$. Jeśli $b : \mathbb{R}^n \to \mathbb{R}$ jest funkcją klasy $\mathscr{C}^2$, która jest ℓ-logarytmicznie μ-silnie wypukła, to dla dowolnego $x \in \mathbb{R}^n$,*

$$\partial_\beta^2 b(x) \prod_{j=0}^{\ell-1} \ln^{[j]} b(x) \geq \mu \prod_{j=0}^{\ell-1} \left(\ln^{[j]} b(x)\right)^2 + (\partial_\beta b(x))^2 + (\partial_\beta b(x))^2 \sum_{k=1}^{\ell-1} \prod_{j=\ell-k}^{\ell-1} \ln^{[j]} b(x).$$

W szczególności,

$$\partial_\beta^2 b(x) \geq \mu \prod_{j=0}^{\ell-1} \ln^{[j]} b(x) \geq \mu b(x) \prod_{j=0}^{\ell-2} \exp^{[\ell-j-1]}(0) \ \textit{oraz}\ \partial_\beta^2 b(x) \geq \sqrt{\mu} b(x) |\partial_\beta b(x)|.$$

Jeśli dodatkowo $g(x) \geq \exp^{[\ell-1]}(1)$ dla $x \in \mathbb{R}^n$, to $\partial_\beta^2 b(x) \geq \mu b(x)$.

Dowód. Dla $\ell = 1$ teza redukuje się do Faktu 1.4.2. Dla $\ell > 1$, Z Faktu 1.5.1 mamy

$$\left(\ln^{[\ell-1]} b(x)\right) \left\{ \frac{\partial_\beta^2 \left(\ln^{[\ell-2]} b(x)\right) \left[\ln^{[\ell-2]} b(x)\right] - \left(\partial_\beta \left[\ln^{[\ell-2]} b(x)\right]\right)^2}{\left[\ln^{[\ell-2]} b(x)\right]^2} \right\}$$
$$\geq \mu \left[\ln^{[\ell-1]} b(x)\right]^2 + \left(\partial_\beta \left[\ln^{[\ell-1]} b(x)\right]\right)^2,$$

i w konsekwencji

$$\begin{aligned} &\left[\ln^{[\ell-1]} b(x)\right] \left[\ln^{[\ell-2]} b(x)\right] \partial_\beta^2 \left[\ln^{[\ell-2]} b(x)\right] \\ &\quad \geq \mu \left[\ln^{[\ell-1]} b(x)\right]^2 \left[\ln^{[\ell-2]} b(x)\right]^2 + \left[\ln^{[\ell-2]} b(x)\right]^2 \left(\partial_\beta \left[\ln^{[\ell-1]} b(x)\right]\right)^2 \\ &\qquad + \left[\ln^{[\ell-1]} b(x)\right] \left(\partial_\beta \left[\ln^{[\ell-2]} b(x)\right]\right)^2. \end{aligned}$$

Postępując dalej indukcyjnie, otrzymujemy

$$\begin{aligned} \partial_\beta^2 b(x) \prod_{j=0}^{\ell-1} \left[\ln^{[j]} b(x)\right] &\geq \mu \prod_{j=0}^{\ell-1} \left[\ln^{[j]} b(x)\right]^2 + \left(\partial_\beta \left[\ln^{[\ell-1]} b(x)\right]\right)^2 \prod_{j=0}^{\ell-2} \left[\ln^{[j]} b(x)\right]^2 \\ &\quad + \sum_{k=2}^{\ell} \left(\partial_\beta \left[\ln^{[\ell-k]} b(x)\right]\right)^2 \prod_{j=\ell-k+1}^{\ell-1} \left[\ln^{[j]} b(x)\right] \prod_{j=0}^{\ell-k-1} \left[\ln^{[j]} b(x)\right]^2 \\ &= \mu \prod_{j=0}^{\ell-1} \left[\ln^{[j]} b(x)\right]^2 + (\partial_\beta b(x))^2 + (\partial_\beta b(x))^2 \sum_{k=2}^{\ell} \prod_{j=\ell-k+1}^{\ell-1} \left[\ln^{[j]} b(x)\right], \end{aligned}$$

co daje tezę. □

Weźmy funkcję $b : \mathbb{R}^n \to \mathbb{R}$ ℓ-logarytmicznie μ-silnie wypukła, $\mu > 0$, klasy $\mathscr{C}^2$. Załóżmy, że $0 = \operatorname{argmin}_{\mathbb{R}^n} b$ i $b(0) = \exp^{[\ell-1]}(1)$. Z Faktu 1.4.4 otrzymujemy

Fakt 1.5.3. *Zgodnie z powyższymi założeniami, dla dowolnego $x \in \mathbb{R}^n$,*

$$\exp^{[j]}\left(\frac{\mu}{2}|x|^2\right) \leq \ln^{[\ell-j]} b(x), \quad j = 1, \ldots, \ell,$$

wówczas, $b(x) \geq \mu \exp^{[\ell]}\left(\frac{\mu}{2}|x|^2\right)$ oraz

$$\partial_\beta^2 b(x) \geq \mu \prod_{j=1}^{\ell} \exp^{[j]}\left(\frac{\mu}{2}|x|^2\right) \geq \mu \exp^{[\ell]}\left(\frac{\mu}{2}|x|^2\right).$$

1.6 Wielomiany

Niech $f \in \mathbb{R}[x]$ będzie wielomianem rzeczywistym zmiennych $x = (x_1, \ldots, x_n)$, postaci

$$f = \sum_{|\nu| \leq d} a_\nu x^\nu, \tag{1.11}$$

gdzie $a_\nu \in \mathbb{R}$, $x^\nu = x_1^{\nu_1} \cdots x_n^{\nu_n}$ oraz $|\nu| = \nu_1 + \cdots + \nu_n$ dla $\nu = (\nu_1, \cdots, \nu_n) \in \mathbb{N}^n$ (przyjmijmy, że $0 \in \mathbb{N}$). Załóżmy, że $d = \deg f$, tj., $d = \max\{|\nu| : a_\nu \neq 0\}$. Wtedy

$$f = f_0 + \cdots + f_d, \tag{1.12}$$

gdzie f_j *jest wielomianem jednorodnym stopnia* j lub wielomianem zerowym, t.j.

$$f_j = \sum_{|\nu| = j} a_\nu x^\nu, \quad 0 \leq j \leq d.$$

Wielomian f_d nazywamy *formą wiodącą* wielomianu f. Połóżmy

$$f_{d*} = \min_{|x|=1} f_d(x).$$

1.6.1 Norma wielomianu

Połóżmy

$$\|f\| := \sum_{|\nu| \leq d} |a_\nu|.$$

Oczywiście funkcja $\| \cdot \|$ określona wzorem $\mathbb{R}[x] \ni g \mapsto \|g\| \in \mathbb{R}$ jest normą w pierścieniu $\mathbb{R}[x]$. Ponadto,

$$||f|| = ||f_0|| + \cdots + ||f_d||$$

oraz

$$|f_j(x)|| \leq \|f_j\| |x|^j \quad \text{dla } x \in \mathbb{R}^n \text{ oraz } j = 0, \ldots, d.$$

Lemat 1.6.1. *Dla dowolnego $x \in \mathbb{R}^n$ oraz dowolnego $\beta \in S^{n-1}$,*

$$
\begin{aligned}
|f(x)| &\le \sum_{j=0}^{d} ||f_j|| \cdot |x|^j, \\
|\partial_\beta f(x)| &\le \sum_{j=0}^{d-1} (j+1)||f_{j+1}|| \cdot |x|^j, \\
|\partial_\beta^2 f(x)| &\le \sum_{j=0}^{d-2} (j+1)(j+2)||f_{j+2}|| \cdot |x|^j
\end{aligned}
$$

Jeśli dodatkowo $f_{d*} > 0$ to $f_d(x) \ge f_{d*}|x|^d$ i w konsekwencji,

$$
f(x) \ge f_{d*}|x|^d - \sum_{j=0}^{d-1} ||f_j|| \cdot |x|^j
$$

Dowód. Pierwsza nierówność tezy jest oczywista. Udowodnimy nierówność drugą i trzecią. Mamy

$$
\partial_\beta f(x) = \sum_{j=1}^{d} \sum_{|\nu|=j} a_\nu \partial_\beta x^\nu \quad \text{oraz} \quad \partial_\beta^2 f(x) = \sum_{j=2}^{d} \sum_{|\nu|=j} a_\nu \partial_\beta^2 x^\nu.
$$

Niech $\beta = (\beta_1, \ldots, \beta_n)$. Weźmy dowolne $\nu = (\nu_1, \cdots, \nu_n) \in \mathbb{N}^n$, $|\nu| = j$. Wówczas

$$
|\partial_\beta x^\nu| \le \sum_{k=1}^{n} \nu_k |x_1^{\nu_1} \cdots x_k^{\nu_k - 1} \cdots x_n^{\nu_n}| \le j|x|^{j-1}
$$

i w konsekwencji,

$$
|\partial_\beta^2 x^\nu| \le \sum_{k=1}^{n} \nu_k |\partial_\beta x_1^{\nu_1} \cdots x_k^{\nu_k - 1} \cdots x_n^{\nu_n}| \le j(j-1)|x|^{j-2}.
$$

To daje drugą i trzecią nierówność tezy.

Jeśli $f_{d*} > 0$, to oczywiście $f_d(x) \ge f_{d*}|x|^d$. Zatem z (1.12) i własności normy,

$$
f(x) \ge f_d(x) - \sum_{j=0}^{d-1} |f_j(x)| \ge f_{d*}|x|^d - \sum_{j=0}^{d-1} ||f_j|| \cdot |x|^j,
$$

co daje drugą część tezy. $\square$

Wniosek 1.6.2. *Weźmy dowolne* $\beta \in S^{n-1}$. *Wówczas dla dowolnego* $x \in \mathbb{R}^n$ *takiego, że* $|x| \geq 1$, *zachodzą nierówności*

$$|\partial_\beta f(x)| \leq d||f|| \cdot |x|^{d-1} \tag{1.13}$$

oraz

$$|\partial_\beta^2 f(x)| \leq d(d-1)||f|| \cdot |x|^{d-2}. \tag{1.14}$$

Dowód. Istotnie, z lematu 1.6.1

$$|\partial_\beta f(x)| \leq d|x|^{d-1}(||f_1|| + \cdots + ||f_d||) \leq d|x|^{d-1} \cdot ||f||.$$

oraz

$$|\partial_\beta^2 f(x)| \leq d(d-1)|x|^{d-2} \cdot ||f||,$$

co daje (1.13) i (1.14) i kończy dowód. □

Z lematu 1.6.1 mamy

Wniosek 1.6.3. *Jeśli* $\nabla f(0) = 0$, *to*

$$|\nabla f(x)| \leq d\sqrt{n}\|f - f_0\| \cdot |x| \quad \textit{dla } |x| \leq 1.$$

1.6.2 Pewne oszacowania

Niech $f \in \mathbb{R}[x]$ będzie wielomianem postaci (1.11), $d = \deg f$. Załóżmy, że $f_{d*} > 0$. Wtedy $\|f\| \geq \|f_d\| \geq f_{d*}$. Niech

$$\mathbb{K}(f) := \frac{2\|f\|}{f_{d*}}$$

oraz

$$m(f) := f_{d*} - \sum_{j=0}^{d-1} \mathbb{K}(f)^{j-d}||f_j||.$$

Oczywiście, $\mathbb{K}(f) \geq 2$.

Będziemy potrzebować następującego lematu (patrz [8][Lemma 3.4]).

Lemat 1.6.4. *Jeśli* $d = \deg f > 0$ *oraz* $f_{d*} > 0$, *wtedy* $m(f) > 0$ *i*

$$f(x) \geq m(f)|x|^d \quad \textit{dla dowolnego } x \in \mathbb{R}^n \textit{ takiego, że } |x| \geq \mathbb{K}(f).$$

Z wniosku 1.6.4 i lematów 1.6.1 1.6.2 i otrzymujemy natychmiast

Wniosek 1.6.5. *Niech f będzie wielomianem postaci* (1.11) *takim, że $f_{d*} > 0$. Weźmy dowolne $\beta \in S^{n-1}$. Wówczas dla dowolnego $x \in \mathbb{R}^n$, $|x| \geq \mathbb{K}(f)$ mamy*

$$\frac{|\partial_\beta f(x)|}{f(x)} \leq \frac{d||f||}{m(f)} \cdot |x|^{-1} \leq \frac{d||f||}{2m(f)} \tag{1.15}$$

oraz

$$\frac{|\partial_\beta^2 f(x)|}{f(x)} \leq \frac{d(d-1)||f||}{m(f)} \cdot |x|^{-2} \leq \frac{d(d-1)||f||}{4m(f)}. \tag{1.16}$$

Dla $R > 0$, kładziemy

$$D_n^0(f,R) := \sum_{j=0}^{d} ||f_j||R^j, \qquad D_n^1(f,R) := \sum_{j=0}^{d} j||f_j||R^{j-1},$$

$$D_n^2(f,R) := \sum_{j=0}^{d} j(j-1)||f_j||R^{j-2}$$

oraz

$$D_n(f,R) := \max\{D_n^1(f,R),\ D_n^2(f,R)\}.$$

Wówczas z lematu 1.6.1 mamy

Wniosek 1.6.6. *Dla każduch $\beta \in S^{n-1}$ oraz $x \in \mathbb{R}^n$ takiego, że $|x| \leq R$, mamy*

$$\begin{gathered} |f(x)| \leq D_n^0(f,R), \qquad |\partial_\beta f(x)| \leq D_n^1(f,R) \leq D_n(f,R), \\ |\partial_\beta^2 f(x)| \leq D_n^2(f,R) \leq D_n(f,R). \end{gathered} \tag{1.17}$$

1.6.3 Oszacowanie zer wielomianu

Niech $f \in \mathbb{R}[x]$ będzie wielomianem postaci (1.11). Załóżmy, że $f_{d*} > 0$ i połóżmy

$$K_f(r) = 2\max\left\{\left(\frac{||f_0||+r}{f_{d*}}\right)^{1/d}, \max_{1\leq j\leq d-1}\left|\frac{||f_{d-j}||}{f_{d*}}\right|^{1/j}\right\} \quad \text{dla } r \geq 0.$$

Kładziemy również $K(f) := K_f(0)$.

Fakt 1.6.7. *Dla dowolnego* $r \geq 0$*,*

$$X_{f\leq r} := \{x \in \mathbb{R}^n : f(x) \leq r\} \subset \{x \in \mathbb{R}^n : |x| \leq K_f(r)\}.$$

Dowód. Oznaczmy $r = r(x) = |x|$ oraz $\theta = \theta(x) = \frac{1}{|x|}x$ dla $x \neq 0$. Wtedy we współrzędnych biegunowych można zapisać $x = r\theta$, $r > 0$, $\theta \in S^{n-1}$ oraz

$$f(x) = f(r\theta) = \sum_{j=0}^{d} f_j(\theta) r^j, \quad x \neq 0. \tag{1.18}$$

Oczywiście,

$$|f_j(\theta)| \leq ||f_j|| \quad \text{dla} \quad \theta \in S^{n-1}.$$

Ponieważ liczba

$$2 \max_{1\leq j\leq d} \left| \frac{a_{d-j}}{a_0} \right|^{1/j}$$

szacuje z góry moduł dowolnego zera wielomianu $a_0 z^d + a_1 z^{d-1} + \cdots + a_d$, gdzie $a_0 \neq 0$, więc z (1.19) widzimy, że wielomian $f - r$ nie ma zer $x \in \mathbb{R}^n$ takich, że $|x| > K_f(r)$. Ponieważ wielomian f ma dodatnie wartości dla $x \in \mathbb{R}^n$ takich, że $|x|$ dąży do nieskończonośc, więc otrzymujemy tezę. $\square$

1.6.4 Gradient wielomianu we współrzędnych biegunowych

Niech $f \in \mathbb{R}[x]$ będzie wielomianem postaci (1.11). Wówczas możemy zapisać

$$f(x) = \sum_{j=0}^{d} f_j\left(\frac{1}{|x|}x\right)|x|^j, \quad x \in \mathbb{R}^n \setminus \{0\}.$$

Przyjmijemy oznaczenia

$$r = r(x) = |x| \quad \text{i} \quad \theta = \theta(x) = \frac{1}{|x|}x \quad \text{dla } x \neq 0.$$

Wtedy $x = r\theta$, $r > 0$, $\theta \in S^{n-1}$ i możemy zapisać wielomian f we współrzędnych biegunowych

$$f(x) = f(r\theta) = \sum_{j=0}^{d} f_j(\theta) r^j, \quad x \neq 0. \tag{1.19}$$

Wówczas

$$\nabla f(x)=\partial_r f(r\theta)\theta+\nabla' f(r\theta), \tag{1.20}$$

dla

$$\partial_r f(r\theta)=\frac{\langle\nabla f(r\theta),r\theta\rangle}{r}=\partial_\theta f(r\theta)=\frac{\partial f(r\theta)}{\partial r}=\sum_{j=1}^{d} jf_j(\theta)r^{j-1} \tag{1.21}$$

oraz

$$\nabla' f(r\theta)=\nabla f(r\theta)-\partial_r f(r\theta)\theta.$$

Oczywiście,

$$\langle\nabla' f(r\theta),\theta\rangle=0\quad\text{dla } x=r\theta\neq 0$$

oraz

$$\nabla' f(x)=\nabla f(x)-\frac{\langle\nabla f(x),x\rangle}{|x|^2}x\quad\text{dla } x\neq 0.$$

Wektor $\partial_r f(r\theta)\theta$ nazywamy się *częścią promieniową* (lub radialną) gradiantu $\nabla f(x)$ a wektor $\nabla' f(r\theta)$ – *częścią sferyczną* gradiantu $\nabla f(x)$.

Z definicji gradiandu $\nabla' f$ otrzymujemy następującą uwagę.

Uwaga 1.6.8. Niech $e_1,\ldots,e_n$ będzie bazą standardową przestrzeni liniowej $\mathbb{R}^n$, t.j., $e_j=(0,\ldots,0,1,0,\ldots,0)$, gdzie 1 jest na $j-$tym miejscu. Weźmy dowolne $x\in\mathbb{R}^n$, $x\neq 0$. Biorąc

$$\alpha_j=\frac{\langle e_j,x\rangle}{|x|^2}x,\qquad v_j=e_j-\alpha_j\qquad\text{dla } j=1,\ldots,n,$$

mamy $|v_j|=1-\frac{x_j^2}{|x|^2}$,

$$v_j=\left(-\frac{x_1x_j}{|x|^2},\ldots,-\frac{x_{j-1}x_j}{|x|^2},1-\frac{x_j^2}{|x|^2},-\frac{x_{j+1}x_j}{|x|^2},\ldots,-\frac{x_nx_j}{|x|^2}\right),\quad j=1,\ldots,n$$

oraz

$$\nabla' f(x)=\sum_{j=1}^{n}\langle\nabla f(x),v_j\rangle v_j.$$

1.7 Wielomianowy wzrost funkcji

Niech $X \subset \mathbb{R}^n$, niech $f : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^k$, $k \geq 2$. Mówiemy, że funkcją f ma *wielomianowy wzrost rzędu* ℓ, jeśli istnieją $D \in \mathbb{R}$ i $\alpha > 0$ takie, że

$$|\partial_\beta^j f(x)| \leq D(1+|x|)^\alpha \quad \text{dla dowolnych } x \in X, \text{ i } \beta \in S^{n-1}, \text{ i } 0 \leq j \leq \ell,$$

gdzie $\partial_\beta^j f(x)$ jest pochodną j-tego rzędu funkcji f w kierunku wektora β w punkcie x, $\ell \leq k$.

Oczywiście, każda funkcja wielomianowa ma wzrost wielomianu ℓ-tego rzędu dla dowolnego ℓ. Przykładami takich funkcji są również: funkcje wymierne zdefiniowane na zbiorze $\mathbb{R}^n$, $\sin(p(x))$, $\cos(p(x))$, gdzie p jestnwielomian, $\ln(1+p(x))$ jeśli p przyjmuje tylko wartości nieujemne. Co więcej, każda funkcja semialgebraiczna klasy $\mathscr{C}^\ell$ zdefiniowana na zbiorze $\mathbb{R}^n$ ma wzrost wielomianu ℓ-tego rzędu dla dowolnego ℓ. Funkcja exp nie ma tej własności.

Rozdział 2

Uwypuklanie funkcji

Jedną z metod stosowanych w algorytmach minimalizacji funkcji jest sprowadzenie problemu do funkcji wypukłych, poprzez deformację danej funkcji do funkcji wypukłej. Powodem jest to, że stosunkowo łatwo jest obliczyć minimum funkcji wypukłej. Klasycznym podejściem do uwypuklenia funkcji f na zbiorach ograniczonych $X \subset \mathbb{R}^n$ jest dodawanie funkcji silnie wypukłej b do funkcji f (patrz np [19], [11] i [20] dla funcji kwadratowej $b(x) = \gamma|x|^2$, $\gamma > 0$). W algorytmach minimalizacji funkcji f ważne jest uzyskanie silnej wypukłości funkcji $x \mapsto b(x - \xi) + f(x)$ dla dowolnego $\xi \in X$ i pewnej funkcji b silnie wypukłej (por. [9]). W tym rozdziale pokażemy, że w przypadku funkcji dodatniej na zbiorze zwartym i wypukłym, można ją zdefommować do funkcji silnie wypukłej przez dodanie do niej lub pomnożenie jej przez funkcję silnie wypukłą i dodatnią.

2.1 Uwagi o uwypukleniu funkcji

Fakt 2.1.1. *Jeśli $b : \mathbb{R}^n \to \mathbb{R}$ jest funkcją klasy $\mathscr{C}^2$ taką, że dla dowolnego zbioru zwartego i wypukłego $X \subset \mathbb{R}^n$ oraz dowolnego wielomianu $f : \mathbb{R}^n \to \mathbb{R}$ przyjmującego tylko wartości dodatnie na zbiorze X, istnieje $N_0 \in \mathbb{N}$ takie, że dla każdego $N \geq N_0$ funkcja $\varphi(x) = b^N(x)f(x)$ jest silnie wypukła na X, to funkcja b przyjmuje tylko wartości dodatnie $\mathbb{R}^n$.*

Dowód. Z założenia dla $f = 1$ i dowolnego zbioru zwartego i wypukłego $X \subset \mathbb{R}^n$, istnieje N_0 takie, że dla każdego $N \geq N_0$, funkcja $\varphi(x) = b^N(x)$ jest silnie wypukła na zbioru X. Ponieważ funkcja b jest klasy $\mathscr{C}^2$, to z Faktu 1.3.2 dla dowolnego wektora $v \in S^{n-1}$, mamy

$$\begin{aligned}\partial_v^2 b^N(x) &= N(N-1)b^{N-2}(x)(\partial_v b(x))^2 + Nb^{N-1}(x)\partial_v^2 b(x)\\ &= Nb^{N-2}(x)\left[(N-1)(\partial_v b(x))^2 + b(x)\partial_v^2 b(x)\right] > 0 \quad \text{dla } x \in X.\end{aligned}$$

A więc, $b(x) \neq 0$ dla $x \in \mathbb{R}^n$. W związku z tym, z uwagi na ciągłość funkcji $x \mapsto b(x)$, $(x, v) \mapsto \partial_v b(x)$, $(x, v) \mapsto \partial_v^2 b(x)$, własność Darboux daje tezę i kończy dowód.

□

Przykład 2.1.2. Zgodnie z założeniami faktu 2.1.1, nie możemy wymagać, by funkcja b była wypukła. Na przykład dla funkcji $b(x) = \sqrt[4]{1 + |x|^2}$, $x \in \mathbb{R}^n$, widzimy że spełnia ona tezę faktu 2.1.1 (patrz [9, Theorem 5.1]) ale funkcja b nie jest nawet wypukła.

W dalszym ciągu pracy będziemy mówić, że funkcja f jest dodatnia na zbiorze X, gdy w tym zbiorze przyjmuje tylko wartości dodatnie.

Fakt 2.1.3. *Jeśli $b : \mathbb{R}^n \to \mathbb{R}$ jest funkcją klasy $\mathscr{C}^2$ taką, że dla dowolnego zbioru zwartego i wypukłego $X \subset \mathbb{R}^n$ oraz dowolnego wielomianu $f : \mathbb{R}^n \to \mathbb{R}$ dodatniego na zbiorze X, istnieje $N_0 \in \mathbb{N}$ takie, że dla każdego $N \geq N_0$, funkcja $\varphi(x) = b^N(x)f(x)$ jest logarytmiczne silnie wypukła na X, to funkcja b jest również logarytmiczne silnie wypukła na zbiorze X.*

Dowód. Ponieważ funkcja logarytmiczne silnie wypukła jest również silnie wypukła, więc zgodnie z faktem 2.1.1, funkcja b jest dodatnia na zbiorze $\mathbb{R}^n$. Z załżenia dla $f = 1$ i dowolnego zbioru zwartego i wypukłego $X \subset \mathbb{R}^n$ istnieje N_0 takie, że dla każdego $N \geq N_0$ funkcja $\varphi(x) = b^N(x)$ jest logarytmiczne silnie wypukłe na zbiorze X. Zatem funkcja $\ln \varphi(x) = N \ln b(x)$ jest silnie wypukłe na zbiorze X. W konsekwencji $\ln \varphi$ jest silnie wypukłe oraz b jest logarytmiczne silnie wypukłe na zbiorze X. □

Uwaga 2.1.4. Powyższy fakt nie zachodzi w przypadku funkcji silnie wypukłych zamiast funkcji logarytmicznie silnie wypukłych (patrz przykład 2.1.2).

2.2 Uwypuklenie funkcji na zbiorach zwartych

Załóżmy, że $b : \mathbb{R}^n \to \mathbb{R}$ jest funkcją klasy $\mathscr{C}^2$, μ-śilnie wypukłą i dodatnią w każdym punkcie zbioru $\mathbb{R}^n$.

Weźmy dowolnym zbiór zwarty i wypukły $X \subset \mathbb{R}^n$. Niech

$$S = \max\{b(x) : x \in X\}.$$

Weźmy dowolną funkcję $f : \mathbb{R}^n \to \mathbb{R}$ klasy $\mathscr{C}^2$ dodatnią na zbiorze X. Niech $m, D \in \mathbb{R}$ będą liczbami dodatnimi takimi, że

$$f(x) \geq m, \quad |\partial_\beta f(x)| \leq D, \quad |\partial_\beta^2 f(x)| \leq D \quad dla\ x \in X \text{ i } \beta \in S^{n-1}.$$

Niech

$$N(\mu, S, m, D) = \frac{S}{\mu}\left(\frac{D}{m} + \frac{D^2}{m^2}\right) + 1.$$

Następujący lemat jest wersją lematu 49 z pracy [18] Klaudii Rosiak.

Lemat 2.2.1. *Dla każdego $N > N(\mu, S, m, D)$ funkcja $\lambda_N(x) = b^N(x)f(x)$ jest silnie wypukła na zbiorze X.*

Dowód. Weźmy dowolne $N \geq N(\mu, S, m, D)$ oraz $x, \beta \in \mathbb{R}^n$, $|\beta| = 1$. Wtedy

$$\begin{aligned}
\partial_\beta \lambda_N(x) &= Nb^{N-1}(x)f(x)\partial_\beta b(x) + b^N(x)b(x)\partial_\beta f(x)\\
\partial_\beta^2 \lambda_N(x) &= N(N-1)b^{N-2}(x)f(x)\left(\partial_\beta b(x)\right)^2 + 2Nb^{N-1}(x)\partial_\beta b(x)\partial_\beta f(x)\\
&\qquad + Nb^{N-1}(x)f(x)\partial_\beta^2 b(x) + b^N(x)\partial_\beta^2 f(x).
\end{aligned}$$

Zatem dla $x \in \mathbb{R}^n$,

$$\begin{aligned}
&\partial_\beta^2 \lambda_N(x)\\
&= b^N(x)\left(N(N-1)f(x)\left(\frac{\partial_\beta b(x)}{b(x)}\right)^2 + 2N\frac{\partial_\beta b(x)}{b(x)}\partial_\beta f(x) + \partial_\beta^2 f(x) + Nf(x)\frac{\partial_\beta^2 b(x)}{b(x)}\right).
\end{aligned}$$

Ponieważ $b(x) > 0$ dla $x \in \mathbb{R}^n$, więc mamy

$$\partial_\beta^2 \lambda_N(x) = b^N(x)\Lambda(x),$$

gdzie

$$\Lambda(x) = N(N-1)f(x)\left(\frac{\partial_\beta b(x)}{b(x)}\right)^2 + 2N\frac{\partial_\beta b(x)}{b(x)}\partial_\beta f(x) + \partial_\beta^2 f(x) + Nf(x)\frac{\partial_\beta^2 b(x)}{b(x)}.$$

Ponieważ f i b są funkcjami klasy $\mathscr{C}^2$, więc funkcja λ jest również klasy $\mathscr{C}^2$. Wystarczy więc pokazać, że

$$\Lambda(x) > 0 \quad \text{dla} \ \ x \in X. \tag{2.1}$$

Istotnie, niech $x \in X$ i połóżmy $t = \frac{\partial_\beta b(x)}{b(x)}$. Z założeń o funkcjach f i b, mamy

$$\Lambda(x) \geq N(N-1)m|t|^2 - 2ND|t| - D + Nm\frac{\mu}{S}.$$

Wyróżnik funkcji kwadratowej po prawej stronie powyższej nierówności ma postać

$$\begin{aligned}
\Delta =& 4N^2D^2 - 4N(N-1)m\left(-D + Nm\frac{\mu}{S}\right) \\
=& 4N^2D^2 + 4N(N-1)mD - 4N^2(N-1)m^2\frac{\mu}{S} \\
=& 4N^2D^2 - \frac{4Nm^2\mu}{S}\left[N(N-1) - (N-1)\frac{SD}{\mu m}\right] \\
=& -\frac{4Nm^2\mu}{S}\left[N(N-1) - N\frac{SD}{\mu m} + \frac{S}{\mu}\frac{D}{m} - N\frac{S}{\mu}\frac{D^2}{m^2})\right] \\
=& -\frac{4Nm^2\mu}{S}\left[N\left(N - 1 - \frac{S}{\mu}\frac{D}{m} - \frac{S}{\mu}\frac{D^2}{m^2}\right) + \frac{S}{\mu}\frac{D}{m}\right].
\end{aligned}$$

Zatem, dla $N \geq N(\mu, S, m, D)$ mamy $\Delta < 0$ i w konsekwencji

$$N(N-1)m|t|^2 - 2ND|t| - D + Nm\frac{\mu}{S} > 0 \quad \text{dla} \ \ t \in \mathbb{R}.$$

To daje (2.1) i kończy dowód. $\square$

Niech

$$S' = \max\{b(x-\xi) : x,\, \xi \in X\}.$$

Z lematu 2.2.1 otrzymujemy

Wniosek 2.2.2. *Dla dowolnego $N \geq N(\mu, S', m, D)$ i dowolnego $\xi \in X$ funkcja*

$$\lambda_{N,\xi}(x) = b^N(x-\xi)f(x) \tag{2.2}$$

jest silnie wypukla na zbiorze X.

Uwaga 2.2.3. Niech $b : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją μ-silnie wypukłą, $\mu > 0$, oraz niech $X \subset \mathbb{R}^n$ będzie zbiorem zwartym i wypukłym. Niech $f : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^2$ i niech $D \in \mathbb{R}$ będzie liczbą dodatnią taką, że

$$|\partial_\beta^2 f(x)| \leq D \quad \textit{dla } x \in X \text{ oraz } \beta \in S^{n-1}.$$

Wówczas dla dowolnego $\xi \in \mathbb{R}^n$ oraz

$$N > \frac{D}{\mu},$$

funkcją $\Psi_{N,\xi} : \mathbb{R}^n \to \mathbb{R}$ określoną wzorem

$$\Psi_{N,\xi}(x) = Nb(x - \xi) + f(x), \quad x \in \mathbb{R}^n,$$

jest silnie wypukłą na X (dokładniej $N\mu - D$-silnie wypukła). Jeśli dodatkowo $N \geq \frac{D}{\mu} + 1$, to funkcja $\Psi_{N,\xi}$ jest μ-silnie wypukła.

Istotnie, weźmy dowolne $\xi \in \mathbb{R}^n$. Ponieważ $N\mu > D$, wówczas dla dowolnego $\beta \in S^{n-1}$, mamy

$$\partial_\beta^2 \Psi_{N,\xi}(x) = N\partial_\beta^2 b(x - \xi) + \partial_\beta^2 f(x) \geq N\eta - D > D - D = 0 \quad \text{for } x \in X.$$

To daje tezę uwagi.

Niech $f : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^2$ przyjmującą tylko wartości dodatnie. Weźmy dowolny zbiór wypukły i zwarty $X \subset \mathbb{R}^n$. Niech $m, D \in \mathbb{R}$ będą liczbami dodatnimi takimi, że

$$f(x) \geq m, \quad |\partial_\beta f(x)| \leq D, \quad |\partial_\beta^2 f(x)| \leq D \quad \textit{dla } x \in X \text{ i } \beta \in S^{n-1}.$$

Niech

$$N_{\exp}(\mu, m, D) = \frac{1}{\mu}\left(\frac{D}{m} + \frac{D^2}{m^2}\right).$$

Lemat 2.2.4. *Dla dwolnego $N > N_{\exp}(\mu, m, D)$ i dowolnego $\xi \in \mathbb{R}^n$ funkcja $\lambda_{N,\xi}(x) = b^N(x - \xi)f(x)$ jest logarytmicznie silnie wypukła na zbiorze X.*

Dowód. Weźmy dowolne $\xi \in \mathbb{R}^n$. Niech $\psi_{N,\xi} = \ln \lambda_{N,\xi}$. Wtedy

$$\psi_{N,\xi}(x) = N \ln b(x - \xi) + \ln f(x), \quad x \in \mathbb{R}^n$$

więc dla dowolnego $\beta \in S^{n-1}$, mamy

$$\partial_\beta \psi_{N,\xi}(x) = N\partial_\beta(\ln b(x-\xi)) + \frac{\partial_\beta f(x)}{f(x)}, \quad x \in \mathbb{R}^n,$$

oraz

$$\partial_\beta^2 \psi_{N,\xi}(x) = N\partial_\beta^2(\ln b(x-\xi)) + \frac{f(x)\partial_\beta^2 f(x) - (\partial_\beta f(x))^2}{f(x)^2}, \quad x \in \mathbb{R}^n.$$

W konsekwencji, dla $N > N_{\exp}(\mu, m, D)$ i $x \in X$, dostajemy

$$\partial_\beta^2 \psi_N(x) \geq N\mu - \frac{D}{m} - \frac{D^2}{m^2} > 0, \quad x \in \mathbb{R}^n.$$

Ponieważ $\partial_\beta^2 \psi_N$ jest funkcją ciągłą oraz zbiór X jest zwartym, więc otrzymujemy tezę. To kończy dowód. $\square$

2.2.1 Uwypuklanie wielomianów

Niech $f \in \mathbb{R}[x]$ będzie wielomianem postaci (1.11). Załóżmy, że $d = \deg f$. Wówczas wielomian f można przedstawić w postaci (1.12), to znaczy

$$f = f_0 + \cdots + f_d,$$

gdzie f_j jest wielomianem jednorodnym stopnia j lub wielomianem zerowym.

Weźmy dowolne $R > 0$. Z wniosku 1.6.6 mamy

Fakt 2.2.5. *Dla dowolnych $\beta \in S^{n-1}$ i $x \in \mathbb{R}^n$ takich, że $|x| \leq R$ mamy*

$$|\partial_\beta^2 f(x)| \leq D_n^2(f, R) \leq D_n(f, R). \tag{2.3}$$

Z uwagi 2.2.3 i faktu 2.2.5 mamy

Wniosek 2.2.6. *Jeśli $b : \mathbb{R}^n \to \mathbb{R}$ jest funkcją μ-silnie wypukłą klasy $\mathscr{C}^2$ i $N > D_n^2(f, R)/\mu$, to dla każdego $\xi \in \mathbb{R}^n$ funkcja*

$$\phi_{N,\xi} = Nb(x-\xi) + f(x), \qquad x \in \mathbb{R}^n,$$

jest silnie wypukła w kuli $B_R := \{x \in \mathbb{R}^n : |x| \leq R\}$.

Niech $b : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją μ-silnie wypukłą klasy $\mathscr{C}^2$, która przyjmuje tylko wartości dodatnie, niech $X \subset \mathbb{R}^n$ będzie zbiorem zwartym i wypukłym i niech

$$S = \max\{b(x) : x \in X\},\ S' = \max\{b(x-\xi) : x, \xi \in X\},\ R = \max\{|x| : x \in X\}.$$

Z lematu 2.2.1 i wniosku 2.2.2 otrzymujemy

Wniosek 2.2.7. *Jeśli*

$$f(x) \geq m \quad dla\ x \in X \tag{2.4}$$

dla pewnej stałej dodatniej m, to dla dowolnego

$$N > N(\mu, S, m, D_n(f, R))$$

funkcja

$$\varphi_N(x) = b^N(x) f(x)$$

jest silnie wypukła na zbiorze X.

Jeśli dodatkowo

$$N \geq N(\mu, S', m, D_n(f, R)),$$

to dla dowolnego $\xi \in X$ funkcja

$$\varphi_{N,\xi}(x) = b^N(x - \xi) f(x)$$

jest silnie wypukła na zbiorze X.

Uwaga 2.2.8. Z wniosku 1.6.6 mamy

$$|f(x)| \leq D_n^0(f, R) \quad \text{oraz} \quad f(x) + D_n^0(f, R) \geq 0 \quad \text{dla } x \in \mathbb{R}^n,\ |x| \leq R.$$

Zatem, jeśli weźmiemy wielomian $\tilde{f} = f + D_n^0(f, R) + 1$ zamiast wielomianu f, to wielomian $\tilde{f}$ spełnia (2.4) z $m = 1$, więc z wniosku 2.2.7 otrzymujemy, że dla każdego $N > N(\mu, S, 1, D_n(f, R) + D_n^0(f, R) + 1)$, funkcja $\varphi_N(x) = b^N(x)\tilde{f}(x)$ jest silnie wypukła na zbiorze X. Analogicznie jak powyżej z tego samego wniosku mamy, że dla dowolnego $N \geq N(\mu, S', 1, D_n(f, R) + D_n^0(f, R) + 1)$ i dowolnego $\xi \in X$ funkcja $\varphi_{N,\xi}(x) = b^N(x - \xi)\tilde{f}(x)$ jest silnie wypukła na zbiorze X. Z punktu widzenia celu jaki nam przyświeca, a mianowicie wyszukiwania punktów krytycznych funkcji f, nie ma znaczenia, czy weźmiemy funkcję f czy $\tilde{f}$, bo obie te funkcje maja ten sam zbiór punktów krytycznych. Zatem założenie $f(x) \geq m$ dla $x \in X$ możemy pominąć, biorąc $\tilde{f}$ zamiast f i $m = 1$.

2.2.2 Uwypuklenie logarytmiczne wielomianów na zbiorach nieograniczonych

Dla wielomianu f postaci (1.11) takiego, że $f_{d*} > 0$ oraz dla każdego $\mu > 0$, kładziemy

$$N_{\exp,\infty}(\mu, f) = \frac{d(d+1)||f||}{4\mu m(f)},$$

gdzie liczba $m(f)$ jest określina przed lematem 1.6.4, to znaczy

$$m(f) := f_{d*} - \sum_{j=0}^{d-1} \mathbb{K}(f)^{j-d}||f_j||,$$

gdzie $\mathbb{K}(f) := \frac{2\|f\|}{f_{d*}}$. Przypomnijmy, że dla $R > 0$,

$$D_n(f, R) := \max\left\{\sum_{j=1}^{d} j||f_j||R^{j-1}; \sum_{j=1}^{d} j(j-1)||f_j||R^{j-2}\right\}.$$

Oczywiście, dla każdych $\beta, x \in \mathbb{R}^n$ takich, że $|\beta| = 1$ i $|x| \leq R$ mamy

$$|\partial_\beta f(x)| \leq D_n(f, R), \quad |\partial_\beta^2 f(x)| \leq D_n(f, R). \tag{2.5}$$

Niech $b : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją logarytmicznie μ-silnie wypukłą. Z lematu 2.2.4 i wniosku 1.6.5 mamy

Wniosek 2.2.9. *Niech $X \subset \mathbb{R}^n$ będzie zbiorem domkniętym i wypukłym. Niech f będzie wielomianem postaci* (1.11) *takim, że $f_{d*} > 0$ oraz istnieje $m > 0$ takie, że $f(x) \geq m$ dla $x \in X$. Wówczas dla każdego*

$$N > \max\{N_{\exp}(\mu, m, D_n(f, \mathbb{K}(f))), N_{\exp,\infty}(\mu, f)\}$$

oraz dowolnego $\xi \in \mathbb{R}^n$ funkcja $\varphi_{N,\xi}(x) = b^N(x - \xi)f(x)$ jest logarytmicznie silnie wypukła na zbiorze X.

Dowód. Weźmy dowolne $\xi \in \mathbb{R}^n$. Niech $\psi_{N,\xi}(x) = \ln \varphi_{N,\xi}(x)$. Weźmy dowolne $\beta \in S^{n-1}$. Z lematu 2.2.4 otrzymujemy, że istnieje $\mu_1 > 0$ takie, że $\partial_\beta^2 \psi_{N,\xi}(x) \geq \mu_1$ dla $x \in X$, $|x| \leq \mathbb{K}(f)$. Ponieważ

$$\partial_\beta^2 \psi_{N,\xi}(x) = N\partial_\beta^2(\ln b(x-\xi)) + \frac{\partial_\beta^2 f(x)}{f(x)} - \left(\frac{\partial_\beta f(x)}{f(x)}\right)^2, \quad x \in \mathbb{R}^n,$$

zatem z wniosku 1.6.5 istnieje $\mu_2 > 0$ takie, że $\partial_\beta^2 \psi_{N,\xi}(x) \geq \mu_2$ dla $x \in X$, $|x| \geq \mathbb{K}(f)$. To kończy dowód. □

2.2.3 Wielomiany o współczynnikach całkowitych

W przypadku zastosowania powyższych wyników ważne jest oszacowanie liczb f_{d*}, $m = \min\{f(x) : x \in X\}$ i $R = \max\{|x| : x \in X\}$ dla wielomianu f oraz zbioru semialgebraicznego, zwartego i wypukłego $X \subset \mathbb{R}^n$. W przypadku, gdy f oraz wielomiany opisujące zbiór X mają współczynniki całkowitych, powyższe liczby można efektywnie oszacować. Dokładniej, niech $X \subset \mathbb{R}^n$, $n \geq 2$, będzie zbiorem semialgebraicznym i zwartym postaci

$$X = \{x \in \mathbb{R}^n : g_1(x) = 0, \ldots, g_l(x) = 0, g_{l+1}(x) \geq 0, \ldots, g_k(x) \geq 0\}, \tag{2.6}$$

gdzie $g_1, \ldots, g_k \in \mathbb{Z}[x]$. Przy powyższych oznaczeniach, G. Jeronimo, D. Perrucci, E. Tsigaridas in [6] udowodnili, że

Twierdzenie 2.2.10. *Niech $f, g_1, \ldots, g_k \in \mathbb{Z}[x]$ będą wielomianami ze stopniami ograniczonymi przez liczbę całkowitą parzystą d i współczynnikami o wartości bezwzględnych nie przekraczających H, i niech $\tilde{H} = \max\{H, 2n+2k\}$. Jeśli $f(x) > 0$ dla $x \in X$ oraz zbiór X postaci* (2.6) *jest zwarty, to*

$$f(x) \geq \left(2^{4-\frac{n}{2}}\tilde{H}d^n\right)^{-n2^nd^n} \qquad \text{dla } x \in X.$$

Z twierdzenia 2.2.10 mamy

Wniosek 2.2.11. *Niech $f \in \mathbb{Z}[x]$ będzie wielomianem jednorodnym o stopniu ograniczonym przez liczbę całkowitą parzystą d i współczynnikach o wartościach bezwzględnych nie przekraczających H, i niech $\tilde{H} = \max\{H, 2n+2\}$. Jeśli $f(x) > 0$ dla $|x| = 1$, to*

$$f(x) \geq \left(2^{4-\frac{n}{2}}\tilde{H}d^n\right)^{-n2^nd^n} \qquad \text{dla } |x| = 1.$$

Z twierdzenia 2.2.10 otrzymujemy natychmiast (patrz [8, Theorem 2.7])

Twierdzenie 2.2.12. *Niech $X \subset \mathbb{R}^n$ będzie zbiorem semialgebraicznym zwartym i wypukłym postaci* (2.6) *i niech $f, g_1, \ldots, g_k \in \mathbb{Z}[x]$ będą wielomianami o stopniach ograniczonych przez liczbą całkowitą parzystą d i współczynnikaach o wartościach bezwzględnych nie przekraczających H. Biorąc*

$$\mathfrak{b}(n, d, H, k) = \left(2^{4-\frac{n}{2}} \max\{H, 2n + 2k\} d^n\right)^{-n2^n d^n}$$

oraz

$$R = \sqrt{\left[\mathfrak{b}(n + 1, \max\{d, 4\}, H, k + 2)\right]^{-1} - 1},$$

mamy

$$\max\{|x| : x \in X\} \leq R. \tag{2.7}$$

Rozdział 3

Uwypuklenie funkcji na zbiorach nieograniczonych

W tym rozdziale przedstawimy uwypuklenie funkcji przez funkcje logarytmicznie wypukłe na zbiorach nieograniczonych. Dokładniej, niech $b : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją logarytmcznie silnie wypukłą klasy $\mathscr{C}^k$, $k \geq 2$ i niech

$$0 = \operatorname{argmin}_{\mathbb{R}^n} b, \qquad b(0) = 1.$$

Niech $f : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^k$ oraz niech $X \subset \mathbb{R}^n$ będzie zbiorem nieograniczonym, wypukłm i domkniętym.

W tym rozdziale pokażemy różnice między uwypukleniem funkcji f na zbiorze X przez dodanie do niej funkcji b i pomnożenie jej przez funkcję b. Pokażemy, że jeśli funkcja f jest wielomianiem dodatnim na zbiorze $\mathbb{R}^n$, którego forma wiodąca jest dodatnia poza początkiem układu współrzędnych, to istnieje jednoznacznie obliczalna liczba $N \geq 1$, taka, że dla dowolnego $\xi \in \mathbb{R}^n$, funkcja $\varphi_{N,\xi}(x) = b(N(x-\xi))f(x)$ jest silnie wypukła na zbiorze X. Zauważymy, że nie można tego uzyskać dla funkcji $\phi_{N,\xi}(x) = b(N(x-\xi)) + f(x)$.

3.1 Uwagi o uwypuklaniu funkcji przez dodanie funkcji silnie wypukłej

Bezpośrednio sprawdzamy, że zachodzi następujący fakt [por. uwaga2.2.3].

Fakt 3.1.1. *Niech $b : \mathbb{R}^n \to (0, +\infty)$ będzie funkcją μ-silnie wypukłą klasy $\mathscr{C}^2$, Niech $X \subset \mathbb{R}^n$ będzie zbiorem zwartym i wypukłym oraz niech $f : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^2$. Niech $C \in \mathbb{R}$ będzie stałą taką, że $\partial_\beta^2 f(x) \geq C$ dla dowolnego $x \in X$ i dowolnego $\beta \in S^{n-1}$. Wówczas dla dowolnych $N > -C/\mu$ i $\xi \in \mathbb{R}^n$, funkcją $\phi_{N,\xi} : \mathbb{R}^n \to \mathbb{R}$ określoną wzorem*

$$\phi_{N,\xi}(x) = Nb(x - \xi) + f(x), \quad x \in \mathbb{R}^n, \tag{3.1}$$

jest silnie wypukłą na zbiorze X (dokładniej $N\mu + C$-silnie wypukła).

Powyższy fakt jest trudno zastosować w przypadku zbiorów nieograniczonych. Mianowicie mamy

Fakt 3.1.2. *Niech $b : \mathbb{R}^n \to (0, +\infty)$ będzie funkcją wypukłą klasy $\mathscr{C}^2$, Niech $X \subset \mathbb{R}^n$ będzie zbiorem wypukłym i domkniętym, niech $f : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^2$ i niech $N > 0$. Jeśli dla każdego $\xi \in X$ funkcja $\phi_{N,\xi}$ określona wzorem* (3.1) *jest wypukła na zbiorze X, to*

$$\partial_\beta^2 \phi_{N,\xi}(\xi) = N\partial_\beta^2 b(0) + \partial_\beta^2 f(\xi) \geq 0 \quad \textit{dla każdego } \beta \in S^{n-1}.$$

W szczególności pochodne $\partial_\beta^2 f$, $\beta \in S^{n-1}$, są wspólnie ograniczone od dołu na zbiorze X.

W związku z powyższym, fakt 3.1.1 można przenieść do przypadku zbiorów nieograniczonych tylko przy założeniu, że pochodne kierunkowe drugiego rzędu $\partial_\beta^2 f$, $\beta \in S^{n-1}$, są ograniczone od dołu na zbiorze X.

Przykład 3.1.3. Łatwo jest sprawdzić, że wielomian

$$f(x, y) = (x^2 - y^2)^2 + x^4 + 1$$

nie spełnia tezy faktu 3.1.2. Istotnie, $\frac{\partial^2 f}{\partial x^2}(0, y) = -4y^2$ nie jest ograniczone od dołu. Tak więc, funkcja $\phi_{N,\xi}$ nie jest wypukła w $\mathbb{R}^2$ dla żadnej funkcji wypukłej b. Nie można tego naprawić nawet przy założeniu, że funkcja b jest logartmicznie μ-silnie wypukła.

Aby uzyskać funkcję silnie wypukłą, na zbiorze nieograniczonym, przez dodanie do niej funkcji silnie wypukłej, musimy dołączyć współczynnik zależny od ξ. Musimy przy tym zakładać, że funkcja b jest logarytmicznie silnie wypukła. Pokazuje to następujący przykład.

Przykład 3.1.4. Rozważmy wielomian f z przykładu 3.1.3 i niech $b(x,y) = x^2+y^2+1$. Funkcja b jest 2-silnie wypukła, natomiast dla dowolnej funkcji $g(\xi)$, funkcja $\Phi_\xi(x,y) = g(\xi)b((x,y)-\xi)+f(x,y)$ nie jest wypukła na zbiorze $\mathbb{R}^2$. Istotnie, dla dowolnego $\xi \in \mathbb{R}^2$ mamy $\frac{\partial^2 \Phi_\xi}{\partial_x^2}(0,y) = 2g(\xi)-4y^2 \to -\infty$ gdy $y \to \infty$.

Przy założeniu, że funkcja b jest logarytmicznie μ-silnie wypukła, a funkcja f jest wielomianem, w paragrafie 3.2, pokażemy że dla dowolnego $\xi \in \mathbb{R}^n$ funkcja $\phi_\xi(x) = Nb(\xi)b(x-\xi)+f(x)$ jest μ-silnie wypukła na zbiorze $\mathbb{R}^n$, dla pewnej efektywnie wyliczalnej liczby N (patrz lemat 3.2.5, por. lemat 3.2.1).

Zauważmy, że wielomian f w powyższych przykładach przyjmuje tylko wartości dodatnie, a forma wiodąca $f_4(x,y) = (x^2-y^2)^2+x^4$ przyjmuje również tylko wartości dodatnie w zbiorze $\mathbb{R}^2 \setminus \{(0,0)\}$. W twierdzeniu 3.3.3 pokażemy, że przy tych założeniach można uzyskać silna wypukłość (nawet logarytmicznie silna wypukłść) funkcji f poprzez pomnożenie jej przez funkcję logarytmicznie silnie wypukłą b (bez dodatkowych czynników zależnych od ξ). Sposób uzyskania funkcji wypukłej z danego wielomianu na zbiorze ograniczonym przez pomnożenie go przez funkcję silnie wypukłą zaproponowano w [9] i w [8] – przez pomnożenie go przez funkcję logarytmicznie silnie wypukłą $b(x) = e^{|x|^2}$. W tym przypadku potrzebujemy również dodatkowych założeń dla wielomianu f. Mianowicie mamy

Fakt 3.1.5. *Niech $b : \mathbb{R}^n \to (0,+\infty)$ będzie funkcją μ-silnie wypukłą klasy $\mathscr{C}^2$ taką, że $0 = \operatorname{argmin}_{\mathbb{R}^n} b$, niech $X \subset \mathbb{R}^n$ będzie zbiorem wypukłym i domkniętym oraz niech $f : \mathbb{R}^n \to (0,+\infty)$ będzie funkcją klasy $\mathscr{C}^2$. Niech*

$$\varphi_{N,\xi}(x) = b(N(x-\xi))f(x),$$

gdzie $N > 0$.

(i) *Jeśli dla dowolnego $\xi \in X$ funkcja $\varphi_{N,\xi}$ jest ściśle wypukła na zbiorze X, to $C\partial_\beta^2 f(x) \geq -f(x)$ dla dowolnych $x \in X$ i $\beta \in S^{n-1}$ i pewnej stałej $C > 0$.*

(ii) *Jeśli fankcja b jest logarytmicznie μ-silnie wypukła i $C\partial_\beta^2 f(x) \geq -f(x)$ oraz $Cf(x) \geq |\partial_\beta f(x)|$ dla dowolnych $x \in X$, $\beta \in S^{n-1}$ i pewnej stałej $C > 0$, to dla $N > \frac{2C}{\sqrt{\mu}} + \frac{1}{C\mu}$, funkcja $\varphi_{N,\xi}$ jest ściśle wypukła na zbiorze X.*

Dowód. Istotnie, przy założeniach punktu (i),

$$(3.2)\quad \partial_\beta^2\varphi_{N,\xi}(x)$$
$$= N^2\partial_\beta^2 b(N(x-\xi))f(x) + 2N\partial_\beta b(N(x-\xi))\partial_\beta f(x) + b(N(x-\xi))\partial_\beta^2 f(x) \geq 0$$

dla dowolnych $x, \xi \in X$ i $\beta \in S^{n-1}$. Ponieważ $0 = \operatorname{argmin}_{\mathbb{R}^n} b$, więc mamy $\partial_\beta b(0) = 0$. Stąd biorąc $\xi = x$ otrzymujemy (i).

Zgodnie z załżeniami (ii), wykorzystując własności funkcji logarytmicznie silnie wypukłych (patrz wniosek 1.4.3 w paragrafie 1.4) otrzymujemy

$$\begin{aligned}\partial_\beta^2\varphi_{N,\xi}(x) \geq& f(x)\left(N^2\partial_\beta^2 b(N(x-\xi)) - 2NC|\partial_\beta b(N(x-\xi))| - \frac{1}{C}b(N(x-\xi))\right)\\ \geq& f(x)\partial_\beta^2 b(N(x-\xi))\left(N^2 - \frac{2C}{\sqrt{\mu}}N - \frac{1}{\mu C}\right) > 0\end{aligned}$$

dla $x, \xi \in X$, $\beta \in S^{n-1}$ oraz $N > \frac{2C}{\sqrt{\mu}} + \frac{1}{C\mu}$. To daje (ii). □

W przypadku tej metody naturalne jest założenie, że funkcja f przyjmuje tylko wartości dodatnie. Istotnie, z dowodu w pnkcie (i) fakt 3.1.5, otrzymujemy

Fakt 3.1.6. *Niech $b : \mathbb{R}^n \to (0, +\infty)$ będzie funkcją μ-silnie wypukłą klasy $\mathscr{C}^2$ taką, że $0 = \operatorname{argmin}_{\mathbb{R}^n} b$, niech $X \subset \mathbb{R}^n$ będzie zbiorem wypukłym oraz niech $f : \mathbb{R}^n \to \mathbb{R}$ będzie klasy $\mathscr{C}^2$. Jeśli funkcja $\varphi_{N,\xi}$ jest wypukła na zbiorze X dla pewnego $\xi \in X$ i wystarczająco dużej liczby $N > 0$, to $f(x) \geq 0$ dla $x \in X$.*

Dowód. Gdyby $f(x) < 0$, to dla dostatecznie dużych N mielibyśmy

$$N^2\partial_\beta^2 b(0)f(x) + b(0)\partial_\beta^2 f(x) < 0,$$

bo $\partial_\beta^2 b(0) \geq \mu$. To jest niemożliwe, bo z (3.2) dla $\xi = x \in X$ oraz $\beta \in S^{n-1}$, mamy $N^2\partial_\beta^2 b(0)f(x) + b(0)\partial_\beta^2 f(x) \geq 0$. Otrzymana sprzeczność kończy dowód. □

Możemy łatwo sprawdzić, czy wielomiany przyjmujące tylko wartości dodatnie i których formy wiodące przyjmują tylko wartości dodatnie poza zerem w $\mathbb{R}^n$, spełniają założenie faktu 3.1.5 (i) oraz tezę (ii). Dla logarytmiczniej

μ-silnie wypukłej funkcji b i wielomianu f przyjmującego tylko dodatnie wartości, którego forma wiodąca przyjmuje tylko wartości dodatnie w $\mathbb{R}^n \setminus \{0\}$, w twierdzeniach 3.3.1 i 3.3.3 pokażemy silna wypukłość (odpowiednio logarytmicznie silna wypukłość) funkcji $\varphi_{N,\xi}(x) = b(N(x-\xi))f(x)$ na zbiorze X dla dowolnego $\xi \in X$ i pewnej efektywnie obliczonych N.

3.2 Uwypuklenie funkcji przez dodanie funkcji logarytmicznie silnie wypukłych

Niech $X \subset \mathbb{R}^n$ będzie zbiorem wypukłym i domkniętym i niech $f : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^k$, $k \geq 2$, posiadającą wzrost wielomianu drugiego rzędu

$$|\partial_\beta^2 f(x)| \leq D(1+|x|)^\alpha \quad \textit{dla } x \in X \text{ i } \beta \in S^{n-1}, \tag{3.3}$$

dla pewnych $D > 0$ i $\alpha \in \mathbb{N}$.

Niech $b : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^k$ i logarytmicznie μ-silnie wypukłą, $\mu > 0$. Załóżmy, że $0 = \operatorname{argmin}_{\mathbb{R}^n} b$ i $b(0) = 1$.

Lemat 3.2.1. *Niech*

$$N(|\xi|) = \frac{D}{\mu}\left(|\xi| + 1 + \sqrt{\frac{\alpha}{\mu}}\right)^\alpha + 1 \tag{3.4}$$

Wówczas dla dowolnego $\xi \in \mathbb{R}^n$ funkcja $\phi_\xi : \mathbb{R}^n \to \mathbb{R}$ określona wzorem

$$\phi_\xi(x) = N(|\xi|)b(x-\xi) + f(x), \quad x \in \mathbb{R}^n,$$

jest silnie wypukła na zbiorze X (dokładniej μ-silnie wypukła).

Dowód. Z faktu 1.4.4, dla dowonych $\beta \in S^{n-1}$ i $x, \xi \in \mathbb{R}^n$ mamy

$$\partial_\beta^2 b(x-\xi) \geq \mu \exp\left(\frac{\mu}{2}|x-\xi|^2\right) \geq \mu \exp\left(\frac{\mu}{2}(|x|-|\xi|)^2\right),$$

więc z założenia o funkcji f, mamy

$$\partial_\beta^2 \phi_\xi(x) \geq N(|\xi|)\mu \exp\left(\frac{\mu}{2}(|x|-|\xi|)^2\right) - D(1+|x|)^\alpha.$$

Wystarczy więc pokazać, że

$$N(|\xi|)\mu \exp\left(\frac{\mu}{2}(|x|-|\xi|)^2\right) - D(1+|x|)^\alpha \geq \mu.$$

Ponieważ $\exp\left(\frac{\mu}{2}(|x|-|\xi|)^2\right) \geq 1$ oraz $\mu > 0$, więc wystarczy pokazać, że

$$N(|\xi|) \geq \frac{D(1+|x|)^\alpha}{\mu \exp\left(\frac{\mu}{2}(|x|-|\xi|)^2\right)} + 1. \tag{3.5}$$

Weźmy funkcję

$$h(r) = \frac{D(1+r)^\alpha}{\mu \exp\left(\frac{\mu}{2}(r-|\xi|)^2\right)}, \quad r \geq 0.$$

Bezpośrednio sprawdzamy, że największą wartością funkcji h jest

$$h(r_1) = \frac{D(1+r_1)^\alpha}{\mu \exp\left(\frac{\mu}{2}(r_1-|\xi|)^2\right)},$$

gdzie

$$r_1 = \frac{|\xi| - 1 + \sqrt{(|\xi|+1)^2 + \frac{4\alpha}{\mu}}}{2} > 0.$$

Ponieważ

$$h(r_1) = \frac{D\left(\frac{|\xi|+1+\sqrt{(|\xi|+1)^2+\frac{4\alpha}{\mu}}}{2}\right)^\alpha}{\mu \exp\left(\frac{\mu}{2}\left(\frac{-|\xi|-1+\sqrt{(|\xi|+1)^2+\frac{4\alpha}{\mu}}}{2}\right)^2\right)} \leq \frac{D}{\mu}\left((|\xi|+1) + \sqrt{\frac{\alpha}{\mu}}\right)^\alpha, \tag{3.6}$$

więc otrzymujemy (3.5) i w konsekwencji dostajemy tezę. □

Uwaga 3.2.2. Ponieważ,

$$0 < -(|\xi|+1) + \sqrt{(|\xi|+1)^2 + \frac{4\alpha}{\mu}} \to 0, \quad \text{as } |\xi| \to \infty,$$

więc widzimy, że oszacowanie w (3.6) jest asymptotycznie optymalne.

Niech $N : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją określoną wzorem (3.4). Określamy funkcję $N_1 : \mathbb{R}^n \to \mathbb{R}$, wzorem

$$N_1(\xi) = N(\sqrt{|\xi|^2+1}) = \frac{D}{\mu}\left(\sqrt{|\xi|^2+1} + 1 + \sqrt{\frac{\alpha}{\mu}}\right)^\alpha + 1. \tag{3.7}$$

Oczywiście, $N_1(\xi) > N(|\xi|)$ dla $\xi \in \mathbb{R}^n$ oraz funkcja N_1 jest klasy $\mathscr{C}^\infty$. Niech $\psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ będzie funkcją określoną wzorem

$$\psi(\xi, x) = N_1(\xi) b(x-\xi) + f(x), \quad (\xi, x) \in \mathbb{R}^n \times \mathbb{R}^n.$$

Biorąc $\psi_\xi(x) = \psi(\xi, x)$. Z lematu 3.2.1 mamy

Wniosek 3.2.3. *Dla dowolnego $\xi \in \mathbb{R}^n$ funkcja ψ_ξ jest μ-silnie wypukła na zbiorze X.*

Ponieważ funkcja b jest logartmicznie silnie wypukła, to istnieje $R > 0$ taka, że $b(\xi) > N(|\xi|)$ dla $|\xi| \geq R$. Tak więc z lematu 3.2.1 łatwo otrzymujemy następujący wniosek.

Wniosek 3.2.4. *Istnieje liczba $N > 0$ taka, że dla dowolnego $\xi \in \mathbb{R}^n$ funkcja $\psi_\xi : \mathbb{R}^n \to \mathbb{R}$ określona wzorem $\psi_\xi(x) = Nb(\xi)b(x-\xi)+f(x)$ jest silnie wypukła na zbiorze X.*

Aby uzyskać efektywne oszacowanie dla stałej N we wniosku 3.2.4, musimy jednak powtórzyć dowód lematu 3.2.1. Mianowicie mamy

Lemat 3.2.5. *Niech*

$$N > \frac{D}{\mu 2^\alpha} \exp\left(\frac{2\alpha^2}{\mu} + \frac{4\alpha}{\mu\sqrt{\mu}} + \alpha + \frac{2}{\sqrt{\mu}} + \frac{2}{\mu^2} + 2\sqrt{\frac{\alpha}{\mu}}\right). \tag{3.8}$$

Wówczas dla dowolnego $\xi \in \mathbb{R}^n$, funkcja $\psi_\xi : \mathbb{R}^n \to \mathbb{R}$ określona wzorem

$$\psi_\xi(x) = Nb(\xi)b(x-\xi) + f(x)$$

jest

$$\mu_1 := \ln \frac{N\mu 2^\alpha}{D} - \left(\frac{2\alpha^2}{\mu} + \frac{4\alpha}{\mu\sqrt{\mu}} + \alpha + \frac{2}{\sqrt{\mu}} + \frac{2}{\mu^2} + 2\sqrt{\frac{\alpha}{\mu}}\right) - \textit{silnie wypukła}$$

na zbiorze X.

Dowód. Weźmy dowone $\beta \in S^{n-1}$. Mamy $\partial_\beta^2[b(\xi)b(x-\xi)] = b(\xi)\partial_\beta^2 b(x-\xi)$, więc z faktu 1.4.4,

$$\begin{aligned} \partial_\beta^2[b(\xi)b(x-\xi)] \geq & \mu \exp\left(\frac{\mu}{2}\left(|x-\xi|^2 + |\xi|^2\right)\right) \\ \geq & \mu \exp\left(\frac{\mu}{2}\left((|x| - |\xi|)^2 + |\xi|^2\right)\right) \end{aligned} \tag{3.9}$$

dla dowolnych $\xi \in \mathbb{R}^n$ i $x \in X$.

Zauważmy, że dla dowolnych $\xi \in \mathbb{R}^n$ i $x \in X$,

$$\partial_\beta \psi_\xi(x) \geq \mu_1. \tag{3.10}$$

Istotnie, z założenia o funkcji f oraz z (3.9) widzimy, że wystarczy pokazać, że dla $(\xi, x) \in \mathbb{R}^n \times X$,

$$(1+|x|)^\alpha < \frac{N\mu}{D} \exp\left(\frac{\mu}{2}\left((|x|-|\xi|)^2+|\xi|^2\right)\right),$$

równoważnie, że

$$\alpha \ln(1+|x|) < \ln\frac{N\mu}{D} + \frac{\mu}{2}\left((|x|-|\xi|)^2+|\xi|^2\right).$$

Niech $r=|x|$. Biorąc

$$h(r) = -\alpha\ln(1+r) + \ln\frac{N\mu}{D} + \frac{\mu}{2}\left((r-|\xi|)^2+|\xi|^2\right), \quad r \geq 0,$$

widzimy, że $h'(r) = \frac{-\alpha}{1+r} + \mu(r-|\xi|)$ a więc bezpośrednio sprawdzamy, że minimalna wartość funkcji h jest osiągnięta w punkcie

$$r_0 = \frac{1}{2}\left(|\xi|-1+\sqrt{(1-|\xi|)^2+\frac{4}{\mu}(|\xi|+\alpha)}\right).$$

Oznaczmy $A = \sqrt{(1-|\xi|)^2+\frac{4}{\mu}(|\xi|+\alpha)}$. Wówczas

$$A \leq 1+|\xi|+\frac{2}{\sqrt{\mu}}(|\xi|+1)+2\sqrt{\frac{\alpha}{\mu}},$$

więc dla każdego $r \geq 0$ mamy

$$\begin{aligned}
h(r) \geq h(r_0) &= \ln\frac{N\mu}{D} + \frac{\mu}{2}\left((r_0-|\xi|)^2+|\xi|^2\right) - \alpha\ln(1+r_0)\\
&= \ln\frac{N\mu}{D} + \frac{\mu}{2}\left(\frac{(-|\xi|-1+A)^2}{4}+|\xi|^2\right) - \alpha\ln\left(\frac{|\xi|+1+A}{2}\right)\\
&\geq \ln\frac{N\mu}{D} + \frac{\mu}{2}|\xi|^2 - \alpha|\xi| - \alpha A + \alpha\ln 2\\
&\geq \ln\frac{N\mu 2^\alpha}{D} + \frac{\mu}{2}|\xi|^2 - \alpha|\xi| - \alpha\left(1+|\xi|+\frac{2}{\sqrt{\mu}}(|\xi|+1)+2\sqrt{\frac{\alpha}{\mu}}\right)\\
&\geq \ln\frac{N\mu 2^\alpha}{D} + \frac{\mu}{2}|\xi|^2 - \left(2\alpha+\frac{2}{\sqrt{\mu}}\right)|\xi| - \alpha - \frac{2}{\sqrt{\mu}} - 2\sqrt{\frac{\alpha}{\mu}}\\
&\geq \ln\frac{N\mu 2^\alpha}{D} - \left(\frac{2\alpha^2}{\mu}+\frac{4\alpha}{\mu\sqrt{\mu}}+\frac{2}{\mu^2}+\alpha+\frac{2}{\sqrt{\mu}}+2\sqrt{\frac{\alpha}{\mu}}\right).
\end{aligned}$$

Stąd dla stałej N spełnionej (3.8) otrzymujemy, że (3.10) zachodzi. Co daje tezę. $\square$

3.3 Uwypuklenie wielomianów przez pomnożenie ich przez funkcję logarytmicznie silnie wypukłą

Niech $f \in \mathbb{R}[x]$ będzie wielomienem postaci (1.11) stopnia d i niech

$$f = f_0 + \cdots + f_d,$$

gdzie f_j jest wielomianem jednorodnym stopnia j lub zerem. Przypomnijmy, że $f_{d*} = \min_{|x|=1} f_d(x)$. Oczywiście $f_{d*} > 0$ wtedy i tylko wtedy, gdy forma wiodąca f_d wielomianu f przyjmuje tylko wartości dodatnie w $\mathbb{R}^n \setminus \{0\}$.

Niech $b : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją logarytmicznie μ-silnie wypukłą klasy $\mathscr{C}^k$, $k \geq 2$. Załóżmy, że

$$0 = \operatorname{argmin}_{\mathbb{R}^n} b \quad \text{i} \quad b(0) = 1.$$

Niech

$$K = 2\max\left\{ \frac{||f_{d-1}|| + \frac{2d}{\sqrt{\mu}}||f_d||}{f_{d*}}, \max_{1\leq j\leq d-2}\left(\frac{||f_j|| + \frac{2(j+2)}{\sqrt{\mu}}||f_{j+1}|| + \frac{(j+1)(j+2)}{\mu}||f_{j+2}||}{f_{d*}} \right)^{\frac{1}{j}} \right\}.$$

Twierdzenie 3.3.1. *Załóżmy, że $f_{d*} > 0$ oraz istnieje $m > 0$ takie, że*

$$f(x) \geq m \quad \text{dla } x \in \mathbb{R}^m. \tag{3.11}$$

Wówczas dla dowolnego

$$N > 2\max\left\{ \frac{\frac{2}{\sqrt{\mu}}\sum_{j=0}^{d-1}(j+1)||f_{j+1}|| \cdot K^j}{m}, \sqrt{\frac{\frac{1}{\mu}\sum_{j=0}^{d-2}(j+1)(j+2)||f_{j+2}|| \cdot K^j + 1}{m}} \right\} \tag{3.12}$$

oraz dla dowolnego $\xi \in \mathbb{R}^n$ funkcja $\varphi_{N,\xi} : \mathbb{R}^n \to \mathbb{R}$ określona wzorem

$$\varphi_{N,\xi}(x) = b(N(x-\xi))f(x)$$

jest μ-silnie wypukła w $\mathbb{R}^n$.

Dowód. Weźmy dowolne $\beta \in S^{n-1}$ i $x \in \mathbb{R}^n$. Mamy

$$\partial_\beta^2 \varphi_\xi(x) = N^2 \partial_\beta^2 b(N(x-\xi)) f(x) + 2N \partial_\beta b(N(x-\xi)) \partial_\beta f(x) + b(N(x-\xi)) \partial_\beta^2 f(x).$$

Wówczs, z lematu 1.6.1 i wniosku 1.4.3,

$$\begin{aligned}
\partial_\beta^2 \varphi_\xi(x) \geq & N^2 \partial_\beta^2 b(N(x-\xi)) f_{d*} |x|^d - N^2 \partial_\beta^2 b(N(x-\xi)) \sum_{j=0}^{d-1} ||f_j|| \cdot |x|^j \\
& - N \frac{2}{\sqrt{\mu}} \partial_\beta^2 b(N(x-\xi)) \sum_{j=0}^{d-1} (j+1) ||f_{j+1}|| \cdot |x|^j \\
& - \frac{1}{\mu} \partial_\beta^2 b(N(x-\xi)) \sum_{j=0}^{d-2} (j+1)(j+2) ||f_{j+2}|| \cdot |x|^j \\
= & \partial_\beta^2 b(N(x-\xi)) g_N(|x|),
\end{aligned}$$

gdzie

$$\begin{aligned}
g_N(t) = & N^2 f_{d*} t^d - \left(N^2 ||f_{d-1}|| + N \frac{2d}{\sqrt{\mu}} ||f_d|| \right) t^{d-1} \\
& - \sum_{j=0}^{d-2} \left(N^2 ||f_j|| + N \frac{2(j+2)}{\sqrt{\mu}} ||f_{j+1}|| + \frac{(j+1)(j+2)}{\mu} ||f_{j+2}|| \right) t^j
\end{aligned}$$

Zatem $\partial_\beta^2 \varphi_\xi(x) \geq \mu$ wtedy i tylko wtedy, gdy $g_N(|x|) \geq 1$. Oznaczmy współczynnik wielomianu $g_N - 1$ przy t^{d-j} przez $g_{N,j}$. Zauważmy, że wielomian $g_N - 1$ nie ma zer t takich, że $|t| > K_N$, gdzie

$$K_N = 2 \max_{1 \leq j \leq d} \left| \frac{g_{N,j}}{g_{N,0}} \right|^{1/j}$$

Ponieważ $g_{N,0} = N^2 f_{d*} > 0$, więc $g_N(t) - 1 > 0$ dla $|t| > K_N$ i w konsekwencji

$$\partial_\beta^2 \varphi_\xi(x) \geq \mu \quad \text{dla } |x| > K_N. \tag{3.13}$$

Zauważmy, że dla każdego $N \geq 1$,

$$K_N \leq K. \tag{3.14}$$

Faktycznie,

$$K_N = 2\max\left\{\frac{N^2||f_{d-1}|| + N\frac{2d}{\sqrt{\mu}}||f_d||}{N^2 f_{d*}},\right.$$
$$\left.\max_{1\le j\le d-2}\left(\frac{N^2||f_j|| + N\frac{2(j+2)}{\sqrt{\mu}}||f_{j+1}|| + \frac{(j+1)(j+2)}{\mu}||f_{j+2}||}{N^2 f_{d*}}\right)^{1/j}\right\} \le K_1 = K.$$

To daje (3.14).

Z (3.13) i (3.14) mamy

$$\partial^2_\beta \varphi_\xi(x) \ge \mu \quad \text{dla } |x| > K. \tag{3.15}$$

Aby zakończyć dowód, wystarczy pokazać, że

$$\partial^2_\beta \varphi_\xi(x) \ge \mu \quad \text{dla } |x| \le K. \tag{3.16}$$

Podobnie jak na początku dowodu dla $|x| \le K$, wobec (3.11), mamy

$$\partial^2_\beta \varphi_\xi(x) \ge \partial^2_\beta b(N(x-\xi))h(N),$$

gdzie h jest funkcją kwadratową postaci

$$h(N) = N^2 m - N\frac{2}{\sqrt{\mu}}\sum_{j=0}^{d-1}(j+1)||f_{j+1}||\cdot K^j - \frac{1}{\mu}\sum_{j=0}^{d-2}(j+1)(j+2)||f_{j+2}||\cdot K^j$$

Zatem analogicznie jak powyższej, dla dowolnej liczby N spełniającej (3.12) mamy $h(N) \ge 1$ i w konsekwencji (3.16). To kończy dowód. □

Uwaga 3.3.2. W Twierdzeniu 3.3.1 wykorzystaliśmy funkcję $b(Nx)$ do uwypuklenia wielomianu f. Podobne wyniki uzyskano w [9], gdzie użyto funkcji $b^N(x)$ dla $b(x) = \exp(|x|^2)$. Z Faktu 1.4.4 zauważmy, że $b(Nx)$ odpowiada wzięciu funkcji $b^{N^2}(x)$. Widać to wyraźnie, gdy weźmiemy funkcje $b(x) = \exp(|x|^2)$. Co więcej, $b(Nx)$ jest $N^2\mu$-silnie wypukła, a $b^N(x)$ jest $N\mu$-silnie wypukła, pod warunkiem, że funkcja b jest logarytmicznie μ-silnie wypukła, a 1 to minimalna wartość funkcji b.

Oznaczmy $K = \frac{||f||+1}{f_{d*}}$. Niech

$$A = \sum_{j=0}^{d}||f_j||K^j\sum_{j=0}^{d-2}(j+1)(j+2)||f_{j+1}||K^{j-2} + \left(\sum_{j=0}^{d-1}(j+1)||f_{j+1}||K^j\right)^2.$$

Podobnie jak w twierdzenie 3.3.1 dowodzimy następującego twierdzenia.

Twierdzenie 3.3.3. *Załóżmy, że $f_{d*} > 0$ oraz istnieje $m > 0$ takie, że zachodzi* (3.11). *Wówczas dla dowolnego*

$$N \geq \max\left\{\sqrt{\frac{A}{m\mu}+1}, \sqrt{\frac{||f||^2 d(2d-1)}{\mu}+1}\right\} \tag{3.17}$$

oraz dowolne $\xi \in \mathbb{R}^n$, funkcja $\varphi_{N,\xi} : \mathbb{R}^n \to \mathbb{R}$ określona wzorem $\varphi_{N,\xi}(x) = b(N(x-\xi))f(x)$ jest logarytmicznie μ-silnie wypukła w zbiorze $\mathbb{R}^n$.

Dowód. Ponieważ b jest funkcją logarytmicznie μ-silnie wypukłą, to dla dowolnych $x \in \mathbb{R}^n$ i $\beta \in S^{n-1}$, mamy

$$\begin{aligned} \partial_\beta^2 \ln \varphi_{N,\xi}(x) = N^2 & \frac{b(N(x-\xi))\partial_\beta^2 b(N(x-\xi)) - \left(\partial_\beta b(N(x-\xi))\right)^2}{b^2(N(x-\xi))} \\ & + \frac{f(x)\partial_\beta^2 f(x) - \left(\partial_\beta f(x)\right)^2}{f^2(x)} \geq N^2\mu - \left|\frac{f(x)\partial^2\beta f(x) - \left(\partial_\beta f(x)\right)^2}{f^2(x)}\right|. \end{aligned} \tag{3.18}$$

Z Faktu 1.6.1,

$$\begin{aligned} \left|\frac{f(x)\partial_\beta^2 f(x) - \left(\partial_\beta f(x)\right)^2}{f^2(x)}\right| \leq & \frac{1}{f^2(x)}\sum_{j=0}^{d}||f_j|||x|^j \sum_{j=0}^{d-2}(j+1)(j+2)||f_{j+2}|||x|^{d-2} \\ & + \frac{1}{f^2(x)}\left(\sum_{j=0}^{d-1}(j+1)||f_{j+1}|||x|^j\right)^2. \end{aligned}$$

Wówczas z (3.11),

$$\left|\frac{f(x)\partial^2 f(x) - \left(\partial_\beta f(x)\right)^2}{f^2(x)}\right| \leq \frac{A}{m} \quad \text{dla } |x| \leq K,$$

a ponieważ $K > 1$, więc z faktu 1.6.1 mamy

$$\left|\frac{f(x)\partial^2 f(x) - \left(\partial_\beta f(x)\right)^2}{f^2(x)}\right| \leq \frac{||f||^2 d(2d-1)}{\left(f_{d*}|x| - ||f||\right)^2} \leq ||f||^2 d(2d-1) \quad \text{dla } |x| \geq K.$$

W konsekwencji, z (3.18), dla N spełniającego (3.17) otrzymujemy $\partial_\beta^2 \ln \varphi_{N,\xi}(x) \geq \mu$. To daje, że funkcja $\varphi_{N,\xi}$ jest logarytmicznie μ-silnie wypukła w $\mathbb{R}^n$. □

Uwaga 3.3.4. W pracy [9], Example 3.6 pokazano, że bez założenia $f_{d*} > 0$, tezy powyższych dwóch twierdzeń nie zachodzą. Dokładniej, dla $b(x) = \exp(|x|^2|)$ oraz $f(x, y, z) = (y^2 + z^2 + 1)\,[(x^2 - 1)^2 + (yz + 1)^2 + y^2]$ tezy nie zachodzą.

Uwaga 3.3.5. Możliwe jest poprawienie oszacowań N w twierdzeniach 3.3.1 i 3.3.3 za pomocą funkcji b logarytmicznie silnie wypukłych ℓ-tego rzędu. Nie możemy jednak pominąć założeń dotyczących funkcji f.

Rozdział 4

Odwzorowanie $\xi \mapsto \operatorname{argmin}_{|x|\le R}(N \cdot b(x-\xi)+f(x))$

Niech $f : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^k$, $k \geq 2$. Weźmy $r \in \mathbb{R}$ i załóżmy, że zbiór

$$X_{f\le r} := \{x \in \mathbb{R}^n : f(x) \le r\}$$

jest ograniczony i niepusty. Niech $R_{f\le r}$ będzie promieniem zbioru $X_{f\le r}$, t.j.,

$$R_{f\le r} = \sup\{|x| : x \in X_{f\le r}\}.$$

Weźmy dowolne $R > R_{f\le r}$ i niech

$$B_R = \{x \in \mathbb{R}^n : |x| \le R\}.$$

Ponieważ $X_{f\le r} \neq \emptyset$, mamy $R_{f\le r} \geq 0$ a więc, $R > 0$.

Niech $D_R \in \mathbb{R}$ będzie liczbą dodatnią taką, że

$$|\partial_\beta^2 f(x)| \le D_R \quad \text{dla} \quad x \in B_R, \quad \beta \in S^{n-1}. \tag{4.1}$$

Przypomnijmy, że $S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$.

Dla uproszczenia zapisu niech $b : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^k$, $k \geq 2$, która jest μ-silnie wypukła, $\mu > 0$, taka, że

$$0 = \operatorname{argmin}_{\mathbb{R}^n} b \quad \text{i} \quad b(0) = 1. \tag{4.2}$$

Niech N będzie liczbą taką, że

$$N \geq \frac{D_R}{\mu} + 1. \tag{4.3}$$

Zgodnie z uwagą 2.2.3 dla dowolnego $\xi \in B_R$ funkcja

$$\phi_{N,\xi}(x) = Nb(x - \xi) + f(x)$$

jest μ-silnie wypukła na zbiorze B_R. Niech $\kappa_N : B_R \to B_R$ będzie odwzorowaniem określonym wzorem

$$\kappa_N(\xi) = \operatorname{argmin}_{B_R} \phi_{N,\xi} \in B_R \quad \text{dla } \xi \in B_R. \tag{4.4}$$

4.1 Własności odwzorowania κ_N

Fakt 4.1.1. $\kappa_N(X_{f\leq r}) \subset X_{f\leq r}$.

Dowód. Weźmy dowolne $\xi \in B_{f\leq r}$ oraz niech $x = \kappa_N(\xi)$. Wówczas $\phi_{N,\xi}(x) \leq \phi_{N,\xi}(\xi)$ i w konsekwencji $Nb(x - \xi) + f(x) \leq Nb(0) + f(\xi)$. Ponieważ, z (4.2), $b(0) \leq b(x - \xi)$, więc mamy $f(x) \leq f(\xi)$, co daje tezę i kończy dowód. □

Lemat 4.1.2. *Funkcja $\kappa_N|_{X_{f\leq r}}$ jest klasy $\mathscr{C}^{k-1}$.*

Dowód. Weźmy dowolne $\xi \in X_{f\leq r}$. Zauważmy, że punkt $x = \kappa_N(\xi)$ spełnia następując układ równań

$$\nabla \phi_{N,\xi}(x) = 0. \tag{4.5}$$

Istotnie, zgodnie z wyborem R mamy $\min\{f(x) : |x| = R\} > r$, więc, $X_{f\leq r} \subset \operatorname{Int} B_R$ a z faktu 4.1.1, $\kappa_N(\xi) \in \operatorname{Int} B_R$. Zatem x spełnia (4.5). Ponieważ jakobian (w odniesieniu do x) układu równań jest równy hessjanowi $\phi_{N,\xi}$ więc Jacobian jest niezerowy w punkcie x, ponieważ macierz Hessego funkcji $\phi_{N,\xi}$ ma tylko dodatnie wartości własne (patrz fakt 1.3.1). Stąd i z twierdzenia o funkcji uwikłanej dostajemy tezę. □

Lemat 4.1.3. *Odwzorowanie*

$$\kappa_N|_{X_{f\leq r}} : X_{f\leq r} \to \kappa_N(X_{f\leq r}) \tag{4.6}$$

jest dyfeomorfizmem klasy $\mathscr{C}^{k-1}$.

Dowód. Z lematu 4.1.2, odwzorowanie (4.6) jest klasy $\mathscr{C}^{k-1}$. Weźmy dowolne $\xi \in X_{f\leq r}$ i niech $x = \kappa_N(\xi)$. Zgodnie z dowodem lematu 4.1.2 z (4.5), mamy

$$N\nabla b(x-\xi) + \nabla f(x) = 0, \tag{4.7}$$

gdzie $\nabla b(x-\xi)$ jest gradientem funkcji $b(x-\xi)$ względem x. Zatem z faktu 1.2.3, punkt ξ jest jednoznaczenie wyznaczony przez x. W konsekwencji odwzorowanie (4.6) jest bijekcją , a więc jest homeomorfizmem, ponieważ zbiór $X_{f,R}$ jest zwarty a odwzorowanie κ_N – ciągłe. Do zakończenia dowodu wystarczy pokazać, że odwzorowanie $\left(\kappa_N|_{X_{f\leq r}}\right)^{-1} : \kappa_N(X_{f\leq r}) \to X_{f\leq r}$ jest klasy $\mathscr{C}^{k-1}$. W tym celu wystarczy pokazać, że jakobian względem ξ układu równań (4.7) jest niezerowy dla dowolnego $(x,\xi) \in X_{f\leq r} \times \kappa_N(X_{f\leq r})$ takiego, że $\xi = \kappa_N(x)$. Wynika to z faktu, że jakobian w odniesieniu do ξ tego układu równań jest równy hessjanowi funkcji $-Nb(x-\xi)$, więc nie zeruje się nigdzie w zbiorze $X_{f\leq r}$. W konsekwencji $\left(\kappa_N|_{X_{f\leq r}}\right)^{-1}$ jest odwzorowaniem klasy $\mathscr{C}^{k-1}$, co kónczy dowód. $\square$

Przypomnijmy, że Σ_f oznacza zbiór punktów krytycznych funkcji f.

Lemat 4.1.4. *Zbiorem punktów stałych odwzorowania* $\kappa_N|_{X_{f\leq r}}$ *jest* $\Sigma_f \cap X_{f\leq r}$.

Dowód. Niech $\xi \in X_{f\leq r}$ będzie punktem stałym odwzorowania $\kappa_N|_{X_{f\leq r}}$. Wówczas, analogicznie jak w dowodzie lematu 4.1.3, mamy $\nabla\phi_{N,\xi}(\xi) = 0$, t.j.

$$N\nabla b(0) + \nabla f(\xi) = 0.$$

Ponieważ $0 = \operatorname{argmin}_{\mathbb{R}^n} b$, to mamy $\nabla b(0) = 0$, więc $\nabla f(\xi) = 0$ i $\xi \in \Sigma_f$. Niech teraz $\xi \in X_{f\leq r}$ będzie punktem krytycznym funkcji f oraz niech $x = \kappa_N(\xi)$. Wówczas x jest jedynym punktem w zbiorze $X_{f\leq r}$ dla którego $\nabla\phi_{N,\xi}(x) = 0$. Ponieważ $\nabla\phi_{N,\xi}(\xi) = 0$, więc mamy $\xi = x$ i ξ jest punktem stałym odwzorowania $\kappa_N|_{X_{f\leq r}}$. $\square$

Wniosek 4.1.5. *Jeśli* $\xi \in X_{f\le r} \setminus \Sigma_f$ *oraz* $x = \kappa_N(\xi)$, *wtedy*

$$\partial_{x-\xi} f(\xi + t(x-\xi)) = \langle \nabla f(\xi + t(x-\xi)), x - \xi \rangle < 0 \quad \text{dla } t \in [0,1], \tag{4.8}$$

$x \notin \Sigma_f$ *i funkcja*

$$f_{\xi,x} : [0,1] \ni t \mapsto f(\xi + t(x-\xi)) \in \mathbb{R}$$

jest ściśle malejąca. W szczególności ciąg $f(\kappa_N^\nu(\xi))$, $\nu \in \mathbb{N}$, *jest ściśle malejący, a ciąg* $\kappa_N^\nu(\xi)$, $\nu = 0, 1, \ldots$, *jest różnowartościowy oraz*

$$\kappa_N^\nu(\xi) \notin \Sigma_f \quad \text{dla } \nu = 0, 1, \ldots.$$

Dowód. Ponieważ $\xi \notin \Sigma_f$, więc zgodnie z lematem 4.1.4 mamy $x \neq \xi$. Ponieważ punkt x jest jedynym punktem w zbiorze $X_{f\le r}$, w którym $\phi_{N,\xi}$ przyjmuje minimalną wartość w zbiorze $X_{f\le r}$, więć (4.7) zachodzi t.j., $N\nabla b(x-\xi) + \nabla f(x) = 0$. Ponieważ $x - \xi \neq 0$, więc mamy $\nabla b(x-\xi) \neq 0$. Zatem

$$\nabla f(x) \neq 0. \tag{4.9}$$

Ponadto, funkcja

$$[0,1] \ni t \mapsto \phi_{N,\xi}(\xi + t(x-\xi)) \in \mathbb{R}$$

jest silnie wypukła z minimalną wartością w punkcie 1, więc jest ściśle malejąca i jej pochodna nie ma zer w zbiorze $(0,1)$. W koncekwencji, dla $\beta = \frac{x-\xi}{|x-\xi|}$ mamy

$$\partial_\beta \phi_{N,\xi}(\xi + t(x-\xi)) < 0 \quad \text{dla } t \in (0,1).$$

Z drugiej strony $\partial_\beta b(t(x-\xi)) > 0$ dla $t \in (0,1]$ oraz

$$\partial_\beta \phi_{N,\xi}(x) = N\partial_\beta b(x-\xi) + \partial_\beta f(x),$$

więc $\partial_\beta f(\xi + t(x-\xi)) < 0$ i w konsekwencji (4.8) zachodzi. W szczególności $x \notin \Sigma_f$. Co więcej, funkcja $f_{\xi,x}$ jest ściśle malejąca. Druga część wniosku jest łatwą konsekwencją powyższego, co kończy dowód. □

Uwaga 4.1.6. Ponieważ funkcja $\phi_{N,\xi}$ jest μ-silnie wypukła dla dowolnego $\xi \in X_{f\le r}$, więc

$$\frac{\mu}{2}|\xi - \kappa_N(\xi)|^2 \le f(\xi) - f(\kappa_N(\xi)) \le C|\xi - \kappa_N(\xi)|$$

dla pewnej stałej $C > 0$.

Istotnie, nierówność po prawej stronie wynika z tego, że funkcji f jest lipschitzowska na zbiorze $X_{f\leq r}$. Ponieważ $\nabla\phi_{N,\xi}(\kappa_N(\xi)) = 0$ oraz $b(0) \leq b(x-\xi)$, więc zgodnie z definicją silnie wypukłej $\phi_{N,\xi}$ otrzymujemy $Nb(0) + f(\xi) \geq Nb(\kappa_N(\xi) - \xi) + f(\kappa_N(\xi)) + \frac{\mu}{2}|\xi - \kappa_N(\xi)|^2$, co daje tezę uwagi.

4.2 Iteracje odwzorowywania κ_N

Wykorzystwając pomysł z [9, Section 7] otrzymujemy następujący algorytm zbliżeniowy dla funkcji klasy $\mathscr{C}^2$ na zbiorach wypukłych (por [17]).

Twierdzenie 4.2.1. *Jeśli funkcja $f : \mathbb{R}^n \to \mathbb{R}$ jest semialgebraiczną klasy $\mathscr{C}^2$ spełniającą (4.1) oraz N spełnia (4.3), wówczas dla dowolnego $\xi \in X_{f\leq r}$:*

(a) *szereg*

$$\sum_{\nu=0}^{\infty} \operatorname{dist}(\kappa_N^\nu(\xi), f^{-1}(f(\kappa_N^{\nu+1}(\xi)))) \tag{4.10}$$

jest zbieżny.

(b) *istnieje stała $C > 0$ taka, że*

$$|\kappa_N^{\nu+1}(\xi) - \kappa_N^\nu(\xi)| \leq C \operatorname{dist}(\kappa_N^\nu(\xi), f^{-1}(f(\kappa_N^{\nu+1}(\xi)))), \quad \nu = 0, 1, \ldots \tag{4.11}$$

w szczególności szereg

$$\sum_{\nu=0}^{\infty} |\kappa_N^{\nu+1}(\xi) - \kappa_N^\nu(\xi)| \tag{4.12}$$

jest zbieżny.

(c) *punkt graniczny $\lim_{\nu\to\infty} \kappa_N^\nu(\xi)$ istnieje i należy do zbioru Σ_f.*

(d) *szereg $\sum_{\nu=0}^{\infty} \nabla f(\kappa_N^\nu(\xi))$ jest zbieżny.*

(e) *krzywa $\gamma_\xi : [0, +\infty) \to X_{f\leq r}$ określona wzorem*

$$\gamma_\xi(t) = \kappa_N^\nu(\xi) + (t - \nu)(\kappa_N^{\nu+1}(\xi) - \kappa_N^\nu(\xi)) \quad \text{dla } t \in [\nu, \nu + 1) \tag{4.13}$$

ma skończoną długość równą (4.12), a funkcja $f \circ \gamma_\xi : [0, +\infty) \to \mathbb{R}$ jest malejąca. Jeśli dodatkowo $\xi \notin \Sigma_f$, to funkcja $f \circ \gamma_\xi$ jest ściśle malejąca.

Dowód. Weźmy dowolne $\xi \in X_{f \le r}$. Część (e) wynika bezpośrednio z (b) i wniosku 4.1.5. Wystarczy więc udowodnić (a), (b), (c) i (d).

Niech $\xi_0 = \xi$ i $\xi_{\nu+1} = \kappa_N^{\nu}(\xi_0)$ dla $\nu = 0, 1, \ldots$. Wówczas $\xi_{\nu+1} = \kappa_N(\xi_\nu)$ dla $\nu = 0, 1, \ldots$.

Wykorzystując monotoniczność ciągu $f(\xi_\nu)$ (patrz wniosek 4.1.5) i zasadę porównania (patrz [9, Lemma 7.7]) otrzymujemy, że szereg (4.10) jest zbieżny (patrz dowód [9, Theorem 7.5]). To daje (a).

Udowodnimy część (b). Niech $a_\nu \in f^{-1}(f(\xi_\nu))$, $\nu = 1, 2, \ldots$, będzie takie, że

$$\mathrm{dist}(\xi_\nu, f^{-1}(f(\xi_{\nu+1})) = |\xi_\nu - a_{\nu+1}|.$$

Wtedy z definicji ξ_ν, mamy

$$Nb(\xi_{\nu+1} - \xi_\nu) + f(\xi_{\nu+1}) \le Nb(a_{\nu+1} - \xi_\nu) + f(a_{\nu+1}).$$

Ponieważ $f(a_{\nu+1}) = f(\xi_{\nu+1}) > 0$, więc mamy

$$b(\xi_{\nu+1} - \xi_\nu) \le b(a_{\nu+1} - \xi_\nu). \tag{4.14}$$

Zgodnie ze zbieżnością szeregu (4.10) mamy $\lim_{\nu\to\infty}(a_{\nu+1} - \xi_\nu) = 0$, a w konsekwencji, $\lim_{\nu\to\infty}(\xi_{\nu+1} - \xi_\nu) = 0$, ponieważ punkt $0 \in \mathbb{R}^n$ jest jedynym punktem, w którym funkcja b przyjmuje minimalną wartość.

Weźmy rozwinięcie Taylora funkcji b w punkcie $0 \in \mathbb{R}^n$ (przypomnijmy, że $\nabla b(0) = 0$),

$$b(x) = b(0) + \frac{1}{2} x^T H_b(0) x + R_3(x),$$

gdzie $H_b(0)$ jest macierzą Hessego funkcji b w punkcie 0 oraz $|R_3(x)| \le M|x|^3$ w pewnym otoczeniu U punktu 0 dla pewnej stałej $M > 0$. Wobec poprzedniego, można założyć, że $a_{\nu+1} - \xi_\nu \in U$ oraz $\xi_{\nu+1} - \xi_\nu \in U$ dla $\nu = 0.1.\ldots$. Wtedy z (4.14), mamy

$$\begin{aligned}(\xi_{\nu+1} - \xi_\nu)^T H_b(0)(\xi_{\nu+1} - \xi_\nu) - 2M|\xi_{\nu+1} - \xi_\nu|^3 \\ \le (a_{\nu+1} - \xi_\nu)^T H_b(0)(a_{\nu+1} - \xi_\nu) + 2M|a_{\nu+1} - \xi_\nu|^3.\end{aligned} \tag{4.15}$$

Niech $\lambda_1 < \ldots < \lambda_k$ będą wszystkimi wartościami własnymi macierzy $H_b(0)$ i niech $X_1, \ldots X_k \subset \mathbb{R}^n$ będą podprzestrzeniami niezmienniczymi tej macierzy odpowiednio dla wartości własnych $\lambda_1, \ldots, \lambda_k$. Ponieważ macierz $H_b(0)$

jest symetryczna i dodatnio określona, więc $0 < \lambda_1 < \ldots < \lambda_k$, przestrzenie $X_1, \ldots, X_k$ są parami prostopadłe i $\mathbb{R}^n$ jest sumą prostą tych podprzestrzeni. W konsekwejcji łatwo dostajemy, że dla dowolnego $y \in \mathbb{R}^n$ mamy

$$\lambda_1 |y|^2 \leq y^T H_b(0) y \leq \lambda_k |y|^2. \tag{4.16}$$

Istotnie, weźmy dowolne $y \in \mathbb{R}^n$. Wówczas $y = y_1 + \cdots + y_k$, gdzie $y_j \in X_j$ dla $j = 1, \ldots, k$. Ponieważ przestrzenie X_j są parami prostopadłe, to mamy

$$\begin{aligned} y^T H_b(0) y &= (y_1 + \cdots + y_k)^T H_b(0)(y_1 + \cdots + y_k) \\ &= (y_1 + \cdots + y_k)^T (\lambda_1 y_1 + \cdots + \lambda_k y_k) = \lambda_1 y_1^T y_1 + \cdots + \lambda_k y_k^T y_k \\ &= \lambda_1 |y_1|^2 + \cdots + \lambda_k |y_k|^2 \end{aligned}$$

Ponieważ $|y|^2 = |y_1|^2 + \cdots + |y_k|^2$, więc z powyższego łatwo otrzymujemy (4.16)

Z (4.14) i (4.16) dostajemy

$$|\xi_{\nu+1} - \xi_\nu|^2 \leq C^2 |a_{\nu+1} - \xi_\nu|^2$$

dla pewnej stałej $C > 0$. Stąd $|\xi_{\nu+1} - \xi_\nu| \leq C|a_{\nu+1} - \xi_\nu|$ co daje (4.11) i razem z (a) daje (b).

Zgodnie ze zbieżnością szeregu $\sum_{\nu=0}^{\infty} |\kappa_N^{\nu+1}(\xi) - \kappa_N^{\nu}(\xi)|$ otrzymujemy, że ciąg $\xi_\nu = \kappa_N^\nu(\xi)$ dąży do pewnego ξ_*. Aby udowodnić, że $\xi_* \in \Sigma_f$, zauważmy, że (analogicznie jak w dowodzie lematu 4.1.3) mamy (4.7), t.j.,

$$N \nabla b(\xi_{\nu+1} - \xi_\nu) + \nabla f(\xi_{\nu+1}) = 0 \quad \text{dla } \nu = 0, 1, \ldots.$$

Ponieważ $\nabla b(0) = 0$ i ∇b jest odwzorowaniem lipshitzowskim na zbiorze $X_{f \leq r}$, więc istnieje taka stałą $L > 0$, że $|\nabla b(\xi_{\nu+1} - \xi_\nu)| \leq L|\xi_{\nu+1} - \xi_\nu|$ dla dowolnego ν, więc

$$|\nabla f(\xi_{\nu+1})| \leq NL|\xi_{\nu+1} - \xi_\nu|.$$

Stąd, przez zbieżność szeregu (4.12), otrzymujemy zbieżność szeregu $\sum_{\nu=0}^{\infty} \nabla f(\xi_{\nu+1})$, co daje (d). Co więcej, ciągłość gradientu ∇f i warunek konieczny zbieżności szeregów daje $\nabla f(\xi_*) = \lim_{\nu \to \infty} \nabla f(\xi_{\nu+1}) = 0$. Co daje $\xi_* \in \Sigma_f$, czyli zachodzi (c). To kończy dowód. □

Uwaga 4.2.2. W dowodzie twierdzenia 4.2.1 pokazaliśmy m.in., że jeśli ∇b jest odwzorowaniem Lipschitza w $X_{f \leq r}$ ze stałą $L > 0$, to $|\xi_{\nu+1} - \xi_\nu|$ można oszacować od dołu w następujący sposób

$$|\xi_{\nu+1} - \xi_\nu| \geq \frac{|\nabla f(\xi_{\nu+1})|}{LN}.$$

4.3 Zbieżność jednostajna ciągu κ_N^ν

Niech $f : \mathbb{R}^n \to \mathbb{R}$ jest funkcją semialgebraiczną klasy $\mathscr{C}^2$ spełniającą (4.1) oraz niech N będzie liczbą spełniającą (4.3). Niech $\kappa_{N,*} : X_{f\leq r} \to \Sigma_f \cap X_{f\leq r}$ będzie odwzorowaniem określonym wzorem

$$\kappa_{N,*}(\xi) = \lim_{\nu\to\infty} \kappa_N^\nu(\xi).$$

Z twierdzenia 4.2.1 (c) widzimy, że odwzorowanie to jest dobrze określone.

Twierdzenie 4.3.1. *Niech $0 \in \operatorname{Int} X_{f\leq r}$ oraz niech $f(0)$ będzie najmniejszą wartością funkcji f. Wtedy istnieje $f(0) < \delta < r$ taki, że ciąg κ_N^ν jest zbieżny jednostajnie do $\kappa_{N,*}$ w zbiorze $U = X_{f\leq\delta}$. W szczególności odwzorowanie*

$$\kappa_{N,*}|_U : U \to U \cap \Sigma_f$$

jest ciągłe oraz $\kappa_{N,*}(\xi) = \xi$ wtedy i tylko wtedy, gdy $\xi \in U \cap \Sigma_f$. W konsekwencji $\kappa_{N,*}|_U$ jest retrakcją, a zbiór $U \cap \Sigma_f$ jest retraktem zbioru U.

Dowód. Oznaczmy $\omega_\nu = \kappa_N^\nu$ i $\omega_* = \kappa_{N,*}$ w zbiorze $X_{f\leq r}$. Ponieważ dowolna funkcja semialgebraiczna ma skończoną liczbę wartości krytycznych, więc jeśli zmniejszymy r, możemy założyć, że $f(0)$ jest jedyną wartością krytyczną f w zbiorze $X_{f\leq r}$. Wtedy $f(\omega_*(\xi)) = f(0)$ dla $\xi \in X_{f\leq r}$ i $f^{-1}(f(0)) = f^{-1}(f(0)) \cap X_{f\leq r} = \Sigma_f \cap X_{f\leq r}$. Oznaczmy $Z = f^{-1}(f(0))$. Oczywiście Z jest zbiorem zwartym.

Potrzebujemy następującego lematu

Lemat 4.3.2. *Przy założeniu twierdzenia 4.3.1 dla dowolnego otoczeniu $W \subset \mathbb{R}^n$ zbioru Z istnieje $c > 0$ taki, że*

$$X_{f\leq f(0)+c} = \{x \in \mathbb{R}^n : |f(x) - f(0)| < c\} \subset W.$$

Co więcej, jeśli $\omega_{\nu_0}(\xi) \in X_{f\leq f(0)+c}$ to $\omega_\nu(\xi) \in X_{f\leq f(0)+c}$ dla dowolnego $\nu \geq \nu_0$.

Pierwsza część lematu 4.3.2 jest oczywista. Dodatkowa część wynika bezpośrednio z twierdzenia 4.2.1 (e).

Z nierówności Łojasiewicza z gradientem (patrz [13, 15]), mamy

$$|f(x) - f(0)|^\varrho \leq C|\nabla f(x)| \tag{Ł1}$$

w pewnym otoczeniu $W \subset \mathbb{R}^n$ zbioru Z dla pewnych stałych $0 < \varrho < 1$ i $C > 0$. Zauważmy, że nierówność Łojasiewicza z gradientem (Ł1) została udowodniona w otoczeniu punktu. Ponieważ zbiór Z jest zwarty, więc łatwo uzyskamy tę nierówność wokół tego zbioru.

Niech C, ϱ i W będą takie jak w (Ł1). Będziemy potrzebować następującego lematu.

Lemat 4.3.3. *Weźmy dowolne $\xi_0 \in X_{f \leq r} \setminus \Sigma_f$ oraz niech $\xi_\nu = \omega_\nu(\xi_0)$ dla $\nu = 1, 2, \ldots$. Niech γ_{ξ_0} będzie krzywą określoną wzorem (4.13) dla $\xi = \xi_0$. Przy założeniach twierdzenia 4.3.1, dla dowolnego $\varepsilon > 0$ istnieje $f(0) < \delta < r$ takie, że dla dowolnego $\xi_0 \in X_{f \leq \delta}$ długość krzywej γ_{ξ_0} nie przekracza ε.*

Dowód lematu 4.3.3. Z twierdzenia 4.2.1 (e) mamy, że krzywa γ_{ξ_0} ma skończoną długość równą $\sum_{\nu=0}^{\infty} |\xi_{\nu+1} - \xi_\nu|$.

Z lematu 4.3.2 istnieje $c > 0$ takie, że $f(0) + 2c \leq r$ oraz

$$\{x \in \mathbb{R}^n : (f(x) - f(0))^{1-\varrho} < 2c\} \subset W.$$

Niech

$$U = \{x \in \mathbb{R}^n : (f(x) - f(0))^{1-\varrho} < c\}.$$

Oczywiście $f(0)$ jest jedyną wartością krytyczną funkcji $f|_W$.

Weźmy dowolne rozwiązanie $\gamma : [0, \beta) \to W \setminus Z$ maksymalne (prawostronne) układu równań

$$x' = -\frac{\nabla f(x)}{|\nabla f(x)|} \qquad \text{w } W \setminus Z \tag{4.17}$$

takie, że

$$\gamma(0) \in U. \tag{4.18}$$

Z (Ł1) otzymujemy *nierówność Kurdyki Łojasiewicza* (por., [10, Proposition 1]):

$$|\nabla(f - f(0))^{1-\varrho}(x)| \geq (1 - \varrho)C \quad \text{dla } x \in W \setminus Z.$$

Bez straty ogólności rozważań możemy założyć, że $(1 - \varrho)C \leq 1$. Stąd wynika, że

$$((f - f(0))^{1-\varrho} \circ \gamma)' = -|\nabla(f - f(0))^{1-\varrho}) \circ \gamma| \leq -(1 - \varrho)C \quad \text{w } [0, \beta)$$

(por. dowód [10, Theorem 1]). W konsekwencji $(f - f(0))^{1-\varrho} \circ \gamma$ i $f \circ \gamma$ są funkcjami malejącymi oraz dla każdego $0 < s < \beta$,

$$\begin{aligned}(f - f(0))^{1-\varrho}(\gamma(0)) - (f - f(0))^{1-\varrho}(\gamma(s)) = \\ = -s((f - f(0))^{1-\varrho} \circ \gamma)'(t) \geq s(1-\varrho)C \geq s\end{aligned}$$

dla pewnego $0 < t < s$. Ponieważ s jest równe długości $\gamma|_{[0,s]}$, więc mamy

$$\operatorname{length} \gamma|_{[0,s]} \leq (f - f(0))^{1-\varrho}(\gamma(0)) - (f - f(0))^{1-\varrho}(\gamma(s)), \tag{4.19}$$

gdzie $\operatorname{length} \gamma$ oznacza długość krzywej γ. Z powyższego otrzymujemy, że długość $\gamma|_{[0,\beta)}$ nie przekracza $(f-f(0))^{1-\varrho}(\gamma(0))$. Ponieważ funkcja $(f-f(0))^{1-\varrho}\circ\gamma$ jest malejąca, przy założeniu (4.18) otrzymujemy, że trajektoria $\gamma|_{[0,\beta)}$ nie może wyjść ze zbioru U i w konsekwencji musi mieć punkt graniczny w zbiorze Z. To daje, że dowolne maksymalne rozwiązanie w prawo $\gamma : [0, \beta) \to U \setminus Z$ układu równań (4.17) z warunkiem początkowym (4.18) działa w zbiorze $U \setminus f^{-1}(f(0))$ oraz, że przecina się dokładnie w jednym punkcie z każdą pozimicą $f^{-1}(y)$, $f(0) < y < f(\gamma(0))$.

Podobnie jak wyżej pokazujemy, że dla każdych $0 \leq s_1 < s_2 < \beta$,

$$\operatorname{length} \gamma|_{[s_1.s_2]} \leq (f - f(0))^{1-\varrho}(\gamma(s_1)) - (f - f(0))^{1-\varrho}(\gamma(s_2)). \tag{4.20}$$

Weźmy dowolne $\varepsilon > 0$. Bez utraty ogólności możemy założyć, że $\varepsilon < c$. Weźmy

$$\delta = f(0) + \left(\frac{\varepsilon}{C}\right)^{1/(1-\varrho)},$$

gdzie $C > 0$ jest takie jak w Twierdzeniu 4.2.1 (c). Załóżmy teraz, że $f(\xi_0) < \delta$. Wtedy $C(f(\xi_0) - f(0))^{1-\varrho} < \varepsilon$ oraz $C(f(\xi_\nu) - f(0))^{1-\varrho} < \varepsilon$ dla dowolnego ν (patrz Twierdzenie 4.2.1 (e)). Weźmy rozwiązanie $\gamma : [0, \beta) \to U \setminus f^{-1}(f(0))$ z (4.17) takie, że $\gamma(0) = \xi_\nu$. Z powyższego wynika że istnieje $s_1 > 0$ takie, że $f(\gamma(s_1)) = f(\xi_{\nu+1})$, a w myśl (4.20) i twierdzenia 4.2.1 (b),

$$|\xi_{\nu+1} - \xi_\nu| \leq C \operatorname{length} \gamma|_{[0,s_1]} \leq C \left[(f - f(0))^{1-\varrho}(\xi_\nu) - (f - f(0))^{1-\varrho}(\xi_{\nu+1})\right].$$

Ponieważ $\lim_{\nu\to\infty}(f(\xi_\nu) - f(0)) = 0$ oraz $1 - \varrho > 0$, więc

$$\sum_{k=\nu}^{\infty} \left[(f(\xi_k) - f(0))^{1-\varrho} - (f(\xi_{k+1}) - f(0))^{1-\varrho}\right] = (f(\xi_\nu) - f(0))^{1-\varrho}.$$

Z tego i twierdzenia 4.2.1 (c) otrzymujemy, że długość γ_{ξ_0} nie przekracza ε. Na tym kończy się dowód lematu 4.3.3. □

Teraz możemy kontynuować dowód twierdzenia 4.3.1. Możemy założyć, że (Ł1) jest spełnione w zbiorze $U = X_{f\leq\delta}$ dla pewnego $f(0) < \delta < r$ oraz teza lematu 4.3.3 zachodzi dla każdego $\xi \in U$.

Przy założeniu, że $f(0)$ jest minimalną wartością funkcji f mamy, że $f(\omega_*(\xi)) = f(0)$ dla $\xi \in U$, więc jest funkcją ciągłą. Stąd, lemat 4.1.2 daje, że

$$(f \circ \omega_\nu - f \circ \omega_*)^{1-\varrho} : U \mapsto \mathbb{R}, \quad \nu = 0, 1, \ldots,$$

jest ciągiem funkcji ciągłych i zgodnie z twierdzeniem 4.2.1 (e) – malejących. Oczywiście, $\lim_{\nu\to\infty} (f \circ \omega_\nu - f \circ \omega_*) = 0$ jest funkcją ciągłą. Tak więc, zgodnie z twierdzeniem Diniego, ciąg

$$(f \circ \omega_\nu - f \circ \omega_*)^{1-\varrho} \quad \text{dąży jednostajnie do 0 na zbiorze } U. \tag{4.21}$$

Wybierając δ, analogicznie jak w dowodzie lematu 4.3.3, dla dowolnego $\xi \in U$ otrzymujemy, że

$$\begin{aligned} |\omega_\nu(\xi) - \omega_*(\xi)| &\leq \sum_{k=\nu}^{\infty} |\omega_{\nu+1}(\xi) - \omega_\nu(\xi)| \\ &\leq \frac{1}{C(1-\varrho)} \sum_{k=\nu}^{\infty} \left[(f(\omega_k(\xi)) - f(\omega_*(\xi)))^{1-\varrho} - (f(\omega_{k+1}(\xi)) - f(\omega_*(\xi)))^{1-\varrho} \right] \\ &= \frac{1}{C(1-\varrho)} (f(\omega_\nu(\xi)) - f(\omega_*(\xi)))^{1-\varrho}. \end{aligned} \tag{4.22}$$

To oraz (4.21) daje tezę twierdzenia 4.3.1. □

Z lematów 4.3.2 i 4.3.3 otrzymujemy natychmiast

Wniosek 4.3.4. *Przy założeniu twierdzenia 4.3.1 dla dowolnego $\varepsilon > 0$ istnieje $f(0) < \delta < r$ takie, że dla dowolnego $\xi \in X_{f\leq\delta}$,*

$$|\omega_\nu(\xi) - \omega_*(\xi)| < \varepsilon \quad \textit{dla dowolnego } \nu.$$

Rozdział 5

Odwzorowanie $\xi \mapsto \operatorname{argmin}_X(N_1(\xi)b(x-\xi)+f(x))$

Niech $b : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^k$, $k \geq 2$, logarytmicznie μ-silnie wypukłą, $k \geq 2$, $\mu > 0$. Załóżmy, że

$$0 = \operatorname{argmin}_{\mathbb{R}^n} b \quad \text{i} \quad b(0) = 1.$$

Niech $X \subset \mathbb{R}^n$ będzie zbiorem wypukłem i domkniętym. Niech $f : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^k$, $k \geq 2$, o wzroście wielomianowym drugiego rzędu.

Niech $\psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ będzie funkcją określoną wzorem

$$\psi(\xi, x) = N_1(\xi)b(x-\xi) + f(x), \quad (\xi, x) \in \mathbb{R}^n \times \mathbb{R}^n,$$

gdzie N_1 jest określone przez (3.7). Przyjmijmy, że $\psi_\xi(x) = \psi(\xi, x)$. Zgodnie z wnioskiem 3.2.3, dla każdego $\xi \in \mathbb{R}^n$ funkcja ψ_ξ jest μ-silnie wypukła na zbiorze X. Możemy więc określić odwzorowanie $\kappa : X \to X$, przez

$$\kappa(\xi) = \operatorname{argmin}_X \psi_\xi \in X \quad \text{dla } \xi \in X.$$

5.1 Własności odwzorowania κ

Pokażemy, że odwzorowanie κ jest dobrze określone, ponadto pokażemy, że $\kappa(\xi)$ jest jedynym dolnym punkem krytycznym funkcji ψ_ξ na zbiorze X. Przypomnijmy, że punkt $a \in X$ jest *dolnym punktem krytycznym funkcji* różniczkowalnej g na zbiorze X, jeśli

$$\langle \nabla g(a), x - a \rangle \geq 0 \quad \text{dla } x \in X \ \text{ w pewnym otoczeniu punktu } a. \tag{5.1}$$

Zbiór dolnych punktów krytycznych funkcji g na zbiorze X oznaczymy przez $\Sigma_{X,g}$.

Fakt 5.1.1. *Dla dowolnego $\xi \in X$ punkt $\kappa(\xi)$ jest jedynym dolnym punktem krytycznym funkcji ψ_ξ na zbiorze X. W konsekwencji $\kappa(\xi)$ jest jedynym punktem krytycznym funkcji ψ_ξ, jeśli $\kappa(\xi) \in \operatorname{Int} X$.*

Dowód. Weźmy dowolne $\xi \in X$ i niech $x = \kappa(\xi)$. Załóżmy, że x nie jest dolnym punktem krytycznym funkcji ψ_ξ. Wtedy istnieje punkt $y \in X$ taki, że $\langle \nabla \psi_\xi(x), y - x \rangle < 0$. Ponieważ X jest zbiorem wypukłym, więc funkcja

$$h_{x,y} : [0, 1] \ni t \mapsto \psi_\xi(x + t(y - x)) \in \mathbb{R}$$

jest dobrze określona oraz $h'_{x,y}(0) = \partial_{y-x}\psi_\xi(x) = \langle \nabla \psi_\xi(x), y - x \rangle < 0$. W konsekwencji istnieje punkt $y_0 = x + t_0(y - x) \in X$, $t_0 \in (0, 1)$, taki, że $\psi_\xi(y_0) < \psi_\xi(x)$. To jest sprzeczne z faktem, że $\psi_\xi(x)$ jest najmniejszą wartością funkcji ψ_ξ na zbiorze X. Zatem x jest dolnym punktem krytycznym funkcji ψ_ξ.

Załóżmy, że istnieje inny dolny punkt krytyczny $y \in X$ funkcji ψ_ξ. Wówczas $\langle \nabla \psi_\xi(x), y - x \rangle \geq 0$ i $\langle \nabla \psi_\xi(y), y - x \rangle \leq 0$ a więc $h'_{x,y}(0) \geq 0 \geq h'_{x,y}(1)$. To jest niemożliwe, ponieważ funkcja $h'_{x,y}$ jest dodatnia jako pochodna funkcji silnie wypukłej $h_{x,y}$. □

Fakt 5.1.2. *Funkcja κ jest ciągła. Ponadto funkcja κ jest klasy $\mathscr{C}^{k-1}$ w otoczeniu dowolnego punktu $\xi \in X$ takiego, że $\kappa(\xi) \in \operatorname{Int} X$. W szczególności, jeśli $X = \mathbb{R}^n$, funkcja κ jest klasy $\mathscr{C}^{k-1}$.*

Dowód. Załóżmy, że κ nie jest odwzorowaniem ciągłym. Wówczas istnieje ciąg $\xi_\nu \in X$, który dąży do jakiegoś $\xi_0 \in X$ oraz $\kappa(\xi_\nu)$ dąży do jakiegoś $x_0 \neq \kappa(\xi_0)$. Zatem $\psi_{\xi_\nu}(\kappa(\xi_\nu)) \leq \psi_{\xi_\nu}(\kappa(\xi_0))$. Stąd z ciągości ψ mamy $\psi_{\xi_0}(x_0) \leq \psi_{\xi_0}(\kappa(\xi_0))$. To jest sprzeczne z faktem, że $\kappa(\xi_0) = \operatorname{argmin}_X \psi_{\xi_0}$.

Weźmy dowolne $\xi \in X$ oraz niech $x = \kappa_N(\xi)$. Jeśli $x \in \operatorname{Int} X$, to x jest jedynym punktem krytycznym funkcji ψ_ξ w zbiorze X, więc

$$\nabla \psi_\xi(x) = 0. \tag{5.2}$$

Ponieważ jakobian (w odniesieniu do x) układu równań (5.2) jest równy Hessian funkcji ψ_ξ, , to jakobian jest różny od zera w x, ponieważ macierz Hessego ma tylko dodatnie wartości własne. Zatem twierdzenie o funkcji uwikłanej daje dodatkową część faktu. $\square$

Fakt 5.1.3. *Dla dowolnego $\xi \in X$ mamy $f(\kappa(\xi)) \leq f(\xi)$. Ponadto równość zachodzi wtedy i tylko wtydy, gdy $\xi = \kappa(\xi) \in X \cap \Sigma_{X,f}$.*

Dowód. Ponieważ $\psi_\xi(\kappa(\xi)) \leq \psi_\xi(\xi)$, mamy

$$N_1(\xi)b(\kappa(\xi) - \xi) + f(\kappa(\xi)) \leq N_1(\xi)b(0) + f(\xi).$$

Z założenia na funkcji b mamy $b(0) \leq b(\kappa(\xi) - \xi)$ a więc $f(\kappa(\xi)) \leq f(\xi)$ i otrzymujemy pierwszą część faktu. Co więcej, jeśli $f(\kappa(\xi)) = f(\xi)$ wtedy $b(\kappa(\xi) - \xi) = b(0)$ oraz $\xi = \kappa(\xi)$. Oczywiście, jeśli $\xi = \kappa(\xi)$, to $f(\kappa(\xi)) = f(\xi)$.

Załóżmy, że $\xi = \kappa(\xi)$. Wówczas dla dowolnego $x \in X \setminus \{\xi\}$ funkcja $t \mapsto \psi_\xi(\xi + t(x - \xi))$ jest rosnąca na zbiorze $[0,1]$. W konsekwencji $\langle x - \xi, \nabla\psi_\xi(\xi + t(x - \xi)\rangle = \partial_{x-\xi}\psi_\xi(\xi + t(x - \xi) \geq 0$ dla $t \in [0,1]$ więc $\langle x - \xi, \nabla\psi_\xi(\xi)\rangle \geq 0$. Z założenia $\xi = \kappa(\xi)$ mamy $\nabla\psi_\xi(\xi) = \nabla f(\xi)$ zatem $\langle x - \xi, \nabla f(\xi)\rangle \geq 0$ dla $x \in X$ oraz $\xi \in \Sigma_{X,f}$.

Załóżmy, że $\xi \in \Sigma_{X,f}$. Wówczas $\langle \nabla f(\xi), x - \xi\rangle \geq 0$ dla $x \in X$ w otoczeniu punktu ξ. Ponieważ $\nabla\psi_\xi(\xi) = N_1(\xi)\nabla b(0) + \nabla f(x) = \nabla f(\xi)$ wtedy punkt ξ jest dolnym krytycznym funkcji ψ_ξ i w konsekwencji $\xi = \operatorname{argmin}_X \psi_\xi = \kappa(\xi)$. $\square$

Z faktu 5.1.3 mamy

Wniosek 5.1.4. *Dla dowolnego $\xi \in X$ ciąg $f(\kappa^\nu(\xi))$, $\nu \in \mathbb{N}$ ma granicę.*

Wniosek 5.1.5. *Zbiór punktów stałych odwzorowania κ jest równy $\Sigma_{X,f}$.*

Uwaga 5.1.6. Można zdefiniować odwzorowanie $\kappa(\xi) = \operatorname{argmin}_X \psi_\xi$ dla funkcji $\psi_\xi(x) = Nb(\xi)b(x - \xi) + f(x)$, gdzie N spełnia (3.8). Łatwo pokazać, że tezy faktów 5.1.1, 5.1.2, 5.1.3 i wniosków 5.1.4, 5.1.5 również zachodzą dla tego odwzorowania.

5.2 Odwzorowanie $\kappa_N : X \ni \xi \mapsto \operatorname{argmin}_X b(N(x-\xi))f(x)$

Przy dodatkowym założeniu, że funkcja f jest wielomianem spełniającym założenia twierdzenia 3.3.1 (lub twierdzenia 3.3.3) dla funkcji

$$\varphi_{N,\xi}(x) = b(N(x-\xi))f(x)$$

można zdefiniować odwzorowanie

$$\kappa_N(\xi) = \operatorname{argmin}_X \varphi_{N,\xi}, \tag{5.3}$$

gdzie N spełnia (3.12) lub (3.17) . Łatwo pokazać, że tezy faktów 5.1.1, 5.1.2, 5.1.3 i wniosków 5.1.4, 5.1.5 również zachodzą dla odwzorowania κ_N .

Z twierdzenia 3.3.1 (lub z twierdzenia 3.3.3) otrzymujemy

Wniosek 5.2.1. *Jeśli $f \in \mathbb{R}[x]$ jest wielomianem takim, że $f_{d*} > 0$ i $f(x) \geq m$ dla $x \in \mathbb{R}^n$ i pewną stałą $m > 0$, wówczas dla dowolnego N spełniającego* (3.12) *(lub* (3.17)*), odwzorowanie $\kappa_N : \mathbb{R}^n \ni \xi \mapsto \operatorname{argmin}_{\mathbb{R}^n} \varphi_{N,\xi} \in \mathbb{R}^n$ jest dyfeomorfizmem klasy $\mathscr{C}^{k-1}$*

Dowód. Z faktu 5.1.2 wynika, że odwzorowanie κ_N jest klasy $\mathscr{C}^{k-1}$. Z definicji κ_N wynika, że dla dowolnego $\xi \in \mathbb{R}^n$ punkt $x = \kappa_N(\xi)$ spełnia równanie $\nabla\varphi_{N,\xi}(x) = 0$, więc $Nf(x)\nabla b(N(x-\xi)) + b(N(x-\xi))\nabla f(x) = 0$ i w konsekwencji

$$N\nabla(\ln b)(N(x-\xi)) + \nabla(\ln f(x)) = 0. \tag{5.4}$$

Ponieważ b jest funkcją logarytmicznie silnie wypukłą, z faktu 1.2.3 mamy, że odwzorowanie $\nabla(\ln b) : \mathbb{R}^n \to \mathbb{R}^n$ jest bijekcją i przez (5.4) odwzorowanie κ_N jest również bijekcją $\mathbb{R}^n \to \mathbb{R}^n$. Jakobian odwzorowania $\xi \mapsto N\nabla(\ln b)(N(x-\xi)) + \nabla(\ln f(x))$ jest równy $N^{2n}h(\ln b)(N(x-\xi))$, gdzie $h(\ln b)$ jest wyznacznikiem macierzy Hessego $\ln b$. Tak więc, jakobian jest niezerowy i w konsekwencji, zgodnie z twierdzeniem o funkcji uwikładnej, κ_N^{-1} jest funkcją klasy $\mathscr{C}^{k-1}$. $\square$

Rozdział 6

Odwzorowanie $\xi \mapsto \operatorname{argmin}_{|x|\leq R} b^N(x-\xi)f(x)$

Niech $f : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^2$. Weźmy $r > 0$ i załóżmy, że zbiór

$$X_{f\leq r} = \{x \in \mathbb{R}^n : f(x) \leq r\}$$

jest zbiorem ograniczonym i niepustym. Niech $R_{f\leq r}$ będzie promieniem zbioru $X_{f\leq r}$, t.j.,

$$R_{f\leq r} = \sup\{|x| : x \in X_{f,r}\}.$$

Weźmy dowolne $R > R_{f\leq r}$ oraz kulę domkniętą

$$B_R = \{x \in \mathbb{R}^n : |x| \leq R\}.$$

Ponieważ $X_{f\leq r} \neq \emptyset$, mamy $R_{f\leq r} \geq 0$ a więc $R > 0$.

Niech $m_R, D_R \in \mathbb{R}$ będą liczbami dodatnimi takimi, że

$$(6.1)\quad f(x) \geq m_R, \quad |\partial_\beta f(x)| \leq D_R, \quad |\partial_\beta^2 f(x)| \leq D_R \qquad \text{dla } x \in B_R,\ \ \beta \in \mathbb{R}^n,\ |\beta| = 1.$$

Nięch $b : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją klasy $\mathscr{C}^2$ która jest μ-silnie wypukła, $\mu > 0$, przyjmującą tylko wartości dodatnie taką, że

$$0 = \operatorname{argmin}_{\mathbb{R}^n} b(x) \quad \text{i} \quad b(0) = 1$$

oraz niech

$$S'_{b,f,r} = \max\{b(x-\xi) : x, \xi \in B_R\}.$$

Niech N będzie liczbą całkowitą taką, że

$$N \geq N(\mu, S'_{b,R}, m_R, D_R). \tag{6.2}$$

Z wniosku 2.2.2 dla dowolnego $\xi \in B_R$ funkcja

$$\lambda_{N,\xi}(x) = b^N(x-\xi)f(x) \tag{6.3}$$

jest silnie wypukła na zbiorze B_R. Niech $\kappa_N : B_R \to B_R$ będzie odwzorowaniem określonym wzorem

$$\kappa_N(\xi) = \operatorname{argmin}_{B_R} \lambda_{N,\xi} \in B_R \quad \text{dla } \xi \in B_R. \tag{6.4}$$

6.1 Własności odwzorowania $\xi \mapsto \operatorname{argmin}_{|x|\leq R} b^N(x-\xi)f(x)$

Poniższy fakt i kolejne dwa lematy dowodzimy analogicznie jak fakt 4.1.1 i lematy 4.1.2 oraz 4.1.3. Ponieważ funkcja κ_N jest teraz wyznaczona przez inną funkcję niż w rozdziale 4, więc dla pełności pracy przytaczamy ich dowody.

Fakt 6.1.1. $\kappa_N(X_{f\leq r}) \subset X_{f\leq r}$.

Dowód. Weźmy dowolne $\xi \in B_{f,r}$. Zauważmy, że $f(\kappa_N(\xi)) \leq f(\xi)$ dla $\xi \in B_{f,r}$. Istotnie, niech $x = \kappa_N(\xi)$. Wtedy $\lambda_{N,\xi}(x) \leq \lambda_{N,\xi}(\xi)$ i w konsekwencji, $b^N(x-\xi)f(x) \leq b^N(0)f(\xi)$. Ponieważ $b(0) \leq b(x-\xi)$, to otrzymujemy $f(x) \leq f(\xi)$. Co daje tezę i kończy dowód. □

Lemat 6.1.2. *Funkcja $\kappa_N|_{X_{f\leq r}}$ jest klasy $\mathscr{C}^1$.*

Dowód. Weźmy dowolne $\xi \in X_{f,r}$. Zauważmy, że $x = \kappa_N(\xi)$ spełnia następujący układ równań

$$\nabla \lambda_{N,\xi}(x) = 0. \tag{6.5}$$

Istotnie, wybierając B_R mamy $\min\{f(x) : |x| = R\} > r$, więc , $X_{f,r} \subset \operatorname{Int} B_R$ oraz z faktu 6.1.1, $\kappa(\xi) \in \operatorname{Int} B_R$. Zatem x spełnia (6.5). Ponieważ jakobian (w odniesieniu do x) układu równań jest równy Hessianowi funkcji $\lambda_{N,\xi}$, więc Jacobian jest różny od zera w punkcie x, ponieważ macierz Hessego funkcji $\lambda_{N,\xi}$ ma tylko wartości własne dodatnie. Zatem twierdzenie o funkcji uwikładnej daje te, że □

Lemat 6.1.3. *Niech b będzie funkcją logarytmicznie μ-silnie wypułą oraz niech $N > N_{\exp}(\mu, m_R, D_R)$. Wówczaz odwzorowanie*

$$\kappa|_{X_{f\le r}} : X_{f\le r} \to \kappa(X_{f\le r}) \tag{6.6}$$

jest dyfeomorfizmem klasy $\mathscr{C}^1$.

Dowód. Weźmy dowolne $\xi \in X_{f,r}$ i niech $x = \kappa_N(\xi)$. Ponieważ $b(x-\xi) > 0$, więc przy oznaczeniach dowodu lematu 6.1.2 z (6.5) mamy

$$N\nabla b(x-\xi) f(x) + b(x-\xi)\nabla f(x) = 0, \tag{6.7}$$

gdzie $\nabla b(x-\xi)$ jest gradientem $b(x-\xi)$ względem x. Wtedy

$$\frac{1}{b(x-\xi)}\nabla b(x-\xi) + \frac{1}{Nf(x)}\nabla f(x) = 0. \tag{6.8}$$

Co więcej, z wniosku 1.4.6 wynika, że punkt ξ jest jednoznaczenie określony przez x. Wobec tego odwzorowanie (6.6) jest bijekcją i w konsekwencji jest homeomorfizmem, ponieważ $X_{f,R}$ jest zwarty oraz funkcja κ jest ciągła.

Do zakończenia dowodu, wystarczy pokazać, że odwzorowanie $\left(\kappa|_{X_{f\le r}}\right)^{-1} : \kappa(X_{f\le r}) \to X_{f\le r}$ jest klasy $\mathscr{C}^1$. W tym celu, w myśl twierdzenia o funkcji uwikłanej, wystarczy pokazać, że jakobian w odniesieniu do ξ układu równań (6.8) jest niezerowy dla dowolnego $(x,\xi) \in X_{f\le r} \times \kappa(X_{f\le r})$ takiego, że $\xi = \kappa(x)$. Wynika to z faktu, że jakobian w odniesieniu do ξ układu równań (6.8) jest równy Hessianowi $\ln(\lambda_{N,\xi})$, więc nie ma zera w dowolnym punkcie zbioru $X_{f\le r}$. W konsekwencji $\left(\kappa|_{X_{f\le r}}\right)^{-1}$ jest odwzorowaniem klasy $\mathscr{C}^1$, co kónczy dowód. □

Z lematu 1.4.8 otrzymujemy analogiczny lemat do lematu 6.1.3 dla funkcji silnie wypukłych. Niestety ta wersja nie jest tak efektywna jak lemat 6.1.3.

Lemat 6.1.4. *Niech b będzie funkcją silnie wypukłą. Wówczas istnieje N_0 takie, że dla dowolnego $N > N_0$, odwzorowanie*

$$\kappa_N|_{X_{f\leq r}} : X_{f\leq r} \to \kappa_N(X_{f\leq r}) \tag{6.9}$$

jest dyfeomorfizmem klasy $\mathscr{C}^1$.

Dowód. Niech $\varepsilon > 0$ oraz $\delta > 0$ będzie jak w lemacie 1.4.8. Wówczas istnieje N_1 takie, że dla dowolnego $N \geq N_1$ mamy

$$\frac{1}{Nf(x)}|\nabla f(x)| < \delta \quad \text{dla } x \in X_{f\leq r}$$

Wtedy dla $N_0 = \max\left\{N_1, N(\mu, S'_{b,R}, m_R, D_R)\right\}$, analogicznie jak w dowodzie lematu 6.1.3 (z lematu 1.4.8) otrzymujemy tezę. □

Przypomnijmy, że Σ_f oznacza zbiór punktów krytycznych funkcji f.

Lemat 6.1.5. *Zbiorem punktów stałych odwzorowania $\kappa|_{X_{f\leq r}}$ jest $\Sigma_f \cap X_{f\leq r}$.*

Dowód. Niech $\xi \in X_{f\leq r}$ będzie punktym stałym odwzorowania $\kappa|_{X_{f\leq r}}$. Wówczas, analogicznie jak w dowodzie lematu 6.1.3, otrzymujemy $\nabla\lambda_{N,\xi}(\xi) = 0$, t.j.,

$$N\nabla b(0)f(\xi) + b(0)\nabla f(\xi) = 0.$$

Ponieważ b przyjmuje wartość minimalna równa zero, więc mamy $\nabla b(0) = 0$, a więc $\nabla f(\xi) = 0$ i $\xi \in \Sigma_f$. Niech teraz $\xi \in X_{f\leq r}$ będzie punktem krytycznym funkcji f i niech $x = \kappa(\xi)$. Wówczas x jest jedynym punktem w zbiorze $X_{f\leq r}$ dla którego $\nabla\lambda_{N,\xi}(x) = 0$. Ponieważ $\nabla\lambda_{N,\xi}(\xi) = 0$, więc otrzymujemy $\xi = x$ oraz ξ jest punktem stałym odwzorowania $\kappa|_{X_{f\leq r}}$. □

6.2 Iteracje odwzorowania $\xi \mapsto \operatorname{argmin}_{|x|\leq R} b^N(x-\xi)f(x)$

Wniosek 6.2.1. *Niech $\xi \in X_{f\leq r} \setminus \Sigma_f$ i niech $x = \kappa_N(\xi)$. Wtedy*

$$\partial_{x-\xi} f(\xi + t(x-\xi)) = \langle \nabla f(\xi + t(x-\xi)), x-\xi\rangle < 0 \quad \textit{dla } t \in [0,1], \tag{6.10}$$

$x \notin \Sigma_f$ oraz funkcja

$$f_{\xi x} : [0,1] \ni t \mapsto f(\xi + t(x-\xi)) \in \mathbb{R}$$

jest ściśle malejąca. W szczególności ciąg $f(\kappa_N^\nu(\xi))$, $\nu \in \mathbb{N}$, jest ściśle malejący a ciąg $\kappa_N^\nu(\xi)$, $\nu = 0, 1, \ldots$, jest różnowartościowy oraz

$$\kappa_N^\nu(\xi) \notin \Sigma_f \quad \text{dla } \nu = 0, 1, \ldots.$$

Dowód. Ponieważ $\xi \notin \Sigma_f$, więc zgodnie z lematem 6.1.5 mamy $x \neq \xi$. Ponieważ punkt x jest jedynym punktem zbioru $X_{f\le r}$, w którym $\lambda_{N,\xi}$ przyjmuje minimalną wartość w zbiorze $X_{f\le r}$, więc $N\nabla b(x-\xi)f(x)+b(x-\xi)\nabla f(x) = 0$. Ponieważ $x - \xi \neq 0$, więc mamy $\nabla b(x-\xi) \neq 0$ a więc

$$\nabla f(x) \neq 0. \tag{6.11}$$

Ponadto funkcja

$$[0,1] \ni t \mapsto \lambda_{N,\xi}(\xi + t(x-\xi)) \in \mathbb{R}$$

jest silnie wypukła z minimalną wartością 1, więc jest ściśle malejąca i jej pochodna nie ma zer w zbiorze $(0,1)$. W konsekwencji, dla $\beta = \frac{x-\xi}{|x-\xi|}$ mamy

$$\partial_\beta \lambda_{N,\xi}(\xi + t(x-\xi)) < 0 \quad \text{dla } t \in (0,1).$$

Z drugiej strony $\partial_\beta b(t(x-\xi)) > 0$ dla $t \in (0,1]$ i

$$\partial_\beta \lambda_{N,\xi}(x) = Nb^{n-1}(x-\xi)\partial_\beta b(x-\xi)f(x) + b^N(x-\xi)\partial_\beta f(x),$$

więc $\partial_\beta f(\xi+t(x-\xi)) < 0$ i w konsekwencji (6.10) zachodzi. W szczególności $x \notin \Sigma_f$. Co więcej, funkcja $f_{\xi,r}$ jest ściśle malejąca. Ostatnia część tego twierdzenia jest łatwą konsekwencją powyższego. To kończy dowód □

Uwaga 6.2.2. Jeśli funkcja $\lambda_{N,\xi}$ jest μ-silnie wypukła, to dla dowolnego $\xi \in X_{f\le r}$,

$$f(\xi) - f(\kappa_N(\xi)) \ge \frac{\mu}{2}|\xi - \kappa_N(\xi)|^2.$$

Jeśli dodatkowo $\lambda_{N,\xi}$ jest μ-wykładniczo silnie wypukła, to dla dowolnego $\xi \in X_{f\le r}$,

$$\frac{f(\xi)}{f(\kappa_N(\xi))} \ge \exp\left(\frac{\mu}{2}|\xi - \kappa_N(\xi)|^2\right).$$

Podobnie jak twierdzenie 4.2.1, korzystając pomysłu z [9, Section 7], dowodzimy następujące

Twierdzenie 6.2.3. *Niech $f : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją semialgebraiczną klasy $\mathscr{C}^2$ spełniającą* (6.1) *oraz N spełnia* (6.2). *Wówczas dla dowolnego $\xi \in X_{f\leq r}$.*

(a) *Punkt graniczny $\lim_{\nu\to\infty} \kappa_N^\nu(\xi)$ istnieje i należy do zbioru Σ_f.*

(b) *Szereg $\sum_{\nu=0}^{\infty} |\kappa_N^{\nu+1}(\xi) - \kappa_N^\nu(\xi)|$ jest zbieżny.*

W szczególności krzywa $\gamma_\xi : [0, +\infty) \to X_{f\leq r}$ określona wzorem

$$\gamma_\xi(t) = \kappa_N^\nu(\xi) + (t-\nu)(\kappa_N^{\nu+1}(\xi) - \kappa_N^\nu(\xi)) \quad \textit{dla } t \in [\nu, \nu+1)$$

ma skończoną długość, a funkcja $f \circ \gamma_\xi : [0, +\infty) \to \mathbb{R}$ jest malejąca. Jeśli dodatkowo $\xi \notin \Sigma_f$, to funkcja $f \circ \gamma_\xi$ jest ściśle malejąca.

Dowód. Weźmy dowolne $\xi \in X_{f\leq r}$. Druga część twierdzenia wynika bezpośrednio z (b) i wniosku 6.2.1, więc wystarczy udowodnić (a) i (b).

Niech $\xi_0 = \xi$ i $\xi_{\nu+1} = \kappa_N^\nu(\xi_0)$ dla $\nu = 0, 1, \ldots$. Wówczas $\xi_{\nu+1} = \kappa_N(\xi_\nu)$ dla $\nu = 0, 1, \ldots$.

Przytoczymy szkic rozumowania zastosowanego w [9] w przypadku $X = X_{f\leq r}$ i $\xi_0 \in X_{f\leq r}$. W [9, Theorem 7.5], twierdzenie zostało pokazane przy założeniu, że funkcja b ma postać $b(x) = 1 + |x|^2$. Oczywiście, funkcja b jest silnie wypukła. W tym przypadku mamy (Patrz [9, Lemma 7.1])

$$|\xi_{\nu+1} - \xi_\nu| = \mathrm{dist}(\xi_\nu, f^{-1}(f(\xi_{\nu+1}))). \quad \nu = 0, 1, \ldots \tag{6.12}$$

oraz ciąg $f(\xi_\nu)$ jest malejący (patrz [9, Lemma 7.2] i wniosek 6.2.1). Wykorzystując monotoniczności ciągu $f(\xi_\nu)$ i zasadę porównania (patrz [9, Lemma 7.7]) otrzymujemy, że szereg

$$\sum_{\nu=0}^{\infty} \mathrm{dist}(\xi_\nu, f^{-1}(f(\xi_{\nu+1}))) \tag{6.13}$$

jest zbieżny. Wtedy z (6.12), szereg

$$\sum_{\nu=0}^{\infty} |\xi_{\nu+1} - \xi_\nu| \tag{6.14}$$

jest zbieżny i w konsekwencji, ciąg ξ_ν dąży do pewnego ξ_*.

Aby pokazać, że $\xi_* \in \Sigma_f$, analogicznie zauważmy, że jak w dowodzie lematu 6.1.3 otrzymujemy(6.8), t.j.,

$$N\nabla b(\xi_{\nu+1} - \xi_\nu) f(\xi_{\nu+1}) + b(\xi_{\nu+1} - \xi_\nu)\nabla f(\xi_{\nu+1}) = 0 \quad \text{dla } \nu = 0, 1, \ldots.$$

Ponieważ $\nabla b(0) = 0$ i ∇b jest odwzorowaniem lipschitzowskim na zbiorze $X_{f \leq r}$, więc istnieje $L > 0$ takie, że $|\nabla b(\xi_{\nu+1} - \xi_\nu)| \leq L|\xi_{\nu+1} - \xi_\nu|$ dla dowolnego ν, zatem

$$|\nabla f(\xi_{\nu+1})| \leq \frac{N f(\xi_{\nu+1})}{b(\xi_{\nu+1} - \xi_\nu)} L|\xi_{\nu+1} - \xi_\nu|.$$

Stąd, zgodnie z zbieżnością szeregu (6.14), otrzymujemy zbieżność szeregu $\sum_{\nu=0}^{\infty} \nabla f(\xi_{\nu+1})$. Ponadto z ciągłości gradientu ∇f i warunku koniecznego zbieżności szeregu dostajemy $\nabla f(\xi_*) = \lim_{\nu\to\infty} \nabla f(\xi_{\nu+1}) = 0$. To daje tezę w przypadku $b(x) = 1 + |x|^2$. Zauważmy, że dowód na to, że $\xi_* \in \Sigma_f$ różni się od tego w artykule [9]. Zostało to przeprowadzone bez żadnych założeń dla postaci funkcji b, więc udowodniliśmy, że zachodz (a), pod warunkiem, że (b) jest spełniony.

Wróćmy do dowodu twierdzenia 6.2.3. Wystarczy udowodnić część (b).

W dowodzie zbieżności szeregu (6.13) postać funkcji b nie była istotna, dowód polegał na zastosowaniu zasady porównania, semialgebraiczności funkcji f oraz monotoniczności ciągu $f(\xi_\nu)$. Stąd szereg (6.13) jest zbieżny. Biorąc zatem pod uwagę powyższe rozważania, wystarczy pokazać zbieżność szeregu (6.14). W tym celu wystarczy pokazać, że istnieje stała $C > 0$ taka, że

$$|\xi_{\nu+1} - \xi_\nu| \leq C \operatorname{dist}(\xi_\nu, f^{-1}(f(\xi_{\nu+1}))), \quad \nu = 0, 1, \ldots \tag{6.15}$$

Nierówność (6.15) dowodzimy analogicznie jak nierówność (4.11) w dowodzie twierdzenia 4.2.1. To kończy dowód twierdzenia 6.2.3. □

Uwaga 6.2.4. W dowodzie twierdzenia 6.2.3 pokazaliśmy między innymi, że jeśli ∇b jest odwzorowanie Lipschitza w zbiorze $X_{f \leq r}$ ze stałą $L > 0$, to $|\xi_{\nu+1} - \xi_\nu|$ można oszacować z dołu w następujący sposób

$$|\xi_{\nu+1} - \xi_\nu| \geq \frac{|\nabla f(\xi_{\nu+1})| b(\xi_{\nu+1} - \xi_\nu|)}{L N f(\xi_{\nu+1})}.$$

6.3 Odwzorowanie κ_N dla $b(x) = \exp|x|^2$

Z punktu widzenia IT ważne jest, aby wiedzieć, jak szybko ciąg κ_N^ν zbliża się do swojej granicy. Jednym z pojawiających się tutaj problemów jest to, czy ciąg zbiega w ustalonym kierunku, czyli czy część sferyczna ciągu (we współrzędnych biegunowych) ma granicę. Wydaje się, że jest to dość trudny problem i należy zastosować metody rozwiązania geadientowej hipotezy Rene Thoma zastosowane w [7]. Prowadzi to do dyskretnej hipotezy R. Thoma: Czy ciąg $\kappa_N^\nu/|\kappa_N^\nu|$ ma granicą, gdy $\nu \to \infty$. Tu od razu napotykamy na trudność. Podczas gdy w przypadku trajektorii pola gradientu obowiązuje właściwość Darboux, nie jest tak w przypadku dyskretnym. Na stosunkowo prostym przykładzie pokażemy, jakie są podobieństwa, a jakie różnice w przypadku trajektorii i ciągu.

Niech $f \in \mathbb{R}[x]$ będzie wielomianem postaci

$$f(x) = f_0 + f_k(x) + \cdots + f_d(x), \tag{6.16}$$

gdzie f_j będzie wielomianem jednorodnym stopnie j lub zerem dla $j = 0, k, \dots, d$, $k > 1$ oraz $f_k \neq 0$. Zakładamy, że

$$f(x) \geq 1 \quad \text{dla } x \in \mathbb{R}^n. \tag{6.17}$$

Przyjmiemy, że

$$b(x) = \exp|x|^2, \qquad x \in \mathbb{R}^n. \tag{6.18}$$

Fakt 6.3.1. *Funkcja b jest 2-logarytmicznie silnie wypukła w $\mathbb{R}^n$. Ponadto*

$$\nabla b^N(x) = 2Nb^N(x) \cdot x \quad \textit{dla } x \in \mathbb{R}^n.$$

Niech $S'_{b,R} = e^{4R^2}$ i niech

$$N \geq N(\nu, S'_{b,R}, 1, D_n(f, R)). \tag{6.19}$$

Zgodnie z wnioskiem 2.2.2 funkcja

$$\lambda_{N,\xi}(x) = f(x)\exp^{N|x-\xi|^2}, x \in \mathbb{R}^n, \xi \in X_{f\leq r},$$

jest silnie wypukła na pewnej kuli zawierającej zbiór $X_{f\leq r}$ w swoim wnętrzu oraz odwzorowanie κ_N określone wzorem (6.4) jest dobrze określone. Z faktu 4.1.1, analogicznie jak w dowodzie lematu 6.1.3, z (6.7) mamy

Fakt 6.3.2. *Odwzorowanie $\kappa_N : X_{f\le r} \to \kappa_N(X_{f\le r})$ jest odwrotnością odwzorowanie*

$$\kappa_N(X_{f\le r}) \ni x \mapsto x + \frac{1}{2Nf(x)}\nabla f(x) \in X_{f\le r}\,, \tag{6.20}$$

więc jest analityczne i semialgebraiczne, t.j., jest odwzorowanem Nasha.

Ponieważ $\frac{1}{2Nf(x)}\nabla f(x) = \nabla\left(\frac{1}{2N}\ln f(x)\right)$, więc dla $g(x) = \frac{1}{2N}\ln f(x)$ mamy

Uwaga 6.3.3. Macierz Jakobiego $J(\kappa_N)$ odwzorowaniu κ_N wyraża się wzorem

$$J(\kappa_N(\xi)) = (I + H(g)(\kappa_N(\xi)))^{-1}\,,$$

gdzie I jest macierzą jednostkową $m \times m$.

Z uwqagi 6.3.3 widziemy, że $J(\kappa_N(\xi))$ jest macierzą symetryczną. Więc mamy

Uwaga 6.3.4. Odwzorowanie $\kappa_N : X_{f,r} \to \kappa_N(X_{f,r})$ jest gradientem pewnej funkcji analitycznej $F : X_{f,r} \to \mathbb{R}$. Ponadto, $\xi = \kappa_N(\xi) + \nabla g(\kappa_N(\xi))$ oraz

$$\nabla\left(F(\xi) - \frac{|\xi|^2}{2}\right) = -\nabla g(\kappa_N(\xi)).$$

Ponieważ założyliśmy (6.17), więz z lematu 1.6.4 i twierdzenia 4.2.1 otrzymujemy

Wniosek 6.3.5. *Niech $R = K_f$. Załóżmu. że $f_{d*} = \min_{|x|=1} f_d(x) > 0$ i niech*

$$N > \max\{N_{\exp}(\mu, 1, D_n(f, \mathbb{K}(f)), N_{\exp,\infty}(\mu, f)\}\,.$$

Wowczas odwzorowanie $\kappa_N : \mathbb{R}^n \to \mathbb{R}^n$ jest dyfeomorfizmem klasy $\mathscr{C}^1$. Ponadto, dla dowolnego $\xi \in \mathbb{R}^n$ punkt graniczny $\lim_{\nu\to\infty}\kappa_N^\nu(\xi)$ istnieje i należy do zbioru Σ_f.

Niech $\omega_0 : X_{f\le r} \ni \xi \mapsto \xi \in X_{f\le r}$ będzie odwzorowaniem tożsamościowym oraz niech $\omega_\nu : X_{f\le r} \to X_{f\le r}$ będzi odwzorowaniem określonym wzorem

$$\omega_{\nu+1} = \kappa_N(\omega_\nu) \quad \text{dla } \nu \ge 0$$

Z faktu 4.1.1 mamy $\omega_\nu(\xi) \in X_{f\le r}$ dla dowolnego $\xi \in X_{f\le r}$ i $\nu = 1, 2, \ldots$, więc odwzorowania ω_ν są dobrze określone. Oczywiście $\omega_\nu = \kappa_N^\nu$ dla $\nu = 0, 1, \ldots$.

Z lematu 6.1.5 i wniosku 6.2.1 mamy

Fakt 6.3.6. *Ciąg $\omega_\nu(\xi_0)$ jest stały wtedy i tylko wtedy, gdy $\xi_0 \in X_{f\le r} \cap \Sigma_f$. Ponadto dla $\xi_0 \in X_{f\le r} \setminus \Sigma_f$ ciąg $\omega_\nu(\xi_0)$ jest rożnowartościowy oraz $\omega_\nu(\xi_0) \neq \omega_*(\xi_0)$ dla dowolnego ν.*

6.3.1 Pewne krzywe o własnościach podobnych do trajektorii pola gradientu

Weźmy dowolne $\xi_0 \in X_{f\le r} \setminus \Sigma_f$ oraz niech $\xi_\nu = \omega_\nu(\xi_0)$ dla $\nu = 1, 2, \ldots$. Niech $\gamma_{\xi_0} : [0, +\infty) \to X_{f\le r}$ będzie krzeywą określoną wzorem

$$\gamma_{\xi_0}(t) = \xi_\nu + (t-\nu)(\xi_{\nu+1} - \xi_\nu) \quad \text{dla } t \in [\nu, \nu+1),\ \nu \in \mathbb{N}. \tag{6.21}$$

Krzywa γ_{ξ_0} ma kilka własności podobnych do własności pola gradientowego. Mianowicie, krzywa γ_{ξ_0} ma następujące własności (por. twierdzenie 6.2.3 i fakt 6.3.2):

Fakt 6.3.7. (i) *Krzywa γ_{ξ_0} ma skończoną długość równą $\sum_{\nu=0}^{\infty} |\xi_{\nu+1} - \xi_\nu|$.*

(ii) *Funkcja $f \circ \gamma_{\xi_0} : [0, +\infty) \to \mathbb{R}$ jest ściśje malejąca (przypomnijmy, że założyliśmy, że $\xi_0 \notin \Sigma_f$).*

(iii) *Dla $t \in (\nu, \nu+1)$, $\nu = 0, 1, \ldots$ mamy*

$$\gamma'(t) = \xi_{\nu+1} - \xi_\nu = -\frac{1}{2Nf(\xi_{\nu+1})} \nabla f(\xi_{\nu+1}).$$

Warunek (iii) nie oznacza, że $\gamma'(t) = \frac{1}{2Nf(\gamma(t))} \nabla f(\gamma(t))$. Jest to jedna z trudności w badaniach ciągu ξ_ν, która nie występuje w badaniach trajektorii pola gradientowego.

Krzywe te mają jeszcze jedną własność, podobną do własności trajektorii pól gradientowych. Mianowicie, wykorzystując fakt 6.3.7, analogicznie jak lematu 4.3.3 dowodzimy następującego faktu.

Fakt 6.3.8. *Niech $0 \in \operatorname{Int} X_{f\le r}$ oraz niech $f(0)$ będzie wartością minimalną funkcji f. Wówczas dla dowolnego $\varepsilon > 0$ istnieje $f(0) < \delta < r$ takie, że dla dowolnego $\xi_0 \in X_{f\le\delta}$ (t.j., $f(\xi_0) < \delta$) długość krzywej γ_{ξ_0} nie przekracza ε.*

Z faktu 6.3.8, podobnie jak w dowodzie lematu 4.3.3, dostajemy

Wniosek 6.3.9. *Niech* $0 \in \operatorname{Int} X_{f\leq r}$ *i niech* $f(0)$ *będzie minimslną wartością funkcji* f. *Wówczaz dla dowolnego* $\varepsilon > 0$ *istnieje* $f(0) < \delta < r$ *takie, że dla dowolnego* $\xi \in X_{f\leq\delta}$ *(t.j.,* $f(\xi) < \delta$*)*

$$|\omega_\nu(\xi) - \omega_*(\xi)| < \varepsilon \quad \textit{dla dowolnego } \nu.$$

6.3.2 Zbieżność jednostajna ciągu ω_ν

Ciąg odwzorowań ω_ν ma pewną własność podobną do własności potoku pola gradientowego (por., [12, 14], patrz także podrozdział 6.3.1). Dopkładniej, zachodzi twierdzenie podobne do twierdzenia 4.3.1 dla ciągu ω_ν.

Twierdzenie 6.3.10. *Niech* $0 \in \operatorname{Int} X_{f\leq r}$ *oraz niech* $f(0)$ *będzie minimalną wartością funkcji* f. *Wówczas istnieje* $f(0) < \delta < r$ *takie, że ciąg* ω_ν *jest jednostajnie zbieżny do* ω_* *w zbiorze* $U = X_{f\leq\delta}$. *W szczególności odwzorowanie* $\omega_* : U \to U \cap \Sigma_f$ *jest ciągłe oraz* $\omega_*(\xi) = \xi$ *dla* $\xi \in U \cap \Sigma_f$, *t.j.,* ω_* *jest retrakcją, a zbiór* $U \cap \Sigma_f$ *jest retraktem zbioru* U.

Szkic dowodu. Dowód twierdzenia przebiega analogicznie jak dowód twierdzenia 4.3.1. Ponieważ ciąg odwzorowań ω_ν jest określony przez iteracje odwzorowania $\operatorname{argmin}_{X_{f\leq r}} \lambda_{N,\xi}$, gdzie $\lambda_{N,\xi}(x) = b^N(x-\xi)f(x)$ a ciąg κ_N^ν, występujący w twierdzeniu 4.3.1 – przy pomocy iteracji odwzorowania $\operatorname{argmin}_{X_{f\leq r}} \phi_{N,\xi}$, gdzie $\phi_{N,\xi}(x) = Nb(x-\xi) + f(x)$, więc podajemy szkic dowodu twierdzenia 6.3.10.

Niech C, ϱ będą takie jak w (Ł1). Załóżmy, że (Ł1) zachodzi w zbiorze

$$U = X_{f\leq\delta} = \{x \in \mathbb{R}^n : f(x) \leq \delta\}$$

dla pewnych $f(0) < \delta < r$ oraz teza faktu 6.3.8 zachodzi dla dowolnego $\xi \in U$.

Przy założeniu, że $f(0)$ jest minimalną wartością funkcji f mamy, że $f(\omega_*(\xi) = f(0)$ dla $\xi \in U$, więc jest to funkcja ciągła. Z lematu 6.1.2 widzimy, że

$$(f \circ \omega_\nu - f \circ \omega_*)^{1-\varrho} : U \mapsto \mathbb{R}$$

jest ciągiem funkcji ciągłych, a z twierdzenia 6.2.3, że jest ciągiem malejącym. Oczywiście, $\lim_{\nu\to\infty}(f \circ \omega_\nu - f \circ \omega_*) = 0$ jest funkcja ciągła. Tak więc, z twierdzenia Diniego, ciąg

$$(6.22) \qquad (f \circ \omega_\nu - f \circ \omega_*)^{1-\varrho} \quad \text{jednostajnie zbieżny do zera na zbieorze } U.$$

Analogicznie jak w dowodzie lematu 4.3.3, dla dowolnego $\xi \in U$ otrzymujemy, że

$$
\begin{aligned}
|\omega_\nu(\xi) - \omega_*(\xi)| &\leq \sum_{k=\nu}^{\infty} |\omega_{\nu+1}(\xi) - \omega_\nu(\xi)| \\
&\leq \frac{1}{C(1-\varrho)} \sum_{k=\nu}^{\infty} \left[(f(\omega_k(\xi)) - f(\omega_*(\xi)))^{1-\varrho} - (f(\omega_{k+1}(\xi)) - f(\omega_*(\xi)))^{1-\varrho} \right] \\
&= \frac{1}{C(1-\varrho)} (f(\omega_\nu(\xi)) - f(\omega_*(\xi)))^{1-\varrho}.
\end{aligned}
$$

To i (6.22) daje tezę. □

Uwaga 6.3.11. Bez założenia, że $f(0)$ jest najmniejszą wartością funkcji, teza twierdzenia 6.3.10 nie zachodzi. Mianowicie, jeśli zbiór $X_{f\leq r}$ jest spójny, a funkcja f ma co najmniej dwie wartości krytyczne w $X_{f\leq r}$, to łatwo otrzymujemy sprzeczność.

Uwaga 6.3.12. Można pokazać, że twierdzenie 6.3.10 jest prawdziwe, gdy ciąg ω_ν wyznaczymy przy pomocy dowolnej innej funkcji silnie wypukłej b, zamiast $\exp(|x|^2)$.

6.3.3 Uwagi o zbieżności ciągu $\xi_\nu = \omega_\nu(\xi_0)$

Niech $d = \deg f$.

Weźmy dowolne $\xi_0 \in X_{f\leq r}$. Wiemy, że ciąg $\xi_\nu = \omega_\nu(\xi_0) = \kappa_N^\nu(\xi_0)$, $\nu = 0, 1, \ldots$ dąży do pewnego $\xi_* \in \Sigma_f \cap X_{f\leq r}$. Oczywiście, ∇b jest odwzorowaniem Lipschitza w zbiorze $X_{f\leq r}$ z pewną stałą $L > 0$. Zatem z uwagi 6.2.4, wielkość skoku i ν-tej iteracji $|\xi_{\nu+1} - \xi_\nu|$ można oszacować od dołu w następujący sposób

$$
|\xi_{\nu+1} - \xi_\nu| \geq \frac{|\nabla f(\xi_{\nu+1})|}{LN}.
$$

Korzystając więc z efektywnej nierówności Łojasiewicza (patrz [10, Remark 4])

$$
|\nabla f(x)| \geq C_1 \operatorname{dist}(x, f^{-1}(f(\xi_*)))^{(d-1)(6d-9)^{n-1}}
$$

w otoczeniu zbioru $f^{-1}(f(\xi_*))$ dla pewnej dodatniej stałej C_1, otrzymujemy

$$
\operatorname{dist}(\xi_{\nu+1}, f^{-1}(f(\xi_*)))^{(d-1)(6d-9)^{n-1}} \leq \frac{LN}{C_1} |\xi_{\nu+1} - \xi_\nu| \tag{6.23}
$$

dla dostatecznie dużych ν. To daje oszacowanie odległości punktu $\xi_{\nu+1}$ od poziomicy $f^{-1}(f(\xi_*))$, do której dąży ciąg ξ_ν, wyrażony w terminach skoku $|\xi_{\nu+1} - \xi_\nu|$, który robimy w ν-tym kroku.

Ponieważ $\lambda_{N,\xi}$ jes funkcjąt μ-silnie wypukłą, to z Uwagi 6.2.2, mamy

$$\frac{\mu}{2}|\xi_\nu - \xi_{\nu+1}|^2 \leq f(\xi_\nu) - f(\xi_{\nu+1}) \leq C|\xi_\nu - \xi_{\nu+1}|$$

dla dowolnego ν i pewnej stałej $C > 0$. Z tego i (6.23) mamy

$$\operatorname{dist}(\xi_{\nu+1}, f^{-1}(f(\xi_*)))^{(d-1)(6d-9)^{n-1}} \leq \frac{LN}{C_1}\sqrt{\frac{2}{\mu}\left[f(\xi_\nu) - f(\xi_{\nu+1})\right]}. \tag{6.24}$$

Z [10, Corollary 7] zauważmy, że (4.22) zachodzi z $\varrho = 1 - \frac{1}{d(6d-3)^{2n-1}}$ oraz pewną stałą $C > 0$. Więc mamy

$$|\xi_\nu - \xi_*| \leq \frac{d(6d-3)^{2n-1}}{C}(f(\xi_\nu) - f(\xi_*))^{1/d(6d-3)^{2n-1}}$$

dla dostatecznie dużego ν.

6.3.4 Dalsze własności ciągu ω_ν

Weźmy dowolne $\xi \in X_{f \leq r}$.

Z twierdzenia 6.2.3, ciąg

$$\omega_\nu(\xi) \quad \text{ma punkt graniczny } \omega_*(\xi) \in \Sigma_f\,, \tag{6.25}$$

szereg

$$\sum_{\nu=0}^{\infty} |\omega_{\nu+1}(\xi) - \omega_\nu(\xi)| \quad \text{jest zbieżny} \tag{6.26}$$

oraz ciąg

$$f(\omega_\nu(\xi)) \quad \text{jest malejący.} \tag{6.27}$$

Z faktu 6.3.2 (lub z faktu 6.3.1), analogicznie jak w dowodzie lematu 6.1.3, mamy

$$\omega_{\nu+1}(\xi) - \omega_\nu(\xi) = -\frac{1}{2Nf(\omega_{\nu+1}(\xi))}\nabla f(\omega_{\nu+1}(\xi)), \quad \nu \in \mathbb{N}. \tag{6.28}$$

W szczególności, z (6.26), szereg

$$\sum_{\nu=0}^{\infty} |\nabla f(\omega_\nu(\xi))| \quad \text{jest zbieżny.} \tag{6.29}$$

Uwaga 6.3.13. Z nierowności Bochnak-Łojasiewicza (patrz [2]),

$$|f(x) - f(\omega_*(\xi))| \leq C|\nabla f(x)||x - \omega_*(\xi)| \tag{BŁ}$$

w otoczeniu punktu $\omega_*(\xi)$ w przestrzeniu $\mathbb{R}^n$ dla pewnej stałej dodatniej C, więc z (6.29) otrzymujemy, że szereg

$$\sum_{\nu=0}^{\infty} \frac{f(\omega_\nu(\xi)) - f(\omega_*(\xi))}{|\omega_\nu(\xi) - \omega_*(\xi)|} \quad \text{jest zbieżny,}$$

pod warunkiem, że $\xi \notin \Sigma_f$.

Uwaga 6.3.14. Z nierówności gradientowej Łojasiewicza (patrz [13, 15]),

$$|f(x) - f(\omega_*(\xi))|^\varrho \leq C|\nabla f(x)| \tag{Ł1}$$

w otoczeniu zbioru $f^{-1}(f(\omega_*(\xi)))$ w przestrzeniu $\mathbb{R}^n$ dla pewnych stałych $0 < \varrho < 1$ i $C > 0$, otrzymujemy, że szereg

$$\sum_{\nu=0}^{\infty} (f(\omega_\nu(\xi)) - f(\omega_*(\xi)))^\varrho \quad \text{jest zbieżny.} \tag{6.30}$$

Z *globalnej nierówności Łojasiewicza* :

$$|f(x) - f(y)| \geq C \left(\frac{\operatorname{dist}(x, f^{-1}(f(y)))}{1 + |x|^2} \right)^{d(6d-3)^{n-1}} \quad \text{dla } x \in \mathbb{R}^n, \tag{Ł2}$$

przy ustalonym y dla pewnej stałej dodatniej C oraz $d = \deg f$ (patrz [10, Corollary 10]), mamy

Fakt 6.3.15. *Dla dowolnego otoczenia $U \subset \mathbb{R}^n$ zbioru $f^{-1}(f(\omega_*(\xi)))$ istnieje $\varepsilon > 0$ takie, że*

$$\{x \in \mathbb{R}^n : |f(x) - f(\omega_*(\xi))| < \varepsilon\} \subset U.$$

Co więcej, jeśli $f(\omega_{\nu_0}(\xi)) - f(\omega_*(\xi)) < \varepsilon$, to $f(\omega_\nu(\xi)) - f(\omega_*(\xi)) < \varepsilon$ oraz $\omega_\nu(\xi) \in U$ dla dowolnego $\nu \geq \nu_0$.

Uwaga 6.3.16. Niech $\xi \in X_{f\leq r}$ Weżmy dowolne $\varepsilon > 0$. Jeśli N spełnia (6.19) i dodatkowo

$$N \geq \frac{d\sqrt{n}}{2\varepsilon}\|f - f_0\|, \tag{6.31}$$

wtedy istnieje ν_0 takie, że dla dowolnego $\nu \geq \nu_0$,

$$|\omega_{\nu+1}(\xi) - \omega_\nu(\xi)| \leq \varepsilon|\omega_{\nu+1}(\xi)|.$$

Istotnie, z (6.28) i wniosek 1.6.3 istnieje ν_0 takie, że dla dowolnego $\nu \geq \nu_0$,

$$|\omega_{\nu+1}(\xi) - \omega_\nu(\xi)| \leq \frac{d\sqrt{n}}{2Nf(\xi_{\nu+1})}\|f - f_0\| \cdot |\omega_{\nu+1}(\xi)| \leq \frac{d\sqrt{n}}{2N}\|f - f_0\| \cdot |\omega_{\nu+1}(\xi)|.$$

Stąd, (6.31) daje tezę uwagi.

Uwaga 6.3.17. Z uwagi 6.2.2, wynika,że istnieje $\mu > 0$ takie, że

$$f(\omega_\nu(\xi)) - f(\omega_{\nu+1}(\xi)) \geq \mu|\omega_\nu(\xi) - \omega_{\nu+1}(\xi)|^2 \quad \text{dla dowolnego } \nu. \tag{6.32}$$

Przy dodatkowym założeniu, że punkt $0 \in \mathbb{R}^n$ jest izolowaną osobliwością funkcji f, to znaczy $\nabla f(0) = 0$ oraz $\nabla f(x) \neq 0$ w pewnym sąsiedztwie punktu $0 \in \mathbb{R}^n$, istnieją dodatnie stałe C, α takie, że

$$|\nabla f(x)| \geq C|x|^\alpha \qquad \text{w otoczeniu zera.} \tag{6.33}$$

Najmniejszy wykładnik α jest nazywany wykładnikiem *Łojasiewicza gradientu w punkcie* $0 \in \mathbb{R}^n$ i jest oznaczony przez $\mathcal{L}_0(\nabla f)$. Wiadomo, że $\mathcal{L}_0(\nabla f) \leq (d-1)(6d-9)^{n-1}$, gdzie $d = \deg f$ (patrz [10, uwaga 4]) i (6.33) zachodzi z $\alpha = \mathcal{L}_0(\nabla f)$. Wówczas (6.29) cokazuje, że szybkość zbieżności ciągu ξ_ν jest dość duża. Mianowicie mamy następujący fakt.

Weżmy dowolne $\xi_0 \in X_{f\leq r} \setminus \Sigma_f$ oraz niech $\xi_\nu = \omega_\nu(\xi_0)$ for $\nu = 0, 1, \ldots,$. Załóżmy, że $\omega_*(\xi_0) = 0$. Wówczas z (6.29) i (6.33) mamy

Fakt 6.3.18. *Jeśli punkt* $0 \in \mathbb{R}^n$ *jest izolowaną osobliwością funkcji* f, *to szereg*

$$\sum_{\nu=0}^{\infty} |\xi_\nu|^\alpha \qquad \textit{jest zbieżny,}$$

gdzie $\alpha = (d-1)(6d-9)^{n-1}$ *i* $d = \deg f$.

6.3.5 Część sferyczna ciągu $\kappa_N^\nu(\xi)$

W związku z wynikami podpunktu 6.3.1, pojawia się problem zbieżności ciągu części sferycznych ciągu $\omega_\nu = \kappa_N^\nu(\xi)$. Jest to przeniesienie problemu Rene Thoma dla trajektorii pola gradientowego (rozwiązanego w [7]) do przypadku dyskretnego. Rozpatrzymy ten problem przy założeniu, że $\omega_\nu(\xi) \to 0$ gdy $\nu \to \infty$ oraz kilku dodatkowych założeniach.

Przy dość mocnych założeniach otrzymujemy, że granica sferycznej części ciągu $\xi_\nu = \omega_\nu(\xi_0)$ istnieje. Mianowicie zachodzi następujący fakt.

Fakt 6.3.19. *Załóżmy, że $f_k(\theta) > 0$ dla $\theta \in S^{n-1}$. Wówczas istnieje następująca granica*

$$\lim_{\nu \to \infty} \frac{1}{|\xi_\nu|} \xi_\nu. \tag{6.34}$$

Co więcej, ciąg $|\xi_\nu|$ jest ściśle malejący od pewnego miejsca.

Dowód. Napiszmy f we współrzędnych biegunowych:

$$f(r\theta) = f_0 + r^k f_k(\theta) + \cdots + r^d f_d(\theta),$$

gdzie $r > 0$ and $\theta \in S^{n-1}$. Wówczas

$$\partial_r f(r\theta) = kr^{k-1} f_k(\theta) + \cdots + dr^{d-1} f_d(\theta), \tag{6.35}$$

$$\nabla' f(r\theta) = r^k \nabla' f_k(\theta) + \cdots + r^d \nabla' f_d(\theta).$$

oraz

$$\nabla f(r\theta) = \partial_r f(r\theta)\theta + \nabla' f(r\theta)$$

Więc, z założeni, że $f_k(\theta) > 0$ dla $\theta \in S^{n-1}$, istnieje $r_0 > 0$,takie, że

$$\frac{|\nabla' f(r\theta)|}{r} \leq C_1 r^{k-1} \leq C_2 \partial_r f(r\theta) \leq C_3 |\nabla f(r\theta)| \quad \text{dla } 0 < r < r_0 \tag{6.36}$$

i pewnych stałych dodatnich C_1, C_2, C_3.

Weźmy krzywą $\gamma = \gamma_{\xi_0}$ określoną wzorem (6.21). Z faktu 6.3.7 (ii) mamy, że funkcja $f \circ \gamma$ jest ściśle malejąca, więc mamy

$$\gamma(t) \neq 0 \quad \text{dla } t \in [0, +\infty),$$

oraz możemy zapisać γ we współrzędnych biegunowych $\gamma(t) = r_\gamma(t)\theta_\gamma(t)$, $r_\gamma(t) = |\gamma(t)| > 0$ amd $\theta_\gamma(t) \in S^{n-1}$. Wtedy

$$\gamma'(t) = r'_\gamma(t)\theta_\gamma(t) + r_\gamma(t)\theta'_\gamma(t) \quad \text{dla } t \in (\nu, \nu+1), \quad \nu = 0, 1, \dots,$$

oraz $\langle \theta_\gamma(t), \theta'_\gamma(t)\rangle = 0$ dla $t \in [0, +\infty) \setminus \mathbb{Z}$. Z drugiej strony, z faktu 6.3.7 (iii),

$$\gamma'(t) = \xi_{\nu+1} - \xi_\nu = -\frac{1}{2Nf(\xi_{\nu+1})}\nabla f(\xi_{\nu+1}) \quad \text{dla } t \in (\nu, \nu+1), \quad \nu = 0, 1, \dots .$$

Ponieważ $\nabla f(\xi_{\nu+1}) \neq 0$, możemy zapisać ∇f we współrzędnych biegunowych $\nabla f(r\theta) = \nabla' f(r\theta) + \partial_r f(r\theta)\theta$. Zatem

$$r'_\gamma(t) = -\frac{1}{2Nf(\xi_{\nu+1})}\partial_r f(\xi_{\nu+1}) \quad \text{for } t \in (\nu, \nu+1), \quad \nu = 0, 1, \dots$$

oraz

$$r_\gamma(t)\theta'_\gamma(t) = -\frac{1}{2Nf(\xi_{\nu+1})}\nabla' f(\xi_{\nu+1}) \quad \text{for } t \in (\nu, \nu+1), \quad \nu = 0, 1, \dots .$$

Stąd i z (6.35),

$$r'_\gamma(t) = -\frac{1}{2Nf(\xi_{\nu+1})}\left[kr_\gamma^{k-1}(t)f_k(\theta_\gamma(t)) + \dots + dr_\gamma^{d-1}(t)f_d(\theta_\gamma(t))\right]$$

dla $t \in (\nu, \nu+1)$, $\nu = 0, 1, \dots$. Przy założeniu, że $f_k(\theta) > 0$ dla $\theta \in S^{n-1}$ zauważmy, że pochodna ma stały znak $r'_\gamma(t) < 0$ dla wystarczająco dużego $t \notin \mathbb{Z}$. W konsekwencji, ciąg $|\xi_\nu|$ jest ściśle malejący od pewnego miejsca i udowodniliśmy dodatkową część tego faktu. Co więcej, $r_\gamma(t)$ dąży do 0 gdy $t \to \infty$ oraz z (6.36), mamy

$$\begin{aligned}|\theta'(t)| &= \frac{1}{2Nf(\xi_{\nu+1})r_\gamma(t)}|\nabla' f(\xi_{\nu+1})| \\ &\leq \frac{C_2}{2Nf(\xi_{\nu+1})}\partial_r f(\xi_{\nu+1}) = C_2|r'_\gamma(t)| \leq C_3|\gamma'(t)|\end{aligned}$$

dla $t \in (\nu, \nu+1)$ i wystarczająco dużego ν. Skoro krzywa γ ma skończoną długość z (patrz fakt 6.3.7 (i)), to z powyższego wynika, że θ_γ również ma skończoną długość. W związku z tym krzywa $\Theta : [0, +\infty) \to \mathbb{R}^n$ określonym wzorem

$$\Theta(t) = \theta(\xi_\nu) + (t-k)\left[\theta(\xi_{\nu+1}) - \theta(\xi_\nu)\right] \quad \text{dla } t \in [\nu, \nu+1), \ \nu = 0, 1, \dots$$

ma skończoną długość. To daje to, że istnieje granica $\lim_{\nu\to\infty}\theta(\xi_\nu)$ t.j., istnieje granica (6.34). $\square$

W dowodzi hipotezy gradientowej R. Thoma w [7] kluczową rolę pełnią zbiory

$$W_\varepsilon = \{x \in \mathbb{R}^n \setminus \{0\} : \varepsilon|\nabla' f(x)| \leq |\partial_r f(x)|\}, \quad \varepsilon > 0,$$

gdzie $\nabla' f(x)$ jest częścią sferyczną, a $\partial_r f(x)\frac{x}{|x|}$ — część radialną gradientu $\nabla f(x)$. Jedną z ważniejszych własności było zachowanie trajektorii pola gradientu podczas przekraczania brzegu takiego zbioru (oraz własności tzw. funkcji sterującej). Dokładniej, trajektoria pola gradientowego musi przebiegać przez ten zbiór od pewnego punktu i nie może go opuszczać. W dyskretnym przypadku ciąg może wskoczyć do tego zbioru lub z niego wyjść bez przekraczania jego brzugu.

Aby metoda z [7] mogła być zastosowana w przypadku dyskretnym, wydaje się, że następujące przypuszczenie musiałoby zachodzić.

Przypuszczenie 1. Istnieje stała $\varepsilon > 0$ i ν_0 taka, że dla każdego $\nu \geq \nu_0$

$$\varepsilon|\omega_{\nu+1} - \omega_\nu| \leq ||\omega_{\nu+1}| - |\omega_\nu||,$$

równoważnie, $\varepsilon|\nabla f(\omega_\nu)| \leq |\partial_r f(\omega_\nu)|$, t.j., $\omega_\nu(\xi) \in W_\varepsilon$.

Uwaga 6.3.20. W rzeczywistości w dowodzie faktu 6.3.19 udowodniliśmy, że W_ε, dla pewnego $\varepsilon > 0$, jest równe jakiemuś otoczeniu początku układu współrzędnych. Ponadto przy założeniu tego faktu udowodniliśmy, że $\varepsilon|\nabla' f(r\theta)| \leq r|\partial_r f(r\theta)|$ w otoczeniu zera. Jest to silniejszy warunek niż fakt, że ξ_ν należy do zbioru W_ϵ. Wydaje się, że udowodnienie przypuszenie 1, nie wystarczy, aby pokazać, że ciąg $\theta(\xi_\nu)$ jest zbieżny. Ciąg ξ_ν powinien spełniać nierówność $\varepsilon|\nabla' f(\xi_\nu)| \leq |\xi_\nu|^\alpha |\partial_r f(\xi_\nu)$ dla pewnych stałych dodatnich ε, α.

Bibliografia

[1] A.N. Abdullah, K. Rosiak, S. Spodzieja, *Convexifying of polynomials by convex factor*, Chapter in Analytic and Algebraic Geometry 4, 31 pp. Łódź University Press 2022.

[2] J. Bochnak, S. Łojasiewicz, *A converse of the Kuiper-Kuo theorem.* Proceedings of Liverpool Singularities?Symposium, I (1969/70), pp. 254?261. Lecture Notes in Math., Vol. 192. Springer, Berlin, 1971.

[3] S. Boyd and L. Vandenberghe. *Convex Optimization.* Cambridge University Press, 2004.

[4] E.K.P. Chong and S.H. Zak. *An Introduction to Optimization.* Fourth edition. Wiley, 2013

[5] Geoffrey A. Jehle and Philip J. Reny, *Advanced Microeconomic Theory*, (New York: Addison-Wesley, 2001).

[6] G. Jeronimo, D. Perrucci, E. Tsigaridas, *On the minimum of a polynomial function on a basic closed semialgebraic set and applications*. SIAM J. Optim. 23 (2013), no. 1, 241–255.

[7] K. Kurdyka, T. Mostowski, A, Parusiński, *Adam Proof of the gradient conjecture of R. Thom.* Ann. of Math. (2) 152 (2000), no. 3, 763?792.

[8] K. Kurdyka, K. Rudnicka, S. Spodzieja, *Exponential convexifying of polynomials.* Bull. Sci. Math. 180 (2022), Paper No. 103197, 23 pp.

[9] K. Kurdyka, S. Spodzieja, *Convexifying positive polynomials and sums of squares approximation.* SIAM J. Optim. 25 (2015), no. 4, 2512–2536.

[10] K. Kurdyka, S. Spodzieja, *Separation of real algebraic sets and the ?ojasiewicz exponent.* Proc. Amer. Math. Soc. 142 (2014), no. 9, 3089–3102.

[11] W.B. Liu, C.A. Floudas, *A remark on the GOP algorithm for global optimization.* (English summary) J. Global Optim. 3 (1993), no. 4, 519?521.

[12] S. Łojasiewicz, *Une propriété topologique des sous-ensembles analytiques réels.* (French) 1963 Les Ėquations aux Dérivées Partielles (Paris, 1962) pp. 87?89 Ėditions du Centre National de la Recherche Scientifique, Paris

[13] S. Łojasiewicz, *Ensembles semi-analytiques*, preprint IHES, 1965.

[14] S. Łojasiewicz, *Sur les trajectoires du gradient d'une fonction analytique.* (French) [Trajectories of the gradient of an analytic function] Geometry seminars, 1982?1983 (Bologna, 1982/1983), 115?117, Univ. Stud. Bologna, Bologna, 1984.

[15] S. Łojasiewicz, *Sur les trajectoires du gradient d?une function analytique*, in: Geometry Seminars, 1982-1983, Univ. Stud. Bologna, Bologna 1984, 115 -117.

[16] Y. Nesterov, *Introductory lectures on convex optimization.* A basic course. Applied Optimization, 87. Kluwer Academic Publishers, Boston, MA, 2004.

[17] R.T. Rockafellar, *Monotone operators and the proximal point algorithm.* SIAM J. Control Optim. 14 (1976), no. 5, 877–898.

[18] K. Rosiak *Convexifying of polynomials* [Uwypuklanie wielomianów] (in Polish), Master Thesis, University of Lodz, Lodz 2020, 37pp.

[19] A.N. Tikhonov, V.Y. Arsenin, *Solutions of ill-posed problems.* Translated from the Russian. Preface by translation editor Fritz John. Scripta Series in Mathematics. V.H. Winston & Sons, Washington, D.C.: John Wiley & Sons, New York-Toronto, Ont.-London, 1977.

[20] S. Zlobec, *Estimating convexifiers in continuous optimization.* (English summary) Math. Commun. 8 (2003), no. 2, 129?137.

Skorowidz

Wykaz oznaczeń

$\sum_f$	zbiór wartości krytycznych	19
$\mathbb{R}$	ciało liczb rzeczywistych	23
$\mathbb{N}$	zbiór liczb naturalnych	23
$\lvert\cdot\rvert$	norma Euklidesowa w $\mathbb{R}^n$	23
$\mathbb{R}^0_+$	zbiór liczb rzeczywistych dodatnich	23
$\mathbb{N}_0$	zbiór liczb całkowitych nieujemnych	**??**
$\partial_v f$	pochodna kierunkowa funkcji f w kierunku pola wektorowego v	24
$\partial_v^2 f(x)$	pochodna drugiego rzędu funkcji f w kierunku v przy x	24
$I_{a,v}$	$I_{a,v} = \{t \in \mathbb{R} : a + tv \in X\}$ przedział lub zbiór jednoelementowego	25
S^{n-1}	$S^{n-1} = \{x \in \mathbb{R}^n : \lvert x\rvert = 1\}$ sfera jednostkowa w $\mathbb{R}^n$	26
∇f	gradient funkcji f	27
$f = \sum_{\lvert\nu\rvert\leq d} a_\nu x^\nu$	wielomian rzeczywisty zmiennych $x = (x_1, \ldots, x_n)$	34
$\mathbb{R}[x]$	$\mathbb{R}[x_1, ..., x_n]$ pierścieniem wielomianów n zmiennych o wspólrzędnych z $\mathbb{R}$	34
$\lVert f\rVert$	$\lVert f\rVert := \sum_{\lvert\nu\rvert\leq d} \lvert a_\nu\rvert$ norma w pierścieniu $\mathbb{R}[x]$	34
f_j	$f_j = \sum_{\lvert\nu\rvert=j} a_\nu x^\nu, \quad 0 \leq j \leq d$ wielomian jednorodny stopnja j	34
$m(f)$	$m(f) := f_{d*} - \sum_{j=0}^{d-1} \mathbb{K}(f)^{j-d}\lvert\lvert f_j\rvert\rvert$	36
$D_n^1(f, R)$	$D_n^1(f, R) := \sum_{j=0}^{d} j\lvert\lvert f_j\rvert\rvert R^{j-1}$	37
$D_n^2(f, R)$	$D_n^2(f, R) := \sum_{j=0}^{d} j(j-1)\lvert\lvert f_j\rvert\rvert R^{j-2}$	37
$D_n(f, R))$	$D_n(f, R) := \max\{D_n^1(f, R),\ D_n^2(f, R)\}$	37
$\varphi(x)$	$\varphi(x) = b^N(x)f(x)$	41
$N(\mu, S, m, D)$	$N(\mu, S, m, D) = \frac{S}{\mu}\left(\frac{D}{m} + \frac{D^2}{m^2}\right) + 1$	43
S	$S = \max\{b(x) : x \in X\}$	43
S'	$S' = \max\{b(x-\xi) : x,\, \xi \in X\}$	44

$\lambda_{N,\xi}(x)$	$\lambda_{N,\xi}(x) = b^N(x-\xi)f(x)$	44
$N_{\exp}(\mu, m, D)$	$N_{\exp}(\mu, m, D) = \frac{1}{\mu}\left(\frac{D}{m} + \frac{D^2}{m^2}\right)$	45
$\Psi_{N,\xi}(x)$	$\Psi_{N,\xi}(x) = Nb(x-\xi) + f(x), \quad x \in \mathbb{R}^n$	45
$\phi_{N,\xi}$	$\phi_{N,\xi} = Nb(x-\xi) + f(x), \qquad x \in \mathbb{R}^n$	46
$\psi_{N,\xi}(x)$	$\psi_{N,\xi}(x) = N \ln b(x-\xi) + \ln f(x), \quad x \in \mathbb{R}^n$	46
$\tilde{f}$	$\tilde{f} = f + D_n^0(f, R) + 1$	47
R	$R = \max\{\lvert x\rvert : x \in X\}$	47
$\varphi_N(x)$	$\varphi_N(x) = b^N(x)f(x)$	47
$\varphi_{N,\xi}(x)$	$\varphi_{N,\xi}(x) = b^N(x-\xi)f(x)$	47
$N_{\exp,\infty}(\mu, f)$	$N_{\exp,\infty}(\mu, f) = \frac{d(d+1)\lvert\lvert f\rvert\rvert}{4\mu m(f)}$	48
f_{d*}, m	f_{d*}, $m = \min\{f(x) : x \in X\}$	49
$\tilde{H}$	$\tilde{H} = \max\{H, 2n + 2k\}$	49
$\mathbb{Z}[x]$	$\mathbb{Z}[x_1, ..., x_n]$ pierścieniem wielomianów n zmiennych o współrzędnych z $\mathbb{Z}$	49
$\varphi_{N,\xi}(x)$	$\varphi_{N,\xi}(x) = b(N(x-\xi))f(x)$	51
$\phi_\xi(x)$	$\phi_\xi(x) = Nb(\xi)b(x-\xi) + f(x)$	53
$\phi_\xi(x)$	$\phi_\xi(x) = N(\lvert\xi\rvert)b(x-\xi) + f(x), \quad x \in \mathbb{R}^n,$	55
$N(\lvert\xi\rvert)$	$N(\lvert\xi\rvert) = \frac{D}{\mu}\left(\lvert\xi\rvert + 1 + \sqrt{\frac{\alpha}{\mu}}\right)^\alpha + 1$	55
$\psi(\xi, x)$	$\psi(\xi, x) = N_1(\xi)b(x-\xi) + f(x), \quad (\xi, x) \in \mathbb{R}^n \times \mathbb{R}^n$	56
$N_1(\xi)$	$N_1(\xi) = N(\sqrt{\lvert\xi\rvert^2 + 1}) = \frac{D}{\mu}\left(\sqrt{\lvert\xi\rvert^2 + 1} + 1 + \sqrt{\frac{\alpha}{\mu}}\right)^\alpha + 1$	56
$X_{f \leq r}$	$X_{f\leq r} := \{x \in \mathbb{R}^n : f(x) \leq r\}$	64
$R_{f\leq r}$	$R_{f\leq r} = \sup\{\lvert x\rvert : x \in X_{f\leq r}\}$	64
B_R	$B_R = \{x \in \mathbb{R}^n : \lvert x\rvert \leq R\}$	64
$\kappa_N(\xi)$	$\kappa_N(\xi) = \operatorname{argmin}_{B_R} \phi_{N,\xi} \in B_R \quad$ dla $\xi \in B_R$	65
$\kappa_{N,*}(\xi)$	$\kappa_{N,*}(\xi) = \lim_{\nu\to\infty} \kappa_N^\nu(\xi)$	71
$\kappa_{N,*}\rvert_U$	$\kappa_{N,*}\rvert_U : U \to U \cap \Sigma_f$	71
$X_{f\leq f(0)+c}$	$X_{f\leq f(0)+c} = \{x \in \mathbb{R}^n : \lvert f(x) - f(0)\rvert < c\} \subset W$	71
$\kappa(\xi)$	$\kappa(\xi) = \operatorname{argmin}_X \psi_\xi \in X \quad$ dla $\xi \in X$	75
$\kappa_N(\xi)$	$\kappa_N(\xi) = \operatorname{argmin}_X \varphi_{N,\xi}$	78
$S'_{b,f,r}$	$S'_{b,f,r} = \max\{b(x-\xi) : x, \xi \in B_R\}$	80
$\gamma_\xi(t)$	$\gamma_\xi(t) = \kappa_N^\nu(\xi) + (t-\nu)(\kappa_N^{\nu+1}(\xi) - \kappa_N^\nu(\xi)) \quad$ dla $t \in [\nu, \nu+1)$	84
$\Theta(t)$	$\Theta(t) = \theta(\xi_\nu) + (t-k)\left[\theta(\xi_{\nu+1}) - \theta(\xi_\nu)\right] \quad$ dla $t \in [\nu, \nu+1),\ \nu = 0, 1, \ldots$	95