dr hab. Maciej P. Denkowski, prof. UJ
Uniwersytet Jagielloński
Wydział Matematyki i Informatyki
Instytut Matematyki
maciej.denkowski@uj.edu.pl

Kraków, 31 sierpnia 2023 r.

Recenzja rozprawy doktorskiej pana mgra Abdulljabara Naji Abdullaha pt. *Logarytmiczne uwypuklenie wielomianów*

Omawiana tutaj przez mnie rozprawa doktorska pana mgra Abdullaha sytuuje się na styku matematyki czystej i stosowanej, a dotyczy ważnego w wielu dziedzinach zagadnienia optymalizacyjnego, polegającego na wyszukiwaniu punktów, w których dana funkcja rzeczywista osiąga ekstremum, czy też nieco ogólniej – punktów krytycznych. Jedną z klasycznych metod, stosowaną m.in. w problemach sterowania optymalnego i algorytmach minimalizacji funkcji, jest uwypuklenie funkcji (określonej na zbiorze zwartym i wypukłym), bądź przez dodanie do niej funkcji silnie wypukłej, bądź też — co jest nowszym podejściem — przez domnożenie jej przez taką funkcję. Jak wiadomo, problem wyznaczenia minimum funkcji silnie wypukłej jest znacznie prostszy.

Punktem wyjścia wszystkich rozważań rozprawy jest w gruncie rzeczy następujący problem: dana jest funkcja rzeczywista f klasy $\mathscr{C}^2$ określona w otoczeniu (domkniętego) zbioru wypukłego $X \subset \mathbb{R}^n$ oraz silnie (lub logarytmicznie silnie) wypukła funkcja $b\colon \mathbb{R}^n \to \mathbb{R}$ tej samej klasy, przy czym możemy założyć, że jedyne jej minimum jest osiągane w zerze: $\mathrm{argmin}_{\mathbb{R}^n} b = 0$, oraz $b(0) = 1$, czyli $b \geqslant 1$. Przypomnijmy, że silna wypukłość b w danym wypadku oznacza szacowanie przyrostu funkcji b w dowolonym punkcie x od dołu przez wartość różniczki w x na przyroście argumentu powiększoną o silnie wypukły składnik $c||x-y||^2$ (gdzie $c > 0$ jest stałą; mówimy wtedy, że funkcja jest $2c$-silnie wypukła; powiemy, że $b > 0$ jest logarytmicznie silnie wypukła, gdy silnie wypukła jest funkcja $\ln b$). W szczególności jest to funkcja ściśle wypukła. Interesuje nas uwypuklenie f postaci $x \mapsto Nb(x-\xi) + f(x)$, gdzie liczba $N > 0$ jest dostatecznie duża (naturalne pytanie: jak duża?), a $\xi \in \mathbb{R}^n$ jest parametrem — i to jest podejście klasyczne; względnie, uwypuklenie postaci $x \mapsto b(x-\xi)^N f(x)$. Od razu przeczuwamy, że ograniczość zbioru X ma duży wpływ na zagadnienie, a różnica pomiędzy obu podejściami jest znaczna. W przypadku, gdy uzyskana funkcja, nazwijmy ją ϕ_ξ, wychodzi faktycznie silnie wypukła, możemy skojarzyć z nią odwzorowanie $\kappa(\xi) = \text{argmin}_{X'}\phi_\xi$, gdzie X' jest odpowiednio dobranym zbiorem do X, w taki sposób, by możliwe było iterowanie κ. Naturalnym jest w tej sytuacji oczekiwanie zbieżności iteracyj $\kappa^\nu(\xi)$ do punktów krytycznych f.

W ostatnich latach ponownie wzrosło zainteresowanie możliwościami, jakie niesie ze sobą dla zastosowań topologia ujarzmiona, czyli w szczególności geometria semi-algebraiczna, subanalityczna, czy wreszcie struktury o-minimalne. Praca pana mgra Abdullaha uwzgłednia właśnie w istotny sposób przypadek wielomianów czy ogólniej funkcyj semi-algebraicznych (a więc jedne z najważniejszych z punktu widzenia zastosowań), choć zawiera również szereg istotnych wyników całkiem ogólnych dla funkcyj klasy przynajmniej $\mathscr{C}^2$. Jej źródeł należy dopatrywać się w uzyskanych nie tak dawno wynikach K. Kurdyki i S. Spodziei [9] (wg numeracji bibliografii rozprawy), a spora jej część oparta jest na wspólnej z K. Rosiak i S. Spodzieją publikacji doktoranta [1] w materiałach konferencyjnych. Tematyka rozprawy wpisuje się w ważne i aktualne nurty badań i ma znaczenie dla zastosowań, a stosowane metody są pomysłowym rozwinięciem technik z pracy [9].

Rozprawa, napisana po polsku na stu trzech stronach, składa się z sześciu rozdziałów poprzedzonych streszczeniami w językach angielskim i polskim, i wzbogacona jest o skorowidz oraz wykaz oznaczeń. Wstęp czyli streszczenie przedstawia motywacje, ogólny zarys i główne wyniki pracy. Rozdział pierwszy zawiera podstawowe pojęcia i fakty dotyczące funkcyj wypukłych, ściśle wypukłych, silnie wypukłych i logarytmicznie silnie wypukłych, tudzież wielomianów i pewnych oszacowań wykorzystywanych dalej w rozprawie. W rozdziale drugim przedstawione są pokrótce podstawowe fakty dotyczące uwypuklania funkcji za pomocą domnażania jej przez potęgę funkcji silnie wypukłej (z wyjątkiem istotnej uwagi 2.2.3 dotyczącej uwypuklenia przez dodanie takiej funkcji). Kolejny rozdział, trzeci, zawiera już kilka nowych wyników dotyczących uwypuklania również przez dodanie funkcji silnie wypukłej. W rozdziale czwartym Autor przeprowadza analizę własności funkcjiv κ_N przyporządkującej parametrowi przesunięcia ξ jedynego punktu, w którym osiągane jest minimum funkcji po uwypukleniu postaci $Nb(x-\xi) + f(x)$ z $N \gg 1$ na kuli $||x|| \leqslant R$ w $\mathbb{R}^n$. Tutaj znajdujemy pierwsze istotne wyniki rozprawy. Następnie Doktorant bada analogiczne odwzorowanie, lecz tym razem na zbiorze wypukłym i domkniętym X w miejsce kuli, co wymaga modyfikacji uwypuklenia do postaci $N_1(\xi)b(x-\xi) + f(x)$, przy czym rozdział kończy się jednym wynikiem dla przypadku wielomianu f i $X = \mathbb{R}^n$ oraz uwypuklenia postaci $b(N(x-\xi))f(x)$. W rozdziale szóstym przenoszone są wyniki rozdziału czwartego na przypadek uwypuklenia postaci $b(x-\xi)^N f(x)$ ponownie na kuli i ten rozdział jest najobszerniejszy. Osobno rozpatrzony jest w nim przypadek szczególnej funkcji uwypuklającej $b(x) = \exp(||x||^2)$. Cały ten rozdział opiera się on na wspomnianej wcześniej publikacji [1].

Rozprawa napisana jest poprawnie, choć Autor nie ustrzegł się pokaźnej liczby błędów zecerskich i pewnych niekonsekwencyj w oznaczeniach. Są to jednak mało istotne dla wartości pracy usterki, które łatwo wyeliminować i nie wpływają na czytelność tekstu.

Przechodząc do omówienia matematycznej strony uzyskanych przez Doktoranta wyników należy podkreślić na wstępie ich częstokroć nader techniczny charakter. Większość dowodów to dość zgrabna, klasyczna analiza i wcale subtelne oszacowania. Istotną rolę odgrywają w szczególności wyniki Kurdyki i Spodziei z [5] oraz Jeronimo, Perrucciego i Tsigardisa z [6], tudzież gradientowa nierówność Łojasiewicza i efektywna nierówność Łojasiewicza w ujęciu Kurdyki i Spodziei [10].

Rozdział pierwszy pełni rolę przygotowawczą. Za najważniejsze wyniki w nim przedstawione, potrzebne w dalszej części pracy, należy uznać oszacowania podane jako fakt 1.4.2, wynikający zeń wniosek 1.4.3, oraz fakt 1.4.4 (gdzie środkowa część drugiej wyró.znionej na stronie 30 nierówności jest niejasna, ale też i niepotrzebna, bo skrajne nierówności wywodzą się z (1.7)) i bardzo istotny lemat 1.4.8 (gdzie błędnie przywołany jest w dowodzie fakt 1.2.1 zamiast 1.3.1). Skądinąd lepiej byłoby te „fakty" przemianować na „stwierdzenia" (w rozprawie się ich zresztą dowodzi), nie są to bowiem obserwacje całkiem oczywiste. Nie rozumiem natomiast, jaki cel przyświecał umieszczeniu bardzo technicznych uwag o funkcjach logarytmicznie silnie wypukłych ℓ-tego rzędu w podrozdziale 1.5, cenny za to jest podrozdział 1.6 dotyczący wielomianów rzeczywistych wielu zmiennych i pewnych oszacowań, mających znaczenie w zastosowaniach w ogóle. Należy zaznaczyć, że we wzorach pomiędzy wnioskami 1.6.5 i 1.6.6 na str. 37 sumowanie winno zaczynać się od $j = 1$. Fakt 1.6.7, jakkolwiek prosty, mógłby mieć dowód zredagowany nieco szczegółowiej (liczba szacująca zera wielomianu to, przy wyborze indeksów przyjętym przez Autora, $2\max_j \left|\frac{a_j}{a_0}\right|^{1/j}$ — w pracy indeks w liczniku jest błedny; błędne jest też odwołanie do wzoru (1.19) zamiast do (1.18)). Natomiast w części dotyczącej gradientu we współrzędnych biegunowych chętnie bym widział (szczególnie w rozprawie doktorskiej, gdzie właśnie jest miejsce na szerokie komentarze) nieco dokładniejsze rozpisanie całej konstrukcji, tym bardziej, że ma ona jasny sens geometryczny. Skądinąd we wzorze (1.21) zapis $\partial_\theta f(r\theta)$ w ogóle nie powinien się pojawić. W uwadze 1.6.8 wszystkie wyrażenia α_j czy v_j zależą od x i wskazane byłoby to zaznaczyć ze względu na wzór pojawiający się na dole str. 39. W trzeciej linijce od dołu na tejże stronie norma $||v_j||$ winna być w kwadracie.

W rozdziale drugim zamieszczono ogólne uwagi nt. uwypuklania funkcji na zbiorach zwartych i wypukłych. Wyniki te są cenne, w szczególności fakty 2.1.1 i 2.1.3 oraz rozważania dotyczące uwypuklania wielomianów. Dwie uwagi na temat oznaczeń: wydaje się, że zapis $b(x)^N$ potęgowania funkcji jest bardziej naturalny niż $b^N(x)$, kojarzący się raczej ze składaniem; po drugie zaś wskazanym byłoby ujednolicić oznaczenia — przykładowo funkcje φ z faktu 2.1.3 i λ_N z lematu 2.2.1 w obrębie tego samego rozdziału różni tylko to, że pierwsza zdefiniowana jest dla wielomianu $f > 0$, druga, tym samym wzorem, dla dowolnej funkcji dodatniej i klasy $\mathscr{C}^2$, przy czym po uwzględnieniu we wzorze przesunięcia o ξ, oznaczenie zmienia się w $\lambda_{N,\xi}$ (wniosek 2.2.2, oznaczenie na szczęście utrzymane konsekwentnie w rozdziale 6, ale nie we wstępie na str. 15). W dowodzie lematu 2.2.1 we wzorze na $\partial_\beta \lambda_N(x)$ pojawia się w drugim składniku nadmiarowy czynnik $b(x)$, dalej jednak rachunki są poprawne. W tym rozdziale znajduje się bardzo istotna uwaga 2.2.3 dotycząca możliwości uwypuklenia f na zwartym zbiorze wypukłym za pomocą $Nb(x-\xi)+f(x)$ z dostatecznie dużym czynnikiem N (niepotrzebnie oznacza się tę funkcję $\Psi_{N,\xi}(x)$, gdy już w rozdziale 3 i w rozdziale 4 to samo wyrażenie będzie zapisane jako $\phi_{N,\xi}$), zdublowana niejako faktem 3.3.1.

Problem uwypuklania na zbiorach nieograniczonych podjęty zostaje w rozdziale trzecim. Bardzo dobry wstęp i przykłady pokazują, jakie problemy napotyka się przy uwypuklaniu przez dodanie funkcji silnie wypukłej. Nieco razi oznaczenie $N(||\xi||)$ z lematu 3.2.1, zważywszy, że funkcja N nie jest określona osobnym wzorem na $[0,+\infty)$ a pojawia się we wzorze (3.7) jako funkcja n zmiennych. Rachunki i szacowania stają się dosyć techniczne, pomocne dla czytelnika byłoby zamieszczenie więcej szczegółów. I tak, w dowodzie lematu 3.2.1 ostatnia nierówność daleka jest od oczywistości w takiej akurat postaci, jak ją podano, a w dowodzie lematu 3.2.5 stwierdzenie, że (3.10) (gdzie omyłkowo jest wpisana pierwsza pochodna zamiast drugiej) wynika z zapowiedzi w pierwszej linijce na stronie 58 jest całkiem niejasne. Co więcej dalej rachunki przestają się zgadzać i wzór na r_0 jest błędny (zamiast wyrażenia $\frac{4}{\mu}(||\xi||+\alpha)$ powinien zawierać $4(||\xi||+\alpha/\mu)$). Faktem jest, że wynik w swojej esencji pozostaje prawdziwy, niemniej jednak konkretna postać μ_1 z tezy winna być prawdopodobnie inna.

Dowód twierdzenia 3.3.1 zawiera na str. 60 błędne sformułowanie $\partial_\beta^2 \varphi_\xi(x) \geqslant \mu$ *wtedy i tylko wtedy, gdy* $g_N(||x||) \geqslant 1$ do zastąpienia przez samo *wtedy, gdy* (co pozostaje bez wpływu na dowód). Uwaga 3.3.2 powinna być rozwinięta i napisana jaśniej, jako że dotyczy bardzo ważnej motywacji rozprawy polegającej na poszerzeniu wyników z artykułu [9], gdzie nie rozważano funkcji $b(x)=\exp(||x||^2)$, jak jest napisane (ta funkcja badana była w [1]), a bardziej nawet szczególną $b(x)=(1+||x||^2)$. Warto przy tym odnotować istotną dyskusję uzupełniającą uwagi wstępne do rozdziału 3, gdzie domnaża się f przez funkcję $b(N(x-\xi))$ a nie tak jak w rozdziale 2 przez $b(x-\xi)^N$, co akurat przy funkcji eksponens z pozoru jest prawie tym samym podejściem. Rozdział zamyka się twierdzeniem 3.3.3, do którego również mam zastrzeżenia natury rachunkowej, oszacowanie N wychodzi bowiem inne niż podaje teza. W dowodzie, z (3.11) szacowanie winno być przez A/m^2, a nie przez A/m, natomiast końcowe szacowanie z faktu 1.6.1 (a chyba należałoby przywołać także i 1.6.4) nie jest jasne, w mianowniku powinno raczej znaleźć się wyrażenie $(f_{d*}||x||-d||f||)^2$ (zgubiono d) a i licznik według moich obliczeń powinien zawierać czynnik nieco bardziej skomplikowany niż $d(2d-1)$. Ta część wymaga od Autora dokładniejszego rozpisania przeliczeń, choć jak poprzednio — twierdzenie w swojej istocie jest prawdziwe, zmienią się co najwyżej liczby.

Ostatnie trzy rozdziały zawierają najważniejsze z mojego punktu widzenia wyniki rozprawy a dotyczą własności odwzorowania przyporządkowującego parametrowi przesunięcia ξ jedynego punktu $\kappa(\xi)$, w którym uwypuklenie funkcji f osiąga minimum, oraz zastosowania do szukania punktów krytycznych f. Dana jest funkcja $f\colon \mathbb{R}^n \to \mathbb{R}$ klasy $\mathscr{C}^k$ z $k \geqslant 2$, zbiór domknięty i wypukły $X \subset \mathbb{R}^n$ oraz funkcja $b\colon \mathbb{R}^n \to \mathbb{R}$ tej samej co f klasy, μ-silnie wypukła z jedynym minimum w środku układu współrzędnych i $b(0) = 1$. Rozważane są następujące przypadki:

Rozdział 4 Zakładamy, że $X_{f\leqslant r} := \{x \in \mathbb{R}^n \mid f(x) \leqslant r\}$ (gdzie $r > 0$ jest ustalone) jest niepusty i zawarty w kuli $B_R := \mathbb{B}(0, R)$, f ma drugie pochodne kierunkowe jednostajnie ograniczone na B_R a N jest dostatecznie duże, aby $\phi_{N,\xi}(x) = Nb(x-\xi)+f(x)$ była silnie wypukła na B_R (co gwarantuje uwaga 2.2.3). Wtedy ma sens odwzorowanie $\kappa_N(\xi) = \mathrm{argmin}_{B_R}\phi_{N,\xi} \in B_R$.

Rozdział 5 Tutaj nie zakładamy zwartości rozważanego zbioru X, od b żądamy logarytmicznej μ-silnej wypukłości, co przy założeniu, że f ma wzrost wielomianowy drugiego rzędu (zgodnie z definicją z 3.2), gwarantuje na podstawie wniosku 3.2.3, że przy $N_1(\xi)$ zadanej wzorem (3.7) funkcja $\psi_\xi(x) = N_1(\xi)b(x-\xi)+f(x)$ jest μ-silnie wypukła na zbiorze X. Pozwala to teraz zdefiniować $\kappa_N(\xi) = \mathrm{argmin}_X\psi_{N,\xi} \in X$. Osobno wspomniany na końcu rozdziału jest przypadek, gdy f jest wielomianem odseparowanym od zera i z dodatnim minimum formy wiodącej na sferze, $X = \mathbb{R}^n$, a uwypuklenie ma postać $b(N \cdot (x-\xi))f(x)$.

Rozdział 6 Z niejasnych przyczyn w tym rozdziale Autor ogranicza się do funkcyj klasy $\mathscr{C}^2$, choć w istocie całość przenosi się na klasę $\mathscr{C}^k$ z $k \geqslant 2$ jak poprzednio (w szczególności funkcja κ_N jest wtedy klasy $\mathscr{C}^{k-1}$, por. lemat 6.1.2). Przyjmujemy założenia rozdziału 4 z dodatkowym wymogiem jednostajnego ograniczenia f z dołu oraz jednostajnego ograniczenia także i pierwszych pochodnych kierunkowych tej funkcji na B_R. Wniosek 2.2.2 gwarantuje wówczas, że dla dużych wartości $N > 0$, funkcja $\lambda_{\xi,N}(x) = b(x-\xi)^N f(x)$ jest silnie wypukła na B_R, ma przeto sens $\kappa_N(\xi) = \mathrm{argmin}_{B_R}\lambda_{N,\xi} \in B_R$. Bardzo interesujący jest podrozdział dotyczący przypadku wielomianu $f \geqslant 1$ i konkretnej funkcji $b(x) = \exp||x||^2$.

Jak widać rozdziały 4 i 6 są bardzo podobne, jeśli chodzi o punkt wyjścia (a także i wyniki) i być może lepiej byłoby ich było nie rozdzielać rozdziałem dotyczącym zgoła całkiem innego przypadku. Jak wspomniałem wyżej, klasa funkcji κ_N jest zawsze o jeden niższa. Ciekawym pytaniem wydaje się być zagadnienie, czy ten spadek rzeczywiście zawsze musi mieć miejsce. W rozdziałach 4 i 6 Autor pokazuje również, że $\kappa_N(X_{f\leqslant r}) \subset X_{f\leqslant r}$, a punkty stałe $\kappa_N|_{X_{f\leqslant r}}$ to dokładnie $X_{f\leqslant r} \cap \Sigma_f$, gdzie Σ_f oznacza punkty krytyczne f. Podobny wynik otrzymany jest też w rozdziale 5, przy czym sytuacja jest delikatniejsza ze względu na pewną dowolność zbioru X. Najważniejsze jednak są rezultaty dotyczące iteracji κ_N, tj. twierdzenia 4.2.1 (algorytm zbliżeniowy tj. *proximity algorithm* dla funkcji semi-algebraicznej) i 4.3.1 (o retrakcji) oraz analogiczne twierdzenia 6.2.3 i 6.3.10. W rozdziale 5 brak tego typu wyników, za to istotne są wnioski 5.1.5 oraz 5.2.1.

Dowody spisano momentami zanadto zwięźle; należałoby też uwypuklić, *nomen omen*, pewne subtelności (np. krzywa γ_ξ z twierdzenia 4.2.1 przyjmuje wartości w $X_{f\leqslant r}$ dzięki monotoniczności f uzyskanej w 4.1.5) czy też uporządkować oznaczenia — iteracje κ_N generują bowiem dwa typy zapisu: w twierdzeniu 4.2.1 mamy $\xi_{\nu+1} = \kappa_N^\nu(\xi_0)$ (odnotujmy przy okazji kolizję oznaczenia ν-tej iteracji κ_N ze stosowanym w innych miejscach oznaczeniem N-tej potęgi $b^N(x)$) a następnie $\omega_\nu(\xi_0) = \kappa_N^\nu(\xi_\nu)$ w twierdzeniu 4.3.1. Z kolei w lemacie 4.1.3 stosuje się pewien skrót myślowy (powielony dalej), gdy mowa o dyfeomorfizmie pomiędzy zbiorami, które nie są otwarte i nie posiadają żadnej struktury, aczkolwiek z kontekstu łatwo domyślić się, o co w istocie chodzi.

Są też nieco poważniejsze usterki. W twierdzeniu 4.2.1 jeden z kluczowych wyników, jakim jest zbieżność szeregu z podpunktu (a) Autor zbywa prostym odwołaniem do pracy [9] („zasada porównania" 7.7 pochodząca w istocie z pracy D. D'Acunto i K. Kurdyki *Bounds for gradient trajectories and geodesic diameter of real algebraic sets*, Bull. London Math. Soc. 38 (2006), no. 6, 951-965 dotycząca krzywych gradientowych, i dowód twierdzenia 7.5 z [9]), gdy tymczasem sprawa jest dosyć delikatna (podobnie rzecz się ma w przypadku analogicznego szeregu z twierdzenia 6.2.3). Należało obowiązkowo zawrzeć dowód zbieżności, tym bardziej że nie jest to jednak standardowe rozumowanie, a dodatkowo w [9] , gdzie prowadzone były podobne rozważania, istotną rolę odgrywał wybór funkcji b (co jest zresztą odnotowane w dowodzie twierdzenia 6.2.3, por. [9] Lemma 7.1). W skrócie rozumowanie mogłoby wyglądać następująco: ciąg $f(\xi_\nu)$ maleje i jest ograniczony, ponadto da się wybrać krzywą (tj. trajektorię) gradientową γ_ν startującą z punktu ξ_ν i przechodzącą przez poziomicę $f^{-1}(f(\xi_{\nu+1}))$, a co za tym idzie $\mathrm{dist}(\xi_\nu, f^{-1}(f(\xi_{\nu+1})) \leqslant \mathrm{length}\gamma_\nu$, gdzie długość γ_ν szacujemy z góry przez długość wycinka danego *talwegu* Γ przez $f^{-1}((f(\xi_{\nu+1}), f(\xi_\nu)))$ (por. [9] Comparison principle 7.7). Stąd szereg szacuje się po prostu przez długość talwegu pomiędzy poziomicami granicy ciągu $f(\xi_\nu)$ a $f(\xi_0)$ w B_R.

Omówmy od razu analogiczne twierdzenie 6.2.3, gdzie dowód początkowo przeprowadzony jest dla funkcji $b(x) = 1 + ||x||^2$, czyli takiej jak w [9], wskutek czego zachodzi wzór (6.12) znacznie upraszczający rozumowanie. Chcąc udowodnić podpunkt (a), Doktorant powołuje się na zbieżność szeregu odległości (6.13) stwierdzając na str. 85, że *w dowodzie zbieżności szeregu (6.12) postać funkcji b nie była istotna, dowód polegał na zastosowaniu zasady porównania, semi-algebraiczności funkcji f oraz monotoniczności ciągu* $f(\xi_\nu)$, tyle tylko, że tego dowodu nigdzie nie podano. Oczywiście, ze zbieżności szeregu (6.13) w połączeniu z nierównością (6.15) otrzymujemy (a), ale i tu skwitowanie problemu słowami: *nierówności (6.15) dowodzimy analogicznie jak nierówności (4.11) w dowodzie twierdzenia 4.2.1* to jednak za mało. Od rozprawy doktorskiej oczekujemy bądź co bądź pełniejszych rozumowań. Mamy tu więc dwa fakty wymagające dowodu. Pierwszy z nich można wykazać rzeczywiście za pomocą talwegów tak, jak omówiłem to wyżej i co jest rozumowaniem inspirowanym pracą [9], drugi natomiast da się uzyskać podobną metodą, jak zastosowana w dowodzie twierdzenia 4.2.1 i to można by streścić w kilku dosłownie linijkach. Wybieramy punkt $a_{\nu+1} \in f^{-1}(f(\xi_{\nu+1}))$ realizujący odległość ξ_ν od włókna, wykorzystując nierówność $\lambda_{N,\xi}(\xi_{\nu+1}) \leqslant \lambda_{N,\xi}(a_{\nu+1})$ i postać funkcji $\lambda_{N,\xi}$ dochodzimy do $b(0) \leqslant b(\xi_{\nu+1} - \xi_\nu) \leqslant b(a_{\nu+1} - \xi_\nu) \to b(0)$ (zbieżność otrzymujemy ze zbieżności szeregu i ciągłości b), skąd wobec jedyności minimum, $||\xi_{\nu+1} - \xi_\nu|| \to 0$; dalej bez problemu powtarzamy rozumowanie z 4.2.1 wykorzystujące rozwinięcie Taylora. Nawiasem mówiąc, na górze strony 85 błędnie przywołano (6.8) zamiast (6.7).

Wracając ponownie do omawiania rozdziału czwartego, formalnie rzecz biorąc, we wzorze (4.8) pochodna winna być wzięta w znormalizowanym kierunku $x - \xi$. W dowodzie lematu 4.3.2 warto podkreślić, że stałe C i ϱ w nierówności (Ł1) można dowolnie zwiększać, przez co m.in. na str. 73 można mieć $(1-\varrho)C \geqslant 1$ — bez czego nie widzę uzasadnienia dla ostatniej nierówności w wyróżnionym wzorze na górze strony — a przy definicji δ moża wziąć tę samą stałą $C > 0$ w oparciu o twierdzenie 4.2.1 (b) (nie (c) jak podano w tekście). Zamiast (4.17) naturalniej jest rozpatrzeć układ $x' = -\frac{\nabla[(f-f(0))^{1-\varrho}]}{||\nabla[(f-f(0))^{1-\varrho}]||}$, co ma sens w $W \setminus Z$. Pod koniec dowodu, we ciągu nierówności (4.22) indeks sumowania w pierwszym pojawiającym się szeregu to k, czyli powinno być $||\omega_{k+1}(\xi) - \omega_k(\xi)||$, szacowanie zaś tej sumy jest z dokładnością do stałej C, a nie $\frac{1}{C(1-\varrho)}$ (por. ostatni ciąg nierówności na dole strony 73).

W rozdziale piątym, na koniec dowodu faktu 5.1.1 należy powołać się na to, że $h_{x,y}$ ma absolutne minimum w $t = 0$, nie jest bowiem prawdą, by pochodna funkcji silnie wypukłej miała być dodatnia, jak pokazuje najprostszy przykład silnie wypukłej t^2 (chyba, że chodziło o dodatniość na prawo od minimum). W związku z wnioskiem 5.1.4 pojawia się naturalne pytanie, czy $f(\kappa_N^\nu(\xi))$ może maleć do $-\infty$ (ciąg iteracyj nie musi być zbieżny), nierozstrzygnięte w pracy. Dowód wniosku 5.2.1 mógłby zawierać więcej szczegółów, nawet jeśli powtarza się argumentacja z 4.1.3 (warto byłoby może rozpisać $\nabla \ln \varphi_{N,\xi}$).

Przejdźmy teraz do omówienia kilku problemów w tym rozdziale. Zaznaczmy przy tym, że mowa tu o jedynej części rozprawy już opublikowanej (w [1]). Pierwsze wątpliwości napotykamy w lemacie 6.1.4, gdzie bezpośrednie zastosowanie lematu 6.1.3 jest niemożliwe, skoro b jest tylko silnie wypukła a nie logarytmicznie silnie wypukła, odwołanie natomiast do lematu 1.4.8 daje zaledwie lokalną injektywność, gdy zapowiadany jest wynik globalny. Krótko mówiąc, dowód wymaga dalszego rozpisania. W podrozdziale 6.3 zakłada się cichcem, że $X_{f\leqslant r}$ jest zwarty, dobrze byłoby to jasno napisać. Na dole strony 86 należy odwołać się nie do faktu 4.1.1, lecz do 6.3.1, ale to błąd zecerski. We wzorze (6.19) w miejsce ν powinno się było raczej znaleźć 2. Uwaga 6.3.4 jest nadto lakoniczna i wymaga dalszego uzasadnienia — nie jest bowiem jasne, ani co rozumiemy pod pojęciem funkcji analitycznej na zbiorze, który nie posiada struktury analitycznej, ani dlaczego właściwie κ_N jest gradientem takiej funkcji. Wydaje się, że chodziło raczej o funkcję κ_N rozpatrywaną na całej kuli B_R.

Wniosek 6.3.5 jest podobnie tajemniczy, odwołanie do twierdzenia 4.2.1 nie ma sensu, jako że dotyczy ono zupełnie innej funkcji b i innej techniki uwypuklania niż aktualnie rozważana, a i lemat 1.6.4 też niewiele wnosi; prawdopodobnie należało przytoczyć wniosek 5.2.1, choć i on nie jest wystarczający. Da się go jednak zastosować, funkcja $b(x) = \exp(||x||^2)$ ma tę własność, że $b(x)^N = b(\sqrt{N}x)$, dzięki czemu na mocy wniosku 5.2.1 można stwierdzić, że dla $N \gg 1$, κ_N określona na $\mathbb{R}^n$ jest dyfeomorfizmem. Problem jednak w tym, że κ_N rozpatruję się w rozdziale 6 jako minimum λ_N na B_R i stąd potrzeba doprecyzowania, czym jest R. W samej wypowiedzi wniosku nie jest jasne, czym jest liczba K_f (zapewne $K_f(0)$ ze strony 37) a i μ w rozpatrywanym przypadku wynosi 2. W pracy [1] argumentacja wykorzystuje pośrednio wyniki przedstawione w twierdzeniu 6.2.3.

W fakcie 6.3.6 pojawia się oznaczenie ω_*, którego znaczenia musimy się domyślać (na podstawie analogii do twierdzenie 4.2.1 (c)). Dalej, w twierdzeniu 6.3.10 zapewne zakłada się nadal zwartość $X_{f\leqslant r}$ choć nie jest to napisane (ale wspomniana jest analogia do twierdzenia 4.3.1). W rozprawie doktorskiej spodziewać by się można w miejsce uwagi 6.3.11 o możliwości podania kontrprzykładu raczej jego samego, tym bardziej, gdy jest określony jako łatwy do otrzymania.

Podrozdział 6.3.3, na szczęście zaniedbywalny z punktu widzenia całej rozprawy, niesie nowe wątpliwości, od mało istotnych (zapewne stopień f winien wynosić co najmniej 2 a założenie zwartości $X_{f\leqslant r}$ jest niewątpliwie konieczne, ale — jak rozumiem — obowiązuje w całym rozdziale 6), przez drobne usterki rachunkowe (zastosowanie uwagi 6.2.4 do oszacowania wielkości skoku ν-tej iteracji prowadzi do nierówności z dodatkowym czynnikiem r w mianowniku: $||\xi_{\nu+1} - \xi_\nu|| \geqslant \frac{||\nabla f(\xi_{\nu+1})||}{LNr}$), po bardziej subtelne, dotyczące zastosowania efektywnej nierówności Łojasiewicza z [10] Remark 4. Cytowana uwaga orzeka (i jest to wspomniane *expressis verbis* na stronie 93), że wykładnik Łojasiewicza gradientu wielomianu f stopnia $d \geqslant 2$ w otoczeniu zera tego ostatniego nie przekracza $(d-1)(6d-9)^{n-1}$, co oznacza, że w otoczeniu $\xi_* \in \Sigma_f$ spełniona jest nierówność

$$||\nabla f(x)|| \geqslant \text{const.dist}(x, \nabla f^{-1}(0))^{(d-1)(6d-9)^{n-1}},$$

a więc inna niż podano na stronie 90 rozprawy. Chcąc, ażeby pojawiła się w niej odległość od włókna $f^{-1}(f(\xi_*))$, musimy przejść przez klasyczną nierówność gradientową $||\nabla f(x)|| \geqslant \mathrm{c(z)}|f(x) - f(\xi_*)|^{\theta_z}$ rozważaną w otoczeniu punktu $z \in f^{-1}(f(\xi_*))$ (gradient znika wzdłuż tego włókna). Szczęśliwie włókno to, na mocy założeń, jest zwarte, a wykładniki $\theta_z \in [0,1]$ w małym otoczeniu możemy zwiększyć bezkarnie do 1, co oznacza, że wykorzystując efektywną nierówność Łojasiewicza [10] Corollary 5 dla f, otrzymamy wspólną stałą $C > 0$, przy której w otoczeniu $f^{-1}(f(\xi_*))$ dostaniemy

$$||\nabla f(x)|| \geqslant \text{const.dist}(x, \nabla f^{-1}(f(\xi_*)))^{d(3d-3)^{n-1}}.$$

Wykładnik wychodzi zatem inny niż podaje Autor rozprawy, co ma wpływ na ostateczne oszacowanie ze strony 91, chociaż nie na jego istotę. Skądinąd końcówka rozumowania na tej stronie (błednie określona w tekście (4.22), gdy chodzi o oszacowanie z góry strony 90 nie posiadające numeru) też nie jest poprawna, skoro Autor powołuje się na [10] Corollary 7, który to wynik dotyczy przypadku zespolonego i nie ma tu zastosowania. Podejrzewam, że należało przywołać oszacowanie D'Acunto-Kurdyki oznaczone (D-K) w [10]. Skądinąd, skoro nadal omawiany jest przypadek funkcji $\exp(||x||^2)$, to $\mu = 2$.

W przypadku podrozdziału 6.3.4 na dole strony 91 należy powołać się nie tyle na dowód lematu 6.1.3, co raczej 6.3.2, ewentualnie 6.3.7 (iii). W uwadze 6.3.14 warto raz jeszcze podkreślić, że mamy do czynienia z włóknem zwartym (Autor robi to faktycznie w kilku innych newralgicznych miejscach, a to jedno z nich). Czytelnikowi mogą się nieco „rozmywać" aktualne w danym podrozdziale założenia (przykładowo w uwadze 6.3.16 warto pamiętać, że f ma minimum w środku układu współrzędnych, choć niekoniecznie tylko w tym jednym punkcie).

Jedyna istotna i być może nie taka łatwa do usunięcia luka w rozdziale 6, znajduje się w rozważaniach nt. zbieżności części sferycznych ciągu $\kappa_N^\nu(\xi)$, czyli w podrozdziale 6.3.5. Problem dotyczy krzywej γ ze strony 94, określonej wzorem (6.21). Po zapisaniu jej we współrzędnych biegunowych $\gamma(t) = r_\gamma(t)\theta_\gamma(t)$, mamy naturalnie $\gamma'(t) = r'_\gamma(t)\theta_\gamma(t)+r_\gamma(t)\theta'_\gamma(t)$. Z definicji γ wiemy, że $\gamma'(t) = \xi_{\nu+1}-\xi_\nu$, gdy $t \in (\nu, \nu+1)$. Z drugiej strony, tak jak jest napisane w pracy, mamy $\xi_{\nu+1}-\xi_\nu = -\frac{1}{2Nf(\xi_{\nu+1})}\nabla f(\xi_{\nu+1})$, oraz, wykorzystując współrzędne biegunowe, $\nabla f(\xi_{\nu+1}) = \nabla' f(\xi_{\nu+1})+\partial_r f(\xi_{\nu+1})\frac{\xi_{\nu+1}}{||\xi_{\nu+1}||}$. Mamy tedy do czynienia z zapisem $\gamma'(t)$ (o stałej wartości $\xi_{\nu+1} - \xi_\nu$ dla $t \in (\nu+1, \nu)$) w dwu układach ortogonalnych: jednym, danym przez parę wektorów $(\xi_{\nu+1}, \nabla' f(\xi_{\nu+1}))$ niezależnych od $t \in (\nu, \nu + 1)$, oraz drugim, wyznaczonym przez parę $(\theta_\gamma(t), \theta'_\gamma(t))$, zależną od t (wektor $\gamma(t)$ sunie bowiem po odcinku $[\xi_\nu, \xi_{\nu+1}]$). W związku z tym podane na stronie 95 zależności na $r'_\gamma(t)$ i $r_\gamma(t)\theta'_\gamma(t)$ (wyrażone za pomocą składowych radialnej i sferycznej gradientu w $\xi_{\nu+1}$) mają sens wyłącznie w punkcie $t = \nu + 1$ (z pochodnymi jednostronnymi oczywiście, dzięki ciągłości), bynajmniej nie dla dowolnego $t \in (\nu, \nu + 1)$. Istotnie, nietrudno zauważyć, że $r'_\gamma(t)$ i $r_\gamma(t)\theta'_\gamma(t)$ mogą być stale równe, dla $t \in (\nu, \nu+1)$, podanym w pracy dwu wyrażenionom tylko wtedy, gdy wektory $\theta_\gamma(t)$ i $\xi_{\nu+1} - \xi_\nu$ są współliniowe (zauważmy, że $r'_\gamma(t) = \langle\theta_\gamma(t), \xi_{\nu+1} - \xi_\nu\rangle$), co poza skrajnym i całkiem trywialnym przypadkiem nie będzie zachodzić. Ten sam błąd znajduje się w [1], skąd zaczerpnięto rozumowanie. Co ciekawe, pożądane zależności uzyskane wyłącznie w punktach $t = \nu + 1$ wystarczają mimo wszystko do pokazania ścisłej monotoniczności ciągu $||\xi_\nu||$ (w punktach ξ_ν znajdują się bowiem węzły łamanej γ). Teraz jednak nie mamy oszacowania na $||\theta'_\gamma(t)||$, a jedyne, co możemy wywieść, to tyle, że dla pochodnych jednostronnych $||\theta'_{\gamma,-}(\nu+1)|| \leqslant C_3||r'_{\gamma,-}(\nu+1)|| \leqslant C_3||\xi_{\nu+1} - \xi_\nu|| = C_3||\gamma'(t)||$, co nie wystarcza do stwierdzenia, że θ_γ ma skończoną długość.

Łamana Θ ze strony 95, wpisana w θ_γ, powinna być zadana wzorem

$$\Theta(t) = \theta_\gamma(\nu) + (t-\nu)[\theta_\gamma(\nu+1) - \theta_\gamma(\nu)], \quad t \in [\nu, \nu+1)$$

(w rozprawie powielony jest omyłkowy zapis z [1]). Dla dokończenia dowodu wystarczyłoby wiedzieć, że Θ ma skończoną długość, nie wynika to jednak natychmiastowo z poprawionej wersji rozumowania. Ostatecznie trzeba by mieć oszacowanie na $||\theta'_\gamma(t)||$ przez normę pochodnej $||\gamma'(t)|| = ||\xi_{\nu+1} - \xi_\nu||$, by skorzystać z tego, że γ ma skończoną długość, lecz przeliczając wprost

$$\begin{aligned}\theta'_\gamma(t) = \left(\frac{\gamma(t)}{||\gamma(t)||}\right)' &= \frac{\gamma'(t)}{||\gamma(t)||} - \frac{\langle \gamma(t), \gamma'(t)\rangle}{||\gamma(t)||^3}\gamma(t) = \\ &= \frac{\xi_{\nu+1} - \xi_\nu}{r_\gamma(t)} - \frac{\langle \theta_\gamma(t), \xi_{\nu+1} - \xi_\nu\rangle}{r_\gamma(t)}\theta_\gamma(t),\end{aligned}$$

widzimy, jak bardzo bruździ czynnik $\frac{1}{r_\gamma(t)} = \frac{1}{||\gamma(t)||}$ rosnący do nieskończoności. Wydaje się, że usunięcie luki w dowodzie wymaga jeszcze pewnego nakładu pracy.

Przechodząc nareszcie do podsumowania należy raz jeszcze podkreślić, że tematyka rozprawy doktorskiej pana mgra Abdullaha doskonale wpisuje się w aktualnie prowadzone badania na pograniczu matematyki stosowanej, a konkretnie zagadnień optymalizacyjnych, oraz geometrii semi-algebraicznej. Autor rozprawy dobrze orientuje się w omawianej problematyce i wykazuje sporą biegłość rachunkową i wcale szeroką wiedzę, tak z teorii funkcyj wypukłych, jak i z geometrii semi-algebraicznej potrzebną do dowodów przedstawianych twierdzeń. Wyniki zawarte w rozprawie są ważne i ciekawie poszerzają stan obecnej wiedzy w zakresie uwypuklania funkcyj i algorytmów szukania punktów krytycznych. Ich dowody wymagały od Autora oryginalnego podejścia i umiejętności adaptacji istniejących technik. Wymienione przeze mnie wyżej usterki nie umniejszają istotnie znaczenia całej rozprawy, którą oceniam jako oryginalny i wartościowy wkład w ważną dziedzinę matematyki.

Stwierdzam tym samym, że recenzowana przez mnie praca pana mgra Abdulljabara Naji Ahmeda Abdullaha spełnia ustawowe wymogi stawiane rozprawom doktorskim i wnoszę o jej przyjęcie oraz o dopuszczenie pana mgra Adbullaha do dalszych etapów przewodu doktorskiego.

dr hab. Maciej Denkowski, prof. UJ