



AGH

Akademia Górniczo-Hutnicza
im. Stanisława Staszica
w Krakowie

Praca magisterska

Aproksymacja funkcji kawałkami regularnych przy
użyciu informacji dokładnej i niedokładnej

Tomasz Czyż

Kierunek: Matematyka
Specjalność: Matematyka obliczeniowa
i komputerowa

Nr albumu: 290565

Promotor
dr Maciej Goćwin



Wydział Matematyki Stosowanej

Kraków 2021

Oświadczenie studenta

Upředzony(-a) o odpowiedzialności karnej na podstawie art. 115 ust. 1 i 2 ustawy z dnia 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych (t.j. Dz.U. z 2018 r. poz. 1191 z późn. zm.): „Kto przywłaszcza sobie autorstwo albo wprowadza w błąd co do autorstwa całości lub części cudzego utworu albo artystycznego wykonania, podlega grzywnie, karze ograniczenia wolności albo pozbawienia wolności do lat 3. Tej samej karze podlega, kto rozpowszechnia bez podania nazwiska lub pseudonimu twórcy cudzy utwór w wersji oryginalnej albo w postaci opracowania, artystyczne wykonanie albo publicznie zniekształca taki utwór, artystyczne wykonanie, fonogram, wideogram lub nadanie.”, a także upředzony(-a) o odpowiedzialności dyscyplinarnej na podstawie art. 307 ust. 1 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2018 r. poz. 1668 z późn. zm.) „Student podlega odpowiedzialności dyscyplinarnej za naruszenie przepisów obowiązujących w uczelni oraz za czyn uchylający godności studenta.”, oświadczam, że niniejszą pracę dyplomową wykonałem(-am) osobiście i samodzielnie i nie korzystałem(-am) ze źródeł innych niż wymienione w pracy. Jednocześnie Uczelnia informuje, że zgodnie z art. 15a ww. ustawy o prawie autorskim i prawach pokrewnych Uczelni przysługuje pierwszeństwo w opublikowaniu pracy dyplomowej studenta. Jeżeli Uczelnia nie opublikowała pracy dyplomowej w terminie 6 miesięcy od dnia jej obrony, autor może ją opublikować, chyba że praca jest częścią utworu zbiorowego. Ponadto Uczelnia jako podmiot, o którym mowa w art. 7 ust. 1 pkt 1 ustawy z dnia 20 lipca 2018 r. — Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2018 r. poz. 1668 z późn. zm.), może korzystać bez wynagrodzenia i bez konieczności uzyskania zgody autora z utworu stworzonego przez studenta w wyniku wykonywania obowiązków związanych z odbywaniem studiów, udostępniać utwór ministrowi właściwemu do spraw szkolnictwa wyższego i nauki oraz korzystać z utworów znajdujących się w prowadzonych przez niego bazach danych, w celu sprawdzania z wykorzystaniem systemu antyplagiatowego. Minister właściwy do spraw szkolnictwa wyższego i nauki może korzystać z prac dyplomowych znajdujących się w prowadzonych przez niego bazach danych w zakresie niezbędnym do zapewnienia prawidłowego utrzymania i rozwoju tych baz oraz współpracujących z nimi systemów informatycznych.

.....
(Podpis czytelny studenta)

Oświadczenie promotora

Oświadczam, że praca spełnia wymogi stawiane pracom magisterskim.

.....
(Podpis promotora)

Spis treści

Streszczenie	2
Abstract	3
Wprowadzenie	4
Rozdział 1. Definicje	6
1.1. Informacja, algorytm, aproksymacja	6
1.2. Model obliczeniowy	8
1.3. Klasy funkcji	9
Rozdział 2. Ograniczenia na błąd	12
2.1. Ograniczenie z dołu	12
2.2. Ograniczenia z góry	15
Rozdział 3. Algorytmy	17
3.1. Algorytm oparty na wielomianach Lagrange’a	17
3.2. Algorytm oparty na różnicach dzielonych	18
Rozdział 4. Analiza algorytmów	21
4.1. Analiza algorytmu opartego o wielomiany Lagrange’a	21
4.2. Analiza algorytmu opartego o różnice dzielone	26
Rozdział 5. Testy numeryczne	34
Bibliografia	41

Streszczenie

Streszczenie

Słowa kluczowe

słowa kluczowe

Abstract

Abstract

Key words

keywords

Wprowadzenie

Aproksymacja funkcji oparta na dostępnej informacji jest problemem badanym od lat. Powstają coraz bardziej zaawansowane algorytmy, działające przy coraz słabszych założeniach o funkcji. Często jednak w rozważaniach teoretycznych pomijany jest czynnik zewnętrzny, który może powodować zaburzenia dostępnych informacji. W tej pracy rozważamy, jaki wpływ na wyniki numeryczne może mieć zaniedbanie tego faktu.

W rozważaniach zakładamy, że mamy dostęp tylko do częściowej informacji o funkcji, a jedynym źródłem informacji jest tzw. *wyrocznia*. To podejście ma praktyczne uzasadnienie w obliczeniach numerycznych, gdzie odwołania do wyroczni odpowiadają ewaluacjom funkcji. Na ogół wartości, które otrzymujemy są wynikami pewnych pomiarów (np. fizycznych), które zawsze są obarczone pewnym błędem. Uwzględnienie zaburzenia danych jest więc intuicyjne.

Dodatkowo, najczęściej spotykane dane cechują się pewnym stopniem nieregularności. Z tego powodu powstaje wiele prac, w których przyjmuje się słabsze założenia na aproksymowaną funkcję. W tej pracy rozważać będziemy funkcje, które zawierają dokładnie jeden, nieznany punkt osobliwy, w którym nie musi być zachowana ciągłości czy różniczkowalność.

Mówimy, że funkcja skalarna g jest (r, ϱ) -regularna na przedziale $[a, b]$, jeśli $g \in C^r([a, b])$ oraz $g^{(r)}$ jest Hölderowsko ciągła z wykładnikiem $\varrho \in (0, 1]$. Rozważmy przestrzeń $F_{r, \varrho}$ T -okresowych funkcji f , które składają się z dwóch (r, ϱ) -regularnych części oddzielonych nieznanym punktem osobliwym \hat{t}_f . Założenie o okresowości funkcji zostało wprowadzone, aby uprościć prezentację problemu. Po technicznych modyfikacjach wszystkie wyniki zachodzą dla funkcji nieokresowych.

W celu porównania wpływu zaburzenia informacji przedstawimy dwa algorytmy z artykułów [5] i [4]. Pierwszy z nich bazuje na wielomianach Lagrange’a i jego analiza nie uwzględnia zaburzenia danych wejściowych. Algorytm z [5] dopuszcza informacje niedokładną, a kluczowy krok opiera się na różnicach dzielonych.

Oba algorytmy stosują podejście adaptacyjne do wybierania dodatkowych ewaluacji funkcji, to znaczy, wybór kolejnych punktów jest uzależniony od wcześniejszych wartości. Skuteczność algorytmów adaptacyjnych i nieadaptacyjnych została szeroko przeanalizowana dla wielu klas funkcji przy założeniu, że informacja jest dokładna. Użycie adaptacji w obu z omawianych algorytmów jest uzasadnione wynikami z m.in. [2]. W artykule pokazano, że błąd L_p -aproksymacji funkcji z jednym punktem osobliwym dla algorytmów nieadaptacyjnych używających n ewaluacji funkcji nie może być

lepszy niż $n^{1/p}$, przy czym algorytmy adaptacyjne osiągają optymalne tempo zbieżności n^{-r} .

W tej pracy udowodniono również, że dla funkcji posiadających wiele punktów osobliwych tempo zbieżności błędu najgorszego przypadku względem normy L_p maleje do $n^{1/p}$. Z tego powodu skupiamy się na funkcjach z jedną osobliwością. Wyniki te przytoczymy w rozdziale 2.

W rozdziale 3 przedstawimy algorytmy z prac [5] i [4].

Rozdział 1

Definicje

1.1. Informacja, algorytm, aproksymacja

W tym rozdziale wyjaśnimy co rozumiemy przez aproksymację i w jaki sposób ją otrzymujemy. W tym celu wprowadzimy fundamentalne pojęcia, takie jak operator rozwiązania, informacja niedokładna oraz algorytm. Szczególną uwagę poświęcimy informacji która, mówiąc w skrócie, jest tym co wiemy o problemie do rozwiązania. Głównym założeniem tej pracy jest warunek, że informację o funkcji otrzymujemy z pewnym błędem. Taką informację nazywamy *niedokładną* lub *zaburzoną*.

Niech F będzie przestrzenią liniową a G przestrzenią unormowaną, obie nad ciałem liczb rzeczywistych. Odwzorowanie

$$S : F \rightarrow G$$

nazywamy *operatorem rozwiązania*. Elementy f z F nazywane są *problemami*, a elementy $S(f)$ *rozwiązaniami*. Dla każdego problemu f chcemy obliczyć aproksymację rozwiązania $S(f)$. Niech $U(f)$ będzie obliczoną aproksymacją.

Niech $\varepsilon \geq 0$. Mówimy, że $U(f)$ jest ε -aproksymacją problemu f wtedy i tylko wtedy, gdy $\|S(f) - U(f)\| \leq \varepsilon$. Celem jest znalezienie takiej aproksymacji $U(f)$, że jest ona ε -aproksymacją dla wszystkich elementów f z F . Aby to zrobić potrzebujemy posiadać pewną wiedzę o f .

Operatorem informacji (lub informacją) nazywamy odwzorowanie

$$N : F \rightarrow 2^Y,$$

gdzie Y jest zbiorem skończonych ciągów liczb rzeczywistych, $Y \subset \bigcup_{n=1}^{\infty} \mathbb{R}^n$, czyli $N(f)$ jest podzbiorem Y . Na ogół nie mamy dostępu do pełnej wiedzy o funkcji, dlatego musimy założyć, że możemy zbierać informacje o f poprzez pewien funkcjonal $L(f)$, gdzie $L : F \rightarrow \mathbb{R}$.

Przez Λ oznaczmy klasę dopuszczalnych operacji L , czyli $L \in \Lambda$ wtedy i tylko wtedy, gdy $L(f)$ może zostać obliczone dla każdego elementu f z F . Rozważmy teraz dwa sposoby doboru informacji. Informację N nazywamy nieadaptacyjną wtedy i tylko wtedy, gdy istnieje $L_1, \dots, L_n \in \Lambda$ takie, że

$$N(f) = [L_1(f), L_2(f), \dots, L_n(f)] \quad \forall f \in F.$$

W tym przypadku poszczególne funkcjonały informacji zależą tylko od elementu f i są obliczane niezależnie. Ogólniejszą klasą jest informacja adaptacyjna, w której możemy wybierać wartości funkcjonałów bazując na poprzednich wynikach. Mówiąc dokładniej, informacja N jest adaptacyjna wtedy i tylko wtedy, gdy

$$N(f) = [L_1(f), L_2(f; y_1), \dots, L_i(f; y_1, \dots, y_{n(f)-1})],$$

gdzie $y_1 = L(f_1)$ i $y_i = L_i(f; y_1, y_2, \dots, y_{i-1})$ dla $i = 2, 3, \dots, n(f)$. Musimy również założyć, że $L_i(\cdot; y_1, \dots, y_{i-1})$ należą do operacji dozwolonych. W przypadku informacji adaptacyjnej nie możemy z góry określić liczby operacji na problemie f , ponieważ jest to dynamicznie ustalane podczas procesu obliczania kolejnych wartości y_i .

Warto zauważyć, że jeśli rozważane zagadnienie wymaga obliczenia bardzo dużej ilości informacji o problemie w krótkim czasie, to zastosowanie informacji nieadaptacyjnej może przyspieszyć proces, ze względu na możliwość zrównoleglenia obliczeń. W przypadku adaptacyjnym kolejność obliczeń ma znaczenie, więc informacje musimy pozyskiwać sekwencyjnie, co zazwyczaj jest wolniejsze.

W niniejszej pracy zakładamy, że f jest funkcją, a jedynymi dostępnymi funkcjonałami informacji są wartości funkcji w punktach. Informację nieadaptacyjną oraz adaptacyjną możemy więc zapisać w postaci $N(f) = [f(t_1), f(t_2), \dots, f(t_{n(f)})]$. Jeżeli $n(f) = n$ i punkty t_i otrzymujemy *a priori*, wtedy N jest nieadaptacyjna. Natomiast, jeżeli $n(f)$ różni się lub wybór punktów t_i jest zależny od $f(t_1), f(t_2), \dots, f(t_{i-1})$, to N jest adaptacyjna.

Kluczowym założeniem jest fakt, że informacje o funkcji uzyskujemy z pewnym błędem. Mówiąc dokładniej, informacje o funkcji przyjmują postać $y_i = f(x_i) + e_i$, dla $1 \leq i \leq n$, gdzie $|e_i| \leq \delta$ to tak zwany *szum*.

Dla przykładu, dla informacji nieadaptacyjnej składającej się z zaburzonych ewaluacji funkcji f w punktach t_1, \dots, t_n z precyzją δ , mamy

$$N(f) = \{y \in \mathbb{R}^n : |y_i - f(t_i)| \leq \delta, \quad 1 \leq i \leq n\}.$$

Każdy element $y \in N(f)$ będziemy nazywać *informacją o f* . Zauważmy, że dla $\delta = 0$, zbiór $N(f)$ ma dokładnie jeden element dla wszystkich $f \in F$, tzn. informacja N jest *dokładna*. Zakładamy, że $N(f)$ jest niepuste dla wszystkich $f \in F$. W przypadku, gdy istnieje f dla którego $N(f)$ ma przynajmniej dwa elementy, wtedy informacja jest *niedokładna*.

Znając informację $N(f)$ możemy obliczyć aproksymację f używając *algorytmu* φ , w postaci odwzorowania

$$\varphi : Y \rightarrow G.$$

Dostajemy, że aproksymacja $U(f) = \varphi(N(f))$, czyli algorytm φ używa informacji $N(f)$, aby wyprodukować element z G , który aproksymuje rozwiązanie $S(f)$.

W tej pracy rozważamy zagadnienie aproksymacji pewnych funkcji skalarnych, używając skończonej liczby wartości funkcji, przy założeniu, że błąd aproksymacji jest mierzony w normie L^p ($1 \leq p \leq \infty$). A zatem operator rozwiązania S jest odwzorowaniem identycznościowym prowadzącym z przestrzeni funkcji F w przestrzeń L^p , natomiast błąd aproksymacji $f \in F$ otrzymanej algorytmem $\varphi : Y \rightarrow L^p(0, T)$ z użyciem informacji $y \in N(f)$ wynosi

$$\|f - \varphi(y)\|_{L^p} = \left(\int_0^T |(f - \varphi(y))(x)|^p dx \right)^{1/p} \quad \text{dla } 1 \leq p < \infty$$

oraz

$$\|f - \varphi(y)\|_{L^\infty} = \operatorname{ess\,sup}_{0 < x \leq T} |(f - \varphi(y))(x)|.$$

Dla przykładu, niech F będzie przestrzenią ciągłych funkcji rzeczywistych dwukrotnie różniczkowalnych $f : [0, 1] \rightarrow \mathbb{R}$, a $G = L_2(0, 1)$. To oznacza, że $S(f) = f$, natomiast dla $0 \leq t_1 \leq \dots \leq t_n \leq 1$, y odpowiada n obserwacjom $f(t_i)$ dla $1 \leq i \leq n$, z precyzją δ każda. Przykładem algorytmu aproksymującego funkcję f może być konstrukcja funkcji sklepanej $l : [0, 1] \rightarrow \mathbb{R}$, czyli takiej funkcji, że l jest wielomianem na każdym podprzedziale (t_i, t_{i+1}) podziału $t_1 < \dots < t_n$.

1.2. Model obliczeniowy

W ogólności, optymalność algorytmu oraz złożoność problemu zależą od przyjętego modelu obliczeniowego. Model jest określony poprzez sposób w jaki błąd i koszt algorytmu są zdefiniowane.

Jeżeli za błąd i koszt przyjmujemy wydajność na najtrudniejszym spośród wszystkich problemów w danej klasie, wtedy mówimy o *modelu najgorszego przypadku*. Innymi często rozważanymi modelami są: model probabilistyczny, średni, mieszany, losowy czy asymptotyczny, jednak nimi nie będziemy zajmować się w tej pracy.

Niech $N : F \rightarrow 2^Y$ będzie operatorem informacji a $S : F \rightarrow G$ operatorem rozwiązania dla $G = L^p$. Poprzez *błąd najgorszego przypadku* algorytmu $\varphi : Y \rightarrow L^p$ na zbiorze $E \subset F$ rozumiemy:

$$e_p^{\text{wor}}(\varphi, N, E) = \sup_{f \in E} \sup_{y \in N(f)} \|S(f) - \varphi(y)\|_{L^p}$$

Celem jest znalezienia algorytmu, który minimalizuje błąd najgorszego przypadku względem wszystkich algorytmów w danej klasie. Algorytm osiągający to minimum nazywamy *optymalnym*.

W tej pracy rozważamy algorytmy aproksymujące funkcje z osobliwością, które wykorzystują określoną liczbę wartości funkcji. Oznaczmy przez $\mathcal{N}(n, \delta)$ klasę wszystkich (adaptacyjnych) informacji N , które używają co najwyżej n ewaluacji funkcji, z precyzją δ każda. Wtedy przez *minimalany błąd najgorszego przypadku* w klasie E , który może zostać osiągnięty przez algorytm używający informacji o co najwyżej n wartościach funkcji z precyzją δ rozumiemy:

$$r_p^{\text{wor}}(n, \delta, E) = \inf \{e_p^{\text{wor}}(\varphi, N, E) : \varphi \text{ używa } N \in \mathcal{N}(n, \delta)\}$$

W tej pracy porównujemy dwa algorytmy aproksymujące funkcje z osobliwością, które osiągają błąd na takim samym, optymalnym poziomie błędu najgorszego przypadku, jednak tylko w przypadku jednego z nich w analizie zostało uwzględnione zaburzenie informacji.

Drugą cechą modelu obliczeniowego jest koszt algorytmu. Może się na niego składać zarówno koszt uzyskania informacji o wartości funkcji, jak i koszt operacji arytmetycznych potrzebnych przy obliczeniach algorytmu. Ewaluacje wartości funkcji w modelach, które zakładają ograniczoną ich liczbę, często odnoszą się do wykonywania jakichś pomiarów. Są one nierzadko pracochłonne i drogie. Oznacza to, że koszt zdobycia informacji o funkcji jest nieporównywalnie większy niż koszt działań arytmetycznych na komputerze. Z tego powodu przyjmujemy, że koszt algorytmu φ jest równy tylko kosztowi uzyskania informacji o wartości funkcji f , czyli

$$\text{cost}(\varphi, f) = n,$$

Koszt algorytmu w całej klasie F ma postać

$$\text{cost}(\varphi, F) = \sup_{f \in F} \text{cost}(\varphi, f)$$

Niech $\varepsilon > 0$. ε -złożoność najgorszego przypadku klasy F mierzymy w następujący sposób

$$\text{comp}(\varepsilon, F) = \inf \{ \text{cost}(\varphi, F), \mid \varphi \text{ - algorytm oraz } \sup_{f \in F} \sup_{t \in [a, b]} \|f(t) - \varphi(t)\| \leq \varepsilon \}.$$

Celem jest otrzymanie ścisłych granic na $\text{comp}(\varphi, F)$, gdy $\varepsilon \rightarrow 0$.

Do określania błędu będziemy używać notacji Ω , Θ , \mathcal{O} , o (wersji Knutha). Gdy dwie funkcje f i g zdefiniowane na \mathbb{R}_+ i przyjmują wartości nieujemne, to piszemy $f(x) = \Omega(g(x))$, jeżeli istnieją dodatnie stałe c_1 i c_2 , takie że $f(x) \geq c_1 g(x)$ dla $x \in [0, c_2]$. Przez $f(x) = \Theta(g(x))$ rozumiemy $f(x) = \Omega(g(x))$ oraz $g(x) = \Omega(f(x))$, czyli istnieją takie stałe c_1 , c_2 i c_3 , że $c_1 g(x) \leq f(x) \leq c_2 g(x)$ dla $x \in [0, c_3]$. Natomiast przez $f(x) = \mathcal{O}(g(x))$ rozumiemy, że $g(x) = \Omega(f(x))$, a przez $f(x) = o(g(x))$ rozumiemy, że $\lim_{x \rightarrow 0} \frac{f(x)}{g(x)} = 0$.

1.3. Klasy funkcji

Wprowadzimy teraz dwie definicje klasy funkcji kawałkami regularnych użyte w artykułach [4] i [5].

Przez funkcje r -regularne rozumiemy funkcje, których pochodne do rzędu r istnieją w całej dziedzinie oraz są one ciągłe.

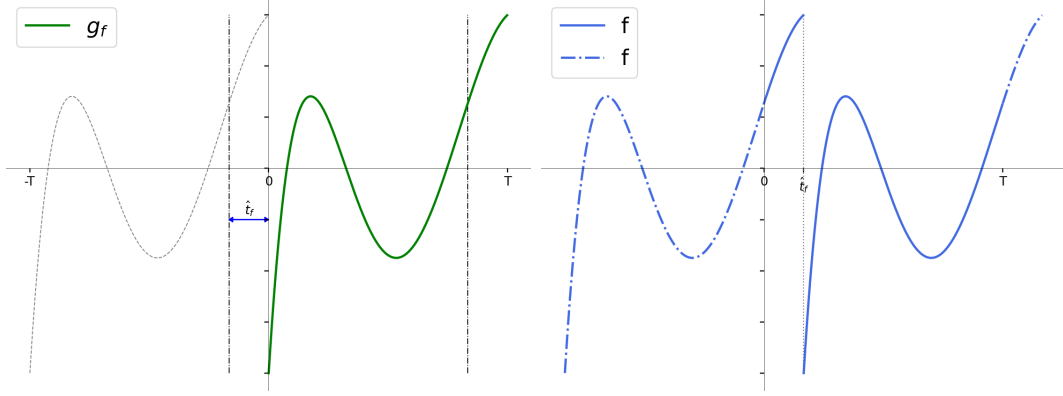
Dla liczby całkowitej $r \geq 0$, $\varrho \in (0, 1]$ oraz $a < b$, przez $H_{r, \varrho}(a, b)$ oznaczamy klasę funkcji $g : [a, b] \rightarrow \mathbb{R}$ takich, że $g \in C^r([a, b])$ i $g^{(r)}$ jest Hölderowsko ciągła z wykładnikiem ϱ , tzn.

$$c(g) := \sup_{a \leq x \leq y \leq b} \frac{|g^{(r)}(x) - g^{(r)}(y)|}{|x - y|^\varrho} < \infty.$$

Dla danego $T > 0$ niech $F_{r, \varrho} = F_{r, \varrho}(T)$ będzie przestrzenią funkcji $f : \mathbb{R} \rightarrow \mathbb{R}$ spełniających następujące warunki: istnieje $\hat{t}_f \in [0, T]$ i $g_f \in H_{r, \varrho}(0, T)$ takie, że

$$f(lT + \hat{t}_f + x) = g_f(x) \quad \text{dla każdego} \quad l = 0, \pm 1, \pm 2, \dots \quad \text{i} \quad x \in [0, T]$$

Można powiedzieć, że f jest 'kopia' g_f na każdym z przedziałów $(lT + \hat{t}_f, (l+1)T + \hat{t}_f]$ i f jest prawostronnie ciągła w $lT + \hat{t}_f$. W związku z tym wszystkie punkty, które różnią się między sobą o wielokrotność T będą uważane za identyczne. Dla przykładu, jeżeli $0 < x_1 \leq T < x_2 \leq 2T$, to przedział $(x_1, x_2]$ będzie utożsamiany z $(x_1, T] \cup (0, x_2 - T] \subset (0, T]$.



Przez $\Delta_f^{(j)}$ oznaczmy *skoki nieciągłości* dla kolejnych pochodnych f w punkcie nieciągłości \hat{t}_f ,

$$\Delta_f^{(j)} = f^{(j)}(\hat{t}_f^+) - f^{(j)}(\hat{t}_f^-) \quad 0 \leq j \leq r,$$

Ze względu na zachowanie się funkcji w punkcie osobliwym, rozróżniamy następujące klasy funkcji

$$\begin{aligned} \mathcal{K} &= \{f = c\mathbb{1}_{\mathbb{R}}, c \in \mathbb{R}\}, \\ \mathcal{H}_{r,\varrho} &= \{f \in F_{r,\varrho} : c(g_f) \leq 1, \Delta_f^{(j)} = 0 \text{ dla każdego } 0 \leq j \leq r\} \\ \mathcal{F}_{r,\varrho}^C &= \{f \in F_{r,\varrho} : c(g_f) \leq 1, \Delta_f^{(0)} = 0\} \\ \mathcal{F}_{r,\varrho}^D &= \{f \in F_{r,\varrho} : c(g_f) \leq 1, |\Delta_f^{(0)}| \leq 1\} \\ \mathcal{F}_{r,\varrho} &= \{f \in F_{r,\varrho} : c(g_f) \leq 1\}. \end{aligned}$$

Oczywiście

$$\mathcal{K} \subset \mathcal{H}_{r,\varrho} \subset \mathcal{F}_{r,\varrho}^C \subset \mathcal{F}_{r,\varrho}^D \subset \mathcal{F}_{r,\varrho}.$$

Algorytm przedstawiony w [4], oryginalnie bazuje na lekko zmodyfikowanych klasach funkcji, co wynika z innej natury problemu rozważanego w pracy. Dla spójności i lepszego przedstawiania problemu wprowadzimy oryginalnie rozważane klasy funkcji i sprecyzujemy różnice między klasami przedstawionymi wcześniej.

Niech $L_0, L_r, D_0, D_1, \dots, D_r$ będą dodatnimi stałymi. Wprowadźmy klasę funkcji regularnych

$$\begin{aligned} G_{r,\varrho}^{\text{reg}}([a, b]) &= \{g : [a, b] \rightarrow \mathbb{R}^d \mid g \in C^{(r)}([a, b]), \|g^{(j)}(x)\| \leq D_j, j = 0, 1, \dots, r \\ &\quad \|g^{(r)}(x) - g^{(r)}(y)\| \leq L_r |x - y|^\varrho, \|g(y) - g(x)\| \leq L_0 |y - x|, x, y \in [a, b]\}. \end{aligned}$$

Rozważmy następującą klasę $\mathcal{G}_{r,\varrho}([a, b])$ funkcji $g : [a, b] \rightarrow \mathbb{R}^d$ z co najwyżej jednym (nieznany) punktem osobliwym \hat{t}_g . Funkcja $g : [a, b] \rightarrow \mathbb{R}^d, g = [g^1, g^2, \dots, g^d]^T$, należy do $\mathcal{G}_{r,\varrho}([a, b])$, jeśli istnieje punkt $\hat{t}_g \in (a, b)$ taki, że

$$g \in G_{r,\varrho}^{\text{reg}}([a, \hat{t}_g]) \cap G_{r,\varrho}^{\text{reg}}([\hat{t}_g, b])$$

oraz lewostronna granica każdej składowej pochodnej $(g^k)^{(j)}$ istnieje w \hat{t}_g . W punkcie osobliwym \hat{t}_g , pochodne są rozumiane jako prawostronne.

Dla funkcji $g \in \mathcal{G}_{r,\varrho}([a, b])$ definiujemy wielomian

$$s_g(t) = \sum_{j=0}^r \frac{1}{j!} \Delta_g^{(j)} (t - \hat{t}_g)^j, \quad t \in [a, b], \quad (1.1)$$

gdzie $\Delta_g^{(j)}$ jest wektorem skoków w punkcie osobliwym zdefiniowanym jak wcześniej. Zauważmy, że $g \in \mathcal{G}_{r,\varrho}$ wtedy i tylko wtedy, gdy

$$g(t) = h(t) + \mathbb{1}_{[\hat{t}, b]}(t) \cdot s_g(t), \quad t \in [a, b], \quad (1.2)$$

gdzie $h \in G_{r,\varrho}^{\text{reg}}$.

Jeżeli $\Delta_g^{(j)} = 0$ dla wszystkich $j = 0, 1, \dots, r$, wtedy g jest regularna, czyli $g \in G_{r,\varrho}^{\text{reg}}([a, b])$. Jeżeli $g \in \mathcal{G}_{r,\varrho}([a, b])$ i $\Delta_g^{(0)} = 0$, wtedy g jest Lipschitzowsko ciągła w $[a, b]$.

Definicja klasy $G_{r,\varrho}^{\text{reg}}$ różni się od definicji $F_{r,\varrho}$ dopuszczeniem wielowymiarowości funkcji g , bardziej ogólnym podejściem do parametrów klasy oraz brakiem założenia o okresowości. Przyjęcie, że funkcje prowadzą w \mathbb{R} ułatwia rozważania praktyczne, a wyniki teoretyczne można łatwo uogólnić. W przypadku drugiej różnicy, ponieważ w tej pracy skupiamy się na wynikach numerycznych, przyjęcie za jedyny parametr klasy $c(g_f) = 1$ jest uzasadnionym uproszczeniem. Funkcję zawsze możemy przemnożyć przez stałą, aby otrzymać odpowiedni parameter. Natomiast założenie dotyczące okresowości ma jedynie na celu uprościć analizę algorytmu i może zostać łatwo uogólnione tak, aby wszystkie dowody dotyczące funkcji okresowych zachodziły również dla funkcji nieokresowych.

A zatem wszystkie różnice da się wyeliminować przy pomocy kilku technicznych zabiegów. Można powiedzieć, że z dokładnością do wymienionych różnic, klasa $\mathcal{G}_{r,\varrho}$ odpowiada kalsie $\mathcal{F}_{r,\varrho}^D$.

Ograniczenia na błąd

2.1. Ograniczenie z dołu

Na początku przytoczymy znane wyniki dotyczące ograniczeń z dołu dla informacji dokładnej, które uzasadniają użycie algorytmów adaptacyjnych. Wiadomo, że w klasie funkcji r -regularnych, najlepszym tempem zbieżności dla błędu jest n^{-r} . Pokazano to m.in. w [2], gdzie skonstruowano algorytm zachowujący tę własność. W tej samej pracy udowodniono również, że wprowadzenie osobliwości, powoduje pogorszenie się tempa zbieżności dla algorytmów nieadaptacyjnych do $n^{1/p}$. Pokazuje to następujące twierdzenie z [2].

Twierdzenie 2.1. *Niech φ_n będzie dowolnym algorytmem nieadaptacyjnym korzystającym z n ewaluacji funkcji oraz niech $\Delta > 0$. Istnieje kawałkami stała funkcja $f \in F_{r,1}$ taka, że $|\Delta_f^{(0)}| \leq \Delta$ oraz*

$$\|f - \varphi_n f\|_{L^p} \geq \frac{1}{2} \Delta \left(\frac{T}{n+1} \right)^{1/p}.$$

Dowód. Załóżmy, że φ_n oblicza f w punktach $x_0 < \dots < x_n$. Niech $x_0 = 0$ i $x_n = T$. Istnieją $0 < a < b < T$ takie, że $b - a \geq T/(n+1)$ i $[a, b] \subset [x_k, x_{k+1}]$ dla pewnego $0 \leq k \leq n-1$. Weźmy teraz funkcje $g_1 = \Delta \mathbb{1}_{(a,T]}$ oraz $g_2 = \Delta \mathbb{1}_{(b,T]}$. Ponieważ g_1 i g_2 używają tej samej informacji, tj. $g_1(x_i) = g_2(x_i)$ dla wszystkich $1 \leq i \leq n$ i $\|g_1 - g_2\|_{L^p} \geq \Delta(T/(n+1))^{1/p}$, to błąd algorytmu nie może być mniejszy niż $\Delta(T/(n+1))^{1/p}/2$ dla przynajmniej jednej z funkcji g_i , co dowodzi tezę. \square

W artykule [2] pokazano również ograniczenia dla algorytmów nieadaptacyjnych w przypadku wielu punktów osobliwych. Wprowadźmy dodatkowe oznaczenia, aby przytoczyć niektóre z tych wyników.

Oznaczmy przez $F_{r,1}^\ell$ klasę funkcji $f : [0, T] \rightarrow \mathbb{R}$ kawałkami gładkich z co najwyżej ℓ punktami osobliwymi. Precyzując, istnieje liczba całkowita ℓ , punkty $0 = \hat{t}_0 < \hat{t}_1 < \dots < \hat{t}_\ell < \hat{t}_{\ell+1} = T$ oraz $g_i(\hat{t}_i, \hat{t}_{i+1}) \in H_{r,1}(\hat{t}_i, \hat{t}_{i+1})$, takie że dla każdego $0 \leq i \leq \ell$ i $t \in (\hat{t}_i, \hat{t}_{i+1})$ zachodzi

$$f(t) = g_i(t)$$

oraz $f(0) = g_0(0)$, $f(T) = g_\ell(T)$ i f jest lewo lub prawostronnie ciągła z każdym \hat{t}_i . Wybierzmy teraz funkcje o interesującej nas regularności. Zdefiniujmy

$$\mathcal{F}_{r,1}^\ell = \mathcal{F}_{r,1}^\ell(L_r, L_1, D_0) := \left\{ f \in F_{r,1}^\ell \mid \|f^{(r)}\|_{L^\infty} \leq L_r, \|f'\|_{L^\infty} \leq L_1, \|\bar{\Delta}_f^{(0)}\|_q \leq D_0 \right\},$$

gdzie $\bar{\Delta}_f^{(0)} = (\Delta_1^{(0)}, \dots, \Delta_{k_f}^{(0)}) \in \mathbb{R}^\ell$ jest wektorem wszystkich skoków nieciągłości funkcji f oraz

$$\|\bar{\Delta}_f^{(0)}\|_q = \begin{cases} \left(\sum_{j=1}^{\ell} |\Delta_j^{(0)}|^q \right)^{1/q} & \text{dla } 1 \leq q < \infty, \\ \max_{1 \leq j \leq \ell} |\Delta_j^{(0)}| & \text{dla } q = \infty. \end{cases}$$

Poniższe stwierdzenie, pochodzące w [2], uzasadnia dlaczego skupiamy się na analizie funkcji z tylko jednym punktem osobliwym.

Stwierdzenie 2.2. *Niech $2 \leq \ell \leq \infty$ i $1 \leq q \leq \infty$. Dla każdego (adaptacyjnego) algorytmu φ_n używającego n ewaluacji funkcji, mamy*

$$\sup_{f \in \mathcal{F}_{r,1}^\ell} \|f - \varphi_n f\|_{L^p} \geq D_0 2^{-1/q} \lfloor \ell/2 \rfloor^{1-1/q} \left(\frac{T}{n+1} \right)^{1/p}$$

(stosując konwencję $\infty^0 = 1$ oraz $\infty^a = \infty$ dla $a > 0$).

Dowód. Niech $x_1 \leq x_2 \leq \dots \leq x_n$ będą punktami, w których obliczamy wartości funkcji $f \equiv 0$. Wybierzmy $0 \leq a < b \leq T$ takie, że $b - a \geq T/(n+1)$ oraz $[a, b] \subset [x_s, x_{s+1}]$ dla pewnego s , gdzie $x_0 = 0, x_{n+1} = T$.

Dla $2 \leq \ell < \infty$, weźmy $s := \lfloor \ell/2 \rfloor$ i $\Delta := D_0(2k)^{-1/q}$ oraz zdefiniujmy

$$f^* := \Delta \sum_{j=1}^s \mathbb{1}_{(a+\omega j, b-\omega j)} \quad \text{dla } 0 < \omega < \frac{b-a}{2s}.$$

Zauważmy, że $f_1 := f^*$ i $f_{-1} := -f^*$ dzielą wspólną informację i obie należą do $\mathcal{F}_{r,1}^\ell$. Dodatkowo zauważmy, że $f_1(x) = \Delta s$ dla każdego $x \in (a + \omega s, b - \omega s)$. Stąd, dla każdego ω , błąd najgorszego przypadku w normie L^p algorytmu φ_n jest ograniczony poprzez

$$\frac{1}{2} \|f_1 - f_{-1}\|_{L^p} \geq \Delta k(b - a - 2\omega k)^{-1/p} \geq D_0 2^{-1/q} \lfloor \ell/2 \rfloor^{1-1/q} \left(\frac{T}{n+1} - 2\omega \left\lfloor \frac{\ell}{2} \right\rfloor \right)^{1/p}.$$

Gdy weźmiemy $\ell \rightarrow \infty$ otrzymamy ograniczenie dla $\ell = \infty$. □

Z stwierdzenia 2.2 wynika, że nawet z użyciem adaptacji, nie jesteśmy w stanie osiągnąć lepszego ograniczenia z dołu na błąd aproksymacji niż $n^{1/p}$.

Zanim przejdziemy do rozważań dotyczących informacji zaburzonej, sformułujmy lemat dotyczący ograniczenia z dołu dla informacji dokładnej. Lemat ten pochodzi z [7], a jego dowód opiera się na własnościach informacji.

Lemat 2.3. *Założmy, że $f_1, \dots, f_{n+1} \in F$ mają rozłączne nośniki i $n \in \mathbb{N}$. Dodatkowo, niech $-f_i \in F$ oraz $\|f_i\| \geq \hat{\kappa}$ dla każdego $i = 0, \dots, n+1$. Wtedy $r(n, 0, F) \geq \hat{\kappa}$.*

Wiemy już jakie minimalne błędy mogą zostać osiągnięte przez algorytmy aproksymujące bazujące na informacji dokładnej. Poniższe stwierdzenie, przedstawione w [5], wprowadza kilka własności problemu aproksymacji przy obecności zaburzenia informacji.

Stwierdzenie 2.4. *Dla każdego n i $\delta \geq 0$ mamy:*

- (i) $r_p^{\text{wor}}(n, \delta, \mathcal{K}) \geq \delta T^{1/p}$,
- (ii) $r_p^{\text{wor}}(n, \delta, \mathcal{H}_{r,\varrho}) \geq a_{r,\varrho} n^{-(r+\varrho)}$, dla pewnego $a_{r,\varrho} > 0$,
- (iii) $r_p^{\text{wor}}(n, \delta, \mathcal{F}_{r,\varrho}) = \infty$, $r \geq 1$.

Dowód. W celu udowodnienia (i) wystarczy zauważyć, że $y = (0, \dots, 0)$ jest informacją o funkcji stałej postaci $f_{\pm} = \pm\delta$ dla każdego N z precyzją δ . Wynika z tego, że błąd dowolnego algorytmu używającego N jest równy co najmniej $\|f_{+\delta} - f_{-\delta}\|_{L^p}/2 = \delta T^{1/p}$.

Oczywiście $r_p^{\text{wor}}(n, \delta, F) \geq r_p^{\text{wor}}(n, 0, F)$, więc aby pokazać nierówność (ii) skorzystamy ze znanych rezultatów dotyczących minimalnego błędu aproksymacji dla informacji dokładnej pochodzących z [7].

Zdefiniujmy funkcję

$$\psi(x) = a(1 - \bar{x}^2)^{r+1},$$

gdzie $|\bar{x}| \leq 1$ oraz $\psi(x) = 0$ dla $x \neq \bar{x}$. Możemy wybrać $a > 0$, takie że $\psi \in \mathcal{H}_{r,\varrho}$. Rozważmy funkcję

$$f_i(x) = \kappa \cdot \psi((x - x_i)\kappa^{-1/(r+\varrho)}),$$

gdzie $i = 1, \dots, n$ oraz x_i są parami oddalone od siebie o co najmniej $2\kappa^{1/(r+\varrho)}$, tzn. $|x_i - x_j| \geq 2\kappa^{1/(r+\varrho)}$ dla $i \neq j$. Możemy wybrać n rzędu $(1/\kappa)^{-(r+\varrho)}$. Ponieważ $f_i \in \mathcal{H}_{r,\varrho}$, to dolną otrzymujemy z lematu 2.3

Aby pokazać (iii), niech $S(M) \subset \mathcal{F}_{r,\varrho}$ będzie rodziną funkcji f_s z parametrem $s \in [0, T]$ zdefiniowaną

$$f_s(x) = \frac{M}{T} (x \mathbb{1}_{[0,s)}(x) + (x - T) \mathbb{1}_{[s,T)}(x)), \quad 0 \leq x \leq T.$$

Niech N będzie dowolną (adaptacyjną) informacją używającą nie więcej niż n ewaluacji funkcji. Ponieważ dla każdego ustalonego x , funkcja $f_s(x)$ może przyjmować tylko dwie wartości w zależności czy $s \leq x$ lub $s > x$, to całkowita liczba punktów użytych przez N dla klasy $S(M)$ wynosi co najwyżej $2^n - 1$. Dlatego istnieje przedział $[s_1, s_2] \subset (0, T)$ o długości $T2^{-(n-1)}$, który nie zawiera żadnego z tych punktów. To oznacza, że $N(f_{s_1}) = N(f_{s_2})$, a więc błąd dowolnego algorytmu używającego informacji N wynosi przynajmniej $\|f_{s_1} - f_{s_2}\|_{L^p}/2 = \delta M(T2^{-(n+p+1)})^{1/p}$. Z uwagi na to, że M jest dowolnie duże, błąd również może być dowolnie duży. \square

Stwierdzenie 2.4 (iii) mówi, że nie możemy uogólnić wyników na klasę kawałkami Hölderowskich funkcji $\mathcal{F}_{r,\varrho}$ z $r \geq 1$. Z tego powodu rozważania będziemy prowadzić głównie na klasie $\mathcal{F}_{r,\varrho}^D$ funkcji kawałkami Hölderowskich z jednostajnie ograniczonymi skokami nieciągłości $\Delta_f^{(0)}$ oraz na klasie $\mathcal{F}_{r,\varrho}^C$ funkcji kawałkami Hölderowskich ciągłych.

Podsumowując, z stwierdzenia 2.4 (i)-(ii) otrzymujemy ograniczenia z dołu

$$r_p^{\text{wor}}(n, \delta, \mathcal{F}_{r,\varrho}^D) \geq r_p^{\text{wor}}(n, \delta, \mathcal{F}_{r,\varrho}^C) \geq \max(\delta T^{1/p}, a_{r,\varrho} n^{-(r+\varrho)}).$$

W dalszej części pracy udowodnimy, że te nierówności są ostre, z wyjątkiem pierwszej dla $p = \infty$. Jest to główny wynik artykułu [5].

2.2. Ograniczenia z góry

Dla uproszczenia, w dalszej części pracy przyjmujemy oznaczenia algorytmów pochodzące od pierwszych liter nazwisk autorów poszczególnych artykułów. Oznaczenie φ^{KP} odnosi się do algorytmu z pracy [4] opartego na wielomianach Lagrange’a. Analogicznie, oznaczenie φ^{MP} odnosi się do algorytmu z pracy [5] opartego na różnicach dzielonych.

Górne ograniczenia na błąd otrzymujemy poprzez analizę skonstruowanych algorytmów, przedstawioną szczegółowo w rozdziale 4.

Omawiane algorytmy osiągają te same ograniczenia z góry z dokładnością do stałej, jednak jak wspomnieliśmy, algorytm przedstawiony w pracy [4] bazuje na informacji dokładnej, w przeciwieństwie do algorytmu z pracy [5], który uwzględnia zaburzenie danych. Ta różnica wpłynęła na to, że do uzyskania tych samych wyników autorzy doszli w odmienny sposób. Aby lepiej przedstawić przebieg rozumowania, nie uogólniamy wyników, które są bardziej szczegółowe niż założenia tej pracy wymagają. Tyczy się to głównie algorytmu φ^{KP} , ponieważ jest on tylko częścią rozwiązania innego problemu, który jest tematem pracy [4].

Poniższe twierdzenie łączy wyniki artykułów [4] i [5] dotyczące górnych ograniczeń na błąd aproksymacji konstruowanych przez algorytmy.

Twierdzenie 2.5. *Niech $r + \varrho \geq 1$ oraz niech $\mathcal{G}_{r,\varrho} = \mathcal{G}_{r,\varrho}([a, b])$ z $\Delta_g^{(0)} = 0$. Wtedy zachodzi*

- (i) $e_p^{\text{wor}}(\varphi^{MP}, N, \mathcal{F}_{r,\varrho}^D) = \mathcal{O}(\max(\delta, n^{-(r+\varrho)}))$, dla $1 \leq p < \infty$,
- (ii) $e_\infty^{\text{wor}}(\varphi^{MP}, N, \mathcal{F}_{r,\varrho}^C) = \mathcal{O}(\max(\delta, n^{-(r+\varrho)}))$,
- (iii) $e_\infty^{\text{wor}}(\varphi^{KP}, N, \mathcal{G}_{r,\varrho}) = \mathcal{O}(n^{-(r+\varrho)})$.

Dodatkowo w rozdziale 4. pokażemy, że koszty algorytmów zachowują się następująco:

Stwierdzenie 2.6.

1. $\text{cost}(\varphi^{MP}, \mathcal{F}_{r,\varrho}^D) = \text{cost}(\varphi^{MP}, \mathcal{F}_{r,\varrho}^C) = \mathcal{O}(n)$,
2. $\text{cost}(\varphi^{KP}, \mathcal{G}_{r,\varrho}) = \mathcal{O}(n)$.

Z powyższych wyników oraz przytoczonych wcześniej rezultatów o ograniczeniach z dołu wynikają wnioski dotyczące minimalnych błędów najgorszego przypadku.

Wniosek 2.7.

- (i) $r_p^{\text{wor}}(n, \delta, \mathcal{F}_{r,\varrho}^D) = \Theta(\max(\delta, n^{-(r+\varrho)})), \quad \text{dla } 1 \leq p \leq \infty,$
- (ii) $r_\infty^{\text{wor}}(n, \delta, \mathcal{F}_{r,\varrho}^C) = \Theta(\max(\delta, n^{-(r+\varrho)})),$
- (iii) $r_\infty^{\text{wor}}(n, 0, \mathcal{G}_{r,\varrho}) = \Theta(n^{-(r+\varrho)}).$

Rozdział 3

Algorytmy

Oba algortmy na wejściu otrzymują siatkę o $m + 1$ równoodległych punktach $t_j = a + (b - a)/m$, przyjmując $a = 0$, $b = T$ dla algorytmu φ^{KP} . Długość przedziału $[t_j, t_{j+1}]$ oznaczamy przez $h = \frac{T}{m}$.

3.1. Algorytm oparty na wielomianach Lagrange’a

Wprowadźmy postać wielomianów Lagrange’a używanych w algorytmie. Niech $g : [\alpha, \beta] \rightarrow \mathbb{R}$. Przez $w_g^s([\alpha, \beta])$ oznaczamy interpolacyjny wielomian Lagrange’a o rzędzie co najwyżej s , oparty na $s + 1$ równoodległych węzłach $x_j = \alpha + (\beta - \alpha)j/s$, dla $j = 0, 1, \dots, s$

$$w_g^s([\alpha, \beta])(x) = \sum_{i=0}^s g(x_i) \Phi_i(x), \quad x \in \mathbb{R}, \quad (3.1)$$

gdzie

$$\Phi_i(x) = \prod_{\substack{k=0 \\ k \neq i}}^s \frac{x - x_k}{x_i - x_k}, \quad i = 0, 1, \dots, s.$$

Algorytm przedstawiony w pracy [4] lokalizuje osobliwość przy pomocy wielomianów Lagrange’a w_g^r . Na wejściu algorytm otrzymuje $g \in \mathcal{G}_{r,\varrho}([a, b])$, przedział $[a, b]$, regularność r oraz współczynnik Höldera ϱ . Kluczowym elementem algorytmu jest zdefiniowana poniżej wielkość (*test*), która jest użyta do wykrycia punktu osobliwego.

$$A_g(\alpha, \bar{\alpha}, \bar{\beta}, \beta) = \max_{0 \leq j \leq r} \frac{\|w_g^r([\bar{\beta}, \beta])(z_j) - w_g^r([\alpha, \bar{\alpha}]) (z_j)\|}{\bar{h}^{r+\varrho}}, \quad (3.2)$$

gdzie $\alpha < \bar{\alpha} < \bar{\beta} < \beta$, $z_j = \bar{\alpha} + (\bar{\beta} - \bar{\alpha})j/r$, dla $j = 0, 1, \dots, r$ oraz $\bar{h} = \beta - \alpha$ jest długością przedziału, na którym *test* jest zdefiniowany.

Algorytm oparty na wielomianach Lagrange'a

- K1:** Niech $\omega := h^{r+\varrho}$ oraz $A_g^i = A_g(t_i, t_i + \omega, t_{i+1} - \omega, t_{i+1})$
 Oblicz $A := \max \{A_g^i \mid t_{j+1} - t_j > 4\omega, j = 0, 1, \dots, m-1\}$
jeżeli istnieją różne k i l takie, że $A = A_g^k \wedge A = A_g^l$ **wtedy**
 idź do **K3**
- K2:** Niech $[t_k, t_{k+1}]$ będzie przedziałem odpowiadającym A z **K1**
 Niech $[\alpha, \beta] := [t_k, t_{k+1}]$ oraz $B := \{\alpha, \beta\}$
dopóki $\beta - \alpha > 4\omega$ **wykonuj:**
 Oblicz $v := (\alpha + \beta)/2$ oraz $B := B \cup \{v\}$
jeżeli $A_g(\alpha, \alpha + \omega, v - \omega, v) = A_g(v, v + \omega, \beta - \omega, \beta)$ **wtedy**
 idź do **K3**
- w p.p.**
 za następny przedział $[\alpha, \beta]$ wybierz podprzedział $[\alpha, v]$ lub $[v, \beta]$,
 dla którego wartość testu była większa
- K3:** Niech $\bar{M} := \{t_0, \dots, t_m\} \cup B$ będzie podziałem. Do oznaczenia kolejnych punktów \bar{M} użyjemy tych samych oznaczeń co dla podziału początkowego, tzn. $\alpha = t_0 < t_1 < \dots < t_k = \beta$, gdzie $k = m + \mathcal{O}(\log m)$.
 Finalna aproksymacja dana jest wzorem

$$\varphi^{KP}(x) = \begin{cases} g(t_i), & \text{gdy } x \in [t_i, t_{i+1}) \wedge t_{i+1} - t_i \leq 4\omega, \\ g(t_i), & \text{gdy } x \in [t_i, t_i + \omega) \wedge t_{i+1} - t_i > 4\omega, \\ w_g^r([t_i + \omega, t_{i+1} - \omega])(x), & \text{gdy } x \in [t_i + \omega, t_{i+1} - \omega) \wedge t_{i+1} - t_i > 4\omega, \\ g(t_{i+1} - \omega), & \text{gdy } x \in [t_{i+1} - \omega, t_{i+1}) \wedge t_{i+1} - t_i > 4\omega, \end{cases}$$

dla $i = 0, 1, \dots, k-1$ z $\varphi^{KP}(b)$ zdefiniowanym przez ciągłość na ostatnim przedziale.

Zauważmy, że liczba ewaluacji wartości funkcji g jest proporcjonalna do liczby przedziałów końcowej siatki, czyli dla podziału początkowego z równoodległymi węzłami obliczenie optymalnej aproksymacji φ^{KP} wymaga $n = \mathcal{O}(m)$ ewaluacji funkcji g .

3.2. Algorytm oparty na różnicach dzielonych

W tym rozdziale opiszemy algorytm bazujący na informacji zaburzonej przedstawiony w artykule [5]. Analizowany algorytm używa co najwyżej n wartości funkcji z precyzją δ oraz w najgorszym przypadku ma błąd proporcjonalny do $\max(\delta, n^{-(r+\varrho)})$ w klasie funkcji $\mathcal{F}_{r,\varrho}^D$ dla $p < \infty$ oraz w klasie $\mathcal{F}_{r,\varrho}^C$ dla $p \leq \infty$. Do wykrycia przedziału z punktem osobliwym, algorytm wykorzystuje różnice dzielone zdefiniowane następująco.

Niech $m \geq 2r + 1$, $h = \frac{T}{m}$ oraz $t_i = ih$ dla każdego i . Przez d_i oznaczmy różnicę dzieloną stopnia $r + 1$ bazującą na wartościach $f(t_i)$:

$$d_i = f[t_i, \dots, t_{i+r+1}] = \sum_{j=1}^{i+r+1} f(t_j) \prod_{\substack{k=1 \\ k \neq j}}^{i+r+1} (t_j - t_k)^{-1}. \quad (3.3)$$

Następnie oznaczmy przez \tilde{d}_i (niedokładną) różnicę dzieloną stopnia $r+1$ bazującą na wartościach $y_j = \tilde{f}(t_j) := f(t_j) + e_j$, gdzie $|e_j| \leq \delta$,

$$\tilde{d}_i = \tilde{f}[t_i, \dots, t_{i+r+1}] = \sum_{j=1}^{i+r+1} y_j \prod_{\substack{k=1 \\ k \neq j}}^{i+r+1} (t_j - t_k)^{-1}. \quad (3.4)$$

Algorytm używa podziału o ilości punktów i długości przedziałów wprowadzonych jak w definicji różnic dzielonych.

Dodatkowo, ze względu na inny sposób wykorzystania, wielomiany interpolacyjne związane z algorytmem φ^{MP} będziemy oznaczać przez p , jeśli bazują na informacji dokładnej, oraz przez \tilde{p} , jeśli bazują na informacji niedokładnej. Ostatnią zmienną potrzebną na wejściu algorytmu jest $\omega = \omega(h)$ spełniająca $0 < \omega < (r+1)h$, pełniącą podobną rolę jak w φ^{KP} .

Na początku, algorytm aproksymuje punkt osobliwy \hat{t}_f . Jest to realizowane w trzech krokach. W pierwszym kroku, przy pomocy siatki o rozmiarze długości h i różnic dzielonych lokalizowany jest punkt \hat{t}_f na przedziale $[u_1, v_1]$ o długości $(r+1)h$. W kroku drugim używamy wielomianów interpolacyjnych \tilde{p}_+ i \tilde{p}_- do zwiężenia tego przedziału do $[u_2, v_2]$. Krok trzeci produkuje przedział $[u_3, v_3] \subseteq [u_2, v_2]$, w którym różnica $|\tilde{p}_+ - \tilde{p}_-|$ jest nierosnąca na $[u_3, \xi]$ i niemalejąca na $[\xi, v_3]$, gdzie ξ jest finalną aproksymacją \hat{t}_f .

Oznaczenia $\arg \max_j \psi_j$ oraz $\arg \min_j \psi_j$ użyte w algorytmie oznaczają argument j maksymalizujący oraz minimalizujący ψ_j względem j .

Algorytm oparty na różnicach dzielonych

K1 Oblicz różnice dzielone $\tilde{d}_i = \tilde{f}[t_i, \dots, t_{i+r+1}]$ dla $1 \leq i \leq m$ oraz znajdź

$$i^* = \arg \max_{1 \leq i \leq m} |\tilde{d}_i|.$$

Niech $u_1 = t_{i^*}$ i $v_1 = t_{i^*+r+1}$.

K2 Oznaczmy przez \tilde{p}_+ i \tilde{p}_- wielomiany stopnia co najwyżej r , które interpolują punkty $(t_j, \tilde{f}(t_j))$ odpowiednio dla $i^* - r \leq j \leq i^*$ oraz dla $i^* + r + 1 \leq j \leq i^* + 2r + 1$. Następnie wykonaj iterację:

$u := u_1, v := v_1$

dopóki $v - u > \omega$ **wykonuj**:

$z_j := u + j(v - u)/(r + 2)$, dla $j = 1, 2, \dots, r + 1$

$j^* := \arg \max_{1 \leq j \leq r+1} |\tilde{p}_+(z_j) - \tilde{p}_-(z_j)|$

jeżeli $|\tilde{f}(z_{j^*}) - \tilde{p}_-(z_{j^*})| \leq |\tilde{f}(z_{j^*}) - \tilde{p}_+(z_{j^*})|$ **wtedy**

$u := z_{j^*}$

w p.p.

$v := z_{j^*}$

koniec

Niech $u_2 := u$ i $v_2 := v$.

K3 Wykonaj iterację:
 $u := u_2, v := v_2$
dopóki istnieje maksimum lokalne $|\tilde{p}_+ - \tilde{p}_-|$ na (u, v) **wykonuj**
 $z :=$ największe maksimum lokalne $|\tilde{p}_+ - \tilde{p}_-|$ na (u, v)
jeżeli $|\tilde{f}(z) - \tilde{p}_-(z)| \leq |\tilde{f}(z) - \tilde{p}_+(z)|$ **wtedy**
 $u := z$
w p.p.
 $v := z$
koniec
Niech $u_3 := u$ i $v_3 := v$.

Finalną aproksymacją \hat{t}_f jest

$$\xi := \arg \min_{u_3 \leq x \leq v_3} |\tilde{p}_+ - \tilde{p}_-|$$

Niech $N_h^*(y_h)$ będzie operatorem informacji odpowiadający powyższemu algorytmowi. Aproksymacja $\varphi_h^*(y_h)$ funkcji f z informacją y_h , tj. dla $y_h \in N_h^*(f)$ jest konstruowana w następujący sposób. Na przedziale $[u_1, v_1)$ ekstrapolujemy

$$\varphi_h^*(y_h) = \begin{cases} \tilde{p}_-(x), & \text{gdy } u_1 \leq x < \xi, \\ \tilde{p}_+(x), & \text{gdy } \xi \leq x < v_1. \end{cases}$$

Poza przedziałem $[u_1, v_1)$ stosujemy interpolacje funkcjami sklejanymi o stopniu r , bazujących na $r+1$ kolejnych punktach t_i, \dots, t_{i+r} , takich że $x \in [t_i, t_{i+r})$ i $t_j \notin (u_1, v_1)$ dla $1 \leq j \leq i+r$. W przypadku gdy $r=0$ bierzmy x , takie że $|x - t_i| \leq h/2$.

Przedstawiony algorytm używa m wartości funkcji w kroku 1 oraz jedną wartość funkcji w każdej iteracji w krokach 2 i 3. Czyli w kroku 2 używamy co najwyżej

$$\left\lceil \frac{\ln \left(\frac{(r+1)h}{\omega(h)} \right)}{\ln \left(\frac{r+2}{r+1} \right)} \right\rceil$$

wartości funkcji i $(r-1)$ w kroku 3. Stąd otrzymujemy, że jeżeli $\omega = \omega(h) \geq kh^\alpha$ dla pewnego ustalonego k i α , wtedy w najgorszym przypadku liczba użytych wartości funkcji równa się asymptotycznie $m = \frac{T}{h}$ dla $h \rightarrow 0^+$.

Analiza algorytmów

4.1. Analiza algorytmu opartego o wielomiany Lagrange’a

Twierdzenia i lematy przedstawione w tym rozdziale pochodzą z artykułu [4] i pokazują, że tempo zbieżności błędu aproksymacji algorytmu φ^{KP} wynosi $\mathcal{O}(n^{-(r+\varrho)})$, a koszt algorytmu zależy od $\mathcal{O}(m)$ wartości funkcji.

Przedstawmy krótko schemat analizy. Aproksymacja konstruowana w ostatnim kroku algorytmu składa się z funkcji stałych na odpowiednio małych przedziałach oraz wielomianów Lagrange’a na pozostałych. Aby udowodnić, że aproksymacja zachowuje wspomniane właściwości musimy pokazać, że zarówno obecność punktu osobliwego na ”małych” przedziałach, jak i na przedziałach z interpolacją nie pogorsza znacząco błędu.

Przypadek dotyczący interpolacji udowodniono, ograniczając błąd interpolacji wartością testu A_g (stwierdzenie 4.5), który z kolei ograniczono przez normę wielomianu s_g , zdefiniowanego w rozdziale 1.3. To oszacowanie wymaga rozważenia błędu interpolacji w zależności czy punkt osobliwy \hat{t}_g należy do przedziału z węzłami interpolacji (lemat 4.1), czy nie (lematy 4.2).

Przypadek, gdy przybliżamy funkcją stałą, ma uzasadnienie w postaci wielomianu s_g w połączeniu z długością przedziału (wniosek 4.6).

Zacznijmy od rozważenia błędu interpolacji Lagrange’a dla funkcji nieciągłej. Błąd ten jest ograniczony za względu na wielomian s_g (1.1).

Lemat 4.1. *Istnieje stała C taka, że dla wszystkich $[\alpha, \beta] \subset [a, b]$, wszystkich $g \in \mathcal{G}_{r,\varrho}([\alpha, \beta])$ oraz $s = 0, 1, \dots, r$, mamy*

$$\sup_{t \in [\alpha, \beta]} \|g(t) - w_g^s([\alpha, \beta])(t)\| \leq C \left(\min \left\{ \sup_{t \in [\alpha, \hat{t}_g]} \|s_g(t)\|, \sup_{t \in [\hat{t}_g, \beta]} \|s_g(t)\| \right\} + \bar{h}^{\min\{s+1, r+\varrho\}} \right).$$

Dowód. Najpierw pokażemy, że

$$\sup_{t \in [\alpha, \beta]} \|g(t) - w_g^s([\alpha, \beta])(t)\| \leq C \left(\sup_{t \in [\hat{t}_g, \beta]} \|s_g(t)\| + \bar{h}^{\min(s+1, r+\varrho)} \right). \quad (4.1)$$

W tym celu zdefiniujemy funkcję

$$\tilde{g}(t) = \begin{cases} g(t), & \text{gdy } t \in [\alpha, \hat{t}_g), \\ g(t) - s_g(t), & \text{gdy } t \in [\hat{t}_g, \beta]. \end{cases} \quad (4.2)$$

Z (1.2) wiemy, że $\tilde{g} \in G_{r, \varrho}^{\text{reg}}([\alpha, \beta])$. Niech $t_k = \alpha + (\beta - \alpha)k/s$ dla $k = 0, 1, \dots, s$ będą węzłami dla interpolacji $w_g^s([\alpha, \beta])$ na przedziale $[\alpha, \beta]$. Dla indeksu

$$j^* = \min \{k = 1, 2, \dots, s \mid t_k \geq \hat{t}_g\}, \quad (4.3)$$

otrzymujemy

$$w_g^s([\alpha, \beta])(t) = \sum_{i=0}^{j^*-1} g(t_i) \Phi_i(t) + \sum_{i=j^*}^s (g(t_i) - s_g(t_i)) \Phi_i(t), \quad (4.4)$$

gdzie

$$\Phi_i(t) = \prod_{\substack{k=0 \\ k \neq i}}^s \frac{t - t_k}{t_i - t_k},$$

czyli

$$w_g^s([\alpha, \beta])(t) = w_g^s([\alpha, \beta])(t) - \sum_{i=j^*}^s s_g(t_i) \Phi_i(t).$$

Z (4.2) i (4.4) otrzymujemy, że dla $t \in [\alpha, \beta]$ mamy

$$\begin{aligned} g(t) - w_g^s([\alpha, \beta])(t) &= (g(t) - \tilde{g}(t)) + (\tilde{g}(t) - w_g^s([\alpha, \beta])(t)) \\ &\quad + (w_g^s([\alpha, \beta])(t) - w_g^s([\alpha, \beta])(t)) \\ &= \mathbb{1}_{[\hat{t}_g, \beta]}(t) s_g(t) + (\tilde{g}(t) - w_g^s([\alpha, \beta])(t)) - \sum_{i=j^*}^s s_g(t_i) \Phi_i(t), \end{aligned}$$

a ponieważ \tilde{g} jest funkcją regularną dla $t \in [\alpha, \beta]$, to zachodzi

$$\|g(t) - w_g^s([\alpha, \beta])(t)\| \leq \sup_{t \in [\hat{t}_g, \beta]} \|s_g(t)\| + \max_{j^* \leq i \leq s} \|s_g(t_i)\| \sum_{i=j^*}^s |\Phi_i(t)| + C \bar{h}^{\min\{s+1, r+\varrho\}},$$

gdzie C zależy tylko od parametrów klasy $\mathcal{G}_{r, \varrho}([\alpha, \beta])$. Ponadto, istnieje stała \bar{C} zależna jedynie od s taka, że dla wszystkich $t \in [\alpha, \beta]$ zachodzi

$$\sum_{i=j^*}^s |\Phi_i(t)| \leq \bar{C}, \quad (4.5)$$

co dowodzi nierówność (4.1)

Teraz musimy pokazać, że

$$\sup_{t \in [\alpha, \beta]} \|g(t) - w_g^s([\alpha, \beta])(t)\| \leq C \left(\sup_{t \in [\alpha, \hat{t}_g]} \|s_g(t)\| + \bar{h}^{\min\{s+1, r+\varrho\}} \right). \quad (4.6)$$

Postępujemy jak wyżej używając regularnej funkcji

$$\tilde{g}(t) = \begin{cases} g(t) + s_g(t), & \text{gdy } t \in [\alpha, \hat{t}_g], \\ g(t), & \text{gdy } t \in [\hat{t}_g, \beta] \end{cases} \quad (4.7)$$

oraz j^* zdefiniowanym jak w (4.3). Nierówności (4.1) i (4.6) udowadniają tezę lematu. \square

Z postaci testu A_g widzimy, że możemy przybliżać wartości na przedziale z osobliwością, przy pomocy wielomianu interpolacyjnego skonstruowanego z użyciem punktów z przedziału niezawierającego osobliwości. Poniższy lemat opisuje zachowanie błędu takiej ekstrapolacji.

Lemat 4.2. *Istnieją stałe C i \bar{C} zależne od r i L_r , takie że dla wszystkich $[\alpha, \beta] \subset [a, b]$, $\bar{\alpha} \in (\alpha, \beta)$ oraz $g \in \mathcal{G}_{r, \varrho}([\alpha, \beta])$, mamy*

$$\hat{t}_g \in (\bar{\alpha}, \beta) \implies g(t) - w_g^r([\alpha, \bar{\alpha}]) (t) = s_g(t) \mathbb{1}_{[\hat{t}_g, \beta]}(t) + R_g(t), \quad t \in [\bar{\alpha}, \beta]$$

oraz

$$\hat{t}_g \in (\alpha, \bar{\alpha}) \implies g(t) - w_g^r([\bar{\alpha}, \beta]) (t) = -s_g(t) \mathbb{1}_{[\alpha, \hat{t}_g]}(t) + \bar{R}_g(t), \quad t \in [\alpha, \bar{\alpha}],$$

gdzie $\|R_g(t)\| \leq C \bar{h}^{r+\varrho}$ dla $t \in [\bar{\alpha}, \beta]$ i $\|\bar{R}_g(t)\| \leq \bar{C} \bar{h}^{r+\varrho}$ dla $t \in [\alpha, \bar{\alpha}]$.

Dowód. Niech $\tilde{g}_1, \tilde{g}_2 \in G_{r, \varrho}^{\text{reg}}([\alpha, \beta])$ będą dane odpowiednio jak w (4.2) i (4.7).

Założmy najpierw, że $\hat{t}_g \in (\alpha, \bar{\alpha})$. Ponieważ $\hat{t}_g > \bar{\alpha}$, to $\tilde{g}_1(t) = g(t)$ dla wszystkich $t \in [\alpha, \bar{\alpha}]$ oraz

$$w_g^r([\alpha, \bar{\alpha}]) (t) = w_{\tilde{g}_1}^r([\alpha, \bar{\alpha}]) (t), \quad t \in [\alpha, \beta],$$

Dlatego, dla $t \in [\bar{\alpha}, \beta]$ mamy

$$g(t) - w_g^r([\alpha, \bar{\alpha}]) (t) = g(t) - \tilde{g}_1(t) + \tilde{g}_1(t) - w_{\tilde{g}_1}^r([\alpha, \bar{\alpha}]) (t) = s_g(t) \mathbb{1}_{[\hat{t}_g, \beta]}(t) + R_g(t), \quad (4.8)$$

gdzie $R_g(t) = \tilde{g}_1(t) - w_{\tilde{g}_1}^r([\alpha, \bar{\alpha}]) (t)$. Ponieważ \tilde{g}_1 i $w_{\tilde{g}_1}^r$ są regularne, to R_g możemy ograniczyć stałymi zależnymi od r i L_r , co pokazuje pierwszą część tezy.

Założmy teraz, że $\hat{t}_g \in (\alpha, \bar{\alpha})$. Ponieważ $\hat{t}_g \leq \bar{\alpha}$, to $\tilde{g}_2(t) = g(t)$ dla $t \in [\bar{\alpha}, \beta]$ oraz

$$w_g^r([\bar{\alpha}, \beta]) (t) = w_{\tilde{g}_2}^r([\bar{\alpha}, \beta]) (t), \quad t \in [\alpha, \beta].$$

Analogicznie, dla $t \in [\alpha, \bar{\alpha}]$ mamy

$$g(t) - w_g^r([\alpha, \bar{\alpha}]) (t) = s_g(t) \mathbb{1}_{[\hat{t}_g, \beta]}(t) + \bar{R}_g(t), \quad \text{gdzie } \bar{R}_g(t) = \tilde{g}_1(t) - w_{\tilde{g}_1}^r([\alpha, \bar{\alpha}]) (t) \quad (4.9)$$

a z regularności \tilde{g}_1 i $w_{\tilde{g}_2}^r$ wynika oszacowanie na \bar{R}_g , co kończy dowód. \square

Lemat 4.2 uzasadnia definicję testu $A_g(\alpha, \bar{\alpha}, \bar{\beta}, \beta)$ w następujący sposób. Niech $g \in \mathcal{G}_{r,\varrho}([\alpha, \beta])$, $\alpha < \bar{\alpha} < \bar{\beta} < \beta$, $\bar{h} = \beta - \alpha$. Jeżeli $\hat{t}_g \in (\bar{\alpha}, \bar{\beta}]$, to z lematu 4.2 otrzymujemy

$$\begin{aligned} g(t) - w_g^r([\alpha, \bar{\alpha}]) &= s_g(t) \mathbb{1}_{[\hat{t}_g, \beta]}(t) + R_g(t), \quad t \in [\bar{\alpha}, \beta], \\ g(t) - w_g^r([\bar{\beta}, \beta]) &= -s_g(t) \mathbb{1}_{[\alpha, \hat{t}_g]}(t) + \bar{R}_g(t), \quad t \in [\alpha, \bar{\beta}]. \end{aligned}$$

Odejmując równania stronami dostajemy

$$w_g^r([\bar{\beta}, \beta])(t) - w_g^r([\alpha, \bar{\alpha}]) &= s_g(t) + R_g(t) - \bar{R}_g(t), \quad \text{dla } t \in [\bar{\alpha}, \bar{\beta}], \quad (4.10)$$

gdzie $\|R_g(t)\| \leq C\bar{h}^{r+\varrho}$ oraz $\|\bar{R}_g(t)\| \leq \bar{C}\bar{h}^{r+\varrho}$ dla $t \in [\bar{\alpha}, \bar{\beta}]$.

Z tego wynika, że gdy $\hat{t}_g \in (\alpha, \beta)$, wtedy możemy zrekonstruować nieznaną wielomian $s_g(t)$ na przedziale $[\alpha, \beta]$ w graniach błędu $\mathcal{O}(\bar{h}^{r+\varrho})$ poprzez wyznaczenie wielomianów Lagrange'a $w_g^r([\bar{\beta}, \beta])$ i $w_g^r([\alpha, \bar{\alpha}])$.

Przytoczmy teraz dwa stwierdzenia pokazujące relacje między położeniem punktu osobliwego \hat{t}_g a wartością testu $A_g(\alpha, \bar{\alpha}, \bar{\beta}, \beta)$. Pierwsze dotyczy przypadku regularnego, tzn. gdy osobliwość znajduje się poza przedziałem $[\alpha, \beta]$.

Stwierdzenie 4.3. *Istnieje stała C^* zależna od r i L_r taka, że dla wszystkich $\alpha < \bar{\alpha} < \bar{\beta} < \beta$ i $[\alpha, \beta] \subset [a, b]$ oraz $g \in \mathcal{G}_{r,\varrho}([\alpha, \beta])$, mamy*

$$\hat{t}_g \text{ z niezerowym wielomianem } s_g \text{ nie jest w } (\alpha, \beta) \implies A_g(\alpha, \bar{\alpha}, \bar{\beta}, \beta) \leq C^*.$$

Dowód. Skoro $\hat{t}_g \notin (\alpha, \beta)$, to funkcja g jest regularna na $[\alpha, \beta]$. Stąd dla $t \in [\alpha, \beta]$ mamy

$$\begin{aligned} \|w_g^r([\bar{\beta}, \beta])(t) - w_g^r([\alpha, \bar{\alpha}])\| &\leq \|w_g^r([\bar{\beta}, \beta])(t) - g(t)\| + \|g(t) - w_g^r([\alpha, \bar{\alpha}])\| \\ &\leq C^* \bar{h}^{r+\varrho}, \end{aligned}$$

gdzie C^* jest stałą. □

Uwaga 4.4. Z stwierdzenia 4.3 oraz z definicji algorytmu φ^{KP} wynika, że dla jakiegokolwiek przedziału $[\alpha, \beta]$, który nie został wybrany podczas algorytmów, mamy $A_g(\alpha, \bar{\alpha}, \bar{\beta}, \beta) \leq C^*$, jeżeli osobliwość \hat{t}_g jest jedyna. Zachodzi to, ponieważ w pierwszym i drugim kroku algorytmu przedziały wybierane są na podstawie wartości testu.

Następna właściwość pokazuje, że w przypadku osobliwym, górne ograniczenie na błąd interpolacji może być wyrażone za pomocą $A_g(\alpha, \bar{\alpha}, \bar{\beta}, \beta)$.

Stwierdzenie 4.5. *Niech $D > 0$. Istnieją stałe C i \bar{N} , zależne tylko od parametrów klasy $\mathcal{G}_{r,\varrho}([\alpha, \beta])$ i D , takie, że dla wszystkich $[\alpha, \beta] \subset [a, b]$, $[\bar{\alpha}, \bar{\beta}] \subset (\alpha, \beta)$, $g \in \mathcal{G}_{r,\varrho}([\alpha, \beta])$ oraz $s = 0, 1, \dots, r$, mamy*

$$\begin{aligned} \hat{t}_g \in (\bar{\alpha}, \bar{\beta}] \wedge \beta - \alpha \leq D(\bar{\beta} - \bar{\alpha}) &\implies \sup_{t \in [\gamma, \zeta]} \|g(t) - w_g^s([\gamma, \zeta])(t)\| \\ &\leq C (1 + A_g(\alpha, \bar{\alpha}, \bar{\beta}, \beta)) \bar{h}^{\min\{s+1, r+\varrho\}}, \end{aligned} \quad (4.11)$$

gdzie $[\gamma, \zeta] = [\alpha, \beta]$ lub $[\gamma, \zeta] = [\bar{\alpha}, \bar{\beta}]$.

Dowód. Załóżmy, że $[\gamma, \zeta] = [\bar{\alpha}, \bar{\beta}]$. Dowód jest analogiczny dla $[\gamma, \zeta] = [\alpha, \beta]$. Z (4.10) mamy

$$\|s_g(z_j)\| \leq (A_g(\alpha, \bar{\alpha}, \bar{\beta}, \beta) + C + \bar{C}) \bar{h}^{r+e}, \quad (4.12)$$

gdzie z_j są z definicji $A_g(\alpha, \bar{\alpha}, \bar{\beta}, \beta)$, natomiast C i \bar{C} zależą wyłącznie od r i L_r . Wielomian s_g możemy wyrazić jako

$$s_g(t) = \sum_{j=0}^r s_g(z_j) \bar{\Phi}_j(t), \quad \text{gdzie } \bar{\Phi}_j(t) = \prod_{\substack{k=0 \\ k \neq j}}^r \frac{t - z_k}{z_j - z_k}, \quad t \in \mathbb{R}. \quad (4.13)$$

Ponieważ $\beta - \alpha \leq D(\bar{\beta} - \bar{\alpha})$, to istnieje stała \bar{K} zależna tylko od r i D taka, że

$$\sum_{j=0}^r |\bar{\Phi}_j(t)| \leq \bar{K}, \quad t \in [\alpha, \beta], \quad (4.14)$$

stąd

$$\|s_g(t)\| = \mathcal{O}((1 + A_g(\alpha, \bar{\alpha}, \bar{\beta}, \beta)) \bar{h}^{r+e}), \quad t \in [\alpha, \beta]. \quad (4.15)$$

Widzimy, że (4.11) wynika z lematu 4.1 dla $[\alpha, \beta] = [\bar{\alpha}, \bar{\beta}]$. \square

Przejdźmy po przypadku, gdy punkt osobliwy znajduje się w pobliżu brzegu przedziału. Poniższy wniosek jest następstwem lematu 4.1 i mówi o tym, że punkt osobliwy znajdujący się blisko brzegu przedziału, nie powoduje to znaczącego wzrostu błędu.

Wniosek 4.6. *Istnieje stała C , taka że dla wszystkich $[\alpha, \beta] \subset [a, b]$, wszystkich $g \in \mathcal{G}_{r,e}([\alpha, \beta])$ z $\Delta_g^{(0)} = 0$, $0 \leq \omega \leq \min\{1, \bar{h}\}$ oraz $s = 0, 1, \dots, r$, mamy*

$$\hat{t}_g \in (\alpha, \alpha + \omega] \cup [\beta - \omega, \beta) \implies \sup_{t \in [\alpha, \beta]} \|g(t) - w_g^s([\alpha, \beta])(t)\| \leq C (\omega + \bar{h}^{\min\{s+1, r+e\}}).$$

Dowód. Wiemy, że $s_g(t) = \sum_{j=1}^r \frac{1}{j!} \Delta_g^j (t - \hat{t}_g)^j$. Jeżeli $\hat{t}_g \in (\alpha, \alpha + \omega]$, wtedy $\|s_g(t)\| = \mathcal{O}(\omega)$ dla $t \in [\alpha, \hat{t}_g]$. To samo zachodzi dla $t \in [\hat{t}_g, \beta]$, jeśli $\hat{t}_g \in [\beta - \omega, \beta)$. Z lematu 4.1 otrzymujemy szukaną nierówność \square

Za pomocą wprowadzonych lematów możemy teraz udowodnić ograniczenie górne dla φ^{KP} .

Dowód twierdzenia 2.5(iii). Niech $g \in \mathcal{G}_{r,e}([a, b])$ i $\Delta_g^0 = 0$. W kroku 3. algorytmu, aproksymacja $\varphi^{KP}(t)$ jest zdefiniowana jako funkcja kawałkami regularna na podziale $t_0 < t_1 < \dots < t_k$, zawierającym przedziały brzegowe.

W każdym podprzedziale, gdzie $\varphi^{KP}(t)$ jest zdefiniowana jako funkcja stała mamy, że $\|g(t) - \varphi^{KP}(t)\| = \mathcal{O}(h^{r+e})$. Wynika to z długości przedziału oraz Lipschitzowskiej ciągłości g na $[t_i, t_{i+1}]$. Wniosek 4.6 gwarantuje, że to ograniczenie zachodzi również, gdy punkt osobliwy \hat{t}_g znajduje się w takim przedziale.

Wobec tego weźmy przedziały $[t_i + h^{r+e}, t_{i+1} - h^{r+e}]$ dla $t_{i+1} - t_i > 4h^{r+e}$. Na takim przedziale aproksymacja ma postać:

$$\varphi^{KP}(t) = w_g^r([t_i + h^{r+e}, t_{i+1} - h^{r+e}]) (t).$$

Jeżeli $\hat{t}_g \in (t_i + h^{r+e}, t_{i+1} - h^{r+e}]$, wtedy z stwierdzenia 4.5, dla $t \in [t_i + h^{r+e}, t_{i+1} - h^{r+e})$, mamy

$$\begin{aligned} & \|g(t) - w_g^r([t_i + h^{r+e}, t_{i+1} - h^{r+e}]) (t)\| \\ &= \mathcal{O}((1 + A_g(t_i, t_i + h^{r+e}, t_{i+1} - h^{r+e}, t_{i+1})) (t_{i+1} - t_i)^{r+e}). \end{aligned}$$

Z konstrukcji algorytmu oraz uwagi 4.4 wiemy, że przedział początkowego podziału, który zawierał punkt osobiwy z niezerowym skokiem w pochodnych (nie w pobliżu brzegu), został podzielony na podprzedziały w kroku 2. algorytmu. Z tego wynika, że wszystkie przedziały na których stosujemy interpolację, nie zawierają istotnego punktu osobiwego, czyli z stwierdzenia 4.3 zachodzi

$$A_g(t_i, t_i + h^{r+e}, t_{i+1} - h^{r+e}, t_{i+1}) \leq C^*, \text{ dla każdego } t_{i+1} - t_i > 4h^{r+e},$$

gdzie C^* jest dane jak w stwierdzeniu 4.3.

Podsumowując otrzymujemy

$$\|g(t) - w_g^r([t_i + h^{r+e}, t_{i+1} - h^{r+e}]) (t)\| = \mathcal{O}(h^{r+e}), \quad t \in [t_i + h^{r+e}, t_{i+1} - h^{r+e}).$$

co dowodzi ograniczenie na błąd przedstawione w twierdzeniu. □

Uwaga 4.7. Twierdzenie 2.5(iii) zachodzi również dla funkcji g , która ma skok w punktach t_i początkowego podziału M oraz ma niezerowy wielomian s_g dla co najwyżej jednego nieznanego punktu t_g , $t_g \neq t_i \forall i$.

4.2. Analiza algorytmu opartego o różnice dzielone

Zanim przejdziemy do analizy algorytmu, przedstawmy kilka własności różnic dzielonych.

Niech t_0, \dots, t_r będą parami rozłącznymi punktami należącymi do przedziału $[0, T]$. Dla $r \geq 1$, różnice dzieloną możemy przedstawić następująco

$$f[t_0, \dots, t_r] = \frac{f[t_1, \dots, t_r] - f[t_0, \dots, t_{r-1}]}{t_r - t_0}.$$

Kolejna własność jest uogólnieniem twierdzenia Lagrange'a i mówi, że jeśli f jest r -krotnie różniczowalną funkcją określoną na przedziale $[0, T]$, wtedy istnieje punkt $\xi \in \text{conv}(t_0, \dots, t_r) := (\min\{t_0, \dots, t_r\}, \max\{t_0, \dots, t_r\})$, taki że

$$f[t_0, \dots, t_r] = \frac{f^{(r)}(\xi)}{r!}.$$

Poniższy lemat przedstawia ograniczenie na wartość różnicy dzielonej. Własność ta została przedstawiona w artykule [3].

Lemat 4.8. Niech $f : [a, b] \rightarrow \mathbb{R}$ oraz $a < c_0 < c_1 < \dots < c_k$ dla $k \geq r$. Wtedy dla podciągu $0 < j_0 < j_1 < \dots < j_r \leq k$ mamy

$$|f[c_{j_0}, c_{j_1}, \dots, c_{j_r}]| \leq \max_{0 \leq i \leq k-r} |f[c_i, c_{i+1}, \dots, c_{i+r}]|.$$

Dowód. Przypuśćmy, że teza nie jest prawdziwa.
Bez straty ogólności możemy założyć, że

$$f[c_{j_0}, c_{j_1}, \dots, c_{j_r}] > 0.$$

Niech p_f będzie wielomianem stopnia r interpolującym f w t_j, \dots, t_{j+r} . Wtedy

$$(f - p_f)[c_i, c_{i+1}, \dots, c_{i+r}] < 0, \quad \text{dla } 0 \leq i \leq k - r.$$

Z tego wynika, że różnica dzielona rzędu $r - 1$ funkcji $(f - p_f)[c_i, c_{i+1}, \dots, c_{i+r}]$, dla $i \leq i \leq k + 1 - r$, zmienia znak co najwyżej raz (przyjęcie wartości równej 0 również jest uznawane za zmianę znaku), różnica dzielona rzędu $r - 2$ zmienia znak co najwyżej dwa razy, i tak dalej.

Końcowo dostajemy, że $(f - p_f)(c_i)$, dla $0 \leq i \leq k$, zmienia znak co najwyżej r razy. Otrzymujemy sprzeczność, ponieważ $f(c_j) = p_f(c_j)$, dla $0 \leq i \leq r$. □

Przypomnijmy również, że błąd interpolacji wielomianowej wyraża się wzorem

$$f(x) - p(x) = f[x, t_0, \dots, t_r] \prod_{j=0}^r (x - t_j),$$

gdzie p jest wielomianem interpolacyjnym opartym na punktach t_0, \dots, t_r .

Analiza algorytmu wykorzystującego różnice dzielone do lokalizacji osłabiwości opiera się na własnościach interpolacji z użyciem zaburzonej informacji oraz punktowej analizie błędów aproksymacji. Analiza ta została opracowana w pracy [5] i w jej skład wchodzi dowody i lematy umieszczone w dalszej części tego rozdziału. Zaczniemy od przedstawiania kilku lematów dotyczących różnic dzielonych i błędów interpolacji, gdy mamy do czynienia z informacją niedokładną.

Poniższy lemat mówi, że wartość zaburzonej różnicy dzielonej dla funkcji regularnej na przedziale jest ograniczona ze względu na precyzję z jaką otrzymujemy wartości funkcji oraz przez stałą zależną od parametrów klasy.

Lemat 4.9. *Jeżeli $f \in H_{r,e}(t_i, t_{i+r+1})$, wtedy*

$$|\tilde{d}_i| \leq \frac{c(g_f)(r+1)^e}{(r+1)!} h^{e-1} + \delta \frac{2^{r+1}}{(r+1)!} h^{-(r+1)}.$$

Dowód. Z nierówności trójkąta mamy $|\tilde{d}_i| \leq |d_i| + |\tilde{d}_i - d_i|$. Wykorzystując własności różnic dzielonych oraz Hölderowskość funkcji f otrzymujemy

$$\begin{aligned} |d_i| &= \frac{|f[t_{i+1}, \dots, t_{i+r+1}] - f[t_i, \dots, t_{i+r}]|}{t_{i+r+1} - t_i} \\ &= \frac{1}{r!} \frac{|f^{(r)}(\xi_1) - f^{(r)}(\xi_2)|}{t_{i+r+1} - t_i} \leq \frac{c(g_f)}{r!} \frac{|\xi_1 - \xi_2|^e}{t_{i+r+1} - t_i}, \end{aligned} \tag{4.16}$$

a ponieważ $\xi_1 \in \text{conv}(t_{i+1}, \dots, t_{i+r+1})$ i $\xi_2 \in \text{conv}(t_i, \dots, t_{i+r})$, to $|\xi_1 - \xi_2| \leq t_{i+r+1} - t_i$, stąd pierwszy człon nierówności trójkąta możemy oszacować przez

$$|d_i| \leq \frac{c(g_f)}{r!} (t_{i+r+1} - t_i)^{e-1} \leq \frac{c(g_f)(r+1)^e}{(r+1)!} h^{e-1}.$$

Oszacowanie na drugi człon wynika z definicji różnicy dzielonej zaburzonej i niezaburzonej, gdzie $y_j = f(t_j) + e_j$ dla $|e_j| \leq \delta$,

$$\begin{aligned} |\tilde{d}_i - d_i| &= \left| \left(\sum_{j=i}^{i+r+1} f(t_j) - \sum_{j=i}^{i+r+1} (f(t_j) + e_j) \right) \prod_{\substack{\ell=i \\ \ell \neq j}}^{i+r+1} (t_\ell - t_j)^{-1} \right| \\ &= h^{-(r+1)} \left| \sum_{j=0}^{r+1} e_j \prod_{\substack{\ell=0 \\ \ell \neq j}}^{r+1} (\ell - j)^{-1} \right| \leq \delta h^{-(r+1)} \sum_{j=0}^{r+1} \prod_{\substack{\ell=0 \\ \ell \neq j}}^{r+1} |\ell - j|^{-1} = \delta \frac{2^{r+1}}{(r+1)!} h^{-(r+1)}, \end{aligned} \quad (4.17)$$

co dowodzi lemat. \square

Teraz oszacujemy błąd interpolacji i ekstrapolacji w obecności zaburzenia wartości funkcji. Niech p_i i \tilde{p}_i odpowiadają wielomianom stopnia co najwyżej r opartych na odpowiednio dokładnych i niedokładnych wartościach funkcji f w punktach $t_i, t_{i+1}, \dots, t_{i+r}$. Dla $r \geq 1$, wprowadźmy oznaczenia:

$$\beta_r = \max_{0 \leq t \leq r} \left| \prod_{k=0}^r (t - k) \right|, \quad \Lambda_r = \max_{0 \leq t \leq r} \sum_{k=0}^r \prod_{\substack{\ell=0 \\ \ell \neq k}}^r \left| \frac{t - \ell}{k - \ell} \right|, \quad \tilde{\Lambda}_r = \sum_{k=0}^r \prod_{\substack{\ell=0 \\ \ell \neq k}}^r \left| \frac{2r+1-\ell}{k-\ell} \right|.$$

Lemat 4.10. Niech $f \in H_{0,\varrho}$, wtedy:

dla $x \in [t_{i-\frac{1}{2}}, t_{i+\frac{1}{2}}]$:

$$|f(x) - \tilde{p}_i(x)| \leq C_{0,\varrho}(f) h^\varrho + \delta, \quad C_{0,\varrho}(f) = c(g_f) 2^{-\varrho},$$

dla $x \in [t_{i-1}, t_{i-\frac{1}{2}}) \cup (t_{i+\frac{1}{2}}, t_{i+1}]$:

$$|f(x) - \tilde{p}_i(x)| \leq C_{0,\varrho}(f) h^\varrho + \delta, \quad \bar{C}_{0,\varrho}(f) = c(g_f).$$

Niech $f \in H_{r,\varrho}$ i $r \geq 1$, wtedy:

dla $x \in [t_i, t_{i+r}]$:

$$|f(x) - \tilde{p}_i(x)| \leq C_{r,\varrho}(f) h^{r+\varrho} + \delta \Lambda_r, \quad C_{r,\varrho}(f) = c(g_f) 2^{-\varrho},$$

dla $x \in [t_{i-r-1}, t_i) \cup (t_{i+r}, t_{i+2r+1}]$:

$$|f(x) - \tilde{p}_i(x)| \leq \bar{C}_{r,\varrho}(f) h^{r+\varrho} + \delta \bar{\Lambda}_r, \quad \bar{C}_{r,\varrho}(f) = c(g_f) \frac{(2r+1)!(2r+1)^\varrho}{r(r!)^2}.$$

Dowód. W przypadku, gdy $r = 0$ interpolujemy funkcją stałą, więc błąd interpolacji dla $x \in [t_{i-\frac{1}{2}}, t_{i+\frac{1}{2}})$ wynosi

$$\begin{aligned} |f(x) - \tilde{p}_i(x)| &= |f(x) - f(t_i) - e_i| \leq |f(x) - f(t_i)| + |e_i| \\ &\leq c(g_f) \left| t_{i-\frac{1}{2}} - t_i \right|^\varrho + |e_i| \leq c(g_f) 2^{-\varrho} h^\varrho + \delta. \end{aligned}$$

Niech $r \geq 1$, wykorzystajmy nierówność trójkąta

$$|f(x) - \tilde{p}_i(x)| \leq |f(x) - p_i(x)| + |\tilde{p}_i(x) - p_i(x)|.$$

Jeżeli $x \in [t_i, t_{i+r}]$, wtedy z wyrażenia na błąd interpolacji Lagrange'a (4.2) dla pierwszego członu powyższej sumy, mamy:

$$\begin{aligned}
 |f(x) - p_i(x)| &= |(x - t_i) \cdots (x - t_{i+r}) f[t_i, \dots, t_{i+r}, x]| \\
 &\leq \beta_r h^{r+1} \frac{|f[t_{i+1}, \dots, t_{i+r}, x] - f[t_i, \dots, t_{i+r-1}, x]|}{t_{i+r} - t_i} \\
 &\leq \beta_r h^{r+1} \frac{1}{r!} \frac{|f^{(r)}(\xi_1) - f^{(r)}(\xi_2)|}{t_{i+r} - t_i} \leq \beta_r h^{r+1} \frac{c(g_f)}{r!} \frac{|\xi_1 - \xi_2|^e}{t_{i+r} - t_i}, \\
 &\leq \beta_r h^{r+1} \frac{c(g_f)}{r^{1-e} r!} h^{e-1} = C_{r,e}(f) h^{r+e}.
 \end{aligned} \tag{4.18}$$

gdzie $\xi_1 \in \text{conv}(t_{i+1}, \dots, t_{i+r})$ i $\xi_2 \in \text{conv}(t_i, \dots, t_{i+r-1})$.

Oszacowanie na drugi człon wynika z definicji wielomianów Lagrange'a

$$\begin{aligned}
 |\tilde{p}_i(x) - p_i(x)| &= \left| \sum_{k=i}^{i+r} f(x_k) \Phi_k(x) - \sum_{k=i}^{i+r} (f(x_k) + e_k) \Phi_k(x) \right| \\
 &\leq \left| \sum_{k=i}^{i+r} e_k \Phi_k(x) \right| \leq \delta \Lambda_r,
 \end{aligned} \tag{4.19}$$

gdzie Φ_i jest zdefiniowane jak w (3.1).

Przypadek dla $x \in [t_{i-r-1}, t_i] \cup (t_{i+r}, t_{i+2r+1}]$ jest analogiczny. \square

Rozważmy teraz błąd interpolacji i ekstrapolacji, przy obecności punktu osobliwego w przedziale zawierającym węzły interpolacji.

Lemat 4.11. Niech $f \in F_{r,e}$ oraz

$$\hat{t}_f \in \begin{cases} (t_{i-\frac{1}{2}}, t_{i+\frac{1}{2}}], & \text{gdy } r = 0, \\ (t_i, t_{i+r}], & \text{gdy } r \geq 1. \end{cases}$$

Przypuśćmy, że

$$|\tilde{d}_k| \leq B h^{e-1} \forall k. \tag{4.20}$$

Wtedy dla każdego $x \in [t_{i-1}, t_{i+1}]$, gdy $r = 0$ lub dla każdego $x \in [t_{i-r-1}, t_{i+2r+1}]$, gdy $r \geq 1$, mamy:

$$|f(x) - \tilde{p}_i(x)| \leq D_r(B, f) h^{r+e} + \delta \Lambda_r,$$

gdzie

$$D_r(B, f) = \begin{cases} c(g_f) + B, & \text{gdy } r = 0, \\ c(g_f) \frac{\beta_r (r+1)^e}{r r!} + B (2^{r+1} - 1) \frac{(2r)!}{(r-1)!}, & \text{gdy } r \geq 1. \end{cases}$$

Dowód. Przypadki, gdy $\hat{t}_f \leq x$ i $\hat{t}_f > x$ są analogiczne. Weźmy $\hat{t}_f \leq x$. Jeżeli $r = 0$, wtedy z nierówności trójkąta mamy

$$|f(x) - \tilde{p}_i(x)| \leq |f(x) - p_{i+1}(x)| + |p_{i+1}(x) - \tilde{p}_{i+1}(x)| + |\tilde{p}_{i+1}(x) - \tilde{p}_i(x)|,$$

gdzie $p_{i+1}(x) = f(t_{i+1})$, $\tilde{p}_{i+1}(x) = \tilde{f}(t_{i+1}) = f(t_{i+1}) + e_{i+1}$ i $\tilde{p}_i(x) = \tilde{f}(t_i)$.
Zauważmy, że

$$|\tilde{p}_{i+1}(x) - \tilde{p}_i(x)| = \frac{|\tilde{f}(t_{i+1}) - \tilde{f}(t_i)|}{t_{i+1} - t_i} (t_{i+1} - t_i) = |\tilde{d}_i| h \leq B h^e.$$

Korzystając dodatkowo z regularności f oraz ograniczenia na zaburzenie, otrzymujemy oszacowanie

$$|f(x) - \tilde{p}_i(x)| \leq c(g_f) h^e + \delta + B h^e = (c(g_f) + B) h^e + \delta,$$

co pokazuje pierwszą część lematu.

Założmy, że $r \geq 1$ i $\hat{t}_f \leq x < t_{i+r}$. Wybierzmy najmniejszy indeks j taki, że $\hat{t}_f \leq t_j$. Oczywiście $i + 1 \leq j \leq i + r$ oraz $x \in [t_{j-1}, t_{j+r}]$. Otrzymujemy

$$|f(x) - \tilde{p}_i(x)| \leq |f(x) - p_j(x)| + |p_j(x) - \tilde{p}_j(x)| + |\tilde{p}_j(x) - \tilde{p}_i(x)|. \quad (4.21)$$

A ponieważ $\hat{t}_f \notin (t_j, t_{j+r}]$, to postępując analogicznie jak w (4.18), otrzymujemy

$$|f(x) - p_j(x)| \leq c(g_f) \beta_r h^{r+1} \frac{1}{r!} \frac{(t_{j+r} - t_{j-1})^e}{t_{j+r} - t_j} = C_{r,e}(f) \left(1 + \frac{1}{r}\right)^e h^{r+e}.$$

Tak jak w równaniu (4.19), mamy

$$|p_j(x) - \tilde{p}_j(x)| \leq \delta \Lambda_r.$$

Możemy teraz oszacować pozostały człon $|\tilde{p}_j(x) - \tilde{p}_i(x)|$. Dla $i + r + 1 \leq k \leq j + r$, mamy

$$(\tilde{f} - \tilde{p}_i)[t_i, \dots, t_{i+r}, t_k] = \frac{y_k - \tilde{p}_i(t_k)}{(k - i)(k - i - 1) \dots (k - i - r) h^{r+1}} \quad (4.22)$$

oraz

$$\left| (\tilde{f} - \tilde{p}_i)[t_i, \dots, t_{i+r}, t_k] \right| = \left| \tilde{f}[t_i, \dots, t_{i+r}, t_k] \right| \leq \max_{i \leq \ell \leq k-r-1} |\tilde{d}_\ell| \leq B h^{e-1}, \quad (4.23)$$

gdzie pierwsza nierówność wynika z lematu 4.8, natomiast druga z (4.20). Biorąc (4.22) oraz (4.23), otrzymujemy:

$$|y_k - \tilde{p}_i(t_k)| \leq \frac{(2r)!}{(r-1)!} B h^{r+e}. \quad (4.24)$$

Także ostatni człon nierówności (4.21) możemy oszacować następująco

$$\begin{aligned} |\tilde{p}_j(x) - \tilde{p}_i(x)| &= \left| \sum_{k=j}^{j+r} (\tilde{p}_j(t_k) - \tilde{p}_i(t_k)) \prod_{\substack{s=j \\ s \neq k}}^{j+r} \frac{x - t_s}{t_k - t_s} \right| \\ &\leq \left(\max_{j \leq k \leq j+r} |y_k - \tilde{p}_i(t_k)| \right) \left(\max_{0 \leq t \leq r+1} \sum_{k=0}^r \prod_{\substack{s=0 \\ s \neq k}}^r \left| \frac{t - s}{k - s} \right| \right). \end{aligned} \quad (4.25)$$

Pierwsze maksimum z powyższego równania jest oszacowane poprzez (4.24). Natomiast drugie maksimum jest osiągane dla $t = r + 1$ i jest równe

$$\sum_{k=0}^r \prod_{\substack{s=0 \\ s \neq k}}^r \left| \frac{r+1-s}{k-s} \right| = \sum_{k=0}^r \binom{r+1}{k} = 2^{r+1} - 1.$$

Stąd

$$|\tilde{p}_j(x) - \tilde{p}_i(x)| \leq \frac{(2r)!}{(r-1)!} (2^{r+1} - 1) B h^{r+e}.$$

□

Przeprowadzimy teraz punktową analizę błędu, tzn. analizę wartości $|f(x) - \varphi_h^*(y_h)|$ dla każdego x . Rozważmy kilka przypadków w zależności od lokalizacji punktu osobliwego \hat{t}_f . Przypomnijmy, że

$$(u_3, v_3] \subseteq (u_2, v_2] \subseteq (u_1, v_1],$$

gdzie u_i, v_i dla $i = 1, 2, 3$ są punktami zlokalizowanymi w krokach 1-3 algorytmu. Dla przypadku I przyjmijmy, że

$$\delta \leq b h^{r+e}, \quad \text{dla pewnej stałej } b > 0. \quad (4.26)$$

Uzasadnienie tego wyjaśnimy później, zobacz uwagę 4.14.

Przypadek I: $\hat{t}_f \notin (u_1, v_1]$. Przy takich założeniach, z lematu 4.9 oraz z (4.26) mamy, że

$$|\tilde{d}_i| \leq B_r(b, f) h^{e-1}, \quad \text{gdzie } B_r(b, f) = \frac{c(g_f)(r+1)^e + b 2^{r+1}}{(r+1)!}.$$

Przypuśćmy, że $\varphi_h^*(y)(x) = \tilde{p}_i(x)$ dla pewnego i . Jeżeli $x \notin [u_1, v_1]$, wtedy $x \in [t_i, t_{i+r})$ (albo $x \in [t_{i-1/2}, t_{i+1/2})$ dla $r = 0$) oraz $[t_i, t_{i+r}] \cap (u_1, v_1) = \emptyset$. Wtedy z lematów 4.10 i 4.11 wiemy, że błąd aproksymacji w punkcie x jest ograniczony z góry przez $C_{r,\varrho}(f)h^{r+e} + \delta\Lambda_r$ dla $t_f \notin (t_i, t_{i+r})$ (albo $\hat{t}_f \notin [x_{i-1/2}, x_{i+1/2})$ dla $r = 0$) lub przez $E_r(f)h^{r+e} + \delta\Lambda_r$, gdy $\hat{t}_f \in (t_i, t_{i+r})$, gdzie $E_r(f) = D_r(B_r(b, f), f)$.

Z drugiej strony, jeżeli $x \in [u_1, v_1]$, wtedy znów z lematu 4.11 wiemy, że błąd jest ograniczony przez $\bar{C}_{r,\varrho}h^{r+e} + \delta\bar{\Lambda}_r$ dla $\hat{t}_f \notin (t_i, t_{i+r})$ (albo $\hat{t}_f \notin (t_{i-1/2}, t_{i+1/2})$ dla $r = 0$), lub przez $E_r(f)h^{r+e} + \delta\Lambda_r$ w przeciwnym przypadku.

Przypadek II: $\hat{t}_f \in (u_1, v_1]$. Bez straty ogólności załóżmy, że

$$u_1 < \hat{t}_f \leq \xi \leq v_1.$$

Z lematu 4.10 dla $x \notin [u_1, v_1]$ wiemy, że błąd jest ograniczony z góry przez $C_{r,\varrho}(f)h^{r+e} + \delta\Lambda_r$ a dla $x \in [u_1, \hat{t}_f) \cup [\xi, v_1]$ przez $\bar{C}_{r,\varrho}h^{r+e} + \delta\bar{\Lambda}_r$. Z tego powodu możemy założyć, że

$$x \in [\hat{t}_f, \xi). \quad (4.27)$$

Mamy trzy możliwości:

Przypadek IIa: $\hat{t}_f \in (u_1, v_1] \setminus (u_2, v_2]$. Z przyjętych założeń wynika, że w kroku 2. algorytmu, w pewnej iteracji musi zachodzić

$$\left| \tilde{f}(z_{j*}) - \tilde{p}_-(z_{j*}) \right| \leq \left| \tilde{f}(z_{j*}) - \tilde{p}_+(z_{j*}) \right| \quad \text{oraz} \quad \hat{t}_f \in (u, z_{j*}],$$

co powoduje, że dla wszystkich $1 \leq j \leq r+1$ zachodzi

$$\begin{aligned} |\tilde{p}_+(z_j) - \tilde{p}_-(z_j)| &\leq |\tilde{p}_+(z_{j*}) - \tilde{p}_-(z_{j*})| \leq \left| \tilde{f}(z_{j*}) - \tilde{p}_-(z_{j*}) \right| + \left| \tilde{f}(z_{j*}) - \tilde{p}_+(z_{j*}) \right| \\ &\leq 2 \left| \tilde{f}(z_{j*}) - \tilde{p}_+(z_{j*}) \right| \leq 2 (C_{r,\varrho}(f)h^{r+\varrho} + \delta\Lambda_r). \end{aligned}$$

Jeżeli wielomian p stopnia co najwyżej r jest $|p(z_j)| \leq a, 1 \leq j \leq r+1$, wtedy dla wszystkich $u_1 \leq x < v_1$ jest $|p(x)| \leq (2^{r+1} - 1)a$, zobacz (4.25). Otrzymaliśmy

$$\begin{aligned} |f(x) - \varphi_h^*(\mathbf{y})(x)| &= |f(x) - \tilde{p}_-(x)| \leq |f(x) - \tilde{p}_+(x)| + |\tilde{p}_+(x) - \tilde{p}_-(x)| \\ &= 2^{r+1} (C_{r,\varrho}(f)h^{r+\varrho} + \delta\Lambda_r). \end{aligned}$$

Przypadek IIb: $\hat{t}_f \in (u_2, v_2] \setminus (u_3, v_3]$. Z przyjętych założeń wynika, że w kroku 3. algorytmu, w pewnej iteracji musi zachodzić

$$\left| \tilde{f}(z) - \tilde{p}_-(z) \right| \leq \left| \tilde{f}(z) - \tilde{p}_+(z) \right| \quad \text{oraz} \quad \hat{t}_f \in (u, z], \quad (4.28)$$

a to wraz z (4.27) wskazują na to, że

$$|\tilde{p}_+(x) - \tilde{p}_-(x)| \leq \max(|\tilde{p}_+(z) - \tilde{p}_-(z)|, |\tilde{p}_+(\hat{t}_f) - \tilde{p}_-(\hat{t}_f)|),$$

a ponieważ w przeciwnym przypadku z nie byłoby największym lokalnym maksimum wyrażenia $|\tilde{p}_+ - \tilde{p}_-|$ na przedziale interval (u, v) . Następnie otrzymujemy

$$\begin{aligned} |\tilde{p}_+(\hat{t}_f) - \tilde{p}_-(\hat{t}_f)| &\leq |f(\hat{t}_f^+) - \tilde{p}_+(\hat{t}_f)| + |f(\hat{t}_f^-) - \tilde{p}_-(\hat{t}_f)| + |f(\hat{t}_f^+) - f(\hat{t}_f^-)| \\ &\leq 2 (\bar{C}_{r,\varrho}(f)h^{r+\varrho} + \delta\bar{\Lambda}_r) + \left| \Delta_f^{(0)} \right|, \end{aligned}$$

a z (4.28) mamy

$$|\tilde{p}_+(z) - \tilde{p}_-(z)| \leq 2 \left| \tilde{f}(z) - \tilde{p}_+(z) \right| \leq 2 (\bar{C}_{r,\varrho}(f)h^{r+\varrho} + \delta\bar{\Lambda}_r).$$

Ostatecznie otrzymujemy

$$\begin{aligned} |f(x) - \varphi_h^*(\mathbf{y})(x)| &= |f(x) - \tilde{p}_-(x)| \leq |f(x) - \tilde{p}_+(x)| + |\tilde{p}_+(x) - \tilde{p}_-(x)| \\ &\leq 3 (\bar{C}_{r,\varrho}(f)h^{r+\varrho} + \delta\bar{\Lambda}_r) + \left| \Delta_f^{(0)} \right|. \end{aligned}$$

Przypadek IIc: $\hat{t}_f \in (u_3, v_3]$ Ponieważ funkcja $|\tilde{p}_+ - \tilde{p}_-|$ nie ma żadnego lokalnego maksimum w przedziale (u_3, v_3) , to wiemy, że jest nierosnąca w przedziale $[\hat{t}_f, \xi)$. Dlatego znów otrzymujemy

$$|\tilde{p}_+(x) - \tilde{p}_-(x)| \leq |\tilde{p}_+(\hat{t}_f) - \tilde{p}_-(\hat{t}_f)| \leq 2 (\bar{C}_{r,\varrho}(f)h^{r+\varrho} + \delta\bar{\Lambda}_r) + \left| \Delta_f^{(0)} \right|$$

oraz

$$\begin{aligned} |f(x) - \varphi_h^*(\mathbf{y})(x)| &= |f(x) - \tilde{p}_-(x)| \leq |f(x) - \tilde{p}_+(x)| + |\tilde{p}_+(x) - \tilde{p}_-(x)| \\ &\leq 3 (\bar{C}_{r,\varrho}(f)h^{r+\varrho} + \delta\bar{\Lambda}_r) + \left| \Delta_f^{(0)} \right|. \end{aligned}$$

Podsumowując analizę błędu dla każdego punktu otrzymujemy, że dla $\delta \leq bh^{r+\varrho}$ mamy

$$\begin{cases} |f(x) - \varphi_h^*(y_h)(x)| = \mathcal{O}(\max(1, c(g_f))h^{r+\varrho}) & \text{dla } x \notin (u_2, v_2], \\ |f(x) - \varphi_h^*(y_h)(x)| = \mathcal{O}(\max(1, c(g_f))h^{r+\varrho} + |\Delta_f^{(0)}|) & \text{dla } x \in (u_2, v_2], \end{cases}$$

gdzie $v_2 - u_2 \leq \omega$.

Wiemy również, że liczba ewaluacji funkcji n jest proporcjonalna do h^{-1} , tak więc $h^{r+\varrho}$ jest proporcjonalne do $n^{-(r+\varrho)}$. Z tej obserwacji oraz przedstawionej analizy wynikają poniższe stwierdzenia dotyczące błędu najgorszego przypadku.

Stwierdzenie 4.12. *Niech $1 \leq p \leq \infty$. Jeżeli $\delta \leq bh^{r+\varrho}$ oraz $\omega(h) = h^{(r+\varrho)p+1}$, wtedy*

$$e_p^{\text{wor}}(\varphi_h^*, N_h^*; \mathcal{F}_{r,\varrho}^D) = \mathcal{O}(n^{-(r+\varrho)}).$$

Przypomnijmy, że powyższe ograniczenia górne nie może zostać spełnione przez algorytmy nieadaptacyjne, co zostało pokazane w [2]. Pokazano tam również, że dla $p = \infty$ nie istnieje algorytm z błędem zbiegającym do zera, dlatego założenia $p < \infty$ jest niezbędne. Dodatkowo, gdy rozważymy klasę $\mathcal{F}_{r,\varrho}^C \subset \mathcal{F}_{r,\varrho}^D$, to możemy uprościć algorytm biorąc $\omega(h) = (r+1)h$ i unikając iteracji w kroku 2. Otrzymujemy w ten sposób algorytm, który dla $r = 0, 1$ jest nieadaptacyjny, a dla $r \geq 2$ używa co najwyżej $r-1$ dodatkowych punktów, niezależnie od tego jak małe jest h . Co więcej, ograniczenie górne zachodzi dla $p = \infty$. Stosując powyższą modyfikację możemy sformułować następujące stwierdzenie.

Stwierdzenie 4.13. *Jeżeli $\delta \leq bh^{r+\varrho}$ i $\omega(h) = (r+1)h$, wtedy:*

$$e_\infty^{\text{wor}}(\varphi_h^*, N_h^*; \mathcal{F}_{r,\varrho}^C) = \mathcal{O}(n^{-(r+\varrho)}). \quad (4.29)$$

Ponownie, dla $r \geq 2$ użycie informacji adaptacyjnej jest konieczne. Łącząc wyniki stwierdzeń 4.12, 4.13 i 2.4(i) otrzymujemy twierdzenie 2.5(i)-(ii). Faktycznie, dla ustalonego δ i n możemy wybrać $h = \frac{T}{m}$ takie, że

$$m = m(n, \delta) = \left\lceil \min \left(\beta n, \frac{1}{T} \left(\frac{b}{\delta} \right)^{\frac{1}{r+\varrho}} \right) \right\rceil = \Theta(\min(n, \delta^{-1/(r+\varrho)})). \quad (4.30)$$

Uwaga 4.14. *Zauważmy, że dla ustalonej precyzji δ nie ma sensu brać m więcej niż $m_{\max} = \Theta(\delta^{-1/(r+\varrho)})$ wartości funkcji, ponieważ dla $m = m_{\max}$ osiągamy maksymalną dokładność dla danego δ .*

Testy numeryczne

W tym rozdziale porównamy działanie algorytmów φ^{KP} i φ^{MP} w praktyce. Oba algorytmy zostały zaimplementowane z użyciem języka Python, wykorzystując m.in. popularne biblioteki `numpy`, `scipy` czy `matplotlib` (wizualizacja). Pewnym ograniczeniem wynikającym z zastosowania tych modułów jest brak natywnego wsparcia dla precyzji większej niż `float64`, która gwarantuje dokładność na poziomie około 15-ego miejsca po przecinku. Okazało się to szczególnie ważne dla algorytmu φ^{KP} ze względu na wybór bardzo małych przedziałów do interpolacji wielomianowej. Zauważmy, że dla stałych zdefiniowanych jak w algorytmie, przy danej precyzji η nie ma sensu brać m większego niż

$$\frac{h^{r+e}}{r+1} = \eta \implies m = T((r+1)\eta)^{-\frac{1}{r+e}},$$

ponieważ odległość między punktami użytymi do interpolacji, jest mniejsza niż błąd reprezentacji.

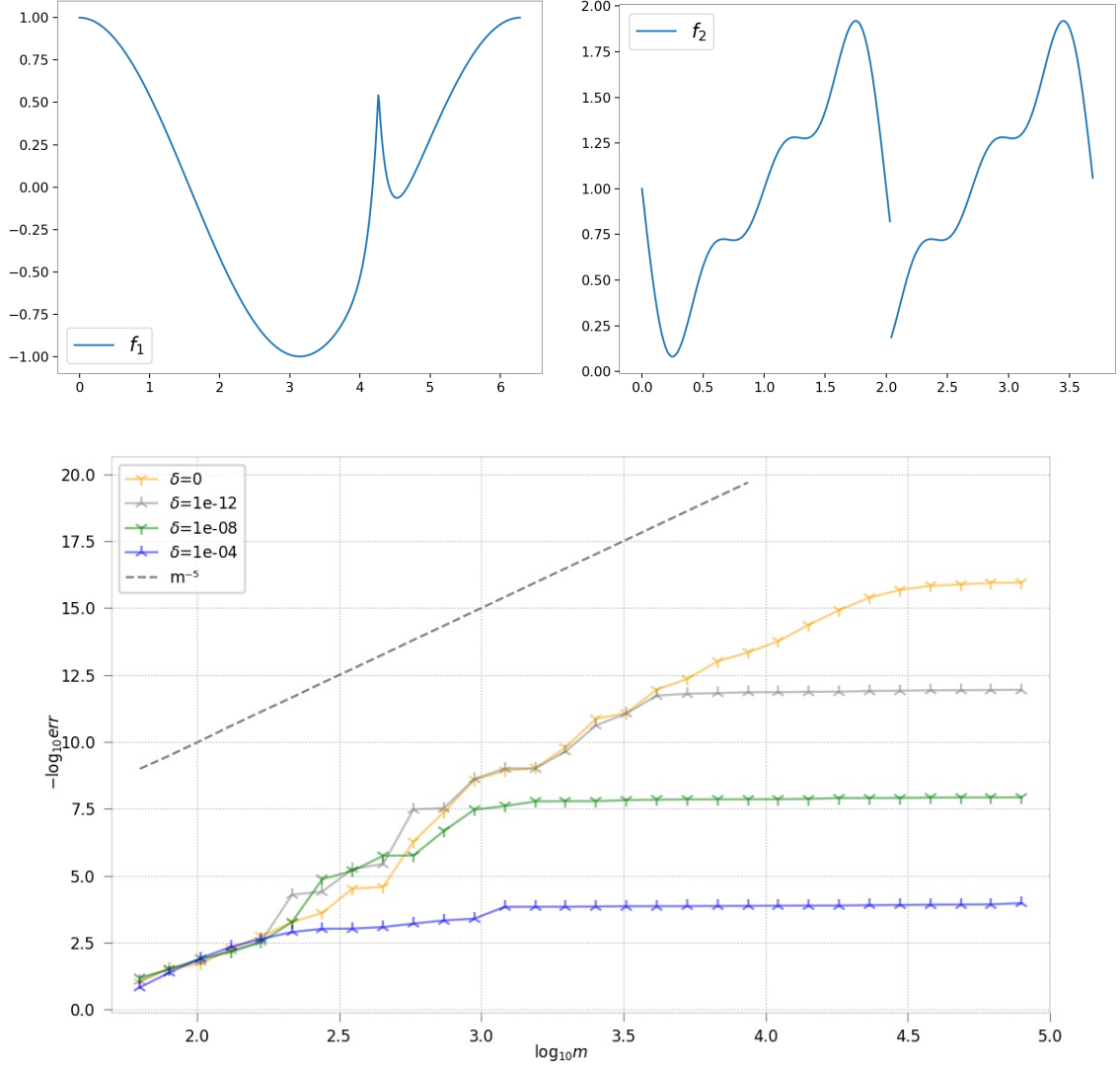
Algorytmy testowaliśmy na funkcjach:

$$f_1(x) = \cos(x) + e^{-8|x-t_f|} \quad x \in [0, 2\pi], \quad t_f \in \left[\frac{2}{3}\pi, \frac{4}{3}\pi\right]$$

$$f_2(x) = \begin{cases} \bar{f}_2(x), & \text{gdy } x \in [0, t_f], \\ \bar{f}_2(x+c), & \text{gdy } x \in [t_f, 4-c], \end{cases}$$

gdzie $\bar{f}_2(x) = 1 - \frac{2}{\pi} \sum_{n=1}^3 \frac{1}{n} \sin(n\pi x)$, $c = 0.3$ i $t_f \in [\frac{T}{2} - \frac{1}{5}, \frac{T}{2} + \frac{1}{5}]$.

Osobliwość w podanych funkcjach występuje w punkcie t_f . Podczas testów punkt ten był losowany z podanych przedziałów, które wraz z stałą c , zostały dobrane tak, aby losowane funkcje reprezentowały różne przypadki. Zauważmy, że funkcja f_2 w pobliżu



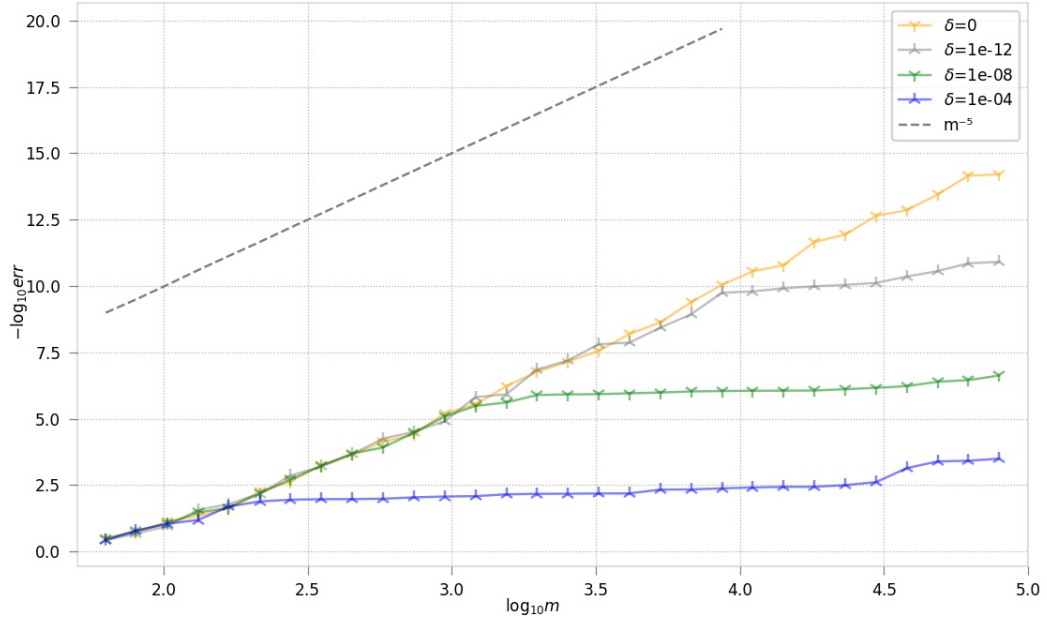
Rys. 5.1. φ^{MP} dla funkcji f_1 z $r = 4$ w normie L^2 z $r = 4$.

osobliwości może być rosnąca lub malejąca, w zależności wyniku losowania. Dzięki temu zabiegowi urozmaicamy aproksymowane funkcji, ale też zmniejszamy liczbę trywialnych przypadków, takich jak położenie osobliwości w węźle czy w jego pobliżu (aproksymacja funkcją stałą w φ^{KP}).

Obie funkcje są okresowe oraz należą zarówno do klasy $\mathcal{F}_{r,\varrho}$, jak i $\mathcal{G}_{r,\varrho}$, jednak funkcja f_2 jest nieciągła, czyli nie zachodzi $\Delta_{f_2}^{(0)} = 0$. Algorytmy były testowane w normach L^2 i L^∞ . Dla każdego zestawu danych wejściowych, tj. dla określonego m , δ i ϱ , błąd najgorszego przypadku obliczaliśmy jako maksimum z kilkuset uruchomień algorytmu. W ramach obliczania pojedynczego błędu t_f było ustalone.

Dla każdego z algorytmów, końcowa aproksymacja jest funkcją opartą na `float64`, z tego powodu błąd aproksymacji dla informacji dokładnej również jest ograniczony przez precyzję reprezentacji.

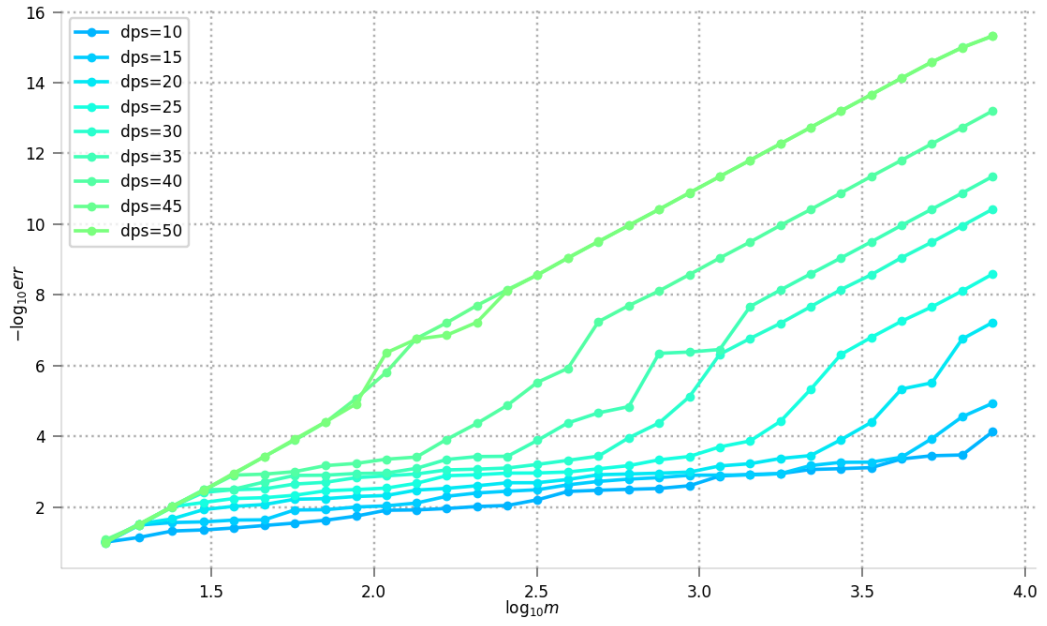
Algorytm φ^{MP} zachowywał teoretyczne tempo zbieżności dla obu funkcji w obu



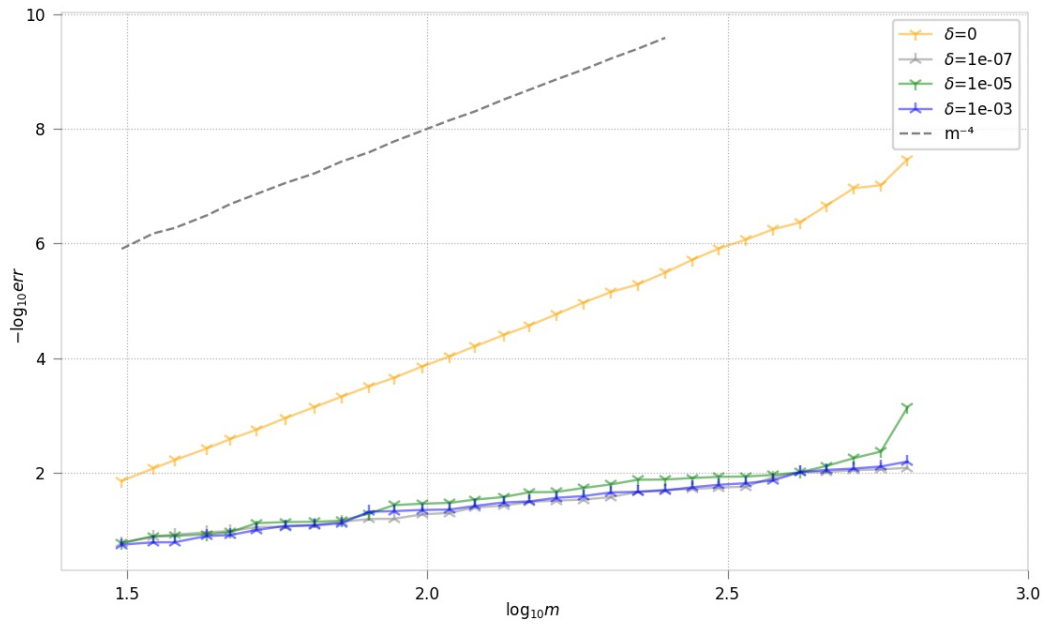
Rys. 5.2. φ^{MP} dla funkcji f_1 z $r = 4$ w normie L^∞ .

normach, zobacz rysunki 5.1 i 5.2. Równanie prostej wyznaczającej teoretyczny błąd aproksymacji nie zawiera żadnego przesunięcia na żadnym z wykresów w tej pracy.

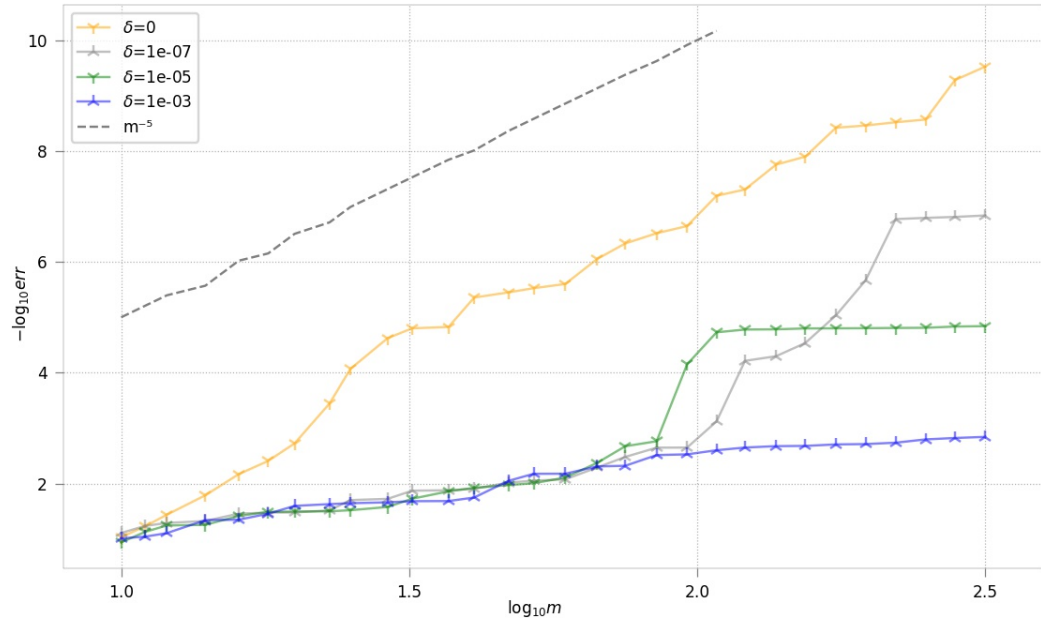
Testując algorytm oparty na wielomianach Lagrange’a, wyniki otrzymywane z użyciem standardowej precyzji, okazały się niestabilne. Algorytm miał trudności z poprawnym wyznaczeniem przedziału z punktem osobliwym. Aby potwierdzić, że za wyniki niezgodne z oczekiwaniami odpowiada precyzja obliczeń, zaimplementowaliśmy część algorytmu z użyciem biblioteki mpmath. Pozwala ona na przeprowadzenie obliczeń na liczbach o dowolnej precyzji. W bibliotece mpmath dokładność obliczeń ustawia się parametrem `dps`, oznaczającego *miejsca dziesiętne* (*ang. decimal places*). Zastosowanie danej precyzji odpowiada użyciu typu `float` z mantysą reprezentowaną na około $3.33 \cdot \text{dps}$ bitach. Z tą zmianą wiążą się dwie ważne kwestie: wydłużenie czasu obliczeń oraz brak kompatybilności z innymi bibliotekami (numpy, itp.). Z początku zwiększyliśmy dokładność obliczeń w pierwszym kroku algorytmu, jednak nie przyniosło to oczekiwanych efektów. Po wnikliwej analizie okazało się, że drugim problemem jest prawie zerowa (zerowa przy niższej precyzji) zmiana wartości funkcji na bardzo małych przedziałach, powodująca duże skoki w wartościach różnic dzielonych. Dopiero zwiększenie dokładności otrzymywanych wartości aproksymowanej funkcji rozwiązało problem. Rysunek 5.3 przedstawia zależność błędów od liczby punktów w początkowym podziale przy zmieniającej się precyzji obliczeń dla informacji bez szumu. Zwiększonej dokładności użyliśmy tylko do lokalizacji punktu osobliwego oraz dodawania nowych punktów w kroku drugim algorytmu. Końcowa aproksymacja nadal bazuje na `float64`. Po wprowadzeniu omówionych usprawnień algorytm zachowywał wszystkie teoretyczne własności dla informacji nie zburzonej. Jednak w przypadku informacji niedokładnej, algorytm wciąż błędnie wyznaczał przedziały z punktem osobliwym, co widać na rysunku 5.4. Jednak w normie L^2 błąd prezentował się dużo lepiej, rysunki 5.5 i 5.6



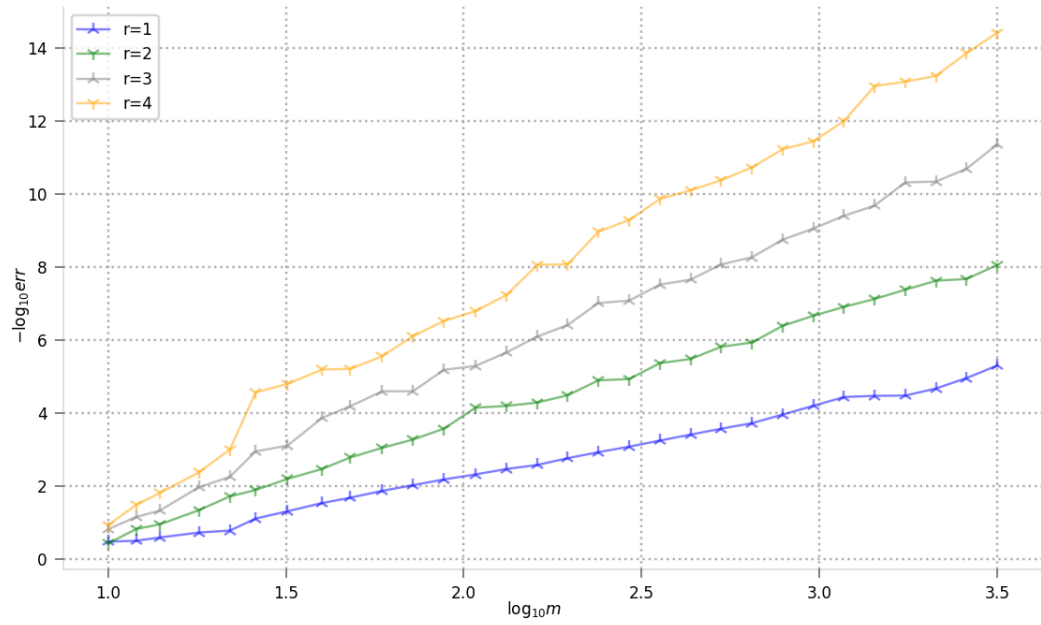
Rys. 5.3. φ^{KP} dla funkcji f_1 z $r = 4$ oraz $\delta = 0$ w normie L^∞ dla różnych precyzji reprezentacji.



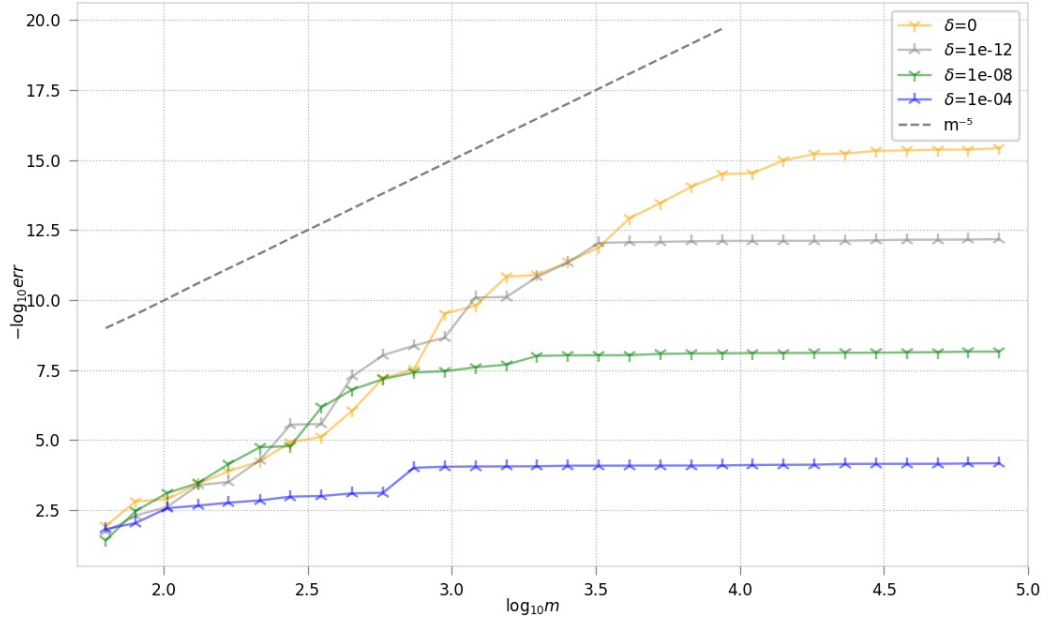
Rys. 5.4. φ^{KP} dla funkcji f_1 z $r = 4$ w normie L^∞ .



Rys. 5.5. φ^{KP} dla funkcji f_1 z $r = 1, 2, 3, 4$ w normie L^2 .



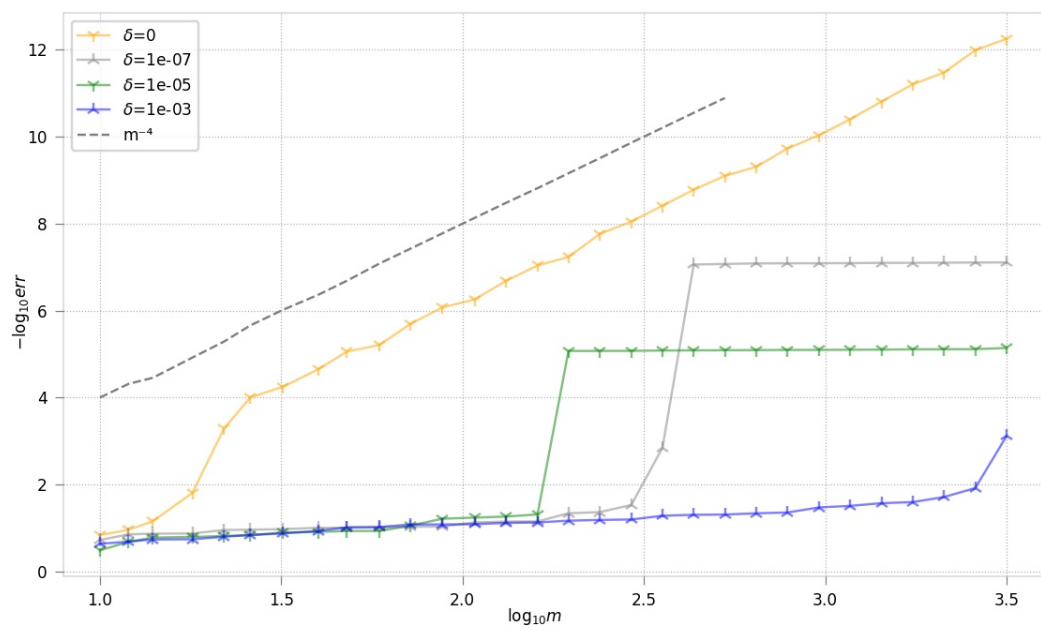
Rys. 5.6. φ^{KP} dla funkcji f_1 z $r = 1, 2, 3, 4$ i $\delta = 0.5h^{r+1}$ w normie L^2 .



Rys. 5.7. φ^{MP} dla funkcji f_2 z $r = 4$ w normie L^2 .

Kolejnym testem była weryfikacja zachowania konstruowanych aproksymacji dla nieciągłej funkcji f_2 w normie L^2 . Aproksymacja skonstruowana algorytmem φ^{MP} utrzymywała własności teoretyczne również dla tej funkcji, zobacz 5.7.

Algorytm φ^{KP} , pomimo że warunek z twierdzenia 2.5 nie był spełniony, osiągał minimalny próg błędu wynikający z obecności szumu, jednak dopiero po przekroczeniu pewnej liczby początkowych przedziałów, zobacz rysunek 5.8.



Rys. 5.8. $-\log_{10}(\text{err})$ vs $\log_{10}(m)$ dla funkcji f_2 z $r = 4$ w normie L^2 .

Bibliografia

- [1] F. Arandiga, A. Cohen, R. Donat, N. Dyn, *Interpolation and approximation of piecewise smooth functions*, SIAM J. Numer. Anal. 43 (2005) 41–57
- [2] L. Plaskota, G. W. Wasilkowski, Y. Zhao, *The power of adaption for approximating functions with singularities*, Mathematics Of Computation 77 2008, p. 2309–2338
- [3] L. Plaskota, G. W. Wasilkowski, *Uniform approximation of piecewise r -smooth and globally continuous functions*, SIAM Journal on Numerical Analysis, Vol. 47, No. 1 (2008/2009)
- [4] B. Kacewicz, P. Przybyłowicz, *Complexity of the derivative-free solution of systems of IVPs with unknown singularity hypersurface*, Journal of Complexity
- [5] P. M. Morkisz, L. Plaskota, *Approximation of piecewise Hölder functions from inexact information*, Journal of Complexity
- [6] J. F. Traub, H. Woźniakowski, G. W. Wasilkowski *Information-Based Complexity*, Academic Press, New York, 1988
- [7] E. Novak *Deterministic and Stochastic Error Bounds in Numerical Analysis*, Springer Verlag, Berlin, 1988