Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie
Wydział Matematyki Stosowanej
Katedra Analizy Matematycznej, Matematyki Obliczeniowej
i Metod Probabilistycznych

Rozprawa doktorska

Optymalne algorytmy randomizacyjne i kwantowe dla problemów nieliniowych

Maciej Goćwin

Promotor: prof. dr hab. Bolesław Kacewicz

Kraków 2012

Dziękuję mojemu Promotorowi,
prof. Bolesławowi Kacewiczowi,
za poświęcony czas oraz cenne rady
związane z niniejszą rozprawą,
a także za wsparcie i wyrozumiałość.

Spis treści

Streszczenie . . . 3
Abstract . . . 4
Wstęp . . . 5
Rozdział 1. Wprowadzenie . . . 9
1.1. Złożoność oparta na informacji . . . 9
1.1.1. Wstęp . . . 9
1.1.2. Informacja i algorytm . . . 10
1.1.3. Modele obliczeń . . . 12
1.2. Model obliczeń kwantowych . . . 14
1.2.1. Wstęp . . . 14
1.2.2. Model obliczeniowy . . . 15
1.2.3. Algorytm kwantowy . . . 18
Rozdział 2. Przydatne wyniki o złożoności wybranych problemów . . . 24
2.1. Wstęp . . . 24
2.2. Problemy dyskretne . . . 25
2.2.1. Przeszukiwanie baz danych . . . 25
2.2.2. Maksymalizacja funkcji dyskretnych . . . 26
2.2.3. Średnia liczb rzeczywistych . . . 28
2.3. Problem ciągły – równania różniczkowe . . . 29
2.3.1. Układy równań różniczkowych . . . 29
2.3.2. Równania różniczkowe wyższych rzędów . . . 35
Rozdział 3. Maksymalizacja funkcji ciągłych . . . 38
3.1. Wstęp . . . 38
3.2. Sformułowanie problemu i znane wyniki . . . 39
3.3. Górne ograniczenia w modelu kwantowym . . . 40
3.4. Dolne ograniczenia w modelu kwantowym . . . 43
3.5. Podsumowanie . . . 45
Rozdział 4. Równania nieliniowe . . . 47
4.1. Wstęp . . . 47
4.2. Sformułowanie problemu i znane wyniki . . . 48

4.3. Górne ograniczenia 50
4.4. Dolne ograniczenia 55
4.5. Podsumowanie 57
Rozdział 5. Problemy brzegowe 58
5.1. Wstęp 58
5.2. Liniowy dwupunktowy problem brzegowy 59
5.2.1. Sformułowanie problemu 59
5.2.2. Znane wyniki w modelu deterministycznym 60
5.2.3. Wstępne wyniki 61
5.2.4. Algorytmy 70
5.2.5. Górne ograniczenia w modelu deterministycznym 74
5.2.6. Górne ograniczenia w modelach randomizacyjnym i kwantowym 77
5.2.7. Dolne ograniczenia 82
5.2.8. Podsumowanie 86
5.3. Nieliniowy dwupunktowy problem brzegowy 87
5.3.1. Sformułowanie problemu i znane wyniki 87
5.3.2. Algorytmy 91
5.3.3. Górne ograniczenia 95
5.3.4. Dolne ograniczenia dla problemów (5.53) i (5.57) 99
5.3.5. Podsumowanie 102
Rozdział 6. Podsumowanie 104
Dodatek A. Fakty pomocnicze 107
Dodatek B. Dowody pomocniczych wyników 111
Bibliografia 118

Streszczenie

W rozprawie zajmujemy się badaniem złożoności oraz poszukiwaniem optymalnych algorytmów dla pewnych problemów nieliniowych w modelach randomizacyjnym i kwantowym. Zajmujemy się takimi problemami jak: maksymalizacja funkcji, szukanie zer funkcji, rozwiązywanie problemów brzegowych dla równań różniczkowych.

Pierwszym z rozważanych problemów jest problem maksymalizacji funkcji z klasy Höldera. Analizujemy złożoność tego problemu w modelu kwantowym. Wyznaczamy górne i dolne ograniczenia na złożoność tego problemu oraz konstruujemy optymalny algorytm. Porównując uzyskane wyniki ze znanymi ograniczeniami w modelu deterministycznym i randomizacyjnym otrzymujemy, że algorytmy kwantowe są kwadratowo szybsze dla tego problemu.

Kolejnym z rozważanych przez nas problemów jest poszukiwanie zer funkcji z klasy Höldera. Wyznaczamy dla tego problemu górne i dolne ograniczenia na złożoność w modelach randomizacyjnym i kwantowym. Z otrzymanych wyników wynika, że algorytmy randomizacyjne nie dają przewagi nad algorytmami deterministycznymi, natomiast algorytmy kwantowe są kwadratowo szybsze.

Ostatnim z rozważanych problemów jest rozwiązywanie problemów brzegowych dla równań różniczkowych z różniczkowalnymi funkcjami prawej strony o ograniczonych pochodnych. Analizujemy dwa problemy: liniowy i nieliniowy. Dla problemu liniowego (który definiuje nieliniowy operator rozwiązania) poszukujemy ograniczeń na złożoność w modelu deterministycznym z informacją standardową oraz liniową, modelu randomizacyjnym i kwantowym. Problem nieliniowy analizujemy w modelu randomizacyjnym i kwantowym. Ograniczenia na złożoność w modelu deterministycznym dla tego problemu są znane. Dla obu problemów pokazujemy, że algorytmy randomizacyjne dają przyspieszenie w stosunku do deterministycznych o $1/2$ w mianowniku wykładnika, a algorytmy kwantowe o 1.

Słowa kluczowe

złożoność, model randomizacyjny, model kwantowy, algorytm optymalny, problemy nieliniowe , maksymalizacja funkcji, równania nieliniowe, problemy brzegowe

Abstract

We study in this thesis the complexity of some nonlinear problems in the randomized and quantum settings. We consider problems such as function maximization, searching the zeros of functions and solving boundary value problems for the differential equations.

The first problem we consider is maximization of functions from the Hölder class in the quantum model. We analyze the complexity of this problem and construct the optimal algorithm. Comparing obtained results with known bounds in the deterministic and randomized settings we see that quantum algorithms are quadratically faster.

We next consider searching of the zeros of functions from the Hölder class. We show for this problem upper and lower complexity bounds in the randomized and quantum settings. From the obtained results we get that randomized algorithms do not yield a speed-up over deterministic algorithms and quantum algorithms are quadratically faster.

At last, we consider solving boundary value problems for differential equations. We analyse two problems: linear (which defines a nonlinear solution operator) and nonlinear. For linear problem we find complexity bounds in the deterministic setting with the standard and the linear information, and the randomized and quantum settings. Nonlinear problem is analysed in the randomized and quantum settings. The complexity bounds in the deterministic setting are known for this problem. For both problems we show, that randomized algorithms yield a speed-up over the deterministic algorithms by 1/2 in the denominator of the exponent and quantum algorithms by 1.

Key words

computational complexity, randomized setting, quantum setting, optimal algorithm, nonlinear problems, function maximization, nonlinear equations, boundary value problems

Wstęp

W rozprawie zajmujemy się numerycznym przybliżaniem rozwiązań pewnych nieliniowych problemów. Będziemy poszukiwać optymalnych algorytmów rozwiązujących te problemy i badać ε–złożoność tych problemów, którą określa się jako minimalny koszt rozwiązania danego problemu z dokładnością co najwyżej ε w danej klasie elementów i w danym modelu obliczeniowym. Dział matematyki zajmujący się złożonością obliczeniową to Information-Based Complexity (IBC). Podstawy teorii IBC powstały w latach 80–tych XX-tego wieku, [TWW88]. Rozważa się tu różne modele obliczeniowe różniące się definicjami błędu oraz kosztu algorytmów, dostępną w obliczeniach informacją i dozwolonymi operacjami. Najlepiej zbadanym modelem obliczeniowym jest model deterministyczny. Wyznaczono tu ograniczenia na złożoność takich problemów jak na przykład całkowanie, maksymalizacja, aproksymacja funkcji, źle postawione problemy, równania różniczkowe (zob. np. [TWW88, Nov88, Kac87]) jak również wielu innych problemów liniowych i nieliniowych. W ostatnich latach wiele uwagi poświęca się złożoności w modelach randomizacyjnym i kwantowym. W modelach tych zbadane zostały takie problemy jak całkowanie, aproksymacja, czy też problemy początkowe dla równań różniczkowych (zob. np. [Nov88, Nov01, Hei04a, Hei04b, Kac06]). Dla większości problemów udało się pokazać przewagę algorytmów randomizacyjnych i kwantowych nad klasycznymi algorytmami deterministycznymi.

W rozprawie zajmujemy się badaniem złożoności pewnych problemów nieliniowych w modelach randomizacyjnym i kwantowym. Rozpatrujemy cztery problemy: maksymalizację funkcji, poszukiwanie zer funkcji oraz rozwiązywanie dwóch problemów brzegowych dla równań różniczkowych zwyczajnych. Pokażemy, że dla tych problemów zastosowanie algorytmów randomizacyjnych lub kwantowych daje przewagę nad algorytmami deterministycznymi.

W rozdziale 1 wprowadzimy modele obliczeniowe rozważane w tej pracy. Przedstawimy niezbędne pojęcia informacji, błędu, kosztu oraz złożoności w poszczególnych modelach. Ponieważ model kwantowy obliczeń jest mniej znany niż pozostałe modele, przedstawimy w podrozdziale 1.2 dozwolone operacje w tym modelu oraz zdefiniujemy podstawowe pojęcia, takie jak bit kwantowy, wyrocznia kwantowa oraz algorytm kwantowy.

W rozdziale 2 prezentujemy pewne znane wyniki potrzebne w dalszej części pracy. Przedstawione są tam ograniczenia na złożoność oraz algorytmy rozwiązujące pewne problemy dyskretne (przeszukiwanie baz danych, maksymalizacja dyskretnych ciągów, obliczanie średniej ciągu liczb) oraz problemy początkowe dla równań różniczkowych rzędu pierwszego i wyższych [Kac04, Kac06, GS08].

Dalsza część rozprawy zawiera wyniki własne. W rozdziale 3 badamy złożoność problemu poszukiwania maksimum funkcji w modelu kwantowym. Rozdział ten oparty jest na pracy [Goć06]. Rozważamy funkcje $f : [0,1]^d \to \mathbb{R}$ z klasy Höldera $F_d^{r,\rho}$ funkcji r-krotnie różniczkowalnych, których r-te pochodne cząstkowe spełniają warunek Höldera z wykładnikiem ρ. Konstruujemy algorytm kwantowy rozwiązujący ten problem, wykorzystujący optymalny algorytm maksymalizujący dyskretne ciągi i otrzymujemy za jego pomocą górne ograniczenia na złożoność w modelu kwantowym rzędu $\varepsilon^{-d/(2(r+\rho))}$. Wykazujemy następnie optymalność tego algorytmu, pokazując ograniczenia z dołu na złożoność tego samego rzędu. Wykorzystujemy w tym celu dolne ograniczenia na złożoność problemu maksymalizacji ciągu binarnego z pracy [NW99]. Porównując otrzymane wyniki z ograniczeniami na złożoność problemu maksymalizacji funkcji w modelach deterministycznym i randomizacyjnym otrzymujemy, że algorytmy kwantowe są kwadratowo szybsze dla tego problemu.
Główne wyniki rozdziału 3 to twierdzenie 14 pokazujące górne ograniczenia na błąd oraz koszt przedstawionego tutaj algorytmu kwantowego, wynikające z niego twierdzenie 15 prezentujące górne ograniczenie na złożoność w modelu kwantowym oraz twierdzenie 16 mówiące o dolnych ograniczeniach na złożoność w modelu kwantowym.

Rozdział 4 poświęcony jest problemowi poszukiwania zer funkcji z klasy Höldera $F_d^{r,\rho}$. Pokazujemy górne i dolne ograniczenia na złożoność tego problemu w modelach randomizacyjnym (rzędu $\varepsilon^{-d/(r+\rho)}$ - także w modelu deterministycznym) i kwantowym (rzędu $\varepsilon^{-d/(2(r+\rho))}$). Górne ograniczenia na złożoność uzyskiwane są przez skonstruowanie optymalnych algorytmów wykorzystujących optymalne algorytmy przeszukiwania dyskretnego. Ich konstrukcja podobna jest do konstrukcji algorytmu kwantowego z rozdziału 3 poszukującego maksimum funkcji. W celu udowodnienia ograniczeń z

dołu sprowadzamy problem przeszukiwania dyskretnego do problemu poszukiwania zer funkcji i wykorzystujemy znane ograniczenia z dołu na złożoność przeszukiwania dyskretnego. Z otrzymanych wyników wynika, że algorytmy randomizacyjne nie dają przewagi nad algorytmami deterministycznymi, natomiast algorytmy kwantowe są kwadratowo szybsze dla tego problemu.
Główne wyniki rozdziału 4 zawarte są w twierdzeniu 18 przedstawiającym górne ograniczenia na złożoność tego problemu we wszystkich modelach oraz w twierdzeniu 19 dotyczącym ograniczeń z dołu.

W rozdziale 5 badamy złożoność pewnych problemów brzegowych dla równań różniczkowych. Rozdział ten składa się z dwóch podrozdziałów. W podrozdziale 5.2 rozważamy liniowy dwupunktowy problem brzegowy dla równań różniczkowych rzędu k o rozdzielonych warunkach brzegowych (z punktu widzenia IBC jest to problem nieliniowy). Podrozdział ten oparty jest na pracy [GS10]. Zakładamy, że funkcje prawej strony należą do klasy F^r funkcji r-krotnie różniczkowalnych o ograniczonych pochodnych. Wyznaczamy ograniczenia na złożoność w modelach randomizacyjnym i kwantowym, a także dla porównania w modelu deterministycznym z informacją standardową i liniową. Okazuje się, że złożoność jest rzędu $\varepsilon^{-1/r}$ w modelu deterministycznym z informacją standardową. W modelu deterministycznym z informacją liniową dostajemy ograniczenia z góry rzędu $\varepsilon^{-1/(r+k-q)}$ (przy założeniu, że prawa strona równania zależy od pochodnych do rzędu q), a ograniczenia z dołu rzędu $\varepsilon^{-1/(r+k)}$ (różnią się więc one o q). W modelu randomizacyjnym otrzymujemy ograniczenia w zasadzie rzędu $\varepsilon^{-1/(r+1/2)}$, a w modelu kwantowym rzędu $\varepsilon^{-1/(r+1)}$ (ograniczenia z góry w modelach randomizacyjnym i kwantowym różnią się od dolnych ograniczeń o dowolnie mały parametr w mianowniku wykładnika - zob. tw. 23).
W podrozdziale 5.3 rozważamy nieliniowy problem brzegowy rzędu pierwszego. Zakładamy, że funkcja prawej strony $f : [a, b] \times \mathbb{R}^d \to \mathbb{R}^d$ należy do klasy F_d^r funkcji r-krotnie różniczkowalnych o ograniczonych pochodnych cząstkowych. Rozważamy również skalarny problem brzegowy rzędu drugiego z prawą stroną zależną jedynie od wartości rozwiązania oraz z warunkami brzegowymi danymi przez wartości na końcach przedziału. Poszukujemy ograniczeń na złożoność tych problemów w modelach randomizacyjnym i kwantowym. Ograniczenia z góry uzyskujemy poprzez skonstruowanie algorytmów opartych na metodzie strzałów z wykorzystaniem metody Newtona do rozwiązywania równań nieliniowych. Problemy początkowe rozwiązujemy przy wykorzystaniu prawie optymalnych algorytmów z pracy [Kac06] oraz ich modyfikacji dla funkcji kawałkami regularnych przedstawionych w rozdziale 2. Ograniczenia z dołu dla układu równań

rzędu pierwszego dostajemy przez sprowadzenie tego zagadnienia do oszacowań z dołu dla problemu początkowego. Dla problemu skalarnego rzędu drugiego ograniczenia z dołu uzyskujemy przez odpowiednie skorzystanie z oszacowań z dołu dla problemu obliczania całki z wagą. Otrzymujemy ograniczenia z dołu w modelu randomizacyjnym rzędu $\varepsilon^{-1/(r+1/2)}$, a w modelu kwantowym rzędu $\varepsilon^{-1/(r+1)}$. Ograniczenia z góry są prawie tego samego rzędu i różnią się od ograniczeń dolnych o dowolnie mały parametr w mianowniku wykładnika. Okazuje się, że algorytmy randomizacyjne dają przyspieszenie w stosunku do algorytmów deterministycznych o $1/2$ w mianowniku wykładnika, a algorytmy kwantowe o 1.
Głównymi wynikami rozdziału 5 dotyczącymi problemu liniowego są: twierdzenia 20 i 22 przedstawiające ograniczenia na błąd i koszt algorytmów w modelu deterministycznym oraz w modelach randomizacyjnym i kwantowym, twierdzenia 21 i 23 podające górne ograniczenia na złożoność oraz twierdzenie 24 przedstawiające dolne ograniczenia na złożoność we wszystkich modelach. Najważniejsze wyniki podrozdziału 5.3 zawarte są w twierdzeniach 27, 28, 29 przedstawiających górne i dolne ograniczenia na złożoność w modelu randomizacyjnym i kwantowym.

Rozdział 6 zawiera podsumowanie wyników zawartych w rozprawie.
Rozprawa zawiera również dwa dodatki. W dodatku A przedstawione są fakty pomocnicze dotyczące pewnych funkcji przydatnych przy uzyskiwaniu dolnych ograniczeń. Podaje on również pewne istotne twierdzenia używane w tej rozprawie. Dodatek B zawiera dowody lematów z rozdziału 5.

Rozdział 1

Wprowadzenie

1.1. Złożoność oparta na informacji

1.1.1. Wstęp

W rozprawie tej zajmujemy się przybliżonym rozwiązywaniem pewnych nieliniowych problemów obliczeniowych, przy wykorzystaniu niepełnej informacji o problemie. Celem jest zbadanie złożoności obliczeniowej tych problemów. Przedstawimy teraz za [TWW88] szkic teorii złożoności opartej na informacji (Information-Based Complexity – IBC), w zakresie potrzebnym w tej rozprawie.

Zdefiniujemy abstrakcyjny problem obliczeniowy oraz modele obliczeń. Niech F będzie pewnym zbiorem, a G przestrzenią liniową unormowaną. Problem obliczeniowy definiuje się za pomocą pewnego operatora $S : F \to G$ zwanego *operatorem rozwiązania*. Naszym celem jest znalezienie przybliżenia wartości $S(f)$ dla dowolnego $f \in F$. Niech $U(f) \in G$ będzie obliczonym przybliżeniem. Jakość przybliżenia $S(f)$ przez $U(f)$ można mierzyć korzystając z różnych kryteriów. Podstawowym kryterium rozważanym w tej pracy jest *błąd bezwzględny* czyli $\|S(f) - U(f)\|$.

Niech $\varepsilon > 0$. Mówimy, że $U(f)$ jest *ε–aproksymacją* $S(f)$ jeśli $\|S(f) - U(f)\| \leqslant \varepsilon$. W pracy tej rozważamy *przypadek najgorszy*, w którym szukamy elementów $U(f)$, które są ε–aproksymacjami dla każdego elementu $f \in F$.

Nie wszystkie problemy można zapisać za pomocą operatora rozwiązania S odwzorowującego zbiór F w przestrzeń unormowaną G. Nie da się w ten sposób zapisać na przykład rozważanego w tej pracy problemu szukania zer funkcji $f : \mathbb{R}^d \to \mathbb{R}^d$. Załóżmy, że chcemy znaleźć element x_ε taki, że $\|f(x_\varepsilon)\| \leqslant \varepsilon$. Tego problemu nie można zapisać używając powyższej definicji operatora rozwiązania. Można go jednakże sformułować za pomocą *uogólnionego operatora rozwiązania* W definiowanego następująco. Niech F i G będą pewnymi zbiorami, a W odwzorowaniem,

$$W : F \times \mathbb{R}_+ \to 2^G.$$

Załóżmy, że W ma następujące własności:

- $W(f, 0) \neq \emptyset, \qquad \forall f \in F,$
- jeśli $\varepsilon_1 \leqslant \varepsilon_2$, to $W(f, \varepsilon_1) \subset W(f, \varepsilon_2)$, $\forall \varepsilon_1, \varepsilon_2 \in \mathbb{R}_+,\ \forall f \in F$.

Przy tych dwóch założeniach możemy zdefiniować ε–aproksymację związaną z elementem f jako element $x_\varepsilon \in G$ taki, że

$$x_\varepsilon \in W(f, \varepsilon).$$

Tak zdefiniowane odwzorowanie W jest uogólnieniem operatora rozwiązania S. Dla problemów zdefiniowanych przez operator rozwiązania można bowiem zdefiniować

$$W(f, \varepsilon) = \{g \in G : \|S(f) - g\| \leqslant \varepsilon\}.$$

Dla problemu szukania zer funkcji, uogólniony operator rozwiązania można zdefiniować następująco

$$W(f, \varepsilon) = \{x \in \mathbb{R}^d : \|f(x)\| \leqslant \varepsilon\}.$$

1.1.2. Informacja i algorytm

W celu obliczenia ε–aproksymacji musimy dysponować pewną informacją o elemencie wejściowym f. Zakładamy, że posiadamy pewną informację częściową otrzymywaną przez obliczenie wartości $L(f)$ pewnego operatora $L : F \to \mathcal{I}$ dla pewnej przestrzeni $\mathcal{I}$. Operator L zwany jest *operatorem informacji*. Zwykle dopuszczamy tylko operatory informacji L z pewnej klasy Λ, np. klasy funkcjonałów liniowych lub klasy operatorów

zdefiniowanych przez wartości funkcji. Dla każdego elementu f możemy obliczyć pewną liczbę wartości $L(f)$ dla $L \in \Lambda$. Rozważać można dwa rodzaje informacji: *nieadaptacyjną* i *adaptacyjną*. Informacja N nazywana jest *nieadaptacyjną* wtedy, gdy istnieją $L_1, L_2, \ldots, L_n \in \Lambda$ takie, że

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

Liczba n operatorów informacji nazywana jest *licznością* informacji N.
Bardziej ogólnym rodzajem informacji jest informacja *adaptacyjna*. W informacji adaptacyjnej zakłada się, że zarówno operatory informacji jak i ich liczba mogą zależeć od f. Informacja nazywa się *adaptacyjną* wtedy gdy

$$N(f) = [L_1(f), L_2(f; y_1), \ldots, L_{n(f)}(f; y_1, y_2 \ldots, y_{n(f)-1})],$$

gdzie $y_1 = L_1(f)$, $y_i = L_i(f; y_1, y_2, \ldots, y_{i-1})$ dla $i = 2, 3, \ldots, n(f)$. Jak widać, wybór operatora L_i zależy od poprzednio obliczonych wartości $y_1, y_2 \ldots, y_{i-1}$. Zakładamy przy tym, że $L_i(\,\cdot\,; y_1, \ldots, y_{i-1}) \in \Lambda$ dla każdych ustalonych $y_1, \ldots, y_{i-1}$.
Liczba $n(f)$ oznacza całkowitą liczbę operatorów informacji o elemencie f i nazywana jest *licznością* informacji N o f. Nie jest ona taka sama dla wszystkich f z F, lecz jest wyznaczana w trakcie obliczeń na podstawie aktualnej wiedzy o f. Może ona być wyznaczona w następujący sposób. Po obliczeniu wartości każdego operatora informacji, na podstawie wcześniej wyznaczonych wartości $y_1 = L_1(f), y_2 = L_2(f; y_1), \ldots, y_i = L_i(f; y_1, \ldots, y_{i-1})$ podejmowana jest decyzja czy zostanie obliczony kolejny, $(i+1)$-szy operator informacji. Decyzja ta podejmowana jest na podstawie wartości pewnej binarnej funkcji $ter_i : \mathcal{I}^i \to \{0, 1\}$. Jeśli na i-tym kroku $ter_i(y_1, y_2, \ldots, y_i) = 0$, to obliczana jest kolejna wartość operatora informacji $L_{i+1}(f; y_1, \ldots, y_i)$ i cały proces jest powtarzany. Jeśli natomiast $ter_i(y_1, y_2, \ldots, y_i) = 1$, to $n(f)$ przyjmuje wartość i, a $y_1, \ldots, y_i$ stanowi ostateczną informację o f, $N(f) = [y_1, \ldots, y_i]$. Tak więc, liczność $n(f)$ jest równa

$$n(f) = \min\{i : ter_i(y_1, y_2, \ldots, y_i) = 1\},$$

przy konwencji, że $\min \emptyset = +\infty$.
Adaptacyjna informacja, w przeciwieństwie do informacji nieadaptacyjnej, wymaga obliczania sekwencyjnego. Musimy czekać na obliczenie y_i zanim obliczymy kolejną wartość informacji y_{i+1}. Informacja nieadaptacyjna może natomiast być obliczana równolegle na wielu procesorach.

Mając obliczoną informację $y = N(f)$ można obliczyć przybliżenie $U(f) \in G$ rozwiązania $S(f)$. Dokonujemy tego za pomocą pewnego odwzorowania

$$\phi : N(F) \to G$$

zwanego *algorytmem*. Poszukiwane przybliżenie jest definiowane jako $U(f) = \phi(N(f))$. Tak więc, algorytm wykorzystuje znaną informację produkując pewną aproksymację elementu $S(f)$. Jest to ogólna idea abstrakcyjnego algorytmu. Czasem w teorii lub praktyce wymaga się, aby algorytm spełniał jakieś dodatkowe warunki, np. liniowość ze względu na wykorzystywaną informację.

1.1.3. Modele obliczeń

W rozprawie tej rozważać będziemy różne modele obliczeniowe, różniące się zarówno dozwolonymi operacjami (operacje arytmetyczne, porównania, losowanie), informacją, jak i definicjami błędu i kosztu. Rozważać będziemy głównie modele randomizacyjny i kwantowy. Będziemy porównywać te modele z modelem deterministycznym w różnych klasach informacji. Model deterministyczny jest najlepiej zbadanym modelem. Rozważać będziemy pewne problemy zadane przez nieliniowy operator rozwiązania $S : F \to G$ lub uogólniony operator rozwiązania $W : F \times \mathbb{R}_+ \to 2^G$, gdzie F będzie pewną klasą funkcji, a G przestrzenią unormowaną.
Zdefiniujemy teraz rozważane przez nas modele: randomizacyjny i kwantowy, jak również model deterministyczny, z którym będziemy porównywać otrzymane wyniki.

W klasycznym modelu deterministycznym do obliczeń dysponujemy tylko operacjami arytmetycznymi oraz porównywaniem liczb. Zakłada się, że wszystkie operacje arytmetyczne wykonywane są dokładnie. Rozważa się dwa typy informacji: informację *standardową* oraz informację *liniową*. Informacja standardowa składa się z wartości funkcji f oraz jej pochodnych w pewnych ustalonych punktach. Informacja liniowa składa się natomiast z dowolnych funkcjonałów liniowych ze względu na element wejściowy f. Tak więc, klasa informacji standardowych jest podklasą klasy informacji liniowych. Niech otrzymanym przybliżeniem w wyniku działania algorytmu ϕ dla elementu wejściowego f będzie $U(f)$.

W modelu randomizacyjnym oprócz operacji arytmetycznych i porównań liczb dopuszczona jest również operacja losowania. Informacja w tym modelu jest podobna do informacji standardowej w modelu deterministycznym. Składa się ona z wartości funkcji f oraz jej pochodnych w pewnych punktach, które mogą być wybierane losowo.

Wynikiem algorytmu ϕ dla elementu f w tym przypadku jest zmienna losowa $U^\omega(f)$ na pewnej przestrzeni probabilistycznej $(\Omega, \Sigma, \mathbf{P})$, $\omega \in \Omega$.

W modelu kwantowym dostępnymi operacjami są wyłącznie operacje unitarne na przestrzeni Hilberta $\underbrace{\mathbb{C}^2 \otimes \ldots \otimes \mathbb{C}^2}_{m \text{ razy}}$. Informacja o problemie natomiast otrzymywana jest przez zastosowanie pewnej operacji unitarnej Q_f zwanej *wyrocznią kwantową*. Dokładny opis modelu kwantowego znajduje się w następnym podrozdziale. Wynikiem algorytmu ϕ, podobnie jak w przypadku randomizacyjnym, jest zmienna losowa $U^\omega(f)$ na pewnej przestrzeni probabilistycznej $(\Omega, \Sigma, \mathbf{P})$.

Każdy algorytm obarczony jest pewnym kosztem. W tej pracy rozważać będziemy wyłącznie koszt informacyjny definiowany jako liczba obliczeń informacji użytej przez algorytm w przypadku deterministycznym i randomizacyjnym oraz jako liczbę odwołań do wyroczni kwantowej w przypadku kwantowym. Koszt algorytmu ϕ dla elementu wejściowego f (przy danej informacji) oznaczać będziemy przez $\mathrm{cost}(S, \phi, f)$ z odpowiednim indeksem górnym ('det', 'rand', 'quant', 'det $-$ lin', 'det $-$ st') oznaczającym model obliczeń (oraz ewentualnie dopuszczalną klasę informacji). Koszt globalny w klasie funkcji F definiujemy jako

$$\mathrm{cost}(S, \phi, F) = \sup_{f \in F} \mathrm{cost}(S, \phi, f)$$

z odpowiednim indeksem górnym zależnym od modelu i klasy informacji.

Błąd algorytmu ϕ na elemencie f w przypadku deterministycznym (przy danej informacji) definiujemy jako odległość pomiędzy rozwiązaniem dokładnym i obliczonym przybliżeniem

$$\mathrm{e}^{\mathrm{det}}(S, \phi, f) = \|S(f) - U(f)\|_G, \tag{1.1}$$

natomiast błąd w modelu randomizacyjnym definiowany jest jako oczekiwana odległość pomiędzy rozwiązaniem dokładnym i przybliżonym

$$\mathrm{e}^{\mathrm{rand}}(S, \phi, f) = \left(\int_\Omega \|S(f) - U^\omega(f)\|_G^2 d\mathbf{P}(\omega) \right)^{1/2}. \tag{1.2}$$

W modelu kwantowym używamy błędu probabilistycznego z prawdopodobieństwem porażki ograniczonym przez $\delta \in (0, 1/4]$ definiowanego następująco

$$\mathrm{e}^{\mathrm{quant}}(S, \phi, f, \delta) = \inf \left\{ \alpha : \mathbf{P} \left(\|S(f) - U^\omega(f)\|_G > \alpha \right) \leqslant \delta \right\}. \tag{1.3}$$

Często przyjmuje się prawdopodobieństwo porażki $\delta = 1/4$. Błąd z takim ograniczeniem prawdopodobieństwa porażki oznaczać będziemy przez

$$\mathrm{e}^{\mathrm{quant}}(S, \phi, f) = \mathrm{e}^{\mathrm{quant}}(S, \phi, f, 1/4). \tag{1.4}$$

W rozprawie tej zajmujemy się przypadkiem najgorszym, więc we wszystkich modelach błąd globalny w klasie funkcji F definiowany jest jako

$$\mathrm{e}^{\#}(S,\phi,F) = \sup_{f\in F} \mathrm{e}^{\#}(S,\phi,f), \quad \# \in \{\mathrm{det},\mathrm{rand},\mathrm{quant}\}. \tag{1.5}$$

W oznaczeniach dla krótkości nie zaznaczamy zależności błędu i kosztu od informacji N.

Ważnym pojęciem w teorii złożoności obliczeniowej jest pojęcie złożoności problemu. Dla ustalonej dokładności $\varepsilon > 0$, ε–złożonością nazywa się minimalny koszt obliczenia przybliżenia, którego błąd nie przekracza ε

$$\mathrm{comp}^{\#}(\varepsilon,S,F) = \inf\left\{\mathrm{cost}^{\#}(S,\phi,F) :\ \phi, N \text{ takie, że } \mathrm{e}^{\#}(S,\phi,F) \leqslant \varepsilon\right\}, \quad \# \in \{\mathrm{det},\mathrm{rand},\mathrm{quant}\}. \tag{1.6}$$

W modelu kwantowym, dla podkreślenia zależności od prawdopodobieństwa porażki δ, będziemy używać oznaczeń $\mathrm{cost}^{\mathrm{quant}}(S,\phi,F,\delta)$ i $\mathrm{comp}^{\mathrm{quant}}(\varepsilon,S,F,\delta)$ oraz będziemy oznaczać

$$\mathrm{comp}^{\mathrm{quant}}(\varepsilon,S,F) = \mathrm{comp}^{\mathrm{quant}}(\varepsilon,S,F,1/4).$$

W naszej pracy zbadamy złożoność kilku problemów nieliniowych. Będziemy chcieli wyznaczyć dolne i górne ograniczenia na złożoność w przypadku randomizacyjnym i kwantowym. Będziemy również chcieli znaleźć n-ty optymalny algorytm rozwiązujący dany problem, to znaczy, dla ustalonego $n \in \mathbb{N}$, algorytm ϕ^* o koszcie nie przekraczającym n, który spełnia

$$\mathrm{e}(S,\phi^*,F) = \Theta\left(\inf\left\{\mathrm{e}(S,\phi,F) :\ \phi, N \text{ takie, że } \mathrm{cost}(S,\phi,F) \leqslant n\right\}\right).$$

1.2. Model obliczeń kwantowych

1.2.1. Wstęp

Idea komputerów kwantowych pochodzi z początku lat osiemdziesiątych dwudziestego wieku. Pierwsze idee przedstawione były przez Manina [Man80] oraz Feynmana [Fey82]. W roku 1985, Deutsch [Deu85] wprowadził podstawy teoretyczne obliczeń kwantowych opierające się na kwantowej maszynie Turinga.
Przełom w obliczeniach kwantowych nastąpił w 1993 za sprawą pracy Shora [Sho94]. Przedstawia ona efektywny algorytm faktoryzacji liczb całkowitych na komputerze kwantowym. Rozkłada on liczbę całkowitą N na czynniki pierwsze w czasie

$O((\log N)^2(\log\log N)(\log\log\log N))$. Wynik ten jest o tyle ważny, że nie są znane żadne wielomianowe algorytmy klasyczne rozwiązujące ten problem. Kolejną fundamentalną pracą był artykuł Grovera [Gro96] przedstawiający kwantowy algorytm przeszukiwania baz danych. Algorytm ten pozwala znaleźć element N-elementowej bazy danych w czasie $O(\sqrt{N})$, kiedy najlepszy algorytm klasyczny potrafi rozwiązać to zadanie w czasie $O(N)$ (por. też [Hir04, GK03]). Na algorytmie tym opiera się wiele innych algorytmów kwantowych rozwiązujących zarówno problemy dyskretne jak i ciągłe.

Od tamtej pory przebadanych zostało wiele problemów na komputerze kwantowym. Początkowo zajmowano się problemami dyskretnymi. Pokazano, że komputery kwantowe znacząco przyspieszają obliczenia dla wielu problemów. Znalezione zostały algorytmy kwantowe sumujące ciągi liczb, obliczające dyskretną średnią, medianę oraz k-ty najmniejszy element (zob. np. [BHT98, DH98, Gro98, NW99]).

Zaczęto również poszukiwać algorytmów kwantowych dla problemów ciągłych, począwszy od pracy Abramsa i Williamsa [AW99], którzy zaproponowali pewne idee kwantowego algorytmu całkowania. Problem ten badał również Novak [Nov01], który wyznaczył dolne i górne ograniczenia na złożoność problemu całkowania funkcji z klasy Höldera na komputerze kwantowym. W pracach [Hei04a, Hei04b] zbadany jest problem aproksymacji funkcji w przestrzeniach L_p oraz w przestrzeni Soboleva. Dla pewnych parametrów klas funkcji pokazano przewagę komputerów kwantowych nad klasycznymi. Zbadano również pewne problemy nieliniowe. W pracy [Kac06], Kacewicz analizował rozwiązywanie problemów początkowych dla układów równań różniczkowych pierwszego rzędu. Znalazł on prawie optymalne algorytmy randomizacyjny i kwantowy rozwiązujące ten problem. Analogiczne algorytmy rozwiązujące równania różniczkowe wyższych rzędów przeanalizowane zostały w pracy [GS08].

W następnym podrozdziale przedstawimy model obliczeń kwantowych. Prezentacja ta opiera się na [Hei02, Hei03a, Hei04a, Hei04b, NC00].

1.2.2. Model obliczeniowy

Na komputerze klasycznym, podstawową informacją jest *bit*. Może on przyjmować dwie wartości, które oznaczymy przez 0 i 1. Odpowiednikiem bitu na komputerze kwantowym jest *bit kwantowy* lub inaczej *qubit*. Qubit reprezentowany jest przez sferę jednostkową w zespolonej przestrzeni Hilberta $H_1 := \mathbb{C}^2$. Niech e_0, e_1 będą wektorami bazowymi tej przestrzeni. Z godnie z notacją Diraca dla mechaniki kwantowej wektor

e_0 zapisujemy jako $|0\rangle$, a e_1 jako $|1\rangle$. Są to tak zwane *stany klasyczne*. Sfera jednostkowa w przestrzeni H_1 stanowi zbiór możliwych stanów qubitów. Stanem qubitu jest więc dowolny wektor ψ z przestrzeni H_1 taki, że $\|\psi\| = 1$. Oprócz stanów klasycznych $|0\rangle$ i $|1\rangle$ może on przyjmować również wartość ich kombinacji liniowych:

$$\psi = \alpha_0|0\rangle + \alpha_1|1\rangle \qquad \alpha_0, \alpha_1 \in \mathbb{C},\ |\alpha_0|^2 + |\alpha_1|^2 = 1.$$

Stanu qubitu nie można odczytać bezpośrednio, to znaczy nie jesteśmy w stanie odczytać wartości parametrów α i β. Można natomiast dokonać pomiaru qubitu. W wyniku pomiaru stanu kwantowego otrzymamy jeden ze stanów klasycznych $|0\rangle$ lub $|1\rangle$ z prawdopodobieństwem odpowiednio $|\alpha_0|^2$ lub $|\alpha_1|^2$.

Układ m wzajemnie oddziałujących na siebie qubitów reprezentowany jest przez iloczyn tensorowy

$$H_m = \underbrace{H_1 \otimes H_1 \otimes \ldots \otimes H_1}_{m},$$

który stanowi 2^m-wymiarową zespoloną przestrzeń Hilberta. Bazę kanoniczną tej przestrzeni tworzą wektory

$$e_{i_0} \otimes e_{i_1} \otimes \ldots \otimes e_{i_{m-1}}, \qquad (i_0, i_1, \ldots, i_{m-1}) \in \{0,1\}^m.$$

W notacji Diraca będziemy zapisywać wektor $e_{i_0} \otimes e_{i_1} \otimes \ldots \otimes e_{i_{m-1}}$ jako $|i_0\rangle|i_1\rangle \ldots |i_{m-1}\rangle$, lub $|i_0 i_1 \ldots i_{m-1}\rangle$ lub $|i\rangle$, gdzie $i := (i_0 i_1 \ldots i_{m-1})_2 := \sum_{k=0}^{m-1} i_k 2^{m-1-k}$. Wektory bazowe $|i\rangle$ $(i = 0, 1, \ldots, 2^m - 1)$ nazywane są stanami klasycznymi. Dowolny stan kwantowy ma postać

$$|\psi\rangle = \sum_{i=0}^{2^m-1} \alpha_i |i\rangle, \qquad \sum_{i=0}^{2^m-1} |\alpha_i|^2 = 1.$$

W wyniku pomiaru stanu $|\psi\rangle$ otrzymujemy jeden ze stanów klasycznych $|i\rangle$ z prawdopodobieństwem $|\alpha_i|^2$ $(i = 0, 1, \ldots 2^m - 1)$.

Z równania Schrödingera wynika, że wszystkie przekształcenia systemu kwantowego muszą być reprezentowane przez operacje unitarne na H_m. Obliczenia kwantowe zakładają, że potrafimy wykonywać pewną liczbę elementarnych operacji unitarnych (bramek kwantowych) na systemie kwantowym. Przedstawimy najpierw bramki jedno-qubitowe. Bramki takie działają tylko na jednym czynniku iloczynu tensorowego $H_m = H_1 \otimes H_1 \otimes \ldots \otimes H_1$. Bramka Walsha-Hadamarda $W : H_1 \to H_1$ definiowana jest następująco (przez wartości na wektorach bazowych)

$$W|0\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \quad W|1\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}.$$

Działanie tej bramki na j-tym czynniku przestrzeni H_m dane jest przez operator unitarny

$$W_m^{(j)} = Id \otimes \ldots Id \otimes \underbrace{W}_{j} \otimes Id \otimes \ldots \otimes Id,$$

gdzie Id oznacza identyczność na H_1.
Kolejną bramką jedno-qubitową jest operacja przesunięcia fazy. Dla parametru $0 \leqslant \theta < 2\pi$ bramka $P_\theta : H_1 \to H_1$ zdefiniowana jest następująco

$$P_\theta|0\rangle = |0\rangle, \quad P_\theta|1\rangle = e^{i\theta}|1\rangle.$$

Operator $P_{\theta,m}^{(j)}$ definiowany jest podobnie jak $W_m^{(j)}$.
Rozważymy teraz bramki dwu-qubitowe, to znacz działające na dowolnych dwóch komponentach $H_1 \otimes H_1 \otimes \ldots \otimes H_1$. Bramka kwantowego XOR (exclusive OR) lub CNOT (controlled NOT), $X : H_1 \otimes H_1 \to H_1 \otimes H_1$, zdefiniowana jest następująco

$$\begin{aligned}
X|0\rangle|0\rangle &= |0\rangle|0\rangle,\\
X|0\rangle|1\rangle &= |0\rangle|1\rangle,\\
X|1\rangle|0\rangle &= |1\rangle|1\rangle,\\
X|1\rangle|1\rangle &= |1\rangle|0\rangle.
\end{aligned}$$

Tak więc, kiedy pierwszy bit jest równy zero, nic się nie dzieje, kiedy zaś jest równy jeden, drugi bit jest zaprzeczony. Niech $X_m^{(k,l)} : H_m \to H_m$ będzie operatorem unitarnym danym przez zastosowanie X do k-tego i l-tego składnika H_m. Zdefiniowany on jest następująco

$$X_m^{(k,l)}|i_0\rangle \ldots |i_{k-1}\rangle \ldots |i_{l-1}\rangle \ldots |i_{m-1}\rangle = |i_0\rangle \ldots |y\rangle \ldots |z\rangle \ldots |i_{m-1}\rangle,$$

gdzie

$$|y\rangle|z\rangle := X|i_{k-1}\rangle|i_{l-1}\rangle.$$

Zdefiniujmy teraz

$$\mathcal{G}_m = \{W_m^{(j)}, P_{\theta,m}^{(j)}, X_m^{(k,l)} : 1 \leqslant j, k \neq l \leqslant m, 0 \leqslant \theta < 2\pi\}.$$

Można pokazać, że zbiór ten stanowi uniwersalny system bramek kwantowych (zob. np. [NC00]). Zachodzi mianowicie

Twierdzenie 1 ([NC00]). *Dowolną operację unitarną na H_m można przedstawić jako złożenie skończonej liczby elementów z $\mathcal{G}_m$ (z dokładnością do skalarnego czynnika zespolonego o module 1).*

Zbiór $\mathcal{G}_m$ jest zbiorem o nieskończonej liczbie elementów, ponieważ parametr θ w operacji przesunięcia fazy może przyjmować nieskończoną liczbę wartości. Stanowi to trudność, ponieważ chcąc korzystać z takiego zestawu bramek trzeba by było do każdego zadania konstruować oddzielny komputer kwantowy, mający zaimplementowane odpowiednie bramki. Rozważmy skończony podzbiór zbioru $\mathcal{G}_m$:

$$\mathcal{G}_m^0 = \{W_m^{(j)}, P_{\pi/4,m}^{(j)}, X_m^{(k,l)} : 1 \leqslant j, k \neq l \leqslant m\}.$$

Można pokazać, że zbiór ten stanowi w przybliżeniu uniwersalny układ bramek kwantowych (zob. np. [NC00]). Zachodzi bowiem twierdzenie, że dowolną operację unitarną na H_m można przybliżyć z dowolną dokładnością w normie operatorowej przez skończoną liczbę elementów z $\mathcal{G}_m^0$ (z dokładnością do skalarnego czynnika zespolonego o module 1).

Twierdzenie 2 ([NC00]). *Niech $\varepsilon > 0$. Niech U będzie dowolną operacją unitarną na H_m. Istnieją takie operacje $V_1, V_2, \ldots, V_k \in \mathcal{G}_m^0$ oraz $\alpha \in \mathbb{R}$, że*

$$\max_{|\psi\rangle} \|(U - e^{i\alpha} V_k \ldots V_2 V_1)|\psi\rangle\| \leqslant \varepsilon.$$

Tak więc, po zaimplementowaniu bramek ze zbioru $\mathcal{G}_m^0$, możemy wykonywać w przybliżeniu dowolne operacje unitarne.

1.2.3. Algorytm kwantowy

Mając zdefiniowane dozwolone operacje kwantowe możemy przedstawić działanie algorytmu kwantowego. Załóżmy, że mamy pewien problem obliczeniowy dany przez operator rozwiązania $S : F \to G$, gdzie F jest pewny niepustym zbiorem funkcji idących z pewnego zbioru D o wartościach w zbiorze K, a G przestrzenią unormowaną. Naszym celem jest znalezienie przybliżenia wartości $S(f)$ dla dowolnego $f \in F$. Wprowadźmy jeszcze pewne oznaczenia. Niech $\mathbb{Z}[0, N) := \{0, 1, \ldots, N-1\}$ dla $N \in \mathbb{N}$. Niech $\mathcal{C}_m = \{|i\rangle : i \in \mathbb{Z}[0, 2^m)$ będzie zbiorem stanów klasycznych na H_m, a $\mathcal{U}(H_m)$ zbiorem wszystkich operacji unitarnych na H_m.

Zdefiniujmy najpierw wyrocznię kwantową za [Hei02]. Wyrocznią kwantową na F nazywamy szóstkę

$$Q = (m, m', m'', Z, \tau, \beta),$$

gdzie $m, m', m'' \in \mathbb{N}$, $m' + m'' \leqslant m$, $Z \subseteq \mathbb{Z}[0, 2^{m'})$ jest niepustym podzbiorem oraz

$$\tau : Z \to D$$

$$\beta : K \to \mathbb{Z}[0, 2^{m''})$$

są pewnymi odwzorowaniami. Później wyjaśnimy znaczenie poszczególnych składników. Taka szóstka definiuje odwzorowanie wyroczni (oznaczone również przez Q)

$$Q : F \ni f \mapsto Q_f \in \mathcal{U}(H_m).$$

Operację unitarną Q_f można jednoznacznie zdefiniować poprzez zadanie wartości na dowolnym wektorze bazowym $h \in \mathcal{C}_m$. Załóżmy, że $h = |i\rangle|x\rangle|y\rangle$, gdzie $|i\rangle \in \mathcal{C}_{m'}, |x\rangle \in \mathcal{C}_{m''}, |y\rangle \in \mathcal{C}_{m-m'-m''}$. Wtedy operację Q_f definiujemy na wektorze h w następujący sposób:

$$Q_f|i\rangle|x\rangle|y\rangle = \begin{cases} |i\rangle|x \oplus \beta(f(\tau(i)))\rangle|y\rangle & \text{jeśli } i \in Z \\ |i\rangle|x\rangle|y\rangle & \text{w przeciwnym wypadku,} \end{cases} \tag{1.7}$$

gdzie $\oplus$ oznacza w tym wypadku dodawanie modulo $2^{m''}$.

Taka definicja wyroczni obejmuje również przypadek dyskretny. Załóżmy, że $D = \mathbb{Z}[0, m')$, a $K = \{0, 1\}$. Wtedy wyrocznię binarną dla funkcji $f : D \to K$ można zdefiniować na wektorze $h = |i\rangle|x\rangle|y\rangle$ jako $Q_f|i\rangle|x\rangle|y\rangle = |i\rangle|x \oplus f(i)\rangle|y\rangle$. Tak zdefiniowana wyrocznia wykorzystywana jest w wielu algorytmach kwantowych dla problemów dyskretnych na przykład w algorytmie Grovera przeszukiwania baz danych (zob. [Gro96]).

Funkcja β służy do kodowania wartości funkcji f w postaci ciągów binarnych. Dzięki temu można przypisać wartości funkcji f odpowiedni stan rejestru kwantowego. Funkcja $\tau : i \to \tau(i) \in D$ opisuje natomiast związek pomiędzy ciągiem binarnym, a odpowiednim elementem z dziedziny funkcji f. Jest to więc funkcja odpowiadająca za dekodowanie elementów dziedziny. Obie te funkcje zadawane są przez projektanta algorytmu kwantowego.

W rozprawie tej potrzebować będziemy nie tylko wyroczni o wartościach funkcji f, ale również o jej pochodnych cząstkowych rzędu od 0 do r dla $r \geqslant 0$ dla funkcji z klasy $C^r(\bar{D})$, gdzie $D \subset \mathbb{R}^d$ jest niepustym zbiorem otwartym. Wyrocznię taką można skonstruować w następujący sposób. Zdefiniujmy funkcję $g : \tilde{D} \times D \to K$, gdzie $\tilde{D} = \{(i_1, i_2, \ldots, i_d) : i_j \geqslant 0, j = 1, 2, \ldots, d,\ i_1 + i_2 + \ldots i_d \leqslant r\}$ w następujący sposób

$$g(i_1, i_2, \ldots, i_d, x) = \frac{\partial^{i_1+i_2+\ldots+i_d} f}{\partial x_1^{i_1} \partial x_2^{i_2} \ldots \partial x_d^{i_d}}(x). \tag{1.8}$$

Wyrocznię $\tilde{Q}$ o wartościach i pochodnych funkcji f zdefiniujemy jako wyrocznię o wartościach funkcji g, czyli

$$\tilde{Q} : F \ni f \mapsto \tilde{Q}_f = Q_g.$$

Przedstawimy teraz działanie algorytmu z k pomiarami. Algorytm taki działa następująco:

- Tworzymy stan początkowy, czyli pewien stan klasyczny b_0.
- Na stanie b_0 wykonujemy pewną liczbę operacji unitarnych, w tym operację wyroczni kwantowej.
- Mierzymy otrzymany stan kwantowy. W wyniku pomiaru otrzymujemy pewien stan klasyczny ξ_0. Stan ten jest zapamiętywany.
- Na podstawie ξ_0 tworzymy nowy stan początkowy b_1.
- Wykonujemy kolejną serię operacji unitarnych (być może z inną wyrocznią i liczbą qubitów).
- Kolejny raz następuje pomiar, w wyniku czego dostajemy pewien stan klasyczny ξ_1, który jest zapamiętywany.
- Na podstawie ξ_0 i ξ_1 ustalany jest kolejny stan początkowy b_2.
- Po k takich cyklach otrzymujemy ciąg $\xi_0, \xi_1, \ldots, \xi_{k-1}$.
- Wynikiem algorytmu jest wartość pewnej funkcji φ na ciągu $\xi_0, \xi_1, \ldots, \xi_{k-1}$.

Przedstawimy teraz formalną definicję za [Hei02]. Algorytmem kwantowym na F bez pomiarów nazywamy parę

$$\phi = (Q, (U_j)_{j=0}^n),$$

gdzie Q jest wyrocznią kwantową na F, $n \in \mathbb{N} \cup \{0\}$, a $U_j \in \mathcal{U}(H_m)$ $(j = 0, 1, \ldots, n)$, gdzie $m = m(Q)$ jest liczbą qubitów wyroczni kwantowej. Mając dane ϕ i $f \in F$, definiujemy operacją unitarną $\phi_f \in \mathcal{U}(H_m)$ następująco

$$\phi_f = U_n Q_f U_{n-1} \ldots U_1 Q_f U_0. \tag{1.9}$$

Oznaczamy przez $n_q(\phi) := n$ liczbę zapytań kwantowych oraz przez $m(\phi) := m = m(Q)$ liczbę qubitów używanych przez ϕ.

Algorytmem kwantowym na F z wartościami w G o k pomiarach nazywamy trójkę

$$\phi = ((\phi_l)_{l=0}^{k-1}, (b_l)_{l=0}^{k-1}, \varphi),$$

gdzie $k \in \mathbb{N}$, a ϕ_l $(l = 0, 1, \ldots, k-1)$ są algorytmami kwantowymi na F bez pomiarów. Ponadto

$$b_0 \in \mathbb{Z}[0, 2^{m(\phi_0)}),$$

a dla $1 \leqslant l \leqslant k-1$, b_l jest funkcją

$$b_l : \prod_{i=0}^{l-1} \mathbb{Z}[0, 2^{m(\phi_i)}) \to \mathbb{Z}[0, 2^{m(\phi_l)}).$$

Wreszcie φ jest funkcją

$$\varphi : \prod_{l=0}^{k-1} \mathbb{Z}[0, 2^{m(\phi_l)}) \to G.$$

Dla danej funkcji $f \in F$ wynikiem algorytmu ϕ jest $\phi(f) = \varphi(\xi_0, \xi_1, \ldots, \xi_{k-1})$, gdzie ξ_0 jest wynikiem pomiaru stanu $\phi_{0,f}|b_0\rangle$, a ξ_i $(i = 1, 2, \ldots, k-1)$ jest wynikiem pomiaru stanu $\phi_{i,f}|b_i(\xi_0, \xi_1, \ldots, \xi_{i-1})\rangle$.

Liczba $n_q(\phi) := \sum_{l=0}^{k-1} n_q(\phi_l)$ jest liczbą zapytań kwantowych użytych przez algorytm ϕ. Koszt algorytmu ϕ z F do G definiujemy jako ilość zapytań kwantowych tego algorytmu, czyli

$$\mathrm{cost}^{\mathrm{quant}}(S, \phi, F) = n_q(\phi).$$

Można pokazać następujący lemat (zob. [Hei02]).

Lemat 1 ([Hei02]). *Dla dowolnego algorytmu ϕ z F do G o k pomiarach istnieje algorytm $\tilde{\phi}$ z F do G o jednym pomiarze taki, że $n_q(\tilde{\phi}) = 2n_q(\phi)$ oraz*

$$\tilde{\phi}(f) = \phi(f)$$

dla każdego $f \in F$.

Zgodnie z tym lematem, bez straty ogólności możemy traktować algorytm kwantowy jako algorytm z jednym pomiarem, kosztem dwukrotnego zwiększenia liczby zapytań kwantowych.

Przypomnijmy, że błąd algorytmu ϕ na elemencie wejściowym $f \in F$ definiujemy przez (1.3).

Niech $G = \mathbb{R}$. Można pokazać, że prawdopodobieństwo porażki δ można zmniejszać wykładniczo poprzez M-krotne powtórzenie algorytm ϕ i wzięcie mediany wyników. Niech $\mu : \mathbb{R}^M \to \mathbb{R}$ będzie odwzorowaniem mediany, czyli $\mu(a_0, a_1, \ldots, a_{M-1})$ jest wartością $\lceil (M+1)/2 \rceil$-tego elementu niemalejącej permutacji ciągu (a_i). Dla dowolnego algorytmu ϕ z F do $\mathbb{R}$ oznaczmy przez $\mu(\phi^M) := \mu(\phi, \ldots, \phi)$. Prawdziwy jest następujący lemat (zob. [Hei02])

Lemat 2 ([Hei02]). *Niech S będzie operatorem rozwiązania z F do $\mathbb{R}$, a ϕ algorytmem kwantowym rozwiązującym problem zdefiniowany przez S. Wtedy dla dowolnego $f \in F$,*

$$\mathrm{e}^{\mathrm{quant}}(S, \mu(\phi^M), f, e^{-M/8}) \leqslant \mathrm{e}^{\mathrm{quant}}(S, \phi, f, 1/4).$$

Dzięki temu lematowi można zwiększać prawdopodobieństwo sukcesu algorytmu kwantowego, logarytmicznie zwiększając jego koszt. Lemat ten odnosi się jedynie do przypadku gdy wynikiem algorytmu jest liczba rzeczywista. Podobnie można postępować, gdy G jest przestrzenią $l_\infty(T)$ ograniczonych funkcji rzeczywistych na pewnym zbiorze T z normą supremum. W tym przypadku powtarzamy algorytm M-krotnie i bierzemy medianę w każdym punkcie. Niech $g_0, g_1, \ldots, g_{M-1}$ będą funkcjami z przestrzeni G. Wtedy dla dowolnego $t \in T$ wartość mediany ciągu funkcji (g_i) w punkcie t definiujemy następująco: $\tilde{\mu}(g_0, g_1, \ldots, g_{M-1})(t) = \mu(g_0(t), g_1(t), \ldots, g_{M-1}(t)))$. Wtedy prawdziwy jest następujący lemat (zob. [Hei04a])

Lemat 3 ([Hei04a]). *Niech S będzie operatorem rozwiązania z F do $l_\infty(T)$, a ϕ algorytmem kwantowym rozwiązującym problem zadany przez S. Wtedy dla dowolnego $f \in F$,*

$$\mathrm{e}^{\mathrm{quant}}(S, \tilde{\mu}(\phi^M), f, e^{-M/8}) \leqslant \mathrm{e}^{\mathrm{quant}}(S, \phi, f, 1/4).$$

Na mocy tych lematów, wystarczy badać błąd algorytmu z prawdopodobieństwem porażki $\delta = 1/4$. Dla parametru $\delta \in (0, 1/4)$ koszt wzrośnie co najwyżej logarytmicznie.

Zachodzi jeszcze pytanie czy da się skonstruować algorytm kwantowy ϕ na F wykorzystując inny znany algorytm $\tilde{\phi}$ na pewnym niepustym zbiorze funkcji $\tilde{F}$ z $\tilde{D}$ do $\tilde{K}$. Chcielibyśmy dla funkcji $f \in F$ skonstruować funkcję $\tilde{f} = \Gamma(f) \in \tilde{F}$ i zastosować do niej algorytm $\tilde{\phi}$. Problem polega na tym, że algorytm ϕ może korzystać tylko z wyroczni Q na F, a postępując w powyższy sposób wykorzystujemy $\tilde{Q}_{\Gamma(f)}$, gdzie $\tilde{Q}$ jest wyrocznią na $\tilde{F}$. Symulowanie wyroczni $\tilde{Q}_{\Gamma(f)}$ przez Q_f będzie potrzebne w dalszej części rozprawy. Rozwiązanie tego problemu można znaleźć w [Hei03b]. Pokazane tam jest w jaki sposób można symulować $\tilde{Q}_{\Gamma(f)}$ przez pewien algorytm bez pomiarów korzystający z wyroczni Q. Wynik tego postępowania zawiera następujący lemat.

Załóżmy, że mamy dane odwzorowanie $\Gamma : F \to \tilde{F}$ następującej postaci:
Niech $\kappa, m^* \in \mathbb{N}$. Istnieją następujące odwzorowania

$$\eta_j : \tilde{D} \to D \quad (j = 0, 1, \ldots, \kappa - 1)$$
$$\beta : K \to \mathbb{Z}[0, 2^{m^*})$$
$$\rho : \tilde{D} \times \mathbb{Z}[0, 2^{m^*})^\kappa \to \tilde{K}$$

takie, że dla dowolnej funkcji $f \in F$ i $s \in \tilde{D}$

$$(\Gamma(f))(s) = \rho(s, \beta \circ f \circ \eta_0(s), \ldots, \beta \circ f \circ \eta_{\kappa-1}(s)). \tag{1.10}$$

Lemat 4 ([Hei03b]). *Dla danego odwzorowania* $\Gamma : F \to \tilde{F}$ *postaci (1.10), przestrzeni unormowanej* G *i algorytmu kwantowego* $\tilde{\phi}$ *z* $\tilde{F}$ *do* G*, istnieje algorytm kwantowy* ϕ *z* F *do* G *wykorzystujący*

$$n_q(\phi) = 2\kappa n_q(\tilde{\phi})$$

odwołań do wyroczni Q_f *danej przez (1.7) taki, że*

$$\phi(f) = \tilde{\phi}(\Gamma(f)).$$

Z lematu tego wynika, że jeśli symulujemy $\tilde{f}$ korzystając z κ wartości funkcji f, to koszt takiego algorytmu rośnie 2κ razy. Lemat ten można również zastosować do wyroczni o pochodnych funkcji f. Wystarczy w definicji funkcji Γ zastąpić f przez g zdefiniowaną w (1.8).

W rozprawie tej zwykle nie będziemy używać wprost czystych algorytmów kwantowych postaci (1.9). Będziemy używać tak zwanych *algorytmów hybrydowych*, które są kombinacją algorytmów kwantowych z wieloma pomiarami i klasycznych algorytmów deterministycznych. Jednak na podstawie lematu 1 każdy algorytm z wieloma pomiarami możemy zastąpić algorytmem z jednym pomiarem zwiększając koszt co najwyżej dwukrotnie. Ponadto, jeśli algorytm kwantowy korzysta z wyników pewnego algorytmu deterministycznego, to na mocy lematu 4, możemy zastąpić taki algorytm czystym algorytmem kwantowym z wyrocznią o f, którego liczba odwołań do wyroczni przekracza co najwyżej dwukrotnie iloczyn kosztu deterministycznego i liczby odwołań do wyroczni tamtego algorytmu. Każdy algorytm hybrydowy można więc zastąpić przez czysty algorytm kwantowy. Koszt algorytmu hybrydowego będziemy liczyć jako iloczyn liczby odwołań do wyroczni kwantowych i kosztu deterministycznego obliczenia jednej wyroczni. Również w dolnych ograniczeniach na złożoność wystarczy ograniczyć się do czystych algorytmów kwantowych, ponieważ koszty algorytmu hybrydowego i symulującego go czystego algorytmu kwantowego różnią się co najwyżej o stałą multiplikatywną.

Rozdział 2

Przydatne wyniki o złożoności wybranych problemów

2.1. Wstęp

Przedstawimy w tym rozdziale pewne znane wyniki, które będą przydatne w dalszej części pracy. Pokażemy znane ograniczenia na złożoność w modelach randomizacyjnym i kwantowym dla problemów dyskretnych takich jak: przeszukiwanie baz danych, maksymalizacja funkcji dyskretnych oraz obliczanie średniej liczb rzeczywistych. Wyniki te wykorzystywane będą przy dowodzeniu ograniczeń na złożoność.
Przedstawimy również wyniki dotyczące rozwiązywania problemów początkowych dla równań różniczkowych. Przedstawimy ograniczenia na złożoność oraz optymalne algorytmy dla układów równań różniczkowych pierwszego rzędu oraz równań różniczkowych wyższych rzędów. Algorytmy te będą wykorzystywane przy dowodzeniu ograniczeń na złożoność dla problemów brzegowych.

2.2. Problemy dyskretne

2.2.1. Przeszukiwanie baz danych

Przedstawimy tutaj znane wyniki dla problemu przeszukiwania dyskretnego. Rezultaty te będą wykorzystane w dalszej części pracy przy badaniu problemu rozwiązywania równań nieliniowych. Rozważmy następujący problem: mając daną funkcję binarną $g : \{0, 1, \ldots, n\} \to \{0, 1\}$, znaleźć punkt $S(g) = i_0 \in \{0, 1, \ldots, n-1\}$ taki, że $g(i_0) = 1$. Zakładamy, że istnieje przynajmniej jeden taki punkt. Rozważmy trzy modele obliczeniowe: deterministyczny, randomizacyjny i kwantowy. Informację i koszt dla tego problemu definiujemy tak jak w rozdziale 1. W zadaniu tym nie rozważamy jednak błędu. W modelu deterministycznym uznajemy, że algorytm rozwiązuje problem, jeśli znajduje jego rozwiązanie dla każdej funkcji binarnej g, a w modelu randomizacyjnym i kwantowym, jeśli znajduje rozwiązanie z prawdopodobieństwem co najmniej $1 - \delta$, dla pewnego parametru $\delta \in (0, 1/4]$. Złożoność definiowana jest jako minimalny koszt algorytmu rozwiązującego ten problem dla dowolnej funkcji wejściowej g i oznaczona jest przez $\mathrm{comp}^{\mathrm{det}}(S, n)$ lub $\mathrm{comp}^{\mathrm{rand/quant}}(S, n, \delta)$.

Łatwo pokazać, że złożoność tego problemu w modelu deterministycznym wynosi

$$\mathrm{comp}^{\mathrm{det}}(S, n) = n - 1. \tag{2.1}$$

Jeśli bowiem przeszukamy $n - 1$ różnych wartości funkcji g to albo wśród nich znajdziemy rozwiązanie, albo rozwiązaniem jest ostatni punkt, w którym nie sprawdzaliśmy wartości funkcji. Nie wystarczy przeglądnąć mniejszej ilości punktów, ponieważ zawsze istnieje taka funkcja binarna, która będzie przyjmować wartość 1, w którymś z punktów, których nie sprawdziliśmy (a w pozostałych 0).
W modelu randomizacyjnym, przy założeniu, że szukamy rozwiązania z prawdopodobieństwem porażki nie większym niż δ, złożoność jest rzędu

$$\mathrm{comp}^{\mathrm{rand}}(S, n, \delta) = \Theta(n(1 - \delta)). \tag{2.2}$$

Optymalny algorytm w tym przypadku polega na sprawdzeniu wartości w podzbiorze o $m = \lceil n(1 - \delta) \rceil$ punktach, wybranym losowo z równomiernego rozkładu na m-elementowych podzbiorach. Każdy punkt znajduje się w wylosowanym podzbiorze z prawdopodobieństwem m/n. Tak więc prawidłowe rozwiązanie znajduje się w wylosowanym podzbiorze z prawdopodobieństwem $m/n \geqslant n(1-\delta)/n = 1-\delta$. Wylosowanie mniejszej liczby elementów nie zapewni prawdopodobieństwa znalezienia prawidłowego rozwiązania na poziomie $1 - \delta$.

Lepsze rezultaty można osiągnąć korzystając z komputerów kwantowych. Ograniczenia na złożoność w tym modelu przedstawione są w następującym twierdzeniu.

Twierdzenie 3 ([Gro96, BBHT98]). *Istnieją dodatnie stałe C_1, C_2 takie, że dla dowolnego n i dowolnej liczby rozwiązań problemu przeszukiwania $t \geqslant 1$, złożoność problemu przeszukiwania w modelu kwantowym spełnia*

$$C_1\sqrt{n/t} \leqslant \mathrm{comp}^{\mathrm{quant}}(S,n) \leqslant C_2\sqrt{n/t}. \tag{2.3}$$

Górne ograniczenia w tym twierdzeniu przy założeniu, że istnieje tylko jedno rozwiązanie udowodnione zostało przez L. K. Grovera w 1996 roku w pracy [Gro96]. M. Boyer *et al.* w pracy [BBHT98] zbadali zachowanie algorytmu Grovera przy większej liczbie rozwiązań. Pokazali oni, że jeśli istnieje t rozwiązań (przy czym liczba rozwiązań jest znana) to koszt algorytmu Grovera jest ograniczony przez $O(\sqrt{n/t})$. Zaprezentowali oni również uogólnienie algorytmu Grovera dla nieznanej liczby rozwiązań. Koszt tego algorytmu jest tego samego rzędu co przy znanej liczbie rozwiązań.
W pracy [BBHT98] przedstawione są również dolne ograniczenia na złożoność.

Z twierdzenia tego wynika, że algorytm Grovera jest optymalny w modelu kwantowym. Ponadto widać, że komputery kwantowe dają kwadratowe przyspieszenie w stosunku do komputerów klasycznych.

2.2.2. Maksymalizacja funkcji dyskretnych

Przedstawimy tutaj rezultaty dotyczące problemu maksymalizacji (minimalizacji) dyskretnego ciągu. Wyniki te będą wykorzystywane w dalszej części pracy. Rozważmy następujący problem: mając daną funkcję dyskretną $f : \{0, 1, \ldots, n-1\} \to [0, 1]$, znaleźć liczbę $S(f) = f(i) = \max(\min)\{f(j) : j = 0, 1, \ldots, n-1\}$. Rozważymy trzy modele obliczeniowe: deterministyczny, randomizacyjny i kwantowy. Definicje informacji i kosztu dla tego problemu są takie jak w rozdziale 1. Powiemy, że algorytm rozwiązuje ten problem, gdy w modelu deterministycznym zwraca punkt, w którym maksymalizuje się funkcja f, a w modelach randomizacyjnym i kwantowym, gdy zwraca taki punkt z prawdopodobieństwem co najmniej $1-\delta$, dla pewnego $\delta \in (0, 1/4]$. Złożoność definiujemy, podobnie jak dla problemu przeszukiwania, jako minimalny koszt algorytmu rozwiązującego ten problem dla dowolnej funkcji f. Łatwo można pokazać, że złożoność tego problemu na klasycznym komputerze jest rzędu:

$$\mathrm{comp}^{\mathrm{det}}(S,n) = \Theta(n) \quad , \quad \mathrm{comp}^{\mathrm{rand}}(S,n) = \Theta(n)$$

W modelu deterministycznym algorytm znajdujący maksimum funkcji dyskretnej musi sprawdzić wszystkie wartości funkcji. Gdyby nie sprawdził wszystkich wartości, to zawsze istnieje taka funkcja dyskretna, która osiąga maksimum w którymś z niesprawdzonych punktów. W modelu randomizacyjnym algorytm optymalny rozwiązujący to zadanie z prawdopodobieństwem porażki mniejszym niż δ sprawdza wartość funkcji w $\lceil n(1-\delta)\rceil$ punktach wybranych losowo z rozkładu równomiernego. Dolne ograniczenia wynikają z ograniczeń dla problemu przeszukiwania.

Lepsze rezultaty mogą być osiągnięte na komputerze kwantowym. Ograniczenia na złożoność w tym modelu przedstawia następujące twierdzenie.

Twierdzenie 4 ([DH98, NW99]). *Istnieją dodatnie stałe C_1, C_2 takie, że dla dowolnego n, złożoność problemu maksymalizacji funkcji dyskretnej w modelu kwantowym spełnia*

$$C_1\sqrt{n} \leqslant \mathrm{comp}^{\mathrm{quant}}(S,n) \leqslant C_2\sqrt{n}. \tag{2.4}$$

Górne ograniczenia w tym twierdzeniu wykazali C. Dürr i P. Høyer. W pracy [DH98] przedstawili oni algorytm kwantowy znajdujący minimum ciągu n liczb z prawdopodobieństwem większym niż $1/2$ w czasie $O(\sqrt{n})$. Algorytm ten oparty jest na uogólnieniu algorytmu Grovera przeszukiwania baz danych.

Dolne ograniczenia dla tego problemu w modelu kwantowym wyznaczone zostały przez A. Nayaka i F. Wu w pracy [NW99]. Rozpatrywali oni ogólniejszy problem: dla $f : \{0,1,\ldots,n-1\} \to [0,1]$ i $\Delta > 0$ znaleźć Δ–aproksymację k-tego najmniejszego elementu, t.j. liczby $f(i)$, która jest j-tym najmniejszym elementem ciągu $(f(0), f(1), \ldots, f(n-1))$ dla pewnego całkowitego $j \in (k-\Delta, k+\Delta)$.

Dla $\Delta = 1$ (lub mniejszej) problem ten sprowadza się do problemu znalezienia k-tego najmniejszego elementu. Dla $k = n-1$, k-ty najmniejszy element jest wartością maksymalną funkcji f.

Nayak i Wu wyznaczyli dolne ograniczenie na złożoność tego problemu na komputerze kwantowym rzędu $\Omega(N)$, gdzie $N = \sqrt{n/\Delta} + \sqrt{k(n-k)}/\Delta$. Ograniczenia te są wyprowadzone przy użyciu metody wielomianowej R. Bealsa *et al.* z [BBR+98]. Dla $\Delta = 1$ i $k = n-1$ otrzymujemy ograniczenie $\Omega(\sqrt{n})$.

W pracy [NW99] przedstawiony jest również algorytm rozwiązujący ten problem o koszcie $O(N \log(N) \log\log(N))$. Zainspirowany on jest algorytmem znajdowania minimum Dürra i Høyera [DH98] i wykorzystuje algorytm przeszukiwania dyskretnego Boyera *et al.* [BBHT98]. Znajduje on Δ–aproksymację k-tego najmniejszego elementu

dla dowolnego $k \in \{0, 1, \ldots, n-1\}$ i $\Delta \geqslant 1$ z prawdopodobieństwem nie mniejszym niż $2/3$.

W pracy [NW99] przedstawione też są dolne ograniczenia na złożoność dla problemu poszukiwania maksimum funkcji binarnej $g : \{0, 1, \ldots, n-1\} \to \{0, 1\}$ czyli wartości logicznej alternatywy. Oznaczmy operator rozwiązania dla tego problemu przez S_1. Z dowodu twierdzenia 1.5 z [NW99] wynika następujące twierdzenie o dolnym ograniczeniu na złożoność problemu maksymalizacji w węższej klasie funkcji dyskretnych niż w twierdzeniu 4.

Twierdzenie 5 ([NW99]). *Istnieje dodatnia stała C taka, że dla dowolnego n, złożoność problemu maksymalizacji funkcji binarnej w modelu kwantowym spełnia*

$$\mathrm{comp}^{\mathrm{quant}}(S_1, n) \geqslant C\sqrt{n}. \tag{2.5}$$

2.2.3. Średnia liczb rzeczywistych

Przedstawimy w tym podrozdziale wyniki dotyczące złożoności problemu obliczania średniej z ciągu liczb rzeczywistych. Rezultaty te będą używane do ustanowienia ograniczeń na złożoność dla problemu rozwiązywania równań różniczkowych (poprzez ograniczenia dla problemu całkowania). Problem ten wygląda następująco: mając daną funkcję dyskretną $f : \{0, 1, \ldots, n-1\} \to [0, 1]$, obliczyć średnią ciągu $(f(0), f(1), \ldots, f(n-1))$, czyli liczbę $S(f) = \dfrac{1}{n}\sum_{i=0}^{n-1} f(i)$ z dokładnością $\varepsilon \in (0, 1/2)$. Rozważymy trzy modele obliczeniowe: deterministyczny, randomizacyjny i kwantowy. Definicje informacji, kosztu, błędu i złożoności dla tego problemu są takie jak w rozdziale 1.

Znana jest złożoność tego problemu na klasycznym komputerze. W modelu deterministycznym wynosi ona

$$\mathrm{comp}^{\mathrm{det}}(\varepsilon, S) = \Theta\left(n(1-2\varepsilon)\right). \tag{2.6}$$

Aby pokazać ograniczenie z góry, przyjmijmy $m = \lceil n(1-2\varepsilon) \rceil$ i zdefiniujmy algorytm jako $\phi(f) = (1/n)(\sum_{i=0}^{m-1} f(i) + (1/2)(n-m))$. Błąd takiego algorytmu jest nie większy niż ε, a koszt wynosi m. Aby pokazać ograniczenie z dołu załóżmy, że algorytm korzysta z m wartości funkcji. Weźmy dwie funkcje f_1 i f_2, które mają takie same wartości w punktach, z których korzysta algorytm, a w pozostałych $n-m$ punktach funkcja f_1 przyjmuje wartość 0, a f_2 wartość 1. Różnica pomiędzy rozwiązaniami dla tych dwóch funkcji wynosi $(n-m)/n$. Błąd algorytmu nie może być mniejszy od połowy tej liczby,

czyli jest nie mniejszy niż $(n-m)/(2n)$. Aby błąd był nie większy niż ε, musi więc zachodzić $(n-m)/(2n) \leqslant \varepsilon$. Stąd, $m \geqslant n(1-2\varepsilon)$.
W modelu randomizacyjnym złożoność tego problemu wynosi (zobacz [Mat92, Nov88])

$$\mathrm{comp}^{\mathrm{rand}}(\varepsilon, S) = \Theta(\min\{n, \varepsilon^{-2}\}). \tag{2.7}$$

Znane są również ograniczenia na złożoność tego problemu na komputerze kwantowym. Zawarte są one w następującym twierdzeniu.

Twierdzenie 6 ([Gro98, NW99]). *Istnieją dodatnie stałe C_1, C_2 takie, że dla dowolnego $n \in \mathbb{N}$ i $\varepsilon > 0$, złożoność problemu obliczania średniej w modelu kwantowym spełnia*

$$C_1 \min\{n, \varepsilon^{-1}\} \leqslant \mathrm{comp}^{\mathrm{quant}}(\varepsilon, S) \leqslant C_2 \min\{n, \varepsilon^{-1}\}. \tag{2.8}$$

Górne ograniczenie dla funkcji binarnych pochodzi z roku 1998 z pracy G. Brassarda *et al.* [BHT98]. Algorytm dla problemu liczenia średniej ciągu liczb z przedziału $[0,1]$ skonstruował L. K. Grover w artykule [Gro98]. Algorytm ten liczy średnią z błędem nie większym niż ε i jego koszt jest rzędu $O(\varepsilon^{-1})$.
Dolne ograniczenia dla tego problemu znaleźli A. Nayak i F. Wu. W pracy [NW99] pokazali oni, że algorytm Grovera jest optymalny.

Jak widać, komputery kwantowe dają kwadratowe przyspieszenie dla problemu liczenia średniej w stosunku do modelu randomizacyjnego.

2.3. Problem ciągły – równania różniczkowe

2.3.1. Układy równań różniczkowych

Przedstawimy tutaj pewne wyniki dotyczące złożoności problemu rozwiązywania problemów początkowych dla układów równań różniczkowych w modelu randomizacyjnym i kwantowym. Wyniki te zostaną wykorzystane w rozdziale 5.
W pracy Kacewicza [Kac06], rozważany jest problem rozwiązywania problemów początkowych postaci

$$\begin{cases} z'(x) = f(z(x)), & x \in [a,b], \\ z(a) = \eta, \end{cases} \tag{2.9}$$

gdzie $a < b$, $f : \mathbb{R}^d \to \mathbb{R}^d$, $z : [a,b] \to \mathbb{R}^d$, a f należy do klasy Höldera

$$F_d^{r,\rho} = \{f : \mathbb{R}^d \to \mathbb{R}^d |\ f \in C^r(\mathbb{R}^d), |D^{(i)} f(y_1)| \leqslant D,\ i = 0, 1, \ldots r, \\ \|D^{(r)} f(y_1) - D^{(r)} f(y_2)\| \leqslant H\|y_1 - y_2\|^\rho,\ \ y_1, y_2 \in \mathbb{R}^d\},$$

gdzie $r \geqslant 0$, $\rho \in (0,1]$, D i H są pewnymi dodatnimi stałymi definiującymi klasę, a $D^{(i)}$ przebiega przez zbiór wszystkich pochodnych cząstkowych rzędu i. Oznaczmy operator rozwiązania dla tego problemu przez S. Układ nieautonomiczny $z'(x) = f(x, z(x))$ również można zapisać w postaci (2.9) poprzez dodanie jednego równania skalarnego

$$\begin{bmatrix} u'(x) \\ z'(x) \end{bmatrix} = \begin{bmatrix} 1 \\ f(u(x), z(x)) \end{bmatrix}.$$

W pracy [Kac06] rozważany jest model randomizacyjny i kwantowy. Przedstawione są tam algorytmy rozwiązujące ten problem, zdefiniowane rekurencyjnie na zmniejszających się przedziałach. Oznaczmy przez $A^K([c,d], n, y, \delta)$ algorytm przybliżający rozwiązanie lokalnego problemu początkowego

$$\begin{cases} u'(x) = f(u(x)), & x \in [c,d], \\ u(c) = y. \end{cases} \tag{2.10}$$

Liczba K jest tu parametrem naturalnym, $[c,d]$ jest podprzedziałem przedziału $[a,b]$, na którym działa algorytm, $n+1$ jest liczbą (równomiernych) punktów siatki w przedziale $[c,d]$, y jest wektorem wartości początkowych, a $\delta \in (0, 1/2)$ jest prawdopodobieństwem porażki. Wynikiem algorytmu jest funkcja będąca wielomianem na M podprzedziałach o równej długości przedziału $[c,d]$, gdzie $M = n^{2^K-1}$ w przypadku randomizacyjnym i $M = n^K$ w przypadku kwantowym. Szczegóły można znaleźć w [Kac06]. Niech L będzie stałą Lipschitza funkcji f (w klasie $F_d^{r,\rho}$ dla $r \geqslant 1$ można przyjąć $L = D$). Algorytmy te mają następujące własności, które będą wykorzystane w rozdziale 5.

Twierdzenie 7 ([Kac06]). *Niech $[c,d] \subset [a,b]$, a $\delta \in (0,1/2)$. Niech l^K będzie przybliżeniem rozwiązania problemu (2.10) zwróconym przez algorytm $A^K([c,d], n, y, \delta)$. Niech $\alpha_K = (r+\rho)(2^K-1) + 2^{K-1} - 1$, $\beta_K = 2^K - 1$ w modelu randomizacyjnym, a $\alpha_K = (r+\rho)K + K - 1$, $\beta_K = K$ w modelu kwantowym. Wtedy istnieją dodatnie stałe C_1^K, $\tilde{C}_1^K$, C_2^K zależne tylko od parametrów klasy, a, b, K i niezależne od warunków początkowych takie, że*

$$\sup_{t\in[c,d]} \|u(t) - l^K(t)\| \leqslant \begin{cases} \tilde{C}_1^K n^{-\alpha_K} & \text{jeśli } L(d-c) \text{ jest dowolne,} \\ & a\ n \geqslant L(d-c)/\ln 2, \\ C_1^K (d-c)^{r+\rho+1} n^{-\alpha_K} & \text{jeśli } L(d-c) \leqslant \ln 2 \ i\ n \in \mathbb{N} \end{cases} \tag{2.11}$$

z prawdopodobieństwem przynajmniej $1-\delta$.
Koszt całkowity ograniczony jest przez

$$\operatorname{cost}^{\text{rand/quant}}(S, A^K, F_d^{r,\rho}) \leqslant C_2^K n^{\beta_K} \left(\beta_K \log n + \log 1/\delta\right). \tag{2.12}$$

Błąd $\sup_{t\in[c,d]} \left\|u(t) - l^K(t)\right\|$ *jest deterministycznie ograniczony z góry przez pewną stałą* Q *zależną tylko od* a, b, K *i parametrów klasy.*

Z twierdzenia tego wynikają następujące ograniczenia na błąd oraz koszt algorytmów.

Wniosek 1 ([Kac06]). *Niech* $\gamma \in (0,1)$. *Istnieją algorytmy randomizacyjny* ϕ^{rand} *i kwantowy* ϕ^{quant} *dla problemu (2.9) i dodatnie stałe* $K_1(\gamma)$, $K_2(\gamma)$ *oraz* $\varepsilon_0(\gamma)$ *(zależne od* γ, *parametrów klasy* $F_d^{r,\rho}$, a *i* b*) takie, że dla każdego* $\varepsilon \in (0, \varepsilon_0(\gamma))$ *i* $\delta \in (0, 1/2)$ *algorytmy* ϕ^{rand} *i* ϕ^{quant} *zwracają przybliżenia* l *rozwiązania z problemu (2.9) spełniające*

$$\sup_{x\in[a,b]} \|z(x) - l(x)\| \leqslant \varepsilon \tag{2.13}$$

z prawdopodobieństwem nie mniejszym niż $1-\delta$ *i kosztem*

$$\text{cost}^{\text{rand}}(S, \phi^{\text{rand}}, F_d^{r,\rho}) \leqslant K_1(\gamma) \left(\frac{1}{\varepsilon}\right)^{1/(r+\rho+1/2-\gamma)} \log\frac{1}{\delta}, \tag{2.14}$$

$$\text{cost}^{\text{quant}}(S, \phi^{\text{quant}}, F_d^{r,\rho}) \leqslant K_2(\gamma) \left(\frac{1}{\varepsilon}\right)^{1/(r+\rho+1-\gamma)} \log\frac{1}{\delta}. \tag{2.15}$$

W dowodzie tego twierdzenia użyto algorytmów A^K przedstawionych wyżej, dla $K = \lceil \log(1/\gamma+1) \rceil$, $n = \lceil (C^{\text{rand}}/\varepsilon)^{1/\alpha_K} \rceil$ w modelu randomizacyjnym oraz $K = \lceil 2/\gamma \rceil$, $n = \lceil (C^{\text{quant}}/\varepsilon)^{1/\alpha_K} \rceil$ w modelu kwantowym, dla pewnych stałych C^{rand} i C^{quant}. Wynikiem algorytmu jest więc funkcja kawałkami wielomianowa na M podprzedziałach o tej samej długości przedziału $[a,b]$, gdzie $M = \Theta\left(\varepsilon^{-1/(r+\rho+1/2-\gamma/2)}\right)$ w modelu randomizacyjnym i $M = \Theta\left(\varepsilon^{-1/(r+\rho+1-\gamma/2)}\right)$ w modelu kwantowym.

Ograniczenia (2.14) i (2.15) prowadzą do górnych oszacowań na złożoność problemu (2.9),

$$\text{comp}^{\text{rand}}(\varepsilon, S, F_d^{r,\rho}) = O\left(\left(\frac{1}{\varepsilon}\right)^{1/(r+\rho+1/2-\gamma)}\right), \tag{2.16}$$

$$\text{comp}^{\text{quant}}(\varepsilon, S, F_d^{r,\rho}, \delta) = O\left(\left(\frac{1}{\varepsilon}\right)^{1/(r+\rho+1-\gamma)} \log\frac{1}{\delta}\right). \tag{2.17}$$

Dolne ograniczenia na złożoność tego problemu przedstawione są w pracy [Kac04]. Ograniczenia te zawarte są w następującym twierdzeniu.

Twierdzenie 8 ([Kac04]). *Istnieją dodatnie stałe k^{rand} i k^{quant} takie, że dla dowolnego dostatecznie małego $\varepsilon > 0$, ε–złożoność problemu (2.9) w klasie $F_d^{r,\rho}$ spełnia*

$$\mathrm{comp}^{\mathrm{rand}}(\varepsilon, S, F_d^{r,\rho}) \geqslant k^{\mathrm{rand}} \left(\frac{1}{\varepsilon}\right)^{1/(r+\rho+1/2)}, \tag{2.18}$$

$$\mathrm{comp}^{\mathrm{quant}}(\varepsilon, S, F_d^{r,\rho}, \delta) \geqslant k^{\mathrm{quant}} \left(\frac{1}{\varepsilon}\right)^{1/(r+\rho+1)}. \tag{2.19}$$

Górne ograniczenia dla tego problemu w modelu randomizacyjnym zostały poprawione w pracy [HM08]. Rozważany tam jest problem postaci

$$\begin{cases} z'(x) = f(x, z(x)), & x \in [a, b], \\ z(a) = \eta, \end{cases} \tag{2.20}$$

gdzie funkcja prawej strony $f : [a, b] \times \mathbb{R}^d \to \mathbb{R}^d$ należy do klasy Höldera $\bar{F}_d^{r,\rho}$ danej przez

$$\begin{aligned} \bar{F}_d^{r,\rho} = \{ & f : [a, b] \times \mathbb{R}^d \to \mathbb{R}^d \mid f \in C^r([a, b] \times \mathbb{R}^d), |D^{(i)} f(x_1, y_1)| \leqslant D, \\ & i = 0, 1, \ldots r, \quad \|D^{(r)} f(x_1, y_1) - D^{(r)} f(x_2, y_2)\| \leqslant H \left(|x_1 - x_2|^\rho + \|y_1 - y_2\|^\rho\right), \\ & x_1, x_2 \in [a, b], \ y_1, y_2 \in \mathbb{R}^d \}, \end{aligned}$$

gdzie $r \geqslant 0$, $\rho \in [0, 1]$, D i H są pewnymi stałymi, a $D^{(i)}$ przebiega przez zbiór wszystkich pochodnych cząstkowych rzędu i. Dodatkowo zakłada się tu, że funkcja f spełnia warunek Lipschitza ze względu na drugą zmienną. Oznaczmy operator rozwiązania dla tego problemu przez $\tilde{S}$. W pracy [HM08] pokazane są następujące ograniczenia z góry na złożoność.

Twierdzenie 9 ([HM08]). *Istnieje dodatnia stała K taka, że dla dowolnego $\varepsilon > 0$, ε–złożoność problemu (2.20) w klasie $\bar{F}_d^{r,\rho}$ spełnia*

$$\mathrm{comp}^{\mathrm{rand}}(\varepsilon, \tilde{S}, \bar{F}_d^{r,\rho}) \leqslant K \left(\frac{1}{\varepsilon}\right)^{1/(r+\rho+1/2)}. \tag{2.21}$$

Ograniczenia te są polepszone w stosunku do górnych ograniczeń z [Kac06] o parametr γ w wykładniku i są ostre.

W rozdziale 5.3 potrzebować będziemy pewnej modyfikacji algorytmu z [Kac06] i oszacowań z góry dla funkcji kawałkami regularnych. Załóżmy, że mamy dany problem początkowy postaci (2.20), gdzie funkcja f jest postaci

$$f(t, y) = f_i(t, y), \quad \text{dla } t \in [t_i, t_{i+1}), \ y \in \mathbb{R}^d, \tag{2.22}$$

gdzie $t_i = a + i(b - a)/M$ dla $i = 0, 1, \ldots, M$, $M \in \mathbb{N}$, a funkcje $f_i : [a, b] \times R^d \to R^d$ należą do klasy Höldera $\bar{F}_d^{r,\rho}$, $r \in \mathbb{N}$, $\rho \in (0, 1]$. Oznaczmy klasę funkcji postaci (2.22)

przez $F_d^{r,\rho}(M)$. Rozwiązaniem tak zdefiniowanego problemu jest funkcja ciągła na $[a,b]$ i $r+1$ razy różniczkowalna w każdym podprzedziale (t_i, t_{i+1}), $i = 0, 1, \ldots, M-1$. Rozwiązanie to istnieje i jest jednoznaczne. Oznaczmy operator rozwiązania tego problemu przez S^*. Przedstawimy modyfikację wniosku 1 dla tak postawionego problemu.

Twierdzenie 10. *Niech $\gamma \in (0,1)$. Istnieją dodatnie stałe $\varepsilon_0(\gamma)$, $K_1(\gamma)$, $K_2(\gamma)$ (zależne tylko od γ, parametrów klasy $F_d^{r,\rho}$, a i b) takie, że dla każdego $\delta \in (0, 1/2)$, $M \in \mathbb{N}$ i $\varepsilon \in (0, \varepsilon_0(\gamma))$ istnieją algorytmy randomizacyjny ϕ^{rand} i kwantowy ϕ^{quant} takie, że przybliżenia l rozwiązania z problemu (2.20) z funkcją prawej strony klasy $F_d^{r,\rho}(M)$ spełniają*

$$\sup_{x\in[a,b]} \|z(x) - l(x)\| \leqslant \varepsilon \tag{2.23}$$

z prawdopodobieństwem nie mniejszym niż $1-\delta$ i kosztem

$$\mathrm{cost}^{\mathrm{rand}}(S^*, \phi^{\mathrm{rand}}, F_d^{r,\rho}(M)) \leqslant K_1(\gamma) M^{1/(2(r+\rho+1/2))} \left(\frac{1}{\varepsilon}\right)^{1/(r+\rho+1/2-\gamma)} \left(\log M + \log\frac{1}{\delta}\right), \tag{2.24}$$

$$\mathrm{cost}^{\mathrm{quant}}(S^*, \phi^{\mathrm{quant}}, F_d^{r,\rho}(M)) \leqslant K_2(\gamma) M^{1/(r+\rho+1)} \left(\frac{1}{\varepsilon}\right)^{1/(r+\rho+1-\gamma)} \left(\log M + \log\frac{1}{\delta}\right). \tag{2.25}$$

Dowód. Rozważmy lokalne problemy początkowe dla $i = 1, 2, \ldots, M-1$

$$\begin{cases} z_i'(x) = f(t, z_i(x)), & x \in [t_i, t_{i+1}], \\ z_i(t_i) = \eta_i, \end{cases} \tag{2.26}$$

gdzie $\eta_0 = \eta$, $\eta_i = z_{i-1}(t_i)$, $i = 0, 1, \ldots, M-1$. Wtedy dla $x \in [t_i, t_{i+1})$ mamy $z(t) = z_i(t)$. Na każdym z przedziałów $[t_i, t_{i+1}]$ algorytmem Kacewicza $A^K([t_i, t_{i+1}], n, Y_i, \delta_1)$ wyznaczamy przybliżenia l_i^K problemu

$$\begin{cases} u_i'(x) = f(t, u_i(x)), & x \in [t_i, t_{i+1}], \\ u_i(t_i) = Y_i. \end{cases} \tag{2.27}$$

Warunki początkowe definiujemy tutaj jako $Y_0 = \eta$, $Y_i = l_{i-1}^K(t_i)$, $i = 1, 2, \ldots, M-1$. Liczby naturalne $n \geqslant L(b-a)/\ln 2$ i K oraz $\delta_1 \in (0, 1/2)$ są parametrami, które sprecyzujemy później. Niech $l(t) = l_i^K(t)$, dla $t \in [t_i, t_{i+1})$ oraz $l(b) = l_{M-1}^K(b)$ będzie poszukiwanym przybliżeniem rozwiązania.

Korzystając ze standardowej zależności rozwiązania od warunków początkowych (zob. np. [HNW00, s. 56]) dla $i = 0, 1, \ldots, M-1$ dostajemy ograniczenie

$$\begin{aligned} \|z - l_i^K\|_{[t_i,t_{i+1}]} &\leqslant \|z - u_i\|_{[t_i,t_{i+1}]} + \|u_i - l_i^K\|_{[t_i,t_{i+1}]} \\ &\leqslant e^{D(b-a)/M} \|\eta_i - Y_i\| + \|u_i - l_i^K\|_{[t_i,t_{i+1}]}, \end{aligned}$$

gdzie $\|\cdot\|_{[t_i,t_{i+1}]}$ oznacza normę supremum na przedziale $[t_i, t_{i+1}]$. Korzystając z twierdzenia 7 dostajemy więc z prawdopodobieństwem co najmniej $1-\delta_1$

$$\|z - l_i^K\|_{[t_i,t_{i+1}]} \leqslant e^{D(b-a)/M}\|\eta_i - Y_i\| + \hat{C}_1^K \frac{(b-a)^{r+\rho+1}}{M^{r+\rho+1}} n^{-\alpha_K}, \tag{2.28}$$

gdzie α_K zdefiniowane jest w twierdzeniu 7, a $\hat{C}_1^K$ jest pewną stałą zależną tylko od a, b, C_1^K, $\tilde{C}_1^K$, L, r, ρ. Mamy więc

$$\|z - l_0^K\|_{[t_0,t_1]} \leqslant \hat{C}_1^K \frac{(b-a)^{r+\rho+1}}{M^{r+\rho+1}} n^{-\alpha_K} \tag{2.29}$$

z prawdopodobieństwem co najmniej $1-\delta_1$. W następnym przedziale, z prawdopodobieństwem co najmniej $(1-\delta_1)^2$, mamy

$$\begin{aligned}\|z - l_1^K\|_{[t_1,t_2]} &\leqslant e^{D(b-a)/M}\|z(t_1) - l_0^K(t_1)\| + \hat{C}_1^K \frac{(b-a)^{r+\rho+1}}{M^{r+\rho+1}} n^{-\alpha_K} \\ &\leqslant \left(1 + e^{D(b-a)/M}\right) \hat{C}_1^K \frac{(b-a)^{r+\rho+1}}{M^{r+\rho+1}} n^{-\alpha_K}.\end{aligned}$$

Postępując indukcyjnie dalej, dla $i = 1, 2, \ldots, M-1$, z prawdopodobieństwem co najmniej $(1-\delta_1)^{i+1}$ otrzymujemy

$$\|z - l_1^K\|_{[t_i,t_{i+1}]} \leqslant \left(1 + e^{D(b-a)/M} + \ldots + e^{D(b-a)i/M}\right) \hat{C}_1^K \frac{(b-a)^{r+\rho+1}}{M^{r+\rho+1}} n^{-\alpha_K}.$$

Z prawdopodobieństwem co najmniej $(1-\delta_1)^M \geqslant 1 - M\delta_1$ zachodzi więc

$$\begin{aligned}\|z - l\| &\leqslant \left(1 + e^{D(b-a)/M} + \ldots + e^{D(b-a)(M-1)/M}\right) \hat{C}_1^K \frac{(b-a)^{r+\rho+1}}{M^{r+\rho+1}} n^{-\alpha_K} \\ &\leqslant \hat{C}_1^K M e^{D(b-a)} \frac{(b-a)^{r+\rho+1}}{M^{r+\rho+1}} n^{-\alpha_K} = C^K M^{-(r+\rho)} n^{-\alpha_K},\end{aligned}$$

gdzie $C^K = \hat{C}_1^K e^{D(b-a)}(b-a)^{r+\rho+1}$. Przyjmijmy $n = \left\lceil \left(\frac{\varepsilon M^{r+\rho}}{C^K}\right)^{-1/\alpha_K} \right\rceil$ i $\delta_1 = \delta/M$. Wtedy, z prawdopodobieństwem co najmniej $1-\delta$ mamy

$$\|z - l\| \leqslant \varepsilon.$$

Koszt wyznaczenia przybliżenia l jest sumą kosztów wyznaczenia wszystkich l_i^K, $i = 0, 1, \ldots, M-1$. Zgodnie z twierdzeniem 7 jest on ograniczony przez

$$\begin{aligned}\mathrm{cost}^{\mathrm{rand/quant}}(S^*, \phi^{\mathrm{rand/quant}}, F_d^{r,\rho}(M)) &\leqslant C_2^K M n^{\beta_K} \left(\beta_K \log n + \log 1/\delta_1\right) \\ &\leqslant \tilde{C}_2^K M^{1-(r+\rho)\beta_K/\alpha_K} \varepsilon^{-\beta_K/\alpha_K} (\log\frac{1}{\varepsilon} + \log M + \log\frac{1}{\delta}),\end{aligned} \tag{2.30}$$

gdzie $\tilde{C}_2^K$ jest stałą niezależną od ε, M i δ (zależną natomiast od K, a, b i parametrów klasy), a β_K zdefiniowane jest w twierdzeniu 7.

Rozważmy model randomizacyjny. Przypomnijmy, że $\alpha_K = (r+\rho)(2^K-1)+2^{K-1}-1$, a $\beta_K = 2^K - 1$. Zauważmy, że $\alpha_K/\beta_K = r + \rho + 1/2 - 1/(2(2^K-1))$. Przyjmijmy $K = \lceil \log(1/\gamma + 1) \rceil$. Wtedy $r + \rho + 1/2 - \gamma/2 \leqslant \alpha_K/\beta_K \leqslant r + \rho + 1/2$. Wracając do (2.30) dostajemy

$$\begin{aligned}\operatorname{cost}^{\mathrm{rand}}(S^*, \phi^{\mathrm{rand}}, F_d^{r,\rho}(M)) &\leqslant \tilde{C}_2^K M^{1/(2(r+\rho+1/2))} \Big(\frac{1}{\varepsilon}\Big)^{1/(r+\rho+1/2-\gamma/2)} \Big(\log\frac{1}{\varepsilon} + \log M + \log\frac{1}{\delta}\Big) \\ &\leqslant K_1(\gamma) M^{1/(2(r+\rho+1/2))} \Big(\frac{1}{\varepsilon}\Big)^{1/(r+\rho+1/2-\gamma)} \Big(\log M + \log\frac{1}{\delta}\Big),\end{aligned}$$

gdzie $K_1(\gamma)$ jest stałą zależną od a, b, γ i parametrów klasy, a niezależną od M i ε. To kończy dowód w przypadku randomizacyjnym.

Przejdźmy do przypadku kwantowego. Przypomnijmy, że $\alpha_K = (r+\rho)K + K - 1$, a $\beta_K = K$. Przyjmijmy $K = \lceil 2/\gamma \rceil$. Wtedy $r + \rho + 1 - \gamma/2 \leqslant \alpha_K/\beta_K \leqslant r + \rho + 1$. Podstawiając to do (2.30) dostajemy ograniczenie

$$\begin{aligned}\operatorname{cost}^{\mathrm{quant}}(S^*, \phi^{\mathrm{quant}}, F_d^{r,\rho}(M)) &\leqslant \tilde{C}_2^K M^{1/(r+\rho+1)} \Big(\frac{1}{\varepsilon}\Big)^{1/(r+\rho+1-\gamma/2)} \Big(\log\frac{1}{\varepsilon} + \log M + \log\frac{1}{\delta}\Big) \\ &\leqslant K_2(\gamma) M^{1/(r+\rho+1)} \Big(\frac{1}{\varepsilon}\Big)^{1/(r+\rho+1-\gamma)} \Big(\log M + \log\frac{1}{\delta}\Big),\end{aligned}$$

gdzie $K_2(\gamma)$ jest stałą zależną od a, b, γ i parametrów klasy, a niezależną od M i ε. To kończy dowód twierdzenia. □

2.3.2. Równania różniczkowe wyższych rzędów

W rozdziale 5 będziemy potrzebować również pewnych wyników dotyczących problemów początkowych dla równań różniczkowych skalarnych wyższych rzędów. Problemy takie w modelu randomizacyjnym i kwantowym rozważane są w pracy [GS08]. Rozważmy następujący problem początkowy

$$\begin{cases} u^{(k)}(x) = g\left(x, u(x), u'(x), \dots, u^{(q)}(x)\right), & x \in [a,b], \\ u^{(j)}(a) = u_a^j, \quad j = 0, 1, \dots, k-1, \end{cases} \tag{2.31}$$

gdzie $0 \leqslant q < k$, $g : [a,b] \times \mathbb{R}^{q+1} \to \mathbb{R}$ należy do klasy F^r funkcji r-krotnie różniczkowalnych z ograniczonymi pochodnymi cząstkowymi, $u : [a,b] \to \mathbb{R}$ $(a < b)$. Operator rozwiązania dla tego problemu jest zdefiniowany przez $S(f) = u$, a jako normę w przestrzeni rozwiązań przyjmuje się normę supremum. W pracy [GS08] przedstawiono algorytmy randomizacyjny i kwantowy rozwiązujące ten problem. Algorytmy te są

definiowane w sposób rekurencyjny na zmniejszających się przedziałach. Oznaczmy przez $A^K([c,d],n,Y,\delta)$ algorytmy rozwiązujące lokalny problem początkowy

$$\begin{cases} u^{(k)}(x) = g\left(x, u(x), u'(x), \ldots, u^{(q)}(x)\right), & x \in [c,d], \\ u^{(j)}(c) = y^j, \quad j = 0,1,\ldots,k-1, \end{cases} \tag{2.32}$$

W notacji tej K jest parametrem naturalnym oznaczającym głębokość rekurencji, $[c,d]$ jest przedziałem zawartym w $[a,b]$, na którym działa algorytm, $n+1$ jest liczbą (równomiernych) punktów siatki w przedziale $[c,d]$, $Y = [y^0, y^1, \ldots, y^{k-1}]$ jest wektorem warunków początkowych, a $\delta \in (0, 1/2)$ jest pewnym parametrem związanym z prawdopodobieństwem porażki. Wynikiem algorytmu $A^K([c,d],n,Y,\delta)$ jest funkcja kawałkami wielomianowa l^K. Szczegóły można znaleźć w [GS08].
Algorytmy te mają następujące własności, które będą wykorzystane w rozdziale 5.

Twierdzenie 11 ([GS08]). *Niech $[c,d] \subset [a,b]$ i $\delta \in (0,1/2)$. Niech l^K będzie przybliżeniem rozwiązania problemu (2.32) zwróconym przez algorytm $A^K([c,d],n,Y,\delta)$, gdzie $Y = [y^0,\ldots,y^{k-1}]$, $n \geqslant 2$. Niech $\alpha_K = r2^K + 2^{K-1} - r - 1$, $\beta_K = 2^K - 1$, $\psi_K(n) = \sum_{i=1}^{K-1} n^{2^i-1}$ w modelu randomizacyjnym, a $\alpha_K = rK + K - 1$, $\beta_K = K$, $\psi_K(n) = \dfrac{n^K - n}{n-1}$ w modelu kwantowym. Wtedy istnieją dodatnie stałe C_1^K, C_2^K zależne tylko od parametrów klasy, a, b, k, K i warunków początkowych takie, że dla dostatecznie dużego n zachodzi*

$$\sum_{j=0}^{k-1} \sup_{x\in[c,d]} \left| u^{(j)}(x) - l^{K\,(j)}(x) \right| \leqslant C_1^K \, (d-c)^{r+1} n^{-\alpha_K}, \tag{2.33}$$

z prawdopodobieństwem co najmniej

$$\mathbf{P}(A^K, n) = (1-\delta)^{\psi_K(n)}. \tag{2.34}$$

Koszt całkowity ograniczony jest przez

$$\mathrm{cost}^{\mathrm{rand/quant}}(S, A^K, F^r) \leqslant C_2^K n^{\beta_K} \log 1/\delta. \tag{2.35}$$

(W (2.33), przez pochodną l^K w w punktach nieciągłości rozumiemy pochodną prawostronną.)
Można dodatkowo pokazać, że dla pewnej stałej $\tilde{L}$ zależnej od parametrów klasy, a, b, k, K i warunków początkowych zachodzą deterministyczne ograniczenia

$$\sum_{j=0}^{k-1} \sup_{x\in[c,d]} \left| u^{(j)}(x) - l^{K\,(j)}(x) \right| \leqslant \tilde{L} \, (d-c)^{r+1} n^{-r(2^K-1)}, \text{ w modelu randomizacyjnym,}$$

$$\sum_{j=0}^{k-1} \sup_{x\in[c,d]} \left| u^{(j)}(x) - l^{K\,(j)}(x) \right| \leqslant \tilde{L} \, (d-c)^{r+1} n^{-rK}, \text{ w modelu kwantowym.}$$

W pracy [GS08] zaprezentowane są również górne i dolne ograniczenia na złożoność problemu (2.31) w modelu randomizacyjnym i kwantowym. Górne ograniczenia zawarte są w następującym twierdzeniu.

Twierdzenie 12 ([GS08]). *Dla dowolnego $\gamma \in (0,1)$ istnieją nieujemne stałe $C_1(\gamma)$, $C_2(\gamma)$, $\varepsilon_0(\gamma)$ (zależne wyłącznie od γ, parametrów klasy F^r, a, b i warunków początkowych) takie, że dla dowolnego $\varepsilon \in (0, \varepsilon_0(\gamma))$ złożoność problemu (2.31) w modelach randomizacyjnym i kwantowym spełnia*

$$\operatorname{comp}^{\mathrm{rand}}(\varepsilon, S, F^r) \leqslant C_1(\gamma) \left(\frac{1}{\varepsilon}\right)^{1/(r+1/2-\gamma)} \tag{2.36}$$

oraz dla $\delta \in (0, 1/2)$

$$\operatorname{comp}^{\mathrm{quant}}(\varepsilon, S, F^r, \delta) \leqslant C_2(\gamma) \left(\frac{1}{\varepsilon}\right)^{1/(r+1-\gamma)} \log \frac{1}{\delta}. \tag{2.37}$$

Następujące twierdzenie przedstawia dolne ograniczenie na złożoność tego problemu.

Twierdzenie 13 ([GS08]). *Istnieją dodatnie stałe C_1 i C_2 zależne wyłącznie od parametrów klasy F^r oraz k takie, że dla dowolnego dostatecznie małego $\varepsilon > 0$*

$$\operatorname{comp}^{\mathrm{rand}}(\varepsilon, S, F^r) \geqslant C_1 \left(\frac{1}{\varepsilon}\right)^{1/(r+1/2)}, \tag{2.38}$$

oraz dla dowolnego $\delta \in (0, 1/4)$

$$\operatorname{comp}^{\mathrm{quant}}(\varepsilon, S, F^r, \delta) \geqslant \operatorname{comp}^{\mathrm{quant}}(\varepsilon, S, F^r) \geqslant C_2 \left(\frac{1}{\varepsilon}\right)^{1/(r+1)}. \tag{2.39}$$

Z powyższych twierdzeń wynika, że algorytmy zaprezentowane w [GS08] są prawie optymalne. Prawie optymalność ma miejsce z dokładnością do dowolnie małego parametru $\gamma \in (0, 1)$.

Rozdział 3

Maksymalizacja funkcji ciągłych

3.1. Wstęp

W rozdziale tym zajmiemy się problemem maksymalizacji funkcji ciągłych. Problem ten jest istotnym zagadnieniem matematyki stosowanej i pojawia się przy rozwiązywaniu wielu zadań ciągłych.

Zajmiemy się tutaj poszukiwaniem maksimum funkcji z klasy Höldera. Wyznaczymy górne i dolne ograniczenia na złożoność tego problemu w modelu kwantowym. Porównamy otrzymane wyniki ze znanymi ograniczeniami w modelu deterministycznym i randomizacyjnym.

Górne ograniczenia na złożoność zostaną wyznaczone przez skonstruowanie pewnego algorytmu kwantowego rozwiązującego ten problem.

W celu wyznaczenia dolnych ograniczeń na złożoność w modelu kwantowym wykorzystamy dolne ograniczenia na złożoność problemu maksymalizacji ciągu zerojedynkowego. Górne i dolne ograniczenia na złożoność okażą się być tego samego rzędu, czyli proponowany przez nas algorytm okaże się optymalny. Pokażemy w rezultacie, że komputery kwantowe są kwadratowo szybsze od klasycznych obliczeń deterministycznych i randomizacyjnych.

W następnym podrozdziale przedstawimy sformułowanie problemu oraz zaprezentujemy w celu porównania znane wyniki dotyczące tego problemu w modelach deterministycznym i randomizacyjnym. W kolejnym podrozdziale skonstruujemy algorytm kwantowy rozwiązujący ten problem i pokażemy górne ograniczenia na złożoność. W podrozdziale 3.4 wyprowadzimy dolne ograniczenia na złożoność w modelu kwantowym.

3.2. Sformułowanie problemu i znane wyniki

W rozdziale tym zajmiemy się problemem znajdowania maksimum funkcji ciągłej $f : [0,1]^d \to \mathbb{R}$. Naszym celem jest obliczenie liczby

$$S(f) = \max_{t \in [0,1]^d} f(t)$$

z pewną dokładnością $\varepsilon > 0$. Rozważać będziemy klasę Höldera funkcji f

$$F_d^{r,\rho} = \Big\{ f : [0,1]^d \to \mathbb{R} \mid \ f \in C^r([0,1]^d),\ \|D^{(i)} f\| \leqslant D,\ i = 0,1,\ldots,r \\ \left| D^{(r)} f(x) - D^{(r)} f(y) \right| \leqslant H \|x-y\|^\rho \ \ \forall x, y \in [0,1]^d \Big\},$$

gdzie $D^{(i)}$ przebiega przez zbiór wszystkich pochodnych cząstkowych rzędu i, $r \in \mathbb{N}_0$, $0 < \rho \leqslant 1$, D i H są nieujemnymi stałymi, a $\|\cdot\| = \|\cdot\|_\infty$ oznacza odpowiednio normę maksimum i supremum dla wektorów i funkcji.
Definicje błędu, kosztu, informacji oraz złożoności wprowadzone są w rozdziale 1. Definicje błędu w różnych modelach zawarte są we wzorach 1.1, 1.2, 1.4, 1.5, natomiast złożoność definiowana jest wzorem 1.6.

Znane są górne i dolne ograniczenia na złożoność tego problemu w modelu deterministycznym i randomizacyjnym. Znaleźć je można np. w [Nov88, s. 34 i 59]. Ograniczenia te są rzędu

$$\mathrm{comp}^{\mathrm{det}}(\varepsilon, S, F_d^{r,\rho}) = \Theta\left(\left(\frac{1}{\varepsilon} \right)^{d/(r+\rho)} \right), \tag{3.1}$$

$$\mathrm{comp}^{\mathrm{rand}}(\varepsilon, S, F_d^{r,\rho}) = \Theta\left(\left(\frac{1}{\varepsilon} \right)^{d/(r+\rho)} \right). \tag{3.2}$$

W modelu deterministycznym dopuszczona jest tutaj wyłącznie informacja standardowa, dana przez wartości funkcji f oraz jej pochodnych cząstkowych w pewnych punktach.

W następnym podrozdziale wyznaczymy górne oszacowanie na złożoność problemu szukania maksimum funkcji z klasy Höldera w modelu kwantowym. Przedstawimy

algorytm wykorzystujący kwantowy algorytm znajdowania maksimum dyskretnego ciągu. Wyznaczymy również oszacowanie błędu oraz kosztu tego algorytmu.

3.3. Górne ograniczenia w modelu kwantowym

Skonstruujemy teraz algorytm kwantowy szukający maksimum funkcji z klasy Höldera. Niech $n \in \mathbb{N}$. Algorytm ϕ^* zdefiniowany jest następująco:

1. Dzielimy kostkę $[0,1]^d$ równomiernie na $N = n^d$ kostek K^i, $i = 1, 2, \ldots, n^d$, o długości krawędzi $1/n$. Oznaczmy środek kostki K^i przez t^i.
2. Na każdej kostce K^i stosujemy rozwinięcie Taylora funkcji f w punkcie t^i (zob. dodatek A). Dla $t \in K^i$ mamy

   $$f(t) = w^i(t) + R^i(t, t^i),$$

   gdzie

   $$w^i(t) = \sum_{k=0}^{r} \frac{1}{k!} f^{(k)}(t^i)(t - t^i)^k,$$

   $$R^i(t, t^i) = \int_0^1 \left(f^{(r)}(\theta t + (1-\theta)t^i) - f^{(r)}(t^i) \right) (t - t^i)^r \frac{(1-\theta)^{r-1}}{(r-1)!} d\theta.$$
3. Niech $m^i(f) = \max_{t \in K^i} w^i(t)$, $i = 1, 2, \ldots, n^d$ (pewne nieznane liczby). Niech $\tilde{m}^i(f)$ będzie przybliżeniem $m^i(f)$ obliczonym za pomocą pewnego klasycznego algorytmu z dokładnością $\varepsilon_1 = (1/n)^{r+\rho}$:

   $$|m^i(f) - \tilde{m}^i(f)| \leqslant \varepsilon_1.$$
4. Wyznaczamy maksimum $\tilde{m}^{i^*}(f)$ ciągu $\tilde{m}^1(f), \ldots, \tilde{m}^N(f)$ z prawdopodobieństwem nie mniejszym niż $3/4$ za pomocą optymalnego algorytmu kwantowego przedstawionego w podrozdziale 2.2.2.
5. Algorytm zwraca $\phi^*(f) = \tilde{m}^{i^*}(f)$.

Następujące twierdzenie przedstawia górne ograniczenia na błąd oraz koszt tego algorytmu.

Twierdzenie 14. *Istnieją dodatnie stałe C_1 i C_2 zależne wyłącznie od parametrów klasy takie, że dla dowolnego $n \in \mathbb{N}$ błąd oraz koszt algorytmu ϕ^* spełniają oszacowania*

$$\mathrm{e}^{\mathrm{quant}}(S, \phi^*, F_d^{r,\rho}) \leqslant C_1 \left(\frac{1}{n}\right)^{r+\rho},$$

$$\mathrm{cost}^{\mathrm{quant}}(S, \phi^*, F_d^{r,\rho}) \leqslant C_2 n^{d/2}.$$

Dowód. Oszacujmy najpierw błąd algorytmu ϕ^*. Niech $f \in F_d^{r,\rho}$. Z prawdopodobieństwem nie mniejszym niż 3/4 zachodzi

$$\begin{aligned}
|S(f) - \phi^*(f)| &= |\max_{t\in[0,1]^d} f(t) - \max_{i=1,\ldots,N} \tilde{m}^i(f)| \\
&= |\max_{i=1,\ldots,N} \max_{t\in K^i} f(t) - \max_{i=1,\ldots,N} \tilde{m}^i(f)| \\
&\leqslant \max_{i=1,\ldots,N} |\max_{t\in K^i} f(t) - \tilde{m}^i(f)| \\
&\leqslant \max_{i=1,\ldots,N} \Big(|\max_{t\in K^i} f(t) - m^i(f)| + |m^i(f) - \tilde{m}^i(f)| \Big) \\
&\leqslant \max_{i=1,\ldots,N} |\max_{t\in K^i} f(t) - \max_{t\in K^i} w^i(t)| + \varepsilon_1 \\
&\leqslant \max_{i=1,\ldots,N} \max_{t\in K^i} |f(t) - w^i(t)| + \varepsilon_1. \qquad (3.3)
\end{aligned}$$

Dla $t \in K^i$, korzystając z warunku Höldera mamy

$$\begin{aligned}
|f(t) - w^i(t)| &= |R^i(t, t^i)| \\
&= \left| \int_0^1 \Big(f^{(r)}(\theta t + (1-\theta)t^i) - f^{(r)}(t^i) \Big) (t - t^i)^r \frac{(1-\theta)^{r-1}}{(r-1)!} d\theta \right| \\
&\leqslant \int_0^1 \Big\| f^{(r)}(\theta t + (1-\theta)t^i) - f^{(r)}(t^i) \Big\| \, \|t - t^i\|^r \frac{(1-\theta)^{r-1}}{(r-1)!} d\theta \\
&\leqslant \int_0^1 d^r H \|\theta(t - t^i)\|^\rho \|t - t^i\|^r \frac{(1-\theta)^{r-1}}{(r-1)!} d\theta \\
&\leqslant d^r H \|t - t^i\|^{r+\rho} \int_0^1 \frac{(1-\theta)^{r-1}}{(r-1)!} d\theta \\
&\leqslant d^r H (1/2)^{r+\rho} (1/n)^{r+\rho} (1/r!) = \tilde{H} (1/n)^{r+\rho}, \qquad (3.4)
\end{aligned}$$

dla każdego i, gdzie $\tilde{H} = d^r H (1/2)^{r+\rho} (1/r!)$ jest stałą niezależną od n. Stąd i z (3.3) dla dowolnej funkcji $f \in F_d^{r,\rho}$ dostajemy górne ograniczenie na błąd

$$\mathrm{e}^{\mathrm{quant}}(S, \phi^*, f) \leqslant \tilde{H} \left(\frac{1}{n}\right)^{r+\rho} + \varepsilon_1.$$

Ponieważ $\varepsilon_1 = (1/n)^{r+\rho}$, otrzymujemy dla dowolnej funkcji $f \in F_d^{r,\rho}$

$$\mathrm{e}^{\mathrm{quant}}(S, \phi^*, f) \leqslant (\tilde{H} + 1) \left(\frac{1}{n}\right)^{r+\rho},$$

co dowodzi pierwszej części twierdzenia.

Przejdźmy teraz do oszacowania kosztu algorytmu ϕ^*. W celu obliczenia $\widetilde{m}^i(f)$ musimy znać wartość funkcji f oraz wartości wszystkich jej pochodnych cząstkowych rzędu aż do r w punkcie t^i. Tak więc, koszt informacyjny obliczenia $\tilde{m}^i(f)$ wynosi

$$\sum_{k=0}^{r} \binom{d+k-1}{k} = \frac{(d+r)!}{d!\, r!},$$

co jest niezależne od n. Zgodnie z (2.4) koszt obliczenia maksimum ciągu $\tilde{m}^1(f), \ldots, \tilde{m}^N(f)$ na komputerze kwantowym wynosi $O(\sqrt{N}) = O(\sqrt{n^d})$ odwołań do wyroczni odnoszącej się do ciągu $\tilde{m}^1(f), \ldots, \tilde{m}^N(f)$. Pokażemy jeszcze jak z wyroczni o ciągu $(\tilde{m}^i(f))$ przejść na wyrocznię o wartościach i pochodnych f. Zastosujemy lemat 4 do funkcji g danej przez (1.8). Funkcje $\eta_j(i)$ zwracają $(i_1^j, i_2^j, \ldots, i_d^j, t^i)$, gdzie $i_1^j, \ldots, i_d^j$ oznacza rząd pochodnej po odpowiednich zmiennych (zależy wyłącznie od j), a t^i jest środkiem kostki K^i. Funkcja $\Gamma(g)(i) = \tilde{m}_i(f) = \rho\left(i, \beta(g(i_1^0, i_2^0, \ldots, i_d^0, t^i)), \ldots, \beta(g(i_1^{M-1}, i_2^{M-1}, \ldots, i_d^{M-1}, t^i))\right)$, gdzie $M = \dfrac{(d+r)!}{d!r!}$ jest kosztem deterministycznym wyznaczenia wielomianu w^i. Funkcja β jest funkcją kodującą liczby rzeczywiste na liczby całkowite (oczywiście z pewną dokładnością, zakładamy jednak, że rejestr kwantowy składa się z wystarczającej liczby qubitów by błąd tego kodowania można było zaniedbać). Zgodnie z lematem 4, koszt całkowity w terminach odwołań do wyroczni odnoszącej się do f i jej pochodnych jest równy dwukrotności iloczynu liczby odwołań do wyroczni odnoszącej się do ciągu $(\tilde{m}^i(f))$ i kosztu deterministycznego takiej wyroczni. Zatem, całkowity koszt jest szacowany przez

$$\operatorname{cost}(S, \phi^*, F_d^{r,\rho}) \leqslant C_2\, n^{d/2}$$

dla pewnej stałej C_2 zależnej wyłącznie od parametrów klasy. □

Porównując górne oszacowania na błąd i koszt algorytmu ϕ^* otrzymujemy następujące oszacowania na złożoność problemu szukania maksimum funkcji z klasy Höldera na komputerze kwantowym.

Twierdzenie 15. *Istnieje dodatnia stała C zależna wyłącznie od parametrów klasy taka, że dla dowolnego $\varepsilon > 0$ złożoność problemu szukania maksimum funkcji z klasy Höldera w modelu kwantowym jest szacowana przez*

$$\operatorname{comp}^{\mathrm{quant}}(\varepsilon, S, F_d^{r,\rho}) \leqslant C \left(\frac{1}{\varepsilon}\right)^{d/(2(r+\rho))}.$$

Dowód. Zgodnie z twierdzeniem 14, aby błąd algorytmu ϕ^* był nie większy od ε wystarczy przyjąć $n = \left\lceil (C_1/\varepsilon)^{1/(r+\rho)} \right\rceil$. Koszt algorytmu ϕ^* dla tak wybranego n jest szacowany przez

$$\operatorname{cost}(\phi^*, f) \leqslant C_2 \left\lceil \left(\frac{C_1}{\varepsilon}\right)^{1/(r+\rho)} \right\rceil^{d/2} \leqslant C \left(\frac{1}{\varepsilon}\right)^{d/(2(r+\rho))},$$

dla stałej C zależnej wyłącznie od parametrów klasy. □

W następnym podrozdziale przedstawimy dolne ograniczenia na złożoność problemu szukania maksimum.

3.4. Dolne ograniczenia w modelu kwantowym

W poprzednim podrozdziale skonstruowaliśmy algorytm rozwiązujący problem poszukiwania maksimum funkcji z klasy Höldera na komputerze kwantowym. Pozostaje pytanie czy algorytm ten jest optymalny. W celu udzielenia odpowiedzi na to pytanie, znajdziemy dolne ograniczenie na złożoność tego problemu i porównamy go z górnym ograniczeniem wyznaczonym przez algorytm z poprzedniego podrozdziału.

Następujące twierdzenie przedstawia dolne ograniczenie na złożoność problemu szukania maksimum funkcji z klasy Höldera na komputerze kwantowym.

Twierdzenie 16. *Istnieją dodatnie stałe $\tilde{C}$ i ε_0 zależne wyłącznie od parametrów klasy takie, że dla dowolnego $\varepsilon \in (0, \varepsilon_0)$ złożoność problemu szukania maksimum funkcji z klasy Höldera w modelu kwantowym szacuje się z dołu przez*

$$\mathrm{comp}^{\mathrm{quant}}(\varepsilon, S, F_d^{r,\rho}) \geqslant \tilde{C} \left(\frac{1}{\varepsilon}\right)^{d/(2(r+\rho))}$$

Dowód. W celu udowodnienia ograniczenia z dołu pokażemy, że dowolny algorytm rozwiązujący problem szukania maksimum funkcji z klasy Höldera rozwiązuje również problem maksymalizacji ciągu binarnego (czyli obliczenia wartości logicznej alternatywy ciągu binarnego). Załóżmy, że $\tilde{\phi}$ jest dowolnym algorytmem obliczającym $S(f) = \max_{t\in[0,1]^d} f(t)$ dla dowolnej funkcji $f \in F_d^{r,\rho}$ z błędem nie przekraczającym $\varepsilon > 0$ z prawdopodobieństwem nie mniejszym niż $3/4$. Niech koszt tego algorytmu będzie równy $c(\varepsilon)$ odwołań do wyroczni dla f. Dla dostatecznie małego $\varepsilon_1 > 0$, klasa $F_d^{r,\rho}$ zawiera $n = \Theta\left(\varepsilon_1^{-d/(r+\rho)}\right)$ funkcji $f_1, \ldots, f_n$ o rozłącznych nośnikach takich, że $\max_{t\in[0,1]^d} f_i(t) = \varepsilon_1$ oraz $|D^{(r)}f_i(x) - D^{(r)}f_i(y)| \leqslant H/2\|x-y\|^\rho$, $\forall x, y \in [0,1]^d$ (zobacz lemat 8 w dodatku A).

Niech ε będzie dostatecznie małe i $\varepsilon_1 = 4\varepsilon$. Niech $X = (x_1, \ldots, x_n)$ będzie dowolnym ciągiem takim, że $x_i \in \{0,1\}, \quad i = 1, 2, \ldots, n$. Wtedy funkcja

$$f_{\varepsilon_1} := \sum_{i=1}^{n} x_i f_i$$

należy do klasy $F_d^{r,\rho}$. Istotnie, jeśli $x \in K_i$, $y \in K_j$, $i \neq j$, gdzie K_i, K_j są nośnikami funkcji f_i, f_j, to $|D^{(r)}f_{\varepsilon_1}(x) - D^{(r)}f_{\varepsilon_1}(y)| = |x_i D^{(r)}f_i(x) - x_j D^{(r)}f_j(y)| \leqslant x_i|D^{(r)}f_i(x) - D^{(r)}f_i(y)| + x_j|D^{(r)}f_j(x) - D^{(r)}f_j(y)| \leqslant H\|x-y\|^\rho$. Pozostałe warunki

też oczywiście zachodzą. Tak więc, algorytm $\tilde{\phi}$ zastosowany do f_{ε_1} oblicza przybliżenie $\max_{t\in[0,1]^d} f_{\varepsilon_1}(f)$ z kosztem $c(\varepsilon) = c(\varepsilon_1/4)$ takie, że

$$|\max_{t\in[0,1]^d} f_{\varepsilon_1}(t) - \tilde{\phi}(f_{\varepsilon_1})| \leqslant \varepsilon = \frac{\varepsilon_1}{4} \tag{3.5}$$

z prawdopodobieństwem nie mniejszym niż 3/4.
Z definicji funkcji f_{ε_1} widać, że

$$\max_{t\in[0,1]^d} f_{\varepsilon_1}(t) = \begin{cases} \varepsilon_1 & \text{wtedy i tylko wtedy, gdy} \quad \max_{i=1,\ldots,n} x_i = 1 \\ 0 & \text{wtedy i tylko wtedy, gdy} \quad \max_{i=1,\ldots,n} x_i = 0 \end{cases} . \tag{3.6}$$

Jeśli $3/4\varepsilon_1 \leqslant \tilde{\phi}(f_{\varepsilon_1}) \leqslant 5/4\varepsilon_1$, to zgodnie z (3.5)

$$\max_{t\in[0,1]^d} f_{\varepsilon_1}(t) \geqslant \tilde{\phi}(f_{\varepsilon_1}) - 1/4\varepsilon_1 \geqslant 1/2\varepsilon_1.$$

Tak więc, w tym przypadku, zgodnie z (3.6) mamy $\max_{t\in[0,1]^d} f_{\varepsilon_1}(t) = \varepsilon_1$ i $\max_{i=1,\ldots,n} x_i = 1$ z prawdopodobieństwem nie mniejszym niż 3/4. Podobnie, z prawdopodobieństwem przynajmniej 3/4, jeśli $-1/4\varepsilon_1 \leqslant \tilde{\phi}(f_{\varepsilon_1}) \leqslant 1/4\varepsilon_1$, to $\max_{t\in[0,1]^d} f_{\varepsilon_1}(t) = 0$ i $\max_{i=1,\ldots,n} x_i = 0$.
Bazując na algorytmie $\tilde{\phi}$, zdefiniujemy teraz algorytm $\tilde{\tilde{\phi}}$, który znajduje maksimum ciągu $X = (x_1, \ldots, x_n)$. Algorytm $\tilde{\tilde{\phi}}$ skonstruowany jest następująco:

jeśli $3/4\varepsilon_1 \leqslant \tilde{\phi}(f_{\varepsilon_1}) \leqslant 5/4\varepsilon_1$, to kładziemy $\tilde{\tilde{\phi}}(X) = 1$,
jeśli $-1/4\varepsilon_1 \leqslant \tilde{\phi}(f_{\varepsilon_1}) \leqslant 1/4\varepsilon_1$, to kładziemy $\tilde{\tilde{\phi}}(X) = 0$.

W pozostałych przypadkach kładziemy $\tilde{\tilde{\phi}}(X) = 0$. Zgodnie z tą definicją oraz własnościami maksimum funkcji f_{ε_1} mamy następujące własności algorytmu $\tilde{\tilde{\phi}}$:

jeśli mamy $\max_i x_i = 0$, to $\tilde{\tilde{\phi}}(X) = 0$ z prawdopodobieństwem co najmniej 3/4,
jeśli mamy $\max_i x_i = 1$, to $\tilde{\tilde{\phi}}(X) = 1$ z prawdopodobieństwem co najmniej 3/4.

Stąd, algorytm $\tilde{\tilde{\phi}}$ oblicza maksimum z n liczb $x_1, \ldots, x_n$ takich, że $x_i \in \{0, 1\}$ (logiczna alternatywa), z prawdopodobieństwem nie mniejszym niż 3/4. Z twierdzenia 5 wiemy, że koszt takiego algorytmu jest $\Omega(\sqrt{n})$ odwołań do wyroczni o ciągu (x_i), czyli do funkcji binarnej $h : \{1, 2, \ldots n\} \ni i \mapsto h(i) = x_i$. Zgodnie z lematem 4, wyrocznię o wartościach i pochodnych f_{ε_1} możemy symulować przez wyrocznię o ciągu (x_i), więc koszt algorytmu $\tilde{\tilde{\phi}}$ czyli $c(\varepsilon)$ (mierzony liczbą odwołań do wyroczni o f_{ε_1}) jest nie mniejszy niż konieczna liczba odwołań do wyroczni o ciągu (x_i), czyli jest on $\Omega(\sqrt{n})$. Pokażemy jak tutaj zdefiniować odpowiednie funkcje w lemacie 4. Niech $\eta_0(i_1, i_2, \ldots, i_d, t) = i$ dla $t \in K_i$, gdzie K_i jest nośnikiem funkcji f_i. Dla

$t \in K_i$ definiujemy $\Gamma(h)(i_1, i_2, \ldots, i_d, t) = \rho(i_1, i_2, \ldots, i_d, t, h(\eta_0(i_1, i_2, \ldots, i_d, t)) = \rho(i_1, i_2, \ldots, i_d, t, x_i) = x_i \frac{\partial^{i_1+i_2+\ldots,i_d} f_i}{\partial t_1^{i_1} \partial t_2^{i+2} \ldots \partial t_d^{i_d}}(t) = \frac{\partial^{i_1+i_2+\ldots,i_d} f_{\varepsilon_1}}{\partial t_1^{i_1} \partial t_2^{i+2} \ldots \partial t_d^{i_d}}(t)$. Jako, że

$$n = \Theta\left((1/\varepsilon_1)^{\frac{d}{r+\rho}}\right) = \Theta\left((1/\varepsilon)^{\frac{d}{r+\rho}}\right),$$

dostajemy z lematu 4

$$c(\varepsilon) = \Omega\left(\sqrt{n}\right) = \Omega\left((1/\varepsilon)^{d/(2(r+\rho))}\right).$$

To kończy dowód twierdzenia. □

3.5. Podsumowanie

W rozdziale tym rozpatrywaliśmy złożoność zadania znajdowania maksimum funkcji z klasy Höldera. Skonstruowaliśmy algorytm kwantowy rozwiązujący ten problem. Znaleźliśmy górne ograniczenia na błąd oraz koszt tego algorytmu, dzięki czemu uzyskaliśmy górne ograniczenie na złożoność. Wyznaczyliśmy również dolne ograniczenie na złożoność tego problemu w modelu kwantowym. Otrzymane górne i dolne ograniczenia na złożoność okazały się być zgodne (równe z dokładnością do stałej multiplikatywnej), więc skonstruowany przez nas algorytm jest optymalny.

Porównamy uzyskane wyniki ze znanymi rezultatami dla modelu deterministycznego i randomizacyjnego. Dolne i górne ograniczenia na złożoność w modelu deterministycznym, randomizacyjnym i kwantowym przedstawia tabela 3.1 (dla większej przejrzystości pominięto stałe).

model	ograniczenia z dołu	ograniczenia z góry
det	$\varepsilon^{-\frac{d}{r+\rho}}$	$\varepsilon^{-\frac{d}{r+\rho}}$
rand	$\varepsilon^{-\frac{d}{r+\rho}}$	$\varepsilon^{-\frac{d}{r+\rho}}$
quant	$\varepsilon^{-\frac{d}{2(r+\rho)}}$	$\varepsilon^{-\frac{d}{2(r+\rho)}}$

Tabela 3.1. Ograniczenia na złożoność

Złożoność w modelu randomizacyjnym jest tego samego rzędu co w modelu deterministycznym. Z wyników zaprezentowanych w tabeli 3.1 wynika, że na komputerze kwantowym uzyskuje się kwadratowe przyspieszenie w stosunku do modelu deterministycznego i randomizacyjnego w całym obszarze wartości d, r, ρ. W modelu kwantowym, podobnie jak w modelach deterministycznym i randomizacyjnym występuje

wykładnicza zależność od wymiaru d (efekt ten zwany jest „przekleństwem wymiaru"), jednak zależność ta jest słabsza w porównaniu z modelami deterministycznym i randomizacyjnym.

Rozdział 4

Równania nieliniowe

4.1. Wstęp

W tym rozdziale zajmiemy się problemem rozwiązywania równań nieliniowych. Równania nieliniowe pojawiają się przy rozwiązywaniu wielu problemów z matematyki stosowanej, fizyki czy też astronomii, począwszy od prostych problemów takich jak szukanie pierwiastka liczby rzeczywistej do rozwiązywania równania Keplera służącego do wyznaczania przyszłego położenia planet. Rozwiązywanie równań nieliniowych potrzebne będzie również przy innych problemach rozważanych w tej rozprawie, między innymi przy rozwiązywaniu problemów brzegowych dla równań różniczkowych.

Skoncentrujemy się tutaj na znajdowaniu zer funkcji d zmiennych należących do klasy Höldera. Poszukamy ograniczeń na złożoność tego problemu w modelu randomizacyjnym i kwantowym. Dla porównania wyznaczymy również ograniczenia dla modelu deterministycznego.

W celu wyznaczenia górnych ograniczeń na złożoność skonstruujemy trzy algorytmy rozwiązujące ten problem w modelach: deterministycznym, randomizacyjnym i kwantowym. Algorytmy te opierać się będą na przybliżaniu funkcji wielomianami oraz wykorzystywać będą algorytmy przeszukiwania dyskretnego ciągu.

Dolne ograniczenia na złożoność uzyskamy poprzez sprowadzenie problemu dyskretnego przeszukiwania do problemu rozwiązywania równań nieliniowych.
Uzyskane górne i dolne ograniczenia na złożoność we wszystkich modelach są zgodne (z dokładnością do stałej multiplikatywnej). Wynika z tego, że konstruowane przez nas algorytmy są optymalne.
Pokażemy, że złożoność w modelu randomizacyjnym jest takiego samego rzędu jak złożoność w modelu deterministycznym, natomiast komputery kwantowe dają kwadratowe przyspieszenie dla tego problemu.

W następnym podrozdziale sformułujemy problem oraz przedstawimy znane wyniki dla tego problemu. W podrozdziale 4.3 przedstawimy górne ograniczenia na złożoność, a w podrozdziale 4.4 wyprowadzimy ograniczenia z dołu.

4.2. Sformułowanie problemu i znane wyniki

W rozdziale tym zajmiemy się rozwiązywaniem równań nieliniowych. Niech $f : [0,1]^d \to \mathbb{R}$ będzie pewną funkcją o wartościach rzeczywistych. Załóżmy, że istnieje taki punkt $t^0 \in [0,1]^d$, że $f(t^0) = 0$. Naszym celem jest znalezienie przybliżonego rozwiązania równania $f(t) = 0$ z dokładnością ε. Rozważać będziemy klasę Höldera funkcji f, czyli klasę zdefiniowaną jako

$$F_d^{r,\rho} = \{f : [0,1]^d \to \mathbb{R} |\ f \in C^r([0,1]^d), |D^{(i)}f(x)| \leqslant D,\ i = 0,1,\dots r,$$
$$|D^{(r)}f(x) - D^{(r)}f(y)| \leqslant H\|x-y\|^\rho\ \forall x,y \in [0,1]^d\},$$

gdzie $r \geqslant 0$, $0 < \rho \leqslant 1$, D i H są dodatnimi stałymi, a $D^{(i)}$ przebiega zbiór wszystkich pochodnych cząstkowych rzędu i. Jak poprzednio, jako normę przyjmujemy tutaj normę nieskończoność, $\|\cdot\| = \|\cdot\|_\infty$, definiowaną jako maksymalna wartość bezwzględna współrzędnych wektora. Zakładamy, że funkcja f ma przynajmniej jedno zero. Klasę takich funkcji będziemy oznaczać przez $F_d^{r,\rho}(0) = \{f \in F_d^{r,\rho} |\ \exists t^0 \in [0,1]^d : f(t^0) = 0\}$. Niech $\varepsilon > 0$. Punkt t^* uważamy za rozwiązanie, jeśli zachodzi kryterium residualne:

$$|f(t^*)| \leqslant \varepsilon.$$

Będziemy używać uogólnionego operatora rozwiązania W (zob. podrozdział 1.1). Niech W będzie uogólnionym operatorem rozwiązania dla tego problemu.

$$W(f,\varepsilon) = \{t \in \mathbb{R}^d : |f(t)| \leqslant \varepsilon\}.$$

Naszym celem jest znalezienie punktu należącego do zbioru $W(f,\varepsilon)$. Rozważać będziemy trzy modele obliczeniowe: model deterministyczny, randomizacyjny i kwantowy. W modelu deterministycznym dopuszczać będziemy wyłącznie informację standardową, czyli informację składającą się z wartości funkcji f lub jej pochodnych cząstkowych do rzędu r w pewnych punktach. Definicje algorytmu, informacji, kosztu i złożoności przedstawione są w rozdziale 1. Przedstawimy tylko definicje błędu dla tego problemu, ponieważ różni się ona nieco od definicji wprowadzonej dla problemów definiowanych przez operator rozwiązania.
Załóżmy, że wynikiem algorytmu ϕ dla funkcji $f \in F_d^{r,\rho}(0)$ jest punkt $t^* \in \mathbb{R}^d$. Błąd lokalny algorytmu ϕ w kryterium residualnym definiujemy jako

$$\mathrm{e}_{\mathrm{res}}(W,\phi,f) = |f(t^*)|.$$

Globalny błąd w klasie $F_d^{r,\rho}(0)$ w modelu deterministycznym definiowany jest jako

$$\mathrm{e}^{\mathrm{det}}(W,\phi,F_d^{r,\rho}(0)) = \sup_{f\in F_d^{r,\rho}(0)} \mathrm{e}_{\mathrm{res}}(W,\phi,f).$$

Błąd globalny w klasie $F_d^{r,\rho}(0)$ w modelu randomizacyjnym dany jest przez

$$\mathrm{e}^{\mathrm{rand}}(W,\phi,F_d^{r,\rho}(0)) = \sup_{f\in F_d^{r,\rho}(0)} \left(\int_\Omega (\mathrm{e}_{\mathrm{res}}(W,\phi,f))^2 d\mathbf{P}(\omega)\right)^{1/2}. \tag{4.1}$$

Błąd w modelu kwantowym definiowany jest jako

$$\mathrm{e}^{\mathrm{quant}}(W,\phi,F_d^{r,\rho}(0)) = \sup_{f\in F_d^{r,\rho}(0)} \inf\{\alpha : \mathbf{P}(\mathrm{e}_{\mathrm{res}}(W,\phi,f) > \alpha) \leqslant 1/4\}.$$

Znane są pewne wyniki dotyczące tego problemu. W pracy [Sik83], Sikorski rozważał problem poszukiwani zer funkcji z klasy Lipschitza, czyli klasy $F_d^{0,1}(0)$, w modelu deterministycznym. Przedstawił on optymalny algorytm rozwiązujący ten problem oraz wyznaczył ścisłe ograniczenia na złożoność tego problemu rzędu

$$\mathrm{comp}^{\mathrm{det}}(\varepsilon,W,F_d^{0,1}(0)) = \Theta\left(\left(\frac{1}{\varepsilon}\right)^d\right).$$

Celem tego rozdziału jest wyznaczenie ograniczeń na złożoność tego problemu w modelu randomizacyjnym oraz kwantowym oraz wyznaczenie optymalnych algorytmów dla ogólnej klasy Höldera $F_d^{r,\rho}(0)$. Dla porównania wyznaczymy również ograniczenia na złożoność w modelu deterministycznym.

4.3. Górne ograniczenia

Rozważmy najpierw klasę funkcji $F_d^{0,\rho}(0)$. Górne ograniczenia na złożoność w tej klasie mogą być wyznaczone przez zastosowanie algorytmu poszukiwania minimum funkcji. Niech f należy do klasy $F_d^{0,\rho}(0)$ oraz niech $g(x) = |f(x)|$. Zauważmy, że dla dowolnych $x, y \in [0,1]^d$ mamy

$$|g(x) - g(y)| = \Big||f(x)| - |f(y)|\Big| \leqslant |f(x) - f(y)| \leqslant H\|x - y\|^\rho.$$

Wynika z tego, że funkcja g również należy do klasy $F_d^{0,\rho}(0)$. Ponadto funkcja g ma minimum w punkcie, w którym funkcja f przyjmuje wartość 0. Korzystając z ograniczeń na złożoność problemu szukania minimum (który jest równoważny problemowi szukania maksimum) z rozdziału 3 otrzymujemy następujący wniosek

Wniosek 2. *Istnieją stałe C_0^{det}, C_0^{rand}, C_0^{quant} takie, że dla dowolnego $\varepsilon > 0$ zachodzi:*

$$\mathrm{comp}^{\mathrm{det}}(\varepsilon, W, F_d^{0,\rho}(0)) \leqslant C_0^{\mathrm{det}} \left(\frac{1}{\varepsilon}\right)^{d/\rho},$$

$$\mathrm{comp}^{\mathrm{rand}}(\varepsilon, W, F_d^{0,\rho}(0)) \leqslant C_0^{\mathrm{rand}} \left(\frac{1}{\varepsilon}\right)^{d/\rho},$$

$$\mathrm{comp}^{\mathrm{quant}}(\varepsilon, W, F_d^{0,\rho}(0)) \leqslant C_0^{\mathrm{quant}} \left(\frac{1}{\varepsilon}\right)^{d/(2\rho)}.$$

Rozważmy teraz ogólną klasę Höldera $F_d^{r,\rho}(0)$ z $r \geqslant 1$. W tej klasie nie możemy użyć bezpośrednio algorytmu minimalizacji, ponieważ funkcja $g(x) = |f(x)|$ nie musi należeć do klasy Höldera $F_d^{r,\rho}$. Z tego powodu skonstruujemy nowe algorytmy poszukiwania pierwiastków funkcji z klasy Höldera w modelach deterministycznym, randomizacyjnym i kwantowym.

Niech $\varepsilon > 0$ i $n \in \mathbb{N}$ będą parametrami algorytmu. Algorytmy $\phi^{\mathrm{det}}(n, \varepsilon)$, $\phi^{\mathrm{rand}}(n, \varepsilon)$, $\phi^{\mathrm{quant}}(n, \varepsilon)$ zdefiniowane są następująco:

1. Dzielimy kostkę $[0,1]^d$ równomiernie na n^d kostek K^i, $i = 1, 2, \ldots, n^d$, o krawędziach długości $1/n$ i o rozłącznych wnętrzach. Oznaczmy środek kostki K^i przez t^i.
2. Na każdej kostce K^i stosujemy rozwinięcie Taylora w^i funkcji f w punkcie t^i (zob. dodatek A). Dla $t \in K^i$ mamy

   $$f(t) = w^i(t) + R^i(t, t^i),$$

   gdzie

   $$w^i(t) = \sum_{k=0}^{r} \frac{1}{k!} f^{(k)}(t^i)(t - t^i)^k,$$

$$R^i(t,t^i) = \int_0^1 \left(f^{(r)}(\theta t + (1-\theta)t^i) - f^{(r)}(t^i)\right)(t-t^i)^r \frac{(1-\theta)^{r-1}}{(r-1)!}d\theta.$$

3. Niech $m^i(f) = \min_{t\in K^i} |w^i(t)|$, $i = 1, 2, \ldots, n^d$ (pewne nieznane liczby). Niech $\tilde{t}^i \in K^i$ będzie punktem takim, że

   $$\left|m^i(f) - |w^i(\tilde{t}^i)|\right| \leqslant \varepsilon/3$$

   wyznaczonym za pomocą pewnego klasycznego algorytmu. Niech $\tilde{m}^i(f) = |w^i(\tilde{t}^i)|$.
4. Zdefiniujmy funkcję binarną

   $$g(i) = \begin{cases} 1 & \text{gdy} \quad \tilde{m}^i(f) \leqslant 2\varepsilon/3 \\ 0 & \text{gdy} \quad \tilde{m}^i(f) > 2\varepsilon/3 \end{cases}.$$

   Niech i^* będzie takie, że $g(i^*) = 1$. Wykażemy w dalszej części rozdziału, że taki punkt i^* istnieje.

   Następnie wyznaczamy i^* za pomocą:
   - w modelu deterministycznym – optymalnego deterministycznego algorytmu przeszukiwania dyskretnego;
   - w modelu randomizacyjnym – optymalnego randomizacyjnego algorytmu przeszukiwania dyskretnego z prawdopodobieństwem porażki nie większym niż $\delta^{\text{rand}} = 1/n^{2(r+\rho)}$;
   - w modelu kwantowym – algorytm Grovera przeszukiwania z nieznaną liczbą rozwiązań z prawdopodobieństwem porażki nie większym niż $\delta^{\text{quant}} = 1/4$.
5. Algorytm jest określony przez $\phi^{\#}(n,\varepsilon) = t^* = \tilde{t}^{i^*}$, $\# \in \{\text{det}, \text{rand}, \text{quant}\}$.

Różnica pomiędzy algorytmami ϕ^{det}, ϕ^{rand} i ϕ^{quant} polega na sposobie wyznaczenia i^* takiego, że $g(i^*) = 1$.

Przeprowadzimy teraz analizę tych algorytmów. Na początek pokażemy, że algorytmy te są dobrze określone oraz, że dają dobre przybliżenia poszukiwanego rozwiązania.

Twierdzenie 17. *Niech $C_{r,\rho} = 1/r!d^r(1/2)^{r+\rho}H$. Jeśli $n \geqslant (3C_{r,\rho}/\varepsilon)^{1/(r+\rho)}$, to funkcja $g(i)$ w definicji algorytmów $\phi^{\text{det}}(n,\varepsilon)$, $\phi^{\text{quant}}(n,\varepsilon)$, $\phi^{\text{quant}}(n,\varepsilon)$ przyjmuje wartość 1 dla pewnego i. Ponadto, algorytm $\phi^{\text{det}}(n,\varepsilon)$ zwraca punkt t^* taki, że $|f(t^*)| \leqslant \varepsilon$ na pewno, a algorytmy $\phi^{\text{rand/quant}}(n,\varepsilon)$ zwracają punkty t^* takie, że $|f(t^*)| \leqslant \varepsilon$ z prawdopodobieństwem co najmniej $1 - \delta^{\text{rand/quant}}$.*

Dowód. Dla $t \in K^i$, korzystając z warunku Höldera, mamy

$$\begin{aligned}
|f(t) - w^i(t)| &= |R(t, t^i)| \\
&= \left| \int_0^1 \left(f^{(r)}(\theta t + (1-\theta)t^i) - f^{(r)}(t^i) \right) (t - t^i)^r \frac{(1-\theta)^{r-1}}{(r-1)!} d\theta \right| \\
&\leqslant \int_0^1 \left\| f^{(r)}(\theta t + (1-\theta)t^i) - f^{(r)}(t^i) \right\| \|t - t^i\|^r \frac{(1-\theta)^{r-1}}{(r-1)!} d\theta \\
&\leqslant \int_0^1 d^r H \|\theta(t - t^i)\|^\rho \|t - t^i\|^r \frac{(1-\theta)^{r-1}}{(r-1)!} d\theta \\
&\leqslant d^r H \|t - t^i\|^{r+\rho} \int_0^1 \frac{(1-\theta)^{r-1}}{(r-1)!} d\theta \\
&\leqslant d^r H (1/2)^{r+\rho} (1/n)^{r+\rho} (1/r!) = C_{r,\rho} (1/n)^{r+\rho}. \qquad (4.2)
\end{aligned}$$

Niech $t^0 \in [0,1]^d$ będzie punktem, w którym $f(t^0) = 0$. Niech K^{i^0} będzie kostką zawierającą t^0. Wtedy, korzystając z (4.2) dostajemy

$$|w^{i^0}(t^0)| = |w^{i^0}(t^0) - f(t^0)| \leqslant C_{r,\rho}(1/n)^{r+\rho}.$$

Stąd, $m^{i^0}(f) \leqslant C_{r,\rho}(1/n)^{r+\rho}$. Ponadto, korzystając z warunku na $\tilde{t}^i$ w definicji algorytmów dostajemy

$$\tilde{m}^{i^0}(f) \leqslant |\tilde{m}^{i^0}(f) - m^{i^0}(f)| + m^{i^0}(f) \leqslant \varepsilon/3 + C_{r,\rho}(1/n)^{r+\rho} \leqslant 2\varepsilon/3. \qquad (4.3)$$

Tak więc, zgodnie z definicją funkcji g mamy $g(i^0) = 1$. Algorytmy są więc dobrze określone.
Niech $t^* = \tilde{t}^{i^*}$ będzie punktem zwróconym przez algorytm $\phi^{\text{det/rand/quant}}$. Jako, że $g(i^*) = 1$ (z prawdopodobieństwem nie mniejszym niż $1 - \delta^{\text{rand/quant}}$ w modelach randomizacyjnym i kwantowym), więc $|w^{i^*}(\tilde{t}^{i^*})| = \tilde{m}^{i^*}(f) \leqslant 2\varepsilon/3$ (z tym samym prawdopodobieństwem). Stąd

$$|f(t^*)| \leqslant |f(t^*) - w^{i^*}(t^*)| + |w^{i^*}(t^*)| \leqslant C_{r,\rho}(1/n)^{r+\rho} + 2\varepsilon/3 \leqslant \varepsilon. \qquad (4.4)$$

W modelu randomizacyjnym i kwantowym zachodzi to z prawdopodobieństwem nie mniejszym niż $1 - \delta^{\text{rand/quant}}$. To kończy dowód twierdzenia. □
Następujące twierdzenie przedstawia górne ograniczenia na złożoność problemu rozwiązywania równań nieliniowych.

Twierdzenie 18. *Istnieją dodatnie stałe C^{det}, C^{rand} i C^{quant} takie, że dla dowolnego $\varepsilon > 0$ złożoność problemu znajdowania pierwiastków funkcji z klasy Höldera spełnia*

$$\text{comp}^{\text{det}}(\varepsilon, W, F_d^{r,\rho}(0)) \leqslant C^{\text{det}} \left(\frac{1}{\varepsilon}\right)^{d/(r+\rho)},$$

$$\mathrm{comp}^{\mathrm{rand}}(\varepsilon, W, F_d^{r,\rho}(0)) \leqslant C^{\mathrm{rand}} \left(\frac{1}{\varepsilon}\right)^{d/(r+\rho)},$$

$$\mathrm{comp}^{\mathrm{quant}}(\varepsilon, W, F_d^{r,\rho}(0)) \leqslant C^{\mathrm{quant}} \left(\frac{1}{\varepsilon}\right)^{d/(2(r+\rho))}.$$

Dowód. Przyjmijmy $n = \left\lceil (3C_{r,\rho}/\varepsilon)^{1/(r+\rho)} \right\rceil$. Zgodnie z twierdzeniem 17

$$\mathrm{e}_{\mathrm{res}}(W, \phi^{\mathrm{det/rand/quant}}, f) = |f(t^*)| \leqslant \varepsilon. \tag{4.5}$$

W przypadku randomizacyjnym zachodzi to z prawdopodobieństwem przynajmniej

$$1 - \delta^{\mathrm{rand}} = 1 - 1/n^{2(r+\rho)} \geqslant 1 - \left(\varepsilon/(3C_{r,\rho})\right)^2, \tag{4.6}$$

natomiast w modelu kwantowym z prawdopodobieństwem przynajmniej $1 - \delta^{\mathrm{quant}} = 3/4$. Z (4.5), w modelach deterministycznym i kwantowym uzyskujemy bezpośrednio ograniczenia na błąd

$$\mathrm{e}^{\mathrm{det}}(W, \phi^{\mathrm{det}}, F_d^{r,\rho}(0)) \leqslant \varepsilon,$$

$$\mathrm{e}^{\mathrm{quant}}(W, \phi^{\mathrm{quant}}, F_d^{r,\rho}(0)) \leqslant \varepsilon.$$

W modelu randomizacyjnym musimy przejść z ograniczeń na błąd probabilistyczny do ograniczeń na błąd zdefiniowany przez (4.1). Zauważmy, że dla dowolnej $f \in F_d^{r,\rho}(0)$ zachodzi $\mathrm{e}_{\mathrm{res}}(W, \phi^{\mathrm{quant}}, f) \leqslant D$ (jako, że wynikiem zawsze jest punkt z kostki $[0,1]^d$). Korzystając z (4.5) oraz (4.6) dostajemy dla dowolnej $f \in F_d^{r,\rho}(0)$

$$\begin{aligned}\left(\mathrm{e}^{\mathrm{rand}}(W, \phi^{\mathrm{rand}}, f)\right)^2 &= \int\limits_{\mathrm{e}_{\mathrm{res}}(W,\phi^{\mathrm{rand}},f)\leqslant\varepsilon} (\mathrm{e}_{\mathrm{res}}(W, \phi^{\mathrm{rand}}, f))^2 d\mathbf{P}(\omega) \qquad (4.7)\\ &\quad + \int\limits_{\mathrm{e}_{\mathrm{res}}(W,\phi^{\mathrm{rand}},f)>\varepsilon} (\mathrm{e}_{\mathrm{res}}(W, \phi^{\mathrm{rand}}, f))^2 d\mathbf{P}(\omega)\\ &\leqslant \varepsilon^2 + \delta^{\mathrm{rand}} D^2 \leqslant \left(1 + D^2/(9C_{r,\rho}^2)\right)\varepsilon^2, \qquad (4.8)\end{aligned}$$

a więc

$$\mathrm{e}^{\mathrm{rand}}(W, \phi^{\mathrm{rand}}, F_d^{r,\rho}(0)) \leqslant \sqrt{1 + D^2/(9C_{r,\rho}^2)}\,\varepsilon. \tag{4.9}$$

Wyznaczymy teraz ograniczenia na koszt algorytmów. Koszt wyznaczenia każdego wielomianu Taylora w^i jest kosztem obliczenia wszystkich pochodnych cząstkowych funkcji f rzędu od 0 do r w pewnych punktach. Wyznaczenie $\tilde{t}^i$ i $\tilde{m}^i(f)$ nie tworzy dodatkowego kosztu informacyjnego. Oczywiście, generuje pewien koszt kombinatoryczny, czyli koszt działań arytmetycznych. Zgodnie z [Ren87, Ren89] istnieje algorytm przybliżający minimum wielomianu z dokładnością ε o koszcie kombinatorycznym rzędu $O(\log\log 1/\varepsilon)$

działań arytmetycznych. Oznaczmy koszt informacyjny obliczania $\tilde{t}^i$ oraz $\tilde{m}^i(f)$ przez $\mathrm{cost}(\tilde{t}^i, \tilde{m}^i(f))$. Koszt ten dla każdego i wynosi

$$\mathrm{cost}(\tilde{t}^i, \tilde{m}^i(f)) = \sum_{k=0}^{r} \binom{d+k-1}{k} = \frac{(d+r)!}{d!\, r!}, \tag{4.10}$$

co stanowi liczbę pochodnych cząstkowych rzędu od 0 do r. Jest to niezależne od n i ε.

Oznaczmy operator rozwiązania dla problemu przeszukiwania dyskretnego przez S. W modelu deterministycznym, całkowity koszt informacyjny jest iloczynem kosztu obliczenia jednej pary $(\tilde{t}^i, \tilde{m}^i(f))$ oraz kosztu przeszukiwania dyskretnego, który zgodnie z (2.1) wynosi $\mathrm{cost}^{\mathrm{det}}(S, n^d) = n^d - 1$. Stąd

$$\mathrm{cost}^{\mathrm{det}}(W, \phi^{\mathrm{det}}, F_d^{r,\rho}(0)) = \mathrm{cost}(\tilde{t}^i, \tilde{m}^i(f))(n^d - 1) = O(n^d) = O\left(\varepsilon^{-d/(r+\rho)}\right). \tag{4.11}$$

W modelu randomizacyjnym, zgodnie z (2.2), istnieje algorytm wyznaczający punkt, w którym $g(i) = 1$ z prawdopodobieństwem przynajmniej $1 - \delta^{\mathrm{rand}}$ o koszcie $\mathrm{cost}^{\mathrm{rand}}(S, n^d) = O(n^d(1 - \delta^{\mathrm{rand}}))$ wartości $\tilde{m}^i(f)$. Tak więc, koszt całkowity w tym modelu wynosi

$$\mathrm{cost}^{\mathrm{rand}}(W, \phi^{\mathrm{rand}}, F_d^{r,\rho}(0)) = \mathrm{cost}(\tilde{t}^i, \tilde{m}^i(f))\mathrm{cost}^{\mathrm{rand}}(S, n^d) = O(n^d) = O\left(\varepsilon^{-d/(r+\rho)}\right). \tag{4.12}$$

W modelu kwantowym, algorytm Grovera, zgodnie z (2.3) i lematem 2, zwraca właściwy punkt z prawdopodobieństwem co najmniej $1-\delta^{\mathrm{quant}}$ używając $\mathrm{cost}^{\mathrm{quant}}(S, n^d) = O(\sqrt{n^d}\log(1/\delta^{\mathrm{quant}}))$ zapytań do wyroczni kwantowej odnośnie ciągu $(\tilde{m}^i(f))$. Stąd, zgodnie z lematem 4 (zastosowanym do funkcji g danej przez (1.8)) koszt całkowity jest ograniczony przez

$$\begin{aligned}\mathrm{cost}^{\mathrm{quant}}(W, \phi^{\mathrm{quant}}, F_d^{r,\rho}(0)) &= \mathrm{cost}(\tilde{t}^i, \tilde{m}^i(f))O(n^{d/2}) = O(n^{d/2}) \\ &= O\left(\varepsilon^{-d/(2(r+\rho))}\right),\end{aligned} \tag{4.13}$$

co kończy wyznaczanie ograniczeń na koszt.

W modelach deterministycznym i kwantowym błąd algorytmu ϕ^{det} i ϕ^{quant} jest ograniczony przez ε, więc wzory (4.11) i (4.13) wyznaczają górne ograniczenia na złożoność. W modelu randomizacyjnym, zgodnie z (4.9), aby błąd algorytmu ϕ^{rand} był nie większy niż $\varepsilon > 0$ wystarczy przyjąć dokładność $\varepsilon_1 = \varepsilon/\sqrt{1 + D^2/(9C_{r,\rho}^2)}$ w (4.9). Zgodnie z (4.12), koszt tego algorytmu jest ograniczony przez

$$\mathrm{cost}^{\mathrm{rand}}(W, \phi^{\mathrm{rand}}, F_d^{r,\rho}(0)) = O\left(\varepsilon_1^{-d/(r+\rho)}\right) = O\left(\varepsilon^{-d/(r+\rho)}\right),$$

co stanowi górne ograniczenie na ε–złożoność w modelu randomizacyjnym. □
W następny podrozdziale wyznaczymy dolne ograniczenia na złożoność. Dzięki temu będziemy mogli stwierdzić, czy algorytmy przedstawione w tej części są optymalne.

4.4. Dolne ograniczenia

W poprzednim podrozdziale skonstruowaliśmy algorytmy poszukujące zer funkcji z klasy Höldera w modelach deterministycznym, randomizacyjnym i kwantowym. Sprawdzimy teraz czy są one optymalne. W tym celu wyznaczymy dolne ograniczenia na złożoność tego problemu. Następujące twierdzenie podaje dolne ograniczenia na złożoność problemu poszukiwania zer funkcji z klasy Höldera.

Twierdzenie 19. *Istnieją dodatnie stałe c^{det}, c^{rand}, c^{quant} i ε_0 takie, że dla dowolnego $\varepsilon \in (0, \varepsilon_0)$ złożoność problemu poszukiwania zer funkcji z klasy Höldera spełnia*

$$\mathrm{comp}^{\mathrm{det}}(\varepsilon, W, F_d^{r,\rho}(0)) \geqslant c^{\mathrm{det}} \left(\frac{1}{\varepsilon}\right)^{d/(r+\rho)},$$

$$\mathrm{comp}^{\mathrm{rand}}(\varepsilon, W, F_d^{r,\rho}(0)) \geqslant c^{\mathrm{rand}} \left(\frac{1}{\varepsilon}\right)^{d/(r+\rho)},$$

$$\mathrm{comp}^{\mathrm{quant}}(\varepsilon, W, F_d^{r,\rho}(0)) \geqslant c^{\mathrm{quant}} \left(\frac{1}{\varepsilon}\right)^{d/(2(r+\rho))}.$$

Dowód. Załóżmy, że algorytmy $\tilde{\phi}^{\mathrm{det}}$, $\tilde{\phi}^{\mathrm{rand}}$, $\tilde{\phi}^{\mathrm{quant}}$ rozwiązują problem znajdowania zer funkcji z klasy Höldera z błędem odpowiednio $\mathrm{e}^{\mathrm{det/rand/quant}}(W, \tilde{\phi}^{\mathrm{det/rand/quant}}, F_d^{r,\rho}(0)) \leqslant \varepsilon$ i kosztem $c^{\mathrm{det/rand/quant}}(\varepsilon)$.
Dla dostatecznie małego $\varepsilon_1 > 0$, klasa $F_d^{r,\rho}$ zawiera $n = \Theta(\varepsilon_1^{-d/(r+\rho)})$ nieujemnych funkcji $f_1, f_2, \ldots, f_n$ o rozłącznych nośnikach, takich, że $\max_{t\in[0,1]^d} f_i(t) = \varepsilon_1$ (zobacz lemat 8 w dodatku A). Oznaczmy nośnik funkcji f_i przez K_i.
Niech ε będzie dostatecznie małe i $\varepsilon_1 = 4\varepsilon$. Niech $X = (x_1, x_2, \ldots, x_n) \in \{0,1\}^n$ będzie dowolnym ciągiem binarnym zawierającym jedną jedynkę. Wtedy funkcja

$$f_{\varepsilon_1} := \varepsilon_1 - \sum_{i=1}^{n} x_i f_i$$

należy do klasy $F_d^{r,\rho}(0)$. Zauważmy, że na zbiorze K_i takim, że $x_i = 1$ funkcja $f_{\varepsilon_1} = \varepsilon_1 - f_i$ ma pierwiastek. Natomiast na zbiorach K_i takich, że $x_i = 0$ funkcja f_{ε_1} ma wartość ε_1.

Rozważmy przypadek deterministyczny. Niech t^* będzie punktem zwróconym przez algorytm $\tilde{\phi}^{\mathrm{det}}$ dla funkcji f_{ε_1}. Niech K_{i^*} będzie zbiorem zawierającym t^*. Wtedy, $|f_{\varepsilon_1}(t^*)| \leqslant \varepsilon = 1/4\varepsilon_1$. Tak więc, zbiór K_{i^*} zawiera punkt, w którym wartość funkcji

f_{ε_1} jest nie większa niż $1/4\varepsilon_1$, skąd wynika, że $x_{i^*} = 1$. W takim razie algorytm $\tilde{\phi}^{\text{det}}$ znajduje również element i^* taki, że $x_{i^*} = 1$. Zgodnie z (2.1) koszt takiego algorytmu jest nie mniejszy niż $n - 1$ odwołań do elementów ciągu. Stąd, dla każdego algorytmu szukania zer funkcji

$$c^{\text{det}}(\varepsilon) = \Omega(n) = \Omega(\varepsilon_1^{-d/(r+\rho)}) = \Omega(\varepsilon^{-d/(r+\rho)}) \tag{4.14}$$

odwołań do wartości funkcji.

Rozważmy przypadek kwantowy. Niech t^* będzie punktem zwróconym przez algorytm $\tilde{\phi}^{\text{quant}}$ dla funkcji f_{ε_1}. Niech K_{i^*} będzie zbiorem zawierającym t^*. Podobnie do przypadku deterministycznego wnioskujemy, że $x_{i^*} = 1$ z prawdopodobieństwem nie mniejszym niż $3/4$. Zgodnie z (2.3), koszt takiego algorytm jest nie mniejszy niż $\Omega(\sqrt{n})$ odwołań do wyroczni o ciągu X. Korzystając z lematu 4 dostajemy więc

$$c^{\text{quant}}(\varepsilon) = \Omega(\sqrt{n}) = \Omega(\varepsilon_1^{-d/(2(r+\rho))}) = \Omega(\varepsilon^{-d/(2(r+\rho))}) \tag{4.15}$$

odwołań do funkcji.

Rozważmy przypadek randomizacyjny. Korzystając z nierówności Markowa (dla zmiennej losowej X, $\mathbf{P}(|X| \geqslant \alpha) \leqslant E(|X|^p)/\alpha^p$, dla $\alpha > 0$, $p > 0$) dostajemy

$$P(\mathrm{e}_{\text{res}}(W, \tilde{\phi}^{\text{rand}}, f_{\varepsilon_1}) \geqslant 2\varepsilon) \leqslant \frac{(\mathrm{e}^{\text{rand}}(W, \tilde{\phi}^{\text{rand}}, F_d^{r,\rho}(0)))^2}{(2\varepsilon)^2} \leqslant \frac{1}{4}.$$

Tak więc, algorytm $\tilde{\phi}^{\text{rand}}$ wyznacza pierwiastki funkcji f_{ε_1} z błędem residualnym nie większym niż 2ε z prawdopodobieństwem nie mniejszym niż $3/4$. Niech t^* będzie punktem zwróconym przez algorytm $\tilde{\phi}^{\text{rand}}$ dla funkcji f_{ε_1}. Niech K_{i^*} będzie zbiorem zawierającym t^*. Mamy, że $|f_{\varepsilon_1}(t^*)| \leqslant 2\varepsilon = 1/2\varepsilon_1$ z prawdopodobieństwem nie mniejszym niż $3/4$, skąd $x_{i^*} = 1$ z prawdopodobieństwem nie mniejszym niż $3/4$. Zgodnie z (2.2), koszt algorytmu $\tilde{\phi}^{\text{rand}}$ nie może być mniejszy niż $\Omega(n)$ odwołań do wyrazów ciągu. Jako, że każda wartość funkcji f_{ε_1} lub jej pochodnej jest wyznaczona przez jedno odwołanie do ciągu (x_i), więc dla dowolnego algorytmu szukającego zer funkcji

$$c^{\text{rand}}(\varepsilon) = \Omega(n) = \Omega(\varepsilon_1^{-d/(r+\rho)}) = \Omega(\varepsilon^{-d/(r+\rho)}). \tag{4.16}$$

Otrzymujemy więc, że koszt wyznaczenia ε–przybliżenia w każdym z modeli wynosi odpowiednio (4.14), (4.16), (4.15). Stąd wynika teza twierdzenia 19. □

4.5. Podsumowanie

W tym rozdziale zbadaliśmy złożoność problemu poszukiwania zer funkcji z klasy Höldera. Przedstawiliśmy algorytmy rozwiązujące ten problem w modelach deterministycznym, randomizacyjnym i kwantowym. Przeanalizowaliśmy błędy oraz koszt tych algorytmów. Na tej podstawie ustaliliśmy górne ograniczenia na ε–złożoność tego problemu. Wyznaczyliśmy również dolne ograniczenia na złożoność w tych trzech modelach. Ograniczenia te zawarte są w tabeli 4.1 (dla większej przejrzystości pominęliśmy stałe).

model	ograniczenia z dołu	ograniczenia z góry
det	$\varepsilon^{-\frac{d}{r+\rho}}$	$\varepsilon^{-\frac{d}{r+\rho}}$
rand	$\varepsilon^{-\frac{d}{r+\rho}}$	$\varepsilon^{-\frac{d}{r+\rho}}$
quant	$\varepsilon^{-\frac{d}{2(r+\rho)}}$	$\varepsilon^{-\frac{d}{2(r+\rho)}}$

Tabela 4.1. Ograniczenia na złożoność

Z wyników przedstawionych w tabeli 4.1 wynika, że górne i dolne ograniczenia na złożoność są zgodne we wszystkich modelach (różnią się wyłącznie o stałą). Tak więc, skonstruowane przez nas algorytmy są (asymptotycznie) optymalne. Złożoność w modelu randomizacyjnym jest tego samego rzędu co złożoność w modelu deterministycznym. W modelu kwantowym otrzymujemy natomiast kwadratowe przyspieszenie w stosunku do pozostałych modeli. We wszystkich trzech modelach występuje wykładnicza zależność od wymiaru przestrzeni d.

Rozdział 5

Problemy brzegowe

5.1. Wstęp

Zajmiemy się w tym rozdziale złożonością problemów brzegowych dla równań różniczkowych zwyczajnych. Analiza złożoności problemów brzegowych jest związana z badaniem złożoności problemów początkowych. Rozwiązywanie problemów początkowych dla układów równań różniczkowych rozważane było w pracy [Kac87] w modelu deterministycznym oraz w [Kac06] w modelach randomizacyjnym i kwantowym. Złożoność problemów wyższych rzędów w modelu deterministycznym zbadana została w [Szc06], a w modelach randomizacyjnym i kwantowym w [GS08]. Rozwiązywanie problemów brzegowych jest zwykle bardziej złożone od rozwiązywania problemów początkowych. Oprócz problemów początkowych wykorzystuje się tu również algorytmy poszukiwania zer funkcji.

W tym rozdziale zajmiemy się pewnymi szczególnymi przypadkami problemów brzegowych. Będziemy badać złożoność dwóch problemów dwupunktowych: dla niejednorodnych równań liniowych wyższych rzędów oraz dla układów równań nieliniowych pierwszego rzędu.

5.2. Liniowy dwupunktowy problem brzegowy

5.2.1. Sformułowanie problemu

W tej części zajmiemy się rozwiązywaniem dwupunktowych problemów brzegowych rzędu k o rozdzielonych warunkach brzegowych postaci

$$\begin{cases} u^{(k)}(x) = \sum_{p=0}^{q} f_p(x)u^{(p)}(x) + f_k(x), & x \in [a,b], \\ u^{(v)}(a) = \alpha_v, \qquad v = 0,1,\ldots,k-1, & v \neq v_0, \\ u^{(v_0)}(b) = \beta_{v_0}, \end{cases} \tag{5.1}$$

gdzie $0 \leqslant q < k$, $0 \leqslant v_0 < k$, $u : [a,b] \to \mathbb{R}$, $a < b$. Zakładamy, że funkcje $f_j : [a,b] \to \mathbb{R}$ dla $j = 0,1,\ldots,q$ i $j = k$ należą do klasy F^r funkcji r-krotnie różniczkowalnych z wszystkimi pochodnymi ograniczonymi zdefiniowanej następująco

$$F^r = \left\{ f : f \in C^r([a,b]), f(x) \geqslant 0, \|f^{(p)}\| \leqslant D \text{ dla } p = 0,1,\ldots,r \right\},$$

gdzie $r \geqslant 1$, a D jest pewną dodatnią stałą. Dla wygody zapisu przyjmiemy $f_j \equiv 0$ dla $j = q+1, q+2, \ldots, k-1$. Będziemy w tej części używać normy nieskończoność dla wektorów, $\|\cdot\| = \|\cdot\|_\infty$. Dla macierzy $A = (a_{i,j})$ normę definiujemy jako $\|A\|_\infty = \max_i \sum_j |a_{i,j}|$. W przestrzeni funkcji będziemy używać normy supremum na odcinku $[a,b]$.

Dla problemu (5.1) istnieje jednoznaczne rozwiązanie w klasie $C^{r+k}([a,b])$, co zostanie pokazane w podrozdziale 5.2.3.

Wartość operatora rozwiązania S dla tego problemu na elemencie wejściowym $f = (f_0, f_1, \ldots, f_k)$ jest równa rozwiązaniu problemu brzegowego (5.1), $S(f) = u$. Mimo, że rozpatrujemy równanie liniowe, operator rozwiązania S dla tego problemu jest nieliniowy, choćby dlatego, że suma rozwiązań dla dwóch funkcji prawej strony nie będzie na ogół spełniać warunków brzegowych. Naszym celem jest przybliżenie wartości $S(f_0, f_1, \ldots, f_k)$ dla dowolnych $f_0, f_1, \ldots, f_q \in F^r$, $f_{q+1}, f_{q+2}, \ldots, f_k \equiv 0$ z dokładnością $\varepsilon > 0$ mierzoną w normie supremum.

Przedstawimy teraz rodzaje rozpatrywanych modeli obliczeniowych. Definicje informacji, błędu, kosztu i złożoności są analogiczne do tych wprowadzonych w rozdziale 1. Rozważać będziemy trzy modele obliczeniowe: deterministyczny, randomizacyjny i kwantowy. W modelu deterministycznym rozważać będziemy dwa rodzaje

informacji: standardową i liniową. Dla kompletności pokażemy jak ogólne definicje z rozdziału 1 przedstawiają się dla tego problemu.
Informacja standardowa składa się z wartości funkcji prawej strony $f_0, f_1, \ldots, f_q, f_k$ oraz ich pochodnych do rzędu r w pewnych deterministycznie ustalonych punktach. Informacja liniowa składa się natomiast z wartości dowolnych funkcjonałów liniowych na funkcjach prawej strony. W modelu randomizacyjnym wartości funkcji prawej strony i ich pochodnych mogą być obliczane w losowo wybranych punktach. W modelu kwantowym informacja dana jest przez wyrocznię kwantową.
W modelu deterministycznym, wynikiem algorytmu ϕ dla tego problemu na pewnym elemencie wejściowym $f = (f_0, f_1, \ldots, f_k)$ jest pewna ograniczona funkcja l. Błąd algorytmu ϕ w klasie F^r w przypadku deterministycznym jest definiowany jako

$$\mathrm{e}^{\mathrm{det-st/det-lin}}(S, \phi, F^r) = \sup_{f_1,\ldots,f_q,f_k \in F^r} \sup_{x\in[a,b]} |u(x) - l(x)|.$$

W modelu randomizacyjnym i kwantowym, wynikiem algorytmu ϕ jest pewna losowa funkcja l^ω na pewnej przestrzeni probabilistycznej $(\Omega, \Sigma, \mathbf{P})$. W modelu randomizacyjnym, błąd algorytmu ϕ w klasie F^r zdefiniowany jest następująco

$$\mathrm{e}^{\mathrm{rand}}(S, \phi, F^r) = \sup_{f_0,\ldots,f_q,f_k \in F^r} \left(\mathbf{E} \left(\sup_{x\in[a,b]} |u(x) - l^\omega(x)| \right)^2 \right)^{1/2}. \tag{5.2}$$

Błąd probabilistyczny w modelu kwantowym dany jest przez

$$\mathrm{e}^{\mathrm{quant}}(S, \phi, F^r) = \sup_{f_0,\ldots,f_q,f_k \in F^r} \inf \left\{ \alpha : \mathbf{P} \left(\sup_{x\in[a,b]} |u(x) - l^\omega(x)| > \alpha \right) \leqslant 1/4 \right\}.$$

Koszt we wszystkich modelach definiowany jest jako liczność informacji i oznaczony jest przez $\mathrm{cost}^{\mathrm{det-st/det-lin/rand/quant}}(S, \phi, F^r)$.
Złożoność definiowana jest wzorem (1.6).
W następnym podrozdziale przedstawimy znane wyniki dotyczące tego problemu.

5.2.2. Znane wyniki w modelu deterministycznym

W literaturze rozważany był dwupunktowy, jednorodny, liniowy problem brzegowy rzędu drugiego postaci

$$\begin{cases} u''(x) = f(x)u(x), \\ u(0) = c, \\ u'(T) = 0, \quad x \in [0, T], \end{cases} \tag{5.3}$$

gdzie $f \in F^r$. Niech S^* będzie operatorem rozwiązania dla tego problemu. W pracy [Kac90] przedstawione są dolne i górne ograniczenia na złożoność tego problemu w modelu deterministycznym z wykorzystaniem informacji standardowej. Są one rzędu

$$\mathrm{comp}^{\mathrm{det-st}}(\varepsilon, S^*, F^r) = \Theta\left(\left(\frac{1}{\varepsilon}\right)^{1/r}\right). \tag{5.4}$$

Górne ograniczenia na złożoność pokazane są w tej pracy poprzez skonstruowanie optymalnego algorytmu. Algorytm ten polega na przybliżeniu funkcji f przez interpolacyjną funkcję sklejaną, a następnie rozwiązanie powstałego problemu brzegowego za pomocą metody strzałów wielokrotnych. Algorytm ten korzysta jedynie z wartości funkcji f (nie korzysta z wartości pochodnych), co dowodzi, że informacja składająca się wyłącznie z wartości funkcji jest optymalna w klasie informacji standardowych.

Znane są również górne ograniczenia na złożoność tego problemu w modelu deterministycznym z informacją liniową, patrz [Drw07]. Są one rzędu

$$\mathrm{comp}^{\mathrm{det-lin}}(\varepsilon, S^*, F^r) = O\left(\left(\frac{1}{\varepsilon}\right)^{1/(r+2)}\right). \tag{5.5}$$

W [Drw07] przedstawiony jest algorytm wykorzystujący metodę strzałów wielokrotnych. Algorytm ten korzysta z informacji całkowej o funkcji f.
Z ograniczeń tych wynika, że informacja liniowa daje przyspieszenie o 2 w wykładniku w stosunku do informacji standardowej.

Naszym celem jest uzyskanie dolnych i górnych oszacowań na złożoność problemu (5.1) w modelach randomizacyjnym i kwantowym. Dla porównania wyznaczymy również ograniczenia na złożoność tego problemu w modelu deterministycznym zarówno z informacja standardową jak i liniową. Porównamy również otrzymane wyniki ze znanymi wynikami (5.4), (5.5).
W następnym podrozdziale przedstawimy pewne pomocnicze wyniki dotyczące rozwiązania problemu (5.1), przydatne przy wyznaczaniu górnych ograniczeń na złożoność.

5.2.3. Wstępne wyniki

W tej części pracy wyprowadzimy pewne własności rozwiązania problemu (5.1), które zostaną użyte w dalszej części pracy.
Rozważmy równomierny podział odcinka $[a, b]$ na n części za pomocą punktów $x_i = a + ih, \quad i = 0, 1, \ldots, n$, gdzie $h = (b-a)/n$. Oznaczmy przez $u_{j,i}$, $j = 0, 1, \ldots, k-1$,

$i = 0, 1, \ldots, n-1$, rozwiązania następujących lokalnych problemów początkowych na przedziale $[x_i, x_{i+1}]$

$$\begin{cases} u_{j,i}^{(k)}(x) = \sum_{p=0}^{q} f_p(x) u_{j,i}^{(p)}(x), & x \in [x_i, x_{i+1}], \\ u_{j,i}^{(v)}(x_i) = \delta_{vj}, & v = 0, 1, \ldots, k-1, \end{cases} \tag{5.6}$$

a przez $u_{k,i}$ rozwiązania problemów

$$\begin{cases} u_{k,i}^{(k)}(x) = \sum_{p=0}^{q} f_p(x) u_{k,i}^{(p)}(x) + f_k(x), & x \in [x_i, x_{i+1}], \\ u_{k,i}^{(v)}(x_i) = 0, & v = 0, 1, \ldots, k-1. \end{cases} \tag{5.7}$$

Rozwiązania tych problemów istnieją i są jednoznaczne, gdyż są to liniowe problemy początkowe o ciągłych funkcjach prawej strony.
Szukamy rozwiązania u problemu (5.1) w postaci

$$u(x) = \sum_{j=0}^{k-1} s_i^j u_{j,i}(x) + u_{k,i}(x) \text{ dla } x \in [x_i, x_{i+1}), \quad i = 0, 1, \ldots, n-1 \tag{5.8}$$

dla pewnych wartości s_i^j. Zauważmy, że dla dowolnych s_i^j, tak określona funkcja u spełnia równanie różniczkowe w (5.1). Zauważmy również, że $s_i^j = u^{(j)}(x_i)$.
Niech $a_{j,i+1}^v = u_{j,i}^{(v)}(x_{i+1})$ dla $i = 0, 1, \ldots, n-1$, $j = 0, 1, \ldots, k$, $v = 0, 1, \ldots, k-1$. Aby funkcja u osiągała odpowiednią regularność oraz aby spełnione były warunki brzegowe w (5.1), musi być spełniony następujący układ równań

- warunki brzegowe w punkcie a

$$s_0^v = \alpha_v, \qquad v = 0, 1, \ldots, k-1, \quad v \neq v_0$$

- ciągłość u oraz jej pochodnych w punkcie x_{i+1}

$$s_{i+1}^v = \sum_{j=0}^{k-1} s_i^j a_{j,i+1}^v + a_{k,i+1}^v, \qquad v = 0, 1, \ldots, k-1, \; i = 0, 1, \ldots, n-2$$

- warunki brzegowe w punkcie b

$$s_n^{v_0} = \sum_{j=0}^{k-1} s_{n-1}^j a_{j,n}^{v_0} + a_{k,n}^{v_0} = \beta_{v_0}.$$

Układ ten można zapisać w postaci macierzowej

$$AS = P, \tag{5.9}$$

gdzie $S = [s_0^0, \ldots, s_0^{v_0-1}, s_0^{v_0+1}, \ldots, s_0^{k-1}, s_1^0 \ldots s_1^{k-1}, \ldots, s_{n-1}^0 \ldots s_{n-1}^{k-1}, s_0^{v_0}]^T \in \mathbb{R}^{kn}$, $P = [\alpha_0, \ldots, \alpha_{v_0-1}, \alpha_{v_0+1}, \ldots, \alpha_{k-1}, a_{k,1}^0, \ldots, a_{k,1}^{k-1}, \ldots, a_{k,n-1}^0, \ldots, a_{k,n-1}^{k-1}, a_{k,n}^{v_0} - \beta_{v_0}]^T \in \mathbb{R}^{kn}$ i A jest macierzą wymiaru $kn \times kn$ postaci

$$A = \left[\begin{array}{c|c|c|c|c|c|c} I_{k-1} & 0 & 0 & \cdots & 0 & 0 & 0 \\ \hline -A_1 & I_k & 0 & \cdots & 0 & 0 & -b_1 \\ \hline 0 & -A_2 & I_k & \cdots & 0 & 0 & 0 \\ \hline & & & \ddots & & & \\ \hline 0 & 0 & 0 & \cdots & I_k & 0 & 0 \\ \hline 0 & 0 & 0 & \cdots & -A_{n-1} & I_k & 0 \\ \hline 0 & 0 & 0 & \cdots & 0 & -A_n & 0 \end{array}\right].$$

Każda macierz A_i, $i = 2, 3, \ldots, n-1$, ma wymiar $k \times k$ i jest postaci $A_i = [a_{j,i}^v]_{v,j=0,1,\ldots,k-1}$. Macierz A_1 ma wymiar $k \times (k-1)$ i jest postaci $A_1 = [a_{j,1}^v]_{v=0,1,\ldots,k-1, j=0,1,\ldots,k-1, j\neq v_0}$, a wektor $b_1 \in \mathbb{R}^k$ jest dany przez $b_1 = [a_{v_0,1}^v]_{v=0,1,\ldots,k-1}$. Macierz A_n jest wymiaru $1 \times k$ i ma postać $A_n = [a_{0,n}^{v_0}, a_{1,n}^{v_0}, \ldots, a_{k-1,n}^{v_0}]$. Przez I_k oznaczamy macierz identyczności wymiaru $k \times k$.

Dla każdego $n \geqslant 1$, problem (5.1) ma jednoznaczne rozwiązanie wtedy i tylko wtedy, gdy macierz A jest nieosobliwa. Istotnie, jeśli macierz A jest nieosobliwa, to istnienie rozwiązania problemu wynika z (5.8). Jednoznaczność wynika natomiast z faktu, że każde rozwiązanie problemu (5.1) można zapisać w postaci (5.8).

Jeśli założymy natomiast, że problem (5.1) ma jednoznaczne rozwiązanie, to jest ono w postaci (5.8), więc układ $AS = P$ ma rozwiązanie. Układ ten nie może mieć dwóch różnych rozwiązań, gdyż dawały by one dwa różne rozwiązania problemu (5.1) w postaci (5.8). Stąd A jest nieosobliwa.

Do zbadania własności macierzy A w (5.9) potrzebna nam będzie ograniczoność z góry rozwiązań problemów lokalnych (5.6), (5.7) przez pewne stałe zależne tylko od przedziału $[a, b]$ i parametrów klasy F^r.

Fakt 1. *Istnieje dodatnia stała Q zależna tylko od a, b, k i D taka, że rozwiązania każdego z problemów (5.6) lub (5.7) są ograniczone przez Q,*

$$|u_{j,i}^{(v)}(x)| \leqslant Q, \ \textit{dla } x \in [x_i, x_{i+1}],$$

$$i = 0, 1, \ldots, n-1, \ j = 0, 1, \ldots, k, \ v = 0, 1, \ldots, k-1.$$

Dowód jest standardowy, zamieszczony jest dla kompletności w dodatku B.

Aby wyprowadzić własności rozwiązania S układu (5.9) pokażemy, że macierz A jest nieosobliwa, a norma jej odwrotności jest $O(n)$.

Lemat 5. *Istnieje dodatnia stała R i $n_0 \in \mathbb{N}$ zależne jedynie od a, b, k i D takie, że dla każdego $n \geqslant n_0$ macierz A jest nieosobliwa oraz*

$$\|A^{-1}\| \leqslant Rn.$$

Dowód. W celu znalezienia macierzy odwrotnej do A zastosujemy rozkład LU macierzy A za pomocą blokowej eliminacji Gaussa. Niech

$$L_i = \begin{bmatrix} I_{k-1} & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & I_k & \cdots & 0 & 0 & \cdots & 0 & 0 \\ & & \ddots & & & & & \\ 0 & 0 & \cdots & I_k & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & A_i & I_k & \cdots & 0 & 0 \\ & & & & & \ddots & & \\ 0 & 0 & \cdots & 0 & 0 & \cdots & I_k & 0 \\ 0 & 0 & \cdots & \underbrace{0}_{i} & 0 & \cdots & 0 & 1 \end{bmatrix}$$

będzie macierzą o tej samej strukturze blokowej co macierz A. Macierz odwrotna do macierzy L_i ma postać

$$L_i^{-1} = \begin{bmatrix} I_{k-1} & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & I_k & \cdots & 0 & 0 & \cdots & 0 & 0 \\ & & \ddots & & & & & \\ 0 & 0 & \cdots & I_k & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & -A_i & I_k & \cdots & 0 & 0 \\ & & & & & \ddots & & \\ 0 & 0 & \cdots & 0 & 0 & \cdots & I_k & 0 \\ 0 & 0 & \cdots & \underbrace{0}_{i} & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

Można to sprawdzić przez mnożenie macierzy. Rzeczywiście $L_i L_i^{-1} = L_i^{-1} L_i = I_{nk}$. Poszukamy teraz macierzy górnej trójkątnej U w rozkładzie LU macierzy A. Uzyskamy

ją przez mnożenie macierzy A lewostronnie kolejno przez macierze $L_1, L_2, \ldots, L_n$. Zauważmy, że

$$L_1 A = \left[\begin{array}{c|c|c|c|c|c|c} I_{k-1} & 0 & 0 & \cdots & 0 & 0 & 0 \\ \hline 0 & I_k & 0 & \cdots & 0 & 0 & -b_1 \\ \hline 0 & -A_2 & I_k & \cdots & 0 & 0 & 0 \\ \hline & & & \ddots & & & \\ \hline 0 & 0 & 0 & \cdots & I_k & 0 & 0 \\ \hline 0 & 0 & 0 & \cdots & -A_{n-1} & I_k & 0 \\ \hline 0 & 0 & 0 & \cdots & 0 & -A_n & 0 \end{array}\right].$$

Po drugim mnożeniu dostajemy

$$L_2 L_1 A = \left[\begin{array}{c|c|c|c|c|c|c} I_{k-1} & 0 & 0 & 0 & \cdots & 0 & 0 \\ \hline 0 & I_k & 0 & 0 & \cdots & 0 & -b_1 \\ \hline 0 & 0 & I_k & 0 & \cdots & 0 & -A_2 b_1 \\ \hline 0 & 0 & -A_3 & I_k & \cdots & 0 & 0 \\ \hline & & & & \ddots & & \\ \hline 0 & 0 & 0 & 0 & \cdots & I_k & 0 \\ \hline 0 & 0 & 0 & 0 & \cdots & -A_n & 0 \end{array}\right].$$

Po kolejnym mnożeniu mamy

$$L_3 L_2 L_1 A = \left[\begin{array}{c|c|c|c|c|c|c|c} I_{k-1} & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \hline 0 & I_k & 0 & 0 & 0 & \cdots & 0 & -b_1 \\ \hline 0 & 0 & I_k & 0 & 0 & \cdots & 0 & -A_2 b_1 \\ \hline 0 & 0 & 0 & I_k & 0 & \cdots & 0 & -A_3 A_2 b_1 \\ \hline 0 & 0 & 0 & -A_4 & I_k & \cdots & 0 & 0 \\ \hline & & & & & \ddots & & \\ \hline 0 & 0 & 0 & 0 & 0 & \cdots & I_k & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & \cdots & -A_n & 0 \end{array}\right].$$

Ostatecznie otrzymujemy macierz górną trójkątną U postaci

$$U = L_n L_{n-1} \cdots L_1 A = \left[\begin{array}{c|c|c|c|c|c} I_{k-1} & 0 & 0 & \cdots & 0 & 0 \\ \hline 0 & I_k & 0 & \cdots & 0 & -b_1 \\ \hline 0 & 0 & I_k & \cdots & 0 & -b_2 \\ \hline & & & \ddots & & \\ \hline 0 & 0 & 0 & \cdots & I_k & -b_{n-1} \\ \hline 0 & 0 & 0 & \cdots & 0 & -b_n \end{array}\right],$$

gdzie $b_i = A_i A_{i-1} \cdots A_2 b_1 \in \mathbb{R}^k$ dla $i = 2, 3, \ldots, n-1$ i $b_n = A_n A_{n-1} \cdots A_2 b_1 \in \mathbb{R}$, $n \geqslant 2$.

Wyznaczymy teraz macierz dolną trójkątną L rozkładu LU macierzy A. Jako, że $U = L_n L_{n-1} \ldots L_1 A$ i istnieją macierze L_i^{-1}, więc

$$A = L_1^{-1} L_2^{-1} \cdots L_n^{-1} U.$$

Przyjmijmy

$$L = L_1^{-1} L_2^{-1} \cdots L_n^{-1}.$$

Wyznaczymy postać tej macierzy. Zauważmy, że

$$L_1^{-1} L_2^{-1} = \left[\begin{array}{c|c|c|c|c|c|c} I_{k-1} & 0 & 0 & \cdots & 0 & 0 & 0 \\ \hline -A_1 & I_k & 0 & \cdots & 0 & 0 & 0 \\ \hline 0 & -A_2 & I_k & \cdots & 0 & 0 & 0 \\ \hline & & & \ddots & & & \\ \hline 0 & 0 & 0 & \cdots & I_k & 0 & 0 \\ \hline 0 & 0 & 0 & \cdots & 0 & I_k & 0 \\ \hline 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{array}\right].$$

W kolejnym mnożeniu uzyskujemy

$$L_1^{-1} L_2^{-1} L_3^{-1} = \left[\begin{array}{c|c|c|c|c|c|c|c} I_{k-1} & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \hline -A_1 & I_k & 0 & 0 & \cdots & 0 & 0 & 0 \\ \hline 0 & -A_2 & I_k & 0 & \cdots & 0 & 0 & 0 \\ \hline 0 & 0 & -A_3 & I_k & \cdots & 0 & 0 & 0 \\ \hline & & & & \ddots & & & \\ \hline 0 & 0 & 0 & 0 & \cdots & I_k & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & \cdots & 0 & I_k & 0 \\ \hline 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{array}\right].$$

Ostatecznie otrzymujemy macierz dolną trójkątną L postaci

$$L = L_1^{-1} L_2^{-1} \cdots L_n^{-1} = \left[\begin{array}{c|c|c|c|c|c|c} I_{k-1} & 0 & 0 & \cdots & 0 & 0 & 0 \\ \hline -A_1 & I_k & 0 & \cdots & 0 & 0 & 0 \\ \hline 0 & -A_2 & I_k & \cdots & 0 & 0 & 0 \\ \hline & & & \ddots & & & \\ \hline 0 & 0 & 0 & \cdots & I_k & 0 & 0 \\ \hline 0 & 0 & 0 & \cdots & -A_{n-1} & I_k & 0 \\ \hline 0 & 0 & 0 & \cdots & 0 & -A_n & 1 \end{array}\right].$$

Wyznaczymy teraz elementy macierzy A_i, $i = 1, 2, \ldots, n$. Dla $v, j = 0, 1, \ldots, k-1$, korzystając z twierdzenia Taylora oraz faktu 1, dostajemy

$$a^v_{j,i+1} = u^{(v)}_{j,i}(x_{i+1}) = u^{(v)}_{j,i}(x_i) + h u^{(v+1)}_{j,i}(x_i) + O(h^2),$$

gdzie stałe w notacji O zależą wyłącznie od a, b, k i D. Zgodnie z warunkami początkowymi problemów (5.6), (5.7), dla $v < k$ wartość $u^{(v)}_{j,i}(x_i)$ może wynosić wyłącznie 1 lub 0 w zależności od tego czy j jest równe, czy różne od v. Natomiast $u^{(k)}_{j,i}(x_i) = \sum_{p=0}^{k-1} f_p(x_i) u^{(p)}_{j,i}(x_i) = f_j(x_i)$. Tak więc,

$$a^v_{j,i+1} = \begin{cases} 1 + O(h^2), & \text{gdy } j = v,\ v \neq k-1 \\ h + O(h^2), & \text{gdy } j = v+1,\ v \neq k-1 \\ O(h^2), & \text{gdy } j \neq v, j \neq v+1,\ v \neq k-1 \\ f_j(x_i)h + O(h^2), & \text{gdy } j \neq k-1,\ v = k-1 \\ 1 + f_{k-1}(x_i)h + O(h^2), & \text{gdy } j = k-1,\ v = k-1 \end{cases} \tag{5.10}$$

dla $i = 0, 1, 2, \ldots, n-1$, $v, j = 0, 1, \ldots, k-1$. Stąd, macierz A_i dla $i = 2, 3, \ldots, n-1$ ma postać

$$A_i = \begin{bmatrix} 1 & h & 0 & \cdots & 0 & 0 \\ 0 & 1 & h & \cdots & 0 & 0 \\ 0 & 0 & 1 & \ddots & 0 & 0 \\ & & & \ddots & \ddots & \\ 0 & 0 & 0 & \cdots & 1 & h \\ f_{0,i}h & f_{1,i}h & f_{2,i}h & \cdots & f_{k-2,i}h & 1 + f_{k-1,i}h \end{bmatrix} + [O(h^2)]^{k-1}_{v,j=0}, \tag{5.11}$$

gdzie $f_{j,i} = f_j(x_{i-1})$.

Macierz L jest nieosobliwa. Pokażemy, że macierz U również jest nieosobliwa oraz, że normy macierzy odwrotnych do L i U są ograniczone przez stałe. Potrzebujemy do tego dolnego ograniczenia na liczbę $b_n = A_n A_{n-1} \cdots A_2 b_1$. Zauważmy, że dla $n \geqslant 3$ każda z macierzy A_i, $i = 2, 3, \ldots, n-1$, może być zapisana w postaci $A_i = I_k + B_i h + C_i$, gdzie $\|B_i\| \leqslant G$, $\|C_i\| \leqslant Hh^2$ dla pewnych stałych G i H zależnych wyłącznie od a, b, k, i D. Ponadto macierz B ma nieujemne elementy (jako, że liczby $f_{j,i}$ są nieujemne). Niech $\tilde{A}_i = A_i A_{i-1} \cdots A_2$, $i = 2, 3, \ldots, n-1$. Pokażemy, że $\tilde{A}_i = I_k + \tilde{B}_i + \tilde{C}_i$, gdzie $\tilde{B}_i$ ma nieujemne elementy oraz $\|\tilde{B}_i\| \leqslant G_i$, $\|\tilde{C}_i\| \leqslant H_i$, dla $G_i = (1 + Gh)^{i-1} - 1$, $H_i = (1 + Gh + Hh^2)^{i-1} - (1 + Gh)^{i-1}$. Istotnie, dla $i = 2$ jest to spełnione. Załóżmy,

że warunek jest spełniony dla i. Wtedy
$\tilde{A}_{i+1} = A_{i+1}\tilde{A}_i = (I + B_{i+1}h + C_{i+1})(I + \tilde{B}_i + \tilde{C}_i) = I + \tilde{B}_{i+1} + \tilde{C}_{i+1}$, gdzie

$$\tilde{B}_{i+1} = \tilde{B}_i + B_{i+1}h + B_{i+1}\tilde{B}_i h,$$
$$\tilde{C}_{i+1} = C_{i+1} + C_{i+1}\tilde{B}_i + C_{i+1}\tilde{C}_i + \tilde{C}_i + B_{i+1}\tilde{C}_i h.$$

Widać, że $\tilde{B}_{i+1}$ ma nieujemne elementy oraz
$\|\tilde{B}_{i+1}\| \leqslant \|\tilde{B}_i\| + \|B_{i+1}\|h + \|B_{i+1}\|\|\tilde{B}_i\|h = G_{i+1}$, gdzie

$$\begin{cases} G_{i+1} = G_i(1 + Gh) + Gh, & i \geqslant 3, \\ G_2 = Gh \end{cases}.$$

Z rozwiązania tego równania otrzymujemy, że $G_i = (1+Gh)^{i-1} - 1$. Ponadto $\|\tilde{C}_{i+1}\| \leqslant \|C_{i+1}\| + \|C_{i+1}\|\|\tilde{B}_i\| + \|C_{i+1}\|\|\tilde{C}_i\| + \|\tilde{C}_i\| + \|B_{i+1}\|\|\tilde{C}_i\|h \leqslant H_{i+1}$, gdzie

$$\begin{cases} H_{i+1} = H_i(1 + Gh + Hh^2) + Hh^2(1 + Gh)^{i-1}, & i \geqslant 3, \\ H_2 = Hh^2 \end{cases}.$$

Z rozwiązania tego równania mamy $H_i = (1 + Gh + Hh^2)^{i-1} - (1 + Gh)^{i-1}$. Mamy więc $\tilde{A}_{n-1} = I_k + \tilde{B}_{n-1} + \tilde{C}_{n-1}$, gdzie $\tilde{B}_{n-1}$ ma nieujemne elementy i

$$\begin{aligned} \|\tilde{C}_{n-1}\| &\leqslant H_{n-1} = (1 + Gh + Hh^2)^{n-2} - (1 + Gh)^{n-2} \\ &= \Big((1 + Gh + Hh^2) - (1 + Gh)\Big) \sum_{j=0}^{n-3} (1 + Gh + Hh^2)^j (1 + Gh)^{n-3-j} \\ &\leqslant Hh^2 \sum_{j=0}^{n-3} e^{j(Gh+Hh^2)} e^{(n-3-j)(Gh)} \leqslant Hh^2 n e^{n(Gh+Hh^2)} e^{nGh} = O(h). \end{aligned}$$

Oznaczmy $\tilde{B}_{n-1} = [B_{v,j}]_{v,j=0}^{k-1}$. Macierz $A_n = [a_{0,n}^{v_0}, a_{1,n}^{v_0}, \ldots, a_{k-1,n}^{v_0}]$ zgodnie z (5.10) ma postać $[0, \ldots, 0, \underbrace{1}_{v_0+1}, 0, \ldots, 0] + O(h)$, a wektor $b_1 = [a_{v_0,1}^0, a_{v_0,1}^1, \ldots, a_{v_0,1}^{k-1}]^T$ zgodnie z (5.10) jest postaci $[0, \ldots, 0, \underbrace{1}_{v_0+1}, 0, \ldots, 0]^T + O(h)$. Stąd, $b_n = A_n\tilde{A}_{n-1}b_1 = 1 + B_{v_0+1,v_0+1} + O(h)$, gdzie stała w notacji O zależy wyłącznie od a, b, k i D. Dla dostatecznie małego h otrzymujemy więc $b_n \geqslant 1/2$. Wynika z tego, że dla dostatecznie dużego n, macierz U jest nieosobliwa.
Macierz U^{-1} ma postać

$$U^{-1} = \left[\begin{array}{c|c|c|c|c|c} I_{k-1} & 0 & 0 & \cdots & 0 & 0 \\ \hline 0 & I_k & 0 & \cdots & 0 & -b_1 b_n^{-1} \\ \hline 0 & 0 & I_k & \cdots & 0 & -b_2 b_n^{-1} \\ \hline & & & \ddots & & \\ \hline 0 & 0 & 0 & \cdots & I_k & -b_{n-1} b_n^{-1} \\ \hline 0 & 0 & 0 & \cdots & 0 & -b_n^{-1} \end{array}\right],$$

co łatwo sprawdzić przez mnożenie przez U. Wyznaczymy jeszcze postać $L^{-1} = L_n L_{n-1} \dots L_1$. Zauważmy, że po pierwszym mnożeniu dostajemy

$$L_2 L_1 = \left[\begin{array}{c|c|c|c|c|c|c} I_{k-1} & 0 & 0 & 0 & \cdots & 0 & 0 \\ \hline A_1 & I_k & 0 & 0 & \cdots & 0 & 0 \\ \hline A_2 A_1 & A_2 & I_k & 0 & \cdots & 0 & 0 \\ \hline 0 & 0 & 0 & I_k & \cdots & 0 & 0 \\ \hline & & & & \ddots & & \\ \hline 0 & 0 & 0 & 0 & \cdots & I_k & 0 \\ \hline 0 & 0 & 0 & 0 & \cdots & A_n & 1 \end{array}\right].$$

Ostatecznie, po kolejnych mnożeniach uzyskujemy

$$L^{-1} = \left[\begin{array}{c|c|c|c|c|c} I_{k-1} & 0 & 0 & \cdots & 0 & 0 \\ \hline A_1 & I_k & 0 & \cdots & 0 & 0 \\ \hline A_2 A_1 & A_2 & I_k & \cdots & 0 & 0 \\ \hline & & & \ddots & & \\ \hline A_{n-1} \dots A_1 & A_{n-1} \dots A_2 & A_{n-1} \dots A_3 & \cdots & I_k & 0 \\ \hline A_n \dots A_1 & A_n \dots A_2 & A_n \dots A_3 & \cdots & A_n & 1 \end{array}\right].$$

Elementy macierzy U^{-1} oraz L^{-1} są ograniczone przez normy odpowiednich bloków. Oszacujemy teraz normy macierzy A_i, $i = 2, 3, \dots, n-1$. Korzystając z (5.11), sumy wartości bezwzględnych elementów w każdym z wierszy od 1 do $k-1$ są ograniczone $1 + h + O(h^2)$, natomiast w wierszu k-tym przez $1 + kDh + O(h^2)$. Stąd $\|A_i\| \leqslant 1 + (kD+1)h + O(h^2) \leqslant 1 + (kD+2)h$, dla dostatecznie małego h. Tak samo ograniczone są normy macierz A_1 i wektora b_1 (jako, że zawierają się one w macierzy $[a^v_{j,1}]_{v,j=0,1,\dots,k-1}$, która ma taką samą strukturę jak macierze A_i). Tak więc, dla $i, j = 1, 2, \dots, n$, $i < j$, mamy

$$\|A_j \cdots A_{i+1} A_i\| \leqslant \|A_j\| \cdots \|A_{i+1}\| \|A_i\| \leqslant (1 + h(kD+2))^n \leqslant C,$$

gdzie $C = e^{(b-a)(kD+2)}$. Podobnie, $\|b_j\| = \|A_j \cdots A_2 b_1\| \leqslant C$. Norma każdego bloku macierzy L^{-1} jest ograniczona przez C, a normy bloków macierzy U^{-1} są ograniczone przez $b_n^{-1}\|b_j\| \leqslant 2C$. Wynika stąd, że elementy macierzy U^{-1} i L^{-1} są rzędu $O(1)$ ze stałymi zależnymi wyłącznie od a, b i D. Jako, że $A^{-1} = U^{-1}L^{-1}$ i każdy wiersz macierzy U^{-1} ma co najwyżej dwa niezerowe elementy, norma macierzy A^{-1} jest rzędu $O(n)$ ze stałą zależną wyłącznie od a, b, k i D. □

Nieosobliwość macierzy A pociąga za sobą istnienie i jednoznaczność rozwiązania problemu (5.1). Z ograniczeń na macierz A^{-1} wynikają następujące (znane skądinąd) ograniczenia na rozwiązanie problemu (5.1).

Fakt 2. *Istnieje dodatnia stała T zależna wyłącznie od a, b, k, D i warunków brzegowych taka, że rozwiązanie problemu (5.1) jest ograniczone przez*

$$|u^{(v)}(x)| \leqslant T, \ \ dla\ x \in [a,b], \ v = 0, 1, \ldots, k-1.$$

Dowód. Zgodnie z (5.8) i faktem 1 na każdym przedziale $[x_i, x_{i+1}]$ mamy

$$|u^{(v)}(x)| \leqslant Q\left(\sum_{j=0}^{k-1} |s_i^j| + 1\right), \tag{5.12}$$

gdzie x_i są punktami równomiernego podziału odcinka $[a,b]$ na n podprzedziałów. Potrzebujemy więc ograniczeń na $|s_i^j|$. Przyjmijmy $n = n_0$, gdzie $n_0 \in \mathbb{N}$ jest stałą z lematu 5 zależną od a, b, k, D. Zgodnie z lematem 5 norma A^{-1} jest ograniczona przez Rn_0. Norma wektora P w równaniu (5.9) dla $n = n_0$ jest ograniczona przez pewną dodatnią stałą W zależną od a, b, k, D i warunków brzegowych. Tak więc, norma wektora $S = A^{-1}P$ jest ograniczona przez RWn_0. Stąd oraz z (5.12) dostajemy, że dla $n = n_0$ i dla każdego $x \in [x_i, x_{i+1}]$, $|u^{(v)}(x)|$ jest ograniczona przez $T = Q(kRWn_0 + 1)$. □

Jako, że $S = [s_0^0, \ldots, s_0^{v_0-1}, s_0^{v_0+1}, \ldots, s_0^{k-1}, s_1^0 \ldots s_1^{k-1}, \ldots, s_{n-1}^0 \ldots s_{n-1}^{k-1}, s_0^{v_0}]^T$ i $s_i^j = u^{(j)}(x_i)$, $i = 0, 1, \ldots, n-1$, $j = 0, 1, \ldots, k-1$, więc z faktu 2 wynika następujące ograniczenie na rozwiązanie układu $AS = P$:
istnieje dodatnia stała T i $n_0 \in \mathbb{N}$ zależne jedynie od a, b, k i D i warunków brzegowych takie, że dla każdego $n \geqslant n_0$ rozwiązanie układu $AS = P$ spełnia

$$\|S\| \leqslant T \tag{5.13}$$

Ograniczenia na $\|A^{-1}\|$ i $\|S\|$ będą używane w następnych podrozdziałach.

5.2.4. Algorytmy

W tej części rozprawy zdefiniujemy algorytmy rozwiązujące (5.1): deterministyczny z informacją standardową i liniową, randomizacyjny i kwantowy. Oznaczmy te algorytmy w odpowiednich modelach przez $\phi_n^{\mathrm{det-st}}$, $\phi_n^{\mathrm{det-lin}}$, ϕ_n^{rand} i ϕ_n^{quant}. Wszystkie te algorytmy oparte są na obliczeniu rozwiązania $\tilde{S}$ układu równań liniowych $\tilde{A}\tilde{S} = \tilde{P}$,

gdzie macierz $\tilde{A}$ i wektor $\tilde{P}$ są przybliżeniami odpowiednio A i P oraz skorzystaniu z zależności (5.8). W celu określenia przybliżeń $\tilde{A}$ i $\tilde{P}$ zastępujemy elementy $a^v_{j,i},\ i = 1,2,\ldots,n,\ j = 0,1,\ldots,k,\ v = 0,1,\ldots,k-1$ przez ich przybliżenia $\tilde{a}^v_{j,i}$. Algorytmy w różnych modelach będą się różniły sposobem wyznaczenia przybliżeń $\tilde{a}^v_{j,i}$. Pokażemy teraz, jak wyznaczyć te przybliżenia w odpowiednich modelach.
Niech $l_{j,i}$ będzie wielomianem Taylora przybliżającym $u_{j,i}$

$$l_{j,i}(x) = \sum_{p=0}^{r+k-1} \frac{u^{(p)}_{j,i}(x_i)(x-x_i)^p}{p!}. \tag{5.14}$$

Rozważmy model deterministyczny z informacją standardową. Przybliżenie $\tilde{a}^v_{j,i}$ jest zdefiniowane jako wartości v-tej pochodnej funkcji $l_{j,i}$ w punkcie x_{i+1}

$$\tilde{a}^v_{j,i+1} = l^{(v)}_{j,i}(x_{i+1}). \tag{5.15}$$

Przejdziemy teraz do modelu deterministycznego z informacją liniową. Całkując równania (5.6), (5.7) $(k-v)$ razy, dostajemy

$$\begin{aligned} a^v_{j,i+1} = u^{(v)}_{j,i}(x_{i+1}) &= \sum_{p=0}^{k-1-v} \frac{u^{(p+v)}_{j,i}(x_i)h^p}{p!} \\ &+ \int_{x_i}^{x_{i+1}} \int_{x_i}^{t_{k-v-1}} \ldots \int_{x_i}^{t_1} \left(\sum_{p=0}^{q} f_p(t) u^{(p)}_{j,i}(t) \right) dt dt_1 dt_2 \ldots dt_{k-v-1} \end{aligned} \tag{5.16}$$

dla $i = 0,1,\ldots,n-1$, $v = 0,1,\ldots,k-1$, $j = 0,1,\ldots,k-1$, i

$$\begin{aligned} a^v_{k,i+1} = u^{(v)}_{k,i}(x_{i+1}) &= \sum_{p=0}^{k-1-v} \frac{u^{(p+v)}_{k,i}(x_i)h^p}{p!} \\ &+ \int_{x_i}^{x_{i+1}} \int_{x_i}^{t_{k-v-1}} \ldots \int_{x_i}^{t_1} \left(\sum_{p=0}^{q} f_p(t) u^{(p)}_{k,i}(t) + f_k(t) \right) dt dt_1 dt_2 \ldots dt_{k-v-1}. \end{aligned} \tag{5.17}$$

Następnie, całkując przez części zamienimy całkę wielokrotną na całkę z wagą. Niech $I = \int_{x_i}^{x_{i+1}} \int_{x_i}^{t_{k-v-1}} \ldots \int_{x_i}^{t_1} \left(\sum_{p=0}^{q} f_p(t) u^{(p)}_{j,i}(t) \right) dt dt_1 dt_2 \ldots dt_{k-v-1}$ będzie całką wielokrotną ze wzoru (5.16). Przyjmijmy $g'(t_{k-v-1}) = 1$, a

$$f(t_{k-v-1}) = \int_{x_i}^{t_{k-v-1}} \int_{x_i}^{t_{k-v-2}} \ldots \int_{x_i}^{t_1} \left(\sum_{p=0}^{q} f_p(t) u^{(p)}_{k,i}(t) \right) dt dt_1 dt_2 \ldots dt_{k-v-2}.$$

Zauważmy, że $g(t_{k-v-1}) = t_{k-v-1}$, a

$$f'(t_{k-v-1}) = \int_{x_i}^{t_{k-v-1}} \int_{x_i}^{t_{k-v-3}} \ldots \int_{x_i}^{t_1} \left(\sum_{p=0}^{q} f_p(t) u^{(p)}_{k,i}(t) \right) dt dt_1 dt_2 \ldots dt_{k-v-3}.$$

Stosując wzór na całkowanie przez części $\int_{x_i}^{x_{i+1}} f(t_{k-v-1})g'(t_{k-v-1})dt_{v-k-1} = [f(t_{v-k-1})g(t_{v-k-1})]_{x_i}^{x_{i+1}} - \int_{x_i}^{x_{i+1}} f'(t_{k-v-1})g(t_{k-v-1})dt_{v-k-1}$ dostajemy

$$
\begin{aligned}
I=&\left[t_{k-v-1}\int_{x_i}^{t_{k-v-1}}\int_{x_i}^{t_{k-v-2}}\dots\int_{x_i}^{t_1}\left(\sum_{p=0}^{q}f_p(t)u_{k,i}^{(p)}(t)\right)dtdt_1dt_2\dots dt_{k-v-2}\right]_{x_i}^{x_{i+1}}\\
&-\int_{x_i}^{x_{i+1}}t_{k-v-1}\int_{x_i}^{t_{k-v-1}}\int_{x_i}^{t_{k-v-3}}\dots\int_{x_i}^{t_1}\left(\sum_{p=0}^{q}f_p(t)u_{k,i}^{(p)}(t)\right)dtdt_1\dots dt_{k-v-3}dt_{k-v-1}\\
=&\int_{x_i}^{x_{i+1}}(x_{i+1}-t_{k-v-1})\int_{x_i}^{t_{k-v-1}}\int_{x_i}^{t_{k-v-3}}\dots\int_{x_i}^{t_1}\left(\sum_{p=0}^{q}f_p(t)u_{k,i}^{(p)}(t)\right)dtdt_1\dots dt_{k-v-3}dt_{k-v-1}.
\end{aligned}
$$

Następnie przyjmujemy
$f(t_{k-v-1}) = \int_{x_i}^{t_{k-v-1}}\int_{x_i}^{t_{k-v-3}}\dots\int_{x_i}^{t_1}\left(\sum_{p=0}^{q}f_p(t)u_{k,i}^{(p)}(t)\right)dtdt_1dt_2\dots dt_{k-v-3}$,
a $g'(t_{k-v-1}) = x_{i+1} - t_{k-v-1}$ i kolejny raz stosujemy wzór na całkowanie przez części. Po zastosowaniu tego wzoru $k-v-1$ razy dostajemy

$$
\begin{aligned}
a_{j,i+1}^{v} = u_{j,i}^{(v)}(x_{i+1}) =& \sum_{p=0}^{k-1-v}\frac{u_{j,i}^{(p+v)}(x_i)h^p}{p!}\\
&+\frac{1}{(k-1-v)!}\int_{x_i}^{x_{i+1}}(x_{i+1}-t)^{k-1-v}\left(\sum_{p=0}^{q}f_p(t)u_{j,i}^{(p)}(t)\right)dt
\end{aligned}
$$

dla $i = 0, 1, \dots, n-1$, $v = 0, 1, \dots, k-1$, $j = 0, 1, \dots, k-1$. Podobnie postępując z całką wielokrotną ze wzoru (5.17) otrzymujemy

$$
\begin{aligned}
a_{k,i+1}^{v} = u_{k,i}^{(v)}(x_{i+1}) =& \sum_{p=0}^{k-1-v}\frac{u_{k,i}^{(p+v)}(x_i)h^p}{p!}\\
&+\frac{1}{(k-1-v)!}\int_{x_i}^{x_{i+1}}(x_{i+1}-t)^{k-1-v}\left(\sum_{p=0}^{q}f_p(t)u_{k,i}^{(p)}(t)+f_k(t)\right)dt.
\end{aligned}
$$

We wzorach tych wartości $u_{j,i}^{(p+v)}(x_i)$ dane są przez wartości początkowe w (5.6), (5.7). Nieznane pochodne $u_{k,i}^{(p)}(t)$ występujące pod całką zastępujemy przez ich przybliżenia Taylora $l_{j,i}^{(p)}(t)$. W ten sposób otrzymujemy przybliżenia liczb $a_{j,i+1}^{v}$ w modelu deterministycznym z informacją liniową

$$
\begin{aligned}
\tilde{a}_{j,i+1}^{v} =& \sum_{p=0}^{k-1-v}\frac{u_{j,i}^{(p+v)}(x_i)h^p}{p!}\\
&+\frac{1}{(k-1-v)!}\int_{x_i}^{x_{i+1}}(x_{i+1}-t)^{k-1-v}\left(\sum_{p=0}^{q}f_p(t)l_{j,i}^{(p)}(t)\right)dt,\quad j=0,1,\dots,k-1,
\end{aligned}
\tag{5.18}
$$

$$\tilde{a}^v_{k,i+1} = \sum_{p=0}^{k-1-v} \frac{u^{(p+v)}_{k,i}(x_i)h^p}{p!} + \frac{1}{(k-1-v)!} \int_{x_i}^{x_{i+1}} (x_{i+1}-t)^{k-1-v} \left(\sum_{p=0}^{q} f_p(t) l^{(p)}_{k,i}(t) + f_k(t) \right) dt. \tag{5.19}$$

Informacja tutaj składa się z informacji standardowej potrzebnej do wyznaczenia wielomianów Taylora $l_{j,i}$ oraz z informacji całkowej danej przez całki we wzorach (5.18) i (5.19), a więc korzystamy z informacji liniowej.

Przejdźmy teraz do modelu randomizacyjnego i kwantowego. W tych przypadkach, liczby $\tilde{a}^v_{j,i+1}$ definiujemy jako wartości pochodnych funkcji zwróconych przez prawie optymalne algorytmy randomizacyjne lub kwantowe rozwiązujące problemy początkowe, zastosowane do problemów (5.6) i (5.7). Algorytmy te zdefiniowane są w pracy [GS08] i przedstawione w podrozdziale 2.3.2. Przypomnijmy, że algorytmy te oznaczone są przez $A^K([c,d],n,Y,\delta)$, gdzie K jest parametrem całkowitym, $[c,d]$ jest rozważanym przedziałem, $n+1$ jest liczbą równoodległych punktów siatki w $[c,d]$, Y jest wektorem wartości początkowych, a δ parametrem prawdopodobieństwa. Zdefiniujemy w następujący sposób potrzebne nam parametry. Niech $\gamma \in (0,1)$ będzie ustaloną liczbą. Przyjmujemy $K = \lceil \log\left((r+3/2)\gamma^{-1}-1\right) \rceil$, $\delta = n^{-(r+1)2^{K+1}+2}$ w modelu randomizacyjnym, a $K = \lceil (r+3)\gamma^{-1} - 2 \rceil$ i $\delta = 1/(4(k+1)n^{K+1})$ w modelu kwantowym.
Niech $\tilde{l}_{j,i}$ będzie funkcją zwróconą przez algorytm $A^K([x_i,x_{i+1}],n,Y,\delta)$, dla K i δ zdefiniowanych wyżej, przybliżającą rozwiązanie $u_{j,i}$ problemu (5.6) lub (5.7). W modelach randomizacyjnym i kwantowym wektor wartości początkowych jest dany przez $Y = [u_{j,i}(x_i), u'_{j,i}(x_i), \ldots, u^{(k-1)}_{j,i}(x_i)]^T$. Wtedy, $\tilde{a}^v_{j,i+1}$ zdefiniowane jest jako v-ta pochodna funkcji $\tilde{l}_{j,i}$ w punkcie x_{i+1}

$$\tilde{a}^v_{j,i+1} = \tilde{l}^{(v)}_{j,i}(x_{i+1}). \tag{5.20}$$

Mając przybliżenia $\tilde{a}^v_{j,i+1}$ zdefiniujemy teraz algorytmy rozwiązujące problem (5.1) we wszystkich modelach. W tym celu zastosujemy wzór (5.8), w którym nie znamy zarówno wartości s^j_i jak i funkcji $u_{j,i}$. W celu wyznaczenia przybliżeń liczb s^j_i zastępujemy liczby $a^v_{j,i+1}$ w definicjach macierzy A i wektora P przez ich przybliżenia $\tilde{a}^v_{j,i+1}$ zdefiniowane przez (5.15), (5.18), (5.19) lub (5.20). W ten sposób otrzymujemy macierz $\tilde{A}$ i wektor $\tilde{P}$. Niech $\tilde{S} = [\tilde{s}^0_0, \ldots, \tilde{s}^{k-1}_{n-1}, \tilde{s}^{v_0}_0]^T \in \mathbb{R}^{nk}$ będzie rozwiązaniem układu $\tilde{A}\tilde{S} = \tilde{P}$ (istnienie rozwiązania tego układu będzie pokazane wkrótce). Nieznane rozwiązania $u_{j,i}$ problemów początkowych (5.6), (5.7) zastępujemy w modelu deterministycznym

ich przybliżeniami Taylora $l_{j,i}$ zdefiniowanymi przez (5.14), a w modelach randomizacyjnym i kwantowym przez $\tilde{l}_{j,i}$. Wynik $\tilde{u}$ algorytmu $\phi_n^{\mathrm{det-st}}$ lub $\phi_n^{\mathrm{det-lin}}$ przybliżającego rozwiązanie u problemu (5.1) w w modelu deterministycznym jest więc zdefiniowany następująco

$$\tilde{u}(x) = \sum_{j=0}^{k-1} \tilde{s}_i^j l_{j,i}(x) + l_{k,i}(x), \qquad x \in [x_i, x_{i+1}]. \tag{5.21}$$

Natomiast wynik $\tilde{u}$ algorytmu ϕ_n^{rand} lub ϕ_n^{quant} jest zdefiniowany jako

$$\tilde{u}(x) = \sum_{j=0}^{k-1} \tilde{s}_i^j \tilde{l}_{j,i}(x) + \tilde{l}_{k,i}(x), \qquad x \in [x_i, x_{i+1}]. \tag{5.22}$$

W następnych podrozdziałach pokażemy górne ograniczenia na błąd oraz koszt tak zdefiniowanych algorytmów, a co za tym idzie, także na złożoność problemu.

5.2.5. Górne ograniczenia w modelu deterministycznym

W tej części rozważać będziemy model deterministyczny. Pokażemy górne ograniczenia na błąd oraz koszt algorytmów $\phi_n^{\mathrm{det-st}}$ i $\phi_n^{\mathrm{det-lin}}$. Ograniczenia te zawarte są w następującym twierdzeniu.

Twierdzenie 20. *Istnieją dodatnie stałe $C_e^{\mathrm{det-st}}$, $C_c^{\mathrm{det-st}}$, $C_e^{\mathrm{det-lin}}$, $C_c^{\mathrm{det-lin}}$ zależne wyłącznie od a, b, k, parametrów klasy i warunków brzegowych takie, że dla każdego $n \in \mathbb{N}$ zachodzi*

$$\mathrm{e}^{\mathrm{det-st}}(S, \phi_n^{\mathrm{det-st}}, F^r) \leqslant C_e^{\mathrm{det-st}} n^{-r}, \tag{5.23}$$

$$\mathrm{cost}^{\mathrm{det-st}}(S, \phi_n^{\mathrm{det-st}}, F^r) \leqslant C_c^{\mathrm{det-st}} n, \tag{5.24}$$

$$\mathrm{e}^{\mathrm{det-lin}}(S, \phi_n^{\mathrm{det-lin}}, F^r) \leqslant C_e^{\mathrm{det-lin}} n^{-(r+k-q)}, \tag{5.25}$$

$$\mathrm{cost}^{\mathrm{det-lin}}(S, \phi_n^{\mathrm{det-lin}}, F^r) \leqslant C_c^{\mathrm{det-lin}} n. \tag{5.26}$$

Dowód. Z (5.8) i (5.21) wynika, że błąd algorytmu może być zapisany jako $\|u - \tilde{u}\|_{[a,b]} = \max\limits_{0 \leqslant i \leqslant n-1} E_i$, gdzie

$$E_i = \|u - \tilde{u}\|_{[x_i, x_{i+1}]} = \left\| \sum_{j=0}^{k-1} s_i^j u_{j,i} + u_{k,i} - \sum_{j=0}^{k-1} \tilde{s}_i^j l_{j,i} - l_{k,i} \right\|_{[x_i, x_{i+1}]}. \tag{5.27}$$

Korzystając z nierówności trójkąta dostajemy

$$\begin{aligned} E_i \leqslant \sum_{j=0}^{k-1} \Big\{ |s_i^j - \tilde{s}_i^j| \left(\|u_{j,i}\|_{[x_i, x_{i+1}]} + \|u_{j,i} - l_{j,i}\|_{[x_i, x_{i+1}]} \right) + |s_i^j| \|u_{j,i} - l_{j,i}\|_{[x_i, x_{i+1}]} \Big\} \\ + \|u_{k,i} - l_{k,i}\|_{[x_i, x_{i+1}]}. \end{aligned} \tag{5.28}$$

Funkcja $l_{j,i}^{(v)}$ jest przybliżeniem Taylora funkcji $u_{j,i}^{(v)}$, stąd (zob. dodatek A) mamy

$$\|l_{j,i}^{(v)} - u_{j,i}^{(v)}\|_{[x_i,x_{i+1}]} = O(h^{r+k-v}) \text{ dla } v = 0, 1, \ldots, k-1. \tag{5.29}$$

Wiadomo, że (zob. np. [KS92, tw. 4.1.3]) jeśli $\|A^{-1}\|\|A - \tilde{A}\| < 1$, to rozwiązanie $\tilde{S}$ istnieje oraz

$$\|S - \tilde{S}\| \leqslant \frac{\|A^{-1}\| \left(\|P - \tilde{P}\| + \|A - \tilde{A}\|\|S\|\right)}{1 - \|A^{-1}\|\|A - \tilde{A}\|}. \tag{5.30}$$

Musimy więc znaleźć ograniczenia na $\|A-\tilde{A}\|$ i $\|P-\tilde{P}\|$. W modelu deterministycznym z informacją standardową z (5.29) mamy $|a_{j,i+1}^v - \tilde{a}_{j,i+1}^v| = |u_{j,i}^{(v)}(x_{i+1}) - l_{j,i}^{(v)}(x_{i+1})| = O(h^{r+k-v})$, $v = 0, 1, \ldots, k-1$. Każdy wiersz macierzy A i $\tilde{A}$ zawiera co najwyżej $k+1$ niezerowych elementów, więc $\|A - \tilde{A}\| = O(h^{r+1})$ oraz $\|P - \tilde{P}\| = O(h^{r+1})$ ze stałymi zależnymi wyłącznie od a, b, k i D. Zgodnie z (5.13), norma $\|S\|$ jest ograniczona przez stałą zależną od a, b, k, D i warunków brzegowych. Korzystając z (5.30) i lematu 5 otrzymujemy, że $\tilde{S}$ istnieje i $\|S - \tilde{S}\| = O(h^r)$. Oszacujemy teraz E_i z (5.28). Wiemy, że $|s_i^j - \tilde{s}_i^j| = O(h^r)$, z faktu 1 mamy $\|u_{j,i}\| = O(1)$, z (5.29) wynika, że $\|u_{j,i} - l_{j,i}\| = O(h^{r+k})$, a z (5.13), że $|s_i^j| = O(1)$. Z (5.28) dostajemy więc, że $E_i = O(h^r)$. Jako, $\|u - \tilde{u}\| = \max_{0 \leqslant i \leqslant n-1} E_i$, więc w modelu deterministycznym z informacją standardową $\|u - \tilde{u}\| = O(h^r)$, gdzie stała w notacji O zależy wyłącznie od a, b, k, D i warunków brzegowych, co dowodzi (5.23).

Koszt algorytmu w modelu deterministycznym z informacją standardową składa się z kosztu wyznaczenia wszystkich $\tilde{a}_{j,i}^v$, $i = 1, 2, \ldots, n$, $j = 0, 1, \ldots, k$, $v = 0, 1, \ldots, k-1$ oraz kosztu wyznaczenia $\tilde{u}$ mając dane $\tilde{s}_i^j$. Zgodnie z (5.15) i (5.21), koszt ten jest równy kosztowi wyznaczenia wszystkich wielomianów Taylora $l_{j,i}$, $i = 1, 2, \ldots, n$, $j = 0, 1, \ldots, k$. Koszt wyznaczenia jednego wielomianu $l_{j,i}$ jest kosztem obliczenia wartości pochodnych funkcji $u_{j,i}$ rzędu od k do $k+r-1$ co jest równoważne obliczeniu wartości funkcji f_p, $p = 0, 1, \ldots q, k$ oraz ich pochodnych do rzędu r w punkcie x_i. Tak więc koszt wyznaczenia jednego wielomianu $l_{j,i}$ jest nie większy niż $(q+2)(r+1) \leqslant (k+1)(r+1)$. Koszt całkowity, czyli koszt wyznaczenia wszystkich wielomianów $l_{j,i}$, $i = 1, 2, \ldots, n$, $j = 0, 1, \ldots, k$, jest nie większy niż

$$\text{cost}^{\text{det}-\text{st}}(S, \phi_n^{\text{det}-\text{st}}, F^r) \leqslant (r+1)(k+1)^2 n,$$

co dowodzi (5.24).

Przejdźmy do analizy algorytmu w modelu deterministycznym z informacją liniową.

Podobnie jak w przypadku informacji standardowej, korzystając z wzorów (5.18), (5.19) i (5.29) dostajemy

$$\begin{aligned}|a^v_{j,i+1} - \tilde{a}^v_{j,i+1}| &= \left| \frac{1}{(k-1-v)!} \int_{x_i}^{x_{i+1}} (x_{i+1}-t)^{k-1-v} \sum_{p=0}^{q} f_p(t) \left(u^{(p)}_{j,i}(t) - l^{(p)}_{j,i}(t) \right) dt \right| \\ &\leqslant \frac{1}{(k-1-v)!} \int_{x_i}^{x_{i+1}} h^{k-1-v} \sum_{p=0}^{q} D \left| u^{(p)}_{j,i}(t) - l^{(p)}_{j,i}(t) \right| dt \\ &= O(h^{r+2k-q-v}).\end{aligned}$$

Stąd, $\|A - \tilde{A}\| = O(h^{r+k-q+1})$ i $\|P - \tilde{P}\| = O(h^{r+k-q+1})$. Podobnie jak w przypadku informacji standardowej można wnioskować, że $\tilde{S}$ istnieje i $\|S - \tilde{S}\| = O(h^{r+k-q})$. Korzystając z (5.27) i (5.28) otrzymujemy ograniczenie $\|u - \tilde{u}\|_{[a,b]} = O(h^{r+k-q})$, co dowodzi (5.25).

Koszt, podobnie jak w przypadku standardowym, składa się z kosztu wyznaczenia wielomianów Taylora $l_{j,i}$, $i = 1, 2, \ldots, n$, $j = 0, 1, \ldots, k$, oraz z kosztu wyznaczenia wszystkich $\tilde{a}^v_{j,i}$, $i = 1, 2, \ldots, n$, $j = 0, 1, \ldots, k$, $v = 0, 1, \ldots, k-1$. Koszt wyznaczenia każdego $\tilde{a}^v_{j,i}$ wynosi 1 (jest to funkcjonał liniowy na funkcjach prawej strony), tak więc koszt wyznaczenia wszystkich $\tilde{a}^v_{j,i}$ wynosi $nk(k+1)$. Koszt wyznaczenia wielomianów $l_{j,i}$ jest nie większy niż w przypadku informacji standardowej, więc koszt wyznaczenia wszystkich $l_{j,i}$ jest nie większy niż $(r+1)(k+1)^2 n$. Tak więc, koszt całkowity jest ograniczony przez

$$\mathrm{cost}^{\mathrm{det-lin}}(S, \phi_n^{\mathrm{det-lin}}, F^r) \leqslant (r+2)(k+1)^2 n,$$

co kończy dowód twierdzenia 20. □

Z twierdzenia 20 wynika następujące twierdzenie dotyczące złożoności.

Twierdzenie 21. *Istnieją stałe $C^{\mathrm{det-st}}$ i $C^{\mathrm{det-lin}}$ zależne wyłącznie od a, b, k, parametrów klasy i warunków brzegowych takie, że dla dowolnego $\varepsilon > 0$ zachodzi*

$$\mathrm{comp}^{\mathrm{det-st}}(\varepsilon, S, F^r) \leqslant C^{\mathrm{det-st}} \varepsilon^{-1/r}, \tag{5.31}$$

$$\mathrm{comp}^{\mathrm{det-lin}}(\varepsilon, S, F^r) \leqslant C^{\mathrm{det-lin}} \varepsilon^{-1/(r+k-q)} \tag{5.32}$$

Dowód. Rozważmy najpierw przypadek informacji standardowej. Przyjmijmy $n = \left\lceil \left(C_e^{\mathrm{det-st}} / \varepsilon \right)^{1/r} \right\rceil$. Wtedy zgodnie z twierdzeniem 20 mamy

$$\mathrm{e}^{\mathrm{det-st}}(S, \phi_n^{\mathrm{det-st}}, F^r) \leqslant \varepsilon,$$

$$\mathrm{cost}^{\mathrm{det-st}}(S, \phi_n^{\mathrm{det-st}}, F^r) \leqslant C_c^{\mathrm{det-st}} n \leqslant C_c^{\mathrm{det-st}} \left((C_e^{\mathrm{det-st}})^{1/r} \varepsilon^{-1/r} + 1 \right) \leqslant C^{\mathrm{det-st}} \varepsilon^{-1/r}$$

dla pewnej stałej $C^{\mathrm{det-st}}$. To dowodzi pierwszej części twierdzenia.
Analogicznie dla informacji liniowej przyjmujemy $n = \left\lceil \left(C_e^{\mathrm{det-st}}/\varepsilon\right)^{1/(r+k-q)} \right\rceil$. Z twierdzenia 20 dostajemy

$$\mathrm{e}^{\mathrm{det-lin}}(S, \phi_n^{\mathrm{det-lin}}, F^r) \leqslant \varepsilon,$$

$$\begin{aligned}\mathrm{cost}^{\mathrm{det-lin}}(S, \phi_n^{\mathrm{det-lin}}, F^r) &\leqslant C_c^{\mathrm{det-lin}} \left((C_e^{\mathrm{det-lin}})^{1/(r+k-q)} \varepsilon^{-1/(r+k-q)} + 1 \right) \\ &\leqslant C^{\mathrm{det-lin}} \varepsilon^{-1/(r+k-q)}\end{aligned}$$

dla pewnej stałej $C^{\mathrm{det-lin}}$, co kończy dowód twierdzenia. □

5.2.6. Górne ograniczenia w modelach randomizacyjnym i kwantowym

Następujące twierdzenie przedstawia górne ograniczenia na błąd oraz koszt algorytmów zdefiniowanych w podrozdziale 5.2.4 w modelach randomizacyjnym i kwantowym.

Twierdzenie 22. *Dla dowolnego $\gamma \in (0,1)$ istnieją dodatnie stałe C_e^{rand}, C_c^{rand}, C_e^{quant}, C_c^{quant} zależne wyłącznie od a, b, k, γ, parametrów klasy i warunków brzegowych takie, że dla każdego $n \geqslant 2$ zachodzi*

$$\mathrm{e}^{\mathrm{rand}}(S, \phi_n^{\mathrm{rand}}, F^r) \leqslant C_e^{\mathrm{rand}} n^{-(r2^K + 2^{K-1} - 1)}, \tag{5.33}$$

$$\mathrm{cost}^{\mathrm{rand}}(S, \phi_n^{\mathrm{rand}}, F^r) \leqslant C_c^{\mathrm{rand}} n^{2^K} \log n, \tag{5.34}$$

$$\mathrm{e}^{\mathrm{quant}}(S, \phi_n^{\mathrm{quant}}, F^r) \leqslant C_e^{\mathrm{quant}} n^{-(r(K+1)+K-1)}, \tag{5.35}$$

$$\mathrm{cost}^{\mathrm{quant}}(S, \phi_n^{\mathrm{quant}}, F^r) \leqslant C_c^{\mathrm{quant}} n^{K+1} \log n, \tag{5.36}$$

gdzie $K = \lceil \log\left((r + 3/2)\gamma^{-1} - 1\right) \rceil$ w modelu randomizacyjnym,
a $K = \lceil (r+3)\gamma^{-1} - 2 \rceil$ w modelu kwantowym.

Dowód. Z (5.8) i (5.22), podobnie jak we wzorach (5.27) i (5.28), dostajemy ograniczenia na błąd algorytmu

$$\begin{aligned}E_i = \|u - \tilde{u}\|_{[x_i, x_{i+1}]} &= \left\| \sum_{j=0}^{k-1} s_i^j u_{j,i} + u_{k,i} - \sum_{j=0}^{k-1} \tilde{s}_i^j \tilde{l}_{j,i} - \tilde{l}_{k,i} \right\|_{[x_i, x_{i+1}]} \\ &\leqslant \sum_{j=0}^{k-1} \left\{ |s_i^j - \tilde{s}_i^j| \left(\|u_{j,i}\|_{[x_i, x_{i+1}]} + \|u_{j,i} - \tilde{l}_{j,i}\|_{[x_i, x_{i+1}]} \right) + |s_i^j| \|u_{j,i} - \tilde{l}_{j,i}\|_{[x_i, x_{i+1}]} \right\} \\ &\quad + \|u_{k,i} - \tilde{l}_{k,i}\|_{[x_i, x_{i+1}]}. \end{aligned} \tag{5.37}$$

Rozważmy model randomizacyjny. Przypomnijmy, że przybliżenie $\tilde{l}_{j,i}$ rozwiązania $u_{j,i}$ w tym modelu otrzymywane jest przez zastosowanie prawie optymalnego algorytmu $A^K([x_i, x_{i+1}], n, Y, \delta)$ z pracy [GS08] do problemów lokalnych (5.6) oraz (5.7), gdzie K jest zdefiniowane powyżej. Przypomnijmy, że Y jest wektorem warunków początkowych i $\delta = n^{-(r+1)2^{K+1}+2}$.

Aby znaleźć ograniczenia na błąd skorzystamy z (5.37). Potrzebujemy więc ograniczeń na $\|S - \tilde{S}\|$. Skorzystamy w tym celu z nierówności (5.30). Musimy wyznaczyć oszacowania na $\|A - \tilde{A}\|$ i $\|P - \tilde{P}\|$.

Korzystamy z twierdzenia 11 zastosowanego do problemów (5.6) i (5.7) na przedziale $[c, d] = [x_i, x_{i+1}]$. Z twierdzenia tego wynika, że dla $i = 0, 1, \ldots, n-1$, $j = 0, 1, \ldots, k$, dla pewnej stałej C_1 zależnej wyłącznie od a, b, k, K i parametrów klasy zachodzi

$$\sum_{v=0}^{k-1} |u_{j,i}^{(v)}(x_{i+1}) - \tilde{l}_{j,i}^{(v)}(x_{i+1})| \leqslant C_1 h^{r+1} n^{-(r(2^K-1)+2^{K-1}-1)}. \tag{5.38}$$

Ograniczenia te zachodzą z prawdopodobieństwem nie mniejszym niż

$$\mathbf{P}(A^K, n) = (1-\delta)^{\sum_{i=1}^{K-1} n^{2^i-1}} \geqslant (1-\delta)^{n^{2^K-1}} \geqslant 1 - n^{2^K-1}\delta = 1 - \delta_1,$$

gdzie $\delta_1 = n^{-(r2^{K+1}+2^K-1)}$. Koszt obliczenia przybliżenia $\tilde{l}_{j,i}$ jest ograniczony na mocy (2.35) przez

$$\operatorname{cost}(\tilde{l}_{j,i}) \leqslant C_2 n^{2^K-1} \left(\log \frac{1}{\delta}\right) \leqslant \tilde{C}_2 n^{2^K-1} \log n \tag{5.39}$$

dla pewnych stałych C_2 i $\tilde{C}_2$ zależnych wyłącznie od a, b, k, K i parametrów klasy. Ponadto, z twierdzenia 11 wynika, że zachodzi deterministycznie

$$\sum_{v=0}^{k-1} |u_{j,i}^{(v)}(x_{i+1}) - \tilde{l}_{j,i}^{(v)}(x_{i+1})| \leqslant \tilde{L} h^{r+1} n^{-r(2^K-1)} \tag{5.40}$$

dla pewnej stałej $\tilde{L}$.

Jako, że $h = (b-a)/n$, z (5.38) i (5.20) dostajemy

$$\sum_{v=0}^{k-1} |a_{j,i+1}^v - \tilde{a}_{j,i+1}^v| = \sum_{v=0}^{k-1} |u_{j,i}^{(v)}(x_{i+1}) - \tilde{l}_{j,i}^{(v)}(x_{i+1})| = O\left(n^{-(r2^K+2^{K-1})}\right).$$

Stanowi to ograniczenie na każdy element macierzy $|A - \tilde{A}|$ oraz $|P - \tilde{P}|$. Jako, że każdy wiersz macierzy $A - \tilde{A}$ zawiera co najwyżej $k+1$ niezerowych elementów, więc $\|A - \tilde{A}\| = O\left(n^{-(r2^K+2^{K-1})}\right)$ i $\|P - \tilde{P}\| = O\left(n^{-(r2^K+2^{K-1})}\right)$. Te ograniczenia są spełnione z prawdopodobieństwem co najmniej $(1-\delta_1)^{(k+1)n} \geqslant 1 - (k+1)n\delta_1$, które jest prawdopodobieństwem przecięcia $(k+1)n$ zdarzeń (5.38). Z lematu 5 mamy,

że $\|A^{-1}\| = O(n)$, a z (5.13) mamy, że $\|S\| = O(1)$. Z (5.30) dostajemy więc, że $\tilde{S}$ istnieje dla n dostatecznie dużych oraz $\|S - \tilde{S}\| = O\left(n^{-(r2^K+2^{K-1}-1)}\right)$ z prawdopodobieństwem nie mniejszym niż $1-(k+1)n\delta_1$. Skorzystamy teraz z (5.37). Z faktu 1 mamy $\|u_{j,i}\|_{[x_i,x_{i+1}]} = O(1)$, z (5.13) mamy $|s_i^j| = O(1)$, a z (5.38) mamy $\|u_{j,i} - \tilde{l}_{j,i}\|_{[x_i,x_{i+1}]} = O(n^{-(r2^K+2^{K-1})})$. Jako, że $\|u-\tilde{u}\| = \max_{0\leqslant i\leqslant n-1} E_i$, dostajemy ograniczenie na odległość pomiędzy rozwiązaniem prawdziwym u a przybliżonym $\tilde{u}$

$$\|u - \tilde{u}\|_{[a,b]} \leqslant \lambda, \tag{5.41}$$

gdzie $\lambda = Cn^{-(r2^K+2^{K-1}-1)}$, dla pewnej stałej C zależnej od a, b, k, K, parametrów klasy i warunków brzegowych, z prawdopodobieństwem nie mniejszym niż $1-(k+1)n\delta_1$.

Oszacujemy teraz błąd zdefiniowany przez (5.2). Dla dowolnych funkcji $f_0, f_1, \ldots, f_q, f_k \in F^r$ mamy

$$\mathbf{E}\|u-\tilde{u}\|^2 = \int\limits_{\Omega} \|u-\tilde{u}\|^2 d\mathbf{P}(\omega) = \int\limits_{\|u-\tilde{u}\|\leqslant\lambda} \|u-\tilde{u}\|^2 d\mathbf{P}(\omega) + \int\limits_{\|u-\tilde{u}\|>\lambda} \|u-\tilde{u}\|^2 d\mathbf{P}(\omega).$$

Pierwsza całka w powyższej nierówności jest ograniczona przez λ^2. Znajdziemy teraz ograniczenie na drugą całkę. Jako, że, zgodnie z (5.40), $|a_{j,i}^v - \tilde{a}_{j,i}^v| = O(h^{r+1}n^{-r(2^K-1)}) = O(n^{-r})$, a każdy wiersz macierzy A oraz $\tilde{A}$ zawiera co najwyżej $k+1$ niezerowych elementów, więc $\|A-\tilde{A}\| = O(n^{-r})$ oraz $\|P-\tilde{P}\| = O(n^{-r})$. Ograniczenia te zachodzą deterministycznie. Korzystając z (5.30) i lematu 5 dostajemy, że $\tilde{S}$ zawsze istnieje oraz $\|S-\tilde{S}\| = O(n^{1-r}) = O(1)$. Jako, że $\|u_{j,i} - \tilde{l}_{j,i}\| = O(n^{-r})$, a $|s_i^j|$ oraz $\|u_{j_i}\|$ są ograniczone przez stałe, więc z (5.37) zachodzi $\|u-\tilde{u}\| = O(1)$. Tak więc, funkcja pod drugą całką jest rzędu $O(1)$, a warunek $\|u-\tilde{u}\| > \lambda$ zachodzi z prawdopodobieństwem nie większym niż $(k+1)n\delta_1$. Mamy więc

$$E\|u-\tilde{u}\|^2 \leqslant c_1\left(\lambda^2 + (k+1)n\delta_1\right)$$

dla pewnej stałej c_1 niezależnej od n. Przypomnijmy, że $\delta_1 = n^{-(r2^{K+1}+2^K-1)}$. Stąd, dla pewnej stałej C_e^{rand} zależnej od a, b, k, K, parametrów klasy i warunków brzegowych otrzymujemy

$$e^{\text{rand}}(S, \phi_n^{\text{rand}}, F^r) \leqslant C_e^{\text{rand}} n^{-(r2^K+2^{K-1}-1)}.$$

Koszt całkowity jest kosztem wyznaczenia wszystkich $\tilde{a}_{j,i}^v$, $i = 1,2,\ldots,n$, $j = 0,1,\ldots,k$, $v = 0,1,\ldots,k-1$, czyli kosztem wyznaczenia wszystkich $\tilde{l}_{j,i}$. Zgodnie z (5.39) jest ograniczony przez

$$\text{cost}^{\text{rand}}(S, \phi_n^{\text{rand}}, F^r) \leqslant n(k+1)\text{cost}(\tilde{l}_{j,i})) \leqslant \tilde{C}_2(k+1)n^{2^K}\log n.$$

To kończy dowód w przypadku randomizacyjnym.

Przejdźmy do dowodu twierdzenia w przypadku kwantowym. Przybliżenia $\tilde{l}_{j,i}$ rozwiązań $u_{j,i}$ w tym modelu, podobnie jak w modelu randomizacyjnym, otrzymywane są poprzez zastosowanie algorytmu $A^K([x_i, x_{i+1}], n, Y, \delta)$ z pracy [GS08] do problemów lokalnych (5.6) oraz (5.7). Liczby K i δ są zdefiniowane w podrozdziale 5.2.4 dla danych γ, r i k, a Y jest wektorem warunków początkowych. Przypomnijmy, że $\delta = 1/(4(k+1)n^{K+1})$.

Z twierdzenia 11 dostajemy ograniczenie

$$\sum_{v=0}^{k-1} |a_{j,i+1}^v - \tilde{a}_{j,i+1}^v| = \sum_{v=0}^{k-1} |u_{j,i}^{(v)}(x_{i+1}) - \tilde{l}_{j,i}^{(v)}(x_{i+1})| \leqslant C_1 h^{r+1} n^{-(rK+K-1)} \tag{5.42}$$

dla $i = 0, 1, \ldots, n-1$, $j = 0, 1, \ldots, k$ i pewnej dodatniej stałej C_1 zależnej wyłącznie od a, b, k, K i parametrów klasy. Zachodzi to z prawdopodobieństwem nie mniejszym niż

$$\mathbf{P}(A^K, n) = (1-\delta)^{(n^K-n)/(n-1)} \geqslant (1-\delta)^{n^K} \geqslant 1 - n^K\delta = 1 - \delta_1,$$

gdzie $\delta_1 = 1/(4(k+1)n)$. Koszt obliczenia przybliżenia $\tilde{l}_{j,i}$ jest ograniczony przez

$$\mathrm{cost}(\tilde{l}_{j,i}) \leqslant C_2 n^K \log\frac{1}{\delta} \leqslant \tilde{C}_2 n^K \log n \tag{5.43}$$

dla pewnych stałych C_2 i $\tilde{C}_2$ zależnych wyłącznie od a, b, k, K i parametrów klasy. Podobnie, jak w modelu randomizacyjnym, aby oszacować błąd, skorzystamy z (5.37) i (5.30). Z (5.42) mamy, że $|a_{j,i+1}^v - \tilde{a}_{j,i+1}^v| = O\left(n^{-(r(K+1)+K)}\right)$, dla $i = 0, 1, \ldots, n-1$, $v = 0, 1, \ldots, k-1$, $j = 0, 1, \ldots, k$. Jako, że każdy wiersz macierzy $A - \tilde{A}$ zawiera co najwyżej $k+1$ niezerowych elementów, więc $\|A - \tilde{A}\| = O\left(n^{-(r(K+1)+K)}\right)$ oraz $\|P - \tilde{P}\| = O\left(n^{-(r(K+1)+K)}\right)$ z prawdopodobieństwem nie mniejszym niż $(1-\delta_1)^{(k+1)n} \geqslant 1-(k+1)n\delta_1$, które jest prawdopodobieństwem przecięcia $k+1$ zdarzeń (5.42). Korzystając z (5.30), lematu 5 i z (5.13) otrzymujemy $\|S - \tilde{S}\| = O\left(n^{-(r(K+1)+K-1)}\right)$. Następnie z (5.37) dostajemy

$$\|u - \tilde{u}\| \leqslant C n^{-(r(K+1)+K-1)}$$

z prawdopodobieństwem nie mniejszym niż $1 - (k+1)n\delta_1$. Jako, że $\delta_1 = \dfrac{1}{4(k+1)n}$, dostajemy ograniczenie na błąd

$$e^{\mathrm{quant}}(S, \phi_n^{\mathrm{quant}}, F^r) \leqslant C_e^{\mathrm{quant}} n^{-(r(K+1)+K-1)}$$

dla pewnej stałej C_e^{quant} zależnej wyłącznie od a, b, k, K, parametrów klasy i warunków brzegowych.

Znajdziemy teraz ograniczenie na koszt w modelu kwantowym. Koszt całkowity jest kosztem wyznaczenia wszystkich $\tilde{a}^v_{j,i}$, $i = 1, 2, \ldots, n$, $j = 0, 1, \ldots, k$, $v = 0, 1, \ldots, k-1$. Z lematu 1 wynika, że koszt wyznaczenia wszystkich $\tilde{a}^v_{j,i}$ jest iloczynem kosztu wyznaczenia jednego przybliżenia $\tilde{l}_{j,i}$ rozwiązania problemu lokalnego (5.6), (5.7) i liczby tych przybliżeń. Zgodnie (5.43) koszt całkowity jest więc ograniczony przez (zobacz (5.43))

$$\mathrm{cost}^{\mathrm{quant}}(S, \phi_n^{\mathrm{quant}}, F^r) = O\left(n\ \mathrm{cost}(\tilde{l}_{j,i})\right) = O\left(n^{K+1}\log n\right) \leqslant C_c^{\mathrm{quant}} n^{K+1}\log n$$

dla pewnej stałej C_c^{quant} zależnej od a, b, k, K i parametrów klasy. To kończy dowód twierdzenia. □

Na podstawie twierdzenia 22 można wywnioskować następujący wynik.

Twierdzenie 23. *Niech $\gamma \in (0,1)$. Istnieją stałe $C^{\mathrm{rand}}(\gamma)$, $C^{\mathrm{quant}}(\gamma)$ zależne od γ, a, b, k, parametrów klasy i warunków brzegowych takie, że dla dostatecznie małego $\varepsilon > 0$ zachodzi*

$$\mathrm{comp}^{\mathrm{rand}}(\varepsilon, S, F^r) \leqslant C^{\mathrm{rand}}\varepsilon^{-1/(r+1/2-\gamma)}, \tag{5.44}$$

$$\mathrm{comp}^{\mathrm{quant}}(\varepsilon, S, F^r) \leqslant C^{\mathrm{quant}}\varepsilon^{-1/(r+1-\gamma)}. \tag{5.45}$$

Dowód. Rozważmy najpierw przypadek randomizacyjny. Przyjmijmy $n = \left\lceil \left(C_e^{\mathrm{rand}}/\varepsilon\right)^{1/(r2^K+2^{K-1}-1)} \right\rceil$. Wtedy zgodnie z twierdzeniem 22 mamy

$$\mathrm{e}^{\mathrm{rand}}(S, \phi_n^{\mathrm{rand}}, F^r) \leqslant \varepsilon,$$

$$\begin{aligned}\mathrm{cost}^{\mathrm{rand}}(S, \phi_n^{\mathrm{rand}}, F^r) &\leqslant C_c^{\mathrm{rand}} n^{2^K}\log n \leqslant C_c^{\mathrm{rand}}\left((C_e^{\mathrm{rand}}/\varepsilon)^{1/(r2^K+2^{K-1}-1)} + 1\right)^{2^K+1} \\ &\leqslant C^{\mathrm{rand}}(K)\varepsilon^{-(2^K+1)/(r2^K+2^{K-1}-1)} \\ &= C^{\mathrm{rand}}(K)\varepsilon^{-1/(r+1/2-(r+3/2)/(2^K+1))}\end{aligned}$$

dla pewnej stałej $C^{\mathrm{rand}}(K)$. Przypomnijmy, że $K = \lceil \log\left((r+3/2)\gamma^{-1} - 1\right)\rceil$. Stąd dostajemy

$$\mathrm{cost}^{\mathrm{rand}}(S, \phi_n^{\mathrm{rand}}, F^r) \leqslant C^{\mathrm{rand}}(\gamma)\varepsilon^{-1/(r+1/2-\gamma)},$$

dla pewnej stałej $C^{\mathrm{rand}}(\gamma)$, co kończy dowód w przypadku randomizacyjnym.
Przejdźmy do modelu kwantowego. Przyjmijmy $n = \left\lceil (C_e^{\mathrm{quant}}/\varepsilon)^{1/(r(K+1)+K-1)} \right\rceil$. Wtedy zgodnie z twierdzeniem 22 mamy

$$\mathrm{e}^{\mathrm{quant}}(S, \phi_n^{\mathrm{quant}}, F^r) \leqslant \varepsilon,$$

$$\begin{aligned}\operatorname{cost}^{\mathrm{quant}}(S,\phi_n^{\mathrm{quant}},F^r) &\leqslant C_c^{\mathrm{quant}} n^{K+1}\log n \leqslant C_c^{\mathrm{quant}}\left((C_e^{\mathrm{quant}}/\varepsilon)^{1/(r(K+1)+K-1)}+1\right)^{K+2}\\ &\leqslant C^{\mathrm{quant}}(K)\varepsilon^{-(K+2)/(r(K+1)+K-1)}\\ &= C^{\mathrm{quant}}(K)\varepsilon^{-1/(r+1-(r+3)/(K+2))}\end{aligned}$$

dla pewnej stałej $C^{\mathrm{quant}}(K)$. Przypomnijmy, że $K=\lceil (r+3)\gamma^{-1}-2\rceil$. Stąd dostajemy

$$\operatorname{cost}^{\mathrm{quant}}(S,\phi_n^{\mathrm{quant}},F^r) \leqslant C^{\mathrm{quant}}(\gamma)\varepsilon^{-1/(r+1-\gamma)},$$

dla pewnej stałej $C^{\mathrm{quant}}(\gamma)$, co kończy dowód twierdzenia. □

W następnej części wyprowadzimy dolne ograniczenia na złożoność tego problemu.

5.2.7. Dolne ograniczenia

W tym rozdziale pokażemy dolne ograniczenia na złożoność problemu (5.1) w modelu deterministycznym z informacją standardową i liniową, modelu randomizacyjnym i kwantowym. Ograniczenia te zawarte są w następującym twierdzeniu.

Twierdzenie 24. *Istnieją dodatnie stałe $c^{\mathrm{det-st}}$, $c^{\mathrm{det-lin}}$, c^{rand}, c^{quant} i ε_0 takie, że dla dowolnego $\varepsilon\in(0,\varepsilon_0)$ złożoność problemu (5.1) jest ograniczona przez*

$$\operatorname{comp}^{\mathrm{det-st}}(\varepsilon,S,F^r) \geqslant c^{\mathrm{det-st}}\varepsilon^{-1/r}, \tag{5.46}$$

$$\operatorname{comp}^{\mathrm{det-lin}}(\varepsilon,S,F^r) \geqslant c^{\mathrm{det-lin}}\varepsilon^{-1/(r+k)}, \tag{5.47}$$

$$\operatorname{comp}^{\mathrm{rand}}(\varepsilon,S,F^r) \geqslant c^{\mathrm{rand}}\varepsilon^{-1/(r+1/2)}, \tag{5.48}$$

$$\operatorname{comp}^{\mathrm{quant}}(\varepsilon,S,F^r) \geqslant c^{\mathrm{quant}}\varepsilon^{-1/(r+1)}. \tag{5.49}$$

Dowód. Przeprowadzimy najpierw dowód w modelu deterministycznym z informacją standardową oraz w modelu randomizacyjnym i kwantowym. W celu udowodnienia dolnych ograniczeń w tych modelach, pokażemy jak można wykorzystać rozwiązanie problemu (5.1) do rozwiązania problemu całkowania, a następnie problemu sumowania liczb rzeczywistych. Następnie skorzystamy ze znanych dolnych ograniczeń dla problemu sumowania z podrozdziału 2.2. Rozważmy problem postaci (5.1) z funkcjami prawej strony $f_j(x)\equiv 0$ dla $j=0,1,\ldots,q$. Rozważamy więc problem postaci

$$\begin{cases} u^{(k)}(x)=f(x), & x\in[a,b],\\ u^{(v)}(a)=\alpha_v, & v=0,1,\ldots,k-1,\quad v\neq v_0,\\ u^{(v_0)}(b)=\beta_{v_0}, & \end{cases}\quad , \tag{5.50}$$

gdzie $f \in F^r$. Niech $v_0 \geqslant 1$. Pokażemy jak przez całkowanie równania w (5.50) k razy dostać wzór na $u(x)$. Po pierwszym całkowaniu równania od a do x dostajemy

$$u^{(k-1)}(x) = \alpha_{k-1} + \int_a^x f(t)dt.$$

Po kolejnych $k - v_0 - 2$ całkowaniach na $[a, x]$ otrzymujemy

$$u^{(v_0+1)}(x) = \sum_{i=v_0+1}^{k-1} \frac{(x-a)^{i-v_0-1}}{(i-v_0-1)!}\alpha_i + \int_a^x \int_a^{t_{v_0+2}} \dots \int_a^{t_{k-1}} f(t)dtdt_{k-1}\dots dt_{v_0+3}dt_{v_0+2}.$$

Następnie całkujemy na przedziale od b do x i dostajemy

$$u^{(v_0)}(x) = \beta_{v_0} + \sum_{i=v_0+1}^{k-1} \left(\frac{(x-a)^{i-v_0}}{(i-v_0)!} - \frac{(b-a)^{i-v_0}}{(i-v_0)!} \right) \alpha_i + \int_b^x \int_a^{t_{v_0+1}} \dots \int_a^{t_{k-1}} f(t)dtdt_{k-1}\dots dt_{v_0+2}dt_{v_0+1}.$$

Następnie całkujemy v_0 razy na przedziale od a do x. Ostatecznie otrzymujemy

$$\begin{aligned} u(x) = &\sum_{i=0}^{v_0-1} \frac{(x-a)^i}{i!}\alpha_i + \frac{(x-a)^{v_0}}{v_0!}\beta_{v_0} \\ &+ \sum_{i=v_0+1}^{k-1} \left(\frac{(x-a)^i}{i!} - \frac{(x-a)^{v_0}(b-a)^{i-v_0}}{v_0!(i-v_0)!} \right) \alpha_i \\ &+ \int_a^x \int_a^{t_1} \dots \int_a^{t_{v_0-1}} \int_b^{t_{v_0}} \int_a^{t_{v_0+1}} \dots \int_a^{t_{k-1}} f(t)dtdt_{k-1}\dots dt_2dt_1. \end{aligned} \tag{5.51}$$

Rozważmy problem obliczenia $u(b)$. Ten problem jest równoważny problemowi obliczenia całki wielokrotnej $I = \int_a^b \int_a^{t_1} \dots \int_a^{t_{v_0-1}} \int_b^{t_{v_0}} \int_a^{t_{v_0+1}} \dots \int_a^{t_{k-1}} f(t)dtdt_{k-1}\dots dt_2dt_1$ występującej we wzorze (5.51). Używając $k-1$ razy formuły całkowania przez części przekształcimy całkę wielokrotną na całkę jednokrotną z wagą. Po zastosowaniu wzoru na całkowanie przez części $v_0 - 1$ razy otrzymujemy

$$I = \int_a^b \frac{(b-t_{v_0})^{v_0-1}}{(v_0-1)!} \int_b^{t_{v_0}} \int_a^{t_{v_0+1}} \dots \int_a^{t_{k-1}} f(t)dtdt_{k-1}\dots dt_{v_0+1}dt_{v_0}.$$

Po kolejnym całkowaniu przez części dostajemy

$$\begin{aligned} I = &-\frac{(b-a)^{v_0}}{v_0!} \int_a^b \int_a^{t_{v_0+1}} \int_a^{t_{v_0+2}} \dots \int_a^{t_{k-1}} f(t)dtdt_{k-1}\dots dt_{v_0+2}dt_{v_0+1} \\ &+ \int_a^b \frac{(b-t_{v_0+1})^{v_0}}{v_0!} \int_a^{t_{v_0+1}} \int_a^{t_{v_0+2}} \dots \int_a^{t_{k-1}} f(t)dtdt_{k-1}\dots dt_{v_0+2}dt_{v_0+1}. \end{aligned}$$

Używając $k - v_0 - 1$ razy wzoru na całkowanie przez części przekształcamy obie całki w powyższym wzorze na całki z wagą. Ostatecznie dostajemy

$$I = -\int_a^b \left(\frac{(b-t)^{k-v_0-1}(b-a)^{v_0}}{(k-v_0-1)!v_0!} - \frac{(b-t)^{k-1}}{(k-1)!} \right) f(t)dt. \tag{5.52}$$

Oznaczmy funkcję wagową przez

$$w(t) = \frac{(b-t)^{k-v_0-1}(b-a)^{v_0}}{(k-v_0-1)!v_0!} - \frac{(b-t)^{k-1}}{(k-1)!}.$$

Pochodna tej funkcji wynosi

$$w'(t) = \frac{(b-t)^{k-v_0-2}\left((b-t)^{v_0}(k-v_0-2)!v_0! - (b-a)^{v_0}(k-2)!\right)}{(k-2)!(k-v_0-2)!v_0!}.$$

Zauważmy, że $(b-t)^{k-v_0-2} > 0$, $(b-t)^{v_0} < (b-a)^{v_0}$ oraz $(k-v_0-2)!v_0! \leqslant (k-2)!$ na przedziale (a,b), więc $w'(t) < 0$ na (a,b). Funkcja $w(t)$ jest więc malejąca na $[a,b]$. Jako, że $w(b) = 0$, więc $w(t) > 0$ dla $t \in (a,b)$ oraz $w(t) \geqslant w((a+b)/2) > 0$ dla $t \in [a,(a+b)/2]$.

Pokażemy, że można problem sumowania sprowadzić do problemu obliczania całki (5.52). Niech $x_1, x_2, \ldots, x_n$ będzie ciągiem liczb rzeczywistych z przedziału $[0,1]$. Chcemy obliczyć $1/n \sum_{i=1}^n x_i$. Niech $\varepsilon_1 > 0$ będzie pewnym dostatecznie małym parametrem, który sprecyzujemy później. Klasa F^r zawiera $n = \Theta\left(\varepsilon_1^{-1/r}\right)$ funkcji f_i, $i = 1, 2, \ldots, n$, o rozłącznych nośnikach zawartych w przedziale $[a,(a+b)/2]$ takich, że $\int_a^b f_i(x)dx = \varepsilon_1^{1+1/r}$ (zobacz lemat 9 w dodatku A). Niech $H = w((a+b)/2)$ i $H_i = \int_a^b w(x)f_i(x)dx \geqslant H\varepsilon_1^{1+1/r}$. Zdefiniujmy funkcje $h_i(x) = H\varepsilon_1^{1+1/r} f_i(x)/H_i$. Jako, że $H\varepsilon_1^{1+1/r}/H_i \leqslant 1$, $i = 1, 2, \ldots, n$, i funkcje f_i należą do klasy F^r, więc funkcje h_i również należą do klasy F^r.

Funkcja dana jako $f_{\varepsilon_1}(x) = \sum_{i=1}^n x_i h_i(x)$ należy do klasy F^r. Załóżmy, że algorytm ϕ oblicza $\int_a^b w(x)g(x)dx$ z błędem nie większym niż ε i kosztem N dla dowolnej funkcji $g \in F^r$. Dla $g = f_{\varepsilon_1}$ mamy, że $\int_a^b w(x)f_{\varepsilon_1}(x)dx = H\varepsilon_1^{1+1/r} \sum_{i=1}^n x_i$. Tak więc, korzystając z algorytmu ϕ można obliczyć również średnią $\frac{1}{n}\sum_{i=1}^n x_i$ z błędem nie większym niż $\frac{\varepsilon}{Hn\varepsilon_1^{1+1/r}}$.

Skorzystamy teraz z dolnych ograniczeń na złożoność problemu obliczania średniej z n liczb z przedziału $[0,1]$ (zob. podrozdział 2.2).

Rozważmy model deterministyczny z informacją standardową. Przypomnijmy, że $n = \Theta\left(\varepsilon_1^{-1/r}\right)$. Niech G będzie stałą w dolnym ograniczeniu w notacji Θ. Przyjmijmy dla dostatecznie małego ε, że $\varepsilon_1 = 4\varepsilon/(GH)$. Korzystając z (2.6) dostajemy, że koszt algorytmu ϕ jest ograniczony przez

$$\begin{aligned} N &= \Omega\left(n\left(1 - \frac{2\varepsilon}{Hn\varepsilon_1^{1+1/r}}\right)\right) = \Omega\left(n - \frac{2\varepsilon}{H\varepsilon_1^{1+1/r}}\right) \\ &= \Omega\left(G\varepsilon_1^{-1/r} - \frac{1}{2}G\varepsilon_1\varepsilon_1^{-(r+1)/r}\right) = \Omega\left(\frac{1}{2}G\varepsilon_1^{-1/r}\right) = \Omega\left(\varepsilon^{-1/r}\right). \end{aligned}$$

To stanowi dolne ograniczenie na koszt dowolnego algorytmu obliczającego całkę z wagą I. Koszt dowolnego algorytmu rozwiązującego problem (5.1) musi wynosić przynajmniej tyle co koszt obliczania takiej całki. Tak więc dostaliśmy dolne ograniczenie na złożoność problemu (5.1) w modelu deterministycznym z informacją standardową.

W przypadku modelu randomizacyjnego przyjmijmy $\varepsilon_1 = \varepsilon^{2r/(2r+1)}$. Korzystając z (2.7) dostajemy

$$N = \Omega\left(\min\left\{n, \left(Hn\varepsilon_1^{(r+1)/r}\varepsilon^{-1}\right)^2\right\}\right) = \Omega\left(\varepsilon^{-2/(2r+1)}\right) = \Omega\left(\varepsilon^{-1/(r+1/2)}\right).$$

Stanowi to dolne ograniczenie na złożoność problemu (5.1) w modelu randomizacyjnym.

Rozważmy model kwantowy. Przyjmijmy $\varepsilon_1 = \varepsilon^{r/(r+1)}$. Z (2.8) wnioskujemy

$$N = \Omega\left(\min\left\{n, Hn\varepsilon_1^{(r+1)/r}\varepsilon^{-1}\right\}\right) = \Omega\left(\varepsilon^{-1/(r+1)}\right)$$

co stanowi dolne ograniczenie na złożoność problemu (5.1) w modelu kwantowym.

Mamy więc żądane ograniczenia na złożoność problemu (5.1) w modelach deterministycznym z informacją standardową, randomizacyjnym i kwantowym. Przejdźmy teraz do oszacowań dolnych w modelu deterministycznym z informacja liniową. Załóżmy, że algorytm ϕ korzysta z M funkcjonałów informacji $L_1(f), \ldots, L_M(f)$ (wybranych adaptacyjnie lub nie). Skonstruujemy dwie funkcje prawej strony f_1 i f_2 z klasy F^r takie, że $L(f_1) = L(f_2),\ i = 1, 2, \ldots, M$ i odległość pomiędzy odpowiednimi rozwiązaniami u_1 i u_2 jest możliwie duża. Niech $h_0, h_1, \ldots, h_M$ będą funkcjami z klasy $C^{r+k}(\mathbb{R})$ o rozłącznych nośnikach zawartych w przedziale $[a,b]$ takimi, że $\sup_{x\in[a,b]} |h_i(x)| = \Theta\left(M^{-(r+k)}\right)$, gdzie stała w notacji Θ nie zależy od i, a tylko od a, b, k, r, D. Ponadto, $h_i^{(j)}(a) = h_i^{(j)}(b) = 0$ i $\|h_i^{(j)}\| \leqslant D/2$ dla $i = 0, 1, \ldots, M$, $j = 0, 1, \ldots, k+r$. Istnienie takich funkcji wynika z lematu 8 w dodatku A. Niech $f_1(x) \equiv D/2 \in F^r$

i $f_2(x) = f_1(x) + \sum_{i=0}^{M} a_i h_i^{(k)}(x)$. Liczby a_i wybrane są tak, że $L_j(f_1) = L_j(f_2)$ dla $j = 1, 2, \ldots, M$ i $\max_{i=0,1,\ldots,M} |a_i| = |a_s| = 1$ dla pewnego s. Takie liczby a_i istnieją ponieważ układ równań liniowych $\sum_{i=0}^{M} a_i L_j(h_i^{(k)}) = 0,\ j = 1, 2, \ldots, M$ ma nietrywialne rozwiązanie $[a_0, a_1, \ldots, a_M]^T$. Przy tych warunkach $f_2(x) \geqslant 0$ dla $x \in [a, b]$ oraz wszystkie pochodne f_2 rzędu $0, 1, \ldots, r$ są ograniczone przez D. Funkcja f_2 należy więc do klasy F^r. Niech u_1 i u_2 będą rozwiązaniami problemu (5.50) z funkcjami prawej strony równymi odpowiednio f_1 i f_2. Zauważmy, że $\left(\sum_{i=0}^{M} a_i h_i\right)^{(j)}(a) = \left(\sum_{i=0}^{M} a_i h_i\right)^{(j)}(b) = 0$ dla $j = 0, 1, \ldots, k-1$. Tak więc, $u_2(x) = u_1(x) + \sum_{i=0}^{M} a_i h_i(x)$. Stąd dostajemy

$$\sup_{x\in[a,b]} |u_2(x) - u_1(x)| = \sup_{x\in[a,b]} \left|\sum_{i=0}^{M} a_i h_i(x)\right| = |a_s| \sup_{x\in[a,b]} |h_s(x)| = \Theta\left(M^{-(r+k)}\right).$$

Wynika z tego, że każdy algorytm rozwiązujący problem (5.1) korzystający z informacji liniowej o koszcie nie większym niż M ma błąd nie mniejszy niż $CM^{-(r+k)}$, dla pewnej dodatniej stałej C niezależnej od M. Jeśli chcemy, by błąd był nie większy niż ε, to musimy użyć $(\varepsilon/C)^{-1/(r+k)}$ funkcjonałów informacji. To kończy dowód w przypadku modelu deterministycznego z informacja liniową. □

W modelu deterministycznym z informacją liniową pozostaje luka pomiędzy ograniczeniami na złożoność z góry i z dołu. Ograniczenia te są zgodne wyłącznie dla $q = 0$. W pozostałych modelach otrzymane ograniczenia są ostre (lub prawie ostre).

5.2.8. Podsumowanie

Zbadaliśmy złożoność dwupunktowego problemu brzegowego postaci (5.1). Rozważyliśmy cztery modele: deterministyczny z informacja standardową i liniową, randomizacyjny i kwantowy. Dla każdego z modeli skonstruowaliśmy algorytm rozwiązujący ten problem i oszacowaliśmy jego błąd. Wyznaczyliśmy również dolne ograniczenia na złożoność tego problemu we wszystkich czterech przypadkach. Daje nam to możliwość porównania złożoności tego problemu w różnych modelach obliczeniowych. Tabela 5.1 przedstawia uzyskane ograniczenia na złożoność. Dla zwiększenia przejrzystości pominęliśmy stałe.

Z otrzymanych rezultatów wynika, że obliczenia w modelu randomizacyjnym są szybsze od obliczeń w modelu deterministycznym z informacją standardową o 1/2 w wykładniku, natomiast w modelu kwantowym o 1. Modele randomizacyjny i kwantowy

model	ograniczenia z dołu	ograniczenia z góry
det-st	$\varepsilon^{-\frac{1}{r}}$	$\varepsilon^{-\frac{1}{r}}$
det-lin	$\varepsilon^{-\frac{1}{r+k}}$	$\varepsilon^{-\frac{1}{r+k-q}}$
rand	$\varepsilon^{-\frac{1}{r+1/2}}$	$\varepsilon^{-\frac{1}{r+1/2-\gamma}}$
quant	$\varepsilon^{-\frac{1}{r+1}}$	$\varepsilon^{-\frac{1}{r+1-\gamma}}$

Tabela 5.1. Ograniczenia na złożoność

nie dają natomiast przewagi nad modelem deterministycznym z informacją liniową. Ostatni model daje przyspieszenie o $k-q$ (czyli przynajmniej o 1) w stosunku do modelu deterministycznego z informacją standardową.

Otrzymane ograniczenia są ostre w modelu deterministycznym z informacją standardową, prawie ostre w modelach randomizacyjnym i kwantowym. W modelu deterministycznym z informacją liniową dolne i górne ograniczenia są zgodne tylko gdy $q=0$, czyli gdy prawa strona nie zależy od pochodnych rozwiązania. W pozostałych przypadkach ograniczenia z góry i z dołu różnią się o q w wykładniku.

W modelu deterministycznym z informacją liniową ograniczenia na złożoność zależą od rzędu równania k. W pozostałych modelach zależą wyłącznie od regularności funkcji prawej strony.

W szczególnych przypadkach uzyskane ograniczenia w modelu deterministycznym są zgodne ze znanymi wynikami przedstawionymi w podrozdziale 5.2.2.

5.3. Nieliniowy dwupunktowy problem brzegowy

5.3.1. Sformułowanie problemu i znane wyniki

W tej części zajmiemy się ogólniejszym problemem niż poprzednio. Jak w pracy [Kac02] będziemy rozważać nieliniowy dwupunktowy problem brzegowy postaci

$$\begin{cases} z'(x) = f(x, z(x)), & x \in [a,b], \\ p(z(a), z(b)) = 0, \end{cases} \tag{5.53}$$

gdzie $a<b$, $f : [a,b] \times \mathbb{R}^d \to \mathbb{R}^d$, $z : [a,b] \to \mathbb{R}^d$, a $p : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$. Zakładamy, że problem (5.53) ma jednoznaczne rozwiązanie. Zakładamy ponadto, że funkcja f należy do klasy F_d^r funkcji r-krotnie różniczkowalnych, o pochodnych cząstkowych

ograniczonych przez stałą danej przez:

$$F_d^r = \{f : [a,b] \times \mathbb{R}^d \to \mathbb{R}^d | f \in C^r([a,b] \times \mathbb{R}^d), \|D^{(i)} f\| \leqslant D,\ i = 0, 1, \ldots r\},$$

gdzie $r \geqslant 1$, D jest nieujemną stałą, $D^{(i)}$ przebiega przez zbiór wszystkich pochodnych cząstkowych rzędu i, a $\|\cdot\| = \|\cdot\|_\infty$. Dodatkowo zakładamy, że

(Zał. 1) $p = p(s,w)$ ma ciągłe, ograniczone pochodne cząstkowe rzędu pierwszego i drugiego na $\mathbb{R}^d \times \mathbb{R}^d$.

W dalszej części przedstawimy dodatkowe założenie o funkcjach p i f.

Problem (5.53) sprowadza się do następującego problemu rozwiązywania równania nieliniowego:

$$F(s) = p(s, y(b,s)) = 0, \tag{5.54}$$

gdzie $y(x,s)$ jest rozwiązaniem problemu początkowe danego przez

$$\begin{cases} y'(x) = f(x, y(x)), & x \in [a,b], \\ y(a) = s. \end{cases} \tag{5.55}$$

Zakładamy, że istnieje jednoznaczne rozwiązanie s^* problemu (5.54). Łatwo zauważyć, że s^* jest rozwiązaniem problemu (5.54) wtedy i tylko wtedy, gdy $z(x) = y(x, s^*)$ jest rozwiązaniem problemu (5.53).

Do rozwiązywania (5.54) będziemy stosować metodę Newtona. Przyjmiemy więc założenia gwarantujące, że metoda Newtona będzie dobrze zdefiniowana i zbieżna kwadratowo do rozwiązania s^*. Zakładamy, że funkcje f są takie, że

(Zał. 2) macierz $F'(s^*)$ jest nieosobliwa i $\|F'(s^*)^{-1}\| \leqslant H$, gdzie H jest pewną dodatnią stałą.

Oznaczmy podklasę klasy F_d^r funkcji spełniających (Zał. 2) przez $\tilde{F}_d^r$.

Wartość operatora rozwiązania S dla tego problemu na elemencie wejściowym f jest równa rozwiązaniu z problemu brzegowego (5.53). Naszym celem jest przybliżenie funkcji $S(f) = z$ dla dowolnej $f \in \tilde{F}_d^r$ z pewną dokładnością $\varepsilon > 0$, przy czym błąd mierzony jest w normie supremum (patrz niżej).

Zadanie to było badane w modelu deterministycznym ([Kac02]). W tym rozdziale rozważać będziemy modele randomizacyjny i kwantowy.

Ogólne definicje informacji, błędu, kosztu i złożoności przedstawione zostały w rozdziale 1. Dla porządku wypiszemy teraz definicje błędów w poszczególnych modelach

dla zadania (5.53). W modelu deterministycznym wynikiem algorytmu jest pewna ograniczona funkcja l. Błąd algorytmu ϕ w klasie $\tilde{F}_d^r$ definiowany jest jako

$$\mathrm{e}^{\mathrm{det-st/det-lin}}(S,\phi,\tilde{F}_d^r) = \sup_{f\in\tilde{F}_d^r}\sup_{x\in[a,b]}|z(x)-l(x)|.$$

W modelu randomizacyjnym i kwantowym wynikiem algorytmu jest pewna losowa funkcja l^ω. Błąd w modelu randomizacyjny definiowany jest następująco

$$\mathrm{e}^{\mathrm{rand}}(S,\phi,\tilde{F}_d^r) = \sup_{f\in\tilde{F}_d^r}\left(\mathbf{E}\left(\sup_{x\in[a,b]}|z(x)-l^\omega(x)|\right)^2\right)^{1/2}. \tag{5.56}$$

W modelu kwantowym rozważamy błąd probabilistyczny z prawdopodobieństwem porażki 1/4 definiowany jako

$$\mathrm{e}^{\mathrm{quant}}(S,\phi,\tilde{F}_d^r) = \sup_{f\in\tilde{F}_d^r}\inf\left\{\alpha : \mathbf{P}\left(\sup_{x\in[a,b]}|z(x)-l^\omega(x)|>\alpha\right)\leqslant 1/4\right\}.$$

We wszystkich modelach koszt definiowany jest podobnie jak w rozdziale 1 jako liczność informacji i oznaczony jest przez $\mathrm{cost}^{\mathrm{det/rand/quant}}(S,\phi,\tilde{F}_d^r)$.

Dla $\varepsilon>0$ złożoność zadania definiowana jest przez (1.6) .

Znane są ograniczenia na złożoność tego problemu w modelu deterministycznym. W pracy [Kac02] udowodnione jest następujące twierdzenie

Twierdzenie 25 ([Kac02]). *Dla problemu (5.53) mamy:*

- *dla dowolnej $p(s,w)$ spełniającej (Zał. 1) oraz $r\geqslant 2$ istnieje dodatnia stała K_1 taka, że dla dowolnego dostatecznie małego $\varepsilon>0$, dla informacji standardowej*

  $$\mathrm{comp}^{\mathrm{det-st}}(\varepsilon,S,\tilde{F}_d^r)\leqslant K_1\left(\frac{1}{\varepsilon}\right)^{1/r}\log\log\frac{1}{\varepsilon},$$

  a dla informacji liniowej

  $$\mathrm{comp}^{\mathrm{det-lin}}(\varepsilon,S,\tilde{F}_d^r)\leqslant K_1\left(\frac{1}{\varepsilon}\right)^{1/(r+1)}\log\log\frac{1}{\varepsilon},$$

- *dla $p(s,w)=s-\eta$ oraz $r\geqslant 1$ istnieje dodatnia stała K_2 taka, że dla dowolnego dostatecznie małego $\varepsilon>0$*

  $$\mathrm{comp}^{\mathrm{det-st}}(\varepsilon,S,\tilde{F}_d^r)\geqslant K_2\left(\frac{1}{\varepsilon}\right)^{1/r},$$

  $$\mathrm{comp}^{\mathrm{det-lin}}(\varepsilon,S,\tilde{F}_d^r)\geqslant K_2\left(\frac{1}{\varepsilon}\right)^{1/(r+1)}.$$

Z tego twierdzenia wynika, że dolne i górne ograniczenia są zgodne z dokładnością do czynnika logarytmicznego, jeśli dopuszczamy dowolną funkcję $p(s, w)$.

W pracy [Kac02] rozważany jest również szczególny skalarny problem brzegowy drugiego rzędu postaci

$$\begin{cases} z''(x) = f(x, z(x)), & x \in [a, b], \\ z(a) = A, \\ z(b) = B, \end{cases} \tag{5.57}$$

gdzie $a < b$, $f : [a, b] \times \mathbb{R} \to \mathbb{R}$, a f ma r ciągłych, ograniczonych pochodnych na $[a, b] \times \mathbb{R}$. Dla tego problemu oszacowania w pewnych przypadkach mogą być poprawione w stosunku do ogólnego problemu (5.53). Mianowicie w pracy [Kac02] pokazane jest, że górne ograniczenia dla informacji liniowej mogą być poprawione o jeden rząd wielkości. Niech $\mathcal{G}$ będzie podklasą F_1^r daną przez

$$\begin{aligned} \mathcal{G} = \{ \;\; & f = f(x, u) : \; f \in C^r([a, b] \times \mathbb{R}), \; \|D^{(i)} f\| \leqslant D, \\ & i = 0, 1, \ldots r, \; r \geqslant 1 \;\; \text{oraz} \;\; \frac{\partial f}{\partial u}(x, u) \geqslant 0 \text{ na } [a, b] \times \mathbb{R}\}. \end{aligned} \tag{5.58}$$

Oznaczmy operator rozwiązania dla tego problemu przez S^*. Ograniczenia na złożoność przedstawione są w następującym twierdzeniu.

Twierdzenie 26 ([Kac02]). *Dla problemu (5.57) istnieją dodatnie stałe K_1 i K_2 takie, że dla każdego dostatecznie małego $\varepsilon > 0$*

- *dla $r \geqslant 2$*

$$\operatorname{comp}^{\mathrm{det-st}}(\varepsilon, S^*, \mathcal{G}) \leqslant K_1 \left(\frac{1}{\varepsilon}\right)^{1/r} \log\log\frac{1}{\varepsilon},$$

$$\operatorname{comp}^{\mathrm{det-lin}}(\varepsilon, S^*, \mathcal{G}) \leqslant K_1 \left(\frac{1}{\varepsilon}\right)^{1/(r+2)} \log\log\frac{1}{\varepsilon},$$

- *dla $r \geqslant 1$*

$$\operatorname{comp}^{\mathrm{det-st}}(\varepsilon, S^*, \mathcal{G}) \geqslant K_2 \left(\frac{1}{\varepsilon}\right)^{1/r},$$

$$\operatorname{comp}^{\mathrm{det-lin}}(\varepsilon, S^*, \mathcal{G}) \geqslant K_2 \left(\frac{1}{\varepsilon}\right)^{1/(r+2)}.$$

Przeanalizujemy złożoność problemów (5.53) oraz (5.57) w modelu randomizacyjnym i kwantowym. Dla (5.57) pokażemy oszacowania z dołu wskazujące, że problem ten nie jest łatwiejszy od ogólnego zagadnienia (5.53). Porównamy otrzymane rezultaty z ograniczeniami z twierdzeń 25 i 26 na złożoność tych problemów w modelu deterministycznym.

W następnym podrozdziale zdefiniujemy algorytmy, które posłużą do wyznaczenia górnych ograniczeń na złożoność w modelach randomizacyjnym i kwantowym.

5.3.2. Algorytmy

Rozpatrujemy ogólny problem (5.53). Skonstruujemy dwa algorytmy, randomizacyjny i kwantowy, które posłużą do udowodnienia górnych ograniczeń na złożoność. Będziemy rozwiązywać problem (5.53) przez rozwiązanie problemu stowarzyszonego (5.54), a do rozwiązania (5.54) użyjemy zaburzonej metody Newtona. Zakładamy, że funkcja f należy do klasy $\tilde{F}_d^r$ zdefiniowanej w poprzednim podrozdziale. W [Kac02, dodatek A2] udowodnione jest, że założenia w definicji klasy $\tilde{F}_d^r$ gwarantują, że funkcja F jest dwukrotnie różniczkowalna w sposób ciągły na $\mathbb{R}^d$ oraz że istnieje stała Z zależna wyłącznie od D, a, b i ograniczeń na pochodne cząstkowe funkcji p pierwszego i drugiego rzędu taka, że

$$\|F''(s)\| \leqslant Z \quad \text{dla każdego } s. \tag{5.59}$$

Niech H będzie stałą zdefiniowaną w (Zał. 2), $C = HZ$ oraz $R = 1/(4C)$. Można pokazać, używając standardowych metod, że dla dowolnego punktu startowego $\tilde{s}^0$ należącego do kuli $K(s^*, R) = \{s \in \mathbb{R}^d : \|s - s^*\| \leqslant R\}$, ciąg Newtona

$$\tilde{s}^{l+1} = \tilde{s}^l - F'(\tilde{s}^l)^{-1} F(\tilde{s}^l), \qquad l = 0, 1, \ldots \tag{5.60}$$

jest dobrze zdefiniowany oraz leży w kuli $K(s^*, R)$. Ponadto, $\lim_{l\to\infty} \tilde{s}^l = s^*$ i

$$\|\tilde{s}^{l+1} - s^*\| \leqslant C\|\tilde{s}^l - s^*\|^2, \quad l = 0, 1, \ldots. \tag{5.61}$$

Tak więc, ciąg ten jest zbieżny co najmniej kwadratowo do rozwiązania.

Aby użyć metody Newtona do problemu (5.54) musimy umieć obliczać $F(s)$ i $F'(s)$ w pewnych punktach s. W celu obliczenia $F(s)$ musimy wyznaczyć przybliżenie rozwiązania $y(b, s)$ problemu (5.55). Niech $\varepsilon, \gamma \in (0, 1)$ będą parametrami algorytmów. Zdefiniujmy $k = \lceil 2 \log\log(1/\varepsilon) \rceil$, $\delta = \varepsilon^2/(2k + 1)$ w modelu randomizacyjnym i $\delta = 1/(8k + 4)$ w modelu kwantowym. Dla ustalonego $s \in K(s^*, R)$ niech $l(x, s)$ będzie przybliżeniem rozwiązania $y(x, s)$ problemu początkowego (5.55) z dokładnością ε z prawdopodobieństwem nie mniejszym niż $1 - \delta$. Przybliżenie to jest wyznaczane za pomocą prawie optymalnych algorytmów randomizacyjnego lub kwantowego do rozwiązywania problemów początkowych przedstawionych w [Kac06]. Własności tych algorytmów opisane są w rozdziale 2.3. Algorytmy te były definiowane dla funkcji z

klasy Höldera $F_d^{r,\rho}$, jednak można je również zastosować do funkcji z klasy F_d^r i w oszacowaniach przyjąć formalnie $\rho = 0$.

Zauważmy, że w modelu randomizacyjnym można by użyć optymalnych algorytmów z pracy [HM08], jednak nie ma podobnych wyników dotyczących optymalności dla modelu kwantowego. Poza tym, w [HM08] nie ma wyniku, który w prosty sposób pozwalałby skonstruować algorytm dla funkcji kawałkami regularnych, a takich algorytmów będziemy potrzebować w dalszej części pracy.

Stosując wniosek 1 z podrozdziału 2.3 z $\gamma := \gamma/2$ otrzymujemy, że z prawdopodobieństwem co najmniej $1 - \delta$,

$$\|l(x,s) - y(x,s)\| \leqslant \varepsilon \quad \text{dla każdego } x \in [a,b]. \tag{5.62}$$

Dla ustalonego s, funkcja $l(x,s)$ jest funkcją kawałkami wielomianową na odcinkach długości $(b-a)/M$, gdzie $M = \Theta\left(\varepsilon^{-1/(r+1/2-\gamma/4)}\right)$ w modelu randomizacyjnym i $M = \Theta\left(\varepsilon^{-1/(r+1-\gamma/4)}\right)$ w modelu kwantowym, a stałe w notacji Θ zależą od a, b, γ i parametrów klasy (zob. podrozdział 2.3). Koszt wyznaczenia l dla ustalonego s można oszacować korzystając z (2.14) i (2.15). Oznaczmy ten koszt przez $\mathrm{cost}(l,s,\varepsilon,\delta)$ z odpowiednim indeksem górnym 'rand' lub 'quant' w zależności od modelu. Zgodnie z (2.14) (dla $\gamma := \gamma/2$ i $\rho = 0$) koszt ten dla dowolnego s jest ograniczony przez

$$\mathrm{cost}^{\mathrm{rand}}(l,s,\varepsilon,\delta) \leqslant K_1 \left(\frac{1}{\varepsilon}\right)^{1/(r+1/2-\gamma/2)} \log\frac{1}{\delta} \tag{5.63}$$

oraz zgodnie z (2.15)

$$\mathrm{cost}^{\mathrm{quant}}(l,s,\varepsilon,\delta)) \leqslant K_2 \left(\frac{1}{\varepsilon}\right)^{1/(r+1-\gamma/2)} \log\frac{1}{\delta}, \tag{5.64}$$

gdzie K_1, K_2 są stałymi zależnymi od a, b, γ i parametrów klasy.

Pokażemy teraz jak wyznaczyć przybliżenia $F'(s)$. Zauważmy, że

$$F'(s) = \frac{\partial p}{\partial s}(s, y(b,s)) + \frac{\partial p}{\partial w}(s, y(b,s))\frac{\partial y}{\partial s}(b,s).$$

Musimy wyznaczyć przybliżenie wielkości $\frac{\partial y}{\partial s}(b,s)$. Niech $Y(x,s) = \frac{\partial y}{\partial s}(x,s)$. Zauważmy, że

$$\begin{aligned}\frac{\partial}{\partial x}Y(x,s) &= \frac{\partial}{\partial s}\frac{\partial y}{\partial x}(x,s) = \frac{\partial}{\partial s}f(x,y(x,s)) = \frac{\partial f}{\partial y}(x,y(x,s))\frac{\partial y}{\partial s}(x,s)\\ &= \frac{\partial f}{\partial y}(x,y(x,s))Y(x,s).\end{aligned}$$

Ponadto, ponieważ $y(a,s) = s$, więc $Y(a,s) = I$, gdzie I oznacza macierz identyczności. Z powyższego wynika, że macierz $Y(x,s)$ jest rozwiązaniem problemu początkowego

$$\begin{cases} Y'(x,s) = \frac{\partial f}{\partial y}(x,y(x,s))Y(x,s), & x \in [a,b],\\ Y(a,s) = I. \end{cases} \tag{5.65}$$

W równaniu tym zastąpimy nieznaną wartość $y(x,s)$ przez jej przybliżenie $l(x,s)$. Dostajemy nowy problem początkowy

$$\begin{cases} \bar{Y}'(x,s) = \frac{\partial f}{\partial y}(x, l(x,s))\bar{Y}(x,s), & x \in [a,b], \\ \bar{Y}(a,s) = I. \end{cases} \tag{5.66}$$

Funkcja $\frac{\partial f}{\partial y}$ jest ograniczona, więc prawa strona równania w (5.66) spełnia warunek Lipschitza ze względu na $\bar{Y}$. Jednak, jako, że funkcja $l(x,s)$ jest kawałkami wielomianowa na przedziałach długości $(b-a)/M$, więc prawa strona równania (5.66) nie jest ciągła ze względu na x. Istnieje jednoznaczne rozwiązanie tego problemu w klasie funkcji ciągłych i przedziałami różniczkowalnych (przez pochodną w punktach nieciągłości będziemy rozumieć pochodną prawostronną).

Funkcja prawej strony $g(x,y) = \frac{\partial f}{\partial y}(x, l(x,s))y$ jest funkcją nieograniczoną na $\mathbb{R} \times \mathbb{R}^{d^2}$. Aby skorzystać z twierdzenia 10 potrzebna jest ograniczoność funkcji prawej strony oraz jej pochodnych cząstkowych na $[a,b] \times \mathbb{R}^{d^2}$. Funkcja ta jest kawałkami regularna. Można ją zapisać w postaci $g(x,y) = g_i(x,y) = \frac{\partial f}{\partial y}(x, l_i(x,s))y$ dla $x \in [t_i, t_{i+1})$, gdzie $t_i = a + i(b-a)/M$, $l_i(x,s) = l(x,s)$ dla $x \in [t_i, t_{i+1})$, $i = 0, 1, \ldots, M$. Funkcje g_i są klasy $C^{r-1}\left([t_i, t_{i+1}) \times \mathbb{R}^{d^2}\right)$.

Pokażemy, podobnie jak w [Dau11], jak funkcje $g_i = g_i(x,y)$ przedłużyć do ograniczonych funkcji z klasy $C^{r-1}\left([a,b] \times \mathbb{R}^{d^2}\right)$. Rozwiązanie $\bar{Y}$ jest ograniczone na $[a,b]$ przez pewną stałą c_0 zależną od a, b i parametrów klasy $\tilde{F}_d^r$. Zdefiniujmy funkcję $\psi : \mathbb{R}^{d^2} \to \mathbb{R}^{d^2}$ klasy $C^r\left(\mathbb{R}^{d^2}\right)$, tak by $\psi(y) = y$ dla $\|y\| \leqslant 2c_0$ oraz $\psi(y) = 0$ dla $\|y\| \geqslant 3c_0$. Definiujemy nowe funkcje prawej strony $\tilde{g}_i(x,y) = \frac{\partial f}{\partial y}(x, l_i(x,s))\psi(y)$. Funkcje te mają ograniczone pochodne rzędu do $r-1$. Dla $x \in [a,b]$, $y \in \mathbb{R}^{d^2}$, niech $\tilde{g}(x,y) = \tilde{g}_i(x,y)$ o ile $x \in [t_i, t_{i+1})$. Wtedy problem

$$\begin{cases} \bar{Y}'(x,s) = \tilde{g}(x, \bar{Y}(x,s)), & x \in [a,b], \\ \bar{Y}(a,s) = I \end{cases} \tag{5.67}$$

ma takie to samo rozwiązanie co problem (5.66). Ponieważ funkcja $\tilde{g}$ ma postać (2.22), możemy skorzystać z twierdzenia 10 z rozdziału 2.3 dla problemu (5.67) (w twierdzeniu tym rozpatrywana jest klasa Höldera, jednak możne je również zastosować do klasy funkcji o ograniczonych pochodnych, przyjmując formalnie $\rho = 0$). Przybliżenie otrzymane przez zastosowanie algorytmów, o których mowa w tym twierdzeniu będzie również przybliżać rozwiązanie problemu (5.66).

Dla ustalonego $s \in K(s^*, R)$ niech $L(x, s)$ będzie przybliżeniem rozwiązania $\bar{Y}(x, s)$ problemu (5.66) z dokładnością $\sqrt[3]{\varepsilon}$, z prawdopodobieństwem co najmniej $1 - \delta$, obliczonym algorytmem, o którym mowa w twierdzeniu 10. Zakładamy, że z prawdopodobieństwem co najmniej $1 - \delta$ zachodzi (w modelu randomizacyjnym lub kwantowym)

$$\|L(x, s) - \bar{Y}(x, s)\| \leqslant \sqrt[3]{\varepsilon} \quad \text{dla każdego } x \in [a, b]. \tag{5.68}$$

Oznaczmy koszt wyznaczenia L dla ustalonego s przez $\mathrm{cost}(L, s, \sqrt[3]{\varepsilon}, \delta)$ z odpowiednim indeksem 'rand' lub 'quant' w zależności od modelu. Koszt wyznaczenia L, zgodnie z twierdzeniem 10, można oszacować przez (przyjmujemy w tym twierdzeniu $r := r - 1$, $\rho = 0$, $\gamma := \gamma/3$ i $\varepsilon := \sqrt[3]{\varepsilon}$)

$$\mathrm{cost}^{\mathrm{rand}}(L, s, \sqrt[3]{\varepsilon}, \delta) \leqslant K_3 M^{1/(2(r-1/2))} \varepsilon^{-1/(3(r-1/2-\gamma/3))} \left(\log M + \log \frac{1}{\delta}\right), \tag{5.69}$$

$$\mathrm{cost}^{\mathrm{quant}}(L, s, \sqrt[3]{\varepsilon}, \delta) \leqslant K_4 M^{1/r} \varepsilon^{-1/(3(r-\gamma/3))} \left(\log M + \log \frac{1}{\delta}\right), \tag{5.70}$$

gdzie K_3, K_4 są stałymi zależnymi od a, b, γ i parametrów klasy.

Algorytmy randomizacyjny i kwantowy dla problemu (5.53) są zdefiniowane następująco. Niech

$$\hat{F}(s) = p(s, l(b, s)) \quad \text{i} \quad A(s) = \frac{\partial p}{\partial s}(s, l(b, s)) + \frac{\partial p}{\partial w}(s, l(b, s)) L(b, s).$$

Niech $R_0 > 0$ będzie parametrem algorytmu. Dla $s^0 \in K(s^*, R)$ zdefiniujmy przybliżony ciąg metody Newtona następująco.
Dla $l = 0, 1, \ldots$, jeśli $A(s^l)$ jest nieosobliwa i $\|s^l - A(s^l)^{-1} \hat{F}(s^l) - s^0\| \leqslant R_0$, to

$$s^{l+1} = s^l - A(s^l)^{-1} \hat{F}(s^l), \tag{5.71}$$

w przeciwnym przypadku $s^{l+1} = s^l$. Warunek na ograniczoność przez R_0 będzie potrzebny do pokazania deterministycznej ograniczoności błędu algorytmu przez stałą i będzie użyty w dowodzie twierdzenia o ograniczeniach z góry na złożoność.

Dla $f \in \tilde{F}_d^r$, wynik algorytmu ϕ^{rand} lub ϕ^{quant} jest definiowany jako

$$\phi^{\mathrm{rand/quant}}(f)(x) = l(x, s^k), \tag{5.72}$$

gdzie k jest parametrem zdefiniowanym wcześniej.

Różnica pomiędzy algorytmami w przypadku randomizacyjnym i kwantowym polega na sposobie obliczania przybliżeń $l(b, s^l)$ i $L(b, s^l)$ rozwiązań problemów początkowych (5.55) oraz (5.66), potrzebnych do obliczenia $\hat{F}(s^l)$ i $A(s^l)$.

W następnym podrozdziale przeanalizujemy działanie tych algorytmów i wyznaczymy ograniczenia na ich błąd oraz koszt. Posłuży to do wyznaczenia górnych ograniczeń na złożoność problemu (5.53).

5.3.3. Górne ograniczenia

Górne ograniczenia na złożoność problemu (5.53) w modelu randomizacyjnym i kwantowym przedstawione są w następującym twierdzeniu.

Twierdzenie 27. *Dla dowolnego $\gamma \in (0,1)$ i dla dowolnej funkcji p spełniającej (Zał. 1) istnieją dodatnie liczby $C_1(\gamma)$, $C_2(\gamma)$ i $\varepsilon_0(\gamma)$ (zależne wyłącznie od γ, a, b, parametrów klasy oraz funkcji p) takie, że dla dowolnego $\varepsilon < \varepsilon_0(\gamma)$, ε–złożoność problemu (5.53) w klasie $\tilde{F}_d^r$ w modelu randomizacyjnym i kwantowym spełnia*

$$\mathrm{comp}^{\mathrm{rand}}(\varepsilon, S, \tilde{F}_d^r) \leqslant C_1(\gamma) \left(\frac{1}{\varepsilon}\right)^{1/(r+1/2-\gamma)}, \tag{5.73}$$

$$\mathrm{comp}^{\mathrm{quant}}(\varepsilon, S, \tilde{F}_d^r) \leqslant C_2(\gamma) \left(\frac{1}{\varepsilon}\right)^{1/(r+1-\gamma)}. \tag{5.74}$$

W celu udowodnienia twierdzenia 27 pokażemy teraz pewne własności ciągu $\{s^l\}$. Wykażemy najpierw, że z wysokim prawdopodobieństwem przybliżony ciąg Newtona $\{s^l\}$ wybierany jest w każdym kroku ze wzoru (5.71). Ponadto pokażemy, że zbliża się on odpowiednio szybko do rozwiązania s^*. Poniższy lemat jest modyfikacją lematu 3.2 z [Kac02], gdzie rozpatrywany był przypadek deterministyczny. Poniżej uwzględniamy prawdopodobieństwo zajścia odpowiednich zdarzeń.

Lemat 6. *Niech p spełnia założenie (Zał. 1), $r \geqslant 2$ i $R_0 \geqslant 2R$. Istnieją dodatnie stałe $\tilde{N}_1$, $\tilde{N}_2$ zależne wyłącznie od D, H, p, a oraz b takie, że dla dostatecznie małego ε i dla dowolnej $f \in \tilde{F}_d^r$, jeśli $s^0 \in K(s^*, R)$, to dla dowolnego $l \in \mathbb{N}$, z prawdopodobieństwem co najmniej $(1-\delta)^{2l}$, ciąg $\{s^j\}_{j=0}^l$ jest wybierany ze wzoru (5.71), $\{s^j\}_{j=0}^l \subset K(s^*, R)$ oraz*

$$\|s^i - s^*\| \leqslant C\|s^{i-1} - s^*\|^2(1 + \tilde{N}_1 \sqrt[3]{\varepsilon}) + \|s^{i-1} - s^*\|\tilde{N}_1 \sqrt[3]{\varepsilon} + \tilde{N}_2\varepsilon, \ i = 1, 2, \ldots, l. \tag{5.75}$$

Powyżej $C = 1/(4R)$ jest stałą zdefiniowaną w (5.61), $\delta = \varepsilon^2/(2k+1)$ w modelu randomizacyjnym, $\delta = 1/(8k+4)$ w modelu kwantowym i $k = \lceil 2\log\log(1/\varepsilon)\rceil$ są parametrami zdefiniowanymi wcześniej.

Dowód, który jest pewną modyfikacją dowodu lematu 3.2 z [Kac02], zamieszczony jest w dodatku B.

Sformułujemy teraz lemat, który ustanawia ograniczenia na błąd zaburzonej metody Newtona. Jest on modyfikacją lematu 3.3 z [Kac02], w którym rozważany był przypadek deterministyczny.

Lemat 7. *Niech p spełnia założenie (Zał. 1), $r \geqslant 2$ i $R_0 \geqslant 2R$. Wtedy istnieje stała N_3 zależna wyłącznie od D, H, Z, p, a oraz b taka, że dostatecznie małego ε i dla dowolnej $f \in \tilde{F}_d^r$ zachodzi: jeśli $s^0 \in K(s^*, R)$ i $l \geqslant k = \lceil 2 \log\log(1/\varepsilon) \rceil$, to z prawdopodobieństwem co najmniej $(1-\delta)^{2l}$ mamy*

$$\|s^l - s^*\| \leqslant N_3 \varepsilon.$$

Dowód, który jest pewną modyfikacją dowodu lematu 3.3 z [Kac02], można znaleźć w dodatku B.

Przedstawimy teraz dowód twierdzenia 27.

Dowód twierdzenia 27. Weźmy $R_0 \geqslant 2R$. Najpierw wyznaczymy ograniczenia na błąd algorytmów. Błędy lokalne algorytmów ϕ^{rand} i ϕ^{quant} są ograniczone przez

$$\begin{aligned} \|z - \phi^{\text{rand/quant}}(f)\| &= \sup_{x \in [a,b]} \|y(x, s^*) - l(x, s^k)\| \\ &\leqslant \sup_{x \in [a,b]} \|y(x, s^*) - y(x, s^k)\| + \sup_{x \in [a,b]} \|y(x, s^k) - l(x, s^k)\|. \end{aligned} \tag{5.76}$$

Pierwszy składnik prawej strony, zgodnie z klasycznymi wynikami dotyczącymi zależności rozwiązania równań różniczkowych od warunku początkowego, z wykorzystaniem lematu 7, jest ograniczony przez

$$\sup_{x \in [a,b]} \|y(x, s^*) - y(x, s^k)\| \leqslant e^{D(b-a)} \|s^* - s^k\| \leqslant e^{D(b-a)} N_3 \varepsilon. \tag{5.77}$$

Zachodzi to z prawdopodobieństwem nie mniejszym niż $(1-\delta)^{2k}$. Drugi składnik prawej strony (5.76), zgodnie z (5.62) jest ograniczony przez ε z prawdopodobieństwem co najmniej $1 - \delta$. Stąd

$$\|z - \phi^{\text{rand/quant}}(f)\| \leqslant C_0 \varepsilon \tag{5.78}$$

z prawdopodobieństwem co najmniej $(1-\delta)^{2k+1} \geqslant 1-(2k+1)\delta$, gdzie $C_0 = e^{D(b-a)} N_3 + 1$. Przypomnijmy, że w modelu kwantowym $\delta = 1/(8k+4)$. Tak więc, w tym modelu (5.78) zachodzi z prawdopodobieństwem co najmniej $3/4$. Dostajemy więc, że

$$\mathrm{e}^{\text{quant}}(S, \phi^{\text{quant}}, \tilde{F}_d^r) \leqslant C_0 \varepsilon. \tag{5.79}$$

Znajdziemy teraz ograniczenie na błąd w modelu randomizacyjnym zdefiniowany przez (5.56). Z definicji algorytmu wynika, że $\|s^* - s^k\| \leqslant \|s^* - s^0\| + \|s^0 - s^k\| \leqslant R + R_0$. Zachodzi to w sposób deterministyczny. Z pierwszej nierówności w (5.77) dostajemy, że $\sup_{x\in[a,b]} \|y(x,s^*) - y(x,s^k)\| \leqslant (R+R_0)e^{D(b-a)}$. Ponadto z twierdzenia 7 mamy, że $\sup_{x\in[a,b]} \|y(x,s^k) - l(x,s^k)\|$ jest ograniczone deterministycznie przez stałą. Z powyższych i (5.76) dostajemy, że błąd $\|z - \phi^{\mathrm{rand}}(f)\|$ jest ograniczony w sposób deterministyczny przez pewną stałą $\hat{K}$. Z definicji błędu (5.56) oraz z (5.78) mamy

$$\begin{aligned}
&\left(\mathrm{e}^{\mathrm{rand}}(S,\phi^{\mathrm{rand}},\tilde{F}_d^r)\right)^2 = \sup_{f\in\tilde{F}_d^r} \int\limits_{\Omega} \|z - \phi^{\mathrm{rand}}(f)\|^2 d\mathbf{P}(\omega) \\
&= \sup_{f\in\tilde{F}_d^r} \left(\int\limits_{\|z-\phi^{\mathrm{rand}}(f)\|\leqslant C_0\varepsilon} \|z - \phi^{\mathrm{rand}}(f)\|^2 d\mathbf{P}(\omega) + \int\limits_{\|z-\phi^{\mathrm{rand}}(f)\|>C_0\varepsilon} \|z - \phi^{\mathrm{rand}}(f)\|^2 d\mathbf{P}(\omega) \right) \\
&\leqslant C_0^2\varepsilon^2 + \sup_{f\in\tilde{F}_d^r} \int\limits_{\|z-\phi^{\mathrm{rand}}(f)\|>C_0\varepsilon} \hat{K}^2 d\mathbf{P}(\omega) \leqslant C_0^2\varepsilon^2 + \hat{K}^2(2k+1)\delta.
\end{aligned}$$

Przypomnijmy, że $\delta = \varepsilon^2/(2k+1)$. Dostajemy więc,

$$\mathrm{e}^{\mathrm{rand}}(S,\phi^{\mathrm{rand}},\tilde{F}_d^r) \leqslant \sqrt{C_0^2 + \hat{K}^2}\ \varepsilon. \tag{5.80}$$

Wyznaczymy teraz ograniczenia na koszt informacyjny. Koszt obliczenia przybliżeń $l = l(x,s)$ i $L = L(x,s)$ problemów początkowych (5.55), (5.66) w modelu randomizacyjnym i kwantowym z dokładnością ε_1 i prawdopodobieństwem porażki δ_1 dla ustalonego s oznaczyliśmy przez odpowiednio $\mathrm{cost}^{\mathrm{rand/quant}}(l,s,\varepsilon_1,\delta_1)$ i $\mathrm{cost}^{\mathrm{rand/quant}}(L,s,\varepsilon_1,\delta_1)$. Koszt całkowity wznosi

$$\begin{aligned}
\mathrm{cost}^{\mathrm{rand/quant}}(S,\phi^{\mathrm{rand/quant}},\tilde{F}_d^r) = \sum_{l=0}^{k-1} \Big(&\mathrm{cost}^{\mathrm{rand/quant}}(l,s^l,\varepsilon,\delta) \\
&+ \mathrm{cost}^{\mathrm{rand/quant}}(L,s^l,\sqrt[3]{\varepsilon},\delta)\Big) + \mathrm{cost}^{\mathrm{rand/quant}}(l,s^k,\varepsilon,\delta).
\end{aligned} \tag{5.81}$$

Przypomnijmy, że koszty znalezienia przybliżeń $l(x,s^l)$ i $L(x,s^l)$ dla dowolnego $f \in \tilde{F}_d^r$ są ograniczone przez (zob. (5.63), (5.64), (5.69), (5.70))

$$\mathrm{cost}^{\mathrm{rand}}(l,s^l,\varepsilon,\delta) \leqslant K_1 \left(\frac{1}{\varepsilon}\right)^{1/(r+1/2-\gamma/2)} \log\frac{1}{\delta}, \tag{5.82}$$

$$\mathrm{cost}^{\mathrm{quant}}(l,s^l,\varepsilon,\delta) \leqslant K_2 \left(\frac{1}{\varepsilon}\right)^{1/(r+1-\gamma/2)} \log\frac{1}{\delta}, \tag{5.83}$$

$$\mathrm{cost}^{\mathrm{rand}}(L,s^l,\sqrt[3]{\varepsilon},\delta) \leqslant K_3 M^{1/(2(r-1/2))} \left(\frac{1}{\varepsilon}\right)^{1/(3(r-1/2-\gamma/3))} \left(\log M + \log\frac{1}{\delta}\right),$$

$$\mathrm{cost}^{\mathrm{quant}}(L,s^l,\sqrt[3]{\varepsilon},\delta) \leqslant K_4 M^{1/r} \left(\frac{1}{\varepsilon}\right)^{1/(3(r-\gamma/3))} \left(\log M + \log\frac{1}{\delta}\right),$$

gdzie K_1, K_2, K_3, K_4 zależą wyłącznie od a, b, γ i parametrów klasy. Przypomnijmy, że $M = \Theta\left(\varepsilon^{-1/(r+1/2-\gamma/4)}\right)$ w modelu randomizacyjnym i $M = \Theta\left(\varepsilon^{-1/(r+1-\gamma/4)}\right)$ w modelu kwantowym. Stąd

$$\begin{aligned}
&\mathrm{cost}^{\mathrm{rand}}(L, s^l, \sqrt[3]{\varepsilon}, \delta) \leqslant \tilde{K}_3 \left(\frac{1}{\varepsilon}\right)^{\alpha_1} \left(\log\frac{1}{\varepsilon} + \log\frac{1}{\delta}\right), \\
&\mathrm{cost}^{\mathrm{quant}}(L, s^l, \sqrt[3]{\varepsilon}, \delta) \leqslant \tilde{K}_4 \left(\frac{1}{\varepsilon}\right)^{\alpha_2} \left(\log\frac{1}{\varepsilon} + \log\frac{1}{\delta}\right),
\end{aligned} \tag{5.84}$$

gdzie $\alpha_1 = \dfrac{1}{(2r-1)(r+1/2-\gamma/4)} + \dfrac{1}{3(r-1/2-\gamma/3)}$ i $\alpha_2 = \dfrac{1}{r(r+1-\gamma/4)} + \dfrac{1}{3(r-\gamma/3)}$, a $\tilde{K}_3$ i $\tilde{K}_4$ są stałymi zależnymi wyłącznie od a, b, γ i parametrów klasy. Przypomnijmy, że w modelu randomizacyjnym $\delta = \varepsilon^2/(2k+1)$, a $k = \lceil 2\log\log(1/\varepsilon)\rceil$. Z powyższych ograniczeń i z (5.81), koszt całkowity w modelu randomizacyjnym spełnia

$$\begin{aligned}
\mathrm{cost}^{\mathrm{rand}}(S, \phi^{\mathrm{rand}}, \tilde{F}_d^r) = O\Bigg(\log\log\frac{1}{\varepsilon}\Bigg(\left(\frac{1}{\varepsilon}\right)^{1/(r+1/2-\gamma/2)} \log\frac{2k+1}{\varepsilon^2} \\
+ \left(\frac{1}{\varepsilon}\right)^{\alpha_1}\left(\log\frac{1}{\varepsilon} + \log\frac{2k+1}{\varepsilon^2}\right)\Bigg)\Bigg).
\end{aligned}$$

Jako, że $r \geqslant 2$ i $\gamma \in (0,1)$, więc $(2r-1)(r+1/2-\gamma/4) \geqslant 3(r+1/2-\gamma/2)$ oraz $3(r-1/2-\gamma/3) \geqslant 3/2(r+1/2-\gamma/2)$. Stąd

$$\alpha_1 \leqslant \frac{1}{r+1/2-\gamma/2}.$$

Ponadto

$$\log\frac{2k+1}{\varepsilon^2} = \log(2k+1) + \log\frac{1}{\varepsilon^2} = O\left(\log\log\log\frac{1}{\varepsilon} + \log\frac{1}{\varepsilon}\right) = O\left(\log\frac{1}{\varepsilon}\right).$$

Stąd, koszt całkowity w modelu randomizacyjnym spełnia

$$\mathrm{cost}^{\mathrm{rand}}(S, \phi^{\mathrm{rand}}, \tilde{F}_d^r) = O\left(\left(\frac{1}{\varepsilon}\right)^{1/(r+1/2-\gamma/2)} \log\frac{1}{\varepsilon}\log\log\frac{1}{\varepsilon}\right) = O\left(\left(\frac{1}{\varepsilon}\right)^{1/(r+1/2-\gamma)}\right) \tag{5.85}$$

dla dostatecznie małego ε.

Przejdźmy do oszacowań kosztu w modelu kwantowym. Przypomnijmy, że $\delta = 1/(8k+4)$, a $k = \lceil 2\log\log(1/\varepsilon)\rceil$. Zgodnie z (5.81), (5.83) i (5.84), koszt całkowity w modelu kwantowym spełnia

$$\begin{aligned}
\mathrm{cost}^{\mathrm{quant}}(S, \phi^{\mathrm{quant}}, \tilde{F}_d^r) = O\Bigg(\log\log\frac{1}{\varepsilon}\Bigg(\left(\frac{1}{\varepsilon}\right)^{1/(r+1-\gamma/2)} \log(8k+4) \\
+ \left(\frac{1}{\varepsilon}\right)^{\alpha_2}\left(\log\frac{1}{\varepsilon} + \log(8k+4)\right)\Bigg)\Bigg).
\end{aligned}$$

Zauważmy, że dla $r \geqslant 2$ i $\gamma \in (0,1)$ mamy $r(r+1-\gamma/4) \geqslant 2(r+1-\gamma/2)$ oraz $3(r-\gamma/3) \geqslant 2(r+1-\gamma/2)$. Stąd

$$\alpha_2 \leqslant \frac{1}{r+1-\gamma/2}.$$

Ponadto $\log(8k+4) = O\left(\log\log\log\frac{1}{\varepsilon}\right)$, więc

$$\mathrm{cost}^{\mathrm{quant}}(S, \phi^{\mathrm{quant}}, \tilde{F}_d^r) = O\left(\left(\frac{1}{\varepsilon}\right)^{1/(r+1-\gamma)}\right) \tag{5.86}$$

dla dostatecznie małego ε.

Ponieważ w obu modelach błąd jest ograniczony przez wielkość rzędu ε (zob. (5.80) i (5.79)), więc zgodnie z definicją złożoności, ograniczenia na koszt (5.85) i (5.86) stanowią (z dokładnością do stałej) górne ograniczenia na złożoność problemu (5.53). □

W następnym podrozdziale ustanowimy dolne ograniczenia na złożoność tego problemu.

5.3.4. Dolne ograniczenia dla problemów (5.53) i (5.57)

Przedstawimy teraz dolne ograniczenia na złożoność dla ogólnego problemu brzegowego (5.53). Rozważmy funkcję $p(s,w) = s - \alpha$. Dla tak zdefiniowanej funkcji p, problem (5.53) staje się problemem początkowym z warunkiem początkowym $z(a) = \alpha$. Tak zdefiniowana funkcja p spełnia założenie (Zał. 1), a $F(s) = s - \alpha$. Funkcja F spełnia (Zał. 2), więc dla takiej p, klasa $\tilde{F}_d^r$ jest równa klasie F_d^r. Użyjemy dolnych ograniczeń na złożoność dla problemów początkowych z pracy [Kac04]. W pracy tej rozważana jest klasa Höldera $F_d^{r,\rho}$, jednak ograniczenia na złożoność z tej pracy są również prawdziwe w klasie F_d^r (kładąc formalnie $\rho = 0$). Z twierdzenia 8 otrzymujemy dolne ograniczenia na na złożoność problemu brzegowego (5.53).

Twierdzenie 28. *Istnieją dodatnie stałe c^{rand} i c^{quant} oraz funkcja p spełniająca (Zał. 1) takie, że ε–złożoność problemu (5.53) w klasie $\tilde{F}_d^r$ spełnia*

$$\mathrm{comp}^{\mathrm{rand}}(\varepsilon, S, \tilde{F}_d^r) \geqslant c^{\mathrm{rand}}\left(\frac{1}{\varepsilon}\right)^{1/(r+1/2)}, \tag{5.87}$$

$$\mathrm{comp}^{\mathrm{quant}}(\varepsilon, S, \tilde{F}_d^r) \geqslant c^{\mathrm{quant}}\left(\frac{1}{\varepsilon}\right)^{1/(r+1)}. \tag{5.88}$$

Zauważmy, że te ograniczenia są prawie zgodne z górnymi ograniczeniami z twierdzenia 27. Ograniczenia z góry różnią się od ograniczeń z dołu o dowolnie mały parametr

γ w mianowniku wykładnika. Wynika z tego, że algorytmy ϕ^{rand} i ϕ^{quant} zdefiniowane w podrozdziale 5.3.3 są prawie optymalne.

Przejdźmy teraz do problemów skalarnych drugiego rzędu (5.57). W pracy [Kac02] pokazane jest, że złożoność tego problemu w modelu deterministycznym różni się od złożoności ogólnego nieliniowego problemu brzegowego (5.53), jeśli dopuszczone jest użycie dowolnej informacji liniowej. Pokażemy, że w modelach randomizacyjnym i kwantowym z informacją standardową szczególny problem (5.57) nie jest łatwiejszy od (5.53). Algorytmy rozwiązujące ogólne problemy brzegowe pierwszego rzędu, przedstawione w podrozdziale 5.3.2, zastosowane do (5.57), po przejściu na układ równań, pozostają prawie optymalne (przejście z równania wyższego rzędu na układ równań zachowujące ograniczoność funkcji prawej strony dla problemów początkowych pokazane jest np. w [Dau11]). Dolne ograniczenia na złożoność tego problemu przedstawione zostaną poniżej.

Rozważmy problem (5.57) z funkcją prawej strony $f : [a, b] \times \mathbb{R} \to \mathbb{R}$ należącą do klasy funkcji $\mathcal{G}$ zdefiniowanej przez (5.58). Operator rozwiązania dla tego problemu oznaczony był przez S^*. Następujące twierdzenie przedstawia ograniczenia z dołu na złożoność tego problemu.

Twierdzenie 29. *Istnieją dodatnie stałe c^{rand}, c^{quant} i ε_0 takie, że dla $\varepsilon \in (0, \varepsilon_0)$ złożoność problemu (5.57) w klasie $\mathcal{G}$ spełnia*

$$\text{comp}^{\text{rand}}(\varepsilon, S^*, \mathcal{G}) \geqslant c^{\text{rand}} \left(\frac{1}{\varepsilon}\right)^{1/(r+1/2)}, \tag{5.89}$$

$$\text{comp}^{\text{quant}}(\varepsilon, S^*, \mathcal{G}) \geqslant c^{\text{quant}} \left(\frac{1}{\varepsilon}\right)^{1/(r+1)}. \tag{5.90}$$

Dowód. Rozważmy problem postaci

$$\begin{cases} u''(x) = g(x), & x \in [a, b], \\ u(a) = A, \\ u(b) = B, \end{cases} \tag{5.91}$$

gdzie g jest funkcją klasy F_1^r funkcji o ciągłych, ograniczonych pochodnych do rzędu r taką, że nośnik g zawarty jest w $[a, (a + b)/2]$. Problem ten jest podklasą klasy problemów (5.57) z $f(x, y) = g(x)$. Pokażemy dolne ograniczenia na złożoność obliczania wartości $u(\frac{a+b}{2})$. Te ograniczenia będą również stanowić ograniczenia na złożoność rozwiązywania (5.91), a co za tym idzie na złożoność problemu (5.57). Zauważmy, że

$$u(x) = u(a) + u'(a)(x - a) + \int_a^x \int_a^{t_1} g(t) dt dt_1. \tag{5.92}$$

Obliczając to wyrażenie w punkcie b, korzystając z warunków brzegowych dostajemy

$$u'(a) = \frac{1}{b-a}\left(B - A - \int_a^b \int_a^{t_1} g(t)dtdt_1\right). \tag{5.93}$$

Obliczając $u(\frac{a+b}{2})$ w (5.92) i korzystając z (5.93) otrzymujemy

$$u\left(\frac{a+b}{2}\right) = \frac{1}{2}\left(A + B + \int_a^{\frac{a+b}{2}} \int_a^{t_1} g(t)dtdt_1 + \int_b^{\frac{a+b}{2}} \int_a^{t_1} g(t)dtdt_1\right). \tag{5.94}$$

Obliczymy teraz ostatnie dwie całki. Używając wzoru na całkowanie przez części otrzymujemy

$$\int_a^{\frac{a+b}{2}} \int_a^{t_1} g(t)dtdt_1 = \int_a^{\frac{a+b}{2}} \left(\frac{a+b}{2} - t\right) g(t)dt \tag{5.95}$$

i

$$\int_b^{\frac{a+b}{2}} \int_a^{t_1} g(t)dtdt_1 = \frac{a+b}{2} \int_a^{\frac{a+b}{2}} g(t)dt - b\int_a^b g(t)dt - \int_b^{\frac{a+b}{2}} tg(t)dt. \tag{5.96}$$

Jako, że nośnik funkcji g zawarty jest w przedziale $[a, (a+b)/2]$, więc

$$u\left(\frac{a+b}{2}\right) = \frac{1}{2}\left(A + B + \int_a^{\frac{a+b}{2}} (a-t)g(t)dt\right).$$

Problem obliczania $u(\frac{a+b}{2})$ jest więc równoważny problemowi obliczania całki z wagą $I = \int_a^{\frac{a+b}{2}} (t-a)g(t)dt$. Zauważmy, że na przedziale $[(3a+b)/4, (a+b)/2]$ funkcja wagowa $w(t) = t - a$ jest ograniczona z dołu przez $(b-a)/4$.

Podobnie jak w podrozdziale 5.2.7 pokażemy jak zredukować problem sumowania do problemu obliczania całki z wagą. Niech $x_1, x_2, \ldots, x_n$ będzie ciągiem liczb rzeczywistych z przedziału $[0, 1]$. Chcemy obliczyć $\frac{1}{n}\sum_{i=1}^n x_i$. Niech $\varepsilon_1 > 0$ będzie pewnym dostatecznie małym parametrem, który sprecyzujemy później. Klasa F_1^r zawiera $n = \Theta\left(\varepsilon_1^{-1/r}\right)$ funkcji f_i, $i = 1, 2, \ldots, n$, o rozłącznych nośnikach zawartych w przedziale $[(3a+b)/4, (a+b)/2]$ takich, że $\int_{(3a+b)/4}^{(a+b)/2} f_i(x)dx = \varepsilon_1^{1+1/r}$ (zobacz lemat 9 w dodatku A). Niech $H_i = \int_{(3a+b)/4}^{(a+b)/2} (t-a)f_i(t)dt \geqslant (b-a)/4\varepsilon_1^{1+1/r}$. Zdefiniujmy funkcje

$h_i(x) = (b-a)/4\varepsilon_1^{1+1/r} f_i(x)/H_i$. Jako, że $(b-a)/4\varepsilon_1^{1+1/r}/H_i \leqslant 1$, $i = 1, 2, \ldots, n$, i funkcje f_i należą do klasy F_1^r, więc funkcje h_i również należą do klasy F_1^r.
Funkcja dana jako $f_{\varepsilon_1}(x) = \sum_{i=1}^{n} x_i h_i(x)$ należy do klasy F_1^r. Załóżmy, że algorytm ϕ oblicza (w danym modelu obliczeniowym) $\int_a^{(a+b)/2} (t-a)g(t)dt$ z błędem nie większym niż ε i kosztem N dla dowolnej funkcji $g \in F_1^r$. Dla $g = f_{\varepsilon_1}$ mamy, że $\int_a^{(a+b)/2} (t-a)f_{\varepsilon_1}(t)dt = (b-a)/4\varepsilon_1^{1+1/r} \sum_{i=1}^{n} x_i$. Tak więc, algorytm ϕ oblicza również średnią $\frac{1}{n}\sum_{i=1}^{n} x_i$ z błędem nie większym niż $\frac{4\varepsilon}{(b-a)n\varepsilon_1^{1+1/r}}$.
Skorzystamy teraz z dolnych ograniczeń na złożoność problemu obliczania średniej z n liczb z przedziału $[0,1]$ (zob. podrozdział 2.2).

Rozważmy najpierw modelu randomizacyjny. Dla dostatecznie małego ε przyjmijmy $\varepsilon_1 = \varepsilon^{2r/(2r+1)}$. Korzystając z (2.7) dostajemy

$$N = \Omega\left(\min\left\{n, \left((b-a)/4n\varepsilon_1^{(r+1)/r}\varepsilon^{-1}\right)^2\right\}\right) = \Omega\left(\varepsilon^{-\frac{1}{r+1/2}}\right).$$

Stanowi to dolne ograniczenie na złożoność problemu (5.57) w modelu randomizacyjnym.

Przejdźmy do modelu kwantowego. Dla dostatecznie małego ε przyjmijmy $\varepsilon_1 = \varepsilon^{r/(r+1)}$. Z (2.8), z wykorzystaniem lematu 4, dostajemy

$$N = \Omega\left(\min\left\{n, (b-a)/4n\varepsilon_1^{(r+1)/r}\varepsilon^{-1}\right\}\right) = \Omega\left(\varepsilon^{-1/(r+1)}\right)$$

co stanowi dolne ograniczenie na złożoność problemu (5.57) w modelu kwantowym. To kończy dowód twierdzenia. □

5.3.5. Podsumowanie

Zbadaliśmy w tym rozdziale złożoność nieliniowego problemu brzegowego postaci (5.53) oraz skalarnego problemu rzędu pierwszego postaci (5.57). Rozważyliśmy modele randomizacyjny i kwantowy. Dla każdego z modeli skonstruowaliśmy algorytm rozwiązujący problem (5.53) (rozwiązują one również problem (5.57), po przejściu na układ równań pierwszego rzędu) i oszacowaliśmy jego błąd. Wyznaczyliśmy również dolne ograniczenia na złożoność tych problemów w obu modelach. Dolne i górne ograniczenia na złożoność problemu (5.53) i problemu (5.57) przedstawia tabela 5.2. Dla zwiększenia przejrzystości pominęliśmy stałe.

problem	układ równań (5.53)		p. skalarny (5.57)	
model	ogr. z dołu	ogr. z góry	ogr. z dołu	ogr. z góry
det-st	$\varepsilon^{-\frac{1}{r}}$	$\varepsilon^{-\frac{1}{r}} \log\log \varepsilon^{-1}$	$\varepsilon^{-\frac{1}{r}}$	$\varepsilon^{-\frac{1}{r}} \log\log \varepsilon^{-1}$
det-lin	$\varepsilon^{-\frac{1}{r+1}}$	$\varepsilon^{-\frac{1}{r+1}} \log\log \varepsilon^{-1}$	$\varepsilon^{-\frac{1}{r+2}}$	$\varepsilon^{-\frac{1}{r+2}} \log\log \varepsilon^{-1}$
rand	$\varepsilon^{-\frac{1}{r+1/2}}$	$\varepsilon^{-\frac{1}{r+1/2-\gamma}}$	$\varepsilon^{-\frac{1}{r+1/2}}$	$\varepsilon^{-\frac{1}{r+1/2-\gamma}}$
quant	$\varepsilon^{-\frac{1}{r+1}}$	$\varepsilon^{-\frac{1}{r+1-\gamma}}$	$\varepsilon^{-\frac{1}{r+1}}$	$\varepsilon^{-\frac{1}{r+1-\gamma}}$

Tabela 5.2. Ograniczenia na złożoność

Z otrzymanych rezultatów wynika, że dla obu problemów obliczenia w modelu randomizacyjnym są szybsze od obliczeń w modelu deterministycznym z informacją standardową o 1/2 w wykładniku, natomiast w modelu kwantowym są szybsze o 1.

Złożoność ogólnego problemu (5.53) różni się od złożoności problemu skalarnego (5.57) tylko w modelu deterministycznym z informacją liniową. W modelach randomizacyjnym i kwantowym nie ma to miejsca.

Otrzymane ograniczenia w modelu randomizacyjnym i kwantowym są prawie ostre. Górne i dolne ograniczenia różnią się o dowolnie mały parametr γ w wykładniku. Tak więc skonstruowane przez nas algorytmy są prawie optymalne. W modelu randomizacyjnym do rozwiązywania problemów początkowych, zamiast używać prawie optymalnych algorytmów z [Kac06] można by użyć optymalnych algorytmów z [HM08]. Wydaje się, że wtedy ograniczenia z góry w modelu randomizacyjnym mogły by się poprawić przez pozbycie się parametru γ w wykładniku kosztem pomnożenia ograniczeń przez czynnik logarytmiczny. Użycie tych algorytmów sprawia jednak pewne problemy dowodowe, gdyż wynik z [HM08] nie daje się w prosty sposób zastosować do problemów z kawałkami regularną funkcją prawej strony (w wynikach nie ma zależności od długości przedziału). Pewną niedoskonałością prezentowanych w tym rozdziale algorytmów jest zależność od stałych c_0 i R_0, które są parametrami klasy funkcji prawej strony.

Rozdział 6

Podsumowanie

W rozprawie zbadaliśmy złożoność kilku nieliniowych problemów w modelach randomizacyjnym i kwantowym. Rozważaliśmy takie problemy jak: maksymalizacja funkcji, poszukiwanie zer funkcji, rozwiązywanie problemów brzegowych dla równań różniczkowych (liniowych i nieliniowych). Celem było skonstruowanie optymalnych (lub prawie optymalnych) algorytmów rozwiązujących te problemy oraz wyznaczenie górnych i dolnych ograniczeń na złożoność.

Zbadaliśmy złożoność problemu maksymalizacji funkcji z klasy Höldera $F_d^{r,\rho}$ w modelu kwantowym (dla modeli deterministycznego i randomizacyjnego wyniki są znane). Skonstruowaliśmy optymalny algorytm rozwiązujący ten problem. Opiera się on na lokalnym przybliżaniu funkcji wielomianami Taylora oraz korzysta z optymalnego algorytmu kwantowego maksymalizującego dyskretne ciągi. Na podstawie tego algorytmu wyznaczyliśmy górne ograniczenie na złożoność tego problemu rzędu $\varepsilon^{-d/(2(r+\rho))}$. Wyznaczyliśmy również dolne ograniczenie na złożoność tego problemu, które okazało się być zgodne co do rzędu z ograniczeniem z góry. Porównując otrzymane ograniczenia na złożoność otrzymaliśmy, że model kwantowy daje kwadratowe przyspieszenie w stosunku do modeli deterministycznego i randomizacyjnego.

Wyznaczyliśmy również ograniczenia na złożoność problemu poszukiwania zer funkcji z klasy Höldera $F_d^{r,\rho}$. Rozważaliśmy modele randomizacyjny i kwantowy oraz

dla porównania model deterministyczny. W każdym z modeli przedstawiliśmy algorytm rozwiązujący ten problem. Algorytmy skonstruowane są podobnie do algorytmu kwantowego poszukującego maksimum funkcji. Opierają się na lokalnym przybliżeniu funkcji wielomianami Taylora oraz wykorzystują optymalny algorytm przeszukiwania dyskretnego w odpowiednim modelu. Algorytmy te wyznaczają górne ograniczenia na złożoność rzędu $\varepsilon^{-d/(2(r+\rho))}$ w modelu kwantowym oraz $\varepsilon^{-d/(r+\rho)}$ w modelach deterministycznym i randomizacyjnym. Wyznaczyliśmy również dolne ograniczenia na złożoność, które są zgodne co do rzędu z ograniczeniami z góry. Wynika z tego, że skonstruowane przez nas algorytmy są optymalne. Z otrzymanych ograniczeń wynika, że dla tego problemu model randomizacyjny nie daje przewagi nad modelem deterministycznym, natomiast model kwantowy daje kwadratowe przyspieszenie.

Zbadaliśmy również złożoność pewnych problemów brzegowych dla równań różniczkowych, liniowego i nieliniowego. W obu problemach rozważaliśmy funkcje prawej strony należące do klasy funkcji r-krotnie różniczkowalnych o ograniczonych pochodnych. Problem liniowy badaliśmy w modelach randomizacyjnym i kwantowym oraz dla porównania również w modelu deterministycznym zarówno z informacją standardową jaki i dowolną liniową. Górne ograniczenia dla tego problemu uzyskaliśmy przez skonstruowanie odpowiednich algorytmów. W modelach deterministycznych lokalne problemy początkowe rozwiązywaliśmy wykorzystując rozwinięcie Taylora natomiast w modelach randomizacyjnym i kwantowym użyliśmy prawie optymalnych algorytmów rozwiązujących problemy początkowe z [Goć06]. Znaleźliśmy również dolne ograniczenia na złożoność tego problemu we wszystkich modelach. Uzyskaliśmy ograniczenia rzędu $\varepsilon^{-1/r}$ w modelu deterministycznym z informacją standardową. W modelu deterministycznym z informacją liniową dostaliśmy ograniczenia z góry rzędu $\varepsilon^{-1/(r+k-q)}$, gdzie k jest rzędem równania, a q najwyższym rzędem pochodnych rozwiązania po prawej stronie równania. Ograniczenia z dołu w tym modelu są rzędu $\varepsilon^{-1/(r+k)}$. Tak więc, ograniczenia z dołu i z góry są zgodne tylko wtedy, gdy prawa strona równania zależy wyłącznie od wartości rozwiązania. Uzyskanie zgodnych ograniczeń z góry i z dołu w modelu deterministycznym z informacją liniową pozostaje problemem otwartym. W modelu randomizacyjnym uzyskaliśmy ograniczenia z dołu rzędu $\varepsilon^{-1/(r+1/2)}$, a w modelu kwantowym rzędu $\varepsilon^{-1/(r+1)}$. Ograniczenia z góry w tych modelach różnią się od ograniczeń z dołu o dowolnie mały parametr w mianowniku wykładnika. Z uzyskanych wyników wynika, że obliczenia w modelu randomizacyjnym są dla tego problemu szybsze od obliczeń w modelu deterministycznym z informacją standardową o $1/2$ w mianowniku wykładnika, natomiast obliczenia w modelu kwantowym o 1.

Zbadaliśmy również nieliniowy problem brzegowy. Skonstruowaliśmy algorytmy rozwiązujące ten problem w modelach randomizacyjnym i kwantowym. Algorytmy te polegają na rozwiązywaniu stowarzyszonego z nim równania nieliniowego za pomocą metody Newtona oraz wykorzystują prawie optymalne algorytmy rozwiązujące problemy początkowe dla równań różniczkowych z [Kac06]. Za pomocą tych algorytmów pokazaliśmy górne ograniczenia na złożoność tego problemu w modelach randomizacyjnym i kwantowym (są one tego samego rzędu co ograniczenia dla liniowego problemu brzegowego). Ograniczenia dolne na złożoność są prawie zgodne z ograniczeniami z góry. Porównując otrzymane ograniczenia ze znanymi ograniczeniami w modelu deterministycznym uzyskaliśmy, że podobnie jak dla problemu liniowego, obliczenia randomizacyjne są szybsze od obliczeń deterministycznych o $1/2$ w wykładniku, a obliczenia kwantowe o 1.

Dodatek A

Fakty pomocnicze

Przedstawimy tutaj za [Nov88] pewne funkcje, które przydatne są przy uzyskiwaniu dolnych ograniczeń na złożoność. Rozważmy dwie klasy funkcji. Pierwsza z tych klas to klasa Höldera $F_d^{r,\rho}$ dana przez

$$F_d^{r,\rho}=\{f:[0,1]^d\to\mathbb{R}|\ f\in C^r([0,1]^d), |D^{(i)}f(x)|\leqslant D,\ i=0,1,\dots r,$$
$$|D^{(r)}f(x)-D^{(r)}f(y)|\leqslant H\|x-y\|^\rho, x,y\in[0,1]^d\},$$

gdzie $r\geqslant 0$, $0<\rho\leqslant 1$, D i H są dodatnimi stałymi, a $D^{(i)}$ przebiega przez zbiór wszystkich pochodnych cząstkowych rzędu i. Druga klasa to klasa funkcji r-krotnie różniczkowalnych o ograniczonych pochodnych cząstkowych dana przez

$$F_d^{r}=\{f:[0,1]^d\to\mathbb{R}|\ f\in C^r([0,1]^d), |D^{(i)}f(x)|\leqslant D,\ i=0,1,\dots r,\ x\in[0,1]^d\},$$

gdzie $r\geqslant 0$, D jest dodatnią stałą. Jako $\|\cdot\|$ rozumieć będziemy wszędzie $\|\cdot\|_\infty$, zarówno jako normę wektorów jak i funkcji. Klasa $F_d^{r,\rho}$ jest podklasą klasy F_d^r.

Niech $x=(x_1,x_2,\dots,x_d)\in\mathbb{R}^d$. Zdefiniujmy funkcję

$$G(x)=\begin{cases} a\prod_{i=1}^{d}(1-x_i^2)^{r+1} & \text{dla } \|x\|\leqslant 1\\ 0 & \text{poza tym}\end{cases}.$$

Można tak dobrać wartość parametru $a>0$, że $G\in F_d^{r,\rho}\subset F_d^r$.

Niech $\varepsilon_1\in(0,1)$. Niech $t_1,t_2,\dots,t_n\in[0,1]^d$ będzie równomierną siatką punktów

odległych od siebie wzdłuż każdej współrzędnej o $2\varepsilon_1^{1/(r+\rho)}$ oraz odległymi od brzegów kostki co najmniej o $\varepsilon_1^{1/(r+\rho)}$. Takich punktów można wybrać $n = \Theta(\varepsilon_1^{-d/(r+\rho)})$. Dla $i = 1, 2, \ldots n$ zdefiniujmy funkcje

$$f_i(x) = \varepsilon_1 G\left((x - t_i)\varepsilon_1^{-1/(r+\rho)}\right). \tag{A.1}$$

Dzięki założeniu o punktach t_i funkcje te mają parami rozłączne nośniki. Funkcje te są nieujemne. Zauważmy, że $D^{(j)} f_i(x) = \varepsilon_1^{1-j/(r+\rho)} D^{(j)} G\left((x - t_i)\varepsilon_1^{-1/(r+\rho)}\right)$, $i = 1, 2, \ldots, n$, $j = 0, 1, \ldots, r$. Wynika z tego, że $\|D^{(j)} f_i\| \leqslant \|D^{(j)} G\| \leqslant D$. Sprawdzimy jeszcze czy r-te pochodne cząstkowe spełniają warunek Höldera. Istotnie,

$$\begin{aligned}|D^{(r)} f_i(x) - D^{(r)} f_i(y)| &= \varepsilon_1^{1-r/(r+\rho)} |D^{(r)} G\left((x - t_i)\varepsilon_1^{-1/(r+\rho)}\right) - D^{(r)} G\left((y - t_i)\varepsilon_1^{-1/(r+\rho)}\right)| \\ &\leqslant H\varepsilon_1^{\rho/(r+\rho)} \left\|(x - t_i)\varepsilon_1^{-1/(r+\rho)} - (y - t_i)\varepsilon_1^{-1/(r+\rho)}\right\|^{\rho} \leqslant H\|x - y\|^{\rho}.\end{aligned}$$

Wynika z tego, że $f_i \in F_d^{r,\rho}$.

Niech $\tilde{t}_1, \tilde{t}_2, \ldots, \tilde{t}_{\tilde{n}} \in [0, 1]^d$ będzie równomierną siatką punktów odległych od siebie wzdłuż każdej współrzędnej o $2\varepsilon_1^{1/r}$ oraz odległymi od brzegów kostki co najmniej o $\varepsilon_1^{1/r}$. Takich punktów można wybrać $\tilde{n} = \Theta(\varepsilon_1^{-d/r})$. Dla $i = 1, 2, \ldots \tilde{n}$ zdefiniujmy funkcje

$$\tilde{f}_i(x) = \varepsilon_1 G\left((x - t_i)\varepsilon_1^{-1/r}\right). \tag{A.2}$$

Funkcje te mają parami rozłączne nośniki. Zauważmy, że $D^{(j)} \tilde{f}_i(x) = \varepsilon_1^{1-j/r} D^{(j)} G\left((x - t_i)\varepsilon_1^{-1/r}\right)$, $i = 1, 2, \ldots, \tilde{n}$, $j = 0, 1, \ldots, r$. Wynika z tego, że $\|D^{(j)} \tilde{f}_i\| \leqslant \|D^{(j)} G\| \leqslant D$. Stąd, $\tilde{f}_i \in F_d^{r,\rho}$.

Funkcje te posłużą do pokazania następujących własności klasy Höldera $F_d^{r,\rho}$ oraz klasy F_d^r.

Lemat 8. *Dla klasy Höldera $F_d^{r,\rho}$ istnieje dodatnia stała ε_0 taka, że dla $\varepsilon \in (0, \varepsilon_0)$, klasa $F_d^{r,\rho}$ zawiera $n = \Theta(\varepsilon^{-d/(r+\rho)})$ nieujemnych funkcji $f_1, f_2, \ldots, f_n$ o rozłącznych nośnikach zawartych w $[0, 1]^d$ takich, że $\max_{x \in [0,1]^d} f_i(x) = \varepsilon$ dla $i = 1, 2, \ldots, n$.*
Dla klasy F_d^r istnieje dodatnia stała $\tilde{\varepsilon}_0$ taka, że dla $\varepsilon \in (0, \tilde{\varepsilon}_0)$, klasa F_d^r zawiera $\tilde{n} = \Theta(\varepsilon^{-d/r})$ funkcji $\tilde{f}_1, \tilde{f}_2, \ldots, \tilde{f}_{\tilde{n}}$ o rozłącznych nośnikach zawartych w $[0, 1]^d$ takich, że $\max_{x \in [0,1]^d} \tilde{f}_i(x) = \varepsilon$ dla $i = 1, 2, \ldots, \tilde{n}$.

Dowód. Zauważmy, że dla funkcji f_i oraz $\tilde{f}_i$ zdefiniowanych przez (A.1) i (A.2) zachodzi

$$\max_{x \in [0,1]^d} f_i(x) = \max_{x \in [0,1]^d} \tilde{f}_i(x) = \varepsilon_1 \max_{x \in [-1,1]^d} G(x) = \varepsilon_1 a.$$

Wystarczy przyjąć $\varepsilon_1 = \varepsilon / a$ aby dowieść lemat. □

Lemat 9. *Dla klasy Höldera $F_d^{r,\rho}$ istnieje dodatnia stała ε_0 taka, że dla $\varepsilon \in (0, \varepsilon_0)$, klasa $F_d^{r,\rho}$ zawiera $n = \Theta(\varepsilon^{-d/(r+\rho)})$ nieujemnych funkcji $f_1, f_2, \ldots, f_n$ o rozłącznych nośnikach zawartych w $[0,1]^d$ takich, że $\int\limits_{[0,1]^d} f_i(x)dx = \varepsilon^{1+d/(r+\rho)}$ dla $i = 1, 2, \ldots, n$. Dla klasy F_d^r istnieje dodatnia stała $\tilde{\varepsilon}_0$ taka, że dla $\varepsilon \in (0, \tilde{\varepsilon}_0)$, klasa F_d^r zawiera $\tilde{n} = \Theta(\varepsilon^{-d/r})$ nieujemnych funkcji $\tilde{f}_1, \tilde{f}_2, \ldots, \tilde{f}_{\tilde{n}}$ o rozłącznych nośnikach zawartych w $[0,1]^d$ takich, że $\int\limits_{[0,1]^d} \tilde{f}_i(x)dx = \varepsilon^{1+d/r}$ dla $i = 1, 2, \ldots, \tilde{n}$.*

Dowód. Weźmy funkcje f_i zdefiniowane przez (A.1). Dla takich funkcji zachodzi

$$\int\limits_{[0,1]^d} f_i(x)dx = \int\limits_{[0,1]^d} \varepsilon_1 G\left((x - t_i)\varepsilon_1^{-1/(r+\rho)}\right) dx = \int\limits_{[-1,1]^d} \varepsilon_1 \varepsilon_1^{d/(r+\rho)} G(t)dt.$$

Niech $b = \int\limits_{[-1,1]^d} G(t)dt$. Wystarczy przyjąć $\varepsilon_1 = \varepsilon / b^{(r+\rho)/(r+\rho+d)}$ aby skończyć pierwszą część dowodu.

Weźmy funkcje $\tilde{f}_i$ zdefiniowane przez (A.2). Dla takich funkcji zachodzi

$$\int\limits_{[0,1]^d} \tilde{f}_i(x)dx = \int\limits_{[0,1]^d} \varepsilon_1 G\left((x - t_i)\varepsilon_1^{-1/r}\right) dx = \int\limits_{[-1,1]^d} \varepsilon_1 \varepsilon_1^{d/r} G(t)dt.$$

Wystarczy przyjąć $\varepsilon_1 = \varepsilon / b^{r/(r+d)}$ aby zakończyć dowód. □

Przedstawimy teraz pewną wersję nierówności Gronwalla, która używana jest w tej pracy.

Lemat 10 (Nierówność Gronwalla). *Niech X będzie przestrzenią Banacha, a $U \subset X$ zbiorem otwartym w X. Niech $f, g : [a,b] \times U \to X$ będą funkcjami ciągłymi ,a $y, z : [a,b] \to U$ będą rozwiązaniami problemów początkowych:*

$$y'(t) = f(t, y(t)), \quad y(a) = y_0$$

$$z'(t) = g(t, z(t)), \quad z(a) = z_0$$

Załóżmy, że istnieją dodatnia stała C i ciągła funkcja $\varphi : [a,b] \to [0, \infty)$ takie, że

$$\|g(t, x_2) - g(t, x_1)\| \leqslant C\|x_2 - x_1\|, \quad \forall x_1, x_2 \in [a,b] \ \ \forall t \in U,$$

$$\|f(t, y(t)) - g(t, y(t))\| \leqslant \varphi(t), \quad \forall t \in [a,b].$$

Wtedy, dla każdego $t \in [a,b]$

$$\|y(t) - z(t)\| \leqslant e^{C|t-a|}\|y_0 - z_0\| + e^{C|t-a|} \int\limits_a^t e^{-C|s-a|}\varphi(s)\mathrm{d}s.$$

Dowód tego lematu można znaleźć w [How].

Przypomnimy teraz postać i oszacowanie na resztę we wzorze Taylora. Niech funkcja $P : \mathbb{R}^d \to \mathbb{R}^d$ będzie n-krotnie różniczkowalna w kuli $K(x_0, r) = \{x \in \mathbb{R}^d : \|x - x_0\| \leqslant r\}, r > 0$, i niech $P^{(n)}$ będzie całkowalne od x_0 do dowolnego $x_1 \in K(x_0, r)$. Wtedy

$$P(x_1) = \sum_{k=0}^{n} \frac{1}{k!} P^{(k)}(x_0)(x_1 - x_0)^k + R_n(x_0, x_1),$$

gdzie

$$R_n(x_0, x_1) = \int_0^1 \left(P^{(n)}(\theta x_1 + (1-\theta)x_0) - P^{(n)}(x_0) \right) (x_1 - x_0)^n \frac{(1-\theta)^{n-1}}{(n-1)!} d\theta.$$

Następujący wniosek wynika bezpośrednio z oszacowania reszty we wzorze Taylora. Jeśli funkcja P spełnia powyższe założenia, to

$$\left\| P(x_1) - \sum_{k=0}^{n-1} \frac{1}{k!} P^{(k)}(x_0)(x_1 - x_0)^k \right\| \leqslant \sup_{\xi \in L(x_0, x_1)} \|P^{(n)}(\xi)\| \frac{\|x_1 - x_0\|^n}{n!},$$

gdzie $L(x_0, x_1)$ jest odcinkiem łączącym x_0 i x_1. Wzór Taylora oraz oszacowanie reszty wraz z dowodami można znaleźć między innymi w [Ral69].

Dodatek B

Dowody pomocniczych wyników

Przedstawimy tutaj dowody pewnych faktów i lematów wykorzystywanych w tej rozprawie. Zaczniemy od dowodu faktu 1 z podrozdziału 5.2.3.

Dowód faktu 1. Przyjmijmy, że $j < k$. Problem lokalny (5.6) można zamienić na układ równań:

$$\begin{cases} z'(t) = g(t, z(t)), & t \in [x_i, x_{i+1}] \\ z_{p+1}(x_i) = \delta_{pj}, & p = 0, 1, \ldots, k-1 \end{cases},$$

gdzie $z = [z_1, z_2, \ldots, z_{k-1}, z_k]^T$, $z_p = u_{j,i}^{(p-1)}$ dla $p = 1, 2, \ldots, k$, a $g(t, z) = [z_2, z_3, \ldots, z_k, \sum_{p=0}^{q} f_p(t) z_{p+1}]^T$. Niech y będzie rozwiązaniem problemu:

$$y'(t) = f(t, y(t)), \qquad t \in [x_i, x_{i+1}],$$

gdzie $y = [y_1, y_2, \ldots, y_{k-1}, y_k]^T$, a $f(t, y) = [y_2, y_3, \ldots, y_k, 0]^T$, z tymi samymi warunkami początkowymi. Sprawdźmy założenia nierówności Gronwalla cytowanej w dodatku A. Zauważmy, że

$$\begin{aligned} \|g(t, x^2) - g(t, x^1)\| &\leqslant \max\left\{ \|x^2 - x^1\|, \left| \sum_{p=0}^{q} f_p(t)(x_{p+1}^2 - x_{p+1}^1) \right| \right\} \\ &\leqslant \max\left\{ 1, \sum_{p=0}^{q} \|f_p\| \right\} \|x^2 - x^1\| \leqslant (kD+1)\|x^2 - x^1\|. \end{aligned}$$

Ponadto,

$$\|f(t,y(t))-g(t,y(t))\|=\left|\sum_{p=0}^{q} f_p(t)y_{p+1}(t)\right|$$
$$\leqslant\sum_{p=0}^{q}\|f_p\|\ \|y_{p+1}\|\leqslant kD\|y\|.$$

Jako, że $\|y\|$ zależy wyłącznie od przedziału oraz warunków początkowych, jest ona ograniczona przez pewną stałą K zależną wyłącznie od przedziału $[a,b]$. Możemy więc przyjąć w lemacie Gronwalla $\varphi(t)=DKk$ i $C=kD+1$. Dostajemy dla każdego $t\in[a,b]$:

$$\|y(t)-z(t)\|\leqslant e^{(kD+1)(t-x_i)}\int_{x_i}^{x_{i+1}} e^{-(kD+1)(s-x_i)}DKk\ ds\leqslant K(e^{(kD+1)h}-1),$$

tak więc,

$$\|z(t)\|\leqslant\|y(t)-z(t)\|+\|y(t)\|\leqslant Ke^{(kD+1)h}\leqslant Ke^{(kD+1)(b-a)}.$$

Stąd, dla każdych $p=0,1,\ldots,k-1$ i $t\in[x_i,x_{i+1}]$

$$|u_{j,i}^{(p)}(t)|=|z_{p+1}(t)|\leqslant Q,$$

gdzie $Q=Ke^{(kD+1)(b-a)}$.
Dowód dla $j=k$ jest analogiczny. □

Przedstawimy teraz dowód lematu 6 z podrozdziału 5.3.3 (patrz lemat 3.2 z [Kac02]).

Dowód lematu 6. Niech $s^0\in K(s^*,R)$. Pokażemy przez indukcję, że dla każdego l z prawdopodobieństwem co najmniej $(1-\delta)^{2l}$ ciąg $\{s^j\}_{j=0}^{l}$ jest wybierany z (5.71), $\{s^j\}_{j=0}^{l}\subset K(s^*,R)$ oraz zachodzi (5.75). Dla $l=1$ krok indukcyjny przebiega jak w ogólnym przypadku, więc opuścimy go. Załóżmy, że $s^0,s^1,\ldots,s^l\in K(s^*,R)$ oraz zachodzi (5.75) z prawdopodobieństwem co najmniej $(1-\delta)^{2l}$. Załóżmy ponadto, że w kroku l-tym algorytmy obliczające przybliżenia rozwiązań problemów (5.55) i (5.66) zwracają rozwiązania z założoną dokładności odpowiednio ε i $\sqrt[3]{\varepsilon}$. Zakładamy więc, że

$$\|y(x,s^l)-l(x,s^l)\|\leqslant\varepsilon\ \text{ i }\ \|\bar{Y}(x,s^l)-L(x,s^l)\|\leqslant\sqrt[3]{\varepsilon},\quad\forall x\in[a,b]. \tag{B.1}$$

Zachodzi to z prawdopodobieństwem co najmniej $(1-\delta)^2$.

Pokażemy najpierw, że $A(s^l)$ jest nieosobliwa. Wiadomo, że (zob. np. [KS92, lemat 4.1.2]) jeśli $\|F'(s^l)^{-1}\|\|F'(s^l)-A(s^l)\|<1$, to $A(s^l)$ jest nieosobliwa i

$$\|A(s^l)^{-1}\|\leqslant\frac{\|F'(s^l)^{-1}\|}{1-\|F'(s^l)^{-1}\|\|F'(s^l)-A(s^l)\|}. \tag{B.2}$$

Potrzebujemy więc ograniczeń na $\|F'(s^l) - A(s^l)\|$ oraz $\|F'(s^l)^{-1}\|$. Znajdziemy najpierw ograniczenia na $\|F'(s^l) - A(s^l)\|$. Korzystając z nierówności trójkąta dostajemy

$$\begin{aligned}\|F'(s^l) - A(s^l)\| = &\left\| \frac{\partial p}{\partial s}(s^l, y(b, s^l)) + \frac{\partial p}{\partial w}(s^l, y(b, s^l))Y(b, s^l) \right. \\ &\left. -\frac{\partial p}{\partial s}(s^l, l(b, s^l)) - \frac{\partial p}{\partial w}(s^l, l(b, s^l))L(b, s^l) \right\| \\ \leqslant &\left\| \frac{\partial p}{\partial s}(s^l, y(b, s^l)) - \frac{\partial p}{\partial s}(s^l, l(b, s^l)) \right\| \\ &+ \left\| \left(\frac{\partial p}{\partial w}(s^l, y(b, s^l)) - \frac{\partial p}{\partial w}(s^l, l(b, s^l)) \right) Y(b, s^l) \right\| \\ &+ \left\| \frac{\partial p}{\partial w}(s^l, l(b, s^l)) \left(Y(b, s^l) - L(b, s^l) \right) \right\|.\end{aligned}$$

Jako, że $\frac{\partial p}{\partial s}$ i $\frac{\partial p}{\partial w}$ są funkcjami Lipschitza, mamy

$$\begin{aligned}\|F'(s^l) - A(s^l)\| \leqslant &\tilde{C}_1 \|y(b, s^l) - l(b, s^l)\| + \tilde{C}_2 \|y(b, s^l) - l(b, s^l)\| \|Y(b, s^l)\| \\ &+\tilde{C}_3 \|Y(b, s^l) - L(b, s^l)\|, \end{aligned} \tag{B.3}$$

gdzie $\tilde{C}_1$ i $\tilde{C}_2$ są stałymi Lipschitza funkcji $\frac{\partial p}{\partial s}$ i $\frac{\partial p}{\partial w}$, a $\tilde{C}_3$ jest stałą ograniczającą $\frac{\partial p}{\partial w}$. Potrzebujemy jeszcze ograniczeń na $\|Y(b, s^l)\|$ i $\|Y(b, s^l) - L(b, s^l)\|$.
Z postaci całkowej problemu (5.65) oraz z definicji klasy $\tilde{F}_d^r$ mamy

$$\begin{aligned}\|Y(x, s^l)\| = &\left\| I + \int_a^x \frac{\partial f}{\partial y}(t, y(t, s^l))Y(t, s^l)dt \right\| \leqslant 1 + \int_a^x \left\| \frac{\partial f}{\partial y}(t, y(t, s^l)) \right\| \|Y(t, s^l)\|dt \\ \leqslant &\, 1 + \int_a^x D\|Y(t, s^l)\|dt.\end{aligned}$$

Z klasycznej nierówności Gronwalla (zob. np. [HNW00, s. 61] lub dodatek A) dostajemy

$$\|Y(x, s^l)\| \leqslant e^{D(x-a)} \leqslant e^{D(b-a)}, \qquad x \in [a, b]. \tag{B.4}$$

Korzystając z (B.1) i nierówności trójkąta mamy

$$\begin{aligned}\|Y(b, s^l) - L(b, s^l)\| &\leqslant \|Y(b, s^l) - \bar{Y}(b, s^l)\| + \|\bar{Y}(b, s^l) - L(b, s^l)\| \\ &\leqslant \|Y(b, s^l) - \bar{Y}(b, s^l)\| + \sqrt[3]{\varepsilon}.\end{aligned} \tag{B.5}$$

Z postaci całkowej problemów (5.65) i (5.66) mamy

$$\begin{aligned}\|Y(x, s^l) - \bar{Y}(x, s^l)\| \leqslant &\int_a^x \left\| \frac{\partial f}{\partial y}(t, l(t, s^l)) \right\| \|Y(t, s^l) - \bar{Y}(t, s^l)\|dt \\ &+ \int_a^x \left\| \frac{\partial f}{\partial y}(t, y(t, s^l)) - \frac{\partial f}{\partial y}(t, l(t, s^l)) \right\| \|Y(t, s^l)\|dt.\end{aligned}$$

Z warunków klasy $\tilde{F}_d^r$, z warunku Lipschitza dla $\frac{\partial f}{\partial y}$ (wynika on z ograniczoności drugich pochodnych cząstkowych f przez D), z (B.1) oraz (B.4) dostajemy

$$\begin{aligned}\|Y(x,s^l)-\bar{Y}(x,s^l)\| &\leqslant \int_a^x D\|Y(t,s^l)-\bar{Y}(t,s^l)\|dt + \int_a^x \tilde{C}_4 \left\|y(t,s^l)-l(t,s^l)\right\| dt \\ &\leqslant D\int_a^x \|Y(t,s^l)-\bar{Y}(t,s^l)\|dt + \tilde{C}_5\varepsilon,\end{aligned}$$

gdzie $\tilde{C}_4 = De^{D(b-a)}$, $\tilde{C}_5 = (b-a)\tilde{C}_4$. Z nierówności Gronwalla dla funkcji $\|Y(\cdot,s^l) - \bar{Y}(\cdot,s^l)\|$ dostajemy więc

$$\|Y(x,s^l)-\bar{Y}(x,s^l)\| \leqslant \tilde{C}_5\varepsilon e^{D(b-a)} = \tilde{C}_6\varepsilon,$$

gdzie $\tilde{C}_6$ jest stałą zależną wyłącznie od a, b, D. Wracając do (B.5) otrzymujemy

$$\|Y(b,s^l)-L(b,s^l)\| \leqslant \tilde{C}_6\varepsilon + \sqrt[3]{\varepsilon} \leqslant \tilde{C}_7\sqrt[3]{\varepsilon}, \tag{B.6}$$

gdzie $\tilde{C}_7$ zależy wyłącznie od a, b, D.
Powracając do (B.3) z wykorzystaniem (B.4) i (B.6) dostajemy

$$\|F'(s^l)-A(s^l)\| \leqslant (\tilde{C}_1 + \tilde{C}_2e^{D(b-a)})\varepsilon + \tilde{C}_3\tilde{C}_7\sqrt[3]{\varepsilon} \leqslant C_1\sqrt[3]{\varepsilon}, \tag{B.7}$$

gdzie C_1 jest stałą zależną wyłącznie od a, b, D i p.

Stosując standardowe metody, wykorzystując twierdzenie Taylora, dostajemy, że $F'(s^l)$ jest nieosobliwa oraz $\|F'(s^l)^{-1}\| \leqslant 4/3H$. Stąd oraz z (B.7) mamy, że $\|F'(s^l)^{-1}\|\|F'(s^l)-A(s^l)\| \leqslant 4/3HC_1\varepsilon \leqslant 1/2$ dla $\varepsilon \leqslant 3/(8HC_1)$. Z (B.2) mamy więc, że $A(s^l)$ jest nieosobliwa i

$$\|A(s^l)^{-1}\| \leqslant N_1 = 8/3H. \tag{B.8}$$

Zachodzi to z prawdopodobieństwem $(1-\delta)^{2(l+1)}$.
Pokażemy teraz, że s^{l+1} jest wybierany ze wzoru (5.71), leży w kuli $K(s^*,R)$ i zachodzi (5.75). Mamy, że

$$\begin{aligned}\beta &:= s^l - A(s^l)^{-1}\hat{F}(s^l) \\ &= \left(s^l - F'(s^l)^{-1}F(s^l)\right) + \left(F'(s^l)^{-1}F(s^l) - A(s^l)^{-1}\hat{F}(s^l)\right).\end{aligned}$$

Pierwszy składnik prawej strony $N(s^l) := s^l - F'(s^l)^{-1}F(s^l)$ jest krokiem iteracyjnym Newtona dla równania $F(s)=0$. Z (5.61) i nierówności trójkąta mamy

$$\begin{aligned}\|\beta - s^*\| &\leqslant \left\|N(s^l)-s^*\right\| + \left\|F'(s^l)^{-1}F(s^l) - A(s^l)^{-1}\hat{F}(s^l)\right\| \\ &\leqslant C\|s^l-s^*\|^2 + \left\|F'(s^l)^{-1}F(s^l) - A(s^l)^{-1}\hat{F}(s^l)\right\|.\end{aligned} \tag{B.9}$$

Drugi składnik prawej strony jest ograniczony przez

$$\begin{aligned}
&\left\|F'(s^l)^{-1}F(s^l) - A(s^l)^{-1}\hat{F}(s^l)\right\| \\
&\quad\leqslant \left\|A(s^l)^{-1}\left(A(s^l) - F'(s^l)\right)F'(s^l)^{-1}F(s^l)\right\| + \left\|A(s^l)^{-1}\left(F(s^l) - \hat{F}(s^l)\right)\right\| \\
&\quad\leqslant \left\|A(s^l)^{-1}\right\|\left\|A(s^l) - F'(s^l)\right\|\left\|s^l - N(s^l)\right\| + \left\|A(s^l)^{-1}\right\|\left\|F(s^l) - \hat{F}(s^l)\right\|. \quad \text{(B.10)}
\end{aligned}$$

Korzystając z warunku Lipschitza dla p i (B.1) dostajemy

$$\|F(s^l) - \hat{F}(s^l)\| = \|p(s, y(b, s^l)) - p(s, l(b, s^l))\| \leqslant C_2\|y(b, s^l) - l(b, s^l)\| \leqslant C_2\varepsilon, \quad \text{(B.11)}$$

gdzie C_2 jest stałą zależną wyłącznie od p. Z (B.10) i z (B.8) dla dostatecznie małego ε mamy

$$\left\|F'(s^l)^{-1}F(s^l) - A(s^l)^{-1}\hat{F}(s^l)\right\| \leqslant N_1C_1\sqrt[3]{\varepsilon}\left(\|s^l - s^*\| + \|N(s^l) - s^*\|\right) + N_1C_2\varepsilon.$$

Wracając do (B.9), używając powyższego, (5.61) i założenia indukcyjnego, dla dostatecznie małego ε mamy

$$\begin{aligned}
\|\beta - s^*\| &\leqslant C\|s^l - s^*\|^2\left(1 + N_1C_1\sqrt[3]{\varepsilon}\right) + N_1C_1\sqrt[3]{\varepsilon}\|s^l - s^*\| + N_1C_2\varepsilon \quad \text{(B.12)} \\
&\leqslant CR^2(1 + N_1C_1\sqrt[3]{\varepsilon}) + N_1C_1\sqrt[3]{\varepsilon}R + N_1C_2\varepsilon \quad \text{(B.13)} \\
&\leqslant \frac{1}{4}R(1 + N_1C_1\sqrt[3]{\varepsilon}) + N_1C_1\sqrt[3]{\varepsilon}R + \frac{1}{2}R \\
&\leqslant R.
\end{aligned}$$

Mamy, więc $\|\beta - s^0\| \leqslant \|\beta - s^*\| + \|s^* - s^0\| \leqslant 2R \leqslant R_0$. Stąd, $s^{l+1} = \beta$.
Dostaliśmy więc, że jeśli $s^l \in K(s^*, R)$, to z prawdopodobieństwem co najmniej $(1-\delta)^2$, macierz $A(s^l)$ jest nieosobliwa, s^{l+1} wybierane jest ze wzoru (5.71) i $s^{l+1} \in K(s^*, R)$. Ponadto z (B.12), przyjmując $\tilde{N}_1 = N_1C_1$ oraz $\tilde{N}_2 = N_1C_2$ dostajemy żądaną nierówność na $\|s^{l+1} - s^*\|$. Wynika z tego, że z prawdopodobieństwem co najmniej $(1-\delta)^{2(l+1)}$, ciąg $\{s^j\}_{j=0}^{l+1}$ leży w kuli $K(s^*, R)$ i zachodzi (5.75) dla $l := l+1$, co kończy dowód indukcyjny. □

Udowodnimy teraz lemat 7 z podrozdziału 5.3.3 (patrz lemat 3.3 z [Kac02]).

Dowód lematu 7. Niech $e^l = \|s^l - s^*\|$. Niech $S = 2(1 + \tilde{N}_1\sqrt[3]{\varepsilon})C$ i $\alpha_0 = 2R\tilde{N}_1\sqrt[3]{\varepsilon} + 2\tilde{N}_2\varepsilon$. Z lematu 6 wynika, że $e^l \leqslant R$ i

$$e^{l+1} \leqslant \frac{1}{2}\left((e^l)^2S + \alpha_0\right). \quad \text{(B.14)}$$

Zachodzi to z prawdopodobieństwem co najmniej $(1-\delta)^{2(l+1)}$.
Zdefiniujmy ciąg

$$\alpha_{n+1} = \sqrt{\frac{\alpha_n}{S}}, \qquad n = 0, 1, \ldots.$$

Zauważmy, że dla dostatecznie małego ε zachodzi

$$S\alpha_0 \leqslant 1. \tag{B.15}$$

Pokażemy teraz przez indukcję, że $\alpha_n \geqslant \alpha_0$, dla $n = 0, 1, \ldots$. Zauważmy najpierw, że z definicji ciągu $\{\alpha_n\}$ oraz z (B.15) wynika, że $\alpha_1 = \sqrt{\alpha_0/S} \geqslant \sqrt{\alpha_0^2} = \alpha_0$. Załóżmy, że $\alpha_n \geqslant \alpha_0$. Dostajemy

$$\alpha_{n+1} = \sqrt{\frac{\alpha_n}{S}} \geqslant \sqrt{\frac{\alpha_0}{S}} = \alpha_1 \geqslant \alpha_0.$$

Dostaliśmy więc, że ciąg $\{\alpha_n\}$ jest ograniczony z dołu przez α_0.
Pokażemy teraz indukcyjnie (ze względu na n), że

$$\text{jeśli } e^0 \leqslant \alpha_l, \text{ to } e^n \leqslant \alpha_{l-n}, \quad \text{dla } n = 0, 1, \ldots, l. \tag{B.16}$$

Załóżmy, że $e^n \leqslant \alpha_{l-n}$ dla pewnego n. Z (B.14) i ograniczoności z dołu ciągu $\{\alpha_n\}$ wynika, że dla dowolnego $n = 0, 1, \ldots, l$

$$e^{n+1} \leqslant \frac{1}{2}\left((e^n)^2 S + \alpha_0\right) \leqslant \frac{1}{2}\left(\alpha_{l-n}^2 S + \alpha_0\right) = \frac{1}{2}\left(\alpha_{l-n-1} + \alpha_0\right) \leqslant \alpha_{l-n-1},$$

co kończy indukcję.
Ponieważ $e^0 \leqslant R$, więc z (B.16) wynika, że

$$\text{jeśli } R \leqslant \alpha_l, \text{ to } e^l \leqslant \alpha_0. \tag{B.17}$$

Z definicji ciągu α_n wynika, że $\alpha_n = S\alpha_{n+1}^2$ czyli $S\alpha_n = (S\alpha_{n+1})^2$. Stąd, mamy

$$\alpha_0 = 1/S(S\alpha_l)^{2^l}. \tag{B.18}$$

Z tego wynika, że $R \leqslant \alpha_l$ wtedy i tylko wtedy, gdy $(SR)^{2^l} \leqslant S\alpha_0$.

Pokażemy teraz, że jeśli $l \geqslant \log\log(1/\varepsilon) - \log\log 4/3$, to dla dostatecznie małego ε zachodzi $e^l \leqslant \tilde{C}\sqrt[3]{\varepsilon}$, dla pewnej stałej $\tilde{C}$ zależnej tylko od a, b, p i parametrów klasy. Rzeczywiście, dla dostatecznie małego ε mamy $S \leqslant 3C$, a więc $SR = S/(4C) \leqslant 3/4$. Jeśli więc $(3/4)^{2^l} \leqslant \varepsilon$, to $(SR)^{2^l} \leqslant \varepsilon \leqslant S\alpha_0$ i z (B.18) mamy $R \leqslant \alpha_l$. Z (B.17) otrzymujemy, że $e^l \leqslant \alpha_0 \leqslant \tilde{C}\sqrt[3]{\varepsilon}$, gdzie $\tilde{C} = 2R\tilde{N}_1 + 2\tilde{N}_2$, co należało pokazać.
Po kolejnym kroku, korzystając z lematu 6 dostajemy

$$e^{l+1} \leqslant C\tilde{C}^2\sqrt[3]{\varepsilon}^2(1 + \tilde{N}_1\sqrt[3]{\varepsilon}) + \tilde{C}\tilde{N}_1\sqrt[3]{\varepsilon}^2 + \tilde{N}_2\varepsilon \leqslant \tilde{K}(\sqrt[3]{\varepsilon}^2),$$

gdzie $\tilde{K}$ jest stałą zależną od a, b ,D, H, Z i p. Po jeszcze jednym kroku dostajemy,

$$e^{l+2} \leqslant C\tilde{K}^2\sqrt[3]{\varepsilon}^4(1 + \tilde{N}_1\sqrt[3]{\varepsilon}) + \tilde{K}\tilde{N}_1\varepsilon + \tilde{N}_2\varepsilon \leqslant N_3\varepsilon,$$

gdzie N_3 jest stałą zależną od a, b ,D, H, Z i p, a niezależną od l i ε.
Te ograniczenia zachodzą przy założeniu, że $\{s^j\}_{j=0}^{l+2} \subset K(s^*, R)$ oraz na każdym kroku zachodzi nierówność rekurencyjna na $\|s^{l+2} - s^*\|$ w lemacie 6. Zgodnie z lematem 6 zachodzi to z prawdopodobieństwem co najmniej $(1-\delta)^{2(l+2)}$. Jest to prawdopodobieństwo, że na każdym z $l+2$ kroków, przybliżenia $l(x,s)$ and $L(x,s)$ są obliczone z dokładnością odpowiednio ε i $\sqrt[3]{\varepsilon}$. Ponieważ nierówność na $\|s^{l+2} - s^*\|$ zachodzi dla $l \geqslant \log\log(1/\varepsilon) - \log\log 4/3$, więc nierówność na $\|s^l - s^*\|$ potrzebna w lemacie zachodzi dla $l \geqslant \lceil 2\log\log(1/\varepsilon)\rceil$, dla ε dostatecznie małych. To kończy dowód lematu. □

Bibliografia

[AW99] Daniel Abrams, Colin Williams. Fast quantum algorithms for numerical integrals and stochastic processes. Raport techniczny, http://arxiv.org/abs/quant-ph/9908083, 1999.

[BBHT98] Michel Boyer, Gilles Brassard, Peter Høyer, Alain Tapp. Tight bounds on quantum searching. *Forschritte Der Physik*, 46:493–505, 1998. zobacz również http://arXiv.org/abs/quant-ph/9605034.

[BBR+98] Robert Beals, Harry Buhrman, Cleve Richard, Mosca Michel, Ronald de Wolf. Quantum lower bounds by polynomials. *Proceeding of the 39th Annual IEEE Symposium on Foundations of Computer Science*, strony 352–361, 1998. zobacz również http://arXiv.org/abs/quant-ph/9802049.

[BHT98] Gilles Brassard, Peter Høyer, Alain Tapp. Quantum counting. *Proceedings of the 25th International Colloquium on Automata, Languages and Programming, Lecture Notes in Computer Science*, 1443:820–831, 1998.

[Dau11] Thomas Daun. On the randomized solution of initial value problems. *Journal of Complexity*, 27:300–311, 2011.

[Deu85] David Deutsch. Quantum theory, the Church-Turing principle and the universal quantum computer. *Proceedings of the Royal Society of London A 400*, strony 97–117, 1985.

[DH98] Christoph Dürr, Peter Høyer. A quantum algorithm for finding the minimum. *Proceeding of the 30th Annual ACM Symposium on Theory of Computing*, strony 1516–1524, 1998. zobacz również http://arXiv.org/abs/quant-ph/9607014.

[Drw07] Tomasz Drwięga. The use of integral information in the solution of a two-point boundary value problem. *Opuscula Mathematica*, 27(2):205–202, 2007.

[Fey82] Richard Phillips Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6/7):467–488, 1982.

[GK03] Krzysztof Giaro, Marcin Kamiński. *Wprowadzenie do algorytmow kwantowych*. Akademicka Oficyna Wydawnicza EXIT, Warszawa, 2003.

[Goć06] Maciej Goćwin. On the complexity of searching for a maximum of a function on a quantum computer. *Quantum Information Processing*, 5(1):31–41, 2006.

[Gro96] Lov Kumar Grover. A fast quantum mechanical algorithm for database search. *Proceedings of the 28th ACM Symposium on Theory of Computing*, strony 212–219, 1996.

[Gro98] Lov Kumar Grover. A framework for fast quantum mechanical algorithms. *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, strony 53–62, 1998.

[GS08] Maciej Goćwin, Marek Szczęsny. Randomized and quantum algorithms for solving initial-value problems in ordinary differential equations of order k. *Opuscula Mathematica*, 28(3):247–277, 2008.

[GS10] Maciej Goćwin, Marek Szczęsny. On the complexity of a two-point boundary value problem in different settings. *International Journal of Computer Mathematics*, 87(15):3370–3386, 2010.

[Hei02] Stefan Heinrich. Quantum summation with an application to integration. *Journal of Complexity*, 18:1–50, 2002. zobacz również `http://arXiv.org/abs/quant-ph/0105116`.

[Hei03a] Stefan Heinrich. From Monte Carlo to quantum computation. *Mathematics and Computers in Simulation*, 62:219–230, 2003. zobacz również `http://arXiv.org/abs/quant-ph/0112152`.

[Hei03b] Stefan Heinrich. Quantum integration in Sobolev classes. *Journal of Complexity*, 19:19–42, 2003. zobacz również `http://arXiv.org/abs/quant-ph/0112153`.

[Hei04a] Stefan Heinrich. Quantum approximation I. Embeddings of finite dimensional L_p spaces. *Journal of Complexity*, 20:5–26, 2004. zobacz również `http://arXiv.org/abs/quant-ph/0305030`.

[Hei04b] Stefan Heinrich. Quantum approximation II. Sobolev embedding. *Journal of Complexity*, 20:27–45, 2004. zobacz również `http://arXiv.org/abs/quant-ph/0305031`.

[Hir04] Mika Hirvensalo. *Algorytmy kwantowe*. Idee, metody i narzędzia informatyki. WSiP, Warszawa, 2004.

[HM08] Stefan Heinrich, Berhnard Milla. The randomized complexity of initial value problems. *Journal of Complexity*, 24:77–88, 2008.

[HNW00] Ernst Hairer, Syvert P. Nørsett, Gerhard Wanner. *Solving ordinary differential equations I.* Springer, Berlin, 2000.

[How] Ralph Howard. Gronwall inequality. `http://www.math.sc.edu/~howard/Notes/gronwall.pdf`.

[Kac87] Bolesław Kacewicz. Optimal solution of ordinary differential equations. *Journal of Complexity*, 3:451–465, 1987.

[Kac90] Bolesław Kacewicz. Optimal solution of some two-point boundary value problem. *Banach Center Publications*, 24:241–256, 1990.

[Kac02] Bolesław Kacewicz. Complexity of nonlinear two-point boundary-value problems. *Journal of Complexity*, 18:702–738, 2002.

[Kac04] Bolesław Kacewicz. Randomized and quantum algorithms yield a speed-up for initial-value problems. *Journal of Complexity*, 20:821–834, 2004. zobacz również `http://arXiv.org/abs/quantph/0311148`.

[Kac06] Bolesław Kacewicz. Almost optimal solution of initial-value problems by randomized and quantum algorithms. *Journal of Complexity*, 22:676–690, 2006. zobacz również `http://arXiv.org/abs/quantph/0510045`.

[KS92] Andrzej Kiełbasiński, Hubert Schwetlick. *Numeryczna algebra liniowa*. WNT, Warszawa, wydanie 2, 1992.

[Man80] Yuri Ivanovich Manin. Computable and uncomputable. Sovetskoye Radio, Moskwa, 1980. w języku rosyjskim.

[Mat92] Peter Mathé. Random approximation of finite sums. Preprint 11, Institute for Applied Analysis and Stochastics, Berlin, 1992.

[NC00] Michael Nielsen, Isaac Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[Nov88] Erich Novak. *Deterministic and Stochastic Error Bounds in Numerical Analysis*, wolumen 1349 serii *Lecture Notes in Mathematics*. Springer-Verlag, 1988.

[Nov01] Erich Novak. Quantum complexity of integration. *Journal of Complexity*, 17:2–16, 2001. zobacz również `http://arXiv.org/abs/quant-ph/0008124`.

[NW99] Ashwin Nayak, Felix Wu. The quantum query complexity of approximating the median and related statistics. *Proceedings of the Thirty-First Annual ACM Symposium on the Theory of Computing*, strony 384–393, 1999. zobacz również `http://arXiv.org/abs/quantph/9804066`.

[Ral69] Louis B. Rall. *Computational solution of nonlinear operator equations*. John Wiley & Sons, Inc., New York, USA, 1969.

[Ren87] James Renegar. On the worst-case arithmetic complexity of approximating zeros of polynomials. *Journal of Complexity*, 3:90–113, 1987.

[Ren89] James Renegar. On the worst-case arithmetic complexity of approximating zeros of systems of polynomials. *SIAM J. Comput*, 18:350–370, 1989.

[Sho94] Peter Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, Santa Fe, NM, Listopad 1994. zobacz również `http://arxiv.org/abs/quant-ph/9508027`.

[Sik83] Krzysztof Sikorski. Optimal solution of nonlinear equations satisfying a Lipschitz condition. *Numerische Mathematik*, 43:225–240, 1983.

[Szc06] Marek Szczęsny. Complexity of initial-value problems for ordinary differential equations of order k. *Journal of Complexity*, 22:514–532, 2006.

[TWW88] Joseph F. Traub, Greg Wasilkowski, Henryk Woźniakowski. *Information-Based Complexity*. Academic Press Professional, Inc., San Diego, CA, USA, 1988.