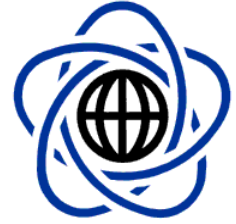




POLITECHNIKA WARSZAWSKA

WYDZIAŁ MATEMATYKI
I NAUK INFORMACYJNYCH



PRACA DYPLOMOWA MAGISTERSKA
MATEMATYKA

**METODY OSZCZĘDNEGO PRÓBKOWANIA
W ALGEBRZE KWATERNIONÓW**

COMPRESSED SENSING METHODS
IN THE QUATERNION ALGEBRA

AUTOR:

ŁUKASZ BŁASZCZYK

PROMOTOR:

DR AGNIESZKA BADEŃSKA

WARSZAWA, CZERWIEC 2016

.....

podpis promotora

.....

podpis autora

Streszczenie

Teoria oszczędnego próbkowania, stworzona przez E. Candès'a około 10 lat temu, jest wciąż dynamicznie rozwijającą się dziedziną nauki, angażującą do pracy zarówno matematyków jak i inżynierów. Jej podstawą jest fakt, że wektory rzadkie (tzn. mające niewiele niezerowych współrzędnych) mogą być zrekonstruowane z małej liczby liniowych pomiarów za pomocą nieliniowych algorytmów rekonstrukcji (np. minimalizacji normy ℓ_1). Jednym z warunków wystarczających powodzenia tej rekonstrukcji jest spełnienie przez macierz pomiarową tzw. *warunku ograniczonej izometrii* (RIP).

Tematem tej pracy jest rozszerzenie metod oszczędnego próbkowania, znanych dla macierzy i wektorów rzeczywistych oraz zespolonych, na algebrę kwaternionów. Kwaterniony, tzn. liczby postaci $r_0 + r_1\mathbf{i} + r_2\mathbf{j} + r_3\mathbf{k}$, stanowią uogólnienie liczb zespolonych. Ze względu na ich liczne zastosowania, m.in. w przetwarzaniu obrazów, rozszerzenie tej teorii wydaje się mieć duże znaczenie praktyczne.

W pierwszych dwóch rozdziałach przytoczone zostaną podstawy teorii oszczędnego próbkowania w ciele liczb rzeczywistych oraz najważniejsze własności algebry kwaternionów i kwaternionowych przestrzeni wektorowych. Rozdział 3 poświęcony jest uogólnieniu teorii oszczędnego próbkowania na przypadek kwaternionowy i zawiera główny wynik naukowy tej pracy, tzn. twierdzenie o stabilnej rekonstrukcji wektorów kwaternionowych z liniowych pomiarów o współczynnikach kwaternionowych. W Rozdziale 4 przeprowadzona zostanie analiza kwaternionowych losowych macierzy gaussowskich z punktu widzenia spełnienia założeń twierdzenia z poprzedniego rozdziału. Na koniec zaprezentowane zostaną wyniki symulacji numerycznych ilustrujących przeprowadzone wcześniej rozważania teoretyczne. Wskazane zostaną również perspektywy praktycznych zastosowań teorii oszczędnego próbkowania w algebrze kwaternionów i plan dalszych prac w tej dziedzinie.

Słowa kluczowe: oszczędne próbkowanie, własność ograniczonej izometrii, algebra kwaternionów

Summary

Compressed sensing theory, initiated about 10 years ago by E. Candès, is a still rapidly developing field of science, involving the work of both mathematicians and engineers. It is based on the fact that sparse vectors (i.e. having few non-zero coordinates) can be reconstructed from a small number of their linear measurements by non-linear reconstruction algorithms (e.g. ℓ_1 -norm minimization). One of the sufficient conditions for the successful reconstruction is so-called *restricted isometry property* (RIP) of the measurement matrix.

The aim of this paper is to generalize the compressed sensing methods, well-known for real and complex matrices and vectors, to the quaternion algebra. Quaternions, i.e. numbers of the form $r_0 + r_1\mathbf{i} + r_2\mathbf{j} + r_3\mathbf{k}$, are the extension of complex numbers. Due to their numerous applications, including in image processing, the extension of this theory seems to be of great practical importance.

In the first two chapters we recall the basis of compressed sensing theory in the field of real numbers and the most important properties of quaternions and quaternionic vector spaces. Chapter 3 is devoted to the generalization of compressed sensing theory to the quaternion algebra and includes the main scientific result of this paper, i.e. concerning the stable reconstruction of quaternion vectors from their linear measurements with quaternion coefficients. In Chapter 4 we analyze quaternionic gaussian random matrices from the point of view of fulfilling the assumptions of the main theorem. Finally, we present results of some numerical simulations to illustrate the previous theoretical considerations. We also point out some practical applications of the compressed sensing theory in the quaternion algebra and show the plan of future work in this field.

Keywords: compressed sensing, restricted isometry property, quaternion algebra

*Składam serdeczne podziękowania dr Agnieszce Badeńskiej
za nieocenioną pomoc, wyrozumiałość, poświęcony czas oraz
wielkie zaangażowanie okazane przy powstawaniu tej pracy.*

Spis treści

1 Wstęp – sformułowanie problemu i definicje	1
1.1. Wprowadzenie	1
1.2. Oszczędne próbkowanie – przypadek rzeczywisty	3
2 Algebra kwaternionów	7
2.1. Podstawowe pojęcia i własności	7
2.2. Kwaternionowe przestrzenie wektorowe	11
3 Uogólnienie metod oszczędnego próbkowania na na algebrę kwaternionów	21
3.1. Podstawowe pojęcia i sformułowanie problemu	21
3.2. Wcześniejsze wyniki dla pomiarów o współczynnikach rzeczywistych	24
3.3. Rekonstrukcja danych kwaternionowych z pomiarów o wsp. kwaternionowych	25
4 Analiza kwaternionowych losowych macierzy gaussowskich	33
4.1. Stałe ograniczonej izometrii	35
4.2. Symulacje numeryczne	42
5 Zastosowania oszczędnego próbkowania	45
5.1. Kilka eksperymentów numerycznych	45
5.2. Perspektywy zastosowań w praktyce	49
6 Podsumowanie i wnioski	51
Bibliografia	52

Rozdział 1

Wstęp – sformułowanie problemu i definicje

1.1. Wprowadzenie

Jednym z najważniejszych etapów badania własności różnych obiektów jest proces wykonywania pomiarów (bezpośrednich lub pośrednich) sygnałów reprezentujących pewne wielkości fizyczne. Wyciąganie wniosków dotyczących tych obiektów i podejmowanie decyzji związanych z nimi jest ściśle uzależnione od procesu rekonstrukcji badanych sygnałów z danych pomiarowych. Jako przykład można tu podać zagadnienie rekonstrukcji obrazów w tomografii rezonansu magnetycznego, gdzie na podstawie sygnałów elektromagnetycznych zarejestrowanych przez cewki odbiorcze można odtworzyć rozkład pewnych wielkości, np. gęstości atomowych, w ciele człowieka.

Proces akwizycji pomiarów jest najczęściej liniowy. Chcąc więc sformalizować rozważany problem, opisuje się go za pomocą układu równań liniowych, tzn.

$$\mathbf{y} = \mathbf{\Phi}\mathbf{x}, \tag{1.1}$$

gdzie $\mathbf{x} \in \mathbb{K}^n$ jest badanym n -wymiarowym sygnałem dyskretnym, a $\mathbf{y} \in \mathbb{K}^m$ to wektor m próbek, jakie zostały zebrane. Macierz $\mathbf{\Phi} \in \mathbb{K}^{m \times n}$ modeluje liniowy proces pomiarowy (i nazywana jest w związku z tym *macierzą pomiarową*). Zazwyczaj \mathbb{K} oznacza ciało liczb rzeczywistych \mathbb{R} lub zespolonych \mathbb{C} .

Widzimy więc, że aby uzyskać informacje na temat sygnału \mathbf{x} , musimy rozwiązać układ równań liniowych (1.1). Algebra liniowa mówi nam, że aby było to możliwe, liczba pomiarów m musi być co najmniej równa wymiarowi n sygnału. Zakładamy przy tym, że pomiary nie dostarczają sprzecznych informacji i są dokładne. Wiemy także, że w przeciwnym przypadku, tzn. gdy $m < n$, jeśli badany układ równań ma co najmniej jedno rozwiązanie, to ma ich nieskończenie wiele, a zatem nie jest możliwe jednoznaczne rozwiązanie problemu rekonstrukcji sygnału \mathbf{x} z pomiarów \mathbf{y} .

Powyższe trudności, a także obserwacja, że sygnały, z jakimi mamy do czynienia w rzeczywistości to zazwyczaj sygnały *rzadkie* lub *kompresowalne* (tzn. takie, których większość współrzędnych jest równa lub bliska zeru), przyczyniły się do powstania teorii *oszczędnego próbkowania* (ang. *compressed sensing* lub *compressive sampling*). Jej autorstwo przypisuje się Emmanuelowi Candèsowi, który wraz m.in. z Justinem Rombergiem i Terencem Tao opisał własności macierzy pomiarowych (rzeczywistych lub zespolonych), które umożliwiają dokładną rekonstrukcję pewnej klasy sygnałów z niewielkiej ilości liniowych pomiarów [7, 9].

Od początku jej istnienia, dziedzina oszczędnego próbkowania dynamicznie się rozwija, angażując do pracy zarówno matematyków jak i inżynierów i znajdując coraz to nowsze zastosowania praktyczne, jak i motywując do rozważań teoretycznych. Dotychczas uwaga naukowców skupiona była przede wszystkim na przypadku sygnałów i pomiarów o współrzędnych rzeczywistych lub zespolonych. Zaczęły pojawiać się jednak prace [27], których autorzy przeprowadzają symulacje numeryczne sugerujące, że metody oszczędnego próbkowania mogą być z powodzeniem stosowane do badania sygnałów o współrzędnych w pierścieniu kwaternionów \mathbb{H} (czasami nazywanym *ciałem nieprzemienne*), tzn. liczb postaci

$$r_0 + r_1\mathbf{i} + r_2\mathbf{j} + r_3\mathbf{k},$$

gdzie \mathbf{i} , \mathbf{j} oraz \mathbf{k} są trzema jednostkami urojonymi. Zgodnie z naszą wiedzą, do tej pory jednak nie pojawiły się prace teoretyczne, które wyjaśniałyby sukces tych eksperymentów.

Uogólnienie wyników na przypadek kwaternionowy może mieć duże znaczenie, ze względu na liczne zastosowania kwaternionów. Obok klasycznych zastosowań w mechanice kwantowej i do opisu obrotów brył sztywnych w grafice komputerowej [1], kwaterniony znalazły zastosowanie również w dziedzinie przetwarzania sygnałów. Struktura kwaternionów (trzy

różne jednostki urojone) pozwala na wykorzystanie ich do opisu obrazów kolorowych [21]. Część urojona kwaternionu może być interpretowana jako trzy składowe kolorowe w obrazie RGB (czerwony, zielony i niebieski). M.in. z tego powodu kwaterniony znalazły zastosowanie w filtrowaniu obrazów [22], dopasowywaniu wzorców [18], wykrywaniu krawędzi [11] czy też znakowaniu wodnym [26].

Celem tej pracy jest uogólnienie teorii oszczędnego próbkowania, znanej dla sygnałów i pomiarów o współrzędnych rzeczywistych lub zespolonych, na przypadek kwaternionowy. Na początek zaprezentujemy podstawowe definicje i twierdzenia, które stanowią podstawę klasycznej teorii oszczędnego próbkowania. Wprowadzimy następnie algebrę kwaternionów i opiszemy jej najistotniejsze własności, ważne z punktu widzenia dalszych rozdziałów, by później móc skupić się na najważniejszym elemencie pracy, tzn. udowodnieniu twierdzeń gwarantujących wierną rekonstrukcję pewnych klas sygnałów kwaternionowych z niewielkiej ilości pomiarów kwaternionowych. Warto dodać, że jest to pierwszy tego typu wynik rozszerzający teorię oszczędnego próbkowania na przypadek kwaternionowy. Zajmiemy się również analizą pewnych macierzy pomiarowych, które z dużym prawdopodobieństwem spełniają założenia dowodzonych twierdzeń, a pracę podsumujemy kilkoma eksperymentami numerycznymi, które zilustrują omawianą teorię i wskażą perspektywy zastosowania jej w praktyce.

1.2. Oszczędne próbkowanie – przypadek rzeczywisty

W pierwszej kolejności zajmiemy się przypadkiem wektorów o współrzędnych rzeczywistych, tzn. wektorami $\mathbf{x} \in \mathbb{R}^n$. Jednym z najważniejszych pojęć teorii oszczędnego próbkowania jest pojęcie *nośnika wektora* (tzn. zbioru indeksów jego niezerowych współrzędnych) i *wektora s-rzadkiego* (tzn. takiego, który ma co najwyżej s niezerowych współrzędnych). Formalne definicje tych pojęć w naturalny sposób uogólniają się na przypadek wektorów o współrzędnych zespolonych i kwaternionowych i zostaną podane w Rozdziale 3.

Rozważmy problem rekonstrukcji sygnału rzadkiego $\mathbf{x} \in \mathbb{R}^n$ z jego liniowych pomiarów

$$\mathbf{y} = \Phi \mathbf{x} \in \mathbb{R}^n,$$

gdzie $\Phi \in \mathbb{R}^{m \times n}$. Interesuje nas w szczególności pytanie, jakie własności powinna mieć macierz Φ , aby istniała możliwość dokładnego zrekonstruowania sygnału, a także jaka jest efek-

tywna metoda rekonstrukcji. Okazuje się [8], że przy pewnych założeniach dotyczących macierzy Φ (o których powiemy w dalszej części rozdziału) dowolny s -rzadki sygnał $\mathbf{x} \in \mathbb{R}^n$ można jednoznacznie odtworzyć z pomiarów $\mathbf{y} = \Phi \mathbf{x} \in \mathbb{R}^m$, $m < n$, rozwiązując zagadnienie minimalizacji ℓ_1 , tzn.

$$\arg \min_{\mathbf{z} \in \mathbb{R}^n} \|\mathbf{z}\|_1 \quad \text{pod warunkiem} \quad \Phi \mathbf{z} = \mathbf{y}, \quad (P_{\mathbb{R}-\mathbb{R},0})$$

gdzie $\|\mathbf{z}\|_1 = \sum_{i=1}^n |z_i|$ dla $\mathbf{z} = (z_1, \dots, z_n)^T$ oraz

$$\arg \min_x f(x) = \{x : \forall_y \quad f(y) \geq f(x)\},$$

tzn. $\arg \min_x f(x)$ oznacza zbiór tych argumentów x , dla których funkcja f osiąga swoje minimum.

W praktycznych zastosowaniach problem $(P_{\mathbb{R}-\mathbb{R},0})$ wymaga pewnej modyfikacji. Dane pomiarowe, jakimi dysponujemy zazwyczaj nie są dokładne, ale obarczone pewnym błędem $\mathbf{e} = \Phi \mathbf{x} - \mathbf{y}$. Zakładamy jednak, że jesteśmy w stanie oszacować normę ℓ_2 tego błędu, tzn. $\|\mathbf{e}\|_2 \leq \eta$ dla pewnego ustalonego $\eta \geq 0$, gdzie $\|\mathbf{e}\|_2 = \left(\sum_{i=1}^m |e_i|^2\right)^{1/2}$, $\mathbf{e} = (e_1, \dots, e_m)^T$. Problem minimalizacji sprowadza się wówczas do rozwiązania zagadnienia

$$\arg \min_{\mathbf{z} \in \mathbb{R}^n} \|\mathbf{z}\|_1 \quad \text{pod warunkiem} \quad \|\Phi \mathbf{z} - \mathbf{y}\|_2 \leq \eta. \quad (P_{\mathbb{R}-\mathbb{R},\eta})$$

Oczywiście w tym przypadku nie mamy możliwości dokładnego zrekonstruowania wektora \mathbf{x} , ale okazuje się, że możemy przybliżyć wartości jego współrzędnych z błędem, którego normę można oszacować w sposób zależny od η . Taki proces nazywamy *stabilną rekonstrukcją wektora z danych pomiarowych*.

Jednym z warunków, który pozwala na stabilną rekonstrukcję wektorów z danych obarczonych błędem pomiarowym jest tzw. *własność ograniczonej izometrii*. Jest to pojęcie wprowadzone przez Emmanuela Candès'a i Terence'a Tao w jednej z pierwszych prac poświęconych teorii oszczędnego próbkowania [8]. Od tamtego czasu jest nieustannie rozwijane – wciąż prowadzone są prace, które prowadzą do uzyskania lepszych oszacowań na stałe, które występują w głównym twierdzeniu tej teorii [6], a także do wykazania, że pewne macierze pomiarowe spełniają ten warunek.

Definicja 1.1. Niech $s \in \mathbb{N}$, $s \geq 1$, oraz $\Phi \in \mathbb{R}^{m \times n}$. Mówimy, że macierz Φ ma *własność s -ograniczonej izometrii* (ang. *s -restricted isometry property* - RIP) ze stałą $\delta_s \geq 0$, jeśli

$$(1 - \delta_s) \|\mathbf{x}\|_2^2 \leq \|\Phi \mathbf{x}\|_2^2 \leq (1 + \delta_s) \|\mathbf{x}\|_2^2 \quad (1.2)$$

dla wszystkich s -rzadkich wektorów $\mathbf{x} \in \mathbb{R}^n$. Najmniejszą liczbę $\delta_s \geq 0$ spełniającą powyższy warunek nazywamy *stałą s -ograniczonej izometrii* (ang. *s -restricted isometry constant*).

Powyższą definicję można uogólnić na wektory o współrzędnych zespolonych i kwaternionowych (Definicja 3.2). Przyjęło się również potocznie mówić, że macierz Φ ma *własność ograniczonej izometrii*, jeśli δ_s jest małe dla odpowiednio dużych s . Okazuje się, że spełnienie warunku ograniczonej izometrii z odpowiednio małą stałą gwarantuje stabilną rekonstrukcję wektorów (niekoniecznie rzadkich) z zaszumionych pomiarów. Mówi o tym poniższe twierdzenie, przytoczone za Candèssem [6], które było obiektem rozważań w pracy licencjackiej Autora.

Twierdzenie 1.1. Niech macierz $\Phi \in \mathbb{R}^{m \times n}$ spełnia *własność $2s$ -ograniczonej izometrii* ze stałą $\delta_{2s} < \sqrt{2} - 1$ i niech $\eta \geq 0$. Wówczas dla dowolnego wektora $\mathbf{x} \in \mathbb{R}^n$ oraz wektora $\mathbf{y} = \Phi \mathbf{x} + \mathbf{e}$ takiego, że $\|\mathbf{e}\|_2 \leq \eta$, rozwiązanie $\mathbf{x}^\#$ problemu

$$\arg \min_{\mathbf{z} \in \mathbb{R}^n} \|\mathbf{z}\|_1 \quad \text{pod warunkiem} \quad \|\Phi \mathbf{z} - \mathbf{y}\|_2 \leq \eta \quad (P_{\mathbb{R}-\mathbb{R}, \eta})$$

spełnia nierówność

$$\|\mathbf{x}^\# - \mathbf{x}\|_2 \leq \frac{C_0}{\sqrt{s}} \|\mathbf{x} - \mathbf{x}_{|s}\|_1 + C_1 \eta \quad (1.3)$$

ze stałymi

$$C_0 = 2 \cdot \frac{1 + (\sqrt{2} - 1) \delta_{2s}}{1 - (\sqrt{2} + 1) \delta_{2s}}, \quad C_1 = \frac{4\sqrt{1 + \delta_{2s}}}{1 - (1 + \sqrt{2}) \delta_{2s}},$$

gdzie $\mathbf{x}_{|s}$ oznacza najlepsze s -rzadkie przybliżenie wektora \mathbf{x} , tzn. wektor powstały z wektora \mathbf{x} poprzez pozostawienie jego s największych co do modułu współrzędnych oraz wyzerowanie pozostałych.

Jak widać z powyższego twierdzenia, jeśli $\delta_{2s} < \sqrt{2} - 1$, wektor $\mathbf{x} \in \mathbb{R}^n$ jest s -rzadki, a pomiary dokładne (tzn. $\eta = 0$), to rekonstrukcja w wyniku minimalizacji ℓ_1 jest jednoznaczna i w jej wyniku odzyskujemy wektor \mathbf{x} . Pozostaje jednak pytanie, jakie macierze pomiarowe Φ

spełniają ten warunek, a także czy założenie $\delta_{2s} < \sqrt{2} - 1$ jest optymalne, tzn. jak duża musi być stała δ_{2s} , by nie była możliwa stabilna rekonstrukcja wszystkich wektorów \mathbf{x} . W dalszej części pracy sformułujemy i udowodnimy analogiczny wynik dla przypadku ogólniejszego, tzn. dla macierzy pomiarowych i wektorów o współrzędnych w algebrze kwaternionów.

Rozdział 2

Algebra kwaternionów

Zanim przejdziemy do uogólnienia przytoczonych w poprzednim rozdziale wyników związanych z oszczędnym próbkowaniem na algebrę kwaternionów, musimy ją zdefiniować oraz pokazać pewne jej własności. Jest ona rozszerzeniem ciała liczb zespolonych, a za datę jej powstania uważa się rok 1843, kiedy Sir William R. Hamilton, uderzony swoim odkryciem, wyrył w kamieniu na moście Broome w Dublinie słynny napis

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1,$$

definiujący reguły mnożenia kwaternionów [25].

2.1. Podstawowe pojęcia i własności

Niech $\mathcal{H} = \{1, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$ oznacza bazę w 4-wymiarowej przestrzeni wektorowej \mathbb{H} nad ciałem \mathbb{R} . Zdefiniujemy mnożenie w \mathbb{H} według równości

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1, \tag{2.1}$$

$$\mathbf{ij} = -\mathbf{ji} = \mathbf{k}, \quad \mathbf{jk} = -\mathbf{kj} = \mathbf{i}, \quad \mathbf{ki} = -\mathbf{ik} = \mathbf{j}. \tag{2.2}$$

Można łatwo zauważyć, że równości (2.2) wynikają wprost z (2.1). Reguły mnożenia wygodnie jest zapisać również w postaci tabeli (patrz Tabela 2.1). Wprost z definicji wynika, że mnożenie elementów z algebry \mathbb{H} jest rozdzielne względem dodawania, a mnożenie przez

skalary (tzn. przez liczby rzeczywiste) jest przemienne, tzn.

$$\forall_{x,y,z \in \mathbb{H}} \quad x(y+z) = xy + xz, \quad (y+z)x = yx + zx, \quad (2.3)$$

$$\forall_{x,y \in \mathbb{H}, \lambda \in \mathbb{R}} \quad x(\lambda y) = (\lambda x)y = \lambda(xy). \quad (2.4)$$

\cdot	1	\mathbf{i}	\mathbf{j}	\mathbf{k}
1	1	\mathbf{i}	\mathbf{j}	\mathbf{k}
\mathbf{i}	\mathbf{i}	-1	\mathbf{k}	$-\mathbf{j}$
\mathbf{j}	\mathbf{j}	$-\mathbf{k}$	-1	\mathbf{i}
\mathbf{k}	\mathbf{k}	\mathbf{j}	$-\mathbf{i}$	-1

Tabela 2.1: Reguły mnożenia w algebrze kwaternionów.

Definicja 2.1. Elementy przestrzeni \mathbb{H} wraz z operacjami algebraicznymi \mathbb{H} jako przestrzeni wektorowej nad ciałem \mathbb{R} oraz mnożeniem zdefiniowanym równościami (2.1) nazywamy (rzeczywistymi) *kwaternionami*. Elementy $\mathbf{i}, \mathbf{j}, \mathbf{k}$ będziemy nazywać kolejnymi *jednostkami urojonymi*.

Symbolu \mathbb{H} używamy ze względów historycznych, dla upamiętnienia W. R. Hamiltona, twórcy kwaternionów. Kwaterniony $q \in \mathbb{H}$ będziemy zapisywać w postaci

$$q = r_0 + r_1\mathbf{i} + r_2\mathbf{j} + r_3\mathbf{k}, \quad r_0, \dots, r_3 \in \mathbb{R}. \quad (2.5)$$

Kwaterniony, dla których przynajmniej jedna z liczb r_0, \dots, r_3 jest równa 0 będziemy zapisywać pomijając odpowiednią jednostkę urojoną (lub liczbę rzeczywistą). Wektor zerowy, tzn. taki że $r_0 = \dots = r_3 = 0$ będziemy zapisywać jako 0.

Kwaterniony stanowią uogólnienie liczb zespolonych, zatem w sposób naturalny rozszerzają się na nie również pojęcia części rzeczywistej i urojonej, a także sprzężenie.

Definicja 2.2. Niech $q \in \mathbb{H}$ będzie postaci (2.5). Liczbę $r_0 \in \mathbb{R}$ nazywamy *częścią rzeczywistą* (skalarną) kwaternionu, natomiast $r_1\mathbf{i} + r_2\mathbf{j} + r_3\mathbf{k}$ nazywamy *częścią urojoną* (wektorową) kwaternionu. Wprowadzamy oznaczenie

$$\operatorname{Re}(q) := r_0, \quad \operatorname{Im}(q) := r_1\mathbf{i} + r_2\mathbf{j} + r_3\mathbf{k}.$$

Kwaterniony $q \in \mathbb{H}$ takie, że $\operatorname{Re}(q) = 0$ nazywamy *czystymi kwaternionami*.

Zazwyczaj, gdy nie będzie budziło to wątpliwości, będziemy stosować zapis wektorowy $\mathbf{q} := \text{Im}(q)$. Wówczas $q = r_0 + \mathbf{q}$. Zauważmy, że \mathbf{q} możemy utożsamiać z wektorem w przestrzeni \mathbb{R}^3 .

Definicja 2.3. Niech $q = r_0 + r_1\mathbf{i} + r_2\mathbf{j} + r_3\mathbf{k} \in \mathbb{H}$. *Sprzężeniem kwaternionowym* nazywamy funkcję

$$\bar{\cdot} : \mathbb{H} \rightarrow \mathbb{H}; \quad \bar{q} := \text{Re}(q) - \text{Im}(q) = r_0 - r_1\mathbf{i} - r_2\mathbf{j} - r_3\mathbf{k}.$$

Z bezpośredniego rachunku wynikają własności zebrane w poniższym fakcie.

Fakt 2.1. Niech $q \in \mathbb{H}$. Zachodzą poniższe wzory:

$$\text{Re}(q) = \frac{q + \bar{q}}{2}, \quad \text{Im}(q) = \frac{q - \bar{q}}{2}, \quad (2.6)$$

$$\bar{q} = -\frac{1}{2}(q + \mathbf{i}qi + \mathbf{j}qj + \mathbf{k}qk). \quad (2.7)$$

Ponieważ utożsamiamy algebrę \mathbb{H} z przestrzenią \mathbb{R}^4 , to naturalnym jest wprowadzenie definicji normy kwaternionu tak jak wprowadza się pojęcie normy dla dowolnych rzeczywistych przestrzeni n -wymiarowych. W ten sam sposób wprowadza się normę liczb zespolonych (utożsamianych z przestrzenią \mathbb{R}^2). Z tego utożsamienia wynika natychmiast prawdziwość Faktu 2.2, a przeprowadzając bezpośredni rachunek otrzymujemy dowód Faktu 2.3.

Definicja 2.4. Niech $q \in \mathbb{H}$ będzie postaci (2.5). *Normą kwaternionową* nazywamy funkcję

$$|\cdot| : \mathbb{H} \rightarrow \mathbb{R}_+ \cup \{0\}; \quad |q| := \sqrt{r_0^2 + r_1^2 + r_2^2 + r_3^2}.$$

Kwaterniony $q \in \mathbb{H}$ takie, że $|q| = 1$ nazywamy *kwaternionami jednostkowymi*.

Fakt 2.2. Niech $q, q_1, q_2 \in \mathbb{H}$, $\lambda \in \mathbb{R}$. Wówczas

$$(a) \quad |q| = 0 \iff q = 0,$$

$$(b) \quad |\lambda \cdot q| = |\lambda| \cdot |q|,$$

$$(c) \quad |q_1 + q_2| \leq |q_1| + |q_2|.$$

Fakt 2.3. Niech $q \in \mathbb{H}$. Wówczas

$$(a) \quad |q|^2 = q \cdot \bar{q} = \bar{q} \cdot q,$$

$$(b) \quad |q| = |\bar{q}| = |-q|.$$

Twierdzenie 2.1. Niech $q \in \mathbb{H}$ będzie taki, że $\operatorname{Re}(q) = 0$ (tzn. $q = \mathbf{q}$). Wówczas

$$q^2 = -|q|^2.$$

Dowód. Niech $q \in \mathbb{H}$, $\operatorname{Re}(q) = 0$. Zauważmy, że wówczas $\bar{q} = -q$. Stąd

$$q^2 = q \cdot q = -\bar{q} \cdot q = -|q|^2,$$

co kończy dowód. □

Z Twierdzenia 2.1 wynika, że równanie

$$q^2 = -1 \tag{2.8}$$

ma nieskończenie wiele rozwiązań. Jest to znaczna różnica w stosunku do ciała liczb zespolonych, w którym każdy wielomian n -tego stopnia ma dokładnie n pierwiastków (licząc z krotnościami). W przypadku algebry kwaternionów, rozwiązaniem równania (2.8) jest cała sfera jednostkowa w \mathbb{R}^3 (tzn. wszystkie kwaterniony jednostkowe o zerowej części rzeczywistej).

Działania w ciele liczb zespolonych mają wiele własności znanych nam z ciała liczb rzeczywistych. Kwaterniony natomiast tworzą *nieprzemianną* algebrę 4-ego rzędu nad ciałem liczb rzeczywistych. W ogólności $q_1 \cdot q_2 \neq q_2 \cdot q_1$. Jest to jednak algebra *łączna*, tzn. dla wszystkich $q_1, q_2, q_3 \in \mathbb{H}$ zachodzi $q_1 \cdot (q_2 \cdot q_3) = (q_1 \cdot q_2) \cdot q_3$. Zachodzi również poniższy fakt, którego dowód wynika z bezpośredniego rachunku.

Fakt 2.4. Niech $q_1, q_2 \in \mathbb{H}$. Wówczas

$$(a) \quad \overline{q_1 \cdot q_2} = \overline{q_2} \cdot \overline{q_1},$$

$$(b) \quad |q_1 \cdot q_2| = |q_1| \cdot |q_2|.$$

Tak jak w przypadku ciała liczb rzeczywistych i zespolonych, w algebrze kwaternionów mamy również zdefiniowane pojęcie elementu odwrotnego, o czym mówi poniższy fakt.

Fakt 2.5. Dla dowolnego niezerowego kwaternionu $q \in \mathbb{H}$ istnieje dokładnie jeden element odwrotny $q^{-1} \in \mathbb{H}$, tzn. taki że

$$q \cdot q^{-1} = q^{-1} \cdot q = 1.$$

Wówczas $q^{-1} = \frac{\bar{q}}{|q|^2}$ i dla $q_1, q_2 \in \mathbb{H}$, $q_1, q_2 \neq 0$, zachodzi wzór

$$(q_1 \cdot q_2)^{-1} = q_2^{-1} \cdot q_1^{-1}. \tag{2.9}$$

Widzimy zatem, że możemy jednoznacznie zdefiniować dzielenie kwaternionów. Nieprzemienność mnożenia sprawia jednak, że należy zdefiniować zarówno dzielenie *prawe* jak i *lewe*, tzn. dla $q_1, q_2 \in \mathbb{H}$, $q_2 \neq 0$, definiujemy

$$q_1/q_2 := q_1 q_2^{-1} = \frac{q_1 \cdot \overline{q_2}}{|q_2|^2}, \quad q_2 \backslash q_1 := q_2^{-1} q_1 = \frac{\overline{q_2} \cdot q_1}{|q_2|^2}.$$

Kwaterniony są więc algebrą z *dzieleniem* (ang. *division algebra*). Podsumowując, algebra kwaternionów spełnia wszystkie aksjomaty ciała, poza przemiennością mnożenia. Czasami nazywa się ją w związku z tym *ciałem nieprzemiennym*.

2.2. Kwanternionowe przestrzenie wektorowe

Podobnie jak w przypadku rzeczywistym i zespolonym, rozważamy n -wymiarowe moduły prawostronne \mathbb{H}^n nad pierścieniem \mathbb{H} , tzn. algebry, w których $(\mathbb{H}^n, +)$ tworzy grubą abelową, a mnożenie $\cdot : \mathbb{H}^n \times \mathbb{H} \rightarrow \mathbb{H}^n$ ma następujące własności:

$$\forall_{\mathbf{x}, \mathbf{y} \in \mathbb{H}^n, q \in \mathbb{H}} \quad (\mathbf{x} + \mathbf{y}) \cdot q = \mathbf{x} \cdot q + \mathbf{y} \cdot q,$$

$$\forall_{\mathbf{x} \in \mathbb{H}^n, q, r \in \mathbb{H}} \quad \mathbf{x} \cdot (q + r) = \mathbf{x} \cdot q + \mathbf{x} \cdot r,$$

$$\forall_{\mathbf{x} \in \mathbb{H}^n, q, r \in \mathbb{H}} \quad \mathbf{x} \cdot (q \cdot r) = (\mathbf{x} \cdot q) \cdot r,$$

$$\forall_{\mathbf{x} \in \mathbb{H}^n} \quad \mathbf{x} \cdot 1 = \mathbf{x}.$$

Ponieważ pierścień \mathbb{H} nie jest ciałem, moduł prawostronny \mathbb{H}^n nad \mathbb{H} nie jest przestrzenią wektorową. Jednak ze względu na to, że pierścień \mathbb{H} czasem nazywa się ciałem nieprzemiennym, przyjęło się również nazywać moduł prawostronny \mathbb{H}^n nad pierścieniem \mathbb{H} *prawostronną przestrzenią wektorową*. W analogiczny sposób definiuje się moduł lewostronny (definiując mnożenie $\cdot : \mathbb{H} \times \mathbb{H}^n \rightarrow \mathbb{H}^n$) i potocznie nazywa go lewostronną przestrzenią wektorową.

Niech \mathbf{x} będzie wektorem w n -wymiarowej przestrzeni \mathbb{H}^n , $n \in \mathbb{N}$. Będziemy interpretować wektory jako jednokolumnowe macierze $\mathbf{x} \in \mathbb{H}^{n \times 1}$, tzn.

$$\mathbf{x} = (x_1, x_2, \dots, x_n)^T, \quad x_1, x_2, \dots, x_n \in \mathbb{H},$$

gdzie T oznacza transpozycję macierzy. Każda kwaternionowa macierz reprezentuje *przekształcenie liniowe* (na przestrzeni prawostronnej nad \mathbb{H})

$$\Phi : \mathbb{H}^n \rightarrow \mathbb{H}^m,$$

które ma następujące własności:

$$\forall_{\mathbf{x}, \mathbf{y} \in \mathbb{H}^n, q \in \mathbb{H}} \quad \Phi(\mathbf{x} + \mathbf{y}) = \Phi\mathbf{x} + \Phi\mathbf{y}, \quad \Phi(\mathbf{x}q) = (\Phi\mathbf{x})q.$$

Stosując notację analogiczną jak dla liczb zespolonych, wprowadzamy również pojęcie *hermitowskiego sprzężenia* macierzy kwaternionowej $\Phi \in \mathbb{H}^{m \times n}$ jako $\Phi^* = \overline{\Phi}^T$. Oczywiście $\Phi^{**} = \Phi$. Ponadto, z bezpośredniego rachunku otrzymujemy również własności sprzężenia hermitowskiego, ujęte w postaci poniższego faktu.

Fakt 2.6. *Niech $\Phi \in \mathbb{H}^{m \times n}$, $\Psi \in \mathbb{H}^{n \times k}$, $\mathbf{x} \in \mathbb{H}^n$ i $q \in \mathbb{H}$. Wówczas*

$$(\Phi q)^* = \bar{q}\Phi^*, \quad (q\Phi)^* = \Phi^*\bar{q}, \quad (\Phi\mathbf{x})^* = \mathbf{x}^*\Phi^*, \quad (\Phi\Psi)^* = \Psi^*\Phi^*.$$

Ponieważ kwaterniony są uogólnieniem liczb zespolonych, naturalnym podejściem jest wprowadzenie funkcji mającej własności iloczynu skalarnego przestrzeni \mathbb{H}^n , tak jak miało to miejsce w \mathbb{R}^n czy \mathbb{C}^n .

Wprowadźmy funkcję

$$\langle \cdot, \cdot \rangle : \mathbb{H}^n \times \mathbb{H}^n \rightarrow \mathbb{H}; \quad \langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^* \mathbf{x} = \sum_{k=1}^n \bar{y}_k x_k, \quad (2.10)$$

gdzie $\mathbf{x} = (x_1, \dots, x_n)^T, \mathbf{y} = (y_1, \dots, y_n)^T \in \mathbb{H}^n$. Stąd również wyprowadzamy wzór na *normę* ℓ_2 wektora kwaternionowego

$$\|\mathbf{x}\|_2 = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{\sum_{k=1}^n \bar{x}_k x_k} = \sqrt{\sum_{k=1}^n |x_k|^2} \quad (2.11)$$

dla dowolnego $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{H}^n$.

Uwaga 2.1. Funkcję, która (jak się okaże) spełnia aksjomaty iloczynu skalarnego, można wprowadzić również dla lewej przestrzeni wektorowej \mathbb{H}^n , definiując

$$(\cdot, \cdot) : \mathbb{H}^n \times \mathbb{H}^n \rightarrow \mathbb{H}; \quad (\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \bar{\mathbf{y}} = \sum_{k=1}^n x_k \bar{y}_k,$$

jednak wiąże się to z trudnościami przy definiowaniu operatora sprzężonego. Dlatego w dalszej części pracy zajmować się będziemy głównie iloczynem skalarnym zdefiniowanym wzorem (2.10).

Twierdzenie 2.2. Funkcja $\langle \cdot, \cdot \rangle$ spełnia aksjomaty iloczynu skalarnego.

Dowód. Niech $\mathbf{x} = (x_1, \dots, x_n)^T, \mathbf{y} = (y_1, \dots, y_n)^T, \mathbf{z} = (z_1, \dots, z_n)^T \in \mathbb{H}^n$ oraz $q \in \mathbb{H}$.

Wówczas

- $\overline{\langle \mathbf{x}, \mathbf{y} \rangle} = \overline{\sum_{k=1}^n \overline{y_k} x_k} = \sum_{k=1}^n \overline{\overline{y_k} x_k} = \sum_{k=1}^n \overline{x_k} y_k = \langle \mathbf{y}, \mathbf{x} \rangle,$
- $\langle \mathbf{x}q, \mathbf{y} \rangle = \sum_{k=1}^n \overline{y_k} x_k q = \left(\sum_{k=1}^n \overline{y_k} x_k \right) q = \langle \mathbf{x}, \mathbf{y} \rangle q,$
- $\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \sum_{k=1}^n \overline{z_k} (x_k + y_k) = \sum_{k=1}^n \overline{z_k} x_k + \sum_{k=1}^n \overline{z_k} y_k = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle,$
- $\langle \mathbf{x}, \mathbf{x} \rangle = \sum_{k=1}^n \overline{x_k} x_k = \sum_{k=1}^n |x_k|^2 = \|\mathbf{x}\|_2^2 \geq 0,$
- $\langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|_2^2 = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}.$

Kończy to dowód. □

Wniosek 2.1. Funkcja $\|\cdot\|_2$ spełnia aksjomaty normy.

Warto zauważyć, że dla $\mathbf{x}, \mathbf{y} \in \mathbb{H}^n$ i $q \in \mathbb{H}$ zachodzi również własność $\langle \mathbf{x}, \mathbf{y}q \rangle = \overline{q} \langle \mathbf{x}, \mathbf{y} \rangle$. Należy zwrócić uwagę, że funkcja $\langle \cdot, \cdot \rangle$ nie jest iloczynem skalarnym w klasycznym sensie ze względu na fakt, że \mathbb{H} nie jest ciałem. Spełnia jednak wszystkie spodziewane własności, ponadto prawdziwe są odpowiedniki nierówności Cauchy'ego-Schwarza i równości polaryzacyjnej, znane z analizy rzeczywistych i zespolonych przestrzeni wektorowych, o czym mówią dwa poniższe twierdzenia.

Twierdzenie 2.3 (Nierówność Cauchy'ego-Schwarza). Dla dowolnych $\mathbf{x}, \mathbf{y} \in \mathbb{H}^n$ zachodzi

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\|_2 \cdot \|\mathbf{y}\|_2. \quad (2.12)$$

Dowód. Dowód przebiega tak samo jak dowód analogicznego twierdzenia dla przestrzeni rzeczywistych i zespolonych. Niech $\mathbf{x}, \mathbf{y} \in \mathbb{H}^n$ oraz $q \in \mathbb{H}$. Jeśli $\mathbf{y} = \mathbf{0}$, to nierówność jest spełniona, zatem założmy, że $\|\mathbf{y}\|_2 > 0$. Z Twierdzenia 2.2 otrzymujemy, że

$$\begin{aligned} 0 &\leq \langle \mathbf{x} - \mathbf{y}q, \mathbf{x} - \mathbf{y}q \rangle = \langle \mathbf{x}, \mathbf{x} \rangle - \langle \mathbf{y}, \mathbf{x} \rangle q - \overline{q} \langle \mathbf{x}, \mathbf{y} \rangle - \overline{q} \langle \mathbf{y}, \mathbf{y} \rangle q \\ &= \|\mathbf{x}\|_2^2 - \overline{\langle \mathbf{x}, \mathbf{y} \rangle} q - \overline{q} \langle \mathbf{x}, \mathbf{y} \rangle + |q|^2 \|\mathbf{y}\|_2^2. \end{aligned}$$

Kładąc $q = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{y}\|_2^2}$ otrzymujemy

$$0 \leq \|\mathbf{x}\|_2^2 - \frac{\overline{\langle \mathbf{x}, \mathbf{y} \rangle} \langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{y}\|_2^2} - \frac{\overline{\langle \mathbf{x}, \mathbf{y} \rangle} \langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{y}\|_2^2} + \frac{|\langle \mathbf{x}, \mathbf{y} \rangle|^2}{\|\mathbf{y}\|_2^2} = \|\mathbf{x}\|_2^2 - \frac{|\langle \mathbf{x}, \mathbf{y} \rangle|^2}{\|\mathbf{y}\|_2^2},$$

co kończy dowód. □

Twierdzenie 2.4 (Równość polaryzacyjna I). *Dla dowolnych $\mathbf{x}, \mathbf{y} \in \mathbb{H}^n$ zachodzi wzór*

$$\begin{aligned} \langle \mathbf{x}, \mathbf{y} \rangle &= \frac{1}{4} \left(\|\mathbf{x} + \mathbf{y}\|_2^2 - \|\mathbf{x} - \mathbf{y}\|_2^2 \right) + \frac{\mathbf{i}}{4} \left(\|\mathbf{x} + \mathbf{y}\mathbf{i}\|_2^2 - \|\mathbf{x} - \mathbf{y}\mathbf{i}\|_2^2 \right) \\ &\quad + \frac{\mathbf{j}}{4} \left(\|\mathbf{x} + \mathbf{y}\mathbf{j}\|_2^2 - \|\mathbf{x} - \mathbf{y}\mathbf{j}\|_2^2 \right) + \frac{\mathbf{k}}{4} \left(\|\mathbf{x} + \mathbf{y}\mathbf{k}\|_2^2 - \|\mathbf{x} - \mathbf{y}\mathbf{k}\|_2^2 \right). \end{aligned} \quad (2.13)$$

Dowód. Niech $\mathbf{x} = (x_1, \dots, x_n)^T$, $\mathbf{y} = (y_1, \dots, y_n)^T \in \mathbb{H}^n$. Rozważmy część rzeczywistą prawej strony wzoru. Otrzymujemy

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2 &= \sum_{k=1}^n \left((\overline{x_k} + \overline{y_k})(x_k + y_k) - (\overline{x_k} - \overline{y_k})(x_k - y_k) \right) \\ &= \sum_{k=1}^n \left(\overline{x_k}x_k + \overline{y_k}x_k + \overline{x_k}y_k + \overline{y_k}y_k - \overline{x_k}x_k + \overline{y_k}x_k + \overline{x_k}y_k - \overline{y_k}y_k \right) \\ &= 2 \sum_{k=1}^n \left(\overline{y_k}x_k + \overline{x_k}y_k \right) \\ &= 2 \left(\langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle \right) \\ &= 2 \left(\langle \mathbf{x}, \mathbf{y} \rangle + \overline{\langle \mathbf{x}, \mathbf{y} \rangle} \right) = 4 \operatorname{Re}(\langle \mathbf{x}, \mathbf{y} \rangle). \end{aligned}$$

Zbadajmy współczynnik przy jednostce urojonej \mathbf{i} po prawej stronie wzoru (2.13). Mamy

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\mathbf{i}\|^2 - \|\mathbf{x} - \mathbf{y}\mathbf{i}\|^2 &= \sum_{k=1}^n \left((\overline{x_k} + \overline{y_k}\mathbf{i})(x_k + y_k\mathbf{i}) - (\overline{x_k} - \overline{y_k}\mathbf{i})(x_k - y_k\mathbf{i}) \right) \\ &= \sum_{k=1}^n \left(\overline{x_k}x_k + \overline{x_k}y_k\mathbf{i} + \overline{y_k}\mathbf{i}x_k + \overline{y_k}\mathbf{i}y_k\mathbf{i} - \overline{x_k}x_k + \overline{x_k}y_k\mathbf{i} + \overline{y_k}\mathbf{i}x_k - \overline{y_k}\mathbf{i}y_k\mathbf{i} \right) \\ &= 2 \sum_{k=1}^n \left(\overline{x_k}y_k\mathbf{i} - \mathbf{i}\overline{y_k}x_k \right) \\ &= -2 \left(\langle \mathbf{x}, \mathbf{y} \rangle \mathbf{i} - \mathbf{i} \langle \mathbf{y}, \mathbf{x} \rangle \right) = -2 \left(\langle \mathbf{x}, \mathbf{y} \rangle \mathbf{i} - \mathbf{i} \overline{\langle \mathbf{x}, \mathbf{y} \rangle} \right). \end{aligned}$$

Analogicznie, dla pozostałych jednostek urojonych dostajemy

$$\|\mathbf{x} + \mathbf{y}\mathbf{j}\|_2^2 - \|\mathbf{x} - \mathbf{y}\mathbf{j}\|_2^2 = -2 \left(\langle \mathbf{x}, \mathbf{y} \rangle \mathbf{j} - \mathbf{j} \overline{\langle \mathbf{x}, \mathbf{y} \rangle} \right)$$

oraz

$$\|\mathbf{x} + \mathbf{y}\mathbf{k}\|_2^2 - \|\mathbf{x} - \mathbf{y}\mathbf{k}\|_2^2 = -2 \left(\langle \mathbf{x}, \mathbf{y} \rangle \mathbf{k} - \mathbf{k} \overline{\langle \mathbf{x}, \mathbf{y} \rangle} \right).$$

Mnożąc lewe strony uzyskanych równości odpowiednio przez $\mathbf{i}, \mathbf{j}, \mathbf{k}$ oraz dodając je stronami

otrzymujemy, że

$$\begin{aligned}
& \mathbf{i} \left(\|\mathbf{x} + \mathbf{y}\mathbf{i}\|_2^2 - \|\mathbf{x} - \mathbf{y}\mathbf{i}\|_2^2 \right) + \mathbf{j} \left(\|\mathbf{x} + \mathbf{y}\mathbf{j}\|_2^2 - \|\mathbf{x} - \mathbf{y}\mathbf{j}\|_2^2 \right) + \mathbf{k} \left(\|\mathbf{x} + \mathbf{y}\mathbf{k}\|_2^2 - \|\mathbf{x} - \mathbf{y}\mathbf{k}\|_2^2 \right) \\
&= -2 \left(\mathbf{i} \langle \mathbf{x}, \mathbf{y} \rangle \mathbf{i} + \mathbf{j} \langle \mathbf{x}, \mathbf{y} \rangle \mathbf{j} + \mathbf{k} \langle \mathbf{x}, \mathbf{y} \rangle \mathbf{k} \right) - 6 \overline{\langle \mathbf{x}, \mathbf{y} \rangle} \\
&= -2 \left(-2 \overline{\langle \mathbf{x}, \mathbf{y} \rangle} - \langle \mathbf{x}, \mathbf{y} \rangle \right) - 6 \overline{\langle \mathbf{x}, \mathbf{y} \rangle} \\
&= 2 \langle \mathbf{x}, \mathbf{y} \rangle - 2 \overline{\langle \mathbf{x}, \mathbf{y} \rangle} = 4 \operatorname{Im}(\langle \mathbf{x}, \mathbf{y} \rangle),
\end{aligned}$$

gdzie skorzystaliśmy z Faktu 2.1. Kończy to dowód. \square

Dla funkcji zdefiniowanej w Uwadze 2.1 mamy analogiczną postać równości polaryzacyjnej dla lewej kwaternionowej przestrzeni wektorowej, tzn.

$$\begin{aligned}
\langle \mathbf{x}, \mathbf{y} \rangle &= \frac{1}{4} \left(\|\mathbf{x} + \mathbf{y}\|_2^2 - \|\mathbf{x} - \mathbf{y}\|_2^2 \right) + \frac{\mathbf{i}}{4} \left(\|\mathbf{x} + \mathbf{i}\mathbf{y}\|_2^2 - \|\mathbf{x} - \mathbf{i}\mathbf{y}\|_2^2 \right) \\
&\quad + \frac{\mathbf{j}}{4} \left(\|\mathbf{x} + \mathbf{j}\mathbf{y}\|_2^2 - \|\mathbf{x} - \mathbf{j}\mathbf{y}\|_2^2 \right) + \frac{\mathbf{k}}{4} \left(\|\mathbf{x} + \mathbf{k}\mathbf{y}\|_2^2 - \|\mathbf{x} - \mathbf{k}\mathbf{y}\|_2^2 \right).
\end{aligned}$$

Równość polaryzacyjną można zapisać również korzystając z innej postaci kwaternionu. Dowolny niezerowy kwaternion $q \in \mathbb{H}$ można zapisać w postaci

$$q = x + \mu y, \tag{2.14}$$

gdzie $x, y \in \mathbb{R}$, $y \geq 0$, a $\mu \in \mathbb{H}$ jest czystym jednostkowym kwaternionem, tzn. takim, że $\operatorname{Re} \mu = 0$ i $\|\mu\| = 1$. Poniższy lemat pozwoli nam na udowodnienie drugiej postaci równości polaryzacyjnej.

Lemat 2.1. *Niech $q = x + \mu y \in \mathbb{H}$, gdzie $x, y \in \mathbb{R}$, $y \geq 0$, a $\mu \in \mathbb{H}$ jest czystym jednostkowym kwaternionem. Wówczas*

$$\bar{\mu} q \mu = q, \quad \bar{\mu} \bar{q} \mu = \bar{q}. \tag{2.15}$$

Dowód. Z przemienności mnożenia kwaternionów przez liczby rzeczywiste mamy

$$\bar{\mu} q \mu = \bar{\mu} (x + \mu y) \mu = \bar{\mu} x \mu + \bar{\mu} \mu y \mu = x |\mu|^2 + |\mu|^2 y \mu = x + \mu y = q.$$

Drugą równość udowadniamy analogicznie. \square

Twierdzenie 2.5 (Równość polaryzacyjna II). *Niech $\mathbf{x}, \mathbf{y} \in \mathbb{H}^n$ i oznaczmy*

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_0 + \mu y_0,$$

gdzie $x_0, y_0 \in \mathbb{R}$, $y_0 \geq 0$ oraz μ jest czystym jednostkowym kwaternionem. Wówczas zachodzi wzór

$$\langle \mathbf{x}, \mathbf{y} \rangle = \frac{1}{4} \left(\|\mathbf{x} + \mathbf{y}\|_2^2 - \|\mathbf{x} - \mathbf{y}\|_2^2 \right) + \frac{\mu}{4} \left(\|\mathbf{x} + \mathbf{y}\mu\|_2^2 - \|\mathbf{x} - \mathbf{y}\mu\|_2^2 \right). \quad (2.16)$$

Dowód. Postać części rzeczywistej otrzymujemy w identyczny sposób jak w przypadku Twierdzenia 2.4. Aby wyznaczyć postać części urojonej zauważmy, że

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\mu\|_2^2 &= \|\mathbf{x}\|_2^2 + \overline{\langle \mathbf{x}, \mathbf{y} \rangle} \mu - \mu \langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|_2^2, \\ \|\mathbf{x} - \mathbf{y}\mu\|_2^2 &= \|\mathbf{x}\|_2^2 - \overline{\langle \mathbf{x}, \mathbf{y} \rangle} \mu + \mu \langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|_2^2, \end{aligned}$$

co wynika z faktu, że dla $\mu \in \mathbb{H}$ takich, że $\operatorname{Re} \mu = 0$ zachodzi $\bar{\mu} = -\mu$. Z Lematu 2.1 otrzymujemy ostatecznie

$$\begin{aligned} \mu \left(\|\mathbf{x} + \mathbf{y}\mu\|_2^2 - \|\mathbf{x} - \mathbf{y}\mu\|_2^2 \right) &= 2\mu \overline{\langle \mathbf{x}, \mathbf{y} \rangle} \mu - 2\mu^2 \langle \mathbf{x}, \mathbf{y} \rangle = -2\bar{\mu} \overline{\langle \mathbf{x}, \mathbf{y} \rangle} \mu + 2 \langle \mathbf{x}, \mathbf{y} \rangle \\ &= -2\overline{\langle \mathbf{x}, \mathbf{y} \rangle} + 2 \langle \mathbf{x}, \mathbf{y} \rangle = 2 \left(\langle \mathbf{x}, \mathbf{y} \rangle - \overline{\langle \mathbf{x}, \mathbf{y} \rangle} \right) \\ &= 4 \operatorname{Im} (\langle \mathbf{x}, \mathbf{y} \rangle). \end{aligned}$$

Kończy to dowód. □

Tak jak w przypadku pierwszej równości polaryzacyjnej, reprezentację daną Twierdzeniem 2.5 można sformułować dla funkcji z Uwagi 2.1. Przyjmuje ona wówczas postać

$$\langle \mathbf{x}, \mathbf{y} \rangle = \frac{1}{4} \left(\|\mathbf{x} + \mathbf{y}\|_2^2 - \|\mathbf{x} - \mathbf{y}\|_2^2 \right) + \frac{\mu}{4} \left(\|\mathbf{x} + \mu\mathbf{y}\|_2^2 - \|\mathbf{x} - \mu\mathbf{y}\|_2^2 \right).$$

Dzięki temu otrzymujemy reprezentację iloczynu skalarnego dla lewych kwaternionowych przestrzeni wektorowych. Reprezentacje dane Twierdzeniami 2.4 i 2.5 można znaleźć w literaturze [19], ale zostały również niezależnie wyprowadzone przez Autora tej pracy.

Pojęciami bezpośrednio związanymi z przekształceniami liniowymi w przestrzeniach rzeczywistych lub zespolonych są wartości własne oraz wektory własne. W przypadku macierzy kwaternionowych pojawia się problem związany z tym, że mnożenie kwaternionów nie jest przemienne. Musimy w związku z tym rozważać dwa różne problemy własne dla macierzy kwaternionowych.

Definicja 2.5. Niech $\Phi \in \mathbb{H}^{n \times n}$. Niezerowy wektor $\mathbf{x} \in \mathbb{H}^n$ nazywamy *prawym wektorem własnym* macierzy Φ , jeśli istnieje $\lambda \in \mathbb{H}$ taka, że

$$\Phi \mathbf{x} = \mathbf{x} \lambda. \quad (2.17)$$

Liczbę λ nazywamy wówczas *prawą wartością własną* macierzy Φ . Analogicznie, niezerowy wektor $\mathbf{x} \in \mathbb{H}$ jest *lewym wektorem własnym* macierzy Φ , jeśli istnieje $\lambda \in \mathbb{H}$ taka, że

$$\Phi \mathbf{x} = \lambda \mathbf{x}$$

i liczbę λ nazywamy *lewą wartością własną* macierzy Φ .

Zanim podamy pewne własności wartości własnych macierzy kwaternionowych, należy wprowadzić jeszcze pojęcie *równoważności kwaternionów*. Będziemy mówić, że dwa kwaterniony $\lambda_1, \lambda_2 \in \mathbb{H}$ są *równoważne*, jeśli istnieje niezerowy kwaternion $q \in \mathbb{H}$ taki, że

$$\lambda_1 = q^{-1} \lambda_2 q.$$

Okazuje się, że dwa wprowadzone wyżej problemy własne muszą być traktowane oddzielnie i w zupełnie inny sposób. W przeciwieństwie do lewych wartości własnych, prawe wartości własne macierzy kwaternionowych są dobrze poznane. Mówią o tym dwa poniższe twierdzenia, przytoczone za [19, 29].

Twierdzenie 2.6. Niech $\Phi \in \mathbb{H}^{n \times n}$. Wówczas

- (a) Φ ma co najmniej jedną lewą wartość własną;
- (b) Φ ma dokładnie n (razem z krotnościami i z dokładnością do równoważności kwaternionów) prawych wartości własnych i wszystkie są liczbami zespolonymi o nieujemnych częściach urojonych.

Pisząc w powyższym twierdzeniu o *liczbach zespolonych* mamy na myśli liczby postaci $r_0 + r_1 \mathbf{i}$, ale także $r_0 + r_1 \mathbf{j}$ lub $r_0 + r_1 \mathbf{k}$. Liczby te są parami równoważne, ponieważ

$$\begin{aligned} r_0 + r_1 \mathbf{j} &= q_{\mathbf{k}} \cdot (r_0 + r_1 \mathbf{i}) \cdot q_{\mathbf{k}}^{-1} \quad \text{dla} \quad q_{\mathbf{k}} = 1 + \mathbf{k}, \\ r_0 + r_1 \mathbf{k} &= q_{\mathbf{j}}^{-1} \cdot (r_0 + r_1 \mathbf{i}) \cdot q_{\mathbf{j}} \quad \text{dla} \quad q_{\mathbf{j}} = 1 + \mathbf{j}. \end{aligned}$$

Szczególnym przypadkiem są tzw. *macierze hermitowskie* tzn. macierze $\Psi \in \mathbb{H}^{n \times n}$ takie, że $\Psi = \Psi^*$. Otrzymujemy wówczas ciekawy rezultat, znany również dla macierzy zespolonych, przytoczony za [19].

Twierdzenie 2.7. *Niech $\Psi \in \mathbb{H}^{n \times n}$ będzie taka, że $\Psi = \Psi^*$. Wówczas*

- (a) Ψ ma dokładnie n (razem z krotnościami) wartości własnych $\lambda_1, \dots, \lambda_n$, które są liczbami rzeczywistymi;
- (b) prawe wektory własne \mathbf{x}_k i \mathbf{x}_ℓ odpowiadające różnym wartościom własnym λ_k i λ_ℓ są ortogonalne, tzn. $\langle \mathbf{x}_k, \mathbf{x}_\ell \rangle = 0$;
- (c) istnieje baza w \mathbb{H}^n , $\mathbf{e}_1, \dots, \mathbf{e}_n$, złożona z wektorów własnych macierzy Ψ , która jest ortonormalna, tzn. $\langle \mathbf{e}_k, \mathbf{e}_\ell \rangle = 0$ dla $k \neq \ell$ oraz $\langle \mathbf{e}_k, \mathbf{e}_k \rangle = 1$.

Zauważmy, że w przypadku macierzy hermitowskich nie ma potrzeby rozróżniania prawych i lewych wartości własnych. Powyższy wynik pozwoli na wyprowadzenie odpowiednich własności norm operatorowych między przestrzeniami ℓ_p i ℓ_q .

Dotychczas rozważaliśmy prawe (lub lewe) kwaternionowe przestrzenie wektorowe \mathbb{H}^n z normą $\|\cdot\|_2$. Podobnie jednak jak w przypadku klasycznych przestrzeni \mathbb{R}^n lub \mathbb{C}^n możemy wprowadzać dowolne normy ℓ_p , zgodnie z poniższą definicją.

Definicja 2.6. Niech $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{H}^n$ i $p \in [1, \infty]$. Normą ℓ_p nazywamy funkcję

$$\|\cdot\|_p : \mathbb{H}^n \rightarrow \mathbb{R}_+ \cup \{0\}; \quad \|\mathbf{x}\|_p := \begin{cases} \left(\sum_{k=1}^n |x_k|^p \right)^{1/p} & \text{dla } p \in [1, \infty), \\ \max_{1 \leq k \leq n} |x_k| & \text{dla } p = \infty. \end{cases} \quad (2.18)$$

Fakt 2.7. *Funkcja zadana wzorem (2.18) spełnia aksjomaty normy.*

Warto zauważyć, że wzór (2.18) można rozszerzyć na $p \in (0, 1)$. Tak zdefiniowana funkcja $\|\cdot\|_p$ jest wówczas quasi-normą (nie spełnia warunku trójkąta). Będziemy również przyjmować standardowe oznaczenie dla liczności nośnika wektora $\mathbf{x} \in \mathbb{H}^n$.

Definicja 2.7. *Nośnikiem wektora $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{H}^n$ nazywamy zbiór indeksów niezerowych współrzędnych tego wektora, tzn.*

$$\text{supp } \mathbf{x} := \{k \in \{1, \dots, n\} : x_k \neq 0\}. \quad (2.19)$$

W dalszej części pracy będziemy stosować następującą notację

$$\|\mathbf{x}\|_0 := \#(\text{supp } \mathbf{x}), \quad (2.20)$$

gdzie $\#A$ oznacza licznosc (moc) zbioru A . Będziemy też często nazywać powyższą wielkość *normą* ℓ_0 . Trzeba mieć na uwadze, że jest to nazwa zwyczajowa (nie jest to norma, ani nawet quasi-norma), która jednak ma swoje uzasadnienie. Zauważmy, że przechodząc do granicy $p \rightarrow 0$ po prawej stronie wyrażenia

$$\|\mathbf{x}\|_p^p = \sum_{k=1}^n |x_k|^p$$

otrzymujemy moc nośnika sygnału \mathbf{x} . Aby być konsekwentnym, należałoby stosować notację $\|\mathbf{x}\|_0^0$, jednak przyjęło się pomijać 0 w indeksie górnym.

Należy zdefiniować jeszcze normę operatorową (macierzową) przekształcenia liniowego $\Phi : \mathbb{H}^n \rightarrow \mathbb{H}^m$, danego macierzą $\Phi \in \mathbb{H}^{m \times n}$.

Definicja 2.8. Niech $\Phi \in \mathbb{H}^{m \times n}$ i $1 \leq p, q \leq \infty$. Normą operatorową (macierzową) przekształcenia Φ między ℓ_p i ℓ_q nazywamy

$$\|\Phi\|_{p \rightarrow q} := \sup_{\|\mathbf{x}\|_p \leq 1} \|\Phi \mathbf{x}\|_q. \quad (2.21)$$

Fakt 2.8. Normę operatorową zadaną wzorem (2.21) można równoważnie zapisać w postaci

$$\|\Phi\|_{p \rightarrow q} = \sup_{\|\mathbf{x}\|_p = 1} \|\Phi \mathbf{x}\|_q = \sup_{\mathbf{x} \in \mathbb{H}^n \setminus \{0\}} \frac{\|\Phi \mathbf{x}\|_q}{\|\mathbf{x}\|_p}.$$

Ostatnim wynikiem, jaki przedstawimy w tym rozdziale, istotnym z punktu widzenia teorii oszczędnego próbkowania, jest charakteryzacja normy operatorowej między ℓ_2 i ℓ_2 dla macierzy hermitowskich. Korzysta ona bezpośrednio z własności prawych wartości i wektorów własnych macierzy hermitowskich omawianych wcześniej.

Twierdzenie 2.8. Niech $\Psi \in \mathbb{H}^{n \times n}$ będzie taka, że $\Psi = \Psi^*$. Wówczas

$$\|\Psi\|_{2 \rightarrow 2} = \sup_{\|\mathbf{x}\|_2 = 1} |\langle \Psi \mathbf{x}, \mathbf{x} \rangle| = \sup_{\mathbf{x} \in \mathbb{H}^n \setminus \{0\}} \frac{|\langle \Psi \mathbf{x}, \mathbf{x} \rangle|}{\|\mathbf{x}\|_2^2}. \quad (2.22)$$

Dowód. Przypomnijmy, że kwaternionowa macierz hermitowska ma rzeczywiste prawe wartości własne $\lambda_1, \dots, \lambda_n$. Ponadto, istnieje ortonormalna baza \mathbb{H}^n złożona z wektorów własnych \mathbf{x}_k odpowiadających wartościom własnym $\lambda_k \in \mathbb{R}$, $k = 1, \dots, n$, tzn.

$$\Psi \mathbf{x}_k = \mathbf{x}_k \lambda_k \quad \text{oraz} \quad \langle \mathbf{x}_k, \mathbf{x}_\ell \rangle = \mathbf{x}_\ell^* \mathbf{x}_k = \delta_{k,\ell} \quad \text{dla} \quad k, \ell = 1, \dots, n,$$

gdzie $\delta_{k,\ell} = 1$, gdy $k = \ell$ i $\delta_{k,\ell} = 0$, gdy $k \neq \ell$.

Wprowadźmy oznaczenie $\lambda_{\max} = \max_{k=1,\dots,n} |\lambda_k|$. Wówczas $\|\Psi\|_{2 \rightarrow 2} = \lambda_{\max}$. Rzeczywiście, dla dowolnego wektora $\mathbf{x} = \sum_{k=1}^n \mathbf{x}_k \alpha_k \in \mathbb{H}^n$ takiego, że $\|\mathbf{x}\|_2^2 = \sum_{k=1}^n |\alpha_k|^2 = 1$, mamy

$$\|\Psi \mathbf{x}\|_2^2 = \left\| \sum_{k=1}^n \Psi \mathbf{x}_k \alpha_k \right\|_2^2 = \left\| \sum_{k=1}^n \mathbf{x}_k \lambda_k \alpha_k \right\|_2^2 = \sum_{k=1}^n |\lambda_k \alpha_k|^2 \leq \lambda_{\max} \sum_{k=1}^n |\alpha_k|^2 = \lambda_{\max},$$

a dla odpowiedniego wektora własnego \mathbf{x}_ℓ dla którego $|\lambda_\ell| = \lambda_{\max}$, otrzymujemy

$$\|\Psi \mathbf{x}_\ell\|_2 = \|\mathbf{x}_\ell \lambda_\ell\|_2 = \lambda_{\max} \|\mathbf{x}_\ell\|_2 = \lambda_{\max}.$$

Z drugiej strony, ponieważ wartości własne są rzeczywiste a wektory \mathbf{x}_k parami ortonormalne, mamy

$$\begin{aligned} \langle \Psi \mathbf{x}, \mathbf{x} \rangle &= \mathbf{x}^* \Psi \mathbf{x} = \left(\sum_{k=1}^n \mathbf{x}_k \alpha_k \right)^* \left(\sum_{\ell=1}^n \Psi \mathbf{x}_\ell \alpha_\ell \right) \\ &= \left(\sum_{k=1}^n \overline{\alpha_k} \mathbf{x}_k^* \right) \left(\sum_{\ell=1}^n \mathbf{x}_\ell \lambda_\ell \alpha_\ell \right) \\ &= \sum_{k=1}^n \overline{\alpha_k} \lambda_k \alpha_k \stackrel{\lambda_k \in \mathbb{R}}{=} \sum_{k=1}^n \lambda_k |\alpha_k|^2. \end{aligned}$$

Stąd

$$|\langle \Psi \mathbf{x}, \mathbf{x} \rangle| \leq \lambda_{\max} \sum_k |\alpha_k|^2 = \lambda_{\max}.$$

Ponownie, dla odpowiedniego wektora własnego w ostatniej nierówności otrzymujemy równość, co daje poszukiwany wynik. \square

Rozdział ten pokazał wiele podobieństw między algebrą kwaternionów a ciałami liczb rzeczywistych i zespolonych. Wiele pojęć i twierdzeń to naturalne uogólnienia tych znanych z algebr niższego rzędu. Własności te znajdują bezpośrednie zastosowanie w dowodach twierdzeń będących podstawą teorii oszczędnego próbkowania.

Rozdział 3

Uogólnienie metod oszczędniego próbkowania na algebrę kwaternionów

3.1. Podstawowe pojęcia i sformułowanie problemu

W dalszej części pracy uogólnimy rozważania przedstawione pod koniec pierwszego rozdziału i sformułujemy zagadnienie, jakim zajmuje się teoria oszczędniego próbkowania w przypadku, gdy sygnał i macierz pomiarowa mają elementy z algebry kwaternionów. Zaczniemy od zdefiniowania podstawowych pojęć tej teorii.

Definicja 3.1. Wektor $\mathbf{x} \in \mathbb{H}^n$ nazywamy s -rzadkim, jeśli ma co najwyżej s niezerowych współrzędnych, tzn.

$$\|\mathbf{x}\|_0 \leq s. \quad (3.1)$$

Tak jak wcześniej wspomnieliśmy, jednym z zadań, jakie stawia sobie teoria oszczędniego próbkowania, jest rekonstrukcja rzadkich wektorów z niewielkiej liczby ich liniowych pomiarów $\mathbf{y} = \Phi \mathbf{x}$, $\Phi \in \mathbb{H}^{m \times n}$, poprzez rozwiązanie zadania optymalizacji wypukłej, a dokładniej minimalizacji normy ℓ_1 , tzn.

$$\arg \min_{\mathbf{z} \in \mathbb{H}^n} \|\mathbf{z}\|_1 \quad \text{pod warunkiem} \quad \Phi \mathbf{z} = \mathbf{y}. \quad (3.2)$$

Dopuszczamy również przypadek ogólniejszy, gdy pomiary nie są dokładne, ale obarczone

pewnym błędem $\mathbf{e} \in \mathbb{H}^m$ (białym szumem), czyli

$$\mathbf{y} = \Phi \mathbf{x} + \mathbf{e}, \quad \text{gdzie } \|\mathbf{e}\|_2 \leq \eta,$$

dla pewnego ustalonego (znanego) $\eta \geq 0$. O ile dokładna rekonstrukcja jest wówczas niemożliwa, chcielibyśmy jednak przybliżyć współrzędne tego wektora w sposób stabilny, tzn. z błędem, którego normę można oszacować w zależności od η . Rozwiązujemy wówczas zmodyfikowane zadanie optymalizacji wypukłej

$$\arg \min_{\mathbf{z} \in \mathbb{H}^n} \|\mathbf{z}\|_1 \quad \text{pod warunkiem} \quad \|\Phi \mathbf{z} - \mathbf{y}\|_2 \leq \eta. \quad (3.3)$$

Podobnie jak w przypadku rzeczywistym i zespolonym, jednym z warunków, który pozwala na stabilną rekonstrukcję wektorów rzadkich z pomiarów obarczonych błędem (lub wierną w przypadku pomiarów dokładnych), jest własność ograniczonej izometrii, wprowadzona przez E. Candès i T. Tao [8], której definicja rozszerza się na przypadek kwaternionowy.

Definicja 3.2. Niech $s \in \mathbb{N}$, $s \geq 1$, oraz $\Phi \in \mathbb{H}^{m \times n}$. Mówimy, że macierz Φ ma *własność s -ograniczonej izometrii* (ang. *s -restricted isometry property* – RIP) ze stałą $\delta_s \geq 0$, jeśli

$$(1 - \delta_s) \|\mathbf{x}\|_2^2 \leq \|\Phi \mathbf{x}\|_2^2 \leq (1 + \delta_s) \|\mathbf{x}\|_2^2 \quad (3.4)$$

dla wszystkich s -rzadkich wektorów $\mathbf{x} \in \mathbb{H}^n$. Najmniejszą liczbę $\delta_s \geq 0$ spełniającą powyższy warunek nazywamy *stałą s -ograniczonej izometrii* (ang. *s -restricted isometry constant*).

Oszacowanie wartości stałej s -ograniczonej izometrii jest niezwykle istotnym aspektem praktycznego wykorzystania teorii oszczędnego próbkowania. Z jednej strony chcemy wiedzieć jak duża może być ta stała by możliwa była stabilna rekonstrukcja, a z drugiej strony poszukujemy macierzy pomiarowych, które takie oszacowania spełniają. Dwa poniższe lematy charakteryzujące stałe ograniczonej izometrii są ważnym narzędziem w wykazaniu Twierdzenia 3.1, głównego wyniku tej pracy. Stanowią one uogólnienie analogicznych wyników dla wektorów i macierzy rzeczywistych oraz zespolonych.

Lemat 3.1. *Stała s -ograniczonej izometrii δ_s macierzy $\Phi \in \mathbb{H}^{m \times n}$ spełnia równość*

$$\delta_s = \max_{S \subseteq \{1, \dots, n\}, \#S \leq s} \|\Phi_S^* \Phi_S - \mathbf{Id}\|_{2 \rightarrow 2},$$

gdzie \mathbf{Id} jest macierzą jednostkową i Φ_S oznacza macierz powstałą z kolumn macierzy Φ o indeksach w zbiorze S .

Dowód. Ustalmy dowolne $s \in \{1, \dots, n\}$ oraz $S \subseteq \{1, \dots, n\}$ takie, że $\#S \leq s$. Zauważmy, że warunek (3.4) można równoważnie zapisać w postaci

$$\left| \|\Phi_S \mathbf{x}\|_2^2 - \|\mathbf{x}\|_2^2 \right| \leq \delta_s \|\mathbf{x}\|_2^2 \quad \text{dla wszystkich } \mathbf{x} \in \mathbb{H}^s,$$

gdzie δ_s jest stałą s -ograniczonej izometrii macierzy Φ . Lewa strona powyższej nierówności jest równa

$$\left| \|\Phi_S \mathbf{x}\|_2^2 - \|\mathbf{x}\|_2^2 \right| = |\langle \Phi_S \mathbf{x}, \Phi_S \mathbf{x} \rangle - \langle \mathbf{x}, \mathbf{x} \rangle| = |\langle (\Phi_S^* \Phi_S - \mathbf{Id}) \mathbf{x}, \mathbf{x} \rangle|$$

i z Twierdzenia 2.8, ponieważ macierz $\Phi_S^* \Phi_S - \mathbf{Id}$ jest hermitowska, otrzymujemy

$$\sup_{\mathbf{x} \in \mathbb{H}^s \setminus \{0\}} \frac{|\langle (\Phi_S^* \Phi_S - \mathbf{Id}) \mathbf{x}, \mathbf{x} \rangle|}{\|\mathbf{x}\|_2^2} = \|\Phi_S^* \Phi_S - \mathbf{Id}\|_{2 \rightarrow 2}.$$

Dowolność wyboru s oraz S kończy dowód. \square

Lemat 3.2. Niech δ_s oznacza stałą s -ograniczonej izometrii macierzy $\Phi \in \mathbb{H}^{m \times n}$ dla dowolnego $s \in \{1, \dots, n\}$. Niech $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{H}^n$ będą wektorami takimi, że $\|\mathbf{x}_1\|_0 \leq s_1$ i $\|\mathbf{x}_2\|_0 \leq s_2$ dla $s_1 + s_2 \leq n$, i ponadto $\text{supp } \mathbf{x}_1 \cap \text{supp } \mathbf{x}_2 = \emptyset$. Wówczas

$$|\langle \Phi \mathbf{x}_1, \Phi \mathbf{x}_2 \rangle| \leq \delta_{s_1+s_2} \|\mathbf{x}_1\|_2 \cdot \|\mathbf{x}_2\|_2.$$

Dowód. Niech $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{H}^n$ będą wektorami spełniającymi założenia lematu i wprowadźmy oznaczenie $S = \text{supp } \mathbf{x}_1 \cup \text{supp } \mathbf{x}_2$. Oczywiście $\#S = s_1 + s_2$. Ponieważ \mathbf{x}_1 i \mathbf{x}_2 mają rozłączne nośniki, to są ortogonalne, tzn. $\langle \mathbf{x}_1, \mathbf{x}_2 \rangle = \langle \mathbf{x}_{1|S}, \mathbf{x}_{2|S} \rangle = 0$, gdzie $\mathbf{x}_{i|S}$ oznacza wektor powstały ze współrzędnych wektora \mathbf{x}_i o indeksach w zbiorze S . Korzystając z nierówności Cauchy'ego-Schwarza (Twierdzenie 2.3) otrzymujemy

$$\begin{aligned} |\langle \Phi \mathbf{x}_1, \Phi \mathbf{x}_2 \rangle| &= \left| \langle \Phi_S \mathbf{x}_{1|S}, \Phi_S \mathbf{x}_{2|S} \rangle \right| = \left| \langle (\Phi_S^* \Phi_S - \mathbf{Id}) \mathbf{x}_{1|S}, \mathbf{x}_{2|S} \rangle \right| \\ &\leq \|\Phi_S^* \Phi_S - \mathbf{Id}\|_{2 \rightarrow 2} \|\mathbf{x}_{1|S}\|_2 \|\mathbf{x}_{2|S}\|_2 \leq \delta_{s_1+s_2} \|\mathbf{x}_{1|S}\|_2 \|\mathbf{x}_{2|S}\|_2, \end{aligned}$$

co kończy dowód, ponieważ $\|\mathbf{x}_{1|S}\|_2 = \|\mathbf{x}_1\|_2$ i $\|\mathbf{x}_{2|S}\|_2 = \|\mathbf{x}_2\|_2$. \square

Powyższe lematy pozwolą już na sformułowanie i udowodnienie twierdzeń, stanowiących podstawę teorii oszczędnego próbkowania w przypadku kwaternionowym.

3.2. Wcześniejsze wyniki dla pomiarów o współczynnikach rzeczywistych

Uogólnienie teorii oszczędnego próbkowania na przypadek kwaternionowy przebiegało w dwóch etapach. W pierwszej kolejności wykazaliśmy odpowiednik Twierdzenia 1.1 dla macierzy pomiarowych $\Phi \in \mathbb{R}^{m \times n}$ o współczynnikach rzeczywistych oraz wektorów $\mathbf{x} \in \mathbb{H}^n$ kwaternionowych. Krok ten miał istotne znaczenie również z punktu widzenia aplikacyjnego, o czym świadczy poniższy lemat (por. [2, Lemma 3.2]).

Lemat 3.3. *Jeśli macierz $\Phi \in \mathbb{R}^{m \times n}$ spełnia warunek s -ograniczonej izometrii dla wektorów o współrzędnych rzeczywistych z pewną stałą $\delta_s \geq 0$, tzn.*

$$(1 - \delta_s) \|\mathbf{z}\|_2^2 \leq \|\Phi \mathbf{z}\|_2^2 \leq (1 + \delta_s) \|\mathbf{z}\|_2^2 \quad (3.5)$$

dla wszystkich s -rzadkich wektorów $\mathbf{z} \in \mathbb{R}^n$, to spełnia warunek s -ograniczonej izometrii dla wektorów o współrzędnych kwaternionowych z tą samą stałą δ_s , tzn.

$$(1 - \delta_s) \|\mathbf{x}\|_2^2 \leq \|\Phi \mathbf{x}\|_2^2 \leq (1 + \delta_s) \|\mathbf{x}\|_2^2$$

dla wszystkich s -rzadkich wektorów $\mathbf{x} \in \mathbb{H}^n$.

Dowód. Niech $\mathbf{x} \in \mathbb{H}^n$ będzie dowolnym s -rzadkim wektorem. Zapiszmy go w postaci

$$\mathbf{x} = \mathbf{x}_r + \mathbf{x}_i \mathbf{i} + \mathbf{x}_j \mathbf{j} + \mathbf{x}_k \mathbf{k},$$

gdzie $\mathbf{x}_i \in \mathbb{R}^n$ dla $i \in \{\mathbf{r}, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$. Zauważmy, że każdy z wektorów $\mathbf{x}_i \in \mathbb{R}^n$ jest s -rzadki, a ponadto

$$\Phi \mathbf{x} = \Phi \mathbf{x}_r + (\Phi \mathbf{x}_i) \mathbf{i} + (\Phi \mathbf{x}_j) \mathbf{j} + (\Phi \mathbf{x}_k) \mathbf{k}$$

i $\Phi \mathbf{x}_i \in \mathbb{R}^m$ dla $i \in \{\mathbf{r}, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$. Stąd

$$\|\Phi \mathbf{x}\|_2^2 = \|\Phi \mathbf{x}_r\|_2^2 + \|\Phi \mathbf{x}_i\|_2^2 + \|\Phi \mathbf{x}_j\|_2^2 + \|\Phi \mathbf{x}_k\|_2^2$$

i oczywiście

$$\|\mathbf{x}\|_2^2 = \|\mathbf{x}_r\|_2^2 + \|\mathbf{x}_i\|_2^2 + \|\mathbf{x}_j\|_2^2 + \|\mathbf{x}_k\|_2^2.$$

Stosując nierówność (3.5) oddzielnie do każdego z wektorów \mathbf{x}_i otrzymujemy tezę. \square

Dzięki Lematowi 3.3 widzimy, że jeśli macierz pomiarowa o elementach rzeczywistych ma własność ograniczonej izometrii dla wektorów o współrzędnych rzeczywistych, to możemy jej z powodzeniem używać w zastosowaniach teorii oszczędnego próbkowania dla wektorów (sygnałów) kwaternionowych. W literaturze znane są pewne klasy macierzy rzeczywistych i zespolonych spełniających tę własność, np. rzeczywiste losowe macierze gaussowskie, czy też częściowa macierz dyskretnej transformacji Fouriera. Szczegółowe rozważania na temat macierzy pomiarowych zaprezentowane zostaną w kolejnym rozdziale.

Pierwszym krokiem w kierunku uogólnienia teorii oszczędnego próbkowania na algebrę kwaternionów był odpowiednik Twierdzenia 1.1 dla przypadku wektorów kwaternionowych i macierzy o elementach w ciele liczb rzeczywistych, tzn.

$$\mathbf{y} = \Phi \mathbf{x} + \mathbf{e}, \quad \text{gdzie } \mathbf{x} \in \mathbb{H}^n, \Phi \in \mathbb{R}^{m \times n}, \mathbf{y} \in \mathbb{H}^m \text{ oraz } \mathbf{e} \in \mathbb{H}^n, \|\mathbf{e}\|_2 \leq \eta$$

dla pewnych $m \leq n$ i $\eta \geq 0$. W dowodzie wspomnianego wyniku [2, Theorem 4.1] korzystaliśmy z kwaternionowej wersji równości polaryzacyjnej (Twierdzenie 2.5), która pozwala na otrzymanie wyniku analogicznego do tego z Lematu 3.2, jednak z gorszą stałą. Główny wynik tej pracy, Twierdzenie 3.1, rozszerza ten rezultat na przypadek macierzy o elementach kwaternionowych i poprawia występujące w tezie oszacowania.

3.3. Rekonstrukcja danych kwaternionowych z pomiarów o współczynnikach kwaternionowych

Narzędzia wprowadzone w tym rozdziale pozwolą nam na udowodnienie podstawowego twierdzenia w teorii oszczędnego próbkowania w przypadku w pełni kwaternionowym, tzn. zarówno dla macierzy pomiarowych jak i sygnałów o współrzędnych kwaternionowych. Przedstawione rozumowanie inspirowane było dowodem zaprezentowanym przez E. Candès'a w [6].

Będziemy stosować notację analogiczną do dotychczasowej, tzn.

$$\mathbf{y} = \Phi \mathbf{x} + \mathbf{e}, \quad \text{gdzie } \mathbf{x} \in \mathbb{H}^n, \Phi \in \mathbb{H}^{m \times n}, \mathbf{y} \in \mathbb{H}^m \text{ oraz } \mathbf{e} \in \mathbb{H}^n, \|\mathbf{e}\|_2 \leq \eta,$$

dla pewnych $m \leq n$ i $\eta \geq 0$. Przez $\mathbf{x}_{|s}$ będziemy oznaczać najlepsze s -rzadkie przybliżenie wektora \mathbf{x} , tzn. wektor powstały z wektora \mathbf{x} przez pozostawienie s największych (co do

normy kwaternionowej) współrzędnych oraz zastąpienie pozostałych współrzędnych zerami. Ponadto, dla dowolnego $\mathbf{h} = (h_1, \dots, h_n)^T \in \mathbb{H}^n$ i zbioru indeksów $T \subseteq \{1, \dots, n\}$, niech $\mathbf{h}_T \in \mathbb{H}^n$ oznacza wektor taki, że

$$(\mathbf{h}_T)_i = \begin{cases} h_i & \text{jeśli } i \in T, \\ 0 & \text{w p.p.} \end{cases}$$

dla $i \in \{1, \dots, n\}$. Dopełnienie zbioru $T \subseteq \{1, \dots, n\}$ będziemy standardowo oznaczać symbolem $T^c = \{1, \dots, n\} \setminus T$.

Twierdzenie 3.1. *Niech macierz $\Phi \in \mathbb{H}^{m \times n}$ spełnia własność $2s$ -graniczonej izometrii ze stałą $\delta_{2s} < \sqrt{2} - 1$ i niech $\eta \geq 0$. Wówczas dla dowolnego $\mathbf{x} \in \mathbb{H}^n$ oraz $\mathbf{y} = \Phi\mathbf{x} + \mathbf{e}$ takich, że $\|\mathbf{e}\|_2 \leq \eta$, rozwiązanie $\mathbf{x}^\#$ problemu*

$$\arg \min_{\mathbf{z} \in \mathbb{H}^n} \|\mathbf{z}\|_1 \quad \text{pod warunkiem} \quad \|\Phi\mathbf{z} - \mathbf{y}\|_2 \leq \eta \quad (P_{\mathbb{H}-\mathbb{H}, \eta})$$

spełnia nierówność

$$\|\mathbf{x}^\# - \mathbf{x}\|_2 \leq \frac{C_0}{\sqrt{s}} \|\mathbf{x} - \mathbf{x}_{|s}\|_1 + C_1 \eta \quad (3.6)$$

ze stałymi

$$C_0 = 2 \cdot \frac{1 + (\sqrt{2} - 1) \delta_{2s}}{1 - (\sqrt{2} + 1) \delta_{2s}}, \quad C_1 = \frac{4\sqrt{1 + \delta_{2s}}}{1 - (1 + \sqrt{2}) \delta_{2s}}.$$

Dowód. Zauważmy na początek, że ponieważ $\mathbf{x}^\#$ jest rozwiązaniem problemu minimalizacji $(P_{\mathbb{H}-\mathbb{H}, \eta})$, a \mathbf{x} spełnia ograniczenie postawione dla tego problemu (tzn. $\|\Phi\mathbf{x} - \mathbf{y}\|_2 \leq \eta$), zatem z nierówności trójkąta wynika, że

$$\|\Phi(\mathbf{x}^\# - \mathbf{x})\|_2 \leq \|\Phi\mathbf{x}^\# - \mathbf{y}\|_2 + \|\mathbf{y} - \Phi\mathbf{x}\|_2 \leq 2\eta. \quad (3.7)$$

Niech $\mathbf{h} := \mathbf{x}^\# - \mathbf{x}$ i rozłóżmy wektor \mathbf{h} na sumę wektorów $\mathbf{h}_{T_0}, \mathbf{h}_{T_1}, \mathbf{h}_{T_2}, \dots$ o rozłącznych nośnikach, gdzie każdy \mathbf{h}_{T_j} jest co najwyżej s -rzadki, w następujący sposób: niech T_0 będzie zbiorem indeksów s największych co do normy kwaternionowej współrzędnych wektora \mathbf{x} , T_1 będzie zbiorem indeksów s największych co do normy kwaternionowej współrzędnych wektora $\mathbf{h}_{T_0^c}$, T_2 będzie zbiorem indeksów s największych współrzędnych wektora $\mathbf{h}_{(T_0 \cup T_1)^c}$, itd.

Zwróćmy uwagę na fakt, że dla $j \geq 2$ zachodzi

$$\|\mathbf{h}_{T_j}\|_2^2 = \sum_{i \in T_j} |h_i|^2 \leq \sum_{i \in T_j} \|\mathbf{h}_{T_j}\|_\infty^2 \leq s \|\mathbf{h}_{T_j}\|_\infty^2,$$

gdzie h_i są kolejnymi współrzędnymi wektora \mathbf{h} oraz $\|\mathbf{h}_{T_j}\|_\infty = \max_{i \in T_j} |h_i|$. Ostatnia nierówność wynika z faktu, że ostatni ze zbiorów T_j może mieć mniej elementów niż s . Ponadto, ponieważ wszystkie niezerowe współrzędne $\mathbf{h}_{T_{j-1}}$ mają normy nie mniejsze niż normy wszystkich niezerowych współrzędnych \mathbf{h}_{T_j} , otrzymujemy

$$\|\mathbf{h}_{T_j}\|_\infty \leq \frac{1}{s} \sum_{i \in T_{j-1}} |h_i| = \frac{1}{s} \|\mathbf{h}_{T_{j-1}}\|_1.$$

Podsumowując powyższe rozważania, dla $j \geq 2$ mamy

$$\|\mathbf{h}_{T_j}\|_2 \leq \sqrt{s} \|\mathbf{h}_{T_j}\|_\infty \leq \frac{1}{\sqrt{s}} \|\mathbf{h}_{T_{j-1}}\|_1, \quad (3.8)$$

a stąd

$$\sum_{j \geq 2} \|\mathbf{h}_{T_j}\|_2 \leq \frac{1}{\sqrt{s}} (\|\mathbf{h}_{T_1}\|_1 + \|\mathbf{h}_{T_2}\|_1 + \dots) \leq \frac{1}{\sqrt{s}} \|\mathbf{h}_{T_0^c}\|_1 \quad (3.9)$$

i ostatecznie, z nierówności trójkąta otrzymujemy

$$\|\mathbf{h}_{(T_0 \cup T_1)^c}\|_2 = \left\| \sum_{j \geq 2} \mathbf{h}_{T_j} \right\|_2 \leq \sum_{j \geq 2} \|\mathbf{h}_{T_j}\|_2 \leq \frac{1}{\sqrt{s}} \|\mathbf{h}_{T_0^c}\|_1. \quad (3.10)$$

Należy zauważyć, że $\|\mathbf{h}_{T_0^c}\|_1$ nie może być duża, ponieważ $\|\mathbf{x} + \mathbf{h}\|_1 = \|\mathbf{x}^\#\|_1$ jest minimalna. Rzeczywiście, zachodzi

$$\|\mathbf{x}\|_1 \geq \|\mathbf{x} + \mathbf{h}\|_1 = \|\mathbf{x}_{T_0} + \mathbf{h}_{T_0}\|_1 + \|\mathbf{x}_{T_0^c} + \mathbf{h}_{T_0^c}\|_1 \geq \|\mathbf{x}_{T_0}\|_1 - \|\mathbf{h}_{T_0}\|_1 - \|\mathbf{x}_{T_0^c}\|_1 + \|\mathbf{h}_{T_0^c}\|_1,$$

co prowadzi do

$$\|\mathbf{x}_{T_0^c}\|_1 = \|\mathbf{x}\|_1 - \|\mathbf{x}_{T_0}\|_1 \geq -\|\mathbf{h}_{T_0}\|_1 - \|\mathbf{x}_{T_0^c}\|_1 + \|\mathbf{h}_{T_0^c}\|_1,$$

a stąd

$$\|\mathbf{h}_{T_0^c}\|_1 \leq \|\mathbf{h}_{T_0}\|_1 + 2\|\mathbf{x}_{T_0^c}\|_1. \quad (3.11)$$

Z (3.10) i (3.11) dostajemy

$$\|\mathbf{h}_{(T_0 \cup T_1)^c}\|_2 \leq \frac{1}{\sqrt{s}} \|\mathbf{h}_{T_0^c}\|_1 \leq \frac{1}{\sqrt{s}} \|\mathbf{h}_{T_0}\|_1 + \frac{2}{\sqrt{s}} \|\mathbf{x}_{T_0^c}\|_1.$$

Zauważmy, że $\|\mathbf{x}_{T_0^c}\|_1 = \|\mathbf{x} - \mathbf{x}_{|s}\|_1$, a z nierówności Cauchy'ego-Schwarza mamy

$$\|\mathbf{h}_{T_0}\|_1 = \sum_{i \in T_0} |h_i| \cdot 1 \leq \sqrt{\sum_{i \in T_0} |h_i|^2} \cdot \sqrt{\sum_{i \in T_0} 1^2} = \|\mathbf{h}_{T_0}\|_2 \cdot \sqrt{s}$$

i oznaczając $e = \frac{1}{\sqrt{s}} \|\mathbf{x} - \mathbf{x}_{|s}\|_1 = \frac{1}{\sqrt{s}} \|\mathbf{x}_{T_0^c}\|_1$, dostajemy

$$\|\mathbf{h}_{(T_0 \cup T_1)^c}\|_2 \leq \|\mathbf{h}_{T_0}\|_2 + 2e, \quad (3.12)$$

co daje pierwsze z oszacowań, jakie posłużą do wykazania tezy.

W kolejnym etapie należy ograniczyć $\|\mathbf{h}_{T_0 \cup T_1}\|_2$. Zauważmy, że

$$\Phi \mathbf{h}_{T_0 \cup T_1} = \Phi \left(\mathbf{h} - \sum_{j \geq 2} \mathbf{h}_{T_j} \right) = \Phi \mathbf{h} - \sum_{j \geq 2} \Phi \mathbf{h}_{T_j}$$

i stąd, korzystając z własności funkcji $\langle \cdot, \cdot \rangle$, otrzymujemy

$$\begin{aligned} \|\Phi \mathbf{h}_{T_0 \cup T_1}\|_2^2 &= \langle \Phi \mathbf{h}_{T_0 \cup T_1}, \Phi \mathbf{h}_{T_0 \cup T_1} \rangle = \langle \Phi \mathbf{h}_{T_0 \cup T_1}, \Phi \mathbf{h} \rangle - \sum_{j \geq 2} \langle \Phi \mathbf{h}_{T_0 \cup T_1}, \Phi \mathbf{h}_{T_j} \rangle \\ &= \langle \Phi \mathbf{h}_{T_0 \cup T_1}, \Phi \mathbf{h} \rangle - \sum_{j \geq 2} \langle \Phi \mathbf{h}_{T_0}, \Phi \mathbf{h}_{T_j} \rangle - \sum_{j \geq 2} \langle \Phi \mathbf{h}_{T_1}, \Phi \mathbf{h}_{T_j} \rangle. \end{aligned} \quad (3.13)$$

Z Twierdzenia 2.3, nierówności (3.7) oraz z własności ograniczonej izometrii wynika, że

$$|\langle \Phi \mathbf{h}_{T_0 \cup T_1}, \Phi \mathbf{h} \rangle| \leq \|\Phi \mathbf{h}_{T_0 \cup T_1}\|_2 \cdot \|\Phi \mathbf{h}\|_2 \leq \sqrt{1 + \delta_{2s}} \cdot \|\mathbf{h}_{T_0 \cup T_1}\|_2 \cdot 2\eta, \quad (3.14)$$

a z Lematu 3.2 dostajemy, że

$$\left| \langle \Phi \mathbf{h}_{T_0}, \Phi \mathbf{h}_{T_j} \rangle \right| \leq \delta_{2s} \cdot \|\mathbf{h}_{T_0}\|_2 \cdot \|\mathbf{h}_{T_j}\|_2, \quad \text{dla } j \geq 2, \quad (3.15)$$

$$\left| \langle \Phi \mathbf{h}_{T_1}, \Phi \mathbf{h}_{T_j} \rangle \right| \leq \delta_{2s} \cdot \|\mathbf{h}_{T_1}\|_2 \cdot \|\mathbf{h}_{T_j}\|_2, \quad \text{dla } j \geq 2. \quad (3.16)$$

Ponieważ T_0 i T_1 są zbiorami rozłącznymi, mamy

$$\|\mathbf{h}_{T_0}\|_2 + \|\mathbf{h}_{T_1}\|_2 \leq \sqrt{2} \|\mathbf{h}_{T_0 \cup T_1}\|_2, \quad (3.17)$$

co wynika z faktu, że $\|\mathbf{h}_{T_0}\|_2^2 + \|\mathbf{h}_{T_1}\|_2^2 = \|\mathbf{h}_{T_0 \cup T_1}\|_2^2$ oraz $2(a^2 + b^2) \geq (a + b)^2$ dla dowolnych $a, b \in \mathbb{R}$.

Stąd, z własności ograniczonej izometrii i kolejno z (3.13), nierówności trójkąta oraz oszacowań (3.14)–(3.16), (3.17) i (3.9),

$$\begin{aligned}
(1 - \delta_{2s}) \|\mathbf{h}_{T_0 \cup T_1}\|_2^2 &\leq \|\Phi \mathbf{h}_{T_0 \cup T_1}\|_2^2 \\
&= \left| \langle \Phi \mathbf{h}_{T_0 \cup T_1}, \Phi \mathbf{h} \rangle - \sum_{j \geq 2} \langle \Phi \mathbf{h}_{T_0}, \Phi \mathbf{h}_{T_j} \rangle - \sum_{j \geq 2} \langle \Phi \mathbf{h}_{T_1}, \Phi \mathbf{h}_{T_j} \rangle \right| \\
&\leq |\langle \Phi \mathbf{h}_{T_0 \cup T_1}, \Phi \mathbf{h} \rangle| + \sum_{j \geq 2} |\langle \Phi \mathbf{h}_{T_0}, \Phi \mathbf{h}_{T_j} \rangle| + \sum_{j \geq 2} |\langle \Phi \mathbf{h}_{T_1}, \Phi \mathbf{h}_{T_j} \rangle| \\
&\leq 2\sqrt{1 + \delta_{2s}} \cdot \eta \cdot \|\mathbf{h}_{T_0 \cup T_1}\|_2 + \delta_{2s} \cdot (\|\mathbf{h}_{T_0}\|_2 + \|\mathbf{h}_{T_1}\|_2) \cdot \sum_{j \geq 2} \|\mathbf{h}_{T_j}\|_2 \\
&\leq \left(2\sqrt{1 + \delta_{2s}} \cdot \eta + \sqrt{2} \delta_{2s} \cdot \sum_{j \geq 2} \|\mathbf{h}_{T_j}\|_2 \right) \cdot \|\mathbf{h}_{T_0 \cup T_1}\|_2 \\
&\leq \left(2\sqrt{1 + \delta_{2s}} \cdot \eta + \frac{\sqrt{2} \delta_{2s}}{\sqrt{s}} \|\mathbf{h}_{T_0^c}\|_1 \right) \cdot \|\mathbf{h}_{T_0 \cup T_1}\|_2.
\end{aligned}$$

Po obustronnym podzieleniu przez $(1 - \delta_{2s}) \|\mathbf{h}_{T_0 \cup T_1}\|_2$ mamy

$$\|\mathbf{h}_{T_0 \cup T_1}\|_2 \leq \frac{2\sqrt{1 + \delta_{2s}}}{1 - \delta_{2s}} \cdot \eta + \frac{\sqrt{2} \delta_{2s}}{1 - \delta_{2s}} \cdot \frac{\|\mathbf{h}_{T_0^c}\|_1}{\sqrt{s}}. \quad (3.18)$$

Z powyższego, i z nierówności (3.11), otrzymujemy

$$\|\mathbf{h}_{T_0 \cup T_1}\|_2 \leq \alpha \cdot \eta + \beta \cdot \frac{\|\mathbf{h}_{T_0}\|_1}{\sqrt{s}} + 2\beta \cdot \frac{\|\mathbf{x}_{T_0^c}\|_1}{\sqrt{s}},$$

gdzie

$$\alpha = \frac{2\sqrt{1 + \delta_{2s}}}{1 - \delta_{2s}}, \quad \beta = \frac{\sqrt{2} \delta_{2s}}{1 - \delta_{2s}}$$

i ponieważ $\|\mathbf{h}_{T_0}\|_1 \leq \sqrt{s} \cdot \|\mathbf{h}_{T_0}\|_2 \leq \sqrt{s} \cdot \|\mathbf{h}_{T_0 \cup T_1}\|_2$, to musi zachodzić

$$\|\mathbf{h}_{T_0 \cup T_1}\|_2 \leq \alpha \cdot \eta + \beta \cdot \|\mathbf{h}_{T_0 \cup T_1}\|_2 + 2\beta \cdot e,$$

gdzie przypomnijmy, że $e = \frac{1}{\sqrt{s}} \|\mathbf{x}_{T_0^c}\|_2$. Zatem

$$\|\mathbf{h}_{T_0 \cup T_1}\|_2 \leq \frac{1}{1 - \beta} (\alpha \cdot \eta + 2\beta \cdot e), \quad (3.19)$$

o ile spełnione będzie założenie $\beta < 1$, tzn.

$$\frac{\sqrt{2} \delta_{2s}}{1 - \delta_{2s}} - 1 = \frac{\delta_{2s} (1 + \sqrt{2}) - 1}{1 - \delta_{2s}} < 0,$$

czyli $\delta_{2s} < \frac{1}{1+\sqrt{2}} = \sqrt{2} - 1$. Podsumowując, z (3.12) oraz (3.19)

$$\begin{aligned} \|\mathbf{h}\|_2 &\leq \|\mathbf{h}_{T_0 \cup T_1}\|_2 + \|\mathbf{h}_{(T_0 \cup T_1)^c}\|_2 \leq \|\mathbf{h}_{T_0 \cup T_1}\|_2 + \|\mathbf{h}_{T_0}\|_2 + 2e \\ &\leq 2\|\mathbf{h}_{T_0 \cup T_1}\|_2 + 2e \leq \frac{2\alpha}{1-\beta} \cdot \eta + \left(\frac{4\beta}{1-\beta} + 2\right) \cdot e = C_1\eta + C_0e, \end{aligned}$$

gdzie

$$\begin{aligned} C_1 &= \frac{2\alpha}{1-\beta} = 2 \frac{2\sqrt{1+\delta_{2s}}}{1-\delta_{2s}} \cdot \left(1 - \frac{\sqrt{2}\delta_{2s}}{1-\delta_{2s}}\right)^{-1} = \frac{4\sqrt{1+\delta_{2s}}}{1-\delta_{2s}} \cdot \frac{1-\delta_{2s}}{1-(1+\sqrt{2})\delta_{2s}} \\ &= \frac{4\sqrt{1+\delta_{2s}}}{1-(1+\sqrt{2})\delta_{2s}}, \end{aligned}$$

oraz

$$\begin{aligned} C_0 &= \frac{4\beta}{1-\beta} + 2 = \frac{2\beta+2}{1-\beta} = 2 \frac{1+\beta}{1-\beta} = 2 \left(1 + \frac{\sqrt{2}\delta_{2s}}{1-\delta_{2s}}\right) \left(1 - \frac{\sqrt{2}\delta_{2s}}{1-\delta_{2s}}\right)^{-1} \\ &= 2 \cdot \frac{1-(1-\sqrt{2})\delta_{2s}}{1-\delta_{2s}} \cdot \frac{1-\delta_{2s}}{1-(\sqrt{2}+1)\delta_{2s}} = 2 \cdot \frac{1+(\sqrt{2}-1)\delta_{2s}}{1-(\sqrt{2}+1)\delta_{2s}}, \end{aligned}$$

co daje tęzę i kończy dowód. □

Zauważmy, że Twierdzenie 3.1 nie zakłada s -rzadkości poszukiwanego wektora \mathbf{x} . Z twierdzenia wynika szereg ciekawych wniosków dotyczących rekonstrukcji wektorów kwaternionowych z dokładnych pomiarów $\mathbf{y} = \Phi\mathbf{x}$, a w szczególności wektorów s -rzadkich.

Wniosek 3.1. *Niech macierz $\Phi \in \mathbb{H}^{m \times n}$ spełnia własność $2s$ -graniczonej izometrii ze stałą $\delta_{2s} < \sqrt{2} - 1$. Wówczas dla dowolnego $\mathbf{x} \in \mathbb{H}^n$ oraz dla dokładnych danych pomiarowych $\mathbf{y} = \Phi\mathbf{x} \in \mathbb{H}^m$ rozwiązanie $\mathbf{x}^\#$ problemu*

$$\arg \min_{\mathbf{z} \in \mathbb{H}^n} \|\mathbf{z}\|_1 \quad \text{pod warunkiem} \quad \Phi\mathbf{z} = \mathbf{y}, \quad (P_{\mathbb{H}-\mathbb{H},0})$$

spełnia nierówności

$$\|\mathbf{x}^\# - \mathbf{x}\|_1 \leq C_0 \|\mathbf{x} - \mathbf{x}_{|s}\|_1, \quad (3.20)$$

$$\|\mathbf{x}^\# - \mathbf{x}\|_2 \leq \frac{C_0}{\sqrt{s}} \|\mathbf{x} - \mathbf{x}_{|s}\|_1 \quad (3.21)$$

ze stałą C_0 jak w Twierdzeniu 3.1. W szczególności, jeśli \mathbf{x} jest wektorem s -rzadkim, to rekonstrukcja w wyniku minimalizacji ℓ_1 jest dokładna.

Dowód. Teza dla sygnałów s -rzadkich jest oczywista i wynika z faktu, że wówczas $\mathbf{x} = \mathbf{x}|_s$. Nierówność (3.21) jest bezpośrednią konsekwencją Twierdzenia 3.1 dla $\eta = 0$.

Wystarczy zatem wykazać jedynie nierówność (3.20). Będziemy stosować oznaczenia takie jak w dowodzie Twierdzenia 3.1. Jak już pokazaliśmy, dla $\mathbf{h} = \mathbf{x}^\# - \mathbf{x}$ i zbiorów T_j prawdziwe są nierówności

$$\|\mathbf{h}_{T_0}\|_1 \leq \sqrt{s} \cdot \|\mathbf{h}_{T_0}\|_2 \leq \sqrt{s} \cdot \|\mathbf{h}_{T_0 \cup T_1}\|_2$$

i łącząc ten wynik z nierównością (3.18) dla $\eta = 0$ otrzymujemy

$$\|\mathbf{h}_{T_0}\|_1 \leq \beta \cdot \|\mathbf{h}_{T_0^c}\|_1, \quad \beta = \frac{\sqrt{2}\delta_{2s}}{1 - \delta_{2s}}. \quad (3.22)$$

Korzystając z powyższej nierówności oraz z (3.11) otrzymujemy

$$\|\mathbf{h}_{T_0^c}\|_1 \leq \beta \cdot \|\mathbf{h}_{T_0^c}\|_1 + 2\|\mathbf{x}_{T_0^c}\|_1$$

i stąd $\|\mathbf{h}_{T_0^c}\|_1 \leq \frac{2}{1-\beta} \cdot \|\mathbf{x}_{T_0^c}\|_1$. Zatem, w przypadku dokładnych pomiarów, tj. gdy $\eta = 0$, wektor $\mathbf{h} = \mathbf{x}^\# - \mathbf{x}$ spełnia, na mocy (3.22), oszacowanie

$$\|\mathbf{h}\|_1 = \|\mathbf{h}_{T_0}\|_1 + \|\mathbf{h}_{T_0^c}\|_1 \leq (1 + \beta) \cdot \|\mathbf{h}_{T_0^c}\|_1 \leq 2 \cdot \frac{1 + \beta}{1 - \beta} \cdot \|\mathbf{x}_{T_0^c}\|_1,$$

co dowodzi tezy, ponieważ $C_0 = 2\frac{1+\beta}{1-\beta}$, zgodnie z obliczeniami z poprzedniego dowodu. \square

Przypuszczamy, że oszacowanie $\delta_{2s} < \sqrt{2} - 1$ nie jest optymalne. W literaturze można znaleźć lepsze oszacowania dla macierzy o współczynnikach rzeczywistych i dla własności ograniczonej izometrii dla wektorów rzeczywistych. Znany jest jednak przykład s -rzadkiego wektora o współczynnikach rzeczywistych, którego nie można jednoznacznie zrekonstruować w wyniku minimalizacji ℓ_1 dla $\delta_s > \frac{1}{3}$, co daje nam również górne oszacowanie dla przypadku kwaternionowego.

Rozdział 4

Analiza kwaternionowych losowych macierzy gaussowskich

Pokazaliśmy w poprzednim rozdziale, że jeśli macierz pomiarowa $\Phi \in \mathbb{H}^{m \times n}$ ma własność ograniczonej izometrii z odpowiednio małą stałą, to możliwa jest dokładna rekonstrukcja wektorów rzadkich z dokładnych pomiarów (czy też stabilna rekonstrukcja dowolnych wektorów z pomiarów obarczonych błędem). Powstaje jednak pytanie, czy istnieją macierze kwaternionowe, które tę własność posiadają. Jak wspomnieliśmy, znane są przykłady macierzy o elementach rzeczywistych czy zespolonych, które mają własność ograniczonej izometrii. Jest to m.in. losowa rzeczywista macierz gaussowska (która spełnia RIP z dużym prawdopodobieństwem) [4, 8] czy też macierz (zespolona) dyskretnej transformacji Fouriera [9].

Rozważmy przykład macierzy losowych. W literaturze istnieje wiele różnych dowodów potwierdzających, że losowe macierze gaussowskie o elementach rzeczywistych mają własność ograniczonej izometrii (z dużym prawdopodobieństwem). W przypadku macierzy zespolonych pojawiły się publikacje, w których autorzy wykonują symulacje numeryczne sugerujące, że macierze te mają tę własność, jednak nie udowadniają tego teoretycznie. Zgodnie z naszą wiedzą, pierwszy (i jedyny) dowód opublikowano w 2013 roku w artykule [28]. Niedawno pojawił się również cykl prac, których autorzy, posługując się prostymi narzędziami probabilistycznymi, twierdzą, że pokazali własność ograniczonej izometrii dla gaussowskich macierzy losowych, zarówno rzeczywistych jak i zespolonych [14, 15, 16]. W naszej opinii dowody te są niepełne i korzystają z nieprawdziwych założeń, co wyjaśnimy w dalszej części rozdziału.

W przypadku losowych macierzy kwaternionowych brakuje jakichkolwiek prac, które dowodziłyby, że mają one własność ograniczonej izometrii. Pojawiły się jednak prace, których autorzy przeprowadzają szereg symulacji numerycznych z wykorzystaniem kwaternionowych losowych macierzy gaussowskich [5, 12, 27]. Sukcesy tych eksperymentów sugerują, że mają one poszukiwaną własność.

W tym rozdziale prześledzimy rozumowanie autorów serii artykułów [14, 15, 16], dostosowując je (gdy jest to uzasadnione) do przypadku macierzy o elementach kwaternionowych, i pokażemy, w których miejscach autorzy wyciągają nieprawdziwe wnioski na temat własności ograniczonej izometrii dla gaussowskich macierzy losowych.

W dalszej części rozdziału przyjmujemy kilka standardowych oznaczeń na rozkłady (rzeczywistych) zmiennych losowych X , tak jak w [13]:

- (a) $X \sim \mathcal{N}(\mu, \sigma^2)$, jeśli X ma rozkład *normalny* (*Gaussa*) o wartości oczekiwanej μ i wariancji σ^2 , tzn. rozkład o gęstości

$$\phi_{\mu, \sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x - \mu)^2}{2\sigma^2}\right\} \quad \text{dla } x \in \mathbb{R}.$$

- (b) $X \sim \Gamma(\alpha, \beta)$, jeśli X ma rozkład *gamma* o wartości oczekiwanej $\frac{\alpha}{\beta}$ i wariancji $\frac{\alpha}{\beta^2}$, tzn. rozkład o gęstości

$$\gamma_{\alpha, \beta}(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad \text{dla } x \in \mathbb{R}_+.$$

- (c) $X \sim \chi^2(m)$, jeśli X ma rozkład χ^2 o m stopniach swobody, tzn. jest sumą kwadratów m niezależnych zmiennych losowych o rozkładzie $\mathcal{N}(0, 1)$; ponadto $\chi^2(m) = \Gamma(\frac{m}{2}, \frac{1}{2})$.

Przyjmujemy również standardowe oznaczenie $\mathbb{E}X$ jako wartość oczekiwaną zmiennej losowej X , $\text{Var} X = \mathbb{E}(X^2) - (\mathbb{E}X)^2$ jako wariancję X i $\text{Cov}(X, Y) = \mathbb{E}(XY) - (\mathbb{E}X) \cdot (\mathbb{E}Y)$ jako kowariancję zmiennych losowych X i Y .

Będziemy rozważać także wielowymiarowe zmienne losowe, tzn. *wektory losowe*. W szczególności będziemy używać uogólnienia rozkładu normalnego na przypadek wielowymiarowy. Mówimy, że (rzeczywisty) wektor losowy $\mathbf{X} = (X_1, \dots, X_n)^T$ ma *wielowymiarowy rozkład normalny*, jeśli istnieje wektor $\boldsymbol{\mu} \in \mathbb{R}^n$ i macierz odwracalna $\mathbf{A} \in \mathbb{R}^{n \times n}$ taka, że $\mathbf{X} = \mathbf{A} \cdot \mathbf{Z} + \boldsymbol{\mu}$, gdzie $\mathbf{Z} = (Z_1, \dots, Z_n)^T$ jest wektorem niezależnych zmiennych losowych o rozkładzie $\mathcal{N}(0, 1)$. Równoważnie, wektor losowy ma wielowymiarowy rozkład normalny, jeśli dowolna kombinacja liniowa jego współrzędnych ma (jednowymiarowy) rozkład normalny.

Powyższe definicje, w tym rozkładu wielowymiarowego, można wykorzystać do zdefiniowania kwaternionowych zmiennych i wektorów losowych. Zmienną losową X będziemy nazywać *kwaternionową zmienną losową*, jeśli

$$X = X_0 + X_1\mathbf{i} + X_2\mathbf{j} + X_3\mathbf{k}$$

i X_0, \dots, X_3 są rzeczywistymi zmiennymi losowymi. W dalszej części pracy będziemy rozważać zmienne X takie, że $X_i \sim \mathcal{N}(0, \frac{\sigma^2}{4})$ i X_i są parami niezależne. Takie zmienne losowe będziemy na potrzeby tej pracy nazywać *kwaternionowymi normalnymi (gaussowskimi) zmiennymi losowymi* o wariancji σ^2 i oznaczać $X \sim \mathcal{N}_{\mathbb{H}}(0, \sigma^2)$. W podobny sposób wprowadzamy definicję *kwaternionowych losowych macierzy gaussowskich*, tzn. macierzy $\Phi = (\phi_{ij}) \in \mathbb{H}^{m \times n}$ takich, że ϕ_{ij} są parami niezależnymi kwaternionowymi zmiennymi losowymi takimi, że $\phi_{ij} \sim \mathcal{N}_{\mathbb{H}}(0, \sigma^2)$.

4.1. Stałe ograniczonej izometrii

Przypomnijmy, że zgodnie z Definicją 3.2 macierz $\Phi \in \mathbb{H}^{m \times n}$ spełnia własność s -ograniczonej izometrii ze stałą $\delta_s \geq 0$, jeśli

$$(1 - \delta_s) \|\mathbf{x}\|_2^2 \leq \|\Phi\mathbf{x}\|_2^2 \leq (1 + \delta_s) \|\mathbf{x}\|_2^2$$

dla wszystkich s -rzadkich wektorów $\mathbf{x} \in \mathbb{H}^n$. Najmniejsza stała δ_s w powyższej definicji nazywana jest stałą s -ograniczonej izometrii.

Zauważmy, że w ogólności najmniejsza stała, która spełnia prawe oszacowanie nie musi być równa najmniejszej stałej spełniającej lewe oszacowanie. Wprowadźmy zatem rozróżnienie na prawą i lewą stałą s -ograniczonej izometrii, tzn. najmniejsze stałe $\delta_s^L, \delta_s^R \geq 0$ takie, że

$$(1 - \delta_s^L) \|\mathbf{x}\|_2^2 \leq \|\Phi\mathbf{x}\|_2^2 \leq (1 + \delta_s^R) \|\mathbf{x}\|_2^2$$

dla wszystkich s -rzadkich wektorów $\mathbf{x} \in \mathbb{H}^n$. Bez utraty ogólności możemy rozważać jedynie wektory \mathbf{x} jednostkowe (dzieląc stronami powyższą nierówność przez $\|\mathbf{x}\|_2^2$ i zakładając, że $\mathbf{x} \neq 0$). Oznaczmy zbiór takich wektorów przez Σ_s . Wówczas

$$1 - \delta_s^L \leq \|\Phi\mathbf{x}\|_2^2 \leq 1 + \delta_s^R$$

dla $\mathbf{x} \in \Sigma_s$ i możemy równoważnie zapisać, że

$$\delta_s^L = 1 - \min_{\mathbf{x} \in \Sigma_s} \|\Phi \mathbf{x}\|_2^2, \quad \delta_s^R = \max_{\mathbf{x} \in \Sigma_s} \|\Phi \mathbf{x}\|_2^2 - 1.$$

Zauważmy, że jeśli Φ jest macierzą losową, to δ_s^L i δ_s^R są rzeczywistymi zmiennymi losowymi. Warto nadmienić, że autorzy prac [14, 15, 16] modyfikują definicję Σ_s biorąc jedynie wektory, które zawierają *dokładnie* s niezerowych współrzędnych. Jednak zbiór Σ_s przestaje być wówczas zwarty i minimum oraz maksimum w powyższej definicji może nie być osiągalne.

Rozważania na temat zmiennych losowych δ_s^L i δ_s^R będziemy przeprowadzać etapami. W pierwszej kolejności, ustalmy pewien zbiór $\mathcal{S} \subseteq \{1, \dots, n\}$, oznaczmy $s = \#\mathcal{S}$ i wybierzmy $\mathbf{x} \in \Sigma_s$ taki, że $\text{supp } \mathbf{x} \subseteq \mathcal{S}$. Wprowadźmy oznaczenie $\mathcal{R}_{\mathcal{S}}(\mathbf{x}) = \|\Phi \mathbf{x}\|_2^2$. Jest to zmienna losowa, której rozkład podany jest w poniższym lemacie.

Lemat 4.1. *Niech $\Phi \in \mathbb{H}^{m \times n}$ będzie kwaternionową losową macierzą gaussowską, której elementy ϕ_{ij} są niezależnymi kwaternionowymi zmiennymi losowymi o rozkładzie $\mathcal{N}_{\mathbb{H}}(0, \frac{1}{m})$ (tzn. część rzeczywista i wszystkie współczynniki części urojonej każdej takiej zmiennej ϕ_{ij} są niezależnymi zmiennymi losowymi o rozkładzie $\mathcal{N}(0, \frac{1}{4m})$) i niech $\mathbf{x} \in \Sigma_s$. Wówczas zmienna losowa*

$$\mathcal{R}_{\mathcal{S}}(\mathbf{x}) = \|\Phi \mathbf{x}\|_2^2$$

ma rozkład $\Gamma(2m, 2m)$ i nie zależy on od \mathbf{x} . W szczególności $\mathbb{E} \mathcal{R}_{\mathcal{S}} = 1$ oraz $\text{Var } \mathcal{R}_{\mathcal{S}} = \frac{1}{2m}$.

Dowód. Zapiszmy macierz Φ w postaci

$$\Phi = \Phi_{\mathbf{r}} + \Phi_{\mathbf{i}}\mathbf{i} + \Phi_{\mathbf{j}}\mathbf{j} + \Phi_{\mathbf{k}}\mathbf{k}, \quad \text{gdzie } \Phi_{\mathbf{r}}, \Phi_{\mathbf{i}}, \Phi_{\mathbf{j}}, \Phi_{\mathbf{k}} \in \mathbb{R}^{m \times n},$$

i analogicznie każdy element tej macierzy w postaci

$$\phi_{ij} = \phi_{\mathbf{r},ij} + \phi_{\mathbf{i},ij}\mathbf{i} + \phi_{\mathbf{j},ij}\mathbf{j} + \phi_{\mathbf{k},ij}\mathbf{k}.$$

Podobnie sposób będziemy oznaczać podział wektora $\mathbf{x} = (x_1, \dots, x_n)^T$ na składowe odpowiadające części rzeczywistej i współczynnikiem części urojonej.

Niech $\mathbf{y} = \Phi \mathbf{x}$. Oznaczając $\mathbf{y} = (y_1, \dots, y_m)^T$ otrzymujemy

$$\begin{aligned} y_k &= \sum_{\ell=1}^n \phi_{k\ell} x_{\ell} = \sum_{\ell=1}^n (\phi_{\mathbf{r},k\ell} + \phi_{\mathbf{i},k\ell}\mathbf{i} + \phi_{\mathbf{j},k\ell}\mathbf{j} + \phi_{\mathbf{k},k\ell}\mathbf{k}) \cdot (x_{\mathbf{r},\ell} + x_{\mathbf{i},\ell}\mathbf{i} + x_{\mathbf{j},\ell}\mathbf{j} + x_{\mathbf{k},\ell}\mathbf{k}) \\ &= \sum_{\ell=1}^n (\phi_{\mathbf{r},k\ell} x_{\mathbf{r},\ell} - \phi_{\mathbf{i},k\ell} x_{\mathbf{i},\ell} - \phi_{\mathbf{j},k\ell} x_{\mathbf{j},\ell} - \phi_{\mathbf{k},k\ell} x_{\mathbf{k},\ell}) \end{aligned}$$

$$\begin{aligned}
& + \sum_{\ell=1}^n (\phi_{\mathbf{i},k\ell} x_{\mathbf{r},\ell} + \phi_{\mathbf{r},k\ell} x_{\mathbf{i},\ell} - \phi_{\mathbf{k},k\ell} x_{\mathbf{j},\ell} + \phi_{\mathbf{j},k\ell} x_{\mathbf{k},\ell}) \mathbf{i} \\
& + \sum_{\ell=1}^n (\phi_{\mathbf{j},k\ell} x_{\mathbf{r},\ell} + \phi_{\mathbf{k},k\ell} x_{\mathbf{i},\ell} + \phi_{\mathbf{r},k\ell} x_{\mathbf{j},\ell} - \phi_{\mathbf{i},k\ell} x_{\mathbf{k},\ell}) \mathbf{j} \\
& + \sum_{\ell=1}^n (\phi_{\mathbf{k},k\ell} x_{\mathbf{r},\ell} - \phi_{\mathbf{j},k\ell} x_{\mathbf{i},\ell} + \phi_{\mathbf{i},k\ell} x_{\mathbf{j},\ell} + \phi_{\mathbf{r},k\ell} x_{\mathbf{k},\ell}) \mathbf{k} \\
& =: y_{\mathbf{r},k} + y_{\mathbf{i},k} \mathbf{i} + y_{\mathbf{j},k} \mathbf{j} + y_{\mathbf{k},k} \mathbf{k}.
\end{aligned}$$

Zauważmy, że dla każdego $e \in \{\mathbf{r}, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$ i każdego $k \in \{1, \dots, m\}$ zmienne $y_{e,k}$ mają rozkład normalny (jako kombinacja liniowa zmiennych gaussowskich). Zachodzi ponadto $\mathbb{E} y_{e,k} = 0$ oraz $\text{Var } y_{e,k} = \frac{1}{4m}$, co wynika z tego, że wszystkie $\phi_{e,k\ell}$ są niezależne oraz

$$\begin{aligned}
\text{Var } y_{\mathbf{r},k} &= \sum_{\ell=1}^n (x_{\mathbf{r},\ell}^2 \text{Var } \phi_{\mathbf{r},k\ell} + x_{\mathbf{i},\ell}^2 \text{Var } \phi_{\mathbf{i},k\ell} + x_{\mathbf{j},\ell}^2 \text{Var } \phi_{\mathbf{j},k\ell} + x_{\mathbf{k},\ell}^2 \text{Var } \phi_{\mathbf{k},k\ell}) \\
&= \frac{1}{4m} \sum_{\ell=1}^n (x_{\mathbf{r},\ell}^2 + x_{\mathbf{i},\ell}^2 + x_{\mathbf{j},\ell}^2 + x_{\mathbf{k},\ell}^2) = \frac{\|\mathbf{x}\|_2^2}{4m} = \frac{1}{4m},
\end{aligned}$$

a dla pozostałych składowych postępujemy analogicznie.

Z niezależności wszystkich $\phi_{e,k\ell}$ wynika również niezależność zmiennych $y_{e,k}$ i $y_{e,\ell}$ dla $e \in \{\mathbf{r}, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$ i $k, \ell \in \{1, \dots, m\}$. Aby sprawdzić niezależność $y_{e_1,k}$ i $y_{e_2,k}$, zauważmy, że

$$\begin{aligned}
\text{Cov}(y_{\mathbf{r},k}, y_{\mathbf{i},k}) &= \mathbb{E}(y_{\mathbf{r},k} \cdot y_{\mathbf{i},k}) - \mathbb{E} y_{\mathbf{r},k} \cdot \mathbb{E} y_{\mathbf{i},k} \\
&= \mathbb{E} \left[\left(\sum_{\ell=1}^n (\phi_{\mathbf{r},k\ell} x_{\mathbf{r},\ell} - \phi_{\mathbf{i},k\ell} x_{\mathbf{i},\ell} - \phi_{\mathbf{j},k\ell} x_{\mathbf{j},\ell} - \phi_{\mathbf{k},k\ell} x_{\mathbf{k},\ell}) \right) \right. \\
&\quad \left. \cdot \left(\sum_{p=1}^n (\phi_{\mathbf{i},kp} x_{\mathbf{r},p} + \phi_{\mathbf{r},kp} x_{\mathbf{i},p} - \phi_{\mathbf{k},kp} x_{\mathbf{j},p} + \phi_{\mathbf{j},kp} x_{\mathbf{k},p}) \right) \right] \\
&= \sum_{\ell=1}^n \mathbb{E} \left((\phi_{\mathbf{r},k\ell} x_{\mathbf{r},\ell} - \phi_{\mathbf{i},k\ell} x_{\mathbf{i},\ell} - \phi_{\mathbf{j},k\ell} x_{\mathbf{j},\ell} - \phi_{\mathbf{k},k\ell} x_{\mathbf{k},\ell}) \right. \\
&\quad \left. \cdot (\phi_{\mathbf{i},k\ell} x_{\mathbf{r},\ell} + \phi_{\mathbf{r},k\ell} x_{\mathbf{i},\ell} - \phi_{\mathbf{k},k\ell} x_{\mathbf{j},\ell} + \phi_{\mathbf{j},k\ell} x_{\mathbf{k},\ell}) \right) \\
&= \sum_{\ell=1}^n (x_{\mathbf{r},\ell} x_{\mathbf{i},\ell} \mathbb{E} \phi_{\mathbf{r},k\ell}^2 - x_{\mathbf{i},\ell} x_{\mathbf{r},\ell} \mathbb{E} \phi_{\mathbf{i},k\ell}^2 - x_{\mathbf{j},\ell} x_{\mathbf{k},\ell} \mathbb{E} \phi_{\mathbf{j},k\ell}^2 + x_{\mathbf{k},\ell} x_{\mathbf{j},\ell} \mathbb{E} \phi_{\mathbf{k},k\ell}^2) \\
&= \frac{1}{4m} \sum_{\ell=1}^n (x_{\mathbf{r},\ell} x_{\mathbf{i},\ell} - x_{\mathbf{i},\ell} x_{\mathbf{r},\ell} - x_{\mathbf{j},\ell} x_{\mathbf{k},\ell} + x_{\mathbf{k},\ell} x_{\mathbf{j},\ell}) = 0,
\end{aligned}$$

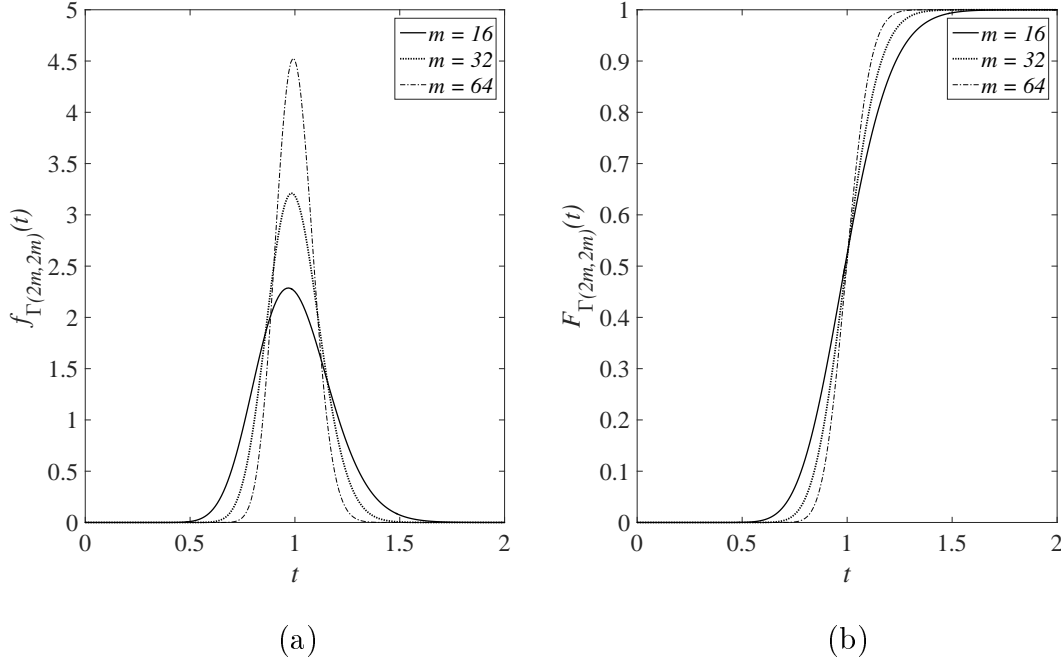
ponieważ $\mathbb{E} \phi_{e,k\ell}^2 = \text{Var } \phi_{e,k\ell} = \frac{1}{4m}$, a dla pozostałych par analogicznie. Przypomnijmy, że w przypadku zmiennych losowych o łącznym rozkładzie normalnym niezależność dwóch zmiennych jest równoważna temu, że ich kowariancja jest zero [13].

Otrzymaliśmy zatem, że wszystkie $y_{e,k}$, $e \in \{\mathbf{r}, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$ i $k \in \{1, \dots, m\}$, są niezależnymi zmiennymi losowymi o rozkładzie $\mathcal{N}(0, \frac{1}{4m})$, a zatem $\sqrt{4m} y_{e,k} \sim \mathcal{N}(0, 1)$. Wynika stąd, że

$$4m\mathcal{R}_{\mathcal{S}}(\mathbf{x}) = \left\| \sqrt{4m} \mathbf{y} \right\|_2^2 = \sum_{k=1}^m \left((\sqrt{4m} y_{\mathbf{r},k})^2 + (\sqrt{4m} y_{\mathbf{i},k})^2 + (\sqrt{4m} y_{\mathbf{j},k})^2 + (\sqrt{4m} y_{\mathbf{k},k})^2 \right)$$

jest sumą $4m$ kwadratów niezależnych zmiennych losowych o rozkładzie $\mathcal{N}(0, 1)$, a zatem ma rozkład $\chi^2(4m) = \Gamma(2m, \frac{1}{2})$. Stąd zmienna $\mathcal{R}_{\mathcal{S}}$ ma rozkład $\Gamma(2m, 2m)$, niezależny od wyboru \mathbf{x} . Zmienna ta ma wartość oczekiwaną 1 i wariancję $\frac{1}{2m}$. \square

Uwaga 4.1. Zauważmy, że rozkład zmiennej $\mathcal{R}_{\mathcal{S}}$ (jak na Rys. 4.1) nie zależy nie tylko od wyboru wektora \mathbf{x} (a więc od wyboru nośnika), ale nawet od $s = \#\mathcal{S}$.



Rys. 4.1. Gęstość (a) i dystrybuanta (b) rozkładu $\Gamma(2m, 2m)$ dla różnych wartości m .

Powyższy wynik ma swój odpowiednik w pracach [14, 15, 16], których autorzy przeprowadzają analogiczne rozważania na temat gaussowskich macierzy losowych o elementach rzeczywistych i zespolonych. Warto zauważyć, że w przypadku macierzy rzeczywistych zmienna $\mathcal{R}_{\mathcal{S}}$ ma rozkład $\Gamma(\frac{m}{2}, \frac{m}{2})$, a dla macierzy zespolonych rozkład $\Gamma(m, m)$. W przypadku macierzy kwaternionowych otrzymujemy więc mniejszą wariancję zmiennej losowej $\mathcal{R}_{\mathcal{S}}$ niż dla macierzy rzeczywistych, tj. $\frac{2}{m}$, czy zespolonych, tj. $\frac{1}{m}$.

Zauważmy, że z definicji prawej i lewej stałej s -ograniczonej izometrii wynika, że

$$1 - \delta_s^L \leq \mathcal{R}_S(\mathbf{x}) \leq 1 + \delta_s^R$$

dla dowolnego nośnika $\mathcal{S} \subseteq \{1, \dots, n\}$ i wektora $\mathbf{x} \in \Sigma_s$ takiego, że $\text{supp } \mathbf{x} \subseteq \mathcal{S}$. Daje nam to pewne oszacowania na prawdopodobieństwo, że stałe ograniczonej izometrii będą odpowiednio duże lub małe, tzn.

$$\forall_{t \geq 0} \quad \mathbb{P}(\delta_s^R \leq t) \leq \mathbb{P}(\mathcal{R}_S - 1 \leq t) = \mathbb{P}(\mathcal{R}_S \leq t + 1) = F_{\Gamma(2m, 2m)}(t + 1), \quad (4.1)$$

$$\forall_{t \geq 0} \quad \mathbb{P}(\delta_s^R > t) \geq \mathbb{P}(\mathcal{R}_S - 1 > t) = \mathbb{P}(\mathcal{R}_S > t + 1) = 1 - F_{\Gamma(2m, 2m)}(t + 1) \quad (4.2)$$

$$\forall_{t \in (0, 1)} \quad \mathbb{P}(\delta_s^L < t) \leq \mathbb{P}(1 - \mathcal{R}_S < t) = \mathbb{P}(\mathcal{R}_S > 1 - t) = 1 - F_{\Gamma(2m, 2m)}(1 - t) \quad (4.3)$$

$$\forall_{t \in (0, 1)} \quad \mathbb{P}(\delta_s^L \geq t) \geq \mathbb{P}(1 - \mathcal{R}_S \geq t) = \mathbb{P}(\mathcal{R}_S \leq 1 - t) = F_{\Gamma(2m, 2m)}(1 - t), \quad (4.4)$$

gdzie $F_{\Gamma(2m, 2m)}$ jest dystrybuantą rozkładu $\Gamma(2m, 2m)$. Potrzebujemy jednak oszacowań w przeciwną stronę.

Autorzy prac [14, 15, 16] sugerują, że rozważania na temat zmiennych losowych δ_s^R i δ_s^L można uprościć. Zajmijmy się stałą δ_s^R . Z definicji, można tę wielkość wyrazić jako

$$\delta_s^R = \max_{\mathbf{x} \in \Sigma_s} \|\Phi \mathbf{x}\|_2^2 - 1 = \max_{\mathcal{S}: \#\mathcal{S}=s} \max\{\mathcal{R}_S(\mathbf{x}): \text{supp } \mathbf{x} \subseteq \mathcal{S}, \|\mathbf{x}\|_2 = 1\} - 1,$$

gdzie $\mathcal{S} \subseteq \{1, \dots, n\}$. W [14] zasugerowano, że wystarczy zbadać rozkład zmiennej losowej

$$\Delta_s^R = \max_{\mathcal{S}: \#\mathcal{S}=s} \{\mathcal{R}_S(\mathbf{x}_S) \text{ dla pewnego } \mathbf{x}_S \text{ takiego, że } \text{supp } \mathbf{x}_S = \mathcal{S}, \|\mathbf{x}_S\|_2 = 1\} - 1,$$

tzn. pomijając maksimum po nieprzeliczalnym zbiorze wektorów \mathbf{x} i badając rozkład dla pojedynczych ustalonych wektorów \mathbf{x}_S o odpowiednim nośniku. Zgodnie z zapewnieniami autorów, słuszność takiego postępowania pozostaje wciąż do udowodnienia, pokażemy jednak że jest to nieprawda.

Zauważmy, że dla dwóch nośników \mathcal{S} i \mathcal{T} , dwie zmienne losowe $\mathcal{R}_S(\mathbf{x}_S)$ i $\mathcal{R}_T(\mathbf{x}_T)$ są zależne jeśli $\mathcal{S} \cap \mathcal{T} \neq \emptyset$, ze względu na użycie tych samych wierszy macierzy Φ w definicji $\mathcal{R}_S(\mathbf{x}_S)$ oraz $\mathcal{R}_T(\mathbf{x}_T)$.

Wyznaczenie rozkładu zmiennej Δ_s^R jest trudne, ze względu na to, że zmienne losowe, których maksimum obliczamy nie są niezależne. Zauważmy jednak, że każda ze zmiennych

$\mathcal{R}_{\mathcal{S}}(\mathbf{x}_{\mathcal{S}})$ ma ten sam rozkład $(\Gamma(2m, 2m))$ i zmiennych tych będzie dokładnie $N := \binom{n}{s}$. Możemy zatem zapisać

$$\Delta_s^R + 1 = \max\{X_i : X_i \sim \Gamma(2m, 2m), i \in \{1, \dots, N\}\},$$

a stąd

$$\mathbb{P}(\Delta_s^R \leq t) = \mathbb{P}(\max_i X_i \leq t + 1) = \mathbb{P}(X_1 \leq t + 1, \dots, X_N \leq t + 1).$$

Każda ze zmiennych X_i ma rozkład gamma, z tymi samymi parametrami, autorzy sugerują więc, nie podając jednak dowodu, że prawdziwe jest oszacowanie

$$\mathbb{P}(X_1 \leq t_1, \dots, X_N \leq t_N) \geq \prod_{i=1}^N \mathbb{P}(X_i \leq t_i), \quad (4.5)$$

tzn. że zmienne losowe o rozkładzie gamma są *dodatnio zależne* (ang. *positively dependent*). Zgodnie z naszą wiedzą, własność ta została udowodniona dla $N = 2$ przez Jensena [17] w ogólniejszej postaci, tzn.

$$\mathbb{P}(X_1 \in A, X_2 \in A) \geq \mathbb{P}(X_1 \in A) \cdot \mathbb{P}(X_2 \in A) \quad (4.6)$$

dla dowolnego mierzalnego zbioru $A \subset \mathbb{R}_+$, oraz w ubiegłym roku rozszerzona na przypadek N -wymiarowy, $N > 2$, przez Royena [20] w postaci (4.5).

Zauważmy, że daje nam to oszacowanie

$$\mathbb{P}(\Delta_s^R \leq t) \geq \prod_{i=1}^N F_{\Gamma(2m, 2m)}(t + 1) = (F_{\Gamma(2m, 2m)}(t + 1))^{\binom{n}{s}}, \quad (4.7)$$

gdzie $F_{\Gamma(2m, 2m)}$ jest dystrybuantą rozkładu $\Gamma(2m, 2m)$. Stąd otrzymujemy

$$\mathbb{P}(\Delta_s^R > t) = 1 - \mathbb{P}(\Delta_s^R \leq t) \leq 1 - (F_{\Gamma(2m, 2m)}(t + 1))^{\binom{n}{s}}. \quad (4.8)$$

Analogiczne rozważania można przeprowadzić dla zmiennej δ_s^L . Możemy wyrazić δ_s^L jako

$$\delta_s^L = 1 - \min_{\mathbf{x} \in \Sigma_s} \|\Phi \mathbf{x}\|_2^2 = 1 - \min_{\mathcal{S}: \#\mathcal{S}=s} \min\{\mathcal{R}_{\mathcal{S}}(\mathbf{x}) : \text{supp } \mathbf{x} \subseteq \mathcal{S}, \|\mathbf{x}\|_2 = 1\},$$

gdzie $\mathcal{S} \subseteq \{1, \dots, n\}$. Wprowadźmy nową zmienną losową

$$\Delta_s^L = 1 - \min_{\mathcal{S}: \#\mathcal{S}=s} \{\mathcal{R}_{\mathcal{S}}(\mathbf{x}_{\mathcal{S}}) \text{ dla pewnego } \mathbf{x}_{\mathcal{S}} \text{ takiego, że } \text{supp } \mathbf{x}_{\mathcal{S}} = \mathcal{S}, \|\mathbf{x}_{\mathcal{S}}\|_2 = 1\},$$

tzn. pomijając maksimum po nieprzeliczalnym zbiorze wektorów \mathbf{x} . Tak jak wcześniej, możemy zapisać

$$1 - \Delta_s^L = \min\{X_i : X_i \sim \Gamma(2m, 2m), i \in \{1, \dots, N\}\},$$

gdzie $N = \binom{n}{s}$. Rozpiszmy

$$\mathbb{P}(\Delta_s^L \leq t) = \mathbb{P}(1 - \Delta_s^L \geq 1 - t) = \mathbb{P}(\min_i X_i \geq 1 - t) = \mathbb{P}(X_1 \geq 1 - t, \dots, X_N \geq 1 - t).$$

Chcielibyśmy oszacować powyższe prawdopodobieństwo, tak jak w rozważaniach dotyczących Δ_s^R , przez iloczyn prawdopodobieństw $\mathbb{P}(X_i \geq 1 - t)$. Wiemy, ze wzoru (4.6), że oszacowanie to zachodzi dla $N = 2$, jednak oszacowanie

$$\mathbb{P}(X_1 \geq t_1, \dots, X_N \geq t_N) \geq \prod_{i=1}^N \mathbb{P}(X_i \geq t_i) \quad (4.9)$$

dla $N > 2$ nie zostało jak dotąd wykazane. Jeśli rzeczywiście tak jest (co zostało założone przez autorów artykułu), to prowadzi to do oszacowania

$$\mathbb{P}(\Delta_s^L \leq t) \geq \prod_{i=1}^N (1 - F_{\Gamma(2m, 2m)}(1 - t)) = (1 - F_{\Gamma(2m, 2m)}(1 - t))^N, \quad (4.10)$$

a w konsekwencji

$$\mathbb{P}(\Delta_s^L > t) = 1 - \mathbb{P}(\Delta_s^L \leq t) \leq 1 - (1 - F_{\Gamma(2m, 2m)}(1 - t))^{\binom{n}{s}}. \quad (4.11)$$

Autorzy artykułów przeprowadzają powyższe rozumowanie dla macierzy Φ o elementach rzeczywistych i zespolonych, ale różni się ono od rozumowania dla macierzy kwaternionowych jedynie parametrami rozkładu gamma. Nie wpływa to na poprawność samego rozumowania.

Według autorów [14, 15, 16], nierówności (4.8) i (4.11) wystarczą, by uzyskać odpowiednie oszacowania na prawdopodobieństwa $\mathbb{P}(\delta_s^R > t)$ i $\mathbb{P}(\delta_s^L > t)$. Zauważmy jednak, że

$$\delta_s^R \geq \Delta_s^R \quad \text{oraz} \quad \delta_s^L \geq \Delta_s^L,$$

co prowadzi do

$$\mathbb{P}(\delta_s^R > t) \geq \mathbb{P}(\Delta_s^R > t) \quad \text{oraz} \quad \mathbb{P}(\delta_s^L > t) \geq \mathbb{P}(\Delta_s^L > t),$$

tzn. mamy oszacowania w niewłaściwą stronę. Wnioski na temat tego, czy losowe macierze gaussowskie mają własność ograniczonej izometrii, wyciągane na podstawie oszacowań (4.8) i (4.11) są wobec tego nieuzasadnione.

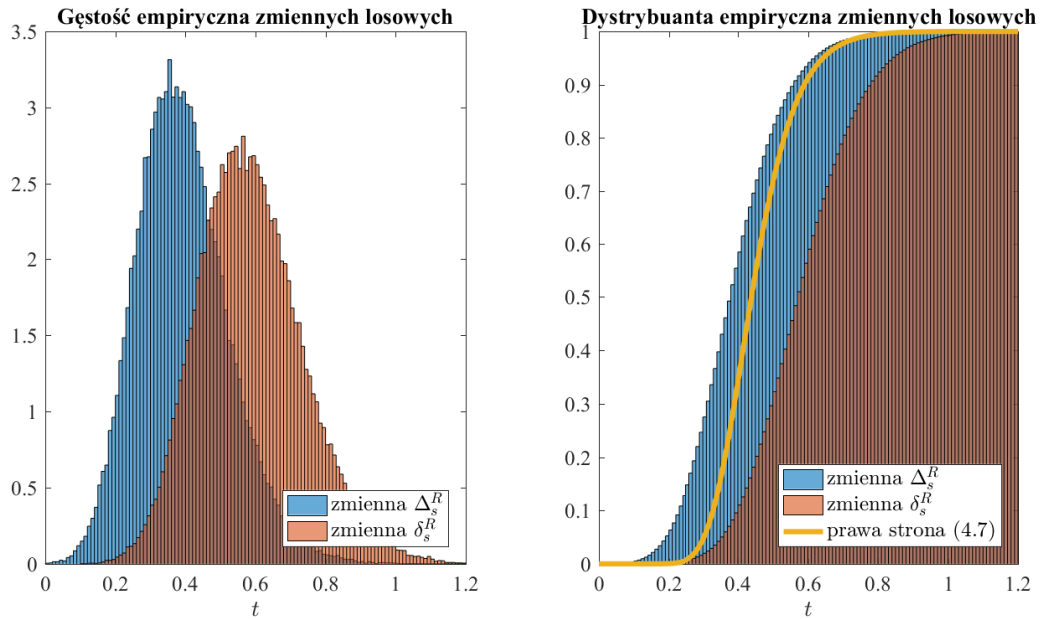
4.2. Symulacje numeryczne

Aby potwierdzić naszą tezę, że autorzy publikacji [14, 15, 16] wyciągnęli niepoprawne wnioski, przeprowadziliśmy serię symulacji numerycznych. Zarówno dla przypadku macierzy rzeczywistych jak i kwaternionowych, wyznaczyliśmy dystrybuanty i gęstości empiryczne zmiennych losowych δ_s^R , δ_s^L , Δ_s^R i Δ_s^L . Eksperyment przeprowadziliśmy w środowisku MATLAB R2016a, na komputerze PC z procesorem Intel(R) Core(TM) i7-4790 CPU (3.60GHz) i 16GB RAM, pracującym pod systemem operacyjnym Microsoft Windows 10 Pro.

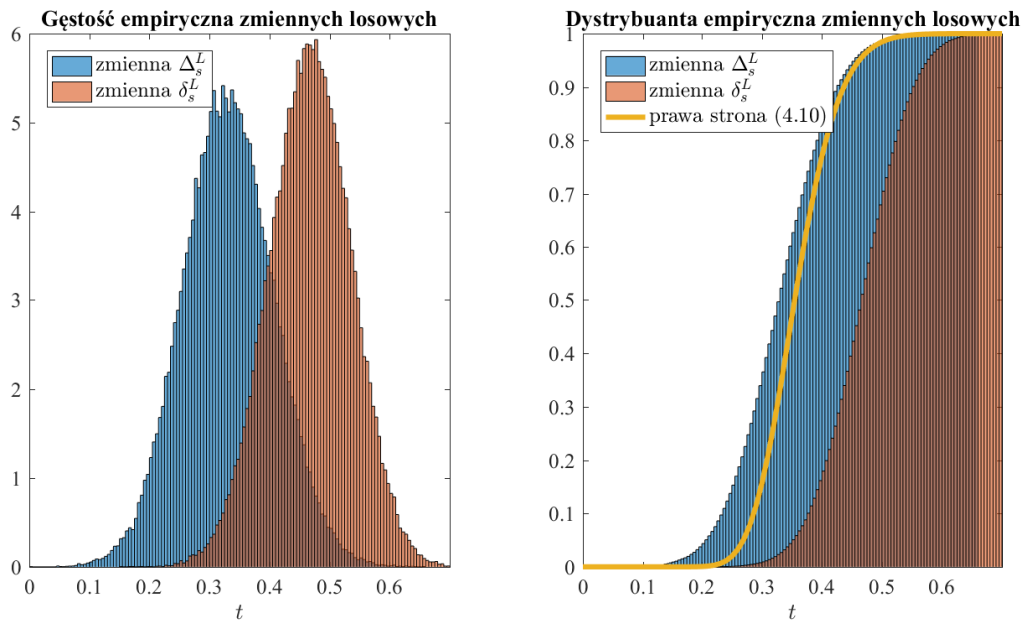
Ze względu na możliwości obliczeniowe, symulację wykonaliśmy dla macierzy $\Phi \in \mathbb{K}^{m \times n}$, gdzie $\mathbb{K} = \mathbb{R}$ i $\mathbb{K} = \mathbb{H}$, $m = 64$, $n = 8$ oraz dla s -rzadkich wektorów $\mathbf{x} \in \mathbb{K}^n$, gdzie $s = 5$. Wyznaczając dystrybuantę empiryczną zmiennych losowych δ_s^R i δ_s^L , dla każdego z $\binom{n}{s} = 56$ różnych nośników $\mathcal{S}_k \subset \{1, \dots, n\}$, takich, że $\#\mathcal{S}_k = s$, wylosowano 10^3 wektorów o normie 1, których nośniki są równe \mathcal{S}_k . Dla danej realizacji zmiennej losowej Φ maksimum jest więc brane po zbiorze $(56 \cdot 10^3)$ -elementowym. W przypadku zmiennych losowych Δ_s^R i Δ_s^L dla każdego nośnika losowany jest jeden wektor o nośniku \mathcal{S}_k i normie 1. Próba losowa składała się z 10^5 realizacji zmiennej losowej Φ .

Wyniki symulacji dla macierzy i wektorów o współrzędnych rzeczywistych zamieszczono na Rys. 4.2 i 4.3, a dla przypadku kwaternionowego na Rys. 4.4 i 4.5. Na wykresy dystrybuant empirycznych naniesiono również prawe strony oszacowań (4.7) i (4.10) (w przypadku macierzy i wektorów rzeczywistych w tych oszacowaniach występują dystrybuanty rozkładu $\Gamma(\frac{m}{2}, \frac{m}{2})$). Zauważmy, że otrzymane wyniki potwierdzają oszacowania (4.7) i (4.10) uzyskane przez autorów [14, 15, 16], pokazują jednak również, że nieuzasadnione jest wyciąganie na tej podstawie wniosków na temat zmiennych δ_s^R i δ_s^L .

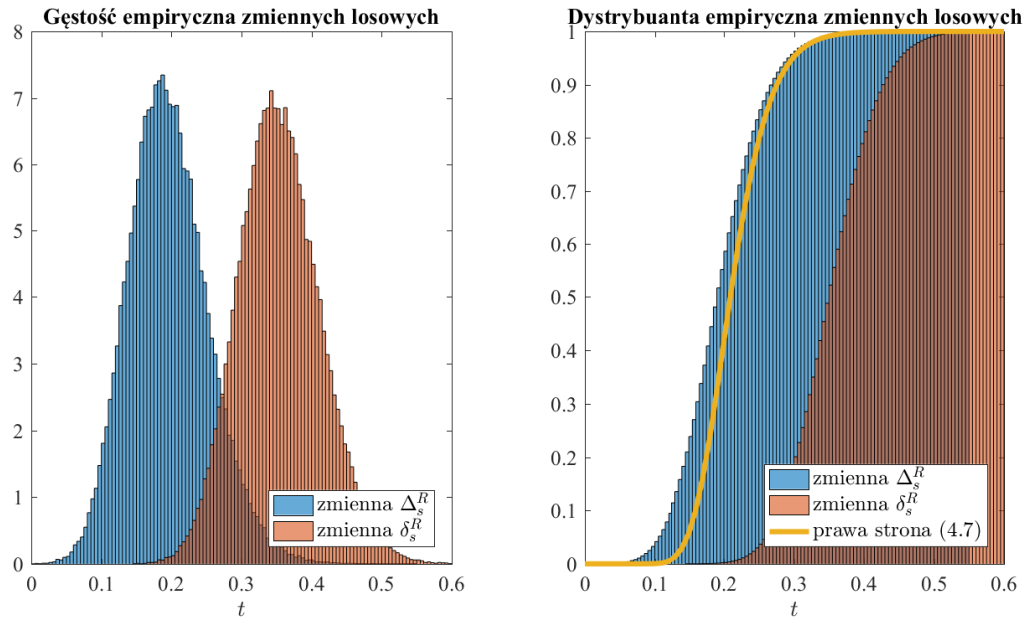
Uzyskane wyniki sugerują, że z dużym prawdopodobieństwem kwaternionowe losowe macierze gaussowskie również mają własność ograniczonej izometrii. Co więcej, można przypuszczać, że dla ustalonej pary (m, n) losowa macierz gaussowska $\Phi \in \mathbb{H}^{m \times n}$ ma statystycznie niższą stałą s -ograniczonej izometrii niż macierz $\Phi \in \mathbb{R}^{m \times n}$. Podejrzewamy, że wynika to z tego, że zmienna $\mathcal{R}_{\mathcal{S}}(\mathbf{x})$ ma w przypadku kwaternionowym mniejszą wariancję niż w przypadku rzeczywistym. Na tę chwilę nie jesteśmy jeszcze w stanie podać ścisłego dowodu, który potwierdziłby naszą tezę. Aktualnie prowadzimy prace nad tym zagadnieniem.



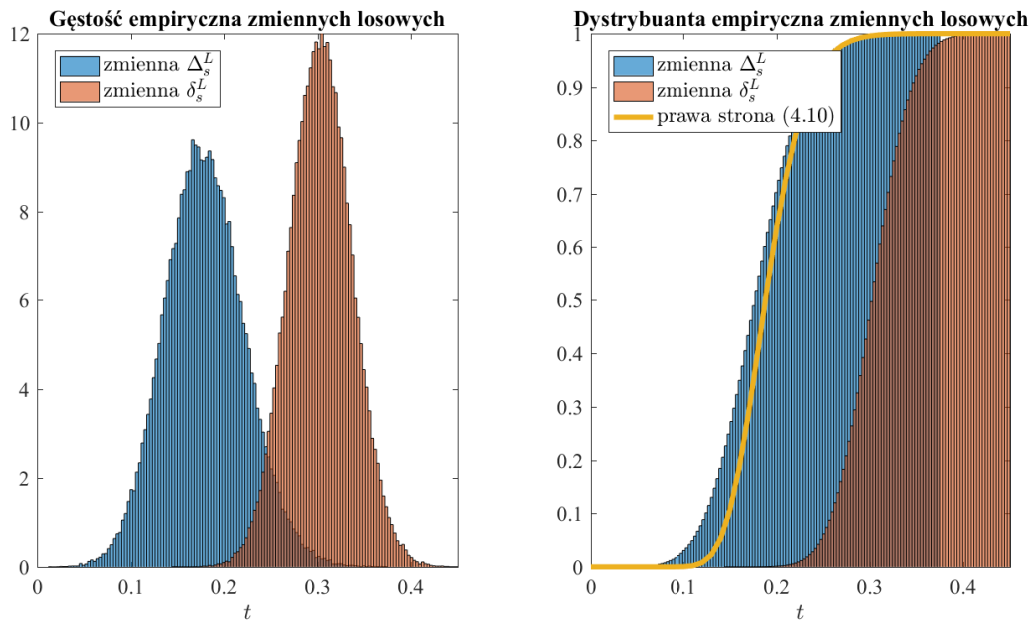
Rys. 4.2. Rozkłady empiryczne zmiennych losowych δ_s^R i Δ_s^R dla gaussowskich macierzy losowych $\Phi \in \mathbb{R}^{m \times n}$, $n = 8$, $m = 64$, $s = 5$.



Rys. 4.3. Rozkłady empiryczne zmiennych losowych δ_s^L i Δ_s^L dla gaussowskich macierzy losowych $\Phi \in \mathbb{R}^{m \times n}$, $n = 8$, $m = 64$, $s = 5$.



Rys. 4.4. Rozkłady empiryczne zmiennych losowych δ_s^R i Δ_s^R dla gaussowskich macierzy losowych $\Phi \in \mathbb{H}^{m \times n}$, $n = 8$, $m = 64$, $s = 5$.



Rys. 4.5. Rozkłady empiryczne zmiennych losowych δ_s^L i Δ_s^L dla gaussowskich macierzy losowych $\Phi \in \mathbb{H}^{m \times n}$, $n = 8$, $m = 64$, $s = 5$.

Rozdział 5

Zastosowania oszczędnego próbkowania

Wiemy już, jakie warunki muszą spełniać macierze pomiarowe, by możliwa była rekonstrukcja rzadkich sygnałów z niewielkiej ilości pomiarów, zarówno w przypadku kwaternionowym jak i rzeczywistym. Wiemy również, że rzeczywiste i zespolone losowe macierze gaussowskie te warunki spełniają (z dużym prawdopodobieństwem). Rozważania przeprowadzone w poprzednim rozdziale sugerują, że ma to miejsce również w przypadku kwaternionowym.

Rozważania teoretyczne z poprzednich rozdziałów zilustrujemy kilkoma przykładami numerycznymi. Inspiracją do przeprowadzenia tych eksperymentów był artykuł [27], w którym autorzy przeprowadzili numeryczną symulację odzyskiwania kwaternionowych sygnałów rzadkich w wyniku minimalizacji normy ℓ_1 . Autorzy artykułu nie poparli jednak swoich rezultatów żadnymi wynikami teoretycznymi. Mając już pewne przesłanki teoretyczne, powtórzmy to doświadczenie.

5.1. Kilka eksperymentów numerycznych

Przypomnijmy, że rozważamy problem minimalizacji normy ℓ_1 , tzn.

$$\arg \min_{\mathbf{z} \in \mathbb{H}^n} \|\mathbf{z}\|_1 \quad \text{pod warunkiem} \quad \Phi \mathbf{z} = \mathbf{y}, \quad (5.1)$$

gdzie $\mathbf{y} = \Phi \mathbf{x} \in \mathbb{H}^m$, $\Phi \in \mathbb{H}^{m \times n}$ i $\mathbf{x} \in \mathbb{H}^n$. Wyrazimy ten problem w języku programowania stożkowego drugiego stopnia (ang. *second-order cone programming* – SOCP).

Zgodnie z rozważaniami przytoczonymi w [27], problem (5.1) jest równoważny

$$\arg \min_{t \in \mathbb{R}_+} t \quad \text{pod warunkiem} \quad \mathbf{y} = \mathbf{\Phi} \mathbf{z}, \quad \|\mathbf{z}\|_1 \leq t. \quad (5.2)$$

Zapiszmy t w postaci $t = \sum_{k=1}^n t_k$, gdzie $t_k \in \mathbb{R}_+$, oraz podobnie jak w poprzednich rozdziałach

$$\mathbf{y} = \mathbf{y}_r + \mathbf{y}_i \mathbf{i} + \mathbf{y}_j \mathbf{j} + \mathbf{y}_k \mathbf{k}, \quad \mathbf{z} = \mathbf{z}_r + \mathbf{z}_i \mathbf{i} + \mathbf{z}_j \mathbf{j} + \mathbf{z}_k \mathbf{k},$$

gdzie $\mathbf{y}_r, \mathbf{y}_i, \mathbf{y}_j, \mathbf{y}_k \in \mathbb{R}^m$, $\mathbf{z}_r, \mathbf{z}_i, \mathbf{z}_j, \mathbf{z}_k \in \mathbb{R}^n$. Będziemy oznaczać

$$\mathbf{z}_r = (z_{r,1}, \dots, z_{r,n})^T, \quad \mathbf{z}_i = (z_{i,1}, \dots, z_{i,n})^T, \quad \mathbf{z}_j = (z_{j,1}, \dots, z_{j,n})^T, \quad \mathbf{z}_k = (z_{k,1}, \dots, z_{k,n})^T.$$

Ponadto, niech $\phi_k \in \mathbb{H}^m$ będzie k -tą kolumną macierzy $\mathbf{\Phi}$, $k \in \{1, \dots, n\}$, i tak jak wcześniej

$$\phi_k = \phi_{r,k} + \phi_{i,k} \mathbf{i} + \phi_{j,k} \mathbf{j} + \phi_{k,k} \mathbf{k},$$

gdzie $\phi_{r,k}, \phi_{i,k}, \phi_{j,k}, \phi_{k,k} \in \mathbb{R}^m$. Zauważmy, że drugie ograniczenie w (5.2) możemy zapisać jako

$$\|(z_{r,k}, z_{i,k}, z_{j,k}, z_{k,k})^T\|_2 \leq t_k \quad \text{dla } k \in \{1, \dots, n\},$$

a to pozwala nam zapisać problem (5.2) w postaci problemu minimalizacji ℓ_1 dla wektorów rzeczywistych:

$$\arg \min_{\tilde{\mathbf{z}} \in \mathbb{R}^n} \mathbf{c}^T \tilde{\mathbf{z}} \quad \text{pod warunkiem} \quad \tilde{\mathbf{y}} = \tilde{\mathbf{\Phi}} \tilde{\mathbf{z}} \quad (5.3)$$

$$\text{oraz} \quad \|(z_{r,k}, z_{i,k}, z_{j,k}, z_{k,k})^T\|_2 \leq t_k \quad \text{dla } k \in \{1, \dots, n\},$$

gdzie

$$\tilde{\mathbf{z}} = (t_1, z_{r,1}, z_{i,1}, z_{j,1}, z_{k,1}, \dots, t_n, z_{r,n}, z_{i,n}, z_{j,n}, z_{k,n})^T \in \mathbb{R}^{5n}, \quad (5.4)$$

$$\mathbf{c} = (1, 0, 0, 0, 0, \dots, 1, 0, 0, 0, 0)^T \in \mathbb{R}^{5n}, \quad (5.5)$$

$$\tilde{\mathbf{y}} = (\mathbf{y}_r^T, \mathbf{y}_i^T, \mathbf{y}_j^T, \mathbf{y}_k^T)^T \in \mathbb{R}^{4m}, \quad (5.6)$$

$$\tilde{\mathbf{\Phi}} = \begin{pmatrix} \mathbf{0} & \phi_{r,1} & -\phi_{i,1} & -\phi_{j,1} & -\phi_{k,1} & \dots & \mathbf{0} & \phi_{r,n} & -\phi_{i,n} & -\phi_{j,n} & -\phi_{k,n} \\ \mathbf{0} & \phi_{i,1} & \phi_{r,1} & -\phi_{k,1} & \phi_{j,1} & \dots & \mathbf{0} & \phi_{i,n} & \phi_{r,n} & -\phi_{k,n} & \phi_{j,n} \\ \mathbf{0} & \phi_{j,1} & \phi_{k,1} & \phi_{r,1} & -\phi_{i,1} & \dots & \mathbf{0} & \phi_{j,n} & \phi_{k,n} & \phi_{r,n} & -\phi_{i,n} \\ \mathbf{0} & \phi_{k,1} & -\phi_{j,1} & \phi_{i,1} & \phi_{r,1} & \dots & \mathbf{0} & \phi_{k,n} & -\phi_{j,n} & \phi_{i,n} & \phi_{r,n} \end{pmatrix} \in \mathbb{R}^{4m \times 5n}. \quad (5.7)$$

Powyższe sformułowanie jest standardową formą problemu SOCP i rozwiążemy je korzystając z pakietu SeDuMi do środowiska MATLAB [23]. Rozwiązanie $\tilde{\mathbf{x}}^\#$ problemu (5.3) można wówczas łatwo zamienić na wektor kwaternionowy $\mathbf{x}^\#$ korzystając z (5.4).

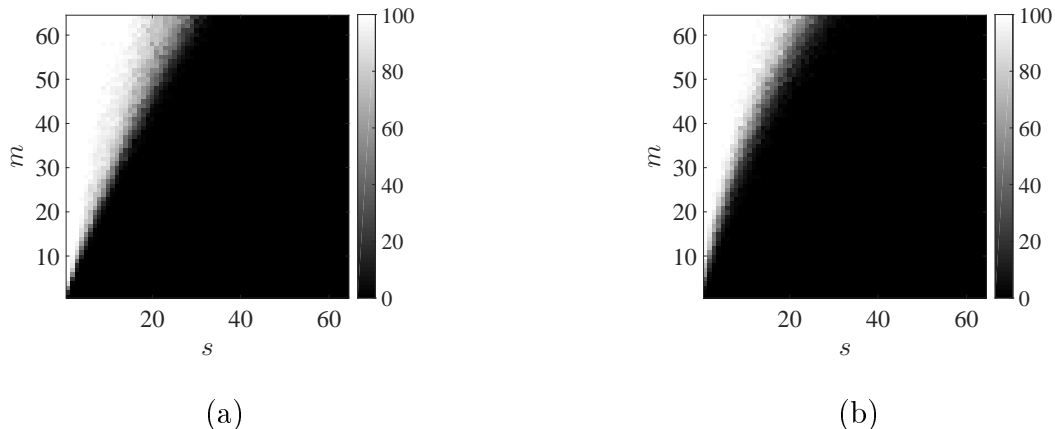
Program został wykonany w środowisku MATLAB R2016a, z wykorzystaniem pakietu SeDuMi 1.3, na komputerze PC z procesorem Intel(R) Core(TM) i7-4790 CPU (3.60GHz) i 16GB RAM, pracującym pod systemem operacyjnym Microsoft Windows 10 Pro.

Zaimplementowano następujący algorytm:

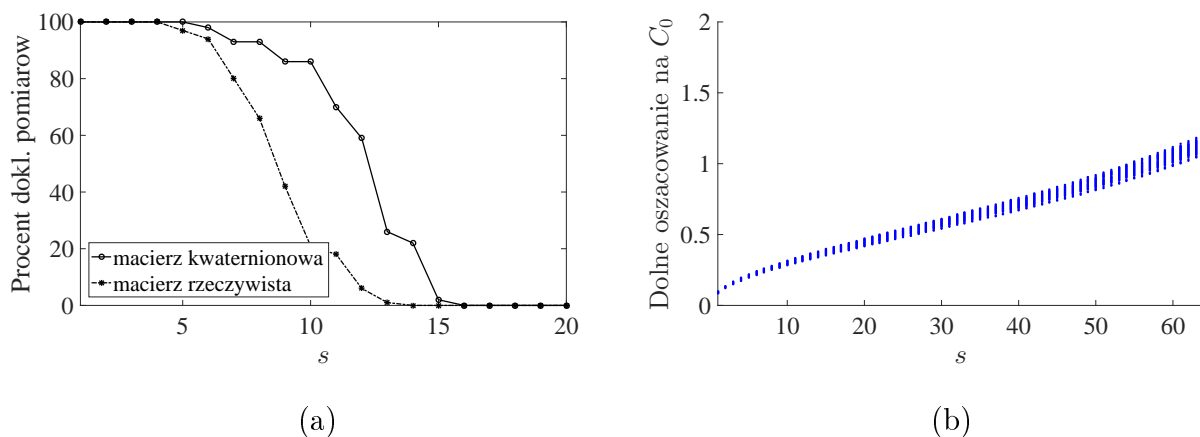
- 1° ustal stałe $n = 128$ (długość wektora \mathbf{x}) i m (liczba pomiarów, tzn. długość wektora \mathbf{y}) i wygeneruj macierz pomiarową $\Phi \in \mathbb{H}^{m \times n}$, której elementy są kwaternionami wygenerowanymi z niezależnych rozkładów normalnych $\mathcal{N}_{\mathbb{H}}(0, \frac{1}{m})$;
- 2° wybierz rzadkość sygnału $s \leq \frac{n}{2}$ i nośnik $\mathcal{S} \subseteq \{1, \dots, n\}$ taki, że $\#\mathcal{S} = s$ (losowo), oraz wygeneruj wektor $\mathbf{x} \in \mathbb{H}^n$ taki, że $\text{supp } \mathbf{x} = \mathcal{S}$, którego współrzędne są kwaternionami wygenerowanymi z niezależnych rozkładów normalnych $\mathcal{N}_{\mathbb{H}}(0, 1)$;
- 3° wyznacz $\mathbf{y} = \Phi \mathbf{x} \in \mathbb{H}^m$;
- 4° wyznacz wektory $\tilde{\mathbf{y}}, \mathbf{c}$ i macierz $\tilde{\Phi}$ jak w (5.4)–(5.7);
- 5° wywołaj program SeDuMi do rozwiązania problemu SOCP sformułowanego w (5.3) i wyznacz rozwiązanie $\tilde{\mathbf{x}}^\#$;
- 6° wyznacz wektor $\mathbf{x}^\#$ i oblicz błąd rekonstrukcji (w sensie normy ℓ_2), tzn. $\|\mathbf{x}^\# - \mathbf{x}\|_2$;
- 7° dla każdej pary (m, s) wykonaj 100 eksperymentów i zapisz błędy każdej rekonstrukcji oraz liczbę dokładnych rekonstrukcji (uznajemy, że rekonstrukcja jest dokładna, jeśli $\|\mathbf{x}^\# - \mathbf{x}\|_2 \leq 10^{-7}$).

Na Rys. 5.1(a) przedstawiono wyniki eksperymentu. Wykres przedstawia procent dokładnej rekonstrukcji w zależności od liczby pomiarów m oraz rzadkości s wektora. Widać, że odpowiedni dobór liczby pomiaru do rzadkości wektora gwarantuje niemal dokładne odtworzenie wyjściowego wektora – dla $m = 32$ i $s \leq \frac{m}{4} = 8$ odzyskujemy ponad 90% wektorów, tak samo dla $m = 64$ i $s \leq 18$. Eksperyment ten przeprowadziliśmy również dla macierzy $\Phi \in \mathbb{R}^{m \times n}$ i wektorów $\mathbf{x} \in \mathbb{R}^n$. Na Rys. 5.1(b) zaprezentowano wyniki dla tego przypadku. Warto zauważyć, że procent dokładnej rekonstrukcji dla danej pary (m, s) jest niższy niż w przypadku kwaternionowym (co porównano na Rys. 5.2(a)) – dla $m = 32$ i $s = 8$ odzyskujemy zaledwie 65% wektorów. Potwierdza to przypuszczenia z poprzedniego rozdziału,

tn. że dla danej pary (m, n) kwaternionowe losowe macierze gaussowskie mają z dużym prawdopodobieństwem niższą stałą s -ograniczonej izometrii niż losowe macierze gaussowskie o elementach rzeczywistych.



Rys. 5.1. Wyniki eksperymentu numerycznego dla $n = 128$ i różnych wartości m i s . Poziom szarości odpowiada procentowi dokładnych pomiarów. (a) $\Phi \in \mathbb{H}^{m \times n}$, (b) $\Phi \in \mathbb{R}^{m \times n}$.



Rys. 5.2. (a) Porównanie wyników eksperymentu numerycznego dla $n = 128$, $m = 32$ i różnych wartości s . (b) Dolne oszacowanie stałej C_0 w Twierdzeniu 3.1 uzyskane z nierówności (3.21) dla $n = 128$ i $m = 32$.

Uzyskane wyniki pozwoliły nam również na uzyskanie oszacowań jak we Wniosku 3.1 dla wektorów, które niekoniecznie są rzadkie. Z nierówności (3.21) uzyskujemy

$$\sqrt{s} \frac{\|\mathbf{x}^\# - \mathbf{x}\|_2}{\|\mathbf{x} - \mathbf{x}_{|s}\|_1} \leq C_0, \quad (5.8)$$

przeprowadzając zatem eksperyment, uzyskujemy empiryczne dolne oszacowanie na wartość stałej C_0 dla różnych wartości m i s . Wynik tego eksperymentu dla $m = 32$ przedstawiono na Rys. 5.2(b). Zgodnie z oczekiwaniami, zależność od s dla ustalonego m jest monotoniczna (co wynika z monotoniczności stałych s -ograniczonej izometrii).

5.2. Perspektywy zastosowań w praktyce

Teoria oszczędnego próbkowania znalazła liczne zastosowania w praktyce, zarówno w przypadku rzeczywistym jak i zespolonym. Jednym z zastosowań budzących największe nadzieje, i który jednocześnie był pierwotną inspiracją do powstania tej teorii, jest tomografia rezonansu magnetycznego [24]. Macierzą pomiarową, która w naturalny sposób powstaje w wyniku zachodzących w tym procesie zjawisk fizycznych, jest macierz współczynników dyskretnej transformacji Fouriera

$$\Phi = \begin{pmatrix} \omega_n^{0 \cdot 0} & \omega_n^{0 \cdot 1} & \dots & \omega_n^{0 \cdot (n-1)} \\ \omega_n^{1 \cdot 0} & \omega_n^{1 \cdot 1} & \dots & \omega_n^{1 \cdot (n-1)} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_n^{(n-1) \cdot 0} & \omega_n^{(n-1) \cdot 1} & \dots & \omega_n^{(n-1) \cdot (n-1)} \end{pmatrix} \in \mathbb{C}^{n \times n}$$

gdzie $\omega_n = \exp\{-\frac{2\pi i}{n}\}$. Macierz powstała z macierzy Φ poprzez losowe wybranie m wierszy spełnia własność ograniczonej izometrii [9]. Można więc przypuszczać, że kwaternionowa dyskretna transformacja Fouriera, w której $\omega_n = \exp\{-\frac{2\pi \mu}{n}\}$ dla pewnego $\mu \in \mathbb{H}$ takiego, że $\text{Re } \mu = 0$ i $\|\mu\| = 1$, również będzie miała dobre własności. Dotychczas nie powstały jednak żadne prace na ten temat.

Prace nad kwaternionową teorią oszczędnego próbkowania wydają się mieć duże znaczenie praktyczne ze względu na wspomniane w Rozdziale 1 zastosowania kwaternionów. Dotychczasowe zastosowania w przetwarzaniu i akwizycji obrazów dotyczyły jedynie obrazów monochromatycznych, które daje się przedstawić w postaci wektorów o współczynnikach rzeczywistych. Użycie kwaternionów pozwoli rozszerzenie zastosowań do przypadku obrazów kolorowych, zawierających trzy składowe kolorowe, które można reprezentować za pomocą wektorów o współrzędnych kwaternionowych.

Rozdział 6

Podsumowanie i wnioski

Definicje i twierdzenia, jakie zostały przytoczone w niniejszej pracy, stanowią fundament nowej gałęzi teorii oszczędnego próbkowania, tzn. dla przypadku kwaternionowego. Są one naturalnym uogólnieniem znanej teorii dla wektorów i macierzy o współrzędnych rzeczywistych i zespolonych, która powstała około 10 lat temu (lata 2004–2006). Pojęcia takie jak rzadkość sygnału czy własność ograniczonej izometrii pozostają takie same, jednak ich analiza wymaga pewnych dodatkowych narzędzi algebraicznych, które pozwalają na radzenie sobie z trudnościami wynikającymi z nieprzemienności mnożenia kwaternionów.

Teoria oszczędnego próbkowania wciąż budzi duże zainteresowanie naukowców na całym świecie zajmujących się różnymi dziedzinami, zarówno matematyki jak i nauk inżynierskich. Powstają wciąż nowe publikacje z tej tematyki (wyszukiwarka bazy ScienceDirect zwraca ok. 2400 publikacji, które są lub zostały przyjęte do publikacji w tym roku), a na konferencjach naukowych organizowane są oddzielne panele poświęcone wyłącznie oszczędnemu próbkowaniu. Interesującym przykładem jest konferencja *International Matheon Conference on Compressed Sensing and its Applications*, która odbywa się w Berlinie, i która gromadzi zarówno matematyków jak i inżynierów. Podczas drugiej edycji tej konferencji w grudniu 2015 roku została zaprezentowana część materiału ujętego w tej pracy [3].

Inspiracją do zajęcia się tą gałęzią teorii oszczędnego próbkowania były publikacje, w których autorzy przeprowadzają eksperymenty numeryczne w algebrze kwaternionów [27]. Brakuje w nich jednak zaplecza teoretycznego, a wszelkie wnioski wyciągane są na podstawie symulacji. Według naszej obecnej wiedzy, do tej pory nie powstały publikacje, które uzasad-

niają te eksperymenty z punktu widzenia matematycznego, teoretycznego. Ta praca stanowi pierwszy krok w kierunku uzupełnienia tych braków.

Niniejsza praca stanowi punkt wyjścia do dalszych rozważań, zarówno teoretycznych jak i eksperymentalnych, na temat oszczędnego próbkowania w algebrze kwaternionów. Wciąż pozostaje otwarta kwestia istnienia macierzy kwaternionowych spełniających własność ograniczonej izometrii, ale dalszej analizie wymagają także optymalne warunki na stałe ograniczonej izometrii gwarantujące dokładną (lub stabilną) rekonstrukcję wektorów. Oddzielną kwestią pozostaje również opracowanie algorytmów szybkiej rekonstrukcji wektorów kwaternionowych w wyniku minimalizacji normy ℓ_1 oraz wdrożenie praktycznych zastosowań tej teorii. Współpraca zespołów z Wydziału Matematyki i Nauk Informacyjnych PW oraz Wydziału Elektroniki i Technik Informacyjnych PW budzi nadzieję na osiągnięcie ciekawych rezultatów, które przyczynią się do dalszego rozwoju tej dziedziny.

Bibliografia

- [1] Andreis D., Canuto E. S., Orbit dynamics and kinematics with full quaternions, *Proc. of the American Control Conference, Boston, Massachusetts*, (2006) 3660–3665.
- [2] Badeńska A., Błaszczuk Ł., Compressed sensing for real measurements of quaternion signals, (2015) (w recenzji), *Dostępne 24.05.2016*: <http://www.mini.pw.edu.pl/~badenska/papers/CS-Rmatrix-Hvectors-submission.pdf>.
- [3] Badeńska A., Błaszczuk Ł., Compressed sensing for quaternion signals. 2. *International Matheon Conference on Compressed Sensing and its Applications, Berlin, Niemcy*, (2015) (poster).
- [4] Baraniuk R., Davenport M., DeVore R., Wakin M., A Simple Proof of the Restricted Isometry Property for Random Matrices, *Constr. Approx.*, 28 (2008), no. 3, 253–263.
- [5] Barthelemy Q., Larue A., Mars J., Color Sparse Representations for Image Processing: Review, Models, and Prospects, *IEEE Trans Image Process*, (2015), 1–12.
- [6] Candès E.J., The restricted isometry property and its implications for compressed sensing. *C R Acad Sci Paris, I* (2008), no. 346, 589–92.
- [7] Candès E.J., Romberg J., Tao T., Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, *IEEE Trans Inform Theory*, 52 (2006), no. 2, 489–509.
- [8] Candès E.J., Tao T., Decoding by linear programming, *IEEE Trans Inform Theory*, 51 (2005), no. 12, 4203–15.

- [9] Candès E.J., Tao T., Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE Trans Inform Theory*, 52 (2006), no. 12, 5406–25.
- [10] Foucart S., Rauhut H., *A mathematical introduction to compressive sensing*, Birkhäuser /Springer, Nowy Jork (2013).
- [11] Gao C., Zhou J., Lang F., Pu Q., Liu C., Novel Approach to Edge Detection of Color Image Based on Quaternion Fractional Directional Differentiation, *Advances in Automation and Robotics*, 1 (2012), 163–170.
- [12] Hawes M. B., Liu W., A Quaternion-Valued Reweighted Minimisation Approach to Sparse Vector Sensor Array Design, *Proceedings of the 19th International Conference on Digital Signal Processing*, (2014), 426–430.
- [13] Jakubowski J., Sztencel R., *Wstęp do teorii prawdopodobieństwa*, Script, Warszawa (2010).
- [14] James O., Lee H.-N., Restricted Isometry Random Variables: Probability Distributions, RIC Prediction and Phase Transition Analysis for Gaussian Encoders, (2015) (w recenzji), arXiv:1410.1956v4.
- [15] James O., On the RIP of Real and Complex Gaussian Sensing Matrices via RIV framework in Sparse Signal Recovery Analysis, *IEEE sponsored 2nd International Conference on Electronics and Communications System* (2015).
- [16] James O., Revisiting the RIP of Real and Complex Gaussian Sensing Matrices Through RIV Framework, *Wireless Pers Commun*, 87 (2016), no. 2, 513–526.
- [17] Jensen D. R., A Note on Positive Dependence and the Structure of Bivariate Distributions, *SIAM J. Appl. Math*, 20 (1971), no. 4, 749–753.
- [18] Pei S.-C., Chang J.-H., Ding J.-J., Color pattern recognition by quaternion correlation, *IEEE Int. Conf. Image Process., Thessaloniki, Greece*, (2010), 894–897.
- [19] Rodman, L., *Topics in Quaternion Algebra*, Princeton University Press, (2014).

-
- [20] Royen T., Some probability inequalities for multivariate gamma and normal distributions, (2015) (w recenzji), arXiv:1507.00528v1.
- [21] Sangwine S. J., Fourier transforms of colour images using quaternion or hypercomplex number. *Electron. Lett*, 32 (1996), no. 21, 1979–1980.
- [22] Sangwine S. J., Ell T. A., Colour image filters based on hypercomplex convolution, *IEEE Proc. Vision, Image and Signal Processing*, 49 (2000), no. 21, 89–93.
- [23] Sturm J. F., Using SeDuMi 1.02, A Matlab toolbox for optimization over symmetric cones, *Optimization Methods and Software*, 11 (1999) no. 1-4, 625–653.
- [24] Vincenti G., Preliminary experiences with compressed sensing multi-slice cine acquisitions for the assessment of left ventricular function: CV_sparse WIP, *MAGNETOM Flash*, 1 (2014), 18–26.
- [25] van der Waerden B. L., Hamilton’s Discovery of Quaternions, *Math. Magazine*, 49 (1976), no. 5, 227–236.
- [26] Witten B., Shragge J., Quaternion-based Signal Processing, *Stanford Exploration Project, New Orleans Annual Meeting* (2006), 2862–2866.
- [27] Wu J., Zhang X., Wang. X., Senhadji L., Shu H., L_1 -norm minimization for quaternion signals, *Journal of Southeast University (English Edition)*, 29 (2012), no. 1.
- [28] Xu K., Wang J., Shim B., The RIP for Random Matrices with Complex Gaussian Entries, *Future Information Technology* (2013), 13–19.
- [29] Zhang F., Quaternions and Matrices of Quaternions, *Linear Algebra Appl*, 251 (1997), 21–57.

Warszawa, dnia

Oświadczenie

Oświadczam, że pracę dyplomową pod tytułem: "Metody oszczędnego próbkowania w algebrze kwaternionów", której promotorem jest dr Agnieszka Badeńska, wykonałem samodzielnie, co poświadczam własnoręcznym podpisem.

.....