Kwaterniony i obroty

Antoni Kościelski

1 Iloczyn skalarny i długość wektora

Iloczynem skalarnym wektorów $\vec{x} = (x_1, \ldots, x_n)$ i $\vec{y} = (y_1, \ldots, y_n)$ nazywamy liczbę

$$\vec{x}\vec{y} = \sum_{i=1}^{n} x_i y_i,$$

zaś liczbę

$$|\,\vec{x}\,| = \sqrt{\vec{x}\vec{x}} = \sqrt{\vec{x}^2}$$

nazywamy długością wektora $\vec{x}$.

Lemat 1.1 *Dla dowolnych wektorów $\vec{x}, \vec{y} \in \mathcal{R}^n$ zachodzi nierówność*

$$|\,\vec{x}\vec{y}\,| \leqslant |\,\vec{x}\,||\,\vec{y}\,|,$$

czyli wartość bezwzględna iloczynu skalarnego wektorów $\vec{x}$ i $\vec{y}$ nie przekracza iloczynu długości tych wektorów.

Dowód. Nierówność ta jest oczywista, gdy $\vec{x} = \vec{0}$. Załóżmy więc, że $\vec{x} \neq \vec{0}$ i rozważmy równanie kwadratowe

$$\sum_{i=1}^{n} (x_i \cdot t - y_i)^2 = 0 \tag{1}$$

z niewiadomą t. Nietrudno zauważyć, że równanie to jest identyczne z równaniem

$$(\sum_{i=1}^{n} x_i^2) \cdot t^2 - 2(\sum_{i=1}^{n} x_i y_i) \cdot t + \sum_{i=1}^{n} y_i^2 = |\,\vec{x}\,|^2 \cdot t^2 - 2\vec{x}\vec{y} \cdot t + |\,\vec{y}\,|^2 = 0$$

oraz jest równoważne układowi równań liniowych

$$x_1 \cdot t = y_1,\ \ldots,\ x_n \cdot t = y_n.$$

Ten układ może mieć najwyżej jedno rozwiązanie. Wobec tego, wyróżnik równania (1)

$$\Delta = 4(\vec{x}\vec{y})^2 - 4|\,\vec{x}\,|^2|\,\vec{x}\,|^2$$

jest niedodatni. Stąd otrzymujemy dowodzoną nierówność. □

Wniosek 1.2 *Dla dowolnych wektorów $\vec{x}, \vec{y} \in \mathcal{R}^n$ zachodzi nierówność*

$$|\,\vec{x} + \vec{y}\,| \leqslant |\,\vec{x}\,| + |\,\vec{y}\,|. \ \square$$

2 Elementy geometrii

Przypuśćmy, że mamy daną płaszczyznę lub przestrzeń trójwymiarową, lub też ogólniejsze pojęcie przestrzeni geometrycznej, o nieustalonym wymiarze. W takiej przestrzeni możemy wprowadzić prostokątny układ współrzędnych. Wtedy każdy punkt da się opisać za pomocą skończonego układu współrzędnych. Przyjmujemy, że współrzędne punktu są liczbami rzeczywistymi.

Jeżeli różne punkty mają różne układy współrzędnych i potrafimy konstruować punkty o zadanych współrzędnych, to taką przestrzeń geometryczną można utożsamić ze zbiorem $\mathcal{R}^n$.

W przestrzeni $\mathcal{R}^n$ odległość punktów $\vec{x} = (x_1, \ldots, x_n)$ i $\vec{y} = (y_1, \ldots, y_n)$ wyrażamy wzorem

$$| \vec{x} - \vec{y} | = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$

Lemat 2.1 *Odległość ma następujące własności:*

1. $| \vec{a} - \vec{b} | \geqslant 0,$
2. $| \vec{a} - \vec{b} | = 0$ *wtedy i tylko wtedy, gdy* $\vec{a} = \vec{b},$
3. $| \vec{a} - \vec{b} | = | \vec{b} - \vec{a} |$
4. $| \vec{a} - \vec{b} | \leqslant | \vec{a} - \vec{x} | + | \vec{x} - \vec{b} |$

dla dowolnych wektorów $\vec{a}, \vec{b}, \vec{x}$. $\square$

Pojęcie odległości pozwala zdefiniować pojęcie prostej. Mówimy, że punkt $\vec{x}$ leży między punktami $\vec{a}$ i $\vec{b}$, jeżeli

$$| \vec{b} - \vec{a} | = | \vec{x} - \vec{a} | + | \vec{b} - \vec{x} |.$$

Trzy punkty $\vec{a}, \vec{b}, \vec{c}$ są współliniowe, jeżeli jeden z tych punktów leży między pozostałymi. Prosta przechodząca przez (różne) punkty $\vec{a}$ i $\vec{b}$ to zbiór

$$\{\vec{x} \in \mathcal{R}^n : \text{punkty } \vec{a}, \vec{b}, \vec{x} \text{ są współliniowe}\}.$$

Odcinek, którego końcami są punkty $\vec{a}$ i $\vec{b}$, to zbiór

$$\{\vec{x} \in \mathcal{R}^n : \vec{x} \text{ leży między } \vec{a}, \vec{b}\}.$$

Lemat 2.2 *Przypuśćmy, że mamy dane dwa punkty* $\vec{a}$ *i* $\vec{b}$ *oraz liczbę* $s \in (0, 1)$. *Istnieje dokładnie jeden punkt* $\vec{c}$ *leżący między* $\vec{a}$ *i* $\vec{b}$ *taki, że* $| \vec{a} - \vec{c} | = s| \vec{a} - \vec{b} |$. *Co więcej, tym punktem jest* $\vec{a} + s(\vec{b} - \vec{a})$.

Dowód. Niech c będzie różne od d. Oczywiście, nierówności $0 < (c - d)^2$ oraz $2cd < c^2 + d^2$ są równoważne. Stąd mamy $2cd < c^2 + d^2$. Proste przekształcenia pozwalają stąd wyprowadzić nierówność

$$\left(\frac{c + d}{2}\right)^2 < \frac{c^2 + d^2}{2}.$$

Załóżmy, że mamy dwa różne punkty $\vec{c} = (c_1, \ldots, c_n)$ i $\vec{d} = (d_1, \ldots, d_n)$ o własnościach podanych w tezie lematu. Wtedy

$$| \vec{a} - \frac{\vec{c} + \vec{d}}{2} |^2 = \sum_{i=1}^{n} \left(\frac{(a_i - c_i) + (a_i - d_i)}{2}\right)^2 < \sum_{i=1}^{n} \frac{(a_i - c_i)^2 + (a_i - d_i)^2}{2} =$$

$$= \frac{|\,\vec{a}-\vec{c}\,|^2 + |\,\vec{a}-\vec{d}\,|^2}{2} = s^2|\,\vec{a}-\vec{b}\,|^2.$$

Tak więc

$$|\,\vec{a} - \frac{\vec{c}+\vec{d}}{2}\,| < s|\,\vec{a}-\vec{b}\,|.$$

Analogicznie dowodzimy, że

$$|\,\frac{\vec{c}+\vec{d}}{2} - \vec{b}\,| < (1-s)|\,\vec{a}-\vec{b}\,|.$$

Sumując otrzymane nierówności stronami otrzymujemy, że

$$|\,\vec{a}-\vec{b}\,| \leqslant |\,\vec{a} - \frac{\vec{c}+\vec{d}}{2}\,| + |\,\frac{\vec{c}+\vec{d}}{2} - \vec{b}\,| < (s+(1-s))|\,\vec{a}-\vec{b}\,| = |\,\vec{a}-\vec{b}\,|,$$

a to nie jest możliwe.

Wykazaliśmy więc jednoznaczność $\vec{c}$. Sprawdzenie, że $\vec{c}$ wyraża się przytoczonym wzorem jest łatwe. □

Wniosek 2.3 *Odcinek łączący punkty $\vec{a}$ i $\vec{b}$ to zbiór*

$$\{\vec{a}+s(\vec{b}-\vec{a}) \in \mathcal{R}^n : 0 \leqslant s \leqslant 1\} = \{t\vec{a}+s\vec{b} \in \mathcal{R}^n : t,s \geqslant 0 \wedge t+s=1\}. \ \Box$$

Wniosek 2.4 *Prosta przechodząca przez punkty $\vec{a}$ i $\vec{b}$ to zbiór*

$$\{\vec{a}+s(\vec{b}-\vec{a}) \in \mathcal{R}^n : s \in \mathcal{R}\} = \{t\vec{a}+s\vec{b} \in \mathcal{R}^n : t,s \in \mathcal{R} \wedge t+s=1\}. \ \Box$$

Przypuśćmy, że mamy trzy punkty $\vec{a}$, $\vec{b}$ i $\vec{c}$ takie, że $\vec{b},\vec{c} \neq \vec{a}$. Niech $p_{\vec{a},\vec{b}}$ będzie prostą przechodzącą przez punkty $\vec{a}$ i $\vec{b}$. Analogicznie definiujemy prostą $p_{\vec{a},\vec{c}}$. Proste $p_{\vec{a},\vec{b}}$ i $p_{\vec{a},\vec{c}}$ są nachylone do siebie pod kątem α, jeżeli $0 \leqslant \alpha \leqslant \pi$ oraz

$$\cos(\alpha) = \frac{(\vec{b}-\vec{a})(\vec{c}-\vec{a})}{|\,\vec{b}-\vec{a}\,||\,\vec{c}-\vec{a}\,|}.$$

Z lematu 1.1 i z własności funkcji cosinus (w tym z ciągłości) wynika, że liczba α jest dobrze określona. Uzasadnienie tej definicji (miary) kąta w $\mathcal{R}^2$ i $\mathcal{R}^3$ jest podawane zwykle na lekcjach matematyki.

Tę samą liczbę nazywamy kątem (miarą kąta) między odcinkami o końcach $\vec{a}$ i $\vec{b}$ oraz o końcach $\vec{a}$ i $\vec{c}$, a także kątem między wektorami $\vec{b}-\vec{a}$ oraz $\vec{c}-\vec{a}$. Zwróćmy jeszcze uwagę, że w geometrii kąt jest często rozumiany inaczej, jako część płaszczyzny między dwoma półprostymi. Oczywiście, dwie półproste o wspólnym początku rozbijają płaszczyznę na dwie części (czyli kąty). Zgodnie z przyjętą definicją, kąt jest miarą „mniejszej” z tych dwóch części.

3 Izometrie

Jednym z najbardziej znanych pojęć geometrycznych jest przystawanie. Figury są przystające, jeżeli jedną z nich można przekształcić na drugą używając izometrii.

Przyjmujemy, że funkcja $f : \mathcal{R}^n \to \mathcal{R}^n$ jest izometrią, jeżeli dla wszystkich $\vec{x},\vec{y} \in \mathcal{R}^n$ zachodzi równość

$$|\,f(\vec{x}) - f(\vec{y})\,| = |\,\vec{x}-\vec{y}\,|.$$

Najprostszym przykladem izometrii jest przesunięcie (o pewien wektor). Przesunięcie o wektor $\vec{a}$ jest to funkcja $p(\vec{x}) = \vec{x}+\vec{a}$.

Lemat 3.1 *Każde przesunięcie jest izometrią. Złożenie izometrii jest izometrią.* $\square$

Lemat 3.2 *Jeżeli f jest izometrią, to funkcja $g(\vec{x}) = f(\vec{x}) - f(\vec{0})$ też jest izometrią. Ponadto, $g(\vec{0}) = \vec{0}$.* $\square$

Z lematu 3.2 wynika, że aby poznać wszystkie izometrie, wystarczy poznać izometrie przekształcające $\vec{0}$ na $\vec{0}$. Pozostałe otrzymujemy stosując składanie z przesunięciami. Jest to konsekwencja oczywistego wzoru $f(\vec{x}) = (f(\vec{x}) - f(\vec{0})) + f(\vec{0})$.

Lemat 3.3 *Jeżeli $f : \mathcal{R}^n \to \mathcal{R}^n$ jest izometrią, to dla dowolnych $\vec{x}, \vec{y} \in \mathcal{R}^n$ oraz dla dowolnych nieujemnych $a, b \in \mathcal{R}$ takich, że $a + b = 1$ mamy*

$$f(a\vec{x} + b\vec{y}) = af(\vec{x}) + bf(\vec{y}).$$

Dowód. Wynika to z lematu 2.2. Punkt $\vec{z} = a\vec{x} + b\vec{y}$ leży między punktami $\vec{x}$ i $\vec{y}$. Mamy bowiem

$$|\, \vec{z} - \vec{x} \,| = |\, a\vec{x} + b\vec{y} - a\vec{x} - b\vec{x} \,| = b|\, \vec{y} - \vec{x} \,|$$

oraz podobnie $|\, \vec{y} - \vec{z} \,| = a|\, \vec{y} - \vec{x} \,|$. Stąd

$$|\, \vec{z} - \vec{x} \,| + |\, \vec{y} - \vec{z} \,| = b|\, \vec{y} - \vec{x} \,| + a|\, \vec{y} - \vec{x} \,| = |\, \vec{y} - \vec{x} \,|.$$

Ponieważ f jest izometrią, więc $f(\vec{z})$ leży między $f(\vec{x})$ i $f(\vec{y})$ oraz

$$|\, f(\vec{z}) - f(\vec{x}) \,| = b|\, f(\vec{y}) - f(\vec{x}) \,|.$$

Ze wspomnianego lematu otrzymujemy, że

$$f(a\vec{x} + b\vec{y}) = f(\vec{z}) = f(\vec{x}) + b(f(\vec{y}) - f(\vec{x})) = af(\vec{x}) + bf(\vec{y}). \quad \square$$

Lemat 3.4 *Jeżeli izometria f spełnia warunek $f(\vec{0}) = \vec{0}$, to jest funkcją jednorodną.*

Dowód. Załóżmy, że $a \in (0, 1)$ oraz weźmy $\vec{x}$. Zauważmy, że

$$a\vec{x} = (1 - a)\vec{0} + a\vec{x}.$$

Stąd i z poprzedniego lematu otrzymujemy, że

$$f(a\vec{x}) = (1 - a)f(\vec{0}) + af(\vec{x}) = af(\vec{x}).$$

Przyjmijmy teraz, że $a > 1$. Wtedy

$$\vec{x} = (1 - \frac{1}{a})\vec{0} + \frac{1}{a}(a\vec{x}).$$

Ponownie korzystamy z poprzedniego lematu:

$$f(\vec{x}) = (1 - \frac{1}{a})f(\vec{0}) + \frac{1}{a}f(a\vec{x})$$

i teraz wystarczy przemnożyć stronami otrzymaną równość przez a.

Jeżeli $a < 0$, to postępujemy podobnie:

$$\vec{0} = \frac{1}{1 - a}(a\vec{x}) + \frac{-a}{1 - a}\vec{x}.$$

Z poprzedniego lematu otrzymujemy, że

$$\vec{0} = f(\vec{0}) = \frac{1}{1 - a}f(a\vec{x}) + \frac{-a}{1 - a}f(\vec{x}).$$

Mnożąc tę równość przez $1 - a$ i przenosząc na drugą stronę jeden ze składników otrzymujemy żądany wzór. $\square$

Lemat 3.5 *Jeżeli f jest izometrią taką, że $f(\vec{0}) = \vec{0}$, to f jest funkcją addytywną.*

Dowód. Z lematów 3.4 i 3.3 otrzymujemy, że

$$\frac{1}{2}f(\vec{x}+\vec{y}) = f(\frac{\vec{x}+\vec{y}}{2}) = f(\frac{1}{2}\vec{x}+\frac{1}{2}\vec{y}) = \frac{1}{2}f(\vec{x})+\frac{1}{2}f(\vec{y}).$$

Stąd otrzymujemy addytywność f. □

Lemat 3.6 *Izometria f taka, że $f(\vec{0}) = \vec{0}$ zachowuje iloczyn skalarny, a więc*

$$\vec{x}\vec{y} = f(\vec{x})f(\vec{y})$$

dla wszystkich $\vec{x}, \vec{y} \in \mathcal{R}^n$.

Dowód. Oczywiście,

$$(\vec{x}-\vec{y})^2 = |\ \vec{x}-\vec{y}\ |^2 = |\ f(\vec{x}) - f(\vec{y})\ |^2 = (f(\vec{x}) - f(\vec{y}))^2.$$

Wobec tego,

$$\vec{x}^2 = (\vec{x}-\vec{0})^2 = (f(\vec{x}) - f(\vec{0}))^2 = (f(\vec{x}))^2$$

oraz

$$\vec{x}^2 - 2\vec{x}\vec{y} + \vec{y}^2 = f(\vec{x})^2 - 2f(\vec{x})f(\vec{y}) + f(\vec{y})^2.$$

Stąd otrzymujemy tezę. □

4 Wyznacznik macierzy izometrii

Pokażemy teraz, że wyznacznik macierzy przekształcenia zależy wyłącznie od przekształcenia, a nie od bazy użytej w definicji macierzy przekształcenia.

Wyprowadzimy to z następującego twierdzenia.

Twierdzenie 4.1 *Przypuśćmy, że V_1, V_2 i V_3 są przestrzeniami liniowymi skończonego wymiaru, B_1, B_2 i B_3 wybranymi bazami w tych (odpowiednio) przestrzeniach oraz $g : V_1 \to V_2$ i $f : V_2 \to V_3$ są przekształceniami liniowymi. Jeżeli A_2 jest macierzą przekształcenia g wyznaczoną przez bazy B_1 i B_2, a A_1 jest macierzą przekształcenia f wyznaczoną przez bazy B_2 i B_3, to A_1A_2 jest macierzą złożenia fg przekształcenia f i g wyznaczoną przez bazy B_1 i B_3.* □

Twierdzenie 4.2 *Niech $f : V \to V$ będzie przekształceniem liniowym, a B_1 i B_2 dwoma skończonymi bazami w V. Niech A_1 i A_2 będą bazami f wyznaczonymi odpowiednio przez bazy B_1 i B_2 (dokładniej, wyznaczając A_i w przestrzeniach argumentów i wartości bierzemy tę samą bazę B_i). Wtedy istnieje macierz C o niezerowym wyznaczniku taka, że*

$$A_2 = CA_1C^{-1}.$$

Dowód. Oczywiście, $f = id \circ f \circ id$. Niech C_2 będzie macierzą przekształcenia identycznościowego wyznaczoną przez bazy B_2 i B_1, a C_1 – wyznaczoną przez bazy B_1 i B_2. Oczywiście, są to macierze odwracalne. Z przytoczonego twierdzenia mamy, że

$$A_2 = C_1A_1C_2.$$

Z tego samego powodu macierz C_1C_2 jest macierzą przekształcenia identycznościowego wyznaczoną przez bazę B_2 (czyli bazy B_2 i B_2). Z definicji macierzy przekształcenia otrzymujemy, że $C_1C_2 = I$. Tak więc $C_2 = C_1^{-1}$. Aby otrzymać tezę, wystarczy przyjąć, że $C = C_1$. □

Wniosek 4.3 *Wyznacznik macierzy przekształcenia nie zależy od wyboru bazy wyznaczającej macierz.*

Dowód. Aby dowieść ten wniosek, wystarczy skorzystać z faktu, że wyznacznik iloczynu macierzy jest iloczynem ich wyznaczników. □

Lemat 4.4 *Jeżeli f jest izometrią taką, że $f(\vec{0}) = \vec{0}$, to f jest funkcją liniową i wyznacznik macierzy funkcji f jest równy ± 1.*

Dowód. Wyznacznik macierzy przekształcenia liniowego nie zależy od bazy użytej w definicji macierzy przekształcenia. Będziemy więc rozważać bazę standardową.

Niech A będzie macierzą funkcji f wyznaczoną przez bazę standardową. Wtedy kolumnami A są wektory $f(\vec{e}_i)$. Z tego powodu wierszami macierzy A^T są też wektory tej postaci Wyliczmy macierz $A^T A$. Jej wyrazami są iloczyny skalarne $f(\vec{e}_i)f(\vec{e}_j)$. Taki iloczyn jest równy iloczynowi $\vec{e}_i\vec{e}_j$. Tak więc macierz $A^T A$ jest macierzą jednostkową. Stąd otrzymujemy, że

$$| A |^2 = | A^T A | = | I | = 1.$$

Ostatecznie, $| A | = \pm 1$. □

5 Obroty w $\mathcal{R}^2$

Z elementarnej geometrii wynika, że obracając punkt o współrzędnych (x, y) o kąt φ dookoła początku układu współrzędnych otrzymujemy punkt o współrzędnych

$$O_\varphi(x, y) = (x \cdot \cos\varphi - y \cdot \sin\varphi, \ \ x \cdot \sin\varphi + y \cdot \cos\varphi).$$

Jest oczywiste, że przekształcenie to jest liniowe. Ponadto zachowuje długość wektora. Liniowe przekształcenia zachowujące długość zachowują odległość. Jest więc to izometria i to przekształcająca $(0, 0)$ na $(0, 0)$. Zauważmy też, że macierz tego przekształcenia w bazie standardowej jest równa

$$\begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}.$$

Wyznacznik tej macierzy jest równy 1.

Pokażemy teraz, że izometria $f : \mathcal{R}^2 \to \mathcal{R}^2$ taka, że $f(0, 0) = (0, 0)$, której macierz ma wyznacznik 1, jest obrotem. Niech

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

będzie macierzą tego przekształcenia w bazie standardowej. Tak więc $f(1, 0) = (a, c)$. Ponieważ izometria zachowuje długość wektora, więc $a^2 + c^2 = 1$. Dla takich liczb istnieje kąt φ taki, że $a = cos\varphi$ i $c = \sin\varphi$. Pozostaje wyliczyć b i d. Oczywiście $f(0, 1) = (b, d)$. Ponieważ izometrie zachowują iloczyn skalarny, więc

$$0 = (1, 0)(0, 1) = f(1, 0)f(0, 1) = (a, c)(b, d) = ab + cd.$$

Ponieważ wyznacznik A jest równy 1, więc

$$ad - bc = 1.$$

Z podanych równań łatwo wyznaczyć, że $b = -c = -\sin\varphi$ oraz $d = a = \cos\varphi$.

W przestrzeni $\mathcal{R}^2$ obroty o środku w początku układu współrzędnych to dokładnie izometrie przekształcające $(0, 0)$ na $(0, 0)$, które mają macierz o wyznaczniku 1.

5.1 Obroty w $\mathcal{R}^2$ a liczby zespolone

Przestrzeń $\mathcal{R}^2$ jest w naturalny sposób izomorficzna z liczbami zespolonymi uważanymi za przestrzeń liniową nad ciałem liczb rzeczywistych. Korzystając z tego izomorfizmu obroty w $\mathcal{R}^2$ można uważać za funkcje przekształcające liczby zespolone w liczby zespolone. Nietrudno zauważyć, że obrót O_φ o kąt φ (dookoła zera) daje się wyrazić wzorem

$$O_\varphi(x+iy) = (\cos\varphi + isin\varphi)(x+iy),$$

a więc jest iloczynem przez pewną liczbę zespoloną o module 1.

Liczby zespolone mogą być rozważane także jako algebra macierzy. Funkcja f zdefiniowana wzorem

$$f(x+iy) = \begin{bmatrix} x & -y \\ y & x \end{bmatrix}$$

przekształca zbiór liczb zespolonych $\mathcal{C}$ na zbiór macierzy o wymiarach 2×2, mających podaną postać. Ten zbiór jest $\mathcal{R}$-algebrą izomorficzną z $\mathcal{C}$, a f jest izomorfizmem tej algebry i $\mathcal{C}$. Zauważmy, że zbiór takich macierzy o wyznaczniku 1 to zbiór macierzy obrotów (dla bazy standardowej). Ponadto zbiór ten przez izomorfizm f odpowiada zbiorowi liczb zespolonych o module 1.

Będziemy teraz tworzyć podobną sytuację dotyczącą obrotów w $\mathcal{R}^3$.

6 Obroty w $\mathcal{R}^3$

Przypuśćmy, że $f : \mathcal{R}^3 \to \mathcal{R}^3$ jest izometrią taka, że $f(0,0,0) = (0,0,0)$, której macierz A (w bazie standardowej) ma wyznacznik 1. Chcemy pokazać, że f jest obrotem.

Jeżeli jest to obrót w $\mathcal{R}^3$, to jest to obrót wokół pewnej osi. Najpierw znajdziemy tę oś. Punkt $\vec{x} \neq \vec{0}$ z tej osi podczas obrotu nie ulega zmianie. Wobec tego, $f(\vec{x}) = \vec{x} = \lambda\vec{x}$ dla $\lambda = 1$. Zaczynamy od szukania λ, dla których jest wektor $\vec{x} \neq \vec{0}$ spełniający podaną równość.

Zauważmy, że

Lemat 6.1 *Przypuśćmy, że A jest macierzą przekształcenia f. Wyznacznik macierzy $A - \lambda I$ jest równy 0 wtedy i tylko wtedy, gdy istnieje niezerowy wektor $\vec{x}$ taki, że $f(\vec{x}) = \lambda\vec{x}$.*

Dowód. Równość $\mid A - \lambda I \mid = 0$ oznacza, że układ równań

$$(A - \lambda I)\vec{x} = \vec{0}$$

ma wiele rozwiązań, w tym niezerowe rozwiązanie $\vec{x}$. Nietrudno zauważyć, że niezerowymi rozwiązaniami tego układu są dokładnie wektory $\vec{x}$ spełniające równość $f(\vec{x}) = \lambda\vec{x}$. □

Jeżeli wyliczymy wyznacznik $\mid A - \lambda I \mid$, to otrzymamy wyrażenie postaci

$$(a_{1,1} - \lambda)(a_{2,2} - \lambda)(a_{3,3} - \lambda) + \ldots,$$

gdzie zamiast ... pojawia się pewien wielomian zmiennej λ drugiego stopnia. Wyrażenie to jest wielomianem trzeciego stopnia i nazywa się wielomianem charakterystycznym przekształcenia f, a jego pierwiastki – wartościami własnymi tego przekształcenia.

Tak więc otrzymaliśmy równanie na λ i jest to równanie trzeciego stopnia. Takie równanie w liczbach rzeczywistych ma rozwiązanie.

Lemat 6.2 *Izometria $f : \mathcal{R}^3 \to \mathcal{R}^3$ taka, że $f(0,0,0) = (0,0,0)$ ma wartości własne i wartości te są równe ± 1.*

Dowód. Niech λ_0 będzie wartością własną izometrii f. Wtedy równanie liniowe $(A-\lambda_0 I)\vec{x} = 0$ ma przynajmniej dwa rozwiązania. Niech $\vec{x_0}$ będzie niezerowym roziwązaniem tego równania. Wobec tego,

$$f(\vec{x}_0) = A\vec{x}_0 = \lambda_0\vec{x}_0.$$

Ponieważ f jest izometrią, więc zachowuje iloczyn skalarny. Stąd,

$$\lambda_0^2\vec{x}_0\vec{x}_0 = (\lambda_0\vec{x}_0)(\lambda_0\vec{x}_0) = f(\vec{x}_0)f(\vec{x}_0) = \vec{x}_0\vec{x}_0.$$

Stąd $\lambda_0^2 = 1$ i ostatecznie otrzymujemy, że pierwiastkami wielomianów charakterystycznych rozważanych przekształceń mogą być tylko liczby ± 1. □

Lemat 6.3 *Niech $\vec{f_1}, \vec{f_2}, \vec{f_1} \in \mathcal{R}^3$ będzie bazą złożoną z wektorów o długości 1 i takich, że $\vec{f_i}\vec{f_j} = 0$ dla $i \neq j$. Przypuśćmy, że*

$$\vec{x} = (x, y, z) = a\vec{f_1} + b\vec{f_2} + c\vec{f_3}$$

oraz

$$\vec{x'} = (x', y', z') = a'\vec{f_1} + b'\vec{f_2} + c'\vec{f_3}.$$

Wtedy

$$\vec{x}\vec{x'} = xx' + yy' + zz' = aa' + bb' + cc'.$$

Dowód. Jest to oczywista konsekwencja dwuliniowości iloczynu skalarnego i przyjętych własności rozważanej bazy. □

Lemat 6.4 *Liczba 1 jest wartością własną każdej izometrii $f : \mathcal{R}^3 \to \mathcal{R}^3$ takiej, że $f(0,0,0) = (0,0,0)$, której macierz A (w bazie standardowej) ma wyznacznik 1.*

Dowód. Niech f będzie izometrią spełniającą założenia lematu. Wiemy, że jedna z liczb ± 1 jest wartością własną f. Jeżeli tą liczbą jest 1, to teza lematu zachodzi w sposób oczywisty. Możemy więc założyć, że -1 jest wartością własną f.

Niech $\vec{v}_3$ będzie wektorem długości 1 takim, że $f(\vec{v}_3) = -\vec{v}_3$. Weźmy teraz dopełnienie ortogonalne przestrzeni generowanej przez $\vec{v}_3$ i jeden z wektorów tej przestrzeni o długości 1. Oznaczmy go symbolem $\vec{v}_2$. Niech $\vec{v}_1$ będzie wektorem długości 1 z dopełnienia ortogonalnego przestrzeni generowanej przez $\vec{v}_2$ i $\vec{v}_3$. Mamy więc trzy wektory takie, że iloczyn skalarny dowolnych dwóch z nich (różnych) jest równy 0, a kwadrat skalarny każdego jest równy 1. Niech teraz A oznacza macierz przekształcenia f względem bazy $\vec{v}_1, \vec{v}_2, \vec{v}_3$, a więc

$$A = [f(\vec{v}_1), f(\vec{v}_2), -\vec{v}_3] = \begin{bmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

(zamiast $f(\vec{v}_i)$ bierzemy kolumnę współrzędnych tego wektora w ustalonej bazie). Zauważmy, że

$$|\, A - \lambda I \,| = \begin{vmatrix} a-\lambda & b & 0 \\ c & d-\lambda & 0 \\ 0 & 0 & -1-\lambda \end{vmatrix} = -\left((a-\lambda)(d-\lambda) - bc\right)(1+\lambda) =$$

$$-\Big(\lambda^2 - (a+d)\lambda + ad - bc\Big)(1+\lambda) = -\Big(\lambda^2 - (a+d)\lambda - 1\Big)(1+\lambda).$$

Tak jest, ponieważ $1 = |\,A\,| = -(ad - bc)$. Wyliczymy do końca wielomian charakterystyczny f. Ponieważ $f(\vec{v}_1)f(\vec{v}_2) = 0$, więc $ab + cd = 0$ na mocy poprzedniego lematu. Zachodzi także równość $a^2 + c^2 = 1$, gdyż długość wektora $f(\vec{v}_1)$ jest równa 1. Tak więc

$$-a = a^2d - (ab)c = a^2d + c^2d = d.$$

Podstawiając otrzymaną równość do wzoru na wielomian charakterystyczny f otrzymujemy, że

$$|\,A - \lambda I\,| = -\left(\lambda^2 - 1\right)(1 + \lambda).$$

Teraz jest oczywiste, że 1 jest wartością własną przekształcenia f i istnieje niezerowy wektor $\vec{x}$ taki, że $f(\vec{x}) = \vec{x}$. $\square$

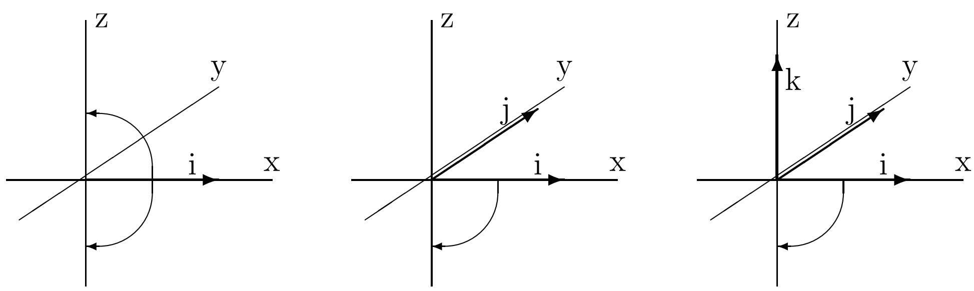
7 Kwaterniony

7.1 Rozważania wstępne

Chcemy skonstruować algebrę kwaternionów, która pozwalałaby na opisywanie obrotów w przestrzeni $\mathcal{R}^3$ podobnie, jak ciało liczb zespolonych opisuje obroty na płaszczyźnie $\mathcal{R}^2$. Najlepiej byłoby, gdyby ta algebra była ciałem, ale to okaże się niemożliwe. Chcemy, aby każdemu elementowi tej algebry (albo np. elementom o module 1) odpowiadał obrót i to w taki sposób, aby mnożeniu odpowiadało składanie obrotów. Wtedy oczywiście kwaternionowi 1 (elementowi neutralnemu mnożenia w algebrze kwaternionów) będzie odpowiadać obrót o kąt 0. Zastanówmy się, jakie własności mają kwaterniony odpowiadające obrotom o kąt półpełny. Jest oczywiste, że jest dużo takich obrotów i takich kwaternionów q też musi być dużo.

Naturalna hipoteza, że $q^2 = 1$ jest nie do przyjęcia. W dowolnych pierścieniach z jednością, bez dzielników zera równanie $x^2 = 1$ ma najwyżej dwa rozwiązania: 1 i -1. Wynika stąd, że odpowiedniość między kwaternionami i obrotami nie może być wzajemnie jednoznaczna. W szczególności, obrót o 0 stopni musi być opisany przez kilka kwaternionów.

Przyjrzyjmy się rysunkom.



Definiując obrót musimy podać oś i kąt obrotu. Z drugiej strony stwierdzenie, że chcemy np. obrócić wektor jednostkowy i (albo koniec tego wektora) o 90 stopni (o $\pi/2$) wokół osi y, nie definiuje jeszcze obrotu. Definiując obrót na płaszczyźnie dodalibyśmy, że ma być on zgodny lub nie zgodny z ruchem wskazówek zegara. W przestrzeni taka informacja też nie będzie wystarczająca. To, czy obrót jest zgodny z ruchem wskazowek zegara, zależy od miejsca obserwacji. Aby opisać obrót, zamiast osi rozumianej jako pewna prosta, dobrze jest opisać oś za pomocą wektora lub prostej o pewnym kierunku. Wtedy podany wyżej obrót f możemy opisać w następujący sposób: obracamy o 90 stopni wokół osi wyznaczonej przez wektor j, i jeżeli patrzymy w kierunku wektora j, to obracamy zgodnie z ruchem wskazówek zegara. Dalej tak będziemy rozumieć obroty: jako obroty wokół wektora, o podany kąt, zgodnie z ruchem wskazówek zegara w podanym sensie.

Obrót f został już precyzyjnie zdefiniowany, ale to nie koniec kłopotów. Obrót ten możemy różnie wyrazić wzorem algebraicznym, zależnie od wyboru wektorów opisujących przestrzeń. Niech k będzie wektorem jednostkowym równoległym do osi z, jednym z dwóch możliwych. Zależnie od wyboru wektora k, albo wzór algebraiczny na obrócony wektor i ma postać $f(i) = k$, albo $f(i) = -k$. Na rysunku wektor k został tak wybrany, aby $f(i) = -k$.

Każdy obrót można opisać w podany sposób podając dwie różne informacje. Bierze się to stąd, że obrót wokół wektora $\vec{v}$ o kąt φ jest także obrotem wokół wektora $-\vec{v}$ o kąt $360 - \varphi$. Ta zależność ma szczególną postać w przypadku obrotów o 180 stopni. Obroty o taki kąt wokół osi $\vec{v}$ i $-\vec{v}$ są identyczne. Kwaterniony oddają przytoczony sposób definiowania obrotów. Kwaterniony odpowiadające obrotom o 180 stopni można utożsamiać z osiami obrotu. Okaże się jak należy się spodziewać, że kwaterniony różniące się znakiem opisują ten sam obrót, także dla dowolnych kątów.

Niech q nadal będzie kwaternionem opisującym obrót wokół jakiegoś wektora o 180 stopni. Przytoczone rozważania sugerują na dwa sposoby, że $q^2 = -1$. Po pierwsze, -1 też opisuje obrót o 0 stopni, a więc jeżeli q^2 nie może być równe 1, to może okazać się równe -1. Kwaternion q^2 może być też interpretowany jako obrót o 360 stopni. Wtedy kwaternion $(-q)q = -q^2$ powinien być interpretowany jako obrót o 0 stopni. Jeżeli $-q^2 = 1$, to oczywiście także $q^2 = -1$. Wszytko wskazuje więc na to, że w algebrze kwaternionów powinno być dużo pierwiastków kwadratowych z -1 i nie może zachodzić twierdzenie Bezout.

Analizując, dlaczego w ciałach są najwyżej dwa pierwiastki z dowolnej liczby, można zauważyć, że bardzo istotna jest przemienność mnożenia. Rezygnując z przemienności można stworzyć algebrę kwaternionów zawierającą bardzo dużo pierwiastków z -1.

7.2 Kwaterniony jako algebra początkowa (lub ciekawostka)

Chcemy więc zbudować algebrę,

- która jest skończenie wymiarową $\mathcal{R}$-algebrą (niekoniecznie przemienną),
- w której każdy różny od zera element ma (obustronny) element odwrotny,
- w której są dwa różne pierwiastki z -1 oznaczane i oraz I takie, że $I \neq -i$ (i oczywiście $i^2 = I^2 = -1$).

Przypuśćmy, że taka algebra istnieje i oznaczmy ją symbolem $\mathcal{H}$. Algebrę $\mathcal{H}$ będziemy nazywać algebrą kwaternionów, a jej elementy - kwaternionami. Czasem nazywa się ją (nieprzemiennym) ciałem kwaternionów. Wtedy należy pamiętać, że w tej algebrze mnożenie nie jest przemienne, a więc nie są spełnione wszystkie warunki wymienione w definicji ciała.

7.3 Pierwsze własności

Zrozumienia tego rozdziału wymaga bardzo dokładnego przypomnienia sobie definicji $\mathcal{R}$-algebry.

Przyjmijmy następujące oznaczenia: $\cdot_w$ i $\cdot_z$ to odpowiednio mnożenia wewnętrzne i zewnętrzne w algebrze $\mathcal{H}$, $1_\mathcal{H}$ i $1_\mathcal{R}$ to jedności w algebrze $\mathcal{H}$ i ciele $\mathcal{R}$, a $0_\mathcal{H}$ i $0_\mathcal{R}$ to zera w tych algebrach. Najpierw pokażemy, że dopisywanie tych indeksów jest niepotrzebne.

Lemat 7.1 *Dla wszystkich $x \in \mathcal{R}$ achodzi równość $x \cdot_z 0_\mathcal{H} = 0_\mathcal{H}$.*

Dowód. Zauważmy, że

$$x \cdot_z 0_{\mathcal{H}} = x \cdot_z (0_{\mathcal{H}} + 0_{\mathcal{H}}) = x \cdot_z 0_{\mathcal{H}} + x \cdot_z 0_{\mathcal{H}}.$$

Odejmując od tej równości stronami $x \cdot_z 0_{\mathcal{H}}$ otrzymujemy, że $x \cdot_z 0_{\mathcal{H}} = 0_{\mathcal{H}}$. □

Lemat 7.2 *Algebra $\mathcal{H}$ zawiera podalgebrę izomorficzną z ciałem liczb rzeczwistych $\mathcal{R}$.*

Dowód. Funkcja $J : \mathcal{R} \to \mathcal{H}$ zdefiniowana wzorem $J(x) = x \cdot_z 1_{\mathcal{H}}$ jest monomorfizmem ciała liczb rzeczywistych $\mathcal{R}$ w algebrę $\mathcal{H}$. Łatwo dowieść, że $J(x +_{\mathcal{R}} y) = J(x) +_{\mathcal{H}} J(y)$ oraz $J(x \cdot_{\mathcal{R}} y) = J(x) \cdot_{\mathcal{H}} J(y)$. Pokażemy tylko różnowartościowość J. Przypuśćmy, że $x \cdot_z 1_{\mathcal{H}} = 0_{\mathcal{H}}$ dla pewnej liczby rzeczywistej $x \neq 0$. Pomnóżmy tę równość stronami przez x^{-1}. Z poprzedniego lematu otrzymujemy, że

$$0_{\mathcal{H}} = (x^{-1}) \cdot_z 0_{\mathcal{H}} = (x^{-1}) \cdot_z (x \cdot_z 1_{\mathcal{H}}) = (x^{-1} \cdot x) \cdot_z 1_{\mathcal{H}} = 1_{\mathcal{R}} \cdot_z 1_{\mathcal{H}} = 1_{\mathcal{H}},$$

(pamiętajmy, że mnożenie zewnętrzne przez 1 jest funkcją identycznościową), a to nie jest możliwe. Uzyskana sprzeczność dowodzi, że równość $x \cdot_z 1_{\mathcal{H}} = 0_{\mathcal{H}}$ zachodzi tylko dla $x = 0$. Tak więc J jest monomorfizmem. □

Udowodniony lemat pozwala utożsamiać liczbę rzeczywistą x z elementem postaci $x \cdot_z 1_{\mathcal{H}}$. Dalej przyjmujemy więc, że

$$x \cdot_z 1_{\mathcal{H}} = x \in \mathcal{H}$$

dla wszystkich liczb rzeczywistych x.

Lemat 7.3 *Jednością algebry $\mathcal{H}$ jest liczba rzeczywista 1, czyli $1_{\mathcal{R}} \cdot_z 1_{\mathcal{H}}$. Zerem w algebrze $\mathcal{H}$ jest liczba rzeczywista 0, czyli $0_{\mathcal{R}} \cdot_z 1_{\mathcal{H}}$.*

Dowód. Mnożenie zewnętrzne przez liczbę 1 jest funkcją identycznościową. Tak więc

$$1_{\mathcal{H}} = 1_{\mathcal{R}} \cdot_z 1_{\mathcal{H}} = 1_{\mathcal{R}}.$$

Druga podana własność dowodzimy podobnie:

$$1_{\mathcal{R}} \cdot_z 1_{\mathcal{H}} = (1_{\mathcal{R}} + 0_{\mathcal{R}}) \cdot_z 1_{\mathcal{H}} = 1_{\mathcal{R}} \cdot_z 1_{\mathcal{H}} + 0_{\mathcal{R}} \cdot_z 1_{\mathcal{H}}.$$

Stąd wynika, że

$$0_{\mathcal{H}} = 0_{\mathcal{R}} \cdot_z 1_{\mathcal{H}} = 0_{\mathcal{R}} \quad \square$$

Lemat 7.4 *Mnożenie zewnętrzne jest mnożeniem wewnętrznym, a więc*

$$x \cdot_z h = (x \cdot_z 1_{\mathcal{H}}) \cdot_w h = x \cdot_w h$$

dla wszystkich $x \in \mathcal{R}$ i $h \in \mathcal{H}$.

Dowód. Zauważmy, że

$$x \cdot_z h = x \cdot_z (1_{\mathcal{H}} \cdot_w h) = (x \cdot_z 1_{\mathcal{H}}) \cdot_w h = x \cdot_w h. \ \square$$

Lemat 7.5 *W algebrze $\mathcal{H}$ mnożenie przez liczby rzeczywiste jest przemienne, a więc*

$$x \cdot_w h = h \cdot_w x$$

dla wszystkich $x \in \mathcal{R}$ i $h \in \mathcal{H}$.

Dowód. Zauważmy, że

$$x \cdot_w h = (x \cdot_z 1_{\mathcal{H}}) \cdot_w h = x \cdot_z (1_{\mathcal{H}} \cdot_w h) = x \cdot_z (h \cdot_w 1_{\mathcal{H}}) = h \cdot_w (x \cdot_z 1_{\mathcal{H}}) = h \cdot_w x. \ \square$$

Dalej nie będziemy już stosować indeksów tak, jak w tym rozdziale.

7.4 $\mathcal{H}$ jest sumą prostą

Lemat 7.6 *Każdy element x algebry $\mathcal{H}$ można przedstawić w postaci*

$$x = \frac{x - ixi}{2} + \frac{x + ixi}{2}.$$

Element $x - ixi$ jest przemienny z i. Element $x + ixi$ spełnia natomiast równość

$$i(x + ixi) = -(x + ixi)i.$$

Dowód. Mamy bowiem

$$(x - ixi)i = xi - ix(i^2) = xi + ix \text{ oraz } i(x - ixi) = ix - (i^2)xi = ix + xi.$$

Drugą równość sprawdzamy równie łatwo, jak poprzednią. W algebrze z przemiennym mnożeniem element $x + ixi$ musiałby być równy 0. □

Przyjmijmy, że

$$\mathcal{H}_+ = \{x \in \mathcal{H} : ix = xi\} \text{ oraz } \mathcal{H}_- = \{x \in \mathcal{H} : ix = -xi\}.$$

Twierdzenie 7.7 *Zbiory $\mathcal{H}_+$ i $\mathcal{H}_-$ są podprzestrzeniami liniowymi $\mathcal{H}$ i $\mathcal{H}$ jest sumą prostą tych podprzestrzeni.*

Dowód. Wobec poprzedniego lematu wystarczy sprawdzić, że jedynym elementem $x \in \mathcal{H}_+ \cap \mathcal{H}_-$ jest wektor zerowy. Zauważmy, że $ix = xi = -ix$. Mnożąc otrzymaną równość stronami przez i otrzymujemy, że $-x = x$, a więc $x = 0$. □

Zauważmy też, że

Lemat 7.8 *Zbiór $\mathcal{H}_+$ jest zamknięty ze względu na mnożenie, a więc jest $\mathcal{R}$-algebrą.* □

7.5 Charakteryzacja $\mathcal{H}_+$

Przyjmijmy, że $\mathcal{C} = \{r + si \in \mathcal{H} : r, s \in \mathcal{R}\}$. Łatwo dowodzi się następujący fakt.

Lemat 7.9 *Zbiór $\mathcal{C}$ jest algebrą izomorficzną z ciałem liczb zespolonych.* □

Będziemy więc utożsamiać elementy $\mathcal{C}$ z liczbami zespolonymi.

Twierdzenie 7.10 *Algebra $\mathcal{H}_+$ jest ciałem liczb zespolonych $\mathcal{C}$.*

Dowód. Oczywiście, $\mathcal{C} \subseteq \mathcal{H}_+$. Wystarczy więc dowieść, że $\mathcal{H}_+ \subseteq \mathcal{C}$. Weźmy więc $h \in \mathcal{H}_+$.

Niech $\mathcal{R}[x]$ będzie pierścieniem wielomianów o współczynnikach rzeczywistych. Przyjmijmy, że $val(x) = h$. Warunek ten pozwala zdefiniować wartość wielomianów w algebrze $\mathcal{H}_+$ przy takim właśnie wartościowaniu. Niech $val : \mathcal{R}[x] \to \mathcal{H}_+$ będzie tak zdefiniowaną wartością wielomianów. Funkcja ta jest homomorfizmem algebry wielomianów $\mathcal{R}[x]$ w algebrę $\mathcal{H}_+$. Pierwsza z tych algebr ma wymiar nieskończony, druga - przeciwnie. Istnieje więc niezerowy wielomian $w(x) \in \mathcal{R}[x]$ taki, że $w(h) = val(w(x)) = 0$. O tym wielomianie możemy założyć, że jest stopnia nie większego niż 2. Wynika ze znanego twierdzenia mówiącego, że każdy wielomian o współczynnikach rzeczywistych jest iloczynem wielomianów stopnia $\leqslant 2$. Jest to tzw. zasadnicze twierdzenie algebry.

Pierwiastki wielomianu stopnia 1 o współczynnikach rzeczywistych są liczbami rzeczywistymi. Możemy więc założyć, że $w(x)$ jest stopnia 2. Możemy także przyjąć, że współczynnik tego wielomianu przy x^2 jest równy 1. Niech

$$w(x) = x^2 + ax + b.$$

Mamy więc

$$w(h) = h^2 + ah + b = 0.$$

Z drugiej strony wiemy, że wielomian w ma dwa pierwiastki zespolone $z, z' \in \mathcal{C}$. Wiedząc to, łatwo wyliczyć współczynniki wielomianu w:

$$a = z + z' \ \text{ oraz } \ b = zz'.$$

Stąd otrzymujemy, że

$$0 = h^2 + (z + z')h + zz' = (h - z)(h - z').$$

Ponieważ w algebrze $\mathcal{H}$ każdy niezerowy element ma element odwrotny, więc nie ma dzielników zera, a także $h = z$ lub też $h = z'$. W obu przypadkach, $h \in \mathcal{C}$. □

7.6 Własności $\mathcal{H}_-$

Przypomijmy, że I oznacza trzeci pierwiastek z -1 w algebrze kwaternionów.

Lemat 7.11 *Pewien pierwiastek z -1 należy do zbioru $\mathcal{H}_-$.*

Dowód. Ten lemat ma rachunkowy dowód. Wprowadźmy następujące oznaczenia:

$$a = \frac{I - iIi}{2} \text{ oraz } b = \frac{I + iIi}{2}.$$

Zachodzą więc następujące równości

$$I = a + b, \qquad ai = ia \ \text{ oraz } \ bi = -ib.$$

Element b spełniający równość $bi = -ib$ nie może być liczbą zespoloną, tym bardziej nie może być liczbą rzeczywistą, ani zerem.

Zauważmy, że

$$(I - iIi)(I + iIi) = I^2 + IiIi - iIiI - iIiiIi = IiIi - iIiI$$

oraz

$$(I + iIi)(I - iIi) = I^2 - IiIi + iIiI - iIiiIi = -(IiIi - iIiI).$$

Stąd otrzymujemy, że

$$ab = -ba.$$

Ta ostatnia równość ma dwie konsekwencje. Po pierwsze,

$$-1 = I^2 = (a + b)^2 = a^2 + ab + ba + b^2 = a^2 + b^2.$$

Ponadto, ponieważ $a \in \mathcal{H}_+ = \mathcal{C}$, więc $a = r + si$ oraz

$$rb + sib = (r + si)b = ab = -ba = -b(r + si) = -rb - sbi = -rb + sib.$$

Oznacza to, że $rb = 0$. Tak więc $r = 0$ oraz $a = si$ dla pewnej liczby rzeczywistej s.

Po połączeniu tych dwóch konsekwencji otrzymujemy, że

$$b^2 = -1 - a^2 = -1 + s^2 \in \mathcal{R}.$$

Łatwo też wykazać, że kwadrat kwaternionu jest nieujemną liczbą rzeczywistą tylko wtedy, gdy ten kwaternion jest liczbą rzeczywistą. Nieujemne liczby rzeczywiste są kwadratami liczb rzeczywistych. Jeżeli $b^2 = r^2$ dla liczby rzeczywistej r, to

$$(b + r)(b - r) = b^2 + rb - br - r^2 = b^2 + rb - rb - r^2 = b^2 - r^2 = 0,$$

gdyż iloczyn przez liczbę rzeczywistą niezależy od kolejności czynników. Stąd wynika, że $b = \pm r \in \mathcal{R}$, a to nie jest możliwe. Ostatecznie,

$$b^2 = -1 + s^2 < 0.$$

Przyjmijmy, że

$$j = \frac{b}{\sqrt{1 - s^2}}.$$

Oczywiście, $j \in \mathcal{H}_-$ oraz $j^2 = -1$. $\Box$

Dalej j będzie oznaczać jeden z pierwiastków z -1 należący do $\mathcal{H}_-$.

Lemat 7.12 *Funkcja $f(x) = xj$ jest liniowa i różnowartościowa, przekształca przestrzeń liniową $\mathcal{H}$ na $\mathcal{H}$, a także przekształca $\mathcal{H}_+$ w $\mathcal{H}_-$ oraz $\mathcal{H}_-$ w $\mathcal{H}_+$.*

Dowód. Podane własności funkcji f wynikają natychmiast z definicji $\mathcal{R}$-algebry. Fakt, że f jest typu „na" jest wynika ze znanych własności funkcji liniowych przekształcających przestrzenie skończonego wymiaru. Pozostałe własności f wynikają z banalnych rachunków. $\Box$

Wniosek 7.13 *Przestrzenie $\mathcal{H}_+$ i $\mathcal{H}_-$ (nad ciałem liczb rzeczywistych) są tego samego wymiaru i mają wymiar 2. Przestrzeń $\mathcal{H}$ ma wymiar 4.*

Dowód. Równość wymiarów podanych przestrzeni wynika z poprzedniego lematu. Oczywiście, wymiar przestrzeni $\mathcal{H}_+ = \mathcal{C}$ jest równy 2. Przestrzeń liniowa $\mathcal{H}$ jest sumą prostą przestrzeni wymiaru 2, ma więc wymiar 4. $\Box$

7.7 Grupa kwaternionów

Rozważmy trzy następujące elementy algebry $\mathcal{H}$: i, j oraz ij. Zauważmy np., że

$$jij = -jji = i.$$

Tak więc prawdziwe są następujące równości:

$$\begin{aligned} i \cdot j &= ij, \\ j \cdot (ij) &= i, \\ (ij) \cdot i &= j. \end{aligned}$$

Wśród kwaternionów są więc trzy elementy a, b i c takie, że

$$ab = c, \quad bc = a \quad \text{oraz} \quad ca = b.$$

Algebrę początkową w klasie grup generowanych przez takie trzy elementy nazywamy grupą kwaternionów. W tym rozdziale opiszemy grupy generowane przez takie trzy elementy.

Lemat 7.14 *Niech G oznacza grupę generowaną przez elementy a, b i c spełniające podane równości. Wtedy kwadraty generatorów są równe, a czwarte potęgi generatorów grupy są równe jedności.*

Dowód. Zauważmy, że

$$a^2 = a(bc) = (ab)c = c^2 \text{ oraz } b^2 = (ca)b = c(ab) = c^2.$$

Ponadto,

$$a^5 = c^4a = ca^2ca = (ca)a(ca) = (ca)(ab) = bc = a.$$

Stąd a^4 jest elementem neutralnym. □

Lemat 7.15 *W grupach rozważanych w tym rozdziale rząd elementu c jest $\leqslant 2$ wtedy i tylko wtedy, gdy $ab = ba$.*

Dowód. Przypuśćmy, że c^2 jest jednością grupy. Wtedy $abab = e$. Pomnóżmy tę równość z lewej strony przez a i z prawej przez b. Wtedy z poprzedniego lematu otrzymujemy, że

$$ab = a^2bab^2 = b^2bab^2 = bb^2ab^2 = ba^2ab^2 = baa^2b^2 = bab^4 = ba,$$

czyli a i b komutują.

Zauważmy też, że jeżeli $ab = ba$, to $c^2 = abab = a^2b^2 = a^4 = e$. □

Jeżeli dwa z generatorów są równe jedności, to trzeci też jest równy jedności i cała grupa jest jednoelementowa.

Jeżeli dokładnie jeden z generatorów jest równy jedności, to pozostałe są sobie równe i grupa składa się z jedności i drugiego generatora.

Jeżeli żaden z generatorów nie jest jednością, to są one parami różne. Jeżeli wśród generatorów jest element rzędu 2, to generatory są przemienne i cała grupa jest przemienna. Taka grupa ma cztery elementy i następującą (wymagającą uzupełnienia) tabelkę :

	e	a	b	c
e	e	a	b	c
a	a	e	c	
b	b		e	a
c	c	b		e

Jeżeli generatory mają rząd 4, to nie mogą komutować i grupa ma rząd 8. Jeżeli przez ε oznaczamy kwadrat np. generatora a, to grupa ma następującą, wymagającą uzupełnienia tabelkę.

	e	a	b	c	ε	εa	εb	εc
e	e	a	b	c	ε	εa	εb	εc
a	a	ε	c		εa	e	εc	
b	b		ε	a	εb		e	εa
c	c	b		ε	εc	εb		e
ε	ε	εa	εb	εc	e	a	b	c
εa	εa	e	εc		a	ε	c	
εb	εb		e	εa	b		ε	a
εc	εc	εb		e	c	b		ε

7.8 Opis algebry kwaternionów $\mathcal{H}$

W przestrzeni liniowej $\mathcal{H}_+ = \mathcal{C}$ łatwo podać bazę: składa się ona z liczb zespolonych 1 oraz i. Funkcja $f(x) = xj$ jest izomorfizmem $\mathcal{H}_+$ i $\mathcal{H}_-$. Wobec tego jedna z baz przestrzeni $\mathcal{H}_-$ składa się z elementów $f(1)$ i $f(i)$, czyli j oraz ij. Oznaczmy iloczyn ij przez k. Wtedy bazą $\mathcal{H}_-$ jest np. para j i k. Natomiast jedna z baz przestrzeni $\mathcal{H}$ składa się z elementów 1, i, j, k. Każdy element $\mathcal{H}$ daje się więc przedstawić w postaci

$$\alpha + \beta i + \gamma j + \delta k$$

dla pewnych $\alpha, \beta, \gamma, \delta \in \mathcal{R}$. Jest to pierwszy ze sposobów reprezentowania kwaternionów. Mnożenie tak reprezentowanych kwaternionów sprowadza się do liczenia iloczynów takich jak ij. Iloczyny te obliczamy zgodnie z następującymi regułami:

$$ij = k = -ji, \quad jk = i = -kj, \quad ki = j = -ik$$

oraz

$$i^2 = j^2 = k^2 = -1.$$

Zbiór $\{\pm 1, \pm i, \pm j, \pm k\} \subseteq \mathcal{H}$ tworzy grupę. Jest to tzw. grupa kwaternionów. Grupa ta ma następującą tabelkę:

$\cdot$	1	i	j	k	-1	$-i$	$-j$	$-k$
1	1	i	j	k	-1	$-i$	$-j$	$-k$
i	i	-1	k	$-j$	$-i$	1	$-k$	j
j	j	$-k$	-1	i	$-j$	k	1	$-i$
k	k	j	$-i$	-1	$-k$	$-j$	i	1
-1	-1	$-i$	$-j$	$-k$	1	i	j	k
$-i$	$-i$	1	$-k$	j	i	-1	k	$-j$
$-j$	$-j$	k	1	$-i$	j	$-k$	-1	i
$-k$	$-k$	$-j$	i	1	k	j	$-i$	-1

Kwaterniony możemy przedstawiać jeszcze w inny sposób. Przypuśćmy, że mamy kwaternion

$$a + bi + cj + dk \in \mathcal{H},$$

gdzie $a, b, c, d \in \mathcal{R}$. Przyjmijmy, że $w = a + bi$ oraz $z = c + di$. Elementy w i z możemy uważać zarówno za kwaterniony, jaki i za liczby zespolone. Zauważmy, że

$$a + bi + cj + dk = w + zj.$$

Jest to inny, skrócony sposób przedstawiania kwaternionów. Tak przedtawione kwaterniony mnożymy w zwykły sposób. Należy jednak pamiętać, że dla $w \in \mathcal{C}$ zachodzą równości

$$jw = j(a + bi) = ja + jbi = aj - bij = (a - bi)j = \overline{w}j, \tag{2}$$

gdzie $\overline{w}$ oznacza liczbę zespoloną sprzężoną z w.

7.9 Istnienie ciała kwaternionów

Na początku założyliśmy, że mamy algebrę spełniającą pewną listę postulatów i staraliśmy się wyobrazić sobie, jak taka algebra wygląda. Nie mamy jednak żadnych gwarancji, że wśród przyjętych aksjomatów nie ma dwóch sprzecznych. Nie udało się nam wyprowadzić z tych postulatów sprzeczności, ale to nie dowodzi, że nigdy nie uda się uzyskać sprzeczności. Jest więc potrzebny dowód istnienia algebry kwaternionów.

Przypuśćmy, że mamy kwaternion

$$a + bi + cj + dk \in \mathcal{H}$$

gdzie $a, b, c, d \in \mathcal{R}$. Niech

$$\varphi(a + bi + cj + dk) = \begin{bmatrix} a + bi & -c - di \\ c - di & a - bi \end{bmatrix}.$$

Jeżeli przyjmijmy, że $w = a + bi$ oraz $z = c + di$, to definicję φ można wyrazić następująco:

$$\varphi(a + bi + cj + dk) = \varphi(w + zj) = \begin{bmatrix} w & -z \\ \overline{z} & \overline{w} \end{bmatrix}.$$

Nietrudno zauważyć, że φ jest funkcją liniową, i to różnowartościową. Trudniej zauważyć, że φ zachowuje mnożenie. Z wzoru (2) wynika, że

$$(w + zj)(v + xj) = wv + zjv + wxj + zjxj = (wv - z\overline{x}) + (z\overline{v} + wx)j. \tag{3}$$

Zauważmy także, że

$$\begin{bmatrix} w & -z \\ \overline{z} & \overline{w} \end{bmatrix} \begin{bmatrix} v & -x \\ \overline{x} & \overline{v} \end{bmatrix} = \begin{bmatrix} wv - z\overline{x} & -(z\overline{v} + wx) \\ \overline{z}v + \overline{wx} & \overline{wv} - \overline{z}x \end{bmatrix}.$$

Wzór ten oznacza, że φ zachowuje mnożenie. Tak więc φ jest izomorfizmem algebry $\mathcal{H}$ i pewnej algebry macierzy

$$\mathcal{H}' = \{ \begin{bmatrix} w & -z \\ \overline{z} & \overline{w} \end{bmatrix} : w, z \in \mathcal{C} \}$$

Twierdzenie 7.16 *$\mathcal{R}$-algebra $\mathcal{H}'$ ma własności wymagane od algebry kwaternionów.*

Dowód. Zbiór $\mathcal{H}'$ jest oczywiście podzbiorem algebry macierzy o wymiarach 2×2 i wyrazach zespolonych. Jest to podalgebra. Zamkniętość ze względu na dodawanie i zewnętrzne mnożenie jest oczywiste, a ze względu na mnożenie – wynika z wyżej przytoczonego wzoru.

Łatwo też wymienić kilka pierwiastków z -1 (z elementu przeciwnego do jedności algebry, czyli macierzy jednostkowej. Np. są to

$$\varphi(i) = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \quad \varphi(j) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \text{oraz} \quad \varphi(k) = \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix}.$$

Aby wykazać istnienie elementów odwrotnych zauważmy, że wyznacznik

$$\begin{vmatrix} w & -z \\ \overline{z} & \overline{w} \end{vmatrix} = |\, w\, |^2 + |\, z\, |^2$$

i jest dodatnią liczbą rzeczywistą dla wszystkich niezerowych elementów $\mathcal{H}'$. Ponadto macierz odwrotna wyrazą się wzorem

$$\begin{bmatrix} w & -z \\ \overline{z} & \overline{w} \end{bmatrix}^{-1} = \frac{1}{|\, w\, |^2 + |\, z\, |^2} \begin{bmatrix} \overline{w} & z \\ -\overline{z} & w \end{bmatrix}. \quad \square$$

Przedstawione dotychczas rozumowania świadczą także o tym, że każda algebra spełniająca podane na początku postulaty (algebra kwaternionów) jest izomorficzna z algebrą $\mathcal{H}'$.

7.10 Podstawowe pojęcia dotyczące kwaternionów

Podobnie jak dla liczb zespolonych, definiujemy kwaternion sprzężony z danym i moduł kwaternionu.

Kwaternionem sprzężonym do kwaternionu $h = a + bi + cj + dk$ nazywamy kwaternion

$$\overline{h} = a - bi - cj - dk.$$

Modułem kwaternionu $h = a + bi + cj + dk$ nazywamy zaś liczbę rzeczywistą

$$|\, h\,| = \sqrt{a^2 + b^2 + c^2 + d^2}.$$

Pojęcia te mają oczekiwane własności. Tak więc

$$\overline{h_1 + h_2} = \overline{h}_1 + \overline{h}_2, \quad \overline{ch} = c\overline{h} \;\; (c \in \mathcal{R}) \quad \text{oraz} \quad \overline{h_1 h_2} = \overline{h}_1 \overline{h}_2.$$

Ponadto

$$|\, h_1 h_2\,| = |\, h_1\,||\, h_2\,| \quad \text{oraz} \quad |\, w + zj\,|^2 = |\, w\,|^2 + |\, z\,|^2 \;\; (w, z \in \mathcal{C})$$

Zachodzi też wzór

$$h\overline{h} = |\, h\,|^2.$$

Wzór ten pozwala łatwo obliczać odwrotność kwaternionu. Wynika bowiem z niego, że

$$h^{-1} = \frac{1}{|\, h\,|^2}\overline{h}.$$

Definiujemy też część rzeczywistą $Re(h)$ i urojoną $Im(h)$ kwaternionu h: jeżeli $h = a + bi + cj + dk$, to

$$Re(h) = a \quad \text{oraz} \quad Im(h) = bi + cj + dk.$$

Zachodzą oczywiste wzory

$$Re(h) = \frac{h + \overline{h}}{2} \quad \text{oraz} \quad Im(h) = \frac{h - \overline{h}}{2}.$$

Cała algebra kwaternionów ma też część rzeczywistą $Re(\mathcal{H})$ identyczną (lub izomorficzną) z ciałem liczb rzeczywistych $\mathcal{R}$ i część urojoną

$$Im(\mathcal{H}) = \{Im(h) : h \in \mathcal{H}\} = \{bi + cj + dk \in \mathcal{H} : a, b, c \in \mathcal{R}\},$$

która – uważana za przestrzeń liniową – jest izomorficzna z przestrzenią $\mathcal{R}^3$.

Algebrę kwaternionów w naturalny spoób można utożsamiać ze zbiorem $\mathcal{R}^4$. Można więc wprowadzić też iloczyn skalarny kwaternionów. Aby nie doszło do kolizji oznaczeń, iloczyn ten powinien być oznaczany inaczej niż w przypadku wektorów. Dla kwaternionów

$$h_1 = a_1 + b_1 i + c_1 j + d_1 k \quad \text{oraz} \quad h_2 = a_2 + b_2 i + c_2 j + d_2 k$$

przyjmijmy więc, że ich iloczyn skalarny $\langle h_1, h_2\rangle$ jest dany wzorem

$$\langle h_1, h_2\rangle = a_1 a_2 + b_1 b_2 + c_1 c_2 + d_1 d_2.$$

7.11 Iloczyn wektorowy

Zauważmy, że

$$\begin{vmatrix} x & y & z \\ a & b & c \\ r & s & t \end{vmatrix} = x \cdot \begin{vmatrix} b & c \\ s & t \end{vmatrix} + y \cdot \begin{vmatrix} c & a \\ t & r \end{vmatrix} + z \cdot \begin{vmatrix} a & b \\ r & s \end{vmatrix}.$$

Przyjmijmy więc, że

$$(a,b,c) \times (r,s,t) = \left(\begin{vmatrix} b & c \\ s & t \end{vmatrix}, \begin{vmatrix} c & a \\ t & r \end{vmatrix}, \begin{vmatrix} a & b \\ r & s \end{vmatrix} \right).$$

Zdefiniowaną operację przekształcającą $\mathcal{R}^3$ w siebie nazywamy iloczynem wektorowym.

Lemat 7.17 *Zachodzi równość*

$$\langle (a,b,c), (a,b,c) \times (r,s,t) \rangle = 0,$$

a więc iloczyn wektorowy dwóch wektorów jest wektorem prostopadłym (ortogonalnym) do tych wektorów. □

Lemat 7.18 *Zachodzi wzór*

$$\begin{vmatrix} b & c \\ s & t \end{vmatrix}^2 + \begin{vmatrix} c & a \\ t & r \end{vmatrix}^2 + \begin{vmatrix} a & b \\ r & s \end{vmatrix}^2 = (a^2+b^2+c^2)(r^2+s^2+t^2) - (ar+bs+ct)^2.$$

Wynika z niego, że długość iloczynu wektorowego jest iloczynem długości mnożonych wektorów i wartości bezwzględnej sinusa kąta między tymi wektorami. □

Iloczyn wektorowy kwaternionów z $Im(\mathcal{H})$ wprowadzany analogicznie, tak aby był zachowywany przez izomorfizm tej przestrzeni i $\mathcal{R}^3$. Przyjmujemy więc, że

$$(ai+bj+ck) \times (ri+sj+tk) = \begin{vmatrix} b & c \\ s & t \end{vmatrix} i + \begin{vmatrix} c & a \\ t & r \end{vmatrix} j + \begin{vmatrix} a & b \\ r & s \end{vmatrix} k.$$

Wniosek 7.19 *Jeżeli $u, v \in Im(\mathcal{H})$, to*

$$|\, u \times v \,|^2 = |\, u \,|^2 |\, v \,|^2 \sin^2 \varphi,$$

gdzie φ jest kątem między wektorami współrzędnych kwaternionów u i v. □

Twierdzenie 7.20 *Jeżeli $u, v \in Im(\mathcal{H})$, to*

$$uv = -\langle u, v \rangle + u \times v.$$

Dowód. Wystarczy napisać wzór na iloczyn dwóch urojonych kwaternionów. □

Wniosek 7.21 *Jeżeli $u, v \in Im(\mathcal{H})$, to*

$$uv + vu = -2\langle u, v \rangle. \ \square$$

7.12 Analiza pewnych przekształceń

Przyjmijmy, że $q = a + bi + cj + dk$ jest kwaternionem o module 1. Po pierwsze zajmiemy się przekształceniami $\alpha_q : \mathcal{H} \to \mathcal{H}$ zdefiniowanymi wzorami

$$\alpha_q(x) = qx.$$

Są to oczywiście przekształcenia liniowe. Zauważmy, że

$$\begin{aligned}
\alpha_q(1) &= a + bi + cj + dk,\\
\alpha_q(i) &= -b + ai + dj - ck,\\
\alpha_q(j) &= -c - di + aj + bk,\\
\alpha_q(k) &= -d + ci - bj + ak.
\end{aligned}$$

Łatwo sprawdzić, że iloczyn skalarny każdej pary złożonej z dwóch różnych takich wektorów jest równy 0, zaś kwadrat skalarny każdego z tych wektorów, czyli kwadrat modułu jest równy $|\, q\,|^2 = a^2 + b^2 + c^2 + d^2$.

Macierzą przekształcenia α_q w bazie $1, i, j, k$ jest

$$A_q = \begin{bmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{bmatrix}.$$

Z faktów przytoczonych w poprzednim akapicie wynika, że

$$|\, A_q\,| = \pm|\, q\,|^4 = \pm 1.$$

Aby się o tym przekonać, wystarczy wyliczyć iloczyn $A_q^T A_q$.

Teraz zajmiemy się przekształceniami $\beta_q : \mathcal{H} \to \mathcal{H}$ zdefiniowanymi wzorami

$$\beta_q(x) = xq^{-1}.$$

Oczywiście, to też są przekształcenia liniowe. Ponieważ $|q| = 1$, więc $q^{-1} = \overline{q} = a - bi - cj - dk$. Zauważmy, że

$$\begin{aligned}
\beta_q(1) &= a - bi - cj - dk,\\
\beta_q(i) &= b + ai + dj - ck,\\
\beta_q(j) &= c - di + aj + bk,\\
\beta_q(k) &= d + ci - bj + ak.
\end{aligned}$$

Macierzą przekształcenia β_q jest

$$B_q = \begin{bmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{bmatrix}.$$

Porównując macierze A_q i B_q można zauważyć, że jedną z nich można otrzymać drugą mnożąc pierwszą kolumnę przez -1, a następnie mnożąc przez -1 pierwszą kolumnę. Te operacje nie zmieniają wyznacznika. Tak więc $|\, A_q\,| = |\, B_q\,|$. Wobec tego, wyznacznik macierzy $A_q B_q$ jest równy 1.

Zauważmy, że

$$A_qB_q = \begin{bmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{bmatrix} \begin{bmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & C_q & \\ 0 & & & \end{bmatrix}.$$

Macierz A_qB_q jest macierzą przekształcenia

$$O_q(x) = qxq^{-1}.$$

Oczywiście, dla każdej liczby rzeczywistej $c \neq 0$ mamy

$$O_{cq}(x) = O_q(x).$$

Pozwala to analizować dalej tylko przekształcenia O_q dla kwaternionów q o module 1.

Lemat 7.22 *Funkcja O_q jest obrotem w przestrzeni $Im(\mathcal{H})$, a więc $O_q : \Im(\mathcal{H}) \to Im(\mathcal{H})$, jest to przekształcenie liniowe i wyznacznik macierzy tego przekształcenia jest równy 1.*

Dowód. Macierzą przekształcenia O_q (wyznaczoną przez bazę $1, i, j, k$) jest A_qB_q. Z wzoru na tę macierz wynika, że funkcja O_q na zbiorze $Im(\mathcal{H})$ przyjmuje wartości należące do $Im(\mathcal{H})$, a macierzą obcięcia tej funkcji do $Im(\mathcal{H})$ jest macierz C_q. Oczywiście, $|\, C_q \,| = |\, A_qB_q \,| = 1$. □

7.13 Oś i kąt obrotu O_q

Przypuśćmy, że $Re(q) = r$ i $Im(q) = u$, a ponadto, że $a \in Im(\mathcal{H})$. Wtedy

$$\begin{aligned} (r+u)a(r-u) &= rar - rau + uar - uau = r^2a + r(ua - au) - uau = \\ &= r^2a + r((-\langle u, a\rangle + u \times a) - (-\langle a, u\rangle + a \times u)) - (-au - 2\langle u, a\rangle)u = \\ &= r^2a + 2r(u \times a) - \langle u, u\rangle a + 2\langle a, u\rangle u. \end{aligned}$$

Równości te wynikają z twierdzenia 7.20 i wniosku 7.21, oraz z oczywistych równości $\overline{u} = -u$ dla $u \in Im(\mathcal{H})$ i $u \times a = -a \times u$.

Twierdzenie 7.23 *Jeżeli moduł $q \in \mathcal{H} \backslash \{\pm 1\}$ jest równy 1 i $Re(q) = \cos\varphi$, to przekształcenie O_q (przestrzeni $Im(\mathcal{H})$ w siebie) jest obrotem o kąt 2φ względem osi $Im(q)$.*

Dowód. Przyjmujemy wyżej wprowadzone oznaczenia i będziemy korzystać z udowodnionego wzoru. Zauważmy, że

$$O_q(u) = (r+u)u(r-u) = r^2u + 2r(u \times u) - \langle u, u\rangle u + 2\langle u, u\rangle u = (r^2 + \langle u, u\rangle)u = |\, q \,|^2 u = u.$$

Oznacza to, że wektor $u = Im(q)$ jest osią obrotu przekształcenia O_q.

Aby ustalić kąt obrotu weźmy dowolne $a \in Im(\mathcal{H})$ prostopadłe do u, a więc spełniające $\langle a, u\rangle = 0$. Wtedy

$$\begin{aligned} O_q(a) &= r^2a + 2r(u \times a) - \langle u, u\rangle a + 2\langle a, u\rangle u(r^2 - \langle u, u\rangle)a + 2r|\, u \,|(\frac{u}{|\, u \,|} \times a) = \\ &= (\cos^2\varphi - \sin^2\varphi)a + 2\cos\varphi\sin\varphi(\frac{u}{|\, u \,|} \times a) = \cos 2\varphi\; a + \sin 2\varphi(\frac{u}{|\, u \,|} \times a) \end{aligned} \qquad (4)$$

(pamiętajmy, że wektor $\frac{u}{|\,u\,|} \times a$ jest prostopadły do u i a). Udowodniony wzór świadczy o tym, że przekształcenie O_q obraca wektory prostopadłe do u o kąt 2φ.

W tej chwili nie potrafimy dokładniej opisać obrotu O_q, głównie dlatego, że nie został wprowadzony aparat pozwalający zbadać, czy O_q obraca zgodnie z „ruchem wskazówek zegara", czy przeciwnie. □

Z wyprowadzonych wzorów można też wywnioskować, że każdy obrót w przestrzeni $Im(\mathcal{H})$ (dookoła zera) jest funkcją postaci O_q. Jeżeli chcemy zdefiniować obrót o kąt φ względem osi u, to obliczamy $r = \cos\frac{\varphi}{2}$, mnożąc u przez odpowiednią liczbę gwarantujemy, aby $r^2 + \langle u, u\rangle = 1$ i definiujemy q jako $r + u$.

7.14 Grupa obrotów w $\mathcal{R}^3$

Niech S oznacza zbiór kwaternionów o module 1. Nietrudno zauważyć, że S jest grupą. Przyjmijmy, że

$$h(q) = O_q | Im(\mathcal{H}).$$

Tak więc h przekształca S w grupę obrotów (dookoła 0) w przestrzeni $Im(\mathcal{H})$, a wartość $h(q)$ to obcięcie O_q do odpowiedniej przestrzeni).

Twierdzenie 7.24 *Funkcja h jest epimorfizmem grupy S na grupę obrotów przestrzeni $Im(\mathcal{H})$ o jądrze złożonym z dwóch elementów ± 1 (a więc każdy obrót jest wyznaczony przez dwa kwaterniony różniące się znakiem).*

Dowód. Prawie cała teza wynika z poprzedniego rozdziału. Udowodnię więc tylko ostatnią część tezy. W tym celu wystarczy wykazać, że jeżeli $qxq^{-1} = x$ dla wszystkich x, to $q = \pm 1$. Przypuśćmy więc, że q spełnia podany warunek. Wtedy w szczególności $qi = iq$. Oznacza to, że $q \in \mathcal{C}$. Przyjmijmy więc, że $q = a + bi$. Teraz równość $qj = jq$ jest równoważna równości $aj + bk = aj - bk$. Wynika z niej, że $b = 0$ i $q = a \in \mathcal{R}$. Ponieważ moduł q jest równy 1, więc $q = \pm 1$. □

8 Przykłady obliczeń

8.1 Pierwiastki z -1

Przedstawione wcześniej rozważania sugerują, że obrót o 180 stopni wyznacza kwaternion, który jest pierwiastkiem z -1. Zbadajmy więc pierwiastki z -1.

Oczywiście,

$$\begin{aligned}
&(a + bi + cj + dk)^2 = \\
&= a^2 - b^2 - c^2 - d^2 + 2abi + 2acj + 2adk + bc(ij + ji) + bd(ik + ki) + cd(jk + kj) = \\
&= a^2 - b^2 - c^2 - d^2 + 2abi + 2acj + 2adk.
\end{aligned}$$

Jeżeli więc $(a + bi + cj + dk)^2 = -1$, to $a^2 - b^2 - c^2 - d^2 = -1$ oraz $ab = ac = ad = 0$. Zauważmy, że założenie $a \neq 0$ prowadzi do sprzeczności. Wtedy bowiem $b = c = d = 0$ i liczba rzeczywista a byłaby pierwiastkiem z -1. Wobec tego, $a = 0$ i $b^2 + c^2 + d^2 = 1$. Pierwiastkami z -1 są więc kwaterniony o części rzeczywistej 0 i module 1. Łatwo też zauważyć, że wszystkie takie kwaterniony są pierwiastkami z -1.

8.2 Przykład obrotu

Spróbujmy znaleźć kwaternion q opisujący obrót o 90 stopni dookoła osi (przechodzącej przez początek układu współrzędnych, tylko takimi obrotami zajmowaliśmy się do tej pory) o kierunku wektora j. Część rzeczywista kwaternionu q to $\cos(45°) = \frac{\sqrt{2}}{2}$. Cały kwaternion ma postać $\frac{\sqrt{2}}{2} + bj$. Aby miał moduł 1, to należy przyjąć, że $b = \pm\frac{\sqrt{2}}{2}$. Przyjęcie, że $b = -\frac{\sqrt{2}}{2}$ zmieni kierunek osi na przeciwny. Interesującym nas kwaternionem jest więc

$$q = \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}j.$$

Spróbujemy teraz ustalić, jak obrót O_q przekształca punkt i. W tym celu należy wyliczyć qiq^{-1}. Dla kwaternionów o module 1 zachodzi wzór $q^{-1} = \overline{q}$. Tak więc

$$\begin{aligned} O_q(i) = qiq^{-1} = qi\overline{q} &= \left(\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}j\right) i \left(\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}j\right) = \frac{1}{2}(1+j)i(1-j) = \\ &= \frac{1}{2}(1+j)(i-k) = \frac{1}{2}(i-k+ji-jk) = \frac{1}{2}(i-k-k-i) = -k. \end{aligned}$$

Okazało się więc, że zgodnie z rysunkiem ze strony 7.1, kwaternion q opisuje obrót o 90 stopni zgodnie z ruchem wskazówek zegara wokół osi o kierunku j.

8.3 Składanie obrotów

Spróbujemy teraz ustalić, czemu jest równe złożenie dwóch obrotów: najpierw obrotu o 90 stopni wokół osi o kierunku j, a następnie też o 90 stopni wokół osi o kierunku i (i osie przechodzą przez 0, a obroty są zgodne z ruchem wskazówek zegara). Obroty te są wyznaczone przez kwaterniony

$$q = \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}j$$

oraz odpowiednio

$$q' = \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}i.$$

Najpierw zbadajmy, co te obroty przekształcają punkt $k+j$. Okazuje się, że

$$O_q(j+k) = \frac{1}{2}(1+j)(j+k)(1-j) = \frac{1}{2}(1+j)(j-j^2+k-kj) = \frac{1}{2}(1+j)(i+j+k+1) =$$

$$= \frac{1}{2}(i+j+k+1+ji+j^2+jk+j) = \frac{1}{2}(i+j+k-k+i+j) = i+j.$$

Podobnie,

$$O_{q'}(O_q(i)) = O_{q'}(i+j) = k+i = i+k.$$

Złożenie tych obrotów jest opisane przez kwaternion

$$q'q = \left(\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}i\right)\left(\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}j\right) = \frac{1}{2}(1+i)(1+j) = \frac{1}{2} + \frac{1}{2}(i+j+k).$$

Oczywiście, $\frac{1}{2} = \cos(60°)$. Tak więc złożenie tych obrotów jest obrotem o 120 stopni wokół osi o kierunku $i+j+k$.

8.4 Orientacja trójki wektorów

8.4.1 Podstawowe definicje

Będziemy zajmować się teraz następującym problemem: mamy dany kwaternion q (na przykład o module 1) i chcemy narysować lub wyobrazić sobie wektor $O_q(x)$ dla danego wektora x. Albo inaczej, chcemy dokładnie opisać obrót wyznaczony przez kwaternion q. Wymaga to wprowadzenia pojęcia orientacji trójki wektorów. Będziemy teraz rozważać wyłącznie trójki wektorów z przestrzeni $\mathcal{R}^3$, które nie są zawarte w pewnej płaszczyźnie, a więc składające się z wektorów liniowo niezależnych. Można je jeszcze inaczej scharakteryzować jako trójki $\vec{a}, \vec{b}, \vec{c}$ takie, że wyznacznik macierzy $[\vec{a}, \vec{b}, \vec{c}]$ (w której współrzędne danych wektorów są zapisane w kolejnych kolumnach) jest różny od zera.

Będziemy mówić, że trójki $\vec{a}_1, \vec{b}_1, \vec{c}_1$ i $\vec{a}_2, \vec{b}_2, \vec{c}_2$ mają tę samą orientację, jeżeli wyznaczniki macierzy $[\vec{a}_1, \vec{b}_1, \vec{c}_1]$ i $[\vec{a}_2, \vec{b}_2, \vec{c}_2]$ są tego samego znaku.

Będziemy też mówić, że trójka wektorów $\vec{a}, \vec{b}, \vec{c}$ ma orientację dodatnią (ujemną), jeżeli macierz $[\vec{a}, \vec{b}, \vec{c}]$ ma wyznacznik dodatni (lub odpowiednio: ujemny).

8.4.2 Trójki tak samo zorientowane

Spróbujemy teraz zrozumieć te pojęcia. Zauważmy najpierw, że trójka wektorów z bazy standardowej $\vec{e}_1, \vec{e}_2, \vec{e}_3$ ma orientację dodatnią.

Przyjmijmy teraz, że mamy dwie trójki $\vec{a}_1, \vec{b}_1, \vec{c}_1$ i $\vec{a}_2, \vec{b}_2, \vec{c}_2$ ortonormalne, a więc spełniające dodatkowe warunki: $\vec{a}_i\vec{b}_i = \vec{a}_i\vec{c}_i = \vec{b}_i\vec{c}_i = 0$ oraz $\vec{a}_i^2 = \vec{b}_i^2 = \vec{c}_i^2 = 1$ dla $i = 1, 2$ (czyli złożone z wektorów długości 1 i parami prostopadłych).

Niech f będzie funkcją liniową zdefiniowaną wzorami $f(\vec{a}_1) = \vec{a}_2$, $f(\vec{b}_1) = \vec{b}_2$ oraz $f(\vec{c}_1) = \vec{c}_2$.

Lemat 8.1 *Funkcja f jest dobrze zdefiniowaną funkcją liniową i macierz tej funkcji ma wyznacznik równy albo* 1, *albo* -1.

Dowód. Ortonormalny układ wektorów jest liniowo niezależny. Wektory $\vec{a}_1, \vec{b}_1, \vec{c}_1$ stanowią więc bazę. Implikuje to, że definicja funkcji f jest poprawna i w rzeczywistości została zdefiniowna bijekcja przestrzeni $\mathcal{R}^3$. Niech A będzie macierzą funkcji f w bazie standardowej.

Wyznacznik macierzy $[\vec{a}, \vec{b}, \vec{c}]$ (której kolumny są wymienione współrzędnymi podanych wektorów w wybranej bazie) nie zależy od wyboru bazy, w której przedstawiamy poszczególne wektory. Także wyznacznik macierzy funkcji f nie zależy od wyboru bazy. Stąd wynika, że wyznacznik macierzy A jest taki sam jak wyznacznik macierzy f w bazie $\vec{a}_1, \vec{b}_1, \vec{c}_1$, a więc jest taki sam jak wyznacznik macierzy $B = [\vec{a}_2, \vec{b}_2, \vec{c}_2]$.

Wyznacznik macierzy B łatwo można obliczyć. Wystarczy zauważyć, że iloczyn $B^T B$ jest macierzą jednostkową. Stąd $\mid B \mid^2 = 1$. □

Wniosek 8.2 *Ortonormalne trójki $\vec{a}_1, \vec{b}_1, \vec{c}_1$ i $\vec{a}_2, \vec{b}_2, \vec{c}_2$ mają tę samą orientację wtedy i tylko wtedy, gdy funkcja f jest obrotem.*

Dowód. Wyznaczniki macierzy o ortonormalnych kolumnach są równe ± 1, jeżeli dla dwóch takich macierzy wyznaczniki są tego samego samego znaku, to są równe. Jeżeli A jest macierzą funkcji f, to także

$$A[\vec{a}_1, \vec{b}_1, \vec{c}_1] = [\vec{a}_2, \vec{b}_2, \vec{c}_2].$$

Stąd, po przejściu do wyznaczników otrzymujemy, łatwo wywnioskować tezę. □

Tak więc dla ortonormalnych trójek wektorów zgodność orientacji oznacza, że jedną z tych trójek można tak obrócić, aby otrzymać drugą. Mamy także oczywisty

Lemat 8.3 *Jeżeli $\alpha, \beta, \gamma > 0$, to trójki $\alpha\vec{a}, \beta\vec{b}, \gamma\vec{c}$ oraz $\vec{a}, \vec{b}, \vec{c}$ mają tę samą orientację.* □

Tak więc skrócenie bądź wydłużenie wektorów nie zmienia orientacji, Co więcej, orientacji nie zmienia ortogonalizacja trójki.

Lemat 8.4 *Trójki $\vec{a}, \vec{b}, \vec{c}$ oraz*

$$\vec{a}, \vec{b} - \frac{\vec{a}\vec{b}}{\vec{a}\vec{a}}\vec{a}, \vec{c} - \frac{\vec{a}\vec{c}}{\vec{a}\vec{a}}\vec{a} - \frac{\vec{d}\vec{c}}{\vec{d}\vec{d}}\vec{d}$$

dla $\vec{d} = \vec{b} - \frac{\vec{a}\vec{b}}{\vec{a}\vec{a}}\vec{a}$, mają tę samą orientację.

Dowód. Ta druga trójka powstaje w wyniku zastosowania do pierwszej metody ortoganalizacji Grama-Schmita. Łatwo sprawdzić bezpośrednio, że ma ortogonalne elementy. Tezę lematu dowodzimy bez trudu korzystając ze znanego faktu, że dodając do wiersza macierzy wielokrotność innego nie zmieniamy wyznacznika. □

Teraz możemy intuicyjnie wyjaśnić pojęcie tej samej orientacji. Dwie trójki mają tę samą orientację, jeżeli po „wyprostowaniu" (czyli ortogonalizacji) i wyrównaniu długości, jedną z nich można tak obrócić, aby otrzymać drugą.

8.4.3 Potrzebne własności kwaternionów

Jak wiemy, kwaterniony urojone z $Im(\mathcal{H})$ można uważać za elementy przestrzeni $\mathcal{R}^3$. Bazą standardową w przestrzeni $Im(\mathcal{H})$ jest układ złożony z wektorów i, j, k. Można więc do kwaternionów urojonych stosować wprowadzone pojęcia. W dalszych rozważaniach będą nam potrzebne dwa fakty dotyczące kwaternionów.

Lemat 8.5 *Jeżeli kwaterniony $u, v \in Im(\mathcal{H})$ utożsamimy z ciągiem współrzednych względem bazy i, j, k, to wyznacznik macierzy $[u, v, u \times v]$ jest równy 0 wtedy i tylko wtedy, gdy u jest wielokrotnością v lub odwrotnie (gdy u i v są współliniowe). Co więcej, jeżeli u i v nie są współliniowe, to trójki $u \times v, u, v$ oraz $u, v, u \times v$ mają orientację dodatnią.*

Dowód. Przyjmijmy, że $u = ai + bj + ck$ oraz $v = ri + sj + tk$. Jeżeli teraz przyjmiemy, że $u \times v = xi + yj + zk$, to zgodnie z definicją iloczynu wektorowego (str. 19) otrzymujemy, że

$$|\,[u \times v, u, v]^T\,| = \begin{vmatrix} x & y & z \\ a & b & c \\ r & s & t \end{vmatrix} = \begin{vmatrix} b & c \\ s & t \end{vmatrix}^2 + \begin{vmatrix} c & a \\ t & r \end{vmatrix}^2 + \begin{vmatrix} a & b \\ r & s \end{vmatrix}^2.$$

Stąd już łatwo otrzymać tezę lematu. □

Lemat 8.6 *Przypuśćmy, że q jest kwaternionem różnym od ± 1, o module 1, o części rzeczywistej $r = \cos\varphi > 0$ ($\varphi \in (0, \pi/2)$) i części urojonej u. Niech a będzie niezerowym kwaternionem urojonym, który nie jest współliniowy z u. Wtedy trójka $u, a, O_q(a)$ ma orientację dodatnią.*

Dowód. Mamy więc kwaternion q opisujący obrót wokół osi u o kąt 2φ i interesuje nas trójka złożona z osi obrotu, obracanego wektora a (nie leżącego na osi obrotu) i wektora $O_q(a)$ otrzymanego z a w wyniku wykonania obrotu opisanego kwaternionem q.

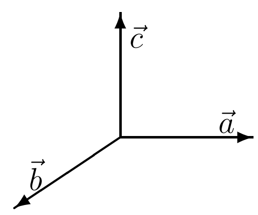
Aby dowieść lemat, wyliczymy wyznacznik $|u, a, O_q(a)|$ macierzy $[u, a, O_q(a)]$ (z kolumnami wypełnionymi współrzędnymi podanych kwaternionów). Zgodnie ze wzorem (4) mamy

$$|\, u, a, O_q(a)\,| = |\, u, a, \cos 2\varphi\; a + \frac{\sin 2\varphi}{|\,u\,|}(u \times a)\,| = \frac{\sin 2\varphi}{|\,u\,|}|\, u, a, u \times a\,|.$$

Teraz wystarczy skorzystać z poprzedniego lematu i oczywistych własności funkcji trygonometrycznych. □

8.4.4 Kwaterniony a przestrzeń fizyczna

Przypuśćmy, że w przestrzeni (fizycznej) wybraliśmy początek układu współrzędnych i trzy osie. Zamierzamy teraz wybrać trzy wektory jednostkowe wyznaczające kierunki osi. Możemy to zrobić na dwa sposoby:



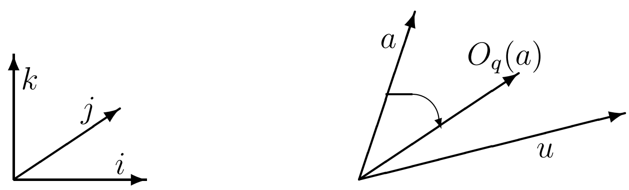
Na pierwszym rysunku wektory $\vec{a}, \vec{b}, \vec{c}$ są ułożone tak, jak kciuk, palec wskazujący i odpowiednio palec środkowy prawej ręki (wektor $\vec{b}$ jest prostopadły do pozostałych i wskazuje kierunek, w którym oddalamy się od osoby obserwującej). Na drugim rysunku, układ wektorów jest taki, jak palców lewej ręki.

O wektorach $\vec{a}, \vec{b}, \vec{c}$ z pierwszego (lewego) rysunku mówimy, że mają orientację prawostronną. O wektorach z drugiego rysunku, że mają orientację lewostronną.

Wektory z bazy standardowej zawsze mają orientację dodatnią. Na rysunku lub w rzeczywistości możemy im nadać albo orientację lewostronną, albo prawostronną.

Mając na dwa sposoby zdefiniowane pojęcie orientacji, także na dwa sposoby możemy definiować pojęcie tej samej orientacji. Oba zdefiniowane pojęcia tej samej orientacji powinny się pokrywać. Po pierwsze dlatego, że opisany w rozdziale 8.4.2 sposób testowania tej samej orientacji („wyprostowanie", wyrównanie i stosowne obrócenie trójki) powinien zachowywać także orientację zdefiniowaną za pomocą układów palców. Po drugie dlatego, że trudno sobie wyobrazić trójkę, która jest jednocześnie prawo- i lewostronna. Oznacza to, że jeżeli wektory bazowe mają orientację prawostronną, to trójki wektorów o orientacji dodatniej pokrywają się z trójkami prawostronnymi. Jeżeli wektory bazowe mają orientację lewostronną, to sytuacja jest analogiczna: trójki o orientacji dodatniej to trójki lewostronne.

Przyjmijmy, że trójka kwaternionów i, j, k (czyli wektorów w trójwymiarowej przestrzeni $Im(\mathcal{H})$) ma orientację prawostronną, a więc taką, jak na poniższym rysunku.



Wtedy wszystkie trójki o orientacji dodatniej też mają orientację prawostronną. Zgodnie z lematem 8.6, zachowując oznaczenia i założenia sformułowane w tym lemacie, trójka kwaternionów $u, a, O_q(a)$ złożona z osi obrotu opisanego kwaternionem q, wybranego wektora (kwaternionu) a i wektora obróconego $O_q(a)$ ma orientację dodatnią. Powinna więc zostać narysowana jako trójka o orientacji prawostronnej. A więc tak, jak na rysunku. Nietrudno zauważyć, że kwaternion $q = \cos\varphi + u$, spełniający założenia lematu 8.6, definiuje obrót wokół osi u, o kąt 2φ, w kierunku zgodnym z ruchem wskazówek zegara.

9 Literatura

[1] Heintz-Dieter Ebbinghaus, Hans Hermes, Friedrich Hirzebruch, Max Koecher, Klaus Mainzer Jürgen Neukirch, Alexandr Prestel, Reinhold Remmert, *Numbers*, Springer-Verlag, 1991,

[2] Witold Więsław, *Grupy, pierścienia, ciała*, Wydawnictwo Uniwersytetu Wrocłaskiego, 1983,

[3] Nathan Jacobson, *Lectures in Abstract Algebra*, D. van Nostrand Company, inc., 1951