Kwaterniony i ich zastosowania

M. W. Kalinowski
e-mail: markwkal@bioexploratorium.pl

I Algebra kwaternionów

Definicja 1. *Algebrą nad ciałem* $\mathscr{C}$ nazywamy ciąg uporządkowany

$$\mathscr{A} = (\mathscr{W}, \mathscr{C}, *), \tag{I.1}$$

gdzie

$$\mathscr{W} = (A, \mathscr{C}, \square, \circ) \tag{I.2}$$

jest przestrzenią wektorową nad ciałem $\mathscr{C}$,

$$\mathscr{C} = (C, +, \cdot) \tag{I.3}$$

jest ciałem, $*$ jest działaniem w zbiorze A, powiązanym z pozostałymi działaniami:

$$\bigwedge_{\lambda\in C} \bigwedge_{a,b\in A} \lambda \circ (a * b) = (\lambda \circ a) * b = a * (\lambda \circ b) \tag{I.4}$$

$$\bigwedge_{a,b,c\in A} c * (a \square b) = (c * a) \square (c * b) \tag{I.5}$$

$$\bigwedge_{a,b,c\in A} (a \square b) * c = (a * c) \square (b * c) \tag{I.6}$$

— rozdzielność prawo- i lewostronna.

$$\bigwedge_{a\in A} \bigwedge_{\mu,\lambda\in C} (\lambda \cdot \mu) \circ (a * b) = (\lambda \circ a) * (\mu \circ b) = \lambda \circ (\mu \circ (a * b)). \tag{I.7}$$

$*$ może być nieprzemienne

$$a * b \neq b * a \tag{I.8}$$

oraz niełączne

$$a * (b * c) \neq (a * b) * c, \tag{I.9}$$

może też mieć dzielniki zera

$$a_1 * a_2 = \circledcirc, \tag{I.10}$$

gdzie $a_1 \neq \circledcirc$, $a_2 \neq \circledcirc$ i $\circledcirc$ jest zerem w grupie abelowej $(A, \square)$.

Działanie $*$ może być odwracalne dla $a \neq \circledcirc$, tj. istnieje b, że

$$a * b = \mathbb{1}, \tag{I.11}$$

gdzie $\mathbb{1}$ jest jedynką w algebrze A,

$$\bigwedge_{a \in A} \mathbb{1} * a = a * \mathbb{1} = a. \tag{I.12}$$

Gdy istnieją dzielniki zera i mamy powyższą własność, elementy odwrotne nie istnieją dla dzielników zera. Wszystkie aksjomaty dla przestrzeni wektorowej $\mathscr{W}$ i ciała $\mathscr{C}$ są spełnione.

Ciąg uporządkowany

$$\mathscr{P} = (A, \square, *)$$

jest pierścieniem w ogólności nieprzemiennym ze względu na mnożenie „$*$". Istnieją uogólnienia tego pojęcia na pierścienie niełączne, aczkolwiek zakładamy często, że pierścień ten posiada jedynkę $\mathbb{1}$ ze względu na mnożenie „$*$". Posiada on także pewien podzbiór elementów odwracalnych ze względu na „$*$". Może jednak mieć dzielniki zera. Pierścień ten może posiadać antyautomorfizm inwolutoryczny. Uogólnienia pojęcia łączności mnożenia mogą być następujące: *alternatywność* $x^2 * y = x * (x * y)$, lub *giętkość* $(x * y) * x = x * (y * x)$.

Definicja 2. *Wymiarem algebry* (*algebraicznym*) nazywamy wymiar przestrzeni wektorowej $\mathscr{W}$.

Wymiar może być skończony lub nieskończony.

Definicja 3. *Algebrą topologiczną* $\overline{\mathscr{A}}$ nad ciałem topologicznym $\overline{\mathscr{C}}$ nazywamy ciąg uporządkowany

$$\overline{\mathscr{A}} = (\mathscr{A}, \overline{\mathscr{C}}, \mathscr{T}), \tag{I.13}$$

gdzie $\overline{\mathscr{C}}$ jest ciałem topologicznym

$$\overline{\mathscr{C}} = (\mathscr{C}, \mathscr{T}_1) \tag{I.14}$$

$$\overline{\mathscr{W}} = (\mathscr{W}, \mathscr{T}), \tag{I.15}$$

a $\overline{\mathscr{W}}$ jest przestrzenią wektorową (liniową) topologiczną. Działanie $*$ jest ciągłe w topologii $\mathscr{T}$. $\mathscr{T}_1$ jest topologią w C, a $\mathscr{T}$ topologią w zbiorze A.

Wszystkie działania są ciągłe w odpowiednich topologiach. Np. „$\circ$" w iloczynie kartezjańskim topologii $\mathscr{T}_1 \times \mathscr{T}$.

Definicja 4. *Algebrą unormowaną* nazywamy algebrę topologiczną, w której topologia $\mathscr{T}$ jest generowana przez normę $\|\cdot\|$, a topologia $\mathscr{T}_1$ jest generowana przez normę $|\cdot|$ (moduł). Jeśli topologia jest zupełna, to algebrę nazywamy *zupełną*, jeśli zwarta, to *zwartą*, *ośrodkową*, *polską*, *spójną* itp.

Zakładamy, że

$$\bigwedge_{a,b \in A} \|a * b\| \leqslant \|a\| \cdot \|b\|. \tag{I.16}$$

Algebrę można przenormować tak, aby

$$\|a * b\| = \|a\| \cdot \|b\|.$$

Można określić też wymiar topologiczny A (pokryciowy, indukcyjny mały lub duży) i jej podzbiorów. Wymiar topologiczny (w ogólności) nie musi być równy algebraicznemu (tj. mocy bazy Hamela przestrzeni wektorowej $\mathscr{W}$). Ciało $\mathscr{C}$ w zastosowaniach jest najczęściej ciałem liczb rzeczywistych $\mathbb{R}$ lub zespolonych $\mathscr{C}$. Możemy jednak rozpatrywać algebry nad ciałem liczb wymiernych lub nad ciałem liczb p-adycznych.

Od tego momentu zakładamy, że rozpatrywana algebra jest nad ciałem liczb rzeczywistych. Będziemy również zakładać, że topologia $\mathscr{T}$ generowana przez normę $\|\cdot\|$ jest zupełna. Jeśli nie jest, to łatwo możemy rozpatrywać jej uzupełnienie. Dla normy $\|\cdot\|$ zakładamy również

$$\bigwedge_{\lambda\in C}\bigwedge_{a\in A}\|\lambda\circ a\| = |\lambda|\,\|a\|. \tag{I.17}$$

Anty-automorfizmem inwolutorycznym w algebrze $\mathscr{A}$ nazywamy odwracalny anty-automorfizm algebry $\mathscr{A}$ taki, że

$$\Phi : A \to A$$
$$\bigwedge_{x\in A}\Phi(\Phi(x)) = x, \quad \Phi^2 = \mathrm{id},$$
$$\bigwedge_{x,y\in A}\Phi(x * y) = \Phi(y) * \Phi(x),\ \Phi(x \mathbin{\square} y) = \Phi(x) \mathbin{\square} \Phi(y),$$
$$\bigwedge_{\lambda\in C, x\in A}\Phi(\lambda\circ x) = \varphi(\lambda)\circ\Phi(x),$$

gdzie φ jest anty-automorfizmem inwolutorycznym w ciele $\mathscr{C}$. Anty-automorfizm taki możemy określić w algebrze topologicznej $\overline{\mathscr{A}}$, wtedy zakładamy, że jest on ciągły w topologii $\mathscr{T}$.

W przypadku algebry unormowanej zupełnej (Banacha) dodatkowo mamy

$$\|\Phi(x)\| \leqslant \|\Phi\|\,\|x\|$$

(Φ jest ciągły, $\|\Phi\|$ istnieje). Dodatkowo zakładamy, że

$$\Phi(x)x = x\Phi(x) = \|x\|^2,$$

wtedy $\|\Phi\| = 1$ i mamy

$$\|\Phi(x)\| = \|x\|$$

(zgodność normy z działaniem i Φ).

W algebrze z dzieleniem (gdzie istnieją elementy odwrotne dla każdego $x \neq 0$) możemy określić

$$\Phi(x) = \|x\|^2\cdot x^{-1},\ x\neq 0,$$
$$\Phi(\circledcirc) = \circledcirc.$$

Łatwo udowodnić, że

$$\Phi(x) * x = x * \Phi(x) = \|x\|^2, \|\Phi\| = 1.$$

Często używamy oznaczenia

$$\Phi(x) = \overline{x}.$$

Definicja 5. *Algebrą kwaternionów* $\mathbb{H}$ nazywamy algebrę topologiczną nad ciałem *liczb rzeczywistych*

$$\mathbb{H} = (\mathscr{R}^4, \mathscr{R}, *, \|\cdot\|, \Phi), \tag{I.18}$$

gdzie $\mathscr{R}^4 = (\mathbb{R}^4, \mathbb{R}, \square, \circ)$ jest czterowymiarową rzeczywistą przestrzenią wektorową nad ciałem liczb rzeczywistych ze zwykłą definicją $\square$ i $\circ$, $\mathscr{R} = (\mathbb{R}, +, \cdot, |\ |)$, $|\cdot|$ oznacza zwykły moduł w zbiorze liczb rzeczywistych.

$$\|a\| = \sqrt{a_0^2 + a_1^2 + a_2^2 + a_3^2} \tag{I.19}$$

oznacza zwykłą normę euklidesową w $\mathbb{R}^4$.

Działanie $*$ możemy określić w następujący sposób. Wprowadzamy w $\mathbb{R}^4$ bazę ortonormalną e_0, e_1, e_2, e_3 w iloczynie skalarnym generującym normę euklidesową (I.19),

$$(a, b) = a_0 b_0 + a_1 b_1 + a_2 b_2 + a_3 b_3 \tag{I.20}$$

$$\|a\| = \sqrt{(a, a)} \tag{I.21}$$

tak, że

$$(e_i, e_j) = \delta_{ij}, \tag{I.22}$$

i oznaczamy

$$\begin{aligned} e_0 &= 1 \\ e_1 &= \mathrm{i} \\ e_2 &= \mathrm{j} \\ e_3 &= \mathrm{k} \end{aligned} \tag{I.23}$$

$\mathrm{i}, \mathrm{j}, \mathrm{k}$ tworzą bazę w przestrzeni $\mathbb{R}^3$. W ten sposób mamy naturalny rozkład $\mathbb{R}^4$ na $\mathbb{R} \times \mathbb{R}^3$. Innymi słowy

$$E^4 = E^1 \times E^3, \tag{I.24}$$

gdzie $\|\cdot\|$ jest zdefiniowana w E^4, a inna norma

$$\|a\|_3 = \sqrt{a_1^2 + a_2^2 + a_3^2}, \quad a = (a_1, a_2, a_3), \tag{I.25}$$

jest normą w E^3.

Definiujemy działanie $*$ na elementach bazy $\mathrm{i}, \mathrm{j}, \mathrm{k}$ tak, że

$$\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = \mathrm{i} * \mathrm{j} * \mathrm{k} = -1. \tag{I.26}$$

W ten sposób każdy kwaternion zapisywany jest w sposób następujący:

$$p = a_0 + \mathrm{i}a_1 + \mathrm{j}a_2 + \mathrm{k}a_3. \tag{I.27}$$

Mnożenie kwaternionów definiujemy następująco:

$$p_1 * p_2 = (a_0 + \mathrm{i}a_1 + \mathrm{j}a_2 + \mathrm{k}a_3) * (a_0' + \mathrm{i}a_1' + \mathrm{j}a_2' + \mathrm{k}a_3'), \tag{I.28}$$

używając zwykłych reguł mnożenia wielomianów i reguł (I.26).

W ten sposób mnożymy wszystkie kwaterniony. Iloczyn tak określony jest łączny, chociaż w ogólności nieprzemienny. Ma on jedynkę $1 \in \mathbb{R}$, jest także odwracalny dla każdego $p \neq 0$. Element odwrotny p' taki, że

$$p * p' = p' * p = 1, \tag{I.29}$$

jest jednoznacznie określony, tak, że wprowadzimy dla niego oznaczenie p^{-1}, $p \neq 0$. Wprowadzimy także pojęcie kwaternionu sprzężonego (anty-automorfizm inwolutoryczny):

$$\Phi(p) = \overline{p} = a_0 - \mathrm{i}a_1 - \mathrm{j}a_2 - \mathrm{k}a_3. \tag{I.30}$$

Używając $\overline{p}$ możemy napisać

$$p^{-1} = \frac{\overline{p}}{\|p\|^2} \tag{I.31}$$

$$\|p\|^2 = p\overline{p} = \overline{p}p. \tag{I.32}$$

Jeśli $\|p\| = 1$, mamy prostszy wzór

$$p^{-1} = \overline{p}, \tag{I.33}$$

„$\overline{}$" jest anty-automorfizmem inwolutorycznym.

Uwaga 1. Definicje mnożenia kwaternionów można wprowadzić w inny sposób, zwany konstrukcją Cayleya-Dicksona. Konstrukcję tę przeprowadzmy w następujący sposób:

Każdą czwórkę liczb definiującą kwaternion możemy zapisać formalnie w postaci pary liczb zespolonych (podkreślam: *kwaterniony są algebrą nad ciałem liczb rzeczywistych, nie zespolonych*):

$$x = (x_1, x_2), \quad y = (y_1, y_2), \quad x_1, x_2, y_1, y_2 \in \mathbb{C}.$$

Definiujemy

$$x * y = (x_1 y_1 - \overline{y}_2 x_2, y_2 x_1 + x_2 \overline{y}_1),$$

$$\overline{x}^2 = (\overline{x}_1, -x_2), \quad \overline{x}_1 \text{ jest sprzężeniem zespolonym.}$$

Tak określony iloczyn jest równoważny wprowadzonemu powyżej. Konstrukcję tę możemy kontynuować dowolną ilość razy w następujący sposób indukcyjny. Niech A_n będzie algebrą zdefiniowaną dla pewnego $n \geqslant 1$. Mnożenie w $A_{n+1} = A_n \times A_n$ definiujemy wzorem

$$x *_{n+1} y = (x_1 *_n y_1 - \overline{y}_2 *_n x_2, y_2 *_n x_1 + x_2 *_n \overline{y}_1),$$

gdzie $*_n$ jest iloczynem w A_n, a $*_{n+1}$ jest iloczynem w A_{n+1},

$$x = (x_1, x_2), \quad y = (y_1, y_2), \quad x_1, x_2, y_1, y_2 \in A_n,$$

$$\overline{x}^{n+1} = \left(\overline{x}_1^n, -x_2\right).$$

$\overline{}^n$ jest sprzężeniem w A_n, $\overline{}^{n+1}$ jest sprzężeniem w A_{n+1}. (Zauważmy, że $\mathscr{C} = (\mathbb{C}, +, *, |\cdot|, \varphi)$ jest ciałem liczb zespolonych, gdzie $\mathbb{C}$ jest zbiorem liczb zespolonych, $|\cdot|$ modułem, φ anty-automorfizmem inwolutorycznym (sprzężeniem zespolonym) takim, że $\bigwedge_{z \in \mathbb{C}} z * \varphi(z) = |z|^2$, a $+, *$ zwykłymi działaniami w $\mathbb{C}$). W tej notacji sprzężenie zespolone w $\mathscr{C}$ ma postać $\overline{}^1$", a w $\mathscr{R}$ „$\overline{}^0$". W ostatnim przypadku jest oczywiście równe tożsamości $\overline{x}^0 = x$.

W ten sposób dla $n = 0$ mamy liczby rzeczywiste, dla $n = 1$ zespolone, dla $n = 2$ kwaterniony, dla $n = 3$ oktoniony, dla $n = 4$ sedeniony. Zauważmy, że dla $n = 3$ iloczyn jest niełączny, aczkolwiek alternatywny. Znaczy to, że w niełącznej algebrze oktonionów mamy następujące prawo:

$$x^2 * y = x * (x * y).$$

Dla $n \geqslant 4$ nawet tego prawa nie ma, mamy inne, słabsze

$$(x * y) * x = x * (y * x).$$

To prawo nazywane jest *giętkością* (*flexibility*). Wszystkie te algebry mają skończony wymiar 2^n, są unormowane w oczywisty sposób za pomocą normy w E^{2^n}. Zaczynając od $n = 4$ (sedeniony) algebry tego typu posiadają dzielniki zera, tak, że prawo

$$\|x *_n y\| = \|x\| \cdot \|y\|$$

jest spełnione tylko dla elementów niebędących dzielnikami zera. Elementy różne od zera i dzielników zera są odwracalne.

Po tej uwadze wracamy do głównego toku.

Każdy kwaternion możemy zapisać w następujący sposób:

$$q = q_s + \vec{q}_v. \tag{I.34}$$

q_s nazywamy częścią skalarną, a $\vec{q}_v$ częścią wektorową. Iloczyn kwaternionów q i p możemy zapisać

$$p * q = p_s q_s - \vec{p}_v \cdot \vec{q}_v + p_s \vec{q}_v + \vec{p}_v \cdot q_s + \vec{p}_v \times \vec{q}_v, \tag{I.35}$$

gdzie

$$p = p_s + \vec{p}_v \tag{I.36}$$

$$\vec{p}_v = \mathrm{i}a_1 + \mathrm{j}a_2 + \mathrm{k}a_3. \tag{I.37}$$

$\vec{p}_v \cdot \vec{q}_v$ jest iloczynem skalarnym w E^3,

$$\vec{p}_v \cdot \vec{q}_v = a_1 a_1' + a_2 a_2' + a_3 a_3' \tag{I.38}$$

$$\vec{p}_v \times \vec{q}_v = \begin{vmatrix} \mathrm{i} & \mathrm{j} & \mathrm{k} \\ a_1 & a_2 & a_3 \\ a_1' & a_2' & a_3' \end{vmatrix} \tag{I.39}$$

jest iloczynem wektorowym.

Zauważmy następujący fakt. Wprowadzając kwaterniony możemy określić dzielenie (mnożenie też) trójwymiarowych wektorów. Wprowadzimy w ten sposób pojęcie kwaternionu wektorowego, tj. takiego, którego część skalarna jest równa zeru. Iloczyn dwóch kwaternionów wektorowych jest równy

$$q \cdot p = -\vec{q}_v \cdot \vec{p}_v + \vec{q}_v \times \vec{p}_v. \tag{I.40}$$

Iloczyn ten pozostanie kwaternionem wektorowym, gdy

$$\vec{q}_v \cdot \vec{p}_v = 0. \tag{I.41}$$

Zatem mnożenie to nie wyprowadza nas poza kwaterniony wektorowe wtedy i tylko wtedy, gdy wektory odpowiadające kwaternionom wektorowym są ortogonalne.

Dzielenie określimy w następujący sposób:

$$q * p^{-1} = \frac{q * \overline{p}}{\|p\|^2} = -\frac{1}{\|p\|_3^2} \vec{q}_v \times \vec{p}_v \tag{I.42}$$

$$\overline{p} = -(\mathrm{i}a_1' + \mathrm{j}a_2' + \mathrm{k}a_3') \tag{I.43}$$

$$\|p\|^2 = \|p\|_3^2 = {a_1'}^2 + {a_2'}^2 + {a_3'}^2. \tag{I.44}$$

$\|\cdot\|_3$ jest normą w E^3.

Wprowadzenie tego działania było marzeniem R. Hamiltona. Sir Rowan uważał to za swoje największe osiągnięcie życiowe. Używając kwaternionów wektorowych możemy reprezentować wektory trójwymiarowe. Wprowadzenie kwaternionów zostało dokonane przed wynalezieniem iloczynu wektorowego w E^3. Iloczyn wektorowy dwóch wektorów możemy wprowadzić tylko w E^3 lub E^7 (w dowolnym E^n możemy zawsze wprowadzić iloczyn wektorowy $(n-1)$ wektorów). W przypadku E^7 kierunek wyniku może być niejednoznaczny.

Reasumując: algebra kwaternionów jest czterowymiarową łączną algebrą nad ciałem liczb rzeczywistych, unormowaną normą w E^4 z dzieleniem. Jest oczywiście algebrą topologiczną, gdzie bazą topologii jest zbiór kul otwartych

$$K(x_0, r) = \{x \in \mathbb{R}^4 : \|x - x_0\| < r\}. \tag{I.45}$$

Posiada ona anty-automorfizm inwolutoryczny. Jest więc algebrą Banacha z gwiazdką.

Kwaterniony mogą także reprezentować obroty w trójwymiarowej przestrzeni wektorowej E^3. W tym celu wprowadzimy kwaterniony jednostkowe, tj. takie, że

$$\|p\| = 1, \tag{I.46}$$

czyli

$$a_0^2 + a_1^2 + a_2^2 + a_3^2 = 1. \tag{I.47}$$

Macierz obrotu (macierz ortogonalna)

$$RR^T, \quad \det R \neq 0, \tag{I.48}$$

może być reprezentowana w następujący sposób:

$$R = \begin{pmatrix} a_0^2 + a_1^2 - a_2^2 - a_3^2 & 2(a_1a_2 - a_0a_3) & 2(a_1a_3 + a_0a_2) \\ 2(a_1a_2 + a_0a_3) & a_0^2 - a_1^2 + a_2^2 - a_3^2 & 2(a_2a_3 - a_0a_1) \\ 2(a_1a_3 - a_0a_2) & 2(a_2a_3 + a_0a_1) & a_0^2 - a_1^2 - a_2^2 + a_3^2 \end{pmatrix}^T, \tag{I.49}$$

^{T} jest transpozycją macierzy.

Jeżeli spełniona jest równość (I.47), to

$$\det R = 1. \tag{I.50}$$

Kwaternion odpowiadający obrotowi o kąt ϑ ma postać

$$a_0 = \cos\frac{\vartheta}{2}, \quad a_i = \gamma_i \sin\frac{\vartheta}{2}, \quad i = 1, 2, 3, \tag{I.51}$$

gdzie γ_i są kosinusami kierunkowymi osi obrotu. Kwaterniony odpowiadające obrotom wokół osi skierowanym wzdłuż układu współrzędnych (x, y, z) mają postać

$$\begin{aligned} q_{R_1} &= \cos\frac{\vartheta_1}{2} + \mathrm{i}\sin\frac{\vartheta_1}{2} \\ q_{R_2} &= \cos\frac{\vartheta_2}{2} + \mathrm{j}\sin\frac{\vartheta_2}{2} \\ q_{R_3} &= \cos\frac{\vartheta_3}{2} + \mathrm{k}\sin\frac{\vartheta_3}{2}, \end{aligned} \tag{I.52}$$

gdzie R_1, R_2, R_3 odpowiadają macierzom obrotu o kąty $\vartheta_1, \vartheta_2, \vartheta_3$ zgodnie z wzorem (I.49).

Możemy w ten sposób zapisać przekształcenie ortogonalne dowolnego wektora w E^3 używając algebry kwaternionów. Mamy

$$q' = q_R \cdot q \cdot q_R^{-1}, \tag{I.53}$$

gdzie q' odpowiada wektorowi po przekształceniu, q przed przekształceniem, q_R kwaternionowi dokonującemu przekształcenia. Kwaterniony q' i q są wektorowe, q_R jest kwaternionem jednostkowym.

Zdefiniujemy rozkład biegunowy kwaternionu podobny do rozkładu biegunowego liczby zespolonej ($z = re^{i\vartheta}$) w następujący sposób

$$q = \|q\| \cdot U(q), \tag{I.54}$$

gdzie $U(q)$ jest elementem grupy $SU(2) \simeq SO(3)$ (macierzą unitarną 2×2). $SU(2)$ jest uniwersalną nakrywającą (dwa razy) grupy $SO(3)$.

Rozważmy następujące równanie kwaternionowe

$$z^2 = -1, \quad z \in \mathbb{H}. \tag{I.55}$$

Równanie to ma nieskończenie wiele rozwiązań kwaternionowych. Jest oczywiste, że rozwiązania są kwaternionami wektorowymi o normie jednostkowej, tj.

$$z_1^2 + z_2^2 + z_3^2 = 1, \tag{I.56}$$

odpowiadającymi kierunkom w E^3 (osiom). Wektory te należą do sfery jednostkowej $S^2 \subset E^3$.

W ten sposób możemy uważać kwaterniony jako sumę nieskończonej liczby kopii płaszczyzn zespolonych dla każdego $I^2 = -1$

$$q = a + Ib, \quad a, b \in \mathbb{R}. \tag{I.57}$$

W każdej identyfikacji mamy $a + Ib$ i identyfikujemy I i $-I$. Każdy kwaternion niebędący liczbą rzeczywistą należy do jedynej kopii $\mathscr{C}$. Identyfikacja I i $-I$ identyfikuje antypodyczne punkty S^2. Dla każdego I mamy

$$I = I_s + I_{\vec{v}}. \tag{I.58}$$

Możemy napisać

$$I = I_s + \|\vec{I_{\vec{v}}}\| U(\vec{I_{\vec{v}}}) = U(\vec{I_{\vec{v}}}). \tag{I.59}$$

Zatem dla $a + b\sqrt{-1}$ mamy $a + bU(\vec{I_{\vec{v}}})$.

Dla każdego kwaternionu q mamy

$$q = q_s + \|q_{\vec{v}}\| U(q_{\vec{v}}). \tag{I.60}$$

Jednostki kwaternionowe możemy reprezentować za pomocą macierzy Pauliego

$$\mathrm{i} \to \sigma_1\sigma_2, \quad \mathrm{j} \to \sigma_3\sigma_1, \quad \mathrm{k} \to \sigma_2\sigma_3, \quad 1 \to J, \tag{I.61}$$

gdzie

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -\sqrt{-1} \\ \sqrt{-1} & 0 \end{pmatrix},$$
$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad J = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{I.62}$$

Równocześnie możemy napisać

$$U(q) = e^{\sqrt{-1}\alpha\sigma_1} e^{\sqrt{-1}\beta\sigma_2} e^{\sqrt{-1}\gamma\sigma_3}, \quad \|q\| = 1, \quad \alpha, \beta, \gamma \in \mathbb{R}. \tag{I.63}$$

Kwaterniony jednostkowe dają nam możliwość przekształcenia ortogonalnego wektorów w E^3. Interesującym problemem byłoby rozszerzenie przekształcenia ortogonalnego ($SO(3)$) do bardziej ogólnego przekształcenia.

Nie możemy zrealizować za pomocą kwaternionów ogólnego afinicznego przekształcenia wektorów w $\mathbb{R}^3$, tj. takiego, że macierz przekształcenia spełnia tylko warunek $\det R \neq 0$. Możemy jednak uogólnić (I.53) do przekształcenia ze zmianą długości wektora, chociaż w dalszym ciągu wiernokątnego

$$q' = q_R \cdot q \cdot \overline{q}_R, \quad \det R \neq 1. \tag{I.64}$$

Co oznacza to przekształcenie dla q będącego kwaternionem wektorowym? Zdefiniujmy

$$\widetilde{q}_R = \frac{q_R}{\|q_R\|}, \quad \|\widetilde{q}_R\| = 1 \tag{I.65}$$

$$\widetilde{q}_R^{-1} = \overline{\widetilde{q}}_R = \frac{\overline{q}_R}{\|q_R\|}, \tag{I.66}$$

otrzymujemy

$$q' = \|q_R\|^2 \widetilde{q}_R \cdot q \cdot \widetilde{q}_R^{-1}. \tag{I.67}$$

W ten sposób

$$\|q'\|_3 = \|q_R\|^2 \cdot \|q\|_3 \tag{I.68}$$

i mamy zmianę długości wektora zadaną przez dowolny kwaternion. $\frac{q_R}{\|q_R\|}$ zadaje przekształcenie ortogonalne w $\mathbb{R}^3$, a $\|q_R\|^2$ określa zmianę długości. Zmiana orientacji może być również określona przez $\frac{q_R}{\|q_R\|}$.

Uwaga 2. Zauważmy, że zbiór $G = \{-1, 1, \mathrm{i}, -\mathrm{i}, \mathrm{j}, -\mathrm{j}, k, -\mathrm{k}\}$ tworzy wraz z mnożeniem $*$ grupę ośmioelementową $\mathscr{G} = (G, *)$. Używając $\mathscr{G}$ możemy zdefiniować kwaterniony jako liniową otoczkę ze względu na $\mathbb{R}$,

$$\mathbb{H} = \mathrm{Lin}_{\mathbb{R}}(G).$$

Zauważmy pewne własności sprzężenia kwaternionów:

$$\overline{q_1 * q_2} = \overline{q_2} * \overline{q_1} \tag{I.69}$$

lub ogólniej

$$\overline{q_1 * q_2 * q_3 * \ldots * q_n} = \overline{q_n} * \overline{q_{n-1}} * \ldots * \overline{q_1}. \tag{I.70}$$

Mamy również

$$\left(q_1 * q_2 * q_3 * \ldots * q_n\right)^{-1} = q_n^{-1} * \cdots * q_1^{-1}. \tag{I.71}$$

Mamy także użyteczne własności łączące algebrę wektorową w E^3 z kwaternionami wektorowymi:

$$\begin{aligned} \vec{q}_{1v} \cdot \vec{q}_{2v} &= -\frac{1}{2}(q_1 * q_2 + q_2 * q_1) \\ \vec{q}_{1v} \times \vec{q}_{2v} &= \frac{1}{2}(q_1 * q_2 - q_2 * q_1), \end{aligned} \tag{I.72}$$

gdzie q_1, q_2 są kwaternionami wektorowymi, „$\cdot$” jest iloczynem skalarnym, „$\times$” iloczynem wektorowym.

Wróćmy teraz do przedstawienia kwaternionów na sferze. Niech

$$q = f_0 + \mathrm{i} f_1 + \mathrm{j} f_2 + \mathrm{k} f_3. \tag{I.73}$$

Mamy

$$q = \|q\|\left(\frac{f_0}{\|q\|} + \mathrm{i}\,\frac{f_1}{\|q\|} + \mathrm{j}\,\frac{f_2}{\|q\|} + \mathrm{k}\,\frac{f_3}{\|q\|}\right) = \|q\|\left(\frac{q_s}{\|q\|} + \frac{q_v}{\|q\|}\right) = \|q\|\left(\frac{q_s}{\|q\|} + \zeta\,\frac{\|q_v\|_3}{\|q\|}\right), \tag{I.74}$$

gdzie $\|q_v\|_3 = \sqrt{f_1^2 + f_2^2 + f_3^2}$, a ζ jest kwaternionem wektorowym o jednostkowej normie euklidesowej w E^3. Zdefiniujmy kąt $0 < \vartheta < \pi$ taki, że

$$\cos\vartheta = \frac{q_s}{\|q\|} \tag{I.75}$$

$$\sin\vartheta = \frac{\|q_v\|_3}{\|q\|}. \tag{I.76}$$

W ten sposób otrzymujemy

$$q = \|q\|(\cos\vartheta + \zeta\sin\vartheta) \tag{I.77}$$

(por. (I.54)). W przypadku kwaternionu jednostkowego

$$q = \cos\vartheta + \zeta\sin\vartheta. \tag{I.78}$$

Niech $\vec{a}, \vec{b} \in E^3$ i niech $\|\vec{a}\|_3 = a$, $\|\vec{b}\| = b$. Niech kąt między wektorami wynosi ϑ, a płaszczyzna, w której leżą wektory $\vec{a}$ i $\vec{b}$ ($\vec{a}$ i $\vec{b}$ nie są kolinearne), niech będzie prostopadła do wektora $\vec{\zeta_v}$ (ζ jest czysto wektorowy i jednostkowy). Załóżmy również, że $\vec{a}$, $\vec{b}$, $\vec{\zeta_v} = \vec{\zeta}$ tworzą prawoskrętny układ współrzędnych. Możemy więc zapisać (I.78) w postaci

$$\begin{aligned} q = \cos\vartheta + \zeta\sin\vartheta = \frac{ab\cos\vartheta + \zeta ab\sin\vartheta}{\|\vec{a}\|_3} &= \frac{\vec{a}\cdot\vec{b} + \vec{a}\times\vec{b}}{\|\vec{a}\|_3} \\ = \frac{-\frac{1}{2}(\vec{a}*\vec{b} + \vec{b}*\vec{a}) + \frac{1}{2}(\vec{a}*\vec{b} - \vec{b}*\vec{a})}{\|\vec{a}\|_3} &= -\frac{\vec{b}*\vec{a}}{\|\vec{a}\|_3} = \vec{b} * \frac{\overline{\vec{a}}}{\|\vec{a}\|_3} = \vec{b} * \vec{a}^{-1}. \end{aligned} \tag{I.79}$$

W powyższych wzorach identyfikujemy wektory $\vec{a}, \vec{b}$ z odpowiadającymi im kwaternionami wektorowymi.

Zgodnie z tymi wzorami każdemu kwaternionowi jednostkowemu możemy przyporządkować łuk koła wielkiego sfery o promieniu $r = a = b$, mający kierunek dodatniego kąta ϑ i łączący wektory $\vec{a}, \vec{b}$. W związku z tym kwaternion q ((I.79)) może być jednoznacznie przedstawiony jako łuk koła wielkiego $\operatorname{arc} q$, którego płaszczyzna jest wyznaczona przez wektor $\vec{\zeta}_v$, a długość przez kąt φ. Kierunek kwaternionu jest zdeterminowany przez kierunek wektora $\vec{\zeta}_v$, położenie łuku na kole wielkim jest dowolne.

Możemy również przedstawić mnożenie kwaternionów na sferze S^2. Mając jednostkowe kwaterniony

$$q_1, q_2, \quad \|q_i\| = 1, \ i = 1, 2,$$

możemy je przedstawić w postaci łuków. Zgodnie z wzorem (I.79) wprowadzamy wektory $\vec{b}, \vec{a}, \vec{c}$ takie, że

$$q_1 = \vec{b} * \vec{a}^{-1}, \quad q_2 = \vec{a} * \vec{c}^{-1}. \tag{I.80}$$

Zatem

$$q_3 = q_1 * q_2 = \vec{b} * \vec{c}^{-1}. \tag{I.81}$$

Oznacza to, że mnożeniu kwaternionów odpowiada operacja geometrycznego sumowania łuków na sferze S^2,

$$\operatorname{arc}(q_3) = \operatorname{arc}(q_1 * q_2) = \operatorname{arc}(q_2) + \operatorname{arc}(q_1). \tag{I.82}$$

Łuki koła wielkiego na S^2 są traktowane jako wektory (ważny jest porządek wykonywania operacji).

Można zinterpretować te operacje używając trygonometrii sferycznej, a także dodawania wektorów w nieeuklidesowej geometrii Riemanna. Zatem suma wektorowa pewnych łuków kół wielkich, określonych przez kwaterniony q_1, q_2 daje łuk koła wielkiego kwaternionu $q_3 = q_2 * q_1$ (a zatem pomnożonych w odwrotnej kolejności). Zauważmy, że możemy rozpatrywać także kwaterniony wektorowe o dowolnej długości: q_1, q_2, $\|q_1\| \neq 1$, $\|q_2\| \neq 1$.

Definiując kwaterniony wektorowe jednostkowe

$$q_i' = \frac{q_i}{\|q_i\|}, \quad i = 1, 2,$$

mamy

$$q_1 * q_2 = \|q_1\| \cdot \|q_2\| \cdot (q_1' * q_2'). \tag{I.83}$$

W ten sposób mamy dodawanie łuków $\operatorname{arc}(q_2') + \operatorname{arc}(q_1')$ na sferze jednostkowej oraz przekształcenie, które określa dylatację:

$$K = \|q_1\| \cdot \|q_3\|. \tag{I.84}$$

Przypominamy, że

$$\|q_i\| = \|q_i\|_3, \quad i = 1, 2.$$

II Analiza kwaternionów

Możemy rozpatrywać funkcje o wartościach kwaternionowych zależnych od zmiennej niezależnej $t \in \mathbb{R}^n$ $(\mathbb{C}^n)$. Funkcje te mogą być ciągłe, różniczkowalne, gładkie lub analityczne, zgodnie z odpowiednimi definicjami, tj.

$$\frac{dq}{dt} = q'(t), \quad t \in \mathbb{R}\ (\mathbb{C}),$$

definiujemy jako granicę ilorazu różnicowego

$$\frac{dq}{dt}(t_0) = \lim_{\Delta t \to 0} \frac{q(t_0 + \Delta t) - q(t_0)}{\Delta t}, \tag{II.1}$$

podobnie wyższe pochodne. Funkcję analityczną za pomocą wzoru

$$q(t) = \sum_{n=0}^{\infty} a_n t^n, \quad t \in \mathbb{R}\ (\mathbb{C}), \quad a_n \in H. \tag{II.2}$$

Zbieżność może być określona zgodnie z normą kwaternionową, np. używając warunku Cauchy'ego

$$\bigwedge_{\varepsilon > 0} \bigvee_{N \in \mathbb{N}_1^\infty} \bigwedge_{n,m > N} \left\| \sum_{p=n}^{m} a_p t^p \right\| < \varepsilon \tag{II.3}$$

lub w sposób słaby.

Możemy również zakładać, że $t \in \mathbb{R}^n$ $(\mathbb{C}^n)$, wtedy stosujemy pojęcie pochodnej Frécheta lub Gateau w zwykły sposób. Rozwinięcie w wielowymiarowy szereg Taylora możemy zdefiniować w podobny sposób. Możemy również używać wielowskaźników, wtedy

$$i = (i_1, \ldots, i_n), \quad t^i = t_1^{i_1} t_2^{i_2} \cdots t_n^{i_n}. \tag{II.4}$$

Interesującą sprawą jest rozpatrywanie funkcji od kwaternionów, tj. takich, które mają dziedzinę $Df = \mathbb{H}$, tj.

$$f(q), \quad f(q) \in \mathbb{H}. \tag{II.5}$$

Niestety nie możemy stworzyć podobnej teorii, jak w przypadku liczb zespolonych, tj. teorii funkcji analitycznych w dziedzinie kwaternionów. Nie znaczy to, że nie możemy rozpatrywać funkcji analitycznych od zmiennej kwaternionowej, np.

$$f(q) = \sum_{n=0}^{\infty} a_n * q^n, \tag{II.6}$$

gdzie $a_n \in \mathbb{H}$. Zbieżność tego szeregu jest względem normy kwaternionowej $\| \cdot \|$. Np.

$$e^q = \sum_{n=0}^{\infty} \frac{1}{n!}\, q^n. \tag{II.7}$$

Niestety, w ogólności

$$e^{q_1 + q_2} \neq e^{q_1} * e^{q_2}. \tag{II.8}$$

Tylko wtedy, gdy $q_1 * q_2 = q_2 * q_1$, mamy

$$e^{q_1+q_2} = e^{q_1} * e^{q_2}. \tag{II.9}$$

Możemy jednak otrzymać wzór

$$e^{q_1} * e^{q_2} = e^{F(q_1,q_2)}, \tag{II.10}$$

gdzie funkcję $F(q_1, q_2)$ możemy jednoznacznie wyznaczyć używając zależności Bakera-Hausdorffa-Campbella. Możemy zdefiniować

$$\sin(q) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!!}\, q^{2n+1} \tag{II.11}$$

$$\cos(q) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!!}\, q^{2n} \tag{II.12}$$

i otrzymać wzór

$$e^{\sqrt{-1}\,q} = \cos q + \sqrt{-1}\sin q. \tag{II.13}$$

Dlaczego nie można stworzyć takiej teorii? Wynika to z bardzo prostego faktu. W przypadku zmiennej zespolonej definicją funkcji holomorficznej jest, aby miała ona pochodną zespoloną w otoczeniu $z_0 \in \mathbb{C}$, $\frac{dH}{dz}(z_0)$. Z tego faktu wynika, że ma ona wszystkie pochodne w tym otoczeniu, oraz że jest rozwijalna w szereg nieskończony Taylora

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n \tag{II.14}$$

$$a_n = \frac{1}{n!}\,\frac{d^n f}{dz^n}(z_0). \tag{II.15}$$

Równocześnie spełnione są równania Cauchy-Riemanna

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}, \tag{II.16}$$

gdzie

$$z = x + \mathrm{i}y, \quad f(z) = u(x,y) + \mathrm{i}v(x,y). \tag{II.17}$$

W przypadku kwaternionów *nie istnieje pochodna* $\frac{df}{dq}$ poza trywialnymi przypadkami

$$\begin{aligned} f(q) &= a = \text{const} \\ f(q) &= a * q, \quad a \in \mathbb{H}. \end{aligned} \tag{II.18}$$

Fakt ten może być łatwo otrzymany ze sprzeczności analogonu równań Cauchy-Riemanna dla

$$\begin{gathered} f(q) = u_0(f_0,f_1,f_2,f_3) + \mathrm{i}u_1(f_0,f_1,f_2,f_3) + \mathrm{j}u_2(f_0,f_1,f_2,f_3) + \mathrm{k}u_3(f_0,f_1,f_2,f_3) \\ q = f_0 + \mathrm{i}f_1 + \mathrm{j}f_2 + \mathrm{k}f_3. \end{gathered} \tag{II.19}$$

Można to również otrzymać badając granicę ilorazu różnicowego

$$\lim_{\|q\|\to 0} \big(f(q+\Delta q) - f(q)\big)(\Delta q)^{-1}. \tag{II.20}$$

Granica ta zależy od drogi, po jakiej Δq osiąga zero. Wystarczy wziąć $f(q) = q^2$. Mimo wszystko możemy rozpatrywać

$$f : \mathbb{H} \to \mathbb{H}$$

jako przekształcenie $\mathbb{R}^4 \to \mathbb{R}^4$ i zdefiniować pochodną Gateau lub Frécheta dla niej, a więc także pochodne cząstkowe względem f_0, f_1, f_2, f_3 i rozpatrywać zagadnienie odwrócenia takiego przekształcenia w tym duchu.

Całkowanie funkcji kwaternionowej zależnej od parametru można również rozpatrzyć. Wtedy

$$F(t) = \int_A q(t)\, d\mu(t), \quad t \in A \subset \mathbb{R}^n \ (\mathbb{C}^n), \quad d\mu(t) = d^n t, \tag{II.21}$$

tak, że μ jest miarą Lebesgue'a w $\mathbb{R}^n$ $(\mathbb{C}^n)$, a A jest zbiorem mierzalnym w $\mathbb{R}^n$ $(\mathbb{C}^n)$ w sensie Lebesgue'a. Całkowalność $q(t)$ określamy w zwykły sposób, jak dla funkcji mierzalnej w sensie Lebesgue'a na $\mathbb{R}^n$ $(\mathbb{C}^n)$ o wartościach w $\mathbb{H}$ traktowanej jak $\mathbb{R}^4$.

W ten sposób możemy określić całki krzywoliniowe, powierzchniowe, wielokrotne, orientowalne, nieorientowalne itp. Możemy również rozpatrywać całki niewłaściwe i badać ich zbieżność w zwykły sposób. Np. zbieżność całki

$$J_q = \int_0^\infty q(t)\, dt$$

jest równoważna zbieżności całki liczbowej

$$J = \int_0^\infty \|q(t)\|\, dt.$$

Rozpatrujemy również dystrybucje o wartościach kwaternionowych u takie, że u jest funkcjonałem liniowym, ciągłym, o wartościach w $\mathbb{H}$, zdefiniowanym na przestrzeni funkcji próbnych $C_0^\infty(\mathbb{R}^n)$ lub $S(\mathbb{R}^n)$, gdzie $C_0^\infty(\mathbb{R}^n)$ oznacza funkcje rzeczywiste klasy C^∞ o zwartych nośnikach na $\mathbb{R}^n$, $S(\mathbb{R}^n)$ funkcje rzeczywiste, szybko malejące na $\mathbb{R}^n$. Topologie w $C_0^\infty(\mathbb{R}^n)$ i $S(\mathbb{R}^n)$ są zdefiniowane w zwykły sposób, prowadząc do przestrzeni $\mathcal{D}(\mathbb{R}^n)$, $\mathcal{S}(\mathbb{R}^n)$. Przestrzenie $\mathcal{D}'_q(\mathbb{R}^n)$ i $\mathcal{S}'_q(\mathbb{R}^n)$ — kwaternionowe przestrzenie dystrybucji i dystrybucji kwaternionowych — pozwalają zdefiniować przekształcenie Fouriera

$$v(t) = Ff(t) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-\sqrt{-1}(x,t)} f(x)\, dx, \tag{II.22}$$

gdzie $dx = d^n x$ — miara Lebesgue'a na $\mathbb{R}^n$,

$$(x,t) = \sum_{j=1}^n x_j t_j,$$

a $f(x)$ jest funkcją kwaternionową zmiennej $x \in \mathbb{R}^n$, dla której istnieje całka, a więc np. $\|f\| \in C_0^\infty(\mathbb{R}^n)$ lub $\|f\| \in S(\mathbb{R}^n)$. Dystrybucyjne przekształcenia określa się następująco:

$$Fu[\varphi] = u[F\varphi], \quad \varphi \in S(\mathbb{R}^n). \tag{II.23}$$

Wprowadzając przestrzeń Hilberta $L^2(\mathbb{H}, \mathbb{R}^n, d^n x)$, tj. takich funkcji kwaternionowych (o wartościach w algebrze kwaternionów), że

$$\int_{\mathbb{R}^n} \|f^2(x)\|\, d^n x < \infty, \tag{II.24}$$

co jest równoważne

$$\int_{\mathbb{R}^n} \|f(x)\|^2 \, d^n x < \infty, \tag{II.25}$$

możemy udowodnić, że przekształcenie F (Fouriera) realizuje izometrię w L^2, tj.

$$\|Fu\|_{L^2} = \|u\|_{L^2}. \tag{II.26}$$

Pozostałe własności przekształcenia Fouriera są również spełnione, tj.

$$D^\beta Fu = (-\sqrt{-1})^{|\beta|} F D^\beta u, \tag{II.27}$$

gdzie $\beta = (\beta_1, \ldots, \beta_n)$ jest wielowskaźnikiem, a

$$|\beta| = \beta_1 + \beta_2 + \ldots + \beta_n.$$

u jest dystrybucją temperowaną o wartościach kwaternionowych. Mamy również zwykłą trójkę Gelfanda

$$\mathcal{S}(\mathbb{R}^n) \otimes \mathbb{H} \subset L^2_q(\mathbb{H}, \mathbb{R}^n, d^n x) \subset \mathcal{S}'(\mathbb{R}^n). \tag{II.28}$$

Miarę w $\mathbb{H}$ określamy jako miarę Lebesgue'a na $\mathbb{R}^4$. W szczególnych przypadkach możemy używać także miary Peano, a także całki Riemanna (przypadek funkcji ciągłych). Uwaga ta dotyczy także poprzednich całek. Możemy również używać innych miar niż miara Lebesgue'a, będących miarami Borela i absolutnie ciągłych względem miary Lebesgue'a. Przypadki miar singularnych mogą być także rozpatrzone zgodnie z twierdzeniem o rozkładzie miary na $\mathbb{R}^4$. Zatem możemy rozpatrzyć miarę

$$\mu_{\mathbb{H}} = \mu_{\text{dyskr}} + \mu_{\text{sing}} + \mu_{\text{abs.ciągła } \mathbb{H}}\,. \tag{II.29}$$

Pozwala nam to rozpatrywać

$$\mu_{\text{abs.ciągła } \mathbb{H}} = g(q)\, d^4 x, \tag{II.30}$$

gdzie

$$g : \mathbb{H} \to \mathbb{R}. \tag{II.31}$$

Zakładamy, że $g(q) > 0$.

Pozwala nam to określić na $\mathbb{H}$ miarę unormowaną, a więc miarę probabilistyczną, i zdefiniować przestrzeń zdarzeń elementarnych $\Omega = (\mathbb{H}, \mu)$. μ jest unormowana i może być rozłożona zgodnie z (II.29). Funkcje mierzalne względem μ określone na $\mathbb{H}$ są zmiennymi losowymi na algebrze kwaternionów.

Istotną rzeczą jest również rozpatrzenie przekształcenia Laplace'a dla funkcji o wartościach kwaternionowych $q(t)$, $t \in \langle 0, +\infty)$,

$$(\mathcal{L}q)(p) = \frac{1}{\sqrt{2\pi}} \int_0^\infty q(t) e^{-pt} \, dt, \quad p \in \mathscr{C}. \tag{II.32}$$

Czasami opuszczamy $\frac{1}{\sqrt{2\pi}}$ w definicji (II.32). Przekształcenie to istnieje, gdy istnieje całka

$$\int_0^\infty \|q(t)\| e^{-pt} \, dt.$$

Można rozpatrywać uogólnienie (II.32) na przypadek tzw. dystrybucji Laplace'a.

Jednowymiarowe przekształcenie Fouriera otrzymujemy dla $p = \mathrm{i}x$, $x \in \mathbb{R}$, $\mathrm{i} = \sqrt{-1}$. Używając przekształcenia Fouriera i Laplace'a możemy określić przekształcenia zmiennych losowych na $\mathbb{H}$ otrzymując momenty centralne i zwykłe momenty dla tych zmiennych. Rozpatrzmy rozkład Gaussa w dziedzinie kwaternionów. Wtedy

$$d\mu = Ae^{-\alpha\|q-q_0\|^2}\, d^4x$$

tak, aby $\int_{\mathbb{R}^4} d\mu = 1$.

Możemy również rozpatrywać zmienne losowe o wartościach kwaternionowych

$$q(t), \quad t \in \mathbb{R}\ (\mathbb{R}^n).$$

t może być także zmienną dyskretną z zadaną miarą na $\mathbb{R}$ ($\mathbb{R}^n$) lub dyskretnym podzbiorze $\mathbb{R}$ (lub $\mathbb{R}^n$). $q(t)$ jest oczywiście funkcją mierzalną (najczęściej będzie ona ciągła, gdy zadamy miarę Lebesgue'a na $\mathbb{R}^n$).

III Zastosowania

Rozpatrzmy model

$$\vec{x}(t) = \sum_{j=1}^{p} A_j \vec{x}(t-j) + \vec{E}(t) \tag{III.1}$$

taki, że A_j są macierzami ortogonalnymi 3×3, a $\vec{x}$ i $\vec{E}$ są wektorami w E^3.

Wprowadzając kwaterniony wektorowe dla $\vec{x}$ i $\vec{E}$ oraz kwaterniony jednostkowe reprezentujące macierze A_j mamy

$$q_{\vec{x}(t)} = \sum_{j=1}^{p} q_{A_j} q_{\vec{x}(t-j)} \cdot q_{A_j}^{-1} + q_{\vec{E}(t)}, \tag{III.2}$$

gdzie mamy korespondencję

$$\begin{aligned} \vec{x}(t) &\to q_{\vec{x}(t)} \\ A_j &\to q_{A_j} \\ \vec{x}(t-j) &\to q_{\vec{x}(t-j)} \\ \vec{E}(t) &\to q_{\vec{E}(t)} \end{aligned} \tag{III.3}$$

Model (III.2) możemy uogólnić w następujący sposób:

$$q_{\vec{x}(t)} = \sum_{j=1}^{p} q_{A_j} * q_{\vec{x}(t-j)} * \overline{q}_{A_j} + q_{\vec{E}(t)}. \tag{III.4}$$

Teraz kwaterniony q_{A_j} odpowiadają macierzom ortogonalnym z jednokładnością $k = \|q_{A_j}\|^{-2}$. W ten sposób możemy zmienić długość wektora o zadaną skalę dla każdego j. Macierz A_j ma postać

$$A_j = k\overline{A}_j, \tag{III.5}$$

gdzie $\overline{A}_j$ jest macierzą ortogonalną. Przekształcenie w ten sposób określone jest oczywiście konforemne (równokątne). Za pomocą kwaternionów nie możemy zdefiniować dowolnego przekształcenia afinicznego w $\mathbb{R}^3$. W dziedzinie częstotliwości mamy

$$\vec{x}(\omega) = A(\omega)\vec{x}(\omega) + \vec{E}(\omega), \quad \omega = 2\pi f. \tag{III.6}$$

Używając kwaternionów mamy

$$q_{\vec{x}(\omega)} = q_{A(\omega)} * q_{\vec{x}(\omega)} * \overline{q}_{A(\omega)} + q_{\vec{E}(\omega)}. \tag{III.7}$$

Rozwiązując równanie kwaternionowe (III.7) ze względu na $q_{\vec{x}(\omega)}$ mamy

$$q_{\vec{x}(\omega)} = T(\omega) * q_{\vec{E}(\omega)}, \tag{III.8}$$

gdzie $T(\omega)$ jest kwaternionową funkcją przejścia spełniającą równanie

$$T(\omega) * q_{\vec{E}(\omega)} = q_{A(\omega)} * T(\omega) * q_{\vec{E}(\omega)} * \overline{q}_{A(\omega)} + q_{\vec{E}(\omega)} \tag{III.9}$$

lub

$$T(\omega) = q_{A(\omega)} * T(\omega) * q_{\vec{E}(\omega)} * \overline{q}_{A(\omega)} * q^{-1}_{\vec{E}(\omega)} + \mathbb{I}, \tag{III.10}$$

gdzie $\mathbb{I}$ jest kwaternionem jednostkowym. Równanie to ma jednoznaczne rozwiązanie.

Załóżmy dodatkowo, że $\vec{x}(t)$, $\vec{E}(t)$ są zmiennymi losowymi dla $t \in \mathbb{R}$ ($t_j \in \mathbb{R}$) i istnieją wartość oczekiwana i wariancja. Mamy model

$$A_0\vec{x}(t) = \sum_{j=1}^{p} A_j\vec{x}(t-j) + \vec{E}(t). \tag{III.11}$$

Mamy więc model wektorowy jednokanałowy (kanałem jest tu zmienna wektorowa w $\mathbb{R}^3$).

Rozpatrujemy model wielokanałowy (k-kanałowy)

$$\vec{x}_i(t), \quad i = 1, 2, \ldots, k.$$

Mamy

$$A_0^i\vec{x}_i(t) = \sum_{j=1}^{p} A_j^i\vec{x}_i(t-j) + \vec{E}_j(t), \tag{III.12}$$

gdzie dla każdego $i = 1, 2, \ldots, k$ A_j^i jest macierzą realizującą przekształcenie ortogonalne z jednokładnością, $j = 0, 1, 2, \ldots, p$,

$$A_j^i = k_j^i\overline{A}_j^i, \quad \overline{A}_j^i\overline{A}_j^{i\,T} = I, \quad I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad k_j^i > 0. \tag{III.13}$$

$\vec{x}_i(t)$ jest wektorem trójwymiarowym dla każdego $t \in \mathbb{R}$ (t może być dyskretne), $i = 1, 2, \ldots, k$.

Używając notacji kwaternionowej możemy zapisać

$$q_{A_0^i} q_{\vec{x}_i(t)} \overline{q}_{A_0^i} = \sum_{j=1}^{p} q_{A_j^i} * q_{\vec{x}_i(t-j)} * \overline{q}_{A_j^i} + q_{\vec{E}_i(t)}, \tag{III.14}$$

gdzie $q_{\vec{x}_i(t-j)}$ jest kwaternionem wektorowym odpowiadającym wektorowi $\vec{x}_i(t-j)$, $q_{\vec{E}_i(t)}$ wektorowi $\vec{E}_i(t)$, $q_{A^i_j}$ $(i=1,\ldots,k,\ j=1,\ldots,p)$ kwaternionem (w ogólności niejednostkowym) odpowiadającym macierzy (III.13) tak, że

$$k^i_j = \|q_{A^i_j}\|^2. \tag{III.15}$$

W ten sposób mamy kwaternionowy zapis modelu opisującego model k-wektorowy wektorów 3-wymiarowych podlegający obrotom i jednokładnościom w $\mathbb{R}^3$. Zakładając, że dla każdego A^i_0 jest odwracalne, możemy ten model sprowadzić do następujących modeli:

$$\vec{x}_i(t) = \sum_{j=1}^{p} A^i_j \vec{x}_i(t-j) + \vec{E}_i(t) \tag{III.16}$$

lub

$$q_{\vec{x}_i(t)} = \sum_{j=1}^{p} q_{A^i_j} q_{\vec{x}_i(t-j)} \overline{q}_{A^i_j} + q_{\vec{E}_i(t)}. \tag{III.17}$$

W dziedzinie częstotliwości mamy

$$\vec{x}_i(\omega) = \sum_{l=1}^{k} A^l_i(\omega)\vec{x}_l(\omega) + \vec{E}_i(\omega), \tag{III.18}$$

w zapisie kwaternionowym

$$q_{\vec{x}_i(\omega)} = \sum_{l=1}^{k} q_{A^l_i(\omega)} * q_{\vec{x}_l(\omega)} * \overline{q}_{A^l_i(\omega)} + q_{\vec{E}_i(\omega)}. \tag{III.19}$$

Definiujemy kwaternionową funkcję przejścia $T_i(\omega) \in \mathbb{H}$, $i = 1, 2, \ldots, k$, w następujący sposób:

$$q_{\vec{x}_i(\omega)} = T_i(\omega) * q_{\vec{E}_i(\omega)}. \tag{III.20}$$

Zatem mamy

$$T_i(\omega) * q_{\vec{E}_i(\omega)} = \sum_{l=1}^{k} q_{A^l_i(\omega)} * T_l(\omega) * q_{\vec{E}_l(\omega)} * \overline{q}_{A^l_i(\omega)} + q_{\vec{E}_i(\omega)}, \quad i = 1, 2, \ldots, k. \tag{III.21}$$

Równanie to ma jednoznaczne rozwiązanie.

Istnieje także inna możliwość zdefiniowania funkcji przejścia

$$q_{\vec{x}_i(\omega)} = T_i(\omega) * q_{\vec{E}_i(\omega)} * \overline{T}_i(\omega). \tag{III.22}$$

Wtedy mamy

$$T_i(\omega) * q_{\vec{E}_i(\omega)} * \overline{T}_i(\omega) = \sum_{l=1}^{k} q_{A^l_i(\omega)} * T_l(\omega) * q_{\vec{E}_l(\omega)} * \overline{T}_i(\omega) * \overline{q}_{A^l_i(\omega)} + q_{\vec{E}_i(\omega)}. \tag{III.23}$$

Równanie to ma również jednoznaczne rozwiązanie. Funkcje przejścia $T_i(\omega) \in \mathbb{H}$ zawierają całą informację na temat układu transformując szum w sygnał.

Można próbować definiować te funkcje także w inny sposób, tj.

$$q_{\vec{x}_i(\omega)} = \sum_{l=1}^{k} T_i^l(\omega) * q_{\vec{E}_l(\omega)} \tag{III.24}$$

lub

$$q_{\vec{x}_i(\omega)} = \sum_{l=1}^{k} T_i^l(\omega) * q_{\vec{E}_l(\omega)} * \overline{T}_i^l(\omega). \tag{III.25}$$

Dwie ostatnie definicje są bardziej skomplikowane, mają jednak pewne dodatkowe zalety, separujące kanały i ich szumy.

Jeśli dokonamy dopasowania $\vec{x}_i(t)$, wyznaczając reszty dopasowania $\vec{E}_i(t)$, możemy wyznaczyć

$$\vec{V}_{ij} = \Big\langle \vec{E}_i(t) - \langle \vec{E}_i(t) \rangle \Big\rangle \Big\langle \vec{E}_j(t) - \langle \vec{E}_j(t) \rangle \Big\rangle, \tag{III.26}$$

gdzie $\vec{V}_{ij}$ jest wektorową macierzą kowariancji (oznacza to, że każdy element $\vec{V}_{ij}$ jest wektorem trójwymiarowym w $\mathbb{R}^3$). $\langle \vec{E}_i(t) \rangle$ jest wartością oczekiwaną (wektorową) dla każdego $i = 1, 2, \ldots, k$. Możemy utworzyć następującą wielkość kwaternionową

$$V_{ij} = \Big\langle q_{\vec{E}_i(t)} - \langle q_{\vec{E}_i(t)} \rangle \Big\rangle * \Big\langle q_{\vec{E}_j(t)} - \langle q_{\vec{E}_j(t)} \rangle \Big\rangle, \tag{III.27}$$

gdzie $\langle q_{\vec{E}_i(t)} \rangle$ jest wartością oczekiwaną zmiennej losowej o wartościach kwaternionowych. W ten sposób mamy kwaternionową macierz kowariancji.

Określmy

$$\sigma_{ij} = \|V_{ij}\|^2, \quad i, j = 1, 2, \ldots, k. \tag{III.28}$$

To samo możemy uczynić rozpatrując model $(k-1)$-kanałowy (usuwając jeden z kanałów wektorowych, np. j_0). Mamy wtedy

$$\big(\sigma_{ij}^{(j_0)}\big)^2 = \|V_{ij}^{(j_0)}\|^2, \quad i, j = 1, 2, \ldots, k, \ i, j \neq i_0, j_0. \tag{III.29}$$

Możemy określić wielkość

$$I_{ij_0} = 1 - \frac{\sigma_{ii}^2}{\sigma_{ii}^{(j_0)}} = 1 - \frac{\|V_{ii}\|^2}{\|V_{ii}^{(j_0)}\|^2} \tag{III.30}$$

$$0 \leqslant I_{ij_0} \leqslant 1. \tag{III.31}$$

Jeśli I_{ij_0} jest bliskie zeru, to kanał j_0 nie ma wpływu na kanał $i \neq j_0$, jeśli $I_{ij_0} \simeq 1$, to sygnał j_0 ma duży wpływ na i.

Zatem kryterium przyczynowości Garniera w przypadku k-kanałowego modelu wektorowego w zapisie kwaternionowym ma postać

$$\|V_{ii}\|^2 < \|V_{ii}^{(j_0)}\|^2. \tag{III.32}$$

Kanał wektorowy j_0 ma wpływ na kanał i.

Kryterium tak sformułowane może być stosowane także w przypadku modeli, które nie są liniowe. W naszej notacji posłużyliśmy się $\langle q_{\vec{E}(t)} \rangle$, co oznacza uśrednienie ze względu na próbę statystyczną. Możemy jednak używać normalnej notacji rachunku prawdopodobieństwa

$$V_{ij} = E\Big(q_{\vec{E}_i(t)} - E(q_{\vec{E}_i(t)})\Big) * E\Big(q_{\vec{E}_j(t)} - E(q_{\vec{E}_j(t)})\Big). \tag{III.33}$$

W konkretnych zastosowaniach warto mieć odwrócenie wzoru (I.49). Niech R będzie macierzą ortogonalną (I.47), (I.50), wtedy

$$\begin{aligned} a_0 &= \pm\frac{1}{2}\sqrt{\mathrm{Tr}(R)+1} \\ a_1 &= \frac{R_{[23]}}{\sqrt{\mathrm{Tr}(R)+1}} \\ a_2 &= \frac{R_{[31]}}{\sqrt{\mathrm{Tr}(R)+1}} \\ a_3 &= \frac{R_{[12]}}{\sqrt{\mathrm{Tr}(R)+1}}\,, \end{aligned} \tag{III.34}$$

gdzie $\mathrm{Tr}(R) = R_{11} + R_{22} + R_{33}$ jest śladem macierzy R. Kwaternion $a = a_0 + \mathrm{i}a_1 + \mathrm{j}a_2 + \mathrm{k}a_3$ jest kwaternionem jednostkowym, $\|a\| = 1$. W przypadku, gdy $\widetilde{R}$ jest przekształceniem z jednokładnością, mamy $\widetilde{R} = kR$, $k > 0$, $\det\widetilde{R} = k^3$, oraz odpowiadający mu kwaternion ma postać $\widetilde{a} = \|\widetilde{a}\| \cdot a$, $\|\widetilde{a}\| = \sqrt{k} = (\det\widetilde{R})^{1/6}$.

W ten sposób mamy

$$\begin{aligned} \widetilde{a}_0 &= \pm\frac{1}{2}\sqrt{\mathrm{Tr}(\widetilde{R})+k} \\ \widetilde{a}_1 &= \frac{\widetilde{R}_{[23]}}{\sqrt{\mathrm{Tr}(\widetilde{R})+k}} \\ \widetilde{a}_2 &= \frac{\widetilde{R}_{[31]}}{\sqrt{\mathrm{Tr}(\widetilde{R})+k}} \\ \widetilde{a}_3 &= \frac{\widetilde{R}_{[12]}}{\sqrt{\mathrm{Tr}(\widetilde{R})+k}}\,, \end{aligned} \tag{III.35}$$

$\mathrm{Tr}(\widetilde{R}) = \widetilde{R}_{11} + \widetilde{R}_{22} + \widetilde{R}_{33}$ jest śladem macierzy $\widetilde{R}$, a k współczynnikiem jednokładności.

Dodatek

Można rozpatrzyć kwaterniony z punktu widzenia twierdzenia Mazura-Gelfanda.

Twierdzenie Mazura-Gelfanda mówi: *Algebra Banacha (nad ciałem $\mathbb{R}$) przemienna z inwolucją, w której każdy element oprócz zera jest odwracalny, jest izomorficzna z ciałem liczb rzeczywistych. W przypadku, gdy jest ona określona nad ciałem liczb zespolonych $\mathbb{C}$, jest ona izomorficzna z ciałem liczb zespolonych.*

Używając twierdzenia Frobeniusa (dowód twierdzenia Mazura-Gelfanda jest oparty na tym twierdzeniu) możemy uogólnić twierdzenie Mazura-Gelfanda, dopuszczając nieprzemienność działania $*$ w algebrze Banacha — w następujący sposób. Jedyną nieprzemienną algebrą Banacha nad $\mathbb{R}$, w której istnieje element odwrotny dla każdego $x \neq 0$, jest algebra kwaternionów (oczywiście z dokładnością do izomorfizmu algebr).

Można postawić następujące hipotezy:

1° Jedynymi algebrami Banacha $\mathcal{A}$ (z dokładnością do izomorfizmu), które są określone nad $\mathbb{R}$ oraz są z dzieleniem, tj. dla każdego $x \neq 0$ istnieje element odwrotny, i z inwolucją, spełniającymi warunek

$$\bigwedge_{x,y\in A} x^2 * y = x * (x * y),$$

są R, $\not{C}$, $\mathbb{H}$, $\mathbb{O}$, tj. ciało liczb rzeczywistych, zespolonych, algebra kwaternionów i algebra oktonionów.

Algebry te są alternujące.

2° Jedynymi algebrami Banacha nad ciałem liczb rzeczywistych, które są giętkie, tj.

$$\bigwedge_{x,y\in A} (x * y) * x = x * (y * x),$$

i posiadają element odwrotny dla $x \neq 0$, są z dokładnością do izomorfizmu algebry R, $\not{C}$, $\mathbb{H}$ i $\mathbb{O}$.

3° Jeśli dopuścimy istnienie dzielników zera, tj. takich elementów, że

$$x * y = 0, \quad x, y \neq 0,$$

(wtedy $\|x * y\| = 0$ i nie mamy $\|x * y\| = \|x\| \cdot \|y\| \bigwedge_{x,y\in A}$) i rozpatrzymy zdegenerowaną algebrę Banacha, tj. taką, że $\|x\| \geqslant 0$, ale taką, że są w niej dzielniki zera, to jedynymi (z dokładnością do izomorfizmu) algebrami, które są giętkie, tj. spełniają

$$\bigwedge_{x,y\in A} (x * y) * x = x * (y * x),$$

i dla wszystkich elementów odwracalnych mamy

$$\|x * y\| = \|x\| \cdot \|y\|,$$

są algebry otrzymane w wyniku konstrukcji Cayleya–Dicksona (patrz Uwaga 1). Oczywiście zakładamy, że algebra jest nad ciałem liczb rzeczywistych z inwolucją.

W literaturze znane jest twierdzenie Gelfanda, które mówi: *Niech $\mathcal{A}$ będzie ciałem* (*także nieprzemiennym*) *takim, że $\mathcal{A}$ jest algebrą Banacha nad ciałem liczb zespolonych. Wtedy $\mathcal{A}$ jest izomorficzne z ciałem liczb zespolonych.*

Podamy także twierdzenie Frobeniusa:

Każde ciało (*nawet niekomutatywne*) *skończeniewymiarowe nad ciałem liczb rzeczywistych* R *jest izomorficzne z* R*,* $\not{C}$ *lub* $\mathbb{H}$*.*

Wymiar jest rozumiany w sensie algebraicznym jako wymiar odpowiedniej przestrzeni wektorowej.

Literatura

[1] Białynicki–Birula A., *Algebra*, PWN, Warszawa, 1976.

[2] Cochrane J. H., *Time Series for Macroeconomics and Finance*, Springer, 1997.

[3] Feller W., *Wstęp do rachunku prawdopodobieństwa*, t. I, II, PWN, Warszawa, 2007, 2009.

[4] Fisz M., *Rachunek prawdopodobieństwa*, PWN, Warszawa, 1967.

[5] Gosiewski Z., Ortyl A., *Algorytmy inercjalnego bezkardanowego systemu orientacji i położenia obiektu o ruchu przestrzennym*, Biblioteka Naukowa Instytutu Lotnictwa 12 (Awionika), Warszawa, 1999.

[6] Granger, C. W. J., *Investigating causal relations by economic models and cross-spectral methods*, Econometrica 37 (1969), p. 424.

[7] Hamilton W. R., *Elements of Quaternions*, vols. I, II, ed. by Ch. J. Joly, Chelsea Publ. Co., New York 1969 (reprint of the second edition, Longmans, Green and Co., London 1899–1901).

[8] Kamiński M., Liang H. L., *Causal Influence: Advances in Neurosignal Analysis*, Critical Reviews in Biomedical Engineering 33 (2005), p. 347.

[9] Kravchenko V., *Applied Quaternion Analysis*, Heldermann, 2003.

[10] Kuipers J. B., *Quaternions and Rotation Sequences. A Primer with Applications to Orbits, Aerospace and Virtual Reality*, Princeton Univ. Press, Princeton, NJ, 1999.

[11] Marcinkowska H., *Dystrybucje, przestrzenie Sobolewa, równania różniczkowe*, PWN, Warszawa, 1993.

[12] Maurin K., *Analiza*, t. I, II, III, PWN, Warszawa, 1991.

[13] Opial Z., *Algebra wyższa*, PWN, Warszawa 1967.

[14] Rudin W., *Podstawy analizy matematycznej*, PWN, Warszawa, 1998.

[15] Szmydt Z., *Transformata Fouriera i równania różniczkowe liniowe*, PWN, Warszawa, 1972.

[16] Żelazko W., *Banach Algebras*, PWN, Warszawa, 1973.