1, 2, 4, 8

Michał ADAMASZEK, Warszawa

Jest to zapis odczytu nagrodzonego medalem Filca na XL Szkole Matematyki Poglądowej, *Matematyczne Obrazki*, styczeń 2008.

Wbrew tytułowi nie jest to artykuł o systemie binarnym. Zamiast tego zainteresujemy się znanym zapewne wielu Czytelnikom sloganem „mnożenie można zdefiniować tylko w przestrzeniach wymiaru $1, 2, 4, 8$".

Algebra to struktura liczbowa udostępniająca dwa działania, dodawanie i mnożenie, spełniające dobrze znaną regułę rozdzielności. Podstawowe przykłady algebr to liczby rzeczywiste ($\mathbb{R}$) i zespolone ($\mathbb{C}$), przy czym ta pierwsza jest jedno-, a druga dwuwymiarowa (jako przestrzeń liniowa nad $\mathbb{R}$). Działanie dodawania w takiej przestrzeni „przychodzi" jako część struktury liniowej i nie będziemy się nim specjalnie interesować. Spróbujemy za to dopasować do niego odpowiednie mnożenie.

Mnożenie w $\mathbb{R}$ i $\mathbb{C}$ ma wiele własności, które chcielibyśmy uznać za charakterystyczne dla „porządnego" działania: ma element neutralny 1, taki że $1 \cdot x = x \cdot 1 = x$; nie ma dzielników zera, to znaczy zachodzi implikacja:

$$\text{jeśli } xy = 0, \text{ to } x = 0 \text{ lub } y = 0$$

oraz zachowuje długość wektorów (normę):

$$|xy| = |x| \cdot |y| \tag{1}$$

Pionowe kreski oznaczają tu wartość bezwzględną liczby rzeczywistej, normę liczby zespolonej lub, bardziej ogólnie, normę wektora zadaną poprzez standardowy iloczyn skalarny w przestrzeni $\mathbb{R}^n$ wzorem:

$$|x|^2 = \langle x, x \rangle$$

Algebry spełniające warunek (1) nazwiemy **algebrami unormowanymi**. Warunek (1) implikuje, że taka algebra nie ma dzielników zera (proszę sprawdzić).

Twierdzenie o algebrach unormowanych, które udowodnimy, brzmi:

Twierdzenie (Hurwitz, 1898) Mnożenie, zadające w $\mathbb{R}^n$ strukturę algebry unormowanej z jedynką można wprowadzić tylko w przestrzeniach $\mathbb{R}^1$, $\mathbb{R}^2$, $\mathbb{R}^4$ i $\mathbb{R}^8$. Co więcej, jeśli mnożenie to ma być łączne ($x(yz) = (xy)z$), to możliwe są jedynie $\mathbb{R}^1$, $\mathbb{R}^2$ i $\mathbb{R}^4$, a jeśli dodatkowo przemienne ($xy = yx$), to tylko $\mathbb{R}^1$ i $\mathbb{R}^2$.

No dobrze, ale jak te wszystkie mnożenia wyglądają?

Ciut historii

Z problemem zdefiniowania mnożenia bez dzielników zera w przestrzeni $\mathbb{R}^n$ zmagał się w połowie XIX w. Hamilton. Przykładów dla $n = 1$ i $n = 2$ dostarczają, co już wiemy, liczby rzeczywiste i zespolone. Podejmowane przez Hamiltona próby skonstruowania mnożenia bez dzielników zera w $\mathbb{R}^3$ były bezowocne, udało się mu za to (w 1843 roku) dla $n = 4$ i tak powstały słynne **kwaterniony**, oznaczane od nazwiska twórcy literą $\mathbb{H}$. Kwaterniony stanowią czterowymiarową algebrę z mnożeniem opisanym w bazie $1, i, j, k$ wzorami

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

i rozszerzonym liniowo na wszystkie wektory. Jeszcze tego samego roku kolega Hamiltona, John T. Graves, podał przykład algebry bez dzielników zera w wymiarze $n = 8$. Historia obeszła się z nim dosyć szorstko, bo jego dzieło nazywa się dziś **oktawami Cayley'a** ($\mathbb{O}$), od nazwiska Artura Cayley'a, który był szybszy z publikacją (choć o oktawach traktował tylko dodatek do jego pracy na nieco inny temat).

Wszystkie wymienione algebry $(\mathbb{R}, \mathbb{C}, \mathbb{H}, \mathbb{O})$ można uzyskać poprzez tzw. **konstrukcję Cayley'a-Dicksona**, stanowiącą uogólnienie znanej konstrukcji liczb zespolonych jako par liczb rzeczywistych. Mając algebrę V, wyposażoną dodatkowo w idempotentną operację sprzężenia $a \mapsto \bar{a}$, definiujemy działania na parach $(a, b) \in V \times V$ wzorami:

$$\overline{(a, b)} = (\overline{a}, -b)$$
$$(a, b) \cdot (c, d) = (ac - b\overline{d}, ad + b\overline{c})$$

otrzymując algebrę $D(V)$ dwukrotnie większego wymiaru. Startując od algebry $\mathbb{R}$ z trywialnym sprzężeniem ($\overline{a} = a$) otrzymamy kolejno $\mathbb{C} = D(\mathbb{R})$, $\mathbb{H} = D(\mathbb{C})$,

$\mathbb{O} = D(\mathbb{H})$. Można to ciągnąć dalej, ale otrzymane algebry przestaną nam się podobać, bo w każdym kolejnym kroku gubimy coraz więcej przyjemnych własności. Zobaczmy, dlaczego. Na przykład, aby udowodnić przemienność tak uzyskanych liczb zespolonych, napisalibyśmy:

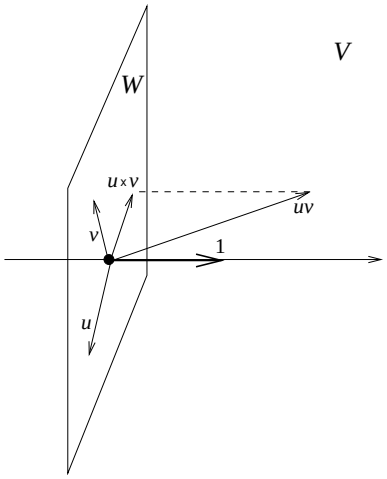
$$(a,b)\cdot(c,d) = (ac - b\overline{d}, ad + b\overline{c}) = (ac - bd, ad + bc) = (ca - d\overline{b}, cb + d\overline{a}) = (c,d)\cdot(a,b)$$

Udało się, ale po drodze skorzystaliśmy z przemienności liczb rzeczywistych i z faktu, że dla liczb rzeczywistych a mamy $\bar{a} = a$. Nie wykażemy zatem analogicznie przemienności kwaternionów, bo sprzężenie w liczbach zespolonych nie jest tożsamościowo identycznością. Podobnie, chcąc wykazać łączność $D(V)$, musimy skorzystać w obliczeniach z przemienności V (proszę przeliczyć). Stąd tabelka:

symbol	wymiar	unormowana	łączna	przemienna
$\mathbb{R}$	1	✔	✔	✔
$\mathbb{C}$	2	✔	✔	✔
$\mathbb{H}$	4	✔	✔	✘
$\mathbb{O}$	8	✔	✘	✘

Kolejne algebry, począwszy od $\mathbb{S} = D(\mathbb{O})$ (tzw. sedeniony) okazują się mieć dzielniki zera, a więc w ogóle nie spełniają naszych kryteriów.

Dodajmy, że o klasyfikacji przyzwoitych algebr są jeszcze co najmniej dwa inne, mocniejsze twierdzenia. Ich przewaga nad twierdzeniem Hurwitza polega na tym, że zachodzą one bez założenia o zachowywaniu normy (1), które dla nas będzie bardzo istotne. Są to twierdzenie Frobeniusa (1878; $\mathbb{R}$, $\mathbb{C}$ i $\mathbb{H}$ to jedyne algebry łączne z jedynką nad $\mathbb{R}$, w których każdy element ma odwrotność lewostronną równą prawostronnej) oraz twierdzenie Kerivare–Botta–Milnora (1958, strukturę algebry bez dzielników zera można wprowadzić w $\mathbb{R}^n$ tylko dla $n = 1, 2, 4, 8$).

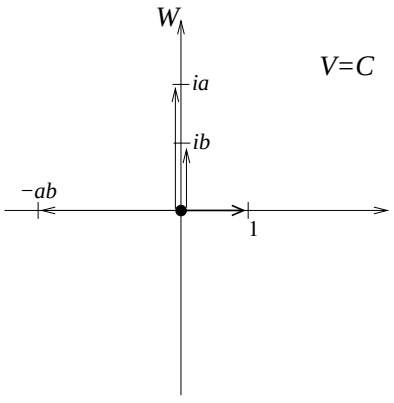


Rys. 1

Dowodu krok pierwszy

Przypuśćmy, że w $V = \mathbb{R}^{d+1}$ mamy mnożenie zadające strukturę algebry unormowanej z jedynką. Na początek skonstruujemy na jego podstawie inne działanie w mniejszej przestrzeni $W = 1^{\perp}$ (przestrzeń prostopadła do wektora $1 \in V$; zatem $W = \mathbb{R}^d$). Będziemy go oznaczać $\times$ i nazywać **iloczynem wektorowym**, zatem dla $u, v \in W$ mamy $u \times v \in W$. Oto definicja (patrz rys.1): aby obliczyć $u \times v$ wykonujemy mnożenie uv a wynik rzutujemy prostopadle na W:

$$u \times v = \pi_W(uv) = uv - \langle uv, 1\rangle \cdot 1$$



Rys. 2

Przykład 1 (trywialny). Jeśli $V = \mathbb{C} \simeq \mathbb{R}^2$, to W jest osią urojoną, a $\pi_W(z) = Im(z)$ (rys.2.). Wektory $u, v \in W$ są postaci $u = ia$, $v = ib$ dla $a, b \in \mathbb{R}$. Wówczas $u \times v = \pi_W(uv) = \pi_W(-ab) = Im(-ab) = 0$. Zatem nasza konstrukcja daje trywialny (zerowy) iloczyn wektorowy w W.

Przykład 2 (ważny). Jeśli $V = \mathbb{H} \simeq \mathbb{R}^4$, to W jest trójwymiarową przestrzenią rozpiętą przez wersory i, j, k. Niech $u = ai + bj + ck$, $v = Ai + Bj + Ck$. Korzystając z tabelki mnożenia kwaternionów liczymy:

$$\begin{aligned} u \times v &= \pi_W((ai + bj + ck)(Ai + Bj + Ck)) = \\ &= \pi_W((-aA - bB - cC) + k(aB - Ab) + i(bC - Bc) + j(cA - Ca)) \\ &= i(bC - Bc) + j(cA - Ca) + k(aB - Ab) \end{aligned}$$

We współrzędnych wyniku dostrzegamy „standardowy" iloczyn wektorowy w $\mathbb{R}^3$, co uzasadnia użycie symbolu i nazwy tego działania także w ogólnym przypadku (wkrótce poznamy jeszcze więcej przesłanek za tym wyborem).

Powinniśmy teraz przetłumaczyć różne własności zachowującego normę mnożenia w V na własności iloczynu wektorowego w W. Ten krok jest czysto rachunkowy (dla zainteresowanych odkładamy go na koniec artykułu). Napiszmy od razu, co wychodzi. Po pierwsze, okazuje się, że spełnione są relacje:

$$u \times v = -v \times u \tag{2}$$

$$\langle u \times v, w\rangle = \langle u, v \times w\rangle \tag{3}$$

Warunek zachowywania normy pociąga następującą, długą tożsamość:

(4) $\langle u \times v, w \times z\rangle + \langle v \times w, z \times u\rangle = 2\langle u, w\rangle\langle v, z\rangle - \langle u, v\rangle\langle w, z\rangle - \langle u, z\rangle\langle v, w\rangle$

Łączność mnożenia jest równoważna warunkowi:

(5) $$\langle v \times w, z \times u\rangle = \langle u, w\rangle\langle v, z\rangle - \langle u, v\rangle\langle w, z\rangle$$

przemienność zaś warunkowi:

(6) $$u \times v \equiv 0$$

(co nie jest takim zaskoczeniem w świetle przykładu 2 powyżej). Pozostaje nam pokazać, że jeśli w $W = \mathbb{R}^d$ istnieje iloczyn wektorowy spełniający warunki (2), (3) i (4) zgodności z iloczynem skalarnym, to $d = 0, 1, 3, 7$ (oraz że zachodzą odpowiednie warianty dla łączności i przemienności). Ale najpierw krótka dygresja poświęcona notacji.

Obrazki

Notacja, którą będziemy się posługiwać, służy głównie do pracy z wyrażeniami zawierającymi iloczyn skalarny i „standardowy" iloczyn wektorowy w $\mathbb{R}^3$, ale okaże się, że równie dobrze nadaje się do każdego działania spełniającego warunki (2) i (3).

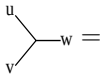
Iloczyn skalarny będziemy oznaczać kreską:

$= \langle u, v\rangle$

O tej kresce można myśleć jak o operatorze, który ma dwa wejścia, a w wyniku daje liczbę. Operator iloczynu wektorowego będzie wyglądał następująco:

$= u \times v$

Ten operator ma dwa wejścia, którymi wchodzą argumenty u i v oraz wyjście, na którym pojawia się wektor $u \times v$. Można jednak spojrzeć na trójnóg jak na operator o trzech argumentach, obliczający iloczyn mieszany:



$= \langle u \times v, w\rangle$

Nabierze to sensu, jeśli zauważymy, że wynik obliczenia $u \times v$ jest połączony kreską (a więc iloczynem skalarnym) z w. Uważny Czytelnik powinien tu zapytać, czy ta interpretacja nie zależy od kolejności wejść. Na szczęście nie, o ile czytamy wejścia zawsze w tym samym kierunku, przeciwnym do ruchu wskazówek zegara. Gwarantuje nam to tożsamość (3):

$= \langle u \times v, w\rangle = \langle v \times w, u\rangle = \langle w \times u, v\rangle$

Z kolei zmiana kolejności w jakimś skrzyżowaniu powoduje, na mocy antysymetrii (2), zmianę znaku całego wyrażenia:

$= \langle v \times u, w\rangle = -\langle u \times v, w\rangle$

Dotychczas pokazaliśmy, że definicja trójnogu ma w ogóle sens dla działań spełniających (2) i (3). Możemy więc używać skomplikowanych operatorów typu:

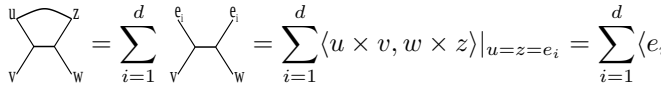
$= \langle u \times v, w \times z\rangle$

Kolejna konwencja dotyczy interpretacji obrazka składającego się z dwóch rozłącznych fragmentów. Piszemy wówczas iloczyn (zwykły, w liczbach rzeczywistych), odpowiednich wyrażeń, np:

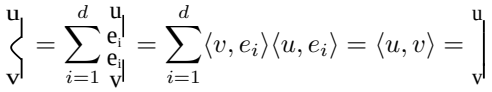
$= \langle u, v\rangle\langle w, z\rangle$

I ostatnia sprawa: konwencja sumacyjna. Ustalmy raz na zawsze bazę ortonormalną $(e_1, \ldots, e_d)$ przestrzeni W. Wtedy połączenie dwóch wejść operatora łukiem oznacza „wpuszczenie" do obu tych wejść kolejno wszystkich

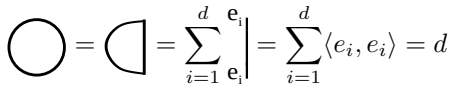
wektorów e_i i zsumowanie uzyskanych wartości (specjaliści powiedzieliby: *zwężenie* tego operatora). Jeden obraz jest tu wart tysiąca słów:



To już koniec definicji. Przyzwyczajenie się do tej notacji i nabranie wewnętrznego przekonania o jej spójności zajmuje trochę czasu. Najważniejsze jest spostrzeżenie, że rysunki można do woli wyginać, bo:



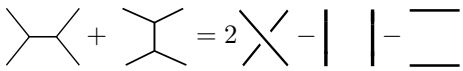
W ramach ćwiczenia zauważmy ponadto, że



jest tym, co chcemy policzyć.

Przepiszmy teraz w naszej nowej notacji warunki (4)–(6). Od teraz zaczynamy pomijać literowe oznaczenia wejść — nie powinno to budzić wątpliwości, wystarczy, że we wszystkich operatorach występujących w jednym wzorze da się dopasować wejścia odpowiadające tej samej zmiennej.

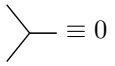
- zachowywanie normy (4):

(4') 

- łączność (5):

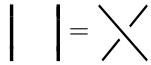
(5') 

- przemienność (6):

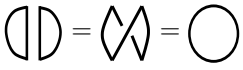
(6') 

Dowód twierdzenia

Zaczniemy od klasyfikacji algebr *przemiennych*. Zakładamy więc, że zachodzi  $\equiv 0$. Wtedy (5') redukuje się do

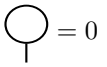


Stosujemy dwukrotnie sumowanie:

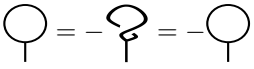


i mamy $d^2 = d$, czyli $d = 0$ lub $d = 1$, tak jak chcieliśmy.

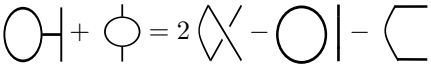
Teraz wykonamy nieco pomocniczych rachunków. Oto twierdzenie o trywialności lizaka:

(7) 

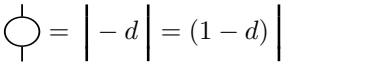
Faktycznie, zmiana kolejności w skrzyżowaniu daje bowiem:



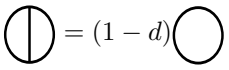
Dalej zastosujmy w (4') sumowanie dwóch wejść z lewej strony:



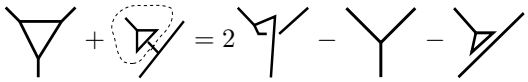
Pierwszy składnik jest zerem, bo zawiera zerowy lizak. Pierwszy i trzeci składnik prawej strony można „wyprostować" i w efekcie dostajemy:

(8) 

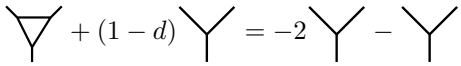
Jeśli w tym wzorze zastosujemy sumowanie pozostałych wejść, to otrzymamy:

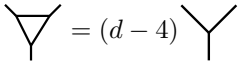
(9) 

Teraz natomiast podłączmy lewe wejścia we wzorze (4') operatorem .
Dostaniemy (po małym rozciągnięciu rysunku):

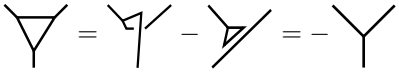


Pierwszy składnik prawej strony wymaga rozplątania (ze zmianą znaku), ostatni jest zerem (zawiera lizak), a w zaznaczonym miejscu można wykonać podstawienie ze wzoru (8). Otrzymamy:



(10) 

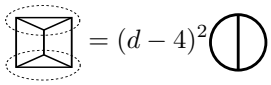
Jesteśmy gotowi do rozpatrzenia przypadku algebr *łącznych*. Zauważmy bowiem, że z warunku łączności (5') wynika, przez to samo podstawienie co poprzednio:

(11) 

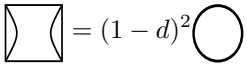
A zatem, jeśli  $\not\equiv 0$ (algebra nie jest przemienna), to z (10) i (11) mamy $d-4=-1$, czyli $d=3$ — znów tak jak chcieliśmy.

Już bez zakładania łączności obliczymy kilka pomocniczych wartości. Na pierwszy ogień pójdzie:

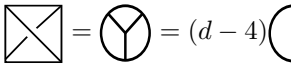
Zauważmy, że u góry i u dołu mamy „trójkąt z trzema nóżkami", taki jak w (10). Zatem, na mocy (10) i (9):

(12) 

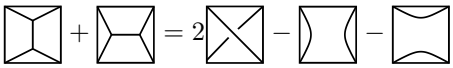
W kolejnym obliczeniu skorzystamy ze wzoru (8):

(13) 

I wreszcie w poniższym rachunku najpierw zmienimy orientację w dwóch skrzyżowaniach, a następnie wyodrębnimy trójkąt z nóżkami (10):

(14) 

To już wszystko, czego nam potrzeba do dokończenia rachunków. Stosujemy teraz sumowanie we wszystkich dostępnych kierunkach we wzorze (4'):



(proszę to sformalizować) i po krótkich przekształceniach otrzymujemy:

$$d(1-d)(d-4)^2 = d(1-d)(d-4) - d(1-d)^2$$

$$d(d-1)(d-3)(d-7) = 0$$

A zatem twierdzenie jest udowodnione.

Dodatek – rachunki

Aby z warunku zachowywania normy (1) wyprowadzić postulowane własności (2)–(6) iloczynu wektorowego $\times$ postępujemy następująco. Najpierw w tożsamości $\langle xy, xy\rangle = \langle x,x\rangle\langle y,y\rangle$ podstawiamy $x=u+e$, $y=v$, gdzie

$u, v \in W$ (tutaj e oznacza jedynkę mnożenia). Pamiętając, że dla $u \in W$ mamy $\langle u, e\rangle = 0$ oraz, że $\langle e, e\rangle = 1$ otrzymujemy:

$$\langle uv + v, uv + v\rangle = \langle u + e, u + e\rangle\langle v, v\rangle$$
$$\langle uv, uv\rangle + 2\langle uv, v\rangle + \langle v, v\rangle = \langle u, u\rangle\langle v, v\rangle + \langle v, v\rangle$$

czyli (ponieważ warunek zachowywania normy zachodzi także dla u, v):

$$\langle uv, v\rangle = 0$$

Podobnie dowodzimy:

$$\langle uv, u\rangle = 0$$

a następnie, podstawiając $x = u + e$, $y = v + e$:

$$\langle uv, e\rangle + \langle u, v\rangle = 0$$

Ponieważ $u \times u = u^2 - \langle u^2, e\rangle e$ więc:

$$|u \times u|^2 = |u^2|^2 + \langle u^2, e\rangle^2 - 2\langle u^2, e\rangle^2 = |u|^4 - \langle u, u\rangle^2 = 0$$

na mocy wzoru $|u^2| = |u|^2$. Zatem $u \times u = 0$ a stąd i z oczywistej dwuliniowości łatwo wynika antysymetria $v \times w = -w \times v$ (stosujemy sztuczkę zwaną *polaryzacją* zmiennej u, czyli podstawiamy $u = v + w$ i korzystamy z dwuliniowości, aby otrzymać wzór zawierający więcej zmiennych niezależnych). Co więcej, mamy wzór

$$u \times v = uv + \langle u, v\rangle e$$

z którego przez zamianę u i v oraz antysymetrię wynika zaskakująco prosta formuła:

$$u \times v = \frac{1}{2}(uv - vu)$$

Wynika z niej od razu, że przemienność mnożenia implikuje $u \times v \equiv 0$. Korzystając z tej formuły i polaryzując równości $\langle uv, u\rangle = \langle uv, v\rangle = 0$ po kilku przekształceniach dostajemy też własność równoległościanu (3).

Teraz w tożsamości $\langle uv, uv\rangle = \langle u, u\rangle\langle v, v\rangle$ podstawiamy $uv = u \times v - \langle u, v\rangle e$ i dostajemy:

$$\langle u \times v, u \times v\rangle + \langle u, v\rangle^2 = \langle u, u\rangle\langle v, v\rangle$$

Polaryzując w tej równości obie zmienne u i v otrzymujemy dokładnie warunek (4). Z kolei aby wyprowadzić wzór (5), będący odpowiednikiem łączności, piszemy:

$$(uv)w = (u \times v - \langle u, v\rangle e) \cdot w = (u \times v)w - \langle u, v\rangle w =$$
$$= (u \times v) \times w - \langle u \times v, w\rangle e - \langle u, v\rangle w$$

Podobnie rozpisujemy $u(vw)$, po czym porównujemy wyrażenia $\langle u(vw), z\rangle$ oraz $\langle (uv)w, z\rangle$ dla $u, v, w, z \in W$. Po przekształceniach z użyciem całego repertuaru środków (reguły równoległoboku, antysymetrii, warunku $\langle e, z\rangle = 0$ oraz wzoru (4)) dostajemy wreszcie upragnioną równość (5).

Podziękowania

Dziękuję bardzo Janowi Dymarze z Wrocławia, od którego po raz pierwszy usłyszałem o algebrach unormowanych i o obrazkowej notacji iloczynu wektorowego i bez którego tego artykułu, ani poprzedzającego go wykładu w Grzegorzewicach zapewne by nie było. Przedstawiony tutaj dowód jest oparty na pracach Markusa Rosta (patrz bibliografia).

Bibliografia

- o oktawach Cayley'a i ich historii: John C.Baez, The Octonions, *Bull. Amer. Math. Soc.* 39 (2002), 145–205.
- oryginalna praca Hurwitza: Adolf Hurwitz, Über die Composition der quadratischen Formen von beliebig vielen Variabeln, *Nachr. Ges. Wiss. Göttingen* (1898), 309–316.
- o diagramach reprezentujących iloczyn wektorowy i skalarny: Elisha Peterson, A Not-so-Characteristic Equation: the Art of Linear Algebra, *http://arxiv.org/abs/0712.2058* oraz zawarte tam odnośniki.
- ten sam dowód, ale bez obrazków...: Markus Rost, On the dimension of a composition algebra, *Documenta Math.* 1:209–214, 1996.
- ...oraz brakujące obrazki: Markus Rost, On Vector Product Algebras, dostępne w Internecie.