

MATEMATYCZNE METODY FIZYKI
ALGEBRA LINIOWA; ELEMENTY RACHUNKU TENSOROWEGO

Andrzej Lenda

Wykład dla pierwszego roku studentów fizyki technicznej
(wszystkie specjalności) – semestr zimowy 2000/2001 i 2001/2002

wersja obecna — 25 stycznia 2002

Wydział Fizyki i Techniki Jądrowej, AGH Kraków

Spis treści

Zamiast przedmowy	vi
Bibliografia – czyli gdzie autor szukał natchnienia	1
1 Początki algebry	2
2 Pojęcie odwzorowania liniowego	10
3 Układy liniowych równań; metoda Gaussa; wyznaczniki	12
3.1 Układ równań liniowych	12
3.2 Metoda Gaussa	13
3.3 Wyznacznik drugiego i trzeciego stopnia	16
3.3.1 Twierdzenie Cramera	17
3.4 Algebra wyznaczników	18
3.5 Układ równań jednorodnych	21
4 Algebra wektorów – dwa i trzy wymiary	23
4.1 Podstawowe definicje; dodawanie wektorów	23
4.2 Iloczyn skalarny	27
4.3 Iloczyn wektorowy	29
4.4 Iloczyn trzech wektorów	32
4.5 Obrót wektora na płaszczyźnie	33
5 Liczby zespolone	37
5.1 Liczby zespolone – trochę historii	37
5.2 Algebra liczb zespolonych	38
5.2.1 Reprezentacja biegunowa liczby zespolonej; liczba zespolona sprzężona; dzielenie liczb zespolonych	40
5.2.2 Wzór de Moivre’a; liczby zespolone i wzory trygonometryczne	43
5.3 Potęga i pierwiastek liczby zespolonej	43
5.4 Obrót wektora na płaszczyźnie	45
6 Przestrzenie wektorowe	47
6.1 Wprowadzenie	47
6.1.1 Formy liniowe	48
6.2 Wektory liniowo zależne	49
6.3 Podprzestrzeń wektorowa; baza	52
6.4 Rząd macierzy	53
6.5 Rząd macierzy inaczej – diagonalizacja macierzy	57
6.6 Układy równań liniowych – podsumowanie	58
6.6.1 Metoda Gaussa-Jordana	60
6.6.2 Układ równań jednorodnych	62
6.6.3 Układ równań niejednorodnych, a układ równań jednorodnych	63
6.7 Aksjomatyczne definicje przestrzeni wektorowej	64

7	Algebra macierzy	66
7.1	Macierze; podstawowe definicje i operacje	66
7.1.1	Dodawanie macierzy	67
7.1.2	Mnożenie macierzy przez liczbę	67
7.1.3	Mnożenie macierzy	68
7.1.4	Macierz transponowana	69
7.1.5	Macierz odwrotna	70
7.1.6	Macierz odwrotna – metoda Gaussa-Jordana	73
7.1.7	Macierz i jej wyznacznik; interpretacja geometryczna wyznacznika	74
7.2	Rząd iloczynu macierzy – twierdzenie Sylvestra	75
8	Formy kwadratowe	77
8.1	Wprowadzenie	77
8.2	Definicja i własności formy kwadratowej	78
8.3	Prawo bezwładności dla form kwadratowych	82
8.4	Formy kwadratowe określone dodatnio	83
8.5	Formy kwadratowe w fizyce	85
9	Odwzorowania	87
9.1	Wprowadzenie	87
9.2	Injekcja, surjekcja, bijekcja. Izomorfizm	87
9.2.1	Złożenie odwzorowań	88
9.3	Odwzorowania i macierze	89
9.4	Jądro i obraz transformacji	91
9.4.1	Baza $\text{Im} T$ i $\ker T$	92
9.5	„Twierdzenie o wymiarach”	94
9.6	$\mathbb{R}^n \rightarrow \mathbb{R}^n$; zmiana bazy	95
9.7	Transformacja liniowa i jej macierz	97
10	Przestrzenie Euklidesowe; iloczyn skalarny; ortogonalność	99
10.1	Najważniejsza forma kwadratowa przestrzeni wektorowej	99
10.2	Ortogonalizacja układu wektorów liniowo niezależnych	100
10.2.1	Reprezentacja macierzy w określonej bazie	102
10.3	Nierówność Schwarz; nierówność trójkąta	103
10.4	Ortogonalność macierzy	104
10.5	Transformacja podobieństwa	106
10.6	Macierzy symetryczne i antysymetryczne; macierze hermitowskie i unitarne	107
10.6.1	Zespolone przestrzenie liniowe	108
11	Problem własny	109
11.1	Pierwiastki charakterystyczne i wartości własne macierzy	109
11.1.1	Niezdegenerowane wartości własne	111
11.2	Problem własny a diagonalizacja macierzy	111
11.3	Odwzorowania symetryczne	112
11.4	Sprowadzanie formy kwadratowej do układu osi głównych	115
12	Trochę fizyki	119
12.1	Potęga odwzorowania liniowego	119
12.2	Diagonalizacja	122
12.2.1	Wektory własne macierzy	122
12.3	Moment bezwładności ciała sztywnego	124
12.4	Przestrzenie funkcyjne i ich ortogonalne bazy	127
12.4.1	Układy współrzędnych krzywoliniowych	127
12.4.2	Przestrzenie funkcyjne – pierwsze kroki	129

13 Wstęp do rachunku tensorowego	135
13.1 Wprowadzenie	135
13.2 Kowariantność i kontrawariantność	136
13.2.1 Forma liniowa	138
13.2.2 Gradient jako wektor kowariantny	139
13.3 Tensory – podstawowe definicje	140
13.4 Podstawowe operacje: dodawanie, mnożenie, kontrakcja	141
13.5 Tensory antysymetryczne i symetryczne; iloczyn wektorowy	142
13.6 Tensor deformacji i tensor naprężeń	143
13.6.1 Tensor deformacji	144
13.6.2 Tensor naprężeń	145
13.7 Różniczkowanie tensorów	149
13.7.1 Gradient	150
13.7.2 Różniczkowanie pola wektorowego	150
13.8 Niezwykłe przygody kropelki wody	151
13.8.1 Operator rotacji a wirowość	154

Spis rysunków

1.1	Iloczyn ab dwóch liczb i kwadrat a^2 .	3
1.2	Kwadrat dwumianu $a + b$ w języku geometrii.	4
1.3	Geometryczna metoda rozwiązywania równania pierwszego stopnia.	5
1.4	Geometryczny zapis równania $x^2 + b^2 = ax$.	5
1.5	Konstrukcja geometryczna dla znalezienia x z Rys.1.4	5
1.6	Uzasadnienie wzoru 1-2.	5
1.7	Pomocnicze rysunki do równania $x^2 + ax = b^2$.	6
1.8	Równanie $x^2 + 10x = 39$.	7
1.9	Równanie $x^2 + 10x = 39$ inaczej.	8
3.1	Wyznacznik trzeciego stopnia – dodatnie i ujemne przyczynki.	18
4.1	Dodawanie wektorów: $\mathbf{A} + \mathbf{B} = \mathbf{C}$ – reguła trójkąta.	23
4.2	Dodawanie wektorów: $\mathbf{A} + \mathbf{B} = \mathbf{C}$ – reguła równoległoboku.	24
4.3	Dodawanie wektorów – zasada przemienności i łączności.	24
4.4	Rozkład wektora \mathbf{V} na współrzędne: $\alpha = \angle(0x, \mathbf{V}), \beta = \angle(0y, \mathbf{V}), \gamma = \angle(0z, \mathbf{V})$.	25
4.5	Iloczyn wektorowy.	29
4.6	Interpretacja geometryczna iloczynu wektorowego – pole równoległoboku.	31
4.7	Interpretacja geometryczna iloczynu mieszanego – objętość równoległościanu.	33
4.8	Współrzędne wektora w dwóch różnych układach współrzędnych.	34
5.1	Dodawanie liczb zespolonych na płaszczyźnie Arganda.	39
5.2	Mnożenie liczby zespolonej przez i .	39
5.3	Mnożenie dwóch liczb zespolonych.	40
5.4	Współrzędne biegunowe na płaszczyźnie zespolonej.	41
5.5	Liczba zespolona sprzężona.	42
5.6	Pierwiastek zespolony: $\sqrt[6]{1}$.	45
6.1	Wektor \mathbf{C} jako różne liniowe kombinacje niewspółliniowych wektorów \mathbf{A} i \mathbf{B} .	49
7.1	Wynik złożenia dwóch obrotów o kąt $\pi/2$ wokół osi $0y$ i $0z$ zależy od ich kolejności.	69
7.2	Przekształcenie wersorów osi realizowane przez macierz \mathcal{M} .	75
7.3	Pole równoległoboku powstałego z kwadratu o boku 1 równe jest $ad - bc$.	76
9.1	Surjekcja, ale nie injekcja: $T(A) = T(B)$ oraz $T(C) = T(D)$.	88
9.2	Injekcja, ale nie surjekcja; nie ma rozwiązań dla $T(x) = A$ i $T(x) = B$.	88
9.3	Rzutowanie wektora na oś $0x$.	90
10.1	Inwersja układu współrzędnych w przestrzeni 3-wymiarowej.	105
12.1	Ruch drgający pod działaniem siły harmoniczej.	119
12.2	Ruch drgający pod działaniem siły harmoniczej – dwie masy.	121
12.3	Nowy układ współrzędnych – osie $0x$ i $0y$ zostały obrócone (w kierunku ujemnym) o 45° .	124
12.4	Punkt $P(x_1^0, x_2^0, x_3^0)$ jako punkt przecięcia trzech prostopadłych płaszczyzn.	127
12.5	Układy współrzędnych: sferycznych (a) i cylindrycznych (b).	128
12.6	Największa długość fali stojącej powstającej w strunie o długości L .	129
12.7	Druga największa długość fali stojącej powstającej w strunie o długości L .	130

12.8 Piąta największa długość fali stojącej powstającej w strunie o długości L .	130
12.9 Pobudzenie struny do drgań poprzez odciągnięcie jej do kształtu trójkątnego.	131
12.10 Pobudzenie struny do drgań poprzez odciągnięcie jej do kształtu parabolicznego.	133
13.1 Inwersja układu współrzędnych w przestrzeni 3-wymiarowej – wektor biegunowy.	142
13.2 Inwersja układu współrzędnych w przestrzeni 3-wymiarowej – wektor osiowy.	143
13.3 Skierowany element powierzchni dS_i ($i = 3$).	146
13.4 Tensor naprężeń σ_{ik} .	147
13.5 Deformacja pręta.	148
13.6 Deformacja ścinania.	149
13.7 Ilustracja obliczeń całkowitego strumienia pola wektorowego dla pewnej powierzchni Gaussa.	152
13.8 Wędrowka kropli.	152
13.9 Wir prędkości cieczy.	155
13.10 Jeszcze jeden wir prędkości cieczy.	155

Zamiast przedmowy

Idea tego opracowania pojawiła się dwa lata temu, kiedy to Wydział Fizyki i Techniki Jądrowej wystąpił z inicjatywą utworzenia niestacjonarnych studiów, dedykowanych w pierwszym rzędzie osobom niepełnosprawnym, a opartych na przygotowanych przez nauczycieli akademickich Wydziału materiałach dydaktycznych, zamieszczonych w Internecie. Opracowany program studiów (komputerowa fizyka techniczna) pozostał programem wirtualnym – narazie nie znalazła się odpowiednio liczna grupa zainteresowanych osób.

Jednocześnie powstała idea pewnej „reformy” sposobu nauczania matematyki studentów pierwszych dwóch lat fizyki technicznej. Opracowany dziesięć lat temu przez PT Wykładowców z Wydziału Matematyki Stosowanej (głównie pp. Kalinowski, Marczyk i Woźniak) znakomity program nauczania miał jedną niedoskonałość – w trakcie pierwszego semestr nasi studenci wysłuchiwali sześciu godzin wykładu z matematyki i odbywali sześć godzin ćwiczeń rachunkowych z tego przedmiotu. Często spędzali więc około 50% swoich „godzin kontaktowych” w tygodniu z jednym i tym samym Wykładowcą. Władze Wydziału Fizyki i Techniki Jądrowej AGH, w konsultacji z Samorządem Studenckim, a także przedstawicielami Wydziału Matematyki Stosowanej, postanowiły powrócić do tradycji wydzielenia z ogółu zagadnień matematycznych elementów algebry wyższej i rachunku tensorowego oraz powierzyć ich nauczanie – zgodnie z tradycją uniwersytecką – fizykom. Oczywiście, tego typu decyzja niesie za sobą zarówno skutki pozytywne jak i negatywne. Fizyk nigdy nie potrafi przedstawić zagadnień matematycznych w sposób tak doskonały jak matematyk; apologię takiego posunięcia stanowi jednak szansa uzmysłowienia młodym studentom fizyki adekwatności języka matematyki (algebry) do konstrukcji pojęć czysto fizycznych. Ułatwia to (mamy nadzieję) naukę obu przedmiotów.

Zamieszczone na tej stronie materiały to treść 30-godzinnego wykładu, realizowanego w ciągu pierwszego semestru. Ponieważ były one początkowo pisane z dedykacją dla studentów niestacjonarnych, dla których miały stanowić podstawę (prawie) samodzielnej nauki są one dość (zbyt?) obszerne i z pewnością nie cechują się zwięzłością. Materiały te poddawane były „próbom ogniowej” w trakcie pierwszego semestru roku akademickiego 2000/20001 (pierwsza realizacja wykładu) i uległy – drobnym w gruncie rzeczy poprawkom i modyfikacjom – w trakcie drugiej realizacji (semestr jesienny 20001/2). Oprócz wykładu studenci mają dwie godziny ćwiczeń, realizowanych pod opieką również fizyków. Wyobrażamy sobie, że po tym semestrze sam wykład ulegnie pewnym modyfikacjom, a ponadto zostanie wzbogacony i obszerny zbiór problemów ilustrujących poszczególne partie materiału.

Autor opracowania wyraża wdzięczność wszystkim osobom, których zainteresowanie przedsięwzięciem i konkretne uwagi pomogły mu w jego realizacji. W pierwszym rzędzie byli to pp. dr hab. Mariusz Woźniak (WMS) i dr hab. Janusz Wolny (WFiTJ), a także koledzy z Wydziału Fizyki, współodpowiedzialni za realizację wykładu i ćwiczeń: pp. dr hab. Antoni Paja, dr Lucjan Pytlik, dr Wojciech Karaś i mgr inż. Bartłomiej Spisak. Dziękuję także wszystkim „użytkownikom” materiałów za zwrócenie mi uwagi na kilka błędów edytorskich, jak i – przede wszystkim – za dyskusje podczas wykładu, które pomogły mi zrewidować pewne sformułowania. Te ostatnie (dyskusje) oczywiście jeszcze trwają i mam nadzieję, że będą w dalszym ciągu owocne. Bardzo liczę na dalszą współpracę z osobami, które – intencjonalnie czy też przypadkowo – zetkną się z tą stroną WWW i , z góry dziękując, proszę o uwagi: lenda@novell.ftj.agh.edu.pl.

Na zakończenie wypadałoby dodać kilka słów o literaturze przedmiotu. Tutaj sprawa jest nieco skomplikowana. W polskiej tradycji nie utrzymała się jeszcze tendencja (obserwowana zwłaszcza w Stanach Zjednoczonych) pisania podręczników z podstaw matematyki przeznaczonych dla studentów kierunków ścisłych (i technicznych). Owszem, istnieją takie książki – zwłaszcza dla studentów fizyki „uniwersyteckiej”, ale ich zasięg jest stosunkowo niewielki, a forma nie do końca spójna z programem „naszego wykładu” (pozycje spisu). Mnie osobiście zachwycają książki pisane *specjalnie* dla fizyków i inżynierów, a te są niestety trudno dostępne. Poniżej podaję dwie grupy podręczników: pierwsza, to książki z których korzystałem przy pisaniu tego opracowania; druga – książki, w których student z pewnością znajdzie wyśmienitą pomoc w utrwaleniu i rozszerzeniu wiadomości prezentowanych na tej stronie WWW.

Skrypt został napisany przy pomocy edytora \LaTeX ($\text{\LaTeX}2\epsilon$), zawierającego konwerter do formatu pdf (konkretnie – \pdf\LaTeX). System \LaTeX 'a jest idealny do pisania matematycznych tekstów. Osobiście uważam, że system \TeX i \LaTeX to rewelacja!

Bibliografia

– czyli gdzie autor szukał natchnienia

- [1] A.I.Kostrikin. *Wstęp do algebry*. PWN Warszawa 1982.
To co w tytule. Czyta się łatwo i przyjemnie.
- [2] A.I.Kostrikin, J.I.Manin. *Algebra liniowa i geometria*. PWN Warszawa 1993.
Solidny, porządnie napisany podręcznik. Raczej dla matematyków niż fizyków. Jest tam (prawie) wszystko o czym się uczyliśmy i jeszcze cała masa innych mądrości.
- [3] A.I.Kostrikin, i inni . *Zbiór zadań z algebry*. PWN Warszawa 1995.
Zbiór zadań do poprzednich dwóch pozycji. Warty przeglądnięcia!!
- [4] Z.Furdzik, J. Maj-Kluskowa, A.Kulczycka, M.Sękowska *Algebra, wyd2*. AGH, UWN-D 1998.
Bardzo porządnie napisany skrypt. Pełen definicji, twierdzeń. Są też przykłady. Język mocno formalny. Dla rzetelnie zainteresowanych rzetelną algebrą.
- [5] A.Strauszkiewicz. *Algebra i geometria*. NKF UJ Kraków 1993.
Świetna, zwięzła książeczka. Choć zwięzła to bogata w treść. Szkoda, że dla naszych potrzeb wystarcza tylko pierwszy rozdział.
- [6] L.Górniewicz, R.S.Ingarden. *Algebra z geometrią dla fizyków*. UMK Toruń 1993.
Dość trudny podręcznik. Mimo, że pisany dla fizyków przez fizyków. Raczej nie polecam.
- [7] S.Przybyło, A. Szlachetowski. *Algebra i wielowymiarowa geometria analityczna w zadaniach; wyd.6..* WNT Warszawa 1994.
Warte grzechu. Sensowne przykłady, poprzedzane krótkimi „ściągamii” z teorii.
- [8] F.W.Byron, R.W.Fuller. *Matematyka w fizyce klasycznej i kwantowej*. PWN Warszawa 1973.
Algebra jest w tomie pierwszym (nie cała). Ale i tak poezja. Książka niezastąpiona w całej edukacji fizyka. Tylko – jak ją znaleźć?
- [9] A.Kurosh. *Higher algebra*. MIR Publishers, Moscow 1980.
Strawnie napisany podręcznik algebry. Mnie się zdecydowanie podoba. Niestety, po angielsku.
- [10] J.H.Hubbard, B.Burke-Hubbard. *Vector Calculus, Linear Algebra and Differential Forms*. Prentice Hall,1999.
No cóż. Tak uczą matematyki Amerykanie i pozostaje tylko żałować, że my tak (jeszcze ?) nie potrafimy. Swobodny i przystępny język, a zarazem „porządny”, usystematyzowany wykład. Wzór słabo dościsły.
- [11] W. Sawyer. *Algebra liniowa dla inżynierów*. WNT Warszawa 1974.
Sympatyczna książeczka. Prekursor stylu jak wyżej, chociaż nieco to rozwlekłe i przegadane. Ale to właśnie stamtąd „ukradłem” opowiadanie o sprężynkach. To już samo powinno zachęcić!

Rozdział 1

Początki algebry

W słowniku wyrazów obcych termin algebra definiowany jest jako: gałąź matematyki, zajmująca się ogólnymi prawami (twierdzeniami) dotyczącymi relacji istniejących pomiędzy elementami pewnych zbiorów (liczb, wektorów, itp.). Język algebry jest językiem symbolicznym – zamiast mówić o konkretnym elemencie zbioru mówimy o jego ogólnym „reprezentancie”, ukrywając go pod symbolem litery.

Algebra rozwijała się jako nauka czysto „użytkowa”. Dlatego – mimo całego szacunku dla wspaniałej struktury dzieł Euklidesa – możemy nazwać grecką algebrą geometryczną ten właśnie spójny i logiczny fundament matematyki, dotyczący linii, odcinków, powierzchni, itp. Z tej bowiem teorii wynikają pewne reguły i konkretne, praktyczne sposoby – na przykład – obliczania pól figur, wysokości lub odległości przedmiotów, albo konstrukcji pewnych elementów. Algebrą będą też dobrze nam znane wzory określające pierwiastki równania kwadratowego poprzez współczynniki poszczególnych potęg zmiennej x występujących w tym równaniu. Algebrą wektorów nazwiemy wszystkie prawa określające operacje dodawania (składania) i mnożenia wielkości wektorowych.

Sam wyraz algebra pojawił się w języku potocznym matematyki już w wieku 13., a jego źródłosłów jest arabski. W 9. wieku, w kalifacie bagdadzkim działał matematyk Mohammed ib Musa al-Chowaliśmy, którego dzieło o oryginalnym tytule *Hisab al-dżaul w’almuqabalah*, czyli „Sposób łączenia i redukcji”, stało się, w kilka wieków później, pierwszym „nowożytnym” podręcznikiem algebry w Europie. Pod pojęciami występującymi w tytule kryły się znane nam, proste operacje, jakich dokonujemy przy rozwiązywaniu równań, na przykład:

$$\begin{array}{ll} 6x^2 - 4x + 1 = 5x^2 + 3 & \text{równanie wyjściowe} \\ 6x^2 + 1 = 5x^2 + 4x + 3 & \text{łączenie} \\ x^2 = 4x + 2 & \text{redukcja.} \end{array}$$

Z powyższego wynika, że „łączenie” to po prostu przenoszenie ujemnych wyrazów na drugą stronę równania, a „redukcja” to – no właśnie, redukowanie identycznych wyrazów, występujących po obu stronach równania. Przetłumaczone w 13. wieku dzieło nosiło łaciński tytuł *Liber Algebrae et Almucabola*. O ile drugi termin nie zagościł w żargonie matematycznym, to pierwszy – tak.¹

Ale algebra pojawiła się oczywiście dużo, dużo wcześniej. Przecież jeżeli wyłączyć astronomię, to matematyka jest z pewnością najstarszą i „o najdłuższym stażu” nauką uprawianą w dziejach ludzkości. Mówi się czasem, że matematyka powstała w starożytnej Grecji. Ale to nie jest prawda. Świadkiem sam Arystoteles, który w swojej *Metaphysica* napisał: „Sztuki matematyczne powstały w Egipcie, gdzie kapłani mieli wiele wolnego czasu.” Coś w tym jest. Stara cywilizacja Egiptu była uzależniona od Nilu – jego wylewów. Ponieważ każdy wylew mógł zatopić (lub odsłonić) kawałek lądu, a faraonowie bardzo wcześnie wynaleźli podatki, problemem numer 1 stało się dość dokładne wyznaczenie pól powierzchni, aby sprawiedliwie podatki obliczać. Przy okazji powstała i algebra – w postaci pewnych technik rachunkowych, potrzebnych do rozwiązywania konkretnych problemów.

W połowie 19. wieku, młody Szkot, p. Henry Rhind, podróżując do Egiptu w celu poratowania zdrowia zakupił w Luksorze papirus – tzw. papirus Rhind’a. Datowany na ca. 1650 p.Ch. jest dość gruntownym i poprawnie skonstruowanym traktatem matematycznym. (Teraz – w British Museum). Podaje reguły (techniki) mnożenia, operacji na ułamkach, ale zawiera także sporo problemów, których rozwiązanie wymaga zastosowania ...no właśnie, algebry. Na przykład: *Jeżeli do pewnej wielkości dodać jej siódma część dostajemy 19. Jaka to wielkość?* N.B. metoda rozwiązania to słynna *regula falsi*. Mam równanie $x + x/7 = 19$, zakładam

¹Nieco zniekształcone nazwisko autora Al-Chowarizmi stało się źródłosłowem terminu algorytm

(błędnie) $x = 7$ (bo łatwo rachować: $7 + 7/7 = 8$). Wynik $- 8$ zamiast 19 – jest więc $19/8$ -razy za mały; przez ten czynnik trzeba pomnożyć pierwotne (fałszywe) założenie: $x = 7 \times \frac{19}{8}$.

Są tam i inne problemy: „W każdym z siedmiu domów jest siedem kotów; każdy kot zabił siedem myszy; każda mysz potrafi zjeść siedem snopków zboża; w każdym snopku jest siedem miar ziarna. Ile miar ziarna zaoszczędziły dzielne koty?” Ciekawostką jest fakt, że w zaproponowanej metodzie rozwiązania można dopatrzeć się wzoru na ... sumę postępu geometrycznego!

Nie tylko cywilizacja starożytnego Egiptu radziła sobie świetnie z całkiem niebanalnymi rachunkami. Na glinianych tabliczkach, który znalazły się w ziemi babilońskiej cztery tysiące lat temu można znaleźć przepisy na rozwiązywanie już równań kwadratowych! Jest tam na przykład taki problem: *Do powierzchni kwadratu dodałem dwie trzecie jego boku. Dostałem 35/60. Jaki jest bok kwadratu?* (Dla nas fizyków, jest to prawdę mówiąc kiepski przykład – co za pomysł żeby dodawać pole do długości, przecież te wielkości mają różne jednostki!); $35/60$ – bo Babilończycy posługiwali się znakomitą systemem 60-tkowym. Równanie jednak można zapisać:

$$x^2 + \frac{2}{3}x = \frac{35}{60}.$$

Do problemu jest dołączona instrukcja – jak znaleźć rozwiązanie. Jest to – dokładnie ! – „wzór” (a raczej przepis) na pierwiastek równania kwadratowego!

Inny babiloński problem: dany jest obwód prostokąta $a = 2(x + y)$ i jego pole $b = xy$. Jakie są boki x i y ? I ten problem sprowadza się do równania drugiego stopnia!

A inne stare cywilizacje? W Chinach, w roku 213 p.Ch. (dynastia Ch'in) cesarz nakazał spalenie książek. Ale coś zostało. Znane jest na przykład dzieło *Dziewięć Rozdziałów Sztuki Matematycznej*, obszerna kompilacja, napisana przez kilku autorów, prawdopodobnie w okresie podobnym do powstania *Elementów* Euklidesa, a więc w pierwszej połowie 3. wieku przed Chrystusem. Oryginały uległy spaleni, ale znane są fragmenty dzieła, zachowane i uzupełniane przed późniejszych matematyków. Wersja najbardziej kompletna pochodzi z 3. wieku, ale już po Chrystusie. Dla rozpoczynających naukę algebry warto przytoczyć jeden z problemów:

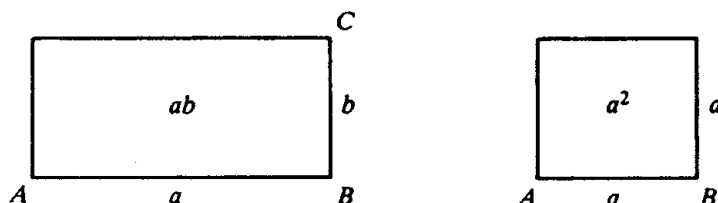
Ziarno może być w jednym z trzech gatunków. Po młóćce uzyskano 3 worki gatunku pierwszego, 2 worki gatunku drugiego i jeden worek gatunku trzeciego. W sumie objętość całego ziarna wyniosła 39 du (jednostka). Dwa worki ziarna pierwszego gatunku, trzy drugiego i jeden trzeciego mają objętość 34 du. A jeden worek pierwszego gatunku, dwa drugiego i trzy trzeciego to 26 du. Ile du zawiera worek każdego z trzech gatunków? Innymi słowy mamy do rozwiązania układ równań:

$$3x + 2y + z = 39$$

$$2x + 3y + z = 34$$

$$x + 2y + 3z = 26.$$

A więc przed Grekami było wielu biegłych matematyków. Skąd inąd, w starożytnej Grecji, a mówiąc dokładniej – w basenie Morza Śródziemnego – w trzecim i czwartym wieku przed Chrystusem matematyka miała się rzeczywiście znakomicie, a algebra – nie najgorzej. We wspomnianych już *Elementach* Euklidesa można znaleźć wiele problemów *par excellence* algebraicznych, rozwiązanych przy pomocy metod ... czysto geometrycznych. Można zaryzykować przypuszczenie, że starożytnym Grekom łatwiej było przeprowadzać pewne rozumowania odniesione do „rzeczywistych” sytuacji, niż atakować bezpośrednio (nieco bardziej abstrakcyjne) problemy czysto rachunkowe. Inaczej mówiąc, zamiast liczb Euklides wolał mieć zawsze do czynienia z odcinkami. Zamiast mówić o iloczynie ab Euklides mówił o prostokącie zbudowanym na odcinkach $AB = a$ i



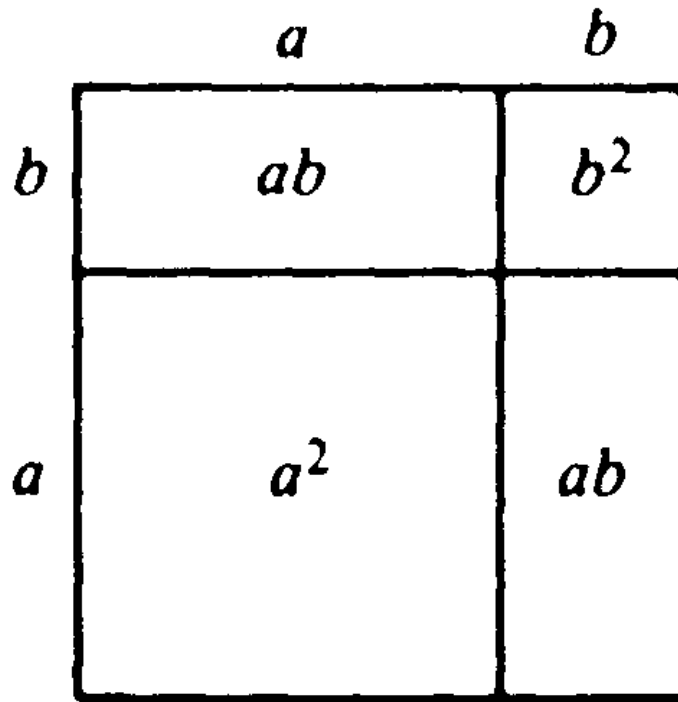
Rysunek 1.1: Iloczyn ab dwóch liczb i kwadrat a^2 .

$BC = b$. Zamiast o kwadracie a^2 o „prawdziwym” kwadracie o krawędzi AB (por. Rys. 1.1). Konsekwentnie,

znany nam dobrze wzór uproszczonego mnożenia (najprostszy przypadek dwumianu Newtona)

$$(a + b)^2 = a^2 + 2ab + b^2$$

miał swoją geometryczną interpretację w takiej postaci jak na rysunku 1.2. W drugiej księdze „*Elementów*”,



Rysunek 1.2: Kwadrat dwumianu $a + b$ w języku geometrii.

tw. propozycja czwarta brzmi: *Jeżeli podzielić odcinek na dwie części, to kwadrat zbudowany na całym odcinku jest równy sumie dwóch kwadratów, zbudowanych na każdej z powstałych części i podwojonego prostokąta zbudowanego z obu odcinków.*

To sprytne, chociaż proste. Ale są i inne, ciekawsze przykłady. Na przykład, równanie pierwszego stopnia (niewiadoma x)

$$ax = bc$$

traktowano jako równość powierzchni ax i bc . Grecy konstruowali najpierw – por. Rys. 1.3 – prostokąt $ABCD$, następnie przedłużali bok BA i odkładali na tym przedłużeniu odcinek a . Prowadząc prostą ED , otrzymujemy – na przecięciu tej prostej z przedłużeniem boku BC – odcinek x , jako odcinek CF . Rzeczywiście, pola dwóch trójkątów EHF i EBF są równe, a po odrzuceniu z dużych trójkątów dwóch par mniejszych (EAD i EKD oraz DCF i DGF) pozostają dwa prostokąty ($KDGH$ i $ABCD$) o równych polach. Równanie pierwszego stopnia to znowu dość trywialny przykład (choć ilustracja geometryczna jest ciekawsza). A równanie kwadratowe? Każde równanie kwadratowe można przekształcić do jednej z trzech postaci:

$$x(x + a) = b^2, \quad x(x - a) = b^2, \quad x(a - x) = b^2, \quad (1-1)$$

a z kolei każda z nich może zostać przełożona na język powierzchni. Na przykład, trzecia postać

$$x^2 + b^2 = ax$$

to konstrukcja zilustrowana na Rys. 1.4. Na odcinku $AB = a$ odkładamy prostokąt $AQFG$ o powierzchni równej b^2 , tak aby powstała figura $QBFL$ była kwadratem; pole $QBLF$ równe jest x^2 . A jak to zrobić? Euklides podaje dokładnie przepis: w środku odcinka AB , w punkcie P wystaw prostopadły odcinek PE o długości b ; następnie narysuj okrąg o środku w E i promieniu $a/2$ (Rys. 1.5). Punkt przecięcia okręgu z odcinkiem AB to punkt Q , przy czym

$$(AQ)(QB) = (PE)^2. \quad (1-2)$$

Widać, że wystarczy już tylko położyć $QB = x$ i spełnione jest trzecie równanie 1-1. Uzasadnienie równości 1-2 nie jest takie trywialne. Trzeba zbudować (por. Rys. 1.6) prostokąt $ABLG$ o bokach a i $x = QB$, dorzucić

języku pól:

$$\begin{aligned} AQFG + HFKC &= (APHG + PQFH) + HFKC \\ &= PBLH + FLDK + HFKC = (PB)^2 \end{aligned}$$

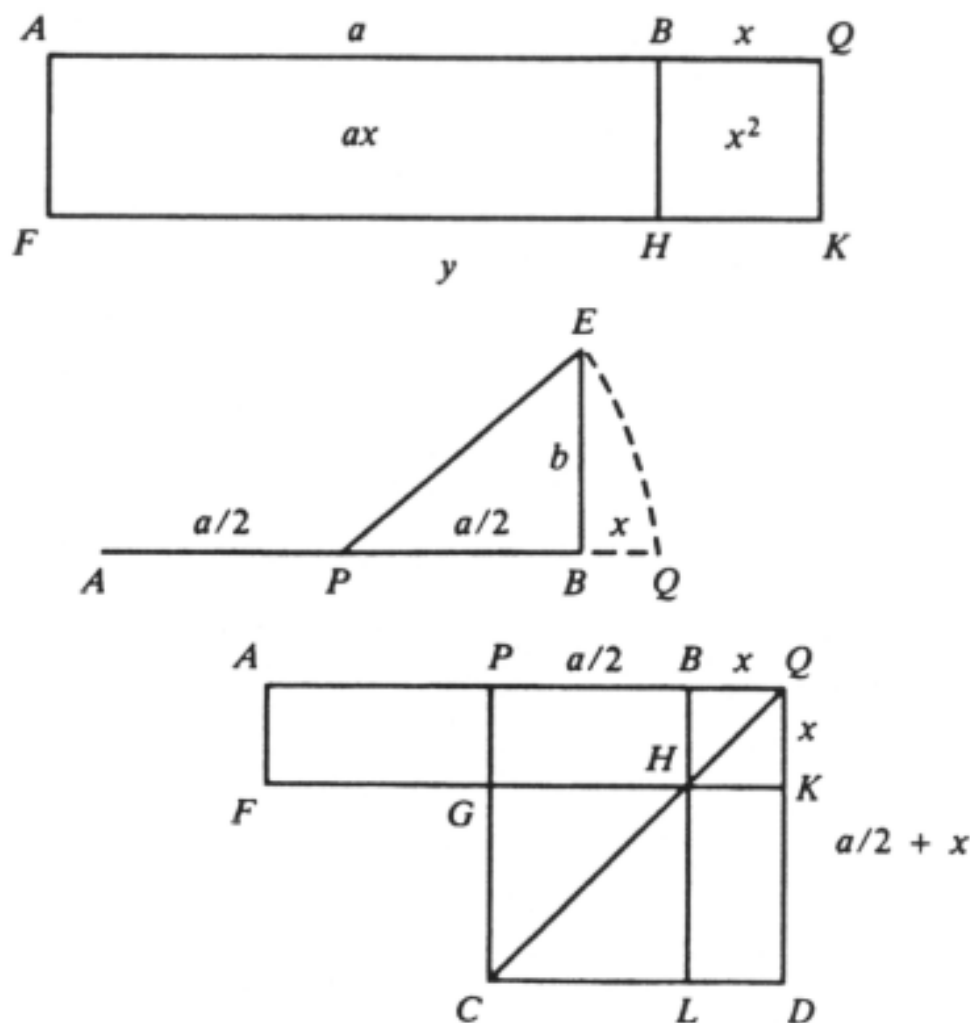
Ale prostokąt $AQFG$ ma pole $(AQ)(QF) = (AQ)(QB)$; kwadrat $HFKC = (PQ)^2$. Czyli

$$(AQ)(QB) + (PQ)^2 = (PB)^2,$$

albo

$$(AQ)(QB) = (PB)^2 - (PQ)^2 = (PE)^2$$

(ukłony od Pitagorasa). Dla tych, którzy są ciekawi swoich możliwości proponuję rozwiązać drugie równanie z trójki 1-1. Jeżeli są kłopoty, to pewną wskazówkę (rysunki analogiczne do przed chwilą prezentowanych) można znaleźć na Rys. 1.7. Warto tylko zauważyć, że jeżeli $b = a^2$ to nasze równanie $(x + a)x = a^2$ to nic



Rysunek 1.7: Pomocnicze rysunki do równania $x^2 + ax = b^2$.

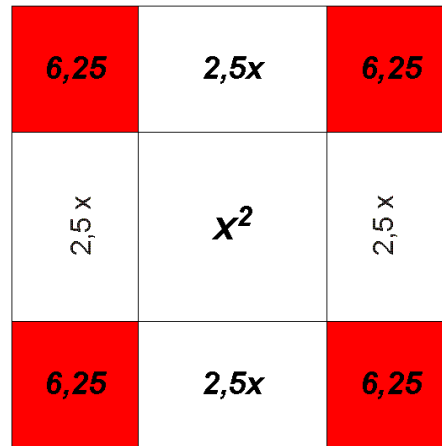
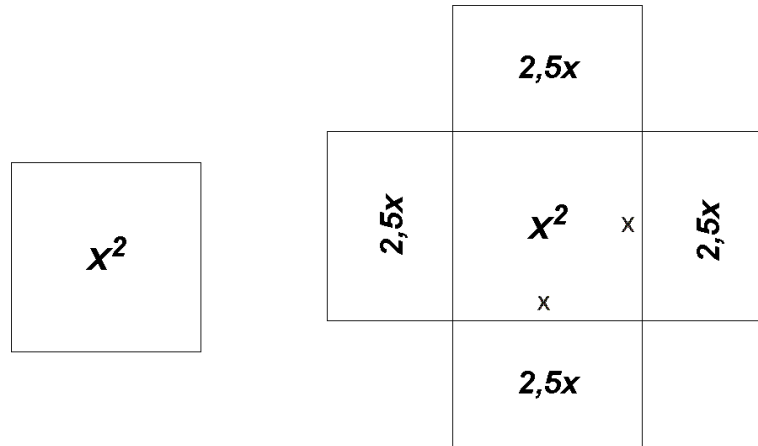
innego jak złoty podział odcinka, o długości $x + a$ na dwie części: x i a , z których dłuższa (a) jest odcinkiem średnim proporcjonalnym pomiędzy krótszą (x) i całym odcinkiem. O złotym podziale odcinka można dużo opowiadać. Jeżeli interesuje Cię wprowadzenie w tę materię to [kliknij tutaj](#).

Te geometryczne przepisy Euklidesa to nic innego jak translacja wzorów babilońskich rachmistrzów, służących do znajdowania pierwiastków równania kwadratowego, na język odcinków i pól figur przy ich pomocy konstruowanych. Ale „przy okazji” ujawnia się ważna różnica: konstrukcja geometryczna jest zawsze możliwa do wykonania, podczas gdy obliczenia algebraiczne wymagają – prędzej czy później – wyciągnięcia pierwiastka. A to ostatnie bardzo często nie jest możliwe, jeżeli mamy do czynienia z liczbami niewymiernymi!

Powróćmy do kalifatu bagdadzkiego. Sam Al-Chowarizmi „czepał pełnymi garściami” z doświadczeń swoich poprzedników. Jego analiza problemu równania kwadratowego przypomina w dużym stopniu podejście geometryczne Euklidesa, chociaż jest już ona nieco bardziej abstrakcyjna. Podejście to

– dotyczące równania $x^2 + 10x = 39$ –

zilustrowane jest Rys. 1.8: kwadrat o powierzchni x^2 (nieznanym boku x) uzupełniamy czterema prostokątami.



Rysunek 1.8: Równanie $x^2 + 10x = 39$.

Każdy z nich ma jeden z boków równy x , a drugi – $2,5$; a więc pole równe $2,5x$. Taki kwadrat + cztery prostokąty to $x^2 + 10x$; teraz dodajemy cztery kwadraciki o polach równych $2,5 \times 2,5 = 6,25$ – powstaje nowy kwadrat o boku: $2,5 + x + 2,5 = x + 5$. Krótko mówiąc:

$$(x + 5)^2 = x^2 + 4 \times 2,5x + 4 \times 6,25 = x^2 + 10x + 25.$$

Ale $x^2 + 10x = 39$. Tak więc:

$$x^2 + 10x + 25 = 39 + 25 = 64,$$

a więc $x + 5 = 8$, $x = 3$. Konstrukcja geometryczna spełnia rolę pomocniczą; w tekście *Liber algebrae* ... można doszukać się bardziej ogólnego algorytmu. Jeżeli startujemy z równania o ogólnej postaci

$$x^2 + px = q$$

to utworzenie pełnego kwadratu po lewej stronie wymaga dodania (do obu stron) wyrazu $4 \left(\frac{p}{4}\right)^2 = p^2/4$:

$$x^2 + px + \frac{p^2}{4} = q + \frac{p^2}{4},$$

albo

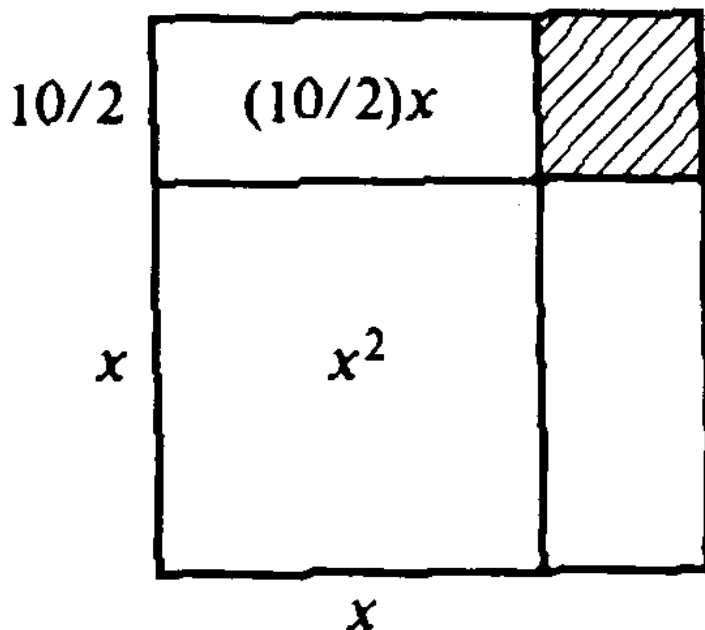
$$\left(x + \frac{p}{2}\right)^2 = q + \left(\frac{p}{2}\right)^2.$$

A z tego wynika już

$$x = \sqrt{q + (p/2)^2} - p/2$$

wzór na pierwiastek równania kwadratowego (a raczej przepis znajdowania jednego (!) pierwiastka) zupełnie nam znajomy.

Warto dodać, że Al-Chowarizmi nie poprzestał na jednej tylko ilustracji; podaje on jeszcze jedną konstrukcję geometryczną, w której zamiast czterech prostokątów mamy dwa, a zamiast czterech kwadracików jeden: (Rys. 1.9). Arabscy matematycy działający w wiekach X–XII (czasy rozkwitu arabskich kalifatów) dokonali



Rysunek 1.9: Równanie $x^2 + 10x = 39$ inaczej.

naprawdę imponującej pracy. Nie tylko zebrali (skompilowali), usystematyzowali i ... przekazali europejskiej cywilizacji całą spuściznę matematyczną Grecji, Bliskiego Wschodu, Egiptu i Azji (Indie!), ale i sami uzyskali cały szereg oryginalnych wyników. I tak, ponieważ przy rozwiązywaniu równań kwadratowych posługiwali się geometrią raczej jako ilustracją, a nie metodą konstrukcyjną, potrafili dostrzec (i przyjąć!) liczby niewymierne. Ba, taki Abû Kâmil (ca. 850-930) nie miał żadnych kompleksów w obliczaniu takich wyrażeń jak:

$$\sqrt{9} - \sqrt{4} = \sqrt{9 - 4 + 2\sqrt{9 \cdot 4}} = \dots 1.$$

W powyższym wzorze wykorzystana jest oczywista równość:

$$\sqrt{a} \pm \sqrt{b} = \sqrt{a + b \pm 2\sqrt{a \cdot b}},$$

choć powyższego wzoru w *Księdze algebry* Abû Kâmil'a ze świecą w ręku szukać. Zamiast tego, jest pełny, werbalny opis postępowania, dla konkretnych wartości a i b .

Ten sam Abû Kâmil proponuje zadanko:

Podziel 10 na dwie części, tak aby jeżeli jedną z nich podzielimy przez drugą, a do wyniku dodamy wynik dzielenia drugiej przez pierwszą, to otrzymamy $4\frac{1}{4}$.

Problem można zapisać w postaci dwóch prostych równań; jedno z nich będzie równaniem drugiego stopnia. Abû Kâmil wprowadza jednak ... trzecią niewiadomą. Zakłada

$$\begin{aligned} x &= 5 - z \\ y &= 5 + z \end{aligned}$$

a warunek na sumę dwóch ilorazów (sprawdź kteczku) staje się prościutkim (choć drugiego stopnia) równaniem dla wyznaczenia niewiadomej z ($z^2 = 9$). Reszta już prosta...

Tak było w czasach starych. A czy można powiedzieć coś o algebrze w czasach mniej starożytnych? Prześledzenie rozwoju tej gałęzi matematyki wykracza poza ramy tej strony. W (bardzo) telegraficznym skrócie

wypada wspomnieć o działającym w 12. wieku Pizańczyku Leonardo [Fibonacci](#), francuskim matematykiem Viète, który był jednym z pierwszych matematyków używających całkiem już współczesnej notacji algebraicznej (pamiętasz wzory Viète'a?). W siedemnastym wieku problemami algebry, w kontekście głównie ogólnej teorii równań algebraicznych, zajmowali się tacy giganci nauki jak Kartezjusz, Newton, d'Alembert i Lagrange. W wieku 18. Cramer i Laplace rozwinęli rachunek wyznaczników. Na przełomie 18. i 19. wieku pojawia się *Princeps Mathematicorum* – [Karol Gauss](#), którego związek z algebrą to teoria liczb zespolonych, a także podstawowe twierdzenie algebry (o istnieniu pierwiastków równania n -tego stopnia).

Wiek 19. to burzliwy rozwój teorii struktur algebraicznych: ciał, grup, ideałów i innych układów o dużym stopniu „złożoności”. Francuz Galois, Niemcy Kummer, Kronecker, kolejni Francuzi: Jordan i Cauchy, Norweg Lie. Pojawia się termin algebra liniowa, której fundamenty budują Anglicy Sylvester i Cayley. W drugiej połowie 19. wieku najważniejsze nazwiska to Niemcy: Hurwitz (algebra wielomianów) i Noether (geometria algebraiczna). Warto poszukać każdego z wymienionych nazwisk, przeczytać o dokonaniach ich właścicieli

....

Rozdział 2

Pojęcie odwzorowania liniowego

Algebra, którą poznaje student pierwszych lat uczelni technicznej, lub uniwersyteckiego wydziału nauk ścisłych nosi zwykle przydomek algebry liniowej. Dlatego też w tym krótkim wstępie dotyczącym prapoczątków warto określić co rozumiemy pod takim właśnie pojęciem, tym bardziej, że termin „liniowy” może mieć nieco inne, mylące konotacje. W naszym przypadku, będziemy mówić o „*odwzorowaniach liniowych*”, czyli o odwzorowaniach posiadających *własność liniowości*. Samo odwzorowanie możemy traktować jako proces, zachodząca według schematu:

$$\text{„coś”} \xrightarrow{\text{proces}} \text{„inne coś”}.$$

Zamiast nieco kolokwialnie „coś” czytelnik może podstawić np. „wejście” (element wejściowy) i „wyjście” (element wyjściowy), lub „sygnał” i „odpowiedź”. Nasuwającym się natychmiast przykładem odwzorowania jest funkcja $y = f(x)$ – odwzorowanie, przyporządkowujące każdemu elementowi zbioru X (dziedziny funkcji) pewien element zbioru Y . Ale nie musimy sięgać do przykładów aż tak sformalizowanych. Odwzorowaniem będzie też np. ilość pieniędzy, jakie nam przyjdzie zapłacić za określoną ilość (masę) określonego (tzn. mającego pewną „cenę jednostkową”) produktu, albo – coś bardziej skomplikowanego – wartości prądów („odpowiedź”) płynących w poszczególnych partiach obwodu elektrycznego w zależności od jego struktury i sił elektromotorycznych („sygnał”).

Własność liniowości jest matematycznym ujęciem dość łatwych do zaakceptowania pojęć proporcjonalności „skutku” do „przyczyny”, a także prostego sumowania się (addytywności) skutków wywoływanych przez różne przyczyny. Innymi słowy: jeżeli x odwzorowuje się na y i dwie przyczyny x_1, x_2 powodują skutki y_1, y_2 , to znaczy

$$x_1 \rightarrow y_1, \quad x_2 \rightarrow y_2,$$

to

$$k_1x_1 + k_2x_2 \rightarrow k_1y_1 + k_2y_2,$$

gdzie k_1, k_2 to pewne liczby. W dalszym ciągu wykładu będziemy wielokrotnie wracać do pojęcia liniowości, ale tu – na początku – warto zwrócić uwagę czytelnika na fakt, że znana nam dobrze *funkcja liniowa*: $y = ax + b$, reprezentująca równanie **linii** prostej na płaszczyźnie xy , **nie** przedstawia sobą – w ogólnym przypadku – odwzorowania liniowego. Jest to logiczne: jeżeli x traktować jako „przyczynę” a y – jako skutek, to nasze odwzorowanie traktowane jako proces powinno dawać zerowy „skutek” ($y = 0$) jeżeli na wejściu „nic się nie dzieje” ($x = 0$). Tak jednak będzie tylko wtedy, jeżeli $b = 0$, a więc nasza prosta przechodzi przez początek układu. Jednocześnie jednak, określenie „układ równań liniowych” (o którym mowa będzie w następnym punkcie) wiąże się właśnie z pojęciem linii prostej na płaszczyźnie.

Termin „liniowy” stanowi często atrybut pewnego *operatora*. Takim operatorem jest na przykład operator pochodnej. Mamy bowiem – dla pochodnej pierwszego rzędu –

$$\frac{d}{dx} \{ \alpha f(x) + \beta g(x) \} = \alpha \frac{df}{dx} + \beta \frac{dg}{dx},$$

($f(x)$ i $g(x)$ – dowolne funkcje; α i β – stałe.) Podobnie dla pochodnej n -tego rzędu.

Ważnym pojęciem będzie też *kombinacja liniowa* elementów należących do pewnego zbioru (np. wektorów). Kombinacja liniowa V układu n elementów V_1, V_2, \dots, V_n to

$$V = a_1V_1 + a_2V_2 + \dots + a_nV_n,$$

gdzie a_1, a_2, \dots, a_n – dowolne stałe.

Algebra liniowa dla potrzeb początkującego użytkownika matematyki w zastosowaniach fizycznych i technicznych ma za zadanie dostarczyć mu dwa podstawowe narzędzia: (1) metody rozwiązywania układów równań liniowych, oraz (2) prawa transformacji wielkości wektorowych, wyrażanych w różnych układach współrzędnych. (Znajomość algebry wektorów jest niesłychanie potrzebna dla omawiania wszystkich problemów fizyki.) Oba te narzędzia oparte są na rachunku macierzowym. Dlatego podstawowy wykład z algebry zaczyna się często od definicji macierzy i podstawowych operacji na tych wielkościach. Łatwiej nam będzie mówić o macierzach i dokonywanych na nich (przy ich pomocy) operacjach jeżeli skojarzymy je z konkretnym zagadnieniem – mianowicie układami równań liniowych. Cały czas będziemy jednak pamiętać, że są one wielce pomocne w przedstawianiu i rozwiązywaniu zagadnień związanych z rachunkiem wektorowym, dlatego następnym zagadnieniem które będziemy omawiać będzie algebra wektorów. Okaże się przy tym, że w opisie pewnych operacji dokonywanych na wektorach okazały się bardzo pomocne *liczby zespolone*, o których mówić będziemy w następnej kolejności. Dopiero po tym wstępie na dobre zajmiemy się macierzami oraz bardziej abstrakcyjną algebrą przestrzeni wektorowych. Zauważmy w tym momencie, że rachunek macierzowy – podstawowe narzędzie algebry – ma znacznie szersze i głębsze aspekty. I tak na przykład, fundament nowoczesnej fizyki, mechanika kwantowa została sama zbudowana na „fundamencie” rachunku macierzowego. Występujące w niej wielkości opisujące stan układów fizycznych (w mikroświecie cząsteczki, atomu czy jądra atomowego) są **macierzami**. Macierzami są też wielkości fizyczne, takie jak energia kinetyczna, moment pędu, itp. Nic więc dziwnego, że mechanika kwantowa w pierwszych dekadach swego rozwoju nazywała się mechaniką macierzową.

Rozdział 3

Układy liniowych równań; metoda Gaussa; wyznaczniki

3.1 Układ równań liniowych

Układ m równań liniowych o n niewiadomych zapisujemy w ogólnej postaci

$$\begin{array}{cccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1, \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & b_2, \\ \dots & + & \dots & + & \dots & + & \dots & = & \dots, \\ a_{m-1,1}x_1 & + & a_{m-1,2}x_2 & + & \dots & + & a_{m-1,n}x_n & = & b_{1,m-1}, \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & b_m. \end{array} \quad (3-1)$$

Wielkości x_i , $i = 1, \dots, n$ to (szukane) niewiadome; współczynniki a_{ki} , $k = 1, 2, \dots, m$, $i = 1, 2, \dots, n$ to (znane) wielkości liczbowe. Stałe (określone) b_i ; $i = 1, 2, \dots, m$ nazywamy wyrazami wolnymi lub niejednorodnościami równań. Jeżeli *wszystkie* te stałe są równe zeru układ nazywamy układem równań jednorodnych, lub – w skrócie – układem jednorodnym. Dodajmy, że tym przypadkiem określenie „liniowy” związane jest (por. poprzedni rozdział) z równaniem opisującym linię prostą – niewiadome występują tylko w pierwszej potędze i nie mogą być pomnożone przez siebie. Konkretnie – najprostszy układ równań:

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 & = & b_1, \\ a_{21}x_1 + a_{22}x_2 & = & b_2 \end{array} \quad (3-2)$$

to równania dwóch prostych (na płaszczyźnie $0x_1x_2$); szukane rozwiązanie to – w zasadzie – współrzędne punktu ich przecięcia.

Rozwiązaniem układu 3-1 jest zbiór n liczb: k_1, k_2, \dots, k_n które podstawione za x -y do układu ($x_i = k_i$; $i = 1, \dots, n$) sprawiają, że każde z równań układu staje się tożsamością. Podkreślmy, że rozwiązanie układu to *zbiór* n liczb.

Układ 3-1 może nie mieć rozwiązań, może mieć nieskończenie wiele rozwiązań, lub może mieć jednoznaczne rozwiązanie – n wielkości x_1, \dots, x_n dla których m (w tym przypadku $m = n$ – patrz niżej) równań staje się równościami typu: lewa strona = prawa strona. Łatwo to zrozumieć, posługując się wspomnianym powyżej (równ. 3-2) przykładem dwóch prostych: mogą one przecinać się w jednym punkcie (jednoznaczne rozwiązanie), mogą być do siebie równoległe (brak rozwiązań), mogą wreszcie pokrywać się (nieskończenie wiele rozwiązań – każdy punkt jednej z prostych należy do drugiej).

Od czego zależy taki a nie inny charakter rozwiązania? Układ 3-1 to n pytań $x_i = ?$, $i = 1, \dots, n$ – udzielenie jednoznacznej odpowiedzi na wszystkie pytania wymaga n niezależnych (to znaczy: nie powtarzających się!) i nie sprzecznych informacji. Dla jednoznacznego charakteru rozwiązania warunkiem koniecznym będzie $n = m$, chociaż nie jest to warunek wystarczający. Mogą być bowiem informacje powtarzające się: na przykład układ, cytowany 250 lat temu przez Leonarda Eulera::

$$\left. \begin{array}{rcl} 3x - 2y & = & 5, \\ 4y - 6x & = & -10. \end{array} \right\} \quad (3-3)$$

Rozwiązując ten układ tradycyjną metodą podstawienia dostajemy

$$x = a, \quad y = (3a - 5)/2.$$

co oznacza, że rozwiązaniem układu jest para: dowolna stała a i $(3a-5)/2$. Układ jest *układem nieoznaczonym*. Jego „nieoznaczoność” – nieskończona liczba możliwych rozwiązań – jest skutkiem pewnej fikcji. „Układ” 3-3 jest tylko formalnie układem *dwóch* równań z dwoma niewiadomymi. Drugie równanie powstało z pierwszego w wyniku pomnożenia go przez -2 . Takie równanie nie wnosi nowej informacji o stosunkach, jakie muszą zachodzić pomiędzy zmiennymi, tylko powieliła informację z pierwszego równania. Mówimy, że oba równania są *liniowo zależne* (por. 6.2).

Możemy też mieć do czynienia z układem sprzecznym:

$$\left. \begin{aligned} 3x + y &= 7, \\ 6x + 2y &= 35. \end{aligned} \right\} \quad (3-4)$$

Nie ma on rozwiązań. Dwa równania – dwie niezależne informacje o x i y są wyraźnie sprzeczne.

Wreszcie układ dwóch równań

$$\left. \begin{aligned} 2x_1 + x_2 &= 7, \\ x_1 - 3x_2 &= -2 \end{aligned} \right\}$$

ma jednoznaczne rozwiązanie – dwie niewiadome x_1 i x_2 są równe $19/7$ i $11/7$ (możesz to sprawdzić stosując znane Ci standardowe metody, np. metodę podstawiania). Układ taki nazywamy układem niesprzecznym i *oznaczonym*.

Macierze

Współczynniki a_{ik} , występujące po lewej stronie równań układu 3-1 można uporządkować w formie tablicy, albo *macierzy*:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1k} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2k} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{l1} & a_{l2} & \dots & a_{lk} & \dots & a_{ln} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mk} & \dots & a_{mn} \end{pmatrix} \equiv \mathcal{A}. \quad (3-5)$$

Macierz \mathcal{A} ma m wierszy – indeks (wskaźnik) wiersza to pierwszy ze wskaźników współczynników a_{lk} oraz n kolumn – indeks (wskaźnik) kolumny to drugi ze wskaźników współczynników a_{lk} . Jeżeli $m = n$ (zazwyczaj najbardziej interesujący nas przypadek!) to macierz nazywamy *macierzą kwadratową* stopnia n .

Algebra macierzy to treść rozdziałów następujących (praktycznie od szóstego wzwyż). Tutaj wprowadźmy tylko pojęcie *macierzy transponowanej* \mathcal{A}^T (w stosunku do danej macierzy \mathcal{A}). Jest to macierz w której wiersze zostały wymienione (transponowane) z kolumnami i *vice versa*:

$$\mathcal{A}^T \equiv \begin{pmatrix} a_{11} & a_{21} & \dots & a_{l1} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{l2} & \dots & a_{m2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{1k} & a_{2k} & \dots & a_{lk} & \dots & a_{mk} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{ln} & \dots & a_{mn} \end{pmatrix}. \quad (3-6)$$

W odniesieniu do *elementów macierzowych* transpozycja oznacza:

$$a_{ik}^T = a_{ki}; \quad k = 1, \dots, m; \quad i = 1, \dots, n. \quad (3-7)$$

3.2 Metoda Gaussa

Układ równań można rozwiązywać metodą podstawienia. Jej formalne „uogólnienie” to tzw. metoda Gaussa. Zanim prześledzimy jej schemat zastanówmy się, jakich operacji możemy dokonywać na równaniach tworzących układ, aby jego charakter nie uległ zmianie. Po pierwsze, na pewno można uszeregować nasze m równań w dowolnej kolejności. Jeżeli chodzi o operacje matematyczne, to obie strony każdego z równań możemy mnożyć przez różną od zera liczbę, do obu stron możemy dodawać (odejmować) te same liczby. A ponieważ traktujemy *a priori* równania jako tożsamości (po podstawieniu za x_i rozwiązania układu – por. wyżej) od

obu stron każdego z równań możemy na przykład odjąć odpowiednie strony innego równania, przemnożone przez pewną liczbę. Powstający – w wyniku zastosowania takich operacji – układ nazywamy *układem równań liniowych równoważnym wyjściowemu* układowi (albo krótko: oba układy nazywamy równoważnymi). Metoda Gaussa opiera się właśnie na sukcesywnym zastępowaniu wyjściowego układu równań układem równoważnym. Schemat metody Gaussa przedstawiamy w formie przykładu.

Przykład 3.1 Do rozwiązania mamy układ równań:

$$\left. \begin{aligned} 3x + 2y + z &= 11, \\ 2x + 3y + z &= 13, \\ x + y + 4z &= 12. \end{aligned} \right\} \quad (3-8)$$

Metoda Gaussa, w swojej najprostszej wersji, sprowadza się do sukcesywnej eliminacji niewiadomych z kolejnych równań, metodą podstawienia. W tym celu dzielimy obie strony równań tworzących układ 3-8 przez współczynniki x -a:

$$\left. \begin{aligned} x + \frac{2}{3}y + \frac{1}{3}z &= \frac{11}{3}, \\ x + \frac{3}{2}y + \frac{1}{2}z &= \frac{13}{2}, \\ x + y + 4z &= 12. \end{aligned} \right\} \quad (3-9)$$

i podstawiamy w drugim i trzecim za x z pierwszego (odejmujemy stronami pierwsze od drugiego i trzeciego):

$$\left. \begin{aligned} x + \frac{2}{3}y + \frac{1}{3}z &= \frac{11}{3}, \\ \frac{5}{6}y + \frac{1}{6}z &= \frac{17}{6}, \\ \frac{1}{3}y + \frac{11}{3}z &= \frac{25}{3}. \end{aligned} \right\} \quad (3-10)$$

Zmienna x zniknęła z drugiego i trzeciego równania. Powtarzamy procedurę w odniesieniu do równania drugiego i trzeciego: dzielimy przez współczynniki przy y

$$\left. \begin{aligned} x + \frac{2}{3}y + \frac{1}{3}z &= \frac{11}{3}, \\ y + \frac{1}{5}z &= \frac{17}{5}, \\ y + 11z &= 25 \end{aligned} \right\} \quad (3-11)$$

i podstawiamy w trzecim za y z drugiego:

$$\left. \begin{aligned} x + \frac{2}{3}y + \frac{1}{3}z &= \frac{11}{3}, \\ y + \frac{1}{5}z &= \frac{17}{5}, \\ z &= 2. \end{aligned} \right\} \quad (3-12)$$

Jedna niewiadoma jest już określona – $z = 2$. Drugie równanie układu 3-12 służy do wyliczenia drugiej niewiadomej: $y = 17/5 - z/5 = 3$, a pierwsze – do wyliczenia x -a: $x = 11/3 - 2y/3 - z/3 = 1$.

Zauważmy, że metoda eliminacji Gaussa sprowadza wyjściową macierz układu równań 3-9

$$\begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 4 \end{pmatrix} \quad (3-13)$$

do tzw. postaci trójkątnej (por. 3-12)

$$\begin{pmatrix} 1 & \frac{2}{3} & \frac{1}{3} \\ 0 & 1 & \frac{1}{5} \\ 0 & 0 & 1 \end{pmatrix}. \quad (3-14)$$

Istnieją liczne warianty metody Gaussa, w tym metoda Gaussa-Jordana, która od metody Gaussa różni się tym, że każde „nowe” równanie służy do wyeliminowania kolejnej nie tylko z następnych, ale także poprzednich

równań. I tak drugie równanie układu 3-11 służy do wyeliminowania y -a nie tylko z trzeciego, ale i pierwszego; wówczas trzecie równanie wykorzystywane jest do eliminacji z -a z pierwszego i drugiego równania. Układ przybiera postać będącą jednocześnie rozwiązaniem:

$$\begin{aligned} x &= 1 \\ y &= 3 \\ z &= 2. \end{aligned} \tag{3-15}$$

Do metody eliminacji Gaussa powrócimy jeszcze w rozdziale szóstym (układy równań) i siódmym (konstrukcja macierzy odwrotnej).

Podany powyżej przykład to scenariusz w pełni „pozytywny”; uzyskaliśmy jednoznaczne rozwiązanie układu. A inne, mniej pozytywne scenariusze?

Przykład 3.2 Weźmy układ

$$\left. \begin{aligned} x_1 - 5x_2 - 8x_3 + x_4 &= 3, \\ 3x_1 + x_2 - 3x_3 - 5x_4 &= 1, \\ x_1 - 7x_3 + 2x_4 &= -5, \\ 11x_2 + 20x_3 - 9x_4 &= 2. \end{aligned} \right\} \tag{3-16}$$

Cztery niewiadome i cztery równania. Zastosowanie metody Gaussa przekształca wyjściową macierz współczynników a_{lk} , uzupełnioną o kolumnę wyrazów wolnych b_l według schematu:

$$\begin{aligned} &\left(\begin{array}{cccc|c} 1 & -5 & -8 & 1 & 3 \\ 3 & 1 & -3 & -5 & 1 \\ 1 & 0 & -7 & 2 & -5 \\ 0 & 11 & 20 & -9 & 2 \end{array} \right) \rightarrow \left(\begin{array}{cccc|c} 1 & -5 & -8 & 1 & 3 \\ 0 & 16 & 21 & -8 & -8 \\ 0 & 5 & 1 & 1 & -8 \\ 0 & 11 & 20 & -9 & 2 \end{array} \right) \\ &\rightarrow \left(\begin{array}{cccc|c} 1 & -5 & -8 & 1 & 3 \\ 0 & -89 & 0 & -29 & 160 \\ 0 & 5 & 1 & 1 & -8 \\ 0 & -89 & 0 & -29 & 162 \end{array} \right) \rightarrow \left(\begin{array}{cccc|c} 1 & -5 & -8 & 1 & 3 \\ 0 & -89 & 0 & -29 & 160 \\ 0 & 5 & 1 & 1 & -8 \\ 0 & 0 & 0 & 0 & 2 \end{array} \right) \end{aligned}$$

Ostatnia linijka to

$$0 = 2;$$

otrzymaliśmy sprzeczność – nasz układ jest układem sprzecznym.

Przykład 3.3 Ostatni scenariusz:

$$\left. \begin{aligned} 4x_1 + x_2 - 3x_3 - x_4 &= 0, \\ 2x_1 + 3x_2 + x_3 - 5x_4 &= 0, \\ x_1 - 2x_2 - 2x_3 + 3x_4 &= 0. \end{aligned} \right\} \tag{3-17}$$

Układ jest układem jednorodnym – cztery niewiadome i tylko trzy równania. Ponieważ liczba równań jest mniejsza od liczby niewiadomych będzie to układ nieoznaczony. Wystarczy przekształcać samą wyjściową macierz współczynników a_{lk} :

$$\left(\begin{array}{cccc} 4 & 1 & -3 & -1 \\ 2 & 3 & 1 & -5 \\ 1 & -2 & -2 & 3 \end{array} \right) \rightarrow \left(\begin{array}{cccc} 0 & 9 & 5 & -13 \\ 0 & 7 & 5 & -11 \\ 1 & -2 & -2 & 3 \end{array} \right) \rightarrow \left(\begin{array}{cccc} 0 & 2 & 0 & -2 \\ 0 & 7 & 5 & -11 \\ 1 & -2 & -2 & 3 \end{array} \right)$$

Nowy (równoważny!) układ to

$$\left. \begin{aligned} 2x_2 - 2x_4 &= 0, \\ 7x_2 + 5x_3 - 11x_4 &= 0, \\ x_1 - 2x_2 - 2x_3 + 3x_4 &= 0. \end{aligned} \right\} \tag{3-18}$$

W takiej sytuacji jedną z niewiadomych – x_2 lub x_4 – musimy przyjąć dowolnie; na przykład, kładąc $x_4 = \alpha$. Pierwsze równanie układu 3-18 daje $x_2 = \alpha$; drugie – $x_3 = \frac{4}{5}\alpha$; trzecie – $x_1 = \frac{3}{5}\alpha$.

Czy możemy mieć do czynienia z układem równań, w którym liczba niezależnych i nie sprzecznych równań jest większa od liczby niewiadomych n ? Z algebraicznego punktu widzenia nie ma to sensu; ale sytuację taką spotykamy – i to bardzo często – w tzw. metodzie regresji liniowej. Nie należy to jednak do tematu tego wykładu, a raczej do wykładu ze statystycznych metod opracowywania danych pomiarowych.

W punkcie 6.6 podając podsumowanie metod rozwiązywania układów równań określimy reguły, które pozwalają zidentyfikować z jakimi typami układów (i ich rozwiązań) mamy do czynienia.

3.3 Wyznacznik drugiego i trzeciego stopnia

Zaprezentowana metoda Gaussa jest prosta, skuteczna i dobrze nadaje się do wykorzystania w komputerowych algorytmach obliczeniowych. Jej podstawową wadą jest to, że *a priori* nie potrafimy określić z jakim typem układu (sprzecznym/niesprzecznym; oznaczonym/nieoznaczonym) mamy do czynienia. Dlatego „obok” metody Gaussa powstała bardziej formalna teoria układów równań liniowych. Zanim sformułujemy ją w sposób bardziej ogólny, prześledźmy raz jeszcze najprostszy i najciekawszy scenariusz: oznaczony układ dwóch równań i zawierających dwie niewiadome:

$$\left. \begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1, \\ a_{21}x_1 + a_{22}x_2 &= b_2. \end{aligned} \right\} \quad (3-19)$$

Współczynniki a_{ik} tworzą macierz drugiego stopnia:

$$\mathcal{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Pierwsze równanie układu 3-19 mnożymy przez a_{22} , drugie – przez a_{12} i odejmujemy stronami; podobnie od drugiego równania, przemnożonego przez a_{11} odejmujemy pierwsze przemnożone przez a_{21} . Układ 3-19 przybiera postać:

$$\left. \begin{aligned} (a_{11}a_{22} - a_{12}a_{21})x_1 &= b_1a_{22} - a_{12}b_2, \\ (a_{11}a_{22} - a_{12}a_{21})x_2 &= b_2a_{11} - a_{21}b_1. \end{aligned} \right\} \quad (3-20)$$

Pod warunkiem, że $a_{11}a_{22} - a_{12}a_{21} \neq 0$, mamy

$$x_1 = \frac{b_1a_{22} - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}}, \quad x_2 = \frac{b_2a_{11} - a_{21}b_1}{a_{11}a_{22} - a_{12}a_{21}}. \quad (3-21)$$

Wspólny mianownik w powyższych wzorach wyraża się jednoznacznie przez współczynniki tworzące macierz \mathcal{A} — jest to różnica iloczynów wyrazów leżących na tzw. *przekątnej głównej* ($a_{11}a_{22}$) i drugiej przekątnej ($a_{12}a_{21}$). Jest to pewna wartość liczbową – *wyznacznik macierzy \mathcal{A}* , albo wyznacznik stopnia drugiego (tak jak macierz):

$$\det \mathcal{A} \equiv |a_{ik}| \equiv \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}. \quad (3-22)$$

(Wszystkie podane wyżej formy zapisu wyznacznika są równoważne i równie często spotykane w literaturze. Skrót „det” pochodzi od słowa łacińskiego *determinare* – oznaczać, wyznaczać. Angielska nazwa wyznacznika to determinant; bardzo podobnie brzmią nazwy w językach romańskich, niemieckim, itp.)

Z definicji wyznacznika wynika, że jest to wielkość liczbową, stowarzyszona z macierzą o tej samej liczbie wierszy i kolumn; z macierzą kwadratową. Wyznacznik jednoelementowej macierzy $\mathcal{A} = (a_{11})$ to liczba a_{11} . Oba liczniki w równaniach 3-21 mają postać analogiczną do mianownika. Są to także wyznaczniki drugiego stopnia. Łatwo sprawdzić, że licznik ułamka określającego x_1 to wyznacznik, w którym kolumna współczynników tej zmiennej (a_{11}, a_{21}) została zastąpiona kolumną wyrazów wolnych (b_1, b_2); podobnie licznik ułamka określającego x_2 to wyznacznik, w którym kolumna wyrazów wolnych została zastąpiona kolumną współczynników tej zmiennej (a_{12}, a_{22}); innymi słowy

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}. \quad (3-23)$$

Tak jest dla układu dwóch równań z dwoma niewiadomymi. Dla układu trzech równań wzory będą analogiczne. Układ taki to

$$\left. \begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3, \end{aligned} \right\} \quad (3-24)$$

a jego macierz współczynników a_{ik} to macierz trzeciego stopnia:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}. \quad (3-25)$$

Jeżeli wzory mają być analogiczne, to:

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}, \quad x_3 = \frac{\begin{vmatrix} a_{11} & a_{23} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}}. \quad (3-26)$$

Powstaje jednak problem – jak (według jakich reguł) obliczać wyznacznik trzeciego stopnia? Można rozwiązać układ 3-24 metodą sukcesywnej eliminacji poszczególnych zmiennych; na przykład mnożąc pierwsze równanie układu przez $a_{22}a_{33} - a_{23}a_{32}$, drugie przez $a_{13}a_{32} - a_{12}a_{33}$, a trzecie przez $a_{12}a_{23} - a_{13}a_{22}$ otrzymujemy wyrażenie typu

$$\text{stała 1} \cdot x_1 = \text{stała 2},$$

gdzie stała 1 wyraża się przez a_{ik} , a stała 2 – przez a_{ik} i b_i . Stała 1 to nic innego jak mianownik wzorów 3-26. Łatwo sprawdzić, że

$$|A| = a_{11}(a_{22}a_{33} - a_{23}a_{32}) + a_{12}(a_{23}a_{31} - a_{21}a_{33}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \quad (3-27)$$

$$= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}. \quad (3-28)$$

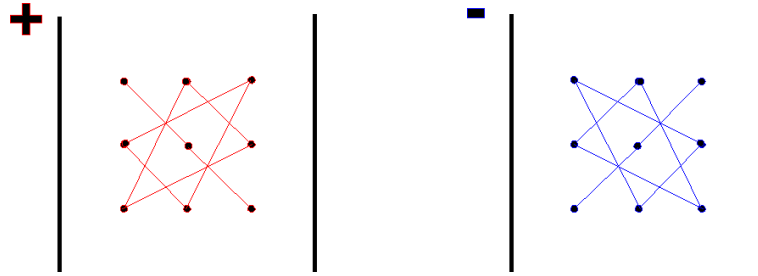
Wyznacznik to suma trzech składników, z których każdy jest iloczynem kolejnego wyrazu pierwszej wiersza a_{1i} , ($i = 1, 2, 3$) oraz wyznacznika, powstałego przez wykreślenie wiersza i kolumny w którym znajduje się wyraz a_{1i} (zawsze będzie to pierwszy wiersz i – kolejno – pierwsza, druga i trzecia kolumna), pomnożonego jeszcze przez $(-1)^{1+i}$. Jest to tzw. rozwinięcie wyznacznika względem pierwszego wiersza. Powrócimy do tego zagadnienia w następnych punktach.

W tym miejscu zauważmy tylko, że przyglądając się postaci wyznacznika trzeciego stopnia (3-25 – macierz 3×3) i powyższym rozwinięciom nietrudno też podać „praktyczne” reguły obliczania takiego wyznacznika. Są one zilustrowane na Rys.3.3. Trzy dodatnie iloczyny w sumie otrzymujemy wymnażając trzy wyrazy leżące na *głównej przekątnej*, biegnącej z lewa na prawo i „w dół”, oraz wyrazy na równoległych do niej mniejszych przekątnych (2 wyrazy), górnej i dolnej, przy czym iloczyn z danej mniejszej przekątnej przemnożony jest jeszcze przez (trzeci) wyraz z narożnika bardziej od niej odległego. Trzy ujemne składniki sumy 3-27 otrzymujemy analogicznie, rozpatrując przekątne biegnące ku górze.

3.3.1 Twierdzenie Cramera

Ten prosty schemat dla układów 2, lub 3 równań może zostać uogólniony na przypadek układu równań o dowolnej liczbie niewiadomych. Zauważmy, że chcąc używać do ich rozwiązywania wyznaczników, *konstrukcji o tej samej liczbie wierszy i kolumn*, musimy założyć, że liczba równań (m) jest równa liczbie niewiadomych (n). Obowiązuje wówczas

Twierdzenie 3.1 (Cramera) *W przypadku układu równań niejednorodnych, których liczba m jest równa liczbie występujących w nich niewiadomych x_1, x_2, \dots, x_n , rozwiązanie układu uzyskujemy w postaci ułamków, takich jak wzory 3-23 i 3-26,*



Rysunek 3.1: Wyznacznik trzeciego stopnia – dodatnie i ujemne przyczynki.

pod warunkiem, że mianownik tych ułamków – wyznacznik główny, utworzony ze współczynników niewiadomych x -ów występujących po lewej stronie równań, jest różny od zera. Liczniki, to wyznaczniki, w którym kolumna współczynników danego x_i została zastąpiona przez kolumnę wyrazów wolnych (wyrazy b_1, b_2, \dots, b_n w 3-1.)

Dowód chwilowo pomijamy. Powrócimy do niego w rozdziale siódmym (punkt 7.1.5).

3.4 Algebra wyznaczników

Przypuśćmy, że mamy do czynienia z wyznacznikiem Δ stopnia n . Wybierzmy liczbę k ; $1 \leq k \leq n - 1$. W wyznaczniku wybieramy teraz k dowolnych kolumn i k dowolnych wierszy; elementy wyznacznika¹ leżące na przecięciu tych wierszy i kolumn tworzą macierz stopnia k . Wyznacznik tej macierzy nazywamy *minorem stopnia k* wyznacznika Δ . Alternatywnie taki minor M możemy uzyskać wykreślając z wyznacznika Δ $n - k$ wierszy i $n - k$ kolumn.

Weźmy minor M stopnia k wyznacznika Δ , który jest wyznacznikiem n -tego stopnia. Wykreślając wszystkie k wierszy i k kolumn, na przecięciach których leżą elementy tworzące minor M , uzyskujemy minor M' , stopnia $n - k$. Jest to *minor dopełniający* minora M . Możemy więc mówić o parach minorów dopełniających wyznacznika Δ . W szczególności, parę takich minorów tworzyć będą: dowolny element a_{ij} i minor stopnia $n - 1$, uzyskany poprzez skreślenie i -tego wiersza i j -ej kolumny w wyznaczniku.

Jeżeli minor M tworzą elementy k pierwszych wierszy i k pierwszych kolumn, to para minorów dopełniających M i M' wygląda tak:

$$\Delta = \left| \begin{array}{ccc|ccc} a_{11} & \dots & \dots & a_{1k} & a_{1,k+1} & \dots & \dots & a_{1n} \\ \dots & \dots & M & \dots & \dots & \dots & \dots & \dots \\ a_{k1} & \dots & \dots & a_{kk} & a_{k,k+1} & \dots & \dots & a_{kn} \\ \hline a_{k+1,1} & \dots & a_{k+1,k} & a_{k+1,k+1} & \dots & a_{k+1,n} \\ \dots & \dots & \dots & \dots & M' & \dots \\ a_{n,1} & \dots & a_{nk} & a_{n,k+1} & \dots & a_{nn} \end{array} \right| \quad (3-29)$$

Jeżeli minor M stopnia k jest utworzony przez elementy, których wskaźniki wierszy przyjmują wartości: i_1, i_2, \dots, i_k , a wskaźniki kolumn: j_1, j_2, \dots, j_k , to *dopełnieniem algebraicznym minora M* nazywamy jego minor dopełniający M' , przemnożony przez znak -1 lub $+1$, w zależności od tego czy suma wszystkich wskaźników wierszy i kolumn elementów M jest parzysta, czy nieparzysta. Innymi słowy, wprowadzamy wielkość s_M :

$$s_M = i_1 + i_2 + \dots + i_k + j_1 + j_2 + \dots + j_k \quad (3-30)$$

i definiujemy dopełnienie algebraiczne minora M , D_M jako

$$D_M \equiv (-1)^{s_M} M'. \quad (3-31)$$

W szczególności, ważnym przypadkiem jest sytuacja, kiedy „minorem” jest pojedynczy element $-a_{pq}$; wówczas jego dopełnienie algebraiczne, oznaczane nieco inaczej, definiujemy jako

$$\text{dopełnienie algebraiczne elementu } a_{pq} \equiv M^{[pq]} = (-1)^{p+q} M',$$

¹Mówiąc o elementach, wierszach i kolumnach wyznacznika myślimy *de facto* o elementach, wierszach i kolumnach *macierzy*, której odpowiada dany wyznacznik.

gdzie minor M' jest minorem, powstałym przez skreślenie p -tego wiersza i q -tej kolumny wyznacznika. Zauważmy teraz, że „przepis” na obliczanie wyznacznika trzeciego stopnia (wzór 3-27)

$$\begin{aligned} |A| &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) + a_{12}(a_{23}a_{31} - a_{21}a_{33}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \\ &= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}. \end{aligned}$$

– tzw. rozwinięcie wyznacznika względem pierwszego wiersza – to suma trzech składników, z których każdy jest iloczynem kolejnego wyrazu pierwszego wiersza a_{1i} , $i = 1, 2, 3$ oraz jego dopełnienia algebraicznego (wyznacznika drugiego stopnia).

Nietrudno sprawdzić, że zamiast rozwijać nasz wyznacznik względem pierwszego wiersza, moglibyśmy rozwijać go względem pierwszej kolumny, bowiem

$$\begin{aligned} |A| &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \\ &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) + a_{21}(a_{13}a_{32} - a_{12}a_{33}) + a_{31}(a_{12}a_{23} - a_{13}a_{22}) \end{aligned} \quad (3-32)$$

– każdy z trzech składników jest iloczynem kolejnego wyrazu pierwszej kolumny a_{1j} , $j = 1, 2, 3$ oraz jego dopełnienia algebraicznego (znowu wyznacznika drugiego stopnia).

Powyższe wzory można zapisać w postaci uogólnionej – dla wyznacznika rzędu n , który możemy rozwijać względem dowolnego wiersza (p) lub kolumny (q):

$$|A| = \sum_{q=1}^n a_{pq} \times M^{[pq]} = \sum_{p=1}^n a_{pq} \times M^{[pq]}. \quad (3-33)$$

Każda kolumna (każdy wiersz) są równorzędne. (Porządek kolumn i wierszy zależy od zupełnie dowolnego ustalenia symboli niewiadomych oraz kolejności wypisywania informacji). Takie *rozwinięcie wyznacznika* względem kolumn i wierszy nazywa się *rozwinięciem Laplace’a*.

Wyznaczniki można obliczać według bardziej „podstawowych” reguł, nie korzystając z rozwinięcia Laplace’a. Taka podstawowa reguła jest nieco skomplikowana: tworzymy n -czynnikowe iloczyny, przy czym każdy czynnik brany jest z kolejnego wiersza (pierwszy symbol $i = 1, 2, \dots, n$) i dowolnie wybranej – ale różnej! – kolumny (drugi symbol $k = k_1, k_2, \dots, k_n$). Liczba takich iloczynów to $n!$ (liczba permutacji w zbiorze k_1, k_2, \dots, k_n). Te $n!$ iloczynów dzielimy na dwie, równoliczne ($n!/2$) klasy – parzystą i nieparzystą. Przynależność do określonej klasy określa nam porządek drugich wskaźników: dla cyklicznych (parzystych) permutacji² $\{1, 2, \dots, n\}$ iloczyn określamy jako parzysty i przypisujemy mu znak dodatni, dla nieparzystej permutacji iloczyn będzie nieparzysty, a jego znak – ujemny. Parzyste (dodatnie) będą więc – dla $n = 3$ – iloczyny $a_{11}a_{22}a_{33}$, $a_{12}a_{23}a_{31}$, $a_{13}a_{21}a_{32}$, nieparzyste (ujemne) – iloczyny $a_{11}a_{23}a_{32}$, $a_{12}a_{21}a_{33}$, $a_{13}a_{22}a_{31}$. Suma algebraiczna tych $n!$ iloczynów to wartość wyznacznika $|A|$:

$$\sum (-1)^p a_{1k_1} a_{2k_2} \dots a_{nk_n}, \quad (3-34)$$

gdzie sumowanie przebiega po wszystkich możliwych permutacjach ciągu drugich wskaźników k_1, k_2, \dots, k_n , a wykładnik potęgi -1 (liczba p) to liczba przestawień potrzebnych do uporządkowania tego ciągu w kolejności naturalnej $1, 2, \dots, n$. Brzmi to dość skomplikowanie, ale w praktyce najczęściej mamy do czynienia z wyznacznikami drugiego i trzeciego rzędu, których podane wyżej reguły obliczania absolutnie trzeba zapamiętać.

Własności wyznaczników

Wnioski dotyczące wartości wyznaczników oraz konsekwencji dokonywanych na nich operacji możemy przedstawić w punktach:

1. Wartość wyznacznika zmienia znak, gdy przestawimy dwie kolumny, lub dwa wiersze.
2. Wartość wyznacznika jest równa zeru, jeżeli:

²Parzystą (nieparzystą) permutacją zbioru elementów (liczb), to permutacja która realizujemy przy pomocy parzystej (nieparzystej) liczby przestawień.

- (a) Wszystkie elementy któregośkolwiek z wierszy, lub którejkolwiek z kolumn, są zerami.
- (b) Wyznacznik zawiera dwie identyczne kolumny lub dwa identyczne wiersze.
- (c) Wszystkie elementy jednego wiersza są wielokrotnościami odpowiednich elementów innego wiersza; to samo w przypadku kolumn.
- (d) Dany wiersz (kolumna) może być wyrażony jako liniowa kombinacja pozostałych wierszy (kolumn) wyznacznika.
3. Przemnożenie wszystkich elementów danego wiersza (danej kolumny) przez stały czynnik powoduje przemnożenie przez ten czynnik wartości wyznacznika.
4. Wartość wyznacznika nie zmienia się, gdy:
- (a) zamienić wiersze na kolumny, a kolumny na wiersze,
- (b) do któregośkolwiek wiersza (którejkolwiek kolumny) dodać kombinację liniową dowolnej liczby wierszy (kolumn)
5. Jeżeli każdy element pewnego wiersza (kolumny) może być zapisany w postaci sumy lub różnicy dwu (lub więcej) wielkości, to wyznacznik można zapisać jako sumę lub różnicę dwóch (lub więcej) wyznaczników. Na przykład dla najprostszego przypadku:

$$\begin{vmatrix} a_{11} \pm b_{11} & a_{12} \pm b_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \pm \begin{vmatrix} b_{11} & b_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

Wymienione powyżej własności można łatwo wykazać. Własność (1) wynika bezpośrednio z wzoru 3-34. Z kolei własność (2b) jest konsekwencją własności (1) – jeżeli przestawianie dwóch *identycznych* wierszy (kolumn) musi powodować zmianę znaku wartości wyznacznika, to wartość ta, taka sama w obu przypadkach, może tylko być zerem. Własność (2c) łatwo wykazać posługując się własnością (3), która z kolei – podobnie jak własność (2a) – wynika z rozwinięcia Laplace’a. Własność (4a) to konsekwencja wzoru 3-34. Własność (4b) to, na przykład:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \stackrel{?}{=} \begin{vmatrix} a_{11} + \alpha a_{12} & a_{12} & a_{13} \\ a_{21} + \alpha a_{22} & a_{22} & a_{23} \\ a_{31} + \alpha a_{32} & a_{32} & a_{33} \end{vmatrix}.$$

Prawy wyznacznik rozwijamy względem pierwszej kolumny, a następnie wyznaczniki drugiego stopnia, przemnożone oddzielnie przez α i przez a_{i1} , „związujemy z powrotem” do wyznaczników trzeciego stopnia (korzystamy *de facto* z własności 5) :

$$\begin{vmatrix} a_{11} + \alpha a_{12} & a_{12} & a_{13} \\ a_{21} + \alpha a_{22} & a_{22} & a_{23} \\ a_{31} + \alpha a_{32} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \alpha \begin{vmatrix} a_{12} & a_{12} & a_{13} \\ a_{22} & a_{22} & a_{23} \\ a_{32} & a_{32} & a_{33} \end{vmatrix}.$$

Drugi z wyznaczników po prawej stronie jest jednak równy zeru (dwie identyczne kolumny) – a więc własność (4b) została wykazana. Wreszcie, własność (5) wynika bezpośrednio z rozwinięcia Laplace’a. W oparciu o tę własność i o własność (2c) dowodzimy (bardzo istotnej!) własności (2d).

Warto zwrócić jeszcze uwagę na jeden fakt, związany z obliczeniami wyznaczników. Rozwijając wyznacznik d :

$$d = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{nn} \end{vmatrix} \quad (3-35)$$

względem j -ej kolumny otrzymujemy

$$d = a_{1j}M^{[1j]} + a_{2j}M^{[2j]} + \dots + a_{nj}M^{[nj]}, \quad (3-36)$$

gdzie $M^{[ij]}$, $i = 1, n$ są odpowiednimi dopełnieniami algebraicznymi elementów j -ej kolumny a_{ij} . Nawiązując do twierdzenia Cramera, zastąpmy elementy tej j -ej kolumny przez n dowolnych liczb b_1, \dots, b_n . Rozwinięcie

$$b_1M^{[1j]} + b_2M^{[2j]} + \dots + b_nM^{[nj]}, \quad (3-37)$$

to rozwinięcie wyznacznika d' , różniącego się od d tą jedną kolumną (dopełnienia algebraiczne elementów obu kolumn są identyczne):

$$d' = \begin{vmatrix} a_{11} & a_{12} & \dots & b_1 & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & b_2 & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & b_n & \dots & a_{nn} \end{vmatrix} \quad (3-38)$$

Co będzie jeżeli za b_1, \dots, b_n podstawić elementy k -tej kolumny: a_{1k}, \dots, a_{nk} , dla $k \neq j$? Taki wyznacznik będzie zawierał dwie identyczne kolumny, a więc będzie równy zero. Wykazaliśmy więc:

6. Suma iloczynów wszystkich elementów danej kolumny przez odpowiednie dopełnienia algebraiczne elementów innej kolumny jest równa zero.

(Analogicznie mamy dla wierszy.) Własność ta będzie pomocna w znalezieniu wzoru na elementy tzw. macierzy odwrotnej w rozdziale siódmym.

3.5 Układ równań jednorodnych

Układ n jednorodnych równań liniowych o n niewiadomych to :

$$\begin{aligned} a_{11}x_1 &+ a_{12}x_2 &+ \dots &+ a_{1n}x_n &= 0, \\ a_{21}x_1 &+ a_{22}x_2 &+ \dots &+ a_{2n}x_n &= 0, \\ \dots &+ \dots &+ \dots &+ \dots &= \dots, \\ a_{n-1,1}x_1 &+ a_{n-1,2}x_2 &+ \dots &+ a_{n-1,n}x_n &= 0, \\ a_nx_1 &+ a_nx_2 &+ \dots &+ a_nx_n &= 0. \end{aligned} \quad (3-39)$$

Jeżeli wyznacznik główny układu 3-39 jest różny od zera, to wzory Cramera dają natychmiast (mało interesujące) *rozwiązanie trywialne*:

$$x_1 = x_2 = \dots = x_{n-1} = x_n = 0. \quad (3-40)$$

(Występujące w licznikach wyznaczniki zawierają wszystkie jedną kolumnę samych zer – własność 2a.) Wniosek – jeżeli mamy mieć jakieś ciekawsze rozwiązanie, to znaczy z przynajmniej jednym $x_i, i = 1, \dots, n$ różnym od zera, *to wyznacznik główny układu 3-39 musi być równy zero*.

Każde z równań układu 3-39 można pomnożyć przez dowolną (różną od zera) liczbę – nie spowoduje to zmiany charakteru wyznacznika głównego (pozostanie równy zero, bądź różny od zera), a samo równanie też pozostanie spełnione, lub nie. Wynika stąd, że istotne są tylko *stosunki* niewiadomych, a nie same ich wartości – jedną z niewiadomych możemy przyjąć jako dowolną, różną od zera liczbę i operować ilorazami pozostałych niewiadomych i tej właśnie wybranej. Każde równanie dla $j = 1, \dots, n$

$$a_{j1}x_1 + \dots + a_{j,k-1}x_{k-1} + a_{jk}x_k + a_{j,k+1}x_{k+1} + \dots + a_{jn}x_n = 0; \quad (3-41)$$

jest równoważne wyrażeniu niewiadomej x_k jako

$$x_k = -(1/a_{jk})(a_{j1}x_1 + \dots + a_{j,k-1}x_{k-1} + a_{j,k+1}x_{k+1} + \dots + a_{jn}x_n). \quad (3-42)$$

A więc dowolna, wybrana niewiadoma może zostać wyrażona poprzez $n - 1$ pozostałych niewiadomych. Zobaczmy w kolejnych rozdziałach, że właśnie to jest przyczyną zerowania się wyznacznika głównego.

W fizyce często używamy pojęcia *stopni swobody*. Liczba stopni swobody to minimalna liczba współrzędnych potrzebnych do opisu stanu układu. Na przykład opis dwu poruszających się niezależnie w przestrzeni cząstek wymaga znajomości $3 + 3 = 6$ współrzędnych. W momencie kiedy cząstki te są połączone ze sobą (np. dwuatomowa cząsteczka gazu) wystarczy już 5 współrzędnych – mogą to być współrzędne środka masy cząsteczki i dwa kąty, określające orientację cząsteczki w przestrzeni. To przejście od sześciu do pięciu stopni swobody, związane jest z dodatkowym warunkiem, który muszą spełniać współrzędne obu cząstek (stała odległość). Formalny zapis tego warunku to *jedno* równanie

$$(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 = l^2 = \text{constans},$$

((x_1, x_2, x_3) to współrzędne jednej, a (y_1, y_2, y_3) – drugiej cząstki; l – ich odległość) i dlatego ($6 - 1 = 5$) liczba stopni swobody redukuje się o jeden. To właśnie to dodatkowe równanie pozwala wyliczyć szóstą

współrzędną, jeżeli znane jest już pozostałych 5. Podobnie znajomość $n - 1$ niewiadomych układu n równań układu jednorodnego określa – poprzez związki typu 3-42 – brakującą niewiadomą.

Jak określić te $n - 1$ niewiadomych? Rozwiązanie nietrywialne układu jednorodnego uzyskamy – w sposób najbardziej ogólny – wykorzystując wykazane powyżej jego własności. Jeżeli rozwiązanie ma być nietrywialne, to przynajmniej jedna z niewiadomych musi być różna od zera. Załóżmy, że jest nią x_n . Przyjmijmy ją jaką pewną dowolną, różną od zera liczbę: $x_n \equiv \alpha$, i podzielmy nią wszystkie równania układu 3-39. Dostaniemy układ

$$\begin{array}{rcccccc}
 a_{11}(x_1/x_n) & + & a_{12}(x_2/x_n) & + & \dots & + & a_{1,n-1}(x_{n-1}/x_n) & = & -a_{1n}, \\
 a_{21}(x_1/x_n) & + & a_{22}(x_2/x_n) & + & \dots & + & a_{2,n-1}(x_{n-1}/x_n) & = & -a_{2n}, \\
 \dots & + & \dots & + & \dots & + & \dots & = & \dots, \\
 a_{n-1,1}(x_1/x_n) & + & a_{n-1,2}(x_2/x_n) & + & \dots & + & a_{n-1,n-1}(x_{n-1}/x_n) & = & -a_{n-1,n}, \\
 a_n(x_1/x_n) & + & a_n(x_2/x_n) & + & \dots & + & a_{n,n-1}(x_{n-1}/x_n) & = & -a_{nn}.
 \end{array} \tag{3-43}$$

Stosunki x_k/x_n , $k = 1, \dots, n - 1$ to „nowe” niewiadome; liczby: $-a_{in}$; $i = 1, \dots, n$ to niejednorodności w powstałym układzie n równań na $n - 1$ niewiadomych. Ponieważ układ będzie oznaczonym tylko wtedy, kiedy liczba równań jest równa liczbie niewiadomych, jedno z równań musimy odrzucić. W dodatku, musi to być tak wybrane równanie, aby wyznacznik główny powstałego układu (stopnia $n - 1$) był różny od zera. Jeżeli uda nam się to osiągnąć, to rozwiązujemy ten układ korzystając np. ze wzorów Cramera (lub metody eliminacji Gaussa).

Do problemu układu równań jednorodnych będziemy jeszcze kilkakrotnie wracać w trakcie wykładu.

Rozdział 4

Algebra wektorów – dwa i trzy wymiary

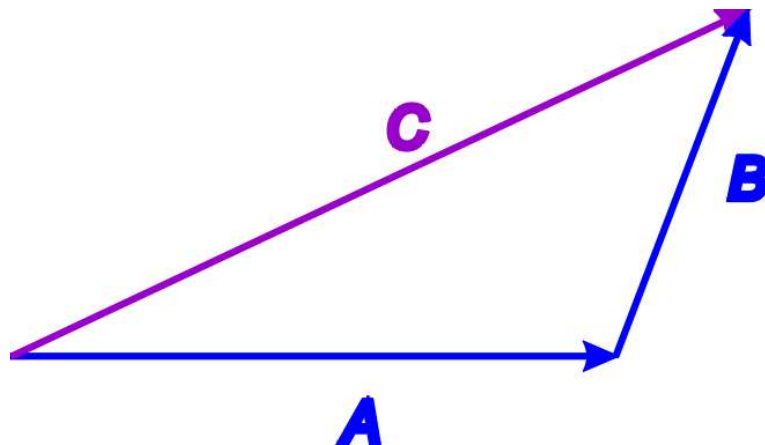
4.1 Podstawowe definicje; dodawanie wektorów

W opisie większości zjawisk fizycznych mamy do czynienia z dwoma typami wielkości. Jedne z nich, takie jak: masa, temperatura, czas, itp. określone są w pełni w momencie podania ich *wartości liczbowej*¹. Nazywamy je wielkościami skalarnymi, lub skalarami. Drugie, do których należą: położenie, prędkość, przyspieszenie, siła, pęd, itp. wymagają do pełnego określenia nie tylko wartości liczbowej, ale podania także związanego z nimi kierunku i zwrotu. Przez kierunek rozumiemy wyróżnioną w przestrzeni prostą, natomiast zwrot to wybór w alternatywie: „do góry” lub „na dół”, albo „w prawo” lub „w lewo”. W przyszłości mówiąc o kierunku będziemy mieli na myśli kierunek **wraz** z wybranym na nim zwrotem. Takie wielkości nazywamy wektorami i — w celu odróżnienia ich od skalarów — oznaczamy je wytłuszczonym drukiem (np. \mathbf{F}), albo przy pomocy umieszczonej nad nimi strzałki (\vec{F}).²

Potoczny sposób myślenia o wektorze polega na wyobrażeniu go sobie jako pewnej „strzałki”, której długość jest proporcjonalna do *wartości liczbowej* wektora (cały czas operujemy określonymi jednostkami!), a której kierunek jest kierunkiem wektora. Taki sposób reprezentacji wektora, w połączeniu z prostymi (i podlegającymi eksperymentalnej weryfikacji) rozważaniami „fizycznymi” (np. ze statyki) pozwala nam określić *prawo dodawania* dwóch wektorów

$$\mathbf{A} + \mathbf{B} = \mathbf{C} \quad (4-1)$$

jako operację, w której początek wektora \mathbf{B} zostaje przesunięty do końca (ostrza) wektora \mathbf{A} . Wypadkowy wektor (suma), to wektor \mathbf{C} – strzałka, łącząca początek wektora \mathbf{A} i koniec wektora \mathbf{B} (Rys. 4.1).



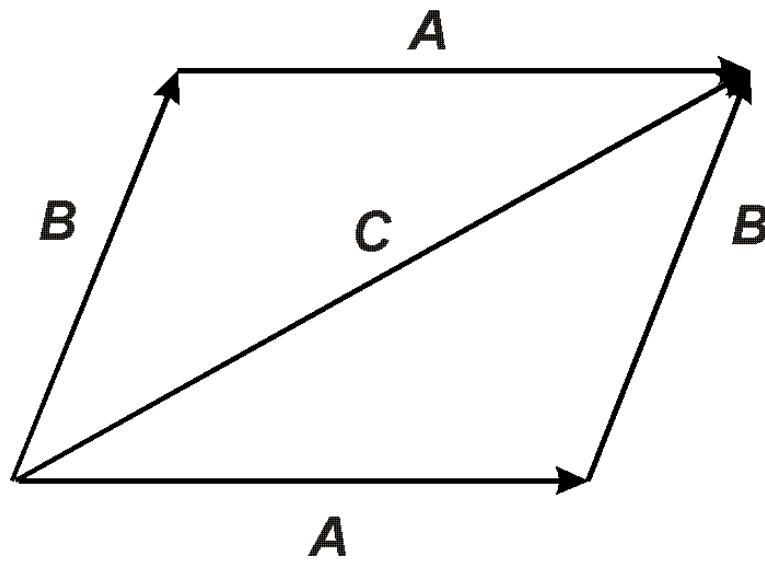
Rysunek 4.1: Dodawanie wektorów: $\mathbf{A} + \mathbf{B} = \mathbf{C}$ – reguła trójkąta.

Jeżeli trójkąt z Rys.4.1 uzupełnić do równoległoboku (Rys.4.2) to z uzyskanej konstrukcji wynika, że dodawanie wektorów jest przemienne:

$$\mathbf{A} + \mathbf{B} = \mathbf{C} = \mathbf{B} + \mathbf{A}, \quad (4-2)$$

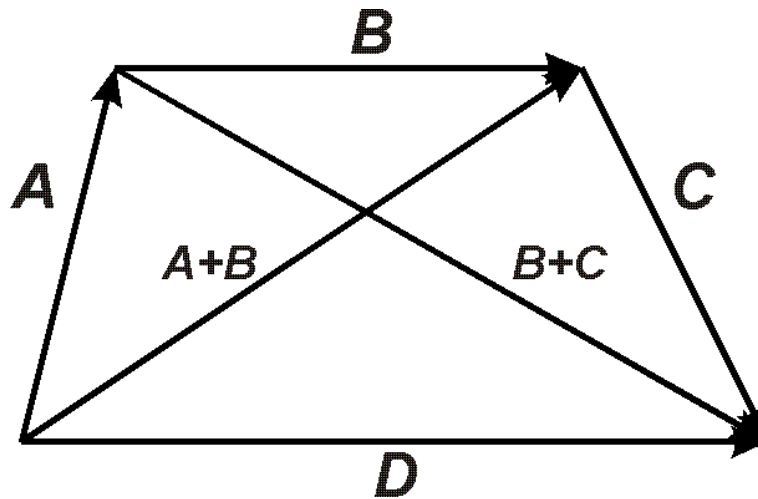
¹i – oczywiście – jednostki!

²Opis niektórych zjawisk fizycznych będzie wymagał wprowadzenia wielkości *tensorowych*. O tensorach będzie mowa w końcowych rozdziałach podręcznika.



Rysunek 4.2: Dodawanie wektorów: $A + B = C$ – reguła równoległoboku.

a rozważając nie trzy, ale cztery wektory (lub ich większą liczbę), łatwo zauważyć, że ma ono także własność



Rysunek 4.3: Dodawanie wektorów – zasada przemienności i łączności.

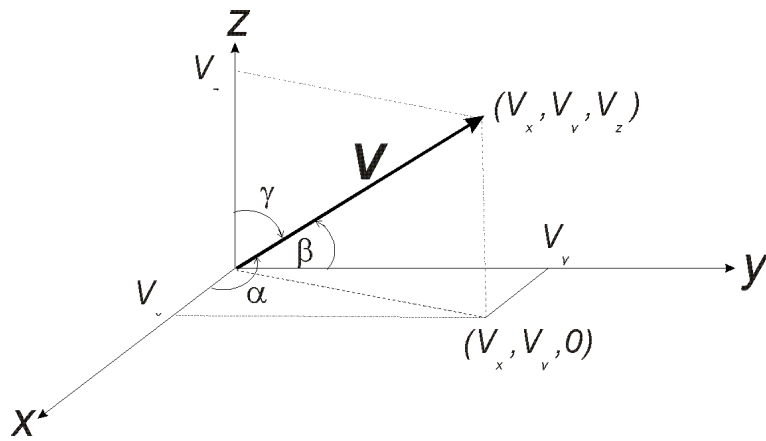
łączności (Rys. 4.3):

$$A + B + C = (A + B) + C = A + (B + C) = D. \quad (4-3)$$

Dodawanie wektorów przy zastosowaniu podanych powyżej reguł nie jest jednak, na dłuższą metę, wygodne. Między innymi dlatego wprowadza się inny sposób mówienia o wektorach, polegający na wyrażaniu wektorów poprzez ich *współrzędne* w określonym *układzie odniesienia*. Wielkość wektorową, którą wyobrażamy sobie jako „skierowaną strzałkę” poddajemy rozkładowi na współrzędne, przy czym rozkład ten będzie zależał od wyboru układu. Ponieważ przeważnie mamy do czynienia z przestrzenią trójwymiarową, to najprostszym, a także – w wielu przypadkach – najpraktyczniejszym układem współrzędnych jest układ trzech osi liczbowych, przecinających się pod kątami prostymi we wspólnym zerze (por. Rys. 4.4). Aby określić w takim układzie współrzędne wektora V umieszczamy jego początek w początku układu współrzędnych; koniec wektora określa nam pewien punkt w przestrzeni, którego współrzędne to trzy liczby: wartości rzutów wektora V na trzy osie naszego kartezjańskiego układu współrzędnych $0x$, $0y$ i $0z$ – V_x , V_y i V_z . (Jak wiadomo położenie każdego punktu w przestrzeni 3-wymiarowej możemy określić podając 3 liczby – 3 współrzędne). Jak wynika z konstrukcji zilustrowanej na Rys.4.4:

$$V \equiv (V_x, V_y, V_z). \quad (4-4)$$

Trzy wielkości (V_x, V_y, V_z) mogą być wyrażone przez długość wektora V , którą oznaczamy będziemy jako $|V|$ lub po prostu V i trzy *kosinusy kierunkowe wektora V* – kosinusy kątów które tworzy on z trzema osiami



Rysunek 4.4: Rozkład wektora \mathbf{V} na współrzędne:

$$\alpha = \angle(0x, \mathbf{V}), \quad \beta = \angle(0y, \mathbf{V}), \quad \gamma = \angle(0z, \mathbf{V}).$$

układu:

$$V_x = V \cos \alpha = V \cos \angle(0x, \mathbf{V}), \quad (4-5)$$

$$V_y = V \cos \beta = V \cos \angle(0y, \mathbf{V}), \quad (4-6)$$

$$V_z = V \cos \gamma = V \cos \angle(0z, \mathbf{V}). \quad (4-7)$$

Jeżeli każdej osi układu przyporządkujemy jednostkowy wektor, *wersor*, o zwrocie zgodnym z dodatnim zwrotem danej osi:

$$\text{oś } 0x \stackrel{\text{def}}{=} \hat{\mathbf{x}}, \quad \text{oś } 0y \stackrel{\text{def}}{=} \hat{\mathbf{y}}, \quad \text{oś } 0z \stackrel{\text{def}}{=} \hat{\mathbf{z}}, \quad (4-8)$$

to iloczyn danego wersora przez odpowiadającą mu współrzędną jest wektorem leżącym wzdłuż odpowiadającej wersorowi osi – wektor taki stanowi odpowiednią *składową wektora* \mathbf{V} :

$$\begin{aligned}\hat{\mathbf{x}}V_x &= \mathbf{V}_x && \text{składowa } x\text{-owa} \\ \hat{\mathbf{y}}V_y &= \mathbf{V}_y && \text{składowa } y\text{-owa} \\ \hat{\mathbf{z}}V_z &= \mathbf{V}_z && \text{składowa } z\text{-owa.}\end{aligned}$$

Sam wektor \mathbf{V} można zapisać jako wektorową sumę jego składowych: x -owej – \mathbf{V}_x , y -owej – \mathbf{V}_y i z -owej – \mathbf{V}_z :

$$\mathbf{V} = \mathbf{V}_x + \mathbf{V}_y + \mathbf{V}_z = \hat{\mathbf{x}}V_x + \hat{\mathbf{y}}V_y + \hat{\mathbf{z}}V_z. \quad (4-9)$$

Trzy współrzędne wektora spełniają oczywistą zależność

$$V_x^2 + V_y^2 + V_z^2 = V^2, \quad (4-10)$$

albo

$$V \stackrel{\text{def}}{=} |\mathbf{V}| = \sqrt{V_x^2 + V_y^2 + V_z^2} = V^2; \quad (4-11)$$

długość wektora jest równa pierwiastkowi kwadratowemu z sumy kwadratów jego współrzędnych. Jest to trójwymiarowy odpowiednik twierdzenia Pitagorasa – kwadrat głównej przekątnej prostopadłościanu zbudowanego na trzech współrzędnych jest równy sumie kwadratów jego trzech boków.

Podkreślmy już w tym miejscu, że wybór trzech prostopadłych osi liczbowych jako naszego *kartezjańskiego* układu współrzędnych jest tylko jednym z wielu możliwych. Nie jest on obowiązkowy – i tak na płaszczyźnie możemy posługiwać się na przykład układem współrzędnych biegunowych, o którym będzie mowa w kolejnym podrozdziale. Z kolei, wyboru takiego a nie innego układu współrzędnych kartezjańskich dokonujemy poprzez określenie jednostkowych wektorów (wersorów): $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ i $\hat{\mathbf{z}}$. To właśnie te trzy wektory jednoznacznie określają naszą *przestrzeń wektorową* – każdy element ten przestrzeni, czyli wektor, może zostać zapisany jako *liniowa kombinacja* trzech (w tym wypadku prostopadłych) wersorów (4-9). Trójkę taką nazywamy *bazą przestrzeni wektorowej*, a o jej trzech składowych mówimy, że „napinają przestrzeń (3-wymiarową) wektorową”. Współczynniki liczbowe występujące w tej liniowej kombinacji to współrzędne wektora: V_x , V_y i V_z . Ale — podkreślmy to raz jeszcze — każdy wektor, który reprezentuje dla nas zwykle pewną określoną wielkość fizyczną, możemy rozkładać na współrzędne w różnych układach współrzędnych, określanymi przez różne trójki wersorów. Co więcej — te trójki nie muszą być (choć w znakomitej większości przypadków są) wektorami wzajemnie prostopadłymi³. Wybrana trójka wersorów, określająca nasz układ współrzędnych, nie musi zajmować określonego położenia w przestrzeni – możemy ją na przykład poddać obrotowi (rotacji) lub przesunięciu (translacji). Zwykle opis zjawiska fizycznego nie zależy od takiego lub innego usytuowania (orientacji) układu – mówimy, że przestrzeń jest *fizycznie jednorodna* (czyli, prostymi słowy, „taka sama”), a także *izotropowa* (czyli: obserwowane w niej zjawiska nie zależą od kierunku obserwacji; to drugie założenie bywa jednak czasami to naruszane). Każda z takich operacji: wybór nowego układu lub transformacja (rotacja i translacja) starego będzie zmieniała współrzędne wektora (odniesione do „nowego” układu), natomiast sam wektor będzie cały czas „tym samym” wektorem, reprezentującym tę samą wielkość fizyczną. Wybór konkretnej bazy (typu układu) i ewentualne jej transformacje (obróć, przesunięcie) dyktowane są zawsze wygodą. Pewne symetrie, występujące przy opisie zjawisk, uwydatniają się lepiej przy użyciu takiego a nie innego układu współrzędnych. Zarówno podstawowe, alternatywne układy współrzędnych, jak i zagadnienie transformacji pomiędzy dwoma różnymi układami i transformacji (obrotu) określonego układu będą omawiane w dalszej części podręcznika.

W momencie wprowadzenia reprezentacji wektorów przy użyciu ich współrzędnych (oczywiście wyrażonych w tym samym układzie), operacji dodawania (odejmowania) wektorów możemy dokonywać przy pomocy rachunku, bez uciekania się do pomocy konstrukcji geometrycznych. Jeżeli bowiem dwa wektory \mathbf{A} i \mathbf{B} są określone, zgodnie z 4-9, jako

$$\begin{aligned}\mathbf{A} &= \hat{\mathbf{x}}A_x + \hat{\mathbf{y}}A_y + \hat{\mathbf{z}}A_z \\ \mathbf{B} &= \hat{\mathbf{x}}B_x + \hat{\mathbf{y}}B_y + \hat{\mathbf{z}}B_z,\end{aligned}$$

to ich suma (różnica) będzie równa

$$\begin{aligned}\mathbf{A} + \mathbf{B} &= \hat{\mathbf{x}}(A_x + B_x) + \hat{\mathbf{y}}(A_y + B_y) + \hat{\mathbf{z}}(A_z + B_z), \\ \mathbf{A} - \mathbf{B} &= \hat{\mathbf{x}}(A_x - B_x) + \hat{\mathbf{y}}(A_y - B_y) + \hat{\mathbf{z}}(A_z - B_z).\end{aligned} \quad (4-12)$$

³W przypadku przestrzeni trójwymiarowej trójka wersorów bazy musi być trójką wektorów nie komplanarnych, czyli nie leżących w jednej płaszczyźnie.

Pozostaje kwestią otwartą, czy taki sposób dodawania wektorów jest rzeczywiście „prostszy” od zilustrowanych na rysunkach konstrukcji trójkątów lub równoległoboków. Dla początkującego studenta fizyki chyba bardziej naturalne wydaje się właśnie „składanie” wektorów (sił, prędkości, itp.) według „reguły trójkąta”, lub „reguły równoległoboku”. Rachunek opisany równaniem 4-12 wymaga bowiem wyobrażenia sobie każdego ze wyrazów sumy (różnicy) jako sumy (różnicy) odpowiednich współrzędnych. Każdy składnik (wektor) musi więc zostać rozłożony na współrzędne – trójkę liczb. Niewątpliwie bardziej naturalne jest myślenie o wektorze (sile, prędkości) jako o pojedynczej „strzałce” niż o trójce liczb. Jeżeli jednak przyswoić sobie ten nieco bardziej abstrakcyjny sposób postrzegania wielkości wektorowej, to niewątpliwie operacje ich składania ulegają znacznemu uproszczeniu.

Operacje „składania” (dodawania lub odejmowania) wektorów nie są jedynymi operacjami, których możemy dokonywać na wielkościach wektorowych. Wektor można też mnożyć przez liczbę (skalar), co sprowadza się do przemnożenia każdej współrzędnej wektora przez tę liczbę:

$$\alpha \mathbf{V} = \alpha \mathbf{V}_x + \alpha \mathbf{V}_y + \alpha \mathbf{V}_z = \hat{x}\alpha V_x + \hat{y}\alpha V_y + \hat{z}\alpha V_z, \quad (4-13)$$

gdzie α jest wielkością skalarną. Jeżeli $\alpha = 0$, to wynik mnożenia będzie *wektorem zerowym*, o długości równej zeru. Wektory możemy także mnożyć „przez siebie”, przy czym możemy mieć do czynienia z trzema różnymi typami mnożenia: iloczynem skalarnym, którego wynikiem jest liczba (skalar), iloczynem wektorowym (wynik – wektor) oraz mnożeniem tensorowym, w wyniku którego dostajemy tensor drugiego rzędu. W następnych podrozdziałach zajmiemy się dwoma pierwszymi typami mnożenia – mnożeniem skalarnym i wektorowym, mającymi zresztą podstawowe znaczenie dla fizyki – i ich kombinacjami, pozostawiając zagadnienie mnożenia „tensorowego” do rozdziału traktującego o tensorach ⁴.

4.2 Iloczyn skalarny

Przypomnijmy raz jeszcze równania, określające współrzędne wektora \mathbf{V} [(4-5) i nast.]:

$$V_x = V \cos \angle(0x, \mathbf{V}), \quad V_y = V \cos \angle(0y, \mathbf{V}), \quad V_z = V \cos \angle(0z, \mathbf{V}).$$

Jak wynika z powyższego równania (i z elementarnej trygonometrii), trzy liczby V_i ; $i = x, y, z$ otrzymujemy rzutując wektor \mathbf{V} na każdą z trzech osi: $0x$, $0y$ i $0z$ – por. Rys.4.4. Wektor rzutujemy na każdy z trzech kierunków, określonych przez (jednostkowe) wektory \hat{x} , \hat{y} i \hat{z} . Taką operację *rzutowania wektora na określony kierunek*, a więc prowadzącą do określenia wielkości skalarnej – współrzędnej wektora wzdłuż tego kierunku, możemy zdefiniować jako *mnożenie skalarne* rzutowanego wektora przez reprezentujący dany kierunek wersor i oznaczyć – zgodnie z powszechnie stosowaną konwencją przy pomocy symbolu kropki:

$$\left. \begin{aligned} V_x &= V \cos \angle(0x, \mathbf{V}) \equiv \mathbf{V} \cdot \hat{x} \\ V_y &= V \cos \angle(0y, \mathbf{V}) \equiv \mathbf{V} \cdot \hat{y} \\ V_z &= V \cos \angle(0z, \mathbf{V}) \equiv \mathbf{V} \cdot \hat{z}. \end{aligned} \right\} \quad (4-14)$$

Z powyższych wzorów wynika więc, że *iloczyn skalarny* dwóch wektorów \mathbf{A} i \mathbf{B} może być zdefiniowany jako

$$\mathbf{A} \cdot \mathbf{B} = AB \cos \angle(\mathbf{A}, \mathbf{B}). \quad (4-15)$$

(Występujące we wzorach 4-14 wersory osi mają długość jednostkową, gdyby je wydłużyć lub skrócić k -krotnie, to ich nowa długość – k – pojawi się także jako mnożnik długości wektora V).

Powiedzieliśmy, że iloczyn skalarny ma podstawowe zastosowanie w fizyce. Na przykład element pracy dL siły \mathbf{F} wykonanej na drodze $d\mathbf{s}$ definiujemy jako iloczyn skalarny tych dwóch wektorów: $dL = \mathbf{F} \cdot d\mathbf{s}$. Zauważmy, że odpowiada to rzutowaniu siły \mathbf{F} na kierunek przesunięcia $d\mathbf{s}$ i przemnożeniu wartości tego rzutu ($F \cos \angle(\mathbf{F}, d\mathbf{s})$) przez długość przesunięcia (ds). Tak więc, bez wprowadzania pojęcia iloczynu skalarnego, pracę musimy definiować jako wielkość skalarną równą iloczynowi przesunięcia (wektor) i składowej siły (wektor) *mającej z nim wspólny kierunek*. Warto o tym pamiętać, gdyż pomaga nam to postrzegać iloczyn skalarny dwóch wektorów jako miarę nie tylko ich wielkości (długości) ale współukierunkowania.

⁴Uwaga : Operacja dzielenie wektorów nie istnieje!

Iloczyn skalarny, jak wynika z wzoru 4-15, w którym występują „zwykłe” operacje mnożenia liczb (wielkości skalarnych), będzie spełniał prawa przemienności i rozdzielności:

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A}, \quad (4-16)$$

$$\mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C}, \quad (4-17)$$

a także będzie operacją liniową, a więc dla dowolnych liczb α i β

$$\mathbf{C} \cdot (\alpha \mathbf{A} + \beta \mathbf{B}) = \mathbf{C} \cdot \alpha \mathbf{A} + \mathbf{C} \cdot \beta \mathbf{B} = \alpha \mathbf{C} \cdot \mathbf{A} + \beta \mathbf{C} \cdot \mathbf{B}. \quad (4-18)$$

Te oczywiste własności pozwalają nam określić reguły wyrażania wartości iloczynu skalarnego dwóch wektorów \mathbf{A} i \mathbf{B} poprzez ich współrzędne. Mamy bowiem

$$\begin{aligned} \mathbf{A} \cdot \mathbf{B} &= \mathbf{A} \cdot (\hat{x}B_x + \hat{y}B_y + \hat{z}B_z) \\ &= B_x \mathbf{A} \cdot \hat{x} + B_y \mathbf{A} \cdot \hat{y} + B_z \mathbf{A} \cdot \hat{z} \\ &= B_x A_x + B_y A_y + B_z A_z = A_x B_x + A_y B_y + A_z B_z. \end{aligned}$$

Tak więc iloczyn skalarnych dwóch wektorów to suma iloczynów ich analogicznych współrzędnych:

$$\mathbf{A} \cdot \mathbf{B} = A_x B_x + A_y B_y + A_z B_z \equiv \sum_{i=x,y,z} A_i B_i. \quad (4-19)$$

Kładąc powyższym wzorze $\mathbf{A} = \mathbf{B}$ mamy

$$\mathbf{A} \cdot \mathbf{A} = A_x^2 + A_y^2 + A_z^2 \equiv \sum_{i=x,y,z} A_i^2 = A^2 \quad (4-20)$$

– skalarny iloczyn wektora przez siebie samego jest równy kwadratowi jego długości (por. 4-11).

Dla czytelnika, który pierwszy raz spotyka się ze znakiem sumy (grecka duża litera sigma – \sum) wyjaśnienie : znak \sum oznacza, że wielkość występująca po nim, przedstawia sumę składników, które powstają przez przyporządkowanie *wskaźnikowi* (na przykład – i) wszystkich możliwych wartości – literowych (jak tutaj : $i = x, y$ i z), lub liczbowych (np. $i = 1, 2$ i 3).

Iloczyn skalarny dwóch prostopadłych wektorów jest równy zero ($\cos \pi/2 = 0$). Iloczyn skalarny może być nie tylko miarą długości danego wektora (równ. 4-20), ale także wskaźnikiem prostopadłości, lub *ortogonalności* dwóch wektorów. Termin ortogonalny, będący greckim tłumaczeniem terminu prostopadły, jest pojęciem szerszym – nie jest bowiem ograniczony do „zwykłej” trójwymiarowej przestrzeni (lub dwuwymiarowej płaszczyzny), ale może się odnosić do przestrzeni wektorowych o dowolnej (także nieskończonej) liczbie wymiarów. W szczególności jednostkowe wersory osi układu kartezjańskiego są też ortogonalne, a nawet *ortonormalne*: iloczyny skalarne różnych, prostopadłych do siebie, wersorów są równe zero, natomiast ich długości (równe pierwiastkowi z iloczynu wersora przez siebie samego) są równe jedności. Właśnie ze względu na to *unormowanie* długości wersorów do jedności mówimy o własności *ortonormalności*. Często wygodnie jest zastąpić oznaczanie wersorów osi układu kartezjańskiego \hat{x} , \hat{y} i \hat{z} ⁵ przez trzy indeksowane (opatrzone wskaźnikami) literki, np. \mathbf{e}_k (zazwyczaj rezygnujemy z symbolu $\hat{}$ – literka \mathbf{e} jest utożsamiana z wektorem jednostkowym), gdzie wskaźnik $k = 1, 2$ i 3 , przy czym $\mathbf{e}_1 = \hat{x}$, $\mathbf{e}_2 = \hat{y}$ i $\mathbf{e}_3 = \hat{z}$. Wówczas relację ortonormalności wersorów możemy zapisać w zwartej postaci

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j. \end{cases} \quad (4-21)$$

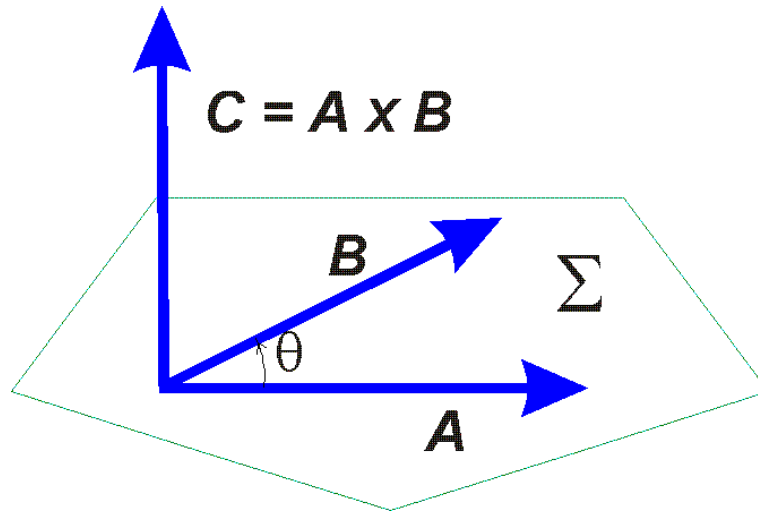
Użyty i jednocześnie zdefiniowany w powyższym wzorze symbol to tzw. delta Kroneckera – równa zero przy różnych i i j , a jedności przy $i = j$. Zauważmy, że wskaźniki i i j w *zasadzie* zmieniają się od 1 do 3, choć – w przypadku przestrzeni n -wymiarowych mogą przebiegać wartości od 1 do n .

Na zakończenie tego krótkiego przypomnienia o własnościach iloczynu skalarnego zwróćmy uwagę, że iloczyn skalarny jest – jak sama jego nazwa wskazuje – liczbą, a więc nie powinien zależeć od wyboru układu współrzędnych. Dwa wektory mogą mieć, w różnych układach współrzędnych, różne „reprezentacje w postaci współrzędnych”, ale ich iloczyn skalarny, obliczony w każdym układzie, musi być taki sam. Formalne uzasadnienie tego wniosku nastąpi w jednym z późniejszych rozdziałów. Tutaj możemy tylko zauważyć, że

⁵Innym, często stosowanym sposobem oznaczania wersorów osi jest użycie trójki \hat{i} , \hat{j} i \hat{k} .

własności prostopadłości (ortogonalności) i współliniowości dwóch wektorów też nie zależą od wyboru układu współrzędnych – w dwóch więc skrajnych przypadkach ich iloczyn skalarny (0 lub iloczyn długości) *musi* być niezależny od wyboru układu. Aby rozszerzyć tę *niezmienniczość* iloczynu skalarnego względem transformacji układu współrzędnych wystarczy pomyśleć, że praca potrzebna do przebycia określonej drogi przy działaniu określonej siły też nie może być uzależniona od wyboru układu (np. od orientacji osi w przestrzeni), w którym przyjdzie nam ją obliczać.

4.3 Iloczyn wektorowy



Rysunek 4.5: Iloczyn wektorowy.

Rozważmy – Rys.4.5 – dwa wektory \mathbf{A} i \mathbf{B} , wychodzące z jednego punktu 0 i tworzące między sobą kąt θ , przy czym $0 < \theta < \pi$ (wektory nie są współliniowe, a kąt między nimi jest mniejszy od kąta półpełnego). Przyjmujemy konwencję, która mówi, że kąt „mierzymy” przechodząc od wektora \mathbf{A} do wektora \mathbf{B} , pamiętając o regule, że obrót przeciwny do ruchu wskazówek zegara odpowiada dodatnim przyrostom kąta.

Takie dwa, nie współliniowe wektory wyznaczają płaszczyznę Σ . W punkcie 0 umieścimy wektor \mathbf{C} , prostopadły do Σ (o zwrocie \mathbf{C} powiemy za chwilę). Ponieważ \mathbf{C} jest prostopadły do obu wektorów \mathbf{A} i \mathbf{B} odpowiednie iloczyny skalarne są równe zero:

$$\begin{aligned}\mathbf{A} \cdot \mathbf{C} &= A_x C_x + A_y C_y + A_z C_z = 0 \\ \mathbf{B} \cdot \mathbf{C} &= B_x C_x + B_y C_y + B_z C_z = 0.\end{aligned}$$

Jest to (por. rozdział poprzedni) układ dwóch równań z trzema niewiadomymi: $x_1 \equiv C_x, x_2 \equiv C_y, x_3 \equiv C_z$. Aby móc z niego utworzyć oznaczony układ równań niejednorodnych podzielmy go stronami przez C_z .⁶ Nowy układ to

$$\left. \begin{aligned}A_x \frac{C_x}{C_z} + A_y \frac{C_y}{C_z} &= -A_z, \\ B_x \frac{C_x}{C_z} + B_y \frac{C_y}{C_z} &= -B_z.\end{aligned} \right\} \quad (4-22)$$

Jego rozwiązanie, uzyskane metodą wyznaczników, to:

$$\frac{C_x}{C_z} = \frac{\begin{vmatrix} -A_z & A_y \\ -B_z & B_y \end{vmatrix}}{\begin{vmatrix} A_x & A_y \\ B_x & B_y \end{vmatrix}}, \quad \frac{C_y}{C_z} = \frac{\begin{vmatrix} A_x & -A_z \\ B_x & -B_z \end{vmatrix}}{\begin{vmatrix} A_x & A_y \\ B_x & B_y \end{vmatrix}}. \quad (4-23)$$

⁶Zakładamy, że $C_z \neq 0$. Zaraz się okaże, że bez tego założenia nasza szanse na sukces są równe zero.

Z powyższych równań wynika, że współrzędne wektora \mathbf{C} są powiązane ze współrzędnymi pozostałej dwójki wektorów wzorami

$$\left. \begin{aligned} C_x &= k(A_y B_z - A_z B_y) \\ C_y &= k(A_z B_x - A_x B_z) \\ C_z &= k(A_x B_y - A_y B_x). \end{aligned} \right\} \quad (4-24)$$

Występująca w powyższych wzorach wielkość k jest dowolną stałą. Połóżmy $k = 1$. (Nasze rozważania nie tracą ze swej ogólności.) Wówczas kwadrat długości wektora \mathbf{C} będzie równy (po elementarnych przekształceniach wzorów 4-24):

$$C^2 = C_x^2 + C_y^2 + C_z^2 = (A_x^2 + A_y^2 + A_z^2)(B_x^2 + B_y^2 + B_z^2) - (A_x B_x + A_y B_y + A_z B_z)^2. \quad (4-25)$$

W pierwszym składniku prawej strony 4-25 rozpoznajemy iloczyn kwadratów długości wektorów \mathbf{A} i \mathbf{B} , w drugim – kwadrat ich iloczynu skalarnego. W takim razie

$$C^2 = A^2 B^2 - (\mathbf{A} \cdot \mathbf{B})^2 = A^2 B^2 - A^2 B^2 \cos^2 \theta = A^2 B^2 \sin^2 \theta. \quad (4-26)$$

Długość wektora \mathbf{C} jest oznaczona jednoznacznie przez wektory \mathbf{A} i \mathbf{B} , podobnie jak jego kierunek (prostopadły do Σ , płaszczyzny utworzonej przez te dwa wektory.) Pozostaje kwestia zwrotu: przyjmujemy „konwencję śruby prawej” – taka śruba, obracając się od wektora \mathbf{A} do wektora \mathbf{B} wysuwa się „ku górze”. Wektor \mathbf{C} jest więc jednoznacznie określony przez \mathbf{A} i \mathbf{B} ; nazywamy go *iloczynem wektorowym* tych wektorów i zapisujemy:

$$\mathbf{C} = \mathbf{A} \times \mathbf{B}. \quad (4-27)$$

Długość \mathbf{C} , jak wynika z 4-26 jest równa $C = AB \sin \theta$. Wynika stąd, że iloczyn dwóch wektorów współliniowych, lub równoległych, jest równy zeru. Możemy też rozszerzyć określenie iloczynu na przypadek $\pi < \theta < 2\pi$. Ponieważ w tym przedziale kątów sinus przybiera wartości ujemne, wektor \mathbf{C} będzie zwrócony „w dół”.⁷ W końcu, ponieważ zmianie kierunku przyrostu kąta odpowiada zmiana jego znaku, a $\sin(-\theta) = -\sin \theta$ to *iloczyn wektorowy nie jest przemienne*:

$$\mathbf{A} \times \mathbf{B} = -\mathbf{B} \times \mathbf{A}. \quad (4-28)$$

Z powyższych rozważań wynika, że dla wersorów osi układu kartezjańskiego mamy:

$$\hat{x} \times \hat{x} = \hat{y} \times \hat{y} = \hat{z} \times \hat{z} = 0, \quad (4-29)$$

natomiast

$$\hat{x} \times \hat{y} = \hat{z}, \quad \hat{y} \times \hat{z} = \hat{x}, \quad \hat{z} \times \hat{x} = \hat{y}, \quad (4-30)$$

a także

$$\hat{y} \times \hat{x} = -\hat{z}, \quad \hat{z} \times \hat{y} = -\hat{x}, \quad \hat{x} \times \hat{z} = -\hat{y}. \quad (4-31)$$

Wzory 4-24 określają (przy $k = 1$) współrzędne wektora \mathbf{C} , a więc:

$$\mathbf{C} = \mathbf{A} \times \mathbf{B} = \hat{x}(A_y B_z - A_z B_y) + \hat{y}(A_z B_x - A_x B_z) + \hat{z}(A_x B_y - A_y B_x). \quad (4-32)$$

Powyższe wyrażenie zapisuje się często przy pomocy wyznacznika:

$$\mathbf{C} = \mathbf{A} \times \mathbf{B} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix}. \quad (4-33)$$

(Wzór 4-32 uzyskujemy rozwijając wyznacznik względem pierwszego wiersza.)

Wyznacznik 4-33 ze względu na symetrię jest łatwy do zapamiętania. Warto też zapamiętać prostą regułę, przy pomocy której możemy zapisywać trzy współrzędne wektora \mathbf{C} jako

$$C_i = A_j B_k - A_k B_j, \quad i, j, k = x, \text{ lub } y, \text{ lub } z - \text{wszystkie wskaźniki różne}, \quad (4-34)$$

⁷To samo otrzymamy, przyjmując konwencję, że wartość iloczynu wektorowego wektorów \mathbf{A} i \mathbf{B} jest obliczana przy założeniu, że obrót od wektora \mathbf{A} (pierwszy czynnik) do wektora \mathbf{B} (drugi czynnik) odbywa się po „krótszej drodze kątowej” ($\theta < \pi$). Obrótom „w prawo” odpowiadają ujemne kąty!

przy czym znak dodatni (pierwszy składnik po prawej stronie równania) odpowiada *parzystym permutacjom* wskaźników (x, y, z) , a więc sekwencjom: $(x, y, z), (y, z, x)$ i (z, x, y) , a ujemny (drugi składnik po prawej stronie równania) odpowiada *nieparzystym permutacjom* wskaźników (x, y, z) , a więc sekwencjom: $(x, z, y), (y, x, z)$ i (z, y, x) . (Permutacje parzyste określa się też mianem przestawień *cyklicznych*.)

Iloczyn wektorowy ma szerokie zastosowanie w fizyce, a także posiada ważną interpretację geometryczną. Z zastosowań fizycznych należałoby wymienić w pierwszym rzędzie wielkości pojawiające się przy opisie ruchu obrotowego, a więc moment pędu \mathbf{L} (kręt) – iloczyn wektorowy wektora położenia \mathbf{r} i wektora pędu ciała \mathbf{p} :

$$\mathbf{L} = \mathbf{r} \times \mathbf{p},$$

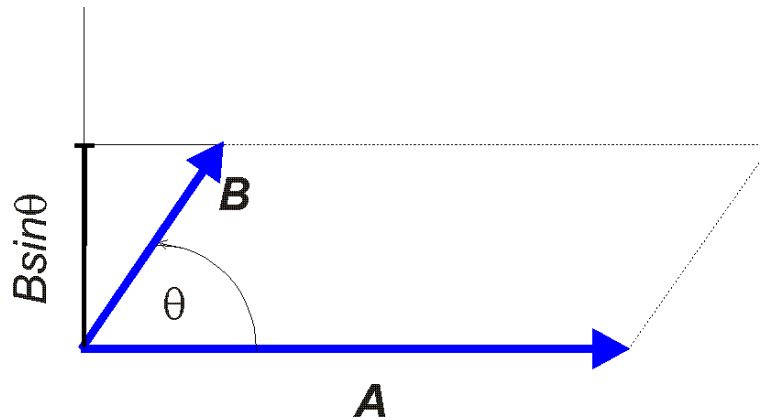
\mathbf{M} – moment siły \mathbf{F}

$$\mathbf{M} = \mathbf{r} \times \mathbf{F},$$

a także prędkość kątową $\boldsymbol{\omega}$, występującą w związku $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$, gdzie \mathbf{v} jest prędkością liniową ciała. Bardziej zaawansowana fizyka, w której występują iloczyny wektorowe, to np. opis siły działającej na ładunek q , poruszający się z prędkością \mathbf{v} w polu magnetycznym, opisanym wektorem indukcji magnetycznej \mathbf{B} (składowa „magnetyczna” siły Lorentza):

$$\mathbf{F} = q\mathbf{v} \times \mathbf{B},$$

a także wszystkie równania, w których występują operatory *rotacji* pola wektorowego. Interpretacja geometryczna



Rysunek 4.6: Interpretacja geometryczna iloczynu wektorowego – pole równoległoboku.

tryczna zilustrowana jest na Rys. 4.6. Pole równoległoboku zbudowanego na wektorach \mathbf{A} i \mathbf{B} jest równe $|\mathbf{A} \times \mathbf{B}| = AB \sin \theta$. Wynika stąd, że długość wektora (będącego iloczynem wektorowym) może reprezentować pewną powierzchnię. W przypadku elementu powierzchni bardzo (nieskończenie) małego kierunku wektora jest prostopadły do tego elementu⁸, tak więc pojawia się możliwość wektorowego traktowania elementów powierzchni. Będzie to miało znakomite znaczenie w opisie pewnych sytuacji i wielkości fizycznych (np. prawo Gaussa w elektrostatyce, pojęcie strumienia wektora).

(W przypadku $\theta > \pi$ miarą powierzchni będzie oczywiście wartość bezwzględna iloczynu $|\mathbf{A} \times \mathbf{B}|$, albo – co jest równoznaczne – długość wektora $\mathbf{B} \times \mathbf{A}$.)

Iloczyn wektorowy spełnia oczywiste prawa rozdzielności względem dodawania i liniowości:

$$(\mathbf{A} + \mathbf{B}) \times \mathbf{C} = \mathbf{A} \times \mathbf{C} + \mathbf{B} \times \mathbf{C}, \quad (4-35)$$

$$\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = \mathbf{A} \times \mathbf{B} + \mathbf{A} \times \mathbf{C}, \quad (4-36)$$

$$\mathbf{A} \times (\alpha \mathbf{B}) = \alpha \mathbf{A} \times \mathbf{B}, \quad (4-37)$$

gdzie α jest dowolną liczbą. Na koniec jeszcze jedna uwaga. Jak wskazuje sama nazwa, także sposób „konstrukcji”, iloczyn wektorowy *powinien* być wielkością wektorową. W fizyce, mówiąc o momencie pędu, lub momencie siły nie mamy wątpliwości, że są to wielkości wektorowe. Z formalnego punktu widzenia sprawa jest jednak nieco bardziej skomplikowana. Jak zobaczymy to w dalszej części wykładu, poświęconej tensorom, iloczyn wektorowy jest formalnie *antysymetrycznym tensorem drugiego rzędu*. Jest właściwie kwestią

⁸Znowu pozostaje kwestia zwrotu. Można też go jednoznacznie określić, poprzez związanie go z kierunkiem obiegu elementu powierzchni – stosując regułę śruby prawej.

„przypadku”, że w „zwykłej”, trójwymiarowej przestrzeni taka wielkość tensorowa jest opisana przy pomocy trójki współrzędnych – stąd możliwość traktowania jej jako wielkości wektorowej. Oczywiście fakt, że dana wielkość może być reprezentowana przy pomocy trójki liczb to jeszcze mało – należałoby sprawdzić, że ta trójka przy zmianie układu odniesienia podlega transformacji identycznej jak współrzędne wektora. Można pokazać, że przy transformacji obrotu i translacji układu tak jest. Do spraw tych powrócimy jeszcze w dalszej części wykładu.

4.4 Iloczyny trzech wektorów

Ponieważ wprowadziliśmy dwie różne reguły mnożenia wektorów należałoby prześledzić sytuację, w której wynik mnożenia dwóch wektorów przemnażamy raz jeszcze – skalarnie lub wektorowo – przez trzeci wektor. Oczywiście, z interesującymi przypadkami będziemy mieli do czynienia tylko wtedy, kiedy „pierwsze” mnożenie jest iloczynem wektorowym. (Gdyby przemnożyć dwa wektory skalarnie to wynik mnożenia jest skalarzem i pomnożenie przez niego trzeciego wektora jest operacją trywialnie prostą.) Tak więc interesować nas będą w zasadzie dwie możliwości:

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) \text{ iloczyn mieszany (skalarny iloczyn mieszany)}$$

oraz

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) \text{ podwójny iloczyn wektorowy.}$$

Iloczyn mieszany (skalarny iloczyn mieszany)

Iloczyn taki jest skalarzem (stąd nazwa), ponieważ drugi jego czynnik jest wektorem: $\mathbf{B} \times \mathbf{C} = \mathbf{D}$, a więc

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = (\mathbf{B} \times \mathbf{C}) \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{D} = \mathbf{D} \cdot \mathbf{A} = \text{skalar.}$$

Rozpisanie czynników iloczynu mieszanego na współrzędne i zastosowanie reguł obliczania iloczynu skalarnego i wektorowego daje:

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = A_x(B_y C_z - B_z C_y) + A_y(B_z C_x - B_x C_z) + A_z(B_x C_y - B_y C_x). \quad (4-38)$$

Na podstawie powyższego wyniku łatwo sprawdzić, że w iloczynie mieszanym można wymienić znaki mnożenia skalarnego i wektorowego, to znaczy

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C} = (\mathbf{B} \times \mathbf{C}) \cdot \mathbf{A} = \mathbf{B} \cdot (\mathbf{C} \times \mathbf{A}). \quad (4-39)$$

(Przestawienie porządku czynników pomiędzy którymi występuje symbol iloczynu skalarnego nie zmienia niczego, natomiast ewentualne przestawienie porządku czynników pomiędzy którymi występuje symbol iloczynu wektorowego powoduje zmianę znaku.)

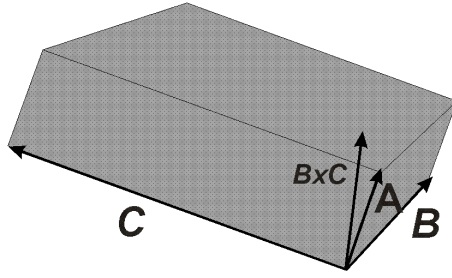
Własność 4-39 wynika także z zapisu iloczynu mieszanego w formie wyznacznika i skorzystania z reguł przestawiania kolumn i wierszy wyznaczników. Zapis relacji 4-38 w postaci wyznacznika to – jak łatwo sprawdzić

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \begin{vmatrix} A_x & A_y & A_z \\ B_x & B_y & B_z \\ C_x & C_y & C_z \end{vmatrix}. \quad (4-40)$$

Podobnie jak w przypadku iloczynu wektorowego, iloczyn mieszany ma interesującą interpretację geometryczną, zilustrowaną na Rys.4.7. Trzy wektory \mathbf{A} , \mathbf{B} i \mathbf{C} tworzą równoległoscian⁹. Iloczyn wektorowy $\mathbf{B} \times \mathbf{C}$ to (por. poprzedni podrozdział) wektor, którego długość jest równa polu podstawy równoległoscianu, napiętej przez wektory \mathbf{B} i \mathbf{C} . Wektor ten przemnażamy skalarnie przez wektor \mathbf{A} – a więc długość wektora $\mathbf{B} \times \mathbf{C}$ (prostopadłego do podstawy) mnożymy przez odpowiedni rzut wektora \mathbf{A} . Wynik mnożenia to objętość równoległoscianu. W tym momencie staje się też intuicyjnie zrozumiałe równanie 4-39 – przecież jako podstawę naszego równoległoscianu możemy wybrać dowolną dwójkę z trzech wektorów, a wysokością będzie wtedy

Podwójny iloczyn wektorowy

⁹Oczywiście pod warunkiem, że żadne dwa wektory z trójki nie są współliniowe. Wtedy jednak – por. 4-26 – ich iloczyn wektorowy, a także iloczyn mieszany są równe zeru – nie ma więc sensu mówić o geometrycznej interpretacji.



Rysunek 4.7: Interpretacja geometryczna iloczynu mieszanego – objętość równoległoscianu.

Iloczyn wektorowy $\mathbf{A} \times (\mathbf{B} \times \mathbf{C})$ jest wektorem. Zauważmy od razu, że nawiasy w nim występujące mają zasadnicze znaczenie. I tak iloczyn trzech wektorów $(\hat{\mathbf{x}} \times \hat{\mathbf{x}}) \times \hat{\mathbf{y}} = 0$ (bo pierwszy czynnik iloczynu jest równy zeru), natomiast $\hat{\mathbf{x}} \times (\hat{\mathbf{x}} \times \hat{\mathbf{y}}) = \hat{\mathbf{x}} \times \hat{\mathbf{z}} = -\hat{\mathbf{y}}$. Zgodnie z definicją iloczynu wektorowego, wektor $\mathbf{A} \times (\mathbf{B} \times \mathbf{C})$ jest prostopadły do obu wektorów: \mathbf{A} oraz $\mathbf{B} \times \mathbf{C}$. Ten ostatni iloczyn jest z kolei prostopadły do płaszczyzny napiętej przez wektory \mathbf{B} i \mathbf{C} – a więc iloczyn $\mathbf{A} \times (\mathbf{B} \times \mathbf{C})$ musi leżeć w tej płaszczyźnie, a więc można go zapisać jako kombinację liniową wektorów \mathbf{B} i \mathbf{C} :

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = x\mathbf{B} + y\mathbf{C}. \quad (4-41)$$

Dwa nieznanne współczynniki tej kombinacji liniowej wyznaczymy mnożąc skalarnie równanie 4-41 przez wektor \mathbf{A} . Lewa strona otrzymanego równania jest równa zeru (skalarny iloczyn wektorów wzajemnie prostopadłych), a więc

$$x\mathbf{A} \cdot \mathbf{B} + y\mathbf{A} \cdot \mathbf{C} = 0, \quad (4-42)$$

co można zapisać jako

$$\frac{x}{y} = \frac{\mathbf{A} \cdot \mathbf{C}}{-\mathbf{A} \cdot \mathbf{B}} \equiv \frac{\alpha \mathbf{A} \cdot \mathbf{C}}{-\alpha \mathbf{A} \cdot \mathbf{B}}, \quad (4-43)$$

gdzie α może być w zasadzie dowolną, różną od zera liczbą. Tak więc $x = \alpha \mathbf{A} \cdot \mathbf{C}$ i $y = -\alpha \mathbf{A} \cdot \mathbf{B}$; podstawiając te wartości do 4-41 otrzymujemy

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \alpha[\mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B})]. \quad (4-44)$$

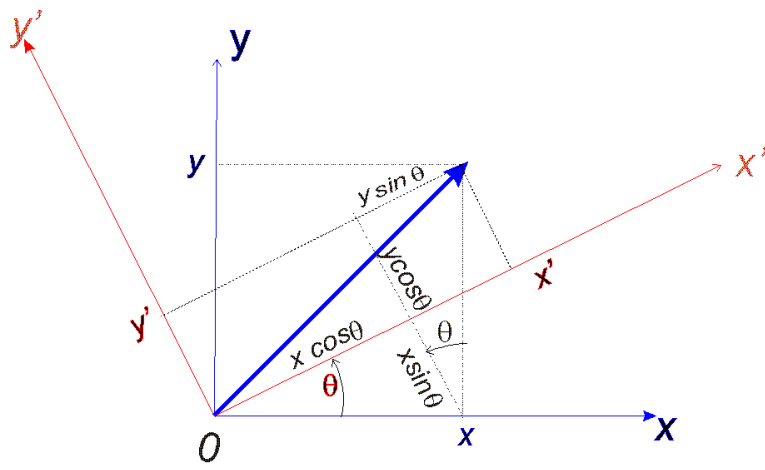
Stosunkowo łatwo wykażemy, że stała α musi być równa 1. Wystarczy rozważyć iloczyn trzech wektorów o długości jednostkowej, na przykład wspomniany już $\hat{\mathbf{x}} \times (\hat{\mathbf{x}} \times \hat{\mathbf{y}}) = -\hat{\mathbf{y}}$. Mamy

$$\hat{\mathbf{x}} \times (\hat{\mathbf{x}} \times \hat{\mathbf{y}}) = \alpha[\hat{\mathbf{x}}(\hat{\mathbf{x}} \cdot \hat{\mathbf{y}}) - \hat{\mathbf{y}}(\hat{\mathbf{x}} \cdot \hat{\mathbf{x}})] = \alpha[0 - \hat{\mathbf{y}}] = -\hat{\mathbf{y}}.$$

Wybraliśmy wprawdzie nieco „specjalne” wektory do wykazania $\alpha = 1$, ale dokładniejsze rozważania pozwalają stwierdzić, że nie ma to wpływu na słuszność wniosku. Zresztą słuszność wzoru 4-44 (z $\alpha = 1$) można także wykazać rozpisując podwójny iloczyn wektorowy na współrzędne, zgodnie z wzorami poprzedniego podrozdziału.

4.5 Obrót wektora na płaszczyźnie

Wektor oznacza dla nas – praktycznie zawsze – pewną realność fizyczną. Siła działająca na ciało i nadająca mu przyspieszenie, pęd poruszającej się cząstki, natężenie pola – takie wielkości istnieją w sposób „obiektywny”, to znaczy ich byt nie może być uzależniony od wyboru układu współrzędnych. Natomiast od wyboru będą zależały współrzędne (składowe). Na Rys. 4.8 mamy przedstawioną sytuację, w której ten sam wektor \mathbf{r} został rozłożony na współrzędne w dwóch układach: układzie Σ ($0xy$) i układzie „primowanym” Σ' ($0x'y'$),



Rysunek 4.8: Współrzędne wektora w dwóch różnych układach współrzędnych.

który w stosunku do Σ jest obrócony o kąt θ^{10} . Przy pomocy rozważań trygonometrycznych (por. rysunek) możemy wyprowadzić wzory wiążące ze sobą współrzędne wektora \mathbf{r} :

$$\begin{aligned} x' &= x \cos \theta + y \sin \theta \\ y' &= -x \sin \theta + y \cos \theta. \end{aligned} \quad (4-45)$$

Obrót układu – wybór nowego układu – powoduje zmianę wartości współrzędnych wektora. Nie zmienia się przy tym *długość wektora*

$$|\mathbf{r}| = r = \sqrt{x^2 + y^2} = \sqrt{x'^2 + y'^2}.$$

Nie jest to dla nas niespodzianką – długość wektora (liczba!) jest skalarą, a definicja formalna skalarów mówi, że jest to wielkość *niezmiennicza* względem transformacji układu współrzędnych, a więc niezależna od układu. Na obecnym poziomie naszych wiadomości o wektorach, możemy wprowadzić „definicję roboczą”¹¹ wektora: **Wektorem \mathbf{A} nazywamy wielkość, której współrzędne A_x, A_y (w układzie Σ) transformują się przy obrocie układu o kąt θ według**

$$\begin{aligned} A'_x &= A_x \cos \theta + A_y \sin \theta \\ A'_y &= -A_x \sin \theta + A_y \cos \theta. \end{aligned} \quad (4-46)$$

(A'_x, A'_y to współrzędne w układzie Σ' , powstałym w wyniku obrotu Σ .)

W praktyce mamy do czynienia z wektorami nie tylko na płaszczyźnie – o dwóch niezależnych współrzędnych. Nasza przestrzeń fizyczna jest przecież przestrzenią trójwymiarową. Co więcej – dość proste zabiegi formalne, stosowane w opisie podstawowych zjawisk fizycznych, zmuszają nas często do korzystania z przestrzeni o dowolnej liczbie wymiarów. Na przykład, opisując układ N cząstek poruszających się w przestrzeni 3-wymiarowej możemy wprowadzić wektor o $3N$ współrzędnych: $x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_N, y_N, z_N$ jako wektor określający położenie wszystkich cząstek – całego układu. Innymi słowy, potrzebne nam jest uogólnienie wzorów 4-46 na przypadek *przestrzeni wektorowej* o dowolnej (ale skończonej) liczby wymiarów. Aby to zrobić, wprowadźmy notację

$$\begin{aligned} \hat{\mathbf{x}} &\rightarrow \mathbf{e}_1 & \hat{\mathbf{y}} &\rightarrow \mathbf{e}_2 \\ \hat{\mathbf{x}}' &\rightarrow \mathbf{e}'_1 & \hat{\mathbf{y}}' &\rightarrow \mathbf{e}'_2 \\ A_x &\rightarrow A_1 & A_y &\rightarrow A_2 \\ A'_x &\rightarrow A'_1 & A'_y &\rightarrow A'_2 \end{aligned} \quad (4-47)$$

a także

$$\left. \begin{aligned} \cos \theta &= \cos \angle(0x', 0x) = \mathbf{e}'_1 \cdot \mathbf{e}_1 \rightarrow a_{11}, \\ \sin \theta &= \cos\left(\frac{\pi}{2} - \theta\right) = \cos \angle(0x', 0y) = \mathbf{e}'_1 \cdot \mathbf{e}_2 \rightarrow a_{12}, \\ -\sin \theta &= \cos\left(\frac{\pi}{2} + \theta\right) = \cos \angle(0y', 0x) = \mathbf{e}'_2 \cdot \mathbf{e}_1 \rightarrow a_{21}, \\ \cos \theta &= \cos \angle(0y', 0y) = \mathbf{e}'_2 \cdot \mathbf{e}_2 \rightarrow a_{22}. \end{aligned} \right\} \quad (4-48)$$

¹⁰Przypominamy – kierunek dodatni przyrostu kąta to kierunek odwrotny do ruchu wskazówek zegara

¹¹Przez „definicję roboczą” rozumiemy określenie pewnej wielkości w kontekście jej własności względem pewnych operacji. Taką operacją jest tutaj wyrażenie współrzędnych wektora w różnych – obróconych względem siebie – układach współrzędnych.

Wówczas równania 4-46 zapiszemy

$$\begin{aligned} A'_1 &= a_{11}A_1 + a_{12}A_2 \\ A'_2 &= a_{21}A_1 + a_{22}A_2. \end{aligned} \quad (4-49)$$

Możemy powyższe wzory zapisać przy pomocy znaku sumy ¹²:

$$A'_n = \sum_{k=1}^2 a_{nk}A_k, \quad n = 1, 2. \quad (4-50)$$

Wzory 4-49 opisują sytuację dwuwymiarową – obrót układu na płaszczyźnie. W przestrzeni n -wymiarowej sytuacja wygląda nieco bardziej skomplikowanie – ale spróbujmy potraktować wzory 4-49 jako „podprzypadek” bardziej ogólnego (odniesionego do przestrzeni trójwymiarowej) układu równań. Ten układ powinien mieć postać:

$$\begin{aligned} A'_1 &= a_{11}A_1 + a_{12}A_2 + a_{13}A_3 \\ A'_2 &= a_{21}A_1 + a_{22}A_2 + a_{23}A_3 \\ A'_3 &= a_{31}A_1 + a_{32}A_2 + a_{33}A_3. \end{aligned} \quad (4-51)$$

W powyższych wzorach (A_1, A_2, A_3) i (A'_1, A'_2, A'_3) to współrzędne wektora \mathbf{A} w układach Σ i Σ' . Współczynniki a_{ik} to – zgodnie z definicjami 4-48 – kosinusy kątów pomiędzy osiami obu układów, albo iloczyny skalarne wersorów tych osi:

$$a_{ik} = \cos \angle(\text{os } i \text{ w układzie } \Sigma', \text{ os } k \text{ w układzie } \Sigma) = \mathbf{e}'_i \cdot \mathbf{e}_k. \quad (4-52)$$

Równanie 4-50 praktycznie nie zmienia się – jedynie wskaźnik sumowania przebiega wartości od 1 do 3 (a nie do 2), definicja transformacji odnosi się do $n = 3$ współrzędnych:

$$A'_n = \sum_{k=1}^3 a_{nk}A_k, \quad n = 1, 2, 3. \quad (4-53)$$

Nie możemy jednak polegać tylko na naszej intuicji. Wzór 4-53 należałoby więc uzasadnić formalnie. Otóż wspomnieliśmy już, że wektor \mathbf{A} może być zapisany w układzie Σ jako kombinacja trzech wektorów niekomplanarnych (nie leżących na jednej płaszczyźnie). Jeżeli za takie trzy wektory wybierzemy wersory $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, to współczynniki liczbowe tej liniowej kombinacji są współrzędnymi wektora w tym właśnie układzie:

$$\mathbf{A} = \mathbf{e}_1A_1 + \mathbf{e}_2A_2 + \mathbf{e}_3A_3. \quad (4-54)$$

Każdą z tych współrzędnych uzyskujemy przemnażając skalarnie wektor \mathbf{A} przez odpowiedni wersor – rzutując go na kierunek tego wersora. Podobnie będzie w układzie Σ' . Współrzędne primowane powstają wyniku przemnożenia \mathbf{A} przez wersory $\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3$. Na przykład A'_1 to $\mathbf{A} \cdot \mathbf{e}'_1$. Podstawiając do tego wzoru za \mathbf{A} z 4-54 mamy:

$$\begin{aligned} A'_1 &= \mathbf{A} \cdot \mathbf{e}'_1 = (\mathbf{e}_1A_1 + \mathbf{e}_2A_2 + \mathbf{e}_3A_3) \cdot \mathbf{e}'_1 \\ &= A_1\mathbf{e}_1 \cdot \mathbf{e}'_1 + A_2\mathbf{e}_2 \cdot \mathbf{e}'_1 + A_3\mathbf{e}_3 \cdot \mathbf{e}'_1 \\ &= A_1a_{11} + A_2a_{12} + A_3a_{13}, \end{aligned} \quad (4-55)$$

w pełnej zgodzie z równaniem 4-53.

Powyższe wzory pozwala obliczyć współrzędne naszego wektora \mathbf{A} w układzie Σ' jeżeli znane są współrzędne w układzie Σ . A gdyby było odwrotnie? Rozważając przypadek dwuwymiarowy, powiedzieliśmy, że układ Σ' powstaje z Σ w wyniku obrotu o kąt θ . W takim razie układ Σ powstaje z Σ' w wyniku obrotu o kąt $-\theta$ i – na przykład – układ równań analogiczny do 4-46 będzie miał postać

$$\begin{aligned} A_x &= A'_x \cos(-\theta) + A'_y \sin(-\theta) = A'_x \cos \theta - A'_y \sin \theta \\ A_y &= -A'_x \sin(-\theta) + A'_y \cos(-\theta) = A'_x \sin \theta + A'_y \cos(\theta). \end{aligned} \quad (4-56)$$

¹²**Uwaga:** dla czytelnika, który pierwszy raz spotkał się z zapisem sumy w tych notatkach: jak już wyjaśnialiśmy znaki sumy to operacje dokonywanych na wielkościach opatrzonych we wskaźniki (wielkości „indeksowane”). Wskaźnik po którym sumujemy – tutaj k ; $k = 1, 2$ – musi się powtórzyć pod znakiem sumy, a po wysumowaniu wielkość od takiego wskaźnika już nie zależy. (Zauważmy, że moglibyśmy zamiast literki k użyć dowolnej litery – jest to tzw. wskaźnik martwy). Natomiast obie strony równania 4-50 **muszą** zależeć od wskaźnika n .

Postępując analogicznie, jak w przypadku rozszerzenia transformacji z układu Σ do Σ' na przypadek trójwymiarowy będziemy mieli *transformację odwrotną*: wyrażenie współrzędnych „nieprimowanych” A_k przez primowane A'_n (por. 4-53)

$$A_k = \sum_{n=1}^3 a'_{kn} A'_n, \quad n = 1, 2, 3, \quad (4-57)$$

gdzie – analogicznie do definicji 4-52

$$a'_{kn} = \cos \angle(\text{ oś } k \text{ w układzie } \Sigma, \text{ oś } n \text{ w układzie } \Sigma') = \mathbf{e}_k \cdot \mathbf{e}'_n = a_{nk}. \quad (4-58)$$

Transformacje pomiędzy dwoma, obróconymi względem siebie, układami współrzędnych są w pełni określone poprzez 9-cio wartościową „tablicę” kosinusów kierunkowych układów. Wszystko jedno czy definiujemy te tablice jako współczynniki a_{kn} , czy a'_{nk} . Jak wynika z określenia tych wielkości – wzory 4-52i 4-58 — znając jedno znamy automatycznie drugie. Macierz transformacji odwrotnej (z układu Σ' do układu Σ) jest macierzą transponowaną w stosunku do macierzy transformacji z układu Σ do układu Σ' . Taką własność – w kontekście macierzy i opisywanych przez nie transformacji – nazywamy *ortogonalnością macierzy (transformacji)*. Będzie o tym mowa w rozdziale dziesiątym.

Rozdział 5

Liczby zespolone

5.1 Liczby zespolone – trochę historii

W połowie 15. wieku problemem No.1 matematyki stało się rozwiązanie równanie trzeciego stopnia. Można powiedzieć, że czas był to najwyższy! Równanie drugiego stopnia umieli już rozwiązywać ...rachmistrze sumeryjscy, 2000 lat przed Chrystusem. Przez niewytłumaczalny kaprys historii rozwoju ludzkiego intelektu problem „o stopień wyższy” czekał na rozwiązanie następne trzy i pół tysiąca lat. Rozwiązanie równanie trzeciego stopnia wiąże się zazwyczaj z nazwiskiem Girolamo Cardano (1501–1576), chociaż wydaje się, że ten niewątpliwie wszechstronny uczoney – prawdziwy „człowiek Renesansu” wykorzystał w swoich dziełach *Practica Mathematicae* (1539) i *Ars Magna* (1545) wyniki uzyskane przez współczesnego mu (i z pewnością nie ustępującego rangą) Nicolo Tartaglia (1500–1557), który zresztą również „inspirował” się wynikami działającego o pół wieku wcześniej Bolończyka Scipione del Ferro (1465–1526).

Technika znalezienia pierwiastków równania

$$x^3 + ax^2 + bx + c = 0$$

polegała na sprowadzeniu równania — poprzez podstawienie

$$x = y - \frac{a}{3}; \quad p = b - \frac{a^2}{3} \quad \text{i} \quad q = -\left(\frac{2a^3}{27} - \frac{ab}{3} + c\right)$$

— do pozbawionego wyrazu z drugą potęgą równania:

$$y^3 = py + q.$$

To właśnie Tartaglia pokazał, że ostatnie równanie ma rozwiązanie

$$y = \sqrt[3]{\frac{q}{2} + \sqrt{\frac{q^2}{4} - \frac{p^3}{27}}} + \sqrt[3]{\frac{q}{2} - \sqrt{\frac{q^2}{4} - \frac{p^3}{27}}}.$$

Jak nietrudno zauważyć, dla $(q/2)^2 < (p/3)^3$ wielkość występująca pod kwadratowym pierwiastkiem staje się ujemna. I tak na przykład „historyczne równanie”, opisywane przez Rafaela Bombelliego, ostatniego z wielkich bolońskich matematyków 16. wieku

$$x^3 = 15x + 4 \tag{5-1}$$

miałoby mieć – zgodnie ze wzorem Cardana-Tartaglii – rozwiązanie

$$x = \sqrt[3]{2 + \sqrt{-121}} + \sqrt[3]{2 - \sqrt{-121}}. \tag{5-2}$$

Występujące pod kwadratowym pierwiastkiem -121 przeczyło zdrowemu (szesnastowiecznemu) rozsądkowi. Ale Bombelli wiedział, że rozwiązaniem równania 5-1 są „prawdziwe” (rzeczywiste liczby): 4 oraz $-2 \pm \sqrt{3}$. Rewolucyjny pomysł Bombelliego polegał na założeniu, że występujące w rozwiązaniu 5-2 pierwiastki trzeciego stopnia to *liczby zespolone*, będące sumą liczby „zwykłej” (rzeczywistej) i „urojonej” – powstającej z przemnożenia pewnej liczby rzeczywistej przez $\sqrt{-1}$. W dodatku oba pierwiastki trzeciego stopnia powinny

się różnić między sobą pojawiającym się w sumie części rzeczywistej i urojonej znakiem, w sposób identyczny do tego w jaki różnią się wielkości występujące pod znakiem pierwiastka, to znaczy

$$\sqrt[3]{2 + \sqrt{-121}} \equiv \alpha + \beta\sqrt{-1} \quad \text{ i } \quad \sqrt[3]{2 - \sqrt{-121}} \equiv \alpha - \beta\sqrt{-1},$$

gdzie α i β należałoby wyznaczyć. Przy takim założeniu:

$$\begin{aligned} 2 + \sqrt{-121} &= 2 + 11\sqrt{-1} = (\alpha + \beta\sqrt{-1})^3 \\ &= \alpha^3 + 3\alpha^2\beta\sqrt{-1} + 3\alpha\beta(\sqrt{-1})^2 + \beta^3(\sqrt{-1})^3 \\ &= \alpha(\alpha^2 - 3\beta^2) + \beta(3\alpha^2 - \beta^2)\sqrt{-1}. \end{aligned}$$

Ostatnia równość będzie spełniona, jeżeli

$$\alpha(\alpha^2 - 3\beta^2) = 2 \quad \text{ oraz } \quad \beta(3\alpha^2 - \beta^2) = 11.$$

Powyższe równania mają – w klasie liczb całkowitych – rozwiązania dla $\alpha = 1$ lub 2 , oraz dla $\beta = 1$ lub 11 . Bliższa analiza wykazuje, że jedyny akceptowalny wybór to $\alpha = 2$ i $\beta = 1$, a więc rozwiązaniem 5-2 będzie

$$\begin{aligned} x &= \sqrt[3]{2 + \sqrt{-121}} + \sqrt[3]{2 - \sqrt{-121}} \\ &= \sqrt[3]{(\alpha + \beta\sqrt{-1})^3} + \sqrt[3]{(\alpha - \beta\sqrt{-1})^3} \\ &= \sqrt[3]{(2 + 1 \times \sqrt{-1})^3} + \sqrt[3]{(2 - 1 \times \sqrt{-1})^3} \\ &= (2 + 1 \times \sqrt{-1}) + (2 - 1 \times \sqrt{-1}) = 4. \end{aligned}$$

W ten sposób właśnie pojawiły się „liczby zespolone”, zawierające w sobie urojoną (a więc nieistniejącą) wielkość – kwadratowy pierwiastek z -1 . Przez przeszło dwieście lat pozostawały pełną abstrakcją matematyczną – abstrakcją, która odpowiednio manipulowana mogła jednak doprowadzić do realnych wyników.

Dopiero na początku osiemnastego wieku powstała nowa koncepcja – wykorzystania tych tworów matematycznych do opisu płaszczyzny. Tak jak zbiór liczb rzeczywistych można w sposób jedno-jednoznaczny przedstawić przy pomocy osi liczbowej x (każdy punkt osi odpowiada pewnej liczbie rzeczywistej od $-\infty$ do ∞ i odwrotnie), tak można wprowadzić jedno-jednoznaczne przyporządkowanie pomiędzy parami liczb i punktami płaszczyzny. Uporządkowaną parę liczb (a, b) traktować możemy jako współrzędne końca wektora, którego początek pokrywa się z początkiem układu współrzędnych. Osie tego układu to dwie „tradycyjne” osie liczbowe, $0x$ i $0y$, z tym, że jednostką osi $0x$ jest 1 , a osi $0y$ – urojona jednostka to $i = \sqrt{-1}$. Te jednostki spełniają jednocześnie role wersorów osi, w tym sensie że dowolny punkt na *płaszczyźnie zespolonej* (zwanej też płaszczyzną Arganda) możemy przedstawić jako $z = 1 \cdot a + i \cdot b$. Współrzędna x -owa, (a) to *część rzeczywista* liczby zespolonej, natomiast współrzędna y -owa – b to jej *część urojona*. Analogicznie mówimy o rzeczywistej osi $0x$ i osi urojonej $0y$ płaszczyzny $0xy$.

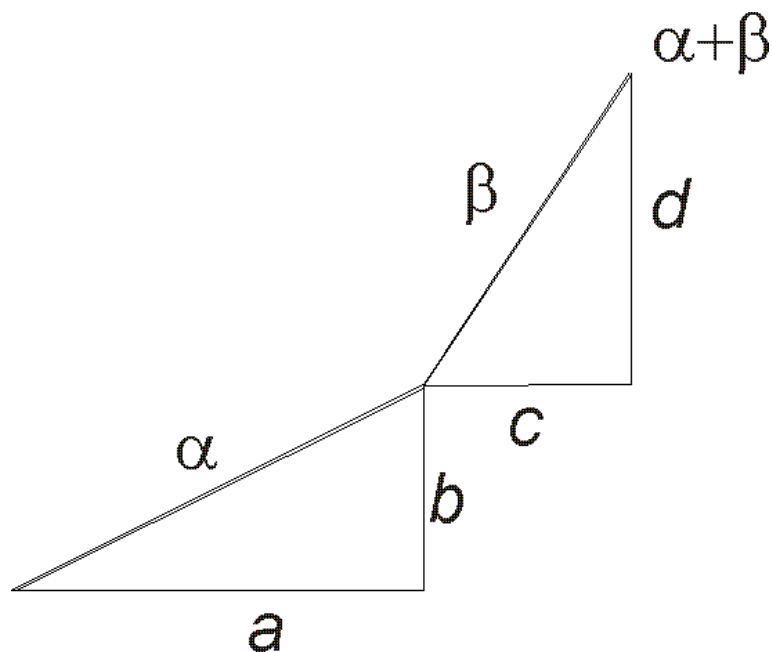
W dalszym jednak ciągu przydatność liczb zespolonych była mało widoczna. Można ich było użyć (zobaczmy to w tym rozdziale) do zgrabnego zapisu pewnych operacji na wektorach w przestrzeni dwuwymiarowej na płaszczyźnie. Dopiero w drugiej połowie 19. wieku zaczęła się objawiać potęga algebry liczb zespolonych, a zwłaszcza teorii funkcji *zmiennej zespolonej*. W fizyce wielkości zespolone mają często znakomitą interpretację formalną – np. *zespolony* współczynnik załamania to wielkość fizyczna składająca się z dwóch części: rzeczywistej – odpowiedzialnej za zjawisko załamania fali padającej na granicę dwóch ośrodków i urojonej – która odpowiada za zjawisko absorpcji.

5.2 Algebra liczb zespolonych

Mamy więc płaszczyznę zespoloną z wybranym układem osi $0x$ i $0y$. Każdy punkt płaszczyzny to wektor (np. α), który możemy zapisać jako sumę dwóch współrzędnych: rzeczywistej (a) i urojonej (b) , przemnożonych przez wersory osi: $\alpha = a + ib$. Dwóm punktom płaszczyzny α i β niech odpowiadają dwie pary liczb: (a, b) i (c, d) . Zwykle prawa algebry, przeniesione na liczby zespolone dają — przy zachowaniu umowy, że $i^2 = -1$:

$$\alpha + \beta = (a, b) + (c, d) \equiv (a + ib) + (c + id) = (a + c) + i(b + d), \quad (5-3)$$

$$\alpha \cdot \beta = (a, b)(c, d) \equiv (a + ib)(c + id) = (ac - bd) + i(ad + bc). \quad (5-4)$$



Rysunek 5.1: Dodawanie liczb zespolonych na płaszczyźnie Arganda.

Tak więc suma punktów, reprezentowanych przez liczby (a, b) i (c, d) to punkt o odciętej $a + c$ i rzędnej $b + d$; dla iloczynu odpowiednie współrzędne to $ac - bd$ i $ad + bc$. Dodawanie liczb zespolonych – wektorów na płaszczyźnie zespolonej ilustruje Rys.5.1.

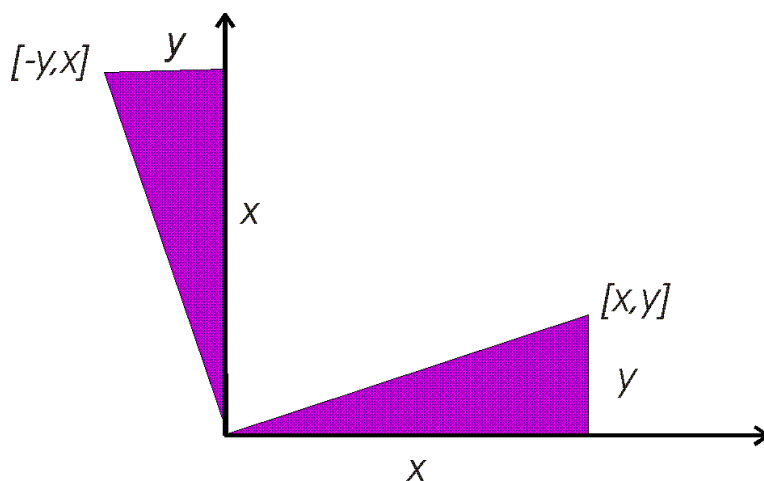
Operację mnożenia łatwiej będzie nam prześledzić dla kilku podprzypadków. I tak przemnożenie liczby zespolonej $(x, y) \equiv x + iy$ przez rzeczywistą liczbę A :

$$A(x + iy) = Ax + iAy \quad (5-5)$$

to przemnożenie współrzędnej x -owej i y -owej punktu przez A . Wektor, któremu odpowiada para (x, y) ulega skróceniu ($A < 1$) lub wydłużeniu ($A > 1$), a kierunek jego pozostaje bez zmian. Z kolei, w wyniku przemnożenia $(x, y) \equiv x + iy$ przez czysto urojoną liczbę iB :

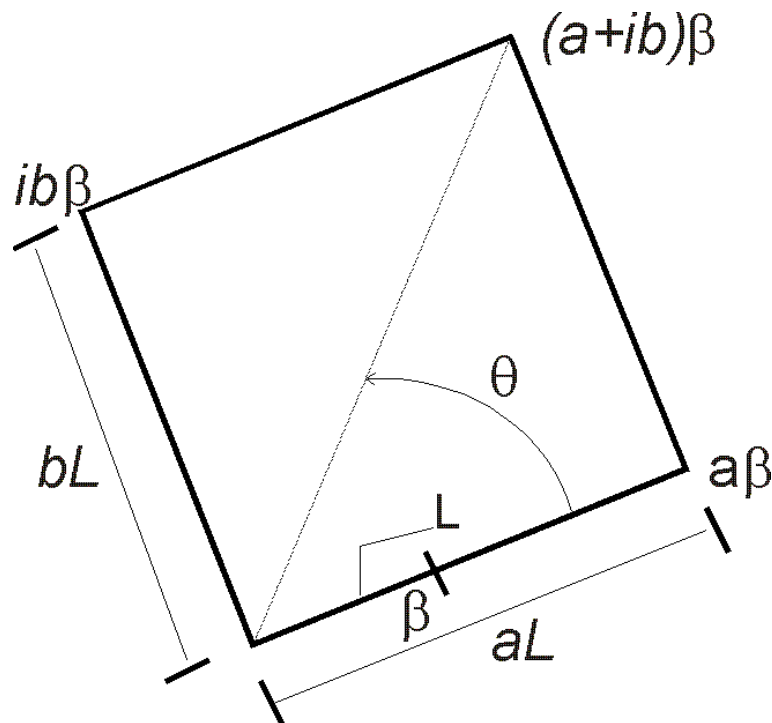
$$iB(x + iy) = -By + iBx \quad (5-6)$$

powstaje wektor, którego współrzędna urojona to „stara” współrzędna rzeczywista – wydłużona ($B > 1$)



Rysunek 5.2: Mnożenie liczby zespolonej przez i .

lub skrócona ($B < 1$), a współrzędna rzeczywista powstaje ze starej współrzędnej urojonej, wydłużonej lub skróconej w tym samym stosunku, po pomnożeniu przez -1 . Z Rys.5.2 (trójkąty podobne) wynika, że mnożenie przez urojoną jedynkę i powoduje obrót wektora o $\pi/2$. Długość tak obróconego wektora to stara długość $\times B$.



Rysunek 5.3: Mnożenie dwóch liczb zespolonych.

Rysunek 5.3 ilustruje przypadek najogólniejszy (równ. 5-4). Wektor $\beta \equiv (c, d)$ mnożymy przez a i – niezależnie – przez ib , a następnie składamy tak uzyskane wektory. Jeżeli długość wektora β oznaczymy przez $L = \sqrt{c^2 + d^2}$, to z rysunku wynika, że długość wektora $\beta \times \alpha = \beta \times (a + ib)$ jest równa $L\sqrt{a^2 + b^2}$, a wektor ten jest obrócony (w stosunku do β) o kąt θ , przy czym:

$$\operatorname{tg} \theta = \frac{bL}{aL} = \frac{b}{a}. \quad (5-7)$$

Ponieważ $\sqrt{a^2 + b^2}$ to nic innego jak długość wektora α , to wynik mnożenia dwóch liczb zespolonych jest wektorem o długości równej iloczynowi długości dwóch wektorów, czynników iloczynu. Wektor ten powstaje w wyniku obrotu jednego z czynników iloczynu (np. wektora β) o kąt, którego tangens określa stosunek części urojonej i rzeczywistej drugiego czynnika.

5.2.1 Reprezentacja biegunowa liczby zespolonej; liczba zespolona sprzężona; dzielenie liczb zespolonych

Dowolna liczba zespolona z – wektor płaszczyzny zespolonej – może być przedstawiona w postaci sumy części rzeczywistej i urojonej

$$z = x + iy. \quad (5-8)$$

Jeżeli zamiast współrzędnych kartezjańskich (układ dwóch osi liczbowych $0x$ i $0y$) użyjemy do opisu położenia końca wektora z współrzędnych biegunowych: odległości (od początku układu) r i kąta (wektora względem osi $0x$) ϕ to (Rys.5.4)

$$z = x + iy = r \cos \phi + i r \sin \phi = r(\cos \phi + i \sin \phi). \quad (5-9)$$

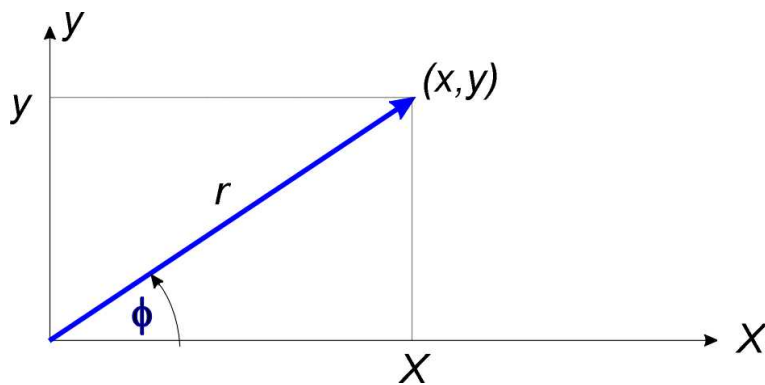
W kontekście algebry liczb zespolonych, wielkość r to *moduł* albo *wartość bezwzględna* liczby z

$$r = |z| = \sqrt{x^2 + y^2}, \quad (5-10)$$

a kąt ϕ to jej *argument*¹. Tangens tego kąta jest równy (por. wzór 5-7)

$$\operatorname{tg} \phi = \arg(z) = \frac{y}{x}. \quad (5-11)$$

¹W żargonie fizyki i techniki używa się także terminu *faza*.



Rysunek 5.4: Współrzędne biegunowe na płaszczyźnie zespolonej.

Wzór 5-9 możemy zapisać w szczególnie wygodnej postaci, jeżeli skorzystamy z tożsamości²

$$\cos x + i \sin x = e^{ix}. \quad (5-12)$$

Wówczas

$$z = r(\cos \phi + i \sin \phi) \equiv re^{i\phi}. \quad (5-13)$$

Powyższy wzór potwierdza zauważone wcześniej operacje, jakich dokonujemy podczas mnożenia dwóch liczb zespolonych:

$$z_1 \cdot z_2 = r_1 e^{i\phi_1} \cdot r_2 e^{i\phi_2} = r_1 r_2 e^{i(\phi_1 + \phi_2)} = |z_1| \cdot |z_2| e^{i[\arg(z_1) + \arg(z_2)]} \quad (5-14)$$

– moduł iloczynu równy jest iloczynowi modułów czynników, a jego argument – sumie argumentów.

Aby zdefiniować dzielenie liczb zespolonych wprowadźmy pojęcie *liczby zespolonej sprzężonej* (do danej liczby zespolonej)³:

$$z^* \stackrel{\text{def}}{=} x - iy = r(\cos \phi - i \sin \phi) = r[\cos(-\phi) + i \sin(-\phi)] = re^{-i\phi}. \quad (5-15)$$

Tak więc liczba zespolona sprzężona to wektor stanowiący zwierciadlane odbicie wektora z w osi rzeczywistej (Rys.5.5). Formalnie, liczbę zespoloną uzyskujemy więc zamieniając i na $-i$ w 5-8 i 5-13. Zauważmy

$$z \cdot z^* = (x + iy) \cdot (x - iy) = re^{i\phi} \cdot re^{-i\phi} = r^2 = x^2 + y^2 = |z|^2 \quad (5-16)$$

— iloczyn liczby zespolonej z i liczby do niej sprzężonej z^* jest równy kwadratowi modułu z . Teraz możemy już łatwo określić operację dzielenia dwóch liczb zespolonych $z_1 = a + ib$ i $z_2 = c + id$. Licznik i mianownik ułamka, przedstawiającego sobą iloraz tych liczb, mnożymy przez zespoloną sprzężoną mianownika

$$\frac{z_1}{z_2} = \frac{a + ib}{c + id} = \frac{a + ib}{c + id} \cdot \frac{c - id}{c - id} = \frac{ac + bd}{c^2 + d^2} + i \frac{bc - ad}{c^2 + d^2}, \quad (5-17)$$

czyli część rzeczywistą i urojoną liczby uzyskanej w wyniku pomnożenia dzielnej przez zespoloną sprzężoną dzielnika dzielimy przez kwadrat modułu dzielnika.

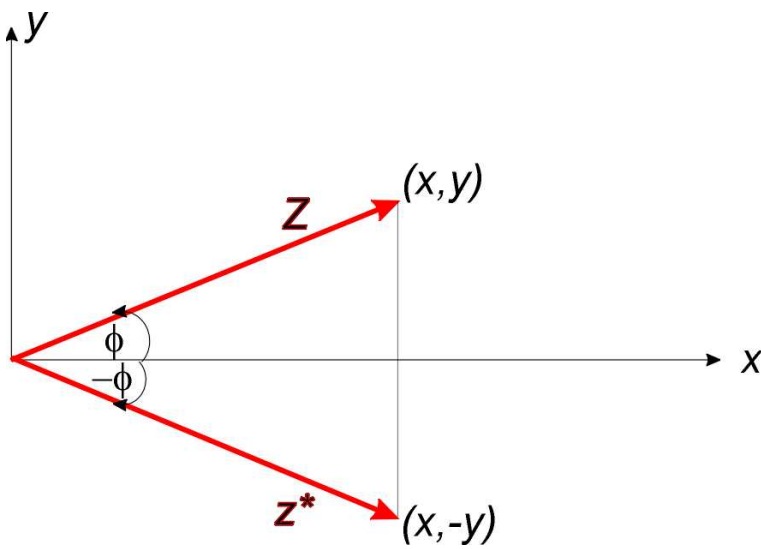
Podsumujmy: opierając się na „pomyśle” Bombelliego $i = \sqrt{-1}$ i stosując prawa algebry odnoszące się do liczb rzeczywistych (takie m.in. jak prawo łączności dodawania, przemienności mnożenia i rozdzielności mnożenia

²Wyprowadzenie tej tożsamości opiera się na możliwości przedstawienia wszystkich występujących w niej funkcji w postaci *szeregów potęgowych* (szeregów Taylora):

$$\begin{aligned} e^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!}, \\ \cos x &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}, \\ \sin x &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}. \end{aligned}$$

Symbol $n!$ to iloczyn $1 \cdot 2 \cdot 3 \cdot \dots \cdot (n-1) \cdot n$. Powyższe wzory pojawiają się w analizie matematycznej.

³Oprócz oznaczenia z^* powszechnie stosowane jest oznaczenie \bar{z} .



Rysunek 5.5: Liczba zespolona sprzężona.

względem dodawania) uzyskujemy logiczną strukturę algebry liczb o dwóch „stopniach swobody”, nazwanych przez nas częścią rzeczywistą i urojoną liczby zespolonej. Z formalnego punktu widzenia należałoby dodać jeszcze dwa szczegóły. Określając takie operacje jak dodawanie i mnożenie elementu zbioru liczb rzeczywistych, wprowadza się pojęcie takich elementów jak zero i jedność. Dodanie zera do danej liczby rzeczywistej pozostawia ją bez zmian, podobnie jak pomnożenie jej przez jedność. Mówimy wówczas o *elementach neutralnych* w odniesieniu do operacji dodawania i mnożenia. Zauważmy, że takie neutralne elementy istnieją także przypadku liczb zespolonych:

$$(a, b) + (0, 0) = (a, b) \quad \text{dodawanie} \quad (5-18)$$

$$(a, b) \cdot (1, 0) = (a, b) \quad \text{mnożenie.} \quad (5-19)$$

Istnieją więc zespolone „zero” i „jedność”. Analogicznie, będą istniały też operacje odwrotne do dodawania liczb zespolonych (odejmowanie – dodawanie liczby zespolonej „ujemnej”, a więc ze zmienionymi na przeciwne znakami części rzeczywistej i urojonej) i mnożenia (dzielenie). W tym drugim przypadku operacja odwrotna do mnożenia będzie istniała dla wszystkich liczb zespolonych za wyjątkiem zera: $(0, 0)$.

O dwóch liczbach zespolonych możemy powiedzieć, że są równe – będzie to oznaczało, że równe są ich części rzeczywiste i urojone. Jeżeli $z_1 = x_1 + iy_1$ i $z_2 = x_2 + iy_2$, to równość

$$z_1 = z_2$$

jest równoważna układowi

$$x_1 = x_2$$

$$y_1 = y_2.$$

Natomiast nie ma sensu mówienie, że jedna liczba zespolona jest większa (mniejsza) od drugiej. Relacje $>$ i $<$ mają sens tylko w przypadku liczb rzeczywistych, „uporządkowanych” na osi liczbowej. Możemy więc porównywać moduły i argumenty liczb zespolonych, ale nie same liczby. Pewne trudności nastrocza też pojęcie nieskończoności na płaszczyźnie zespolonej. W przypadku liczb rzeczywistych nieskończoność kojarzyła nam się z lewym i prawym „końcem” osi liczbowej – mówiliśmy więc o nieskończoności „dodatniej” $+\infty$ i „ujemnej” $-\infty$. Na płaszczyźnie zespolonej możemy sobie wyobrazić nieskończenie wiele sposobów takiego „nieskończonego” oddalania się od początku układu. Nieformalnym, ale pomocnym obrazem sytuacji będzie wyobrażenie sobie płaszczyzny $0xy$ otoczonej w nieskończoności przez – także nieskończone – „morze”, stanowiące odpowiednik dwóch nieskończoności z pojedynczej osi liczbowej.⁴

⁴Pojęcie nieskończoności na płaszczyźnie zespolonej wprowadza się formalnie przy pomocy koncepcji tzw. rzutu stereograficznego. Wykracza to jednak poza ramy tego wykładu. Spotkasz się z nim przy nauce analizy funkcji zmiennej zespolonej.

5.2.2 Wzór de Moivre'a; liczby zespolone i wzory trygonometryczne

Podnosząc wzór 5-13 stronami do potęgi stopnia n otrzymujemy:

$$z^n = [r(\cos \phi + i \sin \phi)]^n = r^n e^{in\phi} = r^n (\cos n\phi + i \sin n\phi) \quad (5-20)$$

i w konsekwencji

$$e^{in\phi} = (\cos \phi + i \sin \phi)^n = \cos n\phi + i \sin n\phi. \quad (5-21)$$

Wzór ten nazywa się zwykle wzorem de Moivre'a lub wzorem Eulera. Wynika z niego, jak i z 5-12, że dobrze nam znane funkcje sinus i kosinus – także i wielokrotności kątów – mogą być przedstawiane jako części funkcji wykładniczej o wykładniku urojonym. Stwarza to znakomite możliwości praktyczne. Być możesz pamiętasz czytelniku, ile mozołu kosztowało Cię kiedyś wyprowadzenie wzorów na sinus i kosinus sumy (różnicy) kątów. Przy pomocy algebry liczb zespolonych można je uzyskać właściwie natychmiast:

$$\begin{aligned} e^{i(\alpha \pm \beta)} &= \cos(\alpha \pm \beta) + i \sin(\alpha \pm \beta) \\ &= e^{i\alpha} \cdot e^{\pm i\beta} = (\cos \alpha + i \sin \alpha) \cdot (\cos \beta \pm i \sin \beta) \\ &= (\cos \alpha \cos \beta \mp \sin \alpha \sin \beta) + i(\sin \alpha \cos \beta \pm \cos \alpha \sin \beta). \end{aligned}$$

Wystarczy porównać części rzeczywiste i urojone w pierwszym i trzecim wierszu, aby otrzymać dobrze znane

$$\begin{aligned} \cos(\alpha \pm \beta) &= \cos \alpha \cos \beta \mp \sin \alpha \sin \beta \\ \sin(\alpha \pm \beta) &= \sin \alpha \cos \beta \pm \cos \alpha \sin \beta. \end{aligned}$$

Tego typu wzorów można – w równie prosty sposób – wykazać wiele (por. przykłady i ćwiczenia). Liczby zespolone są – między innymi – bardzo skutecznym narzędziem w rozwiązywaniu problemów trygonometrycznych. Ale nie tylko.

5.3 Potęga i pierwiastek liczby zespolonej

Wprowadzenie pojęcia *zmiennnej zespolonej* – w ogólności zbioru punktów płaszczyzny zespolonej – nasuwa automatycznie pytanie o *funkcje zmiennnej zespolonej*. Oczywiście naturalnym jest oczekiwać, że będziemy mieli do czynienia z wyrażeniami typu

$$f(z) = f(x + iy) \equiv f(x, y)$$

przyporządkowaniem konkretnym wartości z pewnej wartości (lub kilku wartości?) będących w ogólnym przypadku wartościami zespolonymi, to znaczy składającymi się z części rzeczywistej i urojonej. Tak jest w rzeczywistości; co więcej większość „przepisów funkcyjnych” z osi rzeczywistej, określających $f = f(x)$ ulega dość intuicyjnym „rozszerzeniom”⁵ na płaszczyznę zespoloną. I tak znany nam wzór na „jedynekę trygonometryczną”:

$$\sin^2 x + \cos^2 x = 1$$

wygląda dokładnie tak samo w języku zmiennnej zespolonej:

$$\sin^2 z + \cos^2 z = 1.$$

Sprawy ulegają jednak znacznemu skomplikowaniu w momencie kiedy wchodzimy w obszar rachunku różniczkowego i całkowego. Nie będziemy tutaj o tym mówić; zasygnalizujemy tylko jedną oczywistą „komplikację”. Warunkiem istnienia pochodnej funkcji zmiennej rzeczywistej $f(x)$ w punkcie x_0 jest równość lewo- i prawostronnej granicy ilorazu różnicowego przy $x \rightarrow x_0$. Na płaszczyźnie zespolonej zmierzanie do punktu z_0 może odbywać się na nieskończenie wiele sposobów, wszystkie promieniście zbiegające się w tym punkcie drogi są równouprawnione! Rachunek różniczkowy i całkowy funkcji zmiennnej zespolonej stanowi z jednej strony bogaty w treści i pasjonujący rozdział matematyki, a z drugiej – dostarcza wielu podstawowych narzędzi matematycznych, na przykład metod obliczania całek, sum, a także ... rozwiązywania pewnych, bardzo podstawowych równań fizyki matematycznej.

⁵Proces ten formalnie *nazywa się* rozszerzeniem analitycznym.

W tym momencie ograniczymy się jedynie do bardzo krótkiej dyskusji dwóch funkcji zmiennej zespolonej: funkcji potęgowej, typu

$$f(z) = z^n; \quad z - \text{liczba całkowita}, \quad (5-22)$$

oraz funkcji odwrotnej

$$w(z) = \sqrt[n]{z}. \quad (5-23)$$

Nie ma żadnych problemów z obliczaniem potęgi całkowitej zmiennej z . Wszystko jedno czy korzystamy przy tym z określenia liczby zespolonej w postaci $z = x + iy$, czy reprezentacji biegunowej $z = r \exp(i\theta)$. W pierwszym przypadku wystarczy zastosować wzór Newtona na potęgę dwumianu:

$$z^n = (x + iy)^n = \sum_{k=0}^n \binom{n}{k} x^k (iy)^{n-k} \quad (5-24)$$

i pamiętać, że $i^2 = -1, i^3 = i, i^4 = 1$ itd. W przypadku reprezentacji biegunowej wykorzystujemy wzór de Moivre'a:

$$z^n = (re^{i\theta})^n = r^n e^{in\theta}; \quad (5-25)$$

podnosimy do n -tej potęgi moduł liczby zespolonej, a jej argument zwiększamy n -krotnie.

Pierwiastki zespolone to sprawa nieco bardziej skomplikowana. Rozważmy na początek przypadek pierwiastka kwadratowego:

$$w(z) = \sqrt{z}.$$

Założmy, że taka liczba \sqrt{z} istnieje i jest równa $w = \sqrt{x + iy} \equiv u + iv$. Mamy więc

$$\sqrt{x + iy} = u + iv, \quad (5-26)$$

albo

$$(u + iv)^2 = x + iy \quad (5-27)$$

z czego wynika

$$\begin{aligned} u^2 - v^2 &= x, \\ 2uv &= y. \end{aligned} \quad (5-28)$$

Podnosząc do kwadratu oba równania układu 5-28 i dodając je stronami mamy

$$(u^2 - v^2)^2 + 4u^2v^2 = x^2 + y^2$$

albo

$$u^2 + v^2 = +\sqrt{x^2 + y^2}.$$

Zauważmy, że przed pierwiastkiem mamy znak '+' – suma kwadratów dwóch liczb rzeczywistych musi być nieujemna. Ostatnie równanie, w połączeniu z pierwszym równaniem układu 5-28 daje w końcu

$$\begin{aligned} u^2 &= \frac{1}{2}(x + \sqrt{x^2 + y^2}), \\ v^2 &= \frac{1}{2}(-x + \sqrt{x^2 + y^2}). \end{aligned} \quad (5-29)$$

Pozostaje obliczyć pierwiastki, aby otrzymać po dwie – różniące się znakami – wartości u i v . Aby z tak otrzymanych części: rzeczywistej i urojonej zbudować pierwiastek musimy mieć jednak na uwadze drugie równanie 5-28: znak iloczynu uv musi być taki sam jak znak y . A więc nie cztery, ale dwie możliwe kombinacje u i v dają nam szukane w_1 i w_2 . Na przykład: dla $z = 21 - 20i$ mamy $\sqrt{x^2 + y^2} = \sqrt{21^2 + 20^2} = 29$. Korzystając z 5-29 otrzymujemy $u = \pm 5, v = \pm 2$. Ponieważ $y < 0$ to, znaki u i v muszą być różne. Mamy

$$w_1 = 5 - 2i, \quad \text{oraz} \quad w_2 = -5 + 2i = -w_1.$$

Oba pierwiastki (liczby zespolone!) różnią się znakiem.

Obliczanie pierwiastków wyższego stopnia jest praktycznie możliwe tylko przy skorzystaniu z reprezentacji biegunowej liczby zespolonej $z = re^{i\theta}$. Jeżeli $w = \sqrt[n]{z} \equiv \rho e^{i\psi}$, to (wzór de Moivre'a) $\rho^n = r$, albo $\rho = \sqrt[n]{r}$.

Jeżeli chodzi o związki pomiędzy argumentami to mamy analogicznie $n\psi = \theta$, ale – bardziej dokładnie – $n\psi = \theta + 2k\pi$, ponieważ różnica pomiędzy tymi dwoma kątami typu $2k\pi$ (k całkowite) jest nieistotna. Tak więc

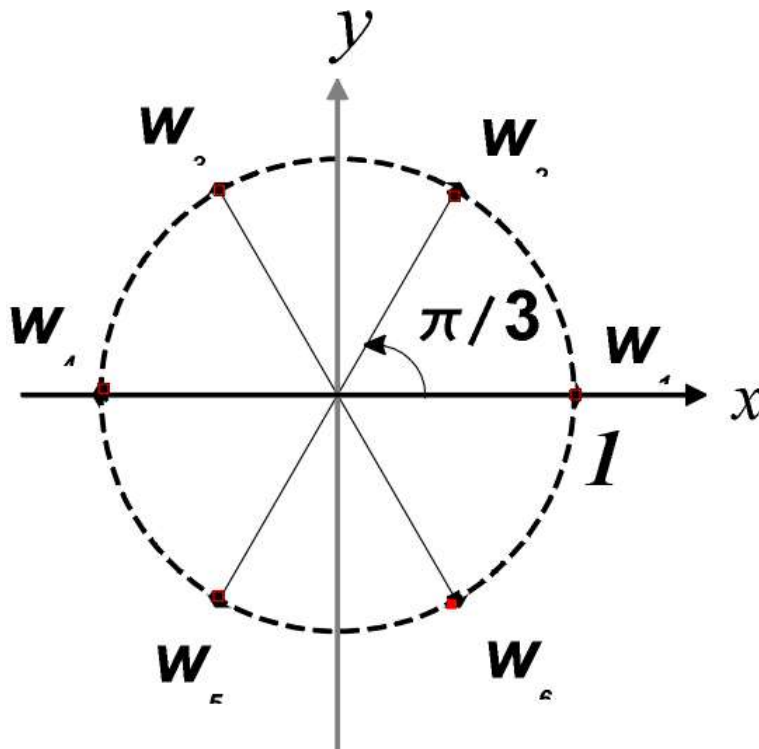
$$\psi = \frac{\theta + 2k\pi}{n}. \quad (5-30)$$

Dla $k = 0, 1, 2, \dots, n - 1$ powyższy wzór daje n różnych wartości argumentu: począwszy od n -tej frakcji argumentu θ aż do kąta $-\theta/n$; kolejne wartości argumentu dzieli kąt $2\pi/n$.

Na przykład $w = \sqrt[3]{-8} \equiv \sqrt[3]{8e^{i\pi}}$. Ponieważ $n = 3$, $k = 0, 1, 2$ i kolejne wartości w_k to:

$$\begin{aligned} w_0 &= \sqrt[3]{8e^{i\pi/3}} = 2(\cos \pi/3 + i \sin \pi/3) = 1 + i\sqrt{3}, \\ w_1 &= \sqrt[3]{8e^{i(\pi/3+2\pi/3)}} = \dots = -2, \\ w_2 &= \sqrt[3]{8e^{i(\pi/3+2\cdot 2\pi/3)}} = \dots = 1 - i\sqrt{3}. \end{aligned}$$

Wszystkie n pierwiastków liczby zespolonej $z = re^{i\theta}$ leży w n wierzchołkach wielokąta foremnego, wpisanego w okrąg o promieniu $\sqrt[n]{r}$. Orientacja wielokąta jest taka, że argument „pierwszego” pierwiastka w_0 (zwanego także wartością główną pierwiastka) to n -ta frakcja kąta θ . W szczególności, wyciągając pierwiastek z dodatniej liczby rzeczywistej ($\theta = 0$) pierwszy pierwiastek (wierzchołek wielokąta) też znajduje się na osi rzeczywistej – por. Rys. 5.6.



Rysunek 5.6: Pierwiastek zespolony: $\sqrt[6]{1}$.

5.4 Obrót wektora na płaszczyźnie

W poprzednim rozdziale, przy pomocy rozważań trygonometrycznych (por. rysunek 4.8) wyprowadziliśmy wzory wiążące ze sobą współrzędne wektora \mathbf{r} :

$$\begin{aligned} x' &= x \cos \theta + y \sin \theta \\ y' &= -x \sin \theta + y \cos \theta. \end{aligned} \quad (5-31)$$

Wzory 5-31 postanowiliśmy traktować jako roboczą definicję wektora: jako wielkości, której współrzędne transformują się przy zmianie (obrocie) Σ właśnie tak jak to z tych wzorów wynika. Zauważmy, że wzory te moglibyśmy bardzo łatwo uzyskać przy pomocy liczb zespolonych. Wektor \mathbf{r} w układzie Σ to

$$\mathbf{r} = x + iy.$$

Układ Σ' powstaje przez obrót o kąt θ . Zamiast obracać układ, obróćmy wektor – o kąt $-\theta$. Obrotowi (o kąt $-\theta$) na płaszczyźnie zespolonej odpowiada mnożenie przez $\exp(-i\theta)$. Nowy wektor to:

$$\begin{aligned}e^{-i\theta} \times \mathbf{r} &= (\cos \theta - i \sin \theta) \times (x + iy) \\ &= x \cos \theta + y \sin \theta + i(-x \sin \theta + y \cos \theta) \equiv x' + iy'\end{aligned}$$

— zamiast rozwiązywać trójkąty z rysunku 4.8 wystarczy przemnożyć dwie liczby zespolone!

Rozdział 6

Przestrzenie wektorowe

6.1 Wprowadzenie

Dwuwymiarowa płaszczyzna – zbiór punktów opisanych dwójką współrzędnych (x, y) , lub (x_1, x_2) , oraz 3-wymiarowa przestrzeń – punkty (x, y, z) , lub (x_1, x_2, x_3) , to najprostsze przykłady *przestrzeni wektorowej*. W geometrii i w fizyce spotykamy jednak często obiekty, których opis wymaga określenia więcej niż trzech wielkości liczbowych. Na przykład, rozważmy układ (rodzinę) sfer w przestrzeni 3-wymiarowej. Każda z nich będzie określona przez podanie trzech współrzędnych jej środka oraz promienia – w sumie czterech liczb. Przestrzeń utworzona przez sfery jest 4-wymiarowa. Inny przykład: zastanówmy się, ile współrzędnych jest potrzebnych do określenia położenia bryły sztywnej w przestrzeni 3-wymiarowej? Położenie środka masy ciała to 3 liczby (jego współrzędne). Do tego trzeba określić kierunek wybranej osi, przechodzącej przez środek masy. Potrafimy to zrobić podając dwie liczby – kosinusy kątów, jakie tworzy ta oś z dwoma (z trzech) osiami kartezjańskiego układu współrzędnych. Wystarczą dwa kąty, bo z faktu, że oś przechodzi przez określony punkt w przestrzeni (środek masy ciała) można wyliczyć trzeci kąt (i jego kosinus). Ostatnia – szósta – liczba, to kąt obrotu względem tej osi. Zbiór wszystkich brył sztywnych, zanurzonych w 3-wymiarowej przestrzeni, tworzy więc przestrzeń 6-wymiarową.

Te nieco abstrakcyjne (przynajmniej dla $n > 3$) struktury stanowią potężne narzędzie matematyczne w rozwiązywaniu wielu problemów fizyki. Używając przestrzeni wektorowej staramy się, aby każdy układ fizyczny można było przedstawić jako element tej przestrzeni, tak zwany *wektor stanu*. W takim wektorze stanu powinna tkwić cała informacja o układzie, chociaż sam wektor nie musi reprezentować określonego „bytu” fizycznego. Na przykład, funkcja ψ w mechanice kwantowej nie ma bezpośredniej interpretacji fizycznej; tą ostatnią posiada natomiast kwadrat jej modułu $|\psi|^2$ – reprezentujący gęstość prawdopodobieństwa znalezienia obiektu kwantowego w sąsiedztwie danego punktu przestrzeni.¹

Taki *wektor stanu* może być (i często jest, jak wynika z powyższej uwagi) pewną wielkością abstrakcyjną, którą matematycznie możemy wyrażać na wiele (nieskończenie wiele) różnych, ale równoważnych sposobów. Konkretny sposób wyrażenia, albo *reprezentacji* wektora stanu zależy od wyboru układu współrzędnych. Ten sam wektor pola elektrycznego będzie miał różne współrzędne (a więc różne reprezentacje) w dwóch różnych układach współrzędnych, z których jeden jest, na przykład, obrocony względem drugiego (por. punkt 4.5). W takich, ogólnie rzecz biorąc, n -wymiarowych² strukturach – uporządkowanych układach (rodzinach) liczb rzeczywistych, reprezentujących *współrzędne punktu przestrzeni wektorowej*, musimy zdefiniować pewne „reguły gry”. Poprzez analogie z rozważanymi już operacjami na wektorach i ich współrzędnymi na płaszczyźnie ($n = 2$) i w „zwykłej” przestrzeni ($n = 3$) *wszystkie* takie układy nazywać będziemy *przestrzeniami wektorowymi*. Formalnie przestrzeń wektorowa \mathbb{R}^n , będzie to pewna struktura algebraiczna, zawierająca w sobie wszystkie własności jej elementów – wektorów \mathbf{V} , o współrzędnych V_i , $i = 1, \dots, n$:

$$\mathbf{V} = (V_1, V_2, \dots, V_k, \dots, V_n). \quad (6-1)$$

Dwa wektory \mathbf{V} i \mathbf{W} możemy dodawać, lub odejmować:

$$\mathbf{V} + \mathbf{W} = (V_1 + W_1, V_2 + W_2, \dots, V_k + W_k, \dots, V_n + W_n), \quad (6-2)$$

¹Przestrzenie wektorów stanów w mechanice kwantowej są przestrzeniami o nieskończonej liczbie wymiarów. Niema to jednak wpływu na sens powyższej uwagi.

²Nie definiujemy jeszcze w tym miejscu *formalnie wymiaru* przestrzeni wektorowej; mówiąc o n -wymiarowych strukturach mamy na myśli tylko liczbę współrzędnych wektora. Definicja wymiaru przestrzeni pojawi się dalej.

$$\mathbf{V} - \mathbf{W} = (V_1 - W_1, V_2 - W_2, \dots, V_k - W_k, \dots, V_n - W_n). \quad (6-3)$$

Z powyższych wzorów wynika, że odejmowanie wektora \mathbf{W} to dodawanie wektora „ujemnego” (o zwrocie przeciwnym) $-\mathbf{W}$; analogicznie do 6-1:

$$-\mathbf{W} = (-W_1, -W_2, \dots, -W_k, \dots, -W_n). \quad (6-4)$$

Dwa wektory \mathbf{V} i \mathbf{W} są sobie równe wtedy i tylko wtedy, jeżeli

$$V_i = W_i, \quad \text{dla każdego } i = 1, \dots, n, \quad (6-5)$$

a ich różnicą jest wtedy wektor

$$\mathbf{V} - \mathbf{W} = (V_1 - W_1, V_2 - W_2, \dots, V_k - W_k, \dots, V_n - W_n) = (0, 0, \dots, 0, \dots, 0), \quad (6-6)$$

tak zwany wektor zerowy.

Tak jak w przestrzeniach 2- i 3-wymiarowych wektor \mathbf{V} można mnożyć przez liczbę (skalar):

$$\alpha \mathbf{V} = (\alpha V_1, \alpha V_2, \dots, \alpha V_k, \dots, \alpha V_n). \quad (6-7)$$

Wszystkie współrzędne wektora ulegają przemnożeniu przez α .

Tak jak powiedzieliśmy już – układ wszystkich wektorów (o n współrzędnych), rozpatrywany w kontekście określonych powyżej operacji dodawania i mnożenia przez liczbę nazywamy przestrzenią wektorową.

Określenie „reguł gry” w przestrzeni wektorowej – operacji dodawania wektorów i mnożenia ich przez skalar nie musi oznaczać, że z operacji tych wynika zawsze pewien określony pożytek. Jeżeli na przykład w 4-wymiarowej przestrzeni dwa wektory:

$$\begin{aligned} \text{Wektor I:} & \quad (x_1, y_1, z_1, R_1) \\ \text{Rodzina II:} & \quad (x_2, y_2, z_2, R_2) \end{aligned}$$

oznaczają odpowiednio dwie sfery: środek pierwszej z nich jest w punkcie (x_1, y_1, z_1) , a sfera ma promień R_1 ; środek drugiej (o promieniu R_2) – w punkcie (x_2, y_2, z_2) , to w pierwszym momencie nie widać pożytku z dodania takich dwóch wektorów. Zauważmy jednak, że można na przykład traktować rodzinę sfer koncentrycznych, o wspólnym środku w punkcie przestrzeni (3-wymiarowej) (x, y, z) i o dowolnym promieniu R_i , jako sumy wektora

$$(x, y, z, 0) \quad \text{i wektorów } (0, 0, 0, R_i).$$

Podobnie, w 6-wymiarowej przestrzeni położenia (3 pierwsze współrzędne) i pędu (kolejne 3 współrzędne) cząstki o masie m kombinacje liniowe typu:

$$(x, y, z, 0, 0, 0) + m \cdot \begin{cases} (0, 0, 0, v_x^{(1)}, v_y^{(1)}, v_z^{(1)}) \\ (0, 0, 0, v_x^{(2)}, v_y^{(2)}, v_z^{(2)}) \\ \dots \quad \dots \quad \dots \\ (0, 0, 0, v_x^{(k)}, v_y^{(k)}, v_z^{(k)}) \\ \dots \quad \dots \quad \dots \end{cases}$$

to wszystkie możliwe stany cząstki, znajdującej się w punkcie o współrzędnych (x, y, z) i mającej trzy składowe pędu równe: $mv_x^{(k)}, mv_y^{(k)}, mv_z^{(k)}$ (znowu mamy do czynienia z nieskończoną liczbą możliwości – wskaźnik (k) , „numerujący” różne stany pędu zmienia się w sposób nieograniczony).

6.1.1 Formy liniowe

Jedną z możliwych sytuacji, kiedy mamy także do czynienia z wektorami – a dokładniej: wielkościami wykazującymi identyczne cechy formalne jak wektory – jest liniowe równanie typu

$$a_1 x_1 + a_2 x_2 + \dots + a_k x_k + \dots + a_n x_n = \begin{cases} 0 \text{ lub} \\ \text{stała.} \end{cases} \quad (6-8)$$

Występujące po lewej stronie wielkości x_1, \dots, x_n to (szukane) niewiadome, natomiast współczynniki a_1, \dots, a_n można traktować jako współrzędne wektora \mathbf{a} . Lewą stronę 6-8

$$f(x) \equiv a_1 x_1 + a_2 x_2 + \dots + a_k x_k + \dots + a_n x_n, \quad (6-9)$$

nazywamy *formą liniową* zmiennych x_1, \dots, x_n ; jest ona jednoznacznie określona przez wektor \mathbf{a} . I odwrotnie: każdemu wektorowi \mathbf{a} , $\mathbf{a} \equiv (a_1, \dots, a_n)$ odpowiada pewna forma liniowa.

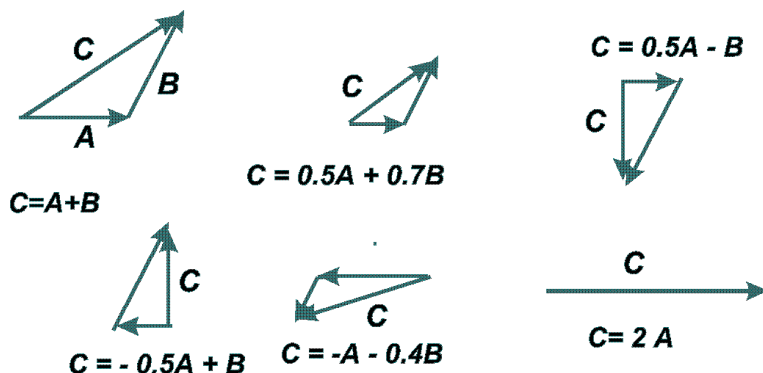
Takimi formami liniowymi są lewe strony wszystkich równań, o których traktuje rozdział trzeci.

6.2 Wektory liniowo zależne

Na rysunku 6.1 pokazany jest schemat dodawania wektorów na płaszczyźnie – wektor C jest sumą wektorów A i B ; podobnie możemy powiedzieć, że $B = C - A$, lub $A = C - B$.

Założmy, że mamy dwa wektory A i B – dowolne, ale *nie współliniowe*. Współliniowość wektorów oznacza, że jeden z nich, np. B , może być zapisany jako $B = \alpha A$. Wówczas *dowolny* wektor C może być zapisany w postaci *liniowej kombinacji*, wektorów A i B – ich sumy z pewnymi współczynnikami liczbowymi a i b :

$$C = aA + bB. \quad (6-10)$$



Rysunek 6.1: Wektor C jako różne liniowe kombinacje niewspółliniowych wektorów A i B .

Kilka prostych przykładów, ilustrujących zasadność tego stwierdzenia można zobaczyć na Rys. 6.1. Zauważ, że w niektórych z nich **nie** występują oba składniki sumy.

Analogicznie możemy wprowadzić pojęcie *liniowej kombinacji* r wektorów ($r \geq 2$):

$$W = a_1 V_1 + a_2 V_2 + \dots + a_r V_r. \quad (6-11)$$

Zwróć uwagę, że w tym wzorze V_i to i -ty wektor (o n współrzędnych!), a nie i -ta współrzędna wektora V – rozpisanie powyższego wzoru na współrzędne zmusza nas do użycia podwójnego wskaźnika:

$$W_k \equiv (W)_k = a_1 (V_1)_k + a_2 (V_2)_k + \dots + a_r (V_r)_k. \quad (6-12)$$

Układ r wektorów nazywa się *układem liniowo zależnym* jeżeli przynajmniej jeden z nich możemy wyrazić w postaci liniowej kombinacji pozostałych; w przeciwnym przypadku mamy do czynienia z układem wektorów liniowo niezależnych. Szczególnym przypadkiem liniowej zależności są dwa wektory współliniowe na płaszczyźnie – $B = \alpha A$.

Przyjętym powszechnie sposobem mówienie o liniowej (nie)zależności jest następujące:

Twierdzenie 6.1 *Układ r wektorów (por. 6-11) jest liniowo zależny wtedy i tylko wtedy gdy równanie*

$$k_1 V_1 + k_2 V_2 + \dots + k_r V_r = 0 \quad (6-13)$$

jest spełnione dla przynajmniej jednej z liczb k_1, k_2, \dots, k_n różnej od zera.

Oba sformułowania są oczywiście równoważne. Jeżeli np. wektor V_r może być wyrażony jako kombinacja liniowa

$$V_r = l_1 V_1 + l_2 V_2 + \dots + l_{r-1} V_{r-1}, \quad (6-14)$$

to mamy równanie

$$l_1 V_1 + l_2 V_2 + \dots + l_{r-1} V_{r-1} - V_r = 0, \quad (6-15)$$

a więc równanie 6-13, w którym na pewno jeden współczynnik jest różny od zera ($a_r = -1$). Z kolei zachodzenie równości 6-13, w której jeden współczynnik jest różny od zera – na przykład $k_r \neq 0$, pozwala przepisać to równanie w postaci:

$$V_r = \left(-\frac{k_1}{k_r}\right) V_1 + \left(-\frac{k_2}{k_r}\right) V_2 + \dots + \left(-\frac{k_{r-1}}{k_r}\right) V_{r-1}, \quad (6-16)$$

Dowód: Przypuśćmy, że nasze s wektorów to

$$\begin{aligned}\mathbf{a}_1 &\equiv (a_{11}, a_{12}, \dots, a_{1n}) \\ \mathbf{a}_2 &\equiv (a_{21}, a_{22}, \dots, a_{2n}) \\ &\vdots \\ \mathbf{a}_s &\equiv (a_{s1}, a_{s2}, \dots, a_{sn})\end{aligned}$$

(pierwszy wskaźnik to wskaźnik wektora, drugi – współrzędnej). Test liniowej zależności, równanie

$$k_1\mathbf{a}_1 + k_2\mathbf{a}_2 + \dots + k_s\mathbf{a}_s = 0 \quad (6-22)$$

rozpisujemy dla poszczególnych współrzędnych wektorów:

$$\begin{aligned}a_{11}k_1 + a_{21}k_2 + \dots + a_{s1}k_s &= 0, \\ a_{12}k_1 + a_{22}k_2 + \dots + a_{s2}k_s &= 0, \\ &\vdots \\ a_{1n}k_1 + a_{2n}k_2 + \dots + a_{sn}k_s &= 0.\end{aligned}$$

Jest to układ n równań na s niewiadomych: k_1, k_2, \dots, k_s . Dla $n < s$ taki układ (por. przykład w punkcie 3.2; będziemy o tym mówić zresztą jeszcze w tym rozdziale) zawsze nietrywialne rozwiązanie; istnieją więc k_1, k_2, \dots, k_s nie wszystkie równe zeru, układ jest liniowo zależny.

Możemy też mówić o bazie przestrzeni wektorowej, jako o układzie n liniowo niezależnych wektorów $\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_n$, który po dołączeniu do układu dowolnego wektora \mathbf{W} staje się układem liniowo zależnym: wektor \mathbf{W} może zostać wyrażony jako liniowa kombinacja wektorów $\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_n$. (Czasem zamiast pojęcia bazy używa się pojęcia maksymalnego – tzn. o maksymalnej liczebności – układu wektorów liniowo niezależnych.) Rozważając dwa układy wektorów: $\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_r$ i $\mathbf{W}_1, \mathbf{W}_2, \dots, \mathbf{W}_s$ możemy powiedzieć, że układ $\{\mathbf{W}\}$ wyraża się liniowo przez elementy układu $\{\mathbf{V}\}$ jeżeli każdy wektor $\mathbf{W}_i, i = 1, \dots, s$ może być przedstawiony jako liniowa kombinacja wektorów $\{\mathbf{V}\}$. To pojęcie jest przechodnie: jeżeli rozważyć np. trzy rodziny (układy) wektorów

$$\mathbf{a}_r, \quad \mathbf{b}_s, \quad \mathbf{c}_t,$$

i każdy wektor \mathbf{c}_i wyraża się jako liniowa kombinacja układu $\{\mathbf{b}\}$, a każdy wektor \mathbf{b}_k wyraża się jako liniowa kombinacja układu $\{\mathbf{a}\}$ to każdy wektor \mathbf{c}_i może być wyrażony jako liniowa kombinacja układu $\{\mathbf{a}\}$. Rzeczywiście, mamy

$$\mathbf{c}_i = C_{i1}\mathbf{b}_1 + \dots + C_{is}\mathbf{b}_s = \sum_{k=1}^s C_{ik}\mathbf{b}_k,$$

(C_{ik} – współczynniki liczbowe kombinacji liniowej); z kolei

$$\mathbf{b}_k = \sum_{m=1}^r B_{km}\mathbf{a}_m,$$

a więc

$$\mathbf{c}_i = \sum_{k=1}^s C_{ik} \left(\sum_{m=1}^r B_{km}\mathbf{a}_m \right) = \sum_{m=1}^r \left(\sum_{k=1}^s C_{ik}B_{km} \right) \mathbf{a}_m \equiv \sum_{m=1}^r A_{im}\mathbf{a}_m,$$

gdzie współczynniki liniowej kombinacji, A_{im} to sumy:

$$A_{im} = \sum_{k=1}^s C_{ik}B_{km}.$$

Mówimy, że dwie rodziny (układy) wektorów są równoważne, jeżeli każdy wektor należący do jednej z nich można wyrazić jako kombinację liniową wektorów drugiej rodziny. Jeżeli dwie rodziny są równoważne i jeżeli pewien wektor (element przestrzeni wektorowej) można wyrazić jako kombinację liniową wektorów jednej rodziny, to będzie on także kombinacją liniową wektorów drugiej rodziny.

Mamy także oczywiste

Twierdzenie 6.4 Jeżeli w n -wymiarowej przestrzeni wektorowej mamy dwie rodziny wektorów:

$$\begin{aligned} \text{Rodzina I:} & \quad (\xi_1, \xi_2, \dots, \xi_r), \\ \text{Rodzina II:} & \quad (\eta_1, \eta_2, \dots, \eta_s) \end{aligned}$$

i rodzina I jest liniowo niezależna, a jej wektory wyrażają się poprzez kombinacje liniowe drugiej, to liczba wektorów w pierwszej rodzinie nie może być większa od liczby wektorów w drugiej: $r \leq s$.

Z ostatniego twierdzenia wynika też, że dwie równoważne rodziny liniowo niezależnych wektorów zawierają te same liczby składników.

Każde dwie bazy przestrzeni wektorowej są rodzinami równoważnymi, a więc muszą zawierać tę samą liczbę składników. Ponieważ wykazaliśmy, że dla $s > n$ układ wektorów musi być liniowo zależny (nie może być bazą), a z drugiej strony zidentyfikowaliśmy w \mathbb{R}^n n -elementową bazę 6-20 wynika z tego, że wszystkie bazy muszą składać się z n wektorów.

6.3 Podprzestrzeń wektorowa; baza

Każdy podzbiór przestrzeni wektorowej \mathbb{R}^n , będący sam przestrzenią wektorową – to znaczy zamknięty (patrz niżej) w stosunku do operacji dodawania wektorów i mnożenia ich przez liczbę – nazywamy *podprzestrzenią wektorową* \mathcal{L} . I tak, w „zwykłej” trójwymiarowej przestrzeni, możemy wyróżnić podprzestrzeń składającą się z wektorów zaczepionych w początku układu współrzędnych i leżących na jednej płaszczyźnie (lub prostej), przechodzącej przez początek układu współrzędnych. Warunki zamknięcia w stosunku do operacji dodawania wektorów i mnożenia ich przez liczbę to:

1. Jeżeli wektory \mathbf{V}_1 i \mathbf{V}_2 należą do \mathcal{L} , to ich suma, wektor $\mathbf{V}_1 + \mathbf{V}_2$ także należy do \mathcal{L} ;
2. Jeżeli wektor \mathbf{V} należy do \mathcal{L} , to wektor $\alpha\mathbf{V}$ (α – dowolna liczba) także należy do \mathcal{L} .

Z drugiego warunku wynika, że do (każdej !) podprzestrzeni \mathcal{L} należy wektor zerowy ($\alpha = 0$), a także wektor przeciwny do każdego wektora \mathbf{V} ($\alpha = -1$). A jeżeli tak, to na mocy warunku pierwszego, do \mathcal{L} należą także różnice dwóch dowolnych jej elementów (wektorów).

Szczególnymi przypadkami podprzestrzeni \mathbb{R}^n będzie sama przestrzeń \mathbb{R}^n oraz wektor zerowy (tzw. podprzestrzeń zerowa). Ciekawszy przykład podprzestrzeni otrzymamy biorąc skończony układ wektorów należących do \mathbb{R}^n

$$\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r \quad r - \text{dowolne} \quad (6-23)$$

i określając \mathcal{L} jako zbiór wszystkich wektorów dających się przedstawić jako liniowe kombinacje 6-23. Dwa takie wektory, \mathbf{b} i \mathbf{c} , to

$$\mathbf{b} = \beta_1\mathbf{a}_1 + \beta_2\mathbf{a}_2 + \dots + \beta_r\mathbf{a}_r, \quad \text{i} \quad \mathbf{c} = \gamma_1\mathbf{a}_1 + \gamma_2\mathbf{a}_2 + \dots + \gamma_r\mathbf{a}_r. \quad (6-24)$$

Suma wektorów

$$\mathbf{b} + \mathbf{c} = (\beta_1 + \gamma_1)\mathbf{a}_1 + (\beta_2 + \gamma_2)\mathbf{a}_2 + \dots + (\beta_r + \gamma_r)\mathbf{a}_r,$$

a także iloczyny ich przez liczbę:

$$\begin{aligned} \alpha\mathbf{b} &= \alpha\beta_1\mathbf{a}_1 + \alpha\beta_2\mathbf{a}_2 + \dots + \alpha\beta_r\mathbf{a}_r, \\ \alpha\mathbf{c} &= \alpha\gamma_1\mathbf{a}_1 + \alpha\gamma_2\mathbf{a}_2 + \dots + \alpha\gamma_r\mathbf{a}_r, \end{aligned}$$

należą do przestrzeni \mathcal{L} .

Mówimy, że dana podprzestrzeń \mathcal{L} jest *generowana* przez układ wektorów 6-23, albo że układ (rodzina) wektorów generuje (albo *napina*) przestrzeń \mathcal{L} . Oczywiście jest, że każda liniowa podprzestrzeń \mathcal{L} przestrzeni \mathbb{R}^n jest generowana przez pewną skończoną rodzinę wektorów i – za wyjątkiem podprzestrzeni zerowej – każda podprzestrzeń posiada pewną bazę. Zauważmy, że wprowadzając układ generujący 6-23 nie żądaliśmy liniowej niezależności wektorów \mathbf{a}_i . Formalnie wymóg taki jest zbyteczny, ale praktycznie wybieramy zawsze układy generujące jako układy liniowo niezależne. Nawiązując do przytoczonego wyżej przykładu – każde dwa niewspółliniowe wektory leżące w płaszczyźnie Σ mogą być rodziną generującą (napinającą) tę płaszczyznę. Do takich dwóch wektorów mogą dodać dowolną liczbę wektorów leżących na Σ – ale ponieważ każdy z nich

może być przedstawiony jako liniowa kombinacja dwóch pierwszych takie powiększanie rodziny generującej nie wnosi nic ciekawego (nowego).

Składająca się z liniowo niezależnych wektorów rodzina generująca (pod)przestrzeni jest bazą tej (pod)przestrzeni. Uzyskujemy nową (równoważną) definicję bazy (pod)przestrzeni wektorowej: *baza to rodzina generująca o minimalnej liczbie składników (wektorów)*.

Wymiar dowolnej podprzestrzeni \mathbb{L} przestrzeni wektorowej \mathbb{R}^n jest określony przez liczbę wektorów bazowych tej podprzestrzeni i nie może być większy od n (dla podprzestrzeni zerowej wymiar jest więc zero.) Jeżeli mamy dwie podprzestrzenie \mathbb{L}_1 i \mathbb{L}_2 to zbiór \mathbb{L}_0 wektorów, które należą do \mathbb{L}_1 i do \mathbb{L}_2 nazywamy iloczynem, albo przecięciem tych dwóch podprzestrzeni. Możemy mieć także sumę \mathbb{L}_+ podprzestrzeni \mathbb{L}_1 i \mathbb{L}_2 – zbiór wektorów, które należą do \mathbb{L}_1 albo do \mathbb{L}_2 . Oznaczając wymiary podprzestrzeni \mathbb{L}_1 , \mathbb{L}_2 , \mathbb{L}_0 i \mathbb{L}_+ odpowiednio przez d_1 , d_2 , d_0 i d_+ mamy:

$$d_+ = d_1 + d_2 - d_0. \quad (6-25)$$

Powyższy wzór wynika z ogólnej teorii zbiorów, ale możemy go łatwo udowodnić. Oznaczmy bazę podprzestrzeni \mathbb{L}_0 przez

$$\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{d_0} \quad (6-26)$$

i uzupełnijmy ją, tak aby mogła tworzyć bazę dla \mathbb{L}_1 :

$$\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{d_0}, \mathbf{b}_{d_0+1}, \dots, \mathbf{b}_{d_1} \quad (6-27)$$

i \mathbb{L}_2

$$\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{d_0}, \mathbf{c}_{d_0+1}, \dots, \mathbf{c}_{d_2}. \quad (6-28)$$

Rodzina generująca dla podprzestrzeni \mathbb{L}_+ będą wszystkie wektory 6-26, 6-27 i 6-28:

$$\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{d_0}, \mathbf{b}_{d_0+1}, \dots, \mathbf{b}_{d_1}, \mathbf{c}_{d_0+1}, \dots, \mathbf{c}_{d_2}. \quad (6-29)$$

Jest ich rzeczywiście $d_+ = d_1 + d_2 - d_0$; pozostaje udowodnić, że są one (baza!) liniowo niezależne. W tym celu rozważmy równanie

$$\begin{aligned} \alpha_1 \mathbf{a}_1 + \dots + \alpha_{d_0} \mathbf{a}_{d_0} + \beta_{d_0+1} \mathbf{b}_{d_0+1} + \dots + \beta_{d_1} \mathbf{b}_{d_1} \\ + \gamma_{d_0+1} \mathbf{c}_{d_0+1} + \dots + \gamma_{d_2} \mathbf{c}_{d_2} = 0, \end{aligned} \quad (6-30)$$

gdzie α_i , β_j i γ_k są pewnymi (niewiadomymi) współczynnikami liczbowymi. Jeżeli to równanie jest spełnione, to możemy wprowadzić wektor \mathbf{D}

$$\begin{aligned} \mathbf{D} &= \alpha_1 \mathbf{a}_1 + \dots + \alpha_{d_0} \mathbf{a}_{d_0} + \beta_{d_0+1} \mathbf{b}_{d_0+1} + \dots + \beta_{d_1} \mathbf{b}_{d_1} \\ &= -\gamma_{d_0+1} \mathbf{c}_{d_0+1} - \dots - \gamma_{d_2} \mathbf{c}_{d_2}. \end{aligned} \quad (6-31)$$

Wyrażenie w lewej części równania to wektor należący do \mathbb{L}_1 ; a wyrażenie w prawej części – wektor należący do \mathbb{L}_2 . Oba te wyrażenia reprezentują jeden i ten sam wektor \mathbf{D} . Wektor ten należy w taki razie do \mathbb{L}_0 i jako taki może być wyrażony jako kombinacja liniowa wektorów bazy 6-26. Z wyrażenia po prawej stronie 6-31 wynika jednak, że \mathbf{D} może być też wyrażony jako kombinacja liniowa wektorów $\mathbf{c}_{d_0+1}, \dots, \mathbf{c}_{d_2}$. Ale ponieważ wektory te, wraz z bazą 6-26 tworzą bazę 6-28 (układ wektorów liniowo niezależnych!) to wszystkie współczynniki $\gamma_{d_0+1}, \dots, \gamma_{d_2}$ muszą być równe zero. Wektor \mathbf{D} jest wektorem zerowym, a jeżeli tak, to jego współrzędne w bazie 6-27 – współczynniki $\alpha_1, \dots, \alpha_{d_0}$ oraz $\beta_{d_0+1}, \dots, \beta_{d_1}$ są też równe zero. Równanie 6-31 ma tylko rozwiązanie trywialne; układ 6-29 jest układem wektorów liniowo niezależnych. (Dla wnikliwego Czytelnika: przeprowadź ten sam dowód dla podprzestrzeni \mathbb{L}_0 będącej podprzestrzenią zerową, tzn. dla $d_0 = 0$.)

6.4 Rząd macierzy

Mając do czynienia z układem n -wymiarowych wektorów nasuwa się pytanie: jak określić czy jest to układ liniowo niezależnych wektorów? Jednym ze sposobów byłoby rozwiązanie układu odpowiednich równań jednorodnych – takiego jak, np. 6-17. W tym podrozdziale zapoznamy się z bardziej uniwersalną metodą, stanowiącą zresztą punkt wyjścia do uogólnienia naszych wiadomości o metodach rozwiązywania układów równań liniowych.

Przypuśćmy, że mamy macierz \mathcal{A} o s wierszach i n kolumnach (liczby s i n zupełnie dowolne!):

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{s1} & a_{s2} & \dots & a_{sn} \end{pmatrix} \quad (6-32)$$

Każda z n kolumn macierzy może być traktowana jako s -wymiarowy wektor. W takim kontekście możemy mówić o liniowej (nie)zależności kolumn macierzy. *Rząd układu kolumn określamy jako maksymalną liczbę liniowo niezależnych kolumn-wektorów \mathcal{A} ; rząd wektorów-kolumn będziemy nazywać też rzędem macierzy \mathcal{A} .* Podobnie moglibyśmy traktować s wierszy macierzy jako s wektorów n -wymiarowych i określać rząd macierzy jako liczbę liniowo niezależnych wierszy-wektorów. Oba podejścia dadzą ten sam wynik!

W rozdziale trzecim zdefiniowaliśmy pojęcie minora macierzy kwadratowej. Pojęcie to można uogólnić na przypadek *dowolnej* macierzy: elementy macierzy, leżące na przecięciu dowolnie wybranych k wierszy i k kolumn ($k \leq \min(s, n)$) tworzą minor stopnia k . Interesują nas teraz minory (wyznaczniki!) które są różne od zera, a konkretnie – minory o największej możliwej wartości k . (Pamiętajmy, że jeżeli wszystkie minory stopnia k są równe zeru, to wszystkie minory wyższych stopni będą też równe zeru – wynika to z techniki rozwinięcia Laplace’a).

Mamy:

Twierdzenie 6.5 (o rzędzie macierzy) *Rząd macierzy jest równy najwyższemu stopniu niezerowych wyznaczników tej macierzy.*

Dowód:

Założmy, że najwyższy niezerowy minor $D \neq 0$ macierzy \mathcal{A} jest stopnia r i – dla prostszego rachunku – jest to minor ulokowany w jej górnym lewym rogu:

$$\mathcal{A} = \begin{pmatrix} \begin{matrix} a_{11} & \dots & a_{1r} \\ \dots & D & \dots \\ a_{r1} & \dots & a_{rr} \end{matrix} & \begin{matrix} a_{1,r+1} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{r,r+1} & \dots & a_{rn} \end{matrix} \\ \begin{matrix} a_{r+1,1} & \dots & a_{r+1,r} \\ \dots & \dots & \dots \\ a_{s1} & \dots & a_{sr} \end{matrix} & \begin{matrix} a_{r+1,r+1} & \dots & a_{r+1,n} \\ \dots & \dots & \dots \\ a_{s,r+1} & \dots & a_{sn} \end{matrix} \end{pmatrix} \quad (6-33)$$

Pierwsze r kolumn macierzy \mathcal{A} tworzy układ liniowo niezależny – gdyby wektory-kolumny były liniowo zależne, to (przynajmniej) jeden z nich byłby liniową kombinacją pozostałych, a więc wyznacznik D byłby równy zeru (własność 2d – punkt 3.4). Natomiast każda l -ta kolumna ($r < l \leq n$) będzie liniową kombinacją pierwszych r kolumn.

Aby to wykazać, utwórzmy wyznacznik Δ_i stopnia $r + 1$:

$$\Delta_i = \begin{vmatrix} a_{11} & \dots & a_{1r} & a_{1l} \\ \dots & \dots & \dots & \dots \\ a_{r1} & \dots & a_{rr} & a_{rl} \\ a_{i1} & \dots & a_{ir} & a_{il} \end{vmatrix}, \quad (6-34)$$

powstały w wyniku „dołożenia” do D odpowiednich wyrazów l -tej kolumny ($r < l \leq n$) i i -tego (i – dowolne) wiersza. Wyznacznik Δ_i jest równy zeru dla każdego i ; jeżeli $i > r$ to wynika to z wyboru r jako kresu górnego stopnia niezerowego wyznacznika; jeżeli $i \leq r$ to w wyznaczniku mamy dwa identyczne wiersze.

Pozostaje zastosować do Δ_i rozwinięcie Laplace’a wg. ostatniego wiersza:

$$\Delta_i = 0 = a_{i1}M^{[i1]} + a_{i2}M^{[i2]} + \dots + a_{ir}M^{[ir]} + a_{il}M^{[il]}, \quad (6-35)$$

gdzie $M^{[ik]}$; $k = 1, \dots, l$ są dopełnieniami algebraicznymi elementów a_{ik} . W szczególności $M^{[il]} = D \neq 0$, a więc

$$a_{il} = -\frac{M^{[i1]}}{D}a_{i1} - \frac{M^{[i2]}}{D}a_{i2} - \dots - \frac{M^{[ir]}}{D}a_{ir}. \quad (6-36)$$

Ponieważ i było dowolne, $i = 1, \dots, s$ to *cała* l -ta kolumna wyraża się poprzez kombinacje liniowe pierwszych r kolumn - *cbdo*.

Twierdzenie o rzędzie macierzy dostarcza więc praktycznego sposobu rozstrzygnięcia kwestii czy dany układ wektorów jest liniowo zależny, czy nie: tworzymy macierz, której kolumnami są wektory i określamy rząd macierzy. Ten ostatni jest równy maksymalnej liczbie liniowo niezależnych wektorów należących do naszego układu.

Dowód twierdzenia daje także pewną wskazówkę praktyczną: w dowodzie interesowały nas tylko minory stopnia $r + 1$ „otaczające” niezerowy minor D . Dlatego określając rząd macierzy staramy się wyszukiwać kolejno niezerowe minory, poczynając od najniższego (tzn. drugiego) stopnia. Jeżeli zlokalizujemy niezerowy minor stopnia k to rozpatrujemy otaczające go minory stopnia $k + 1$; jeżeli wszystkie będą równe zero to rząd macierzy jest równy k .

Przykład 6.1 Aby określić rząd macierzy

$$\begin{pmatrix} 2 & -4 & 3 & 1 & 0 \\ 1 & -2 & 1 & -4 & 2 \\ 0 & 1 & -1 & 3 & 1 \\ 4 & -7 & 4 & -4 & 5 \end{pmatrix} \quad (6-37)$$

zauważmy, że różny od zera minor to np.

$$d = \begin{vmatrix} -4 & 3 \\ -2 & 1 \end{vmatrix} \neq 0. \quad (6-38)$$

Przyległy minor trzeciego rzędu

$$d' = \begin{vmatrix} 2 & -4 & 3 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{vmatrix} \quad (6-39)$$

jest różny od zera ($d' = 1$), ale kolejne, przyległe do d' minory:

$$\begin{vmatrix} 2 & -4 & 3 & 1 \\ 1 & -2 & 1 & -4 \\ 0 & 1 & -1 & 3 \\ 4 & -7 & 4 & -4 \end{vmatrix} = 0, \quad \begin{vmatrix} 2 & -4 & 3 & 0 \\ 1 & -2 & 1 & 2 \\ 0 & 1 & -1 & 1 \\ 4 & -7 & 4 & 5 \end{vmatrix} = 0, \quad (6-40)$$

tak więc rząd macierzy 6-37 jest równy 3.

Przykład 6.2 Aby znaleźć maksymalny układ liniowo niezależnych wektorów – bazę – dla czterech wektorów:

$$\mathbf{V}_1 = (2, -2, -4), \quad \mathbf{V}_2 = (1, 9, 3), \quad \mathbf{V}_3 = (-2, -4, 1), \quad \mathbf{V}_4 = (3, 7, -1),$$

tworzymy macierz

$$\begin{pmatrix} 2 & 1 & -2 & 3 \\ -2 & 9 & -4 & 7 \\ -4 & 3 & 1 & -1 \end{pmatrix},$$

wstawiając wektory jako kolejne kolumny. Rząd macierzy jest 2: lewy górny minor drugiego stopnia jest różny od zera, ale jego przyległe minory trzeciego stopnia są oba równe zero. Wynika stąd, że wektory \mathbf{V}_1 i \mathbf{V}_2 tworzą bazę: wektory \mathbf{V}_3 i \mathbf{V}_4 będą ich liniowymi kombinacjami. To warto (zdecydowanie) sprawdzić. Możemy też z powyższego wysnuć wniosek, że wszystkie cztery wektory (mające po 3 współrzędne, a więc stanowiące elementy „zwykłej”, trójwymiarowej przestrzeni, leżą na jednej płaszczyźnie, którą napinają \mathbf{V}_1 i \mathbf{V}_2 (czy też dowolna dwójka wektorów, wybrana spośród $\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3$ i \mathbf{V}_4).

Często interesuje nas zagadnienie: czy dany wektor ($-y$) należy do podprzestrzeni generowanej przez określone wektory bazowe? W niektórych przypadkach odpowiedź na takie pytanie jest prosta – szczególnie w przypadku odpowiedzi negatywnej.

Przykład 6.3 Czy wektor

$$\mathbf{u} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$

należy do przestrzeni generowanej przez wektor

$$\mathbf{w} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}?$$

Nie, bo nigdy nie uzyskamy w drugim wierszu \mathbf{u} niezerowej wartości współrzędnej (drugi wiersz \mathbf{w} zawiera zero).

Przykład 6.4 Czy wektor

$$\mathbf{v} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

należy do przestrzeni generowanej przez wektory

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \quad \mathbf{w}_2 = \begin{pmatrix} -2 \\ -1 \\ 1 \\ 0 \end{pmatrix} \quad \text{i} \quad \mathbf{w}_3 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix}?$$

Nie. Nie jest możliwe zapisanie \mathbf{v} jako liniowej kombinacji \mathbf{w}_1 , \mathbf{w}_2 i \mathbf{w}_3 . Drugi i trzeci wiersz \mathbf{w}_1 zawiera zera; druga i trzecia współrzędna \mathbf{v} zależy tylko od współrzędnych \mathbf{w}_2 i \mathbf{w}_3 . Zero w drugim wierszu \mathbf{v} wymaga aby współczynniki \mathbf{w}_2 i \mathbf{w}_3 były takie same; to z kolei wyklucza jedynkę w trzecim wierszu \mathbf{v} .

W przypadku gdy odpowiedź nie jest oczywista kwestię dotyczącą przynależności \mathbf{v} do przestrzeni generowanej przez \mathbf{w}_1 , \mathbf{w}_2 i \mathbf{w}_3 rozstrzygamy rozwiązując odpowiedni układ równań.

Przykład 6.5 Dla wektorów:

$$\mathbf{w}_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{w}_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \quad \mathbf{w}_3 = \begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix} \quad \text{i} \quad \mathbf{v} = \begin{pmatrix} 3 \\ 3 \\ 1 \end{pmatrix}$$

pytanie: $\mathbf{v} = x_1\mathbf{w}_1 + x_2\mathbf{w}_2 + x_3\mathbf{w}_3$ jest równoważne układowi równań:

$$\left. \begin{aligned} 2x_1 + x_2 + 3x_3 &= 3, \\ x_1 - x_2 &= 3, \\ x_1 + x_2 + 2x_3 &= 1. \end{aligned} \right\}$$

Pozostawiamy Czytelnikowi sprawdzenie, że \mathbf{v} rzeczywiście daje się wyrazić jako kombinacja liniowa wektorów \mathbf{w}_1 , \mathbf{w}_2 i \mathbf{w}_3 , i to na nieskończenie wiele sposobów.

Oczywistym wnioskiem z twierdzenia o rzędzie macierzy, będzie jego równoważne sformułowanie:

Twierdzenie 6.6 (o rzędzie macierzy) *Maksymalna liczba liniowo niezależnych wierszy danej macierzy jest równa maksymalnej liczbie niezależnych kolumn – a więc określa ona również rząd macierzy.*

Uzasadnienie sprowadza się do rozważenia macierzy transponowanej – wyznacznik takiej macierzy jest taki sam jak macierzy wyjściowej. Wszystkie minory występujące w macierzy wyjściowej pojawiają się w macierzy transponowanej, a ich wartości takie same. Rząd macierzy transponowanej – liczba jej liniowo niezależnych kolumn, a więc liczba liniowo niezależnych wierszy macierzy wyjściowej – jest równy rzędowi macierzy wyjściowej, a w takim razie obie liczby są równe.

W rozdziale trzecim powiedzieliśmy, że wyznacznik jest równy zeru jeżeli (przynajmniej) dwa z jego wierszy (kolumn) są liniowo zależne. Z kolei jeżeli wyznacznik macierzy n -tego stopnia jest równy zeru, to oznacza to, że rząd macierzy jest mniejszy niż n , a więc wiersze (kolumny) muszą być liniowo zależne (liczba liniowo niezależnych elementów musi być mniejsza od n). (Warto powrócić do dyskusji układu równań jednorodnych w punkcie 3.5.)

Fakt zerowania się wyznacznika, w przypadku gdy jego wiersze (kolumny) są liniowo zależne, jest łatwy do zilustrowania, jeżeli przypomnimy sobie, że iloczyn mieszany trzech wektorów – reprezentujący objętość równoległościanu zbudowanego na tych wektorach – jest określony jako wyznacznik 3-ego stopnia, zbudowany ze współrzędnych tych wektorów. Jeżeli jeden z tych wektorów jest liniową kombinacją dwóch pozostałych to znaczy, że leży on w płaszczyźnie, napiętej przez te dwa wektory – „równoległościan” redukuje się do płaskiej figury o zerowej objętości.

6.5 Rząd macierzy inaczej – diagonalizacja macierzy

Rząd macierzy można określić w inny sposób, sprowadzając macierz do tzw. postaci diagonalnej („jedynkowej”). Dla dowolnej macierzy $n \times s$ postać diagonalna oznacza, że wszystkie elementy macierzy są równe zeru, za wyjątkiem elementów $a_{11}, a_{22}, \dots, a_{rr}$, (gdzie $r = \min(n, s)$), które są równe jedności⁴. Rząd takiej macierzy to oczywiście r . Z punktu widzenia zachowania rzędu macierzy operacjami których wolno nam dokonywać na wyjściowej macierzy są operacje analogiczne do tych, które stosowaliśmy w metodzie Gaussa, z tą różnicą, że ze względu na „równouprawnienie” wierszy i kolumn (w kontekście określania rzędu macierzy) operacje te możemy przeprowadzać na wierszach i na kolumnach. Operacje te t:

1. zamiana dwóch dowolnych wierszy (kolumn);
2. pomnożenie elementów danego wiersza (danej kolumny) przez dowolną liczbę różną od zera;
3. dodanie do elementów danego wiersza (kolumny) odpowiednich elementów innego wiersza (kolumny), pomnożonych przez dowolną liczbę

Pierwsze dwie własności są oczywiste; trzecią łatwo wykazać, jeżeli posłużymy się wnioskiem o układach równoważnych wektorów. Rzeczywiście n kolumn macierzy, czyli n wektorów:

$$\mathbf{V}_1, \dots, \mathbf{V}_i, \dots, \mathbf{V}_j, \dots, \mathbf{V}_n \quad (6-41)$$

i n kolumn „nowej” macierzy

$$\mathbf{V}_1, \dots, \mathbf{V}'_i = \mathbf{V}_i + k\mathbf{V}_j, \dots, \mathbf{V}_j, \dots, \mathbf{V}_n \quad (6-42)$$

są układami równoważnymi, a więc ich rząd musi być identyczny.

Stosując trzy wymienione wyżej transformacje można dowolną macierz przeprowadzić do postaci diagonalnej, w której jedyne elementy niezerowe znajdują się na przekątnej głównej i są w dodatku równe 1. Jeżeli macierz wyjściowa ma postać

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{s1} & \dots & a_{sn} \end{pmatrix} \quad (6-43)$$

to – przy założeniu, że $a_{11} \neq 0$ ⁵ – dzielimy pierwszy wiersz przez a_{11} , przekształcając a_{11} do jedności, a od każdej j -tej kolumny poza kolumną pierwszą ($j = 2, \dots, n$) odejmujemy pierwszą pomnożoną przez a_{1j} . Każdy pierwszy element tych kolumn przekształca się w zero. Analogiczną procedurą stosujemy do wierszy (poza pierwszym) nowo-powstałych macierzy. Macierz przyjmuje postać

$$\mathbf{A}' = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & a'_{22} & \dots & a'_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & a'_{s2} & \dots & a'_{sn} \end{pmatrix} \quad (6-44)$$

⁴Dla macierzy kwadratowej $n = s$ i jedynki leżą na przekątnej głównej.

⁵Jeżeli $a_{11} = 0$ przedstawiamy wiersze (kolumny) tak aby uzyskać $a_{11} \neq 0$. Zastanów się, co będzie jeżeli wszystkie pierwsze elementy kolumn i wierszy są $=0$.

(primami oznaczamy elementy nowej macierzy). Następnie opisaną procedurę stosujemy do pod-macierzy, powstałej po odrzuceniu z \mathcal{A}' pierwszego wiersza i pierwszej kolumny – po skończonej liczbie operacji dostaniemy macierz diagonalną.

Liczba niezerowych elementów (jedynek) tej macierzy określa rząd macierzy (tej wyjściowej, jak i wszystkich macierzy powstałych w wyniku kolejnych przekształceń).

Przykład 6.6 Macierz wyjściowa

$$\mathcal{A} = \begin{pmatrix} 0 & 2 & -4 \\ -1 & -4 & 5 \\ 3 & 1 & 7 \\ 0 & 5 & -10 \\ 2 & 3 & 0 \end{pmatrix}.$$

Zamieniamy pierwszą kolumnę z drugą i mnożymy pierwszy wiersz przez $\frac{1}{2}$:

$$\mathcal{A}' = \begin{pmatrix} 1 & 0 & -2 \\ -4 & -1 & 5 \\ 1 & 3 & 7 \\ 5 & 0 & -10 \\ 3 & 2 & 0 \end{pmatrix}.$$

Teraz do trzeciej kolumny dodajemy pierwszą pomnożoną przez 2; następnie do drugiego, trzeciego czwartego i piątego wiersza dodajemy pierwszy pomnożony przez odpowiedni czynnik (odpowiednio: 4, -1 , -5 , -3); nowa macierz to

$$\mathcal{A}'' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & -3 \\ 0 & 3 & 9 \\ 0 & 0 & 0 \\ 0 & 2 & 6 \end{pmatrix}.$$

Drugi wiersz \mathcal{A}'' mnożymy przez -1 ; od trzeciej kolumny odejmujemy drugą pomnożoną przez 3; od trzeciego i piątego wiersza odejmujemy powstały drugi wiersz, pomnożony przez odpowiednie czynniki. Ostateczna postać – diagonalna – to

$$\mathcal{A}_{\text{diag}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Rząd macierzy \mathcal{A} jest więc równy 2.

6.6 Układy równań liniowych – podsumowanie

Pojęcie rzędu macierzy może służyć do sformułowania pewnych podstawowych praw dotyczących rozwiązywalności układów równań liniowych lub ich charakteru (sprzeczny/niesprzeczny; oznaczony/nieoznaczony).

W tym celu przypomnijmy ogólną postać układu m równań liniowych z n niewiadomymi:

$$\begin{aligned} a_{11}x_1 &+ a_{12}x_2 &+ \dots &+ a_{1n}x_n &= b_1, \\ a_{21}x_1 &+ a_{22}x_2 &+ \dots &+ a_{2n}x_n &= b_2, \\ \dots &+ \dots &+ \dots &+ \dots &= \dots, \\ a_{m-1,1}x_1 &+ a_{m-1,2}x_2 &+ \dots &+ a_{m-1,n}x_n &= b_{m-1}, \\ a_{m1}x_1 &+ a_{m2}x_2 &+ \dots &+ a_{mn}x_n &= b_m. \end{aligned} \tag{6-45}$$

Oprócz macierzy współczynników

$$\mathcal{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \tag{6-46}$$

wprowadzamy macierz „uzupełnioną” lub „poszerzoną” o kolumnę wyrazów wolnych

$$\mathcal{A}_{\text{uzup}} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix} \quad (6-47)$$

i określamy rząd macierzy \mathcal{A} i macierzy $\mathcal{A}_{\text{uzup}}$. Jest oczywiste, że $\text{rz}\mathcal{A}_{\text{uzup}}$ może być bądź równy $\text{rz}\mathcal{A}$, bądź równy $\text{rz}\mathcal{A} + 1$.

Twierdzenie 6.7 (Kroneckera-Capelliego): *rozwiązanie układu 6-45 istnieje wtedy i tylko wtedy, gdy rząd macierzy współczynników jest równy rzędowi macierzy uzupełnionej. Ponadto, jeżeli $\text{rz}(\mathcal{A}) = \text{rz}(\mathcal{A}_{\text{uzup}}) = n$ (liczba niewiadomych), to istnieje dokładnie jedno rozwiązanie (układ oznaczony); gdy $\text{rz}\mathcal{A} = \text{rz}\mathcal{A}_{\text{uzup}} = r < n$, to istnieje nieskończenie wiele rozwiązań, zależnych od (dowolnie wybranych) $n - r$ parametrów.*

Zamiast dowodu podajmy „praktyczny” przepis postępowania i kilka przykładów:

W przypadku gdy $\text{rz}(\mathcal{A}) = \text{rz}(\mathcal{A}_{\text{uzup}}) = r < n$, w macierzy \mathcal{A} wybieramy r liniowo niezależnych wierszy i w układzie równań 6-45 pozostawiamy tylko te równania, których współczynniki a_{ik} należą do tych wierszy. W powstałym układzie r równań pozostawimy „po lewej stronie” r niewiadomych; tak aby wyznacznik ich współczynników (stopnia r) był różny od zera, pozostałe $n - r$ niewiadomych przenosimy na prawą stronę i przyporządkowujemy im *dowolnie* wybrane wartości liczbowe. Ostatni krok to rozwiązanie tak utworzonego układu r równań na r niewiadomych metodą Gaussa (lub wzorami Cramera).

Przykład 6.7 Układ

$$\left. \begin{array}{r} x_1 + x_2 - 2x_3 - x_4 + x_5 = 1, \\ 3x_1 - x_2 + x_3 + 4x_4 + 3x_5 = 4, \\ x_1 + 5x_2 - 9x_3 - 8x_4 + x_5 = 0 \end{array} \right\}$$

jest układem posiadającym rozwiązanie bo rząd macierzy współczynników jest równy rzędowi macierzy uzupełnionej (=2). Lewe strony – na przykład – pierwszego i trzeciego równania są liniowo niezależne, bo współczynniki x_1 i x_2 tworzą niezerowy minor drugiego stopnia. Niewiadome x_3, x_4 i x_5 pozostawiamy nieokreślone i przenosimy na prawą stronę; układ dwóch równań rozwiązujemy metodą Cramera:

$$\begin{aligned} x_1 &= \frac{5}{4} + \frac{1}{4}x_3 - \frac{3}{4}x_4 - x_5, \\ x_2 &= -\frac{1}{4} + \frac{7}{4}x_3 + \frac{7}{4}x_4. \end{aligned}$$

Powyższe wzory określają ogólne rozwiązanie układu: konkretne rozwiązania uzyskamy podstawiając za x_3, x_4 i x_5 pewne dowolne wartości. Warto sprawdzić, że podstawiając za x_1 i x_2 z powyższych wzorów do odrzuconego (drugiego) równania układu otrzymujemy tożsamość.

Przykład 6.8 Układ

$$\left. \begin{array}{r} 4x_1 + x_2 - 2x_3 + x_4 = 3, \\ x_1 - 2x_2 - x_3 + 2x_4 = 2, \\ 2x_1 + 5x_2 - x_4 = -1, \\ 3x_1 + 3x_2 - x_3 - 3x_4 = 1 \end{array} \right\}$$

jest układem czterech równań z czterema niewiadomymi. Wyznacznik główny układu jest jednak równy zeru – a więc nie można zastosować wzorów Cramera. Rząd macierzy współczynników jest równy rzędowi macierzy uzupełnionej (=3); (minor 3-ego stopnia w prawym, górnym rogu macierzy współczynników). Rozpatrujemy tylko trzy pierwsze równania (x_4 pozostaje nieokreślone). Rozwiązanie ogólne to

$$x_2 = -\frac{1}{5} - \frac{2}{5}x_1, \quad x_3 = -\frac{8}{5} + \frac{9}{5}x_1, \quad x_4 = 0.$$

Przykład 6.9 Układ

$$\left. \begin{aligned} 5x_1 - x_2 - 2x_3 + x_4 &= 7, \\ 2x_1 + x_2 + 4x_3 - 2x_4 &= 1, \\ x_1 - 3x_2 - 6x_3 + 5x_4 &= 0 \end{aligned} \right\}$$

jest układem trzech równań z czterema niewiadomymi. Rząd macierzy współczynników = 2 (lewy górny róg). Rząd macierzy uzupełnionej o kolumnę wyrazów wolnych = 3 – układ jest sprzeczny i nie posiada rozwiązań.

Przykład 6.10 Układ

$$\left. \begin{aligned} 7x_1 + 3x_2 &= 2, \\ x_1 - 2x_2 &= -3, \\ 4x_1 + 9x_2 &= 11 \end{aligned} \right\}$$

jest układem trzech równań z dwoma niewiadomymi. Rząd macierzy współczynników = 2 (lewy górny róg); rząd macierzy uzupełnionej o kolumnę wyrazów wolnych = 2. Rozwiązujemy dwa pierwsze (liniowo niezależne) równania:

$$x_1 = -\frac{5}{17} \quad \text{i} \quad x_2 = \frac{23}{17}.$$

(Trzecie równanie jest także spełnione.)

6.6.1 Metoda Gaussa-Jordana

Tak jak wspomnieliśmy w rozdziale trzecim, metoda Gaussa-Jordana stanowi pewien wariant metody Gaussa. Metoda ta świetnie nadaje się do obliczeń przy zastosowaniu komputerów. Jej podstawowa zaleta to „jakościowa” możliwość określenia charakteru rozwiązań układu po przeprowadzeniu pełnej eliminacji w wierszach macierzy \mathcal{A} , uzupełnionej kolumną wyrazów wolnych.

Aby rozwiązać układ równań 6-45 – czyli w skrócie: $\mathcal{A}\mathcal{X} = \mathcal{B}$ macierz współczynników, uzupełnioną o kolumnę wyrazów wolnych, poddajemy eliminacji gaussowskiej (podobnie jak w punkcie 3.2, ale tak, aby w poszczególnych wierszach wyeliminować (wyzerować) *maksymalną* liczbę współczynników. Metodę prześledzimy na kilku przykładach:

Przykład 6.11 Nasz układ to:

$$\left. \begin{aligned} 2x + y + 3z &= 1, \\ x - y &= 1, \\ 2x + z &= 1. \end{aligned} \right\} \quad (6-48)$$

Przekształcamy macierz uzupełnioną⁶

$$\begin{aligned} \left(\begin{array}{cccc} 2 & 1 & 3 & 1 \\ 1 & -1 & 0 & 1 \\ 2 & 0 & 1 & 1 \end{array} \right) &\xrightarrow{1} \left(\begin{array}{cccc} 1 & 1/2 & 3/2 & 1/2 \\ 0 & -3/2 & -3/2 & 1/2 \\ 0 & -1 & -2 & 0 \end{array} \right) \xrightarrow{2} \left(\begin{array}{cccc} 1 & 1/2 & 3/2 & 1/2 \\ 0 & 1 & 1 & -1/3 \\ 0 & 0 & -1 & -1/3 \end{array} \right) \\ &\xrightarrow{3} \left(\begin{array}{cccc} 1 & 0 & 1 & 2/3 \\ 0 & 1 & 1 & -1/3 \\ 0 & 0 & 1 & 1/3 \end{array} \right) \xrightarrow{4} \left(\begin{array}{cccc} 1 & 0 & 0 & 1/3 \\ 0 & 1 & 1 & -1/3 \\ 0 & 0 & 1 & 1/3 \end{array} \right) \xrightarrow{5} \left(\begin{array}{cccc} 1 & 0 & 0 & 1/3 \\ 0 & 1 & 0 & -2/3 \\ 0 & 0 & 1 & 1/3 \end{array} \right). \end{aligned}$$

Rozwiązania jak na dłoni: $x = z = 1/3$, $y = -2/3$.

Przykład 6.12 Zmieniając nieco wyjściowy układ, chcemy rozwiązać

⁶Warto sprawdzić; np (1): dzielimy pierwszy wiersz macierzy wyjściowej przez 2, do drugiego wiersza dodajemy pierwszy pomnożony przez $-1/2$, od trzeciego odejmujemy pierwszy. A dalsze?

$$\left. \begin{aligned} 2x + y + 3z &= 1, \\ x - y &= 1, \\ x + y + 2z &= 1. \end{aligned} \right\} \quad (6-49)$$

Przekształcając wyjściową macierz uzupełnioną uzyskujemy postać, w której już nie wszystkie kolumny macierzy współczynników zawierają „samotne” jedynki; w dodatku pojawia się taka jedynka w kolumnie ostatniej:

$$\left(\begin{array}{cccc} 2 & 1 & 3 & 1 \\ 1 & -1 & 0 & 1 \\ 1 & 1 & 2 & 1 \end{array} \right) \rightarrow \left(\begin{array}{cccc} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right). \quad (6-50)$$

Układ jest układem sprzecznym: ostatni wiersz to $0 = 1$.

Przykład 6.13 Kolejny układ to

$$\left. \begin{aligned} 2x + y + 3z &= 1, \\ x - y &= 1, \\ x + y + 2z &= 1/3. \end{aligned} \right\} \quad (6-51)$$

Podobnie jak w poprzednim przypadku nie wszystkie kolumny przekształconej macierzy współczynników zawierają „samotne” jedynki; ostatni wiersz to same zera:

$$\left(\begin{array}{cccc} 2 & 1 & 3 & 1 \\ 1 & -1 & 0 & 1 \\ 1 & 1 & 2 & 1/3 \end{array} \right) \rightarrow \left(\begin{array}{cccc} 1 & 0 & 1 & 2/3 \\ 0 & 1 & 1 & -1/3 \\ 0 & 0 & 0 & 0 \end{array} \right). \quad (6-52)$$

Pierwszy wiersz to: $x + z = 2/3$; drugi: $y + z = -1/3$. Układ ma nieskończenie wiele rozwiązań: dla dowolnie wybranego z obliczamy z tych równań x i y .

Podane trzy przykłady pokrywają wszystkie możliwości. Możemy sformułować reguły, które określają charakter rozwiązań w zależności od końcowej postaci przekształconej macierzy. Po pierwsze uściślijmy definicję tej postaci. Możemy nazwać ją *uogólnioną postacią diagonalną*. Jej charakterystyczne cechy (a zarazem definicja), to:

1. W każdym wierszu pierwszy niezerowy element to 1, tzw. jedynka-dźwignia,⁷ albo „samotna jedynka”.
2. Samotna jedynka w niższym wierszu jest zawsze na prawo w stosunku do samotnej jedynki w wyższych wierszach
3. W każdej kolumnie zawierającej samotną jedynkę wszystkie pozostałe elementy są zerami.
4. Jeżeli mamy wiersze składające się z samych zer to są one u dołu macierzy.
5. *Liczba samotnych jedynek jest równa rzędowi macierzy.*

Po sprowadzeniu macierzy uzupełnionej do uogólnionej postaci diagonalnej będziemy mieli kolumny, w których występują „samotne jedynki” i kolumny, w których takich samotnych jedynek nie ma. Spełnione są następujące reguły:

1. Jeżeli w ostatniej kolumnie (wyrazów wolnych) występuje samotna jedynka układ nie ma rozwiązań (bo mamy równanie: $0 = 1$).
2. W przypadku gdy w ostatniej kolumnie (wyrazów wolnych) nie występuje samotna jedynka to
 - (a) Jeżeli wszystkie kolumny przekształconej macierzy współczynników zawierają samotne jedynki układ ma jednoznaczne rozwiązanie
 - (b) Jeżeli choć jedna kolumna nie zawiera samotnej jedynki układ ma nieskończenie wiele rozwiązań. Wartości zmiennych, którym odpowiadają kolumny pozbawione samotnych jedynek możemy ustalać dowolnie; wartości pozostałych zmiennych wyliczamy z równań.

Powyższe reguły, stanowiące bardziej praktyczne zwerbalizowanie twierdzenia Kroneckera-Capelliego, warto sprawdzić rozwiązując przykłady zamieszczone w punkcie 6.5 metodą Gaussa-Jordana.

⁷Ten niewątpliwie dziwny termin ma swoją genezę w metodzie znajdowania macierzy odwrotnej.

6.6.2 Układ równań jednorodnych

W świetle twierdzenia Kroneckera-Capelliego taki układ ma zawsze rozwiązanie, bo rząd macierzy współczynników i macierzy uzupełnionej są takie same (dodanie kolumny zer nie może zmienić rzędu macierzy). Jeżeli rząd macierzy współczynników jest równy r i $r = n$ to rozwiązaniem może być tylko rozwiązanie trywialne $(0, 0, \dots, 0)$; dla $r < n$ mamy także rozwiązania nietrywialne. Rzeczywiście, aby dla układu n równań z n niewiadomymi istniało nietrywialne rozwiązanie wyznacznik główny musi być równy zeru (oznacza to że $r < n$). Z drugiej strony, jeżeli liczba równań jest mniejsza od liczby niewiadomych to układ *musi* mieć rozwiązanie nietrywialne, bo rząd macierzy współczynników nie może być w tym przypadku równy liczbie niewiadomych ($\text{rz}(\mathcal{A}) < n$).

W rozdziale trzecim dyskutowaliśmy przypadek rozwiązywania układu równań jednorodnych dla n niewiadomych; z dyskusji wynikało, że określone są *stosunki* x_k/x_n dla $k = 1, \dots, n-1$, a nie same niewiadome. Oznacza to, że jeżeli wektor x -ów: $(x_1, x_2, \dots, x_{n-1}, x_n)$ jest rozwiązaniem układu, to wektor $(kx_1, kx_2, \dots, kx_{n-1}, kx_n)$ (k – dowolna liczba) jest także rozwiązaniem. Co więcej, jeżeli wektor $(x_1, x_2, \dots, x_{n-1}, x_n)$ i wektor $(x'_1, x'_2, \dots, x'_n)$ są rozwiązaniami układu, to ich suma – wektor $(x_1 + x'_1, x_2 + x'_2, \dots, x_{n-1} + x'_{n-1}, x_n + x'_n)$ jest także rozwiązaniem; a także suma $(ax_1 + bx'_1, ax_2 + bx'_2, \dots, ax_{n-1} + bx'_{n-1}, ax_n + bx'_n)$ (a, b – dowolne liczby). Okazuje się więc, że jeżeli układ równań jednorodnych ma jakies nietrywialne rozwiązanie, to ma ich nieskończenie wiele. Jeżeli każde z nich traktować jako n -wymiarowy wektor to wiemy, że każdy układ więcej niż n takich wektorów będzie układem wektorów liniowo zależnych. Z nieskończonej mnogości rozwiązań powinno więc być możliwe wybranie *skończonej* liczby rozwiązań, które tworzą tzw. *układ fundamentalny rozwiązań* (bazę), w tym sensie, że każde inne rozwiązanie daje się wyrazić jako ich pewna kombinacja liniowa. Pytanie: ile rozwiązań tworzy układ fundamentalny? Odpowiedź: $n - r$, gdzie n jest liczbą niewiadomych, a r rzędem macierzy współczynników. Na przykład dla $r = n - 1$ (przykład dyskutowany w rozdziale trzecim) mamy *de facto* jedno rozwiązanie fundamentalne, wszystkie inne są jego liniowymi kombinacjami (wielokrotnościami).

Przykład 6.14 Układ

$$\left. \begin{array}{r} 3x_1 + x_2 - 8x_3 + 2x_4 + x_5 = 0, \\ 2x_1 - 2x_2 - 3x_3 - 7x_4 + 2x_5 = 0, \\ x_1 + 11x_2 - 12x_3 + 34x_4 - 5x_5 = 0, \\ x_1 - 5x_2 + 2x_3 - 16x_4 + 3x_5 = 0 \end{array} \right\}$$

jest układem czterech równań jednorodnych z pięcioma niewiadomymi. Rząd macierzy współczynników $r = 2$; każda baza rozwiązań układu składa się z trzech rozwiązań ($n - r = 3$). Ograniczamy się do dwóch pierwszych (liniowo niezależnych) równań układu

$$\left. \begin{array}{r} 3x_1 + x_2 = -8x_3 - 2x_4 - x_5, \\ 2x_1 - 2x_2 = 3x_3 + 7x_4 - 2x_5 \end{array} \right\},$$

w których niewiadome x_3, x_4, x_5 traktujemy jako pewne, dowolne liczby. Rozwiązujemy układ na x_1 i x_2 :

$$\begin{aligned} x_1 &= \frac{19}{8}x_3 + \frac{3}{8}x_4 - \frac{1}{2}x_5, \\ x_2 &= \frac{7}{8}x_3 - \frac{25}{8}x_4 + \frac{1}{2}x_5. \end{aligned}$$

Pozostaje wybrać wartości x_3, x_4, x_5 i z powyższych równań wyliczyć x_1 i x_2 . Przy wyborze x_3, x_4, x_5 kierujemy się zwyczajowo z jednej strony prostotą, z drugiej – chcemy aby wybrane rozwiązania były układem liniowo niezależnym. Dlatego

Wybrane:			Wyliczone:	
x_3	x_4	x_5	x_1	x_2
1	0	0	19/8	7/8
0	1	0	3/8	-25/8
0	0	1	-1/2	1/2

Trzy liniowo niezależne rozwiązania układu to 5-wymiarowe wektory:

$$\mathbf{V}_1 = \left(\frac{19}{8}, \frac{7}{8}, 1, 0, 0 \right), \quad \mathbf{V}_2 = \left(\frac{3}{8}, -\frac{25}{8}, 0, 1, 0 \right), \quad \mathbf{V}_3 = \left(-\frac{1}{2}, \frac{1}{2}, 0, 0, 1 \right).$$

Liniową niezależność gwarantuje nam odpowiedni wybór x_3, x_4, x_5 .

6.6.3 Układ równań niejednorodnych, a układ równań jednorodnych

Kładąc w układzie 6-45, który skrótowo możemy zapisać

$$\sum_{k=1}^n a_{ik}x_k = b_i; \quad i = 1, \dots, m \tag{6-53}$$

wszystkie $b_i = 0$ otrzymujemy stowarzyszony (albo: zredukowany) układ równań jednorodnych:

$$\sum_{k=1}^n a_{ik}x_k = 0; \quad i = 1, \dots, m. \tag{6-54}$$

Pomiędzy rozwiązaniami obu układów istnieje ścisły związek; słuszne są dwa twierdzenia:

Twierdzenie 6.8 (o sumie rozwiązań układu niejednorodnego i jednorodnego) *Suma dowolnego rozwiązania układu 6-53 i dowolnego rozwiązania układu 6-54 jest rozwiązaniem układu 6-53.*

Dowód: jeżeli przez c_1, c_2, \dots, c_n oznaczyć pewne rozwiązanie 6-53, a przez d_1, d_2, \dots, d_n oznaczyć pewne rozwiązanie 6-54 to mamy

$$\sum_{k=1}^n a_{ik}[c_k + d_k] = \sum_{k=1}^n a_{ik}c_k + \sum_{k=1}^n a_{ik}d_k = b_k + 0 = b_k.$$

Twierdzenie 6.9 (o różnicy rozwiązań układu niejednorodnego)

Różnica dwóch dowolnych rozwiązań 6-53 jest rozwiązaniem układu 6-54.

Dowód: jeżeli przez c_1, c_2, \dots, c_n oznaczyć pewne rozwiązanie 6-53, a przez c'_1, c'_2, \dots, c'_n oznaczyć inne rozwiązanie tego układu to mamy

$$\sum_{k=1}^n a_{ik}[c_k - c'_k] = \sum_{k=1}^n a_{ik}c_k - \sum_{k=1}^n a_{ik}c'_k = b_k - b_k = 0.$$

6.7 Aksjomatyczne definicje przestrzeni wektorowej

Grupa i ciało

Zaczynamy od zdefiniowania podstawowych pojęć – grupy i ciała:

Definicja 6.1 Grupę G definiujemy jako zbiór obiektów lub operacji — elementów grupy — na których określona jest pewna operacja nazwana umownie mnożeniem, chociaż lepszym jej określeniem byłoby: składanie (elementów grupy). Operację tą oznaczamy symbolem \cdot . Elementy grupy i operacja mnożenia spełniają cztery warunki:

1. Operacja mnożenia jest zamknięta, tzn. jeżeli a i b są elementami grupy, to ich „iloczyn” ab jest także elementem grupy:

$$a \in G, b \in G \rightarrow (a \cdot b) \in G.$$

(Symbol $a \in G$ czytamy (element) a należy do (zbioru) G .)

2. Operacja mnożenia jest łączna: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.

3. Istnieje element jednostkowy I taki, że $a \cdot I = I \cdot a = a$ dla każdego $a \in G$.

4. Dla każdego elementu grupy a istnieje element odwrotny: a^{-1} taki, że $a \cdot a^{-1} = a^{-1} \cdot a = I$.

Grupę, dla której operacja mnożenia jest przemienna, a więc: $a \cdot b = b \cdot a$ nazywamy grupą abelową.

Przykładami grup mogą być na przykład:

1. Zbiór liczb całkowitych N , z dodawaniem jako operacją składania („mnożenia”). Elementem jednostkowym jest 0 , natomiast elementem odwrotnym $N^{-1} \equiv -N$.
2. Ten sam zbiór nie tworzy grupy, jeżeli za działanie grupowe przyjąć „zwykłe” mnożenie. Element jednostkowy $I = 1$, ale zero (0) nie ma elementu odwrotnego. Jeżeli zero usunąć ze zbioru liczb całkowitych to dostajemy grupę: $N^{-1} = -N; N \neq 0$. Grupa ta, podobnie jak poprzednia, jest abelowa. Obie grupy stanowią przykłady grup nieskończonych (ilość elementów) i dyskretnych (przejście od jednego elementu grupy do następnego następuje w sposób skokowy).
3. Zbiór dwuwymiarowych obrotów R_ϕ o kąt ϕ na płaszczyźnie jest grupą nieskończoną ciągłą (bo kąt ϕ zmienia się w sposób ciągły, przyjmując nieskończenie wiele wartości). Element jednostkowy to R_0 ; element odwrotny to $R_{-\phi}$. Grupa ta jest abelowa.
4. Zbiór obrotów w przestrzeni 3-wymiarowej jest grupą ciągłą, ale nieabelową. (Por. rozdział 4, Rys.4.1).

Zdefiniowanie ciała w oparciu o pojęcie grupy sprowadza się do podania zbioru aksjomatów spełnianych przez elementy ciała:

Definicja 6.2 Ciało jest zbiorem Z , na którym określone są dwie operacje: „ $+$ ” i „ \cdot ”. Ciało $\equiv \{Z, +, \cdot\}$, przy czym:

1. $\{Z, +\}$ jest grupą abelową z elementem jednostkowym \mathcal{O} , tzn.

$$\forall z \in Z; \quad z + \mathcal{O} = \mathcal{O} + z = z.$$

(Symbol $\forall z \in Z$ czytamy dla wszystkich (elementów) z należących do (zbioru) Z .)

2. $\{\mathcal{Z}', \cdot\}$, gdzie $\{\mathcal{Z}'\} = \{\mathcal{Z}\} - \mathcal{O}$, jest grupą abelową z elementem jednostkowym \mathcal{I} , tzn.

$$\forall z \in \mathcal{Z}'; \quad z \cdot \mathcal{I} = \mathcal{I} \cdot z = z.$$

3. operacja $+$ posiada własność rozdzielności względem operacji „ \cdot ”, tzn.

$$\forall z_1, z_2, z_3 \in \mathcal{Z}; \quad z_1 \cdot (z_2 + z_3) = z_1 \cdot z_2 + z_1 \cdot z_3.$$

Najprostsze przykłady ciał są jednocześnie najważniejszymi: zbiór wszystkich liczb wymiernych, zbiór wszystkich liczb rzeczywistych oraz zbiór wszystkich liczb zespolonych, przy czym operacjami „ $+$ ” i „ \cdot ” są we wszystkich przypadkach odpowiednio „zwykłe” dodawanie i mnożenie.

Z powyższych definicji wynika, że bardziej formalna definicja liczb zespolonych może być oparta o pojęcie ciała.

Przestrzeń wektorowa

Definicja 6.3 Przestrzenią wektorową \mathbb{R} rozpiętą nad ciałem \mathcal{Z} nazywamy zbiór $\mathbb{R}(\mathcal{Z})$ elementów — wektorów — spełniający warunki:

1. Istnieje operacja dodawania wektorów „ $+$ ”, taka że $\{\mathbb{R}(\mathcal{Z}), +\}$ jest grupą abelową; elementem jednostkowym operacji „ $+$ ” jest wektor zerowy (0).

2. $\forall \alpha \in \mathcal{Z}$ i $\forall \beta \in \mathcal{Z}$ a także $\forall \mathbf{x}$ i $\forall \mathbf{y} \in \mathbb{R}(\mathcal{Z})$ zachodzi:

$$\begin{aligned} \alpha \mathbf{x}, \alpha \mathbf{y}, \beta \mathbf{x}, \beta \mathbf{y} &\in \mathbb{R}(\mathcal{Z}) \\ \alpha(\beta \mathbf{x}) &= (\alpha\beta) \mathbf{x} \\ \mathcal{I} \mathbf{x} &= \mathbf{x} \\ \alpha(\mathbf{x} + \mathbf{y}) &= \alpha \mathbf{x} + \alpha \mathbf{y} \\ (\alpha + \beta) \mathbf{x} &= \alpha \mathbf{x} + \beta \mathbf{x}. \end{aligned}$$

Przestrzeń wektorowa posiada wymiar: $\mathbb{R}(\mathcal{Z}) = \mathbb{R}^n(\mathcal{Z})$. Przestrzeń może być skończenie wymiarowa lub nieskończenie wielowymiarowa. Decyduje o tym odpowiednio skończoność lub nieskończoność zbioru bazy przestrzeni wektorowej (por. niżej). Przestrzeń wektorowa może być rzeczywistą lub zespoloną, w zależności od tego czy \mathcal{Z} jest ciałem liczb rzeczywistych czy zespolonych.

Definicja 6.4 Zbiór liniowo niezależnych wektorów ξ_i przestrzeni wektorowej $\mathbb{R}^n(\mathcal{Z})$ nazywamy bazą (maksymalną rodziną liniowo niezależną), jeżeli każdy wektor \mathbf{A} przestrzeni \mathbb{R}^n jest liniową kombinacją elementów bazy, tzn.

$$\mathbf{A} = \sum_{i=1}^n A_i \xi_i,$$

gdzie A_i nazywamy współrzędnymi wektora \mathbf{x} w bazie $\{\xi_i\}$.

Górna granica sumy to n – wymiar przestrzeni wektorowej. Mówimy, że zbiór wektorów bazy napina przestrzeń wektorową.

Najprostszym i także najbardziej „praktycznym” przykładem bazy nieskończenie wielowymiarowej jest zbiór $\{\mathcal{X}_n \equiv x^n; n = 0, 1, 2, \dots\}$ napinający przestrzeń wielomianów \mathcal{W}_n , gdzie n może stawać się dowolnie wielkie. Oczywiście są twierdzenia:

Twierdzenie 6.10 Każdy wektor z \mathbb{R}^n ma jednoznaczne przedstawienie w postaci liniowej kombinacji wektorów bazy a także

Twierdzenie 6.11 W przypadku przestrzeni wektorowej \mathbb{R}^n liczba elementów bazy nie zależy od wyboru bazy i jest równa wymiarowi przestrzeni (n).

Rozdział 7

Algebra macierzy

7.1 Macierze; podstawowe definicje i operacje

W podrozdziale 4.5 omawialiśmy transformacje jakim podlegają współrzędne wektora \mathbf{V} (tam używaliśmy symbolu \mathbf{A}), wyrażone w dwóch różnych układach współrzędnych. Przepiszmy teraz zespół trzech równań 4-51 w nieco innej postaci, traktując zarówno współrzędne wektora w „starym” ($\Sigma - V_1, V_2, V_3$) i „nowym” ($\Sigma' - V'_1, V'_2, V'_3$) układzie jak i współczynniki a_{ik} (kosinusy kierunkowe) jako *macierze*, podlegające pewnemu działaniu – *mnożeniu macierzy*. Współrzędne wektora \mathbf{V} zapisujemy w formie macierzy — tablicy o jednej kolumnie i trzech wierszach, współczynniki a_{ik} – w formie *macierzy kwadratowej*, o trzech wierszach i trzech kolumnach. Równanie 4-51 przybiera postać

$$\begin{pmatrix} V'_1 \\ V'_2 \\ V'_3 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \\ V_3 \end{pmatrix}. \quad (7-1)$$

Jednocześnie wiemy, że obowiązują wzory

$$V'_n = \sum_{k=1}^3 a_{nk} V_k; \quad n = 1, 2 \text{ i } 3. \quad (7-2)$$

Zespół dwóch powyższych wzorów traktujemy jako definicję *mnożenia dwóch macierzy*: jednokolumnowa macierz współrzędnych wektora w układzie Σ

$$\mathcal{V} \equiv \begin{pmatrix} V_1 \\ V_2 \\ V_3 \end{pmatrix} \quad (7-3)$$

zostaje pomnożona przez macierz współczynników

$$\mathcal{A} \equiv \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \quad (7-4)$$

a w wyniku powstaje jednokolumnowa macierz współrzędnych wektora w układzie Σ'

$$\mathcal{V}' \equiv \begin{pmatrix} V'_1 \\ V'_2 \\ V'_3 \end{pmatrix}, \quad (7-5)$$

czyli, w skrócie

$$\mathcal{V}' = \mathcal{A} \cdot \mathcal{V}. \quad (7-6)$$

Mimo pozornego podobieństwa, macierze różnią się w sposób zasadniczy od wyznaczników. Te ostatnie były zawsze pewnymi liczbami, natomiast macierze są to pewne *operatory*, a raczej reprezentacje – tych operatorów. Mówiąc o „reprezentacji” mamy na myśli fakt, że zarówno współrzędne (primowane i nie-primowane) wektora, jak i kosinusy kierunkowe zależą od konkretnego wyboru dwóch układów współrzędnych. Macierz

\mathcal{A} współczynników reprezentuje operację obrotu¹, który przeprowadza układ Σ w układ Σ' , ale *konkretna postać współczynników* a_{ik} wynika z konkretnego wyboru tych dwóch układów – takich, a nie innych, kątów pomiędzy ich osiami.

Aby lepiej zrozumieć operację mnożenia macierzy, założymy, że chcemy wyrazić współrzędne wektora \mathbf{V} w układzie Σ'' , który powstaje w wyniku obrotu układu Σ' , podobnie jak ten ostatni powstał z obrotu układu Σ . Nie ma to wielkiego znaczenia czy rozpatrujemy przypadek w dwóch, czy trzech wymiarach. Dlatego możemy zrezygnować z górnej granicy sumowania w poniższych wzorach, pamiętając że – w zależności od sytuacji – będzie on równy 2 albo 3. Analogicznie do równanie 7-6 mamy

$$\mathcal{V}'' = \mathcal{B} \cdot \mathcal{V}', \quad (7-7)$$

gdzie macierz \mathcal{B} to macierz współczynników kosinusów kierunkowych dla pary układów: Σ'' i Σ' . Zgodnie ze wzorem 7-2

$$V_m'' = \sum_{n=1} b_{mn} V_n' = \sum_{n=1} b_{mn} \left(\sum_{k=1} a_{nk} V_k \right) = \sum_{k=1} \left(\sum_{n=1} b_{mn} a_{nk} \right) V_k. \quad (7-8)$$

Przestawienie kolejności sumowania (dozwolone; wartość sumy nie zależy od porządku składników) pozwala nam odczytać „w skrócie” powyższy wzór jako

$$V_m'' = \sum_{k=1} c_{mk} V_k; \quad \text{albo} \quad \mathcal{V}'' = \mathcal{C} \cdot \mathcal{V}, \quad (7-9)$$

gdzie macierz $\mathcal{C} = \mathcal{B}\mathcal{A}$ współczynników c_{mk} to *iloczyn dwóch macierzy*. Reguły mnożenia określa wzór 7-8

$$c_{mk} = \sum_n b_{mn} a_{nk}. \quad (7-10)$$

W powyższym wzorze „ruchomym” (sumacyjnym wskaźnikiem jest n – wskaźnik kolumn dla b_{mn} i wierszy dla a_{nk} ; kolejne wyrazy tworzące m -ty wiersz macierzy \mathcal{B} mnożymy przez odpowiednie wyrazy k -tej kolumny macierzy \mathcal{A} i sumujemy, aby znaleźć c_{mk} . Sposób ten nazywa się potocznie „mnożeniem wierszy przez kolumny” – wynika z niego, że chcąc pomnożyć jedną macierz przez drugą liczba kolumn pierwszej musi być równa liczbie wierszy drugiej.²

7.1.1 Dodawanie macierzy

Suma macierzy: $\mathcal{C} = \mathcal{A} + \mathcal{B}$ może być skonstruowana tylko dla macierzy \mathcal{A} i \mathcal{B} o takich samych liczbach wierszy (np. m) i kolumn (np. n). Element c_{ik} jest równy sumie $a_{ik} + b_{ik}$ dla *wszystkich* $i = 1, \dots, m$ i $k = 1, \dots, n$. Ponieważ elementy macierzy \mathcal{C} powstają w wyniku sumowania dwóch liczb, dodawanie macierzy posiada własności „zwykłego” dodawania: przemienność: $\mathcal{A} + \mathcal{B} = \mathcal{B} + \mathcal{A}$, oraz łączność $(\mathcal{A} + \mathcal{B}) + \mathcal{C} = \mathcal{A} + (\mathcal{B} + \mathcal{C})$. Istnieje także macierz zerowa \mathcal{O} , która dodana do dowolnej macierzy \mathcal{A} pozostawia ją bez zmian: $\mathcal{A} + \mathcal{O} = \mathcal{O} + \mathcal{A} = \mathcal{A}$; elementy macierzy \mathcal{O} są wszystkie zerami:

$$\mathcal{O} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{pmatrix}. \quad (7-11)$$

7.1.2 Mnożenie macierzy przez liczbę

Mnożenie macierzy \mathcal{A} przez liczbę α sprowadza się do przemnożenia *wszystkich elementów macierzy* przez tę liczbę

$$\alpha\mathcal{A} \equiv \begin{pmatrix} \alpha a_{11} & \alpha a_{12} & \dots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \dots & \alpha a_{2n} \\ \dots & \alpha a_{ik} & \dots & \dots \\ \alpha a_{m1} & \alpha a_{m2} & \dots & \alpha a_{mn} \end{pmatrix}. \quad (7-12)$$

¹Na płaszczyźnie, przejście jednego układu w drugi realizujemy przy pomocy obrotu o (jeden) kąt θ . W przestrzeni taki „obróć” to, w najogólniejszym przypadku, złożenie trzech sukcesywnych obrotów, względem trzech osi, o tzw. kąty Eulera.

²W odróżnieniu od wyznaczników macierz nie muszą być „kwadratowe” (o tych samych liczbach kolumn i wierszy). Zresztą – zaczęliśmy od wprowadzenia macierzy jednokolumnowej, której liczba wierszy jest równa wymiarowi naszej przestrzeni.

Takie mnożenie jest oczywiście przemienne

$$\alpha\mathcal{A} = \mathcal{A}\alpha. \quad (7-13)$$

Mnożenie macierzy przez liczbę różni się zasadniczo od analogicznej operacji dla wyznaczników (por. podrozdział 3.4– tu mnożymy wszystkie elementy macierzy, tam wystarczyło pomnożyć elementy jednego wiersza (lub kolumny)).

7.1.3 Mnożenie macierzy

Reguły mnożenia określa wzór 7-10

$$c_{mk} = \sum_n b_{mn}a_{nk}, \quad (7-14)$$

z którego wynika warunek równości liczb wierszy i kolumn w mnożnej i mnożniku. Dla lepszego zilustrowania rozważmy iloczyn dwóch kwadratowych (2×2) macierzy:

$$\begin{pmatrix} p & q \\ r & s \end{pmatrix} \begin{pmatrix} a & b \\ c & c \end{pmatrix} = \begin{pmatrix} pa + qc & pb + qc \\ ra + sc & rb + sc \end{pmatrix}. \quad (7-15)$$

Podobnie jak w przypadku operacji dodawania, mnożenie macierzy posiada także swój „element tożsamościowy”, to znaczy *macierz jednostkowa*, często zapisywaną jako delta Kroneckera (por. wzór 4-21):

$$\mathcal{I} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}. \quad (7-16)$$

Mnożąc przez taką macierz dowolną macierz \mathcal{A} ($m \times n$)³ mamy

$$\mathcal{I}\mathcal{A} = \mathcal{A}\mathcal{I} = \mathcal{A}. \quad (7-17)$$

W powyższym równaniu występuje iloczyn dwóch macierzy: \mathcal{I} oraz \mathcal{A} w różnej kolejności. Wynik – pozostawienie macierzy \mathcal{A} bez zmian nie zależy od tego czy \mathcal{A} jest mnożną czy mnożnikiem. Jest to jednak wypadek szczególny; na ogół, mnożenie macierzy **nie jest przemienne**, to znaczy

$$\mathcal{A}\mathcal{B} \neq \mathcal{B}\mathcal{A}. \quad (7-18)$$

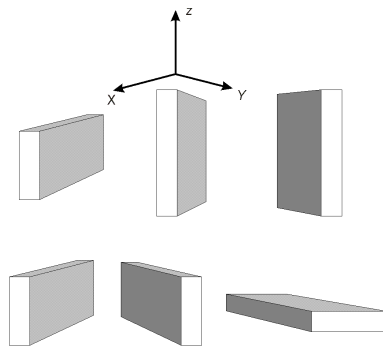
Powyższy postulat może się wydawać na pierwszy rzut oka zaskakujący. Jednak już prosty test pozwala nam przekonać, że tak jest w istocie; zamieniając w prezentowanym przed chwilą przykładzie (por. 7-15) porządek czynników mamy

$$\begin{aligned} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} p & q \\ r & s \end{pmatrix} &= \begin{pmatrix} ap + br & aq + bs \\ cp + dr & cq + ds \end{pmatrix} \\ &\neq \begin{pmatrix} pa + qc & pb + qc \\ ra + sc & rb + sc \end{pmatrix}. \end{aligned}$$

Z powyższego, prostego rachunku wynika potwierdzenie tezy 7-18, a także wniosek, że mnożenie macierzy *może* być przemienne (o ile elementy obu macierzy powiązane są pewnymi, dodatkowymi zależnościami). To, że w ogólnym przypadku $\mathcal{A}\mathcal{B} \neq \mathcal{B}\mathcal{A}$ nie powinno nas jednak dziwić. Powiedzieliśmy na początku tego rozdziału, że macierz może być operatorem obrotu. Jeżeli rozważać przypadek obrotów na płaszczyźnie to kolejność dwóch obrotów, na przykład o kąty α_1 i α_2 (wokół tej samej osi, prostopadłej do płaszczyzny) nie ma wpływu na ostateczny wynik – w obu przypadkach nastąpi obrót o kąt $\alpha_1 + \alpha_2$. Ale w przypadku obrotów w przestrzeni 3-wymiarowej już tak nie jest – wynik dwóch kolejnych obrotów, wokół różnych osi, zależy na ogół od porządku ich wykonania – por. Rys. 7.1. Zresztą – w definicji mnożenia macierzy, stwierdzeniu „mnożymy wiersze przez kolumny” kryje się asymetria – zamieniając kolejność mnożnej i mnożnika zamieniamy ze sobą wiersze i kolumny, których elementy przecież na ogół nie są takie same.

Zależność iloczynu macierzy od kolejności jego czynników nie jest jedynym szczegółem, który odróżnia *algebrę macierzy* od algebry liczb (rzeczywistych, bądź zespolonych). Byliśmy zawsze przyzwyczajeni, że aby iloczyn

³Oczywiście przy spełnieniu warunku: liczba kolumn lewego czynnika = liczbie wierszy prawego czynnika.



Rysunek 7.1: Wynik złożenia dwóch obrotów o kąt $\pi/2$ wokół osi $0y$ i $0z$ zależy od ich kolejności .

dwóch czynników był równy zeru przynajmniej jeden z jego czynników musi być zerem. W przypadku macierzy tak być nie musi; na przykład iloczyn macierzy

$$\mathcal{A} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{oraz} \quad \mathcal{B} = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}$$

jest równy zeru: $\mathcal{A}\mathcal{B} = 0$ (choć $\mathcal{B}\mathcal{A} \neq 0$.) Mnożenie macierzy spełnia „zwykłe” własności łączności: $(\mathcal{A}\mathcal{B})\mathcal{C} = \mathcal{A}(\mathcal{B}\mathcal{C})$ oraz rozdzielności $\mathcal{A}(\mathcal{B} + \mathcal{C}) = \mathcal{A}\mathcal{B} + \mathcal{A}\mathcal{C}$. Kolejność czynników musi być jednak zawsze zachowana.

7.1.4 Macierz transponowana

Jeżeli w macierzy zamienić wiersze na kolumny i odwrotnie powstała macierz nazywana jest *transponowaną* (w stosunku do macierzy wyjściowej). Mieliśmy już okazję spotkać się z takimi macierzami (por. równanie 2-75, wiążące ze sobą współczynniki dwóch macierzy, z których jedna reprezentowała obrót przeprowadzający układ Σ w układ Σ' , a druga – obrót przeprowadzający układ Σ' w układ Σ). Macierz transponowaną względem macierzy \mathcal{A} oznaczyliśmy górnym indeksem T :

$$\mathcal{B} = \mathcal{A}^T \rightarrow b_{ik} = a_{ki}. \quad (7-19)$$

Macierzą transponowaną w stosunku do jednokolumnowej macierzy (np. współrzędnych wektora \mathbf{A}) będzie macierz jednowierszowa; dla przestrzeni 3-wymiarowej

$$\mathcal{A}^T \equiv \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix}^T = (A_1, A_2, A_3). \quad (7-20)$$

Iloczyn powyższych macierzy to kwadrat długości macierzy-wektora \mathbf{A}

$$(A_1, A_2, A_3) \cdot \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} = A_1^2 + A_2^2 + A_3^2 = A^2. \quad (7-21)$$

Z powyższego równania wynika, że chcąc zapisać iloczyn skalarny dwóch wektorów w postaci „macierzowej” musimy jeden z nich przedstawiać – tak jak dotąd – w postaci kolumny, ale drugi w postaci wiersza! W rozdziale dziesiątym będziemy mówić szerzej o tym problemie.

Jeżeli macierz \mathcal{C} jest iloczynem dwóch macierzy: $\mathcal{C} = \mathcal{A}\mathcal{B}$, to macierz transponowana do \mathcal{C} jest iloczynem dwóch macierzy transponowanych, ale w odwrotnej kolejności:

$$\mathcal{C}^T \equiv (\mathcal{A}\mathcal{B})^T = \mathcal{B}^T\mathcal{A}^T. \quad (7-22)$$

Powyższe równanie można sprawdzić bezpośrednim rachunkiem, ale w następnym podrozdziale podamy jego prostsze uzasadnienie.

7.1.5 Macierz odwrotna

Macierz odwrotną (do danej macierzy \mathcal{A}) definiujemy jako macierz \mathcal{A}^{-1} , taką że:

$$\mathcal{A}\mathcal{A}^{-1} = \mathcal{A}^{-1}\mathcal{A} = \mathcal{I}, \quad (7-23)$$

gdzie \mathcal{I} jest macierzą jednostkową (por. 7-13). Sens równania 7-23 jest oczywisty – złożenie dwóch operacji: „wprost” i „odwrotnej” powoduje powrót do stanu wyjściowego. Dlatego iloczyn 7-23 *nie powinien* zależeć od kolejności czynników.

Aby dla macierzy \mathcal{A} istniała macierz odwrotna, \mathcal{A} musi być macierzą kwadratową $n \times n$. (Wynika to z możliwości zamiany kolejności mnożenia – wiersze mogą być zamienione z kolumnami, ale w takim razie liczba jednych i drugich musi być taka sama.) Z kolei, z macierzą kwadratową kojarzymy wyznacznik: $|A|$. Jak zaraz zobaczymy, podstawowym warunkiem aby istniała macierz odwrotna jest, aby **wyznacznik macierzy był różny od zera**. (A więc rząd macierzy jest równy n .) Taką macierz nazywamy *niesobliwą*. Macierz, której wyznacznik jest równy zeru nazywamy *macierzą sobliwą* – dla takiej macierzy nie istnieje macierz odwrotna. Spróbujmy znaleźć przepis na określenie elementów macierzy odwrotnej. W tym celu rozważmy macierz

$$\mathcal{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \quad (7-24)$$

oraz macierz dopełnień algebraicznych elementów \mathcal{A} , tzw. macierz sprzężoną $\tilde{\mathcal{A}}$, w której dopełnienie algebraiczne elementu a_{ij} (i -ty wiersz, j -a kolumna) leży na przecięciu j -ego wiersza i i -tej kolumny⁴:

$$\tilde{\mathcal{A}} = \begin{pmatrix} M^{[11]} & M^{[21]} & \dots & M^{[n1]} \\ M^{[12]} & M^{[22]} & \dots & M^{[n2]} \\ \vdots & \vdots & \vdots & \vdots \\ M^{[1n]} & M^{[2n]} & \dots & M^{[nn]} \end{pmatrix}. \quad (7-25)$$

Iloczyn tych macierzy, macierz $n \times n$, jest równy

$$\mathcal{A}\tilde{\mathcal{A}} = \tilde{\mathcal{A}}\mathcal{A} = \begin{pmatrix} d & 0 & \dots & 0 \\ 0 & d & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & d \end{pmatrix}, \quad (7-26)$$

gdzie d to wyznacznik macierzy \mathcal{A} . Rzeczywiście, mnożąc pierwszy wiersz \mathcal{A} przez pierwszą kolumnę $\tilde{\mathcal{A}}$ mamy

$$a_{11}M^{[11]} + a_{12}M^{[12]} + \dots + a_{1n}M^{[1n]} = d$$

(element (11) macierzy-iloczynu); podobnie mnożąc drugi ($\dots n$ -ty) wiersz \mathcal{A} przez drugą ($\dots n$ -tą) kolumnę $\tilde{\mathcal{A}}$

$$\begin{aligned} a_{21}M^{[21]} + a_{22}M^{[22]} + \dots + a_{2n}M^{[2n]} &= d \\ &\dots \dots \dots \\ a_{n1}M^{[n1]} + a_{n2}M^{[n2]} + \dots + a_{nn}M^{[nn]} &= d \end{aligned}$$

(elementy (22), \dots , (nn) macierzy-iloczynu). Pozostałe elementy iloczynu są równe zeru, na mocy własności wykazanej w punkcie 3.4 (mnożenie i sumowanie elementów danego kolumny przez dopełnienia algebraiczne innego wiersza). Porównując wzory 7-23 i 7-26 widzimy, że element i -tego wiersza i k -tej kolumny macierzy \mathcal{A}^{-1} jest równy dopełnieniu algebraicznemu $M^{[ki]}$ macierzy \mathcal{A} , podzielonemu przez wartość jej wyznacznika

$$a_{ik}^{-1} = \frac{M^{[ki]}}{d}. \quad (7-27)$$

⁴Jest to więc macierz dopełnień algebraicznych elementów macierzy transponowanej \mathcal{A}^T .

Widać, że różny od zera wyznacznik d jest warunkiem koniecznym; łatwo można wykazać, że jest to i warunek wystarczający dla istnienia macierzy odwrotnej.

Prosty na pozór wzór 7-27 jest jednak mało przydatny w praktyce. Podobnie jak w przypadku rozwiązywania układów równań liniowych, tak i tu obliczanie wyznacznika lub minora (zwróćmy uwagę na przestawione wskaźniki!) może okazać się operacją trudną do wykonania. Dlatego w praktyce posługujemy się inną metodą znajdowania macierzy odwrotnej (por. następny podpunkt).

Z kolei, znajomość wzorów na współczynniki macierzy odwrotnej pozwala nam w prosty sposób wykazać prawdziwość wzorów Cramera, na rozwiązanie układu równań niejednorodnych (por. 3.3). Układ taki

$$\begin{aligned} a_{11}x_1 &+ a_{12}x_2 &+ \dots &+ a_{1n}x_n &= b_1, \\ a_{21}x_1 &+ a_{22}x_2 &+ \dots &+ a_{2n}x_n &= b_2, \\ \dots &+ \dots &+ \dots &+ \dots &= \dots, \\ a_{m-1,1}x_1 &+ a_{m-1,2}x_2 &+ \dots &+ a_{m-1,n}x_n &= b_{1,m-1}, \\ a_{m1}x_1 &+ a_{m2}x_2 &+ \dots &+ a_{mn}x_n &= b_m. \end{aligned} \quad (7-28)$$

zapisać możemy w postaci *równania macierzowego*

$$\mathcal{A}\mathcal{X} = \mathcal{B}, \quad (7-29)$$

gdzie macierz \mathcal{A} to macierz współczynników a_{ik} , a macierze \mathcal{X} i \mathcal{B} to jednokolumnowe macierze:

$$\mathcal{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathcal{B} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}. \quad (7-30)$$

Iloczyn $\mathcal{A}\mathcal{X}$ ma sens (odpowiednie liczby wierszy i kolumn; jest to jednokolumnowa macierz).

Zakładając, że wyznacznik macierzy \mathcal{A} , $|\mathcal{A}| = d \neq 0$ możemy skonstruować macierz \mathcal{A}^{-1} ; mnożąc obie strony 7-29 przez \mathcal{A}^{-1} (od lewej) mamy

$$\mathcal{X} = \mathcal{A}^{-1}\mathcal{B}. \quad (7-31)$$

Prawa strona powyższego równania to jednokolumnowa macierz; element jej j -ego wiersza to suma iloczynów elementów j -ego wiersza macierzy \mathcal{A}^{-1} i odpowiednich elementów \mathcal{B} ; konkretnie

$$\begin{aligned} (\mathcal{A}^{-1}\mathcal{B})_j &= \frac{M^{[1j]}}{d}b_1 + \frac{M^{[2j]}}{d}b_2 + \dots + \frac{M^{[nj]}}{d}b_n \\ &= \frac{1}{d} (M^{[1j]}b_1 + M^{[2j]}b_2 + M^{[nj]}b_n). \end{aligned} \quad (7-32)$$

Ale wyrażenie w nawiasie po prawej stronie to rozwinięcie – względem j -ej kolumny – wyznacznika, który otrzymujemy zastępując j -ą kolumnę w wyznaczniku d macierzy \mathcal{A} kolumną wyrazów wolnych – \mathcal{B} (por. punkt 3.4). Uzyskujemy w tym momencie prawie pełny dowód uzasadnienie twierdzenia Cramera; należałoby jeszcze „formalnie” wykazać, że określone wzorem 7-32 wielkości to rzeczywiście rozwiązanie układu 7-28. Wystarczy wyrażenie na \mathcal{X} (wzór 7-31) wstawić do macierzowego równania 7-29, aby uzyskać tożsamość: $\mathcal{B} = \mathcal{B}$.

Powracamy do zagadnienia macierzy odwrotnej. Dla macierzy będącej iloczynem dwóch (lub więcej) czynników mamy:

$$(\mathcal{A}\mathcal{B})^{-1} = \mathcal{B}^{-1}\mathcal{A}^{-1}. \quad (7-33)$$

Ten wzór udowadniamy mnożąc lewą i prawą stronę przez iloczyn $(\mathcal{A}\mathcal{B})$. Lewa strona to $(\mathcal{A}\mathcal{B})(\mathcal{A}\mathcal{B})^{-1} = \mathcal{I}$; prawa to

$$\mathcal{A}\mathcal{B}\mathcal{B}^{-1}\mathcal{A}^{-1} = \mathcal{A}(\mathcal{B}\mathcal{B}^{-1})\mathcal{A}^{-1} = \mathcal{A}\mathcal{I}\mathcal{A}^{-1} = \mathcal{A}\mathcal{A}^{-1} = \mathcal{I}. \quad (7-34)$$

W tym momencie możemy też dostrzec zasadność wzoru 7-22. Z podrozdziału 4.5 bowiem wynika, że *macierzą odwrotną do macierzy obrotu jest macierz transponowana*. Jeżeli tak, to reguła zmiany kolejności czynników przy obliczaniu macierzy odwrotnej (udowodniona powyżej) musi mieć zastosowanie także w przypadku konstrukcji macierzy transponowanej. Ale własność $\mathcal{A}^{-1} = \mathcal{A}^T$, albo (por 2-75)

$$a_{kn}^{-1} = a_{nk} \quad (7-35)$$

nie jest zawsze spełniona! Posiadają ją tylko pewne macierze, albo operacje przez te macierze reprezentowane. Jest to tzw. własność ortogonalności, o której szerzej powiemy w rozdziale dziesiątym.

Pomimo podkreślonej różnicy między macierzą a wyznacznikiem widzimy, że związki pomiędzy nimi (oczywiście, tylko w przypadku macierzy kwadratowych) są bardzo ściśle. Mamy

Twierdzenie 7.1 (o wyznaczniku iloczynu macierzy) *Wyznacznik macierzy będącej iloczynem dwu (kilku) macierzy jest równy iloczynowi wyznaczników poszczególnych macierzy:*

$$\mathcal{C} = \mathcal{A}\mathcal{B}, \quad (7-36)$$

to

$$\det(\mathcal{C}) = \det(\mathcal{A}) \det(\mathcal{B}). \quad (7-37)$$

Dowód tego intuicyjnie łatwego do przyjęcia twierdzenia jest mocno skomplikowany; dla dwóch macierzy $n \times n$, $\mathcal{A} \equiv a_{ij}$ i $\mathcal{B} \equiv b_{ij}$ można go przeprowadzić konstruując „macierz pomocniczą” o wymiarach $2n \times 2n$ i obliczając jej wyznacznik:

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 0 \\ -1 & 0 & \dots & 0 & b_{11} & b_{12} & \dots & b_{1n} \\ 0 & -1 & \dots & 0 & b_{21} & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & -1 & b_{n1} & b_{n2} & \dots & b_{nn} \end{vmatrix}. \quad (7-38)$$

(Pomocnicza macierz powstała przez złożenie czterech „ćwiartek”: dwie z nich (na przekątnej głównej) to macierze \mathcal{A} i \mathcal{B} , trzecia (prawy górny róg) – to macierz zerowa i czwarta (lewy dolny róg) to macierz diagonalna, z -1 na przekątnej.) Można pokazać, stosując z rozwinięcie Laplace’a względem n -tego wiersza, że wartość tego wyznacznika to

$$\Delta = \det \mathcal{A} \det \mathcal{B}. \quad (7-39)$$

Wyznacznik 7-38 poddajemy teraz całemu szeregowi transformacji, które (por. podrozdział 3.4) nie zmieniają jego wartości. Naszym celem jest zastąpienie wszystkich elementów b_{ij} z dolnej prawej ćwiartki zerami. Osiągamy to dodając do $(n+1)$ -ej kolumny kolumnę pierwszą pomnożoną przez b_{11} , drugą pomnożoną przez b_{21} , i tak dalej – aż do n -tej, pomnożonej przez b_{n1} . Następnie do $(n+2)$ -ej kolumny dodajemy: kolumnę pierwszą pomnożoną przez b_{12} , drugą pomnożoną przez b_{22} , i tak dalej – aż do n -tej, pomnożonej przez b_{n2} . Procedurę kontynuujemy – do każdej $(n+j)$ -ej kolumny dodajemy pierwsze n kolumn, przemnożonych przez współczynniki $b_{1j}, b_{2j}, \dots, b_{nj}$. Prawa dolna ćwiartka wyznacznika rzeczywiście wypełnia się zerami, natomiast prawa górna – jak można sprawdzić – wypełnia się elementami $c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \dots + a_{in}b_{nk}$. Po transformacji

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} & c_{11} & c_{12} & \dots & c_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} & c_{21} & c_{22} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} & c_{n1} & c_{n2} & \dots & c_{nn} \\ -1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & -1 & \dots & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & -1 & 0 & 0 & \dots & 0 \end{vmatrix}. \quad (7-40)$$

Pozostaje rozwinąć ten wyznacznik względem n -tej kolumny – jego wartość (po uwzględnieniu znaku) jest $\det \mathcal{C}$. (Powyższy dowód warto prześledzić krok po kroku, w celu nabycia biegłości w „rachunkach wyznaczników”.)

Oczywistym wnioskiem z twierdzenia o wyznaczniku iloczynu macierzy jest

$$\det(\mathcal{A}^{-1}) = \frac{1}{\det(\mathcal{A})}; \quad (7-41)$$

(iloczyn tych dwóch wyznaczników jest równy wyznacznikowi macierzy jednostkowej, którego wartość to oczywiście 1).

7.1.6 Macierz odwrotna – metoda Gaussa-Jordana

Technika Gaussa-Jordana znajdowania macierzy odwrotnej opiera się na równaniach 7-23 i 7-17:

$$A^{-1}A = \mathcal{I} \quad \text{i} \quad (7-33)$$

$$\mathcal{A}^{-1} \cdot \mathcal{I} = \mathcal{A}^{-1} \quad (7-27)$$

Macierz \mathcal{A} i macierz jednostkową poddajemy tej samej sekwencji operacji: z symetrii powyższych równań wynika, że jeżeli ta sekwencja doprowadzi do przekształcenia macierzy \mathcal{A} w macierz \mathcal{I} , to przekształcana równolegle macierz jednostkowa powinna stać się macierzą \mathcal{A}^{-1} . Prześledźmy tę technikę na prostym przykładzie. Macierz \mathcal{A} (3×3), dla której szukamy macierzy odwrotnej i macierz jednostkową poddajemy transformacjom analogicznym do tych, jakie stosowaliśmy w podrozdziale 3.2, a także 6.5.

Przykład 7.1 Macierze „wyjściowe” to:

$$\mathcal{A} = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 4 \end{pmatrix} \quad \text{i} \quad \mathcal{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Pierwszy krok to przekształcenie pierwszej kolumny \mathcal{A} tak, aby zawierała same jedności; w tym celu pierwszy wiersz dzielimy przez 3, a drugi przez 2 – operację przeprowadzamy na *obu* macierzach:

$$\begin{pmatrix} 1 & 2/3 & 1/3 \\ 1 & 3/2 & 1/2 \\ 1 & 1 & 4 \end{pmatrix} \quad \text{i} \quad \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Pierwszy wiersz odejmujemy od drugiego i trzeciego:

$$\begin{pmatrix} 1 & 2/3 & 1/3 \\ 0 & 5/6 & 1/6 \\ 0 & 1/3 & 11/3 \end{pmatrix} \quad \text{i} \quad \begin{pmatrix} 1/3 & 0 & 0 \\ -1/3 & 1/2 & 0 \\ -1/3 & 0 & 1 \end{pmatrix}.$$

Pierwsza kolumna transformowanej macierzy \mathcal{A} ma już szukaną postać: 1 na przekątnej i zera poza nią. Aby uzyskać to samo dla drugiej kolumny, dzielimy drugie wiersze obu macierzy przez 5/6 i odejmujemy tak przetransformowany wiersz, po uprzednim pomnożeniu go przez 2/3, od wiersza pierwszego, oraz – po uprzednim pomnożeniu go przez 1/3 – od wiersza trzeciego. Otrzymujemy

$$\begin{pmatrix} 1 & 0 & 1/5 \\ 0 & 1 & 1/5 \\ 0 & 0 & 18/5 \end{pmatrix} \quad \text{i} \quad \begin{pmatrix} 3/5 & -2/5 & 0 \\ -2/5 & 3/5 & 0 \\ -1/5 & -1/5 & 1 \end{pmatrix}.$$

Druga kolumna uzyskała żadaną postać. Pozostaje wykonać analogiczną sekwencję dla trzeciej: dzielimy trzecie wiersze przez 18/5 i „nowy” trzeci wiersz odejmujemy, po uprzednim pomnożeniu przez 1/5, od pierwszego i drugiego. Macierz \mathcal{A} została przetransformowana do macierzy jednostkowej, a macierz jednostkowa – do macierzy \mathcal{A}^{-1} :

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{i} \quad \begin{pmatrix} 11/8 & -7/18 & 0 \\ -1/18 & -7/18 & -1/18 \\ -1/18 & -1/18 & 5/18 \end{pmatrix}.$$

Podkreślmy raz jeszcze, że metoda eliminacji jest znacznie skuteczniejsza i „mniej niebezpieczna” (z uwagi na mogące pojawiać się problemy numeryczne) niż pozornie „bardziej elegancki” wzór 7-27. Ten ostatni warto stosować w zasadzie tylko do znajdowania macierzy odwrotnej o wymiarach 2×2 .

7.1.7 Macierz i jej wyznacznik; interpretacja geometryczna wyznacznika

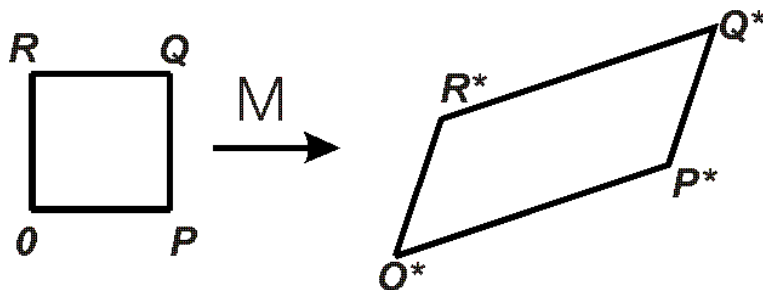
W podrozdziale 4.5 poznaliśmy jawną postać *macierzy obrotu* układu współrzędnych (lub wektora). Ponieważ jest to macierz transformacji, której można przyporządkować transformację odwrotną to wyznacznik jest różny od zera i jest równy jedności:

$$\begin{vmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{vmatrix} = 1. \quad (7-42)$$

Macierz obrotu (układu Σ o kąt θ), której wyznacznik podany jest powyżej, transformuje współrzędne wektora: stare współrzędne $(A_x, A_y, \text{ w układzie } \Sigma)$ zostają zastąpione nowymi $(A'_x, A'_y, \text{ w układzie } \Sigma')$. Powiedzieliśmy, że zamiast obracać układ możemy obracać wektor; równanie 4-56 możemy interpretować jak skutek obrotu wektora \mathbf{A} o kąt $-\theta$.

Jeżeli współczynniki tworzące macierz obrotu – kosinusy kierunkowe, por. wzory punktu 4.5 – przemnożyć przez dodatkową liczbę α , to do obrotu dołączy się wydłużenie (skrócenie) wektora o czynnik α . Czy potrafimy jednak zinterpretować wynik działania na współrzędne wektora *dowolnej* macierzy 2×2 (cały czas rozważamy – dla prostoty – sytuację na płaszczyźnie)? Owszem, jeżeli prześledzimy skutki działania takiej macierzy na wersory osi – wektory bazy kanonicznej (por. 6.2): \mathbf{e}_1 i \mathbf{e}_2 . Ich macierzowy zapis to;

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{oraz} \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (7-43)$$



Rysunek 7.2: Przekształcenie wersorów osi realizowane przez macierz \mathcal{M} .

Jeżeli podziałać na nie macierzą

$$\mathcal{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad (7-44)$$

to

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix} \quad (7-45)$$

oraz

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} b \\ d \end{pmatrix}. \quad (7-46)$$

Dwa *wektory bazowe* płaszczyzny przekształciły się dwa nowe wektory:

$$\mathbf{w}_1 = \begin{pmatrix} a \\ c \end{pmatrix} \quad \text{oraz} \quad \mathbf{w}_2 = \begin{pmatrix} b \\ d \end{pmatrix}. \quad (7-47)$$

Nie będą to już – w ogólnym przypadku – wektory prostopadłe, natomiast będą one liniowo niezależne, a więc nie współliniowe, pod warunkiem, że wyznacznik macierzy 7-44 jest różny od zera. Wynika to z definicji rzędu macierzy, z rachunków wyznaczników, ale może być także uzasadnione prostym rozumowaniem „geometrycznym”: liniowa niezależność dwóch wektorów oznacza różną od zera wartość ich iloczynu wektorowego. Ten zaś, to (por. podrozdział 4.3)

$$\mathbf{w}_1 \times \mathbf{w}_2 = \mathbf{e}_3(ad - bc) = \mathbf{e}_3|A|. \quad (7-48)$$

Nieosobliwa macierz *uogólnionego obrotu* przeprowadza więc wektory bazy kanonicznej w dwa, liniowo niezależne, wektory \mathbf{w}_1 i \mathbf{w}_2 – tak jak przedstawia to Rys.7.2. Przy tym, jednostkowy kwadrat zbudowany na wersorach \mathbf{e}_1 i \mathbf{e}_2 staje się równoległobokiem, o polu równym wyznacznikowi $|A|$ – Rys.7.3. Zauważmy też, że „nowe” wektory to kolumny macierzy \mathcal{M} .

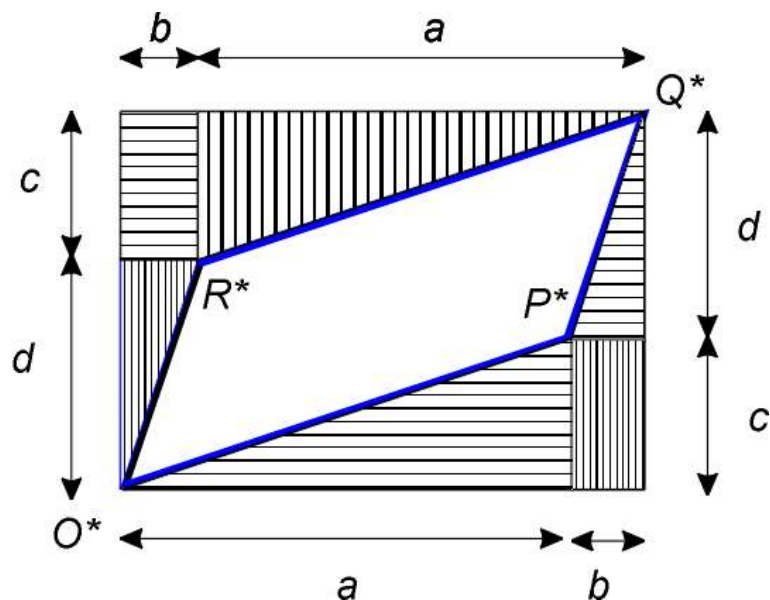
Mnożąc, przez macierz \mathcal{M} , wersory bazy kanonicznej otrzymujemy odpowiednie wektory-kolumny macierzy \mathcal{M} .

Uzyskujemy prostą i użyteczną interpretację wyznacznika: *wyznacznik $|A|$ określa stosunek, w jakim zmieniają się powierzchnie pod wpływem przekształcenie przestrzeni wektorowej, realizowanym przez macierz \mathcal{A} .*⁵ Dla opisywanego w podrozdziale 4.5 przekształcenia czystego obrotu powierzchnie nie ulegały zmianie – obrócony kwadrat jednostkowy miałby w dalszym ciągu powierzchnię równą 1. W przypadku ogólnym mamy dodatkowo do czynienia z tzw. *skalowaniem* – zastąpienie starych wektorów bazowych przez nowy układ wektorów liniowo niezależnych wprowadza nowy „system jednostek” do naszej przestrzeni. Taki nowy system może też na przykład polegać na zastąpieniu centymetrów metrami (w pomiarach długości). Macierz odpowiadająca takiemu przekształceniu będzie miała postać diagonalną, z elementami na przekątnej równym stosunkowi nowych i starych długości.

7.2 Rząd iloczynu macierzy – twierdzenie Sylwestra

Z twierdzenia o wyznaczniku iloczynu macierzy wynika, że macierz będąca iloczynem kilku macierzy, z których choć jedna jest osobliwa, będzie też macierzą osobliwą. Trudniej jest coś powiedzieć o rzędzie macierzy-iloczynu i jego związku z rzędami macierzy-czynników. Na przykład mnożąc przez siebie dwie macierze,

⁵Zauważmy, że zgodnie z punktem 4.3 skorzystaliśmy z iloczynu wektorowego aby określić pewną powierzchnię.



Rysunek 7.3: Pole równoległoboku powstałego z kwadratu o boku 1 równe jest $ad - bc$.

których rząd jest równy jedności:

$$\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

dostajemy w pierwszym przypadku macierz, której rząd jest równy jedności, a w drugim – zeru. Obowiązuje

Twierdzenie 7.2 (Sylvestra) *Rząd macierzy-iloczynu nie może być większy od rzędów poszczególnych czynników iloczynu.*⁶

Z ciekawszym przypadkiem mamy do czynienia, jeżeli rozpatrujemy iloczyn dowolnej macierzy \mathcal{A} i nieosobliwej macierzy \mathcal{Q} . Wówczas:

Twierdzenie 7.3 *Rząd macierzy $\mathcal{A}\mathcal{Q}$ (także: $\mathcal{Q}\mathcal{A}$) jest równy rzędowi macierzy \mathcal{A} .*

Rzeczywiście, jeżeli iloczyn to

$$\mathcal{A}\mathcal{Q} = \mathcal{F}$$

to rząd macierzy \mathcal{F} nie może być większy od rzędu macierzy \mathcal{A} . Mnożąc powyższe równanie od prawej przez \mathcal{Q}^{-1} dostajemy

$$\mathcal{A} = \mathcal{F}\mathcal{Q}^{-1}$$

i – wykorzystując raz jeszcze twierdzenie Sylvestra – wnioskujemy, że rząd macierzy \mathcal{A} nie może być większy od rzędu macierzy \mathcal{F} . Z obu konkluzji wynika teza – rząd macierzy $\mathcal{F} = \mathcal{A}\mathcal{Q}$ jest równy rzędowi macierzy \mathcal{A} .

⁶Bez dowodu. Twierdzenie jest zresztą łatwe do przyjęcia – rząd macierzy jest miarą niezależności wektorów-kolumn; trudno sobie wyobrazić, aby mnożenie układów o określonych stopniach tej niezależności mogło dać w wyniku układ o wyższym stopniu niezależności.

Rozdział 8

Formy kwadratowe

8.1 Wprowadzenie

Jednym z najbardziej chyba fascynujących rozdziałów historii nauk ścisłych są dzieje astronomii. Kilka tysięcy lat obserwacji, wniosków jakościowych, przekształcających się stopniowo w ilościowe. Te pierwsze wnioski ilościowe, choć zupełnie błędne („precyzyjny” system geocentryczny Ptolemeusza, aleksandryjskiego astronoma z 2. wieku po Chrystusie), pobiły swoisty rekord w dziejach nauki – przez prawie półtora tysiąca lat przyjmowane były powszechnie jako opisujące rzeczywistość. Ba, jeszcze w połowie 18. wieku (!) wydawano pseudonaukowe dzieła¹, w których opisywano krążące wokół Ziemi Słońce, wokół którego z kolei krążyły Wenus i Merkury. Co do Marsa i Jowisza sprawa nie była do końca jasna – planety te krążyły według *Informacji* „w zasadzie” wokół Słońca...

Tymczasem już w połowie 16. wieku ukazały się przecież *De revolutionibus orbium coelestium* Kopernika, stanowiące (nie do końca jeszcze poprawne) podwaliny systemu heliocentrycznego, a w początkach 17. wieku, „matematyk cesarski”, Joachim Kepler, następca wielkiego Tychona Brahe, opublikował niewielkie, ale jakże ważne dzieło *Astronomia nova aitiologetos, seu Physica Coelestis, tradita commentariis de motibus stellae Martis, ex observationibus Tychonis Brahe*. W tej *Nowej astronomii, popularnie objaśnionej, czyli Fizyce Niebieskiej, z komentarzami o ruchach gwiazdy Mars, wedle obserwacji Tychona Brahe* Kepler określił dwa postulaty, które znamy obecnie pod nazwą pierwszego i drugiego prawa Keplera. Pierwszym z tych praw było stwierdzenie, że planety poruszają się wokół Słońca po elipsach.²

W tym momencie okazało się, że „akademickie” zainteresowanie (jeszcze starożytnych Greków !) krzywymi stożkowymi, a więc krzywymi jakie uzyskujemy przecinając powierzchnię stożka płaszczyzną stało się w pełni uzasadnione praktycznie. Elipsa to jedna z trzech (pozostałe to hiperbola i parabola) typów krzywych stożkowych. Ich równania, zapisane na płaszczyźnie Oxy , przy założeniu, że środek krzywej stożkowej pokrywa się z początkiem układu współrzędnych, to równanie w których występują drugie potęgi współrzędnych. Najprostszym takim równaniem jest równanie okręgu:

$$x^2 + y^2 = R^2, \quad (8-1)$$

zbioru punktów, których odległość od punktu zwanego środkiem jest wielkością stałą, zwaną promieniem okręgu R . Równanie (i pojęcie) okręgu łatwo uogólnić do przestrzeni 3-wymiarowej; równanie sfery o promieniu R to

$$x^2 + y^2 + z^2 = R^2. \quad (8-2)$$

Z kolei przejście od okręgu do elipsy (od sfery do elipsoidy) polega na pewnym „przeskalowaniu” osi. Załóżmy, że wychodzimy od sfery jednostkowej ($R = 1$) a następnie zmniejszamy jednostkę skali a razy na osi x -ów, b razy na osi y -ów i c razy na osi z -ów. Innymi słowy, trójka $[x, y, z]$ przechodzi w trójkę $[X, Y, Z]$, przy czym zachodzą równości:

$$X = ax, \quad Y = by, \quad Z = cz.$$

¹*Informacja Matematyczna, rozumnie (? – AL) Ciekawego Polaka, świat cały, niebo y ziemię, y co na nich iest, W trudnych Kwestyach y Praktyce iemuż Ułatwiająca*, Przez W.X. Woyciecha Bystrzonowskiego Theologa Soc. Jesu, Roku 1749 drugi raz przedrukowana.

²Drugie prawo K. to prawo zachowania prędkości polowej: promień wodzący planety zakreśla w równych odstępach czasu równe pola. Na trzecie prawo, wiążące kwadraty pól osi eliptycznych torów planet z sześcianami okresów ich obiegu wokół Słońca przyszło czekać jeszcze dobry dziesięć lat.

Wyliczając z powyższych wzorów trójkę x, y, z i podstawiając do równania sfery (8-1) mamy równanie elipsoidy

$$\left(\frac{X}{a}\right)^2 + \left(\frac{Y}{b}\right)^2 + \left(\frac{Z}{c}\right)^2 = 1, \quad (8-3)$$

które dla dwóch wymiarów redukuje się do równania elipsy

$$\left(\frac{X}{a}\right)^2 + \left(\frac{Y}{b}\right)^2 = 1. \quad (8-4)$$

W równaniach tych, a konkretnie w lewych stronach równań, zmienne x, y i z (X, Y , i Z) występują w drugiej potędze. Każda taka „lewa strona” reprezentuje okrąg lub elipsę (bądź sferę lub elipsoidę) – po prawej stronie mamy kwadrat promienia (okręgu lub sfery), albo jedynkę. W tym ostatnim przypadku możemy dokonać prostego przekształcenia – np. równ. 8-4 sprowadzić do postaci

$$\frac{b}{a}X^2 + \frac{a}{b}Y^2 = ab. \quad (8-5)$$

Stosunki a/b i b/a są bezwymiarowe; lewa strona to w dalszym ciągu suma kwadratów zmiennych z pewnymi wagami; natomiast prawą możemy interpretować jako „średnią geometryczną kwadratów półosi elipsy” ($\langle R^2 \rangle = \sqrt{a^2b^2} = ab$).³ Można więc powiedzieć, że lewe strony równań 8-2 i 8-3 reprezentują *rodziny* sfer i elipsoid o tym samym środku (początek układu współrzędnych) i różnych (średnich) promieniach. Wyrażenia takie nazywamy *formami kwadratowymi*. Nie trudno jednak zauważyć, że wybraliśmy je nieco zbyt proste. Równaniem, które traktować będziemy jako definicję ogólnej postaci formy kwadratowej, dla przypadku, kiedy liczba wymiarów naszej przestrzeni wektorowej $n = 2$, będzie

$$Ax^2 + 2Bxy + Cy^2 = D. \quad (8-6)$$

Oprócz wyrazów z kwadratami x -a i y -a mamy w nim wyraz „mieszany”: $2Bxy$, w którym obie zmienne występują w pierwszej potędze. Ich iloczyn więc będzie „wyrazem kwadratowym”. Zauważ, że jeżeli jednostką x -a i y -a jest np. metr, to wszystkie trzy wyrazy: x^2, y^2, xy wyrażone są w metrach kwadratowych.

Czy postać 8-6 różni się w sposób zasadniczy od postaci 8-4? Jeżeli nasz układ współrzędnych, w którym wyrażone są współrzędne x i y obrócić o kąt θ (por. punkt 4.5), to nowe współrzędne x' i y' powiązane są ze „starymi”

$$\left. \begin{aligned} x &= x' \cos \theta - y' \sin \theta \\ y &= x' \sin \theta + y' \cos \theta. \end{aligned} \right\} \quad (8-7)$$

Warto sprawdzić, że dla wprowadzając nowe stałe A' i C' , spełniające

$$\begin{aligned} A' &= A \cos^2 \theta + C \sin^2 \theta + 2B \sin \theta \cos \theta \\ C' &= A \sin^2 \theta + C \cos^2 \theta - 2B \sin \theta \cos \theta, \end{aligned}$$

gdzie $\theta = \frac{1}{2} \arctg[2B/(A - C)]$, równanie 8-6 (po podstawieniu za x i y z 8-7) przyjmuje postać

$$A'x'^2 + C'y'^2 = D, \quad (8-8)$$

– forma kwadratowa, zawierający wyraz mieszany $2Bxy$ przy transformacji liniowej współrzędnych 8-7, może zostać sprowadzona do postaci zawierającej wyłącznie kwadraty poszczególnych zmiennych, tzw. *postaci kanonicznej*.

8.2 Definicja i własności formy kwadratowej

Ogólną postać *formy kwadratowej n zmiennych* (a więc skojarzonej z \mathbb{R}^n – zmienne to współrzędne wektora takiej przestrzeni; $x_i, \quad i = 1, \dots, n$) otrzymujemy wprowadzając oznaczenie a_{ii} na współczynnik wyrazu x_i^2 oraz $2a_{ij}$ dla wyrazu $x_i x_j$, dla $i \neq j$. Ten ostatni wyraz spełnia $x_i x_j = x_j x_i$, a więc zamiast $a_{ij} x_i x_j$ możemy równie dobrze zapisać $a_{ji} x_j x_i$, pod warunkiem, że macierz współczynników a_{ij} jest *macierzą symetryczną* (por.

³Zastanów się, jak przekształcić 8-3, aby po prawej strony wystąpił „średni kwadrat promienia elipsoidy” ($\langle R^2 \rangle = \sqrt[3]{a^2 b^2 c^2}$).

rozdział dziesiąty): $a_{ij} = a_{ji}$, czyli $\mathcal{A} = \mathcal{A}^T$. Dlatego „mieszany” wyraz, ze wskaźnikami $i \neq j$ zapisujemy jako:

$$2a_{ij}x_i x_j = a_{ij}x_i x_j + a_{ji}x_j x_i. \quad (8-9)$$

Przy takich odznaczeniach ogólna i zwarta postać formy kwadratowej to

$$f = \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_i x_j. \quad (8-10)$$

Symetryczna macierz współczynników a_{ij} to macierz kwadratowej formy f , a rzęd tej macierzy r nazywamy rzędem formy kwadratowej. Jeżeli $r = n$, czyli macierz \mathcal{A} jest nieosobliwa, to i kwadratową formę określamy jako nieosobliwą.

Formę kwadratową możemy zapisać w postaci iloczynu trzech macierzy, z których jedna to macierz współczynników a_{ij} , o wymiarach $n \times n$, a dwie pozostałe to jednokolumnowa i n -wierszowa macierz-wektor \mathcal{X} i transponowana do niej, jednowierszowa (n kolumn) macierz \mathcal{X}^T . Mamy:

$$f = \mathcal{X}^T \mathcal{A} \mathcal{X}, \quad (8-11)$$

gdzie

$$\mathcal{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \quad (8-12)$$

oraz

$$\mathcal{X}^T = (x_1, x_2, \dots, x_n), \quad \mathcal{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}. \quad (8-13)$$

Rzeczywiście, iloczyn $\mathcal{A} \mathcal{X}$ to jednokolumnowa macierz

$$\begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \sum_{j=1}^n a_{2j}x_j \\ \vdots \\ \sum_{j=1}^n a_{nj}x_j \end{pmatrix}, \quad (8-14)$$

która po przemnożeniu od lewej przez \mathcal{X}^T daje nam w wyniku „jednoelementową” macierz (jeden wiersz, jedna kolumna) – prawą stronę definicji 8-10.

Jak zaznaczyliśmy na samym początku, postać formy kwadratowej zależy od układu współrzędnych, albo – bardziej ogólnie – od używanej bazy naszej \mathbb{R}^n . Przejście pomiędzy współrzędnymi wektorów \mathbb{R}^n wyrażonymi w jednej i drugiej bazie to liniowa transformacja, reprezentowana przez nieosobliwą macierz \mathcal{Q} (będziemy o tym mówić dokładniej w rozdziale następnym); możemy, na przykład, zapisać:

$$x_i = \sum_{k=1}^n q_{ik}y_k, \quad i = 1, 2, \dots, n. \quad (8-15)$$

Zapisując powyższe równanie w postaci macierzowej:

$$\mathcal{X} = \mathcal{Q} \mathcal{Y}, \quad (8-16)$$

a także (dla macierzy transponowanych

$$\mathcal{X}^T = \mathcal{Y}^T \mathcal{Q}^T, \quad (8-17)$$

i podstawiając z dwóch ostatnich równań do 8-11 otrzymujemy

$$f = \mathcal{Y}^T \mathcal{Q}^T \mathcal{A} \mathcal{Q} \mathcal{Y} \equiv \mathcal{Y}^T \mathcal{B} \mathcal{Y}, \quad (8-18)$$

gdzie macierz \mathcal{B} to

$$\mathcal{B} = \mathcal{Q}^T \mathcal{A} \mathcal{Q}. \quad (8-19)$$

Podobnie jak macierz \mathcal{A} jest to macierz symetryczna

$$\mathcal{B}^T = (\mathcal{Q}^T \mathcal{A} \mathcal{Q})^T = \mathcal{Q}^T \mathcal{A}^T (\mathcal{Q}^T)^T = \mathcal{Q}^T \mathcal{A} \mathcal{Q} = \mathcal{B}.$$

Z powyższych rachunków wynika:

Twierdzenie 8.1 *Forma kwadratowa n współrzędnych, określona macierzą \mathcal{A} , przekształca się – w wyniku poddania współrzędnych transformacji określonej macierzą \mathcal{Q} – w formę kwadratową n nowych współrzędnych; macierz nowej formy to $\mathcal{Q}^T \mathcal{A} \mathcal{Q}$.*

(Transformacja, której podlega macierz \mathcal{A} nazywa się *transformacją podobieństwa*. Powrócimy do niej w rozdziale dziesiątym.)

Jeżeli macierz transformacji \mathcal{Q} była nieosobliwa to – por. punkt 7.2 – rząd macierzy $\mathcal{Q}^T \mathcal{A} \mathcal{Q}$ jest równy rządowi macierzy \mathcal{A} .

Wynika stąd

Twierdzenie 8.2 *Rząd formy kwadratowej nie zmienia się jeżeli poddać ją transformacji, której macierz jest nieosobliwa.*

Powracając do dyskusji na początku tego rozdziału: spośród nieskończenie wielu transformacji szczególnie będą interesowały nas te, które przeprowadzają formę kwadratową 8-10 do postaci kanonicznej (por. 8-8), w której występują wyłącznie kwadraty współrzędnych wektora. Taka postać kanoniczna to

$$f = b_1 y_1^2 + b_2 y_2^2 + \dots + b_n y_n^2, \quad (8-20)$$

gdzie y_1, y_2, \dots, y_n są nowymi współrzędnymi. Nie wszystkie współczynniki b_1, b_2, \dots, b_n muszą być różne od zera. Ale z twierdzenia o rzędzie macierzy formy i z punktu 6.4 wynika, że musimy zawsze mieć r niezerowych współczynników. Zdiagonalizowana macierz formy kwadratowej w jej postaci kanonicznej to

$$\begin{pmatrix} b_1 & 0 & \dots & 0 \\ 0 & b_2 & \dots & 0 \\ & & \ddots & \\ 0 & 0 & \dots & b_n \end{pmatrix} \quad (8-21)$$

– spośród n współczynników r musi być różnych od zera.

Możliwość sprowadzenia formy kwadratowej do postaci kanonicznej określa się jako

Twierdzenie 8.3 (Lagrange’a) *Dla każdej formy kwadratowej istnieje baza, w której forma ta ma postać kanoniczną.*

Dowód tego twierdzenia stanowi zarazem ilustrację techniki uzyskiwania postaci kanonicznej. Aby go przeprowadzić, założmy, że $a_{11} \neq 0$ i przepiszmy 8-10 wydzielając wyrazy zawierające x_1 :

$$f = a_{11} x_1^2 + \sum_{k=2}^n 2a_{1k} x_1 x_k + \sum_{i=2}^n \sum_{j=2}^n a_{ij} x_i x_j. \quad (8-22)$$

Pierwsze dwa wyrazy (uwaga na znak sumy!) po prawej stronie znaku równości uzupełniamy do kwadratu:

$$\begin{aligned} a_{11} x_1^2 + \sum_{k=2}^n 2a_{1k} x_1 x_k &= a_{11} \left\{ \left[x_1^2 + 2 \sum_{k=2}^n \frac{a_{1k}}{a_{11}} x_1 x_k + \left(\sum_{k=2}^n \frac{a_{1k}}{a_{11}} x_k \right)^2 \right] - \left(\sum_{k=2}^n \frac{a_{1k}}{a_{11}} x_k \right)^2 \right\} \\ &= a_{11} \left(x_1 + \sum_{k=2}^n \frac{a_{1k}}{a_{11}} x_k \right)^2 - a_{11} \left(\sum_{k=2}^n \frac{a_{1k}}{a_{11}} x_k \right)^2. \end{aligned} \quad (8-23)$$

Kładąc

$$x'_1 = \sqrt{|a_{11}|} \left(x_1 + \sum_{k=2}^n \frac{a_{1k}}{a_{11}} x_k \right), \quad x'_k = x_k, \quad k = 2, \dots, n \quad (8-24)$$

i podstawiając do 8-22 otrzymujemy⁴

$$f = \text{sign}(a_{11}) x_1'^2 + \dots, \quad (8-25)$$

gdzie wyrazy opuszczone tworzą formę kwadratową nie zawierającą składowej $x_1'^2$. Powyższy schemat stosujemy ponownie, eliminując sukcesywnie wszystkie wyrazy zawierające iloczyny różnych współrzędnych.

Schemat dowodu dla formy w postaci ogólnej wymaga pewnej biegłości w rachunkach, w których mamy do czynienia ze znakami sum. W praktyce mamy do czynienia z formami kwadratowymi w której występują dwie, trzy zmienne. Spróbujmy prześledzić na w miarę prostym przykładzie sprowadzanie formy kwadratowej do postaci kanonicznej.

Zauważmy po pierwsze, że schemat dowodu wymaga aby $a_{11} \neq 0$. Jeżeli tak nie jest, to musi występować w formie pewien wyraz mieszany, zawierający współczynnik $a_{1k}; k = 2, \dots, n$ (inaczej bowiem w formie nie byłoby w ogóle wyrazów z x_1). Przypuśćmy więc, że forma składa się z wyrazu $2a_{12}x_1x_2$ i pozostałych wyrazów, z których każdy zawiera chociaż jedną z wielkości x_3, \dots, x_n . Dokonując nieosobliwej transformacji

$$x_1 = z_1 - z_2, \quad x_2 = z_1 + z_2, \quad x_k = z_k \quad \text{dla } k = 3, \dots, n \quad (8-26)$$

wyraz $2a_{12}x_1x_2$ przekształcamy do

$$2a_{12}x_1x_2 = 2a_{12}(z_1 - z_2)(z_1 + z_2) = 2a_{12}z_1^2 - 2a_{12}z_2^2. \quad (8-27)$$

Podkreślmy: transformacja jest nieosobliwa; jej macierz to

$$\begin{pmatrix} 1 & -1 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}; \quad (8-28)$$

wyznacznik tej macierzy jest równy 2. W formie pojawiają się kwadraty dwóch współrzędnych; łatwo zauważyć, że na pewno nie mogą one zredukować się do zera z pozostałymi składnikami formy (te ostatnie zawierają przynajmniej jedną z wielkości z_3, \dots, z_n).

Przykład 8.1 Mamy przekształcić do postaci kanonicznej formę

$$g = 2x_1x_2 - 6x_2x_3 + 2x_3x_1. \quad (8-29)$$

Pierwsza transformacja to

$$x_1 = y_1 - y_2, \quad x_2 = y_1 + y_2, \quad x_3 = y_3. \quad (8-30)$$

Macierz tej transformacji

$$\mathcal{X} = \mathcal{A}\mathcal{Y} \quad (8-31)$$

to

$$\mathcal{A} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (8-32)$$

Nowa postać formy to

$$g = 2y_1^2 - 2y_2^2 - 4y_1y_3 - 8y_2y_3. \quad (8-33)$$

Współczynnik przy y_1^2 jest różny od zera i aby wyeliminować $-4y_1y_3$ możemy zastosować równ. 8-24. Ale uwaga: w poprzednim kroku zmienne *aktualnej* formy wyrażaliśmy poprzez pewne nowe zmienne $x \rightarrow y$; teraz zmienne aktualne (y) służą do konstrukcji kolejnych zmiennych. Czyli

$$\mathcal{Y} = \mathcal{B}\mathcal{Z}; \quad \mathcal{Z} = \mathcal{B}^{-1}\mathcal{Y}. \quad (8-34)$$

Kładziemy

$$z_1 = \sqrt{2}(y_1 - y_3), \quad z_2 = y_2, \quad z_3 = y_3.$$

⁴Przypominam: $\text{sign}(a_{11})$ to znak (+ lub -) współczynnika a_{11} . Zaraz okaże się, że znaki współczynników formy są istotne!

Jawna postać macierzy tej transformacji (\mathcal{B}^{-1}) i macierzy do niej odwrotnej (\mathcal{B}) to

$$\mathcal{B}^{-1} = \begin{pmatrix} \sqrt{2} & 0 & -\sqrt{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathcal{B} = \begin{pmatrix} 1/\sqrt{2} & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Kolejną postać formy to

$$g = z_1^2 - 2z_2^2 - 2z_3^2 - 8z_2z_3;$$

ostatnia transformacja, analogiczna do poprzedniej to

$$\mathcal{Z} = \mathcal{C}\mathcal{T}; \quad \mathcal{T} = \mathcal{C}^{-1}\mathcal{Z}. \quad (8-35)$$

Kładziemy

$$t_1 = z_1, \quad t_2 = \sqrt{2}(z_2 + 2z_3), \quad t_3 = z_3.$$

Macierz tej transformacji (\mathcal{C}^{-1}) i macierz do niej odwrotna (\mathcal{C}) to

$$\mathcal{C}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2} & 2\sqrt{2} \\ 0 & 0 & 1 \end{pmatrix} \quad \mathcal{C} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & -2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Ostateczną postać formy

$$g = t_1^2 - t_2^2 + 6t_3^2. \quad (8-36)$$

Postać tę możemy uzyskać bezpośrednio, jeżeli do postaci 8-29 zastosować transformację, której (nieosobliwa) macierz jest iloczynem (sprawdź!)

$$\mathcal{A}\mathcal{B}\mathcal{C} = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 3 \\ 1/\sqrt{2} & 1/\sqrt{2} & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Odpowiada to podstawieniu za x_i , $i = 1, 2, 3$ w 8-29

$$\begin{aligned} x_1 &= \frac{1}{\sqrt{2}}t_1 - \frac{1}{\sqrt{2}}t_2 + 3t_3 \\ x_2 &= \frac{1}{\sqrt{2}}t_1 + \frac{1}{\sqrt{2}}t_2 - t_3 \\ x_3 &= t_3. \end{aligned}$$

Do zagadnienia redukcji form kwadratowych powrócimy jeszcze w rozdziale jedenastym.

8.3 Prawo bezwładności dla form kwadratowych

Kanoniczna postać do której sprowadzamy określoną nie jest „jedyną w swoim rodzaju”; każda forma kwadratowa może zostać zredukowana do postaci kanonicznej na szereg sposobów. Na przykład, omawianą przed chwilą formę 8-29 możemy stosując transformację

$$\begin{aligned} x_1 &= t_1 + 3t_2 + 3t_3 \\ x_2 &= t_1 - t_2 - 2t_3 \\ x_3 &= t_3. \end{aligned}$$

przekształcić do postaci kanonicznej

$$g = 2t_1^2 + 6t_2^2 - 8t_3^2,$$

która różni się od 8-30.

Powstaje więc naturalne pytanie: czy różne postaci kanoniczne mają pewne „wspólne” cechy, a także – jakie warunki muszą być spełnione, aby dana forma kwadratowa mogła zostać przekształcona w inną formę, przy

pomocy nieosobliwej transformacji liniowej. W przypadku tej ostatniej ograniczamy się do transformacji, których macierze składają się wyłącznie z elementów będących liczbami rzeczywistymi.

Z twierdzenia Lagrange'a wynika, że każda forma kwadratowa, której rząd równy jest r , a liczba występujących w niej niewiadomych wynosi n może zostać zredukowana do kanonicznej postaci

$$f = c_1x_1^2 + \dots + c_kx_k^2 - c_{k+1}x_{k+1}^2 - \dots - c_r x_r^2, \quad 0 \leq k \leq r, \quad (8-37)$$

gdzie stałe $c_1, \dots, c_k, c_{k+1}, \dots, c_r$ są wszystkie większe od zera (zauważ: w postaci 8-20 nie wypowiadaliśmy się na temat znaku stałych b_i). Co więcej, prosta transformacja

$$z_i = \sqrt{c_i}y_i, \quad \text{dla } i = 1, \dots, r; \quad z_i = y_i, \quad \text{dla } i = r + 1, \dots, n \quad (8-38)$$

srowadza formę 8-37 do tzw. postaci *normalnej* (niektóre podręczniki właśnie taką formę nazywają formą kanoniczną):

$$f = z_1^2 + \dots + z_k^2 - z_{k+1}^2 - \dots - z_r^2. \quad (8-39)$$

Liczby współczynników dodatnich $-k$ i ujemnych $-l = r - k$ to para liczb (k, l) – tzw. *sygnatura* formy kwadratowej.

Mamy

Twierdzenie 8.4 (o bezwładności form kwadratowych) *Sygnatura formy jest niezmiennikiem każdej nieosobliwej transformacji liniowej, tzn. nie zależy od wyboru bazy, w której dana forma ma postać kanoniczną.*

(Dowód pomijamy).

W niektórych podręcznikach sygnaturę formy definiuje się jako różnicę $k - l$. Nazwa „sygnatura” sugeruje (słusznie), że sygnatura formy nie zależy od bazy w jakiej ją wyrażamy, a także (niesłusznie!), że sygnatura jednoznacznie określa formę.

Forma kwadratowa a formy liniowe

W podrozdziale 6.1.1 wprowadziliśmy pojęcie formy liniowej, określonej w wektorowej przestrzeni \mathbb{R}^n . Mnożąc przez siebie dwie formy liniowe

$$\phi = a_1x_1 + a_2x_2 + \dots + a_nx_n \quad \text{i} \quad \psi = b_1x_1 + b_2x_2 + \dots + b_nx_n \quad (8-40)$$

otrzymujemy formę kwadratową $-f = \phi \cdot \psi$. Takie przedstawienie formy kwadratowej w postaci iloczynu dwóch form liniowych nie zawsze jest możliwe. Można wykazać, że

Twierdzenie 8.5 *Forma kwadratowa o współczynnikach rzeczywistych może zostać przedstawiona w postaci iloczynu dwóch form liniowych wtedy i tylko wtedy gdy jej rząd jest nie większy od jedności, a sygnaturą jest para równych sobie liczb $-(k, l = k)$.*

8.4 Formy kwadratowe określone dodatnio

Specjalne miejsce, szczególnie w zastosowaniach „fizycznych” zajmują – jak już zauważyliśmy – formy o współczynnikach rzeczywistych, o postaci normalnej składającej się z n wyrazów (jesteśmy w \mathbb{R}^n !), których wszystkie współczynniki są dodatnie. Formy takie nazywamy formami określonymi dodatnio; ich rząd jest równy n , a ich sygnatura to $(n, 0)$.

Obowiązuje

Twierdzenie 8.6 *Forma kwadratowa o współrzędnych x_1, x_2, \dots, x_n i rzeczywistych współczynnikach jest formą określoną dodatnio wtedy i tylko wtedy jeżeli dla wszystkich rzeczywistych wartości x_i , z których przynajmniej jedna jest różna od zera, wartość formy jest dodatnia.*

W konsekwencji, forma określona dodatnio ma następujące własności:

1. Wyznacznik macierzy formy dodatnio określonej jest różny od zera: $\det |a_{ik}| \neq 0$.

Dowód jest natychmiastowy. Gdyby wyznacznik był równy zeru to układ równań jednorodnych

$$\sum_{k=1}^n a_{ik}x_k = 0 \quad i = 1, \dots, n$$

czyli

$$\mathcal{A}\mathcal{X} = 0$$

miałby nietrywialne rozwiązanie – wektor

$$\mathcal{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

o nie wszystkich $x_i = 0$. W takim razie forma

$$\mathcal{X}^T \mathcal{A} \mathcal{X} = \mathcal{X}^T (\mathcal{A} \mathcal{X}) = (x_1, \dots, x_n) \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = 0,$$

co jest sprzeczne z założeniem o dodatniej określoności formy.

2. Wyznacznik macierzy formy dodatnio określonej jest większy od zera: $\det |a_{ik}| > 0$.

I tutaj dowód jest prosty: na mocy twierdzenia Lagrange'a formę (określoną macierzą współczynników \mathcal{A}) można przedstawić w postaci normalnej; wyznacznik macierzy takiej formy $\det(N) = 1$. Przypuśćmy, że (nieosobliwa) macierz transformacji podobieństwa (por. wzory 8-18, 8-19) to macierz \mathcal{Q} ; macierz naszej formy (w postaci normalnej) to macierz $\mathcal{N} = \mathcal{Q}^T \mathcal{A} \mathcal{Q}$, albo

$$\mathcal{A} = \mathcal{P}^T \mathcal{N} \mathcal{P}, \quad (8-41)$$

gdzie \mathcal{P} to macierz transformacji odwrotnej do transformacji $\mathcal{A} \rightarrow \mathcal{N}$; $\mathcal{P} = \mathcal{Q}^{-1}$. Z twierdzenia o iloczynie wyznaczników mamy

$$\det |A| = \det |P^T| \cdot \det |N| \cdot \det |P| = 1 \cdot (\det |P|)^2 > 0.$$

3. W końcu – bez dowodu – warto poznać prosty sposób identyfikacji „dodatniości” formy zapisanej w dowolnej postaci: kwadratowa forma f o współczynnikach rzeczywistych, we współrzędnych x_1, \dots, x_n jest określona dodatnio, jeżeli wszystkie *minory główne* macierzy formy \mathcal{A} są dodatnie. Minory główne macierzy \mathcal{A} , to minory rzędu $1, 2, \dots, n$, usytuowane w lewym, górnym rogu:

$$a_{11}, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \dots, \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \dots & \dots & \dots & \dots \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{vmatrix}, \dots, \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}. \quad (8-42)$$

Przykład 8.2 Rozpatrzmy formę

$$f = 5x_1^2 + x_2^2 + 5x_3^2 + 4x_1x_2 - 8x_1x_3 - 4x_2x_3. \quad (8-43)$$

Jej minory główne to

$$5, \quad \begin{vmatrix} 5 & 2 \\ 2 & 1 \end{vmatrix} = 1, \quad \begin{vmatrix} 5 & 2 & -4 \\ 2 & 1 & -2 \\ -4 & -2 & 5 \end{vmatrix} = 1,$$

wszystkie są dodatnie – forma jest określona dodatnio.

Przykład 8.3 Forma (bardzo podobna do 8-43)

$$f = 3x_1^2 + x_2^2 + 5x_3^2 + 4x_1x_2 - 8x_1x_3 - 4x_2x_3,$$

nie jest określona dodatnio. Drugi „z rzędu” minor główny

$$\begin{vmatrix} 3 & 2 \\ 2 & 1 \end{vmatrix} = -1$$

jest mniejszy od zera.

8.5 Formy kwadratowe w fizyce

Zanim przejdziemy do dalszych dyskusji formalnych własności form kwadratowych warto uzmysłowić sobie, że struktury te, bardzo istotne w różnych działach matematyki, mają pierwszorzędne znaczenie także w fizyce. Rozważmy na przykład układ dwóch drgających sprężyn⁵ Każda ze sprężyn ma pewien współczynnik sprężystości – odpowiednio k_1 i k_2 , a na końcu każdej sprężyny znajduje się masa: m_1 i m_2 (por. Rys. 12.2). Przesunięcia obu mas względem ich położenia równowagi to x_1 i x_2 . Jeżeli wprowadzić wektor \mathbf{X} :

$$\mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad (8-44)$$

to dwa równania określające siły działające na obie masy – siły harmoniczne: $F_1 = -k_1x_1$ oraz $F_2 = -k_2x_2$ możemy zapisać w postaci macierzowej:

$$\begin{pmatrix} F_1 \\ F_2 \end{pmatrix} = \begin{pmatrix} -k_1 & 0 \\ 0 & -k_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad (8-45)$$

albo:

$$\mathbf{F} = \mathcal{K} \mathbf{X}, \quad (8-46)$$

gdzie \mathbf{F} to wektor określający siły działające na masy, \mathbf{X} – ich przesunięcia, a \mathcal{K} nazwiemy macierzą sprężystości układu. Wiemy, że energia potencjalna układu sprężyn – praca wykonana *przeciwko* siłom sprężystości na przesunięciach x_1 i x_2 , to

$$E_p = \frac{1}{2}k_1x_1^2 + \frac{1}{2}k_2x_2^2, \quad (8-47)$$

albo w zapisie macierzowym:

$$E_p = -\frac{1}{2}(-k_1x_1, -k_2x_2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = -\frac{1}{2}(F_1, F_2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = -\frac{1}{2}(x_1, x_2) \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}. \quad (8-48)$$

Wykorzystując drugą część wzoru 8-48 i podstawiając za wektor \mathbf{F} z równania 8-45 otrzymujemy

$$E_p = -\frac{1}{2}(x_1, x_2) \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} = \frac{1}{2}(x_1, x_2) \begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \quad (8-49)$$

Energia potencjalna układu sprężyn i mas jest formą kwadratową położenia obu mas. Otrzymaliśmy dokładnie wyrażenie typu 8-11

$$\mathbf{X}^T \mathcal{M} \mathbf{X},$$

gdzie macierz \mathcal{M} to macierz opisująca własności układu – macierz współczynników sprężystości). Co więcej energia kinetyczna tego układu będzie też *formą kwadratową prędkości obu mas* – v_1 i v_2 :

$$E_k = \frac{1}{2}m_1v_1^2 + \frac{1}{2}m_2v_2^2 = \frac{1}{2}(v_1, v_2) \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}. \quad (8-50)$$

⁵Zagadnienie to jest szczegółowo omawiane w punkcie 12.1.

Obie energie to formy kwadratowe zmiennych : położenia (E_p) lub prędkości (E_k). Oprócz tych zmiennych argumentów form, występują w nich – w postaci współczynników liczbowych – pewne *parametry fizyczne* układu – współczynniki sprężystości sprężyn lub masy ciał. Warto zauważyć, że obie formy są kanoniczne – występują w nich tylko kwadraty zmiennych, a nie ma wyrazów mieszanych, typu x_1x_2 , czy v_1v_2 . Formalnie wynika to ze specyficznej postaci macierzy \mathcal{K} i \mathcal{M} – są to macierze diagonalne. Ma to też swoje uzasadnienie „fizyczne”, o którym usłyszysz na wykładzie z mechaniki teoretycznej.

Formą kwadratową jest także moment bezwładności bryły sztywnej. Zmiennymi formy są współrzędne (na przykład kartezjańskie) poszczególnych punktów ciała, albo – bardziej konkretnie – elementarnych mas, składających się na bryłę. O momencie bezwładności jako formie kwadratowej można przeczytać w punkcie 12.3.

Rozdział 9

Odwzorowania

9.1 Wprowadzenie

Opisywana i dyskutowana kilkakrotnie operacja obrotu wektora na płaszczyźnie jest prostym przykładem *transformacji liniowej* – obiektem transformacji jest wektor – element \mathbb{R}^2 , natomiast operatorem jest macierz obrotu (2×2). W takim kontekście operacja mnożenia wektora przez macierz-operator określamy jako transformację (odwzorowanie, funkcję określoną dla elementów) \mathbb{R}^2 .

Odwzorowanie polega na przyporządkowaniu elementom jednego zbioru (zbioru wyjściowego – *dziedziny*) elementów drugiego zbioru (zbioru wynikowego – *przeciwdziedziny*). Formalnie możemy zapisać odwzorowanie w postaci np. $T(x) = y$, gdzie x i y to elementy dziedziny i przeciwdziedziny, a T – transformacja (funkcja realizująca odwzorowanie); istnieje szkoła algebraiczna, która preferuje zapis $(x)T = y$. Obrót wektora na płaszczyźnie to przykład odwzorowania typu $\mathbb{R}^2 \rightarrow \mathbb{R}^2$: dwójce liczb (stare współrzędne wektora) przyporządkowujemy inną dwójkę nowych współrzędnych. Ale z odwzorowaniami mamy do czynienia na co dzień¹. – np. powiedzenie: „Jan, ojciec Jasia” jest też odwzorowaniem; rolę funkcji f spełnia słowo „ojciec”. Podobnie, odwzorowaniem $\mathbb{R}^3 \rightarrow \mathbb{R}$ może być końcowa nota zaliczenia ćwiczeń (liczba – element \mathbb{R}) jako średnia trzech not sprawdzianów pisanych przez studenta w trakcie semestru.

Czy każde przyporządkowanie może być uznane za odwzorowanie? W zasadzie tak. Wymagane jest określenie dziedziny i przeciwdziedziny, a także – w pierwszym rzędzie – jednoznacznych reguł, według których elementom pierwszej zostają przyporządkowane elementy drugiej. Nie mniej jednak możemy mieć do czynienia z różnymi typami odwzorowań, których określeniem zajmiemy się a następnym punkcie.

9.2 Iniekcja, suriekcja, biekcja. Izomorfizm

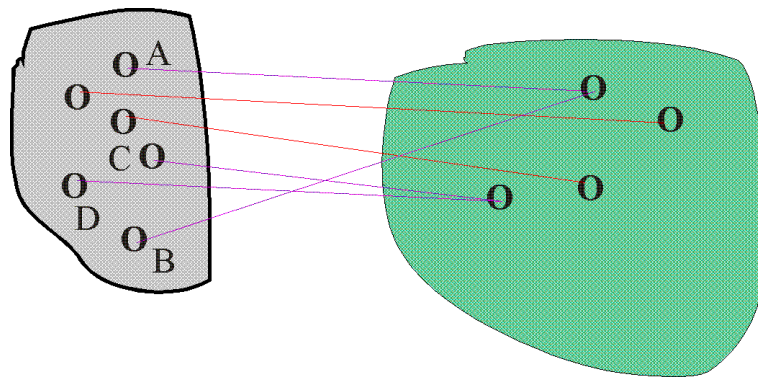
Rozpatrując odwzorowanie $T(x) = b$ zadajmy pytanie: czy to równanie posiada rozwiązanie dla *każdego* b – elementu przeciwdziedziny? Jeżeli tak, to odwzorowanie nazywamy *suriekcją* (odwzorowaniem suriektywnym). Odwzorowanie „ojciec”, którego dziedziną są wszyscy ludzie, a przeciwdziedzina – mężczyźni, nie jest suriekcją – nie każdy mężczyzna jest ojcem. Schematyczny przykład suriekcji daje rysunek 9.1

Kolejne pytanie: czy równanie $T(x) = b$ ma – dla określonego b – tylko jedno rozwiązanie, czy ma ich więcej? Jeżeli rozwiązanie jest jedno i tylko jedno, to odwzorowanie nazywamy *iniekcją* (odwzorowaniem iniektywnym). Równanie „Ojciec x -a to Jan Kowalski” nie będzie iniekcją – Jan K. może mieć kilkoro dzieci. Ale równanie „ y = brat bliźniak x -a”, ma – dla każdego y jedno i tylko jedno rozwiązanie (zauważ, że w tym przypadku dziedzina i przeciwdziedzina to te same zbiory). Iniekcją (ale nie suriekcją) może być np. identyfikator PESEL – nie wszystkie 11-cyfrowe liczby mają swoje odpowiedniki w zbiorze Polaków, ale dany PESEL należy do jednej i tylko jednej osoby. Schematyczny przykład iniekcji daje rysunek 9.2

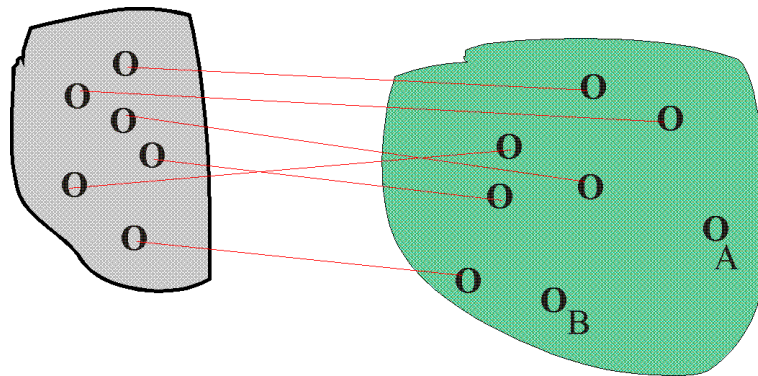
W końcu – najbardziej interesujący w zastosowaniach praktycznych typ – odwzorowanie biekcyjne, albo *biekcja*. Jest to odwzorowanie, które jest jednocześnie iniekcją i suriekcją: każdy element przeciwdziedziny powstaje w wyniku odwzorowania określonego elementu dziedziny.

Rozpatrując dwie (pod)przestrzenie wektorowe \mathbb{L} i \mathbb{L}' będziemy je nazywali *izomorficznymi* (*greckie – o tej samej formie, tej samej strukturze*), (albo będziemy mówili o izomorficznej odpowiedniości ich elementów) jeżeli

¹W sztuce Molière’a „Mieszczanin szlachcicem” bohater, pan Jourdain, dowiaduje się, że „mówił prozą całe życie, nie zdając sobie z tego sprawy”. Podobnie jest z odwzorowaniami w naszej codziennej rzeczywistości.



Rysunek 9.1: Surjekcja, ale nie injekcja: $T(A) = T(B)$ oraz $T(C) = T(D)$.



Rysunek 9.2: Injekcja, ale nie surjekcja; nie ma rozwiązań dla $T(x) = A$ i $T(x) = B$.

1. odwzorowanie między nimi jest bijekcją: każdy wektor \mathbf{a} należący do \mathbb{L} możemy skojarzyć z wektorem \mathbf{a}' z \mathbb{L}' (obrazem \mathbf{a}); różne wektory \mathbb{L} mają różne obrazy, a każdy wektor z \mathbb{L}' jest obrazem pewnego wektora w \mathbb{L} .
2. odwzorowanie jest liniowe: zachodzą zwykłe związki:

$$\begin{aligned}(\mathbf{a} + \mathbf{b})' &= \mathbf{a}' + \mathbf{b}' \\ (\alpha \mathbf{a})' &= \alpha \mathbf{a}'.\end{aligned}$$

Tak więc, dwuwymiarowa przestrzeń wektorów, wychodzących z jednego punktu płaszczyzny (początku układu) jest izomorficzna z przestrzenią \mathbb{R}^2 , której elementami są uporządkowane pary liczb rzeczywistych. Izomorficzne przyporządkowanie obu przestrzeni wymaga określenia na płaszczyźnie konkretnego układu współrzędnych.

Istotną własnością izomorfizmu w przestrzeniach wektorowych jest „wspólnota zera”: obrazem zera przestrzeni \mathbb{L} jest zero przestrzeni \mathbb{L}' . Rzeczywiście, jeżeli \mathbf{a}' jest obrazem \mathbf{a} , to mamy

$$\mathbf{a}' = (\mathbf{a} + \mathbf{0})' = \mathbf{a}' + \mathbf{0}',$$

co oznacza, że wektor $\mathbf{0}'$ jest zerem przestrzeni \mathbb{L}' .

9.2.1 Złożenie odwzorowań

Dwu- lub kilkakrotne zastosowanie odwzorowania, nazywamy złożeniem (lub iloczynem odwzorowań). Wprowadźmy oznaczenie:

$$(f \circ g)(x) \equiv f(g(x)), \tag{9-1}$$

gdzie f i g to dwa odwzorowania. Pozostając w kręgu przykładów „z życia wziętych” oznaczmy przez O odwzorowanie „ojciec”, a przez M – odwzorowania „matka”. Wówczas $O \circ M$ oznacza „dziadka po stronie

matki”, a $M \circ O$ – „babkę po stronie ojca”. Oczywiście jest, że porządek złożenia dwóch odwzorowań jest istotny: $O \circ M \neq M \circ O$; złożenie nie jest przemienne. Jest natomiast łączne:

$$O \circ (O \circ M) = (O \circ O) \circ M, \quad (9-2)$$

czyli: ojciec dziadka (ze strony matki) Jasia to dziadek (ze strony ojca) jego matki.

9.3 Odwzorowania i macierze

Tak jak obiektem odwzorowania może często być wektor – element \mathbb{R}^n , tak matematycznym ujęciem „reguł odwzorowania” może być macierz: mnożenie wektora \mathbf{V} przez macierz \mathcal{A} daje w wyniku nowy wektor \mathbf{V}' – obraz wektora \mathbf{V} . Takie odwzorowania posiadają własność liniowości: odwzorowanie sumy elementów to suma odwzorowań, odwzorowanie wielokrotności elementu to wielokrotność odwzorowania. Zanim przedyskutujemy to dokładnie w następnym punkcie prześledźmy prosty przykład.

Przypuśćmy, że firma produkująca posiłki serwowane w samolotach specjalizuje się w trzech rodzajach dań: (1) boeuf stroganoff, (2) pieczony kurczak z ryżem; (3) pierożki z mięsem. Aby wyprodukować określone liczby tych trzech dań należy zgromadzić odpowiednie „surowce”: pewne ilości (kilogramy, litry, itd.) wołowiny, kurczaków, mąki, śmietany, soli, itd.) Schemat „listy zakupów” mógłby wyglądać tak:

$$\begin{array}{l}
 \text{kg wołowiny} \rightarrow \\
 \text{kg kurczaka} \rightarrow \\
 \text{kg mąki} \rightarrow \\
 \text{kg ryżu} \rightarrow \\
 \vdots \\
 \text{litr śmietany} \rightarrow
 \end{array}
 \begin{array}{c}
 \left(\begin{array}{ccc}
 .15 & .10 & 0 \\
 0 & 0 & 0.20 \\
 \dots & \dots & \dots \\
 \vdots & \vdots & \vdots \\
 \dots & \dots & \dots
 \end{array} \right)
 \end{array}
 \begin{array}{c}
 \left(\begin{array}{c}
 60 \text{ b. stroganoff} \\
 30 \text{ pierożków} \\
 40 \text{ piecz. kurczak}
 \end{array} \right)
 \end{array}
 \begin{array}{c}
 \\
 \\
 \\
 \\
 \\
 \\
 \text{dania}
 \end{array}$$

$$= \begin{array}{c}
 \left(\begin{array}{c}
 12 \text{ kg wołowiny} \\
 6 \text{ kg kurczaka} \\
 \dots \text{ kg mąki} \\
 \dots \text{ kg ryżu} \\
 \vdots \\
 \dots \text{ ltr śmietany}
 \end{array} \right)
 \end{array}$$

W języku macierzy: wektor „zamówienie” \mathcal{Z} (określone liczby dań) odwzorowuje się na wektor „lista zakupów” \mathcal{L} poprzez macierz \mathcal{A} , która zawiera ilości składników, potrzebnych do wyprodukowania jednego dania:

$$\mathcal{L} = \mathcal{A}\mathcal{Z}.$$

Nawiązując do złożenia odwzorowań: zastanów się, jak będzie wyglądała operacja odwzorowania wektora listy zakupów na wektor-liczbę wydatek, a konkretnie: jaka postać będzie miała macierz odwzorowania².

Rozpatrzmy teraz przykład bardziej związany z „prawdziwą” przestrzenią wektorową.

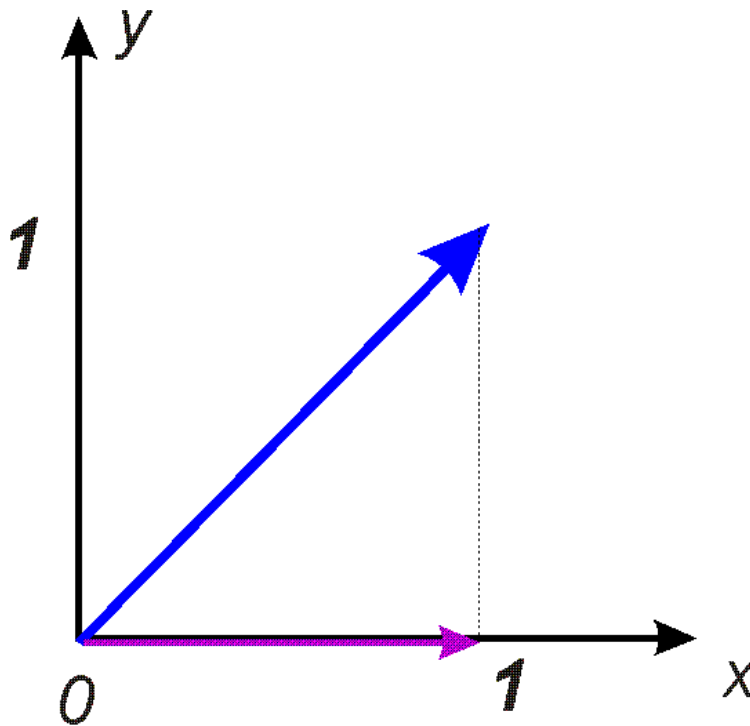
Przykład 9.1 Przypuśćmy, że dokonujemy odwzorowania, w wyniku któremu każdemu wektorowi leżącemu na płaszczyźnie i wychodzącemu z punktu 0, przyporządkowujemy liczbę, będącą x -ową współrzędną wektora – inaczej mówiąc, rzutujemy wektor \mathbf{v} na oś x -ów. (por. rys. 9.3)

Rzutowany wektor – element \mathbb{R}^2 – to $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Jego obraz – element \mathbb{R}^1 – to $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Nietrudno zauważyć, że macierz odwzorowania to $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, ponieważ

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \quad (9-3)$$

Jak znaleźć macierz transformacji w ogólnym przypadku? Obowiązuje (por. punkt 7.1.7)

²Musi to być macierz jednowierszowa o liczbie kolumn równej liczbie wierszy w macierzy \mathcal{L} ; poszczególne jej elementy to ceny jednostkowe artykułów.



Rysunek 9.3: Rzutowanie wektora na oś $0x$.

Twierdzenie 9.1 Każda liniowa transformacja (liniowe odwzorowanie) T , odwzorowująca przestrzeń \mathbb{R}^n w \mathbb{R}^m realizowana jest w wyniku pomnożenia wektorów należących do \mathbb{R}^n przez macierz T o wymiarach $m \times n$;

kolumnami macierzy są wektory-obrazy wektorów bazy kanonicznej w \mathbb{R}^n : $T(\mathbf{e}_i)$.

Dowód: Rozpatrywane odwzorowanie to równanie

$$\mathbf{y} = T \mathbf{v}, \quad (9-4)$$

albo *explicite*

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ t_{m1} & t_{m2} & \cdots & t_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} t_{11}v_1 + t_{12}v_2 + \dots + t_{1n}v_n \\ t_{21}v_1 + t_{22}v_2 + \dots + t_{2n}v_n \\ \vdots \\ t_{m1}v_1 + t_{m2}v_2 + \dots + t_{mn}v_n \end{pmatrix} \quad (9-5)$$

(\mathbf{v} – wektor przestrzeni \mathbb{R}^n ; \mathbf{y} – wektor przestrzeni \mathbb{R}^m ; $T \equiv (t_{ik})$ – macierz odwzorowania.)

Z drugiej strony, postać \mathbf{v} w bazie kanonicznej to

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = v_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + v_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \dots + v_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}. \quad (9-6)$$

(Po prawej stronie występują jednokolumnowe wektory bazy kanonicznej: $\mathbf{e}_1, \dots, \mathbf{e}_n$). Równanie 9-6 możemy zapisać w zwartej postaci

$$\mathbf{v} = v_1 \mathbf{e}_1 + v_2 \mathbf{e}_2 + \dots + v_n \mathbf{e}_n \equiv \sum_{i=1}^n v_i \mathbf{e}_i. \quad (9-7)$$

Z liniowości mamy

$$T(\mathbf{v}) = T\left(\sum_{i=1}^n v_i \mathbf{e}_i\right) = \sum_{i=1}^n v_i T(\mathbf{e}_i). \quad (9-8)$$

Dowód:

(a) jeżeli jądro odwzorowania T zawiera jakiś różny od zera wektor, to równanie $T\mathbf{x} = 0$ ma więcej niż jedno rozwiązanie (tym „drugim” jest zgodnie z definicją jądra wektor zerowy);

(b) jeżeli istnieją dwa różne wektory będące rozwiązaniem układu $T\mathbf{x} = \mathbf{b}$ to znaczy mamy:

$$T(\mathbf{x}_1) = \mathbf{b} = T(\mathbf{x}_2); \quad \mathbf{x}_1 \neq \mathbf{x}_2$$

to na mocy liniowości

$$T(\mathbf{x}_1 - \mathbf{x}_2) = T(\mathbf{x}_1) - T(\mathbf{x}_2) = \mathbf{b} - \mathbf{b} = 0$$

i wektor (różny od zera!) $\mathbf{x}_1 - \mathbf{x}_2$ stanowi element jądra.

Obrazem liniowego odwzorowania T , oznaczanym⁴ przez $\text{Img } T$, nazywamy zbiór wszystkich wektorów \mathbf{b} , dla których istnieje rozwiązanie równania $T\mathbf{x} = \mathbf{b}$.

Zauważmy, że na początku tego rozdziału używaliśmy już określenia obraz w definicji odwzorowania. Obrazem odwzorowania T jest zbiór wszystkich wektorów, które możemy traktować jako obrazy wektorów-elementów dziedziny T .

Przykład 9.2 Jeżeli macierzą T jest

$\begin{pmatrix} 1 & 3 \\ 2 & 0 \end{pmatrix}$ to wektor $\begin{pmatrix} 7 \\ 2 \end{pmatrix}$ jest obrazem T , bo

$$\begin{pmatrix} 1 & 3 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 7 \\ 2 \end{pmatrix}.$$

9.4.1 Baza $\text{Img } T$ i $\ker T$

Z definicji wielkości \ker i Img wynika, że jądro transformacji jest podzbiorem jej dziedziny; jeżeli transformacji T poddana jest przestrzeń wektorowa: $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$ to jądro T jest podprzestrzenią w \mathbb{R}^n . Analogicznie obrazem odwzorowania T jest pewna podprzestrzeń \mathbb{R}^m . Macierz transformacji – \mathcal{A} – to macierz $m \times n$. Powstaje pytanie: czy znając macierz \mathcal{A} możemy zidentyfikować – w możliwie prosty sposób – obie podprzestrzenie? Zidentyfikowanie podprzestrzeni może nastąpić – na przykład – poprzez określenie ich wektorów bazowych. W punkcie 6.6, omawiając metodę Gaussa-Jordana rozwiązywania układów równań liniowych, mówiliśmy o przekształcaniu macierzy do uogólnionej postaci diagonalnej. Okazuje się, że w momencie znalezienia tej postaci dla macierzy transformacji „automatycznie” znajdujemy bazę dla $\text{Img } T$ i dla $\ker T$.

Baza $\text{Img } T$

Zgodnie z twierdzeniem poprzedniego podrozdziału, wszystkie kolumny macierzy \mathcal{A} należą do obrazu odwzorowania. Jeżeli kolumny macierzy \mathcal{A} oznaczymy przez \mathbf{k}_i ; $i = 1, \dots, n$, to mamy:

$$\mathcal{A}\mathbf{e}_i = \mathbf{k}_i; \tag{9-11}$$

– kolumny \mathcal{A} są obrazami wektorów bazowych dziedziny odwzorowania.

Wektory-kolumny \mathcal{A} będą tworzyły bazę pod warunkiem, że są liniowo niezależne. Taki wektorami są z pewnością kolumny zawierające samotne jedynki. Można wykazać, że kolumny pozbawione jedynek-dźwigni (jeżeli istnieją) mogą być wyrażone jako liniowe kombinacje tych pierwszych (por. przykład poniżej). Dlatego mamy

Twierdzenie 9.3 *Bazę podprzestrzeni wektorowej, będącej obrazem odwzorowania T , tworzą kolumny macierzy \mathcal{A} zawierające jedynki-dźwignie.*

Rezygnujemy z pełnego dowodu na korzyść przykładu.

⁴Symbol „ Img ” czytamy *obraz* (łacińskie *imago*). W niektórych podręcznikach używa się skrótu „ Im ”, który można jednak pomylić z identycznym symbolem, określającym część urojonej liczby zespolonej. Dlaczego „ \ker ” jest pisane z małej, a „ Img ” z dużej litery pozostaje słodką tajemnicą matematyków.

Przykład 9.3 Podaną poniżej macierz \mathcal{A} (macierz 4×5 – opisuje więc przejście od \mathbb{R}^5 do \mathbb{R}^4)

$$\mathcal{A} = \begin{pmatrix} 1 & 0 & 1 & 3 & 0 \\ 0 & 1 & 1 & 2 & 0 \\ 1 & 1 & 2 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (9-12)$$

można zdiagonalizować

$$\mathcal{A}_D = \begin{pmatrix} \underline{1} & 0 & 1 & 3 & 0 \\ 0 & \underline{1} & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & \underline{1} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (9-13)$$

Jedynki-dźwignie podkreśliliśmy; występują one w kolumnach 1, 2 i 5; tak więc wektory

$$\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad (9-14)$$

tworzą bazę $\text{Im} T$. Każdy z wektorów ma 4 współrzędne (jesteśmy w \mathbb{R}^4), ale wektorów mamy tylko 3 (a nie 4), a więc $\text{Im} A$ jest podprzestrzenią \mathbb{R}^4 (zwróć uwagę na zera w ostatnich, czwartych wierszach wektorów bazy).

Przykład 9.4 Jeżeli rozpatrzeć np. wektor \mathbf{w} , będący liniową kombinacją wektorów-kolumn, \mathbf{k}_i , macierzy \mathcal{A} :

$$\mathbf{w} = 2\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_3 + 2\mathbf{k}_4 - 3\mathbf{k}_5 = \begin{pmatrix} 7 \\ 4 \\ 8 \\ 0 \end{pmatrix}$$

to (sprawdź!)

$$\mathbf{w} = 7 \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + 4 \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} - 3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Baza $\ker T$

I tym razem punktem wyjścia jest zdiagonalizowana macierz; na przykład może to być macierz z równ. 9-13 – ale interesują nas kolumny nie zawierające jedynek-dźwigni. Obecność takich kolumn oznacza, że (por. 6.6) możemy dowolnie ustalać wartości pewnych niewiadomych. Szukamy bazy dla jądra; z jego definicji wynika, że każdy wektor takiej bazy musi spełniać równanie $\mathcal{A}\mathbf{x} = \mathbf{0}$; układ równań będzie więc układem jednorodnym:

$$\sum a_{ik}x_k = 0; \quad i = 1, \dots, n. \quad (9-15)$$

Przykład 9.5 Posłużmy się dyskutowanym przed chwilą przykładem i podstawmy za macierz \mathcal{A} do 9-15 macierz z równ. 9-13:

$$\begin{pmatrix} \underline{1} & 0 & 1 & 3 & 0 \\ 0 & \underline{1} & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & \underline{1} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad (9-16)$$

albo *explicite*:

$$\left. \begin{aligned} x_1 + x_3 + 3x_4 &= 0 \\ x_2 + x_3 + 2x_4 &= 0 \\ x_5 &= 0. \end{aligned} \right\} \quad (9-17)$$

Jako zmienne dowolne bierzemy x_3 i x_4 – bo trzecia i czwarta kolumna \mathcal{A} nie zawierają samotnych jedynek. Przy ich określonym wyborze pozostałe x -y wyliczamy z 9-17:

$$\left. \begin{aligned} x_1 &= -x_3 - 3x_4 \\ x_2 &= -x_3 - 2x_4 \\ x_5 &= 0 \end{aligned} \right\}. \quad (9-18)$$

Przekładając powyższe operacje na język bazy: x_3 i x_4 mogą być dowolne, a to oznacza, że wektorami bazy mogą być w szczególności

$$\mathbf{u}_1 = \begin{pmatrix} ? \\ ? \\ 1 \\ 0 \\ ? \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} ? \\ ? \\ 0 \\ 1 \\ ? \end{pmatrix}. \quad (9-19)$$

Jedynki i zera – podobnie jak w przypadku bazy kanonicznej – gwarantują liniową niezależność wektorów \mathbf{u}_1 i \mathbf{u}_2 i jednocześnie najprostsza formę. Współrzędne oznaczone pytajnikami wyliczamy ze związków 9-18; to z kolei gwarantuje $\mathcal{A}\mathbf{u}_1 = \mathbf{0} = \mathcal{A}\mathbf{u}_2$. Podstawiając do 9-18 $x_3 = 1, x_4 = 0$ (\mathbf{u}_1) i $x_3 = 0, x_4 = 1$ (\mathbf{u}_2) dostajemy jawną postać bazy jądra transformacji A (albo jądra macierzy \mathcal{A}):

$$\mathbf{u}_1 = \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} -3 \\ -2 \\ 0 \\ 1 \\ 0 \end{pmatrix}. \quad (9-20)$$

Nie ulega wątpliwości, że : (a) należą do jądra, (b) są liniowo niezależne. Do pełnego dowodu, że stanowią one bazę jądra, należałoby jeszcze wykazać, że wektory te są rodziną generującą jądro, tzn. każdy wektor spełniający $\mathcal{A}\mathbf{w} = \mathbf{0}$ można zapisać w postaci kombinacji liniowej \mathbf{u}_1 i \mathbf{u}_2 . Przeprowadzenie tej części dowodu pozostawiamy czytelnikowi.

9.5 „Twierdzenie o wymiarach”

Podsumowaniem materiału rozpatrywanego w tym rozdziale jest tzw. twierdzenie o wymiarach, wynikające bezpośrednio z podanych „przepisów konstrukcji” dwóch baz. Mamy

Twierdzenie 9.4 („o wymiarach”) *Jeżeli T jest liniowym odwzorowaniem $\mathbb{R}^n \rightarrow \mathbb{R}^m$ to:*

$$\dim(\ker T) + \dim(\text{Img } T) = n \quad (9-21)$$

– wymiar jądra odwzorowania + wymiar jego obrazu = wymiarowi dziedziny.

Z samej konstrukcji bazy dla $\text{Img } T$ wynika, że wymiar obrazu odwzorowania jest równy rzędowi macierzy; dlatego $\dim(\text{Img } T)$ nazywamy rzędem odwzorowania. Wymiar jądra – $\dim(\ker T)$ nazywamy natomiast zerowością odwzorowania. Twierdzenie o wymiarach można więc sformułować: rząd odwzorowania + jego zerowość = wymiar dziedziny.

Twierdzenie o wymiarach pozwala w sposób bezpośredni określać wymiary jądra i obrazu odwzorowania, a także jego charakter (surjektywny, iniektywny), jeżeli znamy wymiary i rząd macierzy odwzorowania.

Przykład 9.6 Np. odwzorowanie realizowane przez macierz 3×4 , której rząd jest równy 2, to odwzorowanie $\mathbb{R}^4 \rightarrow \mathbb{R}^3$. Rząd macierzy to wymiar obrazu. Ten ostatni, tak jak i wymiar jądra, wynosi więc 2. Ponieważ wymiar (bazy) obrazu odwzorowania wynosi 2, to odwzorowanie na pewno nie jest surjektywne (w przeciwdziedzinie \mathbb{R}^3 istnieją obiekty, które nie powstają w wyniku tego odwzorowania z elementów \mathbb{R}^4). Czytelnik sam powinien odpowiedzieć na pytanie, czy takie odwzorowanie może być iniektywne.

Podkreślmy: pojęcia jądra i obrazu transformacji, to – w gruncie rzeczy – jeszcze jeden sposób mówienia o istnieniu rozwiązań układów równań liniowych. Jeżeli np. rozpatrywać transformację $\mathbb{R}^n \rightarrow \mathbb{R}^n$, to z twierdzenia o wymiarach wynika, że układ równań $\mathcal{T} \mathbf{x} = \mathbf{b}$ (\mathcal{T} – macierz transformacji; \mathbf{b} dowolny wektor z \mathbb{R}^n) będzie miał rozwiązanie wtedy i tylko wtedy, gdy jedynym rozwiązaniem równania $\mathcal{T} \mathbf{x} = \mathbf{0}$ jest wektor zerowy, tzn. kiedy układ równań jednorodnych ma wyłącznie rozwiązanie trywialne: $x_1, \dots, x_n = 0$. Rzeczywiście, istnienie rozwiązania dla \mathbf{b} będącego dowolnym wektorem z \mathbb{R}^n oznacza, że \mathbb{R}^n jest obrazem T , a więc wymiar jądra to $n - n = 0$. W prostym (prostszy ?) schemacie twierdzenia Cramera oznaczało to, że wyznacznik układu równań musi być różny od zera.

9.6 $\mathbb{R}^n \rightarrow \mathbb{R}^n$; zmiana bazy

Przypuśćmy, że w n -wymiarowej przestrzeni wektorowej mamy dwie bazy: „starą” bazę \mathbf{e}

$$\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n \quad (9-22)$$

i „nową” bazę

$$\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n. \quad (9-23)$$

Każdy wektor nowej bazy 9-23 może zostać przedstawiony (jak każdy wektor należący do \mathbb{R}^n) w postaci liniowej kombinacji wektorów bazowych 9-22

$$\mathbf{e}'_i = \sum_{j=1}^n \tau_{ij} \mathbf{e}_j, \quad i = 1, \dots, n, \quad (9-24)$$

gdzie współczynniki τ_{ij} tworzą macierz \mathcal{T} , nazywaną *macierzą przejścia od starej bazy 9-22 do nowej bazy 9-23*:

$$(\mathcal{T}) \equiv (\tau_{ij}) = \begin{pmatrix} \tau_{11} & \dots & \tau_{1n} \\ \vdots & \vdots & \vdots \\ \tau_{n1} & \dots & \tau_{nn} \end{pmatrix}. \quad (9-25)$$

Równanie wektorowe 9-25 zapiszemy w postaci macierzowej

$$\begin{pmatrix} \mathbf{e}'_1 \\ \mathbf{e}'_2 \\ \vdots \\ \mathbf{e}'_n \end{pmatrix} = \begin{pmatrix} \tau_{11} & \tau_{12} & \dots & \tau_{1n} \\ \tau_{21} & \tau_{22} & \dots & \tau_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \tau_{n1} & \tau_{n2} & \dots & \tau_{nn} \end{pmatrix} \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_n \end{pmatrix}, \quad (9-26)$$

albo w skrócie

$$\mathbf{e}' = \mathcal{T} \mathbf{e}. \quad (9-27)$$

Analogicznie, jeżeli przez \mathcal{T}' oznaczyć macierz przejścia *od nowej bazy do starej* to

$$\mathbf{e} = \mathcal{T}' \mathbf{e}', \quad (9-28)$$

czyli

$$\mathbf{e} = (\mathcal{T}' \mathcal{T}) \mathbf{e} \quad (9-29)$$

$$\mathbf{e}' = (\mathcal{T} \mathcal{T}') \mathbf{e}'. \quad (9-30)$$

Z powyższych równań wynika, że macierz \mathcal{T}' to macierz odwrotna do \mathcal{T} – $\mathcal{T}' \equiv \mathcal{T}^{-1}$:

$$\mathcal{T}' \mathcal{T} = \mathcal{T} \mathcal{T}' = \mathcal{I} = \mathcal{T}^{-1} \mathcal{T} = \mathcal{T} \mathcal{T}^{-1}. \quad (9-31)$$

Ale – jeżeli *mówimy* o macierzy odwrotnej to znaczy, że macierz \mathcal{T} musi być macierzą nieosobliwą. Jest to warunek zarówno konieczny jak i wystarczający! **Każda nieosobliwa macierz, o wymiarach $n \times n$ może służyć jako macierz przejścia w \mathbb{R}^n .**

W rozdziale czwartym omawialiśmy wzory według których przekształcają się współrzędne wektora (na płaszczyźnie, w przestrzeni 3-wymiarowej) przy obrocie układu współrzędnych (zmiana bazy!). Aby zapisać teraz

te wzory w postaci konsystentnej z wzorami tego rozdziału musimy zauważyć, że wektorowe równania – zapis dowolnego wektora \mathbf{a} w różnych bazach:

$$\mathbf{a} = \sum_{j=1}^n \alpha_j \mathbf{e}_j \quad (9-32)$$

$$\mathbf{a} = \sum_{k=1}^n \alpha'_k \mathbf{e}'_k \quad (9-33)$$

w przełożeniu na język macierzowy wymagają (reguły mnożenia macierzy!) przyjęcia konwencji, która traktuje zbiór n wektorów bazowych jako jedno-kolumnową macierz, a zbiór współrzędnych wektora – jako macierz jedno-wierszową:

$$\mathbf{a} = (\alpha_1, \alpha_2, \dots, \alpha_n) \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_n \end{pmatrix} = (\alpha'_1, \alpha'_2, \dots, \alpha'_n) \begin{pmatrix} \mathbf{e}'_1 \\ \mathbf{e}'_2 \\ \vdots \\ \mathbf{e}'_n \end{pmatrix}. \quad (9-34)$$

Korzystając z 9-32 i 9-33, a także z 9-22 otrzymujemy

$$\mathbf{a} = \sum_{k=1}^n \alpha'_k \left(\sum_{j=1}^n \tau_{kj} \mathbf{e}_j \right) = \sum_{j=1}^n \left(\sum_{k=1}^n \alpha'_k \tau_{kj} \right) \mathbf{e}_j, \quad (9-35)$$

a więc

$$\alpha_j = \sum_{k=1}^n \alpha'_k \tau_{kj}; \quad j = 1, \dots, n. \quad (9-36)$$

W zapisie macierzowym:

$$(\alpha_1, \alpha_2, \dots, \alpha_n) = (\alpha'_1, \alpha'_2, \dots, \alpha'_n) \begin{pmatrix} \mathcal{T} \end{pmatrix}. \quad (9-37)$$

Mnożąc powyższe równanie, wyrażające współrzędne wektora \mathbf{a} w starej bazie poprzez współrzędne w bazie nowej, od prawej przez \mathcal{T}^{-1} dostaniemy

$$(\alpha'_1, \alpha'_2, \dots, \alpha'_n) = (\alpha_1, \alpha_2, \dots, \alpha_n) \begin{pmatrix} \mathcal{T}^{-1} \end{pmatrix}, \quad (9-38)$$

albo w skrócie

$$\alpha'_k = \sum_{j=1}^n \alpha_j \tau_{jk}^{-1}; \quad k = 1, \dots, n. \quad (9-39)$$

Wzory 9-38 i 9-39 mają postać, która różni się od „standardowej”. W tej ostatniej wektory przedstawiamy jako *jednokolumnowe* macierze. Przejście do postaci standardowej wymaga poddaniu operacji transpozycji obu stron 9-38 i 9-39. Otrzymamy

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \mathcal{T}^{-1} \end{pmatrix}^T \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}. \quad (9-40)$$

Pełne rozwinięcie powyższego zapisu to

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \tau_{11}^{-1} & \tau_{21}^{-1} & \dots & \tau_{n1}^{-1} \\ \tau_{12}^{-1} & \tau_{22}^{-1} & \dots & \tau_{n2}^{-1} \\ \vdots & \vdots & \vdots & \vdots \\ \tau_{1n}^{-1} & \tau_{2n}^{-1} & \dots & \tau_{nn}^{-1} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}. \quad (9-41)$$

(Przypominam: transpozycja iloczynu macierzy to iloczyn macierzy transponowanych, ale w odwrotnej kolejności; zwróć uwagę na przestawione (transpozycja!) wskaźniki elementów τ^{-1}_{ik} .) Zapis poddanego transpozycji równ. 9-39 to

$$\alpha'_k = \sum_{j=1}^n \alpha_j \tau_{kj}^{-1}; \quad k = 1, \dots, n. \quad (9-42)$$

Macierz transformacji współrzędnych wektora to transponowana macierz odwrotna do macierzy przejścia (macierzy transformacji wersorów bazy). Ale w rozdziale 4 widzieliśmy, że macierz odwrotna *może być* macierzą transponowaną. Transponowana macierz transponowana to oczywiście wyjściowa macierz i okazuje się, że w pewnych sytuacjach macierz przejścia określa zarówno transformację wektorów bazy jak i współrzędnych wektorów (względem tej właśnie bazy). Zagadnienia te omówimy jeszcze dokładniej w rozdziale 10 (macierz ortogonalna – podrozdział 10.4) i rozdziale 10.

Przykład 9.7 Niech w \mathbb{R}^3 istnieje pewna baza $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ i nowa baza

$$\begin{aligned} \mathbf{e}'_1 &= 5\mathbf{e}_1 - \mathbf{e}_2 - 2\mathbf{e}_3 \\ \mathbf{e}'_2 &= 2\mathbf{e}_1 + 3\mathbf{e}_2 \\ \mathbf{e}'_3 &= -2\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3. \end{aligned} \quad (9-43)$$

Macierz przejścia to

$$\mathcal{T} = \begin{pmatrix} 5 & -1 & -2 \\ 2 & 3 & 0 \\ -2 & 1 & 1 \end{pmatrix}$$

macierz odwrotna

$$\mathcal{T}^{-1} = \begin{pmatrix} 3 & -1 & 6 \\ -2 & 1 & -4 \\ 8 & -3 & 17 \end{pmatrix}.$$

Wektor \mathbf{a}

$$\mathbf{a} = \mathbf{e}_1 + 4\mathbf{e}_2 - \mathbf{e}_3$$

będzie miał w „nowej” bazie (9-43) współrzędne

$$(\alpha'_1, \alpha'_2, \alpha'_3) = (1, 4, -1) \begin{pmatrix} 3 & -1 & 6 \\ -2 & 1 & -4 \\ 8 & -3 & 17 \end{pmatrix} = (-13, 6, -27),$$

czyli

$$\mathbf{a} = -13\mathbf{e}'_1 + 6\mathbf{e}'_2 - 27\mathbf{e}'_3.$$

9.7 Transformacja liniowa i jej macierz

Rozważmy przestrzeń wektorową \mathbb{R}^n i odwzorowanie A , w wyniku którego *każdemu* wektorowi $\mathbf{a} \in \mathbb{R}^n$ zostaje przyporządkowany pewien wektor \mathbf{a}' , obraz wektora \mathbf{a} . *Jeżeli odwzorowanie spełnia zwykłe warunki liniowości:*

$$T(\mathbf{a} + \mathbf{b}) = T(\mathbf{a}) + T(\mathbf{b}); \quad T(\alpha\mathbf{a}) = \alpha T(\mathbf{a}), \quad (9-44)$$

to dowolna kombinacja liniowa wektorów \mathbb{R}^n pod wpływem odwzorowania przekształca się w taką samą (z takimi samymi współczynnikami) kombinację liniową wektorów-obrazów:

$$\begin{aligned} T(\alpha_1\mathbf{a}_1 + \alpha_2\mathbf{a}_2 + \dots + \alpha_k\mathbf{a}_k) &= \\ &= T(\alpha_1\mathbf{a}_1) + T(\alpha_2\mathbf{a}_2) + \dots + T(\alpha_k\mathbf{a}_k). \end{aligned} \quad (9-45)$$

W szczególności, ponieważ każdy wektor \mathbf{a} wyraża się, w jednoznaczny sposób, jako pewna kombinacja liniowa wektorów bazowych \mathbf{e}_i ; $i = 1, \dots, n$, to jego obraz pozostaje kombinacją liniową (z tymi samymi współczynnikami) obrazów wektorów bazowych $\mathbf{c}_i \equiv T(\mathbf{e}_i)$; $i = 1, \dots, n$. Podkreślmy w tym miejscu: wektory \mathbf{c}_i już wcale nie muszą być bazą w \mathbb{R}^n , chociaż w przypadku gdy mamy do czynienia z bazą kanoniczną (ortonormalnych wersorów) to nieosobliwa transformacja przeprowadza taką bazę w inną bazę ortonormalną – por. następny rozdział. Z powyższych rozważań wynika:

Twierdzenie 9.5 Każda liniowa transformacja przestrzeni wektorowej \mathbb{R}^n jest jednoznacznie określona przez określenie obrazów wektorów bazowych. (nie zapominajmy – one stanowią kolumny macierzy transformacji!)

Wektory \mathbf{c}_i w dalszym ciągu mogą być wyrażane jako liniowa kombinacja wektorów bazowych \mathbf{e}_i . Zapiszmy

$$\mathbf{c}_i = \sum_{j=1}^n \gamma_{ij} \mathbf{e}_j, \quad (9-46)$$

gdzie współczynniki γ_{ij} tworzą macierz transformacji T – macierz \mathcal{G} ; macierz taka określa, albo specyfikuje, transformację w bazie \mathbf{e}_i .

Znając macierz transformacji bez trudu wyrazimy współrzędne obrazu wektora \mathbf{a} , wektora $T(\mathbf{a})$ w (ciągle tej samej!) bazie \mathbf{e}_i . Dla

$$\mathbf{a} = \sum_{k=1}^n \alpha_k \mathbf{e}_k \quad (9-47)$$

mamy

$$T(\mathbf{a}) = T\left(\sum_{k=1}^n \alpha_k \mathbf{e}_k\right) = \sum_{k=1}^n \alpha_k T(\mathbf{e}_k). \quad (9-48)$$

Macierzowo (por. 9-34, 9-37):

$$T(\mathbf{a}) = \begin{pmatrix} \alpha_1, \alpha_2, \dots, \alpha_n \end{pmatrix} \begin{pmatrix} \mathcal{G} \end{pmatrix} \begin{pmatrix} \mathbf{e}_1 \\ \vdots \\ \mathbf{e}_n \end{pmatrix} \quad (9-49)$$

— macierz-wiersz współrzędnych obrazu wektora \mathbf{a} otrzymujemy mnożąc od prawej współrzędne \mathbf{a} przez macierz transformacji \mathcal{G} . *Wszystko rozpatrujemy w bazie \mathbf{e}_i .*

Przykład 9.8 Rozpatrujemy transformację liniową w \mathbb{R}^3 , której macierz \mathcal{G} ma, w pewnej bazie $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, postać

$$\mathcal{G} = \begin{pmatrix} -2 & 1 & 0 \\ 1 & 3 & 2 \\ 0 & -4 & 1 \end{pmatrix}.$$

W tej samej bazie, wektor \mathbf{a} to:

$$\mathbf{a} = 5\mathbf{e}_1 + \mathbf{e}_2 - 2\mathbf{e}_3.$$

Współrzędne wektora-obrazu \mathbf{a} , to (por. 9-49)

$$(5, 1, -2) \begin{pmatrix} -2 & 1 & 0 \\ 1 & 3 & 2 \\ 0 & -4 & 1 \end{pmatrix} = (-9, 16, 0),$$

czyli

$$T(\mathbf{a}) = -9\mathbf{e}_1 + 16\mathbf{e}_2.$$

Rozdział 10

Przestrzenie Euklidesowe; iloczyn skalarny; ortogonalność

10.1 Najważniejsza forma kwadratowa przestrzeni wektorowej

Określenie n -wymiarowej przestrzeni wektorowej \mathbb{R}^n **nie** stanowi „pełnego” uogólnienia „zwykłej” płaszczyzny, lub przestrzeni 3-wymiarowej. W tych ostatnich, rozważając wektory, posługujemy się z pełnym zrozumieniem takimi pojęciami jak długość wektora i kąt między dwoma wektorami. Oba te pojęcia wyrażamy przy pomocy iloczynu skalarnego dwóch wektorów: długość wektora to nic innego jak pierwiastek (ze znakiem $+$) z iloczynu skalarnego wektora przez samego siebie, a kosinus kąta między dwoma wektorami jest równy ich iloczynowi skalarnemu, podzielonemu przez iloczyn ich długości.

Posługując się przykładem (rozdział 5.) 4-wymiarowej przestrzeni sfer, łatwo zrozumieć, że mówienie o długości takiego wektora nie bardzo ma sens. Dlatego dla dowolnej przestrzeni wektorowej iloczyn skalarny nie zawsze jest (nie zawsze może być) definiowany. Jeżeli jednak definicja jest możliwa (i podana) to nasza przestrzeń wektorowa nazywa się przestrzenią euklidesową. Ściśle mówiąc – przestrzeń wektorowa \mathbb{R}^n , przy podaniu definicji iloczynu skalarnego *zamienia się* w przestrzeń euklidesową, którą oznaczymy przez \mathbb{E}^n . Można wykazać, że takiej zamiany można dokonać jednoznacznie. W zastosowaniach fizycznych (i technicznych) spotykamy się właśnie z takimi przestrzeniami.

Określenie iloczynu skalarnego dokonuje się poprzez podanie jego własności. Dla wektorów w przestrzeni \mathbb{E}^n : \mathbf{U} , \mathbf{V} i \mathbf{W} iloczyn skalarny dwóch z nich, na przykład \mathbf{V} i \mathbf{W} , to pewna wartość liczbowa, którą oznaczamy (\mathbf{V}, \mathbf{W}) , jednoznacznie przyporządkowana tym wektorom. Iloczyn skalarny ma być

$$\begin{array}{ll} \text{I.} & \text{przemienny:} & (\mathbf{U}, \mathbf{W}) = (\mathbf{W}, \mathbf{U}) \\ \text{II.} & \text{rozłączny:} & (\mathbf{U} + \mathbf{V}, \mathbf{W}) = (\mathbf{U}, \mathbf{W}) + (\mathbf{V}, \mathbf{W}) \\ \text{III.} & \text{liniowy:} & (a\mathbf{V}, \mathbf{W}) = a(\mathbf{V}, \mathbf{W}) \quad a - \text{liczba (skalar)} \\ \text{IV.} & \text{określony nieujemnie:} & (\mathbf{V}, \mathbf{V}) \geq 0 \end{array} \quad (10-1)$$

(własność IV: iloczyn skalarny wektora przez samego siebie – kwadrat długości, albo kwadrat *normy* wektora – będzie równy zeru tylko w przypadku wektora zerowego).

Tak jak była o tym mowa w rozdziale czwartym, iloczyn skalarny może służyć do określenia *ortogonalności* dwóch wektorów. Dwa wektory \mathbf{a} i \mathbf{b} są ortogonalne jeżeli $(\mathbf{a}, \mathbf{b}) = 0$. Układ wektorów, z których każde dwa są ortogonalne nazywamy układem ortogonalnym. Takim układem może być baza przestrzeni wektorowej: układ wektorów ξ_1, \dots, ξ_n . Jeżeli dodatkowo długości wektorów bazy są równe 1

$$(\xi_i, \xi_j) = \delta_{ij}; \quad i, j = 1, \dots, n, \quad (10-2)$$

to bazę nazwalibyśmy *ortonormalną*. Taka bazą jest np. baza kanoniczna \mathbb{R}^n .

Jeżeli dwa wektory wyrazić w *dowolnej bazie ortonormalnej* $\xi_1, \xi_2, \dots, \xi_n$ przestrzeni:

$$\mathbf{V} = \sum_{i=1}^n V_i \xi_i, \quad \mathbf{W} = \sum_{i=1}^n W_i \xi_i, \quad (10-3)$$

to definicja, spójna z naszymi określeniami iloczynu skalarnego na płaszczyźnie i w trójwymiarowej przestrzeni i spełniająca powyższe dezyderaty ma postać

$$(\mathbf{V}, \mathbf{W}) = \sum_{i=1}^n V_i W_i. \quad (10-4)$$

Określenie iloczynu skalarnego znakomicie upraszcza mówienie o ortogonalności wektorów, także ortogonalności wektorów bazy przestrzeni wektorowej. W aplikacjach technicznych i fizycznych praktycznie zawsze mamy do czynienia z bazami ortogonalnymi (wyjątek zasługujący na wzmiankę: pewne problemy pojawiające się przy opisie struktur krystalicznych; istniejące struktury fizyczne narzucają czasem wybór układów „ukośnokątnych”).

Jeżeli używamy do opisu wektorów notacji macierzowej, to określenie iloczynu skalarnego wymaga przyjęcia pewnej dodatkowej umowy (konwencji). Nasze wektory w \mathbb{R}^n to n -wierszowe kolumny współrzędnych:

$$\mathbf{V} \equiv \begin{pmatrix} V_1 \\ V_2 \\ \vdots \\ V_n \end{pmatrix}, \quad \mathbf{W} \equiv \begin{pmatrix} W_1 \\ W_2 \\ \vdots \\ W_n \end{pmatrix}. \quad (10-5)$$

Jeżeli mamy jednak *mnożyć* dwie macierze, to liczba kolumn jednej z nich ma być równa liczbie wierszy drugiej. Dlatego w zapisie macierzowym mamy

$$(\mathbf{V}, \mathbf{W}) = (V_1, V_2, \dots, V_n) \begin{pmatrix} W_1 \\ W_2 \\ \vdots \\ W_n \end{pmatrix} \equiv \mathcal{V}^T \mathcal{W}. \quad (10-6)$$

– lewy czynnik iloczynu to macierz jednowierszowa o n kolumnach, a więc macierzy *transponowana* w stosunku do macierzy w 10-5.

10.2 Ortogonalizacja układu wektorów liniowo niezależnych

Obowiązuje:

Twierdzenie 10.1 *Każdy ortogonalny układ k nie-zerowych wektorów – $(\mathbf{a}_i, \mathbf{a}_i) > 0$, dla $i = 1, \dots, k$ – jest układem liniowo niezależnym.*

Dowód: Załóżmy, że mamy układ k niezerowych wektorów ortogonalnych w \mathbb{E}^n :

$$\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k; \quad \mathbf{a}_i \neq 0, \quad i = 1, \dots, k, \quad (10-7)$$

dla których

$$(\mathbf{a}_i, \mathbf{a}_j) = 0; \quad i \neq j. \quad (10-8)$$

Utwórzmy kombinacje liniową tych wektorów i przyrównajmy ją do zera:

$$\alpha_1 \mathbf{a}_1 + \alpha_2 \mathbf{a}_2 + \dots + \alpha_k \mathbf{a}_k = 0. \quad (10-9)$$

Mnożąc obie strony 10-9 przez \mathbf{a}_i , $1 \leq i \leq k$ i wykorzystując odpowiednie własności iloczynu skalarnego, mamy

$$\begin{aligned} 0 = (\mathbf{0}, \mathbf{a}_i) &= (\alpha_1 \mathbf{a}_1 + \alpha_2 \mathbf{a}_2 + \dots + \alpha_k \mathbf{a}_k, \mathbf{a}_i) \\ &= \alpha_1 (\mathbf{a}_1, \mathbf{a}_i) + \alpha_2 (\mathbf{a}_2, \mathbf{a}_i) + \dots + \alpha_k (\mathbf{a}_k, \mathbf{a}_i) \\ &= \alpha_i (\mathbf{a}_i, \mathbf{a}_i). \end{aligned}$$

Ponieważ (z założenia) $(\mathbf{a}_i, \mathbf{a}_i) > 0$, z powyższego wynika, że $\alpha_i = 0$ dla wszystkich i , a więc układ 10-7 jest rzeczywiście układem liniowo niezależnych.

Ortogonalizacja układu wektorów liniowo niezależnych polega na odrzuceniu z danego wektora jego składowych odpowiadających kierunkom bazowym pozostałych $n - 1$ wektorów. Na płaszczyźnie, na przykład, dwa liniowo niezależne wektory to dwa nierównoległe wektory \mathbf{A} i \mathbf{B} . Jeżeli jeden z nich przyjmiemy za wektor

bazowy, dokonując ewentualnie *normalizacji*, tzn. sprowadzającej jego długość do jednostkowej: $\mathbf{A} \rightarrow \mathbf{A}/A$, to drugim wektorem bazowym będzie wektor \mathbf{B} , pozbawiony swojej składowej w kierunku \mathbf{A} czyli

$$\mathbf{B}' = \mathbf{B} + a\mathbf{A}$$

gdzie a wyliczamy z warunku $\mathbf{A} \cdot \mathbf{B}' = 0$:

$$a = -\frac{\mathbf{A} \cdot \mathbf{B}}{A^2}.$$

Wektor \mathbf{B}' też można poddać normalizacji, dzieląc go przez jego długość B' .
W przypadku ogólnym, dysponując układem k wektorów liniowo niezależnych

$$\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k \quad (10-10)$$

przejdzie do układu k wektorów ortogonalnych

$$\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_k \quad (10-11)$$

dokonywane w wyniku procedury sekwencyjnej – aby wyznaczyć kolejny \mathbf{b}_n musimy dysponować wszystkimi poprzednimi: $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{n-1}$. Kolejne \mathbf{b}_n zapisujemy jako sumę wektora \mathbf{a}_n , oraz kombinacji liniowej wektorów $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{n-1}$:

$$\mathbf{b}_n = \mathbf{a}_n + \alpha_{n1}\mathbf{b}_1 + \dots + \alpha_{ni}\mathbf{b}_i + \dots + \alpha_{n,n-1}\mathbf{b}_{n-1}. \quad (10-12)$$

Obecność \mathbf{a}_n gwarantuje niezależność liniową kolejnych \mathbf{b}_n , natomiast ortogonalność wektorów \mathbf{b} osiągniemy poprzez odpowiedni dobór współczynników α_{ni} . Mnożąc obie strony (10-12) przez \mathbf{b}_i ($i = 0, 1, \dots, n-1$) mamy

$$(\mathbf{b}_n, \mathbf{b}_i) = (\mathbf{a}_n, \mathbf{b}_i) \quad (10-13)$$

$$+ \sum_{j=1}^{n-1} \alpha_{nj}(\mathbf{b}_j, \mathbf{b}_i) \quad (10-14)$$

Ze względu na postulowaną ortogonalność \mathbf{b}_i lewa strona (10-13) musi być równa zero, a z sumy po prawej stronie zostaje tylko składnik α_{ni} . Stąd wzory określające współczynniki α_{ni} to

$$\alpha_{ni} = -\frac{(\mathbf{a}_n, \mathbf{b}_i)}{(\mathbf{b}_i, \mathbf{b}_i)}; \quad i = 1, \dots, k. \quad (10-15)$$

Podany powyżej dość banalny przykład z algebry wektorów na płaszczyźnie, powinien uzmysłwić Czytelnikowi to, że procedura ortogonalizacji jest nieskończenie wieloznaczna. Tak jak istnieje nieskończenie wiele par nierównoległych wektorów \mathbf{A} , i \mathbf{B} na płaszczyźnie, z których można konstruować ortogonalną (lub ortonormalną) bazę, tak i wektory w $\mathbb{E}^n - \mathbf{a}_i; i = 1, \dots, n$, służące jako materiał do konstrukcji układu ortogonalnego, mogą być wybrane na nieskończenie wiele sposobów. Jedynym warunkiem jest ich liniowa niezależność. A jeżeli tak, to w wyborze kierujemy się zwykle jego prostotą i adekwatnością do opisywanej sytuacji. Co więcej, względy praktyczne skłaniają nas do korzystania z układów ortonormalnych, które powstają z ortogonalnych w wyniku podzielenia wektorów przez ich długość, lub *normę*¹, czyli pierwiastek z iloczynu skalarnego wektora przez samego siebie:

$$\mathbf{b}_i \xrightarrow{\text{normalizacja}} \frac{\mathbf{b}_i}{\sqrt{(\mathbf{b}_i, \mathbf{b}_i)}}. \quad (10-16)$$

Zauważmy, że proces ortonormalizacji, prowadzony sukcesywnie i konsekwentnie w trakcie procesu ortogonalizacji ułatwia ten ostatni: mianowniki ułamków w 10-15 są wszystkie równe jedność.

Określenie iloczynu skalarnego i wprowadzenie ortonormalnej bazy czyni bardzo przejrzystym język operacji dokonywanych na elementach przestrzeni euklidesowych (wektorach z określoną normą) i operatorach

¹czasem używa się określenia: norma Euklidesowa.

odwzorowań. Ortonormalne wektory kanonicznej bazy to jednokolumnowe macierze

$$\boldsymbol{\xi}_i = \begin{pmatrix} 0_{[1]} \\ \vdots \\ 0_{[i-1]} \\ 1_{[i]} \\ 0_{[i+1]} \\ \vdots \\ 0_{[n]} \end{pmatrix}, \quad (10-17)$$

a transponowane do nich macierze jednowierszowe mają postać:

$$\boldsymbol{\xi}_k^T = (0_{[1]}, \dots, 0_{[k-1]}, 1_{[k]}, 0_{[k+1]}, \dots, 0_{[n]}). \quad (10-18)$$

(W powyższych wzorach dolne wskaźniki $[i], [k]$ określają numer wiersza lub kolumny.)

Jeżeli przedstawienie wektora \boldsymbol{x} w bazie $\boldsymbol{\xi}_i$ ma postać

$$\boldsymbol{x} = \sum_{i=1}^n x_i \boldsymbol{\xi}_i \quad (10-19)$$

to mnożąc² wektor \boldsymbol{x} przez $\boldsymbol{\xi}_k$ (przy ustalonym $k = 1, \dots, n$) mamy

$$(\boldsymbol{\xi}_k, \boldsymbol{x}) = (\boldsymbol{\xi}_k, \sum_{i=1}^n x_i \boldsymbol{\xi}_i) = \sum_{i=1}^n x_i (\boldsymbol{\xi}_k, \boldsymbol{\xi}_i) = \sum_{i=1}^n x_i \delta_{ik} = x_k \quad (10-20)$$

i wektor \boldsymbol{x} możemy zapisać w postaci

$$\boldsymbol{x} = \sum_{i=1}^n x_i \boldsymbol{\xi}_i = \sum_{i=1}^n (\boldsymbol{\xi}_i, \boldsymbol{x}) \boldsymbol{\xi}_i. \quad (10-21)$$

Wielkości $(\boldsymbol{\xi}_i, \boldsymbol{x})$, $i = 1, \dots, n$ to rzuty wektora \boldsymbol{x} na wektory bazy. Iloczyn skalarny dostarcza spójnej i zgrabnej formy zapisu *reprezentacji wektora* \boldsymbol{x} w bazie $\boldsymbol{\xi}_i$

10.2.1 Reprezentacja macierzy w określonej bazie

Podobnie jak mieliśmy reprezentacje wektora w danej bazie, tak możemy mówić i o *reprezentacji macierzy* \mathcal{A} odwzorowania liniowego w określonej bazie, a konkretnie – tej samej bazie kanonicznej $\boldsymbol{\xi}_i$. Wynik działania \mathcal{A} na j -ty wektor bazowy to wektor-kolumna (j -ta) macierzy \mathcal{A} :

$$\mathcal{A} \boldsymbol{\xi}_j = \sum_{i=1}^n a_{ij} \boldsymbol{\xi}_i; \quad j = 1, 2, \dots, n. \quad (10-22)$$

Iloczyn skalarny, służący do wyrażenia elementu macierzy \mathcal{A} w bazie $\boldsymbol{\xi}_i$, zapiszemy *explicitie* w postaci macierzowej:

$$\begin{aligned} & (\boldsymbol{\xi}_k, \mathcal{A} \boldsymbol{\xi}_j) \\ &= \begin{pmatrix} 0, \dots, 1_{[k]}, \dots, 0 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{k1} & a_{k2} & \dots & a_{kj} & \dots & a_{kn} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1_{[j]} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \end{aligned}$$

²Mnożymy „od lewej”. Będzie to miało znaczenie w przypadku, gdy nasza przestrzeń wektorowa jest rozpięta nad ciałem liczb zespolonych; por. następne podrozdziały.

$$= \begin{pmatrix} 0, \dots, 1_{[k]}, \dots, 0 \end{pmatrix} \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{kj} \\ \vdots \\ a_{nj} \end{pmatrix} = a_{kj},$$

albo w wersji „stenograficznej”:

$$(\boldsymbol{\xi}_k, \mathcal{A} \boldsymbol{\xi}_j) = (\boldsymbol{\xi}_k, \sum_{i=1}^n a_{ij} \boldsymbol{\xi}_i) = \sum_{i=1}^n a_{ij} (\boldsymbol{\xi}_k, \boldsymbol{\xi}_i) = \sum_{i=1}^n a_{ij} \delta_{ki} = a_{kj}. \quad (10-23)$$

Gdybyśmy chcieli skorzystać z symetrii iloczynu skalarnego (w przypadku rzeczywistych \mathbb{E}^n) to powyższy wzór przekształcamy pamiętając, że jeżeli wyrażenie 10-22 ma wystąpić jako lewy czynnik iloczynu skalarnego, to musimy utworzyć jego transpozycję; w dodatku transpozycja iloczynu to iloczyn transpozycji czynników w porządku odwrotnym:

$$\begin{aligned} (\mathcal{A} \boldsymbol{\xi}_j, \boldsymbol{\xi}_k) &= (\mathcal{A} \boldsymbol{\xi}_j)^T (\boldsymbol{\xi}_k) = (\boldsymbol{\xi}_j)^T (\mathcal{A})^T (\boldsymbol{\xi}_k) \\ &= \begin{pmatrix} 0, \dots, 1_{[j]}, \dots, 0 \end{pmatrix} \begin{pmatrix} a_{11} & a_{21} & \dots & a_{k1} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{k2} & \dots & a_{n2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{1j} & a_{2j} & \dots & a_{kj} & \dots & a_{nj} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{kn} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{kj} \\ \vdots \\ a_{nj} \end{pmatrix} \\ &= a_{kj}. \end{aligned}$$

Wersja „stenograficzna” niewiele się różni od 10-23:

$$(\mathcal{A} \boldsymbol{\xi}_j, \boldsymbol{\xi}_k) = \left(\sum_{i=1}^n a_{ij} \boldsymbol{\xi}_i, \boldsymbol{\xi}_k \right) = \sum_{i=1}^n a_{ij} (\boldsymbol{\xi}_i, \boldsymbol{\xi}_k) = \sum_{i=1}^n a_{ij} \delta_{ik} = a_{kj}. \quad (10-24)$$

Warto tylko dodać, że w większości „fizycznych” zastosowań mamy do czynienia z macierzami symetrycznymi (albo hermitowskimi (por. punkt 10.5 – w mechanice kwantowej)); wzór 10-24 określa więc a_{jk} jak i a_{kj} (przy ewentualnym zastosowaniu operacji sprzężenia zespolonego).

Zaznaczmy jeszcze, że w przypadku gdy macierz \mathcal{A} jest *macierzą ortogonalną*, działającą na ortonormalne wersory bazy, $\boldsymbol{\xi}_i$, to jest ona macierzą przejścia (do innej bazy ortonormalnej) – por. punkt 10.4.

10.3 Nierówność Schwarz; nierówność trójkąta

Z określenia iloczynu skalarnego wynikają dwie bardzo istotne (i nietrudne do przyjęcia) nierówności. Tzw. nierówność Schwarz to:

$$|\mathbf{u} \cdot \mathbf{v}| \leq |\mathbf{u}| |\mathbf{v}| \quad (10-25)$$

– wartość bezwzględna iloczynu skalarnego dwóch wektorów nie może być większa od iloczynu ich długości; znak równości mamy tylko w przypadku gdy jeden z wektorów jest iloczynem drugiego przez pewną liczbę (oba wektory są liniowo zależne). Nierówność ta obowiązuje w przestrzeniach euklidesowych \mathbb{E}^n (n – dowolne), a także w przestrzeniach wektorowych o nieskończonej liczbie wymiarów (przestrzeniach funkcyjnych z określonym iloczynem skalarnym).³

Konsekwencją tej nierówności jest – jeszcze bardziej oczywista nierówność trójkąta:

$$|\mathbf{x} + \mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}| \quad (10-26)$$

– czyli, że z Krakowa do Warszawy bliżej jest przez Radom niż przez Poznań.

³Dowód nierówności Schwarz pomijamy. W 2- i 3-wymiarowej przestrzeni nierówność ta wynika bezpośrednio z definicji iloczynu skalarnego – por. punkt 4.2.

$$\begin{aligned}
 |\mathbf{x} + \mathbf{y}|^2 &= (\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} + \mathbf{y}) = |\mathbf{x}|^2 + 2\mathbf{x} \cdot \mathbf{y} + |\mathbf{y}|^2 \\
 &\leq |\mathbf{x}|^2 + 2|\mathbf{x} \cdot \mathbf{y}| + |\mathbf{y}|^2 \stackrel{\text{Schwarz}}{\leq} |\mathbf{x}|^2 + 2|\mathbf{x}||\mathbf{y}| + |\mathbf{y}|^2 \\
 &= (|\mathbf{x}| + |\mathbf{y}|)^2.
 \end{aligned}
 \tag{10-27}$$

10.4 Ortogonalność macierzy

Dwie macierze nazywamy ortogonalnymi, jeżeli spełniona jest równość: macierz odwrotna (\mathcal{A}^{-1}) do danej macierz \mathcal{A} jest równa macierzy transponowanej \mathcal{A}^T , czyli

$$a_{ik}^{-1} = a_{ki}. \tag{10-28}$$

Z takimi macierzami spotkaliśmy się już w przypadku obrotów, zarówno w przestrzeni dwuwymiarowej (punkt 4.5), jak i n - wymiarowej. Własność ortogonalności wynikała bezpośrednio z definicji elementów macierzy \mathcal{A} i \mathcal{A}^{-1} .

W dużej grupie problemów – gdy mamy do czynienia np. z transformacjami obrotów – rachunek macierzowy to rachunek macierzy ortogonalnych. Aby łatwiej zrozumieć sens tego określenia rozważmy przypadek obrotu w przestrzeni dwuwymiarowej (ograniczenie się do dwóch wymiarów nie umniejsza uniwersalności przykładu). Niech macierz \mathcal{M} :

$$\mathcal{M} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{10-29}$$

reprezentuje macierz ortogonalną. Wówczas iloczyn $\mathcal{M}^{-1} \cdot \mathcal{M} = \mathcal{M}^T \cdot \mathcal{M} = \mathcal{I}$, czyli

$$\begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} a_{11}^2 + a_{21}^2 & a_{11}a_{12} + a_{21}a_{22} \\ a_{11}a_{12} + a_{21}a_{22} & a_{12}^2 + a_{22}^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{10-30}$$

Powyzsza równość jest równoważna trzem równaniom:

$$\left. \begin{aligned} a_{11}^2 + a_{21}^2 &= 1 \\ a_{12}^2 + a_{22}^2 &= 1 \\ a_{11}a_{12} + a_{21}a_{22} &= 0. \end{aligned} \right\} \tag{10-31}$$

Te trzy równości można zinterpretować geometrycznie: wektory

$$\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} \quad \text{i} \quad \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} \tag{10-32}$$

są jednostkowymi, wzajemnie prostopadłymi wektorami. Współrzędne tych wektorów to kolumny macierzy \mathcal{M} . Z drugiej strony wiemy, że wektory te powstają w wyniku działania macierzy \mathcal{M} na ortonormalne wektory bazy w \mathbb{E}^2 $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ i $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} \tag{10-33}$$

oraz

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}. \tag{10-34}$$

Macierz ortogonalna pozostawia więc bez zmiany długości wersorów bazy i kąt między nimi. **Ortonormalna baza** $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ i $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ pod działaniem macierzy 10-29 przechodzi w inną bazę ortonormalną. Bez zmiany pozostaje też długość dowolnego wektora, który w starej bazie ma przedstawienie

$$\alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{10-35}$$

i długość $\sqrt{\alpha^2 + \beta^2}$, a w nowej (por. 10-33 i 10-34)

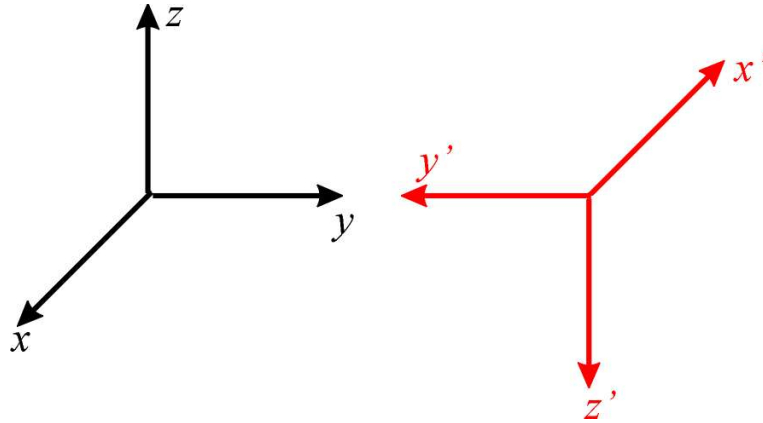
$$\alpha \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} + \beta \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}, \quad (10-36)$$

i kwadrat długości

$$\alpha^2(a_{11}^2 + a_{21}^2) + \beta^2(a_{12}^2 + a_{22}^2) = \alpha^2 + \beta^2.$$

Pozostaje to w zgodzie z naszym wyobrażeniem transformacji obrotu: obrót układu współrzędnych o kąt θ (rozpatrujemy sytuację na płaszczyźnie) to nic innego jak obrót wektora o kąt θ – a to nie zmienia jego długości.

Zauważmy, że macierz ortogonalna nie musi reprezentować tylko transformacji obrotu; transformacją, która pozostawia długości wektorów bazy i kąt zawarty między nimi będzie również transformacja odbicia, nazywana *inwersją* i polegająca na zmianie orientacji jednej lub więcej osi układu współrzędnych. Mówiąc o transformacji odbicia, mamy zwykle na myśli zmianę orientacji wszystkich osi układu – na rys. 10.1 przedstawiona jest „kompletna” transformacja odbicia w przypadku 3-wymiarowym. W wyniku tej transformacji współrzędne



Rysunek 10.1: Inwersja układu współrzędnych w przestrzeni 3-wymiarowej.

wektora zmieniają znak na przeciwny; macierz transformacji odpowiadająca transformacji z rysunku 10.1 ma postać

$$\mathcal{A}_{\text{odbicia}} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad (10-37)$$

choć możemy mieć do czynienia z przypadkami, w których odbiciu podlega tylko dwie osie lub jedna oś. Macierz takiej „ograniczonej” transformacji odbicia różni się od macierzy 10-37 tym, że dla osi pozostawionych bez zmian wartości -1 zostają zastąpione przez $+1$.

W rozdziale trzecim, omawiając własności wyznaczników stwierdziliśmy, że wyznacznik macierzy będącej iloczynem dwóch macierzy jest równy iloczynowi wyznaczników macierzy-czynników. Ponieważ wyznacznik macierzy transponowanej jest równy wyznacznikowi macierzy wyjściowej to z równania 10-30 wynika, że

$$\det \mathcal{A} \cdot \det \mathcal{A}^{-1} = (\det \mathcal{A})^2 = 1, \quad (10-38)$$

a więc wyznacznik macierzy ortogonalnej może być równy $+1$ lub -1 . To, że wyznacznik jest równy jedności nie jest dla nas niespodzianką. W rozdziale siódmym podaliśmy „praktyczną” interpretację wyznacznika – jako „czynnika skalującego element powierzchni (lub – ogólnie – n -wymiarowej „objętości” w \mathbb{R}^n). Transformacja ortogonalna nie zmienia długości wektorów, a jeżeli tak to i budowane z nich objętości muszą pozostać takie same. Z kolei, możliwość pojawienia się wartości wyznacznika -1 jest widoczna już z równania 10-37. Aby zrozumieć lepiej co oznacza ujemny czynnik skalujący zauważmy, że przedstawiona na rys. 10.1 transformacja inwersji trzech osi powoduje zmianę *skrętności* układu współrzędnych. Trójka (x, y, z) jest trójką prawoskrętną, a trójka (x', y', z') – lewoskrętną.⁴

⁴Przypomnijmy: skrętność trójki wektorów (x, y, z) to skrętność śruby, której kierunek ruchu wzdłużnego (górną-dół) określa zwrot np. wektora $z \perp (x, y)$, wynikający z kierunku ruchu obrotowego: $x \rightarrow y$. Logicznym byłoby więc, aby niepełna transformacja inwersji trójki osi (np. $x \rightarrow -x, y \rightarrow -y, z \rightarrow z$), której odpowiada wartość wyznacznika $+1$ (por. 10-37), nie zmieniała skrętności układu, a niepełna transformacja (np. $x \rightarrow -x, y \rightarrow y, z \rightarrow z$) powodowała zmianę skrętności. Sprawdzenie tej hipotezy pozostawiamy Czytelnikowi.

Warto pamiętać:

- Ortogonalna transformacja w przestrzeni Euklidesowej zachowuje iloczyn skalarny.
- Obrazy bazy ortonormalnej to (inna) baza ortonormalna w przestrzeni Euklidesowej. I odwrotnie: jeżeli pod wpływem danej transformacji ortonormalna baza przekształca się w bazę ortonormalną to transformacja jest ortogonalna.
- Każda ortogonalna transformacja jest nieosobliwa; transformacja odwrotna jest również ortogonalna.
- Złożenie (iloczyn) dwóch transformacji ortogonalnych jest transformacją ortogonalną.

10.5 Transformacja podobieństwa

Kilkakrotnie wspominaliśmy już, że obrót układu współrzędnych, w którym wyrażamy współrzędne danego wektora (pozostającego w ustalonej pozycji) \mathbf{V} , jest równoważny obrotowi tego wektora w przeciwnym kierunku (na płaszczyźnie: o przeciwny kąt) w stosunku do nieruchomego układu współrzędnych. Opisy sytuacji „fizycznej” w jednym i drugim przypadku muszą być sobie równoważne⁵.

Rozpatrzmy sytuację, kiedy transformację – na przykład obrotu – zastosujemy zarówno do wektora jak i układu. Oznaczmy przez macierz \mathcal{A} macierz obracającą wektor \mathbf{V} i przeprowadzającą go (w danym układzie współrzędnych $\Sigma \equiv (x_1, x_2, x_3)$) w wektor \mathbf{W} . Mamy

$$\mathbf{W} = \mathcal{A} \mathbf{V}, \quad (10-39)$$

albo macierzowo

$$\begin{pmatrix} W_1 \\ W_2 \\ W_3 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \\ V_3 \end{pmatrix}. \quad (10-40)$$

Teraz poddamy obrotowi sam układ Σ . Macierz obrotu to macierz przejścia \mathcal{T} , która „obraca” stare wersory i przeprowadza je w nowe: $\mathbf{e}' = \mathcal{T} \mathbf{e}$ (por. punkt 9.6).

Wektor \mathbf{W} wyrażony w obróconym układzie Σ' to

$$\begin{aligned} \mathbf{W}' &= \mathcal{T} \mathbf{W} = \mathcal{T} \mathcal{A} \mathbf{V} = (\mathcal{A} \mathbf{V})' \\ &= \mathcal{T} \mathcal{A} (\mathcal{T}^{-1} \mathcal{T}) \mathbf{V} \\ &= (\mathcal{T} \mathcal{A} \mathcal{T}^{-1}) \mathcal{T} \mathbf{V} = (\mathcal{T} \mathcal{A} \mathcal{T}^{-1}) \mathbf{V}'. \end{aligned} \quad (10-41)$$

(Znaki ' oznaczają, że dana wielkość wektorowa jest wyrażana w obróconym układzie współrzędnych Σ' ; alternatywny sposób zapisu tego faktu to $\mathcal{T} \mathbf{V}$. Macierz \mathcal{A} opisywała obrót wektora \mathbf{V} w starym układzie. W nowym (obróconym) układzie ten obrót realizuje macierz $\mathcal{A}' = (\mathcal{T} \mathcal{A} \mathcal{T}^{-1})$:

$$\begin{array}{ccc} \mathcal{T} \mathbf{W} & = & (\mathcal{T} \mathcal{A} \mathcal{T}^{-1}) \mathcal{T} \mathbf{V}. \\ \downarrow & & \downarrow \quad \downarrow \\ \mathbf{W}' & = & \mathcal{A}' \quad \mathbf{V}' \end{array} \quad (10-42)$$

Mamy więc

$$\mathcal{A}' = \mathcal{T} \mathcal{A} \mathcal{T}^{-1} \quad \text{oraz} \quad \mathcal{A} = \mathcal{T}^{-1} \mathcal{A}' \mathcal{T}. \quad (10-43)$$

Ogólnie, kwadratowe macierze, \mathcal{B} i \mathcal{C} nazywamy *macierzami podobnymi*, jeżeli połączone są relacją

$$\mathcal{C} = \mathcal{Q}^{-1} \mathcal{B} \mathcal{Q}, \quad (10-44)$$

gdzie \mathcal{Q} jest pewną nieosobliwą macierzą. Mówimy, że macierz \mathcal{C} to macierz \mathcal{B} , poddana transformacji podobieństwa; macierz tej transformacji to macierz \mathcal{Q} .

Macierze, stanowiące reprezentację tego samego odwzorowania w dwóch różnych bazach są macierzami podobnymi. Reprezentację w bazie \mathbf{e}' uzyskujemy z reprezentacji w bazie \mathbf{e} , poddając tę ostatnią transformacji podobieństwa, której macierzą jest macierz przejścia \mathcal{T}^{-1} (macierz przejścia od bazy \mathbf{e}' do bazy \mathbf{e}).

⁵Pięknym przykładem takiej równoważności takich dwóch formalnych podejść są tzw. „obrazy” lub „podejścia” Schroedingera i Heisenberga – dwa równoważne formalizmy podstaw mechaniki kwantowej.

W przypadku ogólnym dokonująca tej transformacji macierz \mathcal{T}^{-1} nie musi być macierzą ortogonalną; możemy mieć do czynienia z obrotami uogólnionymi układu Σ , przy których np. jednostki poszczególnych osi ulegają skróceniu (wydłużeniu) w różnych stosunkach. Zawsze jednak żądamy nieosobliwości macierzy transformacji. Pierwsze z równań 10-43 rozpisane na współrzędne daje:

$$a'_{ij} = \sum_{k=1}^n \sum_{l=1}^n \tau_{ik} a_{kl} \tau_{lj}^{-1}. \quad (10-45)$$

Górny wskaźnik sum, n , to wymiar przestrzeni wektorowej; przez τ_{ik} oraz τ_{lj}^{-1} oznaczmy odpowiednio elementy macierzy \mathcal{T} oraz \mathcal{T}^{-1} . Jeżeli ta ostatnia jest macierzą ortogonalną to

$$\tau_{lj}^{-1} = \tau_{jl} \quad (10-46)$$

i 10-45 będzie miało postać

$$a'_{ij} = \sum_{k=1}^n \sum_{l=1}^n \tau_{ik} a_{kl} \tau_{jl}. \quad (10-47)$$

Ten wzór można zweryfikować rachunkami opartymi na definicji elementu macierzowego w określonej bazie. Zgodnie z 10-23 mamy

$$a'_{ij} = (\xi'_i, \mathcal{A} \xi'_j) = \left(\sum_{k=1}^n \tau_{ik} \xi_k, \mathcal{A} \sum_{l=1}^n \tau_{jl} \xi_l \right) = \sum_{k=1}^n \sum_{l=1}^n \tau_{ik} \tau_{jl} (\xi_k, \mathcal{A} \xi_l) = \sum_{k=1}^n \sum_{l=1}^n \tau_{ik} \tau_{jl} a_{kl}. \quad (10-48)$$

Powyższe wzory należy interpretować następująco: o macierzy \mathcal{A} możemy myśleć jako o pewnym operatorze, powodującym określone zmiany wektorów (elementów przestrzeni wektorowej). Elementy macierzy (współrzędne operatora – a_{ij}) zależą od wyboru bazy tej przestrzeni i mogą być przedstawione w postaci odpowiednich iloczynów skalarnych, w których występują wersory bazowe, por. 10.2. Wybór taki bywa zwykle podyktowany pewnymi symetriami, występującymi przy opisie obiektu fizycznego. Większość obiektów fizycznych posiada pewne „naturalne” osie symetrii, które stanowią taką „naturalną” bazę. Mogą to być na przykład osie kryształu, osie symetrii bryły sztywnej, której obrót opisujemy, itp. Często jednak złożoność układu fizycznego zmusza nas do stosowania układów współrzędnych (baz przestrzeni wektorowej), które są „naturalne” dla jednej składowej układu fizycznego, natomiast dla pozostałych składowych już nie. Dlatego musimy być przygotowani na konieczność dokonywania odpowiednich *transformacji* naszych operatorów (macierzy) – wyrażanie ich elementów w różnych układach współrzędnych. Równanie 10-45 stanowi taką ogólną „receptę” na dokonywanie takich transformacji.

10.6 Macierzy symetryczne i antysymetryczne; macierze hermitowskie i unitarne

Macierz \mathcal{A} nazywamy symetryczną jeżeli spełniony jest warunek

$$a_{ik} = a_{ki}, \quad (10-49)$$

dla każdego i i k , a więc elementy *symetryczne* względem przekątnej głównej są sobie równe. Z definicji macierzy transponowanej wynika, że macierz symetryczna spełnia

$$\mathcal{A} = \mathcal{A}^T. \quad (10-50)$$

Z kolei macierz będziemy nazywać antysymetryczną jeżeli

$$a_{ik} = -a_{ki}. \quad (10-51)$$

Kładąc w powyższym wzorze $k = i$ mamy $a_{ii} = -a_{ii}$, czyli wszystkie elementy na przekątnej głównej macierzy antysymetrycznej są równe zero. Powyższe własności symetrii mogą mieć miejsce wyłącznie dla macierzy kwadratowych (liczba wierszy = liczba kolumn).

Każda kwadratowa macierz może być przedstawiona w postaci sumy dwóch składników (macierzy), z których jeden jest symetryczny, a drugi – antysymetryczny. Wystarczy zauważyć, że

$$\mathcal{A} = \frac{1}{2}(\mathcal{A} + \mathcal{A}) + 0 = \frac{1}{2}(\mathcal{A} + \mathcal{A}) + \frac{1}{2}(\mathcal{A}^T - \mathcal{A}^T) = \frac{1}{2}(\mathcal{A} + \mathcal{A}^T) + \frac{1}{2}(\mathcal{A} - \mathcal{A}^T). \quad (10-52)$$

Pierwszy składnik tej sumy jest ewidentnie symetryczny, drugi – antysymetryczny. Ponieważ własność symetrii i antysymetrii występuje często w opisie układów fizycznych, warto sobie zdawać sprawę z tego, że używane narzędzia formalne (macierze) są „przygotowane” do uwzględnienia tej własności.

I znowu – w większości problemów fizyki klasycznej macierze z którymi mamy do czynienia są macierzami o elementach czysto rzeczywistych (przestrzenie wektorowe są rozpięte nad ciałem liczb rzeczywistych). Ale w mechanice kwantowej, której cały formalizm zbudowany jest na fundamencie rachunku macierzowego, jesteśmy już w świecie liczb zespolonych i macierze będą posiadały na ogół elementy zespolone. Fakt ten zmusza nas do wprowadzenia pewnych nowych definicji i uogólnienia pewnych poznanych już własności.

1. Macierzą sprzężoną $\bar{\mathcal{A}}$ do macierzy \mathcal{A} nazywamy macierz, której elementy są liczbami zespolonymi sprzężonymi do elementów macierzy \mathcal{A} ⁶

$$\bar{a}_{ik} = (a_{ik})^* = a_{ik}^*. \quad (10-53)$$

2. Macierzą sprzężoną po hermitowsku \mathcal{A}^\dagger do macierzy \mathcal{A} nazywamy macierz transponowaną w stosunku do macierzy $\bar{\mathcal{A}}$

$$\mathcal{A}^\dagger = (\mathcal{A}^*)^T = (\mathcal{A}^T)^*. \quad (10-54)$$

3. Macierz \mathcal{A} nazywamy hermitowską, albo macierzą samosprzężoną, jeżeli jest równa macierzy sprzężonej do niej po hermitowsku

$$\mathcal{A} = \mathcal{A}^\dagger. \quad (10-55)$$

Jeżeli macierz \mathcal{A} jest macierzą rzeczywistą (wszystkie jej elementy są liczbami rzeczywistymi) to macierz sprzężona do niej po hermitowsku jest po prostu macierzą transponowaną: $\mathcal{A}^\dagger = \mathcal{A}^T$; rzeczywiste macierze hermitowskie są macierzami symetrycznymi. Własność „hermitowskości” stanowi więc uogólnienie własności symetrii na przypadek macierzy zespolonych.

4. Macierz \mathcal{U} nazywamy unitarną, jeżeli macierz sprzężona do niej po hermitowsku \mathcal{U}^\dagger jest równa macierzy odwrotnej \mathcal{U}^{-1} :

$$\mathcal{U}^{-1} = \mathcal{U}^\dagger. \quad (10-56)$$

Znowu – jeżeli macierz \mathcal{U} jest rzeczywista to macierz niej po hermitowsku jest macierzą transponowaną: $\mathcal{U}^\dagger = \mathcal{U}^T$ i własność „unitarności” więc jest uogólnieniem własności ortogonalności na przypadek macierzy zespolonych. Z macierzami unitarnymi mamy do czynienia w mechanice kwantowej. Podobnie jak macierze ortogonalne w rzeczywistych przestrzeniach wektorowych pozostawiają bez zmiany długość wektora, macierze unitarne w formalizmie kwantowym pozostawiają bez zmian „długość wektora” (kwadrat modułu funkcji falowej – funkcji w której tkwi cała informacja o stanie układu).

10.6.1 Zespolone przestrzenie liniowe

Przestrzenie wektorowe mogą być przestrzeniami rozpiętymi nad ciałem liczb zespolonych; współrzędne wektora będą wówczas liczbami zespolonymi. Jeżeli dla przestrzeni takich chcemy określić iloczyn skalarny, to musimy poddać drobnej modyfikacji wzór 10-4. Chcemy aby kwadrat normy wektora był zawsze liczbą dodatnią i dlatego

$$(\mathbf{V}, \mathbf{W}) = \sum_{i=1}^n V_i^* W_i \quad (10-57)$$

– współrzędne lewego czynnika iloczynu podlegają operacji sprzężenia zespolonego. Wówczas

$$(\mathbf{x}, \mathbf{x}) = \sum_{i=1}^n x_i^* x_i = \sum_{i=1}^n |x_i|^2 \geq 0. \quad (10-58)$$

(Znak równości ma miejsce tylko w przypadku wektora zerowego.) Z definicji (10-57) wynika, że w przypadku przestrzeni wektorowej rozpiętej nad ciałem liczb zespolonych iloczyn skalarny jest relacją asymetryczną:

$$(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n x_i^* y_i = \left(\sum_{i=1}^n x_i y_i^* \right)^* = (\mathbf{y}, \mathbf{x})^*. \quad (10-59)$$

⁶W tym paragrafie jako znak sprzężenia zespolonego stosujemy „równoległe” * i $\bar{}$.

Rozdział 11

Problem własny

11.1 Pierwiastki charakterystyczne i wartości własne macierzy

Założmy, że mamy *rzeczywistą* przestrzeń wektorową \mathbb{R}^n . Rozpatrujemy nieosobliwe odwzorowanie tej przestrzeni w samą siebie; macierz odwzorowania \mathcal{A} to macierz $n \times n$ (o elementach rzeczywistych). *Macierzą charakterystyczną* macierzy \mathcal{A} nazywamy macierz $\mathcal{A} - \lambda\mathcal{I}$, gdzie λ to pewna niewiadoma liczba, natomiast \mathcal{I} – macierz jednostkowa:

$$\mathcal{A} - \lambda\mathcal{I} = \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{pmatrix}. \quad (11-1)$$

Wyznacznik tej macierzy jest wielomianem zmiennej λ stopnia n : iloczyn wyrazów leżących na głównej przekątnej zawiera wyraz $(-1)^n\lambda$; w pozostałych wyrazach rozwinięcia wyznacznika brak przynajmniej dwóch wyrazów z głównej przekątnej. Współczynniki poszczególnych potęg λ wielomianu łatwo jest określić. Np. współczynnik λ^{n-1} to $(-1)^{n-1}(a_{11} + a_{22} + \dots + a_{nn})$, a wyraz wolny – wyznacznik macierzy \mathcal{A} .

Wielomian $\det(\mathcal{A} - \mathcal{I})$ nazywamy *wielomianem charakterystycznym* macierzy \mathcal{A} , a jego pierwiastki — rzeczywiste, bądź zespolone – jej *pierwiastkami charakterystycznymi*. Termin: charakterystyczny jest uzasadniony. Jeżeli rozpatrzyć \mathcal{B} – macierz podobną do macierzy \mathcal{A} (por. punkt 10.5):

$$\mathcal{B} = \mathcal{Q}^{-1}\mathcal{A}\mathcal{Q},$$

to jej wielomian charakterystyczny jest identyczny jak wielomian charakterystyczny \mathcal{A} :

$$\begin{aligned} \det(\mathcal{B} - \lambda\mathcal{I}) &= \det(\mathcal{Q}^{-1}\mathcal{A}\mathcal{Q} - \lambda\mathcal{I}) = \det(\mathcal{Q}^{-1}[\mathcal{A} - \lambda\mathcal{I}]\mathcal{Q}) \\ &= \det(\mathcal{Q}^{-1})\det(\mathcal{A} - \lambda\mathcal{I})\det(\mathcal{Q}) = \det(\mathcal{A} - \lambda\mathcal{I}). \end{aligned} \quad (11-2)$$

(Wykorzystaliśmy przemienność macierzy jednostkowej z każdą macierzą oraz to, że $\det(\mathcal{Q}^{-1}) = 1/\det(\mathcal{Q})$.) Wynika stąd ważny wniosek: mimo, że transformacja T może być reprezentowana – w różnych bazach – przez różne macierze odwzorowania, to te różne macierze mają wszystkie pierwiastki charakterystyczne takie same – nazywane czasem pierwiastkami charakterystycznymi transformacji. Zbiór tych wszystkich pierwiastków, z uwzględnieniem ich krotności, nazywa się *widmem operatora* transformacji, albo *widmem transformacji*. Podstawowe twierdzenie algebry gwarantuje nam, że jeżeli wielomian charakterystyczny jest wielomianem n -ego stopnia, to zbiór jego pierwiastków – rzeczywistych i urojonych, z uwzględnieniem ich krotności – ma liczebność równą n .

Dla fizyki ten właśnie rozdział ma kapitalne znaczenie. Charakterystyczne pierwiastki macierzy kojarzą się bowiem z odwzorowaniem T typu:

$$T\mathbf{b} = \lambda_0\mathbf{b}, \quad (11-3)$$

w wyniku którego *niezerowy*¹ wektor \mathbf{b} – obiekt transformacji – zostaje po prostu przemnożony przez pewną liczbę. Taki wektor nazywamy *wektorem własnym* odwzorowania T , albo *wektorem własnym macierzy transformacji* \mathcal{T} ; liczba λ_0 to *wartość własna* transformacji (macierzy). Mówimy, że wektor własny \mathbf{b} odpowiada wartości własnej λ_0 .

¹Niezerowy, bo dla wektora zerowego takie równanie jest trywialne: $0 = 0$.

Najprostszym przykładem transformacji, której wektorami własnymi są *wszystkie* wektory to zwykle przeskalowanie osi naszego układu współrzędnych. Jeżeli zmieniamy jednostki osi z cali na centymetry, to wektorami własnymi transformacji są wszystkie wektory naszej przestrzeni, a wartość własna to 2,54. Równie prosty przykład – ale transformacji, która nie posiada wektorów własnych to obrót płaszczyzny euklidesowej, wokół początku układu, o kąt nie będący wielokrotnością π . A w fizyce? Klasycznym przykładem problemu własnego w fizyce jest problem drgającej struny (długość L), której końce są uwięzione.² Wychylenia struny – funkcja $y(x, t)$ spełniają równanie falowe:

$$\frac{\partial^2 y}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 y}{\partial t^2} = 0, \quad (11-4)$$

gdzie c jest prędkością propagacji fali w strunie. Standardowa metoda rozwiązywania równań tego typu³ nakazuje szukać rozwiązania w postaci iloczynu dwóch funkcji $y(x, t) = X(x)T(t)$, co prowadzi do zastąpienia (jednego) równ. 11-4 dwoma równaniami, dla funkcji $X(x)$ i $T(t)$:

$$\left. \begin{aligned} \frac{1}{X} \frac{d^2 X}{dx^2} &= C \\ \frac{1}{c^2 T} \frac{d^2 T}{dt^2} &= C, \end{aligned} \right\}, \quad (11-5)$$

gdzie C jest – w zasadzie dowolną – stałą. Unieruchomienie obu końców struny narzuca na funkcję X warunki:

$$X(0) = X(L) = 0. \quad (11-6)$$

Nietrudno sprawdzić, że pierwsze z równań 11-5 nie posiada rozwiązań, które spełniałyby 11-6 przy $C > 0$ i $C = 0$. Stała C musi być ujemna – $C \equiv -k^2$ i pierwsze z równań 11-5 przybiera postać – analogiczną do równ. 11-3 –

$$\frac{d^2 X}{dx^2} = -k^2 X. \quad (11-7)$$

Jego ogólne rozwiązanie (por. podrozdział 12.1) to

$$X(x) = C_1 \cos kx + C_2 \sin kx, \quad (11-8)$$

ale pierwszy z warunków 11-6 narzuca $C_1 = 0$. Z kolei, drugi warunek to

$$X(L) = \sin kL = 0, \quad (11-9)$$

a więc $kL = n\pi$, albo

$$k \equiv k_n = \frac{n\pi}{L}; \quad X(x) \equiv X_n(x) = C_n \sin \frac{n\pi}{L} x \quad (11-10)$$

– stała C_n nie jest w tym momencie istotna. Nasze równanie własne – 11-7 – posiada więc nieskończony i przeliczalny zbiór wartości własnych i odpowiadających im funkcji własnych.

Powróćmy do macierzy. Z określenia wielomianu charakterystycznego macierzy odwzorowania i wartości własnych wynika:

Twierdzenie 11.1 *Tylko rzeczywiste pierwiastki równania charakterystycznego mogą być wartościami własnymi transformacji.*

Dowód: zapisujemy wektor \mathbf{b} i macierz odwzorowania $\mathcal{T} \equiv (t_{ij})$ w tej samej bazie:

$$\mathbf{b} = \sum_{i=1}^n \beta_i \boldsymbol{\xi}_i; \quad t_{ij} = (\boldsymbol{\xi}_i, \mathcal{T} \boldsymbol{\xi}_j). \quad (11-11)$$

²Problem ten dyskutujemy także w rozdziale 12. Tam też znajdziesz kilka rysunków drgającej struny.

³Metoda separacji zmiennych. Możesz ją znaleźć w drugim rozdziale MMF II i III. W trzecim rozdziale znajdziesz problem struny omówiony szczegółowo.

Równanie 11-3 zapisane macierzowo (por. 9-49)

$$T\mathbf{b} = (\beta_1, \beta_2, \dots, \beta_n) \begin{pmatrix} t_{11} & t_{12} & \dots & t_{1n} \\ t_{21} & t_{22} & \dots & t_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ t_{n1} & t_{n2} & \dots & t_{nn} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix} = \lambda_0 \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} \quad (11-12)$$

jest równoważne układowi n równań

$$\left. \begin{aligned} \beta_1 t_{11} + \beta_2 t_{21} + \dots + \beta_n t_{n1} &= \lambda_0 \beta_1 \\ \beta_1 t_{12} + \beta_2 t_{22} + \dots + \beta_n t_{n2} &= \lambda_0 \beta_2 \\ \vdots & \\ \beta_1 t_{1n} + \beta_2 t_{2n} + \dots + \beta_n t_{nn} &= \lambda_0 \beta_n \end{aligned} \right\}, \quad (11-13)$$

a ponieważ wektor \mathbf{b} jest – z założenia – niezerowy, to warunkiem aby istniało rozwiązanie (nietrywialne!) tego układu

$$\left. \begin{aligned} (t_{11} - \lambda_0)\beta_1 + t_{21}\beta_2 + \dots + t_{n1}\beta_n &= 0 \\ t_{12}\beta_1 + (t_{22} - \lambda_0)\beta_2 + \dots + t_{n2}\beta_n &= 0 \\ \vdots & \\ t_{1n}\beta_1 + t_{2n}\beta_2 + \dots + (t_{nn} - \lambda_0)\beta_n &= 0 \end{aligned} \right\}, \quad (11-14)$$

jest zerowanie się jego wyznacznika:

$$\begin{vmatrix} t_{11} - \lambda_0 & t_{21} & \dots & t_{n1} \\ t_{12} & t_{22} - \lambda_0 & \dots & t_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ t_{1n} & t_{2n} & \dots & t_{nn} - \lambda_0 \end{vmatrix}. \quad (11-15)$$

jeżeli wyznacznik ten jest równy zeru, to i wyznacznik macierzy transponowanej w stosunku do 11-15 jest równy zeru; czyli

$$\det(\mathcal{T} - \lambda_0 \mathcal{I}) = 0, \quad (11-16)$$

czyli wartość własna λ_0 rzeczywiście stanowi pierwiastek równania charakterystycznego. Rzeczywisty charakter λ_0 wynika z 11-13. Pozostawiamy czytelnikowi wykazanie tego twierdzenia „w drugą stronę”.

11.1.1 Niezdegenerowane wartości własne

Mówimy, że wartości własne liniowej transformacji przestrzeni \mathbb{R}^n są niezdegenerowane jeżeli danej wartości własnej odpowiada jeden i tylko jeden wektor własny. Zachodzi to w przypadku, gdy wszystkie pierwiastki wielomianu charakterystycznego macierzy są różne i rzeczywiste. Czasem używa się też określenia: widmo wartości własnych jest widmem prostym.

(Z degeneracją mamy do czynienia, kiedy określonej wartości własnej odpowiada więcej niż jeden wektor własny. Tak będzie w przypadku, gdy wartością własną jest wielokrotny pierwiastek wielomianu charakterystycznego.)

11.2 Problem własny a diagonalizacja macierzy

W rozwiązywaniu problemów fizycznych często szukamy takiej transformacji przestrzeni wektorowej, której macierz – w określonej bazie – ma postać diagonalną, tzn. różne od zera są tylko elementy przekątnej głównej. Nie oznacza to, że *każda* transformacja liniowa może być reprezentowana przez macierz diagonalną! O problemie tym będziemy mówić w następnym podrozdziale.

W tym momencie wykazemy:

Twierdzenie 11.2 *Macierz \mathcal{T} transformacji liniowej T w określonej bazie $\mathbf{e}_1, \dots, \mathbf{e}_n$ może być diagonalna wtedy i tylko wtedy gdy wszystkie wektory bazowe są wektorami własnymi \mathcal{T} .*

Wektory własne $\mathbf{b}_1, \dots, \mathbf{b}_k$ liniowej transformacji T , odpowiadające różnym wartościom własnym tworzą układ wektorów liniowo niezależnych.

Rzeczywiście, macierzowe równanie własne, dla wszystkich wektorów bazowych

$$T(\mathbf{e}_i) = \lambda_i \mathbf{e}_i; \quad i = 1, \dots, n \quad (11-17)$$

czyli *explicite*

$$\begin{pmatrix} t_{11} & t_{12} & \dots & t_{1i} & \dots & t_{1n} \\ t_{21} & t_{22} & \dots & t_{2i} & \dots & t_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ t_{n1} & t_{n2} & \dots & t_{ni} & \dots & t_{nn} \end{pmatrix} \begin{pmatrix} 0_{[1]} \\ \vdots \\ 0_{[i-1]} \\ 1_{[i]} \\ 0_{[i+1]} \\ \vdots \\ 0_{[n]} \end{pmatrix} = \begin{pmatrix} 0_{[1]} \\ \vdots \\ 0_{[i-1]} \\ \lambda_{i[i]} \\ 0_{[i+1]} \\ \vdots \\ 0_{[n]} \end{pmatrix}; \quad i = 1, \dots, n \quad (11-18)$$

będzie spełnione wtedy i tylko wtedy gdy macierz $\mathcal{T} \equiv (t_{ij})$ ma postać

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}. \quad (11-19)$$

Dowód liniowej niezależności \mathbf{b}_i najprościej jest przeprowadzić metodą indukcji matematycznej względem wskaźnika k . Dla $k = 1$ mamy układ liniowo niezależny (jeden wektor). Rozważamy wektory spełniające równanie własne

$$T(\mathbf{b}_i) = \lambda_i \mathbf{b}_i; \quad i = 1, \dots, k, \quad (11-20)$$

przy czym

$$\lambda_i \neq \lambda_j \quad \text{dla} \quad i \neq j. \quad (11-21)$$

Dla $1 \leq l < k$ zakładamy, że układ l wektorów jest liniowo niezależny; równanie

$$\beta_1 \mathbf{b}_1 + \beta_2 \mathbf{b}_2 + \dots + \beta_l \mathbf{b}_l = 0 \quad (11-22)$$

ma rozwiązanie tylko dla wszystkich $\beta_i = 0$. Jeżeli dodanie $(l + 1)$ -ego wektora niszczy liniową niezależność, tzn. równanie

$$\alpha_1 \mathbf{b}_1 + \alpha_2 \mathbf{b}_2 + \dots + \alpha_l \mathbf{b}_l + \alpha_{l+1} \mathbf{b}_{l+1} = 0 \quad (11-23)$$

ma rozwiązanie dla – na przykład – $\alpha_1 \neq 0$, to stosujemy transformację T do obu stron tego równania, otrzymując

$$\lambda_1 \alpha_1 \mathbf{b}_1 + \lambda_2 \alpha_2 \mathbf{b}_2 + \dots + \lambda_l \alpha_l \mathbf{b}_l + \lambda_{l+1} \alpha_{l+1} \mathbf{b}_{l+1} = 0. \quad (11-24)$$

Równanie 11-23 mnożymy przez λ_{l+1} i otrzymane równanie odejmujemy od 11-24. Dostajemy

$$\alpha_1 (\lambda_1 - \lambda_{l+1}) \mathbf{b}_1 + \alpha_2 (\lambda_2 - \lambda_{l+1}) \mathbf{b}_2 + \dots + \alpha_l (\lambda_l - \lambda_{l+1}) \mathbf{b}_l = 0, \quad (11-25)$$

co stanowi zaprzeczenie założenia (równ. 11-22) o niezależności układu l wektorów – współczynnik $\alpha_1 (\lambda_1 - \lambda_{l+1})$ jest różny od zera.

11.3 Odwzorowania symetryczne

Liniowe odwzorowanie Euklidesowej przestrzeni \mathbb{E}^n nazywamy *symetrycznym* albo *samosprzężonym*⁴ jeżeli spełniona jest równość:

$$(S\{\mathbf{a}\}, \mathbf{b}) = (\mathbf{a}, S\{\mathbf{b}\}); \quad (11-26)$$

⁴To określenie odnosi się raczej do operatorów; samosprzężonym może być operator działający na wektory i poddający je odwzorowaniu. Będzie o tym mowa w fizyce!

iloczyn skalarny nie zależy od tego na który z jego czynników działa operator odwzorowania.⁵

Najprostszymi przykładami symetrycznego odwzorowania będzie odwzorowanie tożsamościowe (wektory przechodzą w same siebie) i zerowe (obrazem wszystkich wektorów jest wektor zerowy). Mniej banalne – wspomniane już przemnożenie (wszystkich!) wektorów przez pewną liczbę (skalowanie).

Wykażemy

Twierdzenie 11.3 *Symetryczne odwzorowanie przestrzeni Euklidesowej \mathbb{E}^n jest reprezentowane, w dowolnej bazie ortonormalnej, przez macierz symetryczną. I odwrotnie: jeżeli odwzorowanie jest reprezentowane przez symetryczną (przynajmniej w jednej ortonormalnej bazie) macierz, to odwzorowanie jest symetryczne.*

Dowód: Załóżmy, że symetrycznemu odwzorowaniu S odpowiada, w ortonormalnej bazie $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ macierz $\mathcal{S} \equiv (s_{ij})$. Obliczamy odpowiednie iloczyny skalarne (por. 10-24 i 10-23):

$$(S\{\mathbf{e}_i\}, \mathbf{e}_j) = \left(\sum_{k=1}^n s_{ki} \mathbf{e}_k, \mathbf{e}_j \right) = s_{ji}, \quad (11-27)$$

$$(\mathbf{e}_i, S\{\mathbf{e}_j\}) = \left(\mathbf{e}_i, \sum_{k=1}^n s_{kj} \mathbf{e}_k \right) = s_{ij} \quad (11-28)$$

i korzystamy z 11-26. Mamy:

$$s_{ij} = s_{ji} \quad (11-29)$$

dla wszystkich $i, j = 1, \dots, n$. Macierz \mathcal{S} jest symetryczna.

Druga część dowodu jest równie prosta. Zakładamy, że macierz odwzorowania liniowego w ortonormalnej bazie $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ to pewna symetryczna macierz $\mathcal{S} \equiv (s_{ij})$, czyli dla wszystkich $i, j = 1, \dots, n$ zachodzi $s_{ij} = s_{ji}$. Biorąc dwa dowolne wektory, \mathbf{b} i \mathbf{c} , których reprezentacje w naszej bazie to:

$$\mathbf{b} = \sum_{i=1}^n \beta_i \mathbf{e}_i, \quad \mathbf{c} = \sum_{j=1}^n \gamma_j \mathbf{e}_j \quad (11-30)$$

poddajemy je transformacji S :

$$S\{\mathbf{b}\} = \sum_{i=1}^n \beta_i S\{\mathbf{e}_i\} = \sum_{j=1}^n \left(\sum_{i=1}^n \beta_i s_{ji} \right) \mathbf{e}_j, \quad (11-31)$$

$$S\{\mathbf{c}\} = \sum_{j=1}^n \gamma_j S\{\mathbf{e}_j\} = \sum_{i=1}^n \left(\sum_{j=1}^n \gamma_j s_{ij} \right) \mathbf{e}_i. \quad (11-32)$$

Korzystając z ortonormalności bazy mamy

$$(S\{\mathbf{b}\}, \mathbf{c}) = \sum_{j=1}^n \sum_{i=1}^n \beta_i s_{ji} \gamma_j \quad (11-33)$$

$$(\mathbf{b}, S\{\mathbf{c}\}) = \sum_{i=1}^n \sum_{j=1}^n \beta_i s_{ij} \gamma_j. \quad (11-34)$$

Ponieważ macierz (s_{ij}) jest symetryczna prawe strony obu równań są identyczne; $(S\{\mathbf{b}\}, \mathbf{c}) = (\mathbf{b}, S\{\mathbf{c}\})$. Czytelnik może wykazać dwa, intuicyjnie łatwe do zaakceptowania wnioski:

Twierdzenie 11.4 *Suma dwóch symetrycznych odwzorowań, a także ich iloczyn, jest odwzorowaniem symetrycznym.*

Istotnym – właśnie w kontekście mechaniki kwantowej – jest :

⁵Tego typu odwzorowania są, można to powiedzieć bez przesady, fundamentem mechaniki kwantowej. Przestrzenie wektorowe mechaniki kwantowej są przestrzeniami zespolonymi i nieskończenie wielo-wymiarowymi. Wymaga to innego nieco określenia iloczynu skalarnego; w dodatku (ze względu na zespolony charakter) ten ostatni jest antysymetryczny (por. 10-59). Dlatego, to co będziemy omawiali w tym podrozdziale dla prostszego przypadku rzeczywistych przestrzeni euklidesowych \mathbb{E}^n o skończonej (n) liczbie wymiarów „przenosi się” prosto na przypadek kwantowo-mechaniczny, jeżeli terminy „ortogonalny” i „symetryczny” zastąpić odpowiednio terminami „unitarny” oraz „hermitowski”.

Twierdzenie 11.5 Wszystkie wartości własne symetrycznej macierzy są rzeczywiste.

Dowód: Wystarczy wykazać, że rzeczywiste są wszystkie pierwiastki wielomianu charakterystycznego macierzy symetrycznej $\mathcal{S} \equiv (s_{ij})$, tzn. wszystkie λ_0 spełniające

$$\det(\mathcal{S} - \lambda_0 \mathcal{I}) = 0. \quad (11-35)$$

Równanie jest równoważne układowi n równań jednorodnych:

$$\sum_{j=1}^n s_{ji} x_j - \lambda_0 x_i = 0; \quad i = 1, \dots, n. \quad (11-36)$$

Ponieważ λ_0 (pierwiastek wielomianu charakterystycznego) może być zespolone, to rozwiązania tego układu (istnieją, bo jego wyznacznik = 0 – por. 11-35) $x_i \equiv \beta_i; i = 1, \dots, n$ są (mogą być) też zespolone. Mamy

$$\sum_{j=1}^n s_{ji} \beta_j = \lambda_0 \beta_i; \quad i = 1, \dots, n. \quad (11-37)$$

Mnożymy obie strony każdego z równań 11-37 przez zespoloną sprzężoną $\bar{\beta}_i$ i dodajemy do siebie lewe i prawe strony n równań:

$$\sum_{i=1}^n \sum_{j=1}^n s_{ji} \beta_j \bar{\beta}_i = \lambda_0 \sum_{i=1}^n \beta_i \bar{\beta}_i. \quad (11-38)$$

Występująca po prawej stronie powyższego równania suma to pewna liczba rzeczywista – przynajmniej jedno z β_i musi być różne od zera (układ ma rozwiązanie nietrywialne) i dodajemy kwadraty modułów liczb zespolonych. „Rzeczywista natura” λ_0 wymagałaby więc wykazania, że i lewa strona 11-38 jest liczbą rzeczywistą. Wykażemy to, obliczając jej zespolone sprzężenie (uwaga: współczynniki s_{ji} są rzeczywiste!)⁶

$$\begin{aligned} \left(\sum_{i=1}^n \sum_{j=1}^n s_{ji} \beta_j \bar{\beta}_i \right)^* &= \sum_{i=1}^n \sum_{j=1}^n (s_{ji} \beta_j \bar{\beta}_i)^* = \sum_{i=1}^n \sum_{j=1}^n s_{ji} \bar{\beta}_j \beta_i \\ &\stackrel{s_{ij}=s_{ji}}{=} \sum_{i=1}^n \sum_{j=1}^n s_{ij} \bar{\beta}_j \beta_i \stackrel{i \leftrightarrow j}{=} \sum_{j=1}^n \sum_{i=1}^n s_{ji} \bar{\beta}_i \beta_j. \end{aligned}$$

Czytelnik, który już dzisiaj chciałby poznać podwaliny formalne mechaniki kwantowej, bez trudu prześledzi tok dowodu dla macierzy, której elementy spełniają warunek hermitowskości (por. 10.6) – $s_{ij}^* = s_{ji}$. W końcu mamy bardzo istotne :

Twierdzenie 11.6 Liniowe odwzorowanie \mathbb{E}^n jest symetryczne wtedy i tylko wtedy, jeżeli w \mathbb{E}^n istnieje ortonormalna baza, zbudowana z wektorów własnych macierzy odwzorowania.

Dowód „w jedną stronę”: jeżeli w \mathbb{E}^n istnieje ortonormalna baza $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n$ i dla wektorów bazy mamy

$$S\{\mathbf{f}_i\} = \lambda_i \mathbf{f}_i; \quad i = 1, \dots, n \quad (11-39)$$

to macierz transformacji 11-39 w bazie $\{\mathbf{e}_i\}$ jest diagonalna

$$\begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}. \quad (11-40)$$

Diagonalna macierz jest jednak symetryczna, a więc symetryczna będzie i samo odwzorowanie.

Dowód „w drugą” stronę pozostawiamy czytelnikowi. Można go przeprowadzić metodą indukcji względem n (dla $n = 1$ odwzorowanie to wspomniane na początku podrozdziału skalowanie).

⁶Używamy równoległe symbolów: x^* i \bar{x} dla oznaczenia liczby zespolonej sprzężonej do x .

11.4 Sprowadzanie formy kwadratowej do układu osi głównych

Z rozważań zawartych w poprzednich punktach wynika kolejne twierdzenie, bardzo istotne dla fizyki – i to nie tylko kwantowej, ale i tej klasycznej:

Twierdzenie 11.7 *Dla każdej symetrycznej macierzy \mathcal{S} można znaleźć ortogonalną macierz \mathcal{Q} , która diagonalizuje macierz \mathcal{S} ; tzn. macierz $\mathcal{Q}^{-1}\mathcal{S}\mathcal{Q}$ jest macierzą diagonalną.*

Twierdzenie można właściwie traktować jako pewne podsumowanie naszych wiadomości o odwzorowaniach i ich macierzach. Jeżeli mamy \mathcal{S} , macierz symetryczną rzędu n i ortonormalną bazę $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ w \mathbb{E}^n , to macierz \mathcal{S} reprezentuje w tej bazie pewne symetryczne odwzorowanie S tej przestrzeni euklidesowej. A jeżeli tak, to w \mathbb{E}^n istnieje też baza ortonormalna $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n$, zbudowana z wektorów własnych macierzy odwzorowania S . W tej bazie, macierz odwzorowania S to diagonalna macierz \mathcal{D} :

$$\mathcal{D} = \mathcal{Q}^{-1}\mathcal{S}\mathcal{Q}. \quad (11-41)$$

gdzie \mathcal{Q} jest macierzą przejścia od bazy $\{\mathbf{f}_i\}$ do bazy $\{\mathbf{e}_i\}$:

$$\mathbf{e} = \mathcal{Q}\mathbf{f}. \quad (11-42)$$

Macierz \mathcal{Q} jest ortogonalna, ponieważ stanowi przejście pomiędzy dwoma bazami ortonormalnymi, a ponieważ macierz odwrotna ortogonalnej macierzy \mathcal{Q} jest równa macierzy transponowanej: $\mathcal{Q}^{-1} = \mathcal{Q}^T$ równanie 11-41 można przepisać w postaci

$$\mathcal{D} = \mathcal{Q}^T\mathcal{S}\mathcal{Q}. \quad (11-43)$$

W rozdziale ósmym widzieliśmy jednak, że taką samą postać ma forma kwadratowa: symetryczna macierz \mathcal{S} współczynników formy jest mnożona od prawej i od lewej przez wektor niewiadomych x_i . Na podstawie właśnie wykazanego twierdzenia możemy sformułować kolejne, tzw.

Twierdzenie 11.8 (O osiach głównych:) *Każda rzeczywista forma kwadratowa $f(x_1, x_2, \dots, x_n)$ może zostać sprowadzona do postaci kanonicznej, przy pomocy odwzorowania, którego macierz jest ortogonalna.*

W dodatku:

Twierdzenie 11.9 *Bez względu na to jaką postać ma transformacja ortogonalna, sprowadzająca formę kwadratową, określoną macierzą \mathcal{S} , do postaci kanonicznej, współczynniki tej ostatniej są pierwiastkami wielomianu charakterystycznego macierzy \mathcal{S} (z uwzględnieniem ich krotności.)*

Dowód: Przypuśćmy, że mamy formę f , która w wyniku transformacji ortogonalnej $x_i \rightarrow y_i$; $i = 1, \dots, n$ została sprowadzona do postaci kanonicznej:

$$f(x_1, x_2, \dots, x_n) = \mu_1 y_1^2 + \mu_2 y_2^2 + \dots + \mu_n y_n^2. \quad (11-44)$$

Ponieważ transformacja jest ortogonalna to suma kwadratów x_i , pozostaje bez zmian $\sum_{i=1}^n x_i^2 = \sum_{i=1}^n y_i^2$. Dla dowolnej liczby λ zachodzi

$$f(x_1, x_2, \dots, x_n) - \lambda \sum_{i=1}^n x_i^2 = \sum_{i=1}^n \mu_i y_i^2 - \lambda \sum_{i=1}^n y_i^2. \quad (11-45)$$

Wyznaczniki macierzy, które występują po obu stronach są równe (por. wzory 11-2, a także 11-41 i 11-43) dostajemy

$$\det(\mathcal{S} - \lambda\mathcal{I}) = \begin{vmatrix} \mu_1 - \lambda & 0 & \dots & 0 \\ 0 & \mu_2 - \lambda & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mu_n - \lambda \end{vmatrix} = \prod_{i=1}^n (\mu_i - \lambda). \quad (11-46)$$

Verbalizując jeszcze inaczej wykazane twierdzenie: bez względu na to jaka jest postać macierzy ortogonalnej, która diagonalizuje macierz formy \mathcal{S} , na przekątnej zdiagonalizowanej macierzy pojawiają się pierwiastki charakterystyczne \mathcal{S} (z uwzględnieniem ich krotności).

Znajdowanie odwzorowania, sprowadzającego kwadratową formę do osi głównych

W większości zagadnień interesuje nas nie tylko postać kanoniczna formy kwadratowej, ale transformacja ortogonalna \mathcal{Q} , która tej „redukcji do osi głównych” dokonuje. Praktyczne znalezienie \mathcal{Q} polega na znalezieniu macierzy odwzorowania diagonalizującego symetryczną macierz \mathcal{S} , albo macierzy do niej odwrotnej \mathcal{Q}^{-1} . Z równania 11-42 wynika, że ta ostatnia to macierz przejścia od bazy $\{\mathbf{e}\}$ do bazy $\{\mathbf{f}\}$; jej wiersze to n wektorów własnych symetrycznego odwzorowania, którego macierz – w bazie $\{\mathbf{e}\}$ – to nasza symetryczna macierz \mathcal{S} . Innymi słowy, zadanie sprowadza się do znalezienia układu wektorów własnych macierzy.

Założmy, że λ_0 jest pierwiastkiem charakterystycznym macierzy \mathcal{S} i jego krotność jest równa k_0 . W punkcie 11.2 przekonaliśmy się, że wektory własne macierzy transformacji, odpowiadające danej wartości własnej λ_0 to nietrywialne rozwiązania układu jednorodnych równań (por. 11-12) liniowych

$$(\mathcal{S} - \lambda_0 \mathcal{I})\mathbf{b} = 0. \quad (11-47)$$

Wprawdzie w układzie 11-13 występowała macierz transponowana – \mathcal{T}^T , ale nasza macierz \mathcal{S} jest symetryczna i zamiast \mathcal{S}^T możemy w równ. 11-47 wstawić \mathcal{S} . Jak wynika to z twierdzenia o osiach głównych taki układ posiada k_0 liniowo niezależnych rozwiązań. Szukamy tych rozwiązań w standardowy sposób; po znalezieniu ich poddajemy ich procesowi ortogonalizacji i normalizacji.

Występująca w równ. 11-47 wielkość λ_0 reprezentuje wszystkie różne pierwiastki charakterystyczne; ponieważ suma krotności tych pierwiastków jest równa n to, rozwiązując 11-47 dla kolejnych λ_0 , dostajemy układ n wektorów własnych odwzorowania S , w bazie $\{\mathbf{e}\}$. Jedyne co pozostaje do wykazania to pomocnicze:

Twierdzenie 11.10 *Wektory własne macierzy symetrycznego odwzorowania S przynależne do różnych wartości własnych są wzajemnie ortogonalne.*

Dowód: Weźmy dwa wektory \mathbf{b} i \mathbf{c} :

$$S\{\mathbf{b}\} = \lambda_1 \mathbf{b}, \quad S\{\mathbf{c}\} = \lambda_2 \mathbf{c}, \quad (11-48)$$

przy czym $\lambda_1 \neq \lambda_2$. Obliczmy iloczyny skalarne:

$$(S\{\mathbf{b}\}, \mathbf{c}) = (\lambda_1 \mathbf{b}, \mathbf{c}) = \lambda_1 (\mathbf{b}, \mathbf{c}), \quad (11-49)$$

$$(\mathbf{b}, S\{\mathbf{c}\}) = (\mathbf{b}, \lambda_2 \mathbf{c}) = \lambda_2 (\mathbf{b}, \mathbf{c}). \quad (11-50)$$

Ponieważ odwzorowanie jest symetryczne lewe strony 11-49 i 11-50 są równe, tak więc

$$\lambda_1 (\mathbf{b}, \mathbf{c}) = \lambda_2 (\mathbf{b}, \mathbf{c}), \quad (11-51)$$

lub

$$(\lambda_1 - \lambda_2)(\mathbf{b}, \mathbf{c}) = 0; \quad (11-52)$$

z założenia $\lambda_1 \neq \lambda_2$ wynika ortogonalność \mathbf{b} i \mathbf{c} : $(\mathbf{b}, \mathbf{c}) = 0$.

Przykład 11.1 Sprowadźmy do osi głównych formę

$$f(x_1, x_2, x_3, x_4) = 2x_1x_2 + 2x_1x_3 - 2x_1x_4 - 2x_2x_3 + 2x_2x_4 + 2x_3x_4. \quad (11-53)$$

Symetryczna macierz formy f to

$$\mathcal{S} = \begin{pmatrix} 0 & 1 & 1 & -1 \\ 1 & 0 & -1 & 1 \\ 1 & -1 & 0 & 1 \\ -1 & 1 & 1 & 0 \end{pmatrix}, \quad (11-54)$$

a jej wielomian charakterystyczny to wyznacznik

$$\det(\mathcal{S} - \lambda \mathcal{I}) = \begin{vmatrix} -\lambda & 1 & 1 & -1 \\ 1 & -\lambda & -1 & 1 \\ 1 & -1 & -\lambda & 1 \\ -1 & 1 & 1 & -\lambda \end{vmatrix} = (\lambda - 1)^3(\lambda + 3). \quad (11-55)$$

Macierz \mathcal{S} ma więc potrójny pierwiastek charakterystyczny 1, i (jednokrotny) pierwiastek -3 . Postać kanoniczna formy f to

$$f(y_1, y_2, y_3, y_4) = y_1^2 + y_2^2 + y_3^2 - 3y_4^2. \quad (11-56)$$

Układ 11-47, dla $\lambda_0 = 1$ to:

$$\left. \begin{aligned} -x_1 + x_2 + x_3 - x_4 &= 0, \\ x_1 - x_2 - x_3 + x_4 &= 0, \\ x_1 - x_2 - x_3 + x_4 &= 0, \\ -x_1 + x_2 + x_3 - x_4 &= 0. \end{aligned} \right\} \quad (11-57)$$

Rząd układu jest równy 1; istnieją więc $4 - 1 = 3$ liniowo niezależne rozwiązania. Wybierając je „jak najprostsze” mamy, na przykład ⁷

$$\left. \begin{aligned} \mathbf{b}_1 &= (1, 1, 0, 0), \\ \mathbf{b}_2 &= (1, 0, 1, 0), \\ \mathbf{b}_3 &= (-1, 0, 0, 1), \end{aligned} \right\} \quad (11-58)$$

Dokonujemy ortogonalizacji trzech wektorów

$$\begin{aligned} \mathbf{c}_1 &= \mathbf{b}_1 = (1, 1, 0, 0), \\ \mathbf{c}_2 &= -\frac{1}{2}\mathbf{c}_1 + \mathbf{b}_2 = \left(\frac{1}{2}, -\frac{1}{2}, 1, 0\right), \\ \mathbf{c}_3 &= \frac{1}{2}\mathbf{c}_1 + \frac{1}{3}\mathbf{c}_2 + \mathbf{b}_3 = \left(-\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 1\right). \end{aligned}$$

Z kolei, układ 11-47, dla $\lambda_0 = -3$ to:

$$\left. \begin{aligned} 3x_1 + x_2 + x_3 - x_4 &= 0, \\ x_1 + 3x_2 - x_3 + x_4 &= 0, \\ x_1 - x_2 + 3x_3 + x_4 &= 0, \\ -x_1 + x_2 + x_3 + 3x_4 &= 0. \end{aligned} \right\} \quad (11-59)$$

Rząd układu jest równy 3; istnieje $4 - 3 = 1$ rozwiązanie:

$$\mathbf{c}_4 = (1, -1, -1, 1).$$

Cztery wektory: $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4$ są ortogonalne; dokonując normalizacji uzyskujemy układ czterech wektorów ortonormalnych: Tak więc jawna postać transformacji ortogonalnej, sprowadzającej formę 11-53 do osi głównych to:

$$\begin{aligned} y_1 &= \frac{1}{\sqrt{2}}x_1 + \frac{1}{\sqrt{2}}x_2, \\ y_2 &= \frac{1}{\sqrt{6}}x_1 - \frac{1}{\sqrt{6}}x_2 - \frac{2}{\sqrt{3}}x_3, \\ y_3 &= -\frac{1}{2\sqrt{3}}x_1 + \frac{1}{2\sqrt{3}}x_2 + \frac{1}{2\sqrt{3}}x_3 + \frac{\sqrt{3}}{2}x_4, \\ y_4 &= \frac{1}{2}x_1 - \frac{1}{2}x_2 - \frac{1}{2}x_3 + \frac{1}{2}x_4, \end{aligned}$$

– macierz transformacji ma postać

$$\begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 & 0 \\ 1/\sqrt{6} & -1/\sqrt{6} & -2/\sqrt{3} & 0 \\ -1/2\sqrt{3} & 1/2\sqrt{3} & 1/2\sqrt{3} & \sqrt{3}/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \end{pmatrix} \equiv \mathcal{Q}^{-1}; \quad (11-60)$$

symetryczna macierz \mathcal{S} (11-54) została poddana diagonalizacji:

$$\mathcal{Q}^{-1}\mathcal{S}\mathcal{Q} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix}. \quad (11-61)$$

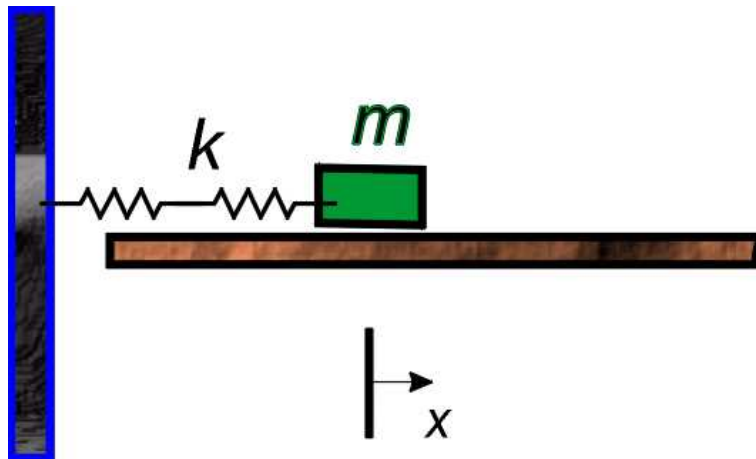
⁷Dlaczego tak? Pomyśl: zależy nam na jak największej ilości zer (na różnych pozycjach). Do spełnienia dowolnego równania układu wystarczą dwa niezerowe x -y (dwa będą zerami); niech będzie to zawsze x_1 i – sukcesywnie – x_2, x_3 i x_4 .

Podkreślmy, że w przypadku wielokrotnej (zdegenerowanej) wartości własnej wybór liniowo niezależnych wektorów własnych jest wieloznaczny; istnieje więc wiele ortogonalnych transformacji, które sprowadzają formę kwadratową f do postaci kanonicznej. Znaleziona powyżej to tylko jedna z wielu możliwych.

Rozdział 12

Trochę fizyki

12.1 Potęga odwzorowania liniowego



Rysunek 12.1: Ruch drgający pod działaniem siły harmoniczej.

Aby uzmysłwić sobie, jak daleko idące skutki wynikają z *liniowości odwzorowań* prześledźmy przykład związany z jednym z pierwszych równań, które rozwiązuje młody adept fizyki. Mam tu na myśli równanie ruchu harmonicznego, to znaczy ruchu zachodzącego pod działaniem siły harmoniczej. Tę ostatnią definiujemy jako siłę wprost proporcjonalną do wychylenia ciała z położenia równowagi i skierowaną „przeciwnie” do wychylenia, a więc starającą się przesunąć ciało do położenia równowagi. Rozważamy przypadek jednowymiarowy: masę m , poruszającą się na idealnie gładkim stole (bez żadnego tarcia), wzdłuż osi x -ów. Źródłem siły harmoniczej jest sprężyna (o zaniedbywalnej masie) o współczynniku sprężystości k ($k > 0$); położenie równowagi (sprężyna ani nie rozciągnięta, ani nie ściśnięta) to punkt $x = 0$. (Por. Rys.12.1.) Definicja siły harmoniczej i drugie prawo dynamiki Newtona dają nam równanie ruchu:

masa \times przyspieszenie = siła działająca, czyli

$$m \frac{d^2 x}{dt^2} = -kx. \quad (12-1)$$

Dzieląc obie strony równania 12-1 przez m i oznaczając

$$\omega^2 \stackrel{\text{def}}{=} \frac{k}{m}, \quad (12-2)$$

mamy

$$\frac{d^2 x}{dt^2} = -\omega^2 x. \quad (12-3)$$

Powyższe równanie jest *różniczkowym równaniem ruchu harmonicznego*.

Nawet jeżeli nie umiemy rozwiązywać takich równań, możemy łatwo sprawdzić¹, że jego rozwiązania to funkcje: $x_1 = \sin \omega t$ i $x_2 = \cos \omega t$. Z teorii równań różniczkowych wynika, że ogólna postać rozwiązania równania 12-1 to

$$x = x(t) = C_1 x_1 + C_2 x_2 = C_1 \sin \omega t + C_2 \cos \omega t, \quad (12-4)$$

gdzie stałe C_1 i C_2 można wyznaczyć, jeżeli dysponujemy pewną dodatkową informacją. Ponieważ do wyznaczenia są *dwie* stałe, to musimy mieć *dwie* niezależne informacje – na przykład położenie i prędkość w chwili $t = 0$ (tzw. warunki początkowe ruchu).

Rozpatrzmy dwa różne warianty takich warunków początkowych. Pierwszy z nich odpowiada sytuacji, gdy w chwili $t = 0$ masa² znajduje się poza położeniem równowagi, w punkcie $x = 1$, a jej prędkość v jest równa zero. (Odciągamy sprężynę przenosząc masę do $x = 1$, a następnie zwalniamy sprężynę.) Oznaczmy taką sytuację symbolem $[x(0), v(0)] \equiv [1, 0]$. Ruch zapoczątkowany takimi warunkami początkowymi będzie odbywał się według równania:

$$x_{[1,0]}(t) = \cos \omega t. \quad (12-5)$$

Łatwo sprawdzić: $x(0) = 1$; prędkość $v(t) = dx/dt = -\omega \sin \omega t$ i $v(0) = 0$.

Drugi wariant to sytuacja, kiedy ciała znajdującemu się w chwili $t = 0$ w położeniu równowagi ($x = 0$) nadajemy pewną jednostkową prędkość; $v(0) = 1$. Taką sytuację możemy oznaczyć jako $[0, 1]$; odpowiadające jej rozwiązanie to

$$x_{[0,1]}(t) = \frac{1}{\omega} \sin \omega t. \quad (12-6)$$

($x(0) = 0$; prędkość $v(t) = dx/dt = \omega(1/\omega) \cos \omega t = \cos \omega t$ i $v(0) = 1$.)

Tak więc dwóm sytuacjom początkowym odpowiadają dwa rozwiązania równania ruchu:

$$\begin{aligned} [1, 0] &\equiv \cos \omega t \\ [0, 1] &\equiv \frac{1}{\omega} \sin \omega t \end{aligned} .$$

W ogólnym przypadku możemy mieć do czynienia, który rozpoczyna się w dowolnym punkcie x_0 i z dowolną prędkością v_0 , a więc w „sytuacji” $[x_0, v_0]$. Ale taka sytuacja to nic innego jak kombinacja liniowa

$$[x_0, v_0] = x_0 \times [1, 0] + v_0 \times [0, 1], \quad (12-7)$$

a jeżeli tak to odpowiadające jej rozwiązanie powinno być analogiczną kombinacją liniową cząstkowych rozwiązań:

$$x(t) = x_0 \times x_{[1,0]} + v_0 \times x_{[0,1]} = x_0 \sin \omega t + \frac{v_0}{\omega} \cos \omega t. \quad (12-8)$$

Powyższy wynik można oczywiście też uzyskać w bardziej standardowy sposób – podstawiając odpowiednie warunki początkowe do równania 12-4 i do równania, jakie otrzymamy różniczkując 12-4 stronami względem czasu, a następnie wyliczając z układu dwóch równań stałe C_1 i C_2 . Ale można też go traktować jako konsekwencję *liniowości samego równania*. Uzyskane rozwiązanie ogólne to liniowe odwzorowanie:

warunki początkowe \rightarrow rozwiązanie opisujące ruch.

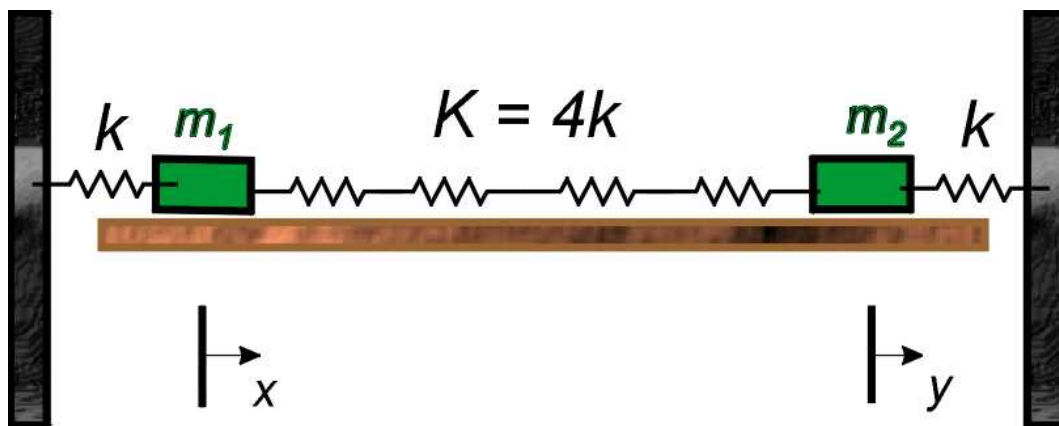
Cząstkowe rozwiązania $x_{[1,0]}$ i $x_{[0,1]}$ łatwo odgadnąć (sprawdzić). Rozważmy teraz ciekawszy przykład, zilustrowany na Rys.12.2. Mamy tu układ dwóch takich samych mas ($m_1 = m_2 = m$) i trzech sprężyn. Dwie skrajne są takie same (mają identyczny współczynnik sprężystości k); sprężyna środkowa jest „mocniejsza” – jej współczynnik sprężystości $K = 4k$. Masy wszystkich sprężyn są zaniedbywalne. Wychylenia mas w stosunku do ich położenia równowagi (wszystkie sprężyny „w spoczynku”) oznaczamy przez x i y . Rozpatrzmy sytuację kiedy $x > 0, y > 0$ i $y > x$. Wówczas środkowa sprężyna jest rozciągnięta i działa na „lewą” masę w prawo, a na „prawą” masę – w lewo. Układ równań Newtona to

$$m \frac{d^2 x}{dt^2} = -kx + K(y - x) = -kx + 4k(y - x) = -5kx + 4ky, \quad (12-9)$$

$$m \frac{d^2 y}{dt^2} = -ky - K(y - x) = -ky - 4k(y - x) = 4kx - 5ky. \quad (12-10)$$

¹Postać rozwiązań problemu ruchu harmonicznego poznałeś zapewne na lekcjach fizyki w szkole średniej.

²Masę traktujemy albo jako bardzo małą – punktową, albo mówiąc o jej położeniu mamy na myśli środek masy przesuwającego się ciała.



Rysunek 12.2: Ruch drgający pod działaniem siły harmoniczej – dwie masy.

W chwili początkowej ($t = 0$) obie masy mogą znajdować się w określonych punktach i mieć pewne prędkości początkowe. Kompletne warunki początkowe to 4 wielkości: $x(0), v_1(0), y(0)$ i $v_2(0)$, gdzie $v_{1,2}$ to prędkości obu mas. Uprościmy na początek analizę i założymy, że $x(0) = y(0) = 0$ (masy są w położeniach równowagi), natomiast $v_1(0) = v_{01}, v_2(0) = v_{02}$. Taki warunek początkowy oznaczmy przez $[0, 0, v_{01}, v_{02}]$.

Położmy $v_{01} = 1$ i $v_{02} = 1$. Obie masy zaczynają więc ruch „w prawo” względem swoich położen równowagi, z takimi samymi prędkościami. Ale w takiej sytuacji łącząca je sprężyna nie podlega deformacji i siły z jakimi oddziałują one na obie masy są równe zeru. Obie masy wykonują, identyczne i niezależne, drgania analogiczne do opisanego powyżej drgania pojedynczej masy. Opis tych drgań to, zgodnie z 12-6

$$x(t) = y(t) = \frac{1}{\omega} \sin \omega t \quad (12-11)$$

i możemy przyjąć, że mamy do czynienia z odwzorowaniem

$$[0, 0, 1, 1] \rightarrow [x(t), y(t)] \equiv \left[\frac{1}{\omega} \sin \omega t, \frac{1}{\omega} \sin \omega t \right]. \quad (12-12)$$

Ogólny przypadek sytuacji początkowej, przy założeniu że $x(0) = y(0) = 0$ to $[0, 0, v_{01}, v_{02}]$. Występują w nim *dwie dowolne* stałe, które nie muszą być równe jedności. Aby zapisać ogólne rozwiązanie dla takiej sytuacji musimy złożyć dwa rozwiązania, uzyskane (wydedukowane) dla *dwóch różnych* sytuacji. Jedno z nich to równanie 12-11. Aby znaleźć drugie, zastanówmy się co stanie się, jeżeli prędkości nadane masom w chwili początkowej będą też takie same i jednostkowe, ale skierowanie przeciwne. Masy startują „naprzeciw sobie” – lewa masa w prawo, a prawa masa w lewo. Wówczas środkowa sprężyna bierze udział w ruchu, ale symetria układu pozwala nam napisać warunek (spełniony podczas trwania ruchu): $x(t) = -y(t)$. Ruch jednej masy jest zwierciadlanym odbiciem ruchu drugiej (lustro umieszczone jest w środku i prostopadle do sprężyny o współczynniku sprężystości K). Przy takim warunku, układ równań 12-9 i 12-10 redukuje się do jednego równania:

$$m \frac{d^2 x}{dt^2} = -9kx, \quad \text{albo} \quad \frac{d^2 x}{dt^2} = -9\omega^2 x, \quad (12-13)$$

którego rozwiązaniem jest $x(t) = -y(t) = \frac{1}{3\omega} \sin 3\omega t$. Funkcja ta spełnia równanie 12-13, $x(0) = y(0) = 0$, a prędkości początkowe

$$\left. \frac{dx}{dt} \right|_{t=0} = 3\omega \frac{1}{3\omega} \cos 3\omega t \Big|_{t=0} = 1; \quad \left. \frac{dy}{dt} \right|_{t=0} = -3\omega \frac{1}{3\omega} \cos 3\omega t \Big|_{t=0} = -1.$$

Tak więc znowu mamy do czynienia z odwzorowaniem

$$[0, 0, 1, -1] \rightarrow [x(t), y(t)] \equiv \left[\frac{1}{3\omega} \sin 3\omega t, -\frac{1}{3\omega} \sin 3\omega t \right]. \quad (12-14)$$

Korzystając z własności proporcjonalności odwzorowań (jeden z elementów liniowości) możemy zapisać równania 12-12 i 12-14 w postaci

$$[0, 0, A, A] \rightarrow \left[\frac{A}{\omega} \sin \omega t, \frac{A}{\omega} \sin \omega t \right] \quad (12-15)$$

$$[0, 0, B, -B] \rightarrow \left[\frac{B}{3\omega} \sin 3\omega t, -\frac{B}{3\omega} \sin 3\omega t \right], \quad (12-16)$$

a korzystając z własności addytywności mamy

$$[0, 0, A + B, A - B] \rightarrow \left[\frac{A}{\omega} \sin \omega t + \frac{B}{3\omega} \sin 3\omega t, \frac{A}{\omega} \sin \omega t - \frac{B}{3\omega} \sin 3\omega t \right]. \quad (12-17)$$

Aby wyrazić powyższe odwzorowanie dla warunków początkowych $[0, 0, v_{01}, v_{02}]$ musimy dokonać prostej operacji, polegającej na wyrażeniu początkowych prędkości v_{01} i v_{02} przez (dowolne) stałe A i B . Kładziemy

$$\begin{aligned} v_{01} &= A + B \\ v_{02} &= A - B \end{aligned}$$

i rozwiązujemy ten układ równań. Dodając i odejmując równania stronami mamy natychmiast

$$A = 0,5(v_{01} + v_{02}) \quad \text{i} \quad B = 0,5(v_{01} - v_{02})$$

i w konsekwencji

$$[0, 0, v_{01}, v_{02}] \rightarrow \left[\begin{aligned} &\frac{0,5(v_{01} + v_{02})}{\omega} \sin \omega t + \frac{0,5(v_{01} - v_{02})}{3\omega} \sin 3\omega t, \\ &\frac{0,5(v_{01} + v_{02})}{\omega} \sin \omega t - \frac{0,5(v_{01} - v_{02})}{3\omega} \sin 3\omega t \end{aligned} \right].$$

Skonstruowaliśmy rozwiązanie układu równań 12-9, opisującego drgania obu mas, dla sytuacji początkowej „typu” $[0, 0, v_{01}, v_{02}]$ (start z położenia równowagi z dowolnymi prędkościami). Rozwiązanie to ma postać

$$x(t)_{[0,0,v_{01},v_{02}]} = \frac{0,5(v_{01} + v_{02})}{\omega} \sin \omega t + \frac{0,5(v_{01} - v_{02})}{3\omega} \sin 3\omega t, \quad (12-18)$$

$$y(t)_{[0,0,v_{01},v_{02}]} = \frac{0,5(v_{01} + v_{02})}{\omega} \sin \omega t - \frac{0,5(v_{01} - v_{02})}{3\omega} \sin 3\omega t. \quad (12-19)$$

Analogicznie powinniśmy rozwiązać zagadnienie drgań dla warunków początkowych typu $[x_0, y_0, 0, 0]$ (start z pewnych wychyleń początkowych z prędkościami równymi zeru). Rozważania są oparte na tym samym pomysle: wyeliminowaniu wpływu środkowej sprężyny (wychylenia takie same i w tym samym kierunku) i drganiu „zwierciadlanym” (takie same wychylenia w przeciwnych kierunkach). Otrzymamy wówczas zespół równań jak 12-18–12-19, opisujących funkcje $x(t)$ i $y(t)$, przy czym występować w nich będą kosinusy, a nie sinusy:

$$x(t)_{[x_{01},x_{02},0,0]} = 0,5(x_{01} + x_{02}) \cos \omega t + 0,5(x_{01} - x_{02}) \cos 3\omega t, \quad (12-20)$$

$$y(t)_{[x_{01},x_{02},0,0]} = 0,5(x_{01} + x_{02}) \cos \omega t - 0,5(x_{01} - x_{02}) \cos 3\omega t. \quad (12-21)$$

W końcu pozostaje kwestia sformułowania rozwiązania w pełni ogólnego, tzn. odpowiadającego warunkom początkowym typu $[x_0, y_0, v_{01}, v_{02}]$. To już jednak właściwie umiemy zrobić – poświęcona była temu pierwsza część tego podrozdziału, a konkretnie wzory 12-7 i 12-8. Aby określić $x(t)$ i $y(t)$ w najbardziej ogólnym przypadku, musimy dodać prawe strony równań 12-18 i 12-20, oraz 12-19 i 12-21.

Podane przykłady uzmysławiają nam potęgę definicji liniowości i uczą nas rozwiązywać równania fizyki w sposób wykorzystujący chyba bardziej pewien „spryt” i inteligencję, niż znajomość formalnych reguł. A także – dla fizyka bardzo jest to istotne – uzmysławiają nam rolę sytuacji początkowej.

12.2 Diagonalizacja

12.2.1 Wektory własne macierzy

Zagadnienie „własne”, a więc znajdowanie wartości i wektorów własnych można bez przesady uznać za fundamentalne zastosowanie algebry liniowej w problemach fizycznych. Aby zilustrować wstępnie co oznacza znalezienie (i wprowadzenie do opisu zjawiska) wektorów własnych powróćmy raz jeszcze do problemu dwóch

mas i trzech sprężyn. Układ równań, opisujący wychylenia pierwszej (x) i drugiej masy (y) z położeń równowagi ma postać (przy uwzględnieniu $K = 4k$, oznaczeniu $\frac{k}{m} \equiv \omega^2$ i redukcji – por. 12-9):

$$\left. \begin{aligned} \frac{d^2x}{dt^2} &= -5\omega^2x + 4\omega^2y, \\ \frac{d^2y}{dt^2} &= 4\omega^2x - 5\omega^2y. \end{aligned} \right\} \quad (12-22)$$

Układ ten rozwiązaliśmy rozpatrując dwa szczególne przypadki sytuacji początkowych: pierwszy – kiedy masy zaczynają ruch z położeń równowagi z takimi samymi (jednostkowymi) prędkościami: $v_{01} = v_{02} = 1$, drugi – kiedy prędkości mas są skierowane przeciwnie $v_{01} = -v_{02} = 1$. Te „sytuacje początkowe” można zapisać w postaci dwóch wektorów:

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ oraz } \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \quad (12-23)$$

Dla tak wybranych wektorów, reprezentujących warunki początkowe, nasz układ równań można było „uproszczyć” korzystając z argumentów fizycznych, wynikających z analizy charakteru ruchu. Ale przecież algebra musi też istnieć niezależnie od fizyki, a więc takie „uproszczenie” układu musi mieć pewne formalne podstawy. Zauważmy, że układ 12-22, zapisany w postaci macierzowej to:

$$\begin{pmatrix} \frac{d^2x}{dt^2} \\ \frac{d^2y}{dt^2} \end{pmatrix} = \begin{pmatrix} -5\omega^2 & 4\omega^2 \\ 4\omega^2 & -5\omega^2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \omega^2 \begin{pmatrix} -5 & 4 \\ 4 & -5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \quad (12-24)$$

Przyspieszenia mas (lewa strona) i ich położenia (jednokolumnowa macierz po prawej stronie) są powiązane macierzą \mathcal{M} :

$$\begin{pmatrix} -5\omega^2 & 4\omega^2 \\ 4\omega^2 & -5\omega^2 \end{pmatrix}. \quad (12-25)$$

Zobaczmy, czy ta macierz pozostaje w jakiejś specjalnej relacji względem wektorów 12-23. Zachodzi:

$$\begin{pmatrix} -5\omega^2 & 4\omega^2 \\ 4\omega^2 & -5\omega^2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -\omega^2 \\ -\omega^2 \end{pmatrix} = -\omega^2 \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad (12-26)$$

oraz

$$\begin{pmatrix} -5\omega^2 & 4\omega^2 \\ 4\omega^2 & -5\omega^2 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -9\omega^2 \\ 9\omega^2 \end{pmatrix} = -9\omega^2 \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \quad (12-27)$$

Wektory 12-23 są wektorami własnymi macierzy \mathcal{M} .

Co z tego wynika? Formułując nasze równania ruchu (12-22) wybraliśmy wektor położeń w postaci

$$\begin{pmatrix} x \\ y \end{pmatrix}, \quad (12-28)$$

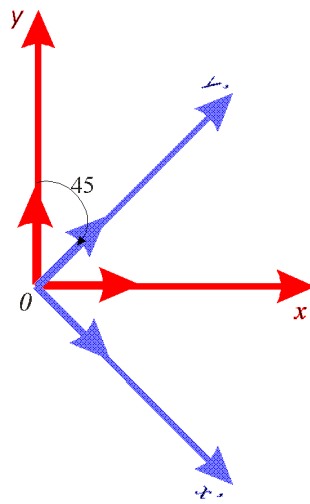
a więc wektor, którego współrzędne wyrażamy w „normalnym”, kartezjańskim układzie współrzędnych. Wersory osi w tym układzie to

$$\hat{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ oraz } \hat{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (12-29)$$

Zauważmy jednak, że wektory 12-23 też mogą określać dwie prostopadłe osie! (Ich iloczyn skalarny jest równy zeru.) Będą to osie – por. Rys.12.3 – układu, który powstaje z układu zdefiniowanego wersorami 12-29 na skutek obrotu o kąt -45° . Aby wektory 12-23 były *wersorami* należy je tylko odpowiednio znormalizować – tak aby ich długość wynosiła 1:

$$\hat{e}'_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ oraz } \hat{e}'_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \quad (12-30)$$

Jeżeli ruch w układzie $0xy$ (por. Rys.12.3) opisany jest parą współrzędnych (x, y) , to – jak wiemy z punktu



Rysunek 12.3: Nowy układ współrzędnych – osie $0x$ i $0y$ zostały obrócone (w kierunku ujemnym) o 45° .

1.2 – w układzie $0x'y'$ (por. rys. 12.3) nowe współrzędne (x', y') mają postać

$$\begin{aligned} x' &= x \cos(-45^\circ) + y \sin(-45^\circ) = \frac{\sqrt{2}}{2}(x - y), \\ y' &= -x \sin(-45^\circ) + y \cos(-45^\circ) = \frac{\sqrt{2}}{2}(x + y). \end{aligned} \quad (12-31)$$

Układ 12-31 możemy rozwiązać ze względu na x i y :

$$\begin{aligned} x &= \frac{\sqrt{2}}{2}(x' + y'), \\ y &= \frac{\sqrt{2}}{2}(x' - y') \end{aligned} \quad (12-32)$$

i podstawić do 12-22. Otrzymujemy

$$\begin{aligned} \frac{d^2 x'}{dt^2} &= -9\omega^2 x', \\ \frac{d^2 y'}{dt^2} &= -\omega^2 y', \end{aligned} \quad (12-33)$$

w pełnej zgodzie z równaniami 12-3 i 12-13. Są to równania o wiele prostsze niż wyjściowe równania 12-22. Tam mamy do czynienia z *układem równań*, w którym obie zmienne występują w obu równaniach; tutaj dostaliśmy dwa zupełnie analogiczne równania, w których występuje pojedyncza zmienna i które rozwiązujemy błyskawicznie. A wszystko dlatego, że wybraliśmy do opisu właściwy układ współrzędnych. Układ, wywodzący się z wektorów własnych macierzy \mathcal{M} , w której tkwi cała informacja o parametrach fizycznych sprężyn (współczynniki sprężystości) i mas (ich wielkości).

12.3 Moment bezwładności ciała sztywnego

Każdy obiekt fizyczny – ciało sztywne – to obiekt o skończonych wymiarach. Określenie sztywne oznacza, że struktura ciała jest „sztywna”, tzn. pod działaniem sił ciało nie deformuje się i odległości pomiędzy punktami ciała pozostają stałe. Z kolei, wprowadzając pojęcie *momentu pędu* ciała, w ruchu obrotowym względem pewnej osi, \mathbf{L} , pierwszym wzorem z jakim spotykamy się to

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} = \mathbf{r} \times m\mathbf{v} = m\mathbf{r} \times \mathbf{v}, \quad (12-34)$$

gdzie \mathbf{v} to prędkość *punktu materialnego* o masie m , a \mathbf{r} jego odległość od osi obrotu. Punkt materialny to pewna abstrakcja; dla ciała sztywnego musimy powyższą definicję zmodyfikować. Robimy to, traktując ciało jako zbiór (sumę) n bardzo małych mas kwazi-punktowych m_i , z których każda posiada pewną prędkość \mathbf{v}_i

i znajduje się w pewnej odległości \mathbf{r}_i od osi obrotu. Jeżeli ciało wykonuje ruch obrotowy wokół pewnej osi, której położenie w przestrzeni jest stałe, to prędkość liniowa każdego „punktu” i

$$\mathbf{v}_i = \boldsymbol{\omega} \times \mathbf{r}_i, \quad (12-35)$$

gdzie $\boldsymbol{\omega}$ jest prędkością kątową obrotu. (Wektor $\boldsymbol{\omega}$ leży na osi obrotu, a jego zwrot związany jest z kierunkiem obrotu regułą śruby prawej). W takim razie

$$\mathbf{L} = \sum_{i=1}^n m_i \mathbf{r}_i \times \mathbf{v}_i = \sum_{i=1}^n m_i [\mathbf{r}_i \times (\boldsymbol{\omega} \times \mathbf{r}_i)] = \sum_{i=1}^n m_i [\boldsymbol{\omega} r_i^2 - \mathbf{r}_i (\boldsymbol{\omega} \cdot \mathbf{r}_i)]. \quad (12-36)$$

(W wyprowadzeniu powyższego wzoru wykorzystaliśmy związek

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B}), \quad (12-37)$$

dla podwójnego iloczynu wektorowego – por. punkt 4.4.) Wzór 12-36 warto rozpisać dla współrzędnych x -, y - i z -owej wektora \mathbf{L}^3 :

$$\left. \begin{aligned} L_x &= \sum_{i=1}^n m_i [\omega_x r_i^2 - x_i (\boldsymbol{\omega} \cdot \mathbf{r}_i)] \\ L_y &= \sum_{i=1}^n m_i [\omega_y r_i^2 - y_i (\boldsymbol{\omega} \cdot \mathbf{r}_i)] \\ L_z &= \sum_{i=1}^n m_i [\omega_z r_i^2 - z_i (\boldsymbol{\omega} \cdot \mathbf{r}_i)]. \end{aligned} \right\} \quad (12-38)$$

($\boldsymbol{\omega} \equiv (\omega_x, \omega_y, \omega_z)$; $\mathbf{r}_i \equiv (x_i, y_i, z_i)$.) Utwórzmy teraz macierz o wymiarach 3×3 ,

$$\mathcal{B} \equiv \begin{pmatrix} B_{xx} & B_{xy} & B_{xz} \\ B_{yx} & B_{yy} & B_{yz} \\ B_{zx} & B_{zy} & B_{zz} \end{pmatrix} \quad (12-39)$$

której wyrazy to

$$\left. \begin{aligned} B_{xx} &= \sum_{i=1}^n m_i [y_i^2 + z_i^2], \\ B_{yy} &= \sum_{i=1}^n m_i [x_i^2 + z_i^2], \\ B_{zz} &= \sum_{i=1}^n m_i [x_i^2 + y_i^2], \\ B_{xy} = B_{yx} &= - \sum_{i=1}^n m_i x_i y_i, \\ B_{yz} = B_{zy} &= - \sum_{i=1}^n m_i y_i z_i, \\ B_{zx} = B_{xz} &= - \sum_{i=1}^n m_i z_i x_i. \end{aligned} \right\} \quad (12-40)$$

Tak zdefiniowana macierz \mathcal{B} to *macierz momentu bezwładności* (tensor momentu bezwładności⁴). Macierz jest symetryczna, a elementy na przekątnej głównej to rzeczywiście wartości momentów bezwładności naszego ciała, obliczanych względem osi $0x$ (B_{xx}), $0y$ (B_{yy}) i $0z$ (B_{zz}), tak zwane *główne momenty bezwładności*. Rzeczywiście, dla dowolnej osi obrotu definicja moment bezwładności bryły sztywnej to

$$B = \sum_{i=1}^n m_i d_i^2, \quad (12-41)$$

³Wzory będą bardziej przejrzyste jeżeli zamiast x_i ; $i = 1, 2, 3$ użyjemy x , y i z .

⁴Formalna definicja tensora pojawi się w następnym rozdziale

gdzie d_i to odległość i -tej masy od osi obrotu. Dla – na przykład – osi $0x$

$$d_i^2 = y_i^2 + z_i^2.$$

Wyraży leżące poza przekątną główna (3 symetryczne pary): B_{xy} , B_{yz} i B_{zx} to tak zwane *momenty dewiacji*. Podobnie jak i główne momenty, momenty dewiacji opisują rozkład masy w ciele.

Równanie 12-38 w zapisie macierzowym to

$$\begin{pmatrix} B_{xx} & B_{xy} & B_{xz} \\ B_{yx} & B_{yy} & B_{yz} \\ B_{zx} & B_{zy} & B_{zz} \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} = \begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix}. \quad (12-42)$$

Formalnie otrzymujemy równość wektora momentu pędu oraz wyniku działania macierzy momentu bezwładności na wektor prędkości kątowej. (Iloczyn macierzy $n \times n$ i jednokolumnowej macierzy o n wierszach to także macierz jednokolumnowa – wektor o n współrzędnych.) Stojąca z lewej strony iloczynu macierz \mathcal{B} – reprezentacja *tensora momentu bezwładności* jest *operatorem* w przestrzeni wektorowej; działając na wektor (element tej przestrzeni) tworzy z niego inny wektor. W ogólnym przypadku, stary i nowy wektor są „zupełnie” różne – tzn., różnią się długościami i kierunkami. Ale macierz \mathcal{B} jest symetryczna, a więc istnieje taki układ współrzędnych, *układ osi głównych ciała sztywnego*, w którym macierz 12-39 ma postać diagonalną, a skomplikowane nieco równanie 12-42 przybiera postać znacznie prostszą

$$\begin{pmatrix} B_{xx} & 0 & 0 \\ 0 & B_{yy} & \\ 0 & 0 & B_{zz} \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} = \begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix}, \quad (12-43)$$

albo

$$\begin{aligned} L_x &= B_{xx}\omega_x, \\ L_y &= B_{yy}\omega_y, \\ L_z &= B_{zz}\omega_z. \end{aligned} \quad (12-44)$$

W takim właśnie, wyróżnionym układzie współrzędnych mamy

$$\begin{aligned} &(\text{moment pędu})_{\text{składowa wzdłuż osi } k} \\ &= (\text{moment bezwładności})_{\text{względem osi } k} \\ &\times (\text{prędkość kątowa})_{\text{składowa wzdłuż osi } k}. \end{aligned}$$

Dla ciał wykazujących symetrię kształtu, osie główne pokrywają się z osiami symetrii ciała. Łatwo to zrozumieć. Jeżeli ciało jest symetryczne względem, na przykład, osi $0z$, to znaczy że „po jednej i po drugiej stronie” tej osi znajduje się jednakowa ilość takich samych quasi-punktowych mas. W wyrażeniu typu

$$\sum_{i=1}^n m_i x_i y_i$$

dodatnie i ujemne współrzędne wektora x_i i y_i – odległości od osi $0z$ – występują więc z taką samą „częstością” (jest tyle samo x_i większych od zera, jak i mniejszych od zera) i taką samą wagą (masą m_i). Jeżeli tak, to powyższe wyrażenie *musi* być równe zeru – zgodnie z 12-43.

Dla ciał o wysokim stopniu symetrii (kula, sześcian) momenty bezwładności liczone względem trzech głównych osi są takie same: $B_{xx} = B_{yy} = B_{zz} = B$. (Zapewne pamiętasz wzór na moment bezwładności kuli, względem *dowolnej* osi, przechodzącej przez jej środek:

$$B = \frac{2}{5}MR^2,$$

gdzie M to masa kuli, a R – jej promień.) W takich przypadkach równanie 12-43 znakomicie się upraszcza:

$$\begin{pmatrix} B & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & B \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} = \begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix}, \quad (12-45)$$

albo – już całkiem prosto –

$$\mathbf{L} = B\boldsymbol{\omega}. \quad (12-46)$$

Moment pędu jest *proporcjonalny* do prędkości kątowej, a współczynnikiem tej proporcjonalności jest miara bezwładności ciała, bezwładności którą musimy pokonać wprowadzając ciało w *ruch obrotowy*.

Dopiero w takim prostym (najprostszym!) przypadku możemy interpretować ten wzór jako analogiczny do związku z mechaniki *ruchu postępowego*

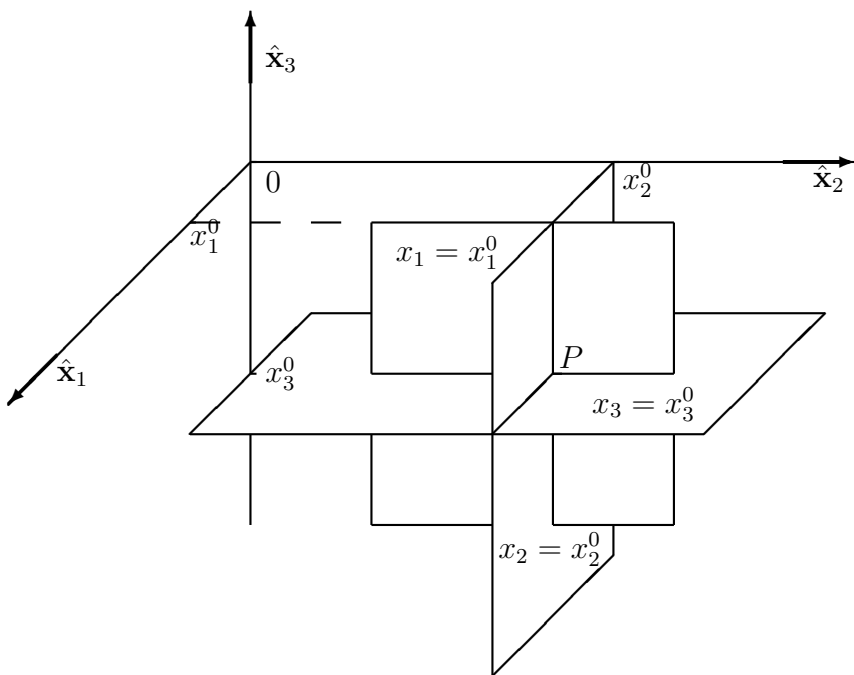
$$\mathbf{p} = m\mathbf{v}, \quad (12-47)$$

którym analogon momentu pędu – pęd \mathbf{p} – jest przedstawiony jako iloczyn masy (miary bezwładności w ruchu postępowym) i jej prędkości liniowej \mathbf{v} . W przypadku ogólnym obowiązuje równanie 12-42.

12.4 Przestrzenie funkcyjne i ich ortogonalne bazy

12.4.1 Układy współrzędnych krzywoliniowych

Aby określić położenie punktu w „zwykłej”, trójwymiarowej przestrzeni posługujemy się zazwyczaj układem trzech wzajemnie prostopadłych osi liczbowych, przecinających się w jednym punkcie — początku układu. Rzuty określonego punktu przestrzeni, P , na poszczególne osie – trzy liczby rzeczywiste zawarte w przedziale $(-\infty, +\infty)$ – stanowią wartości współrzędnych punktu (por. rozdział czwarty). Zwykle liczby te oznaczamy symbolami x, y, z , albo – tak jak tutaj – x_1, x_2, x_3 .



Rysunek 12.4: Punkt $P(x_1^0, x_2^0, x_3^0)$ jako punkt przecięcia trzech prostopadłych płaszczyzn.

Wersory osi układu są ortonormalne

$$\hat{\mathbf{x}}_i \cdot \hat{\mathbf{x}}_k = \delta_{ik}, \quad i, k = 1, 2, 3. \quad (12-48)$$

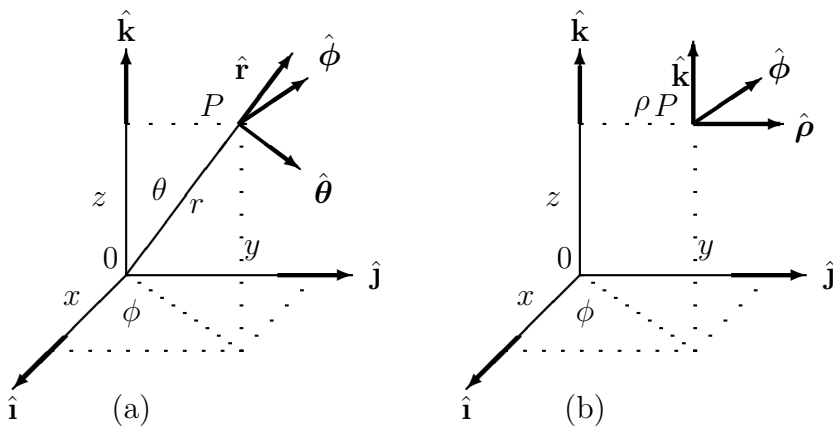
Każdy wersor $\hat{\mathbf{x}}_i$ jest prostopadły do odpowiedniej rodziny powierzchni: $x_i = \text{const}_i$; $i = 1, 2, 3$. Punkt przestrzeni, P , o współrzędnych (x_1^0, x_2^0, x_3^0) to *punkt przecięcia trzech płaszczyzn* o równaniach

$$x_1 = x_1^0; \quad x_2 = x_2^0; \quad x_3 = x_3^0.$$

(Por. Rys.12.4). Układ kartezjański jest układem trzech współrzędnych, których charakter jest identyczny i z których każda występuje w opisie położenia punktu P z identyczną „wagą”. W takim układzie pojawiają się w sposób naturalny takie obszary jak „lewy”, „prawy”, „dolny” i „górny”. Na płaszczyźnie wszystkie cztery ćwiartki układu osi kartezjańskich są „takie same”; w przestrzeni trójwymiarowej „takie same” jest osiem części, zawartych pomiędzy odpowiednimi półpłaszczyznami (Por. Rys.12.4). W niektórych sytuacjach takie „równouprawnienie” dwójki x_1, x_2 , lub trójki x_1, x_2, x_3 kłóci się z oczywistymi symetrami występującymi w naszej przestrzeni albo po prostu jest mało praktyczne. Jeżeli opisujemy np. pole wytworzone przez dipol elektryczny, to z góry wiemy, że będzie ono posiadało symetrię osiową (osią symetrii jest oś dipola) i w takim razie do opisu pola powinny wystarczyć dwie współrzędne. Zamiast dwójki (np.) x_1, x_2 bardziej logiczne będą jednak współrzędne r i θ — odległość od dipola i kąt (mierzony względem osi dipola), pod którym widać punkt pola. Podobnie jeżeli opisywane pole będzie np. zależało tylko od odległości punktu pola od wytwarzającego pole źródła, to jedyną potrzebną do opisu współrzędną będzie ta odległość, a wówczas układ współrzędnych kartezjańskich jest ewidentnie mało praktyczny. Chyba, że ktoś *lubi* wyrażenia typu: $\sqrt{x_1^2 + x_2^2 + x_3^2}$. Układy współrzędnych krzywoliniowych służą właśnie z jednej strony do uwypuklenia, a z drugiej — do wykorzystywania — pewnych „naturalnych” symetrii, a więc i naturalnych współrzędnych, pojawiających się w opisie problemu. Każdy z takich układów określa położenie dowolnego punktu w przestrzeni jako punktu przecięcia trzech *powierzchni*: $q_i = const_i$; $i = 1, 2, 3$. Powierzchnie te już nie muszą (choć mogą) być płaszczyznami. Podobnie powierzchnie (a dokładnie: płaszczyzny styczne do powierzchni w punkcie ich przecięcia) nie przynależne do różnych rodzin nie muszą być wzajemnie prostopadłe, jednak zazwyczaj trzy rodziny: q_1, q_2, q_3 tworzą trójkę wzajemnie ortogonalną. Wersory \hat{q}_1, \hat{q}_2 i \hat{q}_3 — prostopadłe do odpowiednich powierzchni q_i — spełniają więc relację analogiczną do (12-48)

$$\hat{q}_i \cdot \hat{q}_k = \delta_{ik}, \quad i, k = 1, 2, 3. \quad (12-49)$$

Jakie „ortogonalne” rodziny powierzchni spotykamy w przestrzeni 3-wymiarowej? Właściwie tylko dwie zasługują na wymienienie w tym punkcie wykładu: tzw. układ współrzędnych biegunowych sferycznych i układ współrzędnych cylindrycznych.



Rysunek 12.5: Układy współrzędnych: sferycznych (a) i cylindrycznych (b).

Układ współrzędnych sferycznych biegunowych to układ, w którym trzy ortogonalne rodziny powierzchni stanowią:

1. $q_1 = r$ — koncentryczne (o środku w punkcie 0) sfery o promieniu r

$$r = \sqrt{x^2 + y^2 + z^2} = constans \quad (12-50)$$

2. $q_2 = \theta$ — stożki o wierzchołku w punkcie 0, dla których oś $0z$ jest osią symetrii; kąt rozwarcia stożka wynosi 2θ , przy czym

$$\cos \theta = \frac{z}{\sqrt{x^2 + y^2 + z^2}} = \text{constans} \quad (12-51)$$

3. $q_3 = \phi$ — pęk półpłaszczyzn przechodzących przez oś $0z$, tworzących kąt ϕ z płaszczyzną $0xz$;

$$\text{tg } \phi = \frac{y}{x} = \text{constans} \quad (12-52)$$

Natomiast układ współrzędnych cylindrycznych to układ w którym trzy ortogonalne rodziny powierzchni stanowią:

1. $q_1 = \rho$ — cylindry (walce) o wspólnej osi symetrii $0z$ i o promieniu ρ :

$$\rho = \sqrt{x^2 + y^2} = \text{constans} \quad (12-53)$$

2. $q_2 = \phi$ — półpłaszczyzny przechodzące przez oś $0z$, tworzące kąt ϕ z płaszczyzną $0xz$;

$$\text{tg } \phi = \frac{y}{x} = \text{constans} \quad (12-54)$$

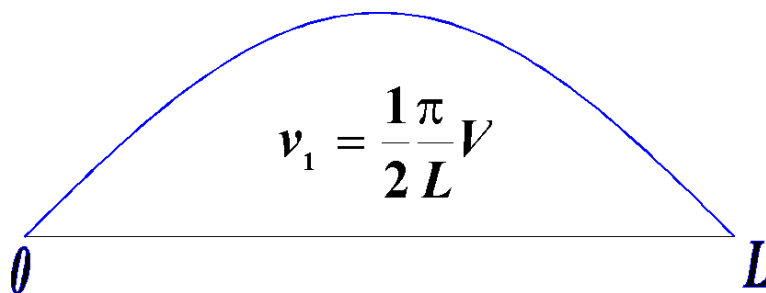
3. $q_3 = z$ — płaszczyzny prostopadłe do osi $0z$

$$z = \text{constans}. \quad (12-55)$$

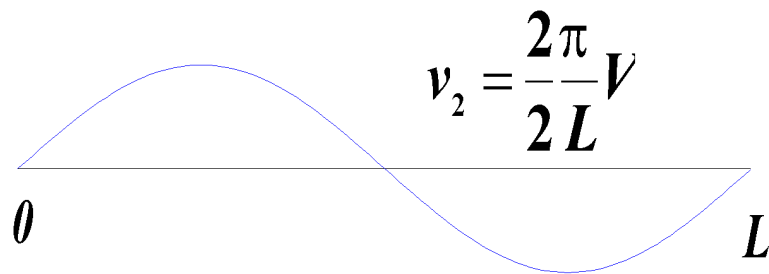
12.4.2 Przestrzenie funkcyjne – pierwsze kroki

Jeżeli przyjdzie nam porównywać (znajdować podobieństwa, różnice) dwa wektory to „wizualnie” możemy tego dokonać na płaszczyźnie, ewentualnie w przestrzeni 3-wymiarowej. Dla wektorów w przestrzeni n -wymiarowej ($n > 3$) pozostaje możliwość porównywania wektorów przez analizę ich „wizytówek” – tzn. zbiorów współrzędnych każdego z nich. Oczywiście współrzędne te muszą być wyrażone względem tej samej bazy w przestrzeni wektorowej. Mówiąc obrazowo: baza przestrzeni wektorowej wprowadza pewien język, w którym opisujemy charakterystyki (własności) obiektów tej przestrzeni.

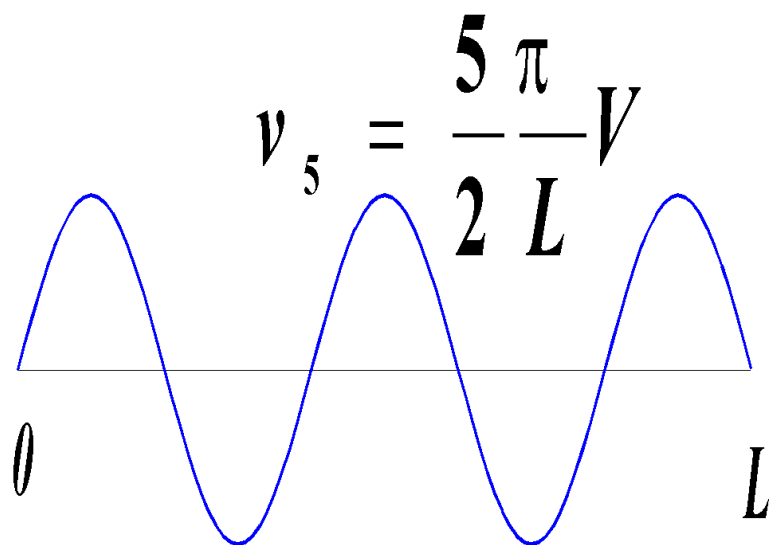
W fizyce i technice mamy do czynienia z problemami, który opis dokonuje się w języku *funkcji*, czyli zależności pewnych wielkości fizycznych (przesunięcie, prędkość, itp.) od innych zmiennych jak położenie (współrzędne przestrzenne), czas, itp. Dla ustalenia uwagi spróbujmy prześledzić stosunkowo prosty problem fizyczny: problem fal stojących, powstających w strunie o długości L , rozpiętej wzdłuż osi $0x$. Oba końce struny są unieruchomione. Strunę pobudzamy do drgań odciągając ją od położenia równowagi (wzdłuż osi $0x$) i puszczając, ewentualnie uderzając (szarpiąc) strunę, a więc nadając jej punktom pewne prędkości w chwili początkowej (ewentualnie stosując oba sposoby jednocześnie). Propagacja drgań wzdłuż struny to fala. Rozchodzące się fale, po odbiciu od zamocowanych końców struny interferują z kolejnymi falami „przychodzącymi” – w efekcie powstaje fala stojąca. Musi to być fala, dla której końce struny są *węzłami*, a więc prędkość wychyleń (drgań poprzecznych) musi tam być równa zero. Innymi słowy mogą być generowane fale o takiej długości, że na całej długości struny może się ułożyć całkowita wielokrotność połówek długości fali. Na przykład $L = \frac{1}{2}\lambda_1$ (Rys.12.6), albo $L = 2\frac{1}{2}\lambda_2$ (Rys.12.7), albo – piąta z rzędu – $L = 5\frac{1}{2}\lambda_5$ (Rys.12.8). Z elementarnej akustyki



Rysunek 12.6: Największa długość fali stojącej powstającej w strunie o długości L .



Rysunek 12.7: Druga największa długość fali stojącej powstającej w strunie o długości L .



Rysunek 12.8: Piąta największa długość fali stojącej powstającej w strunie o długości L .

wiemy, że częstość drgań struny ν , prędkość rozchodzenia się fali c i długość fali λ powiązane są zależnością:

prędkość = częstość \times długość fali, albo

$$v = \nu \times \lambda, \quad \text{czyli } \nu = \frac{v}{\lambda}. \quad (12-56)$$

Przypadek z Rys.12.6 odpowiada więc generowaniu się w strunie o częstości

$$\nu_1 = \frac{v}{\lambda_1} = \frac{v}{2L};$$

przypadek z Rys.12.7 – częstości

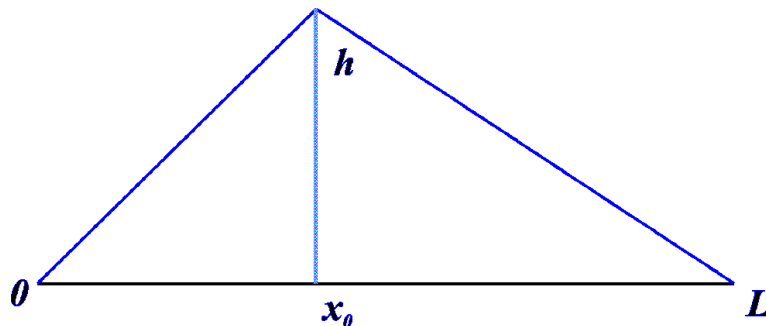
$$\nu_2 = \frac{v}{\lambda_2} = \frac{v}{L} = 2\nu_1;$$

a przypadek z Rys.12.8 – częstości

$$\nu_5 = \frac{v}{\lambda_5} = \frac{5c}{2L} = 5\nu_1.$$

Zbiór wszystkich możliwych częstości jest zbiorem nieskończonym (przeliczalnym); ton (dźwięk) struny to mieszanina tych częstości: podstawowej (ν_1) i tzw. wyższych harmonicznych (ν_2, ν_3, \dots).

Rozważmy przypadek drgań, które powstają tylko na skutek „odciągnięcia” struny w chwili początkowej ($t = 0$) od jej położenia równowagi. Np. strunę odciągamy, nadając jej kształt trójkąta (Rys.12.9) przy



Rysunek 12.9: Pobudzenie struny do drgań poprzez odciągnięcie jej do kształtu trójkątnego.

czym punkty struny nie mają żadnych prędkości początkowych w chwili $t = 0$. Można stosunkowo łatwo wykazać (w momencie kiedy poznamy już równania opisujące rozchodzenia się fal i bardzo podstawowe metody rozwiązywania tych równań), że wychylenie poprzeczne (prostopadłe do osi $0x$) punktów struny to funkcja czasu t i położenia punktu x – $\psi(x, t)$ –

$$\psi(x, t) = \sum_{n=1}^{\infty} C_n \sin \frac{n\pi x}{L} \cos \frac{n\pi vt}{L}. \quad (12-57)$$

Nieskończona suma to odbicie nieskończoności zbioru możliwych częstości drgań; współczynniki (nieznane!) C_n określają wkład tych częstości⁵. Zauważ, że *każdy* wyraz sumy zeruje się dla $x = 0$ i $x = L$ (por. też podrozdział 11). Jak wyznaczyć (znaleźć) nieznane C_n ? Otóż funkcję z Rys.12.9 możemy określić jako

$$\psi(x, t = 0) \equiv f(x) = \begin{cases} h \frac{x}{x_0} & 0 \leq x \leq x_0 \\ h \frac{L-x}{L-x_0} & x_0 \leq x \leq L. \end{cases} \quad (12-58)$$

Z drugiej strony, podstawiając do 12-57 wartość $t = 0$ mamy

$$\psi(x, 0) = \sum_{n=1}^{\infty} C_n \sin \frac{n\pi x}{L} = f(x). \quad (12-59)$$

Wzory 12-58 i 12-59 reprezentują tę samą funkcję! Czyli

$$f(x) = \begin{cases} h \frac{x}{x_0} & 0 \leq x \leq x_0 \\ h \frac{L-x}{L-x_0} & x_0 \leq x \leq L. \end{cases} = \sum_{n=1}^{\infty} C_n \sin \frac{n\pi x}{L}. \quad (12-60)$$

Taki zapis nasuwa nam nieodparte skojarzenie z zapisem wektora – obiektu n -wymiarowej przestrzeni wektorowej \mathbb{R}^n w postaci

$$\mathbf{V} = \sum_{i=1}^n V_i \hat{\boldsymbol{\xi}}_i, \quad (12-61)$$

gdzie wielkości V_i to współrzędne \mathbf{V} w bazie $\{\hat{\boldsymbol{\xi}}_i\}$. Jeżeli tak to w nieskończonym zbiorze funkcji

$$\sin \frac{n\pi x}{L}; \quad n = 1, 2, \dots \quad (12-62)$$

powinniśmy dopatrywać się bazy przestrzeni wektorowej o nieskończonym wymiarze — *bazy przestrzeni funkcyjnej*. Poprawna definicja przestrzeni funkcyjnych jest oczywiście nieco obszerniejsza, ale dla naszych potrzeb wystarczy to – intuicyjne – rozszerzenie wymiaru do nieskończoności i – patrz niżej – wprowadzenie definicji iloczynu skalarnego. Obiektami przestrzeni funkcyjnej są funkcje – takie jak nasza $f(x)$.

⁵Problem struny omówiony szczegółowo znajdziesz w trzecim rozdziale MMF II i III .

Intuicja i tym razem nas nie zawodzi. Tak naprawdę to funkcje 12-62 stanowią połowę bazy; gdybyśmy dopuścili generacje drgań poprzez przekazywanie punktom struny pewnych prędkości w chwili $t = 0$ to musielibyśmy wprowadzić jeszcze funkcje

$$\cos \frac{n\pi x}{L}; \quad n = 0, 1, 2, \dots \quad (12-63)$$

Dopiero zespół funkcji 12-62 i 12-63 to pełna (zupełna) baza. Co więcej jest to baza *ortogonalna*. W przestrzeni funkcyjnej aby mówić o ortogonalności musimy określić definicję iloczynu skalarnego dwóch funkcji. Nie trudno chyba przyjąć, że skutek rozszerzenia wymiaru przestrzeni do nieskończoności to zamiana sumy (w iloczynie skalarnym dwóch wektorów – obiektów przestrzeni wektorowej) całąką dwóch funkcji – obiektów przestrzeni wektorowej:

$$\int_0^L f(x)g(x)dx, \quad (12-64)$$

przy czym całkujemy po pełnym zakresie zmiennej x , z którym mamy do czynienia w naszym problemie. Jeżeli umiesz – choć trochę – całkować to możesz sprawdzić

$$\int_0^L \sin \frac{m\pi x}{L} \sin \frac{n\pi x}{L} dx = \mathcal{S}_n \delta_{nm} \quad (12-65)$$

$$\int_0^L \cos \frac{m\pi x}{L} \cos \frac{n\pi x}{L} dx = \mathcal{S}_n \delta_{nm}, \quad (12-66)$$

gdzie stała \mathcal{S}_n jest równa $L/2$ (niezależnie od wartości n).

Tak więc *ortonormalną* bazę w przestrzeni funkcyjnej, dla $0 \leq x \leq L$ będą tworzyć funkcje

$$\sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L}; \quad n = 1, 2, \dots; \quad \sqrt{\frac{2}{L}} \cos \frac{n\pi x}{L}; \quad n = 1, 2, \dots; \quad \text{oraz} \quad \sqrt{\frac{1}{L}}. \quad (12-67)$$

(Ostatnia funkcja to odpowiednio unormowany kosinus dla $n = 0$.) Jeżeli zamiast przedziału $[0, L]$ przyjąć, bardziej kojarzący się z funkcjami trygonometrycznymi, przedział $[0, 2\pi]$, lub $[-\pi, +\pi]$ to odpowiednikiem 12-67 będą funkcje

$$\sqrt{\frac{1}{\pi}} \sin nx; \quad n = 1, 2, \dots; \quad \sqrt{\frac{1}{\pi}} \cos nx; \quad n = 1, 2, \dots; \quad \text{oraz} \quad \sqrt{\frac{1}{2\pi}}. \quad (12-68)$$

Powyższą bazę nazywamy tradycyjnie „bazą fourierowską”, a szeregi typu 12-60 to szeregi Fouriera (funkcji $f(x)$). Na wykładzie z fizyki spotkasz się niebawem z innymi „bazami przestrzeni funkcyjnych”. Te inne bazy to rodziny ortogonalnych wielomianów. Jest tych rodzin kilka – co można wytłumaczyć prosto różnymi potrzebami, jakie stawia przed nami wybór takiego a nie innego układu współrzędnych do opisu formalnego zjawisk. W poprzednim podpunkcie powiedzieliśmy, że w „zwykłej”, 3-wymiarowej przestrzeni możemy posługiwać się układem współrzędnych kartezjańskich (x, y, z) , współrzędnych biegunowych sferycznych (r, θ, ϕ) i cylindrycznych (ρ, ϕ, z) . Wymienione współrzędne to zmienne, którym odpowiadają różne przedziały zmienności – a z kolei takim przedziałom będą odpowiadały różne ortogonalne bazy. Sytuację ilustruje tabela:

Układ	Współrzędna	Przedział zmienności	Ortogonalna baza
Kartezjański	$x(y, z)$	$(-\infty, +\infty)$	Wielomiany Hermite'a
Sferyczny	r	$[0, +\infty)$	Wielomiany Laguerre'a
Sferyczny	$\cos \theta$	$[-1, +1]$	Wielomiany Legendre'a
Sferyczny (lub cylindryczny)	ϕ	$[0, 2\pi]$	$\sin nx, \cos nx$

Tabela 12.1: Różne zmienne, ich zakresy oraz ortogonalne bazy.

Nasuwa się oczywiście nasze zwykłe pytanie: po co to wszystko? Dlaczego mamy zadawać sobie tyle trudu i proste funkcje 12-58 przekształcać do niewątpliwie bardziej skomplikowanych postaci – 12-60? A w dodatku – jak wyznaczyć występujące w tej ostatniej współczynniki C_n ? Odpowiedź na to drugie pytanie jest prosta:

C_n wyznaczamy tak samo jak wyznaczamy współrzędne V_i wektora \mathbf{V} w bazie $\{\hat{\xi}_i\}$. Te ostatnie to nic innego jak rzuty wektora \mathbf{V} na poszczególne wersory bazy, a więc iloczyny skalarne:

$$V_i = (\hat{\xi}_i, \mathbf{V}). \quad (12-69)$$

Podobnie współczynniki C_n to „rzuty” funkcji $f(x)$ na ortonormalne funkcje bazowe:

$$C_n = \left(\sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L}, f(x) \right) = \int_0^L f(x) \sin \frac{n\pi x}{L} f(x) dx. \quad (12-70)$$

Jeżeli jesteś już biegły w rachunku całkowym to możesz sprawdzić, że

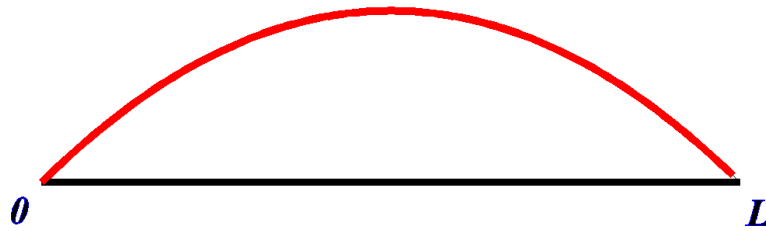
$$\begin{aligned} C_n &= \int_{x_0}^L h \frac{x}{x_0} dx + \int_{x_0}^L \sin \frac{n\pi x}{L} h \frac{L-x}{L-x_0} dx \\ &= \frac{2hL^2}{\pi^2 x_0 (L-x_0)} \sin \frac{n\pi x_0}{L} \frac{1}{n^2}. \end{aligned} \quad (12-71)$$

Powyższy wzór jest dość ciekawy. Na przykład, warto w nim zauważyć, że dla n będącego całkowitą wielokrotnością n^* , określonego z warunku

$$n^* x_0 = LN_{min}$$

gdzie N_{min} jest liczbą całkowitą (najmniejszą z możliwych), współczynniki 12-71) są równe zero. Odpowiada to znikaniu z szeregu Fouriera składowych harmonicznych — podstawowej i wyższych — dla których punkt x_0 jest węzłem fali stojącej. Może spróbujesz się zastanowić, czy powyższy warunek musi być zawsze (tzn. prędzej czy później) spełniony.

No dobrze, ale po co obliczać C_n ? Po to, że C_n obliczone dla rozwinięcia funkcji $f(x)$, reprezentującej *początkowy* kształt struny (dla $t = 0$), to *te same* C_n , które występują w określeniu funkcji $\psi(x, t)$, stanowiącej pełny opis drgań w *dowolnej* chwili t . Formalne rozwiązanie dla funkcji $\psi(x, t)$ — wzór 12-57 — zawiera w sobie współczynniki C_n , których wyznaczenie sprowadza się do rozłożenia funkcji $f(x)$, *pewnej informacji początkowej* w dyskutowanym problemie drgań, w ortonormalnej bazie 12-67. Gdyby początkowe położenie punktów struny było inne, na przykład zamiast trójkąta mielibyśmy parabolę (por. Rys.12.10):



Rysunek 12.10: Pobudzenie struny do drgań poprzez odciągnięcie jej do kształtu parabolicznego.

$$f_1(x) \equiv \psi(x, t = 0) = \frac{1}{L} x(L-x), \quad (12-72)$$

to wzór 12-57 pozostaje słuszny, ale współczynniki C_n należy w nim zastąpić nowymi współczynnikami C'_n , które otrzymamy rozkładając funkcję $f_1(x)$ w bazie 12-67:

$$f_1 = \sum_{n=1}^{\infty} C'_n \sin \frac{n\pi x}{L}. \quad (12-73)$$

Możemy mieć nieskończenie wiele różnych „sytuacji początkowych”, które prowadzą do różnych typów drgań struny. Wszystkie te drgania opisane są jednym i tym samym równaniem 12-57, a raczej nieco ogólniejszym

$$\psi(x, t) = \sum_{n=1}^{\infty} \sin \frac{n\pi x}{L} \left[C_n \cos \frac{n\pi vt}{L} + D_n \sin \frac{n\pi vt}{L} \right], \quad (12-74)$$

w którym współczynniki D_n będą różne od zera, jeżeli w chwili $t = 0$ poszczególne punkty struny uzyskały pewne prędkości początkowe. Wyznamy je w sposób analogiczny do tego jaki służył do wyznaczenia C_n . Funkcje reprezentującą rozkład początkowych prędkości rozłożymy w szereg analogiczny do szeregów 12-60 i 12-73.

Zastosowanie bazy w przestrzeni funkcyjnej to znowu wprowadzenie pewnego języka formalnego, który ma służyć do kodowania całej informacji o rozwiązywanym problemie. Zauważmy, że kształt początkowy w postaci trójkąta to coś zdecydowanie różnego od paraboli; w momencie „przełożenia” obu kształtów na szereg sinusów oba kształty stają się do siebie podobne – przynajmniej ze względu na swoją strukturę. Taki zabieg translatorski – przełożenie *całej* informacji na język pewnej bazy gwarantuje spójność i klarowność opisu całego problemu.

Rozdział 13

Wstęp do rachunku tensorowego

13.1 Wprowadzenie

W drugim rozdziale, mówiąc o własności liniowości podkreśliliśmy oczywiste *proporcjonalności*, zachodzące pomiędzy skutkami wywoływanymi przez proporcjonalne przyczyny. Takie proporcjonalności znajdujemy w większości wzorów fizycznych¹. Siła \mathbf{F} działająca na umieszczony w danym punkcie pola elektrycznego \mathbf{E} ładunek q jest proporcjonalna do wielkości ładunku (proporcjonalność: wektor-skalar) jak i do natężenia pola elektrycznego w tym punkcie (proporcjonalność wektor-wektor):

$$\mathbf{F} = q\mathbf{E}.$$

W przypadku proporcjonalności wektora do skalaru rozumiemy, że zwiększenie (zmniejszenie) wartości skalaru powoduje odpowiedni przyrost (malenie) długości wektora; proporcjonalność wektor-wektor implikuje – w tym przypadku – współliniowość obu wektorów \mathbf{F} i \mathbf{E} . Możemy mieć sytuacje nieco bardziej skomplikowane. Siła Lorentza $\mathbf{F}_L = q\mathbf{v} \times \mathbf{B}$, z jaką pole \mathbf{B} oddziałuje na poruszający się w nim, z prędkością \mathbf{v} , ładunek elektryczny q jest proporcjonalna do dwóch wektorów, ale w sposób bardziej złożony: długość wektora \mathbf{F}_L jest proporcjonalna zarówno do długości wektora \mathbf{v} jak i \mathbf{B} , ale zależy też od kąta pomiędzy oboma wektorami. W dodatku kierunek tej siły nie jest kierunkiem żadnego z dwóch wektorów, choć jest on przez kierunki \mathbf{v} i \mathbf{B} jednoznacznie wyznaczony (prostopadły do płaszczyzny napiętej przez oba wektory).

Opisane powyżej proporcjonalności to jeszcze nie wszystko. Fizyka zna sytuacje, w których zależność: skutek-przyczyna ma charakter bardziej złożony. W większości przypadków mamy z nimi do czynienia w fizyce ośrodków ciągłych i fizyce ciała stałego, a konkretnie – takich ciał stałych, które wykazują własność anizotropowości, to znaczy, w których skala efektu zależy nie tylko od intensywności przyłożonego bodźca, ale również od kierunku w którym ten efekt jest obserwowany. Wyobraźmy sobie dielektryk, wprowadzony do zewnętrznego pola elektrycznego \mathbf{E} . Na skutek uporządkowania (ewentualnie: wytworzenia i uporządkowania) dipoli dielektryka przez pole \mathbf{E} , w dielektryku powstaje wektor polaryzacji elektrycznej \mathbf{P} , który w zdecydowanej większości materiałów jest proporcjonalny do pola elektrycznego:

$$\mathbf{P} = \chi\mathbf{E}, \quad (13-1)$$

gdzie χ to tzw. polaryzowalność elektryczna. Oba wektory mają wspólny kierunek i zwrot, a stała materiałowa χ mówi tylko o „skuteczności” uporządkowania dipoli.

Istnieją jednak materiały dielektryczne, które zachowują się inaczej. Kryształ dielektryka, wprowadzony do zewnętrznego pola może porządkować swoje dipole elektryczne w sposób wynikający z jego wewnętrznej struktury; w efekcie może się zdarzyć, że wektor polaryzacji *nie* jest proporcjonalny w zwykłym sensie do wektora \mathbf{E} . Wektor \mathbf{P} przyjmuje kierunek *różny* od kierunku pola \mathbf{E} . Zwiększenie, na przykład dwukrotne, zewnętrznego pola (\mathbf{E}) powoduje dwukrotne zwiększenie wektora \mathbf{P} (jego długość wzrasta dwukrotnie), ale kierunki obu pól są różne! W ogólności zamiast wektorowego równania 13-1, które rozpisane dla trzech kierunków ($x \equiv x_1, y \equiv x_2, z \equiv z_3$) to

$$\begin{aligned} P_1 &= \chi E_1 \\ P_2 &= \chi E_2, \\ P_3 &= \chi E_3 \end{aligned} \quad (13-2)$$

¹Dziękuję prof. Andrzejowi Maksymowiczowi za interesujące dyskusje, które podsunęły mi pomysł takiego wprowadzenia w temat tensorów.

będziemy mieli – znacznie bardziej skomplikowane –

$$\begin{aligned} P_1 &= \chi_{11}E_1 + \chi_{12}E_2 + \chi_{13}E_3, \\ P_2 &= \chi_{21}E_1 + \chi_{22}E_2 + \chi_{23}E_3, \\ P_3 &= \chi_{31}E_1 + \chi_{32}E_2 + \chi_{33}E_3. \end{aligned} \quad (13-3)$$

Fizycznie oznacza to, że natężenie wektora polaryzacji w kierunku osi $0x$ (\mathbf{P}_1) zależy nie tylko od x -owej składowej natężenia pola elektrycznego (\mathbf{E}_1), ale również od składowych tego pola w kierunku osi $0y$ i $0z$ (\mathbf{E}_2 i \mathbf{E}_3). Matematycznie – zastąpienie jednej stałej materiałowej, polaryzowalności χ , dziewięcioma liczbami χ_{ik} oznacza, że charakter tej stałej jest już inny. Nie jest już ona skalarem, ale tensorem drugiego rzędu.

13.2 Kowariantność i kontrawariantność

Zanim wprowadzimy formalną definicję tensora przypomnijmy sobie jak wyrażają się współrzędne wektora przy przejściu od jednej bazy przestrzeni wektorowej \mathbb{R}^n (wektory: $\mathbf{e}_1, \dots, \mathbf{e}_n$) do innej (wektory: $\mathbf{e}'_1, \dots, \mathbf{e}'_n$), a także prawa transformacji samych wektorów bazy. Wprowadzimy też znaczną modyfikację zapisu. Po pierwsze, „nową” („primowaną”) bazę będziemy oznaczali poprzez znaki „prim” *umieszczone bezpośrednio przy wskaźnikach wielkości wektorowych*. Po drugie *rezygnujemy* z pogrubionego druku dla wyrażania wektorów bazowych. Po trzecie: aby podkreślić różnice pomiędzy współrzędnymi wektora i wektorami bazowymi będziemy używali wskaźników usytuowanych bądź to u góry (współrzędne wektora), bądź to u dołu (wektory bazowe). Po czwarte, wprowadzamy szalenie praktyczną *konwencję sumacyjną² pomijania znaku sumy*, w sytuacjach kiedy sumowanie jest ewidentne. Patrząc wstecz, bez trudu zauważymy, że w przypadku wyrażień stojących pod znakiem sumy, sumowanie *zawsze* wiązało się z koniecznością powtórzenia wskaźnika. Dlatego *powtarzający się wskaźnik traktujemy od tej chwili zawsze jako wskaźnik względem którego przeprowadzana jest operacja sumowania po wszystkich możliwych wartościach*. Zapis wektora \mathbf{x} przyjmuje postać:

$$\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i \rightarrow \sum_{i=1}^n x^i e_i \equiv x^i e_i, \quad (13-4)$$

oraz

$$\mathbf{x} = \sum_{i=1}^n x'_i \mathbf{e}'_i \rightarrow \sum_{i=1}^n x^{i'} e_{i'} \equiv x^{i'} e_{i'}. \quad (13-5)$$

Konwencja sumowania, połączona z dwoma typami wskaźników: górnym i dolnym czyni wzory bardziej przejrzystymi; sumujemy po wskaźniku, który zajmuje górne i dolne położenie. I tak, wzory określające transformację wektorów bazowych (por. punkt 9.7) przyjmują postać:

$$\begin{aligned} e_{1'} &= A_{1'}^1 e_1 + A_{1'}^2 e_2 + \dots + A_{1'}^n e_n, \\ e_{2'} &= A_{2'}^1 e_1 + A_{2'}^2 e_2 + \dots + A_{2'}^n e_n, \\ &\quad \vdots \quad \quad \quad \vdots \\ e_{n'} &= A_{n'}^1 e_1 + A_{n'}^2 e_2 + \dots + A_{n'}^n e_n, \end{aligned} \quad (13-6)$$

albo w skrócie

$$e_{i'} = \sum_{i=1}^n A_{i'}^i e_i \equiv A_{i'}^i e_i. \quad (13-7)$$

Współczynniki $A_{i'}^i$ to elementy macierzy przejścia; przejście następuje od układu rozpiętego przez wersory e_i do układu rozpiętego przez wersory $e_{i'}$, czyli „kierunek przejścia” w odniesieniu do pary wskaźników i (wskaźnik górny – kolumny) oraz i' (wskaźnik górny – wiersza) to „z góry na dół”: $A_{i'}^i \equiv A_{i'}^i \downarrow$. Ponieważ wektory bazowe $e_{i'}$ tworzą układ wektorów liniowo niezależnych, macierz przejścia:

$$\mathcal{A} \equiv \begin{pmatrix} A_{1'}^1 & A_{1'}^2 & \dots & A_{1'}^n \\ A_{2'}^1 & A_{2'}^2 & \dots & A_{2'}^n \\ \vdots & \vdots & \dots & \vdots \\ A_{n'}^1 & A_{n'}^2 & \dots & A_{n'}^n \end{pmatrix} \quad (13-8)$$

²Zwaną czasami – w sposób niezbyt uzasadniony – konwencją Einsteina.

jest macierzą nieosobliwą i istnieje macierz odwrotna

$$\mathcal{A}^{-1} \equiv \begin{pmatrix} A_1^{1'} & A_1^{2'} & \dots & A_1^{n'} \\ A_2^{1'} & A_2^{2'} & \dots & A_2^{n'} \\ \vdots & \vdots & \dots & \vdots \\ A_n^{1'} & A_n^{2'} & \dots & A_n^{n'} \end{pmatrix}. \quad (13-9)$$

Macierz \mathcal{A}^{-1} to macierz przejścia od układu primowanego do nie-primowanego $A_i^{i'} \equiv A_i^{i'} \downarrow$; wzory analogiczne do 13-6 i 13-7 to

$$\begin{aligned} e_1 &= A_1^{1'} e_{1'} + A_1^{2'} e_{2'} + \dots + A_1^{n'} e_{n'}, \\ e_2 &= A_2^{1'} e_{1'} + A_2^{2'} e_{2'} + \dots + A_2^{n'} e_{n'}, \\ &\vdots \\ e_n &= A_n^{1'} e_{1'} + A_n^{2'} e_{2'} + \dots + A_n^{n'} e_{n'} \end{aligned} \quad (13-10)$$

albo w skrócie

$$e_i = \sum_{i'=1}^n A_i^{i'} e_{i'} \equiv A_i^{i'} e_{i'}. \quad (13-11)$$

Iloczyn macierzy \mathcal{A} i \mathcal{A}^{-1} to macierz jednostkowa:

$$\mathcal{A}\mathcal{A}^{-1} = \mathcal{I} = \mathcal{A}^{-1}\mathcal{A}, \quad (13-12)$$

albo, *explicite*

$$A_{k'}^j A_i^{k'} = \delta_i^j, \quad (13-13)$$

$$A_k^{j'} A_{i'}^k = \delta_{i'}^{j'}. \quad (13-14)$$

Zauważmy, że para wskaźników delty Kroneckera zajmuje teraz górne i dolne położenie; wyjaśnimy to bliżej w następnym podrozdziale. Wzór 13-13 można otrzymać też, podstawiając z 13-7 do 13-11 za $e_{i'}$:

$$e_i = A_i^{i'} e_{i'} = A_i^{i'} A_{i'}^j e_j = A_i^j A_{i'}^{i'} e_j = \delta_i^j e_j = e_i. \quad (13-15)$$

Prawa transformacji współrzędnych wektora \mathbf{x} – elementu \mathbb{R}^n :

$$\mathbf{x} = e_i x^i = e_{i'} x^{i'}, \quad (13-16)$$

otrzymujemy podstawiając, na przykład z 13-11 za wersory e_i do powyższego wzoru:

$$\mathbf{x} = x^i e_i = x^i A_i^{i'} e_{i'}. \quad (13-17)$$

Stąd

$$x^{i'} = A_i^{i'} x^i, \quad (13-18)$$

albo *explicite*

$$\begin{aligned} x^{1'} &= A_1^{1'} x^1 + A_1^{2'} x^2 + \dots + A_1^{n'} x^n, \\ x^{2'} &= A_2^{1'} x^1 + A_2^{2'} x^2 + \dots + A_2^{n'} x^n, \\ &\vdots \\ x^{n'} &= A_n^{1'} x^1 + A_n^{2'} x^2 + \dots + A_n^{n'} x^n. \end{aligned} \quad (13-19)$$

Transformacje odwrotne otrzymujemy podstawiając z 13-5 za wersory $e_{i'}$ do 13-16:

$$\mathbf{x} = x^{i'} A_{i'}^i e_i, \quad (13-20)$$

a więc

$$x^i = A_{i'}^i x^{i'}, \quad (13-21)$$

albo *explicitie*

$$\begin{aligned}x^1 &= A_1^1 x^{1'} + A_2^1 x^{2'} + \dots A_n^1 x^{n'}, \\x^2 &= A_1^2 x^{1'} + A_2^2 x^{2'} + \dots A_n^2 x^{n'}, \\&\vdots \\x^n &= A_1^n x^{1'} + A_2^n x^{2'} + \dots A_n^n x^{n'}.\end{aligned}\tag{13-22}$$

Przyglądając się macierzy, która przekształca współrzędne wektora \boldsymbol{x} w układzie Σ we współrzędne układu Σ' (równanie 13-18) widzimy, że macierz ta to

$$\mathcal{A} \text{ dla współrzędnych} \equiv \begin{pmatrix} A_1^{1'} & A_2^{1'} & \dots & A_n^{1'} \\ A_1^{2'} & A_2^{2'} & \dots & A_n^{2'} \\ \vdots & \vdots & \dots & \vdots \\ A_1^{n'} & A_2^{n'} & \dots & A_n^{n'} \end{pmatrix}.\tag{13-23}$$

Macierz ta, to *transponowana macierz odwrotna* (do macierzy przejścia) – por. wzór 13-9. Wynik ten uzyskaliśmy jeszcze w rozdziale dziewiątym – por. wzór 9-38. Występujące w nim współrzędne wektora $x^{i'}$ mają postać wektora jednowierszowego. Chcąc przekształcić wzór 9-38 do „zwykłej” postaci (współrzędne uporządkowane w jednej kolumnie) musimy poddać transpozycji obie strony równania, pamiętając że transpozycja z iloczynu macierzy jest równa iloczynowi transpozycji poszczególnych czynników, ale w porządku odwrotnym. Wzory

$$e_{i'} = A_{i'}^i e_i \tag{13-24}$$

$$x^{i'} = A_i^{i'} x^i \tag{13-25}$$

to fundament rachunku wektorowego. Występująca w 13-24 macierz współczynników $\mathcal{A} \equiv (A_{i'}^i)$ (macierz przejścia $\Sigma \rightarrow \Sigma'$) zawiera w sobie całą informację – jej macierz odwrotna (macierz przejścia $\Sigma' \rightarrow \Sigma$) poddana transpozycji to macierz $\mathcal{A}_{\text{współrzędnych}}$ – por. 13-23.

Wprowadzamy w tym miejscu definicję:

Definicja 13.1 *Wektorem kowariantnym nazywamy wektor, którego prawa transformacji przy przejściu $\Sigma \rightarrow \Sigma'$ dane są równaniem 13-24; prawa transformacji wektora kontrawariantnego określa 13-25.*

Określenia ko- i kontra-wariantny to „współzmienny” i „przeciwzmienny”. Wektory bazowe e_i są kowariantne – przy zmianie układu przekształcają się w „nowe”; wektory bazowe zmieniają się *wraz* z układem. Wektor kontrawariantny natomiast to wektor, który pozostaje zawsze „na swoim miejscu” w przestrzeni wektorowej i nie uczestniczy w żadnych zmianach.

Jeżeli macierz odwrotna do macierzy \mathcal{A} jest macierzą transponowaną: $\mathcal{A}^{-1} = \mathcal{A}^T$ to macierz $\mathcal{A}_{\text{współrz.}} = (\mathcal{A}^{-1})^T = (\mathcal{A}^T)^T = \mathcal{A}$ jest *tą samą macierzą przejścia*. Transformacje ortogonalne mają tę sympatyczną własność, że i wersory osi i współrzędne wektorów transformują się podlegając działaniu jednej i tej samej macierzy.

13.2.1 Forma liniowa

Wprowadzona w rozdziale szóstym (por. punkt 6.1.1) forma liniowa to liniowy operator przyporządkowujący każdemu wektorowi \boldsymbol{x} pewną wartość liczbową

$$f(\boldsymbol{x}) = f(x^i e_i) = \dots \text{liniowość} \dots = x^i f(e_i) \equiv x^i a_i.\tag{13-26}$$

Współczynniki formy $a_i = f(e_i)$ zależą od wyboru bazy. Ich prawo transformacyjne uzyskujemy zapisując formę w układzie Σ' :

$$f(\boldsymbol{x}) = f(x^{i'} e_{i'}) = x^{i'} f(e_{i'}) \equiv x^{i'} a_{i'}.\tag{13-27}$$

Ale

$$f(e_{i'}) = f(A_{i'}^i e_i) = A_{i'}^i f(e_i) = A_{i'}^i a_i,\tag{13-28}$$

albo krótko

$$a_{i'} = A_{i'}^i a_i. \quad (13-29)$$

Współczynniki formy transformują się jak wektor kowariantny.

Warto w tym miejscu uzmysłowić sobie, że forma liniowa w \mathbb{R}^n może reprezentować sobą pewną *hiperpowierzchnię* tej przestrzeni. Na przykład równanie

$$a_i x^i = 1 \quad (13-30)$$

to – w przypadku trójwymiarowej przestrzeni ($x_1 = x$, $x_2 = y$, $x_3 = z$) równanie płaszczyzny

$$a_1 x + a_2 y + a_3 z = 1.$$

Podstawiając za x^i z 13-21 do 13-30 i uwzględniając 13-29 mamy

$$a_i x^i = a_i A_{i'}^i x^{i'} = a_{i'} x^{i'} = 1 \quad (13-31)$$

– równanie hiperpowierzchni pozostaje niezmiennikiem transformacji $\Sigma \rightarrow \Sigma'$. Nie jest to niespodzianką – forma liniowa w dowolnym układzie przyporządkowuje wektorowi pewną liczbę (skalar), której wartość nie może zależeć od układu odniesienia.

13.2.2 Gradient jako wektor kowariantny

Gradientem³ pola skalarnego ϕ – pozostawmy dla ustalenia uwagi w przestrzeni trójwymiarowej – nazywamy funkcję pola $\phi = \phi(x, y, z) = \phi(x_1, x_2, x_3)$.

$$\nabla\phi = \mathbf{e}_1\phi_1 + \mathbf{e}_2\phi_2 + \mathbf{e}_3\phi_3 \equiv \mathbf{e}_1 \frac{\partial\phi}{\partial x^1} + \mathbf{e}_2 \frac{\partial\phi}{\partial x^2} + \mathbf{e}_3 \frac{\partial\phi}{\partial x^3}. \quad (13-32)$$

Zmiana układu odniesienia powoduje transformację współrzędnych gradientu

$$\phi_{i'} \equiv \frac{\partial\phi}{\partial x^{i'}} = \frac{\partial\phi}{\partial x^i} \frac{\partial x^i}{\partial x^{i'}} = \frac{\partial x^i}{\partial x^{i'}} \phi_i. \quad (13-33)$$

Korzystając z 13-21 mamy

$$\frac{\partial x^i}{\partial x^{i'}} = A_{i'}^i. \quad (13-34)$$

i konsekwentnie

$$\phi_{i'} = A_{i'}^i \phi_i \quad (13-35)$$

tak więc współrzędne gradientu transformują się jak wektor kowariantny.

Wzór 13-34 wiąże współczynniki macierzy \mathcal{A} z pochodnymi (cząstkowymi) starych (układ Σ) współrzędnych względem nowych (układ Σ'). Zauważmy, że wskaźniki współczynników $A_{i'}^i$ (górny i dolny) to odpowiednio wskaźniki x -a z licznika i mianownika pochodnej. Analogiczny wzór dla wektora kontrawariantnego uzyskujemy różniczkując wzór 13-25 względem x_i ⁴:

$$\frac{\partial x^{i'}}{\partial x^i} = A_i^{i'}. \quad (13-36)$$

W przypadku kiedy oba układy odniesienia: Σ i Σ' są układami *kartezjańskimi* obie pochodne cząstkowe – występujące w 13-34 i 13-36 są równe

$$\frac{\partial x^{i'}}{\partial x^i} = \frac{\partial x^i}{\partial x^{i'}} = \cos \angle(0x', 0x) = \cos \angle(0x, 0x') \quad (13-37)$$

(porównaj dyskusję w podrozdziale 4.5). Jest to jeszcze jedno sformułowanie własności *ortogonalności* transformacji. W przypadku takiej transformacji (obrót układu) zacierają się różnice między prawami transformacji wektorów ko- i kontra-wariantnych.

³Wkraczamy w obszar analizy wektorowej (tensorowej). Krótka wzmianka – punkt 13.7.

⁴We wzorze 13-25 wskaźnik i jest wskaźnikiem sumacyjnym i jako taki przybiera wszystkie (trzy) wartości. Różniczkujemy względem x_i , przy *ustalonym* i – z sumy trzech składników w 13-25 tylko jeden będzie różny od zera.

13.3 Tensory – podstawowe definicje

Znane przez nas doskonale wielkości skalarne to *tensory zerowego rzędu*; podobnie jak wektory (ko- i kontra-wariantne) są odpowiednimi *tensorami pierwszego rzędu*. Klucz do definicji to prawa transformacji przy przejściu od układu Σ do Σ' : skalary nie doznają żadnych zmian, prawa transformacji tensorów pierwszego rzędu określają wzory 13-24 i 13-25. Konsekwentnie, *tensory drugiego rzędu*: kontra-wariantne, mieszane lub kowariantne to wielkości podlegające prawom transformacji

$$\begin{aligned} T^{i'j'} &= \frac{\partial x^{i'}}{\partial x^i} \frac{\partial x^{j'}}{\partial x^j} T^{ij} \equiv A_i^{i'} A_j^{j'} T^{ij}, \\ U_{j'}^{i'} &= \frac{\partial x^{i'}}{\partial x^i} \frac{\partial x^j}{\partial x^{j'}} U_j^i \equiv A_i^{i'} A_j^{j'} U_j^i, \\ W_{i'j'} &= \frac{\partial x^i}{\partial x^{i'}} \frac{\partial x^j}{\partial x^{j'}} W_{ij} \equiv A_i^{i'} A_j^{j'} W_{ij}. \end{aligned} \quad (13-38)$$

Tensor T^{ij} jest kontra-wariantny względem obu wskaźników; W_{ij} – kowariantny, a tensor U_j^i podlega kontra-wariantnym regułom transformacji dla górnego wskaźnika, a kowariantnym – dla dolnego. Podkreślny raz jeszcze, że wszystkie te różnice znikają w przypadku gdy w naszej „zwykłej” trójwymiarowej przestrzeni stosujemy układy współrzędnych kartezjańskich.

Prawa transformacji – równania 13-38 – ulegają naturalnemu rozszerzeniu na przypadek tensorów wyższych rzędów. Generalnie: tensor rzędu q to wielkość, w której określeniu (zapisie) występuje q wskaźników, w dowolnej kombinacji ko- i kontra-wariantnej; wzór transformacyjny typu 13-38 zawiera natomiast q czynników typu A_j^i lub $A_i^{j'}$. W przestrzeni 3-wymiarowej tensor rzędu q reprezentuje 3^q liczb; w przestrzeni n -wymiarowej liczba składowych tensora to n^q . Dalszą dyskusję tensorów prowadzić będziemy jednak „pragmatycznie” ograniczając się do przestrzeni o trzech (ewentualnie dwóch) wymiarach.

Zilustrowana fizycznym przykładem idea tensora drugiego rzędu mogłaby powodować fałszywe wrażenie, że w zasadzie tensorem drugiego rzędu będzie każde 9 wielkości, uporządkowanych w macierz 3×3 ; działanie takiej macierzy na wektor (mnożenie) daje nam inny (w sensie innej długości i kierunku) wektor – por. punkt 7.1.7. Tak jednak nie jest – aby dziewięć wielkości mogło tworzyć tensor *muszą* one podlegać prawom transformacji jak w 13-38. I tak, dwu-wymiarowy kandydat na tensor drugiego rzędu, macierz

$$T = \begin{pmatrix} -xy & -y^2 \\ x^2 & xy \end{pmatrix} \quad (13-39)$$

(w układzie Σ) i

$$T' = \begin{pmatrix} -x'y' & -y'^2 \\ x'^2 & x'y' \end{pmatrix} \quad (13-40)$$

w obróconym o kąt θ układzie Σ będzie tensorem, jeżeli wszystkie cztery primowane współrzędne powiązane są z nieprimowanymi wzorami jak w 13-38. Na przykład

$$T^{1'1'} = -x'y' \stackrel{?}{=} \frac{\partial x^{1'}}{\partial x^i} \frac{\partial x^{1'}}{\partial x^j} T^{ij} = A_i^{1'} A_j^{1'} T^{ij}.$$

Współczynniki $A_1^{1'}$ i $A_2^{1'}$ to odpowiednio $\cos \theta$ i $\sin \theta$ (por. punkt 4.5); do sprawdzenia pozostaje więc

$$\begin{aligned} & -(x \cos \theta + y \sin \theta)(-x \sin \theta + y \cos \theta) \\ & \stackrel{?}{=} \cos^2 \theta T^{11} + \cos \theta \sin \theta T^{12} + \sin \theta \cos \theta T^{21} + \sin^2 \theta T^{22} \\ & = -xy \cos^2 \theta - y^2 \cos \theta \sin \theta + x^2 \sin \theta \cos \theta + xy \sin^2 \theta. \end{aligned}$$

Powyższa tożsamość weryfikuje – dla $T^{1'1'}$! – prawo transformacji; okazuje się (warto sprawdzić), że weryfikacja wypada pozytywnie i dla trzech pozostałych składowych tensora. Macierz 13-39 reprezentuje rzeczywiście tensor drugiego rzędu.

Tego typu weryfikację własności transformacyjnych dla „kandydatów na tensory” można przeprowadzić i w sposób bardziej ogólny. Na przykład, nasza delta Kroneckera – δ_{kl} okazuje się być tensorem *mieszanym* drugiego rzędu – δ_l^k . Jeżeli tak, to prawo transformacji będzie miało postać

$$\delta_{j'}^{i'} \stackrel{?}{=} A_k^{i'} A_l^{j'} \delta_l^k = \frac{\partial x^{i'}}{\partial x^k} \frac{\partial x^l}{\partial x^{j'}} \delta_l^k. \quad (13-41)$$

Przekształcając prawą stronę mamy (z definicji delty Kroneckera)

$$\frac{\partial x^{i'}}{\partial x^k} \frac{\partial x^l}{\partial x^{j'}} \delta_l^k = \frac{\partial x^{i'}}{\partial x^k} \frac{\partial x^k}{\partial x^{j'}} = (\text{z definicji}) = \frac{\partial x^{i'}}{\partial x^{j'}}.$$

Ostatnie wyrażenie to oczywiście 1 dla $i' = j'$ i zero w innym przypadku; delta Kroneckera δ_l^k jest więc rzeczywiście tensorem mieszanym drugiego rzędu. W dodatku, elementy macierzy tego tensora są identyczne we wszystkich (obróconych względem siebie) układach współrzędnych. Dlatego deltę Kroneckera nazywa się *tensorem izotropowym drugiego rzędu*.

13.4 Podstawowe operacje: dodawanie, mnożenie, kontrakcja

Jest chyba oczywiste, że dodawać do siebie możemy tylko tensory tego samego rzędu i typu. Wynik takiej operacji pozostaje tensorem. Podobnie przemnożenie tensora przez liczbę (skalar) nie zmienia charakteru wielkości tensorowej.

Mnożenie dwóch tensorów przez siebie daje tensor, którego rząd jest równy sumie rzędów mnożonych tensorów. Na przykład

$$T^{ij} U_l^k = V_l^{ijk}. \quad (13-42)$$

Wynik mnożenia, to mieszany tensor czwartego rzędu, ponieważ

$$T^{i'j'} U_{l'}^{k'} = A_i^{i'} A_j^{j'} T^{ij} A_k^{k'} A_{l'}^l U_l^k = A_i^{i'} A_j^{j'} A_k^{k'} A_{l'}^l (T^{ij} U_l^k) \quad (13-43)$$

– spełnione jest równanie, stanowiące uogólnienie 13-38 na przypadek większej niż dwa liczby wskaźników. *Kontrakcja* (po polsku: zwięźenie; termin ten często bywa stosowany) tensora to operacja polegająca na zrównaniu dwóch wskaźników tensora: ko- i kontrawariantnej, a następnie – w zgodzie z konwencją sumacyjną – dodania składowych tensora, w których występują te powtórzone wskaźniki. Na przykład kontrakcja utworzonego przed chwilą tensora czwartego rzędu (jesteśmy w trójwymiarowej przestrzeni) V_l^{ijk} względem pierwszej pary wskaźników to

$$V_l^{iik} = V_l^{11k} + V_l^{22k} + V_l^{33k} \equiv U_l^k; \quad (13-44)$$

otrzymana wielkość nie zależy już od wysumowanych po pełnym zakresie zmienności wskaźników i i j , natomiast zależy od dwóch wskaźników (k, l) , a więc jest tensorem drugiego rzędu. Ogólnie, operacja kontrakcji obniża rząd tensora o dwa. Ma to szczególne znaczenie w przypadku tensora drugiego rzędu:

$$\begin{aligned} U_{j'}^{i'} \xrightarrow{\text{kontrakcja}} U_{i'}^{i'} &= \frac{\partial x^{i'}}{\partial x^k} \frac{\partial x^l}{\partial x^{i'}} U_l^k = \frac{\partial x^l}{\partial x^{i'}} \frac{\partial x^{i'}}{\partial x^k} U_l^k \\ &= \frac{\partial x^l}{\partial x^k} U_l^k = \delta_k^l U_l^k = U_k^k. \end{aligned} \quad (13-45)$$

Poddany kontrakcji tensor drugiego rzędu staje się tensorem zerowego rzędu (skalarem); jego (jedna) wartość nie zależy od układu odniesienia.

Nasuwającym się natychmiast skojarzeniem jest iloczyn skalarny dwóch wektorów:

$$\mathbf{A} \cdot \mathbf{B} = A_i B_i. \quad (13-46)$$

Rzeczywiście: iloczyn dwóch wektorów jest tensorem drugiego rzędu, ale zachodzi pewna niedogodność: kontrakcję zdefiniowaliśmy jako operację, która odbywa się względem pary dwóch różnych wskaźników: ko- i kontrawariantnego. Ponieważ z iloczynem skalarnym typu 13-46 mamy *de facto* do czynienia w układach kartezjańskich to różnica kowariantny–kowariantny jest tu bez znaczenia.⁵ Dlatego od tego momentu „zapominamy” o rozróżnianiu tych dwóch typów tensorów i stosować będziemy zapis z wskaźnikami występującymi zawsze „u dołu” symbolu wielkości tensorowej.

To, że suma elementów przekątnej głównej macierzy – reprezentacji tensora drugiego rzędu – jest skalarem, a więc *niezmiennikiem* transformacji, ma często interesującą interpretację fizyczną. Powrócimy do tego zagadnienia w następnym punkcie. Sumę taką nazywamy *śladem* tensora drugiego rzędu i oznaczamy

$$u_{ii} = u_{11} + u_{22} + u_{33} \equiv \text{Tr} (u_{ik}). \quad (13-47)$$

⁵Przypomnijmy sobie, że przy zapisywaniu iloczynu skalarnego w formie iloczynu dwóch macierzy: jedno-kolumnowej i jedno-wierszowej też musieliśmy nadawać obu czynnikom formy różniące się między sobą.

13.5 Tensory antysymetryczne i symetryczne; iloczyn wektorowy

Kwestię (anty)symetrii tensorów najwygodniej jest związać z ich reprezentacją w postaci macierzy. Wprawdzie ta ostatnia dotyczy tylko tensorów drugiego rzędu, ale pojęcia tensora (anty)symetrycznego dla tensorów wyższych rzędów można wprowadzić drogą prostych uogólnień. Tensor symetryczny to tensor, którego składowe nie zmieniają się przy zmianie porządku wskaźników; antysymetryczny to ten, dla którego zmiana porządku wskaźników powoduje zmianę znaku składowej tensora:

$$S_{ik} = S_{ki} \quad \text{oraz} \quad A_{ik} = -A_{ki}. \quad (13-48)$$

Każdy tensor T_{ik} może być przestawiony jako suma części symetrycznej $T_{ik}^{(s)}$ i antysymetrycznej $T_{ik}^{(a)}$:

$$T_{ik} = \frac{1}{2}(T_{ik} + T_{ki}) + \frac{1}{2}(T_{ik} - T_{ki}) \equiv T_{ik}^{(s)} + T_{ik}^{(a)}. \quad (13-49)$$

Symetria (antysymetria) tensora opisującego pewne wielkości fizyczne stanowi ważny atrybut formalny, odzwierciedlający zawsze pewną treść fizyczną.

W rozdziale czwartym, mówiąc o iloczynie wektorowym, wspomnieliśmy, że nie jest on „do końca wektorem”. Rzeczywiście. Rozważmy antysymetryczny tensor drugiego rzędu:

$$A_{ik} = -A_{ki}.$$

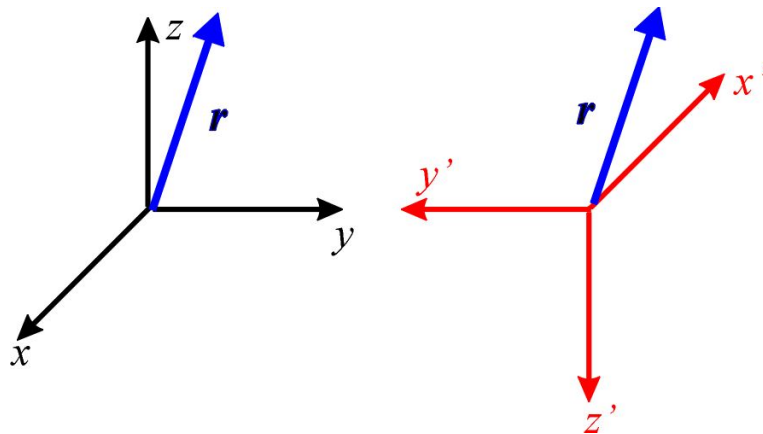
Macierz będąca reprezentacją tego tensora ma na przekątnej głównej zera ($A_{11} = -A_{11} = 0$ i podobnie dla A_{22} i A_{33}), a ze względu na relacje antysymetrii mamy tylko trzy różne elementy, które możemy oznaczyć przez z_1, z_2, z_3 :

$$\begin{pmatrix} 0 & A_{12} & A_{13} \\ A_{21} & 0 & A_{23} \\ A_{31} & A_{32} & 0 \end{pmatrix} \equiv \begin{pmatrix} 0 & -z_3 & z_2 \\ z_3 & 0 & -z_1 \\ -z_2 & z_1 & 0 \end{pmatrix}. \quad (13-50)$$

Tensorowy charakter wielkości z_i (tensor drugiego rzędu) i ich antysymetria pozwala nam zapisać je w postaci antysymetrycznych iloczynów współrzędnych wektorów (tensorów pierwszego rzędu)

$$\left. \begin{aligned} z_1 &= x_2 y_3 - x_3 y_2 \\ z_2 &= x_3 y_1 - x_1 y_3 \\ z_3 &= x_1 y_2 - x_2 y_1, \end{aligned} \right\} \quad (13-51)$$

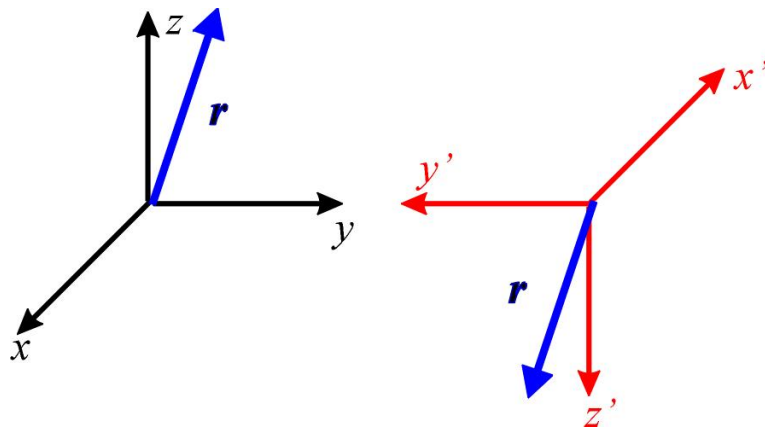
czyli trójka liczb ($z_1, z_2, z_3 \equiv \mathbf{z}$) to iloczyn wektorowy: $\mathbf{z} = \mathbf{x} \times \mathbf{y}$.



Rysunek 13.1: Inwersja układu współrzędnych w przestrzeni 3-wymiarowej – wektor biegunowy.

To, że iloczyn wektorowy nie jest „zwykłym” wektorem łatwo dostrzec rozważając transformację inwersji układu współrzędnych (por. rozdział 10.4). Współrzędne zwykłych (tzw. biegunowych) wektorów \mathbf{x} i \mathbf{y} doznają, przy inwersji układu współrzędnych, oczywistej zmiany

$$x_i \rightarrow -x_i; \quad y_i \rightarrow -y_i; \quad i = 1, 2, 3. \quad (13-52)$$



Rysunek 13.2: Inwersja układu współrzędnych w przestrzeni 3-wymiarowej – wektor osiowy .

Ale w takim razie współrzędne wektora \mathbf{z} – iloczynu wektorowego wektorów \mathbf{x} i \mathbf{y} – pozostają bez zmian; dwa znaki ujemne dają znak dodatni! O ile zwykły (biegunowy) wektor przy inwersji układu pozostaje bez zmian (Rys.13.1), to wektor taki jak \mathbf{z} , *wektor osiowy*, albo pseudo-wektor (Rys.13.2) zmienia zwrot o 180° . Okazuje się więc, że nasza pocziwa przestrzeń dwu- lub trój-wymiarowa, poprzez swój atrybut – skrętność, rozróżnia dwa typy wektorów. Oczywiście jest chyba, że „rozsądne reguły gry” wymagały na przykład, aby nie próbować dodawać do siebie wektora osiowego i biegunowego.

Zapis iloczynu wektorowego dwóch wektorów – tensorów pierwszego rzędu – nabiera formalnej elegancji, jeżeli wprowadzimy pojęcie antysymetrycznego tensora trzeciego rzędu⁶

$$\varepsilon_{ijk} = \begin{cases} 0 & i = j \vee j = k \vee i = k \\ +1 & (i, j, k) = P^+(1, 2, 3) \\ -1 & (i, j, k) = P^-(1, 2, 3) \end{cases}, \quad (13-53)$$

gdzie: $P^\pm(1, 2, 3)$ oznacza bądź parzystą (+) bądź nieparzystą (–) permutację trójki (1, 2, 3). Wówczas równania 13-51 przybierają postać

$$z_i = \varepsilon_{ijk} x_j y_k \quad i = 1, 2, 3. \quad (13-54)$$

Antysymetryczny tensor trzeciego rzędu może także służyć do zapisu iloczynu mieszanego trzech wektorów (por. punkt 4.4). Skalarna wielkość I , uzyskana poprzez kontrakcję (względem trzech par wskaźników)

$$I = \varepsilon_{ijk} x_i y_j z_k \quad (13-55)$$

to objętość równoległościanu zbudowanego na wektorach \mathbf{x} , \mathbf{y} i \mathbf{z} . Ale – i tu musimy być ostrożni. Wzór 13-55 będzie słuszny w prawoskrętnym układzie współrzędnych! W układzie lewoskrętnym definicja antysymetrycznego tensora trzeciego rzędu nakazuje zmienić znak wszystkich jego współrzędnych. Ponieważ objętość równoległościanu (wielkość nieujemna!) nie zależy od skrętności układu wynika stąd, że iloczyn mieszany nie jest zwykłym skalarzem, lecz *pseudoskalarzem*, a więc wielkością, dla której reguły transformacyjne nie są do końca takie same jak dla zwykłych skalarów. Konkretnie: pseudoskalar będzie zmieniał znak przy zmianie skrętności układu.

Symetryczne tensory drugiego rzędu są ściśle związane z zagadnieniem macierzy symetrycznej i symetrycznego operatora odwzorowania (por. punkty 10.6 i 11.3). Istotnym jest fakt, że w tym przypadku mamy zawsze możliwość przedstawienia tensora w szczególnym układzie odniesienia – układzie osi własnych, w którym tensor przybiera postać diagonalną. Te formalne atrybuty reprezentacji tensora niosą z sobą zazwyczaj ciekawe treści fizyczne – widzieliśmy to na przykładzie tensora bezwładności ciała sztywnego w poprzednim rozdziale.

13.6 Tensor deformacji i tensor naprężeń

Wprowadzone w poprzednim punkcie pojęcie ciała sztywnego – ośrodka, którego elementy zachowują stałe odległości, bez względu na działające na ośrodek siły, to pojęcie abstrakcyjne. W rzeczywistości przyłożenie

⁶Jest to tensor antysymetryczny względem dowolnej – czyli każdej – pary wskaźników.

sił powoduje deformację ciała – odległości pomiędzy jego poszczególnymi elementami ulegają zmianom. Jeżeli wyobrazimy sobie ośrodek ciągły jako układ quasi-punktowych mas, których położenia *przed* deformacją określają wektory $\mathbf{r} \equiv (x_1, x_2, x_3)$, a *po przyłożeniu sił deformujących* – $\mathbf{r}' \equiv (x'_1, x'_2, x'_3)$, to *wektorem deformacji* \mathbf{u} nazywamy wektor, którego współrzędne określone są jako:

$$u_i = x'_i - x_i. \quad (13-56)$$

Pole wektorowe \mathbf{u} – wartości wektorów \mathbf{u} w każdym punkcie podlegającego deformacji ciała – to obraz deformacji.

13.6.1 Tensor deformacji

Jeżeli wybierzemy dwa dowolne, nieskończenie blisko siebie położone punkty ciała i oznaczmy ich odległość przed deformacją jako dl to

$$dl^2 = dx_i dx_i = dx_i^2; \quad (13-57)$$

po deformacji, odległość tych samych punktów to dl' :

$$dl'^2 = dx_i'^2 = (dx_i + du_i)^2. \quad (13-58)$$

Występującą w powyższym wzorze różniczkę wektora deformacji możemy wyrazić poprzez różniczki poszczególnych współrzędnych przestrzennych i odpowiednie pochodne cząstkowe:

$$du_i = \frac{\partial u_i}{\partial x_k} dx_k. \quad (13-59)$$

Wzór 13-58 zapiszemy jako

$$dl'^2 = dx_i^2 + 2 \frac{\partial u_i}{\partial x_k} dx_i dx_k + \frac{\partial u_i}{\partial x_k} \frac{\partial u_i}{\partial x_l} dx_k dx_l. \quad (13-60)$$

Drugi wyraz po prawej stronie 13-60 jest ewidentnie symetryczny (uwaga na *podwójną* sumę!):

$$2 \frac{\partial u_i}{\partial x_k} dx_i dx_k = \left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i} \right) dx_i dx_k$$

podobnie zresztą jak trzeci; w tym ostatnim zamieniamy (sumacyjne) wskaźniki $i \leftrightarrow l$:

$$\frac{\partial u_i}{\partial x_k} \frac{\partial u_i}{\partial x_l} dx_k dx_l = \frac{\partial u_l}{\partial x_k} \frac{\partial u_l}{\partial x_i} dx_i dx_k.$$

Po tych zabiegach pozostaje tylko zdefiniować zależną od dwóch wskaźników wielkość

$$u_{ik} \equiv \frac{1}{2} \left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i} + \frac{\partial u_l}{\partial x_k} \frac{\partial u_l}{\partial x_i} \right) \quad (13-61)$$

i wyrażenie 13-60 przyjmuje postać

$$dl'^2 = dx_i dx_i + 2u_{ik} dx_i dx_k = (\delta_{ik} + 2u_{ik}) dx_i dx_k. \quad (13-62)$$

Ponieważ zdający sprawę z deformacji tensor u_{ik} jest symetryczny, to istnieje układ własny⁷, w którym tensor deformacji przyjmuje postać diagonalną:

$$\begin{pmatrix} u_{11} & 0 & 0 \\ 0 & u_{22} & 0 \\ 0 & 0 & u_{33} \end{pmatrix} \equiv \begin{pmatrix} u^{(1)} & 0 & 0 \\ 0 & u^{(2)} & 0 \\ 0 & 0 & u^{(3)} \end{pmatrix}. \quad (13-63)$$

⁷Taki układ własny ma charakter lokalny, to znaczy w otoczeniu pewnego punktu ciała możliwe jest przedstawienie w nim tensora u_{ik} w postaci diagonalnej; układ ten nie będzie już układem własnym dla punktów nie należących do bezpośredniego sąsiedztwa wybranego punktu.

W układzie własnym odległość dl' to

$$\begin{aligned} dl'^2 &= (\delta_{ik} + 2u_{ik})dx_i dx_k \\ &= (1 + 2u^{(1)})dx_1^2 + (1 + 2u^{(2)})dx_2^2 + (1 + 2u^{(3)})dx_3^2. \end{aligned} \quad (13-64)$$

Kwadrat nowej odległości to suma trzech analogicznych składników. Możemy zdefiniować *względne wydłużenie* w kierunku osi x_i

$$\frac{dx'_i - dx_i}{dx_i} = \frac{(\sqrt{1 + 2u^{(i)}} - 1)dx_i}{dx_i} = \sqrt{1 + 2u^{(i)}} - 1. \quad (13-65)$$

W tym miejscu możemy zrobić „praktyczne” założenie o małości deformacji – zazwyczaj⁸ wyrazy typu du_i/dx_k są bardzo małe. Pozwala to uprościć definicję 13-61 tensora deformacji

$$u_{ik} \approx \frac{1}{2} \left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i} \right), \quad (13-66)$$

a także posłużyć się przybliżeniem

$$\sqrt{1 + 2u^{(i)}} \approx 1 + u^{(i)}.$$

Względne wydłużenie w kierunku osi x_i będzie po prostu równe

$$\frac{dx'_i - dx_i}{dx_i} = u^{(i)}. \quad (13-67)$$

Co więcej: rozpatrując „starą” (przed deformacją) i „nową” (po deformacji) objętość elementu ciała, mamy

$$\begin{aligned} dV' &= dx'_1 dx'_2 dx'_3 = (1 + u^{(1)})dx_1 (1 + u^{(2)})dx_2 (1 + u^{(3)})dx_3 \\ &\approx (1 + u^{(1)} + u^{(2)} + u^{(3)})dV, \end{aligned} \quad (13-68)$$

(zachowujemy tylko wyrazy pierwszego rzędu w u_{ik}). *Względna zmiana elementu objętości ciała*

$$\frac{dV' - dV}{dV} = u^{(1)} + u^{(2)} + u^{(3)} = u_{ii} \quad (13-69)$$

– jest równa śladowi macierzy tensora deformacji. Przypomnijmy – ślad (kontrakcja) tensora drugiego rzędu jest skalarem, a więc niezmiennikiem – jego wartość jest taka sama we *wszystkich* układach współrzędnych! Co więcej ślad tensora deformacji – obliczany (dla prostoty!) w układzie osi własnych to, zgodnie z definicją 13-66

$$u_{ii} = u_{11} + u_{22} + u_{33} = \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3} \equiv \operatorname{div} \mathbf{u} \quad (13-70)$$

– czyli dywergencja⁹ wektora deformacji.

13.6.2 Tensor naprężeń

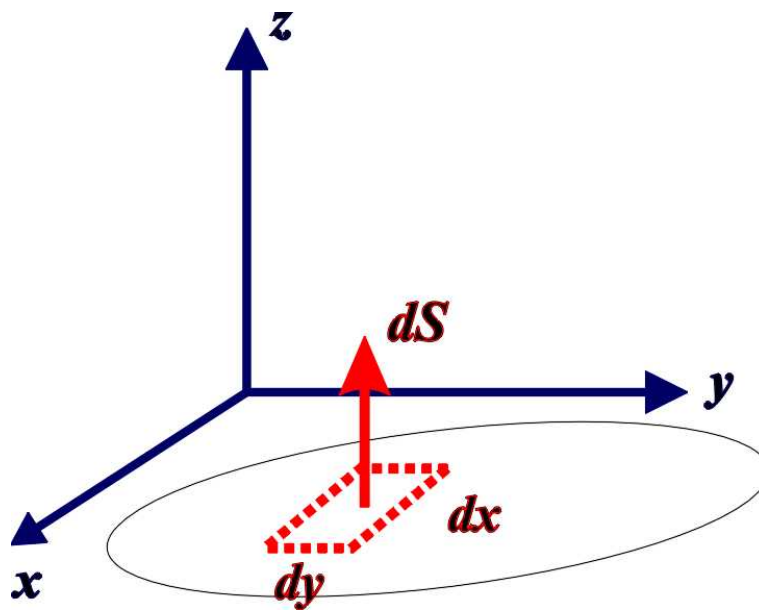
Tak jak tensor deformacji u_{ik} opisuje ilościowo deformację, tak tensor naprężeń stanowi opis jej skutków – sił działających na elementy objętości ciała. Są to tak zwane siły naprężeń – siły wewnętrzne, starające się doprowadzić deformowane ciało do stanu równowagi sprzed deformacji.

Siły naprężeń fizycznie są siłami z jakimi oddziaływują na siebie cząsteczki ciała. Zasięg tych sił jest bardzo niewielki i jest rzędu odległości międzycząsteczkowych. Jeżeli tak, to łatwo pogodzić się z postulatem formalnym: całkowita siła, pochodząca od naprężeń wewnętrznych i działająca na wyróżnioną objętość ciała, powinna dać się wyrazić w postaci całki powierzchniowej (a nie objętościowej), bo siły działające na jakąkolwiek część ciała ze strony otaczających ją cząsteczek działają tylko przez powierzchnię tej części. Ten postulat z kolei, pociąga za sobą następny: siła działająca na nieskończenie mały element powierzchni musi mieć charakter iloczynu tensorowego: tensora pierwszego rzędu (wektora) reprezentującego ten element powierzchni oraz drugiej wielkości tensorowej – tensora drugiego rzędu, zwanego tensorem naprężeń.

Powiedzmy po pierwsze co rozumiemy przez „wektor reprezentujący element powierzchni”. Jest to wektor, którego długość (wartość bezwzględna) jest równa polu powierzchni, a kierunek jest prostopadły (ortogonalny) do tej powierzchni (przy nieskończenie małym elemencie możemy mówić o kierunku prostopadłym do niego jako całości). Takie „zorientowane” powierzchnie bywają w większości przypadków zamknięte (ograniczają pewne objętości) – przyjmujemy konwencję, według której trzeci atrybut tensora – zwrot określamy jako zwrot „na zewnątrz”. Taki skierowany element powierzchni dS_i ($i = 3$) pokazany jest na Rys.13.3. Przy

⁸Odstępstwem od „zwykłych”, tzn. małych deformacji, są na przykład deformacje cienkich prętów i zginanie cienkich płyt.

⁹Porównaj przypis na str. 139.



Rysunek 13.3: Skierowany element powierzchni dS_i ($i = 3$).

takich dezyderatach „formalnych” i -ta składowa siły działająca na element powierzchni

$$F_i = \sigma_{ik} dS_k; \quad (13-71)$$

– sumujemy po wszystkich współrzędnych wektora określającego orientację przestrzenną tego elementu. Wielkość σ_{ik} to – jak wynika z powyższego zapisu – i -ta składowa siły, działająca na *element jednostkowy powierzchni prostopadły do osi $0x_k$* . (Por. Rys.13.4.) Posługując się dość prostą argumentacją fizyczną¹⁰ można wykazać, że tensor naprężeń – σ_{ik} – jest tensorem symetrycznym

$$\sigma_{ik} = \sigma_{ki}. \quad (13-72)$$

Istnieje więc układ osi własnych, w którym σ_{ik} ma postać diagonalną. Aby wyprowadzić związek między dwoma tensorami rozważmy prostą deformację pręta (powierzchnia przekroju S), który ułożony jest wzdłuż osi $0x_1$ i poddany rozciąganiu siłami F (Rys.13.5). Siły działające na jednostkę powierzchni końców pręta to σ_{11} . Pręt ulega względnemu wydłużeniu, którego miarą jest wielkość u_{11} ; jednocześnie wymiary „poprzeczne” (u_{22} i u_{33}) pręta zmniejszają się:

$$\begin{aligned} u_{11} &= \frac{1}{E} \sigma_{11}, \\ u_{22} &= -\frac{\sigma}{E} \sigma_{11}, \\ u_{33} &= -\frac{\sigma}{E} \sigma_{11}. \end{aligned} \quad (13-73)$$

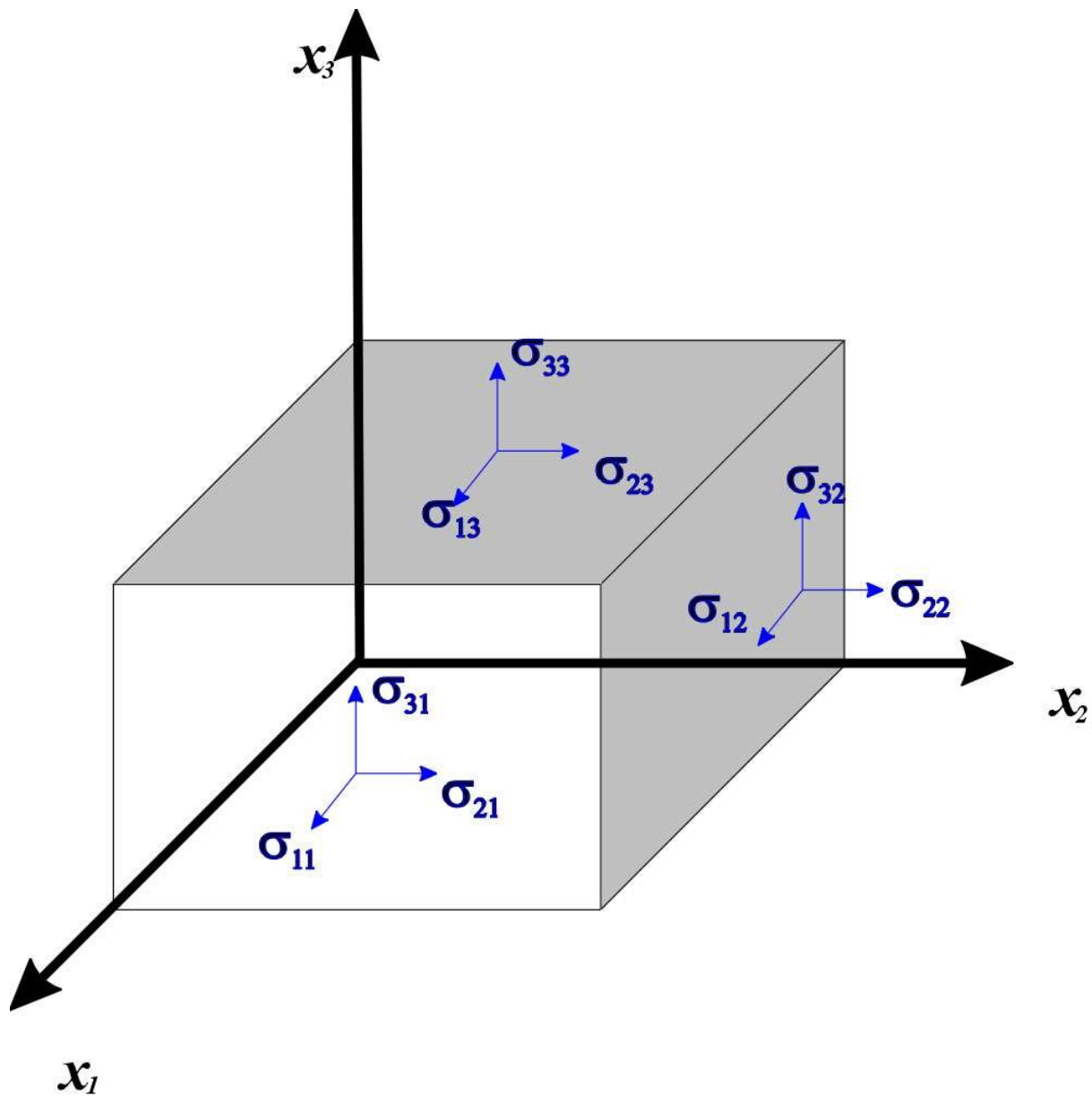
Współczynniki proporcjonalności to moduł Younga (E) i stała Poissona (σ). Gdyby dodatkowo rozciągać (ściskać) pręt w kierunkach prostopadłych do x_1 (siłami σ_{22} i σ_{33}) to pierwsze z równań 13-73 ulega oczywistej modyfikacji

$$u_{11} = \frac{1}{E} \sigma_{11} - \frac{\sigma}{E} \sigma_{22} - \frac{\sigma}{E} \sigma_{33}.$$

Postępując analogicznie z pozostałymi dwoma równaniami przekształcamy 13-73 do postaci bardziej ogólnej i symetrycznej:

$$\begin{aligned} u_{11} &= \frac{1}{E} (1 + \sigma) \sigma_{11} - \frac{\sigma}{E} (\sigma_{11} + \sigma_{22} + \sigma_{33}), \\ u_{22} &= \frac{1}{E} (1 + \sigma) \sigma_{22} - \frac{\sigma}{E} (\sigma_{11} + \sigma_{22} + \sigma_{33}), \\ u_{33} &= \frac{1}{E} (1 + \sigma) \sigma_{33} - \frac{\sigma}{E} (\sigma_{11} + \sigma_{22} + \sigma_{33}). \end{aligned} \quad (13-74)$$

¹⁰Na przykład, postulatem możliwości wyrażenia całkowitego momentu sił działającego na wyróżniony element objętości jako całki po powierzchni zamykającej w sobie tę objętość.



Rysunek 13.4: Tensor naprężeń σ_{ik} .

Występujący w nawiasie ślad tensora naprężeń – σ_{kk} – to pewna stała wielkość liczbowa. Równania 13-74 dają nam już dobre wyobrażenie o zależnościach pomiędzy dwoma tensorami. Byłoby jednak zgrabniej dysponować relacją słuszną w dowolnym układzie odniesienia – nie tylko w układzie osi własnych. Przejście do nowego układu to zwykła transformacja tensora:

$$u_{i'j'} = A_i^{i'} A_j^{j'} u_{ij}, \quad (13-75)$$

$$\sigma_{i'j'} = A_i^{i'} A_j^{j'} \sigma_{ij}. \quad (13-76)$$

Biorąc pod uwagę, że w układzie „nieprimowanym” oba tensory miały postać diagonalną:

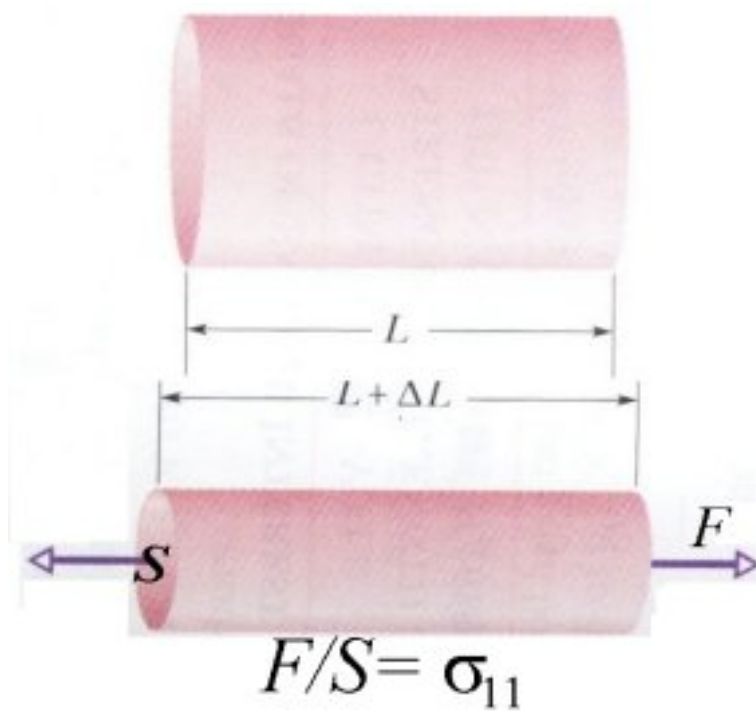
$$u_{ij} = \delta_{ij} u^{(i)}; \quad \sigma_{ij} = \delta_{ij} \sigma^{(i)} \quad (13-77)$$

równania 13-75 i 13-76 przybierają postać

$$u_{i'j'} = A_i^{i'} A_i^{j'} u^{(i)}, \quad (13-78)$$

$$\sigma_{i'j'} = A_i^{i'} A_i^{j'} \sigma^{(i)}. \quad (13-79)$$

Wyjaśnia się, przy okazji, taki a nie inny sposób zapisu składowych diagonalnych tensorów. Sposób ten narzuca nam konwencja Einsteina — ujęty w nawias górny wskaźnik należy traktować specjalnie: bierze on udział w (ewentualnym) sumowaniu, ale nie może — wraz z drugim, identycznym wskaźnikiem — takiego



Rysunek 13.5: Deformacja pręta.

sumowania spowodować. Mnożymy teraz obie strony trzech równań 13-74 odpowiednio: pierwsze przez $A_1^{i'} A_1^{j'}$, drugie – przez $A_2^{i'} A_2^{j'}$, trzecie – przez $A_3^{i'} A_3^{j'}$ i dodajemy stronami. Otrzymujemy

$$A_i^{i'} A_i^{j'} u^{(i)} = \frac{1}{E} (1 + \sigma) A_i^{i'} A_i^{j'} \sigma^{(i)} - \frac{\sigma}{E} (\sigma_{kk}) A_i^{i'} A_i^{j'}. \quad (13-80)$$

Uwzględniając 13-78 i 13-79, a także warunek ortogonalności macierzy współczynników A

$$A_i^{i'} A_i^{j'} = \delta_{i'j'}$$

dostajemy szukane prawo, wiążące składowe obu tensorów

$$u_{i'j'} = \frac{1}{E} (1 + \sigma) \sigma_{i'j'} - \frac{\sigma}{E} (\sigma_{kk}) \delta_{i'j'}. \quad (13-81)$$

W równaniu tym możemy opuścić znaki prim:

$$u_{ij} = \frac{1}{E} (1 + \sigma) \sigma_{ij} - \frac{\sigma}{E} (\sigma_{kk}) \delta_{ij}. \quad (13-82)$$

„Zwyczajowo” przyjęło się raczej wyrażać tensor σ_{ik} przez tensor u_{ik} . Wystarczy dokonać kontrakcji obu stron 13-82:

$$u_{ii} = \frac{1}{E} (1 + \sigma) \sigma_{kk} - 3 \frac{\sigma}{E} (\sigma_{kk}) = \frac{1}{E} (1 - 2\sigma) \sigma_{kk} \quad (13-83)$$

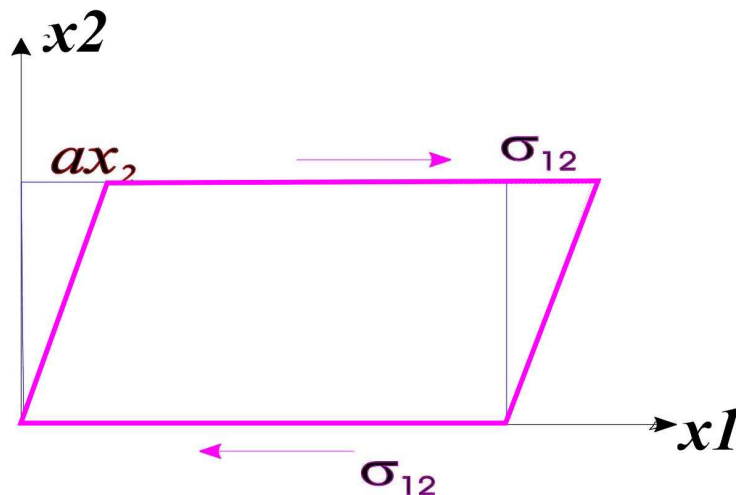
i podstawić z powyższego wzoru za σ_{kk} do 13-82. Otrzymujemy ostatecznie

$$\sigma_{ij} = \frac{E}{1 + \sigma} u_{ij} + \frac{E\sigma}{(1 + \sigma)(1 - 2\sigma)} u_{kk} \delta_{ij}, \quad (13-84)$$

albo – wprowadzając oznaczenia –

$$\lambda = \frac{\sigma E}{(1 + \sigma)(1 - 2\sigma)} \quad \text{i} \quad \mu = \frac{E}{2(1 + \sigma)} \quad (13-85)$$

$$\sigma_{ij} = 2\mu u_{ij} + \lambda u_{kk} \delta_{ij}. \quad (13-86)$$



Rysunek 13.6: Deformacja ścinania.

Dwie nowe stałe – μ i λ – to stałe Lamé'go, mające (podobnie jak stałe E i σ) bezpośrednią interpretację fizyczną. Stała μ to tzw. moduł ścinania – deformacji, której odpowiadają naprężenia wyłącznie *styczne* do powierzchni. Jeżeli (por. Rys.13.6) rozważać tego typu deformację (równoległością z przyłożonym stycznym naprężeniem σ_{12}) to współrzędna wektora deformacji w kierunku osi x_1 jest proporcjonalna do x_2 , $u_1 = ax_2$. Pozostałe współrzędne wektora deformacji są równe zeru: $\mathbf{u} = (ax_2, 0, 0)$. Daje to tensor deformacji o wszystkich $u_{ik} = 0$, za wyjątkiem $u_{12} = u_{21} = \frac{1}{2}a$. Podstawiając do 13-86 mamy

$$\sigma_{12} = 2\mu u_{12} = a\mu; \quad (13-87)$$

wielkość μ jest równa stosunkowi stycznego naprężenia ścinania do miary deformacji – także w kierunku stycznym do powierzchni).

Dla jednorodnego, wszechstronnego ściskania, z ciśnieniem p , tensor naprężeń postać

$$\sigma_{ik} = -p\delta_{ik}, \quad (13-88)$$

a w dodatku

$$u_{11} = u_{22} = u_{33} \quad (13-89)$$

(jeżeli materiał jest jednorodny.) Równanie 13-86 daje natychmiast

$$-p \equiv \sigma_{11} = 2\mu u_{11} + \lambda(3u_{11}) \equiv 3K u_{11}, \quad (13-90)$$

albo

$$3u_{11} = u_{ii} = \frac{\Delta V}{V} = \frac{-p}{K}, \quad (13-91)$$

gdzie

$$K \equiv \lambda + \frac{2}{3}\mu$$

nosi nazwę współczynnika rozszerzalności objętościowej.

13.7 Różniczkowanie tensorów

Ten problem wykracza już poza ramy *algebry* tensorów i należy do *analizy tensorowej*. Ograniczymy się do paru podstawowych pojęć i wzorów, a także pewnych „interpretacji fizycznych”.

Mówimy, że w przestrzeni \mathbb{R}^n jest określone *pole tensorowe* jeżeli każdemu punktowi przestrzeni P przyporządkowany jest tensor określonego rzędu (q), czyli n^q liczb – elementów tensora. Dla tensora drugiego rzędu ($q = 2$) i przestrzeni 3-wymiarowej będzie to 9 liczb; każda z nich będzie funkcją położenia punktu $P \equiv (x_1, x_2, x_3)$:

$$T_{ijk} = T_{ijk}(x_1, x_2, x_3), \quad i, j, k = 1, 2, 3. \quad (13-92)$$

Pole tensorowe rzędu (o walencji) zero to pole skalarne: na przykład temperatura, potencjał pola elektrostatycznego, gęstość ośrodka ciągłego lub bryły sztywnej, ciśnienie, itp. Pole tensorowe pierwszego rzędu to pole *wektorowe*: pole elektrostatyczne, prędkość z jaką przemieszczają się poszczególne cząsteczki płynu (cieczy lub gazu), lub ośrodka sprężystego w którym rozchodzi się fala sprężysta, wektor natężenia prądu w przewodniku, itp.

Przykład pola tensorowego o walencji 2 omówiliśmy w poprzednim punkcie.

Różniczkowanie tensora pola to operacja mająca na celu wyrazić ilościowo zmiany pola tensorowego (poszczególnych elementów T_{ijk} przy przejściu od określonego punktu przestrzeni $P \equiv (x_1, x_2, x_3)$ do punktu znajdującego się w jego bezpośrednim sąsiedztwie $P' \equiv (x_1 + dx_1, x_2 + dx_2, x_3 + dx_3)$). Różniczkę tensora obliczamy jako zwykłą różniczkę funkcji kilku (tu: trzech) zmiennych:

$$dT_{ijk} = \frac{\partial T_{ijk}}{\partial x_l} dx_l \equiv \nabla_l T_{ijk} dx_l. \quad (13-93)$$

Symbol ∇_l z jednej strony upraszcza zapis; z drugiej – sugeruje, że mamy do czynienia z tensorem czwartego rzędu ($3 + 1 = 4$). Rzeczywiście

$$\nabla_{s'} T_{p'q'r'} \equiv \frac{\partial T_{p'q'r'}}{\partial x_{s'}} = \frac{\partial T_{p'q'r'}}{\partial x_l} \frac{\partial x_l}{\partial x_{s'}}. \quad (13-94)$$

Ostatni wyraz to współczynnik $A_{s'}^l$; tensor trzeciego rzędu transformuje się według równania

$$T_{p'q'r'} = A_{p'}^i A_{q'}^j A_{r'}^k. \quad (13-95)$$

Mamy więc

$$\nabla_{s'} T_{p'q'r'} = A_{p'}^i A_{q'}^j A_{r'}^k A_{s'}^l \nabla_l T_{ijk} \quad (13-96)$$

– czyli spełnione jest prawo transformacji dla tensora czwartego rzędu.

13.7.1 Gradient

Gradient to pochodna pola skalarne $\phi(x_1, x_2, x_3)$, a więc tensor pierwszego rzędu – wektor. Współrzędne gradientu określone są wzorem

$$\nabla_i \phi = \phi_i \equiv \frac{\partial \phi}{\partial x_i}, \quad i = 1, 2, 3, \quad (13-97)$$

a w zapisie wektorowym

$$\text{grad } \phi = \nabla \phi = (\nabla_i \phi) e_i. \quad (13-98)$$

Konwencja sumacyjna zostaje tutaj nieco naruszona; jeżeli rozróżniać ko- i kontrawariantne wektory, to w sumę w równ.13-98 przychodzi nam obliczać po wskaźniku i , który jest „dolny” (kowariantny) dla *obu* czynników.

13.7.2 Różniczkowanie pola wektorowego

Z polami wektorowymi (patrz wyżej) spotykamy się bardzo często; pochodna (względem współrzędnych x_i) pola wektorowego oddaje jego zmienność przestrzenną.

Założmy, że nasze pole wektorowe to $\mathbf{a} \equiv \mathbf{a}(P)$ i zgodnie z wywodami tego podrozdziału określmy tensor drugiego rzędu

$$a_{ij} \equiv \nabla_j a_i = \frac{\partial a_i}{\partial x_j}. \quad (13-99)$$

Tak jak każdy tensor drugiego rzędu tensor a_{ik} można rozłożyć na część symetryczną i antysymetryczną (por. 13-49)

$$a_{ij} = \frac{1}{2} (a_{ij} + a_{ji}) + \frac{1}{2} (a_{ij} - a_{ji}) \equiv b_{ij} + c_{ij}, \quad (13-100)$$

gdzie symetryczny i antysymetryczny składnik określone są wzorami

$$b_{ij} = \frac{1}{2} \left(\frac{\partial a_i}{\partial x_j} + \frac{\partial a_j}{\partial x_i} \right), \quad (13-101)$$

$$c_{ij} = \frac{1}{2} \left(\frac{\partial a_i}{\partial x_j} - \frac{\partial a_j}{\partial x_i} \right). \quad (13-102)$$

Obie części możemy zwi z c z podstawowymi operatorami pola wektorowego. Dla cz ści antysymetrycznej mamy – podobnie jak w punkcie 13.5 – trzy niezale ne sk adowe:

$$c_{ij} = \begin{pmatrix} 0 & -c_{12} & c_{13} \\ c_{12} & 0 & -c_{23} \\ -c_{31} & c_{32} & 0 \end{pmatrix} \equiv \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix} \quad (13-103)$$

gdzie wektor $2\mathbf{u} = (2u_1, 2u_2, 2u_3)$ to *rotacja pola wektorowego*, miara jego *wirowo ci*. Zapisana w postaci *explicite* wektorowej rotacja wektora \mathbf{a} to

$$\text{rot } \mathbf{a} = \begin{vmatrix} e_1 & e_2 & e_3 \\ \partial & \partial & \partial \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ a_1 & a_2 & a_3 \end{vmatrix}. \quad (13-104)$$

Zwi zek pomi dzy operatorem rotacji – kt ry, jak wynika z powy szego – polu wektorowemu \mathbf{a} przyporządkowuje inne pole wektorowe $\text{rot } \mathbf{a}$, a *wirowo ci * pola wektorowego \mathbf{a} zilustrujemy w nast pnym punkcie.

Z symetrycznym tensorem, typu b_{ij} spotkali my si  ju  przy omawianiu wst pu do teorii spr ężysto ci. Za-uwa my,  e  lad naszego tensora a_{ij} to

$$a_{ii} = b_{ii} + c_{ii} = b_{ii}, \quad (13-105)$$

bo  lad antysymetrycznego c_{ij} jest r wny zeru (por. 13-103). Z kolei  lad tensora b_{ij}

$$b_{ii} = \frac{\partial a_i}{\partial x_i} \equiv \text{div } \mathbf{a} \quad (13-106)$$

to dywergencja wektora \mathbf{a} – miara  ródlo ci pola wektorowego. Z formalnego punktu widzenia mamy znowu do czynienia z przyporządkowaniem: polu wektorowemu \mathbf{a} pola *skalarne* $\text{div } \mathbf{a}$. Zwi zek pomi dzy operatorem dywergencji a * ródlo ci * pola wektorowego ilustruje znakomicie twierdzenie Gaussa, odnosz ce si  do *całkowitego strumienia pola wektorowego*, przenikaj cego pewn  zamkniet  powierzchnię Σ (por. Rys.13.7). Element strumienia wektora \mathbf{a} przez element powierzchni to iloczyn skalarny: wektora i skierowanego elementu powierzchni elementu powierzchni:

$$d\Phi = \mathbf{a} \cdot d\mathbf{S}, \quad (13-107)$$

a cały strumień to odpowiednia ca ka *powierzchniowa*

$$\Phi = \oint_{\Sigma} \mathbf{a} \cdot d\mathbf{S}. \quad (13-108)$$

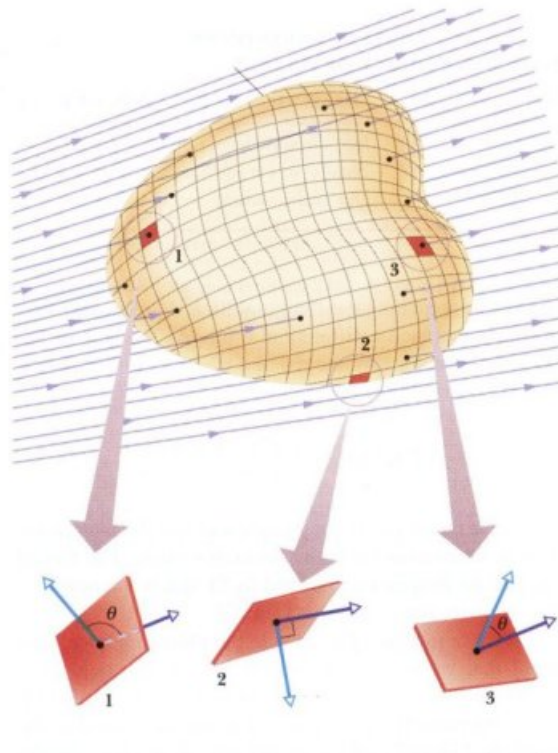
Twierdzenie Gaussa m wi,  e zamiast oblicza  ca k  powierzchniow  z strumienia wektora mo emy obliczy  ca k  obj to ciow  z jego dywergencji:

$$\Phi = \oint_{\Sigma} \mathbf{a} \cdot d\mathbf{S} = \int_V \text{div } \mathbf{a} dV, \quad (13-109)$$

całkuj c po ca ej obj to ci V zamknietej wewn trz Σ . Taka zamkniet  powierzchnia, umieszczona w polu wektorowym \mathbf{a} , mo e by  rozpatrywana w kategoriach „bilansu”: ujemne przyczynki do całkowitego strumienia odpowiadaj  sytuacji kiedy linie wektora wnikaj  do powierzchni (k t pomi dzy \mathbf{a} , a $d\mathbf{S}$ jest wi kszy od $\pi/2$; dodatnie – dla wektora „wychodz cego” na zewn trz (k t mniejszy od $\pi/2$). Je eli „wypadkowy” strumień jest r wny zeru oznacza to,  e zamkniet  w Σ obj to c nie generuje dodatkowych przyczynk w do pola wektorowego \mathbf{a} , ani te  nie „unicestwia” pola – tyle samo wektora wnika do obj to ci, ile ze  wychodzi. Je eli wypadkowy strumień jest r wny od zera oznacza to,  e wewn trz Σ istniej  dodatkowe  ródla pola (ew. „ ródla ujemne”). „Wypadkowa”  ródlo c obj to ci, dla kt rej ca ka z dywergencj  \mathbf{a} jest r wna zeru, jest zerowa; z kolei ca ka *na pewno* b dzie znik c, je eli w ca ej obj to ci tej $\text{div } \mathbf{a} \equiv 0$.

13.8 Niezwykłe przygody kropelki wody

Rozpatrzmy pewien obszar (tr jwymiarowy) Ω , wypełniony ciecz . W obszarze tym mamy pole wektorowe \mathbf{v} – pole pr dko ci małeńkiej cz stki („kropelki”) cieczy, zajmuj cej okre lone poło enie w przestrzeni. Rozpatrujemy tzw. *przepływy stacjonarne*, w kt rych pr dko c kropli cieczy b dzie zale ała od jej poło enia, a

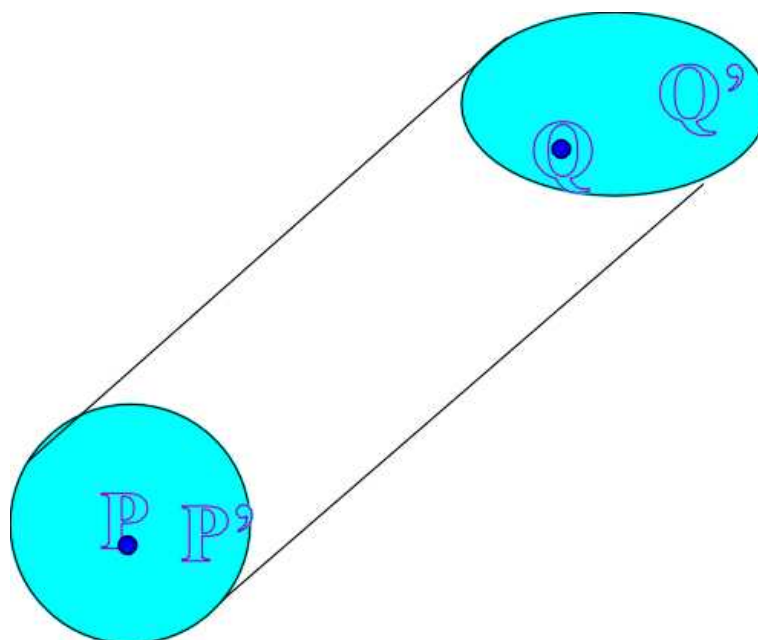


Rysunek 13.7: Ilustracja obliczeń całkowitego strumienia pola wektorowego dla pewnej powierzchni Gaussa.

więc w ogólności będzie zmieniała się od punktu do punktu przestrzeni, ale każda „nowa” cząsteczka cieczy pojawiająca się w określonym punkcie P będzie zachowywała się identycznie, tzn. poruszała się z taką samą prędkością

$$\mathbf{v}(P) = v_i(P)e_i.$$

W momencie początkowym nasza kropelka to sfera, której środek znajduje się w punkcie P ; promień sfery równy jest ρ . Z biegiem czasu kropelka przemieszcza się; jej ruch to złożenie ruchu postępowego (translacji) i obrotowego (rotacji), a ponieważ poszczególne cząsteczki kropli poruszają się nieco inaczej (z nieco innymi prędkościami) to zmienia się – podlega deformacji – początkowy, kulisty kształt kropli – tak jak na rysunku 13.8. Równanie toru kropli to – w pierwszym przybliżeniu – równanie toru, po którym porusza się jej środek.



Rysunek 13.8: Wędrowka kropli.

Jeżeli współrzędne środka kropli oznaczymy jako $x_1(t), x_2(t), x_3(t)$ to wektorowe pole prędkości \mathbf{v} spełnia

równanie różniczkowe:

$$\frac{dx_i}{dt} = v_i, \quad i = 1, 2, 3. \quad (13-110)$$

W przeciągu bardzo małej (praktycznie – nieskończenie małej) chwili czasu Δt środek kropli przesuwa się z punktu P do Q – jego przesunięcie to wektor \overrightarrow{PQ} . Inny punkt kropli – w chwili początkowej zajmujący położenie P' także przesuwa się – do punktu Q' . Mamy (z dokładnością do wyrazów nieskończenie małych pierwszego rzędu)

$$\overrightarrow{PQ} \approx \Delta t \cdot \mathbf{v}(P) \quad (13-111)$$

$$\overrightarrow{P'Q'} \approx \Delta t \cdot \mathbf{v}(P'). \quad (13-112)$$

Wektor $\overrightarrow{PP'}$, łączący dwa wyróżnione punkty kropli w chwili początkowej zmienia się, po czasie Δt , w wektor $\overrightarrow{QQ'}$. Mamy

$$\begin{aligned} \overrightarrow{QQ'} &= \overrightarrow{QP} + \overrightarrow{PP'} + \overrightarrow{P'Q'} \\ &= \overrightarrow{PP'} + (\overrightarrow{P'Q'} - \overrightarrow{PQ}) \\ &\approx \overrightarrow{PP'} + \Delta t[\mathbf{v}(P') - \mathbf{v}(P)]. \end{aligned} \quad (13-113)$$

Występująca w kwadratowym nawiasie różnica prędkości może być, z uwagi na małość naszej kropli, traktowana jako różniczka pola wektorowego prędkości:

$$\mathbf{v}(P') - \mathbf{v}(P) \equiv \Delta \mathbf{v}(P); \quad (13-114)$$

jej i -ta składowa to

$$(\Delta \mathbf{v})_i \equiv \Delta v_i = \frac{\partial v_i}{\partial x_k} dx_k. \quad (13-115)$$

Powyższe wyrażenie poddajemy zabiegowi symetryzacji i antysymetryzacji:

$$\begin{aligned} \frac{\partial v_i}{\partial x_k} dx_k &= \left[\frac{1}{2} \left(\frac{\partial v_i}{\partial x_k} + \frac{\partial v_k}{\partial x_i} \right) + \frac{1}{2} \left(\frac{\partial v_i}{\partial x_k} - \frac{\partial v_k}{\partial x_i} \right) \right] dx_k \\ &\equiv \left[v_{ik}^{(s)} + v_{ik}^{(a)} \right] dx_k. \end{aligned} \quad (13-116)$$

Rozpisując wzór 13-113 na współrzędne, uwzględniając 13-116 oraz kładąc $(\overrightarrow{PP'})_i \equiv dx_i$ otrzymujemy

$$(\overrightarrow{QQ'})_i = [\delta_{ik} + \Delta t \cdot v_{ik}^{(s)} + \Delta t \cdot v_{ik}^{(a)}] dx_k. \quad (13-117)$$

Dwa punkty kropli, zajmujące w chwili początkowej położenia P i P' po upływie czasu Δt zajmują położenia Q i Q' . Transformacja wektora $\overrightarrow{PP'} = dx_i$ w wektor $\overrightarrow{QQ'}$ składa się z trzech członów, którym odpowiadają trzy składniki prawej strony 13-116:

1. $\delta_{ik} dx_k$ – opisuje czysty ruch postępowy (przesunięcie równoległe) wektora $\overrightarrow{PP'}$;
2. $\Delta t v_{ik}^{(s)} dx_k$ – tak jak widzieliśmy to w punkcie 13.6 określa deformację kropli; poszczególne jej cząsteczki poruszają się jednak z nieco różnymi prędkościami i początkowa sfera deformuje się, przyjmując kształt elipsoidy obrotowej. Dla cieczy nieściśliwej objętość takiej elipsoidy równa jest objętości początkowej sfery, a jeżeli tak to $\text{div} \mathbf{v} = v_{ii}^{(s)} = 0$.
3. $\Delta t v_{ik}^{(a)} dx_k$ – ten wyraz powinien być odpowiedzialny za ruch obrotowy kropli. Wyraz ten to pewne przesunięcie (droga); odpowiadająca mu prędkość to $v_{ik}^{(a)} dx_k$. Przyjrzyjmy się tej ostatniej wielkości. W zapisie macierzowym

$$\begin{pmatrix} \Delta v_1 \\ \Delta v_2 \\ \Delta v_3 \end{pmatrix} = \begin{pmatrix} v_{1k}^{(a)} dx_k \\ v_{2k}^{(a)} dx_k \\ v_{3k}^{(a)} dx_k \end{pmatrix}$$

$$\begin{aligned}
&= \begin{pmatrix} 0 & v_{12}^{(a)} & v_{13}^{(a)} \\ v_{21}^{(a)} & 0 & v_{23}^{(a)} \\ v_{31}^{(a)} & v_{32}^{(a)} & 0 \end{pmatrix} \begin{pmatrix} dx_1 \\ dx_2 \\ dx_3 \end{pmatrix} \\
&\equiv \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix} \begin{pmatrix} dx_1 \\ dx_2 \\ dx_3 \end{pmatrix} \\
&\equiv \begin{pmatrix} -u_3 dx_2 + u_2 dx_3 \\ u_3 dx_1 - u_1 dx_3 \\ -u_2 dx_1 + u_1 dx_2 \end{pmatrix} = \begin{pmatrix} (\mathbf{u} \times \mathbf{dx})_1 \\ (\mathbf{u} \times \mathbf{dx})_2 \\ (\mathbf{u} \times \mathbf{dx})_3 \end{pmatrix}, \tag{13-118}
\end{aligned}$$

gdzie wektor \mathbf{u} to

$$\left. \begin{aligned} u_1 &= \frac{1}{2} \left(\frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3} \right) \\ u_2 &= \frac{1}{2} \left(\frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1} \right) \\ u_3 &= \frac{1}{2} \left(\frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \right) \end{aligned} \right\} \tag{13-119}$$

albo (por. 13-104)

$$\mathbf{u} = \frac{1}{2} \text{rot } \mathbf{v}. \tag{13-120}$$

Ten ostatni wzór potwierdza naszą hipotezę: rzeczywiście mamy do czynienia z obrotem kropli wokół swego centrum – wzór 13-119, którego ostatni człon po prawej stronie to iloczyn $\mathbf{u} \times d\mathbf{r}$ jest analogiczny do podstawowego wzoru mechaniki ruchu obrotowego (por. wzór 12-35 w punkcie 12.3) $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$. Prędkością obrotową naszej kropli jest wektor \mathbf{u} , określony wzorem 13-120 – $\frac{1}{2} \times$ rotacja prędkości (ruchu postępowego) kropli.

13.8.1 Operator rotacji a wirowość

Aby ostatnie wzory stały się bardziej czytelne prześledźmy dwa proste przypadki przepływu i możliwych *wirów*. Operator *rotacji* przyporządkowuje określonemu polu wektorowemu – polu prędkości cząsteczek cieczy $\mathbf{v} = \mathbf{v}(x, y, z)$ – inne pole wektorowe. Jeżeli takie pole rotacji jest równe zeru przepływ nazywamy *bezwirowym*. Z definicji rotacji (por. 13-104)

$$\text{rot } \mathbf{v} = \begin{vmatrix} e_1 & e_2 & e_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ v_x & v_y & v_z \end{vmatrix}.$$

wynika, że istotne są tutaj zmiany określonej – np. x -owej – składowej wektora, względem jednego z kierunków ortogonalnych względem osi $0x$. Taką sytuację mamy właśnie na rysunku 13.9, na którym pokazany jest opływ, odbywający się ze stałą prędkością kątową $\boldsymbol{\omega}$ wokół punktu 0. Prędkości liniowe cząsteczek cieczy są równe (por. wzór 12-35)

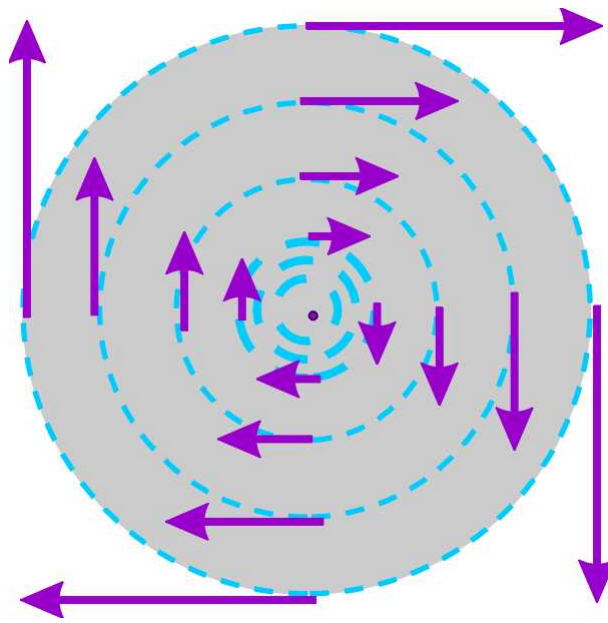
$$\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}, \tag{13-121}$$

gdzie \mathbf{r} to wektor położenia (względem punktu 0). Obliczając powyższy iloczyn wektorowy otrzymujemy

$$\mathbf{v} = \begin{vmatrix} e_1 & e_2 & e_3 \\ 0 & 0 & \omega \\ x & y & z \end{vmatrix} = e_1(-\omega y) + e_2(\omega x), \tag{13-122}$$

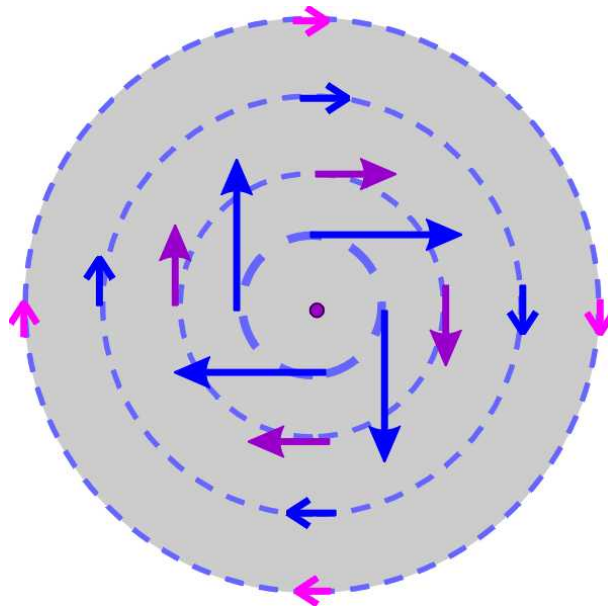
a w takim razie rotacja pola prędkości to po prostu

$$\text{rot } \mathbf{v} = \begin{vmatrix} e_1 & e_2 & e_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ -\omega y & \omega x & 0 \end{vmatrix} = 2\boldsymbol{\omega} \tag{13-123}$$



Rysunek 13.9: Wir prędkości cieczy.

– pole rotacji prędkości to nic innego jak, określający jednoznacznie „okrężny” ruch cieczy (podwojony) wektor $\boldsymbol{\omega}$. (Warto sprawdzić, że dywergencja prędkości jest równa zero – mamy do czynienia z cieczą nieściśliwą.) Z kolei, na rysunku 13.10, pokazany jest opływ, w którym prędkości liniowe cząsteczek płynu maleją w miarę



Rysunek 13.10: Jeszcze jeden wir prędkości cieczy.

oddalania się od punktu 0 (wielkość wektora \boldsymbol{v} jest odwrotnie proporcjonalna do odległości od punktu 0), czyli

$$\boldsymbol{v} = \frac{\hat{\boldsymbol{v}}}{r}, \quad (13-124)$$

gdzie wersor prędkości $\hat{\boldsymbol{v}}$ to

$$\hat{\boldsymbol{v}} = \hat{\boldsymbol{i}} \cos \phi - \hat{\boldsymbol{j}} \sin \phi,$$

(ϕ – kąt między określonym punktem płaszczyzny płynu, a osią x -ów.) Dlatego

$$v_x = \frac{y}{\sqrt{x^2 + y^2}},$$

$$v_y = \frac{-x}{\sqrt{x^2 + y^2}}.$$

Policz sam czytelniku pole wektorowe rotacji takiego przepływu i pole skalarne dywergencji.