

Włodzimierz Salejda
Michał H. Tyc
Marcin Just

ALGEBRAICZNE METODY ROZWIĄZYWANIA RÓWNANIA SCHRÖDINGERA



Instytut Fizyki Politechniki Wrocławskiej
Wrocław 2002

Podręcznik przeznaczony jest dla studentów specjalności *Fizyka komputerowa*, *Fizyka ciała stałego* i podobnych. Omawia metody algebraiczne rozwiązywania stacjonarnego jednowymiarowego równania Schrödingera oraz symetrycznego algebraicznego zagadnienia własnego, jak również przykłady zastosowania opisanych metod w zagadnieniach fizyki ciała stałego i atomowej. Zawiera opis pakietu programów MARRS, służącego do rozwiązywania stacjonarnego jednowymiarowego równania Schrödingera.

Wszystkie zastrzeżone znaki towarowe i firmowe występujące w podręczniku są znakami ich właścicieli.

Fanom fizyki komputerowej

Przedmowa	9
Podziękowania	11
Spis ważniejszych oznaczeń i skrótów	13
1. Wprowadzenie	15
2. Równanie Schrödingera jako algebraiczne zagadnienie własne	21
2.1. Własności stanów związanych. Potencjały wiążące	21
2.2. Algebraiczne zagadnienie własne	22
2.3. Przykłady	24
2.3.1. Oscylator harmoniczny	24
2.3.2. Potencjał Lenarda–Jonesa	26
2.3.3. Potencjał Morse’a	27
2.3.4. Potencjał Konwenta	27
Zadania	27
3. Wyznaczanie wybranych wartości i wektorów własnych trójdzielnej macierzy symetrycznej	29
3.1. Metoda Martina–Deana	29
3.2. Algorytm Martina–Deana	30
3.3. Metody wyznaczania wektorów własnych	31
3.3.1. Proste metody rekurencyjne	32
3.3.2. Metoda DWSZ	32
3.4. Metoda ciągów Sturma	34
Zadania	37
4. Równanie masy efektywnej	39
4.1. Podstawy formalizmu masy efektywnej	39
4.2. Algebraiczne zagadnienie własne	41
Zadania	42
5. Numeryczna analiza równania Schrödingera za pomocą metod różnic skończonych wyższych rzędów	45
5.1. Metoda Guardioli–Rosa	45
5.2. Zagadnienie własne dla macierzy Δ^2	48
5.3. Symetryczne algebraiczne zagadnienie własne w metodzie Guardioli–Rosa	50
5.4. Pięciopunktowe przybliżenie drugiej pochodnej	51
5.5. Metoda Lindberga	52
5.6. Kwazisymetryczna macierz trójdzielna	53
5.7. Ekstrapolacja Richardsona	55
Zadania	57

6.	Metody numeryczne rozwiązywania symetrycznego gęstego algebraicznego zagadnienia własnego	59
6.1.	Elementarne ortogonalne przekształcenia podobieństwa	60
6.1.1.	Obroty Jacobiego i Givensa	60
6.1.2.	Odbicie Householdera	62
6.2.	Metoda QR	63
6.3.	Algorytm QR	65
6.3.1.	Dekompozycja	65
6.3.2.	Rekombinacja	67
6.3.3.	Przesunięcie widma	69
6.3.4.	Niewymierny algorytm QR	70
6.3.5.	Niewymierny algorytm QL	71
6.3.6.	Obliczanie wektorów własnych macierzy trójdzielnej	72
6.3.7.	Wymierne algorytmy QR i QL	72
6.3.8.	Niejawny algorytm QL/QR	74
6.4.	Metoda „dziel i rządź”	76
6.5.	Metoda Householdera	82
6.6.	Zagadnienie własne dla macierzy pasmowej	85
6.7.	Metoda Jacobiego	87
6.8.	Metoda potęgowa	89
6.9.	Metoda iteracji odwrotnych	91
6.10.	Metoda bisekcji i iteracji odwrotnych	93
6.11.	Porównanie wydajności przedstawionych metod	94
6.12.	Zagadnienie własne dla macierzy hermitowskiej	96
	Zadania	97
7.	Zastosowania	99
7.1.	Dynamika sieci łańcuchów atomów	99
7.1.1.	Macierz dynamiczna	99
7.1.2.	Widmo częstości własnych	101
7.2.	Model silnego wiązania	104
7.3.	Dynamika spinowa układów jednowymiarowych	106
7.4.	Zagadnienie radialne	108
7.5.	Atom wodoru w zewnętrznym stałym polu elektrycznym	109
7.5.1.	Przypadek trójwymiarowy	109
7.5.2.	Przypadek dwuwymiarowy	112
7.6.	Kropka kwantowa	114
7.7.	Dynamika elektronów w jednowymiarowych supersieciach w ramach modelu Kroniga–Penneya	115
7.8.	Układy niskowymiarowe w niejednorodnym polu magnetycznym	116
7.8.1.	Supersieć półprzewodnikowa w polu magnetycznym	117
7.8.2.	Dwuwymiarowy gaz elektronowy w niejednorodnym polu magnetycznym	118

Zadania	119
8. Pakiet MARRS	123
8.1. Charakterystyka programu	123
8.2. Graficzny interfejs użytkownika	124
8.3. Schemat wykonywania obliczeń	127
8.4. Dobór algorytmu	127
9. Dokładność i szybkość algorytmów	129
9.1. Wybrane potencjały	129
9.2. Wybrane wyniki obliczeń numerycznych	132
9.3. Analiza dokładności	139
9.3.1. Dokładność operacji zmiennoprzecinkowych	139
9.3.2. Skalowanie wartości własnych	140
9.3.3. Błąd aproksymacji różnicowej drugiej pochodnej	140
9.3.4. Zależność wyników numerycznych od wyboru przedziału całkowania	142
9.4. Dokładność a czas obliczeń	144
10. Niestacjonarne równanie Schrödingera	147
10.1. Ewolucja układu kwantowego	147
10.2. Metody rozwiązywania dyskretnego równania ewolucji	148
Zadania	149
11. Zakończenie	151
A. Problemy do samodzielnego rozwiązania	153
A.1. Zagadnienia jednowymiarowe	153
A.2. Zagadnienia radialne	156
A.3. Zagadnienie jednowymiarowe — porównanie metod	157
A.4. Redukcja Householdera — porównanie podprogramów bibliotecznych	157
A.5. Symetryczne gęste zagadnienie własne — porównanie metod i podprogramów bibliotecznych	158
A.6. Trójdzielne zagadnienie własne — porównanie różnych algorytmów i podprogramów bibliotecznych	158
A.7. Przypadek prawie zwyrodniały (potencjał okresowy)	159
A.8. Przypadek prawie zwyrodniały (potencjał Konwenta)	159
A.9. Trójdzielny układ równań	159
B. Elementy algebry liniowej	161
C. Aproksymacja pochodnych przy pomocy różnic skończonych	169
Zadania	171
D. Klasyczna aproksymacja Padégo	173
Zadania	175
E. Biblioteki podprogramów	177
F. Programy źródłowe	179

Bibliografia	191
Skorowidz	201

Mechanika kwantowa i informatyka to dwie najważniejsze rewolucje naukowe XX wieku [1], które spowodowały spektakularny rozwój techniki komputerowej [2–5]. W dwóch ostatnich dekadach XX wieku przemysł komputerowy odniósł oszałamiający sukces komercyjny¹⁾, zmieniając oblicze naszej cywilizacji, co szczególnie widoczne jest w sferze stosunków społecznych i gospodarczych w krajach wysoko i średnio rozwiniętych. Rosnące wykładniczo, zgodnie z prawami Moore’a [4, 5], moce obliczeniowe²⁾ oraz spadek cen komputerów³⁾ poszerzają ciągle obszary ich zastosowań [5–7]. Jednocześnie wiele wybitnych postaci ze świata nauki wyraża sceptyczne poglądy i snuje pesymistyczne przewidywania dotyczące znaczenia komputerów w przyszłości [8, 9]. Już dzisiaj dostrzega się bariery ograniczające obecny wykładniczy rozwój techniki komputerowej, oparty na zaawansowanych technologiach półprzewodnikowych [4, 7].

Szybko rosnące moce obliczeniowe i powszechna dostępność maszyn cyfrowych oraz postępy w tworzeniu oprogramowania zrewolucjonizowały sposoby rozwiązywania problemów inżynierskich oraz prowadzenie badań w naukach przyrodniczych i technicznych. W ostatnich latach powstała nowa dziedzina fizyki, zwana *fizyką komputerową* [10–16]. Jest to nauka interdyscyplinarna, łącząca fizykę teoretyczną, teorię algorytmów numerycznych [17–27] oraz programowanie. Jej podstawowym narzędziem jest komputer. Warto w tym miejscu zauważyć, że prekursorami fizyki komputerowej byli wybitni matematycy i fizycy zaangażowani w realizację *Projektu Manhattan* [28, 29].

Obszernym działem fizyki komputerowej są metody numeryczne i komputerowe opracowane dla analizowania i rozwiązywania zagadnień mechaniki kwantowej. Podstawowym równaniem nierelatywistycznej mechaniki kwantowej jest niezależne od czasu (stacjonarne) równanie Schrödingera. Analityczne rozwiązania tego równania są znane tylko w szczególnych przypadkach, omawianych szeroko w podręcznikach mechaniki kwantowej i przedstawianych w bieżącej literaturze naukowej [30–40]. W zdecydowanej większości przypadków równanie Schrödingera nie jest rozwiązywalne ściśle. Stosowane są wówczas metody przybliżone lub komputerowe.

W tym podręczniku przedstawiamy algebraiczne metody numerycznego rozwiązywania równania Schrödingera (szczególnie dużo uwagi poświęcimy metodom siatkowym), których istota polega na sprowadzeniu go do *algebraicznego zagadnienia własnego* [41–53]. Uzasadnia to, naszym zdaniem, użytą w tytule i stosowaną w podręczniku terminologię. Prezentowane podejście algebraiczne ma wiele bardzo istotnych

¹⁾ Pierwszy komputer przeznaczony na rynek został wyprodukowany w 1951 roku.

²⁾ Współczesną technikę komputerową charakteryzuje wykładniczy wzrost wydajności mikroprocesorów (mierzonej częstotliwością taktowania mikroprocesora, która obecnie przekracza 1 GHz) oraz pojemności pamięci i nośników danych.

³⁾ Przy rosnących wykładniczo, zgodnie z drugim prawem Moore’a [5], nakładach finansowych na budowę nowych fabryk podzespołów elektronicznych.

zalet (przede wszystkim względną uniwersalność) i nie zostało dotychczas wyczerpująco opisane.

Istotnymi elementami podręcznika są: MARRS⁴⁾ — pakiet programów komputerowych oraz zbiór podprogramów w języku FORTRAN-77, w których zaimplementowano wiele zaawansowanych i efektywnych algorytmów numerycznych.

Jeden z autorów (W. S.) od roku akademickiego 1995/96 prowadzi dla uczestników studiów magisterskich oraz inżynierskich Wydziału Podstawowych Problemów Techniki Politechniki Wrocławskiej wykłady pt. *Metody komputerowe fizyki* oraz *Fizyka komputerowa*. Podręcznik powstał na bazie tych wykładów, które wzbogacili ich byli słuchacze, a obecnie współautorzy: Michał H. Tyc oraz Marcin Just.

Mamy nadzieję, że książka umożliwi zainteresowanym Czytelniczkom i Czytelnikom zdobycie umiejętności posługiwania się najnowszą technologią informatyczną i metodologią fizyki komputerowej oraz pozwoli zapoznać się z aktualnym stanem wiedzy w zakresie zaawansowanych metod i algorytmów numerycznych stosowanych m.in. do rozwiązywania stacjonarnego, jednowymiarowego równania Schrödingera.

Opisany w podręczniku pakiet programów MARRS oraz podprogramy są efektywnymi narzędziami fizyki komputerowej, a także pożytecznymi pomocami dydaktycznymi, które można wykorzystać na wykładach fizyki ogólnej, mechaniki kwantowej, fizyki ciała stałego lub fizyki struktur niskowymiarowych.

Wrocław, maj 2002

W. Salejda

⁴⁾ Metody Algebraiczne Rozwiązywania Równania Schrödingera.

Podziękowania

Autorzy składają podziękowania prof. dr. hab. Tadeuszowi Paszkiewiczowi, pracownikowi naukowo-dydaktycznemu Instytutu Fizyki Uniwersytetu Rzeszowskiego (poprzednio Instytut Fizyki Teoretycznej Uniwersytetu Wrocławskiego) oraz prof. dr. hab. Piotrowi Pierańskiemu, pracownikowi naukowo-dydaktycznemu Wydziału Fizyki Technicznej Politechniki Poznańskiej, za wnikliwe recenzje i cenne uwagi, które przyczyniły się do podniesienia wartości merytorycznej i dydaktycznej podręcznika. Autorzy wyrażają także swoją wdzięczność dr hab. Krystynie Ziętak, prof. nadzw. Politechniki Wrocławskiej, pracownikowi naukowo-dydaktycznemu Instytutu Matematyki Politechniki Wrocławskiej, za przeczytanie manuskryptu i wartościowe uwagi.

Członkom komisji programowych dla specjalności *Fizyka komputerowa* (dzienne studia magisterskie) i *Fizyka komputerowa w nauce i technice* (dzienne studia inżynierskie) na Wydziale Podstawowych Problemów Techniki Politechniki Wrocławskiej: dr. hab. Atoniemu C. Mitusiowi, dr. Jerzemu Peisertowi, dr. inż. Pawłowi Scharochowi i dr. inż. Januszowi Andrzejewskiemu, autorzy wyrażają podziękowania za dyskusje związane z treścią wykładów *Metody komputerowe fizyki* i *Fizyka komputerowa* oraz zawartością niniejszego podręcznika.

Dr. inż. Januszowi Andrzejewskiemu i dr. inż. Pawłowi Scharochowi z Instytutu Fizyki Politechniki Wrocławskiej autorzy dziękują również za rady dotyczące bibliotek procedur numerycznych dostępnych w Internecie.

Jeden z autorów (W. S.) wyraża wdzięczność studentom Wydziału Podstawowych Problemów Techniki Politechniki Wrocławskiej, byłym i obecnym słuchaczom wykładów *Metody komputerowe fizyki* oraz *Fizyka komputerowa*, za ciągłe zachęcanie do wydania wykładów w formie książkowej.

Spis ważniejszych oznaczeń i skrótów

$\mathbf{1}, \mathbf{1}$	wektor jednostkowy, macierz jednostkowa;
$\mathbf{0}, \mathbf{0}$	wektor zerowy, macierz zerowa;
A, B, \dots, Ω	symbole wielkości fizycznych posiadających wymiar;
a, b, \dots, ω	symbole wielkości bezwymiarowych;
$\mathbf{A}, \mathbf{B}, \dots, \mathbf{\Omega}$	wektory kolumnowe posiadające wymiar, $\mathbf{A} = (A_1, \dots, A_n)^T$;
$\mathbf{a}, \mathbf{b}, \dots, \mathbf{\omega}$	wektory kolumnowe bezwymiarowe, $\mathbf{a} = (a_1, \dots, a_n)^T$;
$\mathbf{A}, \mathbf{B}, \dots, \mathbf{\Omega}$	macierze;
$A_{i,j}$	wyraz stojący w i -tym wierszu i j -tej kolumnie macierzy \mathbf{A} ;
$\mathbf{a}^T, \mathbf{A}^T$	wektor transponowany (wierszowy), macierz transponowana;
$\ \mathbf{a}\ _p, \ \mathbf{A}\ _p$	norma p wektora, macierzy;
$\mathbf{a} \cdot \mathbf{b} = \mathbf{a}^T \mathbf{b}$	iloczyn skalarny wektorów \mathbf{a} i \mathbf{b} ;
$\mathbf{a} \otimes \mathbf{b} = \mathbf{a} \mathbf{b}^T$	iloczyn zewnętrzny wektorów \mathbf{a} i \mathbf{b} ;
\mathbb{C}	zbiór liczb zespolonych;
$\mathbb{C}^{n \times m}$	zbiór macierzy zespolonych o n wierszach i m kolumnach;
c^*	liczba zespolona sprzężona z c ;
$D := d/dx$	operator różniczkowania;
$D^2[j/m]$	aproksymanta Padégo rzędu $[j/m]$ operatora drugiej pochodnej;
$\det \mathbf{A} = \mathbf{A} $	wyznacznik macierzy \mathbf{A} ;
$\text{diag}(\mathbf{a})$	macierz diagonalna o wyrazach a_1, \dots, a_n na diagonalu;
\mathcal{E}, ε	energia, bezwymiarowa energia;
$\mathbf{G}_i, \mathbf{G}_i^T$	macierz obrotu Givensa;
\mathbf{L}	macierz trójkątna dolna;
L_c	charakterystyczny dla danego zagadnienia rozmiar liniowy;
M	masa cząstki;
n	liczba punktów siatki, liczba kolumn i wierszy macierzy;
$O(s^k), O(n^k)$	wielkość mała rzędu s^k ($ s \ll 1$), wielkość rzędu n^k ;
\mathbf{Q}	macierz ortogonalna;
\mathbf{R}	macierz trójkątna górna;
\mathbb{R}	zbiór liczb rzeczywistych;
$\mathbb{R}^{n \times m}$	zbiór macierzy rzeczywistych o n wierszach i m kolumnach;
S, s	krok siatki, bezwymiarowy krok siatki;
\mathbf{T}	rzeczywista symetryczna macierz trójdzielna (trójprzekątniowa);
$\text{trid}(\mathbf{b}, \mathbf{a}, \mathbf{c})$	macierz trójdzielna o wyrazach a_1, \dots, a_n na diagonalu, b_1, \dots, b_{n-1} na poddiagonalu i c_1, \dots, c_{n-1} na nadddiagonalu;
$V(X), v(x)$	potencjał o wymiarze energii, bezwymiarowy potencjał;
$v_i = v(x_i)$	wartość potencjału w i -tym punkcie dyskretnej siatki
V_c	charakterystyczna dla danego zagadnienia wartość energii;
X_i, x_i	współrzędna, bezwymiarowa współrzędna i -tego punktu siatki;
$\alpha = 2MV_c L_c^2 / \hbar^2$	bezwymiarowy parametr skali;
Δ_s	operator centralnej różnicy skończonej na siatce o kroku s ;
$\mathbf{\Delta}^2 = \text{trid}[(1, \dots, 1), (-2, \dots, -2), (1, \dots, 1)]$	dyskretna reprezentacja operatora drugiej pochodnej;

$\eta(\mathbf{A})$	liczba ujemnych wartości własnych macierzy \mathbf{A} ;
$\Psi(X)$	funkcja falowa posiadająca wymiar;
$\psi(x)$	bezwymiarowa funkcja falowa;
$\psi_i = \psi(x_i)$	wartość funkcji falowej w i -tym punkcie dyskretnej siatki;
$\boldsymbol{\psi} = (\psi_1, \dots, \psi_n)^T$	wektor określony wartościami funkcji falowej na siatce;
$\lfloor z \rfloor$	nawiększa liczba całkowita nie większa od z (zaokrąglenie w dół);
$\lceil z \rceil$	najmniejsza liczba całkowita nie mniejsza od z (zaokrąglenie w górę).

* * *

AZW	<u>A</u> lgebraiczne <u>Z</u> agadnienie <u>W</u> łasne;
DWSZ	metoda <u>D</u> y, <u>W</u> u, <u>S</u> pratlina i <u>Z</u> henga;
EISPACK	<u>E</u> Igen <u>S</u> ystem <u>P</u> ACKage (nazwa biblioteki podprogramów);
LAPACK	<u>L</u> inear <u>A</u> lgebra <u>P</u> ACKage (nazwa biblioteki podprogramów);
MARRS	<u>M</u> etody <u>A</u> lgebraiczne <u>R</u> ozwiąz y wania <u>R</u> ównania <u>S</u> chrödingera (nazwa pakietu programów);
RS	<u>R</u> ównanie <u>S</u> chrödingera;

Wprowadzenie

W reprezentacji położeń stacjonarne równanie Schrödingera (RS) ma postać zagadnienia własnego dla operatora energii (hamiltonianu) \mathcal{H} [30–40]

$$\mathcal{H}\Psi = \mathcal{E}\Psi. \quad (1.1)$$

Liczbę \mathcal{E} nazywamy *wartością własną* operatora \mathcal{H} (*energiją własną*), a funkcję falową Ψ — *funkcją własną*. W ogólnym przypadku zbiór wartości własnych $\{\mathcal{E}\}$ (*widmo hamiltonianu*) możemy podzielić na dwie części: *widmo dyskretne*, odpowiadające *stanom związanym* opisywanym *normowanymi* funkcjami falowymi, tj. takimi, że $\int |\Psi(\mathbf{R})|^2 d^n\mathbf{R} = 1$, oraz *widmo ciągłe*, odpowiadające *stanom rozproszonym* opisywanym nienormowanymi funkcjami falowymi. W tym podręczniku przedstawiamy metody numeryczne rozwiązywania RS z widmem dyskretnym.

W układach jednowymiarowych równanie (1.1) przyjmuje postać

$$\left[-\frac{\hbar^2}{2M} \frac{d^2}{dX^2} + V(X) \right] \Psi(X) = \mathcal{E}\Psi(X), \quad (1.2)$$

gdzie M oznacza masę cząstki, $\hbar = h/(2\pi)$ — stałą Diraca, X — położenie cząstki, $V(X)$ — potencjał zadany na przedziale $\langle A, B \rangle$.

Analiza numeryczna zagadnienia własnego (1.2) jest prowadzona za pomocą:

- *metod różnic skończonych*, zwanych też *metodami siatkowymi* [54–58], a wśród nich
 - *metod macierzowych*, zwanych także *metodami globalnymi*,
 - *metod strzałów*, zwanych również *metodami Numerowa* [14, 41, 59];
- *metody wariacyjnej* [30–40];
- *metod elementów skończonych*.

Metody te charakteryzujemy poniżej.

Metody macierzowe. Oparte są na aproksymacji drugiej pochodnej za pomocą różnic skończonych [60] wartości $\Psi_i = \Psi(X_i)$ na dyskretnym zbiorze punktów, zwanych dalej siatką, których współrzędne określają związki:

$$X_i = A + i \frac{B-A}{n+1} = A + iS \quad (i = 0, \dots, n+1), \quad (1.3)$$

gdzie $\langle A, B \rangle$ jest przedziałem całkowania równania (1.2). Zbiór liczb Ψ_i (wektor Ψ) stanowi aproksymację funkcji falowej.

Taka dyskretyzacja pozwala sprowadzić RS do postaci algebraicznego zagadnienia własnego [41–53]

$$\mathbf{H}\Psi = \mathcal{E}\Psi$$

z macierzą \mathbf{H} o określonej postaci; zazwyczaj jest to symetryczna macierz pasmowa. O takiej macierzy mówimy, że jest *rozrzedzona*, ponieważ liczba jej niezerowych wyrazów jest rzędu n .

Metody strzałów. W tym podejściu, korzystając z aproksymacji drugiej pochodnej za pomocą różnic skończonych na siatce (1.3), formułujemy jednowymiarowe odwzorowania (w ogólnym przypadku nieliniowe)

$$\Psi_i = \mathcal{F}_{\rightarrow}(\Psi_{i-1}, \dots, \Psi_{i-m}, \mathbf{C}, \mathbf{P}, \mathcal{E}), \quad (1.4)$$

$$\Psi_i = \mathcal{F}_{\leftarrow}(\Psi_{i+1}, \dots, \Psi_{i+m}, \mathbf{C}, \mathbf{P}, \mathcal{E}), \quad (1.5)$$

gdzie $\mathbf{C} = (C_1, \dots, C_c)$ i $\mathbf{P} = (P_1, \dots, P_p)$ oznaczają stałe i parametry modelu, a \mathcal{E} to szukana wartość własna. Obliczenia prowadzimy zazwyczaj według następującego schematu:

1. Dokonujemy próbnego wyboru wartości własnej $\mathcal{E} = \mathcal{E}_0$.
2. Obliczamy wartości Ψ_i funkcji falowej na siatce (1.3) za pomocą dwóch różnych procedur. Jeden ciąg

$$(\Psi_i^{\rightarrow}) = (\Psi_1, \Psi_2, \dots, \Psi_{n-1}, \Psi_n) \quad (1.6)$$

wyznaczamy za pomocą odwzorowania (1.4), rozpoczynając obliczenia od obszaru $\langle A, A + D_A \rangle$, gdzie $D_A \ll B - A$; jest to *procedura wstępująca*. Drugi ciąg

$$(\Psi_i^{\leftarrow}) = (\Psi_n, \Psi_{n-1}, \dots, \Psi_2, \Psi_1) \quad (1.7)$$

wyznaczamy za pomocą odwzorowania (1.5), startując z obszaru $\langle B - D_B, B \rangle$, gdzie $D_B \ll B - A$ (*procedura zstępująca*).

Zauważmy, że do uruchomienia procedur (1.6) i (1.7) wymagana jest znajomość dokładnego lub przybliżonego rozwiązania równania (1.2) na brzegach przedziału całkowania $\langle A, B \rangle$.

3. Sprawdzamy poprawność wyboru wartości własnej \mathcal{E}_0 oraz otrzymanych wartości funkcji falowej za pomocą warunku zszycia w wybranym punkcie siatki X_m . Zazwyczaj żąda się, aby pochodne logarytmiczne funkcji falowej otrzymane za pomocą obu procedur obliczeniowych (1.6) i (1.7) były w punkcie X_m równe z zadaną dokładnością.
4. Kończymy obliczenia, jeśli warunek zszycia jest spełniony. Za wartość własną przyjmujemy wtedy \mathcal{E}_0 , a za wartości odpowiadającej jej funkcji własnej w punktach siatki — obliczone w procesach (1.6) i (1.7) wartości Ψ_i .

W przeciwnym razie obliczenia powtarzamy dla nowej próbnej wartości własnej \mathcal{E}'_0 .

Ten sposób rozwiązywania RS jest stosowany w większości programów dołączonych do podręcznika [40].

Metoda wariacyjna Rayleigha–Ritza–Galerkina. Funkcję falową reprezentujemy jako kombinację liniową n funkcji bazowych, które wybieramy kierując się zasadą prostoty i warunkami symetrii rozpatrywanego zagadnienia [61]. Następnie z zasady wariacyjnej otrzymujemy algebraiczne zagadnienie własne [41–53] z macierzą gęstą, w której liczba wyrazów różnych od zera jest rzędu n^2 .

Funkcję falową Ψ rozwijamy zazwyczaj w ortogonalnej bazie $\{\Phi^{(l)}\}$:

$$\Psi(X) = \sum_{l=1}^n c_l \Phi^{(l)}(X). \quad (1.8)$$

Po podstawieniu skończonego szeregu (1.8) do wyrażenia

$$\langle \Psi | \mathcal{H} | \Psi \rangle = \mathcal{E} \langle \Psi | \Psi \rangle$$

możemy wyznaczyć energię

$$\mathcal{E} = \frac{\langle \Psi | \mathcal{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\sum_{i,l=1}^n c_i^* c_l \langle \Phi^{(i)} | \mathcal{H} | \Phi^{(l)} \rangle}{\sum_{i,l=1}^n c_i^* c_l \langle \Phi^{(i)} | \Phi^{(l)} \rangle} = \frac{\sum_{i,l=1}^n c_i^* c_l H_{i,l}}{\sum_{i,l=1}^n c_i^* c_l S_{i,l}},$$

która jest funkcją nieznanymi współczynników c_l .

Z warunków ekstremum

$$\frac{\partial \mathcal{E}}{\partial c_i} \quad (i = 1, \dots, n)$$

otrzymujemy jednorodny układ równań

$$\sum_{i=1}^n (H_{i,l} - \mathcal{E} S_{i,l}) c_i = 0 \quad (l = 1, \dots, n),$$

który ma niezerowe rozwiązanie dla wartości energii będących rozwiązaniami równania wiekowego

$$\det(\mathbf{H} - \mathcal{E} \mathbf{S}) = 0,$$

gdzie \mathbf{H} oraz \mathbf{S} są macierzami $n \times n$ o wyrazach odpowiednio $H_{i,l}$, $S_{i,l}$. Dodajmy, że jeżeli dokonamy wyboru ortonormalnej bazy funkcji, to $S_{i,l} = \delta_{i,l}$.

Metody elementów skończonych. W tych metodach RS całkujemy (analitycznie lub numerycznie) na skończonej liczbie pododcinków przedziału $\langle A, B \rangle$, a otrzymane rozwiązania zszywamy na granicach podprzedziałów. W tej grupie znajduje się znana i popularna metoda zwana metodą macierzy przejścia¹⁾ [32], w której zazwyczaj przybliżamy potencjał $V(X)$ funkcją przedziałami stałą.

* * *

¹⁾ Ang. *transfer matrix method*.

Podstawową trudnością, jaką napotykamy próbując zastosować do rozwiązania RS metodę strzałów, wariacyjną lub elementów skończonych, jest to, że przed przystąpieniem do obliczeń numerycznych musimy dokonać możliwie trafnego wyboru wartości własnych lub funkcji własnych. Zależy od tego efektywność tego rodzaju metod. W tym sensie nie są one uniwersalne.

W podręczniku pragniemy skupić uwagę Czytelniczek i Czytelników na metodach macierzowych, ponieważ przewyższają pozostałe pod następującymi względami:

- Do uruchomienia obliczeń numerycznych nie jest wymagane określenie na wstępie początkowych wartości własnych lub funkcji własnych.
- W czasie obliczeń numerycznych nie wykonujemy procedury dopasowania.
- Pozwalają wykorzystać zaawansowane i najbardziej efektywne komputerowe algorytmy numerycznej algebry liniowej [41–53, 62, 63].

W zakończeniu wprowadzenia streścimy zawartość podręcznika.

W rozdziale 2 formułujemy stacjonarne jednowymiarowe równanie Schrödingera jako algebraiczne zagadnienie własne. Szczegółowo przedstawiamy oraz ilustrujemy przykładami procedurę dyskretyzacji RS. Pokazujemy sposób sprowadzania RS do postaci algebraicznego zagadnienia własnego z symetryczną macierzą trójdziagonalną, przybliżając operator drugiej pochodnej na dyskretnej siatce punktów za pomocą formuły trójpunktowej.

Rozdział 3 jest poświęcony numerycznym metodom wyznaczania wybranych wartości i wektorów własnych symetrycznej macierzy trójdziagonalnej. Szczególnie dużo miejsca poświęcamy zastosowaniu twierdzenia o ujemnych wartościach własnych macierzy symetrycznej oraz efektywnej metodzie wyznaczania wektorów własnych macierzy trójdziagonalnej, zwanej tu *metodą DWSZ*.

Kolejny rozdział przedstawia uogólnienie zaprezentowanego wcześniej podejścia na przypadek zagadnień opisywanych jednowymiarowym równaniem masy efektywnej [40, 64, 65]. Zaproponowany nowy algorytm pozwala na efektywne wyznaczanie poziomów energetycznych i funkcji falowych elektronów w niskowymiarowych strukturach półprzewodnikowych.

W rozdziale 5 omawiamy najogólniejsze podejście do rozwiązywania stacjonarnego jednowymiarowego RS w ramach metod różnic skończonych, zaproponowane przez R. Guardiolę i J. Rosa [56]. Formalizm ten, przy zastosowaniu dokładniejszych przybliżeń operatora drugiej pochodnej, pozwala sprowadzić RS do algebraicznego zagadnienia własnego z gęstą macierzą symetryczną. Stosunkowo dużo uwagi poświęcamy szczególnemu przypadkowi — metodzie Lindberga, w której mamy do czynienia z kwazisymetryczną macierzą trójdziagonalną.

W rozdziale 6 omawiamy algorytmy numerycznej algebry liniowej, za pomocą których można wyznaczać wszystkie wartości i wektory własne gęstej symetrycznej macierzy rzeczywistej. Najwięcej miejsca poświęcamy najefektywniejszym algorytmom: QR/QL oraz „dziel i rządź” (ang. *divide and conquer*).

Kolejny rozdział zawiera przykłady zastosowań przedstawionych w podręczniku metod i algorytmów numerycznych do badania układów fizycznych będących obiek-

tami zainteresowania fizyki fazy skondensowanej i atomowej.

Opracowany przez dwóch autorów (M.J. i W.S.) i dołączony na płytce CD pakiet MARRS, stworzony na platformie Delphi, jest opisany w rozdziale 8, będącym krótkim przewodnikiem użytkownika.

W następnym rozdziale 9 przedstawione są wyniki obliczeń numerycznych otrzymanych dla wybranych potencjałów za pomocą dołączonych do podręcznika podprogramów. Otrzymane wyniki pozwalają sformułować ważne wnioski dotyczące dokładności, szybkości i zakresu stosowalności omawianych algorytmów numerycznych.

Przedostatni rozdział jest poświęcony odrębnemu zagadnieniu, jakie stanowi niestacjonarne równanie Schrödingera; omawiamy algorytm numeryczny rozwiązywania tego problemu, oparty na metodach macierzowych.

Rozdział 11 zawiera krótkie podsumowanie.

W dodatkach A–F zamieszczone są propozycje projektów do samodzielnej realizacji przez studentów, niezbędne wiadomości z zakresu algebry liniowej, formalizmu różnic skończonych i aproksymant Padégo, wykaz popularniejszych bibliotek procedur numerycznych oraz opis podprogramów, w których zaimplementowano wybrane algorytmy.

Podręcznik zamyka spis literatury źródłowej.

Istotną częścią podręcznika są zamieszczone na końcu rozdziałów zadania do samodzielnego rozwiązania, mające pomóc w zrozumieniu i opanowaniu przedstawionych treści.

Równanie Schrödingera jako algebraiczne zagadnienie własne

W tym rozdziale szczegółowo przedstawiamy metodę sprowadzania stacjonarnego jednowymiarowego równania Schrödingera (1.2) najpierw do dyskretnej postaci bezwymiarowej, a następnie do algebraicznego zagadnienia własnego dla symetrycznej macierzy trójdzielnej — problemu, który będziemy bezpośrednio rozwiązywać numerycznie. Procedurę tę przeprowadzimy dla zagadnienia kwantowego oscylatora harmonicznego, a dodatkowo dla kilku potencjałów znanych z chemii kwantowej i fizyki fazy skondensowanej. Najpierw jednak omówimy krótko własności stanów związanych i warunki ich istnienia [30–40].

2.1. Własności stanów związanych. Potencjały wiążące

Zajmujemy się poszukiwaniem stanów związanych cząstki kwantowej opisanej równaniem Schrödingera (1.2), poruszającej się w obszarze $\langle A, B \rangle$, w którym zadano potencjał w jawnej postaci $V(X)$. Innymi słowy, interesują nas wartości własne \mathcal{E}_j należące do dyskretnego widma energetycznego oraz przynależne do nich zlokalizowane funkcje własne $\Psi^{(j)}$ spełniające warunek unormowania

$$\int_A^B \Psi^*(X) \Psi(X) dX = 1 \quad (2.1)$$

oraz warunki brzegowe

$$\Psi(A) = \Psi(B) = 0. \quad (2.2)$$

Dla uproszczenia zapisu wskaźnik j będziemy odtąd pomijać, o ile nie będzie to prowadzić do niejasności.

Warunkiem koniecznym na to, aby rozpatrywany potencjał był *wiązący*, tj. aby równanie (1.2) posiadało choć jedno rozwiązanie odpowiadające stanowi związanemu, jest to, by funkcja $V(X)$ posiadała lokalne minimum (minima). Potencjały takie nazywamy *jamopodobnymi*. Warunek ten nie jest jednak warunkiem wystarczającym, nawet w przypadku jednowymiarowym. Dokładniejsza analiza [66] prowadzi do następującego wniosku:

Twierdzenie 2-1. Niech $V_-, V_+ > -\infty$ są asymptotycznymi wartościami potencjału odpowiednio dla $X \rightarrow \mp\infty$, a $V_m < \min(V_-, V_+)$ — jego wartością w minimum lokalnym. Wówczas dla $V_- = V_+$ istnieje zawsze co najmniej jeden stan związany, a dla $V_- \neq V_+$ stan związany może istnieć przy spełnieniu dodatkowych warunków.

Równanie Schrödingera rozwiązujemy numerycznie metodą algebraiczną zawsze na *skończonym* przedziale $\langle A, B \rangle$. Przyjęte warunki brzegowe (2.2) fizycznie odpowiadają nieskończonym (nieprzenikalnym) barierom potencjału umieszczonym w punktach $X = A$ i $X = B$. Jak zobaczymy, warunki te będą odgrywać w naszych rozważaniach bardzo istotną rolę. W większości przypadków ten typ warunków brzegowych nie stoi w sprzeczności z rzeczywistym zachowaniem funkcji falowych przy $X \rightarrow A$ i $X \rightarrow B$, o ile tylko przedział całkowania wybierzemy poprawnie.

2.2. Algebraiczne zagadnienie własne

Aby maksymalnie uprościć formuły, przed przystąpieniem do dyskretyzacji równania Schrödingera (1.2), sprowadzimy je do postaci bezwymiarowej. W tym celu przedstawimy potencjał w postaci

$$V(X) = V_c v(X/L_c) = V_c v(x),$$

gdzie $x = X/L_c$ jest bezwymiarową współrzędną, L_c — parametrem o wymiarze długości, określający charakterystyczny rozmiar liniowy układu lub zasięg oddziaływania, natomiast V_c jest charakterystyczną energią¹⁾. Punkty siatki (1.3) mają teraz bezwymiarowe współrzędne

$$x_i = a + is \quad (i = 0, \dots, n+1), \quad (2.3)$$

gdzie bezwymiarowe współrzędne końców przedziału całkowania a i b oraz bezwymiarowy krok s siatki wynoszą

$$a = \frac{A}{L_c}, \quad b = \frac{B}{L_c}, \quad s = \frac{S}{L_c} = \frac{b-a}{n+1}.$$

Wejściowe równanie (1.2) łatwo przekształcamy do postaci

$$\left[-\frac{1}{L_c^2} \frac{d^2}{dx^2} + \frac{2MV_c}{\hbar^2} \frac{V(L_c x)}{V_c} \right] \Psi(L_c x) = \frac{2MV_c}{\hbar^2} \frac{\mathcal{E}}{V_c} \Psi(L_c x).$$

Wprowadzając bezwymiarowy (por. zadanie 2/1) *parametr skali*

$$\alpha = 2MV_c L_c^2 / \hbar^2, \quad (2.4)$$

bezwymiarową funkcję falową²⁾ $\psi(x) = \sqrt{L_c} \Psi(L_c x)$ i bezwymiarową energię własną $\varepsilon = \mathcal{E}/V_c$, otrzymujemy (patrz zadanie 2/2) bezwymiarowe równanie Schrödingera

$$\left[-\frac{d^2}{dx^2} + \alpha v(x) \right] \psi(x) = \alpha \varepsilon \psi(x). \quad (2.5)$$

¹⁾ W rzeczywistości L_c i V_c można wybrać zupełnie dowolnie, np. $L_c = 1$ m i $V_c = 1$ J. Przyjęcie wielkości charakterystycznych dla zagadnienia powoduje jednak, że operujemy dogodnymi w rachunkach liczbami rzędu jedności, co dodatkowo zmniejsza prawdopodobieństwo wystąpienia błędów nadmiaru podczas obliczeń numerycznych.

²⁾ Zgodnie z warunkiem unormowania (2.1), funkcja falowa $\Psi(X)$ ma wymiar $[m^{-1/2}]$.

Przeprowadzimy teraz *dyskretyzację* równania (2.5), przybliżając drugą pochodną funkcji $\psi(x)$ w punktach x_1, \dots, x_n siatki (2.3) z pomocą formuły trójpunktowej (C.7)

$$\begin{aligned} \left. \frac{d^2\psi}{dx^2} \right|_{x=x_i} &= \frac{\psi(x_i+s) - 2\psi(x_i) + \psi(x_i-s)}{s^2} + O(s^2) = \\ &= \frac{\psi_{i+1} - 2\psi_i + \psi_{i-1}}{s^2} + O(s^2). \end{aligned}$$

Użyty tutaj symbol $O(s^2)$ oznacza, że aproksymujemy drugą pochodną w punktach siatki z dokładnością do wyrazów, których wartości są rzędu s^2 . W rezultacie otrzymujemy układ n równań

$$-\psi_{i-1} + (2 + s^2\alpha v_i)\psi_i - \psi_{i+1} = s^2\alpha\varepsilon\psi_i \quad (i = 1, \dots, n), \quad (2.6)$$

gdzie $v_i = v(x_i)$. Jeśli dla uproszczenia zapisu wprowadzimy oznaczenia

$$\tilde{\varepsilon} = s^2\alpha\varepsilon, \quad \tilde{v}_i = s^2\alpha v_i \quad (2.7)$$

i skorzystamy z warunków brzegowych (2.2), które przyjmują postać $\psi_0 = \psi(a) = 0$, $\psi_{n+1} = \psi(b) = 0$, to otrzymamy jednorodny układ równań

$$\begin{cases} (2 + \tilde{v}_1)\psi_1 - \psi_2 & = \tilde{\varepsilon}\psi_1 \\ -\psi_1 + (2 + \tilde{v}_1)\psi_2 - \psi_3 & = \tilde{\varepsilon}\psi_2 \\ \vdots & \vdots \\ -\psi_{i-1} + (2 + \tilde{v}_i)\psi_i - \psi_{i+1} & = \tilde{\varepsilon}\psi_i \\ \vdots & \vdots \\ -\psi_{n-2} + (2 + \tilde{v}_{n-1})\psi_{n-1} - \psi_n & = \tilde{\varepsilon}\psi_{n-1} \\ -\psi_{n-1} + (2 + \tilde{v}_n)\psi_n & = \tilde{\varepsilon}\psi_n, \end{cases}$$

w którym niewiadomymi są zarówno *przeskalowana* energia $\tilde{\varepsilon}$, jak i wartości ψ_i funkcji falowej na siatce. Zapisując powyższy układ równań w formie macierzowej

$$\begin{pmatrix} 2 + \tilde{v}_1 & -1 & & & \\ -1 & 2 + \tilde{v}_2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & 2 + \tilde{v}_{n-1} & -1 \\ & & & -1 & 2 + \tilde{v}_n \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_{n-1} \\ \psi_n \end{pmatrix} = \tilde{\varepsilon} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_{n-1} \\ \psi_n \end{pmatrix}, \quad (2.8)$$

otrzymujemy algebraiczne zagadnienie własne dla *dyskretnego operatora energii* $\tilde{\mathbf{H}}$:

$$\tilde{\mathbf{H}}\boldsymbol{\psi} = \tilde{\varepsilon}\boldsymbol{\psi}, \quad (2.9)$$

którego rozwiązaniami są pary $(\tilde{\varepsilon}, \boldsymbol{\psi})$, gdzie $s^{-2}\tilde{\varepsilon}$ — aproksymacja wartości własnej zagadnienia (2.5), a wektor $\boldsymbol{\psi}$ — *dyskretna* aproksymacja funkcji falowej $\psi(x)$.

W dalszym ciągu będziemy posługiwać się pojęciami wartości własnej i funkcji własnej, mając na myśli ich aproksymacje.

Jak widzimy, macierz $\tilde{\mathbf{H}}$ jest symetryczną trójdagonalną macierzą rzeczywistą, której wyrazy diagonalne określa potencjał $\tilde{v}(x)$:

$$\tilde{\mathbf{H}} = \text{trid}[(-1, \dots, -1), (2 + \tilde{v}_1, \dots, 2 + \tilde{v}_n), (-1, \dots, -1)].$$

W ramach metody siatkowej sprowadziliśmy stacjonarne jednowymiarowe równanie Schrödingera (1.2) do algebraicznego zagadnienia własnego dla symetrycznej macierzy trójdzielnej. Spodziewamy się, że dla rosnącego n (i malejącego s) będzie ono coraz lepszym przybliżeniem wejściowego równania Schrödingera (1.2). Stajemy zatem przed koniecznością rozwiązania metodami numerycznymi AZW dla macierzy trójdzielnej dużych rozmiarów.

Należy tu wspomnieć, że przybliżenie drugiej pochodnej za pomocą różnic skończonych może zawieść w przypadku funkcji wykazujących osobliwości. Znane są zagadnienia własne, w których metoda różnic skończonych w formie tu przedstawionej zawodzi. Typowym przykładem jest dwuwymiarowe zagadnienie radialne³⁾, w którym funkcja falowa stanu podstawowego $\psi^{(0)}(x) \sim \sqrt{x}$ dla $x \rightarrow 0$ (patrz zadania 2/3 i 2/4). W takich przypadkach należy poszukać odpowiednich zmiennych lub zastosować inną metodę rozwiązywania równania Schrödingera.

2.3. Przykłady

Pokażemy teraz, jak doprowadzić stacjonarne jednowymiarowe RS do dyskretnej postaci dla kilku wybranych potencjałów. Warto to wykonać w jawny sposób, z uwagi na konieczność zapisania równania Schrödingera w postaci bezwymiarowej.

Podrozdział ten można pominąć przy pierwszym czytaniu.

2.3.1. Oscylator harmoniczny

Zastosujemy przedstawione powyżej podejście do przypadku ściśle rozwiązywalnego — kwantowego oscylatora harmonicznego. Pozwoli nam to przetestować poprawność procedur komputerowych przez porównanie wyników numerycznych z analitycznymi.

Potencjał oscylatora harmonicznego ma postać

$$V(X) = \frac{1}{2}KX^2 = \frac{1}{2}M\Omega^2X^2,$$

gdzie $\Omega = \sqrt{K/M}$ jest charakterystyczną częstością. Rozwiązaniami RS

$$\left[-\frac{\hbar^2}{2M} \frac{d^2}{dX^2} + \frac{1}{2}M\Omega^2X^2 \right] \Psi(X) = \mathcal{E}\Psi(X) \quad (2.10)$$

są wartości własne energii $\mathcal{E}_j = (1+2j)\mathcal{E}_0$ dla $j = 0, 1, \dots$, gdzie $\mathcal{E}_0 = \frac{1}{2}\hbar\Omega$, a funkcje własne $\Psi^{(j)}(X)$ są funkcjami Hermite'a [30–40].

Z uwagi na parzystość potencjału $V(X) = V(-X)$ wybieramy symetryczny przedział całkowania $\langle -A, A \rangle$, na którym określamy siatkę punktów

$$X_i = -A + \frac{2Ai}{n+1} = -A + iS.$$

³⁾ Np. zagadnienie dwuwymiarowego atomu wodoru Bohra lub, w teorii niskowymiarowych struktur półprzewodnikowych, dwuwymiarowego ekscytanu.

Zastosowanie trójpunktowego przybliżenia (C.7) dla drugiej pochodnej prowadzi do równania

$$-\frac{\hbar^2}{2MS^2}(\Psi_{i-1} - 2\Psi_i + \Psi_{i+1}) + \frac{1}{2}M\Omega^2 X_i^2 \Psi_i = \mathcal{E}\Psi_i \quad (i = 1, \dots, n),$$

a po pomnożeniu obydwu stron powyższego równania przez $2MS^2/\hbar^2$,

$$-\Psi_{i-1} + \left(2 + \frac{M\Omega S^2}{\hbar} \frac{M\Omega X_i^2}{\hbar}\right) \Psi_i - \Psi_{i+1} = \frac{M\Omega S^2}{\hbar} \frac{2\mathcal{E}}{\hbar\Omega} \Psi_i. \quad (2.11)$$

Zauważmy, że wielkość $L_c = \sqrt{\hbar/(M\Omega)}$ ma wymiar długości, a wszystkie ułamki obecne w powyższym równaniu są wielkościami bezwymiarowymi (por. zadanie 2/5).

Dalej będziemy stosować L_c jako jednostkę długości, a występującą w równaniu (2.11) wielkość $\frac{1}{2}\hbar\Omega = \mathcal{E}_0 = V_c$ jako jednostkę energii. Aby otrzymać bezwymiarową dyskretną postać RS (2.10), musimy jeszcze pomnożyć obydwie strony równania (2.11) przez $\sqrt{L_c}$:

$$-\psi_{i-1} + (2 + s^2 v_i) \psi_i - \psi_{i+1} = s^2 \varepsilon \psi_i \quad (i = 1, \dots, n). \quad (2.12)$$

Bezwymiarowa energia równa jest $\varepsilon = \mathcal{E}/\mathcal{E}_0$, potencjał $v_i = x_i^2$, a punkty siatki i jej krok określają wzory

$$x_i = a + is = a \left(\frac{2i}{n+1} - 1 \right), \quad s = \frac{2a}{n+1}.$$

Zauważmy, że $a = A/L_c$ jest bezwymiarową połówkową szerokością przedziału całkowania.

Układ równań (2.12) równoważny jest AZW (2.8), w którym $\tilde{v}_i = s^2 v_i$, $\tilde{\varepsilon} = s^2 \varepsilon$.

W tym miejscu należy zauważyć, że do bezwymiarowego dyskretnego równania (2.12) można również dojść, odgadując postać parametrów L_c i V_c za pomocą analizy wymiarowej. Następnie ze wzoru (2.4) otrzymujemy $\alpha = 1$, a wielkości \tilde{v}_i i $\tilde{\varepsilon}$ z (2.7) podstawiamy do (2.8).

Założmy, że w wyniku numerycznego rozwiązania zagadnienia własnego (2.8) otrzymaliśmy liczby $\tilde{\varepsilon}_j^{\text{num}}(n) = s^2 \varepsilon_j^{\text{num}}(n) = s^2 \mathcal{E}_j^{\text{num}}(n)/\mathcal{E}_0$. Pokażemy teraz, że $\tilde{\varepsilon}_j^{\text{num}}(n) \sim n^{-2}$. Wiedząc z kursu mechaniki kwantowej [30–40], że dokładne wartości energii wynoszą $\mathcal{E}_j = \frac{1}{2}\hbar\Omega(2j+1)$, po prostych przeliczeniach otrzymujemy

$$s^{-2} \tilde{\varepsilon}_j^{\text{num}}(n) = \frac{\mathcal{E}_j^{\text{num}}(n)}{\mathcal{E}_0} \simeq \frac{\mathcal{E}_j}{\mathcal{E}_0} = 2j+1 \quad (j = 0, 1, \dots). \quad (2.13)$$

Spróbujmy oszacować rząd wielkości $\tilde{\varepsilon}^{\text{num}}(n)$. Dla $j = 0$ mamy $s^{-2} \tilde{\varepsilon}_0^{\text{num}}(n) \simeq 1$, więc

$$\tilde{\varepsilon}_j^{\text{num}}(n) \sim s^2 = \left(\frac{2a}{n+1} \right)^2 \ll 1$$

dla ustalonego a , $j \ll n$ i $n \rightarrow \infty$.

O poprawności algorytmu numerycznego (i zapisanych w wybranym języku programowania podprogramów komputerowych) możemy się przekonać w prosty sposób,

tj. sprawdzając przybliżoną równość (2.13), a o ich dokładności będą świadczyć wartości względnych różnic

$$\left| \frac{\mathcal{E}_j^{\text{num}}(n) - \mathcal{E}_j}{\mathcal{E}_0} \right| = \left| \left(\frac{n+1}{2a} \right)^2 \tilde{\varepsilon}_j^{\text{num}}(n) - (2j+1) \right| \quad (j = 0, 1, \dots).$$

Oszacujemy dokładność wyznaczonych numerycznie energii $\mathcal{E}_j^{\text{num}}(n)$. Oznaczmy przez $\tilde{\sigma}_j^{\text{num}}(n)$ dokładność⁴⁾ wyznaczenia wartości $\tilde{\varepsilon}_j^{\text{num}}(n)$. Wtedy błąd $\Sigma_j^{\text{num}}(n)$ obliczonej wartości $\mathcal{E}_j^{\text{num}}(n)$ wynosi

$$\Sigma_j^{\text{num}}(n) = \left(\frac{n+1}{2a} \right)^2 \mathcal{E}_0 \tilde{\sigma}_j^{\text{num}}(n),$$

co oznacza, że wartość błędu numerycznie wyznaczonej wartości własnej $\mathcal{E}_j^{\text{num}}(n)$ rośnie wraz ze wzrostem n jak n^2 . Nakłada to określone ograniczenia na wartość n (tj. na rozmiar macierzy w AZW).

Jak widzimy, gdy używamy formuły trójpunktowej, wzrost n pozwala z jednej strony poprawić dokładność przybliżenia drugiej pochodnej, ale z drugiej strony — co wynika z powyższej analizy — przyczynia się do zmniejszenia dokładności wyznaczanych numerycznie wartości własnych $\mathcal{E}_j^{\text{num}}$. Wnioski te słuszne są dla wszystkich metod różnic skończonych. Dokładniejszą analizę tego problemu, popartą wynikami obliczeń numerycznych, przedstawiamy w rozdziale 9.

2.3.2. Potencjał Lenarda–Jonesa

Potencjał Lenarda–Jonesa, używany do modelowania oddziaływań międzycząsteczkowych, ma postać

$$V(Y) = V_c [(C/Y)^{12} - (D/Y)^6], \quad (2.14)$$

gdzie V_c jest stałą o wymiarze energii, a C i D — stałymi o wymiarze długości⁵⁾. Jedną z tych stałych możemy przyjąć jako jednostkę długości, wprowadzając bezwymiarową współrzędną (na przykład) $y = Y/D$. W bezwymiarowym RS (2.5) występuje wówczas potencjał

$$v(y) = \gamma^{12} y^{-12} - y^{-6}, \quad (2.15)$$

gdzie $\gamma = C/D$ jest bezwymiarowym parametrem, a za jednostkę energii przyjmieśmy V_c . W równaniu (2.5) stała $\alpha = 2MV_c D^2 / \hbar^2$.

Zbadajmy przebieg zmienności funkcji (2.15). Obliczając pochodną dv/dy , znajdujemy minimum w punkcie $y_m = 2^{1/6} \gamma^2$ z wartością $v_m = v(y_m) = -\frac{1}{4} \gamma^{-12}$. Może się okazać wygodne *znormalizowanie* rodziny potencjałów Lenarda–Jonesa względem położenia minimum i głębokości jamy. Przyjmując za bezwymiarowe jednostki długości i energii y_m i $|v_m|$, otrzymujemy

$$v(x) = x^{-12} - 2x^{-6}, \quad (2.16)$$

⁴⁾ Jest ona nie lepsza od dokładności maszynowej (*epsilon maszynowego*), który przy posługiwaniu się zmiennymi typu `REAL*8`, `double` lub równoważnego jest rzędu 10^{-15} .

⁵⁾ Spotyka się też inną postać $V(Y) = V_c(C/Y^{12} - D/Y^6)$, wówczas stałe C i D mają wymiary odpowiednio m^{12} i m^6 .

gdzie $x = y/y_m$ jest nową współrzędną, a stała α wynosi $2^{-2/3}\gamma^{-8}MV_cD^2/\hbar^2$.

2.3.3. Potencjał Morse'a

Potencjał Morse'a używany jest do modelowania oddziaływania atomów w dwuatomowych cząsteczkach. Opisany jest równaniem

$$V(X) = V_c \left[e^{-2CX} - 2e^{-CX} \right], \quad (2.17)$$

gdzie V_c jest stałą o wymiarze energii, $1/C$ — stałą o wymiarze długości. Stałe te można przyjąć za jednostki. Wówczas w bezwymiarowym RS (2.5) obecny jest bezwymiarowy potencjał

$$v(x) = e^{-2x} - 2e^{-x}, \quad (2.18)$$

gdzie $x = CX$ — bezwymiarowa współrzędna, a stała $\alpha = 2MV_c/(C^2\hbar^2)$.

Badając przebieg funkcji (2.18) stwierdzamy, że ma ona minimum w punkcie $x = 0$ i osiąga tam wartość $-V_c$. Innym charakterystycznym punktem, do którego moglibyśmy odnieść jednostkę długości, jest miejsce zerowe $x_z = -\ln 2$. Wartość ta nie zależy od parametru C . Nową zmienną $y = x/x_z$ uzyskuje się przez przeskalowanie o czynnik rzędu jedności, dlatego jej wprowadzenie jest niecelowe.

2.3.4. Potencjał Konwenta

Potencjał Konwenta [67–70] znalazł zastosowanie w modelowaniu wiązań wodorowych. Dany jest wzorem

$$V(X) = V_c(c \cosh DX - 1)^2, \quad (2.19)$$

gdzie, jak poprzednio, V_c jest stałą o wymiarze energii, $1/D$ — stałą o wymiarze długości; $c > 0$ jest bezwymiarowym parametrem decydującym o charakterze potencjału (jednojamowy dla $c \geq 1$; dwujamowy dla $c < 1$).

Ze względu na zmianę charakteru potencjału przy zmianie wartości c trudno jest znaleźć opisujące go uniwersalne parametry (jak położenie minimów, miejsc zerowych, wartości minimalne itp.) inne niż V_c , c , D . Dlatego przyjmiemy V_c za jednostkę energii i $1/D$ za jednostkę długości, co daje bezwymiarowy potencjał

$$v(x) = (c \cosh x - 1)^2, \quad (2.20)$$

gdzie $x = DX$ — nowa współrzędna oraz $\alpha = 2MV_c/(\hbar D)^2$.

Zadania

2/1. Pokazać, że współczynnik α dany wzorem (2.4) jest wielkością bezwymiarową.

2/2. Wyprowadzić równanie (2.5).

2/3. Pokazać bezpośrednim rachunkiem (przy użyciu kalkulatora lub komputera), że trójpunktowe przybliżenie drugiej pochodnej funkcji \sqrt{x} jest numerycznie niepoprawne w granicy $x \rightarrow 0$. Wskazówka: rozważyć punkt $x_1 = s$.

2/4. Znaleźć przykłady innych różniczkowalnych dwukrotnie funkcji, dla których trójpunktowe przybliżenie drugiej pochodnej funkcji jest numerycznie niepoprawne.

2/5. Wyznaczyć wymiar stałej L_c z równania (2.11). Pokazać, że stojące we wzorze (2.11) wyrażenie $S^2 X_i^2 / L_c^4$ jest bezwymiarowe.

Wyznaczanie wybranych wartości i wektorów własnych trójdzielnej macierzy symetrycznej

W tym rozdziale zajmujemy się bliżej niektórymi algorytmami numerycznego rozwiązywania AZW typu (2.8). Zaprezentujemy algorytmy wyznaczania wartości własnych (algorytm Martina–Deana) i wektorów własnych (algorytm DWSZ) trójdzielnej macierzy symetrycznej.

Podstawowe wiadomości z zakresu algebry liniowej, dotyczące rozpatrywanych zagadnień, zamieszczone są w Dodatku B.

3.1. Metoda Martina–Deana

Zastosujemy twierdzenie Martina–Deana (B-17) do macierzy trójdzielnej. Z równania (B.1) wynika, że liczba $l_{\text{ww}}(z)$ wartości własnych macierzy

$$\mathbf{T} = \begin{pmatrix} d_1 & e_1 & & & \\ e_1 & \ddots & \ddots & & \\ & \ddots & \ddots & e_{n-1} & \\ & & e_{n-1} & d_n & \end{pmatrix} \quad (3.1)$$

mniejszych od danej liczby z jest równa liczbie ujemnych wyrazów w ciągu liczbowym

$$u_1 = d_1 - z, \quad u_i = d_i - z - \frac{e_{i-1}^2}{u_{i-1}} \quad (i = 2, \dots, n). \quad (3.2)$$

Jak widzimy, $l_{\text{ww}}(z)$ wynosi

$$l_{\text{ww}}(z) = \eta(\mathbf{T} - z\mathbf{1}) = \sum_{i=1}^n \eta(u_i),$$

gdzie $\eta(\mathbf{A})$ oznacza liczbę ujemnych wartości własnych \mathbf{A} .

Te same wzory można wyprowadzić stosując do macierzy trójdzielnej twierdzenie Sylwestera (B-14)

$$\mathbf{T} - z\mathbf{1} = \begin{pmatrix} d_1 - z & e_1 & & & \\ e_1 & \ddots & \ddots & & \\ & \ddots & \ddots & e_{n-1} & \\ & & e_{n-1} & d_n - z & \end{pmatrix} = \mathbf{L}\mathbf{D}\mathbf{L}^T, \quad (3.3)$$

gdzie \mathbf{L} jest macierzą trójkątną (a ściślej *bidiagonalną*) dolną¹⁾, a $\mathbf{D} = \text{diag}(u_1, \dots, u_n)$ jest macierzą diagonalną²⁾. Z twierdzenia Sylwestera wynika, że bezwładności macierzy $\mathbf{T} - z\mathbf{1}$ i \mathbf{D} są jednakowe.

Ostatnią równość możemy zapisać w postaci

$$\mathbf{T} - z\mathbf{1} = \begin{pmatrix} 1 & & & \\ l_1 & \ddots & & \\ & \ddots & \ddots & \\ & & l_{n-1} & 1 \end{pmatrix} \begin{pmatrix} u_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & u_n \end{pmatrix} \begin{pmatrix} 1 & l_1 & & \\ & \ddots & \ddots & \\ & & \ddots & l_{n-1} \\ & & & 1 \end{pmatrix},$$

gdzie wartości l_i , u_i są nieznanne. Wykonując odpowiednie działania, otrzymujemy związki

$$d_1 - z = u_1, \quad u_{i-1}l_{i-1} = e_{i-1}, \quad d_i - z = l_{i-1}^2 u_{i-1} + u_i \quad (i = 2, \dots, n),$$

a podstawiając drugą równość do trzeciej, znaną już rekurencję

$$u_1 = d_1 - z, \quad u_i = d_i - z - \frac{e_{i-1}^2}{u_{i-1}} \quad (i = 2, \dots, n).$$

Podany algorytm wyznaczania liczby ujemnych wartości własnych symetrycznej macierzy trójdzielnej jest stabilny i wykonywalny nawet przy dzieleniu przez zero³⁾, ponieważ jeśli $u_{i-1} = 0$, to $u_i = -\infty$ i $u_{i+1} = d_{i+1} - z$.

3.2. Algorytm Martina–Deana

Sformułujemy algorytm Martina–Deana, wykorzystujący ciąg (3.2) do znajdowania wartości własnej $\tilde{\epsilon}_j$ trójdzielnej macierzy symetrycznej \mathbf{T} zadaną dokładnością $\tilde{\sigma}$ przy użyciu metody bisekcji.

Algorytm 3-1. (Martina–Deana)

Procedura *Wartość_wł_MD*($\mathbf{d}, \mathbf{e}, j, z_-, z_+, \sigma, \tilde{\epsilon}$):

znajduje j -tą wartość własną $\tilde{\epsilon}_j$ macierzy $\mathbf{T} = \text{trid}(\mathbf{e}, \mathbf{d}, \mathbf{e})$ z dokładnością σ

(z_-, z_+) — wstępne oszacowanie przedziału zawierającego pojedynczą wartość własną

Powtarzaj:

$$z_p := (z_- + z_+)/2;$$

Jeśli $z_+ - z_- \leq \tilde{\sigma}$: $\tilde{\epsilon} := z_p$; Koniec;

Jeśli $\eta(\mathbf{T} - z_p \mathbf{1}) = j - 1$: $z_- := z_p$,

w przeciwnym wypadku $z_+ := z_p$.

Koniec.

¹⁾ Pod pojęciem macierzy trójkątnej dolnej \mathbf{L} rozumiemy macierz, która ma niezerowe wyrazy $l_{i,j}$ na diagonalu i poniżej niej, tj. $l_{i,j} \neq 0$ dla $j \leq i$ oraz $l_{i,j} = 0$ dla $j > i$.

²⁾ W ogólnym przypadku rozkład ma postać $\mathbf{S}(\mathbf{T} - z\mathbf{1})\mathbf{S}^T = \mathbf{L}\mathbf{D}\mathbf{L}^T$, gdzie \mathbf{S} — macierz permutacji. Dla nieosobliwej macierzy trójdzielnej rozkład może być wykonany bez permutacji ($\mathbf{S} = \mathbf{1}$), a algorytm jest numerycznie stabilny [50].

³⁾ O ile komputer będzie stosownie zaprogramowany, tj. zamiast generować błąd dzielenia przez zero, przyjmie $a/0 = \text{sgn } a \cdot \infty$, zgodnie z arytmetyką IEEE [49, 50, 71, 72].

Podprogram BISECT wykorzystujący rozwinięcie powyższego algorytmu zamieszczony jest w Dodatku F.

W jaki sposób wyznaczyć wstępnie przedział $(z_-^{(0)}, z_+^{(0)})$? Możemy posłużyć się twierdzeniami B-7 i B-8. Z drugiego z nich wynika, że wszystkie wartości własne $\tilde{\varepsilon}_j$ macierzy \mathbf{H} występującej w AZW (2.8) należą do przedziału $\langle \min_{i=1}^n \tilde{v}_i, 4 + \max_{i=1}^n \tilde{v}_i \rangle$. Warunek ten pozostaje w zgodzie z faktem, że wartości własne hamiltonianu nie mogą leżeć poniżej globalnego minimum potencjału. Zazwyczaj będziemy się też ograniczać do wartości własnych $\tilde{\varepsilon}_j < \max_{i=1}^n \tilde{v}_i$, ponieważ wartości własne leżące powyżej maksimum potencjału pojawiają się na skutek nałożenia warunków brzegowych (2.2).

Dogodną metodą poszukiwania wszystkich wartości własnych w zadanym przedziale jest *wyszukiwanie binarne* (por. wspomniany podprogram BISECT), polegające na sukcesywnym połowieniu przedziału aż do uzyskania zbioru podprzedziałów $(z_{-,j}^{(0)}, z_{+,j}^{(0)})$ zawierających pojedyncze wartości własne $\tilde{\varepsilon}_j$, tj. spełniających warunki $\eta(\mathbf{T} - z_{-,j}^{(0)} \mathbf{1}) = j - 1$, $\eta(\mathbf{T} - z_{+,j}^{(0)} \mathbf{1}) = j$.

Opisana wyżej metoda bisekcji jest stabilna numerycznie i pozwala wyznaczyć wszystkie wartości własne $\tilde{\varepsilon}_1, \dots, \tilde{\varepsilon}_n$ macierzy \mathbf{T} z zadaną dokładnością. Metoda ta jest dość czasochłonna, zwłaszcza w przypadku dużych macierzy. Jej zbieżność można przyspieszyć, wykorzystując własności wielomianu charakterystycznego macierzy. Będziemy o tym mówić w podrozdziale 3.4.

3.3. Metody wyznaczania wektorów własnych

Ponieważ nieznana jest uniwersalna metoda wyznaczania wektorów własnych macierzy (3.1), poniżej przedstawiamy najprostsze metody rekurencyjne oraz „stosunkowo uniwersalną” metodę zaproponowaną przez Dy, Wu, Spratlina [73] i Zhenga [74], którą będziemy dalej nazywali *metodą DWSZ*.

Założmy, że wyznaczyliśmy z zadaną dokładnością j -tą wartość własną $\tilde{\varepsilon}_j$ algebraicznego zagadnienia własnego (2.8) przy pomocy opisanej uprzednio metody Martina–Deana. W dalszym ciągu rozdziału dla uproszczenia zapisu pominiemy zostanie indeks j . Tym niemniej należy pamiętać, że nasze rozważania dotyczą j -tego wektora własnego odpowiadającego j -tej wartości własnej, tzn. $\boldsymbol{\psi} := \boldsymbol{\psi}^{(j)}$ oraz $\tilde{\varepsilon} := \tilde{\varepsilon}_j$.

Zajmiemy się problemem numerycznego obliczenia współrzędnych⁴⁾ wektora własnego $\boldsymbol{\psi} = (\psi_1, \dots, \psi_n)^T$ AZW postaci

$$\begin{pmatrix} d_1 - \tilde{\varepsilon} & e_1 & & & \\ e_1 & \ddots & \ddots & & \\ & \ddots & \ddots & e_{n-1} & \\ & & e_{n-1} & d_n - \tilde{\varepsilon} & \end{pmatrix} \begin{pmatrix} \psi_1 \\ \vdots \\ \vdots \\ \psi_n \end{pmatrix} = \mathbf{0}. \quad (3.4)$$

⁴⁾ Przypominamy, że $\psi_i = \psi(x_i)$.

3.3.1. Proste metody rekurencyjne

Jeśli założymy $\psi_1 \neq 0$ (możemy przyjąć $\psi_1 = 1$), to pozostałe współrzędne wektora w oparciu o równanie (3.4) można wyrazić poprzez ψ_1 w następujący sposób:

Algorytm 3-2. (Prosta metoda rekurencji do przodu)

Procedura *Wektor_wl_w_przód*($\mathbf{d}, \mathbf{e}, n, \tilde{\epsilon}, \sigma_\psi, \boldsymbol{\psi}, sukces$):

$\psi_1 := 1; \psi_2 := -(d_1 - \tilde{\epsilon})/e_1;$

Dla $i = 3, \dots, n$: $\psi_i := -(e_{i-2}\psi_{i-2} + (d_{i-1} - \tilde{\epsilon})\psi_{i-1})/e_{i-1};$

$\boldsymbol{\psi} := \boldsymbol{\psi}/\|\boldsymbol{\psi}\|_2;$ normowanie wektora

Jeśli $|e_{n-1}\psi_{n-1} + (d_n - \tilde{\epsilon})\psi_n| < \sigma_\psi$: $sukces := \text{TAK}$

w przeciwnym wypadku: $sukces := \text{NIE};$ metoda zawiodła

Koniec.

Liczba σ_ψ jest wymaganą dokładnością obliczeń (gdyby komputer wykonywał obliczenia z nieskończoną dokładnością, mielibyśmy $e_{n-1}\psi_{n-1} - (d_n - \tilde{\epsilon})\psi_n = 0$). W przypadku, gdy ostatnia nierówność nie jest spełniona, musimy wyznaczyć wektor własny inną metodą.

Druga, bardzo podobna, prosta metoda rekurencji polega na założeniu, że $\psi_n \neq 0$:

Algorytm 3-3. (Prosta metoda rekurencji do tyłu)

Procedura *Wektor_wl_w_tyl*($\mathbf{d}, \mathbf{e}, n, \tilde{\epsilon}, \sigma_\psi, \boldsymbol{\psi}, sukces$):

$\psi_n := 1; \psi_{n-1} := -(d_n - \tilde{\epsilon})/e_{n-1};$

Dla $i = n-2, \dots, 1$: $\psi_i := -(e_{i+1}\psi_{i+2} + (d_{i+1} - \tilde{\epsilon})\psi_{i+1})/e_i;$

$\boldsymbol{\psi} := \boldsymbol{\psi}/\|\boldsymbol{\psi}\|_2;$ normowanie wektora

Jeśli $|e_1\psi_2 + (d_1 - \tilde{\epsilon})\psi_1| < \sigma_\psi$: $sukces := \text{TAK}$

w przeciwnym wypadku: $sukces := \text{NIE};$ metoda zawiodła

Koniec.

Przedstawione proste algorytmy rekurencyjne 3-2 i 3-3 nie zawsze pozwalają numerycznie wyznaczyć wektor własny, na co zwrócił uwagę Wilkinson [45]. Poniżej prezentujemy metodę ogólniejszą.

3.3.2. Metoda DWSZ

Przedstawimy obecnie, za pracami [73, 74], zmodyfikowany dla naszych celów, algorytm DWSZ obliczania wektorów własnych macierzy trójdzielnej.

Zapiszmy układ równań (3.4) w następującej postaci:

$$\begin{pmatrix} d_1 - \tilde{\epsilon} & e_1 & & & \\ e_1 & \ddots & \ddots & & \\ & \ddots & \ddots & e_{k-2} & \\ & & e_{k-2} & d_{k-1} - \tilde{\epsilon} & \\ & & & & e_k \end{pmatrix} \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_{k-2} \\ \psi_{k-1} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -e_{k-1}\psi_k \end{pmatrix}, \quad (3.5)$$

$$e_{k-1}\psi_{k-1} + (d_k - \tilde{\epsilon})\psi_k + e_k\psi_{k+1} = 0, \quad (3.6)$$

$$\begin{pmatrix} d_{k+1} - \tilde{\varepsilon} & e_{k+1} & & \\ e_{k+1} & \ddots & \ddots & \\ & \ddots & \ddots & e_{n-1} \\ & & e_{n-1} & d_n - \tilde{\varepsilon} \end{pmatrix} \begin{pmatrix} \psi_{k+1} \\ \psi_{k+2} \\ \vdots \\ \psi_n \end{pmatrix} = \begin{pmatrix} -e_k \psi_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \quad (3.7)$$

Zdefiniujemy $(n-1)$ liczb $\Theta_1^-, \dots, \Theta_{n-1}^-$ oraz $(n-1)$ liczb $\Theta_2^+, \dots, \Theta_n^+$:

$$\Theta_1^- = (d_1 - \tilde{\varepsilon})^{-1}, \quad \Theta_i^- = (d_i - \tilde{\varepsilon} - e_{i-1}^2 \Theta_{i-1}^-)^{-1} \quad (i = 2, \dots, n-1), \quad (3.8)$$

$$\Theta_n^+ = (d_n - \tilde{\varepsilon})^{-1}, \quad \Theta_i^+ = (d_i - \tilde{\varepsilon} - e_i^2 \Theta_{i+1}^+)^{-1} \quad (i = n-1, \dots, 2). \quad (3.9)$$

Przy pomocy ciągów (Θ_i^+) bądź (Θ_i^-) możemy równania (3.5)–(3.7) zapisać na dwa równoważne sposoby:

$$\begin{array}{ll} \psi_2 &= -e_1 \Theta_2^+ \psi_1 & \psi_{n-1} &= -e_{n-1} \Theta_{n-1}^- \psi_n \\ \vdots & & \vdots & \\ \psi_k &= -e_{k-1} \Theta_k^+ \psi_{k-1} & \psi_k &= -e_k \Theta_k^- \psi_{k+1} \\ \hline \psi_{k+1} &= -e_k \Theta_{k+1}^+ \psi_k & \psi_{k-1} &= -e_{k-1} \Theta_{k-1}^- \psi_k \\ \vdots & & \vdots & \\ \psi_n &= -e_{n-1} \Theta_n^+ \psi_{n-1} & \psi_1 &= -e_1 \Theta_1^- \psi_2. \end{array} \quad (3.10)$$

Równania w lewej kolumnie wyprowadzamy kolejno od ostatniego do pierwszego, korzystając z równań (3.5)–(3.7) i definicji Θ_i^+ (3.9). Podobnie równania w prawej kolumnie otrzymujemy od pierwszego do ostatniego, korzystając z definicji Θ_i^- (3.8).

Wzory znajdujące się pod kreską w (3.10) pozwalają sformułować następujący algorytm znajdowania współrzędnych wektora własnego odpowiadającego wyznaczonej wcześniej wartości własnej:

Algorytm 3-4. (DWSZ)

Procedura *Wektor_wł_DWSZ*($\mathbf{d}, \mathbf{e}, n, \tilde{\varepsilon}, \sigma_\psi, \boldsymbol{\psi}$):

$\Theta_1^- = 1/(d_1 - \tilde{\varepsilon})$;

Dla $i = 2, \dots, n-1$: $\Theta_i^- = 1/(d_i - \tilde{\varepsilon} - e_{i-1}^2 \Theta_{i-1}^-)$;

$\Theta_n^+ = 1/(d_n - \tilde{\varepsilon})$;

Dla $i = n-1, \dots, 2$: $\Theta_i^+ = 1/(d_i - \tilde{\varepsilon} - e_i^2 \Theta_{i+1}^+)$;

$k := \text{Wybierz_punkt_pocz}$

1: $\psi_k := 1$;

Dla $i = k+1, \dots, n$: $\psi_i = -e_{i-1} \Theta_i^+ \psi_{i-1}$;

Dla $i = k-1, \dots, 1$: $\psi_i = -e_i \Theta_i^- \psi_{i+1}$;

$\boldsymbol{\psi} := \boldsymbol{\psi} / \|\boldsymbol{\psi}\|_2$; normowanie wektora

$r_k := |e_{k-1} \psi_{k-1} + (d_k - \tilde{\varepsilon}) \psi_k + e_k \psi_{k+1}|$;

Jeśli $r_k > \sigma_\psi$:

$k := \text{Wybierz_inny_punkt_pocz}$;

idź do 1.

Koniec.

Uwagi praktyczne dotyczące algorytmu DWSZ:

1. Jeżeli musimy powtórzyć obliczenia dla nowej wartości k , wygodnie jest znaleźć wskaźnik j_{\max} współrzędnej wyznaczonego wektora ψ , której moduł jest największy. Praktyka pokazuje, że przyjmując $k = j_{\max}$ możemy znacznie zmniejszyć (nawet o kilka rzędów) wartość r_k .
2. Efektywniejszym sposobem wyboru indeksu k jest znalezienie wskaźnika j , dla którego r_j osiąga minimum. Ponieważ

$$r_j = |e_{j-1}\psi_{j-1} + (d_j - \tilde{\varepsilon})\psi_j + e_j\psi_{j+1}| = |-e_{j-1}^2\Theta_{j-1}^- + d_j - \tilde{\varepsilon} - e_j^2\Theta_{j+1}^+|\psi_j,$$

poszukiwanie k można przeprowadzić już na etapie obliczania ciągów (Θ_i^\pm) .

W ten sposób unikamy wielokrotnego wyznaczania współrzędnych wektora ψ .

Omówiony tu algorytm równoważny jest zastosowaniu specyficznego rozkładu macierzy $\mathbf{T} - \tilde{\varepsilon}\mathbf{1}$, zwanego w języku angielskim *twisted factorization*. Rozkład ten wymaga jej przedstawienia w postaci iloczynów \mathbf{LDL}^T oraz \mathbf{UDU}^T . Szczegółowa dyskusja wykorzystania tego rozkładu, łącznie z analizą dokładności, jest przedmiotem prac [75, 76].

Warto podkreślić, że algorytm ten pozwala wyznaczyć pojedynczy wektor własny macierzy \mathbf{T} kosztem $O(n)$ operacji, a wszystkie jej wektory kosztem $O(n^2)$ operacji. Wektor $\psi^{(j)}$ jest poprawnie wyznaczany pod warunkiem, że wartość własna $\tilde{\varepsilon}_j$ jest *izolowana*, tj. nie leży bardzo blisko innych wartości własnych; zadowalającym kryterium jest [76]

$$\frac{|\tilde{\varepsilon}_j^{(\text{num})} - \tilde{\varepsilon}_j|}{\min_{i \neq j} |\tilde{\varepsilon}_i - \tilde{\varepsilon}_j|} \leq \frac{1}{2(n+1)},$$

gdzie $\tilde{\varepsilon}_j^{(\text{num})}$ oznacza numerycznie wyznaczoną wartość własną, a pozostałe symbole — dokładne wartości własne.

3.4. Metoda ciągów Sturma

Przy wyznaczaniu wartości własnych macierzy \mathbf{T} możemy, zamiast z własności ciągu (3.2), skorzystać ze szczególnych własności jej wielomianu charakterystycznego $W(z) = \det(\mathbf{T} - z\mathbf{1})$ stopnia n . Wartości własne $\tilde{\varepsilon}_1, \dots, \tilde{\varepsilon}_n$ macierzy \mathbf{T} są jego miejscami zerowymi.

Dla danej liczby z wartość wielomianu charakterystycznego $W(z)$ możemy obliczyć w stosunkowo prosty sposób, wykorzystując strukturę macierzy \mathbf{T} . Rozwijając wyznacznik względem kolejnych wierszy, otrzymujemy

$$\begin{aligned} p_0(z) &= 1, & p_1(z) &= d_1 - z, \\ p_i(z) &= (d_i - z)p_{i-1}(z) - e_{i-1}^2 p_{i-2}(z) \quad (i = 2, \dots, n) \end{aligned} \tag{3.11}$$

oraz $W(z) = p_n(z)$.

Ciąg (3.11) nosi nazwę *ciągu Sturma*. Ma on następujące właściwości [22, 25, 45]:

1. Jeśli wszystkie $e_i \neq 0$, to miejsca zerowe wielomianu $p_i(z)$ rozdzielają miejsca zerowe wielomianu $p_{i+1}(z)$. Twierdzenie to łatwo jest udowodnić przez spro-

wadzenie do sprzeczności. Z faktu $p_i(z) = p_{i-1}(z) = 0$ wynikałoby, że $p_0 = 0$, co przeczy założeniu.

2. Niech $L(z)$ oznacza liczbę par sąsiednich wyrazów ciągu (3.11), które mają jednakowe znaki; przyjmujemy, że zero ma znak dodatni. Wtedy liczba wartości własnych macierzy \mathbf{T} nie mniejszych od z (tzn. większych lub równych z) wynosi $L(z)$.
3. Niech $N(z)$ oznacza liczbę zmian znaków sąsiednich wyrazów ciągu (3.11). Wtedy liczba wartości własnych macierzy \mathbf{T} mniejszych od z wynosi $N(z)$.
4. Jeśli $p_i(z) = 0$, to $p_{i+1}(z)p_{i-1}(z) < 0$. Jak łatwo zauważyć, w takim przypadku $p_{i+1}(z) = -e_i^2 p_{i-1}(z)$, czyli znaki $p_{i+1}(z)$ i $p_{i-1}(z)$ są przeciwne.

Podane właściwości wykorzystamy do sformułowania odmiennej od metody bisekcji procedury znajdowania wartości własnych \mathbf{T} , zwanej *metodą styczna-sieczna* [17]. Przytoczymy najpierw założenia, przy których można ją stosować.

Niech $W(z)$ wykazuje na przedziale $\langle z_1, z_2 \rangle$ następujące właściwości:

- $W(z)$ zmienia znak, tj. $W(z_1)W(z_2) < 0$.
- $W(z)$ jest monotoniczny, tj. pochodna $W'(z) = \frac{d}{dz}W(z)$ ma stały znak.
- $W(z)$ jest funkcją wypukłą lub wklęsłą, tj. druga pochodna $W''(z) = \frac{d^2}{dz^2}W(z)$ ma stały znak.

Do weryfikacji powyższych założeń będą nam potrzebne efektywne procedury obliczania wartości pierwszej i drugiej pochodnej $W(z)$.

Różniczkując wyrazy ciągu (3.11), otrzymujemy

$$\begin{aligned} p'_0(z) &= 0, & p'_1(z) &= -1, \\ p'_i(z) &= (d_i - z)p'_{i-1}(z) - p_{i-1}(z) - e_{i-1}^2 p'_{i-2}(z) \quad (i = 2, \dots, n) \end{aligned}$$

oraz $W'(z) = p'_n(z)$. Przez ponowne różniczkowanie obliczamy $W''(z) = p''_n(z)$:

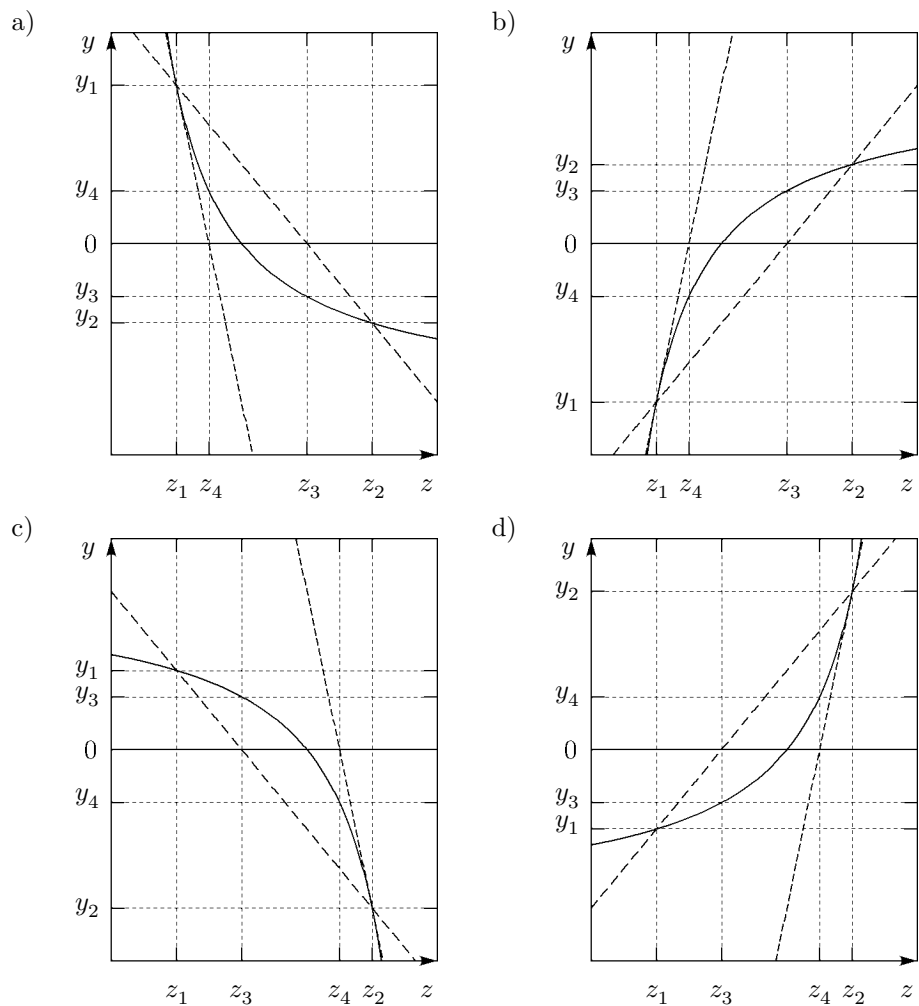
$$\begin{aligned} p''_0(z) &= 0, & p''_1(z) &= 0, \\ p''_i(z) &= (d_i - z)p''_{i-1}(z) - 2p'_{i-1}(z) - e_{i-1}^2 p''_{i-2}(z) \quad (i = 2, \dots, n). \end{aligned}$$

Przedstawimy teraz szczegółowo metodę styczna-sieczna przyjmując, że $W(z)$ spełnia powyższe założenia. Zajmiemy się najpierw przypadkiem przedstawionym na rys. 3.1-a. Niech, dla ustalenia uwagi, $y_1 = W(z_1) > 0$, $y_2 = W(z_2) < 0$ oraz $y_3 = W(z_3) < 0$, gdzie z_3 jest odciętą punktu, w którym prosta przechodząca przez punkty (z_1, y_1) i (z_2, y_2) (czyli *sieczna*) przecina oś OX :

$$z_3 = \frac{y_1 z_2 - y_2 z_1}{y_1 - y_2} = z_2 + \frac{y_2(z_2 - z_1)}{y_1 - y_2}.$$

Obliczmy również odciętą z_4 punktu, w którym styczna do $W(z)$ w punkcie (z_1, y_1) przecina oś OX :

$$z_4 = z_1 - \frac{y_1}{y'_1},$$



Rys. 3.1. Metoda styczna–sieczna dla funkcji wypukłych (a, d) oraz wklęsłych (b, c). Oznaczenia: — funkcja $y(z)$, ---- styczna, --- sieczna

gdzie $y'_1 = W'(z_1)$ jest wartością pochodnej wielomianu charakterystycznego $W(z)$ w punkcie $z = z_1$. Dokonujemy teraz podstawień

$$\begin{aligned} z_1 &:= z_4, & y_1 &:= W(z_4), \\ z_2 &:= z_3, & y_2 &:= W(z_3) \end{aligned}$$

i powtarzamy całą procedurę na nowym przedziale $\langle z_1, z_2 \rangle$. Obliczenia kończymy, gdy $|z_2 - z_1| < \sigma$, gdzie σ jest wymaganą dokładnością wyznaczenia wartości własnej.

Pozostały jeszcze do rozważenia trzy przypadki:

- $W(z_1) < 0$, $W(z_3) > 0$, $W(z_2) > 0$ (rys. 3.1-b),
- $W(z_1) > 0$, $W(z_3) > 0$, $W(z_2) < 0$ (rys. 3.1-c),
- $W(z_1) < 0$, $W(z_3) < 0$, $W(z_2) > 0$ (rys. 3.1-d).

W przypadku pierwszego z nich postępujemy podobnie jak powyżej, co pozostawiamy Czytelniczkom i Czytelnikom do samodzielnego zweryfikowania.

Dwa ostatnie przypadki są także podobne do siebie, dlatego przedstawimy metodę styczna-sieczna jedynie dla $W(z_1) < 0$, $W(z_3) < 0$, $W(z_2) > 0$. Po określeniu przedziału $\langle z_1, z_2 \rangle$ i wartości y_1 , y_2 , y'_2 oraz

$$z_3 = \frac{y_1 z_2 - y_2 z_1}{y_1 - y_2} = z_1 + \frac{y_1(z_2 - z_1)}{y_1 - y_2},$$

łatwo znajdujemy nowy przedział $\langle z_3, z_4 \rangle$ zawierający poszukiwany pierwiastek, gdzie

$$z_4 = z_2 - \frac{y_2}{y'_2}.$$

Dokonując podstawień

$$\begin{aligned} z_2 &:= z_4, & y_2 &:= W(z_4), \\ z_1 &:= z_3, & y_1 &:= W(z_3), \end{aligned}$$

kończymy algorytm.

Zadania

3/1. Wyprowadzić postać ciągu (3.2) w oparciu o twierdzenie Martina–Deana.

3/2. Pokazać, że $\det \mathbf{L} = \det \mathbf{L}^T = 1$ (udowodnić równość $\det(\mathbf{T} - z\mathbf{1}) = \det(\mathbf{D} - z\mathbf{1})$).

3/3. W oparciu o algorytm Martina–Deana napisać program i wyznaczyć za jego pomocą dziesięć najniższych wartości własnych energii oscylatora harmonicznego (patrz paragraf 2.3.1). Wyniki numeryczne porównać z rezultatami analitycznymi.

3/4. Wyznaczyć numerycznie 7 najniższych wartości własnych energii cząstki kwantowej w polu potencjału Lenarda–Jonesa (paragraf 2.3.2).

3/5. Wyznaczyć numerycznie 8 najniższych wartości własnych energii cząstki kwantowej w polu potencjału Morse’a (paragraf 2.3.3). Wyniki numeryczne porównać z rezultatami analitycznymi [30] (zob. też rozdział 9).

3/6. Wyznaczyć numerycznie 6 najniższych wartości własnych energii cząstki kwantowej w polu potencjału Konwenta (paragraf 2.3.4). Wyniki numeryczne porównać z rezultatami analitycznymi [67–70] (zob. też rozdział 9).

3/7. Wyjaśnić, dlaczego warunki poprawności otrzymanego wektora własnego w algorytmach 3-2, 3-3 i 3-4 sprawdzamy po uprzednim unormowaniu wektora.

3/8. Wyprowadzić wzory (3.10).

Równanie masy efektywnej

Własności fizyczne niskowymiarowych układów półprzewodnikowych, takich jak druty i studnie kwantowe czy supersieci, są często badane w przybliżeniu masy efektywnej [40, 64, 65], który pokrótce przedstawimy poniżej. Pokażemy dalej, że równanie masy efektywnej można sprowadzić do algebraicznego zagadnienia własnego z symetryczną macierzą trójdziagonalną.

4.1. Podstawy formalizmu masy efektywnej

Rozważmy kryształ, w którym prawo dyspersji elektronów (lub dziur) jest paraboliczne. Zależność energii \mathcal{E} elektronu od wektora falowego \mathbf{Q} w ogólnym przypadku jest anizotropowa:

$$\mathcal{E}(\mathbf{Q}) = \frac{\hbar^2}{2} \mathbf{Q} \cdot \mathbf{M}^{-1} \mathbf{Q} = \frac{\hbar^2}{2} \sum_{i,j} W_{i,j} Q_i Q_j \quad (i, j = x, y, z), \quad (4.1)$$

gdzie \mathbf{M} jest tensorem masy efektywnej; dla uproszczenia zapisu wprowadziliśmy tensor odwrotności masy efektywnej $\mathbf{W} = \mathbf{M}^{-1}$, reprezentowanego w wybranym układzie współrzędnych przez symetryczną macierz 3×3 o wyrazach $W_{i,j}$. Nie będziemy tu rozważać praw dyspersji wyższych rzędów ani innych, bardziej skomplikowanych wyrażań dla energii kinetycznej [77, 78].

Jeżeli elektron znajduje się w pobliżu dna pasma (tzn. jego energia nie jest zbyt duża), możemy jego funkcję falową $\Phi(\mathbf{R})$ przedstawić w postaci paczki falowej [79]

$$\Phi(\mathbf{R}) = \sum_{\mathbf{Q}} c_{\mathbf{Q}} e^{i\mathbf{Q} \cdot \mathbf{R}} U_0(\mathbf{R}) = U_0(\mathbf{R}) \Psi(\mathbf{R}),$$

gdzie U_0 jest funkcją Blocha na dnie pasma, a Ψ funkcją falową obwiedni¹⁾. Jeżeli zewnętrzny potencjał V zmienia się mało na odległościach porównywalnych z rozmiarem komórki elementarnej kryształu i jest zbyt słaby, by powodować przejścia do innych pasm energetycznych, funkcja obwiedni spełnia stacjonarne²⁾ równanie Schrödingera

$$[\mathcal{E}(-i\nabla) + V(\mathbf{R})] \Psi(\mathbf{R}) = \mathcal{E} \Psi(\mathbf{R}),$$

gdzie $\mathcal{E}(-i\nabla)$ — operator utworzony z $\mathcal{E}(\mathbf{Q})$ przez formalne podstawienie operatora reprezentującego \mathbf{Q} . Wykorzystując zależność (4.1), możemy napisać ogólną postać równania masy efektywnej:

$$\left[-\frac{\hbar^2}{2} \nabla \mathbf{W}(\mathbf{R}) \nabla + V(\mathbf{R}) \right] \Psi(\mathbf{R}) = \mathcal{E} \Psi(\mathbf{R}) \quad (4.2)$$

¹⁾ Ang. *envelope function*.

²⁾ Przypadek niestacjonarny można rozpatrywać w analogiczny sposób.

albo, równoważnie,

$$\left[-\frac{\hbar^2}{2} \sum_{i,j} \frac{\partial}{\partial R_i} W_{i,j}(\mathbf{R}) \frac{\partial}{\partial R_j} + V(\mathbf{R}) \right] \Psi(\mathbf{R}) = \mathcal{E} \Psi(\mathbf{R}).$$

W równaniu tym tensor odwrotności masy efektywnej może być funkcją położenia. Odpowiada to przestrzennej zmienności własności materiału tworzącego układ, np. składu chemicznego supersieci.

W dalszym ciągu będziemy rozpatrywać tylko jednowymiarowy przypadek równania (4.2); wtedy

$$\left[-\frac{\hbar^2}{2} \frac{d}{dX} W(X) \frac{d}{dX} + V(X) \right] \Psi(X) = \mathcal{E} \Psi(X). \quad (4.3)$$

W przypadku, gdy masa efektywna jest przedziałami stała (np. gdy stykają się ze sobą w punkcie X_0 różne materiały A i B)

$$W(X) = \begin{cases} W_A & (X < X_0) \\ W_B & (X > X_0), \end{cases}$$

możemy na sąsiadujących przedziałach osobno rozwiązać „zwykłe” równania Schrödingera dla cząstek o masach $M_i = 1/W_i$ ($i = A, B$):

$$\left[-\frac{\hbar^2}{2M_i} \frac{d^2}{dX^2} + V(X) \right] \psi_i(X) = \mathcal{E} \psi_i(X),$$

a następnie zszyć rozwiązania, nakładając odpowiednie warunki na funkcję obwiedni Ψ i jej pochodną $\frac{d\Psi}{dX}$. Prowadzi to do różnych odmian opisanej pokrótce we wstępie *metody macierzy przejścia*, którą nie będziemy się tu szczegółowo zajmować. W literaturze dyskutowane są różne postaci tych warunków [80–83]. Najprostsze i najczęściej używane to

$$\Psi_A(X_0-) = \Psi_B(X_0+), \quad W_A \frac{d\psi_A}{dX} \Big|_{X_0-} = W_B \frac{d\psi_B}{dX} \Big|_{X_0+}, \quad (4.4)$$

tj. funkcje Ψ i $W \frac{d\Psi}{dX}$ są ciągłe. Łatwo się przekonać, że warunki te prowadzą do zachowania ciągłości strumienia prawdopodobieństwa

$$J = \frac{\hbar}{2iM} \left(\Psi^* \frac{d\Psi}{dX} - \Psi \frac{d\Psi^*}{dX} \right).$$

Należy jednak pamiętać, że warunki (4.4) nie są dokładne, ponieważ w rzeczywistości ciągła musi być pełna funkcja falowa elektronu $U_Q \Psi$, tj. funkcje U_Q musiałyby być jednakowe w obu materiałach. Nawet dla podobnych do siebie pod względem składu chemicznego i struktury krystalicznej materiałów jest to tylko przybliżenie.

4.2. Algebraiczne zagadnienie własne

Omówimy teraz sposób rozwiązywania jednowymiarowego równania masy efektywnej (4.3) dla dowolnej funkcji $M(X)$. Postępując podobnie jak w podrozdziale 2.2, zapisujemy wyjściowe równanie w postaci bezwymiarowej, analogicznej do (2.5),

$$\left[-\frac{d}{dx}w(x)\frac{d}{dx} + \alpha v(x) \right] \psi(x) = \alpha \varepsilon \psi(x), \quad (4.5)$$

gdzie x , w , v , ε są odpowiednio bezwymiarowym położeniem, odwrotną masą, potencjałem i energią, wyrażonymi w jednostkach L_c , $1/M_c$ i V_c , a $\psi(x) = \sqrt{L_c}\Psi(L_c x)$ jest bezwymiarową funkcją falową. Stała α ma wartość $2M_c V_c L_c^2 / \hbar^2$.

Aby otrzymać dyskretną postać równania (4.5), przybliżymy złożony operator różniczkowy $\frac{d}{dx}w(x)\frac{d}{dx}$ za pomocą różnic skończonych. Poniżej opiszemy najprostsze spośród możliwych podejść [55], dające dokładność $O(s^2)$ i sprowadzające problem do symetrycznego trójdziagonalnego zagadnienia własnego.

Zdefiniujmy $f(x) = w(x)\frac{d}{dx}\psi(x)$. Równanie (4.5) zapiszemy w postaci

$$-\frac{d}{dx}f(x) + \alpha v(x)\psi(x) = \alpha \varepsilon \psi(x). \quad (4.6)$$

Wprowadźmy na siatce pomocnicze punkty o współrzędnych

$$x_{i\pm 1/2} = a + (i + \frac{1}{2})s. \quad (4.7)$$

Będziemy przybliżać wartości $f(x)$ w punktach (4.7) na siatce o kroku $\frac{1}{2}s$, stosując różnice centralne (C.3). Otrzymamy

$$f_{i\pm 1/2} = w \frac{d\psi}{dx} \Big|_{x_{i\pm 1/2}} = \pm w_{i\pm 1/2} \frac{\psi_{i\pm 1} - \psi_i}{s} + O(s^2).$$

W celu przybliżenia pochodnej $\frac{df}{dx}$ w punktach o całkowitych wskaźnikach, zastosujmy wzór (C.3) na różnicę centralną dla siatki o kroku $\frac{1}{2}s$:

$$\begin{aligned} \frac{df}{dx} \Big|_{x_i} &= \frac{f_{i+1/2} - f_{i-1/2}}{s} + O(s^2) = \\ &= \frac{w_{i-1/2}\psi_{i-1} - (w_{i-1/2} + w_{i+1/2})\psi_i + w_{i+1/2}\psi_{i+1}}{s^2} + O(s^2). \end{aligned} \quad (4.8)$$

Jak widzimy, w punktach siatki o niecałkowitych wskaźnikach są określone tylko wartości funkcji w , natomiast funkcja falowa ψ jest reprezentowana na punktach o wskaźnikach całkowitych, tak jak w dyskretniej postaci (2.6) zwykłego równania Schrödingera. W ten sposób, pomimo zastosowania siatki złożonej z $2n + 1$ punktów, otrzymamy AZW (4.9) rozmiaru $n \times n$.

Zauważmy, że ciągłość funkcji $f = w \frac{d\psi}{dx}$ jest drugim z warunków (4.4). O tym, że jest on spełniony, przekonujemy się mnożąc równanie (4.6) przez s i podstawiając zależność (4.8). Otrzymujemy

$$\lim_{s \rightarrow 0} [f(z - \frac{1}{2}s) - f(z + \frac{1}{2}s)] = \lim_{s \rightarrow 0} \{\alpha s [\varepsilon - V(x)] \psi(x)\} = 0.$$

Wykorzystując schemat dyskretyzacji (4.8), możemy zapisać równanie masy efektywnej (4.5) dla n kolejnych punktów siatki x_i :

$$-w_{i-1/2}\psi_{i-1} + (w_{i-1/2} + w_{i+1/2} + \tilde{v}_i)\psi_i - w_{i+1/2}\psi_{i+1} = \tilde{\varepsilon}\psi_i,$$

gdzie $\tilde{v}_i = \alpha s^2 v_i$, $\tilde{\varepsilon} = \alpha s^2 \varepsilon$. Nakładając warunki brzegowe $\psi_0 = \psi_{n+1} = 0$, możemy zapisać powyższy układ równań w postaci macierzowego zagadnienia własnego

$$\tilde{\mathbf{H}}\psi = \tilde{\varepsilon}\psi \quad (4.9)$$

z trójdziagonalną macierzą postaci

$$\tilde{\mathbf{H}} = \begin{pmatrix} w_{1/2} + w_{3/2} + \tilde{v}_1 & -w_{3/2} & & & \\ & -w_{3/2} & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & -w_{n-1/2} \\ & & & -w_{n-1/2} & w_{n-1/2} + w_{n+1/2} + \tilde{v}_n \end{pmatrix}.$$

Jest to symetryczna macierz trójdziagonalna, której wartości własne można wyznaczać np. metodą bisekcji z wykorzystaniem twierdzenia Martina–Deana, a wektory własne — przy pomocy metody DWSZ.

Zadania

4/1. Napisać program pozwalający znaleźć numeryczne rozwiązania stacjonarnego jednowymiarowego równania masy efektywnej (tj. energie i funkcje falowe stanów związanych) dla dowolnych zależności $V(X)$ i $M(X)$, zadanych przez tablice wartości.

4/2. W przypadku heterostruktur półprzewodnikowych GaAs-Al_cGa_{1-c}As możemy przyjąć³⁾, że w pasmie przewodnictwa

$$V(c) = c \cdot 873 \text{ meV}, \quad m(c) = (0,067 + 0,083c)m_e;$$

w obszarze wypełnionym materiałem o większym ułamku molowym Al (oznaczonym przez c) znajduje się bariera studni kwantowej.

Znaleźć energie i funkcje falowe stanów własnych w strukturze Al_{0,3}Ga_{0,7}As-GaAs-Al_{0,3}Ga_{0,7}As, tworzącej studnię kwantową o szerokości 500 Å, tj. dla zależności składu materiału od położenia postaci

$$c(x) = \begin{cases} 0 & (|x| < 250 \text{ Å}) \\ 0,3 & (|x| \geq 250 \text{ Å}). \end{cases}$$

4/3. Poczynione w poprzednim zadaniu założenie o idealnie schodkowym charakterze zależności $c(x)$, a co za tym idzie, $V(X)$ i $M(X)$, nie jest uzasadnione. Zbadać, jak zmienia się poziomy energetyczne i funkcje falowe, jeśli przyjmimy zależność postaci

$$c(x) = \begin{cases} 0 & (|x| \leq 250 \text{ Å} - d) \\ 0,15 + \frac{0,15}{d}(|x| - 250) & (250 \text{ Å} - d < |x| < 250 \text{ Å} + d) \\ 0,3 & (|x| \geq 250 \text{ Å} + d), \end{cases}$$

³⁾ Dla $c < 0,45$. Podane parametry pochodzą z pracy [84], a tzw. *band offset* dla pasma przewodnictwa przyjęto równy 70%

gdzie parametr d opisuje *rozmycie* granic studni? Jak zależą wyniki od wartości parametru d ?

4/4. Znaleźć stany związane elektronów w strukturze GaAs-Al_cGa_{1-c}As, której zależność składu od współrzędnej zadana jest funkcją

$$c(x) = 0,3 \left[1 - e^{-(x/d)^2} \right],$$

gdzie $d = 250 \text{ \AA}$.

Numeryczna analiza równania Schrödingera za pomocą metod różnic skończonych wyższych rzędów

W tym rozdziale zajmiemy się metodami różnic skończonych umożliwiającymi przybliżanie drugiej pochodnej w równaniu Schrödingera (1.2) z dokładnością wyższą niż $O(s^2)$ w sensie, o którym jest mowa w Dodatku C. Przedstawimy ogólną metodę Guardioli–Rosa [56] oraz jej przypadki szczególne, zwane metodą pięciopunktową oraz metodą Lindberga [57].

5.1. Metoda Guardioli–Rosa

Niech funkcja $f(x)$ będzie zadana na przedziale $\langle a, b \rangle$, a $f_i = f(x_i)$ oznaczają jej wartości na siatce punktów (2.3) o kroku s . Rozwinięcie Taylora pozwala otrzymać związek pomiędzy wartościami f w sąsiednich punktach siatki:

$$f(x+s) = \left(1 + sD + \frac{s^2}{2!}D^2 + \frac{s^3}{3!}D^3 + \dots\right) f(x),$$

gdzie $D = \frac{d}{dx}$. Ostatnią równość możemy zapisać formalnie w postaci

$$f(x+s) = e^{sD} f(x).$$

Operator Δ_s centralnej różnicy zdefiniujemy teraz jako

$$\Delta_s f(x) = f(x + \frac{1}{2}s) - f(x - \frac{1}{2}s),$$

co w zapisie symbolicznym ma postać

$$\Delta_s f(x) = [\exp(\frac{1}{2}sD) - \exp(-\frac{1}{2}sD)] f(x)$$

i pozwala przyjąć

$$\Delta_s = \exp(\frac{1}{2}sD) - \exp(-\frac{1}{2}sD) = 2 \sinh(\frac{1}{2}sD).$$

Rozwikłanie powyższej relacji prowadzi do związków

$$sD = 2 \sinh^{-1}(\frac{1}{2}\Delta_s) \tag{5.1}$$

$$s^2 D^2 = 4 [\sinh^{-1}(\frac{1}{2}\Delta_s)]^2. \tag{5.2}$$

Po rozwinięciu funkcji odwrotnej do $\sinh x$ (tj. $\operatorname{ar sinh} x$) w szereg Taylora [85], z relacji (5.1) i (5.2) otrzymujemy

$$\begin{aligned} s^2 D^2 = \Delta_s^2 & \left(1 - \frac{1}{12}\Delta_s^2 + \frac{1}{90}\Delta_s^4 - \frac{1}{560}\Delta_s^6 + \frac{1}{3150}\Delta_s^8 - \frac{1}{16632}\Delta_s^{10} + \right. \\ & \left. + \frac{1}{84084}\Delta_s^{12} - \frac{1}{411840}\Delta_s^{14} + \frac{1}{1969110}\Delta_s^{16} + \dots \right). \end{aligned} \tag{5.3}$$

Zauważmy, że we wzorze (5.3) występują tylko parzyste potęgi Δ_s . Zapewnia to, że w naszych rozważaniach pojawiają się wartości f w punktach x_i siatki (2.3). Nieparzyste potęgi Δ_s określone są przez wartości f w punktach $x_{i\pm 1/2}$ spoza siatki (2.3).

Dodajmy, że wartość kroku s jest uwzględniona w definicji Δ_s , chociaż w jawnej postaci nie występuje po prawej stronie równości (5.3). Uwzględnienie w (5.3) wszystkich wyrazów do Δ_s^{2n} włącznie oznacza, że druga pochodna jest wyznaczona z dokładnością $O(s^{2n+2})$ [86].

Szereg (5.3) jest zbieżny dla $|\frac{1}{2}s| < 1$. Aby zwiększyć promień zbieżności przybliżenia s^2D^2 , zastępujemy go aproksymantą Padégo [87]. Zastosowanie procedury z Dodatku D prowadzi do następujących aproksymant Padégo dla operatora s^2D^2 :

$$s^2D^2[1/0] = \Delta_s^2 - \frac{s^4D^4}{12} \quad (5.4)$$

$$s^2D^2[2/0] = \Delta_s^2 \left(1 - \frac{\Delta_s^2}{12}\right) + \frac{s^6D^6}{90}, \quad (5.5)$$

$$s^2D^2[1/1] = \frac{\Delta_s^2}{1 + \frac{1}{12}\Delta_s^2} + \frac{s^6D^6}{240}, \quad (5.6)$$

$$s^2D^2[3/0] = \Delta_s^2 \left(1 - \frac{\Delta_s^2}{12} + \frac{\Delta_s^4}{90}\right) + \frac{s^8D^8}{560}, \quad (5.7)$$

$$s^2D^2[2/1] = \frac{\Delta_s^2(1 + \frac{1}{20}\Delta_s^2)}{1 + \frac{2}{15}\Delta_s^2} - \frac{23s^8D^8}{75600}, \quad (5.8)$$

$$s^2D^2[1/2] = \frac{\Delta_s^2}{1 + \frac{1}{12}\Delta_s^2 - \frac{1}{240}\Delta_s^4} - \frac{31s^8D^8}{60480}, \quad (5.9)$$

$$s^2D^2[4/0] = \Delta_s^2 \left(1 - \frac{\Delta_s^2}{12} + \frac{\Delta_s^4}{90} - \frac{\Delta_s^6}{560}\right) + \frac{s^{10}D^{10}}{3150}, \quad (5.10)$$

$$s^2D^2[3/1] = \frac{\Delta_s^2(1 + \frac{13}{168}\Delta_s^2 - \frac{23}{10080}\Delta_s^4)}{1 + \frac{9}{56}\Delta_s^2} + \frac{3s^{10}D^{10}}{1411200}, \quad (5.11)$$

$$s^2D^2[2/2] = \frac{\Delta_s^2(1 + \frac{31}{252}\Delta_s^2)}{1 + \frac{13}{63}\Delta_s^2 + \frac{23}{3780}\Delta_s^4} + \frac{79s^{10}D^{10}}{4762800}, \quad (5.12)$$

$$s^2D^2[1/3] = \frac{\Delta_s^2}{1 + \frac{1}{12}\Delta_s^2 - \frac{1}{240}\Delta_s^4 + \frac{31}{60480}\Delta_s^6} + \frac{289s^{10}D^{10}}{725760}, \quad (5.13)$$

$$s^2D^2[5/0] = \Delta_s^2 \left(1 - \frac{\Delta_s^2}{12} + \frac{\Delta_s^4}{90} - \frac{\Delta_s^6}{560} + \frac{\Delta_s^8}{3150}\right) - 6,01 \cdot 10^{-5} s^{12}D^{12}, \quad (5.14)$$

$$s^2D^2[4/1] = \frac{\Delta_s^2(1 + \frac{17}{180}\Delta_s^2 - \frac{1}{270}\Delta_s^4 + \frac{43}{226800}\Delta_s^6)}{1 + \frac{8}{45}\Delta_s^2} - 3,688 \cdot 10^{-6} s^{12}D^{12}, \quad (5.15)$$

$$s^2D^2[3/2] = \frac{\Delta_s^2(1 + \frac{49}{276}\Delta_s^2 + \frac{79}{28980}\Delta_s^4)}{1 + \frac{6}{23}\Delta_s^2 + \frac{43}{3220}\Delta_s^4} - 1,156 \cdot 10^{-6} s^{12}D^{12}, \quad (5.16)$$

$$s^2 D^2[2/3] = \frac{\Delta_s^2(1 - \frac{641}{6360}\Delta_s^2)}{1 - \frac{37}{2120}\Delta_s^2 - \frac{959}{76320}\Delta_s^4 + \frac{29891}{32054400}\Delta_s^6} - 3,287 \cdot 10^{-5} s^{12} D^{12}, \quad (5.17)$$

$$s^2 D^2[1/4] = \frac{\Delta_s^2}{1 + \frac{1}{12}\Delta_s^2 - \frac{1}{240}\Delta_s^4 + \frac{31}{60480}\Delta_s^6 - \frac{289}{725760}\Delta_s^8} - \quad (5.18)$$

$$- 1,256 \cdot 10^{-5} s^{12} D^{12}, \quad (5.19)$$

gdzie

$$\Delta_s^{2n} f(x) = \sum_{k=0}^{2n} \binom{2n}{k} (-1)^k f(x + (n-k)s). \quad (5.20)$$

Sens matematyczny wprowadzonych aproksymant (5.4)–(5.19), mających ogólną postać

$$s^2 D^2[j/m] = \frac{a_2 \Delta_s^2 + a_4 \Delta_s^4 + \dots + a_{2j} \Delta_s^{2j}}{1 + b_2 \Delta_s^2 + b_4 \Delta_s^4 + \dots + b_{2m} \Delta_s^{2m}} = \frac{J(\Delta_s^2)}{M(\Delta_s^2)} := M^{-1}(\Delta_s^2) J(\Delta_s^2),$$

staje się jasny po ich zastosowaniu do równania Schrödingera

$$s^2 D^2 \psi(x)|_{x=x_i} = \alpha s^2 [v(x_i) - \varepsilon] \psi(x_i). \quad (5.21)$$

Podstawiając do równania (5.21)

$$s^2 D^2 \psi(x)|_{x=x_i} \simeq M^{-1}(\Delta_s^2) J(\Delta_s^2) \psi(x)|_{x=x_i}$$

otrzymujemy

$$M^{-1}(\Delta_s^2) J(\Delta_s^2) \psi(x)|_{x=x_i} = \alpha s^2 [v(x) - \varepsilon] \psi(x)|_{x=x_i}. \quad (5.22)$$

Działając na obie strony równania (5.22) operatorem $M(\Delta_s^2)$, dostajemy

$$J(\Delta_s^2) \psi(x)|_{x=x_i} = \alpha s^2 M(\Delta_s^2) [v(x) - \varepsilon] \psi(x)|_{x=x_i}. \quad (5.23)$$

Jak widzimy, operatory $M(\Delta_s^2)$ i $J(\Delta_s^2)$ działają odpowiednio na lewą i prawą równania (5.21).

W praktyce obliczeń numerycznych stosujemy obcięty szereg Taylora (5.3) lub odpowiednio aproksymanty Padégo przytoczone wyżej.

Formuły Padégo (5.4)–(5.19) opatrzone standardowymi symbolami typu $[j/m]$, które oznaczają to, że w liczniku i mianowniku występują odpowiednio wielomiany j -tego i m -tego stopnia względem Δ_s^2 .

Zauważmy, że aproksymanty Padégo typu $[j/0]$ odpowiadają j -punktowemu przybliżeniu drugiej pochodnej za pomocą różnic skończonych. Aproksymanty Padégo typu $[1/m]$ są m -punktowymi metodami Numerowa. Uogólnione metody Numerowa odpowiadają diagonalnym aproksymantom Padégo. Jak widzimy, uogólniona metoda Numerowa (5.6) odpowiadająca aproksymancie $s^2 D^2[1/1]$ jest dokładniejsza od metody pięciopunktowej Taylora (5.5). Ogólnie, aproksymanta $s^2 D^2[j/m]$ jest poprawna do wyrazów rzędu $s^{2(j+m+1)}$.

Dla praktycznych zastosowań najlepszymi aproksymantami Padégo są te, które przy danej sumie $j+m$ zawierają najniższe potęgi s (s^{2n} odpowiada $(2n+1)$ -punktowemu przybliżeniu). Odpowiada to aproksymantom $[j/j]$ oraz $[(j+1)/j]$.

Przystąpimy do wyprowadzenia algebraicznego zagadnienia własnego w metodzie Guardiola–Rosa. W tym celu wybieramy jedną z aproksymant Padégo (5.4)–(5.19). Stosujemy wzór (5.23) do każdego punktu siatki. W rezultacie otrzymujemy uogólnione algebraiczne zagadnienie własne [24, 45, 50, 53]

$$\mathbf{J}\psi = \mathbf{M}(\tilde{\mathbf{V}} - \tilde{\varepsilon}\mathbf{1})\psi, \quad (5.24)$$

gdzie $\tilde{\mathbf{V}} = \text{diag}(v_1, \dots, v_n)$ — macierz diagonalna,

$$\mathbf{J} = \sum_{k=1}^j \nu_k \mathbf{\Delta}^{2k}, \quad \mathbf{M} = \mathbf{1} + \sum_{l=1}^m \mu_l \mathbf{\Delta}^{2l}, \quad (5.25)$$

a $\mathbf{\Delta}^2$ oznacza symetryczną macierz trójdagonalną o postaci

$$\mathbf{\Delta}^2 = \begin{pmatrix} -2 & 1 & & & \\ 1 & \ddots & \ddots & & \\ & \ddots & \ddots & 1 & \\ & & & 1 & -2 \end{pmatrix}. \quad (5.26)$$

Przy wyprowadzaniu (5.24) posłużono się wzorem (5.20) oraz zastosowano dla operatorów $\Delta^4, \Delta^6, \dots$ następujące warunki brzegowe:

$$f_{-i} = -f_i, \quad f_{n+i} = -f_{n-i+2}. \quad (5.27)$$

Powodują one, że wektory własne odpowiednich macierzy są periodyczne z okresem $2n+2$, co we współrzędnej x odpowiada odległości $2(b-a)$.

Rozwiążemy teraz zagadnienie własne dla macierzy $\mathbf{\Delta}^2$, której diagonalizacja będzie konieczna w naszych rozważaniach.

5.2. Zagadnienie własne dla macierzy $\mathbf{\Delta}^2$

Zastosowanie aproksymanty (5.4) w równaniu (5.21), przy dodatkowym założeniu $v(x) = \text{const}$ i przy zadanych poprzednio warunkach brzegowych (2.2), prowadzi do zagadnienia własnego postaci

$$\mathbf{\Delta}^2 \mathbf{q} = \lambda \mathbf{q}, \quad (5.28)$$

gdzie $\mathbf{\Delta}^2$ jest macierzą (5.26). W dziedzinie, którą się zajmujemy, zagadnienie (5.28) jest paradygmatem¹⁾. Macierz tę diagonalizujemy za pomocą ortogonalnego przekształcenia podobieństwa [20, 45, 88]:

$$\mathbf{Q} \mathbf{\Delta}^2 \mathbf{Q}^T = \mathbf{\Lambda} = \text{diag}(\lambda, \dots, \lambda_n), \quad (5.29)$$

gdzie $\mathbf{\Lambda}$ jest macierzą diagonalną, macierz ortogonalna \mathbf{Q} ma wyrazy²⁾

$$Q_{i,k} = \sqrt{\frac{2}{n+1}} \sin \frac{ik\pi}{n+1}, \quad (5.30)$$

¹⁾ Słowa *paradygmat* używamy tutaj w celu podkreślenia znaczenia problemu (5.28), odgrywającego rolę ogólnie przyjętego wzorca rozwiązywania zagadnienia własnego.

²⁾ Analityczne rozwiązanie zagadnienia (5.28) podane jest w rozdziale 3 książki [88].

a wartości własne w_k macierzy (5.26) są równe

$$\lambda_k = -4 \sin^2 \frac{k\pi}{2(n+1)} \quad (k = 1, \dots, n).$$

Dodajmy, że wektory własne macierzy $\mathbf{\Delta}^2$ są kolumnami macierzy \mathbf{Q} .

W celu uzasadnienia podanych formuł napiszmy ciąg równości słuszny dla dowolnego k [20]:

$$\sin \frac{(i-1)k\pi}{n+1} + \sin \frac{(i+1)k\pi}{n+1} = 2 \cos \frac{k\pi}{n+1} \sin \frac{ik\pi}{n+1} \quad (i = 1, \dots, n),$$

wynikający ze znanego wzoru dla sumy sinusów. Od i -tej równości odejmujemy obustronnie $2 \sin \frac{ik\pi}{n+1}$ i otrzymujemy

$$\sin \frac{(i-1)k\pi}{n+1} - 2 \sin \frac{ik\pi}{n+1} + \sin \frac{(i+1)k\pi}{n+1} = 2 \left(\cos \frac{k\pi}{n+1} - 1 \right) \sin \frac{ik\pi}{n+1}.$$

Ponieważ $\cos \frac{k\pi}{n+1} - 1 = -2 \sin^2 \frac{k\pi}{2(n+1)}$, to

$$\sin \frac{(i-1)k\pi}{n+1} - 2 \sin \frac{ik\pi}{n+1} + \sin \frac{(i+1)k\pi}{n+1} = -4 \sin^2 \frac{k\pi}{2(n+1)} \sin \frac{ik\pi}{n+1}.$$

Macierzowy zapis ostatniego układu równań dla $j = 1, \dots, n$ jest postaci (5.28), co kończy nasze uzasadnienie.

Dodajmy jeszcze, że macierzową reprezentację operatora Δ_s^{2k} na siatce punktów (2.3) przy założonych warunkach brzegowych (5.27) zadają macierze $\mathbf{\Delta}^{2k}$. Dla $k = 2$ nieskomplikowane obliczenia prowadzą do wzoru

$$\mathbf{\Delta}^4 = \begin{pmatrix} 5 & -4 & 1 & & & \\ -4 & 6 & \ddots & \ddots & & \\ 1 & \ddots & \ddots & \ddots & 1 & \\ & \ddots & \ddots & 6 & -4 & \\ & & & 1 & -4 & 5 \end{pmatrix}.$$

W tym przypadku warunki brzegowe mają postać

$$f_{-1} = -f_1, \quad f_0 = f_{n+1} = 0, \quad f_{n+2} = -f_n$$

i wtedy

$$\Delta_s^4 f(x)|_{x=x_1} = f_{-1} - 4f_0 + 6f_1 - 4f_2 + f_3 = 5f_1 - 4f_2 + f_3.$$

5.3. Symetryczne algebraiczne zagadnienie własne w metodzie Guardiola–Rosa

W ogólnym przypadku równanie (5.24) jest niesymetrycznym uogólnionym zagadnieniem własnym, ponieważ $\mathbf{M}(\tilde{\mathbf{V}} - \varepsilon \mathbf{1})$ jest macierzą niesymetryczną. Sprowadzimy je do symetrycznego AZW. W tym celu mnożymy (5.24) obustronnie przez macierz \mathbf{Q} , diagonalizującą przez podobieństwo macierz $\mathbf{\Delta}^2$ — patrz wzory (5.29)–(5.30):

$$\mathbf{QJ}\psi = \mathbf{QM}(\tilde{\mathbf{V}} - \varepsilon \mathbf{1})\psi. \quad (5.31)$$

Zajmiemy się oddzielnie lewą i prawą stroną ostatniego równania. Z jawnej postaci macierzy \mathbf{J} (5.25) oraz z własności macierzy \mathbf{Q} wynika, że

$$\begin{aligned} \mathbf{QJ}\psi &= \mathbf{Q} \sum_{k=1}^j \nu_k \mathbf{\Delta}^{2k} \psi = \mathbf{Q} \sum_{k=1}^j \nu_k \left(\mathbf{\Delta}^2 \mathbf{Q}^T \mathbf{Q} \right)^k \psi = \\ &= \sum_{k=1}^j \nu_k \underbrace{\mathbf{Q} \mathbf{\Delta}^2 \mathbf{Q}^T \cdots \mathbf{Q} \mathbf{\Delta}^2 \mathbf{Q}^T}_{k \text{ razy}} \mathbf{Q} \psi = \sum_{k=1}^j \nu_k \left(\mathbf{Q} \mathbf{\Delta}^2 \mathbf{Q}^T \right)^k \mathbf{Q} \psi = \\ &= \left(\sum_{k=1}^j \nu_k \mathbf{\Lambda}^k \right) \mathbf{Q} \psi = \mathbf{\Gamma} \mathbf{Q} \psi, \end{aligned} \quad (5.32)$$

gdzie $\mathbf{\Gamma}$ jest macierzą diagonalną $\mathbf{\Gamma} = \text{diag}(\gamma_1, \dots, \gamma_n)$ o elementach

$$\gamma_i = \sum_{k=1}^j \nu_k (-4)^k \sin^{2k} \frac{i\pi}{2(n+1)}. \quad (5.33)$$

W przypadku prawej strony równania (5.31) podobne przekształcenia dają:

$$(\mathbf{QM}\tilde{\mathbf{V}} - \varepsilon \mathbf{QM})\psi = (\mathbf{QM}\mathbf{Q}^T)(\mathbf{Q}\tilde{\mathbf{V}}\mathbf{Q}^T\mathbf{Q})\psi - \varepsilon(\mathbf{QM}\mathbf{Q}^T)\mathbf{Q}\psi = \mathbf{K}(\tilde{\mathbf{V}} - \varepsilon \mathbf{1})\mathbf{Q}\psi, \quad (5.34)$$

gdzie macierz diagonalna $\mathbf{K} = \text{diag}(\kappa_1, \dots, \kappa_n)$ i ma elementy

$$\kappa_i = 1 + \sum_{l=1}^m \mu_l (-4)^l \sin^{2l} \frac{i\pi}{2(n+1)}, \quad (5.35)$$

natomiast $\tilde{\mathbf{V}}$ jest macierzą gęstą o wyrazach

$$\tilde{V}_{i,k} = \frac{2}{n+1} \sum_{l=1}^n \sin \frac{il\pi}{n+1} \tilde{v}_l \sin \frac{kl\pi}{n+1}. \quad (5.36)$$

Ze wzorów (5.32) i (5.34) wynika relacja

$$\mathbf{\Gamma}\phi = \mathbf{K}(\tilde{\mathbf{V}} - \varepsilon \mathbf{1})\phi,$$

gdzie $\phi = \mathbf{Q}\psi$. Mnożymy obustronnie ostatnią równość przez \mathbf{K}^{-1} i grupujemy wyrazy, otrzymując

$$(-\mathbf{K}^{-1}\mathbf{\Gamma} + \tilde{\mathbf{V}})\phi = \varepsilon\phi. \quad (5.37)$$

Ostatnie równanie jest algebraicznym zagadnieniem własnym dla *symetrycznej macierzy gęstej*, ponieważ macierz $\mathbf{K}^{-1}\mathbf{\Gamma}$ jest diagonalna i $\tilde{\mathbf{V}}^T = \tilde{\mathbf{V}}$, co wynika ze wzoru (5.36).

Numeryczne rozwiązywanie AZW z macierzą gęstą będzie dyskutowane w następnym rozdziale. Zaprezentowane tutaj podejście, wykorzystujące diagonalizację macierzy \mathbf{J} i \mathbf{M} , ma zastosowanie dla aproksymant Padégo $s^2D^2[j/m]$ przy $m > 1$ — patrz wzory (5.4)–(5.19). Dalej zajmiemy się przypadkami $[2/0]$ (5.5) oraz $[1/1]$ (5.6), które można analizować w prostszy sposób. W metodach tych druga pochodna jest lokalnie przybliżana z dokładnością do wyrazów $O(s^4)$.

5.4. Pięciopunktowe przybliżenie drugiej pochodnej

Najprostszą pojęciowo metodą wyższego rzędu jest metoda wykorzystująca pięciopunktowe przybliżenie drugiej pochodnej (5.5) albo, równoważnie, (C.9).

Po sprowadzeniu RS do postaci bezwymiarowej, równanie różnicowe dla i -tego punktu siatki (2.3) przyjmuje postać

$$\psi_{i-2} - 16\psi_{i-1} + (30 + \tilde{v}_i)\psi_i - 16\psi_{i+1} + \psi_{i+2} = \tilde{\varepsilon}\psi_i, \quad (5.38)$$

gdzie tym razem

$$\tilde{\varepsilon} = 12\alpha s^2 \varepsilon, \quad \tilde{v}_i = 12\alpha s^2 v_i.$$

Pokażemy, że (5.38) ma postać algebraicznego zagadnienia własnego dla symetrycznej macierzy pasmowej, której wyrazy potrafimy wyznaczyć. Wypiszmy jawne postaci równań (5.38) dla poszczególnych wartości i :

$$\left\{ \begin{array}{lcl} (30 + \tilde{v}_1)\psi_1 - 16\psi_2 + \psi_3 & & = \tilde{\varepsilon}\psi_1 \\ & \ddots & \vdots \\ \psi_{i-2} - 16\psi_{i-1} + (30 + \tilde{v}_i)\psi_i - 16\psi_{i+1} + \psi_{i+2} & & = \tilde{\varepsilon}\psi_i \\ & \ddots & \vdots \\ \psi_{n-2} - 16\psi_{n-1} + (30 + \tilde{v}_n)\psi_n & & = \tilde{\varepsilon}\psi_n, \end{array} \right.$$

gdzie uwzględniono warunki brzegowe $\psi_0 = \psi(a) = \psi_{n+1} = \psi(b) = 0$, $\psi_{-1} = \psi_{n+2} = 0$.

Powyższy układ równań można zapisać w postaci AZW

$$\begin{pmatrix} 30 + \tilde{v}_1 & -16 & 1 & & \\ -16 & 30 + \tilde{v}_2 & \ddots & \ddots & \\ 1 & \ddots & \ddots & \ddots & 1 \\ & \ddots & \ddots & 30 + \tilde{v}_{n-1} & -16 \\ & & 1 & -16 & 30 + \tilde{v}_n \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_{n-1} \\ \psi_n \end{pmatrix} = \tilde{\varepsilon} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_{n-1} \\ \psi_n \end{pmatrix}.$$

Otrzymaliśmy więc algebraiczne zagadnienie własne

$$\mathbf{P}\psi = \tilde{\varepsilon}\psi$$

dla symetrycznej macierzy pentadiagonalnej

$$\mathbf{P} = \begin{pmatrix} d_1 & e_1 & f_1 & & \\ e_1 & d_2 & \ddots & \ddots & \\ f_1 & \ddots & \ddots & \ddots & f_{n-2} \\ & \ddots & \ddots & d_{n-1} & e_{n-1} \\ & & f_{n-2} & e_{n-1} & d_n \end{pmatrix}, \quad (5.39)$$

której wyrazy diagonalne określa potencjał $v(x)$.

W celu wyznaczenia wartości własnych macierzy (5.39) możemy ponownie zastosować twierdzenia Martina–Deana (B-17). Postępując analogicznie, jak w przypadku macierzy trójdzielnej (podrozdział 3.1), otrzymujemy dwa wzajemnie zależne ciągi

$$\begin{aligned} u_1 &= d_1 - z, \quad u_2 = d_2 - z - \frac{p_1^2}{u_1}, \quad u_i = d_i - z - \frac{f_{i-2}^2}{u_{i-2}} - \frac{p_{i-1}^2}{u_{i-1}} \quad (i = 3, \dots, n), \\ p_1 &= e_1, \quad p_i = e_i - \frac{p_{i-1}f_{i-2}}{u_{i-2}} \quad (i = 2, \dots, n-1). \end{aligned} \quad (5.40)$$

Liczba $l_{\text{ww}}(z)$ wartości własnych macierzy \mathbf{P} mniejszych od zadanej liczby z jest równa liczbie ujemnych wyrazów w ciągu (u_i) :

$$l_{\text{ww}}(z) = \eta(\mathbf{P} - z\mathbf{1}) = \sum_{i=1}^n \eta(u_i). \quad (5.41)$$

Wartości własne macierzy pentadiagonalnej \mathbf{P} możemy teraz wyznaczać za pomocą metody bisekcji opisanej w podrozdziale 3.2.

5.5. Metoda Lindberga

Pokażemy za Lindbergiem [57], że aproksymanta Padégo $s^2D^2[1/1]$ postaci (5.6) prowadzi do zagadnienia własnego z *kwazisymetryczną* [45] macierzą trójdzielną.

Zastosowanie przybliżenia

$$s^2D^2 \simeq \left(1 + \frac{1}{12}\Delta_s^2\right)^{-1} \Delta_s^2$$

w równaniu (5.21), na podstawie (5.23), prowadzi do związku

$$\left(1 + \frac{1}{12}\Delta_s^2\right) s^2\alpha[v(x) - \varepsilon]\psi(x) \Big|_{x=x_i} = \Delta_s^2\psi(x) \Big|_{x=x_i}. \quad (5.42)$$

Jeśli wprowadzimy oznaczenia

$$\tilde{v}(x) = \frac{1}{12}s^2\alpha v(x), \quad \tilde{\varepsilon} = \frac{1}{12}s^2\alpha\varepsilon,$$

to równanie (5.42) przyjmie postać

$$(12 + \Delta_s^2)[\tilde{v}(x) - \tilde{\varepsilon}]\psi(x) \Big|_{x=x_i} = \Delta_s^2\psi(x) \Big|_{x=x_i}.$$

Rozpiszmy lewą stronę powyższego równania:

$$\begin{aligned} (12 + \Delta_s^2)[\tilde{v}(x) - \tilde{\varepsilon}]\psi(x)|_{x=x_i} &= \\ &= 12(\tilde{v}_i - \tilde{\varepsilon})\psi_i + [(\tilde{v}_{i-1} - \tilde{\varepsilon})\psi_{i-1} - 2(\tilde{v}_i - \tilde{\varepsilon})\psi_i + (\tilde{v}_{i+1} - \tilde{\varepsilon})\psi_{i+1}] \end{aligned} \quad (5.43)$$

i prawą:

$$\Delta_s^2\psi(x)|_{x=x_i} = \psi_{i-1} - 2\psi_i + \psi_{i+1}. \quad (5.44)$$

Stosując wzory (5.43) i (5.44) do równania Schrödingera we wszystkich punktach siatki (2.3), otrzymujemy uogólnione zagadnienie własne

$$\tilde{\mathbf{G}}\psi = \tilde{\varepsilon}\mathbf{F}\psi, \quad (5.45)$$

gdzie $\tilde{\mathbf{G}}$ jest *niesymetryczną*, a \mathbf{F} — symetryczną macierzą trójdagonalną:

$$\tilde{\mathbf{G}} = \begin{pmatrix} 2 + 10\tilde{v}_1 & -1 + \tilde{v}_2 & & & \\ -1 + \tilde{v}_1 & \ddots & \ddots & & \\ & \ddots & \ddots & -1 + \tilde{v}_n & \\ & & -1 + \tilde{v}_{n-1} & 2 + 10\tilde{v}_n & \end{pmatrix}, \quad \tilde{\mathbf{F}} = \begin{pmatrix} 10 & 1 & & & \\ 1 & \ddots & \ddots & & \\ & \ddots & \ddots & 1 & \\ & & 1 & 10 & \end{pmatrix}.$$

Zajmiemy się teraz wyznaczeniem wartości i wektorów własnych uogólnionego zagadnienia własnego (5.45).

5.6. Kwazisymetryczna macierz trójdagonalna

Definicja 5-1. Niesymetryczną rzeczywistą macierz trójdagonalną postaci

$$\mathbf{T}' = \begin{pmatrix} d'_1 & f'_1 & & & \\ e'_1 & \ddots & \ddots & & \\ & \ddots & \ddots & f'_{n-1} & \\ & & e'_{n-1} & d'_n & \end{pmatrix} \quad (5.46)$$

nazywamy *kwazisymetryczną*, jeśli $e'_i f'_i > 0$ dla $i = 1, \dots, n-1$.

Dowolną trójdagonalną macierz kwazisymetryczną \mathbf{T}' można za pomocą przekształcenia podobieństwa określonego przez macierz diagonalną \mathbf{S} przekształcić w macierz symetryczną \mathbf{T} [45]:

$$\mathbf{S}\mathbf{T}'\mathbf{S}^{-1} = \mathbf{T}, \quad (5.47)$$

gdzie $\mathbf{S} = \text{diag}(s_1, \dots, s_n)$ oraz

$$s_1 = 1, \quad s_i = \sqrt{\prod_{l=1}^{i-1} e'_l / \prod_{l=1}^{i-1} f'_l} \quad (i = 2, \dots, n). \quad (5.48)$$

Wyrazy macierzy $\mathbf{T} = \text{trid}[(e_1, \dots, e_{n-1}), (d_1, \dots, d_n), (e_1, \dots, e_{n-1})]$ są równe

$$d_i = d'_i, \quad e_i = \sqrt{e'_i f'_i}. \quad (5.49)$$

Lindberg udowodnił następujące twierdzenie [57]:

Twierdzenie 5-2. Wszystkie wartości własne równania (5.45) są rzeczywiste i należą do przedziału $\langle \tilde{v}_{\min} + w_1, \tilde{v}_{\max} + w_n \rangle$, gdzie

$$\tilde{v}_{\min} = \min_{i=1}^n \tilde{v}_i, \quad \tilde{v}_{\max} = \max_{i=1}^n \tilde{v}_i, \quad w_i = \frac{\sin^2 \frac{i\pi}{2(n+1)}}{2 + \cos^2 \frac{i\pi}{2(n+1)}}.$$

Założmy, że spełniony jest warunek $z > \tilde{v}_{\min}$. Wówczas wszystkie wyrazy poza-diagonalne macierzy

$$\tilde{\mathbf{G}} - z\mathbf{F} = \begin{pmatrix} 2 + 10(\tilde{v}_1 - z) & -1 + \tilde{v}_2 - z & & & \\ -1 + \tilde{v}_1 - z & \ddots & & \ddots & \\ & \ddots & \ddots & \ddots & -1 + \tilde{v}_n - z \\ & & & -1 + \tilde{v}_{n-1} - z & 2 + 10(\tilde{v}_n - z) \end{pmatrix}$$

są ujemne. Zgodnie z wzorami (5.46)–(5.49), jest ona wtedy kwazisymetryczna, a zatem podobna do symetrycznej macierzy trójdagonalnej

$$\mathbf{T}(z) = \mathbf{S}(z)(\tilde{\mathbf{G}} - z\mathbf{F})\mathbf{S}^{-1}(z),$$

gdzie $\mathbf{T}(z) = \text{trid}[(e_1(z), \dots, e_{n-1}(z)), (d_1(z), \dots, d_n(z)), (e_1(z), \dots, e_{n-1}(z))]$ oraz

$$d_i(z) = 2 + 10(\tilde{v}_i - z), \quad e_i^2(z) = (\tilde{v}_i - z - 1)(\tilde{v}_{i+1} - z - 1).$$

Na mocy twierdzenia Martina–Deana, liczba ujemnych wartości własnych $\mathbf{T}(z)$, będąca zarazem liczbą wartości własnych wejściowego zagadnienia własnego (5.45), jest równa liczbie ujemnych wyrazów w ciągu

$$u_1 = d_1(z), \quad u_i = d_i(z) - \frac{e_{i-1}^2(z)}{u_{i-1}} \quad (i = 2, \dots, n). \quad (5.50)$$

W oparciu o ciąg (5.50), wartości własne równania (5.45) można wyznaczać za pomocą opisanego wcześniej algorytmu 3-1.

Wektory własne ψ w (5.45) można wyznaczyć metodą DWSZ przedstawioną w paragrafie 3.3.2 w wersji dla macierzy symetrycznych. Nadaje się ona do wyznaczania wektorów własnych niesymetrycznych macierzy trójdogonalnych \mathbf{T}' postaci (5.46) [55, 73, 74].

Niech $\tilde{\varepsilon}$ będzie znaną wartością własną. W celu wyznaczenia przynależnego do niej wektora własnego, wprowadźmy $(n-1)$ liczb $\Theta_1^-, \dots, \Theta_{n-1}^-$ oraz $(n-1)$ liczb $\Theta_n^+, \dots, \Theta_2^+$, zdefiniowanych następująco:

$$\begin{aligned} \Theta_1^- &= (d'_1 - \tilde{\varepsilon})^{-1}, & \Theta_i^- &= (d'_i - \tilde{\varepsilon} - e'_{i-1}f'_{i-1}\Theta_{i-1}^-)^{-1} \quad (i = 2, \dots, n-1), \\ \Theta_n^+ &= (d'_n - \tilde{\varepsilon})^{-1}, & \Theta_i^+ &= (d'_i - \tilde{\varepsilon} - e'_if'_i\Theta_{i+1}^+)^{-1} \quad (i = n-1, \dots, 2). \end{aligned}$$

Przy pomocy ciągów (Θ^+) bądź (Θ^-) wyznaczamy współrzędne wektora własnego. Kładziemy $\psi_k = 1$ i obliczamy

$$\psi_i = \begin{cases} -\Theta_i^- f'_i \psi_{i+1} & (i = k-1, \dots, 1) \\ -\Theta_i^+ e'_i \psi_{i+1} & (i = k+1, \dots, n). \end{cases}$$

Wszystkie obliczenia, łącznie z weryfikacją poprawności otrzymanego wektora własnego, prowadzić należy podobnie do postępowania określonego w algorytmie 3-4.

Zamiast stosować metodę DWSZ do macierzy niesymetrycznej $\tilde{\mathbf{G}} - z\mathbf{F}$, możemy zauważyć, że z równania $(\tilde{\mathbf{G}} - \tilde{\varepsilon}\mathbf{F})\psi = \mathbf{0}$ wynika

$$\mathbf{0} = \mathbf{S}(\tilde{\varepsilon})\mathbf{0} = \mathbf{S}(\tilde{\varepsilon})(\tilde{\mathbf{G}} - \tilde{\varepsilon}\mathbf{F})\mathbf{S}^{-1}(\tilde{\varepsilon})\mathbf{S}(\tilde{\varepsilon})\psi = \mathbf{T}(\tilde{\varepsilon})\mathbf{S}(\tilde{\varepsilon})\psi.$$

Oznaczmy $\mathbf{S}(\tilde{\varepsilon})\psi = \phi$. Z uwagi na symetryczność macierzy \mathbf{T} możemy zastosować symetryczny algorytm DWSZ do wektora ϕ , a następnie obliczyć wektor $\psi = \mathbf{S}^{-1}(\tilde{\varepsilon})\phi$.

5.7. Ekstrapolacja Richardsona

Dokładność wyznaczania wartości własnych oraz współrzędnych wektorów własnych AZW można w prosty sposób zwiększyć, rozwiązując RS (1.2) jedną z podanych metod dla dwóch różnych wartości s i $s' < s$ kroku siatki (2.3). W tym celu należy posłużyć się metodą zwaną *ekstrapolacją Richardsona* [54], którą teraz krótko przedstawimy.

Przyjmijmy, że druga pochodna występująca w RS jest aproksymowana za pomocą centralnych różnic skończonych, których dokładność jest rzędu $O(s^2)$ (jak w metodzie trójpunktowej). Dodatkowe obliczenia pozwalają osiągnąć dokładność rzędu $O(s^4)$.

Rozpatrzmy funkcję f , której przybliżone wartości $\{f_i(s)\}$ oraz $\{f'_j(s')\}$ są znane na dwóch siatkach $x_i = a + is$ oraz $x'_j = a + js'$. Dokładna wartość funkcji w punkcie $x = x_i = x'_j$, wspólnym dla obu siatek³⁾, wynosi

$$f_{\text{dokł}}(x) = f_i(s) + Cs^2 + O(s^4) = f'_j(s') + C(s')^2 + O[(s')^4]. \quad (5.51)$$

Zauważmy, że przyjęto założenie, iż lokalnie błąd przybliżenia (5.51) jest postaci wielomianu $Cs^2 + Ds^4 + Es^6 + \dots$.

Rozwiązując układ równań (5.51) względem C , otrzymujemy

$$f_{\text{dokł}}(x) \simeq f'_j(s') + \frac{f'_j(s') - f_i(s)}{(s/s')^2 - 1} = \frac{(s/s')^2 f'_j(s') - f_i(s)}{(s/s')^2 - 1},$$

gdzie tym razem dokładność wyrażenia stojącego po prawej stronie ostatniej równości jest rzędu $O(s^4)$. Jak widać, znając C możemy twierdzić, że dokładność otrzymanych wartości jest rzędu $O(s^4)$. Dla typowego przypadku $s' = \frac{1}{2}s$ otrzymujemy

$$f_{\text{dokł}}(x) = \frac{2^2 f'_j - f_i}{2^2 - 1} + O(s^4) = f'_j + \frac{f'_j - f_i}{2^2 - 1} + O(s^4). \quad (5.52)$$

Przyjmijmy teraz, że aproksymujemy drugą pochodną za pomocą centralnych różnic skończonych z dokładnością rzędu $O(s^{2n})$. Dodatkowe obliczenia pozwalają osiągnąć dokładność rzędu $O(s^{2n+2})$.

Niech $s' = \frac{1}{2}s$, wtedy

$$f_{\text{dokł}}(x) = f_i + Cs^{2n} + O(s^{2n+2}) = f'_j + C\frac{s^{2n}}{2^{2n}} + O[(\frac{1}{2}s)^{2n+2}]. \quad (5.53)$$

³⁾ Jeśli $s' = \frac{1}{2}s$, to jedna siatka jest podzbiorem drugiej. Taki przypadek będziemy dalej rozważać.

Rozwiązując układ równań (5.53) względem C , otrzymujemy

$$f_{\text{dokł}}(x) = \frac{2^{2n} f'_j - f_i}{2^{2n} - 1} + O(s^{2n+2}) = f'_j + \frac{f'_j - f_i}{2^{2n} - 1} + O(s^{2n+2}). \quad (5.54)$$

Wzór (5.54) można stosować do dowolnej wielkości wyznaczanej numerycznie, także do wartości własnych oraz współrzędnych wektorów własnych zagadnienia własnego, obliczanych którąś z metod przedstawionych w poprzednich rozdziałach.

Zilustrujemy to postępowanie na prostym przykładzie. W tabeli zawarte są przybliżone wartości pewnej funkcji $f(x)$ wyznaczone z podaną dokładnością:

krok s	dokładność $O(s^2)$	ekstrapolowana z dokł. $O(s^4)$	ekstrapolowana z dokł. $O(s^6)$
0,2	23,9944	\searrow \nearrow \searrow \nearrow	\searrow \nearrow
0,4	24,2328		
0,8	25,1768		
		23,9149	23,9147
		23,9181	

Poszczególne wartości w tabeli otrzymano, przyjmując

$$f_{\text{dokł}} = \text{obliczona numerycznie wartość} + \text{błąd}.$$

Pierwszą liczbę w trzeciej kolumnie obliczamy, zakładając

$$f_{\text{dokł}} \simeq f_{\text{ekstr}} = 23,9944 + C(0,2)^2 = 24,2328 + C(0,4)^2,$$

i dalej, ze wzoru (5.52),

$$f_{\text{ekstr}} = 23,9944 + \frac{23,9944 - 24,2328}{3} = 23,9149,$$

gdzie $f_{\text{dokł}} = f_{\text{ekstr}} + O[(0,4)^4]$. Podobnie obliczamy drugą liczbę z kolumny trzeciej:

$$f_{\text{ekstr}} = 24,2328 + C(0,4)^2 = 25,1768 + C(0,8)^2,$$

skąd

$$f_{\text{ekstr}} = 24,2328 + \frac{24,2328 - 25,1768}{3} = 23,9181,$$

gdzie tym razem $f_{\text{dokł}} = f_{\text{ekstr}} + O[(0,8)^4]$.

Jeśli teraz rozwiążemy układ równań

$$f_{\text{ekstr}} = 23,9149 + C(0,4)^4 = 23,9181 + C(0,8)^4,$$

to ze wzoru (5.54) dla $n = 2$ otrzymamy stojącą w czwartej kolumnie tabeli liczbę

$$f_{\text{ekstr}} = 23,9149 + \frac{23,9149 - 23,9181}{15} = 23,9147,$$

będącą wartością funkcji $f(x)$ z dokładnością do $O[(0,8)^6]$. Dodajmy, że dokładna wartość funkcji wynosi 23,9144.

Zadania

5/1. Korzystając z szeregu Taylora

$$\sinh^{-1} x = x - \frac{1 \cdot x^3}{2 \cdot 3} + \frac{1 \cdot 3 \cdot x^5}{2 \cdot 4 \cdot 5} - \frac{1 \cdot 3 \cdot 5 \cdot x^7}{2 \cdot 4 \cdot 6 \cdot 7} + \dots,$$

wyprowadzić trzy pierwsze wyrazy rozwinięcia (5.3).

5/2. Udowodnić za pomocą indukcji matematycznej wzór (5.20).

5/3. Wyprowadzić wyrażenia na następujące aproksymanty Padégo dla $s^2 D^2$: $[2/0]$, $[1/1]$, $[3/0]$, $[2/1]$, $[1/2]$, $[4/0]$.

5/4. Wyprowadzić bezpośrednim rachunkiem postać macierzy Δ^4 .

5/5. Uzasadnić równość (5.24). Posłużyć się warunkami brzegowymi (5.27).

5/6. Korzystając ze związku $\Delta = Q^T \Lambda Q$, wyprowadzić wzór (5.32).

5/7. Pokazać słuszność formuł (5.34)–(5.36), pozwalających sprowadzić uogólnione AZW (5.24) do postaci symetrycznego AZW.

5/8. Wyprowadzić wzór (5.38) posługując się aproksymantą Padégo (5.5). Wskazówka: można skorzystać ze wzoru (5.20).

5/9. Stosując twierdzenie Martina–Deana (B-17) wyprowadzić formuły (5.40).

5/10. Pokazać, że siedmiopunktowe przybliżenie drugiej pochodnej ma postać

$$\left. \frac{d^2 \psi}{dx^2} \right|_{x_i} = \frac{2\psi_{i-3} - 27\psi_{i-2} + 270\psi_{i-1} - 490\psi_i + 270\psi_{i+1} - 27\psi_{i+2} + 2\psi_{i+3}}{180s^2} + O(s^6).$$

5/11. Zapisać jednowymiarowe stacjonarne równanie Schrödingera w postaci algebraicznego zagadnienia własnego, wykorzystując siedmiopunktowe przybliżenie drugiej pochodnej z zadania 5/10.

5/12. Wykorzystując twierdzenie Martina–Deana, sformułować algorytm pozwalający wyznaczać wartości własne AZW z poprzedniego zadania metodą bisekcji.

5/13. Wyprowadzić uogólnione macierzowe zagadnienie własne w metodzie Lindberga — wzory (5.43)–(5.45).

5/14. Pokazać, że wyrazy macierzy symetrycznej T podobnej do T' wyrażają się wzorami (5.49).

5/15. Wyznaczyć numerycznie za pomocą metody Guardioli–Rosa dla aproksymanty Padégo $h^2 D^2[4/0]$ dziesięć najniższych poziomów energetycznych i odpowiadających im funkcji falowych cząstki kwantowej o masie M poruszającej się w polu potencjału $V(X) = \frac{1}{2}CX^2 = \frac{1}{2}M\Omega^2 X^2$, gdzie parametr $\Omega = \sqrt{C/M}$.

5/16. Zastosować ekstrapolację Richardsona do zwiększenia dokładności wyznaczania energii i funkcji falowych oscylatora harmonicznego metodą trójpunktową. Porównać otrzymane wyniki z wynikami wybranej metody wyższego rzędu.

Metody numeryczne rozwiązywania symetrycznego gęstego algebraicznego zagadnienia własnego

Przedmiotem naszego zainteresowania będzie w tym rozdziale symetryczne gęste algebraiczne zagadnienie własne (AZW)

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}, \quad (6.1)$$

którego macierz $\mathbf{A} = \mathbf{A}^T$ jest gęsta i ma rozmiar $n \times n$. Liczba niezerowych wyrazów tej macierzy jest rzędu n^2 .

Przedstawimy i szczegółowo omówimy najefektywniejsze z obecnie znanych algorytmów stosowanych do numerycznego rozwiązywania zagadnienia (6.1). Wydaje się, że stanowią one najważniejszą dziedzinę współczesnej numerycznej algebry liniowej [41–53]. Ich rozumienie i sprawne posługiwanie się nimi powinno należeć do kanonu wiedzy i umiejętności absolwentów uniwersyteckich kierunków przyrodniczych oraz wyższych uczelni technicznych. W szczególności dotyczy to studentów kierunków lub specjalności *Fizyka komputerowa*.

Prawie wszystkie numeryczne metody rozwiązywania AZW (6.1) wykonywane są w dwóch etapach. Są nimi:

1. *Redukcja* za pomocą odpowiednich transformacji wejściowej macierzy \mathbf{A} do postaci macierzy specjalnej zawierającej dużą liczbę wyrazów zerowych (*macierzy rozrzedzonej*). Przykładem macierzy specjalnej jest symetryczna macierz trójdzielna oraz *macierz Hessenberga* [21, 41].
2. Numeryczne rozwiązanie AZW z macierzą specjalną.

Wynikiem redukcji symetrycznej rzeczywistej macierzy gęstej \mathbf{A} jest symetryczna macierz trójdzielna \mathbf{T} :

$$\mathbf{Q}^T \mathbf{A} \mathbf{Q} = \mathbf{T} = \begin{pmatrix} d_1 & e_1 & & & \\ e_1 & \ddots & \ddots & & \\ & \ddots & \ddots & e_{n-1} & \\ & & e_{n-1} & d_n & \end{pmatrix}, \quad (6.2)$$

gdzie \mathbf{Q} oznacza macierz ortogonalną. Wynikiem redukcji macierzy niesymetrycznej jest macierz Hessenberga [41–53].

Widmo macierzy \mathbf{T} (wartości własne $\lambda_1, \dots, \lambda_n$) jest takie samo, jak macierzy \mathbf{A} , ponieważ są one podobne (patrz twierdzenie B-11). Wektory własne \mathbf{x}_i macierzy \mathbf{A} możemy łatwo otrzymać, znając wektory własne \mathbf{y}_i macierzy \mathbf{T} : $\mathbf{x}_i = \mathbf{Q}\mathbf{y}_i$.

Konieczność wydzielenia dwóch etapów podyktowana jest specyfiką rozwiązywanego zagadnienia własnego, sprowadzającego się do wyznaczenia zer wielomianu charakterystycznego. W ogólnym przypadku (tj. dla $n \geq 5$) nie istnieją zamknięte ana-

lityczne formuły na pierwiastki wielomianu n -tego stopnia. Dlatego pierwiastki te wyznaczamy za pomocą odpowiednich procedur iteracyjnych wykonywanych na macierzy \mathbf{A} .

Działania na macierzach gęstych wymagają dużej liczby operacji zmiennoprzecinkowych (LOZ). Wystarczy zauważyć, że przy operacjach takich jak mnożenie czy odwracanie macierzy, LOZ jest rzędu n^3 . Operacje na macierzach specjalnych są natomiast o wiele mniej kosztowne w sensie LOZ, np. dla macierzy trójdzielnych LOZ jest rzędu n . Odnosimy jednak, że redukcja macierzy gęstej do postaci specjalnej wymaga $O(n^3)$ operacji.

W rozdziale tym przedstawimy najpierw algorytmy numeryczne, za pomocą których wyznaczamy wszystkie wartości własne rzeczywistej symetrycznej macierzy trójdzielnej. Jak się okaże, najefektywniej realizują to algorytmy QR/QL oraz „dziel i rządź”, które szczegółowo scharakteryzujemy. Następnie omówimy metodę Householdera stosowaną do redukcji macierzy gęstej do postaci trójdzielnej oraz metodę Givensa stosowaną do macierzy pasmowych. Opiszemy również kilka innych metod diagonalizacji macierzy.

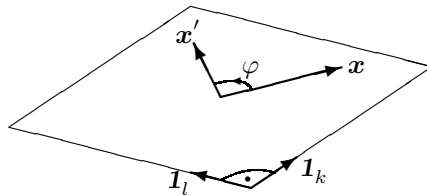
Zanim zajmiemy się bliżej wspomnianymi algorytmami, opiszemy przekształcenia ortogonalnego podobieństwa, którymi będziemy się posługiwać.

6.1. Elementarne ortogonalne przekształcenia podobieństwa

Redukcję macierzy \mathbf{A} do postaci specjalnej, jaką jest macierz trójdzielna (6.2) oraz numeryczne wyznaczanie jej wartości własnych prowadzi się z wykorzystaniem ortogonalnych przekształceń podobieństwa zdefiniowanych w Dodatku B. Poniżej przedstawiamy najprostsze takie przekształcenia, którymi są *obroty Jacobiego* i *Givensa* oraz *odbicia Householdera*. Transformacje te będziemy dalej nazywali elementarnymi ortogonalnymi przekształceniami podobieństwa.

6.1.1. Obroty Jacobiego i Givensa

Transformacje Jacobiego i Givensa odpowiadają obrotowi wektora o kąt φ w płaszczyźnie $X_k O X_l$ n -wymiarowej przestrzeni (\mathbb{R}^n), jak pokazuje rys. 6.1.



Rys. 6.1. Obrót w płaszczyźnie $X_k O X_l$

Obrót Jacobiego reprezentuje macierz \mathbf{J} postaci

$$\mathbf{J} = \begin{matrix} & \begin{matrix} 1 & \cdots & k & \cdots & l & \cdots & n \end{matrix} \\ \begin{matrix} 1 \\ \vdots \\ k \\ \vdots \\ l \\ \vdots \\ n \end{matrix} & \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & \cos \varphi & & -\sin \varphi & & \\ & & & \ddots & & & \\ & & \sin \varphi & & \cos \varphi & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \end{matrix}. \quad (6.3)$$

Jak widać, jej niezerowymi wyrazami są (wybieramy $k < l$):

$$J_{k,k} = J_{l,l} = \cos \varphi, \quad J_{k,l} = -J_{l,k} = -\sin \varphi, \quad J_{i,i} = 1 \quad (i \neq k, l).$$

Łatwo jest sprawdzić, że macierz \mathbf{J} jest ortogonalna, tj.

$$\mathbf{J}^T \mathbf{J} = \mathbf{J} \mathbf{J}^T = \mathbf{1}, \quad (6.4)$$

Ogólna postać macierzy obrotu Givensa \mathbf{G} odpowiada macierzy Jacobiego (6.3). Rozróżnianie obrotów Jacobiego i Givensa ma charakter tradycyjny i wiąże się z ich odmiennym zastosowaniem w różnych algorytmach.

Przytoczmy postać macierzy Givensa \mathbf{G}_i , którą będziemy posługiwali się w celu wyzerowania za pomocą transformacji podobieństwa wyrazu pozadiagonalnego $T_{i+1,i}$ macierzy trójdzielnej (6.2). Odpowiednia macierz Givensa ma postać

$$\mathbf{G}_i = \begin{matrix} & \begin{matrix} 1 & \cdots & i & i+1 & \cdots & n \end{matrix} \\ \begin{matrix} 1 \\ \vdots \\ i \\ i+1 \\ \vdots \\ n \end{matrix} & \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & c_i & -s_i & & \\ & & s_i & c_i & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix} \end{matrix}, \quad (6.5)$$

gdzie $c_i = \cos \varphi$ oraz $s_i = \sin \varphi$.

Do dalszych celów potrzebna nam będzie znajomość wyniku mnożenia $\mathbf{G}_i^T \mathbf{T}$:

$$\begin{pmatrix} \ddots & & & & & \\ & 1 & & & & \\ & & c_i & s_i & & \\ & & -s_i & c_i & & \\ & & & & 1 & \\ & & & & & \ddots \end{pmatrix} \begin{pmatrix} \ddots & \ddots & & & & \\ \ddots & d_{i-1} & e_{i-1} & & & \\ & e_{i-1} & d_i & e_i & & \\ & & e_i & d_{i+1} & e_{i+1} & \\ & & & e_{i+1} & d_{i+2} & \ddots \\ & & & & \ddots & \ddots \end{pmatrix}.$$

W rezultacie otrzymujemy macierz

$$\begin{matrix} & i-1 & i & i+1 & i+2 \\ \begin{matrix} i-1 \\ i \\ i+1 \\ i+2 \end{matrix} & \begin{pmatrix} \ddots & & & \\ & \ddots & & \\ & & d_{i-1} & e_{i-1} \\ & & e'_{i-1} & d'_i & e'_i & g'_i \\ & & f'_{i-1} & E'_i & d'_{i+1} & e'_{i+1} \\ & & & & e_{i+1} & d_{i+2} & \ddots \\ & & & & & \ddots & \ddots \end{pmatrix} \end{matrix},$$

gdzie

$$\begin{aligned} e'_{i-1} &= c_i e_{i-1}, & d'_i &= c_i d_i + s_i e_i, & e'_i &= c_i e_i + s_i d_{i+1}, & g'_i &= s_i e_{i+1}, \\ f'_{i-1} &= -s_i e_{i+1}, & E'_i &= -s_i d_i + c_i e_i, & d'_{i+1} &= -s_i e_i + c_i d_{i+1}, & e'_{i+1} &= c_i e_{i+1}. \end{aligned}$$

Kąt obrotu Givensa φ wybieramy w taki sposób, aby otrzymać $E'_i = 0$, co prowadzi do następujących wartości funkcji trygonometrycznych kąta φ :

$$s_i = \frac{e_i}{\sqrt{d_i^2 + e_i^2}}, \quad c_i = \frac{d_i}{\sqrt{d_i^2 + e_i^2}}. \quad (6.6)$$

Zauważmy, że tak dokonana transformacja *psuje* trójdziagonalną strukturę macierzy \mathbf{T} (6.2), ponieważ pojawiają się dodatkowe wyrazy poza głównymi nad- i poddiagonalami, którymi są $g'_i = s_i e_{i+1}$, $f'_{i-1} = -s_i e_{i+1}$.

6.1.2. Odbicie Householdera

Niech \mathbf{w} będzie wektorem wyznaczającym pewną hiperpłaszczyznę Π w przestrzeni \mathbb{R}^n (wektorem normalnym). Odbicie względem tej hiperpłaszczyzny (rys. 6.2) jest przekształceniem ortogonalnym opisanym macierzą

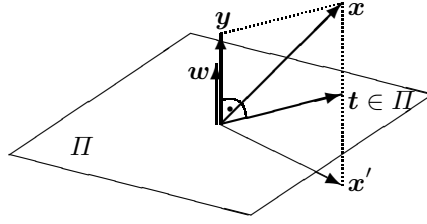
$$\mathbf{Z} = \mathbf{I} - 2\mathbf{w} \otimes \mathbf{w} = \mathbf{I} - 2\mathbf{w}\mathbf{w}^T = \begin{pmatrix} 1 - 2w_1^2 & -2w_1w_2 & \cdots & -2w_1w_n \\ -2w_2w_1 & 1 - 2w_2^2 & \cdots & -2w_2w_n \\ \vdots & \vdots & \ddots & \vdots \\ -2w_nw_1 & -2w_nw_2 & \cdots & 1 - 2w_n^2 \end{pmatrix}, \quad (6.7)$$

gdzie wektor $\mathbf{w} = (w_1, \dots, w_n)^T$ jest unormowany, tj. $\mathbf{w}^T \mathbf{w} = 1$, a symbol \otimes oznacza iloczyn zewnętrzny (patrz dodatek B).

Tak zdefiniowana macierz \mathbf{Z} jest symetryczna ($\mathbf{Z} = \mathbf{Z}^T$), co wynika bezpośrednio z jej definicji (6.7), oraz ortogonalna ($\mathbf{Z}^{-1} = \mathbf{Z}^T$), co uzasadniamy bezpośrednim rachunkiem

$$\mathbf{Z}^T \mathbf{Z} = \mathbf{Z}^2 = \mathbf{I} - 4\mathbf{w} \otimes \mathbf{w} + 4(\mathbf{w} \otimes \mathbf{w})(\mathbf{w} \otimes \mathbf{w}) = \mathbf{I} - 4\mathbf{w}\mathbf{w}^T + 4\mathbf{w}(\mathbf{w}^T \mathbf{w})\mathbf{w}^T = \mathbf{I},$$

gdzie uwzględniono relację $\mathbf{w}^T \mathbf{w} = 1$.



Rys. 6.2. Przekształcenie Householdera: $y = (w \otimes w)x = w(w \cdot x)$ — rzut x na kierunek w , $t = (1 - w \otimes w)x = x - y$ — rzut x na Π , $x' = (1 - 2w \otimes w)x = x - 2y$ — odbicie x względem Π

6.2. Metoda QR

Przedstawimy teraz iteracyjną metodę wyznaczania wszystkich wartości własnych macierzy \mathbf{T} (6.2), noszącą nazwę *metody QR* [41–53]. Rozpocniemy od kilku wprowadzających uwag o charakterze matematycznym.

Jeśli $\mathbf{A} = \mathbf{QR}$ i $\mathbf{B} = \mathbf{RQ}$, gdzie \mathbf{Q} jest macierzą ortogonalną, a \mathbf{R} — macierzą trójkątną górną, to $\mathbf{R} = \mathbf{Q}^T \mathbf{A}$ i

$$\mathbf{B} = \mathbf{RQ} = \mathbf{Q}^T \mathbf{A} \mathbf{Q}.$$

Oznacza to, że macierz \mathbf{B} jest podobna do macierzy \mathbf{A} .

Wyobraźmy sobie teraz, że dla danej wejściowej macierzy $\mathbf{A} = \mathbf{A}_1$ tworzymy ciąg macierzy do niej podobnych zgodnie z następującą regułą:

$$\mathbf{A}_l = \mathbf{Q}_l \mathbf{R}_l, \quad \mathbf{A}_{l+1} = \mathbf{R}_l \mathbf{Q}_l = \mathbf{Q}_l^T \mathbf{A}_l \mathbf{Q}_l \quad (l = 1, 2, \dots). \quad (6.8)$$

U podstaw takiej procedury leży fakt, że każdą symetryczną macierz rzeczywistą można przedstawić w postaci iloczynu $\mathbf{A} = \mathbf{QR}$ (patrz twierdzenia B-21 i B-22).

W praktyce, podczas obliczeń numerycznych za pomocą algorytmu QR dokonuje się dodatkowych zabiegów polegających na:

- Przesuwaniu widma macierzy wejściowej o wartość z_l , co odpowiada zamianie $\mathbf{A}_l \mapsto \mathbf{A}_l - z_l \mathbf{1}$. Bez przesunięcia wyrazy niediagonalne $(\mathbf{A}_l)_{i,j}$ dążą do zera jak $\sim (\lambda_i / \lambda_j)^l$, gdzie λ_i i λ_j są wartościami własnymi macierzy \mathbf{A} . Dla $\lambda_i \simeq \lambda_j$ zbieżność iteracji QR jest bardzo powolna. Wartość z_l dobieramy tak, aby iloraz

$$\left(\frac{\lambda_i - z_l}{\lambda_j - z_l} \right)^l \quad (6.9)$$

zapewniał szybszą zbieżność algorytmu. Problem wyboru z_l omówimy dalej.

- Deflacji macierzy \mathbf{A}_l , czyli sukcesywnym zmniejszaniu jej rozmiaru, a więc podziale na bloki, w miarę jak wyrazy niediagonalne stają się dostatecznie małe.

Interesuje nas granica ciągu \mathbf{A}_l przy $l \rightarrow \infty$, tj. własności macierzy \mathbf{A}_∞ . Wilkinson [45] udowodnił twierdzenie dotyczące tej granicy, które jest nazywane *twierdzeniem QR Wilkinsona* (jego dowód można znaleźć w podręcznikach [21, 24]). Nie

będziemy przytaczać odpowiedniego twierdzenia dotyczącego algorytmu QR dla dowolnych (gęstych symetrycznych lub niesymetrycznych) macierzy. W przedstawianej tu strategii rozwiązywania symetrycznego gęstego AZW, algorytm QR z przesunięciem widma stosowany jest w drugim etapie, tj. do macierzy trójdzielnej, w postaci

$$\mathbf{T}_l - z_l \mathbf{1} = \mathbf{Q}_l \mathbf{R}_l, \quad \mathbf{T}_{l+1} = \mathbf{R}_l \mathbf{Q}_l + z_l \mathbf{1}.$$

Dlatego przytoczymy tylko uproszczoną wersję twierdzenia Wilkinsona dla symetrycznej macierzy trójdzielnej.

Twierdzenie 6-1. (QR Wilkinsona) Jeśli \mathbf{T}_1 ma wszystkie wyrazy pozadiagonalne niezerowe¹⁾, to

$$\lim_{l \rightarrow \infty} \mathbf{T}_l = \text{diag}(\lambda_1, \dots, \lambda_n),$$

gdzie λ_i są wartościami własnymi \mathbf{T}_1 .

Pokazuje się [21, 24], że jeżeli wykonujemy przesuwanie widma w sposób zaproponowany przez Wilkinsona, wyrazy pozadiagonalne $e_i^{(l)}$ macierzy \mathbf{T}_l zawsze dążą do zera. Zbieżność ta jest co najmniej kwadratowa, tj.

$$|e_i^{(l+1)}| \leq c |e_i^{(l)}|^2, \quad c > 0.$$

W większości konkretnych zagadnień zbieżność jest co najmniej sześcienna.

Jeśli macierz \mathbf{T}_1 ma co najmniej jeden zerowy wyraz pozadiagonalny, to zagadnienie własne rozpada się na dwa lub więcej zagadnień własnych mniejszych rozmiarów, do których stosuje się już twierdzenie 6-1. Odpowiada to dokonaniu deflacji macierzy \mathbf{T}_1 .

W praktyce obliczeń numerycznych postępujemy w ten sposób, że w każdym kroku procesu iteracji (6.8) z $\mathbf{A}_l = \mathbf{T}_l$ wyznaczamy macierz \mathbf{Q}_l^T spełniającą zależność

$$\mathbf{Q}_l^T \mathbf{T}_l = \mathbf{R}_l,$$

tj. wykonujemy takie przekształcenia ortogonalne, aby macierz \mathbf{T}_l doprowadzić do postaci macierzy trójkątnej górnej \mathbf{R}_l . Ten etap obliczeń nazywany jest *dekompozycją*. Nie obliczamy przy tym macierzy \mathbf{Q}_l^T w jawnej postaci, ale przedstawiamy ją jako iloczyn elementarnych transformacji, którymi zazwyczaj są obroty Givensa (6.5).

Następnie, znając postać \mathbf{R}_l , wyznaczamy macierz $\mathbf{T}_{l+1} = \mathbf{R}_l \mathbf{Q}_l = \mathbf{Q}_l^T \mathbf{T}_l \mathbf{Q}_l$, mnożąc prawostronnie \mathbf{R}_l przez \mathbf{Q}_l . Ten etap obliczeń nosi nazwę *rekombinacji*.

Procedurę tę powtarzamy wielokrotnie aż do momentu, gdy wyrazy pozadiagonalne macierzy \mathbf{T}_{l+1} nie staną się dostatecznie małe (tj. są co do wartości bezwzględnej mniejsze od liczby zadanej przez użytkownika lub epsilon maszynowego). Wówczas przerywamy obliczenia i za wartości własne przyjmujemy wyrazy stojące na diagonalu macierzy \mathbf{T}_{l+1} .

Wykonanie jednego kroku iteracji QR z przesuwaniem widma dla macierzy \mathbf{T} wymaga $6n$ operacji zmiennoprzecinkowych. Dla wyznaczenia metodą QR wszystkich wartości własnych \mathbf{T} wymagane jest $6n^2$ operacji. Jeżeli wyznaczamy również wektory własne, to całkowita LOZ wynosi $\sim 6n^3$. Redukcja macierzy \mathbf{A} do postaci

¹⁾ A zatem niezdegenerowane (pojedyncze) wartości własne — por. twierdzenie B-18.

trójdzielnej wymaga wykonania $\sim \frac{4}{3}n^3$ lub $\sim \frac{8}{3}n^3$ operacji zmiennoprzecinkowych, w zależności od tego, czy są wyznaczane tylko wartości własne, czy także wektory własne. Całkowity koszt wyznaczenia wszystkich wartości i wektorów własnych rzeczywistej symetrycznej macierzy gęstej wynosi zatem $\sim \frac{26}{3}n^3$ operacji, a samych wartości własnych — jedynie $\sim \frac{4}{3}n^3$ operacji. Występujące w powyższych oszacowaniach stałe współczynniki mogą zależeć od szczegółów implementacji algorytmów; podajemy je za podręcznikiem [50].

Gdybyśmy rozwiązywali symetryczne gęste AZW bez wykonywania pierwszego etapu strategii, to liczba operacji zmiennoprzecinkowych byłaby rzędu n^4 , ponieważ pojedyncza iteracja QR dla macierzy gęstej wymaga wykonania $O(n^3)$ operacji [50].

6.3. Algorytm QR

Przedstawimy obecnie szczegółowo pojedynczy krok algorytmu QR w przypadku trójdzielnej rzeczywistej macierzy symetrycznej postaci (6.2).

Zgodnie z tym, co powiedzieliśmy poprzednio, jeden krok algorytmu QR polega na wykonaniu:

1. Ortogonalno-trójkątnej *dekompozycji* wejściowej macierzy do postaci macierzy górnej trójkątnej \mathbf{R} :

$$\mathbf{Q}^T \mathbf{T} = \mathbf{R}. \quad (6.10)$$

2. *Rekombinacji* macierzy \mathbf{R} w odwrotnym porządku:

$$\tilde{\mathbf{T}} = \mathbf{R}\mathbf{Q}. \quad (6.11)$$

6.3.1. Dekompozycja

Dekompozycję macierzy \mathbf{T} wykonujemy za pomocą $n - 1$ obrotów Givensa

$$\mathbf{G}_{n-1}^T \cdots \mathbf{G}_1^T \mathbf{T},$$

gdzie macierze \mathbf{G}_i mają postać (6.5).

Rezultatem wykonania wszystkich $n - 1$ obrotów Givensa jest macierz trójkątna górna \mathbf{R} postaci

$$\mathbf{R} = \begin{pmatrix} r_1 & q_1 & v_1 & & & \\ & r_2 & q_2 & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & q_{n-2} & v_{n-2} \\ & & & & r_{n-1} & q_{n-1} \\ & & & & & r_n \end{pmatrix}. \quad (6.12)$$

W pierwszym kroku procedury dekompozycji wybieramy kąt obrotu \mathbf{G}_1^T tak, aby zniknął wyraz $(\mathbf{G}_1^T \mathbf{T})_{2,1}$. W wyniku obrotu dostajemy

$$\mathbf{G}_1^T \mathbf{T} = \begin{pmatrix} c_1 & s_1 & & \\ -s_1 & c_1 & & \\ & & 1 & \\ & & & \ddots \end{pmatrix} \begin{pmatrix} d_1 & e_1 & & \\ e_1 & d_2 & e_2 & \\ & e_2 & d_3 & \ddots \\ & & \ddots & \ddots \end{pmatrix} = \begin{pmatrix} r_1 & q_1 & v_1 & \\ & p_2 & w_2 & \\ & e_2 & d_3 & \ddots \\ & & \ddots & \ddots \end{pmatrix},$$

gdzie, przyjmując

$$p_1 = d_1, \quad c_0 = 1,$$

po prostych przekształceniach algebraicznych otrzymujemy:

$$\begin{aligned} r_1 &= c_1 p_1 + s_1 e_1 = \sqrt{p_1^2 + e_1^2}, & c_1 &= p_1/r_1, & s_1 &= e_1/r_1, \\ q_1 &= c_1 c_0 e_1 + s_1 d_2, & v_1 &= s_1 e_2, \\ p_2 &= c_1 d_2 - s_1 c_0 e_1, & w_2 &= c_1 e_2. \end{aligned} \quad (6.13)$$

Kąt obrotu \mathbf{G}_2^T wybieramy tak, aby zniknął wyraz $(\mathbf{G}_2^T \mathbf{G}_1^T \mathbf{T})_{3,2}$. Wynikiem działania jest macierz $\mathbf{G}_2^T \mathbf{G}_1^T \mathbf{T}$ postaci

$$\begin{pmatrix} 1 & & & \\ & c_2 & s_2 & \\ & -s_2 & c_2 & \\ & & & 1 \\ & & & & \ddots \end{pmatrix} \begin{pmatrix} r_1 & q_1 & v_1 & & \\ & p_2 & w_2 & & \\ & e_2 & d_3 & e_3 & \\ & & e_3 & d_4 & \ddots \\ & & & \ddots & \ddots \end{pmatrix} = \begin{pmatrix} r_1 & q_1 & v_1 & & \\ & r_2 & q_2 & v_2 & \\ & & p_3 & w_3 & \\ & & e_3 & d_4 & \ddots \\ & & & \ddots & \ddots \end{pmatrix},$$

gdzie, jak łatwo pokazać,

$$\begin{aligned} r_2 &= c_2 p_2 + s_2 e_2 = \sqrt{p_2^2 + e_2^2}, & c_2 &= p_2/r_2, & s_2 &= e_2/r_2, \\ q_2 &= c_2 c_1 e_2 + s_2 d_3, & v_2 &= s_2 e_3, \\ p_3 &= c_2 d_3 - s_2 c_1 e_2, & w_3 &= c_2 e_3. \end{aligned} \quad (6.14)$$

W wyniku obrotu \mathbf{G}_i^T otrzymujemy $\mathbf{G}_i^T \cdots \mathbf{G}_1^T \mathbf{T}$, tj.

$$\begin{pmatrix} \ddots & & & \\ & c_i & s_i & \\ & -s_i & c_i & \\ & & & \ddots \end{pmatrix} \begin{pmatrix} \ddots & \ddots & \ddots & \\ & p_i & w_i & 0 \\ & e_i & d_{i+1} & e_{i+1} \\ & & \ddots & \ddots & \ddots \end{pmatrix} = \begin{pmatrix} \ddots & \ddots & \ddots & \\ & r_i & q_i & v_i \\ & 0 & p_{i+1} & w_{i+1} \\ & & \ddots & \ddots & \ddots \end{pmatrix},$$

gdzie

$$\begin{aligned} r_i &= c_i p_i + s_i e_i = \sqrt{p_i^2 + e_i^2}, & c_i &= p_i/r_i, & s_i &= e_i/r_i \\ q_i &= c_i c_{i-1} e_i + s_i d_{i+1}, & v_i &= s_i e_{i+1}, \\ p_{i+1} &= c_i d_{i+1} - s_i c_{i-1} e_i, & w_{i+1} &= c_i e_{i+1}. \end{aligned} \quad (6.15)$$

Postępując analogicznie w $(n-1)$ -ym kroku, wyzerujemy ostatni wyraz poddiagonalny i otrzymamy macierz trójkątną górną ($\mathbf{G}_{n-1}^T \cdots \mathbf{G}_1^T \mathbf{T} = \mathbf{R}$):

$$\begin{pmatrix} \ddots & & & \\ & 1 & & \\ & c_{n-1} & s_{n-1} & \\ & -s_{n-1} & c_{n-1} & \end{pmatrix} \begin{pmatrix} \ddots & \ddots & \ddots & \\ & r_{n-2} & q_{n-2} & v_{n-2} \\ & & p_{n-1} & w_{n-1} \\ & & e_{n-1} & d_n \end{pmatrix} = \begin{pmatrix} \ddots & \ddots & \ddots & \\ & r_{n-2} & q_{n-2} & v_{n-2} \\ & & r_{n-1} & q_{n-1} \\ & & & r_n \end{pmatrix}.$$

Otrzymana macierz ma następujące elementy:

$$\begin{aligned} r_{n-1} &= c_{n-1}p_{n-1} + s_{n-1}e_{n-1} = \sqrt{p_{n-1}^2 + e_{n-1}^2}, \\ c_{n-1} &= p_{n-1}/r_{n-1}, \quad s_{n-1} = e_{n-1}/r_{n-1}, \\ q_{n-1} &= c_{n-1}c_{n-2}e_{n-1} + s_{n-1}d_n, \\ p_n &= c_{n-1}d_n - s_{n-1}c_{n-2}e_{n-1} = r_n. \end{aligned}$$

Ostatecznie algorytm dekompozycji (6.10), sprowadzający macierz \mathbf{T} (6.2) do postaci macierzy trójkątnej górnej \mathbf{R} (6.12), w pojedynczym kroku metody QR można sformułować następująco:

Algorytm 6-2. (dekompozycji)

Procedura *Dekompozycja-QR*($\mathbf{d}, \mathbf{e}, \mathbf{p}, \mathbf{q}, \mathbf{r}, \mathbf{c}, \mathbf{s}, n$):

$p_1 := d_1; c_0 := 1;$

Dla $i = 1, \dots, n-1$:

$r_i := \sqrt{p_i^2 + e_i^2}; c_i := p_i/r_i; s_i := e_i/r_i;$

$q_i := c_i c_{i-1} e_i + s_i d_{i+1}; p_{i+1} := c_i d_{i+1} - s_i c_{i-1} e_i; v_i := s_i e_{i+1};$

$r_n := p_n;$

Koniec.

Odnajdujemy, że w rzeczywistości w trakcie obliczeń według algorytmu QR nie musimy zapamiętywać wszystkich liczb p_i, q_i, c_i, s_i , ani w ogóle wyznaczać wyrazów v_i . Ich wartości podano tu dla ułatwienia zrozumienia pojedynczego kroku algorytmu.

6.3.2. Rekombinacja

Przejdziemy obecnie do przedstawienia procesu rekombinacji (6.11), tj. do przywracania macierzy trójkątnej górnej \mathbf{R} (6.12) pierwotnej postaci macierzy trójdzielnej $\tilde{\mathbf{T}} = \mathbf{R}\mathbf{Q}$. W tym celu będziemy sukcesywnie mnożyć macierz \mathbf{R} (6.12) prawostronnie przez macierze Givensa \mathbf{G}_i o postaci (6.5).

Rozpatrzmy obecnie kilka początkowych obrotów. Napiszmy wynik pierwszego mnożenia:

$$\mathbf{R}\mathbf{G}_1 = \begin{pmatrix} r_1 & q_1 & v_1 & \\ & r_2 & q_2 & \ddots \\ & & r_3 & \ddots \\ & & & \ddots \end{pmatrix} \begin{pmatrix} c_1 & -s_1 & & \\ s_1 & c_1 & & \\ & & 1 & \\ & & & \ddots \end{pmatrix} = \begin{pmatrix} \tilde{d}_1 & q'_1 & v_1 & \\ \tilde{e}_1 & \tilde{d}'_2 & q_2 & \ddots \\ & & r_3 & \ddots \\ & & & \ddots \end{pmatrix},$$

gdzie

$$\tilde{d}_1 = r_1 c_1 c_0 + q_1 s_1, \quad q'_1 = q_1 c_1 - s_1 r_1, \quad \tilde{e}_1 = r_2 s_1, \quad d'_2 = r_2 c_1.$$

Wynikiem obrotu \mathbf{G}_2 jest macierz $\mathbf{RG}_1\mathbf{G}_2$ postaci

$$\begin{pmatrix} \tilde{d}_1 & q'_1 & v_1 & & \\ \tilde{e}_1 & d'_2 & q_2 & v_2 & \\ & & r_3 & q_3 & \ddots \\ & & & r_4 & \ddots \\ & & & & \ddots \end{pmatrix} \begin{pmatrix} 1 & & & & \\ & c_2 & -s_2 & & \\ & s_2 & c_2 & & \\ & & & 1 & \\ & & & & \ddots \end{pmatrix} = \begin{pmatrix} \tilde{d}_1 & \tilde{q}_1 & \tilde{v}_1 & & \\ \tilde{e}_1 & \tilde{d}_2 & q'_2 & v_2 & \\ & \tilde{e}_2 & d'_3 & q_3 & \ddots \\ & & & r_4 & \ddots \\ & & & & \ddots \end{pmatrix},$$

o wyrazach

$$\begin{aligned} \tilde{q}_1 &= (q_1 c_1 - s_1 r_1) c_2 + s_2 v_1 = \tilde{e}_1, & \tilde{v}_1 &= c_2 v_1 - s_2 (q_1 c_1 - s_1 r_1) = 0, \\ \tilde{d}_2 &= r_2 c_1 c_2 + q_2 s_2, & q'_2 &= -r_2 c_1 s_2 + q_2 c_2 \\ \tilde{e}_2 &= r_3 s_2, & d'_3 &= r_3 c_2. \end{aligned}$$

Równości $\tilde{v}_1 = 0$ i $\tilde{q}_1 = \tilde{e}_1$ wynikają z warunku zachowywania symetryczności macierzy przez transformację podobieństwa.

W wyniku obrotu \mathbf{G}_3 otrzymujemy macierz $\mathbf{RG}_1\mathbf{G}_2\mathbf{G}_3$ postaci

$$\begin{pmatrix} \tilde{d}_1 & \tilde{e}_1 & & & \\ \tilde{e}_1 & \tilde{d}_2 & q'_2 & v_2 & \\ & \tilde{e}_2 & d'_3 & q_3 & \ddots \\ & & & r_4 & \ddots \\ & & & & \ddots \end{pmatrix} \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & c_3 & -s_3 & \\ & & s_3 & c_3 & \\ & & & & \ddots \end{pmatrix} = \begin{pmatrix} \tilde{d}_1 & \tilde{e}_1 & & & \\ \tilde{e}_1 & \tilde{d}_2 & \tilde{q}_2 & \tilde{v}_2 & \\ & \tilde{e}_2 & \tilde{d}_3 & q'_3 & \ddots \\ & & \tilde{e}_3 & d'_4 & \ddots \\ & & & & \ddots \end{pmatrix},$$

gdzie

$$\begin{aligned} \tilde{q}_2 &= q'_2 c_3 + v_2 s_3 = q_2 c_2 c_3 - r_2 c_1 s_2 c_3 + s_3 v_2 = \tilde{e}_2 \\ \tilde{v}_2 &= v_2 c_3 - q'_2 s_3 = c_3 v_2 - s_3 (q_2 c_2 - r_2 c_1 s_2) = 0, \\ \tilde{d}_3 &= d'_3 c_3 + q_3 s_3 = r_3 c_2 c_3 + q_3 s_3, \\ q'_3 &= c_3 q_3 - d'_3 s_3 = c_3 q_3 - r_3 c_2 s_3, \\ \tilde{e}_3 &= r_4 s_3, & d'_4 &= r_4 c_3. \end{aligned}$$

W i -tym kroku procedury rekombinacji dostajemy macierz $\mathbf{RG}_1 \cdots \mathbf{G}_i$, tj.

$$\begin{pmatrix} \ddots & & & & \\ \ddots & q'_{i-1} & v_{i-1} & & \\ \ddots & d'_i & q_i & \ddots & \\ & & r_{i+1} & \ddots & \\ & & & \ddots & \end{pmatrix} \begin{pmatrix} \ddots & & & & \\ & c_i & -s_i & & \\ & s_i & c_i & & \\ & & & \ddots & \end{pmatrix} = \begin{pmatrix} \ddots & & & & \\ \ddots & \tilde{q}_{i-1} & \tilde{v}_{i-1} & & \\ \ddots & \tilde{d}_i & q'_i & \ddots & \\ & \tilde{e}_i & d'_{i+1} & \ddots & \\ & & & \ddots & \end{pmatrix},$$

której wyrazy dane są wzorami

$$\begin{aligned}\tilde{q}_{i-1} &= q'_{i-1}c_i + v_{i-1}s_i = q_{i-1}c_{i-1}c_i - r_{i-1}c_{i-2}s_{i-1}c_i + s_iv_{i-1}, \\ \tilde{v}_{i-1} &= v_{i-1}c_i - q'_{i-1}s_i = c_iv_{i-1} - s_i(q_{i-1}c_{i-1} - r_{i-1}c_{i-2}s_{i-1}), \\ \tilde{d}_i &= d'_ic_i + q_is_i = r_ic_{i-1}c_i + q_is_i, \\ q'_i &= c_iq_i - d'_is_i = c_iq_i - r_ic_{i-1}s_i, \\ \tilde{e}_i &= r_{i+1}s_i, & d'_{i+1} &= r_{i+1}c_i.\end{aligned}$$

Można sprawdzić, że dla $i = 2, \dots, n$ spełnione są relacje

$$\tilde{q}_{i-1} = \tilde{e}_{i-1}, \quad \tilde{v}_{i-1} = 0. \quad (6.16)$$

Na podstawie przeprowadzonych rozważań proces rekombinacji (6.11) w pojedynczym kroku metody QR możemy sformułować w postaci algorytmu:

Algorytm 6-3. (rekombinacji)

Procedura *Rekombinacja-QR*($\mathbf{d}, \mathbf{e}, \mathbf{r}, \mathbf{q}, \mathbf{c}, \mathbf{s}, n$):

Dla $i = 1, \dots, n-1$:

$$\begin{aligned}\tilde{d}_i &:= r_ic_{i-1}c_i + q_is_i; \quad \tilde{e}_i := r_{i+1}s_i; \\ \tilde{d}_n &:= r_nc_{n-1};\end{aligned}$$

Koniec.

6.3.3. Przesunięcie widma

W celu przyspieszenia procesu zbieżności algorytmów QR i QL zastosujemy podejście zaproponowane przez Wilkinsona [41, 45]. Przypuśćmy, że lewy górny narożnik diagonalizowanej metodą QR macierzy ma postać

$$\begin{pmatrix} d_1 & e_1 & & \\ e_1 & d_2 & \ddots & \\ & \ddots & \ddots & \ddots \end{pmatrix}.$$

Widmo poddajemy przesunięciu o wartość z_l (6.9), będącą pierwiastkiem równania

$$\begin{vmatrix} d_1 - \lambda & e_1 \\ e_1 & d_2 - \lambda \end{vmatrix} = 0,$$

leżącym bliżej d_1 . Jego rozwiązaniami są:

$$\lambda_1 = \frac{d_1 + d_2 - \sqrt{(d_2 - d_1)^2 + 4e_1^2}}{2} = \frac{d_1 + d_2 - 2|e_1|\sqrt{g^2 + 1}}{2}$$

oraz

$$\lambda_2 = \frac{d_1 + d_2 + \sqrt{(d_2 - d_1)^2 + 4e_1^2}}{2} = \frac{d_1 + d_2 + 2|e_1|\sqrt{g^2 + 1}}{2},$$

gdzie $g = (d_2 - d_1)/(2e_1)$.

Przekształćmy nieco otrzymane wyrażenia na pierwiastki:

$$\lambda_1 = d_1 + \frac{d_2 - d_1 - 2|e_1|\sqrt{g^2 + 1}}{2}, \quad \lambda_2 = d_1 + \frac{d_2 - d_1 + 2|e_1|\sqrt{g^2 + 1}}{2},$$

co prowadzi do wzorów

$$\lambda_1 = d_1 - \frac{e_1}{g \operatorname{sgn} e_1 + \sqrt{g^2 + 1}}, \quad \lambda_2 = d_1 - \frac{e_1}{g \operatorname{sgn} e_1 - \sqrt{g^2 + 1}}.$$

Uzasadnimy teraz, że jako wartość przesunięcia widma należy wybrać liczbę

$$z_l = d_1 - \frac{|e_1|}{g + \operatorname{sgn} g \sqrt{g^2 + 1}}. \quad (6.17)$$

W tym celu przyjrzyjmy się następującej tabeli:

$\operatorname{sgn} e_1$	$\operatorname{sgn} g$	z_l
+	+	$d_1 - e_1 / [g + \sqrt{g^2 + 1}]$
+	-	$d_1 - e_1 / [g - \sqrt{g^2 + 1}]$
-	+	$d_1 - e_1 / [g + \sqrt{g^2 + 1}]$
-	-	$d_1 - e_1 / [g - \sqrt{g^2 + 1}]$

Wynika z niej, że wartości z_l (6.17) określają jeden z pierwiastków λ_1 i λ_2 leżący bliżej d_1 . Opisany sposób przesuwania widma jest zaimplementowany w podprogramach TQL1 i TQL2 biblioteki EISPACK.

6.3.4. Niewymierny algorytm QR

Sformułujemy pełny krok algorytmu QR na podstawie algorytmów dekompozycji (6-2) i rekombinacji (6-3). Zauważmy, że można go wykonać w pojedynczej pętli. Zamiast przeprowadzać kolejno dekompozycję i rekombinację, możemy zastosować obroty Givensa w następującej kolejności:

$$\mathbf{A}^{(n-1)} = \mathbf{G}_{n-1}^T \mathbf{T}_l \mathbf{G}_{n-1}, \quad \mathbf{A}^{(n-2)} = \mathbf{G}_{n-2}^T \mathbf{A}^{(n-1)} \mathbf{G}_{n-2}, \quad \dots;$$

$$\mathbf{T}_{l+1} = \mathbf{A}^{(1)} = \mathbf{G}_1^T \mathbf{A}^{(2)} \mathbf{G}_1.$$

Zatem nie musimy przechowywać w pamięci wartości funkcji trygonometrycznych dla wszystkich obrotów Givensa oraz liczb p_i , r_i , q_i . Teraz możemy sformułować algorytm:

Algorytm 6-4. (niewymierny QR)

Procedura *Transformacja_QR*($\mathbf{d}, \mathbf{e}, n$):

$p := d_1$; $c := 1$;

Dla $i = 1, \dots, n-1$:

$g := ce_i$; $h := cp$; $r := \sqrt{p^2 + e_i^2}$;

Jeśli $i > 1$: $e_{i-1} := sr$;

$c := p/r$; $s := e_i/r$; $p := cd_{i+1} - sg$; $d_i := h + s(CG + sd_i)$;

$e_{n-1} := sp$; $d_n := cp$;

Koniec.

Przedstawiony algorytm jest nazywany *niewymiernym algorytmem QR* z uwagi na konieczność obliczania w nim pierwiastków kwadratowych. Zwróćmy uwagę, że po wykonaniu algorytmu 6-4 w miejsce wyrazów macierzy \mathbf{T}_l wpisane zostają wyrazy macierzy \mathbf{T}_{l+1} .

Dodajmy, że w jednym kroku algorytmu QR musimy wykonać $4n$ dodawań, $10n$ mnożeń, $2n$ dzieleni oraz n pierwiastkowań.

6.3.5. Niewymierny algorytm QL

Przeformułujemy nieco nasz algorytm, nadając mu postać *algorytmu QL*, który jest częściej implementowany w bibliotekach procedur numerycznych z uwagi na mniejsze błędy zaokrągleń w trakcie obliczeń (wniosek taki wyciągnięto na podstawie wieloletniej praktyki). W algorytmie QL dekompozycję i rekombinację wykonujemy następująco:

1. Wynikiem dekompozycji $\mathbf{G}_1 \mathbf{G}_2 \cdots \mathbf{G}_{n-1} \mathbf{T}$ macierzy trójdagonalnej jest macierz trójkątna dolna

$$\mathbf{L} = \begin{pmatrix} r_1 & & & & & \\ w_1 & r_2 & & & & \\ z_1 & w_2 & \ddots & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & w_{n-2} & r_{n-1} & \\ & & & z_{n-2} & w_{n-1} & r_n \end{pmatrix}.$$

Tym razem dekompozycję rozpoczynamy od prawego dolnego narożnika macierzy trójdagonalnej, zerując kolejne wyrazy naddiagonalne.

2. Rekombinacji macierzy \mathbf{L} do postaci trójdagonalnej dokonujemy za pomocą przekształceń Givensa

$$\tilde{\mathbf{T}} = \mathbf{L} \mathbf{G}_{n-1}^T \cdots \mathbf{G}_1^T.$$

Postępując w sposób analogiczny do przedstawionego w przypadku niewymiernego algorytmu QR, otrzymujemy ostateczną postać *niewymiernego algorytmu QL*:

Algorytm 6-5. (niewymierny QL)

Procedura *Transformacja_QL*($\mathbf{d}, \mathbf{e}, n$):

$p := d_n; c := 1;$

Dla $i = n, \dots, 2$:

$g := ce_{i-1}; h := cp; r := \sqrt{p^2 + e_{i-1}^2};$

Jeśli $i < n$: $e_i := sr;$

$c := p/r; s := e_{i-1}/r; p := cd_{i-1} - sg; d_i := h + s(CG + sd_{i-1});$

$e_1 := sp; d_1 := cp;$

Koniec.

Opisany wyżej niewymierny algorytm QL jest zaimplementowany (w postaci wykonującej również przesunięcia widma) m.in. w podprogramach biblioteki EISPACK TQL1 i TQL2 (podprogram TQL2 wyznacza również wszystkie wektory własne macierzy).

6.3.6. Obliczanie wektorów własnych macierzy trójdzielnej

Wyznaczanie wektorów własnych można przeprowadzić na dwa sposoby.

1. Znając widmo $\{\lambda_1, \dots, \lambda_n\}$ macierzy trójdzielnej wyznaczone metodą QL lub QR, możemy obliczać współrzędne kolejnych wektorów własnych, stosując metodę DWSZ.
2. Kolumny macierzy \mathbf{Q}^T przekształcenia podobieństwa

$$\mathbf{Q}\mathbf{T}\mathbf{Q}^T = \mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_n),$$

gdzie $\mathbf{Q}^T = \mathbf{G}_{n-1}^T \mathbf{G}_{n-2}^T \dots \mathbf{G}_1^T$, są wektorami własnymi \mathbf{T} (i -ta kolumna odpowiada wartości własnej λ_i , co jest konsekwencją związku $\mathbf{T}\mathbf{Q}^T = \mathbf{Q}^T\mathbf{\Lambda}$).

Pierwszy sposób wymaga użycia tylko $3n$ dodatkowych komórek pamięci operacyjnej (dwa n -wyrazowe ciągi (Θ_i^\pm) oraz n współrzędnych wektora). Obliczony wektor własny można zapisać w pamięci zewnętrznej i zwolnić pamięć operacyjną. Drugi sposób wymaga n^2 komórek pamięci do przechowywania macierzy \mathbf{Q} (względnie \mathbf{Q}^T).

6.3.7. Wymierne algorytmy QR i QL

Konieczność obliczania w każdym kroku pierwiastków kwadratowych wydłuża znacznie czas numerycznego wyznaczania wartości własnych macierzy trójdzielnej za pomocą niewymiernych algorytmów QR lub QL. Okazuje się, że pierwiastkowania można uniknąć. Opiszemy teraz zmodyfikowane w odpowiedni sposób algorytmy, zwane *wymiernymi algorytmami QL i QR*. Przedstawimy wersję opublikowaną w artykule [89]; inne realizacje opisane są w rozdziale 8 podręcznika [48].

Wprowadźmy nowe zmienne:

$$g_i = \frac{p_i}{c_{i-1}}, \quad h_i = p_i c_{i-1}.$$

Zauważmy, że

$$g_{i+1} = \frac{p_{i+1}}{c_i} = d_{i+1} - \frac{s_i c_{i-1} e_i}{c_i} = d_{i+1} - \frac{s_i c_{i-1} e_i^2}{s_i p_i} = d_{i+1} - \frac{e_i^2}{g_i},$$

gdzie uwzględniono związek $p_i s_i = e_i c_i$. Znając g_i możemy wyznaczyć

$$h_{i+1} = p_{i+1} c_i = g_{i+1} c_i^2 = g_{i+1} \frac{p_i^2}{r_i^2},$$

a ponadto

$$\tilde{d}_i = p_i c_{i-1} + q_i s_i = h_i + (c_i c_{i-1} e_i + s_i d_{i+1}) s_i = h_i + s_i^2 (h_i + d_{i+1}),$$

ponieważ $c_{i-1} = h_i / p_i$, $c_i e_i / p_i = s_i$.

Wykorzystując wyżej wyprowadzone wzory możemy obecnie sformułować wymierny algorytm QR:

Algorytm 6-6. (wymierny QR)

Procedura *Wymierna_transformacja_QR*(d, e, n):

$g := h := d_1$;

Dla $i = 1, \dots, n-1$:

$p^2 := gh$; $r^2 := p + e_i^2$;

Jeśli $i > 1$: $e_{i-1}^2 := r^2 s^2$;

$s^2 := e_i^2 / r^2$; $d_i := h + s_i^2(h + d_{i+1})$;

$g := d_{i+1} - e_i^2 / g$; $h := gp^2 / r^2$;

$e_{n-1}^2 := ghs^2$; $d_n := h$;

Koniec.

Zaprezentowany algorytm wymaga $4n$ dodawań, $4n$ mnożeń oraz $3n$ dzieleni w każdym kroku.

W podobny sposób sformułujemy teraz wymierny algorytm QL. Wprowadzamy zmienne

$$g_i = \frac{p_i}{c_i}, \quad h_i = p_i c_i$$

i obliczamy kolejno

$$g_{i-1} = \frac{p_{i-1}}{c_{i-1}} = d_{i-1} - \frac{s_{i-1}c_i e_{i-1}}{c_{i-1}} = d_{i-1} - \frac{e_{i-1}^2 c_i}{p_i} = d_{i-1} - \frac{e_{i-1}^2}{g_i},$$

$$h_i = p_i c_i = g_i c_i^2 = g_i \frac{p_{i+1}}{r_{i+1}^2}.$$

Uwzględniając, że $c_{i+1} = h_{i+1}/p_{i+1}$ i $p_{i+1}s_i = e_i c_i$, mamy

$$\tilde{d}_{i+1} = c_{i+1}p_{i+1} + s_i q_{i+1} = h_{i+1} + s_i(c_i c_{i+1} e_i + s_i d_i) = h_{i+1} + s_i^2(h_{i+1} + d_i).$$

Opierając się na powyższych wzorach oraz niewymiernym algorytmie QL (6-5), możemy sformułować

Algorytm 6-7. (wymierny QL)

Procedura *Wymierna_transformacja_QL*(d, e, n):

$g := h := d_n$;

Dla $i = n, \dots, 2$:

$p^2 := gh$; $r^2 := p^2 + e_{i-1}^2$;

Jeśli $i < n$: $e_i^2 := r^2 s^2$;

$s^2 := e_{i-1}^2 / r^2$; $d_i := h + s^2(h + d_{i-1})$;

$g := d_{i-1} - e_{i-1}^2 / g$; $h := gp^2 / r^2$;

$e_1^2 := ghs^2$; $d_1 := h$.

Koniec.

Przedstawiony algorytm jest zaimplementowany w podprogramach TQLRAT biblioteki EISPACK, a zmodyfikowane wersje powyższych algorytmów QL i QR — w podprogramie SSTERF biblioteki LAPACK [53].

6.3.8. Niejawny algorytm QL/QR

Praktyka obliczeń pokazuje, że przedstawione wyżej algorytmy nie dają poprawnych wyników dla macierzy, wartości wyrazów których różnią się o wiele rzędów. W takich przypadkach stosujemy *niejawny algorytm QL/QR* [41, 45, 50]. Sens słowa *niejawny* polega na tym, że nie obliczamy jawnie rozkładu \mathbf{QL} macierzy \mathbf{T} . Przekształcenie ortogonalne \mathbf{Q} wyznaczamy jako iloczyn obrotów Givensa i innych prostych macierzy ortogonalnych. Opisany algorytm oparty jest na podanym niżej twierdzeniu:

Twierdzenie 6-8. Jeśli \mathbf{A} jest niezdegenerowaną macierzą symetryczną i $\mathbf{T} = \mathbf{Q}^T \mathbf{A} \mathbf{Q}$ — macierzą trójdziagonalną o dodatnich wyrazach pozadiagonalnych, a \mathbf{Q} — macierzą ortogonalną, to macierze \mathbf{T} i \mathbf{Q} są jednoznacznie określone przez ostatni wiersz macierzy \mathbf{Q}^T .

Dowód. Niech \mathbf{q}_i oznacza i -tą kolumnę macierzy \mathbf{Q} . Zapiszmy macierz \mathbf{Q}^T jako n -kolumnowy wektor, którego ostatni wiersz \mathbf{q}_n^T jest znany.

$$\mathbf{Q}^T = \begin{pmatrix} \mathbf{q}_1^T \\ \vdots \\ \mathbf{q}_n^T \end{pmatrix}, \quad \mathbf{Q} = (\mathbf{q}_1, \dots, \mathbf{q}_n), \quad \mathbf{q}_i = \begin{pmatrix} Q_{1,i} \\ \vdots \\ Q_{n,i} \end{pmatrix}.$$

Rozpiszmy obecnie równość $\mathbf{T} \mathbf{Q}^T = \mathbf{Q}^T \mathbf{A}$:

$$\begin{pmatrix} d_1 & e_1 & & & & \\ e_1 & d_2 & \ddots & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & d_{n-1} & e_{n-1} & \\ & & & e_{n-1} & d_n \end{pmatrix} \begin{pmatrix} \mathbf{q}_1^T \\ \mathbf{q}_2^T \\ \vdots \\ \mathbf{q}_{n-1}^T \\ \mathbf{q}_n^T \end{pmatrix} = \begin{pmatrix} \mathbf{q}_1^T \\ \mathbf{q}_2^T \\ \vdots \\ \mathbf{q}_{n-1}^T \\ \mathbf{q}_n^T \end{pmatrix} \mathbf{A}.$$

Wtedy n -ty wiersz tej macierzowej równości ma w zapisie wektorowym postać

$$e_{n-1} \mathbf{q}_{n-1}^T + d_n \mathbf{q}_n^T = \mathbf{q}_n^T \mathbf{A}. \quad (6.18)$$

Z uwagi na ortogonalność \mathbf{Q} , zachodzi $\mathbf{q}_i^T \mathbf{q}_j = \delta_{i,j}$, więc po wymnożeniu ostatniej równości przez \mathbf{q}_n otrzymujemy

$$d_n = \mathbf{q}_n^T \mathbf{A} \mathbf{q}_n.$$

Jeśli znamy wektor \mathbf{q}_n , możemy wyznaczyć d_n . Ponadto z równości (6.18) wynika

$$e_{n-1} \mathbf{q}_{n-1}^T = \mathbf{v}_{n-1}^T, \quad (6.19)$$

gdzie $\mathbf{v}_{n-1}^T = \mathbf{q}_n^T \mathbf{A} - d_n \mathbf{q}_n^T$ jest także znanym wektorem. Podnosząc do kwadratu obie strony równości (6.19), otrzymujemy

$$e_{n-1}^2 = \mathbf{v}_{n-1}^T \mathbf{v}_{n-1}, \quad e_{n-1} = \|\mathbf{v}_{n-1}\|_2,$$

co pozwala wyznaczyć $(n-1)$ -y wiersz macierzy \mathbf{Q}^T :

$$\mathbf{q}_{n-1}^T = \frac{\mathbf{v}_{n-1}^T}{e_{n-1}}.$$

Jak widać, znając ostatni wiersz \mathbf{q}_n^T macierzy \mathbf{Q}^T , wyznaczamy wyrazy e_{n-1} i d_n macierzy \mathbf{T} oraz wiersz \mathbf{q}_{n-1}^T macierzy \mathbf{Q}^T . Można sprawdzić, że jeśli znamy \mathbf{q}_{n-j} i \mathbf{q}_{n-j+1} oraz $(n-j)$ -y wiersz macierzy \mathbf{Q} , to jesteśmy w stanie wyznaczyć \mathbf{q}_{n-j-1} oraz $(n-j)$ -y wiersz \mathbf{T} . \square

Zastosujmy powyższe twierdzenie do macierzy $\mathbf{A} = \mathbf{T}_l$. Niech

$$\bar{\mathbf{T}}_{l+1} = \bar{\mathbf{Q}}_l^T \mathbf{T}_l \bar{\mathbf{Q}}_l,$$

gdzie $\bar{\mathbf{Q}}_l^T$ jest macierzą ortogonalną i ma taki sam ostatni wiersz jak \mathbf{Q}_l^T w algorytmie QL. Wtedy

$$\bar{\mathbf{Q}}_l = \mathbf{Q}_l, \quad \bar{\mathbf{T}}_{l+1} = \mathbf{T}_{l+1}.$$

Przypomnijmy, że w l -tym kroku algorytmu QL (wskaźnik l będziemy dalej pomijać z uwagi na prostotę zapisu)

$$\mathbf{Q}^T = \mathbf{G}_1 \cdots \mathbf{G}_{n-1};$$

ostatni wiersz $\bar{\mathbf{Q}}^T$ jest taki sam, jak ostatni wiersz macierzy Givensa \mathbf{G}_{n-1} , zerującej wyraz $(n-1, n)$ macierzy $\mathbf{T} - z\mathbf{I}$, której wyrazy są równe

$$c_{n-1} = \frac{d_n - z}{\sqrt{e_{n-1}^2 + (d_n - z)^2}}, \quad s_{n-1} = \frac{-e_{n-1}}{\sqrt{e_{n-1}^2 + (d_n - z)^2}}.$$

W ten sposób przesunięcie widma wprowadzamy *niejawnie* poprzez parametry c i s .

Zauważmy, że macierz $\mathbf{G}_{n-1} \mathbf{T} \mathbf{G}_{n-1}^T$ jest postaci

$$\begin{pmatrix} \ddots & & & & \\ & \ddots & & & \\ & & d_{n-3} & e_{n-3} & \\ & & e_{n-3} & d_{n-2} & e_{n-2} & f_{n-2} \\ & & & e_{n-2} & d_{n-1} & e_{n-1} \\ & & & & f_{n-2} & e_{n-1} & d_n \end{pmatrix}, \quad (6.20)$$

w której dodatkowe niezerowe wyrazy f_{n-2} stoją na drugiej pod- i naddiagonali. Musimy ją zredukować do postaci trójdagonalnej za pomocą macierzy \mathbf{S} , której ostatnim wierszem jest $(0, \dots, 0, 1)$. W ten sposób ostatni wiersz $\bar{\mathbf{Q}}^T = \mathbf{S} \mathbf{G}_{n-1}$ będzie równy ostatniemu wierszowi \mathbf{G}_{n-1} .

W celu przywrócenia macierzy (6.20) postaci trójdagonalnej wykonujemy $n-2$ obrotów Givensa $\bar{\mathbf{G}}_j$ w płaszczyznach $(j, j+1)$ dla $j = n-2, \dots, 1$. Obrót $\bar{\mathbf{G}}_j$ zeruje wyraz $(j, j+2)$ i generuje niezerowy wyraz $(j-1, j+1)$; dla $j = 1$ dodatkowy wyraz nie jest generowany. W ten sposób niezerowy wyraz przemieszcza się w górę po drugiej pod- i naddiagonali, aż „wypada poza macierz”, co kończy jedno przekształcenie omawianego tu niejawnego algorytmu QL:

$$\mathbf{Q}^T = \bar{\mathbf{Q}}^T = \mathbf{S} \mathbf{G}_{n-1} = \bar{\mathbf{G}}_1 \cdots \bar{\mathbf{G}}_{n-2} \mathbf{G}_{n-1}.$$

Omówiony algorytm zaimplementowany jest m.in. w podprogramach IMTQL2 z biblioteki EISPACK oraz TQLI z *Numerical Recipes*, jak również w podprogramie SSTEQR biblioteki LAPACK, wywoływanym przez SSTEVE. Dodajmy, że podprogram

SSTEV wybiera wymierny algorytm QR/QL, jeśli nie żądamy obliczania wektorów własnych.

Praktyka pokazuje, że średnia liczba iteracji niejawnego algorytmu QL przypadających na jedną wyznaczaną wartość własną wynosi 1,3–1,6. Liczba operacji zmien-noprzecinkowych w jednej iteracji wynosi $\sim 20n$ [44 (r. 11)], zatem do diagonalizacji macierzy \mathbf{T} potrzeba $\sim 30n^2$ operacji.

6.4. Metoda „dziel i rządź”

Najszybszym algorytmem numerycznym pozwalającym wyznaczać wszystkie wektory i wartości własne symetrycznej macierzy trójdzielnej o rozmiarze większym od 25×25 ²⁾ jest metoda „dziel i rządź”³⁾ (MDiR) [50, 90, 91], którą przedstawimy poniżej.

Zapiszemy macierz trójdzielną jako sumę trzech macierzy rozrzedzonych w następujący sposób:

$$\mathbf{T} = \left(\begin{array}{cccc|cccc} d_1 & e_1 & & & & & & \\ e_1 & \ddots & \ddots & & & & & \\ & \ddots & \ddots & \ddots & & & & \\ & & e_{m-1} & d_m & e_m & & & \\ \hline & & & e_m & d_{m+1} & e_{m+1} & & \\ & & & & e_{m+1} & \ddots & \ddots & \\ & & & & & \ddots & \ddots & e_{n-1} \\ & & & & & & e_{n-1} & d_n \end{array} \right) =$$

$$= \begin{pmatrix} \mathbf{T}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_2 \end{pmatrix} + \left(\begin{array}{cccc|cccc} & & & & |e_m| & e_m & & \\ & & & & e_m & |e_m| & & \\ \hline & & & & & & & \end{array} \right) = \begin{pmatrix} \mathbf{T}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_2 \end{pmatrix} + |e_m| \mathbf{v} \otimes \mathbf{v},$$

gdzie \mathbf{T}_1 i \mathbf{T}_2 są macierzami trójdzielnymi

$$\mathbf{T}_1 = \begin{pmatrix} d_1 & e_1 & & & \\ e_1 & \ddots & \ddots & & \\ & \ddots & \ddots & d_{m-1} & e_{m-1} \\ & & e_{m-1} & d_m - |e_m| & \end{pmatrix}, \quad \mathbf{T}_2 = \begin{pmatrix} d_{m+1} - |e_m| & e_{m+1} & & & \\ e_{m+1} & d_{m+2} & \ddots & & \\ & \ddots & \ddots & e_{n-1} & \\ & & e_{n-1} & d_n & \end{pmatrix},$$

a wektor $\mathbf{v} = (0, \dots, 0, \text{sgn } e_m, 1, 0, \dots, 0)^T$ ma niezerowe współrzędne o wskaźnikach m i $(m+1)$.

²⁾ Rozmiar ten zależy od typu komputera.

³⁾ Ang. *Divide-and-Conquer*.

Założmy, że znane są rozkłady macierzy \mathbf{T}_1 i \mathbf{T}_2 względem wartości własnych

$$\mathbf{T}_1 = \mathbf{Q}_1^T \mathbf{\Lambda}_1 \mathbf{Q}_1, \quad \mathbf{T}_2 = \mathbf{Q}_2^T \mathbf{\Lambda}_2 \mathbf{Q}_2,$$

gdzie $\mathbf{\Lambda}_1$ i $\mathbf{\Lambda}_2$ są macierzami diagonalnymi (innymi słowy, znamy przekształcenie ortogonalne diagonalizujące macierze \mathbf{T}_1 i \mathbf{T}_2). Wtedy

$$\begin{aligned} \mathbf{T} &= \begin{pmatrix} \mathbf{T}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_2 \end{pmatrix} + \varrho \mathbf{v} \otimes \mathbf{v} = \begin{pmatrix} \mathbf{Q}_1^T \mathbf{\Lambda}_1 \mathbf{Q}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2^T \mathbf{\Lambda}_2 \mathbf{Q}_2 \end{pmatrix} + \varrho \mathbf{v} \otimes \mathbf{v} = \\ &= \begin{pmatrix} \mathbf{Q}_1^T & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2^T \end{pmatrix} \left[\begin{pmatrix} \mathbf{\Lambda}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_2 \end{pmatrix} + \varrho \mathbf{u} \otimes \mathbf{u} \right] \begin{pmatrix} \mathbf{Q}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2 \end{pmatrix}, \end{aligned}$$

gdzie $\varrho = |e_m|$ oraz

$$\mathbf{u} = \begin{pmatrix} \mathbf{Q}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2 \end{pmatrix} \mathbf{v} = \begin{pmatrix} \text{ostatnia kolumna } \mathbf{Q}_1 \\ \text{pierwsza kolumna } \mathbf{Q}_2 \end{pmatrix}. \quad (6.21)$$

Zatem wartości własne \mathbf{T} są takie same, jak podobnej do niej macierzy $\mathbf{D} + \varrho \mathbf{u} \otimes \mathbf{u}$, gdzie

$$\mathbf{D} = \begin{pmatrix} \mathbf{\Lambda}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_2 \end{pmatrix} = \text{diag}(r_1, \dots, r_n).$$

Założmy, że wartości własne r_i macierzy \mathbf{D} są uporządkowane: $r_n < \dots < r_1$, oraz że macierz $\mathbf{D} - \lambda \mathbf{1}$ jest nieosobliwa. Jeśli λ jest wartością własną macierzy $\mathbf{D} + \varrho \mathbf{u} \otimes \mathbf{u}$, to wtedy

$$\det(\mathbf{D} + \varrho \mathbf{u} \otimes \mathbf{u} - \lambda \mathbf{1}) = \det(\mathbf{D} - \lambda \mathbf{1}) \det[\mathbf{1} + \varrho(\mathbf{D} - \lambda \mathbf{1})^{-1} \mathbf{u} \otimes \mathbf{u}],$$

a zatem

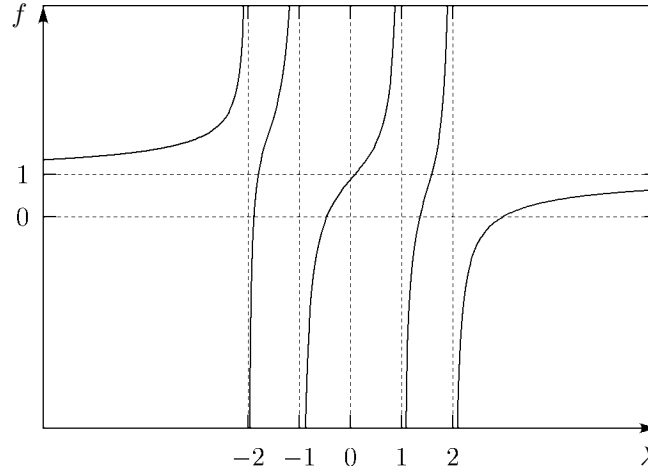
$$\det[\mathbf{1} + \varrho(\mathbf{D} - \lambda \mathbf{1})^{-1} \mathbf{u} \otimes \mathbf{u}] = 0.$$

Na mocy twierdzenia B-24, dla dowolnych dwóch wektorów \mathbf{x}, \mathbf{y} zachodzi równość $\det(\mathbf{1} + \mathbf{x} \otimes \mathbf{y}) = 1 + \mathbf{x} \cdot \mathbf{y}$, więc

$$f(\lambda) := 1 + [\varrho(\mathbf{D} - \lambda \mathbf{1})^{-1} \mathbf{u}] \cdot \mathbf{u} = 1 + \varrho \sum_{i=1}^n \frac{u_i^2}{r_i - \lambda} = 0, \quad (6.22)$$

co oznacza, że wartości własne macierzy \mathbf{T} są pierwiastkami równania $f(\lambda) = 0$. Ponieważ $\varrho > 0$, pochodna $f'(\lambda) = \varrho \sum_{i=1}^n \frac{u_i^2}{(r_i - \lambda)^2} > 0$ i funkcja $f(\lambda)$ jest ściśle rosnąca poza punktami $\lambda = r_i$ (patrz rys. 6.3). W tym miejscu rozważamy ogólny przypadek, gdy wszystkie r_i są różne (wartości własne są niezdegenerowane) oraz $u_i \neq 0$. Wykres $f(\lambda)$ posiada asymptoty pionowe w punktach $\lambda = r_i$, skąd wynika, że kolejne wartości własne macierzy \mathbf{T} leżą w przedziałach (r_i, r_{i+1}) . Z uwagi na monotoniczność $f(\lambda)$ na tych przedziałach, jej pierwiastki można wyznaczać za pomocą szybko zbieżnej metody Newtona lub metody styczna-sieczna (patrz podrozdział 3.4). Ponieważ koszt wyznaczenia $f(\lambda)$ i $f'(\lambda)$ jest rzędu n , to do wyznaczenia wszystkich n wartości własnych \mathbf{T} potrzeba $O(n^2)$ operacji arytmetycznych.

Przy wyznaczaniu wektorów własnych stosujemy następujące twierdzenie [50]:



Rys. 6.3. Przebieg przykładowej funkcji $f(\lambda) = 1 + \frac{0.25}{-2-\lambda} + \frac{0.75}{-1-\lambda} + \frac{0.5}{1-\lambda} + \frac{0.5}{2-\lambda}$

Twierdzenie 6-9. Jeśli λ jest wartością własną macierzy $\mathbf{D} + \varrho \mathbf{u} \otimes \mathbf{u}$, gdzie ϱ — stała, to $(\mathbf{D} - \lambda \mathbf{1})^{-1} \mathbf{u}$ jest wektorem własnym odpowiadającym wartości własnej λ .

Ponieważ macierz $\mathbf{D} - \lambda \mathbf{1}$ jest diagonalna, to koszt wyznaczenia jednego wektora własnego jest rzędu n operacji arytmetycznych.

Podsumujemy nasze dotychczasowe rozważania w postaci rekurencyjnego algorytmu, którego struktura jest przedstawiona na rysunku 6.4.

Algorytm 6-10. (dziel i rządź)

Procedura *Dziel_i_rzadz*($\mathbf{T}, \mathbf{Q}, \mathbf{\Lambda}, n$):

Jeśli $n = 1$:

$\mathbf{Q} := \mathbf{1}; \mathbf{\Lambda} := \mathbf{T}$; Koniec;

$m := \lceil n/2 \rceil$;

$\mathbf{T}_1 := \text{trid}((e_1, \dots, e_{m-1}), (d_1, \dots, d_{m-1}, d_m - \varrho), (e_1, \dots, e_{m-1}))$;

$\mathbf{T}_2 := \text{trid}((e_{m+1}, \dots, e_{n-1}), (d_{m+1} - \varrho, d_{m+2}, \dots, d_n), (e_{m+1}, \dots, e_{n-1}))$;

$\mathbf{v} = (0, \dots, 0, \text{sgn } e_m, 1, 0, \dots, 0)^T$;

Wywołaj *Dziel_i_rzadz*($\mathbf{T}_1, \mathbf{Q}_1, \mathbf{\Lambda}_1, m$);

Wywołaj *Dziel_i_rzadz*($\mathbf{T}_2, \mathbf{Q}_2, \mathbf{\Lambda}_2, n - m$);

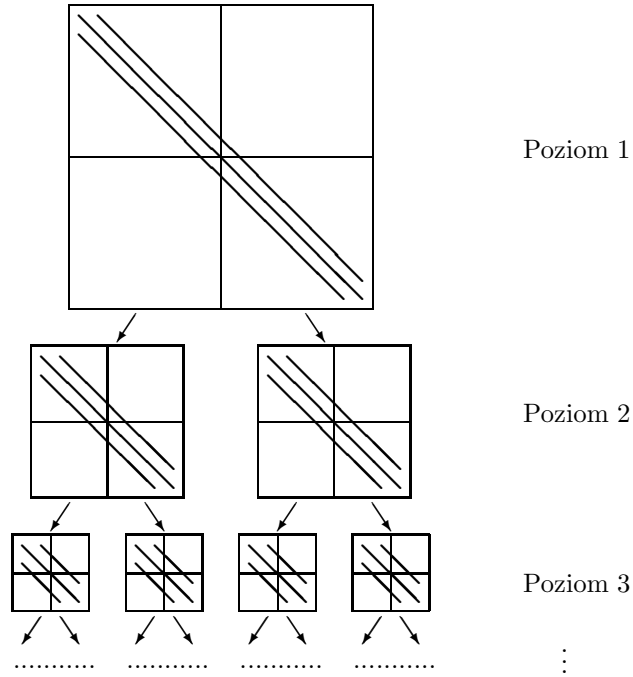
$\mathbf{u} := \begin{pmatrix} \mathbf{Q}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2 \end{pmatrix} \mathbf{v}; \mathbf{D} := \begin{pmatrix} \mathbf{\Lambda}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_2 \end{pmatrix}$;

$\lambda := \text{miejsca_zerowe}(1 + \varrho \sum_{i=1}^n (u_i^2 / (r_i - x)), x)$;

Dla $i = 1, \dots, n$: $\mathbf{q}'_i := (\mathbf{D} - \lambda_i \mathbf{1})^{-1} \mathbf{u}$;

$\mathbf{Q}^T := (\mathbf{q}'_1, \dots, \mathbf{q}'_1)^T \begin{pmatrix} \mathbf{Q}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2 \end{pmatrix}$;

Koniec.



Rys. 6.4. Schemat działania algorytmu „dziel i rządź” na kolejnych poziomach rekurencji

Koszt wyznaczenia wartości własnych jest rzędu n^2 , wektorów własnych zaś jest znacznie większy: $\sim \frac{4}{3}cn^3$ (tak duża liczba operacji wynika z konieczności mnożenia macierzy w ostatnim kroku algorytmu). Tym niemniej, z uwagi na stosowanie deflacji, $c \ll 1$, a ponadto jest to liczba istotnie mniejsza niż $\sim 6n^3$ dla metody QR [92 (r. 30)]. Tak duży zysk wynika z tego, że iteracje w algorytmie „dziel i rządź” wykonywane są na funkcji skalarnej.

Należy zwrócić uwagę, że przechowywanie pośrednich macierzy na kolejnych poziomach rekurencji wymaga użycia dużej liczby komórek pamięci operacyjnej. Przy wyznaczaniu samych wartości własnych jest to $2n^2$, a przy wyznaczaniu wartości i wektorów własnych — aż $3n^2$ komórek [91], co ogranicza rozmiar wejściowego zagadnienia własnego.

Zajmiemy się obecnie metodami rozwiązywania równania wielowego $f(\lambda) = 0$. Jeśli wartość u_i jest mała (ale jeszcze zbyt duża jeszcze do wykonania deflacji), to prosta metoda iteracji Newtona staje się nieprzydatna. Wtedy, zamiast przybliżać $f(\lambda)$ za pomocą funkcji liniowej, w charakterze aproksymanty stosujemy funkcję

$$g(\lambda) = \frac{c_1}{r_j - \lambda} + \frac{c_2}{r_{j+1} - \lambda} + c_3,$$

gdzie c_k ($k = 1, 2, 3$) są nieznanymi stałymi współczynnikami. Ten sposób postępo-

wania, nieco zmodyfikowany, jest wykorzystywany w podprogramie **SLAED4** biblioteki LAPACK. Załóżmy, że znane są nam wartości c_k . Wtedy możemy rozwiązać równanie kwadratowe $g(\lambda) = 0$:

$$c_1(r_{j+1} - \lambda) + c_2(r_j - \lambda) + c_3(r_j - \lambda)(r_{j+1} - \lambda) = 0. \quad (6.23)$$

Niech $\mu^{(l)}$ będzie przybliżoną wartością pierwiastka λ_j w l -tym kroku iteracji Newtona; naturalnym wyborem dla $\mu^{(1)}$ jest środek przedziału $\langle r_j, r_{j+1} \rangle$. W pobliżu tego pierwiastka mamy

$$\begin{aligned} g(\lambda) &= \frac{c_1}{r_j - \lambda} + \frac{c_2}{r_{j+1} - \lambda} + c_3 \simeq f(\lambda) = 1 + \varrho \sum_{i=1}^n \frac{u_i^2}{r_i - \lambda} = \\ &= 1 + \varrho \sum_{i=1}^j \frac{u_i^2}{r_i - \lambda} + \varrho \sum_{i=j+1}^n \frac{u_i^2}{r_i - \lambda} = 1 + f_1(\lambda) + f_2(\lambda). \end{aligned}$$

Zauważmy teraz, że dla $\lambda \in (r_j, r_{j+1})$ funkcja $f_1(\lambda)$ jest sumą składników dodatnich, a funkcja $f_2(\lambda)$ — ujemnych. Obie mogą być zatem policzone dokładnie, w przeciwieństwie do ich sumy.

Wybieramy teraz liczby c_1 oraz \bar{c}_1 , tak aby funkcja

$$g_1(\lambda) = \bar{c}_1 + \frac{c_1}{r_j - \lambda}$$

spełniała relacje

$$g_1(\mu^{(l)}) = f_1(\mu^{(l)}), \quad g_1'(\mu^{(l)}) = f_1'(\mu^{(l)}).$$

Wtedy to hiperbola $g_1(\lambda)$ jest styczna do wykresu funkcji $f_1(\lambda)$ dla $\lambda = \mu^{(l)}$. Można pokazać, że

$$c_1 = f_1'(\mu^{(l)})(r_j - \mu^{(l)})^2, \quad \bar{c}_1 = f_1(\mu^{(l)}) - f_1'(\mu^{(l)})(r_j - \mu^{(l)}).$$

Podobnie wyznaczamy liczby c_2 i \bar{c}_2 dla funkcji

$$g_2(\lambda) = \bar{c}_2 + \frac{c_2}{r_{j+1} - \lambda},$$

która powinna spełniać relacje

$$g_2(\mu^{(l)}) = f_2(\mu^{(l)}), \quad g_2'(\mu^{(l)}) = f_2'(\mu^{(l)}).$$

Ostatecznie szukana funkcja $g(\lambda)$ ma postać

$$g(\lambda) = 1 + g_1(\lambda) + g_2(\lambda) = \frac{c_1}{r_j - \lambda} + \frac{c_2}{r_{j+1} - \lambda} + (\bar{c}_1 + \bar{c}_2 + 1). \quad (6.24)$$

Możemy teraz rozwiązać równanie kwadratowe (6.23) względem λ , przyjmując $\mu^{(l+1)} = \lambda$ za kolejne przybliżenie pierwiastka i powtórzyć całą procedurę.

Przedstawiony wcześniej sposób obliczania wektorów własnych, oparty na twierdzeniu 6-9, nie jest numerycznie stabilny, jeśli poszczególne wartości własne niewiele się różnią. Wtedy obliczone wektory $(\mathbf{D} - \lambda_j \mathbf{1})^{-1} \mathbf{u}$ i $(\mathbf{D} - \lambda_{j+1} \mathbf{1})^{-1} \mathbf{u}$ mogą być bardzo bliskie sobie, obarczone dużymi błędami i nieortogonalne. Można wybrnąć z tej trudności przez zastosowanie formatu zmiennoprzecinkowego o większej liczbie cyfr

znaczących przy obliczaniu różnic $d_j - \lambda_j$ i $d_j - \lambda_{j+1}$. Niestety, w wielu maszynach układy arytmetyki zmiennoprzecinkowej nie obsługują formatu dłuższego niż 64-bitowy (DOUBLE PRECISION w Fortranie). Rozwiązania czysto programowe (tzw. arytmetyka wielokrotnej precyzji) są znacznie mniej efektywne.

Numerycznie stabilny algorytm wyznaczania wektorów własnych dla macierzy typu $\mathbf{D} + \varrho \mathbf{u} \otimes \mathbf{u}$ można sformułować na podstawie twierdzenia Löwnera (B-23), które przytaczamy za książką [50] w Dodatku B:

Algorytm 6-11. (stabilny dziel i rządź)

Procedura *Dziel_i_rzqdź_stab*($\mathbf{T}, \mathbf{Q}, \mathbf{\Lambda}, n$):

Jeśli $n = 1$:

$\mathbf{Q} := \mathbf{1}; \mathbf{\Lambda} := \mathbf{T}$; Koniec;

$m := \lceil n/2 \rceil$;

$\mathbf{T}_1 := \text{trid}((e_1, \dots, e_{m-1}), (d_1, \dots, d_{m-1}, d_m - \varrho), (e_1, \dots, e_{m-1}))$;

$\mathbf{T}_2 := \text{trid}((e_{m+1}, \dots, e_{n-1}), (d_{m+1} - \varrho, d_{m+2}, \dots, d_n), (e_{m+1}, \dots, e_{n-1}))$;

$\mathbf{v} = (0, \dots, 0, \text{sgn } e_m, 1, 0, \dots, 0)^T$;

Wywołaj *Dziel_i_rzqdź*($\mathbf{T}_1, \mathbf{Q}_1, \mathbf{\Lambda}_1, m$);

Wywołaj *Dziel_i_rzqdź*($\mathbf{T}_2, \mathbf{Q}_2, \mathbf{\Lambda}_2, n - m$);

$\mathbf{u} := \begin{pmatrix} \mathbf{Q}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2 \end{pmatrix} \mathbf{v}$; $\mathbf{D} := \begin{pmatrix} \mathbf{\Lambda}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_2 \end{pmatrix}$;

$\lambda := \text{miejsca_zerowe} \left(1 + \varrho \sum_{i=1}^n (u_i^2 / (r_i - x)) \right), x$;

Dla $i = 1, \dots, n$: $w_i := \sqrt{\prod_{j=1}^n (\lambda_j - r_i) / \prod_{j=1, j \neq i}^n (r_j - r_i)}$;

Dla $i = 1, \dots, n$: $\mathbf{q}'_i := (\mathbf{D} - \lambda_i \mathbf{1})^{-1} \mathbf{w}$;

$\mathbf{Q}^T := (\mathbf{q}'_1, \dots, \mathbf{q}'_n)^T \begin{pmatrix} \mathbf{Q}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2 \end{pmatrix}$;

Koniec.

Jak widzimy, w tej metodzie wyznaczamy najpierw wartości własne macierzy $\mathbf{D} + \varrho \mathbf{v} \otimes \mathbf{v}$, a następnie znajdujemy macierz postaci $\mathbf{D} + \mathbf{w} \otimes \mathbf{w}$, która ma takie same wartości własne, i dla niej obliczamy wektory własne (jako $(\mathbf{D} - \lambda_i \mathbf{1})^{-1} \mathbf{w}$).

Można pokazać [50], że różnica między macierzami $\mathbf{D} + \varrho \mathbf{v} \otimes \mathbf{v}$ i $\mathbf{D} + \mathbf{w} \otimes \mathbf{w}$ jest dostatecznie mała, co oznacza, że wektory własne macierzy $\mathbf{D} + \mathbf{w} \otimes \mathbf{w}$ są stabilnymi numerycznie przybliżeniami wektorów własnych $\mathbf{D} + \varrho \mathbf{v} \otimes \mathbf{v}$.

Wyznaczając wektor \mathbf{w} z twierdzenia Löwnera, obliczamy tylko różnice $r_j - r_i$, $\lambda_j - r_i$, iloczyny, ilorazy tych liczb, oraz pierwiastki. Wykonywanie tych operacji z odpowiednią dokładnością zapewnia wysoką stabilność algorytmu.

Istnieje jeszcze szybsza implementacja algorytmu „dziel i rządź”, oparta na szybkiej metodzie multipolowej [50], używanej w zagadnieniach elektrostatyki. Wykorzystuje ona analogię pomiędzy sumą występującą we wzorze (6.22) a wyrażeniem na siłę elektrostatyczną pochodzącą od układu ładunków umieszczonych w przestrzeni dwuwymiarowej. Zastosowanie metody multipolowej zmniejsza liczbę operacji do $\sim n \log n$,

a nawet do $\sim n$. Nie będziemy jej tu omawiać; zaimplementowana jest ona w podprogramie SLAED3 biblioteki LAPACK.

6.5. Metoda Householdera

Omawiając strategię rozwiązywania symetrycznego gęstego zagadnienia własnego stwierdziliśmy, że macierz gęstą \mathbf{A} można (i należy) sprowadzić do macierzy trójdzielnej \mathbf{T} za pomocą przekształceń podobieństwa.

Redukcję macierzy gęstej do postaci trójdzielnej wykonuje się zwykle *metodą Householdera*, zwaną także *metodą odbić Householdera*. Można użyć w tym celu również obrotów Givensa, co jednak w przypadku gęstej macierzy \mathbf{A} jest bardziej czasochłonne.

Zapiszmy gęstą macierz symetryczną $\mathbf{A}_1 := \mathbf{A}$ w postaci blokowej:

$$\mathbf{A}_1 = \left(\begin{array}{c|ccc} * & * & \cdots & * \\ * & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & * \end{array} \right) = \begin{pmatrix} c_1 & \mathbf{b}_1^T \\ \mathbf{b}_1 & \mathbf{D}_1 \end{pmatrix},$$

gdzie gwiazdki oznaczają wyrazy niezerowe; c_1 jest skalar, \mathbf{b}_1 — wektorem o $(n-1)$ współrzędnych, a \mathbf{D}_1 — macierzą rozmiaru $(n-1) \times (n-1)$.

Pierwszym etapem sprowadzania \mathbf{A}_1 do postaci trójdzielnej jest wyzerowanie $n-2$ wyrazów pierwszego wiersza i pierwszej kolumny, tj. doprowadzenie jej za pomocą przekształcenia podobieństwa do postaci

$$\mathbf{A}_2 = \left(\begin{array}{c|ccc} * & \sigma_2 & & \\ \sigma_2 & * & \cdots & * \\ & \vdots & \ddots & \vdots \\ & * & \cdots & * \end{array} \right). \quad (6.25)$$

Zażądamy, aby dokonywało się to w wyniku następujących przekształceń:

$$\mathbf{A}_2 = \bar{\mathbf{Z}}_1^T \mathbf{A}_1 \bar{\mathbf{Z}}_1 = \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{Z}_1^T \end{pmatrix} \begin{pmatrix} c_1 & \mathbf{b}_1^T \\ \mathbf{b}_1 & \mathbf{D}_1 \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{Z}_1 \end{pmatrix} = \begin{pmatrix} c_1 & \mathbf{b}_1^T \mathbf{Z}_1 \\ \mathbf{Z}_1^T \mathbf{b}_1 & \mathbf{Z}_1^T \mathbf{D}_1 \mathbf{Z}_1 \end{pmatrix}, \quad (6.26)$$

gdzie $\mathbf{0}$ oznacza $(n-1)$ -wymiarowy wektor zerowy, \mathbf{Z}_1 — macierz ortogonalną rozmiaru $(n-1) \times (n-1)$ ⁴⁾ postaci

$$\mathbf{Z}_1 = \mathbf{1} - 2\mathbf{w}_1 \otimes \mathbf{w}_1 = \mathbf{1} - 2\mathbf{w}_1 \mathbf{w}_1^T,$$

a $\mathbf{w}_1 \in \mathbb{R}^{n-1}$ — pewien unormowany wektor, którego składowe wyznaczymy dalej.

Z porównania wzorów (6.25) i (6.26) wynika równość wektorowa

$$\mathbf{Z}_1^T \mathbf{b}_1 = \mathbf{b}_1 - 2\mathbf{w}_1 \mathbf{w}_1^T \mathbf{b}_1 = \mathbf{b}_1 - 2\mathbf{w}_1 (\mathbf{w}_1 \cdot \mathbf{b}_1) = (\sigma_2, 0, \dots, 0)^T \quad (6.27)$$

⁴⁾ Macierze $\bar{\mathbf{Z}}_m$ (z kreską) mają rozmiar $n \times n$, a macierze \mathbf{Z}_m — rozmiar $(n-m) \times (n-m)$.

(skorzystaliśmy z własności łączności mnożenia macierzy). Zwróćmy uwagę, że ortogonalne przekształcenie zachowuje długość wektora ($\|\mathbf{Z}_1^T \mathbf{b}_1\|_2 = \|\mathbf{b}_1\|_2$), co pozwala wyznaczyć

$$\sigma_2^2 = \sum_{i=1}^{n-1} b_i^2, \quad \sigma = -\operatorname{sgn} b_1 \sqrt{\sum_{i=1}^{n-1} b_i^2}, \quad (6.28)$$

gdzie, dla uproszczenia zapisu, liczby b_i oznaczają składowe wektora \mathbf{b}_1 . Znak liczby σ wybrano w powyższy sposób dla zmniejszenia błędów zaokrągleń w dalszych obliczeniach.

Oznaczmy $\mathbf{w}_1 \cdot \mathbf{b}_1 = K_1$ i porównajmy współrzędne wektorów w równości (6.27):

$$\begin{array}{rclcl} b_1 & -2w_1 K_1 & = \sigma & b_1 & -\sigma_2 = 2w_1 K_1 \\ b_2 & -2w_2 K_1 & = 0 & b_2 & = 2w_2 K_1 \\ & \vdots & & & \vdots \\ b_{n-1} & -2w_{n-1} K_1 & = 0 & b_{n-1} & = 2w_{n-1} K_1, \end{array} \quad (6.29)$$

gdzie w_i oznacza i -tą współrzędną wektora \mathbf{w}_1 . Po podniesieniu do kwadratu otrzymujemy

$$\begin{array}{rcl} b_1^2 & -2b_1 \sigma_2 + \sigma_2^2 & = 4w_1^2 K_1^2 \\ b_2^2 & & = 4w_2^2 K_1^2 \\ & \vdots & \\ b_{n-1}^2 & & = 4w_{n-1}^2 K_1^2, \end{array}$$

a po dodaniu wszystkich równań stronami,

$$\sum_{i=1}^{n-1} b_i^2 + \sigma_2^2 - 2b_1 \sigma_2 = 4K_1^2 \sum_{i=1}^{n-1} w_i^2 = 4K_1^2 \|\mathbf{w}_1\|_2^2 = 4K_1^2.$$

Korzystając z równości (6.28), otrzymujemy

$$\sigma_2(\sigma_2 - b_1) = 2K_1^2.$$

W tym miejscu widać powód wyboru postaci (6.28 σ : odejmowanie liczb o jednakowych znakach mogłoby prowadzić do utraty cyfr znaczących [71].

Obliczoną wartość $K_1 = \sqrt{\sigma_2(\sigma_2 - b_1)}/2$ podstawiamy do (6.29) i otrzymujemy współrzędne wektora \mathbf{w}_1 :

$$\mathbf{w}_1 = \frac{1}{\sqrt{2\sigma_2(\sigma_2 - b_1)}}(b_1 - \sigma_2, b_2, \dots, b_{n-1})^T.$$

Pozostaje nam jeszcze do obliczenia prawy dolny blok macierzy \mathbf{A}_2 :

$$\begin{aligned} \mathbf{Z}_1^T \mathbf{D}_1 \mathbf{Z}_1 &= (\mathbf{1} - 2\mathbf{w}_1 \otimes \mathbf{w}_1) \mathbf{D}_1 (\mathbf{1} - 2\mathbf{w}_1 \otimes \mathbf{w}_1) = \\ &= \mathbf{D}_1 - 2\mathbf{w}_1 \otimes \mathbf{w}_1 \mathbf{D}_1 - 2\mathbf{D}_1 \mathbf{w}_1 \otimes \mathbf{w}_1 + 4\mathbf{w}_1 \otimes \mathbf{w}_1 \mathbf{D}_1 \mathbf{w}_1 \otimes \mathbf{w}_1 = \\ &= \mathbf{D}_1 - 2\mathbf{w}_1 \mathbf{w}_1^T \mathbf{D}_1 - 2\mathbf{D}_1 \mathbf{w}_1 \mathbf{w}_1^T + 4\mathbf{w}_1 \mathbf{w}_1^T \mathbf{D}_1 \mathbf{w}_1 \mathbf{w}_1^T = \\ &= \mathbf{D}_1 - 2\mathbf{w}_1 \mathbf{u}_1^T - 2\mathbf{u}_1 \mathbf{w}_1^T + 4\mathbf{w}_1 \mathbf{u}_1^T \mathbf{w}_1 \mathbf{w}_1^T = \\ &= \mathbf{D}_1 - \mathbf{w}_1 \mathbf{v}_1^T - \mathbf{v}_1 \mathbf{w}_1^T = \mathbf{D}_1 - \mathbf{w}_1 \otimes \mathbf{v}_1 - \mathbf{v}_1 \otimes \mathbf{w}_1, \end{aligned} \quad (6.30)$$

gdzie wprowadziliśmy pomocnicze wektory

$$\mathbf{u}_1 = \mathbf{D}_1 \mathbf{w}_1, \quad \mathbf{v}_1 = 2[\mathbf{u}_1 - (\mathbf{w}_1 \cdot \mathbf{u}_1) \mathbf{w}_1].$$

Prowadząc obliczenia według tego schematu, nie wykonujemy mnożenia pełnych macierzy (koszt $O(n^3)$ operacji arytmetycznych), lecz ograniczamy się do iloczynów skalarnych i zewnętrznych wektorów (koszt $O(n^2)$ operacji). Oszczędzamy również na wielkości obszaru roboczego pamięci operacyjnej.

W kolejnych krokach algorytmu zerujemy dalsze wiersze i kolumny macierzy \mathbf{A} . W m -tym kroku macierz ma postać

$$\mathbf{A}_m = \left(\begin{array}{cccc|cccc} * & * & & & & & & \\ * & \ddots & \ddots & & & & & \\ & \ddots & \ddots & \ddots & & & & \\ & & \ddots & \ddots & * & & & \\ & & & * & * & * & \cdots & * \\ \hline & & & & * & * & \cdots & * \\ & & & & \vdots & \vdots & \ddots & \vdots \\ & & & & * & * & \cdots & * \end{array} \right) = \begin{pmatrix} \mathbf{C}_m & \mathbf{B}_m^T \\ \mathbf{B}_m & \mathbf{D}_m \end{pmatrix},$$

gdzie blok $\mathbf{C}_m \in \mathbb{R}^{m \times m}$ ma już formę trójdagonalną, a blok $\mathbf{B}_m \in \mathbb{R}^{(n-m) \times m}$ zawiera niezerowe wyrazy tylko w ostatniej kolumnie:

$$\mathbf{B}_m = (\mathbf{0}, \dots, \mathbf{0}, \mathbf{b}_m).$$

W kroku m -tym przekształcenie podobieństwa ma postać

$$\begin{aligned} \mathbf{A}_{m+1} &= \bar{\mathbf{Z}}_m^T \mathbf{A}_m \bar{\mathbf{Z}}_m = \begin{pmatrix} \mathbf{1}_m & \mathbf{0}_m^T \\ \mathbf{0}_m & \mathbf{Z}_m^T \end{pmatrix} \begin{pmatrix} \mathbf{C}_m & \mathbf{B}_m^T \\ \mathbf{B}_m & \mathbf{D}_m \end{pmatrix} \begin{pmatrix} \mathbf{1}_m & \mathbf{0}_m^T \\ \mathbf{0}_m & \mathbf{Z}_m \end{pmatrix} = \\ &= \begin{pmatrix} \mathbf{C}_m & \mathbf{B}_m^T \mathbf{Z}_m \\ \mathbf{Z}_m^T \mathbf{B}_m & \mathbf{Z}_m^T \mathbf{D}_m \mathbf{Z}_m \end{pmatrix}, \end{aligned} \quad (6.31)$$

gdzie $\mathbf{1}_m \in \mathbb{R}^{m \times m}$ jest macierzą jednostkową, $\mathbf{0}_m \in \mathbb{R}^{(n-m) \times m}$ jest macierzą zerową, zaś $\mathbf{Z}_m \in \mathbb{R}^{(n-m) \times (n-m)}$ — macierzą Householdera. W lewym dolnym (i analogicznie w prawym górnym) bloku otrzymujemy

$$\mathbf{Z}_m^T \mathbf{B}_m = \mathbf{Z}_m^T (\mathbf{0}, \dots, \mathbf{0}, \mathbf{b}_m) = (\mathbf{0}, \dots, \mathbf{0}, \mathbf{Z}_m^T \mathbf{b}_m) = \begin{pmatrix} \sigma_{m+1} \end{pmatrix}.$$

Wynik działania $\mathbf{Z}_m^T \mathbf{D}_m \mathbf{Z}_m$ obliczamy analogicznie, jak w pierwszym kroku.

Po wykonaniu $n-2$ kroków otrzymujemy macierz trójdagonalną $\mathbf{T} = \mathbf{A}_{n-1}$, której wartości własne możemy wyznaczyć np. metodą QL.

Do wyznaczenia wektorów własnych wejściowej macierzy \mathbf{A} potrzebna jest znajomość macierzy \mathbf{Q} ortogonalnej transformacji podobieństwa

$$\mathbf{T} = \mathbf{Q}^T \mathbf{A} \mathbf{Q} = \bar{\mathbf{Z}}_{n-2}^T \cdots \bar{\mathbf{Z}}_1^T \mathbf{A} \bar{\mathbf{Z}}_1 \cdots \bar{\mathbf{Z}}_{n-2}.$$

Jej wyrazy znajdujemy, dokonując akumulacji elementarnych przekształceń $\bar{\mathbf{Z}}_m$:

$$\mathbf{Q}_0 = \mathbf{1}, \quad \mathbf{Q}_m = \mathbf{Q}_{m-1} \bar{\mathbf{Z}}_m \quad (m = 1, \dots, n-2), \quad \mathbf{Q} = \mathbf{Q}_{n-2}.$$

Również w tym postępowaniu nie musimy wykonywać kosztownego mnożenia pełnych macierzy. Zauważmy, że

$$\mathbf{Q}_{m-1} \bar{\mathbf{Z}}_m = \mathbf{Q}_{m-1} (\mathbf{1} - 2\bar{\mathbf{w}}\bar{\mathbf{w}}^T) = \mathbf{Q}_{m-1} - \mathbf{q} \otimes \bar{\mathbf{w}},$$

gdzie $\bar{\mathbf{w}}^T = (0, \dots, 0, w_1, \dots, w_{n-m})$ to wektor \mathbf{w}^T uzupełniony zerami do n współrzędnych, a pomocniczy wektor $\mathbf{q} = \mathbf{Q}_{m-1} \mathbf{w}$.

Powyższe rozważania zapiszemy teraz w postaci algorytmu:

Algorytm 6-12. (Householdera)

Procedura *Redukcja_Householdera*($\mathbf{T}, \mathbf{Q}, n$):

$\mathbf{Q} := \mathbf{1}$; inicjalizacja macierzy $\mathbf{Q}_0 = \mathbf{1}$

Dla $m = 1, \dots, n-2$: główna pętla: $n-2$ odbić Householdera

$$\sigma := -\operatorname{sgn} A_{m+1,m} \sqrt{\sum_{i=m+1}^n A_{i,m}^2};$$

$$K := \sqrt{2\sigma(\sigma - b_1)}; w_1 := (A_{m+1,m} - \sigma)/K;$$

Dla $i = 2, \dots, n-m$: $w_i := A_{m+i,m}/K$; współrzędne wektora \mathbf{w}

$$A_{m+1,m} := A_{m,m+1} := \sigma;$$

Dla $i = 2, \dots, n-m$: $A_{m+i,m} := 0$; zerowanie wyrazów w m -tej kolumnie

Dla $i = 1, \dots, n-m$: $u_i := \sum_{j=1}^{n-m} A_{m+i,m+j} w_j$; współrzędne wektora \mathbf{u}

$$t := \sum_{i=1}^{n-m} w_i u_i;$$

Dla $i = 1, \dots, n-m$: $v_i := u_i - t w_i$; współrzędne wektora \mathbf{v}

Dla $i = 1, \dots, n-m$: transformacja $\mathbf{Z}_m^T \mathbf{D}_{m-1} \mathbf{Z}_m$

$$\text{Dla } j = 1, \dots, i: A_{m+i,m+j} := A_{m+i,m+j} - v_j u_i - u_j v_i;$$

Dla $i = 1, \dots, n$: $q_i := \sum_{j=1}^{n-m} Q_{i,m+j} w_j$; współrzędne wektora \mathbf{q}

Dla $i = 1, \dots, n$: akumulacja przekształceń podobieństwa $\mathbf{Q}_m := \mathbf{Q}_{m-1} \bar{\mathbf{Z}}_m$

$$\text{Dla } j = m+1, \dots, n: Q_{i,j} := Q_{i,j} - q_i w_{j-m};$$

Koniec.

Przedstawiony algorytm jest zaimplementowany m.in. w podprogramach TRED1 i TRED2 biblioteki EISPACK oraz TRED2 z *Numerical Recipes*. Jediną różnicą jest to, że podprogramy te wykonują trójdagonalizację zaczynając od ostatniego (a nie od pierwszego) wiersza. Odbicia Householdera wykorzystuje także podprogram SSYTRD, wywoływany przez wspomniane wcześniej podprogram SSYEV biblioteki LAPACK.

6.6. Zagadnienie własne dla macierzy pasmowej

Wiele zagadnień fizycznych opisywanych przez równania różniczkowe prowadzi do algebraicznego zagadnienia własnego z macierzą pasmową (inaczej wstęgową albo war-koczną), tj. mającą niezerowe wyrazy tylko na przekątnej głównej oraz na m poddia-gonalach i m' nadddiagonalach; zazwyczaj są to macierze symetryczne, więc $m = m'$. Z macierzą taką mieliśmy do czynienia analizując pięciopunktowe przybliżenie drugiej

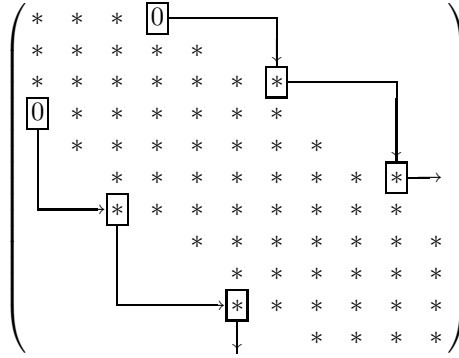
pochoďnej w równaniu Schrödingera, a także w zadaniu 5/11 (przybliżenie siedmipunktowe).

Łatwo zauważyć, że do przechowywania takiej macierzy potrzebujemy załedwie $\sim(m+1)n$ komórek pamięci. Można się także spodziewać, że koszt redukcji takiej macierzy do postaci trójdagonalnej będzie mniejszy niż dla macierzy pełnej. Jest tak rzeczywiście, nie możemy jednak stosować w tym przypadku metody Householdera, gdyż „psuje” ona pasmową strukturę macierzy.

Redukcję macierzy pasmowej do postaci trójdagonalnej wykonujemy za pomocą obrotów Givensa. Najpierw „trójdagonalizujemy” pierwszy wiersz (a zarazem pierwszą kolumnę, ze względu na zachowywanie symetrii macierzy przez ortogonalne przekształcenia podobieństwa), następnie drugi, trzeci itd. Redukcję każdego wiersza rozpoczynamy od zerowania najbardziej zewnętrznego wyrazu; dla ustalenia uwagi niech będzie to $A_{1,m+1}$. W macierzy poddanej transformacji podobieństwa

$$\mathbf{A}' = \mathbf{G}_{m,m+1}^T \mathbf{A} \mathbf{G}_{m,m+1},$$

zerującej $A_{m+1,1}$, pojawia się wyraz niezerowy $A'_{m,2m+1}$ w $(m+1)$ -ej poddiagonali (poza pasmem wejściowej macierzy). Stosując teraz kolejno obroty $\mathbf{G}_{2m,2m+1}$, $\mathbf{G}_{3m,3m+1}$, $\mathbf{G}_{4m,4m+1}, \dots$, „przesuwamy” go w dół, aż „wypadnie poza macierz”, jak na rys. 6.5 (ostatni obrót nie generuje nowego niezerowego wyrazu). Obrotów tych jest $\lfloor \frac{n-1}{m} \rfloor$.



Rys. 6.5. Schemat algorytmu redukcji macierzy pasmowej do postaci trójdagonalnej. Strzałki pokazują działanie kolejnych obrotów Givensa

Analogiczną procedurę wykonujemy dla innych wyrazów w poszczególnych wierszach, skąd liczba elementarnych transformacji podobieństwa potrzebnych do redukcji macierzy wynosi $n_{tr} \leq \frac{m-1}{m} n^2$, co przekłada się na liczbę operacji zmiennoprzecinkowych $n_{op} \leq (4 + \frac{13}{2m}) n^2$ [48]. Widać stąd, że dla $m < \frac{1}{6}n$ metoda ta jest szybsza od metody Householdera.

Nie przedstawiamy tu jawnej postaci algorytmu redukcji macierzy pasmowej do postaci trójdagonalnej. Osoby zainteresowane odsyłamy do podręczników [46, 48] oraz

opisów bibliotek EISPACK [51] i LAPACK [53] (podprogramy odpowiednio BANDR i SSBTRD).

6.7. Metoda Jacobiego

Poddamy macierz symetryczną \mathbf{A} przekształceniu podobieństwa przy użyciu obrotów Jacobiego (6.3):

$$\mathbf{A}' = \mathbf{J}^T \mathbf{A} \mathbf{J}. \quad (6.32)$$

W tym przekształceniu zmieniają się jedynie wyrazy macierzy \mathbf{A} znajdujące się w k -tej i l -tej kolumnie oraz k -tym i l -tym wierszu, a ich wartości wynoszą

$$\begin{aligned} A'_{i,k} &= A'_{k,i} = cA_{i,k} + sA_{i,l} & (i \neq k, l), \\ A'_{i,l} &= A'_{l,i} = cA_{i,l} - sA_{i,k} & (i \neq k, l), \\ A'_{k,k} &= c^2 A_{k,k} + s^2 A_{l,l} + 2csA_{k,l}, \\ A'_{l,l} &= s^2 A_{k,k} + c^2 A_{l,l} - 2csA_{k,l}, \\ A'_{k,l} &= A'_{l,k} = (c^2 - s^2)A_{k,l} + cs(A_{l,l} - A_{k,k}), \end{aligned} \quad (6.33)$$

gdzie dla uproszczenia zapisu wprowadziliśmy oznaczenia $c = \cos \varphi$, $s = \sin \varphi$.

Przekształceń Jacobiego używa się do diagonalizacji macierzy niewielkich rozmiarów. Spośród innych przekształceń ortogonalnych obroty Jacobiego wyróżniają się najszybszym zmniejszaniem sumy kwadratów wyrazów pozadiagonalnych macierzy wejściowej [41].

Transformację Jacobiego wykonuje się w taki sposób, aby wyraz $A'_{k,l}$ macierzy \mathbf{A}' był równy zero, tj.

$$(c^2 - s^2)A_{k,l} + cs(A_{l,l} - A_{k,k}) = 0 \quad \Rightarrow \quad \operatorname{ctg} 2\varphi = \frac{c^2 - s^2}{2cs} = \frac{A_{k,k} - A_{l,l}}{2A_{k,l}}.$$

Transformacja ta odpowiada diagonalizacji macierzy 2×2 :

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} A_{k,k} & A_{k,l} \\ A_{l,k} & A_{l,l} \end{pmatrix} \begin{pmatrix} c & -s \\ s & c \end{pmatrix} = \begin{pmatrix} A'_{k,k} & 0 \\ 0 & A'_{l,l} \end{pmatrix}.$$

Kąt φ oraz wartości jego odpowiednich funkcji trygonometrycznych wyznaczamy za pomocą algorytmu Rutishausera [45]. Niech

$$z = \operatorname{ctg} 2\varphi, \quad t = \operatorname{tg} \varphi = s/c,$$

a poszukiwany kąt $\varphi \in \langle -\frac{\pi}{4}, \frac{\pi}{4} \rangle$. Wartość parametru t wyznaczamy rozwiązując równanie kwadratowe

$$t^2 + 2tz - 1 = 0. \quad (6.34)$$

Rozwiązaniami są

$$t_1 = -(z + \sqrt{z^2 + 1}), \quad t_2 = -(z - \sqrt{z^2 + 1}). \quad (6.35)$$

Wybieramy ten pierwiastek, którego wartość jest mniejsza:

$$t = \frac{\operatorname{sgn} z}{|z| + \sqrt{z^2 + 1}} = \begin{cases} -z + \sqrt{z^2 + 1} = \frac{1}{z + \sqrt{z^2 + 1}} & (z > 0) \\ -z - \sqrt{z^2 + 1} = \frac{1}{z - \sqrt{z^2 + 1}} & (z < 0). \end{cases} \quad (6.36)$$

W przypadku gdy $|z| \gg 1$, przyjmujemy $t = 1/(2z)$.

Można pokazać, że

$$c = \frac{1}{\sqrt{1+t^2}}, \quad s = tc, \quad \tau = \operatorname{tg} \frac{\varphi}{2} = \frac{s}{1+c}. \quad (6.37)$$

W ten sposób wyeliminowaliśmy konieczność czasochłonnego obliczania funkcji trygonometrycznych.

Korzystając z wprowadzonych zmiennych, wzorom (6.33) nadajemy postać numerycznego algorytmu Rutishausera:

$$\begin{aligned} A'_{i,k} &= A_{i,k} + s(A_{i,l} - \tau A_{i,k}) & (i \neq k, l), \\ A'_{i,l} &= A_{i,l} - s(A_{i,k} + \tau A_{i,l}) & (i \neq k, l), \\ A'_{k,k} &= A_{k,k} + tA_{k,l}, \\ A'_{l,l} &= A_{l,l} - tA_{k,l}, \\ A'_{k,l} &= 0. \end{aligned} \quad (6.38)$$

Przytoczmy teraz algorytm odpowiadający pojedynczej transformacji (6.32):

Algorytm 6-13. (Rutishausera)

Procedura *Obrót_Jacobiego*(**A**, **Q**, **A'**, **Q'**, n, k, l , *wektory_wlasne*)

$z := (A_{k,k} - A_{l,l})/(2A_{k,l})$; $t := \operatorname{sgn} z / (|z| + \sqrt{z^2 + 1})$;

$c := 1/\sqrt{1+t^2}$; $s := tc$;

$A'_{k,k} := A_{k,k} + tA_{k,l}$; $A'_{l,l} := A_{l,l} - tA_{k,l}$; $A'_{l,k} := 0$;

Dla $i = 1, \dots, k-1, k+1, \dots, l-1, l+1, \dots, n$:

$A'_{i,k} := A_{i,k} + s(A_{i,l} - \tau A_{i,k})$; $A'_{i,l} := A_{i,l} - s(A_{i,k} + \tau A_{i,l})$;

Jeśli *wektory_wlasne* = TAK:

Dla $i = 1, \dots, n$: $Q'_{i,k} := Q_{i,k} + s(Q_{i,l} - \tau Q_{i,k})$; $Q'_{i,l} := Q_{i,l} - s(Q_{i,k} + \tau Q_{i,l})$.

Koniec.

Koszt pojedynczego obrotu Jacobiego jest rzędu n operacji zmiennoprzecinkowych. Jeżeli wyznaczamy również wektory własne, musimy akumulować kolejne transformacje:

$$\mathbf{A} = \mathbf{J}_m^T \dots \mathbf{J}_1^T \mathbf{A} \mathbf{J}_1 \dots \mathbf{J}_m = \mathbf{Q}^T \mathbf{A} \mathbf{Q}.$$

Pełną diagonalizację metodą Jacobiego wykonujemy, wywołując powyższą procedurę dla wybranych wskaźników (k, l) odpowiednią liczbę razy, aż macierz **A'** będzie różnić się dostecznie mało od macierzy diagonalnej. Miara tej różnicy może być suma kwadratów wyrazów pozadiagonalnych

$$S^2(\mathbf{A}') = \sum_{i < j} (A'_{i,j})^2.$$

Można pokazać, że

$$S^2(\mathbf{A}') = S^2(\mathbf{A}) - A_{k,l}^2 < S^2(\mathbf{A}). \quad (6.39)$$

Implementacje algorytmu Rutishausera mają różną postać [41]:

- W postępowaniu opartym na klasycznym algorytmie Jacobiego kolejne obroty wykonujemy tak, aby wyzerować największy co do modułu wyraz $A_{k,l}$. Po każdym obrocie spełniona jest nierówność $S^2(\mathbf{A}') \leq (1 - \frac{2}{n(n-1)})S^2(\mathbf{A})$, a po m obrotach — nierówność $S^2(\mathbf{A}^{(m)}) \leq (1 - \frac{2}{n(n-1)})^m S^2(\mathbf{A})$. Taka procedura jest najbardziej czasochłonna ze względu na konieczność wyszukiwania przy każdym obrocie największego spośród $\frac{1}{2}n(n-1)$ wyrazów.
- W innej odmianie algorytmu przeglądamy kolejno (kolumnami) pozadiagonalne wyrazy $A_{k,l}$. Obrót w celu wyzerowania wyrazu $A'_{k,l}$ wykonujemy tylko dla dostatecznie dużych wartości $|A_{k,l}|$. W podprogramie JACOBI z *Numerical Recipes* [41] realizowane jest to następująco:
 - Podczas pierwszych trzech procesów przeglądania macierzy \mathbf{A} obroty wykonywane są tylko dla tych wskaźników (k, l) , dla których spełniona jest nierówność

$$|A_{k,l}| > \frac{1}{5n^2} \sum_{j < i} |A_{i,j}|.$$

- Podczas kolejnych procesów przeglądania obroty są pomijane, jeśli zachodzą nierówności $|A_{k,l}| \ll |A_{k,k}|$ i $|A_{k,l}| \ll |A_{l,l}|$. W takich przypadkach od razu przyjmowane jest $A'_{k,l} = 0$.

Praktyka obliczeń numerycznych pokazuje, że metoda Jacobiego jest zazwyczaj 2,5–5 razy wolniejsza od metody QR (włączając czas potrzebny na redukcję macierzy do postaci trójdzielnej). Zaletą metody Jacobiego jest duża dokładność wyznaczania małych wartości własnych i odpowiadających im wektorów własnych.

6.8. Metoda potęgowa

Metodą znacznie prostszą od opisanych powyżej, ale o stosunkowo małym zakresie stosowności, jest metoda potęgowa, pozwalająca wyznaczać największą co do modułu wartość własną λ_{\max} macierzy i odpowiadający jej wektor własny. Odpowiedni algorytm ma następującą postać [50]:

Algorytm 6-14. (potęgowy)

Procedura *Iteracja_potęgowa*($\mathbf{A}, n, \sigma, \mathbf{x}, \lambda_{\max}$)

$\mathbf{x} := \text{dowolny_wektor}(n);$

Powtarzaj

$\mathbf{x}' := \mathbf{x}; \lambda'_{\max} := \lambda_{\max};$ poprzednie wartości

$\mathbf{x} := \mathbf{A}\mathbf{x}; \mathbf{x} := \mathbf{x}/\|\mathbf{x}\|_2; \lambda_{\max} := \mathbf{x}^T \mathbf{A}\mathbf{x};$

aż kryterium_zbieżności(...) $\leq \sigma$;

Koniec.

Jako wartość funkcji *kryterium_zbieżności(...)* przyjmujemy różnicę $|\lambda_{\max} - \lambda'_{\max}|$ lub $\|\mathbf{x} - \mathbf{x}'\|_2$.

Pokażemy najpierw zbieżność metody potęgowej na przykładzie macierzy diagonalnej $\mathbf{A} = \mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_n)$, gdzie $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$. Zauważmy, że $\mathbf{x}^{(i)} = \mathbf{\Lambda}^i \mathbf{x}^{(0)} / \|\mathbf{\Lambda}^i \mathbf{x}^{(0)}\|_2$, gdzie $\mathbf{x}^{(i)}$ oznacza wartość \mathbf{x} w i -tym obiegu pętli, oraz

$$\mathbf{\Lambda}^i \mathbf{x}^{(0)} = \mathbf{\Lambda}^i (s_1, \dots, s_n)^T = (\lambda_1^i s_1, \dots, \lambda_n^i s_n)^T = s_1 \lambda_1^i \left(1, \left(\frac{\lambda_2}{\lambda_1}\right)^i \left(\frac{s_2}{s_1}\right), \dots, \left(\frac{\lambda_n}{\lambda_1}\right)^i \left(\frac{s_n}{s_1}\right)\right)^T,$$

gdzie założono, iż $s_1 \neq 0$. Ponieważ $|\lambda_j/\lambda_1| < 1$ dla $j \neq 1$, to wektor $\mathbf{x}^{(i)}$ dla rosnących i staje się coraz bliższy wektorowi jednostkowemu $(\pm 1, 0, \dots, 0)^T$, tj. wektorowi własnemu odpowiadającemu największej wartości własnej λ_1 . Zatem

$$\lim_{i \rightarrow \infty} (\mathbf{x}^{(i)})^T \mathbf{\Lambda} \mathbf{x}^{(i)} = (\pm 1, 0, \dots, 0) \lambda_1 (\pm 1, 0, \dots, 0)^T = \lambda_1.$$

W przypadku ogólnym zakładamy, że macierz jest diagonalizowalna: $\mathbf{A} = \mathbf{Q}^T \mathbf{\Lambda} \mathbf{Q}$. Zapiszmy $\mathbf{Q}^T = (\mathbf{q}_1, \dots, \mathbf{q}_n)$, gdzie kolumny \mathbf{q}_j są unormowanymi wektorami własnymi przynależnymi do wartości własnych λ_j , oraz $\mathbf{x}^{(0)} = \mathbf{Q}^T \mathbf{Q} \mathbf{x}^{(0)} = \mathbf{Q}^T (s_1, \dots, s_n)^T$. Ponieważ $\mathbf{Q}^T \mathbf{Q} = \mathbf{1}$, i -ta potęga macierzy \mathbf{A} wynosi

$$\mathbf{A}^i = \underbrace{(\mathbf{Q}^T \mathbf{\Lambda} \mathbf{Q}) \dots (\mathbf{Q}^T \mathbf{\Lambda} \mathbf{Q})}_{i \text{ razy}} = \mathbf{Q}^T \mathbf{\Lambda}^i \mathbf{Q},$$

co pozwala napisać ciąg równości

$$\begin{aligned} \mathbf{A}^i \mathbf{x}^{(0)} &= (\mathbf{Q}^T \mathbf{\Lambda}^i \mathbf{Q}) \mathbf{Q}^T (s_1, \dots, s_n)^T = \mathbf{Q}^T (\lambda_1^i s_1, \dots, \lambda_n^i s_n)^T = \\ &= s_1 \lambda_1^i \mathbf{Q}^T \left(1, \left(\frac{\lambda_2}{\lambda_1}\right)^i \left(\frac{s_2}{s_1}\right), \dots, \left(\frac{\lambda_n}{\lambda_1}\right)^i \left(\frac{s_n}{s_1}\right)\right)^T. \end{aligned}$$

Jak widzimy, ostatni wektor ponownie dąży do $(1, 0, \dots, 0)^T$. Mamy

$$\lim_{i \rightarrow \infty} \mathbf{x}^{(i)} = \lim_{i \rightarrow \infty} \frac{\mathbf{A}^i \mathbf{x}^{(0)}}{\|\mathbf{A}^i \mathbf{x}^{(0)}\|_2} = \mathbf{Q}^T (1, 0, \dots, 0)^T = \mathbf{q}_1.$$

Ostatecznie

$$\lim_{i \rightarrow \infty} (\mathbf{x}^{(i)})^T \mathbf{A} \mathbf{x}^{(i)} = \mathbf{q}_1^T \mathbf{Q}^T \mathbf{\Lambda} \mathbf{Q} \mathbf{q}_1 = (\pm 1, 0, \dots, 0)^T \mathbf{\Lambda} (\pm 1, 0, \dots, 0)^T = \lambda_1.$$

Wadą tej metody jest to, że nie działa ona przy takim wyborze $\mathbf{x}^{(0)}$, przy którym $s_1 = 0$. Jej zbieżność zależy od wartości ilorazów λ_j/λ_1 . Uniemożliwia to m.in. stosowanie tego podejścia do macierzy ortogonalnych, dla których $|\lambda_j| = 1$.

6.9. Metoda iteracji odwrotnych

W celu wyznaczenia wartości własnych macierzy \mathbf{A} innych niż wartość maksymalna, wykonujemy przesunięcie widma o wartość z i stosujemy metodę potęgową do macierzy $(\mathbf{A} - z\mathbf{1})^{-1}$:

Algorytm 6-15. (iteracji odwrotnych)

Procedura *Iteracja_odwrotna*($\mathbf{A}, n, z, \sigma, \mathbf{x}, \lambda$)

$\mathbf{x} := \text{dowolny_wektor}(n)$;

Powtarzaj

$\mathbf{x}' := \mathbf{x}$; $\lambda' := \lambda$; poprzednie wartości

$\mathbf{x} := (\mathbf{A} - z\mathbf{1})^{-1}\mathbf{x}$; $\mathbf{x} := \mathbf{x}/\|\mathbf{x}\|_2$; $\lambda := \mathbf{x}^T \mathbf{A} \mathbf{x}$;

aż *kryterium_zbieżności*(...) $\leq \sigma$;

Koniec.

Zbieżność metody iteracji odwrotnych pokazuje się w sposób analogiczny, jak dla metody potęgowej.

Obliczanie macierzy odwrotnej na początku pętli jest procesem czasochłonnym i wprowadzającym stosunkowo duże błędy zaokrągleń [24]. Zamiast obliczać macierz odwrotną, można wyznaczyć \mathbf{x} rozwiązując układ równań

$$(\mathbf{A} - z\mathbf{1})\mathbf{x} = \mathbf{x}',$$

co jest szybsze i dokładniejsze. Nie będziemy tu przedstawiać algorytmów rozwiązywania ogólnego układu równań liniowych, opisanych szczegółowo w podręcznikach [21, 24, 25, 53]. Opiszemy tylko metodę rozwiązywania trójdziagonalnego układu równań

$$(\mathbf{T} - z\mathbf{1})\mathbf{x} = \mathbf{b}, \tag{6.40}$$

z jakim mamy do czynienia, jeśli uprzednio sprowadziliśmy gęstą macierz \mathbf{A} do postaci trójdziagonalnej za pomocą ortogonalnych przekształceń podobieństwa (6.2). Po obliczeniu metodą iteracji odwrotnych wektora własnego \mathbf{x} macierzy \mathbf{T} , znajdujemy odpowiadający mu wektor własny \mathbf{w} macierzy \mathbf{A} z relacji $\mathbf{w} = \mathbf{Q}\mathbf{x}$.

Metoda rozwiązywania trójdziagonalnego układu równań (6.40) opiera się na rozkładzie macierzy $\mathbf{T} - z\mathbf{1}$ na iloczyn $\mathbf{LD}'\mathbf{L}^T$, gdzie

$$\mathbf{L} = \begin{pmatrix} 1 & & & \\ l_1 & \ddots & & \\ & \ddots & \ddots & \\ & & l_{n-1} & 1 \end{pmatrix}, \quad \mathbf{D}' = \begin{pmatrix} d'_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & d'_n \end{pmatrix}.$$

Łatwo pokazać, że

$$d'_1 = d_1 - z, \quad d'_i = d_i - z - \frac{e_{i-1}^2}{d'_{i-1}} \quad (i = 2, \dots, n),$$

a wyrazy pozadiagonalne l_i macierzy \mathbf{L} są równe

$$l_i = \frac{e_i}{d'_i} \quad (i = 1, \dots, n-1).$$

Układ równań (6.40) jest równoważny układowi

$$\mathbf{LD}'\mathbf{L}^T \mathbf{x} = \mathbf{b}, \quad (6.41)$$

który rozwiązujemy w trzech etapach:

- Rozwiązujemy $\mathbf{L}\mathbf{w} = \mathbf{b}$ względem \mathbf{w} .
- Rozwiązujemy $\mathbf{D}'\mathbf{y} = \mathbf{w}$ względem \mathbf{y} . Ponieważ \mathbf{D}' jest macierzą diagonalną, mamy $y_i = w_i/u_i$ i krok ten wygodnie jest wykonać razem z następnym.
- Rozwiązujemy $\mathbf{L}^T \mathbf{x} = \mathbf{y}$ względem \mathbf{x} .

Schemat ten można zapisać w postaci następującego algorytmu:

Algorytm 6-16. (rozwiązywania trójdagonalnego układu równań)

Procedura *Rozwiąż-trójdialg*($\mathbf{d}, \mathbf{e}, r, \mathbf{b}, \mathbf{x}$)

$d'_1 := d_1 - z$;

Dla $i = 2, \dots, n$: rozkład $\mathbf{T} - z\mathbf{1} = \mathbf{LD}'\mathbf{L}^T$

$d'_i := d_i - z - e_i^2/d'_{i-1}$; $l_{i-1} := e_{i-1}/d'_{i-1}$;

$w_1 := b_1$;

Dla $i = 2, \dots, n$: $w_i := b_i - l_{i-1}w_{i-1}$; $\mathbf{L}\mathbf{w} = \mathbf{b}$ — podstawianie w przód

$x_n := w_n/d'_n$;

Dla $i = n-1, \dots, 1$: $x_i := w_i/d'_i - l_i x_{i+1}$; $\mathbf{D}'\mathbf{L}^T \mathbf{x} = \mathbf{w}$ — podstawianie wstecz

Koniec.

Algorytm ten pozwala na oszczędne wykorzystywanie pamięci operacyjnej, ponieważ obliczane wartości mogą być zapisywane na miejscu starych, jeśli te nie są już potrzebne: d'_i na miejscu d_i , l_i na miejscu e_i , a w_i oraz x_i na miejscu b_i .

W serii prac [93] wyprowadzono stosunkowo proste analityczne wyrażenia dla wyrazów macierzy odwrotnej do macierzy trójdagonalnej. Niech $\mathbf{T}' = \text{trid}(\mathbf{e}, \mathbf{d}, \mathbf{f})$ będzie dowolną (niesymetryczną) macierzą trójdagonalną $n \times n$. Wówczas $\mathbf{B} = (\mathbf{T}')^{-1}$ jest macierzą gęstą o wyrazach

$$B_{j,j} = \left(d_j - e_{j-1}f_{j-1} \frac{z_{j-2}}{z_{j-1}} - e_j f_j \frac{y_{j+2}}{y_{j+1}} \right)^{-1}$$

$$B_{i,j} = \begin{cases} -f_i \frac{z_{i-1}}{z_i} B_{i+1,j} = (-1)^{j-i} \prod_{k=1}^{j-i} f_{j-k} \frac{z_{i-1}}{z_{j-1}} B_{j,j} & (i < j) \\ -e_{i-1} \frac{y_{i+1}}{y_i} B_{i-1,j} = (-1)^{i-j} \prod_{k=1}^{i-j} e_{j+k-1} \frac{y_{i+1}}{y_{j+1}} B_{j,j} & (i > j), \end{cases} \quad (6.42)$$

gdzie (z_i) i (y_i) są ciągami zdefiniowanymi następująco:

$$z_0 = 1, \quad z_1 = d_1, \quad z_i = d_i z_{i-1} - e_{i-1} f_{i-1} z_{i-2} \quad (i = 2, \dots, n),$$

$$y_{n+1} = 1, \quad y_n = d_n, \quad y_i = d_i y_{i+1} - e_i f_i y_{i+2} \quad (i = n-1, \dots, 1),$$

a $e_0 = f_n = 0$.

Obliczenie macierzy odwrotnej według powyższych formuł wymaga $n^2 + 7n - 7$ operacji arytmetycznych. Algorytm nie załamuje się (nie występuje dzielenie przez zero) dla macierzy diagonalnie dominujących, tj. gdy $|d_i| > |e_{i-1}| + |f_i|$. Wzory (6.42)

mogą być wykorzystane do stworzenia algorytmu numerycznego rozwiązującego trójdziagonalny układ równań (patrz projekt A.9).

Odmianą metody iteracji odwrotnych stanowi metoda ilorazów Reyleigha [50], która jest zbieżna kubicznie, co oznacza, że liczba miejsc znaczących zwiększa się trzykrotnie w każdym kroku; dowód jej zbieżności podany jest w podręczniku [50].

Algorytm 6-17. (ilorazów Reyleigha)

Procedura *Iteracja_Reyleigha*($\mathbf{A}, n, \varrho, \sigma, \mathbf{x}, \lambda$)

$\mathbf{x} := \text{dowolny_wektor}(n); \mathbf{x} := \mathbf{x}/\|\mathbf{x}\|_2;$

Powtarzaj

$\mathbf{y} := (\mathbf{A} - \varrho \mathbf{1})^{-1} \mathbf{x};$ w rzeczywistości rozwiązujemy układ równań

$\mathbf{x} := \mathbf{y}/\|\mathbf{y}\|_2; \varrho := \mathbf{x}^T \mathbf{A} \mathbf{x};$

aż $\|\mathbf{A} \mathbf{x} - \varrho \mathbf{x}\|_2 \leq \sigma;$

$\lambda := \varrho;$

Koniec.

6.10. Metoda bisekcji i iteracji odwrotnych

W tej metodzie do obliczania zadanej liczby wartości własnych macierzy \mathbf{A} z określonego przedziału wykorzystujemy algorytm bisekcji oparty na twierdzeniu Sylwestera (B-14). Opisaliśmy go dokładnie w rozdziale 3; jeżeli \mathbf{A} jest macierzą gęstą, musimy ją najpierw sprowadzić do postaci trójdziagonalnej. Wektory własne są wyznaczane za pomocą algorytmu iteracji odwrotnych opisanym w poprzednim podrozdziale. Tym razem za przesunięcia z_j przyjmujemy kolejne wartości własne λ_j .

Wyznaczenie k wartości własnych metodą bisekcji wymaga tylko $\sim nk$ operacji zmiennoprzecinkowych, a zatem dla $k \ll n$ metoda ta jest znacznie szybsza od innych. Obliczenie k wektorów własnych metodą iteracji odwrotnych wymaga również $\sim nk$ operacji. Jednakże gdy poszczególne wartości własne leżą zbyt blisko siebie, wyznaczone wektory własne nie są ortogonalne i wymagają reortogonalizacji zmodyfikowaną metodą Grama–Schmidta [50] kosztem $\sim nk^2$ operacji.

Jeśli kolumny macierzy $\mathbf{Q} = (\mathbf{q}_1, \dots, \mathbf{q}_m)$ są obliczonymi wektorami własnymi wymagającymi reortogonalizacji, wówczas odpowiedni algorytm przedstawia się następująco:

Algorytm 6-18. (Gram–Schmidta)

Procedura *Ortogonalizacja_GS*(\mathbf{Q}, n, m)

Dla $i = 1, \dots, m$:

Dla $j = 1, \dots, i - 1$: $\mathbf{q}_i := \mathbf{q}_i - (\mathbf{q}_j \cdot \mathbf{q}_i) \mathbf{q}_j$;

$r := \|\mathbf{q}_i\|_2$;

Jeśli $r = 0$: Przerwij; wektory liniowo zależne

$\mathbf{q}_i := \mathbf{q}_i/r$;

Koniec;

Innym sposobem reortogonalizacji jest wykonanie rozkładu macierzy \mathbf{Q} względem wartości szczególnych (*Singular Value Decomposition, SVD*). Rozkład ten można wykonać np. za pomocą podprogramów *SGESVD* z LAPACK-a lub *SVDCMP* z *Numerical Recipes*.

Metoda bisekcji i iteracji odwrotnych jest zaimplementowana jako jeden z wariantów w podprogramie *SSYEVX* biblioteki LAPACK. Jej kolejne etapy są wykonywane przez podprogramy *SSYTRD* (redukcja trójdagonalna), *SSTEBZ* (bisekcja) i *SSTEIN* (iteracje odwrotne i, w razie potrzeby, reortogonalizacja Grama–Schmidta).

Czasochłonnej reortogonalizacji można uniknąć, stosując podejście zaproponowane przez Dhillona i Parletta [75], opierające się na reprezentacji macierzy $\mathbf{T} - \mu\mathbf{1}$ za pomocą iloczynu \mathbf{LDL}^T , gdzie \mathbf{L} jest macierzą bidiagonalną, a μ — odpowiednio dobranym przesunięciem widma. Otrzymane w tej metodzie wektory własne mają zagwarantowaną ortogonalność.

Podejście to prowadzi do pierwszego stabilnego algorytmu pozwalającego wyznaczać wszystkie wartości i wektory własne trójdagonalnej macierzy symetrycznej kosztem $O(n^2)$ operacji zmiennoprzecinkowych. Dla wartości własnych o dużych względnych odległościach, tzn.

$$\left| \frac{\varepsilon_j - \varepsilon_{j\pm 1}}{\max_i(\varepsilon_i, \varepsilon_{i\pm 1})} \right| = O(1) \quad (i = 1, \dots, n),$$

istnieje matematyczny dowód poprawności algorytmu. Jeśli powyższy warunek nie jest spełniony (wartości własne grupują się), to — jak się wydaje — podejście to daje również poprawne wyniki, lecz brak jest na to matematycznego uzasadnienia.

6.11. Porównanie wydajności przedstawionych metod

Przedstawimy porównanie czasów [50, 53, 91] potrzebnych na rozwiązanie pełnego symetrycznego zagadnienia własnego (wyznaczenie wszystkich wektorów i wartości własnych) za pomocą następujących podprogramów biblioteki LAPACK:

1. Transformacji Householdera i iteracji QR (*SSYTRD* + *SSTEQR*),
2. Transformacji Householdera i bisekcji z odwrotnymi iteracjami (*SSYTRD* + *SSTEBZ* + *SSTEIN*),
3. Transformacji Householdera i metody „dziel i rządź” (*SSYEVD*).

W tabeli podane są stosunki czasów t_{AZW} rozwiązywania przy użyciu podprogramu *SSYEVD* zagadnienia własnego rozmiaru $n \times n$ do czasów $t_{\text{mnoż}}$ mnożenia macierzy $n \times n$ dla stacji roboczej IBM RS600/590 o szczytowej wydajności $\sim 250 \cdot 10^6$ operacji zmiennoprzecinkowych na sekundę:

Rozmiar macierzy	Tylko wartości własne $t_{\text{AZW}}/t_{\text{mnoż}}$	Wartości i wektory własne $t_{\text{AZW}}/t_{\text{mnoż}}$
100	3,1	9,3
1000	1,1	2,8

Widać, że czas wyznaczania wartości własnych dla dostatecznie dużych macierzy jest porównywalny z czasem mnożenia macierzy, natomiast czas wyznaczania wartości i wektorów własnych jest około trzykrotnie dłuższy.

Porównanie metod 1–3 przeprowadzono dla dwóch szczególnych typów macierzy:

- macierzy losowych, których wartości własne leżą (przeciętnie) daleko od siebie,
- macierzy, których wartości własne tworzą ciąg geometryczny $\lambda_i = (\frac{1}{2})^{i-1}$.

Testy przeprowadzono dla macierzy trójdzielnych (bez pierwszego etapu, tj. transformacji Householdera) oraz dla macierzy pełnych.

Metoda „dziel i rządź” (MDiR) okazała się zdecydowanie najszybsza. W przypadku losowych macierzy trójdzielnych metoda bisekcji i iteracji odwrotnych była porównywalna z MDiR (~ 3 razy wolniejsza dla $n = 100$, $\sim 1,5$ razy wolniejsza dla $n = 1000$), natomiast metoda QR okazała się być najwolniejsza (~ 3 razy wolniejsza od MDiR dla $n = 100$ i ~ 15 razy wolniejsza dla $n = 1000$). W przypadku macierzy trójdzielnych o widmie tworzącym ciąg geometryczny, metoda QR była wolniejsza od MDiR ~ 5 razy (dla $n = 100$) i ~ 55 razy (dla $n = 1000$), a metoda bisekcji i iteracji odwrotnych ~ 8 razy (dla $n = 100$) i ~ 73 razy (dla $n = 1000$). Jak widać, przewaga metody „dziel i rządź” nad pozostałymi rośnie dla rosnącego rozmiaru macierzy.

Powyższe proporcje ulegają znacznemu spłaszczeniu przy rozwiązywaniu gęstego zagadnienia własnego, ponieważ każda z użytych metod zawiera czasochłonny etap redukcji macierzy symetrycznej do postaci trójdzielnej. Dla macierzy losowych 100×100 czas rozwiązywania gęstego AZW metodami 1 i 2 jest $\sim 1,8$ razy dłuższy niż metodą 3, natomiast dla macierzy 1000×1000 czasy dla metod 2 i 3 są porównywalne⁵⁾, a metoda 1 jest $\sim 2,6$ razy wolniejsza. Dla macierzy o widmie tworzącym ciąg geometryczny metoda 1 jest $\sim 1,7$ razy wolniejsza, a metoda 2 $\sim 2,3$ razy wolniejsza od metody 3. Proporcje te nie zmieniają się istotnie dla $n = 100$ – 1000 .

Przypomnijmy tu jednak, że algorytm „dziel i rządź” wymaga dużego obszaru roboczego pamięci operacyjnej w związku z jego rekurencyjną strukturą, co stanowi jego wadę.

Z uwagi na to, że dla $n \lesssim 25$ metoda QR jest najszybsza, w praktycznych implementacjach algorytmu „dziel i rządź” dokonuje się rekurencyjnego podziału macierzy trójdzielnej na bloki o rozmiarach nie przekraczających 25×25 , do których następnie stosuje się algorytm QR/QL.

Więcej szczegółowych informacji dotyczących wydajności poszczególnych metod można znaleźć w podręczniku [53] i pracach [75, 91].

⁵⁾ Metoda bisekcji i iteracji odwrotnych jest w tym przypadku bardzo szybka, ponieważ wartości własne leżą daleko od siebie i nie jest potrzebna czasochłonna reortogonalizacja wektorów własnych.

6.12. Zagadnienie własne dla macierzy hermitowskiej

Algebraiczne zagadnienie własne

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \quad (6.43)$$

dla zespolonej macierzy hermitowskiej $\mathbf{A}^H = (\mathbf{A}^*)^T = \mathbf{A} \in \mathbb{C}^{n \times n}$, gdzie $\lambda \in \mathbb{R}$ (wartości własne macierzy hermitowskiej są rzeczywiste), można rozwiązywać za pomocą odpowiednich wersji algorytmów QL i QR, działających na liczbach zespolonych.

Innym rozwiązaniem jest sprowadzenie problemu (6.43) do zagadnienia własnego z symetryczną macierzą rzeczywistą. Niech

$$\mathbf{A} = \mathbf{B} + i\mathbf{C}, \quad \mathbf{B}, \mathbf{C} \in \mathbb{R}^{n \times n}, \quad (6.44)$$

oraz

$$\mathbf{x} = \mathbf{u} + i\mathbf{v}, \quad \mathbf{u}, \mathbf{v} \in \mathbb{R}^n;$$

zauważmy, że

$$\mathbf{B}^T = \mathbf{B}, \quad \mathbf{C}^T = -\mathbf{C}. \quad (6.45)$$

Rozdzielając hermitowskie zagadnienie własne na część rzeczywistą i urojoną

$$(\mathbf{B} + i\mathbf{C})(\mathbf{u} + i\mathbf{v}) = (\mathbf{B}\mathbf{u} - \mathbf{C}\mathbf{v}) + i(\mathbf{C}\mathbf{u} + \mathbf{B}\mathbf{v}) = \lambda\mathbf{u} + i\lambda\mathbf{v},$$

możemy je zapisać w postaci rzeczywistego zagadnienia własnego dla macierzy rozmiaru $2n \times 2n$:

$$\begin{pmatrix} \mathbf{B} & -\mathbf{C} \\ \mathbf{C} & \mathbf{B} \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} = \lambda \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}. \quad (6.46)$$

Macierz występująca w równaniu (6.46) jest symetryczna ze względu na związki (6.45).

Unormowany wektor własny macierzy jest określony z dokładnością do czynnika fazowego (tj. liczby zespolonej o module 1). Zatem wektor $e^{-i\pi/2}\mathbf{x} = -i\mathbf{x} = \mathbf{v} - i\mathbf{u}$ jest także rozwiązaniem zagadnienia własnego (6.43) przynależnym do wartości własnej λ :

$$\begin{pmatrix} \mathbf{B} & -\mathbf{C} \\ \mathbf{C} & \mathbf{B} \end{pmatrix} \begin{pmatrix} \mathbf{v} \\ -\mathbf{u} \end{pmatrix} = \lambda \begin{pmatrix} \mathbf{v} \\ -\mathbf{u} \end{pmatrix}.$$

Oznacza to, że każdej wartości własnej λ macierzy rzeczywistej $2n \times 2n$ z równania (6.46) odpowiadają dwa *ortogonalne* wektory własne

$$\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}, \quad \begin{pmatrix} \mathbf{v} \\ -\mathbf{u} \end{pmatrix}.$$

Innymi słowy, wartości własne macierzy z równania (6.46) są dwukrotnie zdegenerowane.

Podkreślmy, że rozwiązując numerycznie zagadnienie (6.46) otrzymujemy $2n$ wartości własnych, spośród których tylko n jest różnych, oraz $2n$ wektorów własnych (n par). Wektory z każdej pary opisują ten sam stan fizyczny i rozwiązanie wejściowego zespolonego zagadnienia własnego możemy zbudować z dowolnego z nich.

Zadania

- 6/1.** Sprawdzić własność ortogonalności macierzy obrotu Jacobiego, tj. relację (6.4).
- 6/2.** Pokazać, że wzory (6.6) określają funkcje trygonometryczne kąta obrotu Givensa.
- 6/3.** Wyprowadzić formuły (6.13)–(6.15) określające wyrazy macierzy \mathbf{R} w procesie dekompozycji.
- 6/4.** Udowodnić, że w procesie rekombinacji spełniona jest relacja (6.16).
- 6/5.** Pokazać, że przekształcenia podobieństwa wykonane za pomocą obrotów Givensa zachowują własność trójdzielności macierzy symetrycznej.
- 6/6.** Zbadać średnią liczbę iteracji jawnego, niejawnego i wymiennego algorytmu QL, potrzebnych do wyznaczenia jednej wartości własnej symetrycznej macierzy trójdzielnej. Przeprowadzić odpowiednie symulacje na PC.
- 6/7.** Uzasadnić postać (6.21) wektora \mathbf{u} w metodzie „dziel i rządź”.
- 6/8.** Wyprowadzić jawne postaci współczynników c_1 , c_2 , \bar{c}_1 i \bar{c}_2 występujących we wzorze (6.24).
- 6/9.** Sprawdzić słuszność wzoru (6.30). Przeprowadzić analogiczne rozważania dla prawego dolnego bloku macierzy \mathbf{A}_3 .
- 6/10.** Wyprowadzić wzory (6.33) dla wyrazów macierzy poddanej przekształceniu Jacobiego.
- 6/11.** Uzasadnić sformułowanie algorytmu Rutishausera (6.34)–(6.38).
- 6/12.** Wykazać prawdziwość nierówności (6.39) opisującej zmniejszanie się wyrazów pozadiagonalnych macierzy w metodzie Jacobiego.
- 6/13.** Zbadać czas wyznaczania m wartości własnych symetrycznej macierzy trójdzielnej rozmiaru $n \times n$ metodą bisekcji. Dla jakich m i n metoda ta jest szybsza od pozostałych?
- 6/14.** Zbadać, dla jakich rozmiarów symetrycznych macierzy trójdzielnych metoda „dziel i rządź” (podprogram SSTEDC z LAPACK-a) jest szybsza od metody QL/QR niewymiernej (SSTEQR) oraz wymiernej (SSTERF).
- 6/15.** Pokazać, że macierz w równaniu (6.46) jest symetryczna.
- 6/16.** Wyznaczyć macierze \mathbf{B} i \mathbf{C} (6.44) dla macierzy hermitowskiej

$$\mathbf{A} = \begin{pmatrix} 2 & 3 - 2i \\ 3 + 2i & 4 \end{pmatrix}.$$

Zapisać jawną postać odpowiedniego rzeczywistego zagadnienia własnego.

Zastosowania

Przedstawimy wybrane zagadnienia fizyki ciała stałego i fizyki atomowej, do których mogą być zastosowane omówione dotychczas metody numeryczne. Uwagę skupimy na dynamice drgań atomów, falach spinowych i widmach elektronowych układów jednowymiarowych. Zajmiemy się także rozwiązywaniem dwu- i trójwymiarowego równania Schrödingera z potencjałem o symetrii odpowiednio osiowej i sferycznej oraz wybranymi problemami o niższej symetrii: dwu- i trójwymiarowym atomem wodoru w zewnętrznym polu elektrycznym. W tych szczególnych przypadkach symetria potencjału pozwala na rozdzielenie zmiennych i sprowadzanie problemu do jednowymiarowego zagadnienia własnego lub układu takich zagadnień. Podamy także przykład zastosowania omawianych metod w fizyce kropek kwantowych oraz innych niskowymiarowych układów półprzewodnikowych.

7.1. Dynamika sieci łańcuchów atomów

Hamiltonian opisujący w przybliżeniu harmonicznym dynamikę drgań jednowymiarowego łańcucha atomów ma postać [94–96]

$$\mathcal{H}_f = \frac{1}{2} \sum_{i=1}^n \frac{P_i^2}{M_i} + \frac{1}{2} \sum_{i=1}^n K_{i+1,i} (U_{i+1} - U_i)^2, \quad (7.1)$$

gdzie P_i i U_i są odpowiednio pędami (w podejściu kwantowym — operatorami pędu) i wychyleniami z położenia równowagi (w podejściu kwantowym — operatorami wychyleń) i -tego atomu o masie M_i , a $K_{i\pm 1,i}$ są stałymi siłowymi zadającymi oddziaływanie pomiędzy najbliższymi sąsiadami.

7.1.1. Macierz dynamiczna

Niezależnie od podejścia (klasycznego lub kwantowego) widmo energetyczne hamiltonianu (7.1) określa układ równań ruchu

$$M_i \frac{d^2 U_i}{dT^2} = K_{i,i-1} U_{i-1} - (K_{i,i-1} + K_{i+1,i}) U_i + K_{i+1,i} U_{i+1} \quad (i = 1, \dots, n). \quad (7.2)$$

Sprowadzimy obecnie układ równań (7.2) do postaci zagadnienia własnego dla macierzy symetrycznej, którą nazywamy *macierzą dynamiczną*.

Z uwagi na postać prawych stron układu równań (7.2) przyjmujemy, że w stanie stacjonarnym

$$\mathbf{U}(T) = (U_1(T), \dots, U_n(T)) = \mathbf{Z} e^{-i\Omega T} = (Z_1, \dots, Z_n) e^{-i\Omega T},$$

gdzie Z_i — amplituda drgań i -tego atomu. Po podstawieniu do równań (7.2) otrzymujemy

$$-M_i \Omega^2 Z_i = K_{i,i-1} Z_{i-1} - (K_{i,i-1} + K_{i+1,i}) Z_i + K_{i+1,i} Z_{i+1}. \quad (7.3)$$

Nadamy równaniom (7.3) postać bezwymiarową. Wprowadźmy parametry modelu — wielkości K_c , M_c , A_c charakteryzujące oddziaływania sprężyste, masy atomów i odległości między nimi, a także częstość $\Omega_c^2 = K_c/M_c$. Wtedy z równania (7.3) otrzymujemy

$$\begin{aligned} -K_c A_c \frac{K_{i,i-1}}{K_c} \frac{Z_{i-1}}{A_c} + K_c A_c \left(\frac{K_{i,i-1}}{K_c} + \frac{K_{i+1,i}}{K_c} \right) \frac{Z_i}{A_c} - K_c A_c \frac{K_{i+1,i}}{K_c} \frac{Z_{i+1}}{A_c} = \\ = \left(\frac{\Omega}{\Omega_c} \right)^2 \Omega_c^2 \frac{M_i}{M_c} M_c A_c \frac{Z_i}{A_c} \quad (i = 1, \dots, n). \end{aligned}$$

Jak widać, bezwymiarowymi parametrami modelu są

$$k_{i,i\pm 1} = \frac{K_{i,i\pm 1}}{K_c}, \quad z_i = \frac{Z_i}{A_c}, \quad \omega = \frac{\Omega}{\Omega_c}, \quad m_i = \frac{M_i}{M_c}.$$

Zatem możemy zapisać powyższe równania w bezwymiarowej postaci

$$-k_{i,i-1} z_{i-1} + (k_{i,i-1} + k_{i+1,i}) z_i - k_{i+1,i} z_{i+1} = \omega^2 m_i z_i \quad (i = 1, \dots, n). \quad (7.4)$$

Otrzymany układ równań (7.4) nie odpowiada symetrycznemu AZW. Jeśli wprowadzimy nowe bezwymiarowe zmienne dynamiczne $\mathbf{w} = (w_1, \dots, w_n)$, gdzie

$$w_i = z_i \sqrt{m_i},$$

to, po prostych przekształceniach, otrzymujemy symetryczne AZW

$$-\frac{k_{i,i-1}}{\sqrt{m_i m_{i-1}}} w_{i-1} + \frac{k_{i,i-1} + k_{i+1,i}}{m_i} w_i - \frac{k_{i+1,i}}{\sqrt{m_i m_{i+1}}} w_{i+1} = \omega^2 w_i \quad (i = 1, \dots, n),$$

które możemy zapisać jako równanie macierzowe

$$\mathbf{T} \mathbf{w} = \omega^2 \mathbf{w},$$

gdzie \mathbf{T} jest macierzą trójdagonalną postaci (3.1). Jej diagonalne i pozadiagonalne wyrazy są odpowiednio równe

$$d_i = \frac{k_{i,i-1} + k_{i+1,i}}{m_i}, \quad e_i = \frac{-k_{i+1,i}}{\sqrt{m_i m_{i+1}}}.$$

Jeśli na rozważany układ nałożymy szczególny rodzaj warunków brzegowych: *swobodne* ($k_{0,1} = k_{n,n+1} = 0$) lub *ustalone* ($w_0 = w_{n+1} = 0$), to macierz dynamiczna \mathbf{T} jest rzeczywista, symetryczna i trójdagonalna. Wartości i wektory własne takiej macierzy wyznaczamy metodami opisanymi w poprzednich rozdziałach.

Po nałożeniu na układ swobodnych bądź ustalonych warunków brzegowych oraz uwzględnieniu oddziaływań pomiędzy dalszymi sąsiadami ($k_{i,i+j} \neq 0$, $j \geq 2$), macierz dynamiczna staje się symetryczną rzeczywistą macierzą pasmową o j nad- i poddiagonalach (szerokość pasma wynosi $2j + 1$). Zagadnienie własne dla takiej macierzy rozwiązujemy metodami opisanymi w podrozdziale 6.6. W przypadku cyklicznych warunków brzegowych ($w_{n+1} = w_1 e^{iqn}$, gdzie $q = QA_c$) i gdy uwzględniamy tylko

oddziaływania najbliższych sąsiadów, macierz dynamiczna staje się macierzą hermitowską i kwazitrójdagonalną postaci

$$\bar{\mathbf{T}} = \begin{pmatrix} d_1 & e_1 & & e_0 \\ e_1 & \ddots & \ddots & \\ & \ddots & \ddots & e_{n-1} \\ e_0^* & & e_{n-1} & d_n \end{pmatrix}, \quad (7.5)$$

gdzie

$$e_0 = \frac{-k_{n,n+1}}{\sqrt{m_n m_1}} e^{iqn}.$$

Wartości własne macierzy (7.5) możemy wyznaczać numerycznie metodą bisekcji, posługując się twierdzeniem Martina–Deana [97]. Liczba $\eta(\bar{\mathbf{T}} - x\mathbf{1})$ wartości własnych macierzy $\bar{\mathbf{T}}$ mniejszych od danej liczby rzeczywistej x wynosi

$$\eta(\bar{\mathbf{T}} - x\mathbf{1}) = \sum_{i=1}^n \eta(u_i),$$

gdzie (u_i) jest następującym ciągiem liczbowym:

$$\begin{aligned} u_1 &= d_1 - x, & u_i &= d_i - x - \frac{e_{i-1}^2}{u_{i-1}} \quad (i = 2, \dots, n-1), \\ u_n &= d_n - \sum_{i=1}^{n-2} \frac{s_i^2}{u_i} - x - \frac{(e_{n-1} + s_{n-1})^2}{u_{n-1}}, \\ s_1 &= e_0, & s_i &= -\frac{e_{i-1}s_{i-1}}{u_{i-1}} \quad (i = 2, \dots, n-1), \end{aligned} \quad (7.6)$$

W przypadku uwzględnienia oddziaływań dalszych sąsiadów i nałożenia cyklicznych warunków brzegowych, w lewym dolnym i prawym górnym narożniku macierzy dynamicznej pojawiają się bloki rozmiaru $j \times j$ (odpowiednio macierz trójkątna górna i dolna). Macierz dynamiczna jest wtedy zespolona i hermitowska. Do wyznaczania jej wartości własnych można również zastosować twierdzenie Martina–Deana [97].

7.1.2. Widmo częstości własnych

Funkcja rozkładu częstości $g(\omega)$ [95, 98] określa liczbę różnych częstości własnych translacyjnie niezmienniczego układu przypadających na jednostkowy przedział częstości. W celu otrzymania jawnej postaci $g(\omega)$ zauważmy, że n wartości wektorów q z pierwszej strefy Brillouina przypada na odcinek o długości 2π . Zatem

$$g(\omega) d\omega = 2 \frac{n}{2\pi} dq \quad \Rightarrow \quad g(\omega) = \frac{n}{\pi} \frac{dq}{d\omega}.$$

O funkcji $g(\omega)$ mówimy także jako o *gęstości częstości własnych kryształu* lub *funkcji gęstości stanów fononowych*.

Wyznamy jawną postać $g(\omega)$ dla monoatomowego okresowego łańcucha atomów. Zajmiemy się najpierw funkcją $g(\omega^2)$. Zauważmy, że

$$g(\omega^2) = \frac{n}{\pi} \frac{1}{d\omega^2/dq}, \quad 0 \leq \omega^2 \leq 4$$

(patrz zadanie 7/3). Ponieważ

$$\omega^2(q) = 2(1 - \cos q), \quad \frac{d\omega^2}{dq} = 2 \sin q,$$

to

$$\cos q = 1 - \frac{\omega^2}{2}, \quad 2 \sin q = \sqrt{\omega^2(4 - \omega^2)}.$$

Ostatecznie otrzymujemy

$$g(\omega^2) = \frac{n}{\pi} \frac{1}{\sqrt{\omega^2(4 - \omega^2)}}. \quad (7.7)$$

Jak widać, funkcja gęstości $g(\omega^2)$ wykazuje istnienie osobliwości typu $(\omega^2)^{-1/2}$ i $(4 - \omega^2)^{-1/2}$ odpowiednio dla $\omega^2 \rightarrow 0$ i $\omega^2 \rightarrow 4$. Osobliwości te nazywane są osobliwościami Van Hove'a.

Za pomocą przedstawionych metod można badać dynamikę sieci nieuporządkowanych łańcuchów atomów. Układy takie można modelować na różne sposoby [95, 99–101], zadając:

- losowy rozkład stałych sprężystych albo mas atomów,
- niezależne losowe rozkłady stałych sprężystych i mas atomów.

W takich układach spodziewamy się punktowego charakteru widma energetycznego i *zlokalizowanych* wektorów własnych [100, 101], tj. takich, w których amplitudy wychyleń maleją wykładniczo wraz z odległością od określonego punktu, zwanego *punktem lokalizacji*.

W drugiej połowie lat 80. i w latach 90. XX wieku przedmiotem intensywnych badań w fizyce ciała stałego były struktury kwaziperiodyczne, nazywane *kwazikryształami* albo *kryształami aperiodycznymi* [102–112]. Jednowymiarowy kwazikryształ typu Fibonacciego K_l rzędu l jest typowym przykładem, dlatego będziemy mówili o nim jako o paradygmacie. Otrzymujemy go, stosując do dwóch składników (komórek elementarnych) A i B regułę podstawiania

$$B \mapsto A, \quad A \mapsto AB,$$

co równoważne jest konstrukcji rekurencyjnej

$$K_0 = B, \quad K_1 = A, \quad K_2 = K_1 K_0 = AB, \quad \dots, \quad K_l = K_{l-1} K_{l-2}.$$

Podobnie jak w przypadku nieuporządkowanych łańcuchów, modulacji mogą podlegać położenia lub masy atomów. W pierwszym przypadku (łańcuch monoatomowy) atomy położone są w węzłach jednowymiarowej kwazisieci

$$x_i = i + \frac{1}{\zeta} \left[\frac{i}{\zeta} \right] \quad (i = 1, 2, \dots),$$

gdzie $\zeta = \frac{1}{2}(1 + \sqrt{5})$ jest liczbą niewymierną określającą złoty podział odcinka. Oddziaływanie najbliższych sąsiadów zadają stałe siłowe

$$k_{i,i+1} = k_0[1 + \xi(1 - d_{i,i+1})]$$

zależne od bezwymiarowych odległości

$$d_{i,i+1} = x_{i+1} - x_i = \begin{cases} 1 & (\lfloor \frac{i+1}{\zeta} \rfloor - \lfloor \frac{i}{\zeta} \rfloor = 0) \\ 1 + \frac{1}{\zeta} & (\lfloor \frac{i+1}{\zeta} \rfloor - \lfloor \frac{i}{\zeta} \rfloor = 1) \end{cases} \quad (7.8)$$

między sąsiednimi węzłami sieci. Parametr ξ jest *parametrem kwaziperiodyczności*.

W drugim przypadku (tzw. binarny łańcuch Fibonacciego) zakłada się, że okresowa sieć jest *dekorowana* kwaziperiodycznie dwoma rodzajami atomów o masach m_A i m_B , natomiast $k_{i,i+1} = k_0 = \text{const}$. Zapelnianie węzłów sieci dokonuje się zgodnie z regułą

$$m_i = \begin{cases} m_A & (\lfloor \frac{i+1}{\zeta} \rfloor - \lfloor \frac{i}{\zeta} \rfloor = 0) \\ m_B & (\lfloor \frac{i+1}{\zeta} \rfloor - \lfloor \frac{i}{\zeta} \rfloor = 1). \end{cases}$$

Dokładniejsza analiza dynamiki drgań łańcucha Fibonacciego, w którym jednocześnie modulowane są masy atomów i stałe sprężyste [97, 104, 111–115], pokazała, że widmo częstości własnych jest *fraktalem*¹⁾, a funkcja rozkładu częstości własnych

$$\bar{g}(z) = \frac{1}{n} \eta(\mathbf{T} - z\mathbf{1}),$$

jest *multifraktalem*²⁾ (przypomnijmy, że $\eta(\mathbf{T} - z\mathbf{1})$ jest liczbą wartości własnych macierzy \mathbf{T} mniejszych od liczby rzeczywistej z). Funkcja gęstości częstości własnych nie jest określona w punktach ω_w należących do widma, ponieważ nie istnieje granica

$$\lim_{r \rightarrow 0} \frac{\bar{g}(\omega_w + r) - \bar{g}(\omega_w)}{r}.$$

Widmo energetyczne ma charakter zbioru Cantora, natomiast funkcja $\bar{g}(z)$ ma charakter funkcji Cantora zwanej *diabelską drabiną* [116]. Więcej informacji na ten temat można znaleźć w cytowanych wyżej pracach oraz podręcznikach matematycznych [123, 124].

Znając widmo drgań łańcucha atomów, możemy zbadać jego własności termodynamiczne [125], ponieważ energia wewnętrzna \mathcal{E} układu w temperaturze Θ wynosi

$$\mathcal{E}(\Theta) = \sum_{i=1}^n \frac{\hbar \omega_i}{\exp\left(\frac{\hbar \omega_i}{k_B \Theta}\right) - 1}.$$

¹⁾ Wymiar zbioru będącego widmem fononowym jest ułamkowy [116–120]

²⁾ Multifraktalem nazywamy obiekt fraktalny, którego wymiar fraktalny przyjmuje wartości z określonego przedziału [119–122]

7.2. Model silnego wiązania

Rozpatrzmy dynamikę elektronu w jednowymiarowym układzie, poruszającego się w potencjale [94, 96]

$$V(X) = \sum_{i=1}^n V^a(X - X_i). \quad (7.9)$$

Potencjał $V(X)$ nie musi być funkcją okresową; jego źródłem są atomy ułożone na prostej w punktach X_i ($i = 1, \dots, n$).

Zakładamy, że znane jest rozwiązanie *atomowego* zagadnienia własnego w przypadku translacyjnie niezmienniczego, jednowymiarowego układu:

$$\frac{\hbar^2}{2M} \frac{d^2 \Phi^{(i)}(X)}{dX^2} + [\mathcal{E}_i^a - V_i^a(X - X_i)] \Phi^{(i)}(X) = 0 \quad (i = 1, \dots, n),$$

tj. znamy energie \mathcal{E}_i^a elektronów w atomowych stanach związanych, których funkcje falowe $\Phi^{(i)}(X) = \Phi(X - X_i)$ są także znane.

Naszym zadaniem jest rozwiązanie równania Schrödingera z potencjałem (7.9):

$$\frac{\hbar^2}{2M} \frac{d^2 \Psi(X)}{dX^2} + [\mathcal{E} - V(X)] \Psi(X) = 0. \quad (7.10)$$

Zastosujemy w tym celu podejście wariacyjne. Zakładamy, że funkcja falowa $\Psi(X)$ jest kombinacją liniową funkcji atomowych

$$\Psi(X) = \sum_i c_i \Phi(X - X_i) = \sum_i c_i \Phi^{(i)}(X).$$

Wyznamy współczynniki c_i z warunku stacjonarności wartości oczekiwanej operatora energii. Mnożymy równanie (7.10) lewostronnie przez $\sum_i c_i^* \Phi^{(i)}$, a prawostronnie przez $\sum_j c_j \Phi^{(j)}$ i następnie całkujemy po X . Otrzymujemy

$$\mathcal{E} = \frac{\sum_{i,j} c_i^* c_j H_{i,j}}{\sum_{i,j} c_i^* c_j J_{i,j}}, \quad (7.11)$$

gdzie

$$H_{i,j} = \int_{-\infty}^{\infty} \Phi^{(i)}(X) \left[-\frac{\hbar^2}{2M} \frac{d^2}{dX^2} + V(X) \right] \Phi^{(j)}(X) dX \quad (7.12)$$

oraz

$$J_{i,j} = \int_{-\infty}^{\infty} \Phi^{(i)}(X) \Phi^{(j)}(X) dX.$$

Wielkości $J_{i,j}$ określa się mianem *całek przekrywania*. Założyliśmy tu, że atomowe funkcje falowe $\Phi^{(i)}$ są rzeczywiste.

Z równania (7.12) wynika następujący związek:

$$H_{i,j} = \mathcal{E}_i^a J_{i,j} + \int_{-\infty}^{\infty} \Phi^{(i)} [V(X) - V^a(X - X_j)] \Phi^{(j)} = \mathcal{E}_i^a J_{i,j} + V'_{i,j},$$

gdzie $V'_{i,j}$ są elementami macierzowymi potencjału $V(X) - V^a(X - X_j)$.

Dla równania (7.11) warunek stacjonarności ma postać

$$\frac{\partial \mathcal{E}}{\partial c_i^*} = 0 \quad (i = 1, \dots, n).$$

Otrzymujemy z niego jednorodny układ równań

$$\sum_j [H_{i,j} c_j - J_{i,j} \mathcal{E} c_j] = \sum_j [V'_{i,j} - (\mathcal{E} - \mathcal{E}_i^a) J_{i,j}] c_j = 0 \quad (i = 1, \dots, n). \quad (7.13)$$

Ma on niezerowe rozwiązanie $\mathbf{c} = (c_1, \dots, c_n)$ pod warunkiem, że jego wyznacznik jest równy zeru, co prowadzi do zagadnienia wiekowego.

Zapisując otrzymany jednorodny układ równań (7.13) w nieco innej postaci

$$\sum_j V'_{i,j} c_j + \sum_j J_{i,j} \mathcal{E}_i^a c_j = \mathcal{E} \sum_j J_{i,j} c_j \quad (i = 1, \dots, n),$$

łatwo zauważamy, że jest to uogólnione zagadnienie własne

$$(\mathbf{V}' + \mathbf{E}^a \mathbf{J}) \mathbf{c} = \mathcal{E} \mathbf{J} \mathbf{c},$$

gdzie \mathbf{V}' i \mathbf{J} są macierzami o wyrazach odpowiednio $V'_{i,j}$ i $J_{i,j}$, a $\mathbf{E}^a = \text{diag}(\mathcal{E}_1^a, \dots, \mathcal{E}_n^a)$. Wprowadzając nowe zmienne $\mathbf{c}' = \mathbf{J} \mathbf{c}$ otrzymujemy standardowe algebraiczne zagadnienie własne

$$(\mathbf{V}' \mathbf{J}^{-1} + \mathbf{E}^a) \mathbf{c}' = \mathcal{E} \mathbf{c}'. \quad (7.14)$$

W przypadku okresowego układu jednowymiarowego możemy założyć, że:

- $V'_{i,j} = V_c \neq 0$ dla $i = j \pm 1$,
- całki przekrywania $J_{i,j} = J_c \delta_{i,j}$, $J_c \neq 0$.

Wtedy układ równań (7.14) przyjmuje postać

$$\begin{pmatrix} \varepsilon_1^a & v'_1 & & \\ v'_1 & \ddots & \ddots & \\ & \ddots & \ddots & v'_{n-1} \\ & & v'_{n-1} & \varepsilon_n^a \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ \vdots \\ c_n \end{pmatrix} = \varepsilon \begin{pmatrix} c_1 \\ \vdots \\ \vdots \\ c_n \end{pmatrix}, \quad (7.15)$$

gdzie

$$\varepsilon_i^a = \frac{\mathcal{E}_i^a}{\mathcal{E}_c}, \quad v'_i = \frac{V'_{i,i+1}}{J_c \mathcal{E}_c}, \quad \varepsilon = \frac{\mathcal{E}}{\mathcal{E}_c}.$$

Dla monoatomowego łańcucha wszystkie wielkości v'_i oraz ε_i^a przyjmują ustalone wartości v' i ε^a . Rozwiązań w tym przypadku możemy poszukiwać w postaci

$$c_i = c e^{iqx_i} = c e^{iqi}$$

(przyjęliśmy tu jednostkowe odległości międzyatomowe, $x_i = i$). Po podstawieniu do (7.13) otrzymujemy równanie

$$\varepsilon(q) - \varepsilon^a = 2v' \cos q,$$

gdzie q jest wektorem z pierwszej strefy Brillouina. Szerokość pasma energetycznego wynosi $|4v'|$.

W ogólnym przypadku wartości v'_i oraz ε_i^a należy traktować jako parametry modelu. Podobnie jak w przypadku drgań łańcuchów atomów, spodziewamy się, że charakter widma energetycznego elektronów i funkcji falowej będzie zależał od tych parametrów.

Jednowymiarowy model silnego wiązania jest przedmiotem wielu prac [103–109]. Rozpatrywano zarówno modele nieuporządkowanych, jak i kwaziperiodycznych struktur jednowymiarowych. Dynamikę elektronów w łańcuchach nieuporządkowanych modeluje się w następujący sposób:

- Wyrazy diagonalne w równaniu (7.15) są zadawane w określony (losowy lub deterministyczny) sposób, a wyrazy pozadiagonalne są ustalone. Model taki zwany jest modelem z *nieuporządkowaniem typu diagonalnego*.
- Wyrazy diagonalne w równaniu (7.15) są stałe, zaś pozadiagonalne są zadawane losowo lub deterministycznie. Model taki zwany jest modelem z *nieuporządkowaniem typu pozadiagonalnego*.
- Zarówno wyrazy diagonalne, jak i pozadiagonalne macierzy w równaniu (7.15) są zadawane losowo lub deterministycznie.

W modelach takich bada się charakter elektronowego widma energetycznego oraz właściwości elektronowych funkcji falowych. Głównym przedmiotem badań było zagadnienie lokalizacji i delokalizacji elektronów, odpowiadające przejściu metal–izolator.

7.3. Dynamika spinowa układów jednowymiarowych

Zajmiemy się teraz innym interesującym zagadnieniem dotyczącym widma wzbudzeń elementarnych łańcucha atomów obdarzonych własnymi momentami magnetycznymi (spinami).

Hamiltonian takiego układu ma w modelu Heisenberga [96 (r. 33), 126, 127], postać

$$\mathcal{H} = -\frac{1}{2} \sum_{i,j; i \neq j} J_{i,j} \mathbf{s}_i \cdot \mathbf{s}_j - \mu_B \sum_i B_z s_i^z,$$

gdzie μ_B jest magnetonem Bohra, B_z jest zewnętrznym polem magnetycznym skierowanym wzdłuż osi OZ , a składowe operatory spinu spełniają reguły komutacji

$$[s_i^x, s_j^y] = i s_i^z \delta_{i,j}, \quad [s_i^y, s_j^z] = i s_i^x \delta_{i,j}, \quad [s_i^z, s_j^x] = i s_i^y \delta_{i,j}. \quad (7.16)$$

Wygodnie jest wprowadzić operatory kreacji i anihilacji (*podnoszenia i opuszczania*)

$$s_i^+ = s_i^x + i s_i^y, \quad s_i^- = (s_i^+)^* = s_i^x - i s_i^y,$$

które spełniają reguły komutacyjne

$$[s_i^+, s_j^-] = 2 s_i^z \delta_{i,j}, \quad [s_i^z, s_j^\pm] = \pm s_i^\pm \delta_{i,j}. \quad (7.17)$$

Stosując operatory kreacji i anihilacji oraz wykorzystując reguły komutacji możemy łatwo sprawdzić, że

$$\mathbf{s}_i \cdot \mathbf{s}_j = \frac{1}{2} (s_i^+ s_j^- + s_i^- s_j^+) + s_i^z s_j^z = s_i^+ s_j^- + s_i^z s_j^z. \quad (7.18)$$

Zajmiemy się teraz widmem wzbudzeń elementarnych rozważanego układu. W tym celu napiszemy równania ruchu dla operatorów kreacji

$$i\hbar \frac{ds_j^+}{dT} = -[\mathcal{H}, s_j^+] = \mu_B B_z s_j^+ - \sum_i J_{i,j} (s_i^+ s_j^z - s_i^z s_j^+) \quad (j = 1, \dots, n). \quad (7.19)$$

Stacjonarnych rozwiązań układu równań (7.19) będziemy poszukiwali w postaci

$$s_i^+(T) = s_i^+ e^{-i\Omega T},$$

zakładając, że

- spiny są uporządkowane równolegle do osi OZ ,
- wartości składowych z -owych operatorów spinu spełniają relację $s_i^z \simeq s$, gdzie s jest wartością spinu,
- całka wymiany $J_{i,j} \neq 0$ dla $i = j \pm 1$ (uwzględniamy jedynie oddziaływanie pomiędzy najbliższymi sąsiadami).

Wtedy układ równań (7.19) przyjmuje postać

$$\hbar\Omega s_j^+ = \mu_B B_z s_j^+ + s(J_{j-1,j} + J_{j+1,j})s_j^+ - J_{j-1,j}ss_{j-1}^+ - J_{j+1,j}ss_{j+1}^+, \quad (7.20)$$

równoważną zagadnieniu własnemu

$$\begin{pmatrix} d_1 & e_1 & & \\ e_1 & \ddots & \ddots & \\ & \ddots & \ddots & e_{n-1} \\ & & e_{n-1} & d_n \end{pmatrix} \begin{pmatrix} s_1^+ \\ \vdots \\ \vdots \\ s_n^+ \end{pmatrix} = \hbar\Omega \begin{pmatrix} s_1^+ \\ \vdots \\ \vdots \\ s_n^+ \end{pmatrix},$$

gdzie

$$\begin{aligned} d_1 &= \mu_B B_z + sJ_{1,2}, \\ d_j &= \mu_B B_z + s(J_{j-1,j} + J_{j,j+1}) \quad (j = 2, \dots, n-1), \\ d_n &= \mu_B B_z + sJ_{n-1,n}, \\ e_j &= sJ_{j,j+1} \quad (j = 1, \dots, n-1). \end{aligned}$$

W przypadku okresowego łańcucha monoatomowego możemy założyć następującą postać rozwiązania równania (7.20)

$$s_j^+ = se^{iqj},$$

które prowadzi do znanego wyrażenia na energię wzbudzeń magnonowych

$$\hbar\Omega(q) = \mu_B B_z + 2sJ(1 - \cos q) = \mu_B B_z + 4sJ \sin^2(\frac{1}{2}q),$$

gdzie J jest wartością całki wymiany pomiędzy najbliższymi sąsiadami.

Widmo spinowych wzbudzeń elementarnych w układach nieuporządkowanych oraz kwaziperiodycznych można badać w sposób analogiczny do przedstawionego w podrozdziałach dotyczących dynamiki sieci i dynamiki elektronów.

7.4. Zagadnienie radialne

W tym podrozdziale pokazujemy, w jaki sposób można sprowadzić trójwymiarowe równanie Schrödingera ze sferycznie symetrycznym potencjałem $V(\mathbf{R}) = V(R)$ do postaci jednowymiarowego zagadnienia własnego. Wejściowe równanie Schrödingera jest następujące [30 (r. 5)]:

$$\left[-\frac{\hbar^2}{2M} \nabla^2 + V(R) \right] \Psi(\mathbf{R}) = \mathcal{E} \Psi(\mathbf{R}). \quad (7.21)$$

We współrzędnych sferycznych laplasjan ma postać

$$\nabla^2 = \frac{1}{R^2} \left\{ \frac{\partial}{\partial R} \left(R^2 \frac{\partial}{\partial R} \right) + \left[\frac{1}{\sin^2 \vartheta} \frac{\partial^2}{\partial \varphi^2} + \frac{1}{\sin \vartheta} \frac{\partial}{\partial \vartheta} \left(\sin \vartheta \frac{\partial}{\partial \vartheta} \right) \right] \right\}.$$

Będziemy szukali rozwiązań równania (7.21) w postaci

$$\Psi(\mathbf{R}) = U(R) Y_{l,m}(\vartheta, \varphi),$$

gdzie $Y_{l,m}$ ($l = 0, 1, \dots$; $m = -l, \dots, l$) — harmoniki sferyczne, a funkcja radialna $U(R)$ spełnia równanie

$$\left[\frac{1}{R^2} \frac{d}{dR} \left(R^2 \frac{d}{dR} \right) - \frac{l(l+1)}{R^2} + \frac{2M}{\hbar^2} [\mathcal{E} - V(R)] \right] U(R) = 0.$$

Nadamy ostatniemu równaniu postać podobną do rozpatrywanej w poprzednich przypadkach. W tym celu wprowadzimy nową funkcję χ , taką że

$$U(R) = \chi(R)/R.$$

Mamy

$$\frac{dU}{dR} = \frac{1}{R} \frac{d\chi}{dR} - \frac{\chi}{R^2}, \quad R^2 \frac{dU}{dR} = R \frac{d\chi}{dR} - \chi, \quad \frac{d}{dR} \left(R \frac{d\chi}{dR} - \chi \right) = \frac{d^2 \chi}{dR^2}$$

oraz

$$\frac{1}{R^2} \frac{d\chi}{dR} \left(R^2 \frac{dU}{dR} \right) = \frac{1}{R} \frac{d^2 \chi}{dR^2}.$$

Z powyższych relacji otrzymujemy

$$\left[-\frac{d^2}{dR^2} + \frac{2M}{\hbar^2} \left(V(R) + \frac{\hbar^2 l(l+1)}{2MR^2} \right) \right] \chi(R) = \frac{2M\mathcal{E}}{\hbar^2} \chi(R). \quad (7.22)$$

Sprecyzujemy obecnie postać potencjału. Niech energię potencjalną zadaje funkcja

$$V(R) = \frac{K}{R}$$

(np. potencjał kulombowski), gdzie parametr K ma wymiar $[\text{J} \cdot \text{m}]$. Wprowadzimy atomowe jednostki masy, długości, czasu i energii:

$$M_c = M, \quad L_c = \frac{\hbar^2}{M|K|}, \quad T_c = \frac{\hbar^3}{MK^2}, \quad V_c = \frac{MK^2}{2\hbar^2}. \quad (7.23)$$

Używając tych jednostek nadamy równaniu (7.22) postać bezwymiarową. Za-uważmy, że współczynniki wszystkich jego wyrazów mają wymiar $[\text{m}^{-2}]$. Dlatego

mnożymy je obustronnie przez L_c^2 , co prowadzi do poszukiwanego zagadnienia własnego

$$\left[-\frac{d^2}{dr^2} + \left(\frac{l(l+1)}{r^2} + \frac{2 \operatorname{sgn} K}{r} \right) \right] \chi(r) = \varepsilon \chi(r), \quad (7.24)$$

gdzie $r = R/L_c$, $\varepsilon = \mathcal{E}/V_c$.

W przypadku gdy $K < 0$ (tak jest dla atomu wodoru), potencjał $V(R)$ jest wiążący i równanie (7.24) można rozwiązywać numerycznie za pomocą metod macierzowych omówionych w rozdziałach 2–5. Znane są także analityczne wyrażenia na wartości własne w równaniu (7.24):

$$\varepsilon_n = -\frac{1}{n^2}, \quad (n = 1, 2, \dots). \quad (7.25)$$

Dokładnie rozwiązywalne zagadnienie atomu wodoru może być wykorzystane do testowania programów komputerowych używanych następnie do numerycznego rozwiązywania równania Schrödingera w przypadkach, gdy nie są znane rozwiązania analityczne.

7.5. Atom wodoru w zewnętrznym stałym polu elektrycznym

Zaprezentujemy sposób sprowadzenia problemu trój- i dwuwymiarowego atomu wodoru poddanego działaniu zewnętrznego stałego pola elektrycznego [30 (r. 5 i 10), 128] do układu dwóch jednowymiarowych zagadnień własnych. Pozwoli to nam badać efekt Starka w atomie wodoru oraz wpływ zewnętrznego pola elektrycznego na energię wiązania trój- i dwuwymiarowego ekscytanu Wanniera–Motta za pomocą metod macierzowych.

Z uwagi na umieszczenie rozpatrywanych układów w polu elektrycznym, rozważane zagadnienia mają charakter kwazistacjonarny, co oznacza, że czas życia każdego poziomu energetycznego jest skończony. Innymi słowy, potencjał wchodzący w równanie Schrödingera nie wykazuje globalnego minimum w sensie, o którym była mowa w rozdziale 2.

7.5.1. Przypadek trójwymiarowy

Równanie Schrödingera opisujące rozpatrywany układ ma postać [30]

$$\left[-\frac{\hbar^2}{2M} \left(\frac{\partial^2}{\partial X^2} + \frac{\partial^2}{\partial Y^2} + \frac{\partial^2}{\partial Z^2} \right) + \frac{K}{R} + eEZ \right] \Phi(X, Y, Z) = \mathcal{E} \Phi(X, Y, Z), \quad (7.26)$$

gdzie $K = -e^2/(4\pi\epsilon_0\epsilon)$, a E jest natężeniem zewnętrznego pola elektrycznego równoległego do osi OZ . Obok jednostek atomowych (7.23) w postaci

$$V_c = \frac{Me^4}{32\pi^2\epsilon_0^2\epsilon^2\hbar^2}, \quad L_c = \frac{4\pi\epsilon_0\epsilon\hbar^2}{Me^2}, \quad M_c = M,$$

będziemy używać jednostki natężenia pola elektrycznego

$$E_c = V_c/(eL_c).$$

Wtedy wektorem wodzącym jest $\mathbf{r} = \mathbf{R}/L_c$, a natężeniem pola $f = E/E_c$. Jest to duża jednostka, której wartość jest porównywalna z natężeniem pola powodującego jonizację atomu.

W powyższych jednostkach równanie (7.26) przyjmuje bardzo prostą postać:

$$\left[- \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) - \frac{2}{r} + fz \right] \phi(x, y, z) = \varepsilon \phi(x, y, z). \quad (7.27)$$

W celu rozdzielania zmiennych w równaniu (7.27), wprowadzimy ortogonalne współrzędne paraboliczne (ξ, η, φ) , związane ze współrzędnymi kartezjańskimi następującymi relacjami:

$$\begin{aligned} x &= \sqrt{\xi\eta} \cos \varphi, & \xi &= r + z, \\ y &= \sqrt{\xi\eta} \sin \varphi, & \eta &= r - z, \\ z &= \frac{1}{2}(\xi - \eta), & \varphi &= \arg(x + iy), \\ r &= \sqrt{x^2 + y^2 + z^2} = \frac{1}{2}(\xi + \eta). \end{aligned}$$

Zauważmy, że $\xi, \eta > 0$. Wyrazimy element objętości i laplasjan przez współrzędne paraboliczne:

$$\begin{aligned} d^3r &= \frac{1}{4}(\xi + \eta) d\xi d\eta d\varphi, \\ \nabla^2 &= \frac{4}{\xi + \eta} \left[\frac{\partial}{\partial \xi} \left(\xi \frac{\partial}{\partial \xi} \right) + \frac{\partial}{\partial \eta} \left(\eta \frac{\partial}{\partial \eta} \right) \right] + \frac{1}{\xi\eta} \frac{\partial^2}{\partial \varphi^2}. \end{aligned}$$

W nowych zmiennych równanie (7.27) przyjmuje postać

$$\left\{ - \frac{4}{\xi + \eta} \left[\frac{\partial}{\partial \xi} \left(\xi \frac{\partial}{\partial \xi} \right) + \frac{\partial}{\partial \eta} \left(\eta \frac{\partial}{\partial \eta} \right) \right] - \frac{1}{\xi\eta} \frac{\partial^2}{\partial \varphi^2} - \frac{4}{\xi + \eta} + \frac{\xi - \eta}{2} f \right\} \phi = \varepsilon \phi. \quad (7.28)$$

Pomnożymy równanie (7.28) przez czynnik $[-\frac{1}{4}(\xi + \eta)]$. Otrzymamy

$$\frac{\partial}{\partial \xi} \left(\xi \frac{\partial \phi}{\partial \xi} \right) + \frac{\partial}{\partial \eta} \left(\eta \frac{\partial \phi}{\partial \eta} \right) + \frac{\xi + \eta}{4\xi\eta} \frac{\partial^2 \phi}{\partial \varphi^2} + \phi - \frac{\xi^2 - \eta^2}{8} f \phi + \frac{\xi + \eta}{4} \varepsilon \phi = 0. \quad (7.29)$$

Rozwiązań równania (7.29) poszukamy w postaci $\phi(\xi, \eta, \varphi) = g_1(\xi)g_2(\eta)e^{im\varphi}$ [30]. Po podzieleniu przez ϕ otrzymujemy

$$\frac{1}{g_1} \frac{d}{d\xi} \left(\xi \frac{dg_1}{d\xi} \right) + \frac{1}{g_2} \frac{d}{d\eta} \left(\eta \frac{dg_2}{d\eta} \right) - \frac{m^2}{4\xi} - \frac{m^2}{4\eta} + 1 - \frac{\xi^2 - \eta^2}{8} f + \frac{\xi + \eta}{4} \varepsilon = 0.$$

Równanie to zawiera stałe oraz wyrazy zależne tylko od ξ i tylko od η . Może ono być spełnione dla dowolnych (ξ, η) wtedy i tylko wtedy, gdy

$$\begin{cases} \frac{1}{g_1} \frac{d}{d\xi} \left(\xi \frac{dg_1}{d\xi} \right) - \frac{m^2}{4\xi} + \beta_1 - \frac{f\xi^2}{8} + \frac{\varepsilon\xi}{4} = 0 \\ \frac{1}{g_2} \frac{d}{d\eta} \left(\eta \frac{dg_2}{d\eta} \right) - \frac{m^2}{4\eta} + \beta_2 + \frac{f\eta^2}{8} + \frac{\varepsilon\eta}{4} = 0, \end{cases}$$

gdzie parametry separacji β_1, β_2 spełniają związek

$$\beta_1 + \beta_2 = 1.$$

Z powyższego układu równań wynika para związków

$$\begin{cases} \left[\frac{d}{d\xi} \left(\xi \frac{d}{d\xi} \right) + \left(\beta_1 - \frac{m^2}{4\xi} - \frac{f\xi^2}{8} + \frac{\varepsilon\xi}{4} \right) \right] g_1(\xi) = 0 \\ \left[\frac{d}{d\eta} \left(\eta \frac{d}{d\eta} \right) + \left(\beta_2 - \frac{m^2}{4\eta} + \frac{f\eta^2}{8} + \frac{\varepsilon\eta}{4} \right) \right] g_2(\eta) = 0. \end{cases} \quad (7.30)$$

W celu pozbycia się wyrazów z pierwszymi pochodnymi w równaniach (7.30) wprowadzimy nowe funkcje χ_1, χ_2 , takie że

$$g_1(\xi) = \chi_1(\xi)/\sqrt{\xi}, \quad g_2(\eta) = \chi_2(\eta)/\sqrt{\eta}.$$

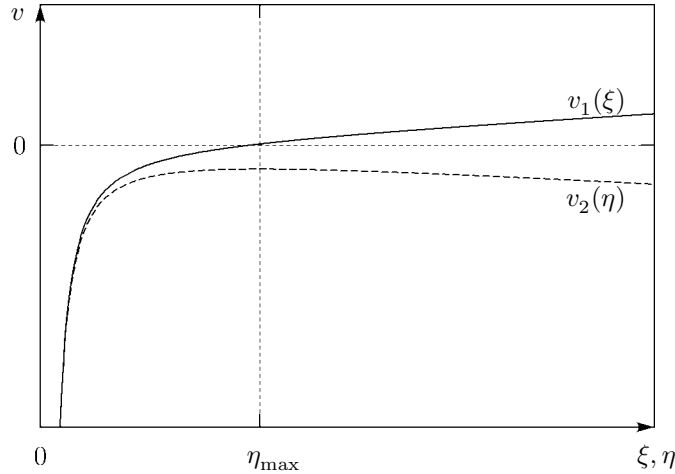
Z definicji funkcji χ_1, χ_2 wynikają warunki brzegowe $\chi_1(0) = \chi_2(0) = 0$.

Po prostych przekształceniach otrzymujemy układ dwóch sprzężonych równań różniczkowych

$$\begin{cases} \left[-\frac{d^2}{d\xi^2} + \left(\frac{m^2-1}{4\xi^2} - \frac{\beta_1}{\xi} + \frac{f\xi}{8} \right) \right] \chi_1(\xi) = \frac{\varepsilon}{4} \chi_1(\xi) \\ \left[-\frac{d^2}{d\eta^2} + \left(\frac{m^2-1}{4\eta^2} - \frac{\beta_2}{\eta} - \frac{f\eta}{8} \right) \right] \chi_2(\eta) = \frac{\varepsilon}{4} \chi_2(\eta). \end{cases} \quad (7.31)$$

Mają one postać równań Schrödingera, w których rolę potencjałów odgrywają funkcje

$$v_1(\xi) = \frac{m^2-1}{4\xi^2} - \frac{\beta_1}{\xi} + \frac{f\xi}{8}, \quad v_2(\eta) = \frac{m^2-1}{4\eta^2} - \frac{\beta_2}{\eta} - \frac{f\eta}{8}. \quad (7.32)$$



Rys. 7.1. Efektywne potencjały jednowymiarowe v_1, v_2 w zagadnieniu trójwymiarowego atomu w zewnętrznym polu elektrycznym dla stanu podstawowego ($m = 0$)

Zauważmy, że przy $f > 0$ potencjał $v_1(\xi)$ jest zawsze wiążący (jamopodobny), czego nie można powiedzieć o potencjale $v_2(\eta)$, który ma charakter bariery (rys. 7.1).

Ta własność jest źródłem wspomnianej kwazistacjonarności rozpatrywanego zagadnienia.

Numeryczną analizę układu równań (7.31) wygodnie jest prowadzić przy założeniu, że parametry separacji mają postać $\beta_1 = \frac{1}{2} + \delta$ oraz $\beta_2 = \frac{1}{2} - \delta$. W obliczeniach, zmieniając δ , wyznaczamy wartości własne $\frac{1}{4}\varepsilon(\beta_1)$ i $\frac{1}{4}\varepsilon(\beta_2)$ z równań (7.31). Kiedy znajdujemy takie δ_0 , że $\varepsilon(\beta_1) = \varepsilon(\beta_2) = \varepsilon$, będziemy uważać ε za szukaną wartość własną równania (7.27).

Dodatkowego komentarza wymaga sposób nakładania na badany układ warunków brzegowych. W przypadku wiążącego potencjału $v_1(\xi)$ nie napotykamy na żadne trudności kładąc $\chi_1(\xi_{\max}) = 0$ dla dostarcznie dużej wartości ξ_{\max} . Drugie równanie układu (7.31) rozwiązujemy, umieszczając nieskończoną barierę potencjału w punkcie o współrzędnej odpowiadającej jego maksimum: $v_2(\eta_{\max}) = \infty$. Ten sposób numerycznej analizy został zastosowany w pracach [128].

7.5.2. Przypadek dwuwymiarowy

Rozpatrzmy teraz dwuwymiarowy model atomu wodoru lub ekscytanu Wanniera–Motta, w którym ruch związanych cząstek ograniczony jest do płaszczyzny XOY . Potencjał oddziaływania elektrostatycznego przyjmujemy w postaci $1/r$, odpowiadającej przestrzeni trójwymiarowej³⁾.

Jeśli zewnętrzne pole elektryczne skierowane jest wzdłuż osi OX , to po wprowadzeniu jednostek (7.23) otrzymujemy bezwymiarowe równanie Schrödingera

$$\left[- \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) - \frac{2}{r} + fx \right] \psi(x, y) = \varepsilon \psi(x, y).$$

Aby rozdzielić zmienne tego równania, zastosujemy ortogonalne współrzędne paraboliczne płaskie (ξ, η) , związane z kartezjańskimi (x, y) wzorami

$$\begin{aligned} x &= \frac{1}{2}(\xi^2 - \eta^2) & \xi &= \operatorname{sgn} x \sqrt{r+x} \\ y &= \xi\eta & \eta &= \sqrt{r-x} \\ r &= \sqrt{x^2 + y^2} = \frac{1}{2}(\xi^2 + \eta^2). \end{aligned} \quad (7.33)$$

Zauważmy, że przy takim określeniu $\eta \geq 0$. Element powierzchni we współrzędnych (7.33) ma postać

$$d^2r = (\xi^2 + \eta^2) d\xi d\eta,$$

a laplasjan

$$\nabla^2 = \frac{1}{\xi^2 + \eta^2} \left(\frac{\partial^2}{\partial \xi^2} + \frac{\partial^2}{\partial \eta^2} \right).$$

Równanie (7.33) we współrzędnych parabolicznych przyjmuje postać

$$\left[- \frac{1}{\xi^2 + \eta^2} \left(\frac{\partial^2}{\partial \xi^2} + \frac{\partial^2}{\partial \eta^2} \right) - \frac{4}{\xi^2 + \eta^2} + \frac{\xi^2 - \eta^2}{2} f \right] \psi(\xi, \eta) = \varepsilon \psi(\xi, \eta), \quad (7.34)$$

³⁾ Jest to najprostszy model ekscytanu w studni kwantowej.

a jego rozwiązań poszukujemy w postaci iloczynu $\psi(\xi, \eta) = g_1(\xi)g_2(\eta)$. Po obustronnym pomnożeniu (7.34) przez $-(\xi^2 + \eta^2)/\psi$ otrzymujemy

$$\frac{1}{g_1} \frac{\partial^2 g_1}{\partial \xi^2} + \frac{1}{g_2} \frac{\partial^2 g_2}{\partial \eta^2} + 4 - \frac{\xi^4 - \eta^4}{2} f + \xi^2 \varepsilon + \eta^2 \varepsilon = 0. \quad (7.35)$$

Podobnie jak w przypadku trójwymiarowym, mamy tu wyrazy stałe oraz zależne tylko od ξ i tylko od η . Równanie (7.35) może być spełnione dla dowolnych (ξ, η) wtedy i tylko wtedy, gdy

$$\begin{cases} \frac{1}{g_1} \frac{\partial^2 g_1}{\partial \xi^2} + \beta_1 - \frac{\xi^4 f}{2} + \xi^2 \varepsilon = 0 \\ \frac{1}{g_2} \frac{\partial^2 g_2}{\partial \eta^2} + \beta_2 + \frac{\eta^4 f}{2} + \eta^2 \varepsilon = 0 \end{cases}$$

oraz

$$\beta_1 + \beta_2 = 4, \quad (7.36)$$

gdzie β_1, β_2 są parametrami rozdzielania zmiennych. Ostatecznie otrzymujemy układ dwóch sprzężonych równań różniczkowych

$$\begin{cases} \left[-\frac{d^2}{d\xi^2} + \left(-\xi^2 \varepsilon + \frac{\xi^4 f}{2} \right) \right] \chi_1(\xi) = \beta_1 \chi_1(\xi) \\ \left[-\frac{d^2}{d\eta^2} + \left(-\eta^2 \varepsilon - \frac{\eta^4 f}{2} \right) \right] \chi_2(\eta) = \beta_2 \chi_2(\eta). \end{cases} \quad (7.37)$$

Mają one postać równań Schrödingera, w których rolę potencjałów odgrywają funkcje

$$v_1(\xi) = -\varepsilon \xi^2 + \frac{1}{2} f \xi^4, \quad v_2(\eta) = -\varepsilon \eta^2 - \frac{1}{2} f \eta^4.$$

Zwróćmy uwagę, że w tym przypadku poszukiwanymi wartościami własnymi są stałe separacji β_1 i β_2 , natomiast energia wiązania atomu (ekscytanu) ε jest parametrem funkcji v_1 i v_2 . Dla danego natężenia pola elektrycznego f rozwiązanie układu równań (7.37) polega na znalezieniu takiej energii wiązania ε , dla której wartości własne obu równań spełnią związek (7.36). Podejście to zostało zastosowane w pracach [129].

Kwestia warunków brzegowych wymaga dodatkowego komentarza. Energia wiązania ε jest ujemna, więc potencjał $v_1(\xi)$ (rys. 7.2) jest zawsze jamopodobny (wiązący) i możemy położyć $g_1(-\xi_{\max}) = g_1(\xi_{\max}) = 0$. Potencjał $v_2(\eta)$ jest określony tylko na półprostej, gdyż $\eta \geq 0$, i nie ma żadnych podstaw do położenia $v_2(0) = 0$. Można jednak rozszerzyć dziedzinę zmiennej η na cały zbiór \mathbb{R} , zauważając, że parom $(-\xi, -\eta)$ i (ξ, η) odpowiada we wzorach (7.33) ta sama para (x, y) . W ten sposób „rozszerzona” funkcja $\psi(\xi, \eta)$ wykazuje własność symetrii

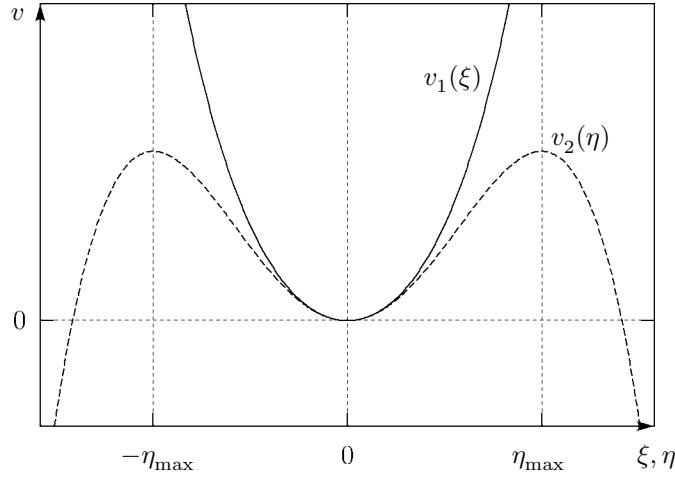
$$\psi(-\xi, -\eta) = \psi(\xi, \eta),$$

równoważną warunkom

$$g_1(-\xi) = g_1(\xi), \quad g_2(\eta) = g_2(-\eta)$$

albo

$$g_1(-\xi) = -g_1(\xi), \quad g_2(\eta) = -g_2(-\eta).$$



Rys. 7.2. Potencjały jednowymiarowe v_1, v_2 w zagadnieniu dwuwymiarowego atomu w zewnętrznym polu elektrycznym

Warunki te nie są trudne do spełnienia. Potencjały $v_1(\xi)$ i $v_2(\eta)$ są funkcjami parzystymi, więc $g_1(\xi)$ i $g_2(\eta)$ nie mogą być funkcjami dowolnymi, a tylko parzystymi bądź nieparzystymi. Obliczenia prowadzimy zatem na rozszerzonym przedziale $-\eta_{\max} < \eta < \eta_{\max}$, stawiając nieskończone bariery potencjału w punktach $\pm\eta_{\max}$.

7.6. Kropka kwantowa

Rozważania nad kropką kwantową z parabolicznym potencjałem wiążącym w zewnętrznym jednorodnym polu magnetycznym, zawierającą dwa elektrony, prowadzą do jednowymiarowego równania Schrödingera [130–132]

$$\left[-\frac{d^2}{dz^2} + v(z) \right] u(z) = \varepsilon u(z). \quad (7.38)$$

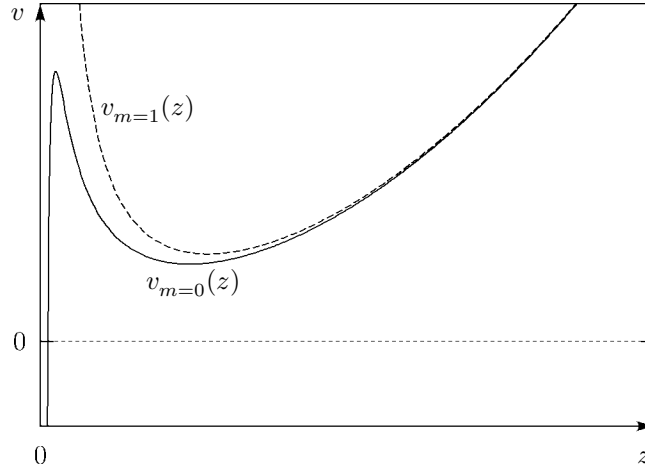
Potencjał w tym równaniu ma postać (patrz rys. 7.3)

$$v(z) = z^2 + \frac{2\lambda}{z} + \frac{m^2 - \frac{1}{4}}{z^2}, \quad (7.39)$$

gdzie $z = R/L_c$ jest bezwymiarową współrzędną radialną,

$$\lambda = \frac{L_c}{A}, \quad A = \frac{8\pi\epsilon_0\epsilon\hbar^2}{KM^*e^2}, \quad L_c = \sqrt{\frac{2\hbar}{M^*\Omega}},$$

Ω jest parametrem charakteryzującym potencjał wiążący, a M^* — masą efektywną elektronu. Ze względu na sposób wyprowadzenia równania (7.38), należy dodać warunek brzegowy $u(0) = 0$ dla funkcji falowej.



Rys. 7.3. Potencjał kropki $v(z)$ kwantowej przy $\lambda = 2$ dla liczby kwantowej $m = 0$ i $m = 1$

Kształt potencjału (7.39) zależy od liczby kwantowej $m = 0, \pm 1, \pm 2, \dots$. Przy $m = 0$ potencjał (dla parametru $\lambda > 0,62$) posiada, oprócz minimum globalnego $v(0) = -\infty$, minimum lokalne dla $z = z_{\min} > 0$. Z powodów fizycznych interesują nas stany związane zlokalizowane wokół $z = z_{\min}$, tj. stany o energiach $\varepsilon > v(z_{\min})$. Liczba tych stanów zależy od wartości parametru λ ; pierwszy stan pojawia się dla $\lambda \simeq 1,1$.

7.7. Dynamika elektronów w jednowymiarowych supersieciach w ramach modelu Kroniga–Penneya

Prostym i adekwatnym modelem używanym do opisu dynamiki elektronów w jednowymiarowych supersieciach jest model Kroniga–Penneya. W pracach [133–135] zbadano przewodnictwo elektryczne supersieci aperiodycznych. Przy założeniu $M_s = M_b$, potencjał w hamiltonianie przyjęto w postaci układu prostokątnych barier (rys. 7.4).

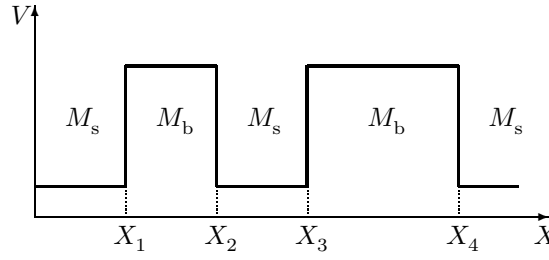
Hamiltonian układu ma w tym przypadku postać hamiltonianu modelu Kroniga–Penneya [133, 134]

$$\mathcal{H}(X) = -\frac{\hbar^2}{2M} \frac{d^2}{dX^2} + \sum_{i=1}^n V_i(X), \quad (7.40)$$

gdzie V_i jest potencjałem i -tej prostokątnej bariery; wartość potencjału w obszarze studni przyjęto równą zero. Stacjonarne równanie Schrödingera dla badanego układu ma formę

$$\mathcal{H}\Psi^{(j)} = \varepsilon_j \Psi^{(j)}, \quad (7.41)$$

gdzie \mathcal{H} zadany jest przez (7.40).



Rys. 7.4. Jednowymiarowy potencjał w modelu Kroniga–Penneya

Wartości własne \mathcal{E}_j i funkcje własne $\Psi^{(j)}$ dla zagadnienia (7.41) można wyznaczać odpowiednio za pomocą metody bisekcji Martina–Deana i metody DWSZ.

W pracy [135] przeanalizowano jednowymiarową supersieć półprzewodnikową, w której masa efektywna nośników prądu zależała od położenia. Założono, że masa nośników prądu pod barierą jest równa M_b , a w obszarach studni kwantowych M_s (rys. 7.4). Hamiltonian ma wtedy postać

$$\mathcal{H}(X) = \sum_{i=1}^n \left[-\theta(X - A_i)\theta(A_{i+1} - X) \frac{\hbar^2}{2M_i} \frac{d^2}{dX^2} + V_i(X) \right], \quad (7.42)$$

gdzie θ jest funkcją Heaviside’a.

Zagadnienie własne z hamiltonianem (7.42) sprowadzono do AZW za pomocą schematu dyskretyzacji opisanego w rozdziale 4. Wartości własne i wektory własne ponownie możemy wyznaczać za pomocą metody bisekcji Martina–Deana i metody DWSZ.

7.8. Układy niskowymiarowe w niejednorodnym polu magnetycznym

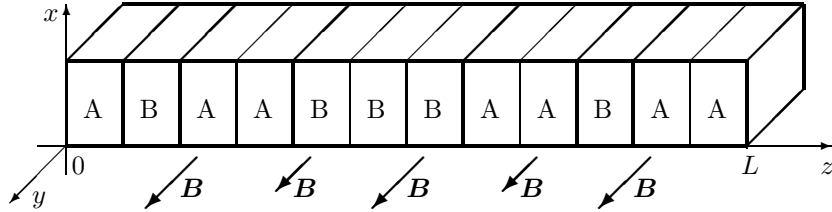
W ostatnim czasie prowadzone są badania transportu nośników ładunku elektrycznego w układach niskowymiarowych poddanych działaniu niejednorodnego pola magnetycznego [136–138]. W tym rozdziale przedstawimy krótko dwa szczególne przypadki, którymi są: supersieć w prostopadłym do jej *kierunku wzrostu*⁴⁾ niejednorodnym polu magnetycznym oraz dwuwymiarowy gaz nieoddziałujących elektronów poddany działaniu prostopadłego niejednorodnego pola magnetycznego.

⁴⁾ Mianem tym określamy kierunek, wzdłuż którego kolejne warstwy struktury układane są w stosownym procesie technologicznym.

7.8.1. Supersieć półprzewodnikowa w polu magnetycznym

Rozważmy jednowymiarową supersieć umieszczoną w polu magnetycznym prostopadłym do jej kierunku wzrostu OZ (tzw. *konfiguracja Voigta*, rys. 7.5). Zajmiemy się przypadkiem, gdy pole jest jednorodne bądź modulowane tylko wzdłuż osi OZ :

$$\mathbf{B} = [0, B(Z), 0].$$



Rys. 7.5. Schemat supersieci w polu magnetycznym

Wybierając cechowanie Landaua, otrzymujemy potencjał wektorowy postaci

$$\mathbf{A} = [A(Z), 0, 0],$$

gdzie, przypomnijmy, $\mathbf{B} = \nabla \times \mathbf{A}$. Jeżeli ruch elektronu w kierunku OX jest swobodny, tj. opisany falą płaską o wektorze falowym K , to jego funkcja falowa ma postać iloczynu

$$\Phi(X, Z) = e^{iKX} \Psi(Z).$$

Po prostych przekształceniach dostajemy kwazijednowymiarowe równanie Schrödingera

$$\left[-\frac{\hbar^2}{2} \frac{d}{dZ} \frac{1}{M(Z)} \frac{d}{dZ} + \frac{[\hbar K + \frac{e}{c} A(Z)]^2}{M(Z)} + V(Z) + g(Z) \mu_B B(Z) \sigma_Z \right] \Psi(Z) = \mathcal{E} \Psi(Z)$$

z dodatkowym składnikiem potencjału pochodzącym od pola magnetycznego, sparametryzowanym przez K :

$$V_m(K; Z) = \frac{[\hbar K + \frac{e}{c} A(Z)]^2}{M(Z)}.$$

Ponieważ kierunek wektora \mathbf{B} jest stały w przestrzeni, powyższe równanie ze spinem rozpada się na dwa równania postaci

$$\left[-\frac{\hbar^2}{2} \frac{d}{dZ} \frac{1}{M(Z)} \frac{d}{dZ} + V_m(K; Z) + V(Z) \pm g(Z) \mu_B B(Z) \right] \Psi_{\pm}(Z) = \mathcal{E}_{\pm} \Psi_{\pm}(Z) \quad (7.43)$$

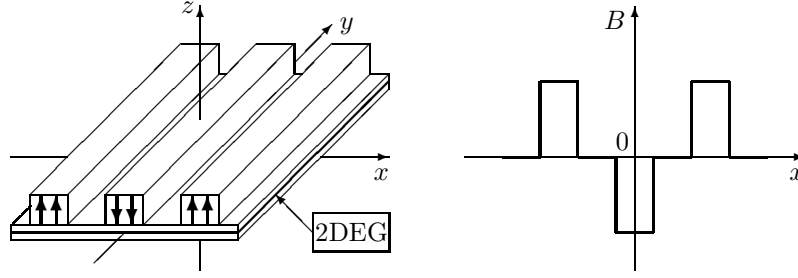
dla dwóch składowych spinowych funkcji falowej $\Psi = (\Psi_+, \Psi_-)^T$.

7.8.2. Dwuwymiarowy gaz elektronowy w niejednorodnym polu magnetycznym

Zajmijmy się teraz heterostrukturą półprzewodnikową zawierającą warstwę dwuwymiarowego gazu elektronowego (w płaszczyźnie XOY), umieszczoną w polu magnetycznym równoległym do kierunku wzrostu OZ (tzw. konfiguracja Faradaya) i modulowanym wzdłuż osi OX :

$$\mathbf{B} = [0, 0, B(X)].$$

Układ taki można otrzymać np. nanosząc na strukturę półprzewodnikową szereg paszków z materiału ferromagnetycznego lub nadprzewodzącego (rys. 7.6).



Rys. 7.6. Konfiguracja jednowymiarowo modulowanego pola magnetycznego. Dwuwymiarowy gaz elektronowy (2DEG) znajduje się w warstwie zaznaczonej grubszą linią

Wybierając cechowanie Landaua otrzymujemy potencjał wektorowy postaci

$$\mathbf{A} = [0, A(X), 0].$$

Jeżeli ruch elektronu w kierunku OY jest swobodny opisany falą płaską o wektorze falowym K , tj.

$$\Phi(X, Y) = e^{iKY} \Psi(X),$$

wówczas — podobnie jak poprzednio — otrzymujemy kwazijednowymiarowe równanie Schrödingera

$$\left[-\frac{\hbar^2}{2M} \frac{d^2}{dX^2} + \frac{[\hbar K + \frac{e}{c} A(X)]^2}{M} + g\mu_B B(X) \sigma_Z \right] \Psi(X) = \mathcal{E} \Psi(X).$$

Tym razem występuje tylko potencjał pochodzący od pola magnetycznego

$$V_m(K; X) = \frac{1}{M} \left[\hbar K + \frac{e}{c} A(X) \right]^2,$$

a masa efektywna nie zależy od położenia. Potencjał ten może tworzyć układ barier o złożonym kształcie.

Równanie ze spinem rozpada się na dwa równania dla składowych spinowych funkcji falowej Ψ , które szczególnie prostą postać

$$\left[-\frac{d^2}{dx^2} + 2v_m(k; x) \pm gb(x) \right] \psi_{\pm}(x) = 2\varepsilon_{\pm} \psi_{\pm}(x) \quad (7.44)$$

przyjmują w „naturalnym” dla pola magnetycznego układzie jednostek, opartym na częstości cyklotronowej $\Omega_c = |e\langle B \rangle|/(Mc)$, energii cyklotronowej $\hbar\Omega_c$ oraz długości magnetycznej $l_B = \sqrt{\hbar c/|e\langle B \rangle|}$; $\langle B \rangle$ jest charakterystyczną (np. średnią) wartością indukcji magnetycznej.

Jak widzimy, równania (7.43) i (7.44) można analizować numerycznie metodami algebraicznymi w celu otrzymania poziomów energetycznych w rozważanych układach oraz zbadania własności magnetotransportu elektronowego (przewodnictwo Landauera [112, 133–135, 139, 140], tunelowanie).

Zadania

7/1. W oparciu o klasyczne równania Hamiltona wyprowadzić równanie (7.2)

7/2. Wyprowadzić równanie (7.2) w przypadku kwantowym.

7/3. Dla okresowego łańcucha n identycznych atomów wyprowadzić relację dyspersji

$$\Omega(Q) = 2\sqrt{\frac{K}{M}} \left| \sin \frac{QA}{2} \right|,$$

gdzie $Q \in \langle -\pi/A, \pi/A \rangle$ jest wektorem falowym z pierwszej strefy Brillouina, A — odległością międzyatomową (rozmiarem komórki elementarnej), K — stałą sprężystą, M — masą atomów. Wskazówka: w równaniu (7.3) przyjąć $Z_l \sim e^{iQlA}$.

7/4. Wyprowadzić równania (7.4). Wskazówka: podzielić i -te równanie przez m_i .

7/5. Wyprowadzić wzory (7.6) korzystając z twierdzenia Martina–Deana.

7/6. Opierając się na twierdzeniu Martina–Deana, wyprowadzić algorytm (analogiczny do (7.6)) wyznaczania wartości własnych macierzy dynamicznej dla układu jednowymiarowego, w którym uwzględniamy oddziaływania najbliższych i drugich z kolei sąsiadów.

7/7. Pokazać, że funkcja gęstości stanów fononowych $g(\omega)$ jednowymiarowego kryształu ma postać

$$g(\omega) = \frac{4n}{\pi} \frac{1}{\sqrt{(2-\omega)(2+\omega)}}.$$

Jakiego typu osobliwości Van Hove’a wykazuje funkcja $g(\omega)$?

7/8. Obliczyć funkcje rozkładu (dystrybuanty) stanów fononowych jednowymiarowego kryształu [98]

$$\bar{g}(\omega^2) = \int_0^{\omega^2} g(z^2) dz^2, \quad \bar{g}(\omega) = \int_0^{\omega} g(z) dz,$$

gdzie $g(\omega^2)$ dane jest wzorem (7.7), a $g(\omega)$ jest podane w zadaniu poprzednim.

7/9. Wyznaczyć numerycznie wektory własne odpowiadające 5. i 25. wartości własnej monoatomowego okresowego łańcucha atomów. Przyjąć swobodne warunki brzegowe i $n = 10^3$. Jaki jest charakter tych wektorów własnych (tj. zdelokalizowany czy zlokalizowany)?

7/10. Wyznaczyć numerycznie unormowaną funkcję rozkładu $\bar{g}(\omega^2) = \eta(\mathbf{T} - \omega^2 \mathbf{1})/n$ częstości własnych monoatomowego okresowego łańcucha $n = 2 \cdot 10^3$ atomów. Wyznaczyć histogramy funkcji $\bar{g}(\omega^2)$ dla różnych szerokości podprzedziałów.

7/11. Wyznaczyć numerycznie wartości własne macierzy dynamicznej (7.5) dla łańcucha z zadania 7/10, przyjmując cykliczne warunki brzegowe. Obliczenia przeprowadzić dla $q = 0$ i $q = \pi$. Otrzymane wyniki porównać z wyrażeniami analitycznymi (patrz zadanie 7/3).

7/12. Wyznaczyć numerycznie unormowaną funkcję rozkładu $\bar{g}(\omega) = \eta(\mathbf{T} - \omega \mathbf{1})/n$ częstości własnych monoatomowego periodycznego łańcucha $n = 2 \cdot 10^3$ atomów. Wyznaczyć histogram funkcji $\bar{g}(\omega)$ dla różnych szerokości podprzedziałów.

7/13. Wyznaczyć numerycznie widmo energetyczne łańcucha $n = 10^3$ atomów, w którym znajduje się lekki lub ciężki atom domieszki o masie m_d . Przyjąć: (a) $m_d = 0,1$; (b) $m_d = 10$. Pozostałe masy $m_i = 1$. Założyć swobodne warunki brzegowe. Jaki jest tym razem charakter wektorów własnych?

7/14. Wyznaczyć numerycznie widmo energetyczne łańcucha $n = 100$ atomów, w którym stałe sprężyste są ustalone, a masy atomów zadaje rozkład jednostajny na przedziale $\langle \frac{1}{2}, 5 \rangle$. Założyć swobodne warunki brzegowe. Jaki jest charakter wektorów własnych (zlokalizowany, zdelokalizowany)?

7/15. Wyznaczyć numerycznie widmo energetyczne łańcucha $n = 100$ atomów, w którym stałe sprężyste zadaje rozkład jednostajny na przedziale $\langle 0,3; 2 \rangle$, a masy atomów — rozkład jednostajny na przedziale $\langle 0,1; 2 \rangle$. Założyć swobodne warunki brzegowe. Jaki jest charakter wektorów własnych (zlokalizowany, zdelokalizowany)?

7/16. Wyznaczyć numerycznie widmo drgań monoatomowego łańcucha kwaziperiodycznego, gdy $n = 987$ (rząd $l = 15$) atomów. Przyjąć swobodne warunki brzegowe i parametr kwaziperiodyczności $\xi = 1$.

7/17. Wyznaczyć numerycznie widmo drgań binarnego kwaziperiodycznego łańcucha $n = 1597$ (rząd $l = 16$) atomów. Przyjąć swobodne warunki brzegowe i parametr kwaziperiodyczności $\xi = 2$.

7/18. Wyznaczyć numerycznie unormowaną funkcję rozkładu $\bar{g}(\omega^2) = \eta(\mathbf{T} - \omega^2 \mathbf{1})/n$ częstości własnych dla układu z zadania 7/16. Wyznaczyć histogramy funkcji $\bar{g}(\omega^2)$ dla różnych szerokości podprzedziałów.

7/19. Wyznaczyć numerycznie unormowaną funkcję rozkładu $\bar{g}(\omega^2) = \eta(\mathbf{T} - \omega^2 \mathbf{1})/n$ częstości własnych dla układu z zadania 7/17. Wyznaczyć histogramy funkcji $\bar{g}(\omega^2)$ dla różnych szerokości podprzedziałów.

7/20. Wyznaczyć zależność pojemności cieplnej $C_V(\Theta) = d\mathcal{E}/d\Theta$ od temperatury dla periodycznego łańcucha atomów z zadania 7/10.

7/21. Wyznaczyć poziomy energetyczne elektronów w periodycznym łańcuchu 900 atomów. Obliczyć funkcję rozkładu wartości własnych i odpowiedni histogram (odpowiadający gęstości stanów elektronowych).

7/22. Wyznaczyć 7., 474. i 847. wektor własny w łańcuchu z zadania 7/21.

7/23. Wyznaczyć numerycznie elektronowe widmo energetyczne dla łańcucha 200 atomów, w którym pozadiagonalne wyrazy macierzy (7.15) są ustalone, a wyrazy diagonalne zadaje rozkład jednostajny na przedziale $\langle \frac{1}{2}, \frac{3}{2} \rangle$.

7/24. Wyznaczyć numerycznie elektronowe widmo energetyczne łańcucha $n = 300$ atomów, w którym wyrazy diagonalne macierzy (7.15) zadaje rozkład jednostajny na przedziale $\langle 0, 7; 1, 3 \rangle$, a wyrazy pozadiagonalne są ustalone. Jaki jest charakter wektorów własnych (zdelokalizowany czy zlokalizowany)?

7/25. Wyprowadzić reguły komutacji (7.17) korzystając z (7.16).

7/26. Wyprowadzić związek (7.18).

7/27. Wyznaczyć numerycznie widmo magnonowe dla periodycznego łańcucha 500 spinów. Obliczyć unormowaną funkcję rozkładu $\bar{g}(z) = \eta(\mathbf{T} - z\mathbf{1})/n$ energii własnych magnonów, gdzie $z = \hbar\Omega/(k_B\Theta)$. Wyznaczyć histogramy funkcji $\bar{g}(z)$ dla różnych szerokości podprzedziałów.

7/28. Wyznaczyć numerycznie widmo magnonowe dla periodycznego łańcucha 600 spinów, w którym cała wymiany między przyjmuje wartości $J_{i,i+1} = (-1)^i J_0$. Obliczyć unormowaną funkcję rozkładu $\bar{g}(z) = \eta(\mathbf{T} - z\mathbf{1})/n$ energii własnych magnonów, gdzie $z = \hbar\Omega/(k_B\Theta)$. Wyznaczyć histogramy funkcji $\bar{g}(z)$ dla różnych szerokości podprzedziałów.

7/29. Wyznaczyć numerycznie widmo magnonowe łańcucha $n = 10^3$ spinów, w którym są losowo umieszczone dwa atomy domieszek o spinach s_d . Przyjąć $s_d = 3$; pozostałe spiny $s_i = 1$. Założyć swobodne warunki brzegowe.

7/30. Wyznaczyć numerycznie widmo magnonowe monoatomowego kwaziperiodycznego łańcucha $n = 987$ (rzęd $l = 15$) spinów. Założyć zależność całki wymiany pomiędzy najbliższymi sąsiadami postaci $J_{i,i+1} = J_0[1 + \xi(1 - d_{i,i+1})]$, gdzie $d_{i,i+1} = x_{i+1} - x_i$ (patrz wzór (7.8)). Przyjąć swobodne warunki brzegowe i parametr kwaziperiodyczności $\xi = 2$. Wyznaczyć numerycznie unormowaną funkcję rozkładu $\bar{g}(z) = \eta(\mathbf{T} - z\mathbf{1})/n$ energii własnych, gdzie $z = \hbar\Omega/(k_B\Theta)$. Wyznaczyć histogramy funkcji $\bar{g}(z)$ dla różnych szerokości podprzedziałów.

7/31. Wyznaczyć numerycznie widmo magnonowe łańcucha $n = 100$ spinów, w którym całki wymiany są zadane rozkładem jednostajnym na przedziale $\langle \frac{1}{2}J_0, \frac{3}{2}J_0 \rangle$. Założyć swobodne warunki brzegowe. Jaki jest charakter wektorów własnych (zdelokalizowany czy zlokalizowany)?

7/32. Wyznaczyć zależność pojemności cieplnej $C_V(\Theta) = d\mathcal{E}/d\Theta$ od temperatury dla periodycznego łańcucha spinów z zadania 7/27.

7/33. Wyprowadzić równanie radialne (7.24).

7/34. Rozwiązać numerycznie równanie (7.24) za pomocą kilku wybranych metod i porównać rezultaty z rozwiązaniem analitycznym (7.25).

7/35. W oparciu o algorytm Martina–Deana napisać program i wyznaczyć za jego pomocą osiem najniższych wartości własnych elektronu w atomie wodoru. Zbadać, jak na wyniki numeryczne wpływa szerokość przedziału, liczba n punktów siatki oraz liczba kwantowa l . Wyniki numeryczne porównać z rezultatami analitycznymi.

7/36. Wyprowadzić bezwymiarowe równanie Schrödingera (7.27) dla atomu wodoru w polu elektrycznym.

7/37. Wyprowadzić układ zagadnień własnych (7.31).

7/38. Wyznaczyć dla $m = 0$ i $\lambda = 2$ wszystkie stany związane cząstki opisanej równaniem (7.38) leżące powyżej minimum lokalnego potencjału (7.39) przy warunku brzegowym $u(0) = 0$.

7/39. Wyprowadzić równania (7.43) i (7.44).

Pakiet MARRS

Program MARRS — *Metody Algebraiczne Rozwiązywania Równania Schrödingera* został napisany w języku Delphi. Stanowi uniwersalne narzędzie pozwalające użytkownikowi w prosty sposób rozwiązywać stacjonarne równanie Schrödingera dla dowolnego kształtu jednowymiarowych studni potencjału. Jest przyjaznym użytkownikowi instrumentem badawczo-dydaktycznym w dziedzinie mechaniki kwantowej i fizyki struktur niskowymiarowych, przydatnym zarówno dla pracowników naukowych, jak i studentów. Pierwsi mogą przy pomocy programu MARRS wyznaczyć wstępne rozwiązania wielu problemów, oszacować parametry doświadczalne, ocenić wyniki eksperymentu, czy też sprawdzić poprawność rozwiązań analitycznych. Studenci w prosty i pobudzający do myślenia sposób mają szansę sprawdzić swą wiedzę z dziedziny mechaniki kwantowej, poznać wydajne metody rozwiązywania równania Schrödingera i przeprowadzić samodzielnie proste eksperymenty.

Pakiet MARRS zapisany jest na dołączonej do książki płytce CD oraz dostępny w Internecie pod adresem <http://www.if.pwr.wroc.pl/~ssalejda/programy/>.

Szczegółowe informacje o programie zawarte są w pracach [141] oraz w dołączonych do programu plikach pomocy.

8.1. Charakterystyka programu

Omówimy krótko najważniejsze możliwości pakietu programów MARRS i przedstawimy jego zalety.

Graficzny interfejs użytkownika. Program wykorzystuje w pełni możliwości systemów Windows 95/NT i zgodnych. Posługiwanie się programem jest proste dzięki zgodnemu z intuicją rozmieszczeniu na ekranie elementów sterujących (przycisków, suwaków, menu itp.) Natychmiastowa prezentacja graficzna wyników na ekranie w sposób znaczący przyspiesza i ułatwia pracę. Stała kontrola wizualna postępu obliczeń (m.in. wyświetlanie częściowych rezultatów) z możliwością przerwania ich w dowolnej chwili bez utraty dotychczasowych wyników stwarza komfortowe warunki pracy z programem. Interfejs użytkownika opisujemy szczegółowo dalej.

Integracja z systemem operacyjnym. Program spełnia funkcję *serwera OLE*, dzięki czemu udostępnia swoje możliwości obliczeniowe wszystkim innym programom w środowisku Windows. Wielowątkowość i obsługa schowka systemowego pozwala na prostą wymianę danych z pozostałymi aplikacjami.

Współpraca z aplikacjami Microsoft Office. Użytkownik może łatwo wykorzystywać otrzymane wyniki we własnych opracowaniach i publikacjach bez potrzeby dokonywania żmudnej konwersji danych.

Bogaty zestaw procedur obliczeniowych. Zaimplementowano metody: trójpunktową, pięciopunktową, Lindberga, Guardioli–Rosa i ich zoptymalizowane odmiany, co zapewnia dużą szybkość obliczeń i pozwala wyznaczać wartości własne z określoną przez użytkownika dokładnością oraz porównywać efektywności różnych algorytmów. Do wyznaczania wektorów własnych zastosowano metodę DWSZ.

Różne sposoby określania potencjału. Użytkownik może zadać potencjał:

- wzorem;
- rysując wykres na ekranie za pomocą myszki;
- importując tabelę wartości z programu Microsoft Excel 97.

Funkcje porównywania. Program pozwala na szybkie badanie wpływu zmian kształtu potencjału na widmo wartości własnych. Proste przełączanie między arkuszami roboczymi daje możliwość wizualnej oceny zmiany położenia poziomów energetycznych i kształtu funkcji falowych.

Pliki pomocy. Podstawowe zasady posługiwania się programem, elementarne wiadomości na temat problematyki fizycznej i szczegółowy opis zastosowanych algorytmów numerycznych dostępne są w języku polskim.

Zapis i odtwarzanie stanu pracy. Pracę z programem można przerwać w dowolnym momencie, zapisując całość arkusza roboczego bądź wybrane wyniki obliczeń i parametry, a następnie odtworzyć dane i kontynuować obliczenia.

Funkcje testowe. Dołączone narzędzia, w tym arkusz w formacie MS Excela zawierający specjalne makrodefinicje, pozwalają ocenić efektywność każdej z metod dla różnych typów potencjałów.

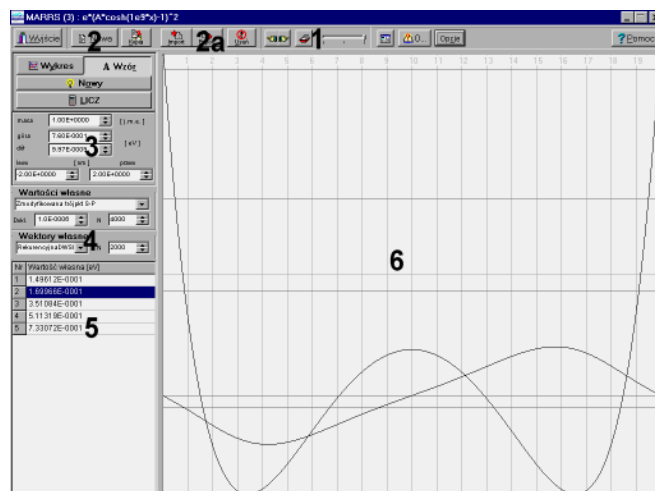
8.2. Graficzny interfejs użytkownika

Praca z programem oparta jest na tzw. *kartach roboczych*. Każda karta zawiera komplet informacji opisujących dany problem fizyczny i niezbędnych do wykonania obliczeń.

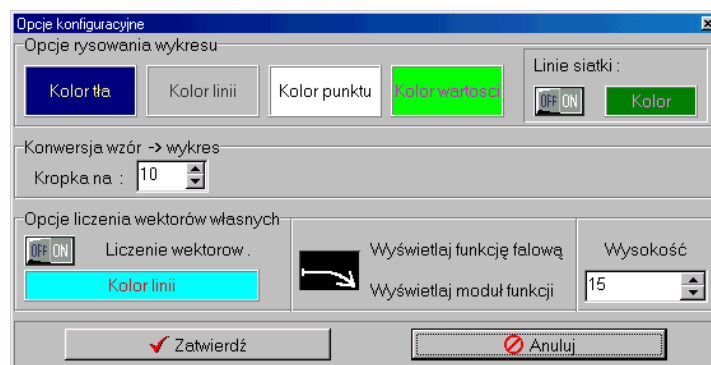
W głównym oknie programu można wydzielić obszary oznaczone na rys. 8.1 cyframi 1–6. U góry znajduje się poziomy pasek menu, pozwalający na zarządzanie kartami roboczymi, tj. nawigację (1), tworzenie nowej karty (2), odczyt lub zapis aktywnej karty roboczej lub potencjału (2a).

Poniżej paska menu wyświetlana jest bieżąca karta robocza, podzielona na dwie części: kontrolki (3–5) służące do określania parametrów obliczeń po lewej i obszar graficznej prezentacji wyników (6).

Opcje związane z pracą programu i sposobem wyświetlania wyników (m.in. sposób wyświetlania funkcji falowej — wartość lub kwadrat modułu, kolorystyka ekranu; rys. 8.2) dostępne są po wciśnięciu odpowiedniego przycisku z górnego paska menu.



Rys. 8.1. Okno główne programu MARRS



Rys. 8.2. Okno opcji programu MARRS

Obszary 3 i 4 są przedstawione w powiększeniu na rysunkach 8.4 i 8.5.

W obszarze 3 użytkownik ustala

- masę cząstki;
- przedział energii, w którym program wyznacza wartości własne i który jest przedstawiany w obszarze graficznym 6 (na pionowej osi);
- granice przedziału całkowania (i zarazem przedziału pokazywanego w obszarze graficznym na osi poziomej).

masa	1.00E+0000	[j.m.e.]	masa cząstki zakres energii
górną	4.69E+0000	[eV]	
dolną	-1.13E+0001		
lewo	-2.00E+0000	[nm]	rozmiary układu
prawo	2.00E+0000		

Rys. 8.3. Parametry zagadnienia

W obszarze 4 użytkownik

- dokonuje wyboru metody obliczania wartości własnych spośród dziesięciu zaimplementowanych;
- określa liczbę punktów siatki w przedziale całkowania (*stopień podziału*), oddzielnie dla wyznaczania wartości własnych i wektorów własnych;
- podaje dokładność obliczeń.

Wartości własne			wybór metody obliczania wartości wł. stopień podziału przedziału całkowania dokładność obliczeń
Zmodyfikowana trójpkt S-P			
Dokł.	1.0E-0006	N 4000	
Wektory własne			wybór metody obliczania wekt. wł. stopień podziału przedziału całkowania
Rekurencyjna DWSI			
		N 2000	

Rys. 8.4. Wybór metody obliczeniowej i określenie jej parametrów

W obszarze 5 pokazane są wyznaczone wartości własne. Wskazanie jednej z nich (przez kliknięcie myszką) powoduje obliczenie odpowiadającej jej funkcji falowej bądź kwadratu jej modułu i wyświetlenie go w obszarze 6, w którym pokazany jest także wykres potencjału oraz poziome linie reprezentujące wyznaczone poziomy energetyczne.

8.3. Schemat wykonywania obliczeń

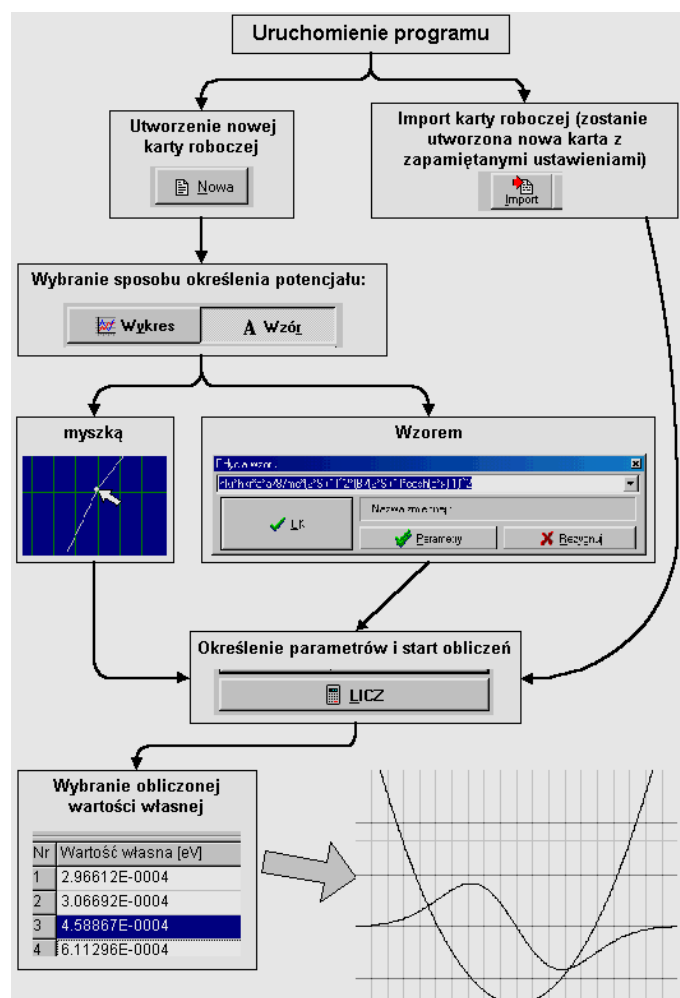
Aby przeprowadzić obliczenia dla danego zagadnienia fizycznego, należy wykonać następujące kroki, przedstawione schematycznie na rys. 8.5:

1. Wprowadzić postać potencjału jednym z dostępnych sposobów.
2. Ustalić parametry dla badanego zagadnienia (masę cząstki, przedział całkowania).
3. Wyznaczyć energie z odpowiedniego przedziału za pomocą wybranej metody obliczeniowej.
4. Wyznaczyć odpowiadające im funkcje falowe.

8.4. Dobór algorytmu

Użytkownik programu może wyznaczyć dokładność zaimplementowanych w pakiecie algorytmów dla zmieniającej się w przedziale od $n = 10$ do $n = 64\,000$ liczby punktów siatki. Obliczenia przeprowadzone są za pomocą arkusza kalkulacyjnego Microsoft Excel (procedury obliczeniowe programu MARRS wywoływane są za pomocą makrodefinicji Visual Basica). Wyniki testów są automatycznie przedstawiane w postaci czytelnych tabel i wykresów bez dodatkowej ingerencji użytkownika. Przykładowe rezultaty zamieszczone są w Dodatku pracy [141].

Otrzymane w ten sposób wyniki stanowią użyteczne źródło informacji na temat optymalnego wykorzystania procedur numerycznych pakietu MARRS. Umożliwiają porównanie zachowania się poszczególnych metod dla podstawowych klas potencjałów (porównaj także rozdział 7). Ułatwiają również dobór do konkretnego zagadnienia optymalnej metody obliczeniowej i jej parametrów (liczba punktów siatki, dokładność obliczeń), zapewniając uzyskanie wymaganej dokładności przy względnie krótkim czasie obliczeń.



Rys. 8.5. Schemat rozwiązywania stacjonarnego jednowymiarowego równania Schrödingera za pomocą programu MARRS

Dokładność i szybkość algorytmów

Ten rozdział poświęcamy przedstawieniu wyników otrzymanych za pomocą opisanych dotychczas algorytmów dla wybranych potencjałów reprezentujących kilka charakterystycznych klas. Są to potencjały wiążące, posiadające pojedyncze lub podwójne minimum lokalne i wykazujące różne zachowania asymptotyczne na granicach przedziałów określoności. Wszystkie one posiadają rozwiązania analityczne, co umożliwia weryfikację poprawności wyników numerycznych.

Zaprezentujemy wyniki obliczeń poziomów energetycznych i funkcji falowych dla wybranych potencjałów. Pokażemy, że postać potencjału ma istotny wpływ na dokładność otrzymywanych rezultatów. Przedstawimy także zależności pomiędzy liczbą punktów siatki, osiąganą dokładnością i czasem trwania obliczeń. Na tej podstawie można określić zachowanie się poszczególnych metod dla różnych klas potencjałów, wyznaczyć optymalne (z punktu widzenia efektywności) wartości parametrów wejściowych procedur i wybrać najwłaściwszy algorytm do konkretnego zagadnienia fizycznego.

Przedstawione wyniki numeryczne otrzymane zostały za pomocą podprogramów, których kody źródłowe są zawarte w Dodatku F i na dołączonej płycie CD. Podobne wyniki [141] otrzymać można, używając narzędzi pakietu MARRS. Zachęcamy również Czytelników do samodzielnego zbadania tej problematyki.

9.1. Wybrane potencjały

Poniżej charakteryzujemy krótko potencjały (rys. 9.1), które wykorzystaliśmy do badania dokładności i efektywności procedur numerycznych. Ponieważ porównywanie wyników numerycznych i analitycznych jest dogodniejsze w postaci bezwymiarowej (analityczne wartości energii własnych są zwykle prostymi ułamkami), podajemy obie postaci: bezwymiarową i wymiarową (w programie MARRS dane wejściowe i wyniki są liczbami mianowanymi).

Oscylator harmoniczny. Równanie Schrödingera określone jest wzorem

$$\left[-\frac{\hbar^2}{2M} \frac{d^2}{dX^2} + \frac{M\Omega^2 X^2}{2} \right] \Psi(X) = \mathcal{E} \Psi(X),$$

a jego postać bezwymiarowa — wzorem

$$\left[-\frac{d^2}{dx^2} + x^2 \right] \psi(x) = \varepsilon \psi(x).$$

Energie kolejnych stanów własnych

$$\mathcal{E}_j = \hbar\Omega(j - \frac{1}{2}), \quad \varepsilon_j = 2j - 1 \quad (j = 1, 2, \dots)$$

są oddalone o stałą wartość, co czyni ten potencjał (rys. 9.1-a) idealnym do weryfikacji wyników — można jej dokonać wręcz naocznie.

Potencjał Konwenta. Najprostszy przykład symetrycznego potencjału podwójnej studni, czyli oscylator anharmoniczny typu $V(X) = AX^4 - BX^2$ jest nierozwiązalny analitycznie. Z tego względu użyliśmy potencjału Konwenta [67–70] (rys. 9.1-b), charakteryzującego się względnie prostą postacią i posiadającego rozwiązania analityczne dla pewnych szczególnych wartości parametrów. Równanie Schrödingera z tym potencjałem ma postać

$$\left[-\frac{\hbar^2}{2M} \frac{d^2}{dX^2} + \frac{\hbar^2 D^2}{8M} (2k+1)^2 \left(\frac{b}{2k+1} \cosh(DX) - 1 \right)^2 \right] \Psi(X) = \mathcal{E} \Psi(X)$$

lub, w zmiennych bezwymiarowych,

$$\left[-\frac{d^2}{dx^2} + \alpha \left(\frac{b}{2k+1} \cosh x - 1 \right)^2 \right] \psi(x) = \alpha \varepsilon \psi(x), \quad \alpha = \frac{1}{4}(2k+1)^2.$$

Dla $k = 1$ ($\alpha = \frac{9}{4}$) znane są rozwiązania analityczne [67–70]:

$$\begin{aligned} \mathcal{E}_1 &= \frac{\hbar^2 D^2}{8M} \left[b^2 + 7 - 4\sqrt{\frac{1}{4} + b^2} \right], & \varepsilon_1 &= \frac{1}{9} \left[b^2 + 7 - 4\sqrt{\frac{1}{4} + b^2} \right], \\ \mathcal{E}_2 &= \frac{\hbar^2 D^2}{8M} (b^2 + 5), & \varepsilon_2 &= \frac{1}{9} (b^2 + 5), \\ \mathcal{E}_3 &= \frac{\hbar^2 D^2}{8M} \left[b^2 + 7 + 4\sqrt{\frac{1}{4} + b^2} \right], & \varepsilon_3 &= \frac{1}{9} \left[b^2 + 7 + 4\sqrt{\frac{1}{4} + b^2} \right]. \end{aligned}$$

Do obliczeń przyjęliśmy $b = 0,03$.

Potencjał Morse’a. Potencjał ten (rys. 9.1-c) wybrany został ze względu na swoje zastosowania w chemii fizycznej do opisu widm cząsteczkowych. Odpowiednie równanie Schrödingera ma postać¹⁾

$$\left[-\frac{\hbar^2}{2M} \frac{d^2}{dX^2} + V_c \left(e^{-2CX} - 2e^{-CX} \right) \right] \Psi(X) = \mathcal{E} \Psi(X),$$

gdzie C , V_c — parametry, a jego postać bezwymiarowa to

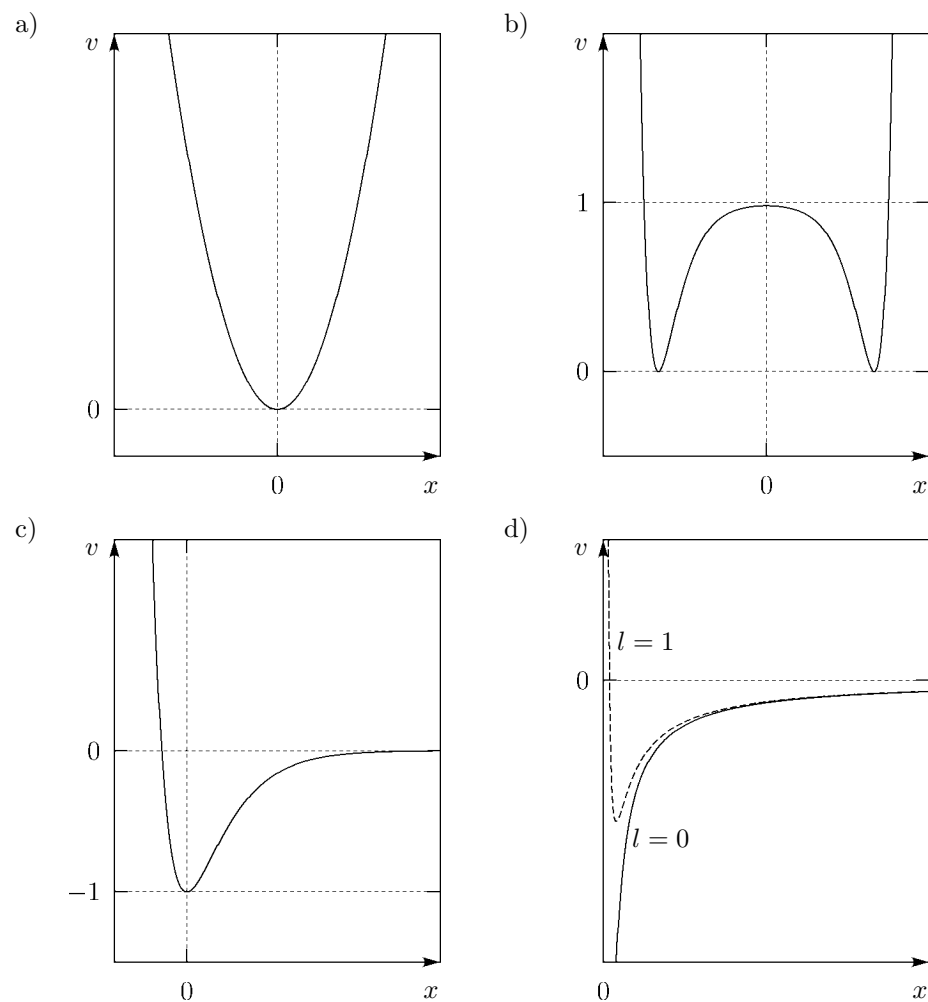
$$\left[-\frac{d^2}{dx^2} + \alpha(e^{-2x} - 2e^{-x}) \right] \psi(x) = \alpha \varepsilon \psi(x).$$

Dokładne wartości najniższych energii dla tych potencjałów to odpowiednio [30]

$$\mathcal{E}_j = -V_c \left[1 - \frac{j - \frac{1}{2}}{\sqrt{\alpha}} \right]^2, \quad \varepsilon_j = - \left[1 - \frac{j - \frac{1}{2}}{\sqrt{\alpha}} \right]^2 \quad (j = 1, \dots, j_{\max}),$$

gdzie $\alpha = 2MV_c/(\hbar C)^2$, $j_{\max} = \lceil \sqrt{\alpha} - \frac{1}{2} \rceil$.

¹⁾ W przykładowym arkuszu roboczym programu MARRS zapisana jest nieco inna postać potencjału Morse’a, przesunięta o stałą: $V(X) = V_c(1 - e^{-CX})^2$.



Rys. 9.1. Poglądowe wykresy potencjałów wybranych do testowania procedur numerycznych: (a) potencjał oscylatora harmonicznego; (b) potencjał Konwent; (c) potencjał Morse'a; (d) potencjał kulombowski dla dwóch wartości orbitalnej liczby kwantowej l

Potencjał kulombowski. Trójwymiarowe równanie Schrödingera — zagadnienie atomu wodoru — sprowadzamy, dzięki symetrii sferycznej, do postaci jednowymiarowej, korzystając z procedury opisanej w podrozdziale 7.4:

$$\left[-\frac{\hbar^2}{2M} \frac{d^2}{dX^2} + \frac{\hbar^2 l(l+1)}{2MX^2} - \frac{e^2}{4\pi\epsilon_0 X} \right] \Psi(X) = \mathcal{E} \Psi(X);$$

e oznacza ładunek elektronu, M — masę zredukowaną układu elektron–proton, zaś $l = 0, 1, 2, \dots$ jest orbitalną liczbę kwantową. W jednostkach atomowych (7.23) otrzymujemy

$$\left[-\frac{d^2}{dx^2} + \frac{l(l+1)}{x^2} - \frac{2}{x} \right] \psi(x) = \varepsilon \psi(x).$$

Rozwiązaniami są [33]

$$\mathcal{E}_j = -\frac{Me^4}{32\pi^2\epsilon_0^2\hbar^2(j+l)^2}, \quad \varepsilon_j = -\frac{1}{(j+l)^2} \quad (j = 1, 2, \dots).$$

Oddzielnie rozpatrujemy przypadki $l = 0$ i $l \neq 0$. Otrzymujemy w ten sposób potencjały należące do dwóch różnych klas (rys. 9.1-d) — dla $l = 0$ potencjał nie wykazuje skończonego minimum (osobliwość typu $-\frac{1}{X}$), dla $l \neq 0$ potencjał ma minimum lokalne i osobliwość typu $+\frac{1}{X^2}$. (Z uwagi na te osobliwości należy przypomnieć, że procedury numeryczne nie wyznaczają wartości potencjału w krańcowych punktach przedziału całkowania $\langle 0, r_{\max} \rangle$.)

9.2. Wybrane wyniki obliczeń numerycznych

W tabelach 9.1–9.5 zestawione są wyniki otrzymane dla wybranych potencjałów za pomocą programu PREC napisanego w języku FORTRAN-77 z wykorzystaniem podprogramów opisanych w Dodatku F; kod źródłowy programu PREC znajduje się na dołączonej płytce CD. Program został skompilowany kompilatorem GNU Fortran g77²⁾ w wersji dla środowiska DOS (opcje optymalizacji -O2) i był uruchomiany, jeśli nie zaznaczono inaczej, na komputerze klasy PC z mikroprocesorem AMD Athlon taktowanym zegarem o częstotliwości 1200 MHz.

Każda z tabel zawiera obliczone trzema metodami (trójpunktową, Lindberga i Guardioli–Rosa) energie trzech najniższych stanów własnych ($\varepsilon_1, \varepsilon_2, \varepsilon_3$) dla różnej liczby punktów siatki n , a także ich wartości dokładne, dane formułami analitycznymi. Liczba punktów siatki zmienia się od $2^6 - 1 = 63$ do $2^{15} - 1 = 32767$ dla metody trójpunktowej i metody Lindberga, oraz od $2^4 - 1 = 15$ do $2^8 - 1 = 255$ dla metody Guardioli–Rosa (GR). Zróżnicowanie to jest spowodowane znacznie większą złożonością obliczeniową metody GR ($O(n^3)$ operacji zmiennoprzecinkowych plus $O(n)$ na każdą wartość własną) i jej znacznie większym zapotrzebowaniem na pamięć ($O(n^2)$ komórek); dla metod trójpunktowej i Lindberga mamy odpowiednio $O(n)$ operacji na każdą wartość własną i $O(n)$ komórek pamięci.

²⁾ Dostępnym nieodpłatnie (patrz dodatek F).

Tabele zawierają dodatkowo dokładności obliczonych energii. Jako miarę dokładności obliczeń j -tej energii własnej przyjęliśmy niepewność względną

$$r_j = \left| \frac{\varepsilon_j^{\text{num}} - \varepsilon_j}{\varepsilon_j} \right|, \quad (9.1)$$

gdzie $\varepsilon_j^{\text{num}}$ oznacza wartość własną wyznaczoną numerycznie, a ε_j — jej wartość dokładną (analityczną). Niepewność względną charakteryzuje zamiennie inny parametr — *liczba cyfr znaczących*

$$l_j = -\log_{10} r_j,$$

którego wartości są zamieszczone w tabelach 9.1–9.5.

Program PREC umożliwia dodatkowo pomiar czasu obliczeń. Zauważmy, że algorytmy bisekcji (i niektóre inne algorytmy znajdowania wartości własnych macierzy) posiadają parametr wejściowy, określony jako *zadana dokładność* (np. σ w algorytmie 3-1). Zadanie zbyt małej wartości σ powoduje niepotrzebne wydłużenie obliczeń (reprezentacja macierzowa (2.9) równania Schrödingera jest przybliżona, więc dokładniejsze wyznaczenie wartości własnych macierzy $\hat{\mathbf{H}}$ nie musi oznaczać poprawy dokładności obliczanych wartości własnych \mathcal{E}_j hamiltonianu \mathcal{H}). Aby uwolnić się od tego problemu i znaleźć rzeczywiste, minimalne czasy obliczeń, każdy z algorytmów badamy w dwóch krokach:

1. Wyznaczamy maksymalną dokładność $r = \frac{1}{3}(r_1 + r_2 + r_3)$ osiąganą dla danej liczby punktów siatki n poprzez wykonanie obliczeń dla wartości parametru *zadana dokładność* rzędu epsilon maszynowego.
2. Mierzmy czas trwania obliczeń t , zadając dokładność uzyskaną w kroku pierwszym.

Zmierzone czasy t oraz dokładności, reprezentowane przez średnią liczbę cyfr znaczących

$$l = \frac{1}{3}(l_1 + l_2 + l_3)$$

przedstawione są na wykresach $l(n)$ i $l(t)$ (rys. 9.2–9.8) i zostaną dokładniej omówione w kolejnych podrozdziałach.

Tabela 9.1. Wyniki obliczeń dla potencjału oscylatora harmonicznego (przedział całkowania $\langle a, b \rangle = \langle -7, 7 \rangle$)

n	Energie			Cyfry znaczące		
	ε_1	ε_2	ε_3	l_1	l_2	l_3
Metoda trójpunktowa						
255	0,99981304487523	2,99906508442331	4,99756881334243	3,7	3,5	3,3
1023	0,99998831735314	2,99994158621972	4,99984812258788	4,9	4,7	4,5
4095	0,99999926984257	2,99999634921071	4,99999050794166	6,1	5,9	5,7
16383	0,99999995436512	2,99999977182590	4,99999940674745	7,3	7,1	6,9
65535	0,99999999714782	2,99999998573912	4,99999996292172	8,5	8,3	8,1
Metoda Lindberga						
255	0,99999993009267	2,99999951059048	4,99999825185098	7,2	6,8	6,5
1023	0,9999999972703	2,9999999808920	4,99999999317566	9,6	9,2	8,9
4095	0,9999999999894	2,9999999999254	4,9999999997335	12,0	11,6	11,3
16383	0,9999999999997	3,00000000000003	4,9999999999985	13,5	14,0	13,5
65535	0,9999999999952	3,00000000000048	4,9999999999952	12,3	12,8	13,0
Metoda Guardioli-Rosa (aproksymanta $[3/2]$)						
31	0,99999994205930	2,99999922574831	4,99999477771376	7,2	6,6	6,0
63	0,9999999995068	2,9999999935431	4,9999999574412	10,3	9,7	9,1
127	0,9999999999990	2,9999999999939	4,9999999999607	13,0	12,7	12,1
255	1,00000000000000	3,00000000000006	4,9999999999978	15,0	13,7	13,4
511	1,00000000000024	2,9999999999976	4,9999999999648	11,6	13,1	12,2
Wartości dokładne						
—	1,00000000000000	3,00000000000000	5,00000000000000	∞	∞	∞

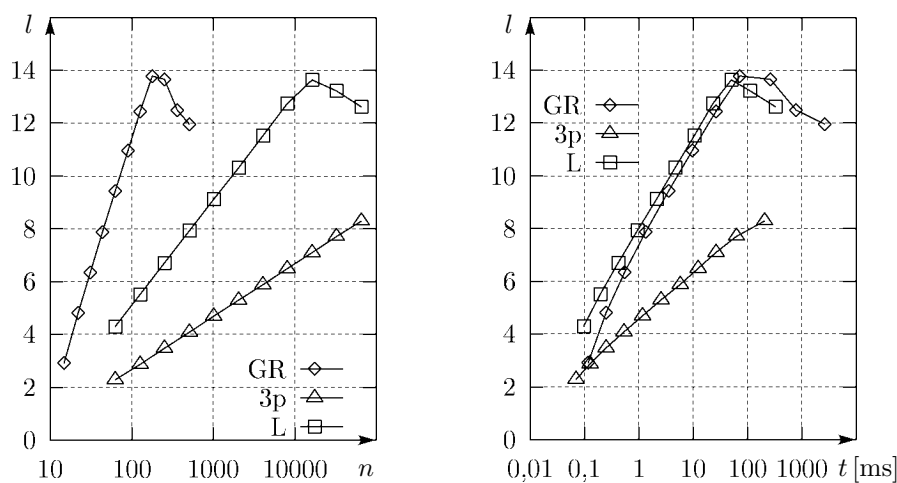
Rys. 9.2. Wyniki dla potencjału oscylatora harmonicznego. Zależność dokładności (liczby cyfr znaczących l) wyników numerycznych od liczby punktów siatki n (po lewej) i czasu obliczeń t (po prawej) w metodach: Guardioli-Rosa (GR), trójpunktowej (3p) i Lindberga (L)

Tabela 9.2. Wyniki obliczeń dla potencjału Konwenta (parametr $b = 0,03$, przedział całkowania $\langle a, b \rangle = \langle -8, 8 \rangle$)

n	Energie			Cyfry znaczące		
	ε_1	ε_2	ε_3	l_1	l_2	l_3
Metoda trójpunktowa						
255	0,55511148050809	0,55551068298819	1,00043762482779	3,6	3,6	4,2
1023	0,55524689212571	0,55564650538980	1,00049576752221	4,8	4,8	5,4
4095	0,55525535100209	0,55565498993725	1,00049939858744	6,0	6,0	6,6
16383	0,55525587966490	0,55565552020448	1,00049962551790	7,2	7,2	7,8
65535	0,55525591270625	0,5556555334612	1,00049963970101	8,4	8,4	9,0
Metoda Lindberga						
255	0,55525579508393	0,55565543508465	1,00049959232021	6,7	6,7	7,3
1023	0,55525591444144	0,55565555508547	1,00049964045797	9,1	9,1	9,7
4095	0,55525591490719	0,55565555555372	1,00049964064581	11,5	11,5	12,1
16383	0,55525591490900	0,55565555555555	1,00049964064654	14,0	14,0	14,7
65535	0,55525591490901	0,55565555555555	1,00049964064651	14,4	14,4	13,5
Metoda Guardioli-Rosa (aproksymanta $[3/2]$)						
31	0,55524668257101	0,55564627607349	1,00049599418856	4,8	4,8	5,4
63	0,55525591165999	0,55565555229003	1,00049963939015	8,2	8,2	8,9
127	0,55525591490628	0,55565555555281	1,00049964064549	11,3	11,3	12,0
255	0,55525591490904	0,55565555555554	1,00049964064664	13,3	13,4	13,0
511	0,55525591490890	0,55565555555556	1,00049964064614	12,7	15,4	12,4
Wartości dokładne						
—	0,55525591490901	0,55565555555556	1,00049964064655	∞	∞	∞

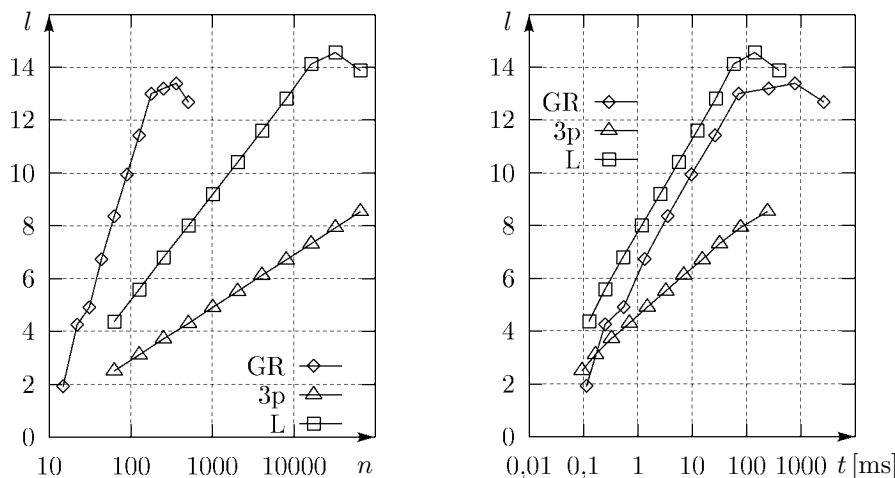
Rys. 9.3. Wyniki dla potencjału Konwenta. Zależność dokładności (liczby cyfr znaczących l) wyników numerycznych od liczby punktów siatki n (po lewej) i czasu obliczeń t (po prawej) w metodach: Guardioli-Rosa (GR), trójpunktowej (3p) i Lindberga (L)

Tabela 9.3. Wyniki obliczeń dla potencjału Morse'a (parametr $\alpha = 25$, przedział całkowania $\langle a, b \rangle = \langle -3, 9 \rangle$)

n	Energie			Cyfry znaczące		
	$-\varepsilon_1$	$-\varepsilon_2$	$-\varepsilon_3$	l_1	l_2	l_3
Metoda trójpunktowa						
255	0,81011956716462	0,49039474387133	0,25060819286672	3,8	3,1	2,6
1023	0,81000746762408	0,49002463578098	0,25003791499769	5,0	4,3	3,8
4095	0,81000046670575	0,49000153959724	0,25000236930979	6,2	5,5	5,0
16383	0,81000002916903	0,49000009622428	0,25000014808039	7,4	6,7	6,2
65535	0,81000000182306	0,49000000601401	0,25000000925502	8,6	7,9	7,4
Metoda Lindberga						
255	0,81000017013349	0,490000068276847	0,25000123634919	6,7	5,9	5,3
1023	0,81000000066353	0,49000000266228	0,25000000481983	9,1	8,3	7,7
4095	0,81000000000259	0,49000000001040	0,25000000001882	11,5	10,7	10,1
16383	0,81000000000001	0,49000000000004	0,25000000000007	13,9	13,1	12,5
65535	0,81000000000000	0,49000000000000	0,25000000000000	14,9	15,1	14,6
Metoda Guardioli-Rosa (aproksymanta [3/2])						
31	0,80999920738317	0,49007058827919	0,25034848303217	6,0	3,8	2,9
63	0,81000001152400	0,49000006149961	0,25000014160756	7,8	6,9	6,2
127	0,81000000000952	0,49000000005044	0,25000000011537	10,9	10,0	9,3
255	0,81000000000002	0,49000000000004	0,25000000000010	13,7	13,1	12,4
511	0,81000000000002	0,49000000000002	0,24999999999996	13,6	13,4	12,8
Wartości dokładne						
—	0,81000000000000	0,49000000000000	0,25000000000000	∞	∞	∞

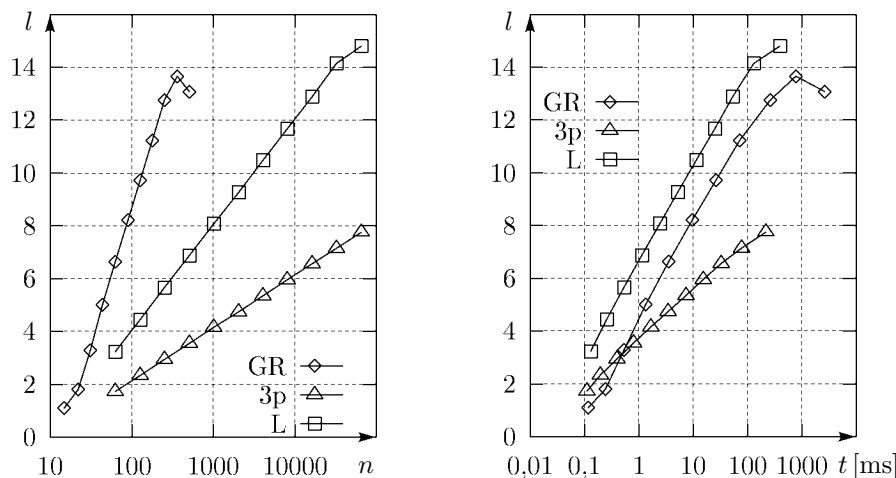
Rys. 9.4. Wyniki dla potencjału Morse'a. Zależność dokładności (liczby cyfr znaczących l) wyników numerycznych od liczby punktów siatki n (po lewej) i czasu obliczeń t (po prawej) w metodach: Guardioli-Rosa (GR), trójpunktowej (3p) i Lindberga (L)

Tabela 9.4. Wyniki obliczeń dla potencjału kulombowskiego ($l = 0$, przedział całkowania $\langle a, b \rangle = \langle 0, 75 \rangle$)

n	Energie			Cyfry znaczące		
	$-\varepsilon_1$	$-\varepsilon_2$	$-\varepsilon_3$	l_1	l_2	l_3
Metoda trójpunktowa						
255	0,97941658125694	0,24867309387079	0,11084745728657	1,7	2,3	2,6
1023	0,99866248059997	0,24991623712626	0,11109455919781	2,9	3,5	3,8
4095	0,99991619501660	0,24999476153006	0,11111007632751	4,1	4,7	5,0
16383	0,99999476136540	0,24999967258276	0,11111104643600	5,3	5,9	6,2
65535	0,99999967258212	0,24999997953637	0,11111110706890	6,5	7,1	7,4
Metoda Lindberga						
255	0,95923088912473	0,24482902422861	0,10957159897739	1,4	1,7	1,9
1023	0,99669665770281	0,24958660351682	0,11098857593844	2,5	2,8	3,0
4095	0,99978066168023	0,24997258060187	0,11110298663611	3,7	4,0	4,1
16383	0,99998609459576	0,24999826181599	0,11111059609278	4,9	5,2	5,3
65535	0,99999912788673	0,24999989098581	0,11111107881059	6,1	6,4	6,5
Metoda Guardioli–Rosa (aproksymanta [3/2]) — zawodzi zupełnie						
31	0,17210022821496	0,08563925563819	0,05118873718940	0,1	0,2	0,3
63	0,21072160467530	0,09891711658181	0,05723422630341	0,1	0,2	0,3
127	0,23408504577355	0,10630909191360	0,06045598160774	0,1	0,2	0,3
255	0,24468906908917	0,10952812287810	0,06183022023684	0,1	0,3	0,4
511	0,24845334293636	0,11065205584705	0,06230615688024	0,1	0,3	0,4
Wartości dokładne						
—	1,00000000000000	0,25000000000000	0,11111111111111	∞	∞	∞

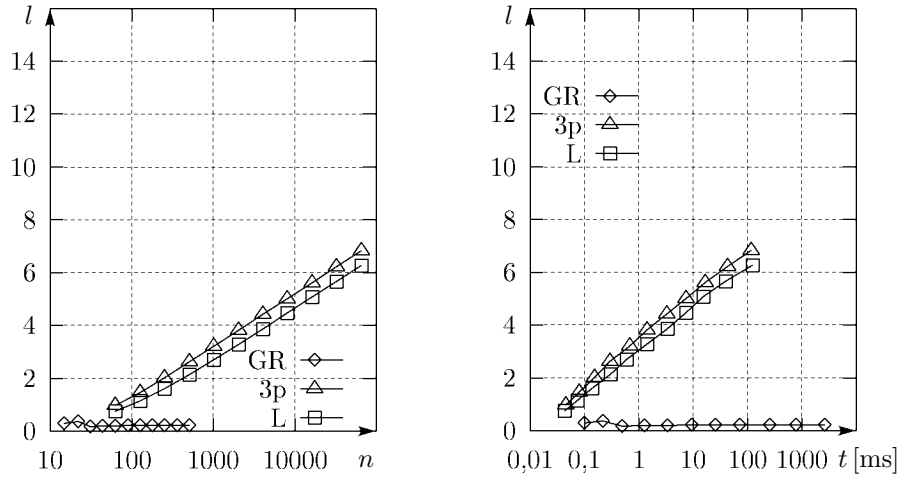
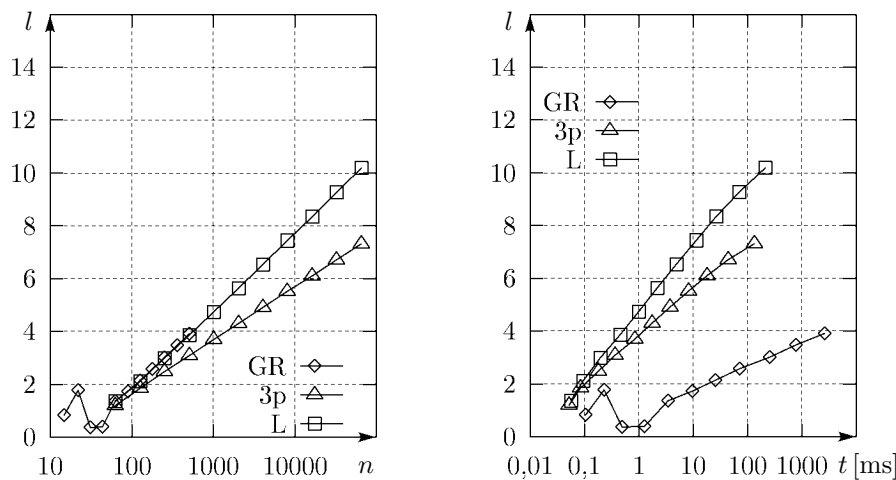
Rys. 9.5. Wyniki dla potencjału kulombowskiego ($l = 0$). Zależność dokładności (liczby cyfr znaczących l) wyników numerycznych od liczby punktów siatki n (po lewej) i czasu obliczeń t (po prawej) w metodach: Guardioli–Rosa (GR), trójpunktowej (3p) i Lindberga (L)

Tabela 9.5. Wyniki obliczeń dla potencjału kulombowskiego ($l = 1$, przedział całkowania $\langle a, b \rangle = \langle 0, 100 \rangle$)

n	Energie			Cyfry znaczące		
	$-\varepsilon_1$	$-\varepsilon_2$	$-\varepsilon_3$	l_1	l_2	l_3
Metoda trójpunktowa						
255	0,25081340883922	0,11148505728578	0,06269241346022	2,5	2,5	2,5
1023	0,25004974296261	0,11113403401616	0,06251181085621	3,7	3,7	3,7
4095	0,25000310469129	0,11111254206471	0,06250073735154	4,9	4,9	4,9
16383	0,25000019402664	0,11111120053897	0,06250004608075	6,1	6,1	6,1
65535	0,25000001212660	0,11111111670033	0,06250000287954	7,3	7,3	7,3
Metoda Lindberga						
255	0,25032370816289	0,11122554418979	0,06255105512988	2,9	3,0	3,1
1023	0,25000550801243	0,11111304802612	0,06250086234361	4,7	4,8	4,9
4095	0,25000008794270	0,11111114200383	0,06250001374707	6,5	6,6	6,7
16383	0,25000000138156	0,11111111159631	0,06250000021536	8,3	8,4	8,5
65535	0,25000000002162	0,11111111111870	0,06250000000285	10,1	10,2	10,3
Metoda Guardioli–Rosa (aproksymanta $[3/2]$)						
31	0,11724684018119	0,06499770337333	0,04125195951349	0,3	0,4	0,5
63	0,26324598870983	0,11585002489644	0,06461035532841	1,3	1,4	1,5
127	0,25221425876731	0,11189728771417	0,06285085992361	2,1	2,2	2,3
255	0,25029504443227	0,11121502529631	0,06254627862663	2,9	3,0	3,1
511	0,25003770010715	0,11112435961408	0,06250589626926	3,8	3,9	4,0
Wartości dokładne						
—	0,25000000000000	0,11111111111111	0,06250000000000	∞	∞	∞

Rys. 9.6. Wyniki dla potencjału kulombowskiego ($l = 1$). Zależność dokładności (liczby cyfr znaczących l) wyników numerycznych od liczby punktów siatki n (po lewej) i czasu obliczeń t (po prawej) w metodach: Guardioli–Rosa (GR), trójpunktowej (3p) i Lindberga (L)

9.3. Analiza dokładności

Wyznaczone numerycznie energie własne i wartości funkcji falowej w punktach siatki są tylko przybliżeniami wartości dokładnych. Źródłami błędów w obliczeniach numerycznych są:

- skończona dokładność reprezentacji liczb rzeczywistych w komputerze oraz wykonywanych operacji zmiennoprzecinkowych,
- skalowanie wartości własnych macierzy $\tilde{\mathbf{H}}$ wraz z kwadratem kroku siatki s (paragraf 2.3.1),
- ograniczona dokładność aproksymacji pochodnych za pomocą różnic skończonych (podrozdział 5.1, dodatek C),
- skończona długość przedziału całkowania.

Znajomość tych przyczyn pozwala zminimalizować ich wpływ, zostaną więc one kolejno omówione.

9.3.1. Dokładność operacji zmiennoprzecinkowych

Reprezentacja liczby zmiennoprzecinkowej w komputerze zajmuje ściśle określoną i skończoną liczbę komórek pamięci, zatem jej dokładność (liczba cyfr znaczących) jest także skończona. W tabeli 9.6 zestawione są dokładności poszczególnych typów zmiennoprzecinkowych, dostępnych w języku Delphi.

Tabela 9.6. Typy danych zmiennoprzecinkowych dostępne w językach Delphi i FORTRAN-77. Uwaga: liczba dziesiętnych cyfr znaczących nie może być podana jako liczba całkowita, gdyż komputer używa wewnętrznie reprezentacji dwójkowej.

Nazwa typu		Długość reprezentacji (bajty)	Liczba cyfr znaczących
Delphi	FORTAN-77		
SINGLE	REAL	4	6–7
REAL	—	6	11–12
DOUBLE	DOUBLE PRECISION	8	16–17
EXTENDED	—	10	19–20

Typy SINGLE, DOUBLE, EXTENDED (w nazewnictwie Delphi) są zgodne z normą IEEE 754 oraz architekturą wewnętrzną jednostek arytmetyki zmiennoprzecinkowej mikroprocesorów (FPU) rodziny Intel x86/x87 i zgodnych [142, 143] (typ REAL występuje głównie ze względów historycznych). W programie MARRS używany jest typ DOUBLE, za wyjątkiem fragmentów kodu, w których wymagana była największa dostępna dokładność i użyty został typ EXTENDED. W programie prec wykorzystywany jest typ DOUBLE PRECISION, co może skutkować nieco mniejszą dokładnością; tym niemniej kompilator g77 potrafi niejawnie umieszczać niektóre zmienne lokalne i wyniki pośrednie w rejestrach FPU x87, co odpowiada użyciu typu EXTENDED.

Większość algorytmów zaimplementowanych w programach MARRS i PREC wykorzystuje długie, rekurencyjnie zdefiniowane ciągi, których liczba wyrazów jest równa liczbie punktów siatki. Może to prowadzić do kumulowania się błędów zaokrągleń, a co za tym idzie, nie można dowolnie poprawiać dokładności poprzez zwiększanie liczby punktów siatki n .

9.3.2. Skalowanie wartości własnych

Przyjmijmy, że kumulacja błędów numerycznych nie pogarsza istotnie dokładności wyznaczania poszczególnych wartości własnych macierzy \mathbf{H} przy wzroście n (założenie to jest spełnione np. dla algorytmu bisekcji [25]). Niech dokładność ta wynosi $\tilde{\sigma}$. Obliczone wartości własne $\tilde{\varepsilon}_j^{\text{num}}$ macierzy \mathbf{H} związane są z przybliżeniami $\mathcal{E}_j^{\text{num}}$ energii własnych hamiltonianu \mathcal{H} wzorem

$$\mathcal{E}_j^{\text{num}} = \frac{V_c \tilde{\varepsilon}_j^{\text{num}}}{s^2 \alpha}. \quad (9.2)$$

Błędy skalują się analogicznie:

$$|\mathcal{E}_j^{\text{num}} - \mathcal{E}_j| = \frac{V_c \tilde{\sigma}}{s^2 \alpha} \approx K_j^{(1)} n^2,$$

gdzie stała $K_j^{(1)} = V_c \tilde{\sigma} / [(b-a)^2 \alpha]$ (por. wzory (2.3)–(2.9)). Ze wzoru (9.1) otrzymujemy zależność

$$r_j^{(1)} \approx K_j^{(1)} n^2 \approx \frac{K_j^{(1)}}{\mathcal{E}_j} n^2 = k_j^{(1)} n^2. \quad (9.3)$$

Widzimy zatem, że skalowanie wartości własnych (9.2) prowadzi do spadku dokładności wyznaczania energii własnych hamiltonianu przy zagęszczeniu siatki.

9.3.3. Błąd aproksymacji różnicowej drugiej pochodnej

Na dokładność obliczonych energii własnych hamiltonianu istotnie wpływa skończona dokładność aproksymacji różnicami skończonymi operatora drugiej pochodnej. Niech $\bar{\Delta}_s^2$ oznacza wybrany operator różnicowy na siatce o kroku s . Wówczas

$$\left. \frac{d^2}{dx^2} \psi(x) \right|_{x=x_i} = \bar{\Delta}_s^2 \psi(x) \Big|_{x=x_i} + O(s^p),$$

gdzie $O(s^p)$ oznacza błąd aproksymacji, $p \geq 2$ — liczba parzysta (w przypadku przybliżenia trójpunktowego (rozdział 3) mamy $p = 2$, w metodzie Lindberga (podrozdział 5.5) i przybliżeniu pięciopunktowym (podrozdział 5.4) $p = 4$, a w metodzie Guardiola–Rosa (podrozdział 5.1) z aproksymantą Padégo [j/m] — $p = 2(j+m)$). Wstawiając operator $\bar{\Delta}_s^2$ do równania Schrödingera (2.5) i przenosząc błąd $O(s^p)$ na jego prawą stronę, otrzymujemy

$$[-\bar{\Delta}_s^2 + s^2 \alpha v(x)] \psi(x) = s^2 \alpha [\varepsilon + O(s^p)] \psi(x).$$

Wnosimy stąd, że obliczona bezwymiarowa energia własna jest obciążona błędem rzędu $O(s^p)$: $\varepsilon_j^{\text{num}} = \varepsilon_j + O(s^p)$, a zatem

$$r_j^{(2)} \approx \frac{|\varepsilon_j^{\text{num}} - \varepsilon_j|}{\varepsilon_j} = \frac{V_c}{\varepsilon_j} |\varepsilon_j^{\text{num}} - \varepsilon_j| = O(s^p) = k_j^{(2)} n^{-p}, \quad (9.4)$$

gdzie stała $k_j^{(2)}$ zależy od wyboru operatora różnicowego oraz od wartości $(p+2)$ -ej pochodnej funkcji falowej $\psi^{(j)}(x)$ na przedziale $\langle a, b \rangle$ (por. wzory (5.4)–(5.19)).

Uwzględniając składniki błędu pochodzące ze skalowania wartości własnych oraz aproksymacji różnicowej (wzory (9.3) i (9.4)), otrzymujemy łączne oszacowanie błędu:

$$r_j = r_j^{(1)} + r_j^{(2)} \approx k_j^{(1)} n^2 + k_j^{(2)} n^{-p}. \quad (9.5)$$

Funkcja (9.5) ma dwa składniki — rosnący i malejący, zatem posiada minimum dla pewnej wartości $n = n_0$. Wynika stąd, że zwiększając liczbę punktów siatki powyżej n_0 pogarszamy, a nie poprawiamy dokładność!

Oszacowanie (9.5) ma charakter jedynie jakościowy. Teoretyczne wyznaczenie wartości n_0 jest bardzo trudne (np. własności przybliżenia drugiej pochodnej zależą od postaci funkcji falowej, a zatem od postaci potencjału). Rzeczywistą dokładność osiąganą za pomocą podprogramów z Dodatku F możemy znaleźć, analizując wyniki testów zestawione w tabelach 9.1–9.5 i na rysunkach 9.2–9.6. Podobna analiza przeprowadzona została w publikacjach [141] dla procedur numerycznych programu MARRS.

Zgodnie z przewidywaniami dokładność obliczeń poprawia się przy wzroście n aż do osiągnięcia najlepszej wartości, a dalsze zwiększanie n nie poprawia, a nawet pogarsza wyniki (punkty przegięcia na wykresach — rys. 9.2–9.6). Efekt ten najlepiej widać na gładkich, nieosobliwych potencjałach (oscylatora harmonicznego, Konwenta i Morse’a — rys. 9.2–9.4). Wyniki przeprowadzonych testów są istotnie różne dla metody Guardioli–Rosa (GR) oraz dla metod trójpunktowej i Lindberga. Różnice te wynikają z fundamentalnych różnic między tymi metodami: w metodzie GR mamy do czynienia z algebraicznym zagadnieniem własnym z macierzą gęstą, w pozostałych — z macierzą pasmową. Co prawda metoda GR zapewnia znacznie lepsze przybliżenie drugiej pochodnej (dla aproksymanty Padégo [3/2] — z dokładnością $O(s^{10})$), wykazuje jednak znacznie większą złożoność obliczeniową i jest bardziej podatna na kumulację błędów zaokrągleń. Skutkiem tego jest spadek dokładności już dla $n \sim 200$ dla potencjałów nieosobliwych. Tymczasem dla metody Lindberga efekt ten występuje przy $n \sim 2 \cdot 10^4$, a dla najmniej złożonej obliczeniowo metody trójpunktowej nie został zaobserwowany nawet przy $n \sim 6 \cdot 10^5$.

Dla osobliwego potencjału kulombowskiego z $l = 0$ metoda GR zawodzi zupełnie (tabela 9.4, rys. 9.5) i nie działa zadowalająco dla $l = 1$ (tabela 9.5, rys. 9.6); także metoda Lindberga nie wypada w tych przypadkach tak dobrze w porównaniu z metodą trójpunktową, jak dla potencjałów nieosobliwych. Otrzymane rezultaty wskazują, że dla potencjałów zbliżonych do osobliwych należy stosować metodę GR co najmniej z dużą ostrożnością.

Liniowa zależność liczby cyfr znaczących od logarytmu n , widoczna na rys. 9.2–9.6 (aż do osiągnięcia najlepszej dokładności danej metody) świadczy o istnieniu zależ-

ności potęgowej, wyrażającej się równoważnymi formułami

$$r = Cn^{-q} \quad \Longleftrightarrow \quad l = q \log_{10} n - \log_{10} C.$$

Spodziewamy się, że wartość wykładnika q będzie bliska liczbie p , określającej dokładność przybliżenia drugiej pochodnej $O(s^p)$ w równaniu Schrödingera. Stała C natomiast może zależeć od kształtu potencjału. Otrzymane przez nas wyniki, zebrane w tabeli 9.7, potwierdzają słuszność tych oczekiwań w przypadku potencjałów nieosobliwych. Natomiast dla potencjałów z osobliwościami metody wyższych rzędów ($p > 2$) sprawują się poniżej oczekiwań.

Tabela 9.7. Wartości parametrów C , q i p w wyrażeniach $r = Cn^{-q}$ i $O(s^p)$ dla poszczególnych potencjałów i metod

Potencjał	Metoda	p	$q \pm 0,1$	$\log_{10} C \pm 0,1$
oscylator harmoniczny	trójpunktowa	2	2,0	1,2
	Lindberga	4	3,9	2,9
	GR [3/2]	10	10,0	8,6
Morse'a	trójpunktowa	2	2,0	1,6
	Lindberga	4	4,0	3,6
	GR [3/2]	10	10,3	11,4
kulombowski ($l = 1$)	trójpunktowa	2	2,0	2,2
	Lindberga	4	2,9	3,7
	GR [3/2]	10	2,9	3,3

9.3.4. Zależność wyników numerycznych od wyboru przedziału całkowania

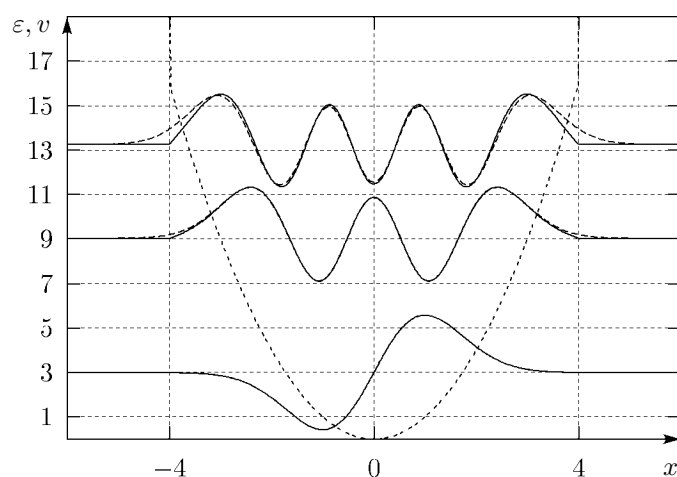
Stosowane podejście algebraiczne zakłada zerowe wartości funkcji falowej na końcach przedziału całkowania $\langle a, b \rangle$ (i poza nim), przy czym przedział ten musi być skończony ($|a|, |b| < \infty$). Fizyczny sens tych warunków brzegowych sprowadza się do umieszczenia na końcach przedziału całkowania nieskończonych barier potencjału. Widać stąd, że metody algebraiczne szczególnie nadają się do rozwiązywania równania Schrödingera z potencjałami zadanymi na skończonych przedziałach, w których nieskończone bariery występują explicite (np. potencjał Pöschla [35]).

Zastosowane warunki brzegowe mają wpływ na wyniki numeryczne dla potencjałów określonych na nieskończonych przedziałach. W dyskusji uwzględnimy dwa typy potencjałów, posiadających:

- skończoną liczbę stanów związanych (np. skończona studnia prostokątna, potencjały Morse'a i Lennarda–Jonesa);
- nieskończoną liczbę stanów związanych (np. oscylator harmoniczny, nieskończona studnia prostokątna lub trójkątna, potencjały Konwenta, Tody [144] i kulombowski).

Tabela 9.8. Wpływ granic całkowania na dokładność obliczania energii własnych na przykładzie oscylatora harmonicznego. Wyniki dla metody trójpunktowej, $n = 2^{10} - 1 = 1023$

j	Energie obliczone $\varepsilon_j^{\text{num}}$	Energie dokładne ε_j	Cyfry znaczące l_j
1	0,99999716725788	1,00000000000000	5,5
2	3,00001013660034	3,00000000000000	5,5
3	5,00035278427913	5,00000000000000	4,2
4	7,00328754084526	7,00000000000000	3,3
5	9,01912323729884	9,00000000000000	2,7
6	11,07859709737236	11,00000000000000	2,1
7	13,24187219537654	13,00000000000000	1,7
8	15,58678715583260	15,00000000000000	1,4
9	18,18116369745394	17,00000000000000	1,2



Rys. 9.7. Wpływ granic całkowania na dokładność obliczania funkcji falowych na przykładzie potencjału oscylatora harmonicznego (---- potencjał z nieskończonymi barierami na końcach przedziału, — dokładne funkcje falowe, --- funkcje obliczone numerycznie metodą DWSZ). Pokazane są funkcje ψ_2 , ψ_5 oraz ψ_7 (ψ_1 — stan podstawowy)

Dla obu typów potencjałów musimy koniecznie zadbać o to, aby granice przedziału całkowania leżały w obszarach, gdzie dokładna funkcja falowa poszukiwanego stanu własnego jest dostatecznie bliska zeru. Zilustrujemy to na przykładzie oscylatora harmonicznego. Tabela 9.8 oraz rysunek 9.7 przedstawiają porównanie wyników numerycznych i analitycznych (energii i funkcji falowych) otrzymanych przy właściwym i niewłaściwym wyborze granic całkowania. Widać, że granice $\langle a, b \rangle = \langle -4, 4 \rangle$ są dość dobrze dobrane tylko dla niższych stanów własnych. Zwróćmy uwagę na to, że przy prawidłowo dobranych granicach obliczona numerycznie funkcja falowa dąży

do zera asymptotycznie, podobnie jak rozwiązanie analityczne.

Oczywiste jest, że obliczenia numeryczne prowadzone dla potencjałów o nieskończonej liczbie stanów związanych dają nam tylko skończoną liczbę najniżej położonych poziomów energetycznych. Szerokość przedziału całkowania musimy dobrać tak, aby na jego końcach potencjał miał wartości większe od energii poszukiwanego stanu własnego. Inaczej mówiąc, spośród otrzymanych numerycznie energii, sens fizyczny mogą mieć mniejsze od $V_g = \min(V(a), V(b))$; na rys. 9.7 tą wartością jest $V_g = 4^2 = 16$. Warunek ten wiąże się z warunkiem poprzednim: jeśli poziom energetyczny znajduje się dostatecznie głęboko poniżej V_g , to funkcja falowa zanika w barierze dostatecznie szybko.

Podkreślmy, że wybór właściwego (tj. dostatecznie szerokiego) przedziału całkowania nie może polegać na jego niekontrolowanym poszerzaniu — w takim przypadku albo pogarszamy dokładność obliczeń zwiększając skok siatki, albo nadmiernie zwiększamy ilość punktów siatki, co niepotrzebnie wydłuża obliczenia i również może wpłynąć ujemnie na ich dokładność. Także bardzo małe wartości funkcji falowej w pobliżu końców nazbyt szerokiego przedziału (być może nawet poniżej najmniejszej liczby reprezentowalnej w stosowanym formacie zmiennoprzecinkowym) mogą prowadzić do niedokładności lub istotnych trudności przy numerycznym wyznaczaniu funkcji falowych.

9.4. Dokładność a czas obliczeń

W praktycznych zastosowaniach, obok wymaganej dokładności obliczeń, istotnym czynnikiem jest ich czas trwania i niezbędny rozmiar pamięci operacyjnej³⁾. Dokładność obliczeń rzędu 12–14 cyfr jest w praktyce rzadko potrzebna, szczególnie gdy porównujemy wyniki obliczeń z danymi eksperymentalnymi, zazwyczaj obciążonymi znacznie większymi niepewnościami.

Analizując zależności osiąganey liczby cyfr znaczących wyniku numerycznego od czasu obliczeń (rys. 9.2–9.6), możemy wyciągnąć następujące wnioski:

- Dla nieosobliwych potencjałów metody Lindberga i Guardioli–Rosa [3/2] wykazują się porównywalną dokładnością maksymalną (10–12 cyfr znaczących), a czasy obliczeń są tego samego rzędu (rys. 9.2–9.4).
- Osiągnięcie danej dokładności metodą trójpunktową wymaga w przypadku potencjałów nieosobliwych znacznie dłuższego czasu obliczeń — około 10-krotnie dłuższego dla 5 cyfr znaczących, 100-krotnie dłuższego dla 8 i (z ekstrapolacji wyników) 1000-krotnie dłuższego dla 12 cyfr znaczących.
- W przypadku potencjałów osobliwych metoda Guardioli–Rosa [3/2] nie pozwala wyznaczyć energii własnych z zadowalającą dokładnością. Dla takich potencjałów istotniejsza wydaje się być duża lokalna gęstość siatki w pobliżu

³⁾ Problemy tego rodzaju nie zawsze można rozwiązać poprzez zakup szybszego (zazwyczaj droższego) komputera. Innym rozwiązaniem jest zastosowanie (lub odkrycie nowego) efektywniejszego algorytmu.

osobliwości, niż duża dokładność $O(s^p)$ przybliżenia drugiej pochodnej. Metoda Lindberga w przypadku potencjałów osobliwych działa dość dobrze, lecz jej dokładność może być mniejsza od teoretycznej wartości $O(s^4)$, a nawet zbliżona do dokładności $O(s^2)$ metody trójpunktowej.

- Dla zadanej liczby punktów siatki najszybsza jest metoda trójpunktowa, a dla danej dokładności — zazwyczaj metoda Lindberga.
- Najbardziej stabilna i uniwersalna, ze względu na zachowanie się dla zbadanych przez nas klas potencjałów, wydaje się być metoda trójpunktowa, choć metoda Lindberga niewiele jej ustępuje.

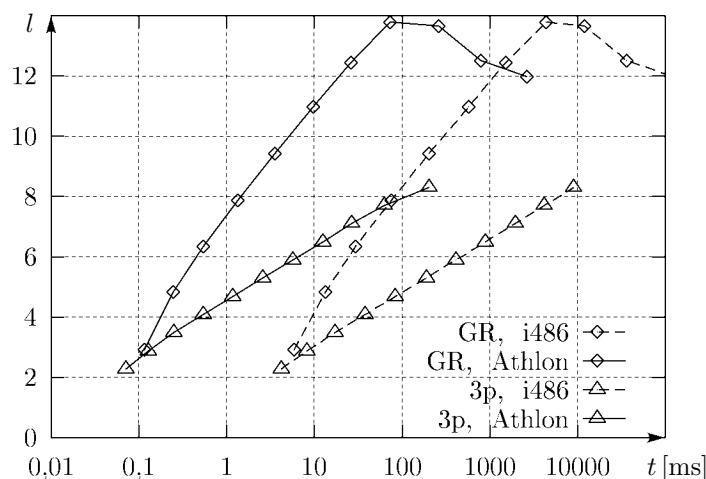
Pogłębiona analiza algorytmów oraz nasze dotychczasowe doświadczenia w pracy z nimi pozwalają wnosić dodatkowo, że:

- Metody GR i Lindberga zachowują się dla potencjałów nieciągłych (np. prostokątne studnie potencjału; program PREC zawiera również taki przykład) podobnie, jak dla potencjałów osobliwych. W tym przypadku dodatkowym czynnikiem wpływającym na dokładność wszystkich metod jest położenie punktów siatki względem punktów nieciągłości potencjału.
- Metoda pięciopunktowa (rzędu $O(s^4)$) zachowuje się podobnie do metody Lindberga. Osiąga nieznacznie gorszą dokładność i jest wolniejsza, zatem jej stosowanie nie wydaje się być celowe.
- Jeżeli oprócz energii obliczamy także funkcje falowe, to metoda GR staje się względnie mało efektywna: w procesie redukcji macierzy gęstej do postaci trójdzielnej musimy akumulować ortogonalne transformacje podobieństwa, co wymaga dodatkowej pamięci i istotnie wydłuża czas obliczeń (por. podrozdział 6.5; opcja ta nie jest zaimplementowana w programie PREC). Natomiast metoda trójpunktowa lub Lindberga umożliwia wyznaczanie funkcji falowych stosunkowo niedużym kosztem, odpowiednio metodą DWSZ lub metodą DWSZ uogólnioną na kwazisymetryczne macierze trójdzielne.
- Wyznaczanie wartości własnych macierzy, występujących w metodach: trójpunktowej i Lindberga, można przyspieszyć, stosując metodę bisekcji (wyszukiwanie binarne) tylko do lokalizacji przedziałów zawierających pojedyncze wartości własne. Następnie w tych podprzedziałach stosujemy metody oparte na wartościach wielomianu charakterystycznego i jego pochodnych (ciąg Sturma, patrz podrozdział 3.4); są one jednak nieco mniej dokładne, niż metoda bisekcji. Algorytmy takie zaimplementowane są w programie MARRS.
- Rozmiar pamięci operacyjnej potrzebny do osiągnięcia danej dokładności jest tego samego rzędu w metodach GR i Lindberga, i mniejszy niż w metodzie trójpunktowej.
- Przy tym samym rozmiarze siatki, zapotrzebowanie na pamięć operacyjną dla metod: trójpunktowej, Lindberga i pięciopunktowej jest znacznie mniejsze niż dla metody GR, która wymaga $O(n^2)$ komórek zamiast $O(n)$. Kosztem istotnego spowolnienia obliczeń, można nawet zredukować zapotrzebowanie na

pamięć metod: trójpunktowej i Lindberga do $O(1)$, obliczając wielokrotnie $\tilde{v}_i = \tilde{v}(x_i)$.

Powyższe wnioski można podsumować następująco: w ogólnym przypadku najbardziej warta polecenia wydaje się być metoda Lindberga — jest dokładna, szybka i względnie uniwersalna. Metoda trójpunktowa, jako najprostsza i zarazem najbardziej uniwersalna i stabilna, także jest godna uwagi (przypomnijmy, że można ją łatwo uogólnić na przypadek równania masy efektywnej — rozdział 4). Metoda Guardioli–Rosa działa bardzo dobrze, ale tylko w przypadku potencjałów nieosobliwych.

Na koniec przedstawiamy porównanie zależności $l(t)$ dla programu PREC uruchomionego na dwóch mikrokomputerach należących do różnych generacji (rys. 9.8). Zmiana komputera powoduje w zasadzie tylko wzrost szybkości obliczeń o stały czynnik.⁴⁾ Pewne niewielkie różnice w wydajności poszczególnych algorytmów pojawiają się też po zastosowaniu różnych kompilatorów. Z drugiej strony, przeprowadzenie testów na maszynach o architekturze odmiennej od PC, do czego Czytelników zachęcamy, może dać nieco inną klasyfikację algorytmów ze względu na ich wydajność.



Rys. 9.8. Zależność liczby cyfr znaczących l wyników od czasu obliczeń t (oscylator harmoniczny, metoda Guardioli–Rosa (GR) oraz metoda trójpunktowa (3p)). Wyniki dla: PC z mikroprocesorem Intel i486DX2 80 MHz, rok prod. 1994 (i486), PC z mikroprocesorem AMD Athlon 1200 MHz, rok prod. 2001 (Athlon). Liczba punktów siatki n — jak na rys. 9.2–9.6

⁴⁾ Dokładniejsza analiza wykazuje, że ów czynnik jest nieco inny dla każdej z metod, choć tego samego rzędu. Poszczególne algorytmy nieco inaczej zmieniają swoje zachowanie po zastosowaniu mikroprocesora z tej samej rodziny, lecz o nowocześniejszej architekturze.

Niestacjonarne równanie Schrödingera

Do tej pory zajmowaliśmy się wyłącznie stacjonarnym (bezczasowym) równaniem Schrödingera. Warto także poświęcić nieco miejsca równaniu Schrödingera z czasem, opisującym ewolucję układu kwantowego i będącym odrębnym i ważnym zagadnieniem.

10.1. Ewolucja układu kwantowego

Przyjmujemy, że na cząstkę działa zależne od czasu T i zmiennej przestrzennej X pole potencjalne $V(X, T)$. Dynamikę jednowymiarowego układu kwantowego określa w *obrazie Schrödingera* równanie falowe

$$i\hbar \frac{\partial}{\partial T} \Psi(X, T) = \mathcal{H}(X, T) \Psi(X, T) = \left[-\frac{\hbar^2}{2M} \frac{\partial^2}{\partial X^2} + V(X, T) \right] \Psi(X, T), \quad (10.1)$$

gdzie $\Psi(X, T)$ — zależna od czasu funkcja falowa.

Rozpatrzmy przypadek, gdy hamiltonian nie zależy jawnie od czasu. Całkując powyższe równanie, otrzymujemy

$$\Psi(X, T + \Delta T) = \exp\left(\frac{\mathcal{H}\Delta T}{i\hbar}\right) \Psi(X, T), \quad (10.2)$$

gdzie $e^{\mathcal{H}/(i\hbar)\Delta T}$ jest unitarnym *operatorem ewolucji*, co zapewnia spełnienie warunku

$$\int_{-\infty}^{\infty} \Psi^*(X, T) \Psi(X, T) dX = 1.$$

Jak widać, znając funkcję falową w pewnej chwili T_0 , możemy wyznaczyć jej wartości dla dowolnej chwili T .

Postępując analogicznie, jak w rozdziale 2, możemy równanie (10.1) sprowadzić do postaci bezwymiarowej:

$$i\alpha \frac{\partial}{\partial t} \psi(x, t) = \left[-\frac{\partial^2}{\partial x^2} + \alpha v(x) \right] \psi(x, t). \quad (10.3)$$

Obok jednostek długości L_c , masy M_c i energii V_c , wprowadziliśmy dodatkowo jednostkę czasu $T_c = \hbar/V_c$.

Ewolucję układu będziemy rozważać w skończonym przedziale czasowym $\langle t_p, t_k \rangle$ oraz na skończonym przedziale $\langle a, b \rangle$ zmienności współrzędnej przestrzennej. Sformułujemy dyskretną postać równania (10.3), posługując się siatką punktów

$$(x_i, t_j) = (a + is, t_p + j\tau), \quad s = \frac{b-a}{n_x+1}, \quad \tau = \frac{t_k-t_p}{n_t+1}.$$

Stan układu (przybliżoną wartość funkcji falowej) w chwili t_j reprezentuje wektor

$$\boldsymbol{\psi}^{(j)} = (\psi_1^{(j)}, \dots, \psi_{n_x}^{(j)})^T,$$

gdzie $\psi_i^{(j)} = \psi(x_i, t_j)$.

Operatorowi energii nadajemy także dyskretną (przybliżoną) postać $\tilde{\mathbf{H}}$, stosując jedną z aproksymacji operatora drugiej pochodnej przedstawionych w rozdziałach 2 i 5. W dalszych rozważaniach zastosujemy przybliżenie trójpunktowe (macierz we wzorze (2.8)).

Dyskretną postacią formalnego rozwiązania (10.2) czasowego równania Schrödingera jest

$$\boldsymbol{\psi}^{(j+1)} = \exp[-i\tilde{\mathbf{H}}\tau/\alpha]\boldsymbol{\psi}^{(j)} \quad (10.4)$$

Macierz unitarna $\mathbf{U} = \exp(-i\tilde{\mathbf{H}}\tau/\alpha)$ określa nieskończony szereg. W obliczeniach numerycznych używamy jej przybliżonej postaci. Stosując aproksymanty Padégo najniższych rzędów, otrzymujemy

$$\mathbf{U}[1/0] = \mathbf{1} - i\tilde{\mathbf{H}}\tau/\alpha, \quad (10.5)$$

$$\mathbf{U}[0/1] = (\mathbf{1} + i\tilde{\mathbf{H}}\tau/\alpha)^{-1}, \quad (10.6)$$

$$\mathbf{U}[1/1] = (\mathbf{1} + \frac{1}{2}i\tilde{\mathbf{H}}\tau/\alpha)^{-1}(\mathbf{1} - \frac{1}{2}i\tilde{\mathbf{H}}\tau/\alpha). \quad (10.7)$$

Można pokazać, że aproksymanty (10.5) i (10.6) nie są macierzami unitarnymi (por. zadanie 10/2; unitarność operatora ewolucji własność $\mathbf{U}^H = \mathbf{U}^{-1}$) oznacza, iż spełnione jest relacja $\|\boldsymbol{\psi}^{(j+1)}\|_2 = \|\boldsymbol{\psi}^{(j)}\|_2$). Dla unitarnej aproksymanty (10.7) otrzymamy dyskretną postać czasowego równania Schrödingera

$$(\mathbf{1} + \frac{1}{2}i\tilde{\mathbf{H}}\tau/\alpha)\boldsymbol{\psi}^{(j+1)} = (\mathbf{1} - \frac{1}{2}i\tilde{\mathbf{H}}\tau/\alpha)\boldsymbol{\psi}^{(j)}. \quad (10.8)$$

10.2. Metody rozwiązywania dyskretnego równania ewolucji

Przedstawimy krótko metody numerycznego rozwiązywania równania (10.8).

Stosując trójpunktowe przybliżenie dla drugiej pochodnej (C.7), otrzymujemy z równania (10.8) dla każdego z punktów siatki x_1, \dots, x_n

$$\begin{aligned} -i\beta\psi_{i-1}^{(j+1)} + [1 + i\beta(2 + \tilde{v}_i)]\psi_i^{(j+1)} - i\beta\psi_{i+1}^{(j+1)} = \\ = i\beta\psi_{i-1}^{(j)} + [1 - i\beta(2 + \tilde{v}_i)]\psi_i^{(j)} + i\beta\psi_{i+1}^{(j)} \quad (i = 1, \dots, n), \end{aligned} \quad (10.9)$$

gdzie $\beta = \tau/(2\alpha s^2)$, a $\tilde{v}_i = \alpha s^2 v_i$. Nakładając warunki brzegowe $\psi_0 = \psi_n = 0$, oznaczające znikanie funkcji falowej na brzegach przedziału $\langle a, b \rangle$ w dowolnej chwili czasu, otrzymujemy z (10.9) równanie macierzowe

$$\begin{pmatrix} 1 + i\beta(2 + \tilde{v}_1) & -i\beta & & \\ -i\beta & \ddots & \ddots & \\ & \ddots & \ddots & -i\beta \\ & & -i\beta & 1 + i\beta(2 + \tilde{v}_n) \end{pmatrix} \begin{pmatrix} \psi_1^{(j+1)} \\ \vdots \\ \vdots \\ \psi_n^{(j+1)} \end{pmatrix} =$$

$$= \begin{pmatrix} 1 - i\beta(2 + \tilde{v}_1) & i\beta & & \\ & i\beta & \ddots & \\ & & \ddots & \ddots \\ & & & \ddots & -i\beta \\ & & & i\beta & 1 - i\beta(2 + \tilde{v}_1) \end{pmatrix} \begin{pmatrix} \psi_1^{(j)} \\ \vdots \\ \vdots \\ \psi_n^{(j)} \end{pmatrix}. \quad (10.10)$$

Po prawej stronie znaku równości występuje iloczyn macierzy trójdzielnej i wektora, który można łatwo wyznaczyć kosztem $\sim n$ operacji zmiennoprzecinkowych; zwróćmy jednak uwagę, że każde dodawanie lub mnożenie liczb zespolonych wymaga kilku elementarnych operacji zmiennoprzecinkowych.

Pozostaje nam do rozwiązania trójdzielny układ równań

$$\mathbf{T}\psi^{(j+1)} = \phi, \quad (10.11)$$

gdzie $\phi = (\mathbf{1} - \frac{1}{2}i\tilde{\mathbf{H}}\tau/\alpha)\psi^{(j)}$, a $\mathbf{T} = \mathbf{1} + \frac{1}{2}i\tilde{\mathbf{H}}\tau/\alpha$ jest zespoloną symetryczną macierzą trójdzielną. Układ równań tej postaci zazwyczaj rozwiązuje się przy wykorzystaniu rozkładu trójkątnego macierzy \mathbf{T} w postaci $\mathbf{T} = \mathbf{SLR}$, gdzie \mathbf{L} — macierz trójkątna dolna z jedynkami na diagonalu, \mathbf{R} — macierz trójkątna górna i \mathbf{S} — macierz przestawień wierszy. Nie będziemy tu omawiać szczegółowo algorytmu takiego rozkładu. Zaimplementowany jest on np. w podprogramie **CGTSV** biblioteki LAPACK, który można wykorzystać do rozwiązywania układu równań (10.11).

Można pokazać, że zastosowanie przybliżeń operatora drugiej pochodnej różnicami skończonymi wyższych rzędów prowadzi do układu równań z zespoloną macierzą pasmową.

Zadania

10/1. Wyprowadzić aproksymanty Padégo (10.5)–(10.7).

10/2. Wykazać, że aproksymanty (10.5) i (10.6) nie są unitarne, natomiast (10.7) jest.

10/3. Znaleźć postać równania (10.10) dla układu opisanego hamiltonianem masy efektywnej, którego postać bezwymiarowa znajduje się w równaniu (4.5).

Zakończenie

Mamy nadzieję, że po dokładnym zapoznaniu się z treścią tej książki, Czytelniczka lub Czytelnik nauczą się rozwiązywać samodzielnie zagadnienia fizyczne prowadzące do stacjonarnego jednowymiarowego równania Schrödingera lub, ogólniej, symetrycznego AZW. Podręcznik ten będzie stanowił pomoc przy tworzeniu własnych programów komputerowych i korzystaniu z pakietu MARRS lub podprogramów bibliotecznych.

Jesteśmy przekonani, że zawartość tego podręcznika będzie cennym materiałem dydaktycznym służącym studentom, jak i nauczycielom akademickim prowadzącym wykłady, ćwiczenia rachunkowe lub zajęcia w laboratoriach komputerowych z fizyki, fizyki kwantowej, mechaniki kwantowej, fizyki ciała stałego, fizyki komputerowej czy fizyki struktur niskowymiarowych.

W tym wydaniu podręcznika nie omawialiśmy metod numerycznych rozwiązywania dwu- i trójwymiarowego równania Schrödingera (poza przypadkami, które dają się sprowadzić do problemów jednowymiarowych), ani też — bardziej szczegółowo — czasowego równania Schrödingera. Autorzy mają nadzieję, że przedstawione one zostaną w następnych wydaniach.

Będziemy wdzięczni wszystkim Czytelniczkom i Czytelnikom, którzy zauważą niedociągnięcia tego wydania i podzielą się z nami swoimi uwagami i spostrzeżeniami. Najłatwiej będzie to uczynić za pomocą poczty elektronicznej — oto adresy autorów:

Włodzimierz Salejda

E-mail: ssalejda@if.pwr.wroc.pl

Web: <http://www.if.pwr.wroc.pl/~ssalejda/>

Michał H. Tyc

E-mail: mhtyc@if.pwr.wroc.pl

Web: <http://www.if.pwr.wroc.pl/~mhtyc/>

Marcin Just

E-mail: mjust@if.pwr.wroc.pl

Web: <http://www.if.pwr.wroc.pl/~mjust/>

Problemy do samodzielnego rozwiązania

W celu nabycia przez studentów umiejętności posługiwania się przedstawionymi w tym podręczniku metodami, poniżej zamieszczamy zagadnienia do samodzielnego rozwiązania w formie studenckich projektów. Z podręczników mechaniki kwantowej, zbiorów zadań i bieżącej literatury wybraliśmy kilkanaście zagadnień. Są to jednowymiarowe równania Schrödingera z różnymi potencjałami lub problemy trójwymiarowe sprowadzające się do jednowymiarowych.

A.1. Zagadnienia jednowymiarowe

Dla potencjału wybranego z listy zamieszczonej poniżej, napisać (w dowolnym języku programowania) program umożliwiający:

- Obliczanie energii własnych kilku najniższych stanów związanych cząstki kwantowej zadaną przez użytkownika dokładnością;
- Badanie wpływu na wyznaczone numerycznie wartości i wektory własne:
 - parametrów określających potencjał — należy rozpatrzyć możliwe postaci podanych potencjałów $V(X)$ w zależności od parametrów; parametry należy dobierać tak, aby $V(X)$ posiadał minimum globalne;
 - szerokości przedziału całkowania $\langle A, B \rangle$;
 - liczby punktów siatki n .

Jeśli dostępne są wyniki analityczne, należy porównać je z rezultatami numerycznymi. Otrzymane należy wyniki przedstawić w postaci odpowiednich wykresów i tabel, przeanalizować je i sformułować wnioski.

Lista potencjałów jednowymiarowych

Należy rozważyć dowolne dodatnie wartości parametrów C , c , D , d , chyba że zostały one zadane.

1. Studnia trójkątna

$$V(X) = \begin{cases} 10^n V_c & (x < 0) \\ V_c C X & (x \geq 0). \end{cases}$$

2. Studnia trójkątno-paraboliczna

$$V(X) = \begin{cases} V_c C |X| & (x < 0) \\ V_c D X^2 & (x \geq 0). \end{cases}$$

3. Studnia półparaboliczna

$$V(X) = \begin{cases} 10^n V_c & (x < 0) \\ V_c C X^2 & (x \geq 0). \end{cases}$$

4. Niesymetryczna studnia liniowa

$$V(X) = \begin{cases} V_c C |X| & (x < 0) \\ V_c D X & (x \geq 0). \end{cases}$$

5. Oscylator asymetryczny

$$V(X) = \begin{cases} V_c C X^2 & (x < 0) \\ V_c D X^2 & (x \geq 0). \end{cases}$$

6. Symetryczna studnia prostokątna

$$V(X) = \begin{cases} -V_c & (|X| < C) \\ 0 & (|X| \geq C). \end{cases}$$

7. Niesymetryczna studnia prostokątna

$$V(X) = \begin{cases} D V_c & (X \leq -C) \\ -V_c & (|X| < C) \\ 0 & (X \geq C). \end{cases}$$

8. Potencjał Morse'a

$$V(X) = V_c (e^{-2CX} - 2e^{-CX});$$

wartości dokładne [30]

$$\mathcal{E}_n = -V_c \left[1 - (n + \frac{1}{2}) \frac{\hbar C}{\sqrt{2MV_c}} \right]^2 \quad (n = 0, \dots, n_{\max}),$$

gdzie $n_{\max} = \left\lceil \frac{\sqrt{2MV_c}}{\hbar C} - \frac{3}{2} \right\rceil$; jeśli $\sqrt{2MV_c} < \frac{1}{2}\hbar C$, nie ma stanów związanych.

9. Potencjał Tellera

$$V(X) = \frac{-V_c}{\cosh^2(X/C)};$$

wartości dokładne [30]

$$\mathcal{E}_n = -\frac{\hbar^2 C^2}{8M} \left[\sqrt{1 + \frac{8MV_c}{\hbar^2 C^2}} - (2n + 1) \right] \quad (n = 0, \dots, n_{\max}),$$

gdzie $n_{\max} = \left\lceil \frac{1}{2} \sqrt{1 + \frac{8MV_c}{\hbar^2 C^2}} - \frac{3}{2} \right\rceil$.

10. Potencjał Tody [144]

$$V(X) = V_c (e^{-CX} + DX).$$

11. Potencjał wykładniczy

$$V(X) = -V_c e^{-|X|/C}.$$

12. Potencjał Gaussa

$$V(X) = -V_c e^{-CX^2}.$$

13. Potencjał Lenarda–Jonesa

$$V(X) = V_c [(C/X)^{12} - (D/X)^6] \quad (X > 0).$$

14. Potencjał Kratzera

$$V(X) = V_c \left(\frac{C}{X} - \frac{X}{C} \right)^2 \quad (X > 0);$$

wartości dokładne [145]

$$\mathcal{E}_n = \hbar\Omega \left[n + \frac{1}{2} + \frac{1}{4} \left(\sqrt{\frac{M\Omega C^2}{\hbar}} - 1 - \sqrt{\frac{M\Omega C^2}{\hbar}} \right) \right] \quad (n = 0, 1, \dots),$$

gdzie $\Omega = \sqrt{8V_c/(MC^2)}$.

15. Potencjał Pöschla

$$V(x) = V_c \left(\frac{1}{\cos^2 CX} + \frac{d}{\sin^2 CX} \right), \quad 0 < X < \frac{\pi}{2C}, \quad d \in \{\frac{1}{2}, 1, 2\};$$

wartości dokładne [145]

$$\mathcal{E}_n = \frac{\hbar^2 C^2}{2M} (2\lambda + 2\nu + 2n)^2 \quad (n = 0, 1, \dots),$$

gdzie liczby λ i ν spełniają równania:

$$2\lambda(2\lambda - 1) = \frac{2MV_c}{\hbar^2 C^2}, \quad 2\nu(2\nu - 1) = d \frac{2MV_c}{\hbar^2 C^2}.$$

16. Potencjał

$$V(X) = V_c \operatorname{ctg}^2 \frac{\pi X}{C}, \quad 0 < X < C;$$

wartości dokładne [145]

$$\mathcal{E}_n = \frac{\hbar^2 \pi^2}{2MC^2} (n^2 - 4\lambda n + 2\lambda) \quad (n = 1, 2, \dots), \quad \lambda = \frac{1}{4} \left[1 - \sqrt{1 + \frac{8MC^2 V_c}{\hbar^2 \pi^2}} \right].$$

17. Potencjał Konwenta [67–70]

$$V(X) = V_c (c \cosh DX - 1)^2, \quad c \in \{\frac{1}{2}, 1, 5\};$$

w przypadku tego potencjału rozwiązania analityczne można wyznaczyć tylko dla pewnych szczególnych wartości parametrów, np. dla $V_c = 9\hbar^2 D^2/(8M)$ i $c \leq 1$ trzy najniższe poziomy energetyczne określają wzory

$$\mathcal{E}_{0,2} = \frac{1}{9} V_c \left[7 + 9c^2 \mp 4\sqrt{\frac{1}{4} + 9c^2} \right], \quad \mathcal{E}_1 = \frac{1}{9} V_c (5 + 9c^2).$$

18. Studnia podwójna

$$V(X) = V_c (-\frac{1}{2}CX^2 + \frac{1}{4}DX^4);$$

przyjąć różne wartości parametru $C^{1/2}/D^{1/4}$.

19. Studnia podwójna

$$V(X) = V_c C(|X| - D)^2;$$

z badać przypadki $D \rightarrow 0$ i $D \rightarrow \infty$.

20. Oscylator anharmoniczny

$$V(X) = V_c(CX^2 + DX^4 + EX^6 + FX^8);$$

przyjąć różne wartości parametrów C, D, E, F .

A.2. Zagadnienia radialne

Wyznaczyć numerycznie podaną liczbę k energii i funkcji własnych dla kulistosymetrycznego potencjału wybranego z poniższej listy.

Lista potencjałów trójwymiarowych

1. Studnia prostokątna

$$V(R) = \begin{cases} -V_c & (R < C) \\ 0 & (R \geq C) \end{cases} \quad k = 5,$$

2. Potencjał Morse'a

$$V(R) = V_c \left[\exp\left(-2c\frac{R-D}{D}\right) - 2 \exp\left(-c\frac{R-D}{D}\right) \right], \quad k = 10;$$

wartości dokładne [35]

$$\mathcal{E}_n = -V_c + \frac{\hbar c}{D} \sqrt{\frac{2V_c}{M}} \left(n + \frac{1}{2}\right) - \frac{\hbar^2 c^2}{2MD^2} \left(n + \frac{1}{2}\right)^2 \quad (n = 0, 1, \dots).$$

3. Kulistosymetryczny potencjał Kratzera

$$V(R) = 2V_c \left(\frac{1}{2} \frac{C^2}{R^2} - \frac{C}{R} \right), \quad k = 4;$$

wartości dokładne [35]

$$\mathcal{E}_n = -\frac{2MC^2V_c^2}{\hbar^2} \left[n + \frac{1}{2} + \sqrt{\left(l + \frac{1}{2}\right)^2 + 2MC^2V_c/\hbar^2} \right]^{-2} \quad (n = 0, 1, \dots).$$

4. Potencjał Hulthena (
- $l = 0$
-)

$$V(R) = -V_c \frac{e^{-R/C}}{1 - e^{-R/C}}, \quad k = 8;$$

wartości dokładne [35]

$$\mathcal{E}_n = -V_c \left(\frac{\beta^2 - n^2}{2n\beta} \right)^2 \quad (n = 1, 2, \dots), \quad \beta = \sqrt{\frac{2MC^2V_c}{\hbar^2}}.$$

5. Potencjał

$$V(R) = \frac{C}{R^2} + DR^2, \quad k = 4;$$

wartości dokładne [30]

$$\mathcal{E}_n = \hbar \sqrt{D/2M} \left[4n + 2 + \sqrt{(2l+1)^2 + 8mC/\hbar^2} \right] \quad (n = 0, 1, \dots).$$

6. Potencjał

$$V(R) = \frac{C}{R^2} - \frac{D}{R}, \quad k = 3;$$

wartości dokładne [30]

$$\mathcal{E}_n = -\frac{2MD^2}{\hbar^2} \left[2n + 1 + \sqrt{(2l+1)^2 + 8mC/\hbar^2} \right]^{-2} \quad (n = 0, 1, \dots).$$

A.3. Zagadnienie jednowymiarowe — porównanie metod

Wyznaczyć numerycznie 10 pierwszych energii i funkcji własnych kwantowego oscylatora harmonicznego o masie M z potencjałem $V(X) = \frac{1}{2}KX^2 = \frac{1}{2}M\Omega^2X^2$ za pomocą metody:

1. Trójpunktowej (rozdział 3);
2. Pięciopunktowej (podrozdział 5.4);
3. Lindberga;
4. Guardioli–Rosa z aproksymantą Padégo $s^2D^2[3/0]$;
5. Guardioli–Rosa z aproksymantą Padégo $s^2D^2[4/0]$;
6. Guardioli–Rosa z aproksymantą Padégo $s^2D^2[5/0]$;
7. Guardioli–Rosa z aproksymantą Padégo $s^2D^2[2/2]$;
8. Guardioli–Rosa z aproksymantą Padégo $s^2D^2[3/2]$.

Otrzymane wyniki porównać z analitycznymi. Porównać metody pod względem efektywności (dokładności, szybkości, wymaganego rozmiaru pamięci operacyjnej).

A.4. Redukcja Householdera — porównanie podprogramów bibliotecznych

Zredukować gęstą macierz symetryczną do postaci trójdzielnej za pomocą transformacji Householdera. Zbadać zależność czasu wykonania zadania od rozmiaru macierzy. Wykorzystać różne procedury biblioteczne (SSYTRD z LAPACK-a, TRED1, TRED2 z EISPACK-a, TRED2 z *Numerical Recipes*, ewentualnie inne).

Przeanalizować przypadki, gdy akumulowane są elementarne transformacje podobieństwa (potrzebne później do wyznaczania wektorów własnych wejściowej macierzy gęstej), oraz gdy krok ten jest pomijany.

A.5. Symetryczne gęste zagadnienie własne — porównanie metod i podprogramów bibliotecznych

Zbadać zależność czasu rozwiązywania gęstego symetrycznej zagadnienia własnego od jego rozmiaru za pomocą następujących metod:

- transformacji Householdera i iteracji QR/QL — podprogramy
 - SSYTRD + SSTEQR (LAPACK),
 - TRED2 + IMTQL2 (EISPACK),
 - TRED2 + TQLI (*Numerical Recipes*);
- transformacji Householdera i bisekcji z iteracjami odwrotnymi — podprogramy
 - SSYTRD + SSTEBS + SSTEIN (LAPACK),
 - TRED2 + TSTURM (EISPACK);
- transformacji Householdera i bisekcji z metodą DWSZ — podprogramy
 - SSYTRD + SSTEBS (LAPACK) + DWSZ (patrz dodatek F),
 - TRED2 + TRIDIB (EISPACK) + DWSZ;
- transformacji Householdera i metody „dziel i rządź” — podprogram SSYEVD (LAPACK).
- metody Jacobiego — podprogram JACOBI (*Numerical Recipes*).

Można użyć także podprogramów z innych źródeł, również własnych.

A.6. Trójdzielne zagadnienie własne — porównanie różnych algorytmów i podprogramów bibliotecznych

Wyznaczyć zależność czasu obliczania wszystkich wartości własnych trójdzielnej macierzy symetrycznej od jej rozmiaru za pomocą algorytmów QR/QL:

- niewymiernego z jawnym przesunięciem widma — podprogram TQL1 (EISPACK);
- niewymiernego z niejawnym przesunięciem widma — podprogramy
 - IMTQL1 (EISPACK),
 - SSTEQR (LAPACK),
 - TQLI (*Numerical Recipes*);
- wymiernego — podprogramy:
 - TQLRAT (EISPACK),
 - SSTERF (LAPACK).

A.7. Przypadek prawie zwyrodniały (potencjał okresowy)

Wyznaczyć energie własne z przedziału $\langle -V_C, V_C \rangle$ i odpowiadające im funkcje własne cząstki kwantowej o masie m poruszającej się w potencjale

$$V(X) = \begin{cases} V_c \cos(AX) & (|AX| < 6\pi) \\ \infty & (|AX| \geq 6\pi). \end{cases}$$

Rozwiązać odpowiednie algebraiczne zagadnienie własne, wykorzystując stabilną wersję metody „dziel i rządź” (podprogram SSTEDC z LAPACK-a) oraz metody: bisekcji i DWSZ. Zweryfikować numerycznie ortogonalność obliczonych wektorów własnych w obu przypadkach. Jak otrzymane wyniki zależą od parametru skali $\alpha = 2mV_c/(A^2\hbar^2)$?

Zadanie dodatkowe: zbadać ortogonalność wektorów własnych wyznaczonych niestabilną wersją metody „dziel i rządź” (wymaga samodzielnego zaprogramowania lub przeróbki odpowiedniego podprogramu bibliotecznego).

A.8. Przypadek prawie zwyrodniały (potencjał Konwenta)

Wyznaczyć energie własne dwóch najniższych położonych poziomów energetycznych i odpowiadające im funkcje własne cząstki kwantowej o masie m poruszającej się w potencjale Konwenta (2.20) dla $c = \frac{1}{2}$ oraz $\alpha = 1$ i $\alpha = 10$. Odpowiednie algebraiczne zagadnienie własne rozwiązać za pomocą stabilnej wersji metody „dziel i rządź” (podprogram SSTEDC z LAPACK-a) oraz bisekcji i metody DWSZ. Zweryfikować numerycznie ortogonalność obliczonych wektorów własnych w obu przypadkach.

Zadanie dodatkowe — jak w poprzednim projekcie.

A.9. Trójdzielny układ równań

Korzystając z wyników prac [93] zaprogramować w dowolnym języku algorytm wyznaczania macierzy odwrotnej do symetrycznej macierzy trójdzielnej. Wykonać testy dla macierzy o wyrazach losowych z przedziału $\langle -1, 1 \rangle$. Wyniki porównać z rezultatami otrzymanymi przy pomocy standardowych procedur biblioteki LAPACK. Porównać czasy wykonywania i dokładność wyników (tj. wyznaczyć $\mathbf{T}\mathbf{T}^{-1} - \mathbf{1}$). Zbadać efektywność zastosowania zaprogramowanej procedury do rozwiązywania trójdzielnego układu równań.

Elementy algebry liniowej

W tym dodatku przytaczamy najważniejsze definicje i twierdzenia matematyczne, którymi posługujemy się przy formułowaniu algorytmów pozwalających numerycznie rozwiązywać algebraiczne zagadnienie własne $\mathbf{Ax} = \lambda \mathbf{x}$.

Na wstępie wprowadzimy definicje normy wektora, wartości własnej i wektora własnego, promienia spektralnego macierzy oraz wybranych norm macierzowych.

Definicja B-1. Normą $\|\mathbf{x}\|_2$ wektora $\mathbf{x} \in \mathbb{R}^n$ nazywamy liczbę

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n x_i^2} = \sqrt{\mathbf{x} \cdot \mathbf{x}},$$

gdzie liczby x_i są składowymi wektora \mathbf{x} , a $\mathbf{x} \cdot \mathbf{x}$ oznacza iloczyn skalarny (B-16).

Definicja B-2. Wartością własną i wektorem własnym macierzy $\mathbf{A} \in \mathbb{R}^{n \times n}$ nazywamy odpowiednio liczbę λ i wektor $\mathbf{x} \neq \mathbf{0}$, spełniające równość

$$\mathbf{Ax} = \lambda \mathbf{x}.$$

Jeżeli $\mathbf{A} = \mathbf{A}^T$, ma ona dokładnie n wartości własnych λ_i i dokładnie n niezależnych liniowo wektorów własnych \mathbf{x}_i . Wszystkie wartości i wektory własne są rzeczywiste, a wektory własne odpowiadające różnym wartościom własnym są ortogonalne, tj. $\mathbf{x}_i \cdot \mathbf{x}_j = 0$, jeśli $\lambda_i \neq \lambda_j$.

Definicja B-3. Promieniem spektralnym macierzy $\mathbf{A} \in \mathbb{R}^{n \times n}$ nazywamy liczbę

$$\varrho(\mathbf{A}) = \max_{i=1}^n |\lambda_i|,$$

gdzie λ_i są wartościami własnymi \mathbf{A} .

Definicja B-4. Normą $\|\mathbf{A}\|_1$ macierzy $\mathbf{A} \in \mathbb{R}^{n \times n}$ nazywamy liczbę

$$\|\mathbf{A}\|_1 = \max_{j=1}^n \sum_{i=1}^n |A_{i,j}|,$$

gdzie $A_{i,j}$ są wyrazami j -tej kolumny \mathbf{A} .

Definicja B-5. Normą $\|\mathbf{A}\|_\infty$ macierzy $\mathbf{A} \in \mathbb{R}^{n \times n}$ nazywamy liczbę

$$\|\mathbf{A}\|_\infty = \max_{i=1}^n \sum_{j=1}^n |A_{i,j}|,$$

gdzie $A_{i,j}$ są wyrazami i -tego wiersza \mathbf{A} .

Definicja B-6. Normę $\|\mathbf{A}\|_2$ macierzy $\mathbf{A} \in \mathbb{R}^{n \times n}$ nazywamy liczbę

$$\|\mathbf{A}\|_2 = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{Ax}\|_2}{\|\mathbf{x}\|_2};$$

Można pokazać, że $\|\mathbf{A}\|_2 = \sqrt{\varrho(\mathbf{A}^T \mathbf{A})} = \sqrt{\lambda_{\max}(\mathbf{A}^T \mathbf{A})}$, gdzie $\lambda_{\max}(\mathbf{A}^T \mathbf{A})$ — największa wartość własna dodatnio określonej macierzy $\mathbf{A}^T \mathbf{A}$.

Przytoczymy teraz twierdzenia, pozwalające szacować położenie widma macierzy na osi liczb rzeczywistych.

Twierdzenie B-7. Dla macierzy \mathbf{A} spełniona jest nierówność $\varrho(\mathbf{A}) \leq \|\mathbf{A}\|_p$, gdzie $p = 1, \infty$.

Twierdzenie B-8. (Gershgorina) Wszystkie wartości własne rzeczywistej macierzy symetrycznej¹⁾ \mathbf{A} leżą w zbiorze $\mathbb{S} = \bigcup_{i=1}^n \mathbb{P}_i$, gdzie

$$\mathbb{P}_i = \langle A_{i,i} - r_i, A_{i,i} + r_i \rangle, \quad r_i = \sum_{\substack{j=1 \\ j \neq i}}^n |A_{i,j}|.$$

Jeżeli k przedziałów \mathbb{P}_i tworzy zbiór rozłączny z pozostałymi, to w zbiorze tym leży dokładnie k wartości własnych macierzy \mathbf{A} .

Zdefiniujemy transformację macierzy przez podobieństwo oraz podamy niektóre jej własności.

Definicja B-9. Niech \mathbf{S} będzie niesobliwą macierzą kwadratową. *Przekształceniem podobieństwa* nazywamy odwzorowanie

$$\mathbf{A} \mapsto \mathbf{B} = \mathbf{S}^{-1} \mathbf{A} \mathbf{S}.$$

Definicja B-10. Macierze \mathbf{A} i \mathbf{B} nazywamy *podobnymi*, jeśli istnieje nieosobliwa macierz \mathbf{S} taka, że $\mathbf{B} = \mathbf{S}^{-1} \mathbf{A} \mathbf{S}$.

Twierdzenie B-11. Macierze podobne mają takie same wartości własne. Jeśli wektor \mathbf{q} jest wektorem własnym macierzy \mathbf{A} odpowiadającym jej wartości własnej λ , to wektor $\mathbf{p} = \mathbf{S}^{-1} \mathbf{q}$ jest wektorem własnym macierzy \mathbf{B} odpowiadającym również wartości własnej λ .

Dowód. Jeśli $\mathbf{A} \mathbf{q} = \lambda \mathbf{q}$, to $\mathbf{S}^{-1} \mathbf{A} \mathbf{q} = \lambda \mathbf{S}^{-1} \mathbf{q}$. Ponieważ $\mathbf{p} = \mathbf{S}^{-1} \mathbf{q}$, to $\mathbf{q} = \mathbf{S} \mathbf{p}$ i mamy $\mathbf{S}^{-1} \mathbf{A} \mathbf{S} \mathbf{p} = \mathbf{B} \mathbf{p} = \lambda \mathbf{S}^{-1} \mathbf{S} \mathbf{p} = \lambda \mathbf{p}$. \square

Do klasy przekształceń podobieństwa należą w szczególności:

1. Przekształcenia *ortogonalne* z rzeczywistymi macierzami \mathbf{S} , które spełniają relację

$$\mathbf{S}^T = \mathbf{S}^{-1}$$

gdzie \mathbf{S}^T oznacza macierz transponowaną. Macierze takie przyjęto oznaczać w literaturze za pomocą symbolu \mathbf{Q} . Dla macierzy ortogonalnej wyznacznik $\det \mathbf{Q} = \pm 1$; jeśli $\det \mathbf{Q} = 1$, macierz tę nazywamy *właściwą*.

¹⁾ Istnieje również ogólniejsza wersja twierdzenia Gershgorina, dotycząca macierzy niesymetrycznych [21, 25].

2. Przekształcenia *unitarne*, z zespolonymi macierzami \mathbf{S} , które spełniają relację

$$\mathbf{S}^{-1} = \mathbf{S}^H$$

gdzie $\mathbf{S}^H = (\mathbf{S}^T)^*$ oznacza macierz sprzężoną po hermitowsku. Macierze takie przyjęto oznaczać symbolem \mathbf{U} .

Twierdzenie B-12. Ortogonalne przekształcenia podobieństwa zachowują własność symetryczności macierzy.

Dowód. Transponujemy macierz \mathbf{B} : $\mathbf{B}^T = (\mathbf{Q}^T \mathbf{A} \mathbf{Q})^T = \mathbf{Q}^T \mathbf{A}^T \mathbf{Q} = \mathbf{B}$. \square

Podobnie unitarne przekształcenia podobieństwa zachowują własność hermitowskości macierzy ($\mathbf{A}^H = \mathbf{A}$).

W podręczniku odwołujemy się często do twierdzenia Sylwestera, nazywanego w matematyce *twierdzeniem o bezwładności macierzy symetrycznej* lub *formy kwadratowej*, oraz związanego z nim twierdzenia Martina–Deana.

Definicja B-13. *Bezwładności²⁾* macierzy symetrycznej \mathbf{A} nazywamy trójkę liczb całkowitych (η, ν, ξ) , gdzie η , ν i ξ są liczbami odpowiednio ujemnych, zerowych i dodatnich wartości własnych macierzy \mathbf{A} .

Twierdzenie B-14. (Sylwestera) Niech \mathbf{A} będzie macierzą symetryczną, a \mathbf{X} — macierzą nieosobliwą ($\det \mathbf{X} \neq 0$). Wtedy macierze \mathbf{A} i $\mathbf{X}^T \mathbf{A} \mathbf{X}$ mają tę samą bezwładność.

Dowód twierdzenia Sylwestera można znaleźć w podręczniku [50].

Do sformułowania twierdzenia Martina–Deana potrzebne nam będzie jeszcze jedno pojęcie:

Definicja B-15. *Iloczynem zewnętrznym* wektorów $\mathbf{v} \in \mathbb{R}^n$ i $\mathbf{w} \in \mathbb{R}^m$ w wybranej bazie nazywamy macierz $n \times m$

$$\mathbf{v} \otimes \mathbf{w} = |\mathbf{v}\rangle\langle\mathbf{w}| = \mathbf{v}\mathbf{w}^T = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} (w_1, \dots, w_m) = \begin{pmatrix} v_1 w_1 & \cdots & v_1 w_m \\ \vdots & \ddots & \vdots \\ v_n w_1 & \cdots & v_n w_m \end{pmatrix}.$$

Podkreślmy w tym miejscu, że iloczynu zewnętrznego nie należy mylić z iloczynem skalarnym:

Definicja B-16. *Iloczynem skalarnym* wektorów $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ nazywamy liczbę

$$\mathbf{v} \cdot \mathbf{w} = \langle \mathbf{v} | \mathbf{w} \rangle = \mathbf{v}^T \mathbf{w} = (v_1, \dots, v_n) \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \sum_{i=1}^n v_i w_i.$$

²⁾ W języku angielskim stosowany jest termin *inertia*.

Twierdzenie B-17. (Martina–Deana) Niech $\mathbf{A} \in \mathbb{R}^{n \times n}$ będzie symetryczną macierzą blokowo-trójdziagonalną postaci

$$\mathbf{A} = \begin{pmatrix} \mathbf{D}_1 & \mathbf{E}_1^T & & \\ \mathbf{E}_1 & \ddots & \ddots & \\ & \ddots & \ddots & \mathbf{E}_{m-1}^T \\ & & \mathbf{E}_{m-1} & \mathbf{D}_m \end{pmatrix},$$

gdzie $\mathbf{D}_i \in \mathbb{R}^{l_i \times l_i}$ są macierzami symetrycznymi, a $\mathbf{E}_i \in \mathbb{R}^{l_{i+1} \times l_i}$; rozmiary klatek spełniają zależność $\sum_{i=1}^m l_i = n$. Niech $\eta(\mathbf{X})$ oznacza liczbę ujemnych wartości własnych macierzy \mathbf{X} . Wtedy

$$\eta(\mathbf{A} - z\mathbf{1}) = \sum_{i=1}^m \eta(\mathbf{U}_i),$$

gdzie (\mathbf{U}_i) jest ciągiem macierzy postaci

$$\mathbf{U}_1 = \mathbf{D}_1 - z\mathbf{1}_1, \quad \mathbf{U}_i = \mathbf{D}_i - z\mathbf{1}_i - \mathbf{E}_{i-1}\mathbf{U}_{i-1}^{-1}\mathbf{E}_{i-1}^T \quad (i = 2, \dots, m), \quad (\text{B.1})$$

a $\mathbf{1}_i \in \mathbb{R}^{l_i \times l_i}$ jest macierzą jednostkową stosownego rozmiaru.

Dowód powyższego twierdzenia podano w pracy [100].

Zastosujmy twierdzenie Martina–Deana do dowolnej (gęstej) rzeczywistej macierzy symetrycznej \mathbf{A} . W tym celu zajmijmy się macierzą $\mathbf{A}_1 = \mathbf{A} - z\mathbf{1}$, którą dzielimy na cztery klatki:

$$\mathbf{A}_1 = \begin{pmatrix} u_1 & \mathbf{y}_1^T \\ \mathbf{y}_1 & \mathbf{D}_1 \end{pmatrix}, \quad (\text{B.2})$$

gdzie $u_1 = A_{1,1} - z$ (jest to skalar) oraz

$$\mathbf{y}_1 = \begin{pmatrix} A_{2,1} \\ \vdots \\ A_{n,1} \end{pmatrix}, \quad \mathbf{D}_1 = \begin{pmatrix} A_{2,2} - z & \cdots & A_{2,n} \\ \vdots & \ddots & \vdots \\ A_{n,2} & \cdots & A_{n,n} - z \end{pmatrix};$$

\mathbf{D}_1 jest macierzą $(n-1) \times (n-1)$, a \mathbf{y}_1 — wektorem o $n-1$ współrzędnych.

Na podstawie twierdzenia Martina–Deana otrzymujemy

$$\eta(\mathbf{A}_1) = \eta(u_1) + \eta(\mathbf{D}_1 - \mathbf{y}_1 \otimes \mathbf{y}_1 / u_1) = \eta(u_1) + \eta(\mathbf{A}_2), \quad (\text{B.3})$$

gdzie wynikiem iloczynu zewnętrznego $\mathbf{y}_1 \otimes \mathbf{y}_1$ jest macierz $(n-1) \times (n-1)$. Jak widzimy, liczba $\eta(\mathbf{A}_1)$ jest określona za pomocą znaku składowa u_1 oraz liczby ujemnych wartości własnych macierzy $\mathbf{A}_2 = \mathbf{D}_1 - \mathbf{y}_1 \otimes \mathbf{y}_1 / u_1$ rozmiaru $(n-1) \times (n-1)$.

W drugim kroku dzielimy na klatki macierz \mathbf{A}_2 :

$$\mathbf{A}_2 = \begin{pmatrix} u_2 & \mathbf{y}_2^T \\ \mathbf{y}_2 & \mathbf{D}_2 \end{pmatrix}, \quad (\text{B.4})$$

gdzie u_2 jest wyrazem stojącym w pierwszym wierszu i pierwszej kolumnie macierzy \mathbf{A}_2 , \mathbf{y} jest wektorem o współrzędnych określonych przez pozadiagonalne wyrazy pierwszej kolumny \mathbf{A}_2 , zaś \mathbf{D}_2 — macierzą rozmiaru $(n-2) \times (n-2)$, otrzymaną

z \mathbf{A}_2 przez skreślenie pierwszego wiersza i pierwszej kolumny. Ponowne zastosowanie twierdzenia B-17 daje

$$\eta(\mathbf{A}_2) = \eta(u_2) + \eta(\mathbf{D}_2 - \mathbf{y}_2 \otimes \mathbf{y}_2 / u_2) = \eta(u_2) + \eta(\mathbf{A}_3), \quad (\text{B.5})$$

W kolejnych krokach (dla $i = 3, \dots, n-1$) powtarzamy powyższą procedurę:

$$\mathbf{A}_i = \mathbf{D}_{i-1} - \frac{\mathbf{y}_{i-1} \otimes \mathbf{y}_{i-1}}{u_{i-1}} = \begin{pmatrix} u_i & \mathbf{y}_i^T \\ \mathbf{y}_i & \mathbf{D}_i \end{pmatrix} \quad (\text{B.6})$$

oraz

$$\eta(\mathbf{A}_i) = \eta(u_i) + \eta(\mathbf{A}_{i+1}), \quad (\text{B.7})$$

gdzie \mathbf{A}_{i+1} — macierz rozmiaru $(n-i) \times (n-i)$; w ostatnim kroku \mathbf{A}_n jest skalar, co kończy postępowanie.

Równanie (B.7) jest podstawowym wynikiem umożliwiającym efektywne wyznaczanie wartości własnych macierzy. Pokazuje ono, że liczba $l_{\text{ww}}(z)$ wartości własnych symetrycznej macierzy rzeczywistej \mathbf{A} mniejszych od danej liczby rzeczywistej z jest określona za pomocą znaków n wielkości skalar, u_i obliczanych rekurencyjnie (patrz (B.2)–(B.6)).

W naszych rozważaniach najczęściej mamy do czynienia z macierzą trójdziagonalną postaci

$$\mathbf{T} = \begin{pmatrix} d_1 & e_1 & & & \\ e_1 & \ddots & \ddots & & \\ & \ddots & \ddots & e_{n-1} & \\ & & e_{n-1} & d_n & \end{pmatrix}, \quad (\text{B.8})$$

dla której słuszne jest następujące twierdzenie:

Twierdzenie B-18. Jeśli wszystkie wyrazy e_2, e_3, \dots, e_n macierzy \mathbf{T} są niezerowe, to jej wartości własne są jednokrotne³⁾ (w fizyce używa się terminu *niezdegenerowane*).

Ostatnie twierdzenie ma istotne znaczenie w obliczeniach numerycznych, które prowadzone są zawsze ze *skończoną* dokładnością (określoną liczbą σ_m , zwaną *epsilon maszynowym*). Może się tak zdarzyć, że wyniki obliczeń numerycznych będą wskazywały na degenerację (tj. równość z dokładnością σ_m) dwóch lub większej liczby wartości własnych rozpatrywanego AZW z macierzą (B.8), mimo że wszystkie $e_i \neq 0$. Wtedy degeneracja ta ma charakter *artefaktu komputerowego*.

Należy także zauważyć, że jeśli macierz \mathbf{T} otrzymaliśmy jako numeryczny wynik przekształcenia ortogonalnego $\mathbf{T} = \mathbf{Q}^T \mathbf{A} \mathbf{Q}$ pewnej macierzy symetrycznej \mathbf{A} , to wartości własne \mathbf{A} są równe wartościom własnym \mathbf{T} tylko w przybliżeniu. W szczególności wynik $e_i = 0$ może być skutkiem błędów zaokrągleń i degeneracja wartości własnych \mathbf{T} niekoniecznie oznacza degenerację wartości własnych \mathbf{A} . I odwrotnie, możemy otrzymać macierz \mathbf{T} z wyrazami $e_i \neq 0$ (z dokładnością maszynową), nawet jeśli \mathbf{A} ma wielokrotne wartości własne.

³⁾ Ich krotności algebraiczne oraz geometryczne [21] wynoszą 1.

Podstawą wszystkich algorytmów diagonalizacji macierzy jest następujące twierdzenie:

Twierdzenie B-19. (Schura) Dla każdej macierzy hermitowskiej \mathbf{A} rozmiaru $n \times n$ istnieje macierz unitarna \mathbf{U} spełniająca równość

$$\mathbf{U}^H \mathbf{A} \mathbf{U} = \text{diag}(\lambda_1, \dots, \lambda_n),$$

gdzie liczby rzeczywiste λ_i są wartościami własnymi macierzy \mathbf{A} , a i -ta kolumna macierzy \mathbf{U} jest wektorem własnym odpowiadającym λ_i . Macierz \mathbf{A} ma n liniowo niezależnych, różnych wektorów własnych [21, 50].

Na mocy tego twierdzenia każdą macierz hermitowską można zdiagonalizować za pomocą przekształcenia unitarnego podobieństwa.

Wniosek B-20. Dla każdej rzeczywistej macierzy symetrycznej \mathbf{A} rozmiaru $n \times n$ istnieje macierz ortogonalna \mathbf{Q} (rzeczywista) spełniająca równość

$$\mathbf{Q}^T \mathbf{A} \mathbf{Q} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n),$$

gdzie liczby rzeczywiste λ_i są wartościami własnymi macierzy \mathbf{A} , a i -ta kolumna macierzy \mathbf{Q} jest wektorem własnym odpowiadającym λ_i . Macierz \mathbf{A} ma n liniowo niezależnych, różnych wektorów własnych.

Zatem każdą symetryczną macierz rzeczywistą można zdiagonalizować przy pomocy ortogonalnego przekształcenia podobieństwa.

Przytoczymy jeszcze (bez dowodu) twierdzenie [25 (r. 6), 50 (r. 3)], będące podstawą metod QR i QL.

Twierdzenie B-21. Każdą rzeczywistą macierz prostokątną $\mathbf{A} \in \mathbb{R}^{m \times n}$ ($m \geq n$), mającą $k \leq n$ liniowo niezależnych kolumn, można przedstawić w postaci $\mathbf{A} = \mathbf{Q}\mathbf{R}$, gdzie $\mathbf{Q} \in \mathbb{R}^{m \times k}$ jest macierzą ortogonalną, a $\mathbf{R} \in \mathbb{R}^{k \times n}$ jest macierzą trapezoidalną górną.

Rozkład ten można otrzymać stosując metodę Gramma–Schmidta (może być ona numerycznie niestabilna, ale na niej najprościej oprzeć dowód [21 (r. 4), 25 (r. 6), 45 (r. 4), 50 (r. 3)], Householdera [21 (r. 4 i 6), 25 (r. 6), 45 (r. 4 i 8)] lub Givensa [21 (r. 6), 25 (r. 6), 45 (r. 4 i 8)]).

Twierdzenie B-22. Jeżeli macierz kwadratowa $\mathbf{A} \in \mathbb{R}^{n \times n}$ jest nieosobliwa, to istnieje *jednoznaczny* (z dokładnością do znaków) rozkład $\mathbf{A} = \mathbf{Q}\mathbf{R}$.

Wyznaczanie wektorów własnych w metodzie „dziel i rządź” (podrozdział 6.4) jest oparte na twierdzeniu, które przytaczamy poniżej.

Twierdzenie B-23. (Löwnera) Niech $\mathbf{D} = \text{diag}(d_1, \dots, d_n)$ będzie macierzą diagonalną i $d_n < \dots < d_1$, a liczby α_i spełniają nierówności

$$d_n < \alpha_n < \dots < d_{i+1} < \alpha_{i+1} < d_i < \alpha_i < \dots < d_1 < \alpha_1.$$

Wtedy istnieje wektor \mathbf{w} taki, że α_i są dokładnymi wartościami własnymi macierzy $\mathbf{M} = \mathbf{D} + \mathbf{w} \otimes \mathbf{w}^T$ i

$$w_i^2 = \prod_{j=1}^n (\alpha_j - d_i) \bigg/ \prod_{\substack{j=1 \\ j \neq i}}^n (d_j - d_i).$$

Dowód. Wykorzystamy twierdzenie B-24. Wtedy wielomian charakterystyczny macierzy \mathbf{M} przyjmuje postać

$$\begin{aligned}\det(\mathbf{M} - \lambda \mathbf{1}) &= \det(\mathbf{D} + \mathbf{w} \otimes \mathbf{w}^T - \lambda \mathbf{1}) = \\ &= \det(\mathbf{D} - \lambda \mathbf{1}) \det[\mathbf{1} + (\mathbf{D} - \lambda \mathbf{1})^{-1} \mathbf{w} \otimes \mathbf{w}^T] = \\ &= \prod_{j=1}^n (d_j - \lambda) \left[1 + \sum_{j=1}^n \frac{w_j^2}{d_j - \lambda} \right] = \\ &= \prod_{j=1}^n (d_j - \lambda) \left[1 + \sum_{\substack{j=1 \\ j \neq i}}^n \frac{w_j^2}{d_j - \lambda} \right] + w_i^2 \prod_{\substack{j=1 \\ j \neq i}}^n (d_j - \lambda).\end{aligned}$$

Położmy teraz $\lambda = d_i$ i porównajmy wynik z wielomianem charakterystycznym \mathbf{M} rozłożonym na czynniki:

$$\det(\mathbf{M} - \lambda \mathbf{1}) = \prod_{j=1}^n (\alpha_j - d_i) = w_i^2 \prod_{\substack{j=1 \\ j \neq i}}^n (d_j - d_i).$$

Stąd już łatwo otrzymujemy wartości w_i . Są one dodatnie ze względu na uporządkowanie liczb α_j i d_j . \square

Twierdzenie B-24. Niech $\mathbf{1} \in \mathbb{R}^{n \times n}$ będzie macierzą jednostkową, a $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ — dowolnymi wektorami. Wówczas

$$\det(\mathbf{1} + \mathbf{x} \otimes \mathbf{y}) = 1 + \mathbf{x} \cdot \mathbf{y}.$$

Dowód. Niech $\mathbf{x}' = \mathbf{x} / \|\mathbf{x}\|_2$ oraz $\mathbf{X} = (\mathbf{x}', \mathbf{X}')$ jest macierzą ortogonalną o pierwszej kolumnie \mathbf{x}' ($\mathbf{X}' \in \mathbb{R}^{n \times (n-1)}$). Wtedy

$$\det(\mathbf{1} + \mathbf{x} \otimes \mathbf{y}) = \det[\mathbf{X}^T (\mathbf{1} + \mathbf{x} \otimes \mathbf{y}) \mathbf{X}] = \det(\mathbf{1} + \mathbf{X}^T \mathbf{x} \mathbf{y}^T \mathbf{X}).$$

Ponieważ $\mathbf{X}^T \mathbf{x} = (\|\mathbf{x}\|_2, 0, \dots, 0)^T$ oraz $\mathbf{y}^T \mathbf{X} = (\mathbf{y}^T \mathbf{x}', \mathbf{y}^T \mathbf{X}')$, mamy

$$\mathbf{1} + (\mathbf{X}^T \mathbf{x})(\mathbf{y}^T \mathbf{X}) = \begin{pmatrix} 1 + \|\mathbf{x}\|_2 \mathbf{y}^T \mathbf{x}' & \|\mathbf{x}\|_2 \mathbf{y}^T \mathbf{X}' \\ \mathbf{0} & \mathbf{1} \end{pmatrix}.$$

Jest to macierz trójkątna i jej wyznacznik wynosi $1 + \|\mathbf{x}\|_2 \mathbf{y}^T \mathbf{x}' = 1 + \mathbf{x}^T \mathbf{y}$. \square

Aproksymacja pochodnych przy pomocy różnic skończonych

Za podręcznikiem [60] przedstawimy tu krótko zasady numerycznego różniczkowania, które stanowią podstawę naszych rozważań.

Niech funkcja rzeczywista $f(x)$ będzie określona na przedziale $\langle a, b \rangle$. Załóżmy, że znane są nam wartości tej funkcji $f_i = f(x_i)$ w równo oddalonych punktach

$$x_i = a + is \quad (i = 0, \dots, n+1), \quad s = \frac{b-a}{n+1}. \quad (\text{C.1})$$

Zbiór punktów (C.1) będziemy dalej nazywać jednowymiarową siatką. Zauważmy, że $x_i \in \langle a, b \rangle$ i $f(a) = f_0$, $f(b) = f_{n+1}$.

Podamy teraz wzory pozwalające przybliżać wartości pochodnych funkcji f za pomocą znanych wartości f_i .

Wzory dla pierwszej pochodnej

- formuła różniczkowania do przodu

$$f^{(\text{I})}(x_i) = f'_i = \frac{f_{i+1} - f_i}{s} + O(s); \quad (\text{C.2})$$

- różnica centralna

$$f^{(\text{I})}(x_i) = f'_i = \frac{f_{i+1} - f_{i-1}}{2s} + O(s^2); \quad (\text{C.3})$$

- formuła różniczkowania do przodu

$$f^{(\text{I})}(x_i) = f'_i = \frac{-f_{i+2} + 4f_{i+1} - 3f_i}{2s} + O(s^2); \quad (\text{C.4})$$

- różnica centralna

$$f^{(\text{I})}(x_i) = f'_i = \frac{-f_{i+2} + 8f_{i+1} - 8f_{i-1} + f_{i-2}}{12s} + O(s^4). \quad (\text{C.5})$$

Wzory dla drugiej pochodnej

- formuła różniczkowania do przodu

$$f^{(\text{II})}(x_i) = f''_i = \frac{f_{i+2} - 2f_{i+1} + f_i}{s^2} + O(s); \quad (\text{C.6})$$

- różnica centralna

$$f^{(\text{II})}(x_i) = f''_i = \frac{f_{i+1} - 2f_i + f_{i-1}}{s^2} + O(s^2); \quad (\text{C.7})$$

- formuła różniczkowania do przodu

$$f^{(\text{II})}(x_i) = f_i^{(\text{II})} = \frac{-f_{i+3} + 4f_{i+2} - 5f_{i+1} + 2f_i}{s^2} + O(s^2); \quad (\text{C.8})$$

- różnica centralna

$$f^{(\text{II})}(x_i) = f_i^{(\text{II})} = \frac{-f_{i+2} + 16f_{i+1} - 30f_i + 16f_{i-1} - f_{i-2}}{12s^2} + O(s^4); \quad (\text{C.9})$$

Wzory dla trzeciej pochodnej

- formuła różniczkowania do przodu

$$f^{(\text{III})}(x_i) = f_i^{(\text{III})} = \frac{f_{i+3} - 3f_{i+2} + 3f_{i+1} - f_i}{s^3} + O(s); \quad (\text{C.10})$$

- różnica centralna

$$f^{(\text{III})}(x_i) = f_i^{(\text{III})} = \frac{-f_{i+2} - 2f_{i+1} + 2f_{i-1} - f_{i-2}}{2s^3} + O(s^2). \quad (\text{C.11})$$

Wzory dla czwartej pochodnej

- formuła różniczkowania do przodu

$$f^{(\text{IV})}(x_i) = f_i^{(\text{IV})} = \frac{f_{i+4} - 4f_{i+3} + 6f_{i+2} - 4f_{i+1} + f_i}{s^4} + O(s); \quad (\text{C.12})$$

- różnica centralna

$$f^{(\text{IV})}(x_i) = f_i^{(\text{IV})} = \frac{f_{i+2} - 4f_{i+1} + 6f_i - 4f_{i-1} + f_{i-2}}{s^4} + O(s^2). \quad (\text{C.13})$$

Wyprowadzenie podanych wzorów polega na rozwinięciu w szereg Taylora funkcji $f(x)$ w odpowiednio wybranych punktach siatki (C.1):

$$\begin{aligned} f_{i+2} &= f_i + 2sf_i^{(\text{I})} + 2s^2f_i^{(\text{II})} + \frac{4}{3}s^3f_i^{(\text{III})} + \frac{2}{3}s^4f_i^{(\text{IV})} + O(s^5) \\ f_{i+1} &= f_i + sf_i^{(\text{I})} + \frac{1}{2}s^2f_i^{(\text{II})} + \frac{1}{6}s^3f_i^{(\text{III})} + \frac{1}{24}s^4f_i^{(\text{IV})} + O(s^5) \\ f_{i-1} &= f_i - sf_i^{(\text{I})} + \frac{1}{2}s^2f_i^{(\text{II})} - \frac{1}{6}s^3f_i^{(\text{III})} + \frac{1}{24}s^4f_i^{(\text{IV})} - O(s^5) \\ f_{i-2} &= f_i - 2sf_i^{(\text{I})} + 2s^2f_i^{(\text{II})} - \frac{4}{3}s^3f_i^{(\text{III})} + \frac{2}{3}s^4f_i^{(\text{IV})} - O(s^5) \end{aligned}$$

oraz wykonaniu kilku prostych przekształceń algebraicznych. Pokażemy to na przykładach.

Obliczając różnicę $f_{i+1} - f_{i-1}$ dostajemy

$$f_i^{(\text{I})} = \frac{f_{i+1} - f_{i-1}}{2s} + O(s^2),$$

gdzie $O(s^2)$ oznacza pominięte wyrazy, będące rzędu s^2 .

W celu uzasadnienia wzoru z różnicą centralną dla drugiej pochodnej należy wyznaczyć sumę $f_{i+1} + f_{i-1}$, z której po prostych przekształceniach otrzymujemy

$$f_i^{(\text{II})} = \frac{f_{i+1} - 2f_i + f_{i-1}}{s^2} + O(s^2),$$

ponieważ $f^{(I)}$ została wyznaczona z dokładnością $O(s^2)$.

Obliczmy teraz drugą pochodną z dokładnością do wyrazów $O(s^4)$. W tym celu zauważmy, że

$$f_{i+1} + f_{i-1} = 2f_i + s^2 f_i^{(II)} + \frac{s^4}{12} f_i^{(IV)},$$

skąd wynika, że

$$s^4 f_i^{(IV)} = 12(f_{i+1} + f_{i-1} - 2f_i - s^2 f_i^{(II)}).$$

Podstawienie powyższej równości do wzoru

$$f_{i+2} + f_{i-2} = 2f_i + 4s^2 f_i^{(II)} + \frac{4s^4}{3} f_i^{(IV)},$$

prowadzi do wyniku

$$f_i^{(II)} = \frac{-f_{i+2} + 16f_{i-1} - 30f_i + 16f_{i-1} - f_{i-2}}{12s^2} + O(s^4).$$

Zadania

C/1. Uzasadnić podane wyżej wzory dla pierwszej (C.5) i trzeciej (C.11) pochodnej, odpowiadające *różnicom centralnym*.

Wskazówka: Korzystając z odpowiednich szeregów Taylora, zapisać wyrażenia dla $(f_{i+1} - f_{i-1})$ oraz $(f_{i+2} - f_{i-2})$ i rozwiązać otrzymany układ dwóch równań ze względu na $f_i^{(I)}$ oraz $f_i^{(III)}$.

C/2. Pokazać, że podane wyżej formuły dla drugiej (C.9) i czwartej (C.13) pochodnej odpowiadające *różnicom centralnym* są rozwiązaniami układu równań otrzymanego po obliczeniu sum $f_{i+1} + f_{i-1}$ oraz $f_{i+2} + f_{i-2}$ (przy wykorzystaniu odpowiednich szeregów Taylora).

C/3. Wyprowadzić podane wzory dotyczące aproksymacji *do przodu*: (C.2), (C.4), (C.6), (C.8), (C.10) i (C.12).

C/4. Jaką postać będą miały odpowiednie wzory dla pochodnych, jeżeli w miejsce aproksymacji *do przodu* zastosujemy aproksymację *do tyłu* (tj. wykorzystującą tylko wartości f_i, f_{i-1}, \dots)?

Klasyczna aproksymacja Padégo

Jednym z bardziej efektywnych sposobów aproksymacji funkcji rzeczywistej jednej zmiennej jest aproksymacja Frobeniusa–Padégo [87], którą tu krótko przedstawimy.

Niech funkcja rzeczywista $f(z)$ ma w otoczeniu punktu $z = 0$ znane rozwinięcie w szereg potęgowy Taylora

$$f(z) = \sum_{i=0}^{\infty} c_i z^i. \quad (\text{D.1})$$

Definicja D-1. *Aproksymantą Padégo* $[j/m]_f(z)$ nazywamy funkcję wymierną postaci

$$[j/m]_f(z) = \frac{a_0 + a_1 z + \cdots + a_j z^j}{1 + b_1 z + \cdots + b_m z^m},$$

spełniającą relację $[j/m]_f(0) = f(0)$, której szereg Taylora wokół $z = 0$ pokrywa się ze skończoną liczbą początkowych wyrazów szeregu (D.1).

Zauważmy, że w aproksymancie $[j/m]_f$ mamy $j + 1$ współczynników w liczniku oraz m w mianowniku. W sumie daje to liczbę $(m + j + 1)$ dowolnych współczynników, które łatwo można wyznaczyć.

Z definicji aproksymanty wynika, że

$$f(z) = [j/m]_f(z) + O(z^{j+m+1}) = \frac{a_0 + a_1 z + \cdots + a_j z^j}{1 + b_1 z + \cdots + b_m z^m} + O(z^{j+m+1}). \quad (\text{D.2})$$

Wynika stąd równość

$$f(z)(1 + b_1 z + \cdots + b_m z^m) = a_0 + a_1 z + \cdots + a_j z^j + O(z^{j+m+1}),$$

z której otrzymujemy (z dokładnością do wyrazów $O(z^{j+m+1})$)

$$(c_0 + c_1 z + \cdots)(1 + b_1 z + \cdots + b_m z^m) = a_0 + a_1 z + \cdots + a_j z^j. \quad (\text{D.3})$$

Porównanie współczynników przy odpowiednich potęgach po obu stronach ostatniej równości prowadzi do układu równań

$$\begin{pmatrix} c_{j-m+1} & \cdots & c_j \\ \vdots & \ddots & \vdots \\ c_j & \cdots & c_{j+m-1} \end{pmatrix} \begin{pmatrix} b_m \\ \vdots \\ b_1 \end{pmatrix} = - \begin{pmatrix} c_{j+1} \\ \vdots \\ c_{j+m} \end{pmatrix}, \quad (\text{D.4})$$

którego rozwiązanie daje szukane współczynniki b_1, \dots, b_m . Przyjęto tu umownie, że $c_i = 0$ dla $i < 0$. Jeśli $j = m$, to powyższy układ równań przyjmuje postać

$$\begin{pmatrix} c_1 & \cdots & c_j \\ \vdots & \ddots & \vdots \\ c_j & \cdots & c_{2j-1} \end{pmatrix} \begin{pmatrix} b_j \\ \vdots \\ b_1 \end{pmatrix} = - \begin{pmatrix} c_{j+1} \\ \vdots \\ c_{2j} \end{pmatrix},$$

Pozostaje jeszcze wyznaczyć nieznane współczynniki a_0, \dots, a_j . Z relacji (D.3), porównując współczynniki przy z^0, \dots, z^j , otrzymujemy kolejno

$$\begin{aligned} a_0 &= c_0, & a_1 &= c_1 + b_1 c_0, \\ a_k &= c_k + \sum_{i=1}^{\min(k,m)} b_i c_{k-i}. \end{aligned} \quad (D.5)$$

Jeśli $j = m$, to ostatnie równanie przyjmuje postać

$$a_k = c_k + \sum_{i=1}^k b_i c_{k-i}.$$

Równania (D.4) i (D.5) pozwalają jednoznacznie określić współczynniki a_i, b_i aproksymanty Padégo $[j/m]_f$.

Każde rozwinięcie $f(z)$ w szereg potęgowy ma swój promień zbieżności $|z| < R$. Aproksymanty Padégo są zbieżne zazwyczaj w nieco szerszym obszarze niż odpowiedni szereg Taylora, co decyduje o ich przydatności.

Przykład. Niech

$$f(z) = \sqrt{\frac{1 + \frac{1}{2}z}{1 + 2z}} = 1 - \frac{3}{4}z + \frac{39}{32}z^2 - \dots$$

Wyznamy aproksymantę $[1/1]$ tej funkcji, tj.

$$\frac{a_0 + a_1 z}{1 + b_1 z} = 1 - \frac{3}{4}z + \frac{39}{32}z^2 + O(z^3).$$

Z powyższej równości wynika, że

$$1 + b_1 z - \frac{3}{4}z - \frac{3}{4}b_1 z^2 + \frac{39}{32}z^2 + O(z^3) = a_0 + a_1 z,$$

skąd znajdujemy $\frac{3}{4}b_1 = \frac{39}{32}$ i $b_1 = \frac{13}{8}$ oraz $a_0 = c_0 = 1, a_1 = c_1 + b_1 c_0 = \frac{7}{8}$. Ostatecznie

$$[1/1]_f(z) = \frac{1 + \frac{7}{8}z}{1 + \frac{13}{8}z} = \frac{8 + 7z}{8 + 13z}.$$

Zauważmy, że $\lim_{z \rightarrow \infty} f(z) = \frac{1}{2}$, a wartość $\lim_{z \rightarrow \infty} [1/1]_f(z) = \frac{7}{13}$, co oznacza dokładność 8%.

Na zakończenie wspomnijmy, że klasyczna aproksymanta Padégo $[j/m]_f$ (D.2) istnieje (w sensie równości (D.3)) pod warunkiem [87], że dla $z = 0$ zachodzi

$$\begin{vmatrix} c_{j-m+1} & \cdots & c_j & c_{j+1} \\ \vdots & \ddots & \vdots & \vdots \\ c_j & \cdots & c_{j+m-1} & c_{j+m} \\ z^m & \cdots & z & 1 \end{vmatrix} \neq 0,$$

gdzie ponownie położono $c_i = 0$ dla $i < 0$.

Przykład. Aproksymanta Padégo $[1/1]_f$ funkcji $f(z) = 1 + z^2$ jest równa 1, ale $1 + z^2 \neq 1$, ponieważ w tym przypadku wyznacznik

$$\begin{vmatrix} c_1 & c_2 \\ z & 1 \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ z & 1 \end{vmatrix}$$

znika dla $z = 0$.

Przedstawione podejście Frobeniusa–Padégo może być uogólnione na funkcje operatorów, co wykorzystujemy w rozdziale 5.

Zadania

D/1. Wyprowadzić układy równań (D.4) i (D.5).

D/2. Wyznaczyć początkowe aproksymanty Padégo jednej z następujących funkcji: $\sin z$, $\cos z$, $\operatorname{tg} z$, $\frac{\sin z}{z}$, $\frac{\operatorname{tg} z}{z}$, $\arcsin z$, $\operatorname{arctg} z$, $\sinh z$, $\cosh z$, $\operatorname{tgh} z$, e^z . Oszacować dokładność przybliżenia danej funkcji za pomocą aproksymanty Padégo w ograniczonym przedziale. Posłużyć się tablicami matematycznymi (np. [85]) zawierającymi odpowiednie rozwinięcia wymienionych funkcji wokół $z = 0$. Wskazówka: Można użyć programu PADE z *Numerical Recipes*.

Biblioteki podprogramów

Opisane w tym podręczniku algorytmy są zaimplementowane w wielu bibliotekach podprogramów, zarówno komercyjnych, jak i powszechnie dostępnych, tj. należących do *public domain*. Większość z nich jest zaprogramowana w językach Fortran 77/90 i C/C++, niektóre także w innych (Pascal, Java).

Biblioteki komercyjne sprzedawane są najczęściej w postaci binarnej (tj. skompilowanej dla określonej platformy sprzętowej, kompilatora i systemu operacyjnego). Zazwyczaj ceny kodów źródłowych takich bibliotek są wysokie.

Biblioteki niekomercyjne dostępne są w postaci kodów źródłowych, co czyni je praktycznie uniwersalnymi, tj. możliwymi do zastosowania w dowolnym środowisku. Dostęp do kodów źródłowych pozwala ponadto na zapoznanie się ze szczegółami implementacji i technikami efektywnego programowania.

Spośród bibliotek niekomercyjnych na szczególną uwagę zasługują EISPACK oraz LAPACK, dostępne w Internecie pod adresem <http://www.netlib.org/>.

Biblioteka EISPACK¹⁾ zawiera podprogramy pozwalające rozwiązywać zagadnienia własne dla macierzy gęstych (symetrycznych i niesymetrycznych rzeczywistych oraz hermitowskich i niehermitowskich zespolonych), pasmowych i trójdzielnych, podprogramy pomocnicze oraz podprogramy wykonujące rozkład macierzy względem wartości szczególnych. Opis biblioteki znajduje się w pracach [51, 52].

Biblioteka LAPACK²⁾, będąca rozwinięciem EISPACK-a oraz innej biblioteki, LINPACK-a, zawiera praktycznie pełny zestaw podprogramów z zakresu numerycznej algebry liniowej. Pozwalają one m.in. na rozwiązywanie układów równań liniowych i zagadnień własnych dla rzeczywistych i zespolonych macierzy gęstych i specjalnych (pasmowych, trójdzielnych, trójkątnych itp.) oraz liniowego zagadnienia najmniejszych kwadratów, wykonanie rozkładu macierzy względem wartości szczególnych. Podprogramy LAPACK-a napisane są w sposób umożliwiający obliczenia równoległe na dużych wieloprocessorowych maszynach, ale działają dobrze również na zwykłym PC. Podręcznik użytkownika [53] jest dostępny w Internecie wraz z całą biblioteką w postaci elektronicznej.

Spośród bibliotek komercyjnych wymienimy kilka, cieszących się większą popularnością:

- *Numerical Recipes* (dostępne w postaci kodów źródłowych w Fortranie, C lub Pascalu, zamieszczonych w książkach [41–44] wraz z obszernymi opisami). Teksty książek (za wyjątkiem wersji dla Pascala) są dostępne bezpłatnie w Internecie pod adresem <http://www.nr.com/>, jednakże za kody podprogramów

¹⁾ Ang. *EIGensystem PACKage*.

²⁾ Ang. *Linear Algebra PACKage*.

wciąż należy płacić;

- bibliotekę NAG³⁾ (adres internetowy: <http://www.nag.com/>);
- bibliotekę IMSL⁴⁾ firmy Visual Numerics Inc. (<http://www.vni.com/>).
- bibliotekę ESSL⁵⁾ firmy IBM, przeznaczoną dla produkowanych przez nią stacji roboczych rodzin RS/6000 i 370 (<http://www.austin.ibm.com/software/Apps/essl.html>).

Biblioteki NAG i IMSL są dostępne na najpopularniejsze platformy sprzętowe i programowe.

³⁾ *Numerical Algorithms Group*

⁴⁾ *International Mathematical and Statistical Libraries.*

⁵⁾ *Engineering and Scientific Subroutine Library*

Programy źródłowe

Dołączona do książki płytką CD zawiera kody źródłowe podprogramów w języku FORTRAN-77, realizujących algorytmy rozwiązywania równania Schrödingera omówione w rozdziałach 2–5. Zapisane na niej są również kody programów: **PREC** (służącego do badania dokładności i szybkości działania poszczególnych algorytmów; wyniki prezentowane są w rozdziale 9) oraz **OSCFN** (obliczającego funkcje falowe oscylatora harmonicznego przedstawione na rys. 9.7).

Tabela F.1 zawiera spis najważniejszych podprogramów i krótki opis ich działania. Szczegółowe opisy znajdują się w komentarzach wewnątrz kodów źródłowych.

Tabela F.1. Podprogramy dołączone na płytce CD

Podprogram	Typ macierzy	Działanie
VGRIDV	diagonalna $\text{diag}(\tilde{v})$	obliczenie wartości potencjału na siatce
GUARDROS	\tilde{H} dla metody Guardioli–Rosa	obliczenie wyrazów macierzy
TRED1 (EISPACK)	symetryczna	redukcja do postaci trójdzielnej metodą Householdera
DEAN	dowolna $T = \text{trid}(e, d, e)$	zliczanie wartości własnych mniejszych od zadanej liczby
DEAN3	\tilde{H} dla metody trójpunktowej	
DEAN1	\tilde{H} dla metody Lindberga	
DEAN5	\tilde{H} dla metody pięciopunktowej	
EIGVL1	\tilde{H} dla wybranej metody	przygotowanie macierzy
BISECT	dowolna	znajdowanie wartości własnych w zadanym przedziale (metoda bisekcji)
EIGVL2	\tilde{H} dla wybranej metody	diagonalizacja
EIGVAL	\tilde{H} dla wybranej metody	<code>eigvl1 + eigvl1</code>
DWSZ	dowolna $T = \text{trid}(e, d, e)$	wyznaczanie wektora własnego
DWSZ3	\tilde{H} dla metody trójpunktowej	przynależnego do danej wartości
DWSZL	\tilde{H} dla metody Lindberga	własnej (metoda DWSZ)

Poniżej zamieszczamy kody źródłowe kilku najważniejszych podprogramów, realizujących opisane w rozdziale 3 algorytmy Martina–Deana i DWSZ. Podprogramy te działają dla symetrycznych macierzy trójdzielnych $T = \text{trid}(e, d, e)$ o niezerowych wyrazach pozadiagonalnych. Jeżeli któryś z wyrazów $e_i = 0$, wtedy macierz rozpada się na bloki, do których można już stosować poniższe podprogramy.

Podprogram DEAN znajduje liczbę wartości własnych macierzy T mniejszych od z ,

tj. $\eta(\mathbf{T} - z\mathbf{1})$. W celu przyspieszenia obliczeń, do funkcji przekazywany jest wektor $\mathbf{e2}$ zawierający kwadraty wyrazów pozadiagonalnych¹⁾. Dodatkowe przyspieszenie możemy uzyskać, jeżeli wiemy, że dla danej liczby z zachodzi nierówność $\eta(\mathbf{T} - z\mathbf{1}) \leq k$, gdzie $k \leq n$. Wówczas po znalezieniu k ujemnych wyrazów ciągu (u_i) możemy przerwać główną pętlę. Jest to szczególnie istotne przy najniższych leżących wartościach własnych.

Funkcja działa prawidłowo w arytmetyce IEEE 754 (tj. gdy dzielenie przez zero daje wynik „ ∞ ”) przy wyłączonym zgłaszaniu wyjątków [71, 72, 142, 143]. Jeżeli dany kompilator nie jest zgodny z normą IEEE 754, w pewnych przypadkach może nastąpić przerwanie wykonywania programu z błędem dzielenia przez zero. Można się przed tym zabezpieczyć, dodając instrukcję w rodzaju „if (u.eq.0.) u=1e-37” przed instrukcją dzielenia.

```

integer function dean(d,e2,n,w,k)
*   ... zwraca liczbę wartości własnych symetrycznej macierzy
*   ... trójdzielnej mniejszych od zadanej liczby rzeczywistej
implicit none
*   ... parametry wywołania
*   d()   przekątna główna macierzy
*   e2()  kwadraty wyrazów pozadiagonalnych, niezerowe
*   n     rozmiar macierzy
*   w     zadana liczba
*   k     górne ograniczenie na wartość dean(...), jeśli znane
double precision d(*),e2(*),w
integer n,k

*   ... zmienne lokalne
double precision u
integer i,j

j=0
u=d(1)-w
if (u.lt.0.) j=1
do 281 i=2,n
    u=d(i)-w-e2(i-1)/u
*   ... licz ujemne wyrazy ciągu
    if (u.lt.0.) j=j+1
*   ... zakończ przeszukiwanie tak szybko, jak to możliwe
    if (j.eq.k) goto 282
281 continue
282 dean=j
return
end

```

¹⁾ Eliminacja jednego mnożenia nie powoduje jednak bardzo dużego wzrostu szybkości działania funkcji, gdyż zdecydowanie najbardziej czasochłonną operacją w pętli jest dzielenie zmiennoprzecinkowe (przynajmniej na mikroprocesorach Intel i486 i nowszych [142, 143]).

Podprogramy DEAN3, DEANL oraz DEAN5 wykonują analogiczne zadania; ciągi (u_i) obliczane są dla szczególnych postaci macierzy trójdzielnych lub pentadiagonalnych, występujących odpowiednio w metodach: trójpunktowej, Lindberga i pięciopunktowej dyskretyzacji równania Schrödingera. Ciągi te zadane są wzorami (3.2), (5.50) i (5.40), przy czym z uwagi na zmniejszenie błędów zaokrągleń, wyrazy macierzy zależne od potencjału są za każdym razem obliczane (tj. parametrem wejściowym jest wektor \tilde{v} , a nie wektory d i e). Zwróćmy uwagę, że wartość u w wyrażeniu typu

$$u = 2. + (v(i) - w) - 1. / u$$

ma szansę być obliczone dokładniej niż w sekwencji

$$\begin{aligned} d(i) &= 2. + v(i) \\ u &= d(i) - w - 1. / u \end{aligned}$$

w przypadku, gdy $v(i), w \ll 2$, a tak zwykle jest, gdyż potencjał i energia skalują się ze skokiem siatki s jak s^2 . Ponadto w podprogramie dla metody pięciopunktowej obliczane są odwrotności liczb u_i , gdyż zmiennoprzecinkowe mnożenie jest zazwyczaj szybsze od dzielenia²⁾

Podprogram BISECT znajduje metodą bisekcji wszystkie wartości własne macierzy T leżące w przedziale (z_{\min}, z_{\max}) , wykorzystując zewnętrzną funkcję FCNT do wyznaczania $\eta(T - z1)$. Umożliwia to zastosowanie podprogramu nie tylko do symetrycznej macierzy trójdzielnej, ale — po podstawieniu pod $fcnt$ funkcji innej niż powyżej opisana funkcja DEAN — również do macierzy pentadiagonalnych itp.

Podanie zbyt małej dokładności tol nie powoduje zapętlania się podprogramu z uwagi na warunek $z = z1 \vee z = z2$, gdzie $z = \frac{1}{2}(z1 + z2)$; będzie on prawdziwy, gdy $z1$ i $z2$ będą się różnić w reprezentacji zmiennoprzecinkowej tylko najmniej znaczącym bitem. Z uwagi na to, że na niektórych maszynach (np. na PC) wyniki częściowe mogą być obliczane z dokładnością większą niż DOUBLE PRECISION, konieczne są dodatkowe środki ostrożności (komentarz za etykietą 220).

W celu przyspieszenia działania podprogramu, informacje uzyskane podczas sukcesywnego połowienia początkowego przedziału są wykorzystywane do szacowania granic podprzedziałów, w których leżą kolejne wartości własne (patrz komentarze).

```

subroutine bisect(d1,d2,n,fcnt,wmin,wmax,tol,m,wk1,wk2,w,m1,m2)
*   ... znajduje metodą bisekcji wartości własne macierzy pasmowej
*   ... nieokreślonego typu, leżące w zadanym przedziale
implicit none
*   ... parametry wywołania
*   d1(),d2()
*       (<) zadają w dowolny, znany funkcji fcnt(...) sposób,
*       macierz pasmową (np. d1 -- wyrazy diagonalne, itd.)
*   n       (<) rozmiar macierzy
*   fcnt(...)
*   funkcja zliczająca wartości własne mniejsze od zadanej liczby
*   wmin,wmax
```

²⁾ Na wielu mikroprocesorach z rodziny Intel 80x86/x87 nawet *wielokrotnie* szybsze.

```

*      (<) granice przedziału zawierającego wartości własne
*      tol      (<) dokładność lokalizacji wartości własnych;
*      UWAGA: tol=0d0 pozwala osiągnąć maksymalną dokładność
*      m      (<) liczba poszukiwanych najmniejszych wartości własnych
*      wk1(),wk2()
*      (#) tablice robocze
*      w()      (>) kolejne wartości własne, jeśli zostały znalezione
*      m1      (>) numer pierwszej wartości własnej w przedziale, minus 1
*      m2      (>) numer ostatniej wartości własnej w przedziale,
*      jeśli m2<=m1, brak wartości własnych w przedziale,
*      jeśli m2-m1>m, zapisane tylko m najmniejszych wartości własnych
integer n,m,wk2(m),m1,m2
double precision d1(*),d2(*),w(m),wmin,wmax,tol,wk1(2,m)
integer fcnt
external fcnt

*      ... zmienne robocze i lokalne
*      wk1(1,?),wk1(2,?)
*      aktualne oszacowania z dołu i z góry kolejnych wartości własnych
*      wk2() wartości funkcji fcnt(...) dla liczb wk1(?,2)
*      z      aktualne przybliżenie bieżącej wartości własnej
*      z1,z2 oszacowania z dołu i z góry bieżącej wartości własnej
*      k,k1,k2 wartość fcnt(...) dla dla z,z1,z2
*      k20     wartość k2 przed zawężeniem bieżącego przedziału
*      m2a     numer ostatniej wartości własnej mieszczącej się w tablicy w()
double precision z,z1,z2
integer i,j,k,m2a,k1,k2,k20

*      ... podprogram zewnętrzny
logical bischk
external bischk

*      ... liczba wartości własnych w przedziale <wmin,wmax>
m1=fcnt(d1,d2,n,wmin,n)
m2=fcnt(d1,d2,n,wmax,n)
if (m2.le.m1) return
m2a=min(m2-m1,m)
*      ... początkowe oszacowania z dołu i z góry wartości własnych
do 205 i=1,m2a
    wk1(1,i)=wmin
    wk1(2,i)=wmax
    wk2(i)=m2
205 continue
k1=m1
i=1
*      ... pętla główna -- kolejne wartości własne

```

```

210 continue
*   ... ustal przedział poszukiwań dla w[m1+i]
    z1=wk1(1,i)
    z2=wk1(2,i)
    k2=wk2(i)
    k20=k2
*   ... pętla -- bisekcja
220 continue
    z=0.5*(z1+z2)
*   ... wywołanie funkcji zmusza kompilatory optymalizujące do zapisania
*   ... zmiennych w pamięci i zabezpiecza przed wejściem w nieskończoną
*   ... pętlę dla zbyt małych wartości tol, jeśli wyniki częściowe
*   ... mogą być obliczane z dokładnością większą niż "double precision"
    if ((z2-z1.le.tol).or.bischk(z1,z,z2)) goto 225
    k=fcnt(d1,d2,n,z,k2)
*   ... wybór prawego albo lewego podprzedziału
    if (k.eq.k1) then
        z1=z
    else
        z2=z
        j=k-m1
*   ... jeśli w przedziale wciąż więcej niż jedna wartość własna,
*   ... poprawa oszacowań dla w[k] (z góry) oraz dla w[k+1] (z dołu)
        if ((k2-k1.gt.1).and.(j.le.m2a)) then
            wk2(j)=k
            wk1(2,j)=z2
            if (j.lt.m2a) wk1(1,j+1)=z2
        end if
        k2=k
    end if
    goto 220
*   ... koniec pętli -- bisekcja
*   ... zapamiętanie znalezionej wartości własnej w[m1+i]
*   ... lub ciągu numerycznie nierozróżnialnych wartości własnych
*   ... w[m1+i],...,w[k2] leżących w przedziale o szerokości tol
225 do 226 j=i,min(k2-m1,m2a)
226 w(j)=z
*   ... przejście do następnej wartości własnej lub zakończenie
    i=k2-m1+1
    if (i.gt.m2a) return
    k1=k2
*   ... uporządkowanie oszacowań dla wartości własnych leżących
*   ... w początkowym przedziale <wm[i],wp[i]>:
*   ... oszacowania z dołu muszą tworzyć ciąg rosnący
230 j=i,min(k20-m1,m2a)
    if (wk1(1,j).lt.z2) then

```

```

        wk1(1,j)=z2
    else
        z2=wk1(1,j)
    end if
230  continue
*    ... to samo dla oszacowań z góry
    z2=wmax
    k2=m2
    do 235 j=min(k20-m1,m2a),i,-1
        if (wk1(2,j).gt.z2) then
            wk1(2,j)=z2
            wk2(j)=k2
        else
            z2=wk1(2,j)
            k2=wk2(j)
        end if
    235  continue
    goto 210
*    ... koniec głównej pętli
end
*    -----
logical function bischk(x1,x,x2)
*    ... sprawdza, czy x leży na brzegu przedziału <x1,x2>
implicit none
double precision x1,x,x2

    bischk=(x.eq.x1).or.(x.eq.x2)
    return
end

```

Podprogram DWSZ wyznacza unormowany wektor własny macierzy **T**, znajdując punkt początkowy w sposób opisany w podrozdziale 3.3.2. W podanej postaci podprogram działa poprawnie w arytmetyce IEEE 754.

Warto w tym miejscu zwrócić uwagę, że popularne koprocesory arytmetyczne rodziny Intel 80x87 oraz jednostki arytmetyki zmiennoprzecinkowej mikroprocesorów i486 i nowszych wykonują obliczenia zgodnie z normą IEEE 754 [142, 143]. Wyłączenie obsługi wyjątków (ich *zamaskowanie*) może jednak wymagać napisania krótkiego podprogramu w języku maszynowym. W tym kontekście wart polecenia jest kompilator g77, dostępny bezpłatnie dla środowisk DOS/Windows³⁾, Linux i wielu innych, który domyślnie nie włącza obsługi wyjątków.

Pętla obliczająca współrzędne wektora własnego **y** zawiera zabezpieczenie przed sytuacją, gdy jedna z numerycznie obliczonych współrzędnych y_i jest *dokładnie* równa zeru. W takim przypadku z i -tego wiersza układu równań (3.4) otrzymujemy, że $e_{i-1}y_{i-1} = -e_i y_{i+1}$. Przykładem takiego „złośliwego” wektora własnego jest wektor

³⁾ M.in. pod adresem internetowym <ftp://ftp.leo.org/pub/comp/os/os2/leo/gnu/emx+gcc/g77/>.

własny macierzy Δ^2 (5.26) przynależny do wartości własnej $z = 2$ (dla dowolnego n nieparzystego).

Podprogram przystosowany jest do obliczania unormowanych funkcji falowych zgodnie z warunkiem $\int_0^l |y(x)|^2 dx \simeq \sum_{i=1}^n |y_i|^2 l/n = 1$.

```

subroutine dwsz(d,e,n,w,l,wk,y,k,rem)
*   ... oblicza metodą DWSZ unormowany wektor własny przynależny do
*   ... danej wartości własnej symetrycznej macierzy trójdzielnej
implicit none
*   ... parametry wywołania
*   d()      (<) przekątna główna macierzy
*   e()      (<) wyrazy pozadiagonalne, niezerowe
*   n        (<) rozmiar macierzy
*   w        (<) wartość własna
*   l        (<) długość przedziału (do normowania);
*           jeśli l=dbl(n), wektor jest normowany do jedności
*   wk()     (#) tablica robocza
*   y()      (>) wektor własny
*   k        (>) wybrany punkt startowy
*   rem      (>) unormowana reszta (błąd) w punkcie x[k]
integer n,k
double precision d(n),e(n),w,l,wk(n,2),y(n),rem

*   ... zmienne lokalne
double precision rmin,r,t
integer i

rmin=1d+308
*   ... ciąg -e[i]Theta^-[i]
wk(1,1)=-e(1)/(d(1)-w)
do 310 i=2,n-1
310 wk(i,1)=-e(i)/(e(i-1)*wk(i-1,1)+d(i)-w)
*   ... ciąg -e[i-1]Theta^+[i]
wk(n,2)=-e(n-1)/(d(n)-w)
do 315 i=n-1,2,-1
    t=e(i)*wk(i+1,2)+d(i)-w
    wk(i,2)=-e(i-1)/t
    r=abs(e(i-1)*wk(i-1,1)+t)
*   ... wybór punktu początkowego z najmniejszą resztą
    if (r.lt.rmin) then
        rmin=r
        k=i
    end if
315 continue
y(k)=1.
t=1.
do 320 i=k-1,1,-1

```

```

        if (y(i+1).ne.0.) then
            y(i)=y(i+1)*wk(i,1)
        else
*           ... zabezpieczenie przed 0 razy nieskończoność
            y(i)=-e(i+1)*y(i+2)/e(i)
        end if
*           ... norma wektora
            t=t+y(i)**2
320 continue
        do 325 i=k+1,n
            if (y(i-1).ne.0.) then
                y(i)=y(i-1)*wk(i,2)
            else
*           ... zabezpieczenie przed 0 razy nieskończoność
                y(i)=-e(i-2)*y(i-2)/e(i-1)
            end if
*           ... norma wektora
                t=t+y(i)**2
325 continue
*           ... normowanie takie, by SUM[i=1,n] (y(x[i]))^2 1/n=1
            t=sqrt(n/(1*t))
            do 330 i=1,n
330 y(i)=y(i)*t
*           ... reszta w punkcie x[k]
            rem=rmin*t
            return
        end

```

Podprogramy DWSZ3 i DWSZL wykonują analogiczne zadanie dla szczególnych postaci macierzy trójdzielnych, występujących w metodach: trójpunktowej i Lindberga (dostosowane są do współpracy z DEAN3 i DEANL).

Podprogram GUARDROS oblicza według wzorów (5.33), (5.35) i (5.36) wyrazy macierzy symetrycznej $(-\mathbf{K}^{-1}\mathbf{\Gamma} + \mathbf{\bar{V}})$ występującej w równaniu (5.37), dla której rozwiązujemy zagadnienie własne w metodzie Guardioli–Rosa. Podprogram posiada wbudowane tabele współczynników dla 15 różnych aproksymant Padégo, danych wzorami (5.4)–(5.19).

Przeprowadzane w nim obliczenia są dość czasochłonne ($O(n^3)$ operacji zmiennoprzecinkowych), dlatego zadbane o minimalizację ilości wywołań funkcji SIN — jej wartości dla wszystkich możliwych $2n + 2$ argumentów są przechowywane w tablicy roboczej.

```

        subroutine guardros(ap,v,n,uplo,wk,h,ldh)
*           ... wyznacza macierz symetryczną do diagonalizacji
*           ... w metodzie Guardioli--Rosa z aproksymantą Padégo [j/m]
        implicit none
*           ... parametry wywołania
*           ap      (<) wybrana aproksymanta:

```

```

*          ap=1 [1/0], ap=2 [1/1], ap=3 [1/2], ap=4 [1/3], ap=5 [1/4],
*          ap=6 [2/0], ap=7 [2/1], ap=8 [2/2], ap=9 [2/3], ap=10 [3/0],
*          ap=11 [3/1], ap=12 [3/2], ap=13 [4/0], ap=14 [4/1], ap=15 [5/0].
*  v()      (<) przeskalowany potencjał
*  n        (<) liczba punktów siatki
*  uplo     (<) do wypełnienia górny trójkąt h ('U'), dolny ('L') lub cała
*  wk()     (#) tablica robocza
*  h()      (>) obliczona macierz
integer ap,n,ldh
double precision v(n),wk(0:2*n+1),h(ldh,n)
character*1 uplo

*  ... zmienne lokalne
integer i,k,l,il,kl,n2
double precision gamma,kappa,q,vv,s,ss
logical up,lo

double precision pi
parameter (pi=3.14159265358979d+0)

*  ... tablice współczynników aproksymant
double precision nu(4,15),mu(4,15)
*  ... liczba wyrazów w liczniku i mianowniku
integer apj(15),apm(15)
save nu,mu,apj,apm
data nu /
& 0d0,0d0,0d0,0d0,
& 0d0,0d0,0d0,0d0,
& 0d0,0d0,0d0,0d0,
& 0d0,0d0,0d0,0d0,
& 0d0,0d0,0d0,0d0,
& (-1d0/12),0d0,0d0,0d0,
& (1d0/20),0d0,0d0,0d0,
& (31d0/252),0d0,0d0,0d0,
& (-641d0/6360),0d0,0d0,0d0,
& (-1d0/12),(1d0/90),0d0,0d0,
& (13d0/168),(-23d0/10080),0d0,0d0,
& (49d0/276),(79d0/28980),0d0,0d0,
& (-1d0/12),(1d0/90),(-1d0/560),0d0,
& (17d0/180),(-1d0/270),(43d0/226800),0d0,
& (-1d0/12),(1d0/90),(-1d0/560),(1d0/3150) /
data mu /
& 0d0,0d0,0d0,0d0,
& (1d0/12),0d0,0d0,0d0,
& (1d0/12),(-1d0/240),0d0,0d0,
& (1d0/12),(-1d0/240),(31d0/60480),0d0,

```

```

& (1d0/12),(-1d0/240),(31d0/60480),(-289d0/725760),
& 0d0,0d0,0d0,0d0,
& (2d0/15),0d0,0d0,0d0,
& (13d0/63),(23d0/3780),0d0,0d0,
& (-37d0/2170),(-959d0/76320),(29891d0/32054400),0d0,
& 0d0,0d0,0d0,0d0,
& (9d0/56),0d0,0d0,0d0,
& (6d0/23),(43d0/3220),0d0,0d0,
& 0d0,0d0,0d0,0d0,
& (8d0/45),0d0,0d0,0d0,
& 0d0,0d0,0d0,0d0 /
data apj / 1,1,1,1,1,2,2,2,2,3,3,3,4,4,5 /
data apm / 0,1,2,3,4,0,1,2,3,0,1,2,0,1,0 /

*   ... rodzaj wypełnienia
up=((uplo.ne.'l').and.(uplo.ne.'L'))
lo=((uplo.ne.'u').and.(uplo.ne.'U'))
*   ... tablica sinusów dla przyspieszenia działania
q=pi/(n+1)
n2=2*n+2
do 260 l=0,n2-1
260 wk(l)=dsin(l*q)
*   ... elementy macierzy
do 275 i=1,n
do 270 k=1,i
*   ... transformata
vv=0.
il=0
kl=0
do 263 l=1,n
*   ... indeksy w tablicy sinusów rosną modulo 2*n+2
il=il+i
if (il.ge.n2) il=il-n2
kl=kl+k
if (kl.ge.n2) kl=kl-n2
vv=vv+wk(il)*wk(kl)*v(l)
263 continue
vv=2.*vv/(n+1)
if (i.ne.k) then
*   ... wstaw wyrazy pozadiagonalne na właściwe pozycje
if (lo) h(i,k)=vv
if (up) h(k,i)=vv
else
*   ... wyrazy diagonalne gamma[i]/kappa[i]
s=-4.*sin(0.5*i*q)**2
*   ... wyraz gamma[i]: suma szeregu

```

```

        ss=s
        gamma=s
        do 265 l=1,apj(ap)-1
            ss=ss*s
            gamma=gamma+nu(l,ap)*ss
265      continue
*      ... wyraz kappa[i]: suma szeregu
        ss=1d0
        kappa=1d0
        do 267 l=1,apm(ap)
            ss=ss*s
            kappa=kappa+mu(l,ap)*ss
267      continue
*      ... wstaw do macierzy
        h(i,i)=vv-gamma/kappa
        endif
270    continue
275  continue
      return
    end

```

Diagonalizacja otrzymanej macierzy polega na zredukowaniu jej do postaci trójdzielnej (np. podprogramem TRED1 z biblioteki EISPACK, a następnie zastosowaniu np. algorytmu bisekcji (BISECT, DEAN) — jest on efektywniejszy od algorytmu QL/QR, gdy wyznaczamy tylko część widma macierzy (w przypadku równania Schrödingera — kilka lub kilkanaście najmniejszych wartości własnych).

- [1] M. Kaku, *Wizje, czyli jak nauka zmieni świat w XXI wieku*, Prószyński i S-ka, Warszawa 2000.
- [2] R. Kurzweil, *Świat Nauki*, nr 12(100), 72 (1999).
- [3] A. Toffler, *Polityka*, nr 2(2227), 26 (2000).
- [4] W. Salejda, *Co wiedzieć powinien inżynier o fizycznej naturze informacji i procesach jej przetwarzania?*, wykład habilitacyjny przedstawiony Radzie Naukowej Wydziału Podstawowych Problemów Techniki Politechniki Wrocławskiej 16 grudnia 1999;
<http://www.if.pwr.wroc.pl/~ssalejda/>.
- [5] J. Birnbaum, R. S. Williams, *Physics Today*, I 2000, 38.
- [6] G.J. Milburn, *Inżynieria kwantowa*, Prószyński i S-ka, Warszawa 1999.
- [7] G.J. Milburn, *Procesor Feynmana*, Wydawnictwo CIS, Warszawa 2000.
- [8] R. Penrose, *Nowy umysł cesarza*, PWN, Warszawa 1995.
- [9] J. Horgan, *Koniec nauki czyli o granicach wiedzy u schyłku ery naukowej*, Prószyński i S-ka, Warszawa 1999.
- [10] P.L. DeVries, *A First Course in Computational Physics*, Wiley, New York 1994.
- [11] P.L. DeVries, *Am. J. Phys.*, **64**, 364 (1996) — przewodnik po literaturze, czasopiśmie naukowych oraz adresach internetowych dotyczących fizyki komputerowej do roku 1995.
- [12] S.E. Koonin, D.C. Meredith, *Computational Physics: Fortran Version*, Addison-Wesley, Reading 1995.
- [13] A.L. Garcia, *Numerical Methods for Physics*, Prentice-Hall, New York 1994.
- [14] R.H. Landau, M.J.P. Mejia, *Computational Physics*, Wiley, New York 1997.
- [15] N.J. Giordano, *Computational Physics*, Prentice-Hall, New York 1997.
- [16] T. Pang, *Metody obliczeniowe w fizyce. Fizyka i komputery*, PWN, Warszawa 2001.
- [17] Б.П. Демидович, И.А. Марон, *Основы вычислительной математики*, wyd. 4 popr., Наука, Москва 1970.
- [18] A. Ralston, *Wstęp do analizy numerycznej*, wyd. 3, PWN, Warszawa 1983.
- [19] G. Dahlquist, A. Björck, *Metody numeryczne*, PWN, Warszawa 1983.
- [20] J.M. Ortega, W.G. Poole, *An Introduction to Numerical Methods for Differential Equations*, Pitman Publishing, 1981;
Дж. Ортега, У. Пул, *Введение в численные методы решения дифференциальных уравнений*, Наука, Москва 1986.
- [21] J. Stoer, R. Bulirsch, *Wstęp do analizy numerycznej*, PWN, Warszawa 1987.

- [22] J. i M. Jankowscy, *Przegląd metod i algorytmów numerycznych*, część 1, WNT, Warszawa 1981.
M. Dryja, J. i M. Jankowscy, *Przegląd metod i algorytmów numerycznych*, część 2, WNT, Warszawa 1988.
- [23] G. Hammerlin, K.-H. Hoffman, *Numerical Mathematics*, seria *Undergraduate Texts in Mathematics*, Springer-Verlag, 1991.
- [24] A. Kiełbasiński, H. Schwetlick, *Numeryczna algebra liniowa*, WNT, Warszawa 1992.
- [25] Z. Fortuna, B. Macukow, J. Wąsowski, *Metody numeryczne*, seria *Podręczniki akademickie: Elektronika, Informatyka, Telekomunikacja*, WNT, Warszawa 1993.
- [26] E. Majchrzak, B. Mochnacki, *Metody numeryczne. Podstawy teoretyczne, aspekty praktyczne i algorytmy*, Wydawnictwo Politechniki Śląskiej, Gliwice 1996.
- [27] A.A. Samarski, J.S. Nikołajew, *Metody rozwiązywania równań siatkowych*, PWN, Warszawa 1988.
- [28] R.W. Seidel, *From Mars to Minerva: The Origin of Scientific Computing in the Atomic Energy Commission Labs*, Physics Today, X 1996, 33.
- [29] A.E. Brenner, Physics Today, X 1996, 24.
- [30] L.D. Landau, E.M. Lifszyc, *Mechanika kwantowa*, PWN, Warszawa 1985;
Л.Д. Ландау, Е.М. Лифшиц, *Теоретическая физика*, Т. 3, *Квантовая механика. Нерелятивистская теория*, Наука, Москва 1989.
- [31] E. Merzbacher, *Quantum Mechanics*, wyd. 3, Wiley, New York 1998.
- [32] R.L. Liboff, *Introductory Quantum Mechanics*, wyd. 3, Addison-Wesley Longman, New York 1998.
- [33] I. Białynicki-Birula, M. Cieplak, J. Kamiński, *Teoria kwantów*, PWN, Warszawa 1991.
- [34] W. Greiner, *Quantum Mechanics. An Introduction*, wyd. 2 popr., Springer-Verlag, Berlin 1993.
- [35] S. Flügge, *Practical Quantum Mechanics*, Springer-Verlag, Berlin 1999.
- [36] R. Shankar, *Principles of Quantum Mechanics*, wyd. 2, Plenum Press, New York-London 1994.
- [37] F. Mandl, *Quantum Mechanics*, Wiley, Chichester 1992.
- [38] J.J. Sakurai, *Modern Quantum Mechanics*, Addison-Wesley, Reading 1994.
- [39] F. Schwabl, *Quantum Mechanics*, wyd. 2 popr., Springer-Verlag, Berlin 1995.
- [40] P. Harrison, *Quantum Wells, Wires and Dots. Theoretical and Computational Physics*, Wiley, Chichester 2000.
- [41] B.P. Flannery, W.H. Press, *Numerical Recipes in FORTRAN 77. The Art of Scientific Computing Volume 1, Volume 1 of Fortan Numerical Recipes*, Cambridge University Press, Cambridge 1993.

- [42] B.P. Flannery, M. Metcalf, S.A. Teukolsky, W.H. Press, W.T. Vetterling, *Numerical Recipes in Fortran 90. The Art of Parallel Scientific Computing Volume 2, Volume 2 of Fortran Numerical Recipes*, Cambridge University Press, Cambridge 1996.
- [43] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, *Numerical Recipes in C. The Art of Scientific Computing. Second Edition*, Cambridge University Press, Cambridge 1993.
- [44] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, *Numerical Recipes in Pascal. The Art of Scientific Computing*, Cambridge University Press, Cambridge 1992.
- [45] J.H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford 1965;
Дж.Х. Уилкинсон, *Алгебраическая проблема собственных значений*, Наука, Москва 1970.
- [46] J.H. Wilkinson, C. Reinsch (red.), *Handbook for Automatic Computations*, t. 2, *Linear Algebra*, Springer-Verlag, Heidelberg 1971.
- [47] F.S. Acton, *Numerical Methods that Work*, The Mathematical Association of America, Washington 1990.
- [48] B.N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice-Hall, New York 1980;
Б. Парлетт, *Симметричная проблема собственных значений. Численные методы*, Мир, Москва 1983.
- [49] N.J. Higham, *Accuracy and Stability of Numerical Algorithms*, SIAM, Philadelphia 1996.
- [50] J.W. Demmel, *Applied Numerical Linear Algebra*, SIAM, Philadelphia 1997.
- [51] B.T. Smith et al., *Matrix eigensystems routines — EISPACK guide*, seria *Lecture Notes in Computer Science*, t. 6, wyd. 2, Springer-Verlag, 1990.
- [52] B.W. Garbow et al., *Matrix eigensystem routines — EISPACK guide extension*, seria *Lecture Notes in Computer Science*, t. 51, wyd. 2, Springer-Verlag, 1990.
- [53] E. Anderson et al., *LAPACK Users' Guide*, wyd. 3, SIAM, Philadelphia 2000;
http://www.netlib.org/lapack/lug/lapack_lug.html.
- [54] J.F. Van der Maelen Uría, S. García-Granda, A. Menéndez-Velázquez, *Am. J. Phys.*, **64**, 327 (1996).
- [55] W. Salejda, M.H. Tyc, J. Andrzejewski, M. Kubisa, J. Misiewicz, M. Just, K. Ryczko, *Acta Phys. Pol. A*, **95**, 881 (1999).
- [56] R. Guardiola, J. Ros, *J. Comput. Phys.*, **45**, 374 (1982).
- [57] B. Lindberg, *J. Chem. Phys.*, **88**, 3805 (1988).
- [58] G.C. Groenenboom, H.M. Buck, *J. Chem. Phys.*, **92**, 4374 (1990).

- [59] J. Killingbeck, *Microcomputer Algorithms*, Hilger, Bristol, 1991;
J. Killingbeck, G. Jolicard, Phys. Lett. A, **172**, 313 (1993);
J. Killingbeck, Phys. Lett. A, **261**, 40 (1999).
- [60] C.F. Gerald, P.O. Wheatley, *Applied Numerical Analysis*, Addison-Wesley, Reading 1989.
- [61] M.A. Núñez, Phys. Rev. A, **47**, 3620 (1992);
M.A. Núñez, G. Izquierdo B., Int. J. Quantum Chem., **47**, 405 (1993);
M.A. Núñez, Int. J. Quantum Chem., **50**, 113 (1993);
M.A. Núñez, Int. J. Quantum Chem., **51**, 57 (1994).
- [62] B. Baron, *Metody numeryczne w Turbo Pascalu*, Helion, Gliwice 1995.
- [63] A. Marciniak, D. Gregulec, J. Kaczmarek, *Numerical Procedures in Turbo Pascal for Your PC*, seria *Biblioteka Użytkownika Mikrokomputerów*, t. 25, Nakom, Poznań 1992;
A. Marciniak, D. Gregulec, J. Kaczmarek, *Podstawowe procedury numeryczne w języku Turbo Pascal*, seria *Biblioteka Użytkownika Mikrokomputerów*, t. 29, wyd. 3, Nakom, Poznań 1997.
- [64] W.A. Harrison, *Teoria ciała stałego*, PWN, Warszawa 1976.
- [65] G. Bastard, *Wave Mechanics Applied to Semiconductor Heterostructures*, Les Editions de Physique, Les Ulis Cedex 1988.
- [66] C.A. Kocher, Am. J. Phys., **45**, 71 (1977).
- [67] H. Konwent, Phys. Lett. A, **118**, 467 (1986).
- [68] H. Konwent, Acta Phys. Pol. A, **71**, 637 (1987).
- [69] H. Konwent, Phys. Stat. Sol. B, **138**, K7 (1986).
- [70] H. Konwent, P. Machnikowski, P. Magnuszewski, A. Radosz, J. Phys. A: Math. Gen., **31**, 7541 (1998).
- [71] D. Goldberg, ACM Computing Surveys 23, 5 (1991);
http://www.sns.ias.edu/Main/computing/compilers_html/common-tools/numerical_comp_guide/goldberg1.doc.html.
- [72] W. Kahan, *Lecture notes on the status of IEEE standard 754 for binary floating-point arithmetics*, Berkeley 1995; e-print
<http://HTTP.CS.Berkeley.EDU/~wkahan/ieee754status/ieee754.ps>.
- [73] K.S. Dy, S.Y. Wu, T. Spratlin, Phys. Rev. B, **20**, 4237 (1979).
- [74] Z. Zheng, J. Phys. Cond. Matt., **19**, L689 (1986).
- [75] I.S. Dhillon, *A New $O(n^2)$ Algorithm for the Symmetric Tridiagonal Eigenvalue/Eigenvector Problem*, praca doktorska, University of California, Berkeley 1997;
<http://www.cs.utexas.edu/users/inderjit/>;
I.S. Dhillon, B.N. Parlett, *Orthogonal eigenvectors and relative gaps*,
<http://www.cs.utexas.edu/users/inderjit/>.

- [76] B.N. Parlett, I.S. Dhillon, *Linear Algebra and Its Applications*, **267**, 247 (1997); <http://www.cs.utexas.edu/users/inderjit/>.
- [77] B.A. Foreman, Phys. Rev. B, **52**, 12241 (1995).
- [78] R. Balian, D. Bessis, Phys. Rev. B, **51**, 17624 (1995).
- [79] O. Madelung, *Introduction to Solid-State Theory*, Springer-Verlag, Berlin–Heidelberg–New York, 1978.
- [80] T. Ando, S. Mori, Surf. Sci. **113**, 124 (1982).
- [81] T. Ando, S. Wakahara, H. Akera, Phys. Rev. B, **40**, 11609 (1989).
- [82] B. Laikhtman, Phys. Rev. B, **46**, 4769 (1992).
- [83] B.A. Foreman, Phys. Rev. Lett., **80**, 3823 (1998).
- [84] S. Adachi, J. Appl. Phys., **58**, R1 (1985).
- [85] M. Abramowitz, I.A. Stegun, *Handbook of Mathematical Functions*, seria *Applied Mathematics*, t. 55, National Bureau of Standards 1964.
- [86] F.B. Hildebrand, *Introduction to Numerical Analysis*, rozdział 5, McGraw–Hill, New York 1974.
- [87] G.A. Baker, Jr., P. Graves-Morris, *Padé Approximants*. seria *Encyclopedia of Mathematics and its Applications*, t. 13 i 14, Addison-Wesley, Reading 1981; Дж. Бейкер, мл., П. Грейвс-Моррис, *Аппроксимации Падэ*, Мир, Москва 1986.
- [88] G.D. Smith, *Numerical solution of partial differential equations*, wyd. 3, Clarendon Press, Oxford 1985.
- [89] C.H. Reinsch, Num. Math., **25**, 591 (1971).
- [90] J.J.M. Cuppen, Numer. Math., **36**, 177 (1981).
- [91] J. Rutter, *A Serial Implementation of Cuppen's Divide and Conquer Algorithm for the Symmetric Eigenvalue Problem*, Report No. UCB/CSD 97/799, Computer Science Division, University of California, Berkeley 1994; <http://sunsite.berkeley.edu/NCSTRL/>.
- [92] L.L. Trefethen, D. Bau III, *Numerical linear algebra*, SIAM, Philadelphia 1997.
- [93] G.Y. Hu, R.F. O'Connell, J. Phys. A: Math. Gen., **29**, 1511 (1996); S. Simons, J. Phys. A: Math. Gen., **30**, 755 (1997); H.A. Yamani, M.S. Abdelmonem, J. Phys. A: Math. Gen., **30**, 2889 (1997); D.A. Lavis, B.W. Southern, I.F. Wilde, J. Phys. A: Math. Gen., **30**, 7229 (1997); Y. Huang, W.F. McColl, J. Phys. A: Math. Gen., **30**, 7919 (1997).
- [94] R.E. Peierls, *Quantum Theory of Solids*, rozdział 1, Clarendon Press, Oxford 1955.
- [95] A. Czachor, *Drgania atomów w ciele stałym*, rozdział 6, PWN, Warszawa 1982.
- [96] N.W. Ashcroft, N.D. Mermin, *Fizyka ciała stałego*, rozdział 22, PWN, Warszawa 1986.
- [97] W. Salejda, Communication of Joint Inst. Nucl. Res., **E17–90–241**, Dubna 1990.

- [98] A.M. Kosewicz, *Mechanika fizyczna nieidealnych krystalicznych ciał stałych*, Wydawnictwo Uniwersytetu Wrocławskiego, Wrocław 2000.
- [99] F.J. Dyson, Phys. Rev., **92**, 1331 (1953).
- [100] P. Dean, Rev. Mod. Phys., **44**, 127 (1972);
П. Дин, *Колебательные спектры неупорядоченных систем. Численные результаты*, w: *Вычислительные методы в теории твердого тела*, rozdział 8, Мир, Москва 1975.
- [101] H. Böttger, *Principles of the Theory of Lattice Dynamics*, Akademie-Verlag, Berlin 1983.
- [102] D. Shechtman, I. Blech, D. Gratias, J. W. Cahn, Phys. Rev. Lett., **53**, 1951 (1984).
- [103] *Workshop on Aperiodic Crystals*, J. Physique Coll. C, **3** (1986).
- [104] P.J. Steinhardt, S. Ostlund, *The Physics of Quasicrystals*, World Scientific, Singapore 1987.
- [105] P. Guyot, P. Krammer, M. de Boissieu, Rep. Prog. Phys., **54**, 1373 (1991).
- [106] S.J. Poon, Adv. Phys., **41**, 303 (1992).
- [107] *Fourth International Conference on Quasicrystals, St. Louis, USA, 31 V–5 VI 1992*, J. Non-Cryst. Solids, **153–4** (1993); materiały tej konferencji zawierają reprezentatywne pozycje dotyczące struktur aperiodycznych.
- [108] A.I. Goldman, R.F. Kelton, Rev. Mod. Phys., **65**, 213 (1993).
- [109] D.P. DiVincenzo, P.J. Steinhardt (red.), *Quasicrystals: The State of the Art*, World Scientific, Singapore 1991.
- [110] D. Levine, P.J. Steinhardt, Phys. Rev. Lett., **53**, 2477 (1984);
D. Levine, P.J. Steinhardt, Phys. Rev. B, **34**, 596 (1986);
J.E.S. Socolar, P.J. Steinhardt, Phys. Rev. B, **34**, 617 (1986);
G.Y. Onoda, P.J. Steinhardt, D.P. DiVincenzo, J.E.S. Socolar, Phys. Rev. Lett., **60**, 2563 (1988);
J.E.S. Socolar, Phys. Rev. B, **39**, 10519 (1989);
H.-C. Jeng, P.J. Steinhardt, Phys. Rev. Lett., **73**, 1943 (1994);
H. Gähler, Phys. Rev. Lett., **74**, 334 (1995);
H.-C. Jeng, P.J. Steinhardt, Phys. Rev. B, **55**, 3520 (1997).
- [111] W. Salejda, Int. J. Mod. Phys. B, **9**, 1429 (1995);
W. Salejda, Int. J. Mod. Phys. B, **9**, 1453 (1995);
W. Salejda, Int. J. Mod. Phys. B, **9**, 1475 (1995).
- [112] W. Salejda, *Dynamika sieci i przewodnictwo elektryczne kwazijednowymiarowych struktur aperiodycznych. Autoreferat habilitacyjny*, e-print
http://www.if.pwr.wroc.pl/~ssalejda/habil_ps.zip.
- [113] W. Salejda, Acta Phys. Pol. A, **78**, 323 (1990).
- [114] W. Salejda, Int. J. Mod. Phys., **5**, 825 (1991).

- [115] W. Salejda, materiały konferencji *Komputerowe Obliczenia Dużej Skali. 1. Szkoła Fizyki Komputerowej*, 365, Instytut Fizyki Jądrowej im. H. Niewodniczańskiego w Krakowie, Kraków 1994.
- [116] B.B. Mandelbrot, *The Fractal Geometry of Nature*, W.H. Freeman and Company, New York 1983.
- [117] M. Barnsley, *Fractals Everywhere*, Academic Press, Boston 1988.
- [118] K. Falconer, *Fractal Geometry*, John Wiley & Sons, Chichester 1990.
- [119] T.C. Halsey, M.H. Jensen, L.P. Kadanoff, I. Procaccia, B.I. Shraiman, Phys. Rev. A, **33**, 1141 (1986).
- [120] А.И. Олемский, А.Я. Флат, Усп. Физ. Наук, **163**, 1 (1993).
- [121] G. Paladin, A. Vulpiani, Phys. Rep., **156**, 147 (1987).
- [122] A.P. Siebesma, *Multifractals in Condensed Matter*, praca doktorska, The Cheese Press, Edam 1989.
- [123] M. Reed, B. Simon, *Methods of Modern Mathematical Physics. 1. Functional Analysis*, Academic Press, New York 1972;
М. Рид, Б. Саймон, *Методы современной математической физики*, Т. 1, *Функциональный анализ*, Мир, Москва 1977.
- [124] P. Billingsley, *Prawdopodobieństwo i miara*, PWN, Warszawa 1987.
- [125] W. Salejda, Communication of Joint Inst. Nucl. Res., **E17-88-880**, Dubna 1988;
W. Salejda, Int. J. Mod. Phys., **3**, 1109 (1989).
- [126] С.В. Тябликов, *Методы квантовой теории магнетизма*, Наука, Москва 1975.
- [127] D.C. Mattis, *The theory of magnetism I: statics and dynamics*, wyd. 2 popr., Springer-Verlag, Berlin 1988.
- [128] K. Ryczko, *Wpływ ekscytonów na podstawowe przejścia optyczne w półprzewodnikach w obecności zewnętrznego pola elektrycznego. Praca magisterska*, Instytut Fizyki Politechniki Wrocławskiej, Wrocław 1996;
S.I. Pokutnyi, W. Salejda, J. Misiewicz, K. Ryczko, Укр. фіз. журн., **43**, 1259 (1998).
- [129] М.Н. Тус, *Dwuwymiarowy ekscyton w zewnętrznym polu elektrycznym. Praca magisterska*, Instytut Fizyki Politechniki Wrocławskiej, Wrocław 1998;
S.I. Pokutnyi, M.H. Tys, W. Salejda, J. Misiewicz, Укр. фіз. журн., **46**, 113 (2001);
S.I. Pokutnyi, M.H. Tys, W. Salejda, J. Misiewicz, Физика Твёрдого Тела, **43**, 888 (2001).
- [130] L. Jacak, P. Hawrylak, A. Wójs, *Kropki kwantowe*, seria *Prace naukowe Instytutu Fizyki Politechniki Wrocławskiej. Monografie*, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 1996.
- [131] L. Jacak, P. Hawrylak, A. Wójs, *Quantum Dots*, Springer-Verlag, Berlin–Heidelberg–New York, 1998.

- [132] J. Krasnyj, A. Mituś, W. Salejda, L. Jacak, *Modelling of qubits in multielectron quantum dots; exploiting the singlet-triplet transition*, praca prezentowana na *SQID Meeting*, Torino, 23–24 VI 2000;
L. Jacak, J. Krasnyj, D. Jacak, W. Salejda, A. Mituś, *Acta Phys. Pol. A*, **99**, 277 (2001).
- [133] W. Salejda, *Proceedings of the VI-th Max Born Symposium on The Nature of Crystalline States, Kudowa-Zdrój, Poland, 21–24 IX 1995*, *Physica A*, **232**, 769 (1996).
- [134] W. Salejda, P. Szyszuk, *Physica A*, **252**, 547 (1998).
- [135] W. Salejda, M. Kubisa, J. Misiewicz, K. Ryczko, M. Tyc, *Materiały XXVII Międzynarodowej Szkoły Materiałów Półprzewodnikowych „Jaszowiec ’98”*, Ustroń-Jaszowiec, 7–12 VI 1998, *Acta Phys. Pol. A*, **94**, 514 (1998).
- [136] J.Q. You, L. Zhang, Q.B. Yang, *Phys. Rev. B*, **55**, 1314 (1997).
- [137] F.M. Peeters, P. Vasilopoulos, *Phys. Rev. B*, **47**, 7508 (1993);
F.M. Peeters, A. Matulis, *Phys. Rev. B*, **48**, 15166 (1993);
A. Matulis, F.M. Peeters, *Phys. Rev. Lett.*, **72**, 1518 (1994);
I.S. Ibrahim, F.M. Peeters, *Am. J. Phys.*, **63**, 171 (1995);
J.Q. You, L. Zhang, P.K. Ghosh, *Phys. Rev. B*, **52**, 17243 (1995);
I.S. Ibrahim, F.M. Peeters, *Phys. Rev. B*, **52**, 17321 (1995);
Y. Guo, B. Gu, W. Duan, Y. Zhang, *Phys. Rev. B*, **55**, 9314 (1997);
J.Q. You, L. Zhang, Q.B. Yang, *Phys. Rev. B*, **55**, 15757 (1997);
Z.L. Ji, D.W.L. Sprung, *Phys. Rev. B*, **56**, 1045 (1997);
F.M. Peeters, I.S. Ibrahim, V.A. Schweigert, *Phys. Rev. B*, **56**, 7508 (1997);
Y. Guo, Z.-Q. Li, B.-L. Gu, Q. Sun, J.-Z. Yu, Y. Kawazoe, *Eur. Phys. J. B*, **3**, 263 (1998);
Y. Guo, B.-L. Gu, Z.-Q. Li, Y. Kawazoe, *Phys. Lett. A*, **238**, 185 (1998);
O.M. Yevtushenko, K. Richter, *Phys. Rev. B*, **58**, 14839 (1998);
F.M. Peeters, J. de Boeck, *Hybrid magnetic-semiconductor nanostructures*, w: H.S. Nalwa (red.), *Nanostructured Materials and Nanotechnology*, Academic Press, San Diego 2002.
- [138] P. Templin, *Dwuwymiarowy gaz elektronowy w niejednorodnym prostopadłym polu magnetycznym. Praca magisterska*, Raport SPR-334/98, Instytut Fizyki Politechniki Wrocławskiej, Wrocław 1998.
- [139] R. Landauer, *Phil. Mag.*, **21**, 863 (1970);
M. Büttiker, Y. Imry, R. Landauer, S. Pinhas, *Phys. Rev. B*, **31**, 6207 (1985);
P. Erdős, R. C. Hendron, *Adv. Phys.*, **31**, 65 (1982);
R. Landauer, *Z. Phys. B*, **68**, 217 (1987).
- [140] W. Salejda, M.H. Tyc, *Physica A*, **303**, 493 (2002).
- [141] M. Just, *Numeryczne metody rozwiązywania równania Schrödingera. Program MARRS. Praca magisterska*, Raport SPR-333/1998, Instytut Fizyki Politechniki Wrocławskiej, Wrocław 1998;

- W. Salejda, M. Just, M.H. Tyc, Computational Methods in Science and Technology, **6**, 73 (2000).
- [142] M. Tuszyński, R. Goczyński, *Koprocesory 80287, 80387 oraz i486*, Komputerowa Oficyna Wydawnicza „Help”, Warszawa 1992.
- [143] H. Małysiak, B. Pochopień, E. Wróbel, *Procesory arytmetyczne*, WNT, Warszawa 1993.
- [144] M. Toda, *Theory of Nonlinear Lattices*, wyd. 2, Springer-Verlag, Berlin 1989.
- [145] J.B. Brojan, J. Mostowski, K. Wódkiewicz, *Zbiór zadań z mechniki kwantowej*, wyd. 2 popr., PWN, Warszawa 1978.

- algebraiczne zagadnienie własne, 15, 17, 21, 29, 42, 48, 51, 54, 55, 59, 100, 105, 107, 116, 151, 161, 165, 177
- algebraiczne zagadnienie własne uogólnione, 48, 50, 53, 105
- algorytm dekompozycji, 67
- algorytm DWSZ, 33
- algorytm dziel i rządź, 78
- algorytm dziel i rządź stabilny, 81
- algorytm Grama–Schmidta, 93
- algorytm Householdera, 85
- algorytm ilorazów Reyleigha, 93
- algorytm iteracji odwrotnych, 91
- algorytm Martina–Deana, 30
- algorytm potęgowy, 89
- algorytm prostej rekurencji do tyłu, 32
- algorytm prostej rekurencji w przód, 32
- algorytm QL niejawny, 74
- algorytm QL niewymierny, 71
- algorytm QL wymierny, 72
- algorytm QR, 65
- algorytm QR niejawny, 74
- algorytm QR niewymierny, 70
- algorytm QR wymierny, 72
- algorytm rekombinacji, 69
- algorytm rozwiązywania trójdziagonalnego układu równań, 92
- algorytm Rutishausera, 88
- aproksymanta Padégo, 46, 52, 148, 173
- arytmetyka IEEE, 30, 180, 184
- atom wodoru, 24, 109, 132
- bezwładność macierzy, 163
- C (język programowania), 177
- ciąg Sturma, 34, 145
- czas obliczeń, 132, 134–138, 144
- dekompozycja, 64, 65, 70, 71
- Delphi, 19, 123
- dokładność obliczeń, 132, 134–140, 144
- dwuwymiarowy gaz elektronowy, 118
- dyskretyzacja, 15, 23, 42, 116
- EISPACK, 70, 72, 73, 75, 85, 87, 177, 179, 189
- ekscyton, 24, 109
- ekstrapolacja Richardsona, 55
- epsilon maszynowe, 26, 133, 165
- formuła pięciopunktowa, 51, 86, 170, 171
- formuła siedmiopunktowa, 57, 86
- formuła trójpunktowa, 23, 148, 169, 170
- Fortran, 81, 132, 177, 179
- fraktal, 103
- funkcja falowa, 15–17, 22, 39, 104, 114, 117, 118, 125, 127, 147, 185
- funkcja falowa obwiedni, 39
- funkcja gęstości stanów, 101
- funkcja własna, 15, 21, 116
- Java, 177
- kropka kwantowa, 114
- kwazikryształ, 102
- LAPACK, 73, 75, 80, 82, 85, 87, 94, 149, 177
- macierz diagonalna, 30, 48, 50, 53, 166
- macierz dynamiczna, 99
- macierz hermitowska, 96, 101, 166, 177
- macierz kwazitrójdziagonalna, 101
- macierz odwrotna, 91, 92
- macierz pasmowa, 16, 85, 100, 149, 177
- macierz pentadiagonalna, 52, 181
- macierz rozrzedzona, 16, 59
- macierz symetryczna, 50, 59, 82, 87, 95, 96, 162–166, 177
- macierz trójdziagonalna, 23, 24, 29, 30, 32, 42, 48, 53, 59, 65, 72, 76, 82, 92, 95, 100, 149, 164, 165, 177, 179
- macierz trójkątna, 30, 63, 65, 71, 149, 167, 177
- magnony, 107

- MARRS (program), 10, 123, 151
 metoda bisekcji, 30, 42, 52, 93, 95, 101, 116, 181
 metoda DWSZ, 32, 42, 54, 55, 72, 116, 124, 143, 145, 184
 metoda dziel i rządź, 76, 95, 166
 metoda Grama–Schmidta, 93, 166
 metoda Guardiolì–Rosa, 45, 50, 124, 132, 134–138, 141, 144, 146
 metoda Householdera, 82, 86, 95
 metoda iteracji odwrotnych, 91, 93, 95
 metoda Jacobiego, 87
 metoda Lindberga, 52, 124, 132, 134–138, 141, 144–146, 181, 186
 metoda macierzy przejścia, 17, 40
 metoda Numerowa, 15, 47
 metoda pięciopunktowa, 51, 124, 145, 181
 metoda potęgowa, 89
 metoda QL, 72, 166
 metoda QR, 63, 72, 95, 166
 metoda strzałów, 15, 16, 18
 metoda styczna–sieczna, 35
 metoda trójpunktowa, 23, 55, 124, 132, 134–138, 141, 143, 146, 181, 186
 metoda wariacyjna, 15, 17, 18, 104
 metody elementów skończonych, 15, 17, 18
 metody siatkowe, 9, 15, 24
 model Heisenberga, 106
 model Kroniga–Penneya, 115
 multifraktal, 103

 norma macierzy, 161, 162
 norma wektora, 161
 Numerical Recipes, 75, 85, 89, 94, 177

 obrót Givensa, 61, 64, 65, 67, 70, 71, 74, 86
 obrót Jacobiego, 61, 87
 odbicie Householdera, 62, 82
 operator kreacji/anihilacji, 106
 oscylator harmoniczny, 24, 129, 134, 143

 Pascal, 177
 pole elektryczne, 109
 pole magnetyczne, 106, 114, 116
 potencjał Konwenta, 27, 130, 135
 potencjał kulombowski, 108, 132, 137, 138
 potencjał Lenarda–Jonesa, 26
 potencjał Morse’a, 27, 130, 136
 potencjał wektorowy, 117, 118
 promień spektralny, 161
 przekształcenie ortogonalne, 48, 60, 64, 74, 91, 162, 166
 przesunięcie widma, 63, 64, 69, 75, 91, 93

 redukcja, 59, 82, 86, 89
 rekombinacja, 64, 65, 67, 70, 71
 rekurencja, 30, 32, 78, 79, 102, 165
 rozmiar pamięci operacyjnej, 145
 równanie masy efektywnej, 39, 116
 równanie Schrödingera niestacjonarne, 147
 równanie Schrödingera stacjonarne, 9, 15–18, 21, 47, 53, 86, 104, 108, 109, 112, 114, 115, 118, 123, 151
 równanie wiekowe, 17, 79, 105
 różnice skończone, 15, 16, 24, 26, 41, 45, 149, 169

 siatka, 15, 16, 22–25, 41, 45, 46, 51, 53, 55, 126, 132, 147, 169
 spin, 106, 117, 118
 stan związany, 15, 21, 104, 115
 strumień prawdopodobieństwa, 40
 supersieć, 39, 115, 117
 szereg Taylora, 45, 47, 170, 173, 174

 trójdziagonalny układ równań, 91
 twierdzenie Gershgorina, 31, 162
 twierdzenie Löwnera, 81, 166
 twierdzenie Martina–Deana, 29, 52, 54, 101, 164
 twierdzenie QR Wilkinsona, 63, 64
 twierdzenie Schura, 166
 twierdzenie Sylwestera, 29, 93, 163

 wartość własna, 15, 16, 21, 23, 29, 30, 34, 42, 49, 52, 54, 55, 59, 63, 72, 76, 78, 89, 91, 93, 94, 96, 101, 103, 112, 113, 116, 126, 161–166, 179–181
 warunki brzegowe, 21–23, 31, 42, 48, 49, 51, 100, 101, 111–114, 142, 148
 wektor własny, 29, 31, 42, 49, 54, 55, 59, 72, 76, 77, 81, 84, 88, 89, 91, 94, 96, 102, 116, 124, 126, 161, 162, 166, 184

widmo częstości własnych, 101
wielomian charakterystyczny, 31, 34, 37,
59, 167
wyznacznik macierzy, 34, 105, 162, 167,
175

zagadnienie radialne, 24, 108
zagadnienie własne, 15, 104, 108, 116