

Uniwersytet Śląski w Katowicach
Wydział Matematyki, Fizyki i Chemii
Kierunek matematyka
Praca dyplomowa magisterska

AUTOR: DOMINIKA DĄBROWSKA

SUMY KWADRATÓW WIELOMIANÓW WYMIERNYCH

Promotor: dr Paweł Gładki

Katowice 2013

ROZDZIAŁ 1

Wstęp

W niniejszej pracy zajmujemy się przedstawianiem wielomianów o współczynnikach wymiernych f , które są sumami kwadratów wielomianów o współczynnikach rzeczywistych $f_1, \dots, f_m$, jako sum kwadratów wielomianów o współczynnikach wymiernych $g_1, \dots, g_l$. Wstępne prace z zakresu omawianego zagadnienia pochodzą z początku XX wieku, kiedy Landau [6] wykazał, że wielomian jednej zmiennej o powyższych własnościach daje się zapisać w postaci sumy kwadratów wielomianów o współczynnikach wymiernych. Twierdzenie Landaua ma też aspekt ilościowy, a mianowicie w latach siedemdziesiątych Pourchet [9] udowodnił, że można wybrać wielomiany $g_1, \dots, g_l \in \mathbb{Q}[x]$ tak, aby $l \leq 5$.

Rezultat Landaua w naturalny sposób prowadzi do następującego pytania, postawionego po raz pierwszy przez Sturmfelsa [12] w 2000 roku, a mianowicie czy analogiczne twierdzenie zachodzi nie tylko dla wielomianów jednej zmiennej $f \in \mathbb{Q}[x]$, ale ogólnie dla $f \in \mathbb{Q}[x_1, \dots, x_n]$? Odpowiedź w znacznej mierze nie jest znana. Okazuje się jednak, że jeżeli współczynniki wielomianów $f_1, \dots, f_m \in \mathbb{R}[x_1, \dots, x_n]$ nie są wybrane całkiem dowolnie, ale należą do pewnych klas rozszerzeń ciała $\mathbb{Q}$, twierdzenie Landaua daje się uogólnić. W szczególności Hillar [3] pokazał, że jeśli $f_1, \dots, f_m \in K[x_1, \dots, x_n]$, gdzie K jest totalnie rzeczywistym rozszerzeniem Galois ciała $\mathbb{Q}$ (to znaczy rozszerzeniem Galois $K = \mathbb{Q}(\theta) \supset \mathbb{Q}$ takim, że dla wszystkich $\sigma \in G(K/\mathbb{Q})$ zachodzi $\sigma(\theta) \in \mathbb{R}$, gdzie $G(K/\mathbb{Q})$ oznacza grupę Galois rozszerzenia $K \supset \mathbb{Q}$; na przykład rozszerzenie Galois $\mathbb{Q}(\sqrt{2}) \supset \mathbb{Q}$ jest totalnie rzeczywiste, gdyż jego grupa Galois składa się z dwóch automorfizmów, $id_{\mathbb{Q}(\sqrt{2})}$ oraz $\sigma : \mathbb{Q}(\sqrt{2}) \rightarrow \mathbb{Q}(\sqrt{2})$ danego wzorem $\sigma(a + b\sqrt{2}) = a - b\sqrt{2}$ i jest jasne, że $id_{\mathbb{Q}(\sqrt{2})}(\sqrt{2}) = \sqrt{2} \in \mathbb{R}$ oraz $\sigma(\sqrt{2}) = -\sqrt{2} \in \mathbb{R}$; z kolei rozszerzenie $\mathbb{Q}(\sqrt{-2}) \supset \mathbb{Q}$ nie jest totalnie rzeczywistym rozszerzeniem ciała $\mathbb{Q}$ ponieważ $id_{\mathbb{Q}(\sqrt{-2})}(\sqrt{-2}) = \sqrt{-2} \notin \mathbb{R}$), to f jest sumą nie więcej niż $2^{[K:\mathbb{Q}]+1} \cdot \binom{[K:\mathbb{Q}]+1}{2} \cdot 4m$ kwadratów w $\mathbb{Q}[x_1, \dots, x_n]$.

Przedmiotem mojej pracy będzie analiza jednego z uogólnień twierdzenia Hillara. Mianowicie w pracy [10] Quarez zauważył, że w powyższym twierdzeniu totalnie rzeczywiste rozszerzenie Galois ciała $\mathbb{Q}$ można zastąpić domknięciem Galois totalnie rzeczywistego rozszerzenia $\mathbb{Q}$. Zauważmy, że rozszerzenie $\mathbb{Q}(\sqrt[3]{2}) \supset \mathbb{Q}$ nie jest Galois (ponieważ, jak łatwo sprawdzić, $G(\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}) = \{id_{\mathbb{Q}(\sqrt[3]{2})}\}$, ale $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$), ale

trywialnie jest totalnie rzeczywiste. Z kolei rozszerzenie $\mathbb{Q}(i, \sqrt[3]{2}) = \mathbb{Q}(i + \sqrt[3]{2}) \supset \mathbb{Q}$, w którym znajdują się wszystkie pierwiastki wielomianu $x^3 - 2$ i które wobec tego jest najmniejszym rozszerzeniem Galois ciała $\mathbb{Q}$ zawierającym podciało $\mathbb{Q}(\sqrt[3]{2})$, oczywiście nie jest totalnie rzeczywiste; tym samym widzimy, że rezultat Quareza istotnie uogólnia wcześniejszą pracę Hillara. Quarez poprawia również oszacowanie podane przez Hillara dotyczące liczby składników w reprezentacji f jako sumy kwadratów wielomianów wymiennych i pokazuje, że jeśli $f \in \mathbb{Q}[x_1, \dots, x_n]$ jest sumą m kwadratów wielomianów o współczynnikach z domknięcia Galois K pewnego totalnie rzeczywistego rozszerzenia ciała $\mathbb{Q}$, to f jest sumą nie więcej, niż $m + 4\lceil \frac{m}{4} \rceil ([K : \mathbb{Q}] - 1)$ kwadratów w $\mathbb{Q}[x_1, \dots, x_n]$.

Niniejsza praca składa się z czterech rozdziałów. W pierwszym z nich zapoznajemy się z kilkoma kanonicznymi rozkładami macierzy o szczególnych postaciach, w szczególności macierzy symetrycznych i dodatnio określonych oraz macierzy symetrycznych o nieosobliwych minorach głównych. Punktem wyjścia do prowadzonych tam rozważań jest rozkład LU (*ang. lower-triangular, upper-triangular*) macierzy jako iloczynu macierzy dolnotrójkatnej i górnotrójkatnej. Pokazujemy, że przy bardzo ogólnym założeniu nieosobliwości minorów głównych danej macierzy, istnienie jej jednoznacznie wyznaczony rozkład LU – jest to dość powszechnie znany rezultat, który, ze względu na późniejsze zastosowania, udowadniamy w szczególnym przypadku dotyczącym macierzy o współczynnikach wymiennych. W pozostałej części tego rozdziału wykorzystujemy udowodnione twierdzenie do otrzymania innych, bardziej specjalizowanych rozkładów macierzy, w szczególności pokazujemy, iż każda dodatnio określona macierz o współczynnikach wymiennych daje się jednoznacznie rozłożyć na iloczyn macierzy LDL^T , gdzie L jest macierzą dolnotrójkatną o samych jedynkach na głównej przekątnej, zaś D jest macierzą diagonalną o dodatnich współrzędnych – będzie to najważniejsze twierdzenie z tego rozdziału wykorzystywane później w dowodzie głównego rezultatu pracy. Przy okazji dowodzimy też kilku pobocznych faktów dotyczących macierzy, na przykład pokazujemy, iż macierz symetryczna o dodatnich wartościach własnych jest dodatnio określona. Prezentowane dowody są całkowicie elementarne i w zasadzie można je traktować jako wariację na temat metody eliminacji Gaussa, a najbardziej zaawansowanym technicznie twierdzeniem na jakie się powołujemy jest twierdzenie spektralne w wersji dla rzeczywistych macierzy symetrycznych znanej z kursowego wykładu algebry liniowej.

Rozdział drugi poświęcony jest pierwiastkom rzeczywistym wielomianów i metodzie Hermite’a ich zliczania. Zastosowanie metody Hermite’a pozwala udowodnić twierdzenie dotyczące rzędu i sygnatury macierzy będącej sumą kwadratów wielomianów wymiennych. Rozdział

trzeci przedstawia algebraiczną metodę Hillara określenia liczby kwadratów wielomianów o współczynnikach wymiernych użytych w dowodzie głównego twierdzenia. W rozdziale tym wykorzystujemy twierdzenie Lagrange'a o 4 kwadratach. Dowód Quareza w twierdzeniu na liczbę kwadratów wielomianów wymiernych użytych przy zapisie wielomianu o współczynnikach wymiernych, który jest sumą kwadratów wielomianów z pewnego totalnie rzeczywistego rozszerzenia Galois jest zmodyfikowaną wersją dowodu Hillara, jednak zapewnia on lepsze oszacowanie liczby składników. Rozdział ten jest najważniejszym elementem całej pracy i stanowi podsumowanie rozważań dotyczących sum kwadratów wielomianów wymiernych. Ostatnia część pracy przedstawia przykłady oszacowań przedstawiania sum kwadratów wielomianów wymiernych bazujących na dekompozycji LU oraz na metodzie Hillara.

ROZDZIAŁ 2

Rozkłady macierzy kwadratowych

W rozdziale tym przedstawiamy szereg twierdzeń dotyczących rozkładu macierzy symetrycznych na iloczyn macierzy o pewnych szczególnych postaciach. Pierwszy z prezentowanych wyników dotyczy rozkładu macierzy na iloczyn macierzy górnotrójkątnej i dolnotrójkątnej; pokażemy, że przy pewnych dodatkowych założeniach rozkład taki jest jednoznaczny. W dalszej kolejności pokażemy, w jaki sposób rezultat ten można poprawić i ostatecznie uzyskać rozkład danej macierzy G na iloczyn macierzy dolnotrójkątnej o samych jedynkach na przekątnej L , macierzy diagonalnej D i macierzy L^T , przy czym współrzędne macierzy L i D mogą być wybrane z tego samego ciała, z którego pochodzą współrzędne macierzy G . Tak sformułowane twierdzenie jest w istocie wariantem znanego z literatury rozkładu macierzy Cholesky'ego, przy czym interesować nas będzie wyłącznie przypadek, w którym współrzędne wszystkich rozważanych macierzy pochodzą z ciała liczb wymiernych $\mathbb{Q}$. Opisane w niniejszym rozdziale rozkłady zostaną wykorzystane w późniejszych rozważaniach, gdzie sumy kwadratów będziemy zapisywali jako wartości pewnych funkcjonałów dwuliniowych zadanych macierzowo – aby móc wygodnie wykonywać obliczenia na macierzach takich funkcjonałów, niezbędne okażą się rozkłady dyskutowane w niniejszym rozdziale. Prezentowane twierdzenia zostały zaczerpnięte z podręcznika Watkinsa [11], a ich dowody nieznacznie dostosowane do naszych zastosowań.

2.1. TWIERDZENIE ([11, Theorem 1.7.19]). *Niech $G \in \mathbb{Q}_n^n$ będzie macierzą, której wszystkie minory główne są nieosobliwe. Wówczas istnieją jednoznacznie wyznaczone macierze dolnotrójkątna $L \in \mathbb{Q}_n^n$ o samych jedynkach na głównej przekątnej i górnotrójkątna $U \in \mathbb{Q}_n^n$ takie, że:*

$$G = LU.$$

Rozkład macierzy G opisany w powyższym twierdzeniu będziemy nazywać rozkładem LU (z ang. *lower-triangular, upper-triangular*).

DOWÓD. Ustalmy wektor $b \in \mathbb{Q}^n$ i rozważmy układ równań:

$$(2.1) \quad Gx = b, x \in \mathbb{Q}^n.$$

Przypomnijmy, że chcąc rozwiązać układ (2.1) za pomocą znanego z kursowego wykładu algebry algorytmu eliminacji Gaussa, wybierając

w kolejnych wierszach niewiadome bazowe i eliminując je z pozostałych wierszy, sprowadzamy układ (2.1) do postaci:

$$(2.2) \quad Ux = y$$

dla pewnego $y \in \mathbb{Q}^n$, gdzie $U \in \mathbb{Q}_n^n$ jest macierzą górnątrójkątną. Ponieważ wszystkie minory główne macierzy G są nieosobliwe, możemy przyjąć, że macierz U ma same jedynki na głównej przekątnej. Zauważmy, że wektor $y \in \mathbb{Q}^n$ został otrzymany z wektora $b \in \mathbb{Q}^n$ przez zastosowanie tych samych operacji elementarnych, jakie zastosowaliśmy do wierszy macierzy G aby otrzymać macierz U . Wykonanie każdej z takich operacji elementarnych jest równoważne pomnożeniu wektora b z lewej strony przez pewną elementarną macierz górnątrójkątną. Z kolei iloczyn elementarnych macierzy górnątrójkątnych jest macierzą górnątrójkątną o samych jedynkach na głównej przekątnej, możemy więc zapisać:

$$(2.3) \quad y = U_1 b$$

dla pewnej macierzy górnątrójkątnej $U_1 \in \mathbb{Q}_n^n$ o samych jedynkach na głównej przekątnej. Dalej, chcąc rozwiązać równanie (2.3) względem b , wykonujemy szereg operacji elementarnych na wierszach macierzy U_1 począwszy od ostatniego a skończywszy na pierwszym, a więc innymi słowy mnożymy (2.3) lewostronnie przez pewną macierz dolnątrójkątną $L \in \mathbb{Q}_n^n$ o samych jedynkach na głównej przekątnej i otrzymujemy:

$$(2.4) \quad Ly = b.$$

Łącząc równania (2.2) i (2.4) otrzymujemy $LUx = b$, a zatem $LUx = Gx$ i w konsekwencji $G = LU$. Dowodzi to istnienia stosownego rozkładu.

Aby dowieść jednoznaczności rozważmy równość $G = LU$ w szczególności:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ l_{21} & 1 & \cdots & 0 \\ l_{31} & l_{32} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \cdots & 1 \end{bmatrix} \cdot \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ 0 & 0 & \cdots & u_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{bmatrix}$$

Pierwszy rząd macierzy L ma tylko jeden niezerowy wyraz. Mnożąc pierwszy rząd macierzy L przez j -tą kolumnę macierzy U otrzymujemy:

$$a_{1j} = 1u_{1j} + 0u_{2j} + 0u_{3j} + \cdots + 0u_{nj},$$

a zatem $a_{1j} = u_{1j}$. Widzimy więc, że pierwszy rząd macierzy U jest jednoznacznie wyznaczony przez u_{1j} .

Zauważmy, że przy okazji wyznaczyliśmy też pierwszą kolumnę macierzy U , a jedynym niezerowym wyrazem jest u_{11} . Mnożąc i -ty wiersz macierzy L przez pierwszą kolumnę U otrzymujemy, że:

$$a_{i1} = l_{i1}u_{11}, \quad i = 2, \dots, n.$$

Założenie, że macierz G jest nieosobliwa powoduje, że U jest również nieosobliwa. Stąd $u_{kk} \neq 0$ dla $k = 1, \dots, n$ i w szczególności $u_{11} \neq 0$. To powoduje, że pierwsza kolumna L jest wyznaczona jednoznacznie według wzorów:

$$l_{i1} = \frac{a_{i1}}{u_{11}}, i = 2, \dots, n$$

Postępując analogicznie można wyznaczyć kolejne rzędy U i kolumny L . Istotnie, założmy, że jednoznacznie wyznaczone jest pierwsze $k-1$ rzędów macierzy U i kolumn macierzy L . Indukcyjnie pokażemy, że k -ty rząd macierzy U i k -ta kolumna L też są wyznaczone jednoznacznie. Zauważmy, że k -ty rząd macierzy L jest równy

$$[l_{k1}, l_{k2}, \dots, l_{k,k-1}, 1, 0, \dots, 0].$$

Mnożąc powyższy rząd przez j -tą kolumnę U , $j \geq k-1$, otrzymujemy kolejno:

$$\begin{aligned} a_{k1} &= l_{k1}u_{11} \\ a_{k2} &= l_{k1}u_{12} + l_{k2}u_{22} \\ &\vdots \\ a_{k,k-1} &= l_{k1}u_{1,k-1} + l_{k2}u_{2,k-1} + \dots + l_{k,k-1}u_{k-1,k-1}. \end{aligned}$$

Ponieważ u_{mj} , $m, j \in \{1, \dots, k-1\}$, są już jednoznacznie wyznaczone, powyższe wzory pozwalają jednoznacznie wyznaczyć $l_{k1}, l_{k2}, \dots, l_{k,k-1}$. Dalej, zauważmy, że k -ta kolumna macierzy U jest równa:

$$\begin{bmatrix} u_{1k} \\ \vdots \\ u_{kk} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Mnożąc j -ty wiersz macierzy L , $j \geq k-1$, przez powyższą kolumnę otrzymujemy kolejno:

$$\begin{aligned} a_{1k} &= 1u_{1k} \\ a_{2k} &= l_{21}u_{1k} + u_{2k} \\ &\vdots \\ a_{k-1,k} &= l_{k-1,1}u_{1k} + l_{k-1,2}u_{2k} + \dots + l_{k-1,k-2}u_{k-2,k} + u_{k-1,k}. \end{aligned}$$

Ponieważ l_{mj} , $m, j \in \{1, \dots, k-1\}$, są jednoznacznie wyznaczone, powyższe wzory pozwalają jednoznacznie wyznaczyć $u_{1k}, u_{2k}, \dots, u_{k-1,k}$. Na koniec, mnożąc k -ty wiersz macierzy L i k -tą kolumnę macierzy U otrzymujemy:

$$a_{kk} = l_{k1}u_{1k} + l_{k2}u_{2k} + \dots + l_{k,k-1}u_{k-1,k} + 1u_{kk}.$$

Ponieważ wyznaczyliśmy już jednoznacznie współrzędne $l_{k1}, \dots, l_{k,k-1}$ oraz $u_{1k}, \dots, u_{k-1,k}$, powyższa równość jednoznacznie określa u_{kk} . $\square$

Powyższe twierdzenie jest punktem wyjścia do dalszych rozważań, w których będziemy modyfikować rezultat Twierdzenia 2.1 zależnie od naszych potrzeb.

2.2. PRZYKŁAD. Rozpatrzmy macierz $G = \begin{bmatrix} 2 & -1 & 2 \\ 2 & 3 & -1 \\ 2 & 2 & -3 \end{bmatrix}$, której rozkładem LU będzie iloczyn macierzy odpowiednio dolnotrójkątnej o samych jedynkach na głównej przekątnej $L = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & \frac{3}{4} & 1 \end{bmatrix}$ i górnortrójkątnej $U = \begin{bmatrix} 2 & -1 & 2 \\ 0 & 4 & -3 \\ 0 & 0 & -\frac{11}{4} \end{bmatrix}$

2.3. TWIERDZENIE ([11, Theorem 1.7.28]). Niech $G \in \mathbb{Q}_n^n$ będzie macierzą, której wszystkie minory główne są nieosobliwe. Wówczas istnieją jednoznacznie wyznaczone macierze dolnotrójkątne $L \in \mathbb{Q}_n^n$ o samych jedynkach na głównej przekątnej, diagonalna $D \in \mathbb{Q}_n^n$ i górnortrójkątna $V \in \mathbb{Q}_n^n$ o samych jedynkach na głównej przekątnej takie, że:

$$G = LDV.$$

Rozkład macierzy G opisany w powyższym twierdzeniu będziemy nazywać rozkładem LDV .

DOWÓD. Wobec Twierdzenia 2.1 istnieją dolnotrójkątna macierz L o samych jedynkach na głównej przekątnej i górnortrójkątna macierz $U = [u_{ij}]$, $i, j \in \{1, \dots, n\}$, takie, że $G = LU$. Macierz G jest nieosobliwa więc, argumentując jak w poprzednim dowodzie, $u_{kk} \neq 0$, dla $k \in \{1, \dots, n\}$. Niech D będzie macierzą diagonalną o współczynnikach $u_{11}, \dots, u_{nn}$ na głównej przekątnej. Wtedy D jest nieosobliwa, a macierz D^{-1} jest macierzą diagonalną, której współczynniki na głównej przekątnej to $u_{11}^{-1}, \dots, u_{nn}^{-1}$. Niech $V = D^{-1}U$. Wówczas $G = LDV$.

Pozostaje wykazać, że rozkład ten jest jedyny. Przypuśćmy, że $G = L_1 D_1 V_1 = L_2 D_2 V_2$. Niech $U_1 = D_1 V_1$ oraz $U_2 = D_2 V_2$. Wtedy U_1 i U_2 są macierzami górnortrójkątnymi, a $G = L_1 U_1 = L_2 U_2$. Z jednoznaczności rozkładu LU mamy, że $L_1 = L_2$ oraz $U_1 = U_2$. Zatem $D_1 V_1 = D_2 V_2$ i dlatego:

$$D_2^{-1} D_1 = V_2 V_1^{-1}.$$

Ponieważ V_1 jest macierzą górnortrójkątną o samych jedynkach na głównej przekątnej, więc V_1^{-1} też. Ponieważ V_2 oraz V_1^{-1} są górnortrójkątnymi macierzami o samych jedynkach na głównych przekątnych, to ich iloczyn również. Z drugiej strony, iloczyn $D_2^{-1} D_1$ jest macierzą diagonalną. Wtedy $V_2 V_1^{-1}$ jest również macierzą górnortrójkątną i diagonalną

o samych jedynkach na głównej przekątnej, a więc po prostu $V_2 V_1^{-1} = I$, czyli $V_2 = V_1$ i w konsekwencji $D_2 = D_1$. $\square$

2.4. PRZYKŁAD. Weźmy macierz G i L jak w poprzednim przykładzie. Macierz G ulega rozkładowi LDV gdy macierz diagonalna D jest

$$\text{postaci } D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -\frac{11}{4} \end{bmatrix} \text{ a macierz górnotrójkątna } V \text{ jest następu-}$$

$$\text{jąca: } V = \begin{bmatrix} 1 & -\frac{1}{2} & 1 \\ 0 & 1 & -\frac{3}{4} \\ 0 & 0 & 1 \end{bmatrix}$$

2.5. WNIOSEK ([11, Theorem 1.7.30]). Niech $G \in \mathbb{Q}_n^n$ będzie macierzą symetryczną, której wszystkie minory główne są nieosobliwe. Wówczas istnieją jednoznacznie wyznaczone macierze dolnotrójkątna $L \in \mathbb{Q}_n^n$ o samych jedynkach na głównej przekątnej i diagonalna $D \in \mathbb{Q}_n^n$ takie, że:

$$G = LDL^T.$$

DOWÓD. Mając dany rozkład $G = LDV$ z Twierdzenia 2.3 wystarczy wykazać, że przy założeniu symetryczności macierzy G zachodzi $V = L^T$. Mamy, że $G = G^T = (LDV)^T = V^T D^T L^T$. V^T jest macierzą dolnotrójkątną o samych jedynkach na głównej przekątnej, D^T jest diagonalna, L^T jest górnotrójkątna o samych jedynkach na głównej przekątnej, a zatem $V^T D^T L^T$ jest rozkładem LDV macierzy G . Wobec jednoznaczności takiego rozkładu $V = L^T$. $\square$

2.6. PRZYKŁAD. Przykładem ilustrującym rozkład jest LDL^T jest

$$\text{dekompozycja macierzy } G = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \text{ na iloczyn macierzy dolno-}$$

$$\text{trójkątnej } L = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ i macierzy diagonalnej } D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{3}{2} & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Przeniesiemy obecnie część uzyskanych rezultatów na przypadek macierzy dodatnio określonych. Będziemy tu milcząco zakładać, że wszystkie macierze dodatnio określone są symetryczne i nie będziemy zajmować się patologicznymi przypadkami macierzy dodatnio określonych i niesymetrycznych.

2.7. UWAGA. Niech $G \in \mathbb{Q}_n^n$ będzie macierzą dodatnio określoną. Wówczas jej minory główne również są macierzami dodatnio określonymi.

DOWÓD. Przypuśćmy nie wprost, że G jest macierzą dodatnio określoną, a jej minor główny B rzędu $m < n$ nie jest dodatnio określony. Oznacza to, że istnieje wektor $x \in \mathbb{R}^m$ taki że $x^T B x \leq 0$ oraz $x \neq 0$. Rozważmy wektor $y \in \mathbb{R}^n$, który ma pierwsze m współrzędnych takie

same jak wektor x i pozostałe $n - m$ współrzędnych równych 0. Wówczas y jest wektorem niezerowym. Ponieważ $y^T G y = x^T B x \leq 0$, stąd G nie jest dodatnio określona, co daje sprzeczność. $\square$

2.8. UWAGA. Niech $G \in \mathbb{Q}_n^n$ będzie macierzą dodatnio określoną. Wówczas G jest nieosobliwa.

DOWÓD. Przypuśćmy nie wprost, że G jest macierzą osobliwą. Wówczas istnieje taki niezerowy wektor $x \in \mathbb{Q}^n$, że $Gx = 0$. Ale wówczas tym bardziej $x^T G x = 0$, więc G nie może być dodatnio określona, co doprowadza nas do sprzeczności. $\square$

Kluczowy do dalszych zastosowań okazuje się następujący rezultat, którego dowód łatwo otrzymujemy łącząc dowiedzione powyżej twierdzenia.

2.9. TWIERDZENIE ([11, Theorem 1.7.31]). Niech $G \in \mathbb{Q}_n^n$ będzie macierzą dodatnio określoną. Wówczas istnieją jednoznacznie wyznaczone macierze dolnotrójkątna $L \in \mathbb{Q}_n^n$ o samych jedynkach na głównej przekątnej i diagonalna $D \in \mathbb{Q}_n^n$ o dodatnich współczynnikach na głównej przekątnej takie, że:

$$G = LDL^T.$$

DOWÓD. Wobec Wniosku 2.5 i Uwag 2.7 i 2.8 wystarczy wykazać, że współczynniki macierzy diagonalnej D , które – zgodnie z notacją z dowodu Twierdzenia 2.3 oznaczamy przez $u_{11}, \dots, u_{nn}$ – są dodatnie. Oznaczmy przez e_i wektor, który ma na i -tej współrzędnej jedynkę, a na pozostałych zera. Ponieważ $G = LDL^T$, więc $D = L^{-1}G(L^T)^{-1}$ i wówczas:

$$\begin{aligned} u_{ii} &= e_i^T L^{-1} G (L^T)^{-1} e_i = e_i^T L^{-1} G (L^{-1})^T e_i \\ &= \underbrace{[(L^{-1})^T e_i]^T}_{x_i^T} G \underbrace{[(L^{-1})^T e_i]}_{x_i} = x_i^T G x_i. \end{aligned}$$

Ponieważ G jest dodatnio określona, więc $u_{ii} = x_i^T G x_i > 0$, dla $i \in \{1, \dots, n\}$. $\square$

Powyższe twierdzenie poprzemy następującym przykładem:

2.10. PRZYKŁAD. Macierz G postaci $G = \begin{bmatrix} 10 & -1 & 2 \\ -1 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$ jest macierzą dodatnio określoną. Wówczas ulega ona rozkładowi na macierz $L = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{10} & 1 & 0 \\ \frac{1}{5} & \frac{2}{9} & 1 \end{bmatrix}$ o samych jedynkach na głównej przekątnej oraz macierz diagonalną $D = \begin{bmatrix} 10 & 0 & 0 \\ 0 & \frac{9}{10} & 0 \\ 0 & 0 & \frac{5}{9} \end{bmatrix}$ o dodatnich współczynnikach na głównej przekątnej.

Na marginesie powyższych rezultatów odnotujmy tu jeszcze jedno twierdzenie dotyczące własności macierzy symetrycznych, które wykorzystamy w późniejszych rozważaniach, aczkolwiek którego dowód jest całkowicie niezależny od przedstawionego powyżej wykładu.

2.11. TWIERDZENIE. *Niech $G \in \mathbb{Q}_n^n$ będzie macierzą symetryczną, której wszystkie wartości własne są dodatnie. Wówczas G jest macierzą dodatnio określoną.*

DOWÓD. Macierz G możemy rozpatrywać jako macierz funkcjonału dwuliniowego symetrycznego $\xi : \mathbb{Q}^n \times \mathbb{Q}^n \rightarrow \mathbb{Q}$. Wobec twierdzenia spektralnego w wersji znanej z kursowego wykładu algebry liniowej (por. na przykład [4, Twierdzenie IX.3.2, str. 268]) istnieje baza ortogonalna przestrzeni $\mathbb{Q}^n$ złożona z wektorów własnych macierzy G , a więc istnieją macierze ortogonalna $P \in \mathbb{Q}_n^n$ oraz diagonalna $D \in \mathbb{Q}_n^n$, której współrzednymi na głównej przekątnej są wartości własne $d_1, \dots, d_n$ macierzy G takie, że $G = PDP^T$. Ustalmy teraz wektor $x \in \mathbb{Q}^n$ i policzmy:

$$\begin{aligned} x^T G x &= x^T P D P^T x \\ &= \underbrace{(P^T x)^T}_{y^T} D \underbrace{P^T x}_y \\ &= y_1^2 d_1 + \dots + y_n^2 d_n, \end{aligned}$$

gdzie $y = [y_1, \dots, y_n]^T \in \mathbb{Q}^n$. Ponieważ $d_1 > 0, \dots, d_n > 0$, więc jest jasne, że $x^T G x > 0$ i tym samym G jest dodatnio określona. $\square$

ROZDZIAŁ 3

Metoda Hermite’a zliczania pierwiastków wielomianów rzeczywistych

W niniejszym rozdziale udowodnimy twierdzenie Hermite’a pozwalające obliczać liczbę rzeczywistych i zespolonych pierwiastków wielomianu jednej zmiennej o współczynnikach rzeczywistych bez ich *explicit* wyznaczania. Metoda użyta w dowodzie będzie zarazem ilustracją do metod, których będziemy używać w następnych rozdziale, gdzie dość często będziemy posługiwać się rachunkiem macierzowym do skrócenia i uproszczenia notacji w obliczeniach o typowo algebraicznym charakterze. Prezentowany materiał został opracowany w oparciu o notatki do wykładu Coste’a [2] z rzeczywistej geometrii algebraicznej.

Niech $u \in \mathbb{R}[X]$ będzie wielomianem stopnia r , niech $\theta_1, \dots, \theta_r$ oznaczają wszystkie jego pierwiastki zespolone (liczone wraz z krotnościami). Niech

$$\alpha_{st} = \sum_{j=1}^r \theta_j^{s+t-2},$$

gdzie $s, t \in \mathbb{N}$, będą sumami kolejnych potęg pierwiastków u . Zauważmy, że na sumy te możemy patrzeć jak na wielomiany symetryczne zmiennych $\theta_1, \dots, \theta_r$ o współczynnikach całkowitych. Zauważmy też, że przy przyjętych oznaczeniach $\alpha_{st} = \alpha_{s't'}$, o ile tylko $s + t = s' + t'$ – wybór tej, na pozór dość ekstrawaganckiej, notacji znajdzie swe uzasadnienie w kolejnym rozdziale, w którym będziemy przeprowadzać obliczenia z wykorzystaniem sum α_{st} .

3.1. UWAGA. $\alpha_{st} \in \mathbb{R}$, dla $s, t \in \mathbb{N}$.

DOWÓD. Niech $u(x) = a_0 x^r + \dots + a_r$. Wówczas $u(x) = a_0(x - \theta_1) \cdot \dots \cdot (x - \theta_r)$ i wobec wzoru na pochodną iloczynu otrzymujemy:

$$\begin{aligned} u'(x) &= a_0 [(x - \theta_2)(x - \theta_3) \dots (x - \theta_r) + \\ &\quad + (x - \theta_1)(x - \theta_3) \dots (x - \theta_r) \\ &\quad \vdots \\ &\quad + (x - \theta_2)(x - \theta_3) \dots (x - \theta_{r-1})] \end{aligned}$$

a zatem wykorzystując znany wzór na sumę szeregu geometrycznego otrzymujemy

$$\begin{aligned}
 (3.5) \quad \frac{u'(x)}{u(x)} &= \frac{1}{x - \theta_1} + \frac{1}{x - \theta_2} + \dots + \frac{1}{x - \theta_r} \\
 &= \frac{1}{x} \cdot \frac{1}{1 - \frac{\theta_1}{x}} + \frac{1}{x} \cdot \frac{1}{1 - \frac{\theta_2}{x}} + \dots + \frac{1}{x} \cdot \frac{1}{1 - \frac{\theta_r}{x}} \\
 &= \frac{1}{x} \sum_{i=0}^{\infty} \left(\frac{\theta_1}{x}\right)^i + \frac{1}{x} \sum_{i=0}^{\infty} \left(\frac{\theta_2}{x}\right)^i + \dots + \frac{1}{x} \sum_{i=0}^{\infty} \left(\frac{\theta_r}{x}\right)^i \\
 &= \sum_{i=0}^{\infty} \theta_1^i \left(\frac{1}{x}\right)^{i+1} + \sum_{i=0}^{\infty} \theta_2^i \left(\frac{1}{x}\right)^{i+1} + \dots + \sum_{i=0}^{\infty} \theta_r^i \left(\frac{1}{x}\right)^{i+1} \\
 &= \sum_{i=0}^{\infty} \left(\frac{1}{x}\right)^{i+1} \cdot (\theta_1^i + \theta_2^i + \dots + \theta_r^i) \\
 &= \sum_{i=0}^{\infty} \left(\frac{1}{x}\right)^{i+1} \cdot \alpha_{s_i t_i}
 \end{aligned}$$

dla pewnych $s_i, t_i \in \mathbb{N}$ takich, że $s_i + t_i = i$. Z drugiej strony

$$u'(x) = a_0 r x^{r-1} + a_1(r-1)x^{r-2} + \dots + a_{r-1}$$

skąd możemy w inny sposób wyrazić iloraz $\frac{u'(x)}{u(x)}$:

$$(3.6) \quad \frac{u'(x)}{u(x)} = \frac{a_0 r x^{r-1} + a_1(r-1)x^{r-2} + \dots + a_{r-1}}{a_0 x^r + a_1 x^{r-1} + \dots + a_r}.$$

Porównując teraz równania (3.5) i (3.6) otrzymujemy:

$$\frac{a_0 r x^{r-1} + a_1(r-1)x^{r-2} + \dots + a_{r-1}}{a_0 x^r + a_1 x^{r-1} + \dots + a_r} = \sum_{i=0}^{\infty} \left(\frac{1}{x}\right)^{i+1} \cdot \alpha_{s_i t_i}$$

dla pewnych $s_i, t_i \in \mathbb{N}$ takich, że $s_i + t_i = i$. Wobec tego

$$\begin{aligned}
 &a_0 r x^{r-1} + a_1(r-1)x^{r-2} + \dots + a_{r-1} \\
 &= (a_0 x^r + a_1 x^{r-1} + \dots + a_r) \cdot \sum_{i=0}^{\infty} \left(\frac{1}{x}\right)^{i+1} \cdot \alpha_{s_i t_i} \\
 &= \sum_{i=0}^{\infty} a_0 x^{r-i-1} \alpha_{s_i t_i} + \sum_{i=0}^{\infty} a_1 x^{r-i-2} \alpha_{s_i t_i} + \dots + \sum_{i=0}^{\infty} a_r x^{-i-1} \alpha_{s_i t_i} \\
 &= (a_0 x^{r-1} \alpha_{s_0 t_0} + a_0 x^{r-2} \alpha_{s_1 t_1} + a_0 x^{r-3} \alpha_{s_2 t_2} + \dots) \\
 &+ (a_1 x^{r-2} \alpha_{s_0 t_0} + a_1 x^{r-3} \alpha_{s_1 t_1} + a_1 x^{r-4} \alpha_{s_2 t_2} + \dots) \\
 &\vdots \\
 &+ (a_r x^{-1} \alpha_{s_0 t_0} + a_r x^{-2} \alpha_{s_1 t_1} + a_r x^{-3} \alpha_{s_2 t_2} + \dots)
 \end{aligned}$$

$$\begin{aligned}
&= a_0 \alpha_{s_0 t_0} x^{r-1} \\
&+ (a_0 \alpha_{s_1 t_1} + a_1 \alpha_{s_0 t_0}) x^{r-2} \\
&+ (a_0 \alpha_{s_2 t_2} + a_1 \alpha_{s_1 t_1} + a_2 \alpha_{s_0 t_0}) x^{r-3} \\
&\vdots \\
&+ (a_0 \alpha_{s_{r-1} t_{r-1}} + a_1 \alpha_{s_{r-2} t_{r-2}} + \dots + a_{r-1} \alpha_{s_0 t_0}) \\
&+ (a_0 \alpha_{s_r t_r} + a_1 \alpha_{s_{r-1} t_{r-1}} + \dots + a_r \alpha_{s_0 t_0}) x^{-1} \\
&+ (a_0 \alpha_{s_{r+1} t_{r+1}} + a_1 \alpha_{s_r t_r} + \dots + a_r \alpha_{s_1 t_1}) x^{-2} \\
&\vdots \\
&+ (a_0 \alpha_{s_{r+i} t_{r+i}} + a_1 \alpha_{s_{r+i-1} t_{r+i-1}} + \dots + a_r \alpha_{s_i t_i}) x^{-i-1} \\
&\vdots
\end{aligned}$$

Porównując współczynniki przy kolejnych potęgach x po obydwu stronach powyższego równania otrzymujemy zatem:

$$\begin{aligned}
a_0 r &= a_0 \alpha_{s_0 t_0} \\
a_1(r-1) &= a_0 \alpha_{s_1 t_1} + a_1 \alpha_{s_0 t_0} \\
a_2(r-2) &= a_0 \alpha_{s_2 t_2} + a_1 \alpha_{s_1 t_1} + a_2 \alpha_{s_0 t_0} \\
&\vdots \\
a_{r-1} &= a_0 \alpha_{s_{r-1} t_{r-1}} + a_1 \alpha_{s_{r-2} t_{r-2}} + \dots + a_{r-1} \alpha_{s_0 t_0} \\
0 &= a_0 \alpha_{s_{r+i} t_{r+i}} + a_1 \alpha_{s_{r+i-1} t_{r+i-1}} + \dots + a_r \alpha_{s_i t_i} \text{ dla } i \geq 0.
\end{aligned}$$

skąd bezpośrednio dostajemy następujące wzory rekurencyjne na $\alpha_{s_i t_i}$, dla pewnych $s_i, t_i \in \mathbb{N}$ takich, że $s_i + t_i = i$:

$$\begin{aligned}
(3.7) \quad \alpha_{s_0 t_0} &= r \\
\alpha_{s_1 t_1} &= -\frac{a_1}{a_0} \\
\alpha_{s_2 t_2} &= -\frac{\alpha_{s_1 t_1} a_1 + 2a_2}{a_0} \\
&\vdots \\
\alpha_{s_i t_i} &= -\frac{\alpha_{s_{i-1} t_{i-1}} a_1 + \alpha_{s_{i-2} t_{i-2}} a_2 + \dots + \alpha_{s_1 t_1} a_{i-1} + i a_i}{a_0} \text{ dla } i \leq r \\
\alpha_{s_i t_i} &= -\frac{\alpha_{s_{i-1} t_{i-1}} a_1 + \alpha_{s_{i-2} t_{i-2}} a_2 + \dots + \alpha_{s_{i-r} t_{i-r}} a_r}{a_0} \text{ dla } i > r.
\end{aligned}$$

Ponieważ współczynniki $a_0, \dots, a_r$ wielomianu u są liczbami rzeczywistymi, ze wzorów (3.7) wynika w szczególności, że $\alpha_{s_i t_i} \in \mathbb{R}$, dla pewnych $s_i, t_i \in \mathbb{N}$ takich, że $s_i + t_i = i$ oraz $i \in \mathbb{N}$. Z kolei ponieważ $s_i, t_i \in \mathbb{N}$ zostały wybrane dowolnie spośród liczb naturalnych spełniających warunek $s_i + t_i = i$, oznacza to, że $\alpha_{st} \in \mathbb{R}$, dla $s, t \in \mathbb{N}$, czego należało dowieść. $\square$

Rozważmy formę kwadratową $Q : \mathbb{R}^r \rightarrow \mathbb{R}$ daną wzorem

$$Q(U) = U^T G U, \quad U \in \mathbb{R}^r,$$

gdzie $G \in \mathbb{R}_r^r$ jest następującą macierzą:

$$G = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1r} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{r1} & \alpha_{r2} & \cdots & \alpha_{rr} \end{bmatrix}.$$

Zauważmy, że w świetle Uwagi 3.1, Q jest dobrze określoną formą o współczynnikach rzeczywistych.

Głównym rezultatem niniejszego rozdziału jest następujące twierdzenie:

3.2. TWIERDZENIE ([2, Theorem 1.14]). *Sygnatura formy Q jest równa liczbie różnych rzeczywistych pierwiastków wielomianu u , a jej rząd równy jest liczbie różnych pierwiastków u (rzeczywistych i zespolonych).*

Występujący w wypowiedzi twierdzenia rząd formy kwadratowej Q jest oczywiście równy rządowi macierzy G . Zauważmy, że rząd formy kwadratowej możemy wygodnie obliczać zgodnie z następującą obserwacją:

3.3. UWAGA. Forma kwadratowa $Q : \mathbb{R}^r \rightarrow \mathbb{R}$ jest rzędu s wtedy i tylko wtedy, gdy istnieją formy liniowe $L_1, \dots, L_s \in F[x_1, \dots, x_r]$, gdzie F jest pewnym rozszerzeniem ciała $\mathbb{R}$, liniowo niezależne nad F i takie, że

$$Q(U) = L_1^2(U) + \dots + L_s^2(U), \quad U \in \mathbb{R}^r.$$

DOWÓD. Załóżmy najpierw, że Q jest formą rzędu s . Wówczas, na mocy twierdzenia Lagrange'a o diagonalizacji formy kwadratowej, w pewnej bazie $\mathcal{B}$ przestrzeni $\mathbb{R}^r$ Q przyjmuje postać:

$$Q(x_1, \dots, x_r) = x_1^2 + \dots + x_p^2 - x_{p+1}^2 + \dots - x_{p+q}^2$$

gdzie $p + q = s$. Przyjmijmy (ciągle w bazie $\mathcal{B}$)

$$L_j(x_1, \dots, x_r) = x_j$$

dla $1 \leq j \leq p$, oraz

$$L_j(x_1, \dots, x_r) = ix_j.$$

Wówczas formy $\{L_j\}_{j=1}^s$ są liniowo niezależne, oraz Q ma postać jak w tezie uwagi.

Załóżmy teraz, że forma Q jest postaci

$$Q(U) = L_1^2(U) + \dots + L_s^2(U)$$

dla s liniowo niezależnych form liniowych $L_1, \dots, L_s \in F[x_1, \dots, x_r]$, gdzie F jest pewnym rozszerzeniem ciała $\mathbb{R}$. Ponieważ wszystkie odwzorowania L_j są formami liniowymi, to każde z nich da się zapisać w

postaci:

$$L_j = a_{j1}x_1 + \dots + a_{jr}x_r$$

gdzie a_{jk} są pewnymi elementami F . Z liniowej niezależności form L_j wynika, że macierz

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1r} \\ \vdots & \vdots & \ddots & \vdots \\ a_{s1} & a_{s2} & \dots & a_{sr} \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

wymiaru $r \times r$ (gdzie ostatnie $r-s$ wierszy zawiera same zera) jest rzędu s . Zauważmy, że macierz $A^T A$ jest macierzą formy Q . Pokażemy, że macierze A i $A^T A$ mają ten sam rząd. W tym celu wystarczy pokazać, że ich jądra są takie same. Jeżeli $Ax = 0$, to oczywiście $A^T Ax = 0$. Jeżeli zaś $A^T Ax = 0$ to również $x^T A^T Ax = 0$, czyli $|Ax|^2 = 0$, stąd $Ax = 0$. Q ma rząd równy rzędowi swojej macierzy $A^T A$, czyli s . $\square$

Przechodzimy teraz do dowodu Twierdzenia 3.2.

DOWÓD. Zauważmy, że:

$$\begin{aligned} G &= \begin{bmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1,r-1} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{r-1,1} & \alpha_{r-1,2} & \dots & \alpha_{r-1,r-1} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 & \dots & 1 \\ \theta_1 & \theta_2 & \dots & \theta_r \\ \vdots & \vdots & \ddots & \vdots \\ \theta_1^{r-1} & \theta_2^{r-1} & \dots & \theta_r^{r-1} \end{bmatrix} \cdot \begin{bmatrix} 1 & \theta_1 & \dots & \theta_1^{r-1} \\ 1 & \theta_2 & \dots & \theta_2^{r-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \theta_r & \dots & \theta_r^{r-1} \end{bmatrix}. \end{aligned}$$

Niech

$$G_1 = \begin{bmatrix} 1 & \theta_1 & \dots & \theta_1^{r-1} \\ 1 & \theta_2 & \dots & \theta_2^{r-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \theta_r & \dots & \theta_r^{r-1} \end{bmatrix}.$$

Wówczas

$$G = G_1^T G_1.$$

Oznaczmy

$$L_j(U) = U_1 + \theta_j U_2 + \dots + \theta_j^{r-1} U_r \in \mathbb{C}[U_1, \dots, U_r], \quad j \in \{1, \dots, r\}.$$

Tym samym

$$\begin{aligned} (3.8) \quad Q(U) &= U^T G U = U^T G_1^T G_1 U \\ &= (G_1 U)^T (G_1 U) \\ &= L_1^2(U) + \dots + L_r^2(U), \quad U \in \mathbb{R}^r. \end{aligned}$$

Zmieniając ewentualnie numerację możemy założyć, że $\theta_1, \dots, \theta_s$ są wszystkimi parami różnymi pierwiastkami u , $s \leq r$. Zauważmy, że formy liniowe

$$L_1, L_2, \dots, L_s$$

są liniowo niezależne nad $\mathbb{C}$. Istotnie, założmy, że dla pewnych skalarów $c_1, \dots, c_s \in \mathbb{C}$ zachodzi:

$$c_1 L_1(U) + c_2 L_2(U) + \dots + c_s L_s(U) = 0$$

dla wszystkich $U \in \mathbb{C}^r$. W szczególności biorąc za U kolejno wektory e_j o jedynce na j -tej współrzędnej i zerach na pozostałych, $j \in \{1, \dots, s\}$, otrzymujemy układ równań:

$$\begin{cases} c_1 & + & c_2 & + & \dots & + & c_r & = & 0 \\ c_1 \theta_1 & + & c_2 \theta_2 & + & \dots & + & c_r \theta_r & = & 0 \\ \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ c_1 \theta_1^{r-1} & + & c_2 \theta_2^{r-1} & + & \dots & + & c_r \theta_r^{r-1} & = & 0 \end{cases}$$

o macierzy

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ \theta_1 & \theta_2 & \dots & \theta_s \\ \vdots & \vdots & \ddots & \vdots \\ \theta_1^{s-1} & \theta_2^{s-1} & \dots & \theta_s^{s-1} \end{bmatrix},$$

której wyznacznik jest równy $\prod_{1 \leq j_1 < j_2 \leq s} (\theta_{j_2} - \theta_{j_1})$, a więc $\neq 0$, gdyż $\theta_{j_1} \neq \theta_{j_2}$ dla $j_1 \neq j_2$. Wobec Uwagi 3.3 rząd formy Q jest więc równy liczbie s parami różnych pierwiastków wielomianu u .

Przypomnijmy, że sygnatura formy kwadratowej $Q(U)$ zmiennych $U = (U_1, \dots, U_r)$ o współczynnikach rzeczywistych jest obliczana następująco: rozkładamy formę $Q(U)$ jako różnicę sum kwadratów

$$Q(U) = \sum_{i=1}^p L_i(U)^2 - \sum_{i=p+1}^{p+q} L_i(U)^2,$$

gdzie L_i są niezależnymi formami liniowymi o współczynnikach w $\mathbb{R}$, i wówczas sygnatura $Q(U)$ jest równa różnicy $p - q$. Dalej zauważmy, że

jeśli θ_{j_1} oraz θ_{j_2} oznaczają 2 sprzężone pierwiastki wielomianu u , to:

$$\begin{aligned}
 (3.9) \quad L_{j_1}^2(U) + L_{j_2}^2(U) &= (U_1 + \theta_{j_1}U_2 + \dots + \theta_{j_1}^{r-1}U_r)^2 \\
 &+ (U_1 + \theta_{j_2}U_2 + \dots + \theta_{j_2}^{r-1}U_r)^2 \\
 &= \sum_{l=1}^r U_l^2 (\theta_{j_1}^{l-1})^2 + \sum_{k=1, l=1, k \neq l}^r U_l \theta_{j_1}^{l-1} U_k \theta_{j_1}^{k-1} \\
 &- \sum_{l=1}^r (U_l \theta_{j_2}^{l-1})^2 + 2 \sum_{k \neq l}^r U_l \theta_{j_2}^{l-1} U_k \theta_{j_2}^{k-1} \\
 &= \sum_{l=1}^r 2U_l^2 \operatorname{Re}(\theta_{j_1}^{l-1})^2 + 2 \sum_{k \neq l}^r U_l \operatorname{Re}(\theta_{j_1}^{l-1}) U_k \operatorname{Re}(\theta_{j_1}^{k-1}) \\
 &= 2 \sum_{i=1}^r U_i^2 \operatorname{Re}(\theta_{j_1}^{i-1})^2 - 2 \sum_{i=1}^r U_i^2 \operatorname{Im}(\theta_{j_1}^{i-1})^2 \\
 &- \sum_{k \neq l}^r U_l \operatorname{Im}(\theta_{j_1}^{l-1}) \operatorname{Im}(\theta_{j_1}^{k-1}) \\
 &+ \sum_{k \neq l}^r U_l \operatorname{Re}(\theta_{j_1}^{l-1}) U_k \operatorname{Re}(\theta_{j_1}^{k-1}) \\
 &= \sum_{k \neq l}^r U_l \operatorname{Re}(\theta_{j_1}^{l-1}) U_k \operatorname{Re}(\theta_{j_1}^{k-1}) + 2(U_1 U_2 \operatorname{Re}(\theta_{j_1})) \\
 &+ U_1 U_3 \operatorname{Re}(\theta_{j_1}^2) + \dots + U_1 U_r \operatorname{Re}(\theta_{j_1}^{r-1}) \\
 &= 2[U_1 + \operatorname{Re}(\theta_{j_1})U_2 + \dots + \operatorname{Re}(\theta_{j_1}^{r-1})U_r]^2 \\
 &- [\operatorname{Im}(\theta_{j_1})U_2 + \dots + \operatorname{Im}(\theta_{j_1}^{r-1})U_r]^2.
 \end{aligned}$$

i ponieważ współczynniki form liniowych po prawej stronie równości (3.9) są rzeczywiste, sumy kwadratów form L_{j_1}, L_{j_2} w rozkładzie (3.8) odpowiadające parom sprzężonych pierwiastków zespolonych wielomianu u nie mają wpływu na sygnaturę. Z drugiej strony każdy rzeczywisty pierwiastek wielomianu u , wobec równości (3.8), zwiększa sygnaturę o 1. Tym samym sygnatura Q jest równa liczbie różnych pierwiastków rzeczywistych u . $\square$

3.4. PRZYKŁAD. Niech wielomian $u = (X - 2)(X^2 - 2x + 3)$. Pierwiastkami wielomianu u są $-i\sqrt{2} + 1$, $i\sqrt{2} + 1$, 2. Wtedy macierz $G = \begin{bmatrix} 3 & 4 & 2 \\ 4 & 2 & -2 \\ 2 & -2 & 2 \end{bmatrix}$, natomiast macierz $G_1 = \begin{bmatrix} 1 & -i\sqrt{2} + 1 & -2i\sqrt{2} - 1 \\ 1 & i\sqrt{2} + 1 & 2i\sqrt{2} - 1 \\ 1 & 2 & 4 \end{bmatrix}$

ROZDZIAŁ 4

Metoda Hillara przedstawienia wielomianów w postaci sumy kwadratów

W tym rozdziale, który jest główną częścią pracy, zaprezentujemy zmodyfikowaną przez Quareza metodę Hillara przedstawiania wielomianów o współczynnikach wymiernych, które są sumami kwadratów wielomianów o współczynnikach rzeczywistych, jako sum kwadratów wielomianów o współczynnikach wymiernych. Prezentowane dowody pochodzą z pracy Quareza [10] i korzystają z przygotowanego we wcześniejszych rozdziałach warsztatu pojęciowego. Zaczniemy od doprecyzowania kilku podstawowych definicji:

4.1. DEFINICJA. *Rozszerzenie algebraiczne* $K = \mathbb{Q}(\theta) \supset \mathbb{Q}$, *spełniające warunek*

$$\forall \sigma \in G(K/\mathbb{Q}) \quad [\sigma(\theta) \in \mathbb{R}],$$

gdzie $G(K/\mathbb{Q})$ oznacza grupę Galois rozszerzenia $K \supset \mathbb{Q}$ nazywamy rozszerzeniem totalnie rzeczywistym.

4.2. UWAGA. Zauważmy, że wobec twierdzenia Abela o elemencie pierwotnym, każde skończenie generowane rozszerzenie algebraiczne ciała $\mathbb{Q}$ jest postaci $K = \mathbb{Q}(\theta)$.

Na potrzeby dalszych rozważań ustalmy totalnie rzeczywiste rozszerzenie Galois $K = \mathbb{Q}(\theta) \supset \mathbb{Q}$. Niech $r = [K : \mathbb{Q}]$; wówczas oczywiście

$$K = \{a_0 + a_1\theta + \dots + a_{r-1}\theta^{r-1} : a_i \in \mathbb{Q}, i \in \{0, \dots, r-1\}\}.$$

4.3. UWAGA. Zauważmy, iż dowolny automorfizm $\sigma \in G(K/\mathbb{Q})$ przedłuża się jednoznacznie do automorfizmu pierścienia $K[x_1, \dots, x_n]$ danego wzorem

$$\zeta \left(\sum_{(\epsilon_1, \dots, \epsilon_n) \in \mathbb{N}^n} a_{(\epsilon_1, \dots, \epsilon_n)} x_1^{\epsilon_1} \dots x_n^{\epsilon_n} \right) = \sum_{(\epsilon_1, \dots, \epsilon_n) \in \mathbb{N}^n} \sigma(a_{(\epsilon_1, \dots, \epsilon_n)}) x_1^{\epsilon_1} \dots x_n^{\epsilon_n}$$

dla którego, dla uproszczenia notacji, również będziemy używać symbolu σ .

Ponadto dla $\epsilon = (\epsilon_1, \dots, \epsilon_n) \in \mathbb{N}^n$ będziemy pisali x^ϵ dla oznaczenia jednomianu $x_1^{\epsilon_1} \cdot x_2^{\epsilon_2} \cdot \dots \cdot x_n^{\epsilon_n}$. Tym samym

$$\sigma \left(\sum_{\epsilon \in \mathbb{N}^n} a_\epsilon x^\epsilon \right) = \sum_{\epsilon \in \mathbb{N}^n} \sigma(a_\epsilon) x^\epsilon.$$

Przy przyjętych oznaczeniach zaczniemy od udowodnienia następującej prostej obserwacji.

4.4. STWIERDZENIE ([10, 2 Hillar's method]). *Niech $f \in \mathbb{Q}[x_1, \dots, x_n]$ będzie sumą m kwadratów:*

$$f = f_1^2 + f_2^2 + \dots + f_m^2,$$

gdzie $f_i \in K[x_1, \dots, x_n]$, $i \in \{1, \dots, m\}$. Wówczas:

$$f = \frac{1}{|G(K/\mathbb{Q})|} \sum_{k=1}^m \sum_{\sigma \in G(K/\mathbb{Q})} [\sigma(f_k)]^2$$

Przed podaniem dowodu zilustrujmy powyższe stwierdzenie prostym przykładem:

4.5. PRZYKŁAD. Niech $K = \mathbb{Q}(\sqrt{2})$. Jak już sygnalizowaliśmy we wstępie, jest to prosty przykład totalnie rzeczywistego rozszerzenia Galois, którego grupa Galois składa się z dwóch automorfizmów, $id_{\mathbb{Q}(\sqrt{2})}$ i $\sigma : \mathbb{Q}(\sqrt{2}) \rightarrow \mathbb{Q}(\sqrt{2})$ danego wzorem

$$\sigma(a + b\sqrt{2}) = a - b\sqrt{2}.$$

Wówczas wielomian

$$f(x, y) = 2x^2 + 2y^2 = (\sqrt{2}x)^2 + (\sqrt{2}y)^2$$

da się zapisać w postaci

$$f = \frac{1}{2} [(id_{\mathbb{Q}(\sqrt{2})}(\sqrt{2})x)^2 + (\sigma(\sqrt{2})x)^2 + (id_{\mathbb{Q}(\sqrt{2})}(\sqrt{2})y)^2 + \sigma(\sqrt{2})y)^2].$$

Przechodzimy obecnie do dowodu Stwierdzenia 4.4.

DOWÓD. Zauważmy, że dla dowolnego wielomianu o współczynnikach wymiernych $g \in \mathbb{Q}[x_1, \dots, x_n]$, zachodzi:

$$(4.10) \quad \sum_{\sigma \in G(K/\mathbb{Q})} \sigma(g) = \sum_{\sigma \in G(K/\mathbb{Q})} g = |G(K/\mathbb{Q})| \cdot g.$$

Z drugiej strony:

$$\begin{aligned} \sum_{\sigma \in G(K/\mathbb{Q})} \sigma(f) &= \sum_{\sigma \in G(K/\mathbb{Q})} \sigma \left(\sum_{k=1}^m f_k^2 \right) \\ &= \sum_{\sigma \in G(K/\mathbb{Q})} \sum_{k=1}^m \sigma(f_k^2) \\ &= \sum_{k=1}^m \sum_{\sigma \in G(K/\mathbb{Q})} [\sigma(f_k)]^2. \end{aligned}$$

W szczególności podstawiając wielomian f we wzorze (4.10) i wstawiając rezultat z lewej strony powyższego równania, a następnie dzieląc obustronnie przez $|G(K/\mathbb{Q})|$, otrzymujemy tezę. $\square$

Głównym twierdzeniem, które będziemy chcieli udowodnić, jest następujący rezultat:

4.6. TWIERDZENIE ([10, Theorem 2]). *Niech $f \in \mathbb{Q}[x_1, \dots, x_n]$ będzie sumą m kwadratów:*

$$f = f_1^2 + f_2^2 + \dots + f_m^2,$$

gdzie $f_i \in K[x_1, \dots, x_n]$, $i \in 1, \dots, m$. Wówczas f jest sumą

$$m + 4\left\lceil \frac{m}{4} \right\rceil ([K : \mathbb{Q}] - 1)$$

kwadratów w $\mathbb{Q}[x_1, \dots, x_n]$.¹

DOWÓD. Wobec Stwierdzenia 4.4 wielomian f możemy zapisać jako:

$$(4.11) \quad f = \frac{1}{|G(K/\mathbb{Q})|} \sum_{k=1}^m \sum_{\sigma \in G(K/\mathbb{Q})} [\sigma(f_k)]^2.$$

Ponieważ

$$K = \{a_0 + a_1\theta + a_2\theta^2 + \dots + a_{r-1}\theta^{r-1} : a_0, \dots, a_{r-1} \in \mathbb{Q}\},$$

więc wielomiany $f_k \in K[x_1, \dots, x_n]$, $k \in \{1, \dots, m\}$, można zapisać jako:

$$\begin{aligned} f_k &= \sum_{\epsilon \in \mathbb{N}^n} a_{\epsilon}^{(k)} x^{\epsilon} = \sum_{\epsilon \in \mathbb{N}^n} (a_{\epsilon,0}^{(k)} + a_{\epsilon,1}^{(k)}\theta + \dots + a_{\epsilon,r-1}^{(k)}\theta^{r-1}) x^{\epsilon} \\ &= \left(\sum_{\epsilon \in \mathbb{N}^n} a_{\epsilon,0}^{(k)} x^{\epsilon} \right) + \left(\sum_{\epsilon \in \mathbb{N}^n} a_{\epsilon,1}^{(k)} x^{\epsilon} \theta \right) + \dots + \left(\sum_{\epsilon \in \mathbb{N}^n} a_{\epsilon,r-1}^{(k)} x^{\epsilon} \theta^{r-1} \right) \\ &= \sum_{i=0}^{r-1} q_{ik} \theta^i, \end{aligned}$$

gdzie $q_{ik} = \sum_{\epsilon \in \mathbb{N}^n} a_{\epsilon,i}^{(k)} x^{\epsilon} \in \mathbb{Q}[x_1, \dots, x_n]$, $i \in \{0, \dots, r-1\}$, $k \in \{1, \dots, m\}$.

¹Symbolom $\lceil \cdot \rceil$ oznaczamy funkcję "sufit" $\lceil \cdot \rceil : \mathbb{R} \rightarrow \mathbb{Z}$ taką, że $\lceil x \rceil = \min\{k \in \mathbb{Z} : k \geq x\}$

Dla uproszczenia notacji powiedzmy, że $G(K/\mathbb{Q}) = \{\sigma_1, \dots, \sigma_r\}$. Wówczas:

$$\begin{aligned}
 (4.12) \quad \sum_{\sigma \in G(K/\mathbb{Q})} [\sigma(f_k)]^2 &= \sum_{j=1}^r [\sigma_j(f_k)]^2 = \sum_{j=1}^r \left[\sigma_j \left(\sum_{i=0}^{r-1} q_{ik} \theta^i \right) \right]^2 \\
 &= \sum_{j=1}^r \left[\sum_{i=0}^{r-1} \sigma_j(q_{ik} \theta^i) \right]^2 \\
 &= \sum_{j=1}^r \left[\sum_{i=0}^{r-1} \sigma_j(q_{ik}) \sigma_j(\theta^i) \right]^2 \\
 &= \sum_{j=1}^r \left[\sum_{i=0}^{r-1} q_{ik} (\sigma_j(\theta))^i \right]^2.
 \end{aligned}$$

Zauważmy, że prawą stronę równości (4.12) możemy przedstawić jako iloczyn skalarny dwóch wektorów, do których w dalszej kolejności zastosujemy proste operacje macierzowe:

$$\begin{aligned}
 (4.13) \quad &\sum_{j=1}^r \left[\sum_{i=0}^{r-1} q_{ik} (\sigma_j(\theta))^i \right]^2 \\
 &= \begin{bmatrix} \sum_{i=0}^{r-1} q_{ik} (\sigma_1(\theta))^i & \cdots & \sum_{i=0}^{r-1} q_{ik} (\sigma_r(\theta))^i \end{bmatrix} \\
 &\quad \cdot \begin{bmatrix} \sum_{i=0}^{r-1} q_{ik} (\sigma_1(\theta))^i \\ \vdots \\ \sum_{i=0}^{r-1} q_{ik} (\sigma_r(\theta))^i \end{bmatrix} \\
 &= \begin{bmatrix} \sum_{i=0}^{r-1} q_{ik} (\sigma_1(\theta))^i \\ \vdots \\ \sum_{i=0}^{r-1} q_{ik} (\sigma_r(\theta))^i \end{bmatrix}^T \cdot \begin{bmatrix} \sum_{i=0}^{r-1} q_{ik} (\sigma_1(\theta))^i \\ \vdots \\ \sum_{i=0}^{r-1} q_{ik} (\sigma_r(\theta))^i \end{bmatrix} \\
 &= \left(\begin{bmatrix} 1 & \sigma_1(\theta) & \cdots & \sigma_1(\theta)^{r-1} \\ 1 & \sigma_2(\theta) & \cdots & \sigma_2(\theta)^{r-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \sigma_r(\theta) & \cdots & \sigma_r(\theta)^{r-1} \end{bmatrix} \cdot \begin{bmatrix} q_{0k} \\ q_{1k} \\ \vdots \\ q_{r-1,k} \end{bmatrix} \right)^T \\
 &\quad \cdot \left(\begin{bmatrix} 1 & \sigma_1(\theta) & \cdots & \sigma_1(\theta)^{r-1} \\ 1 & \sigma_2(\theta) & \cdots & \sigma_2(\theta)^{r-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \sigma_r(\theta) & \cdots & \sigma_r(\theta)^{r-1} \end{bmatrix} \cdot \begin{bmatrix} q_{0k} \\ q_{1k} \\ \vdots \\ q_{r-1,k} \end{bmatrix} \right)
 \end{aligned}$$

$$\begin{aligned}
&= \begin{bmatrix} q_{0k} \\ q_{1k} \\ \vdots \\ q_{r-1,k} \end{bmatrix}^T \begin{bmatrix} 1 & \sigma_1(\theta) & \cdots & (\sigma_1(\theta))^{r-1} \\ 1 & \sigma_2(\theta) & \cdots & (\sigma_2(\theta))^{r-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \sigma_r(\theta) & \cdots & (\sigma_r(\theta))^{r-1} \end{bmatrix}^T \\
&\cdot \begin{bmatrix} 1 & \sigma_1(\theta) & \cdots & (\sigma_1(\theta))^{r-1} \\ 1 & \sigma_2(\theta) & \cdots & (\sigma_2(\theta))^{r-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \sigma_r(\theta) & \cdots & (\sigma_r(\theta))^{r-1} \end{bmatrix} \begin{bmatrix} q_{0k} \\ q_{1k} \\ \vdots \\ q_{r-1,k} \end{bmatrix}
\end{aligned}$$

Oznaczmy

$$G_1 = \begin{bmatrix} 1 & \sigma_1(\theta) & \cdots & (\sigma_1(\theta))^{r-1} \\ 1 & \sigma_2(\theta) & \cdots & (\sigma_2(\theta))^{r-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \sigma_r(\theta) & \cdots & (\sigma_r(\theta))^{r-1} \end{bmatrix}.$$

Macierz $G = G_1^T G_1$ występującą powyżej nazywać będziemy **macierzą Grama sumy kwadratów** (4.12). Niech $u(x)$ będzie wielomianem minimalnym elementu $\theta \in K = \mathbb{Q}(\theta)$ i niech α_{st} oznacza współrzędną s -tego wiersza i t -tej kolumny macierzy G . Zauważmy, że:

$$(4.14) \quad \alpha_{st} = \sum_{j=1}^r (\sigma_j(\theta))^{s+t-2},$$

a zatem współrzędna s -tego wiersza i t -tej kolumny macierzy G jest sumą $s + t - 2$ -gich potęg pierwiastków wielomianu u . Wobec Twierdzenia 3.2, rząd macierzy G jest równy liczbie różnych pierwiastków wielomianu u , a więc r .

Ustalmy $\sigma \in \{\sigma_1, \dots, \sigma_r\}$. Ponieważ $\{\sigma_1, \dots, \sigma_r\} = \{\sigma \circ \sigma_1, \dots, \sigma \circ \sigma_r\}$, więc:

$$\begin{aligned}
\sigma(\alpha_{st}) &= \sigma \left(\sum_{j=1}^r (\sigma_j(\theta))^{s+t-2} \right) \\
&= \sum_{j=1}^r (\sigma \circ \sigma_j(\theta))^{s+t-2} \\
&= \sum_{j=1}^r (\sigma_j(\theta))^{s+t-2} \\
&= \alpha_{st}
\end{aligned}$$

i tym samym α_{st} należy do ciała elementów stałych rozszerzenia $K \supset \mathbb{Q}$, czyli do $\mathbb{Q}$. Innymi słowy, $G \in \mathbb{Q}_r^r$

Zauważmy, że macierz G jest symetryczna. Istotnie, wobec równości (4.14) otrzymujemy:

$$\alpha_{st} = \sum_{j=1}^r (\sigma_j(\theta))^{s+t-2} = \sum_{j=1}^r (\sigma_j(\theta))^{t+s-2} = \alpha_{ts}.$$

Zauważmy też, że wszystkie wartości własne macierzy G są dodatnie. Faktycznie, niech $v = (v_1, \dots, v_r), v \neq 0$ będzie wektorem własnym macierzy G , zaś λ odpowiadającą mu wartością własną. Wówczas:

$$\lambda v^T = G v^T,$$

a zatem

$$\lambda v^T = G_1^T G_1 v^T.$$

Mnożąc powyższą równość obustronnie z lewej strony przez wektor v otrzymujemy

$$v \lambda v^T = v G_1^T G_1 v^T$$

czyli

$$(4.15) \quad \lambda |v|^2 = |G_1 v^T|^2.$$

Dalej, podobnie jak w dowodzie Twierdzenia 3.2 sprawdzamy, że wyznacznik macierzy G_1 jest równy $\prod_{1 \leq j_1 < j_2 \leq r} (\sigma_{j_2}(\theta) - \sigma_{j_1}(\theta))$, a więc $\neq 0$, gdyż zgodnie z założeniami wielomian u jest rozdzielnym i wszystkie jego pierwiastki są parami różne. Tym samym równość (4.15) oznacza, że $\lambda > 0$.

Wobec Twierdzenia 2.11 macierz G jest więc dodatnio określona. Dalej, wobec Twierdzenia 2.9 istnieją jednoznacznie wyznaczone macierze dolnotrójkątna $L \in \mathbb{Q}_r^r$ o samych jedynkach na głównej przekątnej i diagonalna $D \in \mathbb{Q}_r^r$ o dodatnich współczynnikach na głównej przekątnej takie, że $G = LDL^T$. Wobec równości (4.12) oraz (4.13):

$$(4.16) \quad \sum_{\sigma \in G(K/\mathbb{Q})} [\sigma(f_k)]^2 = \begin{bmatrix} q_{0k} & q_{1k} & \cdots & q_{r-1,k} \end{bmatrix} LDL^T \begin{bmatrix} q_{0k} \\ q_{1k} \\ \vdots \\ q_{r-1,k} \end{bmatrix}$$

Niech $g_{0k}, g_{1k}, \dots, g_{r-1,k} \in \mathbb{Q}[x_1, \dots, x_n]$ będą takimi wielomianami, że:

$$L^T \begin{bmatrix} q_{0k} \\ q_{1k} \\ \vdots \\ q_{r-1,k} \end{bmatrix} = \begin{bmatrix} g_{0k} \\ g_{1k} \\ \vdots \\ g_{r-1,k} \end{bmatrix}.$$

Oznaczmy ponadto przez $d_1, \dots, d_r \in \mathbb{Q}$ współrzędne na głównej przekątnej macierzy D . Równanie (4.16) przepisze się wówczas jako:

$$(4.17) \quad \sum_{\sigma \in G(K/\mathbb{Q})} [\sigma(f_k)]^2 = \begin{bmatrix} g_{0k} \\ g_{1k} \\ \vdots \\ g_{r-1,k} \end{bmatrix}^T \begin{bmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_r \end{bmatrix} \begin{bmatrix} g_{0k} \\ g_{1k} \\ \vdots \\ g_{r-1,k} \end{bmatrix} \\ = g_{0k}^2 d_1 + g_{1k}^2 d_2 + \dots + g_{r-1,k}^2 d_r$$

Podstawmy teraz rezultat równania (4.17) do wyjściowego oszacowania (4.11). Otrzymujemy:

$$(4.18) \quad f = \frac{1}{|G(K/\mathbb{Q})|} \sum_{k=1}^m \sum_{\sigma \in G(K/\mathbb{Q})} [\sigma(f_k)]^2 \\ = \frac{1}{r} \sum_{k=1}^m (g_{0k}^2 d_1 + g_{1k}^2 d_2 + \dots + g_{r-1,k}^2 d_r) \\ = \sum_{k=1}^m \left(g_{0k}^2 \frac{d_1}{r} + g_{1k}^2 \frac{d_2}{r} + \dots + g_{r-1,k}^2 \frac{d_r}{r} \right)$$

Ponieważ $L \in \mathbb{Q}_r^r$ jest macierzą dolnotrójkątną o samych jedynkach na głównej przekątnej, wobec równości $G = LDL^T$ otrzymujemy, że współrzędna α_{11} z pierwszego wiersza i pierwszej kolumny macierzy G spełnia równość $\alpha_{11} = d_1$. Z drugiej strony, wobec równania (4.14):

$$\alpha_{11} = \sum_{j=1}^r (\sigma_j(\theta))^{1+1-2} = \sum_{j=1}^r (\sigma_j(\theta))^0 = \sum_{j=1}^r 1 = r,$$

a więc $d_1 = r$ i równanie (4.18) możemy uprościć, jak następuje:

$$(4.19) \quad f = \sum_{k=1}^m \left(g_{0k}^2 \frac{d_1}{r} + g_{1k}^2 \frac{d_2}{r} + \dots + g_{r-1,k}^2 \frac{d_r}{r} \right) \\ = \sum_{k=1}^m \left(g_{0k}^2 + g_{1k}^2 \frac{d_2}{r} + \dots + g_{r-1,k}^2 \frac{d_r}{r} \right) \\ = \sum_{k=1}^m g_{0k}^2 + \frac{d_2}{r} \sum_{k=1}^m g_{1k}^2 + \dots + \frac{d_r}{r} \sum_{k=1}^m g_{r-1,k}^2.$$

Pierwszy składnik końcowej sumy we wzorze (4.19) jest oczywiście sumą m kwadratów wielomianów o współczynnikach wymiernych. Jeżeli chodzi o pozostałe $r - 1$ składników, to każdy z występujących w nich współczynników $\frac{d_j}{r}$, $j \in \{2, \dots, r\}$, jest dodatnią liczbą wymierną, a więc wobec twierdzenia Lagrange'a o czterech kwadratach [13]², daje

²Każda nieujemna liczba całkowita jest sumą kwadratów czterech liczb całkowitych

się zapisać jako suma co najwyżej czterech kwadratów liczb wymiernych. Z kolei $\sum_{k=1}^m g_{jk}^2$, $j \in \{2, \dots, r\}$, jest sumą m kwadratów wielomianów o współczynnikach wymiernych. Iloczyn sumy 4 kwadratów liczb wymiernych i sumy 4 kwadratów wielomianów wymiernych jest sumą 4 kwadratów wielomianów wymiernych ponieważ mamy równość:

$$\begin{aligned}
 (4.20) \quad & (a_1^2 + a_2^2 + a_3^2 + a_4^2)(b_1^2 + b_2^2 + b_3^2 + b_4^2) \\
 &= (a_1b_1 - a_2b_2 - a_3b_3 - a_4b_4)^2 \\
 &+ (a_1b_2 + a_2b_1 + a_3b_4 - a_4b_3)^2 \\
 &+ (a_1b_3 - a_2b_4 + a_3b_1 + a_4b_2)^2 \\
 &+ (a_1b_4 + a_2b_3 - a_3b_2 + a_4b_1)^2.
 \end{aligned}$$

Wobec tego $\frac{d_j}{r} \sum_{k=1}^m g_{jk}^2$, $j \in \{2, \dots, r\}$, jest sumą $4\lceil \frac{m}{4} \rceil$ kwadratów wielomianów o współczynnikach wymiernych.

W konsekwencji, pamiętając o tym, że $r = [K : \mathbb{Q}]$, równanie (4.19) dowodzi reprezentacji wielomianu f jako sumy $m + 4\lceil \frac{m}{4} \rceil([K : \mathbb{Q}] - 1)$ kwadratów wielomianów wymiernych. $\square$

ROZDZIAŁ 5

Przykłady oszacowań

W rozdziale tym przedstawimy dwa przykłady ilustrujące najważniejsze twierdzenia z całej pracy. Pierwszy przykład przedstawia wielomian i totalnie rozszerzenie Galois, który jest sumą $m + 4\lceil \frac{m}{4} \rceil$ wielomianów wymiernych. Drugi przykład przedstawia zapis wielomianu o współczynnikach wymiernych w postaci sumy kwadratów wielomianów wymiernych wraz z dekompozycją LU macierzy Grama.

5.1. PRZYKŁAD. *Niech $f = (X - \theta)^2 + (X + \theta)^2 + (2 - \sqrt{3})^2 = 2X^2 + 15$, gdzie $\theta = 1 + \sqrt{3}$. Weźmy za totalnie rzeczywiste rozszerzenie Galois $k = \mathbb{Q}(\sqrt{3})$. Wówczas $m = 3$, $2X^2 + 15$ jest sumą $m + 4\lceil \frac{m}{4} \rceil \in \mathbb{Q}[X]$ i $2X^2 + 15 = X^2 + X^2 + 3^2 + 2^2 + 1 + 1 + 0$.*

5.2. PRZYKŁAD ([10, Example 1]). *Rozważmy następujący wielomian:*

$$f = 3 - 12y - 6x^3 + 18y^2 + ex^6 + 12x^3y - 6xy^3 + 6x^2y^4,$$

który jest sumą kwadratów postaci:

$$\begin{aligned} f = & (x^3 + \alpha^2y + \beta xy^2 - 1)^2 + \\ & (x^3 + \beta^2y + \gamma xy^2 - 1)^2 + \\ & (x^3 + \gamma^2x + \alpha xy^2 - 1)^2, \end{aligned}$$

w

$$\mathbb{Q}(\alpha)[x, y]$$

gdzie α, β, γ są rzeczywistymi pierwiastkami wielomianu $u(x) = x^3 - 3x + 1$. Macierz G przyjmuje postać:

$$\begin{bmatrix} 3 & 0 & 6 \\ 0 & 6 & -3 \\ 6 & -3 & 18 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & \frac{9}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}.$$

Współczynniki tej macierzy są obliczane przy wykorzystaniu wzorów Viete'a.

$\beta = 2 - \alpha - \alpha^2$, $\gamma = 2\alpha^2 - 2$, a wektor $q = (q_0, q_1, q_2)^T$ wynosi $q = (x^3 + 2xy^2 - 1, -xy^2, y - xy^2)^T$.

Stąd otrzymujemy, że

$$f = ((x^3 + 2xy^2 - 1) + 2(y - xy^2))^2 + 6(-xy^2 - \frac{1}{2}(y - xy^2))^2 + \frac{9}{2}(y - xy^2)^2,$$

które jest sumą 3 kwadratów.

Początkowo przykład ten został przedstawiony przez Hillara, jednak wynik, który otrzymano to 6 kwadratów. W miejscu $m + 4\lceil \frac{m}{4} \rceil ([K : \mathbb{Q}] - 1)$ pojawiała się wtedy suma postaci $: 4m \cdot 2^{[L:\mathbb{Q}]+1} \binom{[L:\mathbb{Q}]+1}{2}$, gdzie L to domknięcie totalnie rzeczywistego rozszerzenia Galois.

Bibliografia

- [1] Basu S., Pollack R., Roy M.F., *Algorithms in Real Algebraic Geometry*, Heidelberg: Springer (2006)
- [2] Coste M., *An introduction to semialgebraic geometry*, Rennes (2002)
- [3] Hillar C.J., *Sums of squares over totally real fields are rational sums of squares*, Proc. Amer. Math. Soc., 137 (2009), 921-930
- [4] Jefimow N.W., Rozendorn E.R., *Algebra liniowa wraz z geometrią wielowymiarową*, Warszawa: PWN (1974)
- [5] Kostrikin A.I., *Wstęp do algebry*, Warszawa: PWN (2005)
- [6] Landau E., *Elementary Number Theory*, New York: Chelsea Publishing Company (1958)
- [7] Mirsky L., *An introduction to linear algebra*, Oxford: Oxford University Press (1955)
- [8] Parrilo P., Sturmfels B., *Minimizing polynomial functions, Algorithmic and quantitative real algebraic geometry*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, Vol. 60, pp. 83-99, AMS
- [9] Pourchet Y., *Sur la representation en somme de carres des polynomes a une indeterminee sur un corps de nombres algebriques*, Acta Arith. 19 (1971), 89-104
- [10] Quarez R., *Tight bounds for rational sums of squares over totally real fields*, Rendiconti del Circolo Matematico di Palermo, 59 (2010), 377-388
- [11] Watkins D. S., *Fundamentals of Matrix Computations*, New York: Wiley Interscience (2002)
- [12] Sturmfels B., *Solving Systems of Polynomial Equations*, Amer.Math.Soc., CBMS Regional Conferences Series, No 97, Providence, Rhode Island, (2002)
- [13] Sierpiński W., *Arytmetyka teoretyczna*, Warszawa: PWN (1959)