ACTA ARITHMETICA
XXIV (1973)

The factorization of an integral matrix into a product of two integral symmetric matrices I*

by

Olga Taussky (Pasadena, Cal.)

Dedicated to C. L. Siegel

The problem of factorizing a matrix whose elements are rational integers into two integral symmetric matrices is studied here in a special case only. It is assumed that the matrix A is a 2×2 matrix and that its characteristic polynomial is x^2-m, where m is a square free integer which is congruent to 2 or $3 \bmod 4$. Hence $\pm\sqrt{m}$, which are the characteristic roots of A, together with the integer 1, form a Z-basis for the maximal order of $Q(\sqrt{m})$. This allows the special case to be discussed particularly easily.

It is known that for any integral A the equation

$$A = S_1 S_2 \tag{1}$$

with $S_i = S_i'$ (the transpose) can be solved over the rationals (for references see e.g. Taussky [9]).

However (1) cannot always be solved over the integers. The study of this problem will be linked to two facts, the first of which is known (see e.g. Taussky [5]):

I. Let A be any integral matrix with irreducible characteristic polynomial $f(x)$ and let a be a zero of $f(x)$. Then there is a 1-1 correspondence between the ideal classes in $Z[a]$ and the matrix classes $\{S^{-1}XS\}$. Here X is a matrix root of $f(x) = 0$ and S runs through all unimodular matrices.

It follows that there is a unique ideal class $\mathfrak{C}_A$ in $Q(\sqrt{m})$ associated with our given matrix A.

II. There is a binary quadratic form $a(\lambda, \mu)$ associated with (1) in the 2×2 case. This form turns up for matrices over any field F and has certain invariant features. In the number theoretic case its significance

* This work was carried out (in part) under an NSF grant.

becomes more intricate. This form can be derived from (1) for

$$AS_1 = S_1 A'. \tag{2}$$

Putting then $S_1 = \begin{pmatrix} x & y \\ y & z \end{pmatrix}$ one obtains a linear relation between x, y, z:

$$(a_{22} - a_{11})y + a_{21}x - a_{12}z = 0. \tag{3}$$

This equation has (apart from trivial cases) a 2-parameter solution. Hence every element in S_1 is a linear form in two parameters λ, μ. Hence $\det S_1$ is a quadratic form in λ, μ. This is our $a(\lambda, \mu)$. The discriminant of this form, after a normalization, is the discriminant of the characteristic polynomial of A. We can now formulate the main result.

THEOREM. *Let A be an integral 2×2 matrix with characteristic polynomial $x^2 - m$, where m is a square free integer congruent to* 2 *or* 3 mod 4. *Then* (1) *can be solved in integers iff the associated quadratic form represents integrally a factor of m. Excepting the case where $a(\lambda, \mu)$ represents factors of $2m$, but not of m, it follows that* (1) *can be solved integrally iff the ideal class $\mathfrak{C}_A$ corresponding to A is of order* 1, 2, 4.

Proof. In the number theoretic case under consideration the two independent solutions of (3) have to be chosen with more care. Denote by $\delta = \gcd[(a_{22} - a_{11}), a_{12}]$, and let y_1 and z_1 satisfy [1]

$$(a_{22} - a_{11})y_1 + a_{21}\delta - a_{12}z_1 = 0. \tag{4}$$

Then

$$x_0 = 0, \quad y_0 = a_{12}/\delta, \quad z_0 = (a_{22} - a_{11})/\delta,$$
$$x_1 = \delta, \quad y_1, \quad z_1$$

are two independent solutions of (3), since the 2×2 minors of their matrix has gcd equal to 1. Hence every integral solution of (3) is an integral linear combination of these and the most general integral S satisfying (2) is of the form

$$S_1(\lambda, \mu) = \begin{pmatrix} \mu\delta & \lambda[a_{12}/\delta] + \mu y_1 \\ \lambda[a_{12}/\delta] + \mu y_1 & \lambda[(a_{22} - a_{11})/\delta] + \mu z_1 \end{pmatrix}. \tag{5}$$

The determinant of $S_1(\lambda, \mu)$ is

$$a(\lambda, \mu) = \lambda^2[-a_{12}^2/\delta^2] + \lambda\mu[(a_{22} - a_{11}) - 2a_{12}y_1/\delta] + \mu^2(\delta z_1 - y_1^2). \tag{6}$$

Using the assumptions that trace $A = 0$ and $\det A = -m$ it can be checked by direct computation that $a(\lambda, \mu)$ has discriminant $4m$. Further $a(\lambda, \mu)$ is primitive. Since

$$\det A = -m = \det S_1 \det S_2 \tag{7}$$

both determinants are divisors of m and have to be represented by $a(\lambda, \mu)$. This immediately restricts the nature of $a(\lambda, \mu)$ in the case of an integral solution of (1). For not every binary form represents divisors of its discriminant.

The remainder of the proof will now be split into several sections.

A. If $a(\lambda, \mu)$ represents a factor d_1 of m then (1) can be satisfied integrally with $\det S_1 = d_1$.

B. The ideal class in $Z[\alpha]$ which corresponds to the class of $a(\lambda, \mu)$ is $\mathfrak{C}_A^{-2}$.

C. $a(\lambda, \mu)$ lies in a form class of order 1 or 2 if (1) holds.

D. The exceptional forms.

Proof of A. Let $\lambda = \lambda_1$, $\mu = \mu_1$ be the integers for which $a(\lambda, \mu) = d_1$. Then there exist integers λ_2, μ_2 such that $a(\lambda_2, \mu_2) = -m/d_1 = d_2$. For, by a unimodular congruence transformation $a(\lambda, \mu)$ can be transformed into a form $a_1(\lambda, \mu)$ in which the coefficient of λ^2 is d_1:

$$a_1(\lambda, \mu) = d_1\lambda^2 + b\lambda\mu + c\mu^2.$$

Next apply the unimodular congruence transformation with matrix $\begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix}$ where r is so chosen that the middle coefficient in the transformed form $a_2(\lambda, \mu)$ is zero, i.e.

$$b + 2d_1 r = 0.$$

This can be achieved since $b^2 - 4cd_1 = 4m$ implies $2d_1 | b$ in both cases $m \equiv 2$ or $m \equiv 3(4)$. The form $a_2(\lambda, \mu)$ then has d_2 as coefficient of μ^2.

We then apply the transformation with matrix $\begin{pmatrix} \lambda_1 & \lambda_2 \\ \mu_1 & \mu_2 \end{pmatrix}$ also to the elements of $S_1(\lambda, \mu)$ obtaining

$$S_1^{(1)}(\lambda, \mu) = \tag{8}$$
$$\begin{pmatrix} (\mu_1\lambda + \mu_2\mu)\delta & (\lambda_1\lambda + \lambda_2\mu)[a_{12}/\delta] + (\mu_1\lambda + \mu_2\mu)y_1 \\ (\lambda_1\lambda + \lambda_2\mu)[a_{12}/\delta] + (\mu_1\lambda + \mu_2\mu)y_1 & (\lambda_1\lambda + \lambda_2\mu)[(a_{22} - a_{11})/\delta] + (\mu_1\lambda + \mu_2\mu)z_1 \end{pmatrix}.$$

Since $\det S_1^{(1)}$ is a quadratic form with no middle term the following relation holds:

$$\lambda_1\lambda_2[-2(a_{12}/\delta)^2] + (\mu_1\lambda_2 + \mu_2\lambda_1)(a_{22} - a_{11} - [2a_{12}y_1/\delta]) + \\ + \mu_1\mu_2(2z_1\delta - 2y_1^2) = 0 \tag{9}$$

[1] Since $m \equiv 2, 3(4)$ and A is of the form $\begin{pmatrix} a & b \\ \frac{m - a^2}{b} & -a \end{pmatrix}$ it can be checked that the coefficients in equation (3) have gcd = 1.

which we may replace by

$$\text{(9A)}\quad \lambda_1\lambda_2[-(a_{12}/\delta)^2]+(\mu_1\lambda_2+\mu_2\lambda_1)(-a_{11}-[a_{12}y_1/\delta])+ \\ +\mu_1\mu_2(z_1\delta-y_1^2) = 0.$$

We next put

$$S_1 = S_1^{(1)}(1,0) \quad \text{and} \quad S_2 = [S_1^{(1)}(0,1)]^{-1}\cdot\det S_1^{(1)}(0,1)$$

and form the product

$$S_1S_2 = \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix}.$$

First compute the elements s_{12} and s_{21} in this product:

$$s_{12} = a_{12}(\lambda_1\mu_2-\mu_1\lambda_2),$$
$$s_{21} = -[\{y_1(a_{22}-a_{11})/\delta\}-z_1(a_{12}/\delta)](\lambda_1\mu_2-\mu_1\lambda_2) = a_{21}(\lambda_1\mu_2-\mu_1\lambda_2)$$

in virtue of (4).

Next compute s_{11}

$$s_{11} = \lambda_1\lambda_2[-(a_{12}/\delta)^2]+\mu_1\mu_2(\delta z_1-y_1^2)+ \\ +\mu_1\lambda_2\big(-2a_{11}-y_1(a_{12}/\delta)\big)-\lambda_1\mu_2 y_1(a_{12}/\delta).$$

Using (9A) this can be rewritten as

$$\mu_1\lambda_2\big(-2a_{11}-y_1(a_{12}/\delta)+a_{11}+y_1(a_{12}/\delta)\big)-\lambda_1\mu_2\big(y_1(a_{12}/\delta)-a_{11}-y_1(a_{12}/\delta)\big) \\ = a_{11}(\lambda_1\mu_2-\lambda_2\mu_1).$$

Similarly

$$s_{22} = a_{22}(\lambda_1\mu_2-\lambda_2\mu_1).$$

Hence

$$S_1S_2 = \pm A$$

since $\lambda_1\mu_2-\lambda_2\mu_1 = \pm 1$(2).

Proof of B. In Taussky [8] all integral S with the property

$$\text{(10)}\qquad S^{-1}AS = A'$$

are characterized. Here A is an $n\times n$ integral matrix with irreducible characteristic polynomial $f(x)$, a a zero of $f(x)$. If S_0 is one of them then all others are of the form $p(A)S_0$ where $p(x)$ runs through all polynomials with rational coefficients such that $p(A)S_0$ is an integral matrix. The corresponding polynomials $p(a)$ then form a fractional ideal. In the case that $Z[a]$ is the maximal order in $Q(a)$ this ideal is in the class of $\mathfrak{C}_A^{-2}$ (see [6]) (3).

[2] In a sequel to this paper a less computational proof will be given.

[3] Actually in [8] the matrix S^{-1} is the one assumed integral, in this case the ideal class discussed above is $\mathfrak{C}_A^2$.

In Taussky [7] it was shown that for $n = 2$ a rational matrix S_0 satisfying (10) has $\det S_0 = -\mathrm{norm}(\lambda)$, $\lambda \in Q(a)$.

Hence the determinants of all the integral S's satisfying (10) are the product of $-\mathrm{norm}\,\lambda$ and the norms of the elements in the ideal discussed above, for $\det p(A) = \mathrm{norm}\,p(a)$.

On the other hand, the form $a(\lambda,\mu)$ corresponds to an ideal in $Z[a]$. The integers represented by $a(\lambda,\mu)$ are—apart from a common factor which is a norm—the norms of the elements in this ideal. But the integers represented by $a(\lambda,\mu)$ are the determinants of the matrices S_1 and the S_1 are the integral matrices which satisfy (10). These integers determine the form apart from the sign of the middle term (see e.g. H. Cohn [1]). Hence the form class of $a(\lambda,\mu)$ corresponds to the ideal class of $\mathfrak{C}_A^{-2}$ or its conjugate.

Proof of C. A quadratic form of discriminant $4m$, m a square free integer, represents divisors of its discriminant iff it is in a form class of order 1 or 2 (see Gauss [0], Scholz [4], Rédei [3]). Hence $\mathfrak{C}_A^2$ must be of order 1 or 2, hence $\mathfrak{C}_A$ is necessarily of order 1 or 2 or 4 if (1) is to hold with integral factors.

Proof of D. The condition in C is also sufficient provided the divisors of $4m$ represented by $a(\lambda,\mu)$ are divisors of m itself in virtue of (7).

It can, however, happen that none of the divisors of $4m$ represented by the form $a(\lambda,\mu)$ are divisors of m. An example of such an occurrence is

$$m = 17\cdot 67, \qquad A = \begin{pmatrix} -33 & 5 \\ 10 & 33 \end{pmatrix}$$

the corresponding form $a(\lambda,\mu)$ is $2x^2+66xy-25y^2$ representing

$$2,\ -2\cdot 17\cdot 67,\ 2\cdot 17,\ -2\cdot 67;$$

2 is represented for $x = 1$, $y = 0$, 34 is represented for $x = 1$, $y = 2$.

Examples. The following two examples illustrate some of the theory.

1. $m = 79$, $A = \begin{pmatrix} -8 & 3 \\ 5 & 8 \end{pmatrix}$, $a(\lambda,\mu) = -9\lambda^2+10\lambda\mu+6\mu^2$.

$a(\lambda,\mu)$ corresponds to an ideal class of order 3 and A cannot be factorized. However the matrix

$$A = \begin{pmatrix} -7 & 3 \\ 5 & 9 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 7 \end{pmatrix}\begin{pmatrix} -9 & 2 \\ 2 & 1 \end{pmatrix}.$$

This A has the characteristic roots $1\pm\sqrt{79}$ and differs from $\begin{pmatrix} -8 & 3 \\ 5 & 8 \end{pmatrix}$ by $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$(4).

[4] In a sequel to this paper it will be shown that to every 2×2 integral matrix there exist rational integers r such that $A+rI$ satisfies (1) for integral factors.

2. $m = 291$, $A = \begin{pmatrix} -16 & 5 \\ 7 & 16 \end{pmatrix}$, $a(\lambda, \mu) = -25\lambda^2+25\lambda\mu-6\mu^2$.

This $a(\lambda, \mu)$ represents -6 which is a divisor of $4m$, but not of m; however the form also represents 3 for $\lambda = 3$, $\mu = 19$ leading to the factorization $\begin{pmatrix} 19 & -4 \\ -4 & 1 \end{pmatrix}\begin{pmatrix} 4 & 23 \\ 23 & 108 \end{pmatrix}$. The matrix A corresponds to an ideal class of order 4.

The author has been guided by computations based on Ince's tables on ideasl in quadratic fields and by a program carried out by G. Hayward at California Institute of Technology. The example given in D was constructed by D. Estes and H. Kisilevsky with whom the author also had helpful discussions on earlier parts.

References

[0] C. F. Gauss, *Disquisitiones arithmeticae*, 1801.

[1] H. Cohn, *A Second Course in Number Theory*, New York 1962.

[2] E. L. Ince, *Cycles of Reduced Ideals in Quadratic Fields*, British Association for the Advancement of Science Tables IV, 1934.

[3] L. Rédei and H. Reichardt, *Die durch vier teilbaren Invarianten der Klassengruppe des quadratischen Zahlkörpers*, J. Reine Angew. Math. 170 (1933), pp. 69–74.

[4] A. Scholz, *Über die Lösbarkeit der Gleichung $t^2-Du^2 = -4$*, Math. Zeitschr. 39 (1934), pp. 95–111.

[5] O. Taussky, *On a theorem of Latimer and MacDuffee*, Canad. J. Math. 1 (1949), pp. 300–302.

[6] — *On matrix classes corresponding to an ideal and its inverse*, Illinois J. Math. 1 (1957), pp. 108–113.

[7] — *Ideal Matrices I*, Archiv d. Math. 13 (1962), pp. 275–282.

[8] — *On the similarity transformation between an integral matrix with irreducible characteristic polynomial and its transpose*, Math. Ann. 166 (1966), pp. 60–63.

[9] — *Symmetric matrices and their role in the study of general matrices*, Linear Algebra and Appl. 5 (1972), pp. 147–159.

[10] — and H. Zassenhaus, *On the similarity transformation between a matrix and its transpose*, Pacific J. Math. 9 (1959), pp. 893–896.

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

Received on 30. 8. 1972 (320)

ACTA ARITHMETICA
XXIV (1973)

Quadratische Formen über Zahlringen

von

Meinhard Peters (Münster)*

Carl Ludwig Siegel zum 75. Geburtstag

Jede natürliche Zahl ist bekanntlich Summe von 4 Quadraten ganzer rationaler Zahlen. In dieser Note erhalten wir Verallgemeinerungen auf algebraische Zahlkörper: z. B. ist in nicht total-reellen Zahlkörpern mit ungerader Diskriminante jede total-positive ganze Zahl Summe von 4 Quadraten ganzer Zahlen (s. Satz 1). Allgemeiner lassen sich in nicht total-reellen Zahlkörpern die durch Summen von Quadraten ganzer Zahlen darstellbaren Zahlen bereits durch eine Summe von 4 ganzen Quadraten darstellen; dieses folgt aus einem Lokal-Global-Prinzip, das der starke Approximationssatz vermittelt (s. Satz 1). In Ordnungen reell-quadratischer Zahlkörper lassen sich die durch Quadratsummen darstellbaren Elemente bereits durch eine Summe von 5 Quadraten darstellen; diese Elemente kann man charakterisieren (s. Satz 2). Es ergeben sich dabei elementare Beweise für einige bekannte Aussagen über die Anzahl der Klassen im Geschlecht der Form $x_1^2+x_2^2+x_3^2+x_4^2$ (s. Satz 3 und Bemerkung 3). Für Verallgemeinerungen von Quadratsummen auf andere quadratische Formen F im Falle der reell-quadratischen Zahlkörper bedarf es bei der hier angewandten Beweismethode der Voraussetzung, daß das Geschlecht von F über Z nur eine Klasse enthält. In einem Anhang ist eine Tabelle der primitiven definiten quadratischen Formen in Diagonalgestalt der Dimensionen $\geqslant 4$ mit einklassigem Geschlecht über Z aufgeführt; diese Tabelle wird ausgenutzt für die eben genannten Verallgemeinerungen (s. Satz 2).

Satz 1. *In einem nicht total-reellen Zahlkörper sind genau diejenigen ganzen Zahlen durch eine quadratische Form F, F regulär und mindestens 4-dimensional, darstellbar, für die das „lokal überall" der Fall ist; insbesondere sind genau diejenigen ganzen Zahlen als Summen von Quadraten ganzer Zahlen darstellbar, für die das ,,lokall überall" der Fall ist. In einer*

* Herrn M. Kneser und Herrn C. L. Siegel danke ich sehr für Hinweise und kritische Bemerkungen.