Automorphs and Generalized Automorphs of Quadratic Forms Treated as Characteristic Value Relations*†

OLGA TAUSSKY
California Institute of Technology
Pasadena, California

1. INTRODUCTION

In two recent papers by Butts and Pall [1] and Butts and Estes [2]‡ the equation $T'AT = eB$ is studied, where A, B are the matrices of integral binary quadratic forms, T an integral 2×2 matrix, and e is an integer. The special case where $A = B$ is given particular consideration. For a fixed A the set of T's and e's that can occur are of interest, as a generalization of the concept of automorphs of quadratic forms. We assume $\det T \neq 0 \det A \neq 0$. It is clear that either $e = \det T$ or $e = -\det T$. For each case the corresponding T's form a two-dimensional Z-module, Z the ring of rational integers. Here the results are given a new interpretation which sheds further light on the difference between the two cases. This is done in Section 2. In Section 3, the relation is obtained from a relation concerning elements in a quadratic field. This was already done in [1] and is reformulated here. The ideas developed in Section 2 are then translated into this point of view. In Section 4, a generalized characteristic value problem is studied which may lead to further interesting generalizations.

---

* Dedicated to Professor A. M. Ostrowski on his 75th birthday.

† This work was carried out in part under NSF grant 3909.

‡ Thanks are due to these authors for information concerning their not yet published results. Thanks are also due to E. C. Dade for assistance with parts of Section 3. Further, D. Estes contributed helpful remarks, in particular by pointing out that the ω's in (7) are not necessarily integers.

2. THE RELATION $TAT' = eA$ AS A CHARACTERISTIC VALUE PROBLEM

The equation

$$TAT' = eA \tag{1}$$

with $T = \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix}$, $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ can be interpreted as a characteristic value relation in which the matrix A is a characteristic vector of the 3×3 matrix

$$\begin{pmatrix} t_{11}^2 & 2t_{11}t_{12} & t_{12}^2 \\ t_{11}t_{21} & t_{12}t_{21} + t_{11}t_{22} & t_{12}t_{22} \\ t_{21}^2 & 2t_{21}t_{22} & t_{22}^2 \end{pmatrix},$$

which is the symmetric product $P_2(T)$ of T and its transpose T' (also called the second induced or power matrix of T; see, e.g., Wedderburn, [3, p. 76]) and e is a corresponding characteristic root. Let τ_1, τ_2 be the characteristic roots of T. Then it is known that $P_2(T)$ has the characteristic roots τ_1^2, $\tau_1\tau_2$, τ_2^2. Hence e is one of these numbers. Since det $T = \tau_1\tau_2$, we immediately see that $e = \det T$ is one possibility. In the only other case where $e = -\det T$ we then have

$$\tau_1^2 = -\tau_1\tau_2$$

or

$$\tau_2^2 = -\tau_1\tau_2.$$

Hence trace $T = \tau_1 + \tau_2 = 0$. In this case $e = \tau_1^2 = \tau_2^2$ and $P_2(T)$ has a double characteristic root and all characteristic roots are rational, while in the case that $e = \tau_1\tau_2$ there is in general, only one rational characteristic root and the two others belong to a quadratic field and are conjugate with respect to this field. The same holds about the corresponding characteristic vectors A.

The case that $\tau_1^2 = \tau_1\tau_2$ or $\tau_2^2 = \tau_1\tau_2$ implies that $\tau_1 = \tau_2$ and det T is a square. The product $P_2(T)$ has then a triple root.

The elementary divisors of $P_2(T)$ (considered as a matrix over the field Q of rational numbers) compared with those of T have not been studied much until recently. But the following result has now been communicated by M. Marcus and S. Pierce [4]: T has linear elementary divisors if and only if $P_2(T)$ has the same property.

There is no doubt that the method of this chapter can be used for n-dimensional problems.

3a. *A representation of an order in a quadratic field which leads to* $TAT' = eA$

Here we explain the connection between Eq. (1) and a representation of the elements of an order in a quadratic number field.

Let F be a quadratic number field, $\mathfrak{a}$ a Z-module with α_1, α_2 as basis, $\alpha_i \in F$. Let $x, y \in Z$. The function which maps the pair x, y into Q via

$$x, y \to \operatorname{norm}_{F/Q}(x\alpha_1 + y\alpha_2)$$

is a quadratic form in x, y. Its matrix differs from A by an integral factor.

Next we associate a matrix $T_\omega = \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix}$ with elements in Z to every $\omega \in (\mathfrak{a} : \mathfrak{a})$ by the relations

$$\omega \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}. \tag{2}$$

On one hand*

$$\operatorname{norm}(x\omega\alpha_1 + y\omega\alpha_2) = \operatorname{norm}\omega \operatorname{norm}(x\alpha_1 + y\alpha_2), \tag{3}$$

and on the other hand by (2)

$$\operatorname{norm}[(t_{11}x + t_{21}y)\alpha_1 + (t_{12}x + t_{22}y)\alpha_2] = \operatorname{norm}\omega \operatorname{norm}(x\alpha_1 + y\alpha_2). \tag{4}$$

Putting $e = \operatorname{norm}\omega$ and $T = T_\omega$, relation (4) leads to relation (1) for the matrix A which corresponds to the form $\operatorname{norm}(x\alpha_1 + y\alpha_2)$. The matrix T has as characteristic roots ω and its conjugate $\bar{\omega}$ and $e = \det T$.† The ω's form an integral order.

---

* Norm will always mean $\operatorname{norm}_{F/Q}$.

† The same process could be carried out for $\omega \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = T \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$ where β_1, β_2 form a Z-basis for another ideal $\mathfrak{b}$ with $\omega \in (\mathfrak{a} : \mathfrak{b})$. This then leads to $TAT' = eB$ where B is the matrix which corresponds to the form $\operatorname{norm}(x\beta_1 + y\beta_2)$. In this case $e^2 \det B = (\det T)^2 \det A$.

3b. *The converse of Section* 3a *and the two types of representations obtained from* $TAT' = eA$

Conversely, let A be the matrix of an integral quadratic form $f(x, y)$ which we assume to have irrational linear factors

$$(x\alpha_1 + y\alpha_2), (x\bar{\alpha}_1 + y\bar{\alpha}_2)$$

when the bar denotes conjugation in the quadratic field to which α_1, α_2 belong. Let (1) hold. This is equivalent to

$$\begin{aligned}[(t_{11}x + t_{21}y)\alpha_1 + (t_{12}x + t_{22}y)\alpha_2][(t_{11}x + t_{21}y)\bar{\alpha}_1 + (t_{12}x + t_{22}y)\bar{\alpha}_2]\\ = e(x\alpha_1 + y\alpha_2)(x\bar{\alpha}_1 + y\bar{\alpha}_2). \qquad (5)\end{aligned}$$

Use the unique factorization property in the polynomial ring $F[x, y]$ with $F = Q(\alpha_1, \alpha_2)$, Q the rationals. It follows that for some ω either

$$(t_{11}x + t_{21}y)\alpha_1 + (t_{12}x + t_{22}y)\alpha_2 = \omega(x\alpha_1 + y\alpha_2) \qquad (6)$$

or

$$(t_{11}x + t_{21}y)\alpha_1 + (t_{12}x + t_{22}y)\alpha_2 = \bar{\omega}(x\bar{\alpha}_1 + y\bar{\alpha}_2). \qquad (7)$$

In case (6) ω is an integer in F and even more, ω is an element of the integral order $(\mathfrak{a} : \mathfrak{a})$. This follows because (6) is equivalent to (2).* In case (7) we obtain a new representation (see also [5]). Further (7) implies $\omega \in (\bar{\mathfrak{a}} : \mathfrak{a})$. Conversely, for every such $\bar{\omega}$ there exists a unique integral T such that (7) holds. Hence, summarizing, we obtain

THEOREM 1. *There is a* $1 - 1$ *correspondence between the integral* T*'s such that* (6) *holds and the elements of* $(\mathfrak{a} : \mathfrak{a})$ *and between the* T*'s such that* (7) *holds and the elements of* $(\bar{\mathfrak{a}} : \mathfrak{a})$.

Although the element ω in (7) lies in F, it is not necessarily an integer, e.g., take $A = \begin{pmatrix} 3 & 4 \\ 4 & 3 \end{pmatrix}$, $T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\alpha_1 = 3, \alpha_2 = 4 - \sqrt{7}, \bar{\omega} = (4 - \sqrt{7})/3$.

THEOREM 2. *While the* T*'s corresponding to* (6) *have as characteristic roots* $\omega, \bar{\omega}$, *the* T*'s corresponding to* (7) *have* trace $T = 0$, det $T = -\omega\bar{\omega}$, *characteristic roots* $\sqrt{\omega\bar{\omega}}, -\sqrt{\omega\bar{\omega}}$.

* Again this could be generalized for $TAT' = eB$.

Proof. The vectors $\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}, \begin{pmatrix}\bar{\alpha}_1\\ \bar{\alpha}_2\end{pmatrix}$ form a basis of the space V of 2×1 column vectors. Since

$$T\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix} = \bar{\omega}\begin{pmatrix}\bar{\alpha}_1\\ \bar{\alpha}_2\end{pmatrix}, \qquad T\begin{pmatrix}\bar{\alpha}_1\\ \bar{\alpha}_2\end{pmatrix} = \omega\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix},$$

the linear transformation of V defined by T has the matrix $\begin{pmatrix}0 & \omega\\ \bar{\omega} & 0\end{pmatrix}$ with respect to this basis. This implies Theorem 2. The characteristic vectors corresponding to the roots are

$$v_1 = \sqrt{\omega}\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix} + \sqrt{\bar{\omega}}\begin{pmatrix}\bar{\alpha}_1\\ \bar{\alpha}_2\end{pmatrix}, \tag{8}$$

$$v_2 = \sqrt{\omega}\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix} - \sqrt{\bar{\omega}}\begin{pmatrix}\bar{\alpha}_1\\ \bar{\alpha}_2\end{pmatrix}. \tag{9}$$

3c. *The characteristic value treatment for the representation*

Let V be a vector space with elements $v_1, v_2, \ldots$. By $V \otimes_c V$ we denote the symmetric part of $V \otimes V$. It is the set which is fixed by the automorphism which transforms $v_1 \otimes v_2$ into $v_2 \otimes v_1$.

Let V be the space of all 2×1 column vectors with entries in F. Let $v_1 = \begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}$, $v_2 = \begin{pmatrix}\bar{\alpha}_1\\ \bar{\alpha}_2\end{pmatrix}$. They form a basis for V. We then have for the 2×2 matrix T_ω studied in Sections 3a and 3b

$$\begin{aligned} T_\omega v_1 &= \omega v_1,\\ T_\omega v_2 &= \bar{\omega} v_2,\\ T_\omega \bar{v}_1 &= \bar{\omega}\bar{v}_1. \end{aligned} \tag{10}$$

If we define the matrix S by the relations

$$\begin{aligned} Sv_1 &= v_2,\\ Sv_2 &= v_1, \end{aligned} \tag{11}$$

we have

$$\begin{aligned} T_\omega S v_1 &= \bar{\omega} v_2,\\ T_\omega S v_2 &= \omega v_1. \end{aligned} \tag{12}$$

The product $\mathfrak{a} \otimes \mathfrak{a}$ has as basis

$$v_1 \otimes_c v_1, v_1 \otimes_c v_2 = v_2 \otimes_c v_1, v_2 \otimes_c v_2.$$

The effect of the operator $P_2(T_\omega)$ on these vectors in the case $|T_\omega| = e$ is as follows:

$$\begin{aligned} v_1 \otimes_c v_1 &\to \omega^2 v_1 \otimes_c v_1 \\ v_1 \otimes_c v_2 &\to \omega\bar{\omega} v_1 \otimes_c v_2 \\ v_2 \otimes_c v_2 &\to \bar{\omega}^2 v_2 \otimes_c v_2. \end{aligned} \tag{13}$$

Hence all three characteristic vectors are independent of ω and the 3×3 operators can be diagonalized simultaneously. This agrees with the theorem of Marcus and Pierce mentioned earlier.

The case $|T_\omega| = -e$ is covered by replacing T_ω by $T_\omega S$. The operator $P_2(T_\omega S)$ has the following effect:

$$\begin{aligned} v_1 \otimes_c v_1 &\to \bar{\omega}^2 v_2 \otimes_c v_2, \\ v_1 \otimes_c v_2 &\to \omega\bar{\omega} v_1 \otimes_c v_2, \\ v_2 \otimes_c v_2 &\to \omega^2 v_1 \otimes_c v_1. \end{aligned} \tag{14}$$

This shows that $v_1 \otimes_c v_2$ is a characteristic vector for $\omega\bar{\omega}$ which is independent of ω. Hence it follows that the 3×3 operators corresponding to all ω's can be transformed to block (upper) triangular form simultaneously with a 1×1 block in the lower end of the diagonal. Summarizing we have

THEOREM 3. *For* det $T = e$ *the matrices* $P_2(T)$ *can be transformed to diagonal form simultaneously; for* det $T = -e$ *the matrices* $P_2(T)$ *can be transformed to block (upper) triangular form simultaneously.*

4. THE GENERALIZED CHARACTERISTIC VALUE RELATIONS (7)

In matrix theory the problem of solving

$$(A - \lambda B)x = 0$$

is usually referred to as a generalized characteristic value problem. Relation (7) is of a similar nature. In this case the following fact holds.

THEOREM 4. *Let T be a fixed 2×2 matrix over Z. Let (α_1, α_2) and ω satisfy* (7), *i.e.*,

$$T\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix} = \bar{\omega}\begin{pmatrix}\bar{\alpha}_1\\ \bar{\alpha}_2\end{pmatrix}.$$

Then so do $(\gamma\alpha_1, \gamma\alpha_2)$ and $\omega\bar{\gamma}\gamma^{-1}$ for each $\gamma \in F$. Conversely, to every λ with norm $\lambda =$ norm ω there exists γ such that $(\gamma\alpha_1, \gamma\alpha_2)$ and λ satisfy (7).

Proof. We call $\bar{\omega}$ a semicharacteristic root of T because the mapping $\theta\colon (\beta_1, \beta_2) \to (\bar{\beta}_1, \bar{\beta}_2)$ is called a semilinear transformation of the vector space of V of 2×1 column vectors over F, associated with the automorphism $\gamma \to \bar{\gamma}$ of F. This is because $\theta(\gamma v) = \bar{\gamma}\theta(v)$ for all $\gamma \in F$, $v \in V$. Let $\gamma \in F$. Then

$$T\begin{pmatrix}\gamma\alpha_1\\ \gamma\alpha_2\end{pmatrix} = \bar{\omega}\begin{pmatrix}\gamma\bar{\alpha}_1\\ \gamma\bar{\alpha}_2\end{pmatrix} = \bar{\omega}\frac{\gamma}{\bar{\gamma}}\begin{pmatrix}\bar{\gamma}\bar{\alpha}_1\\ \bar{\gamma}\bar{\alpha}_2\end{pmatrix}.$$

Then norm $\bar{\omega}(\gamma/\bar{\gamma}) =$ norm $\bar{\omega}$. Conversely, let norm $\bar{\omega} =$ norm $\bar{\lambda}$. Then $\alpha = \bar{\omega}/\bar{\lambda}$ has norm $\alpha = 1$. By Hilbert's theorem 90 [6] we have $\alpha = \bar{\gamma}/\gamma$ where γ is another element in the same field. We then have

$$T\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix} = \bar{\lambda}\alpha\begin{pmatrix}\bar{\alpha}_1\\ \bar{\alpha}_2\end{pmatrix}$$

or

$$T\begin{pmatrix}\gamma\alpha_1\\ \gamma\alpha_2\end{pmatrix} = \bar{\lambda}\begin{pmatrix}\bar{\gamma}\bar{\alpha}_1\\ \bar{\gamma}\bar{\alpha}_2\end{pmatrix}.$$

This proves Theorem 4.

We now ask: What other semicharacteristic vectors does T have? This can partly be studied from relation (1), using the method of Section 2. For $e = -\det T$ the solutions A of (1) form a Z-module with two linearly independent ones among them, since the operator $P_2(T)$ has a double root e with two independent vectors (in virtue of the results in [4]). The two linearly independent A's can correspond to ω's belonging to different fields, e.g., for

$$T = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \qquad e = 1,$$

we have

$$A = \begin{pmatrix} 3 & 4 \\ 4 & 3 \end{pmatrix}, \qquad \alpha_1 = 3, \qquad \alpha_2 = 4 - \sqrt{7}, \qquad \bar{\omega} = \frac{4 - \sqrt{7}}{3};$$

but we also have

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \qquad \alpha_1 = 1, \qquad \alpha_2 = 2 - \sqrt{3}, \qquad \bar{\omega} = 2 - \sqrt{3}.$$

The sum of these A's is $\begin{pmatrix} 4 & 6 \\ 6 & 4 \end{pmatrix}$ corresponding to $\alpha_1 = 2$, $\alpha_2 = 3 - \sqrt{5}$, $\bar{\omega} = (3 + \sqrt{5})/2$.

The vector $\begin{pmatrix} 4 & -3 \\ -3 & 4 \end{pmatrix}$ is orthogonal to $\begin{pmatrix} 3 & 4 \\ 4 & 3 \end{pmatrix}$, i.e., the inner product is zero. Here

$$\alpha_1 = 2, \qquad \alpha_2 = \frac{-3 + \sqrt{-7}}{2}, \qquad \bar{\omega} = \frac{-3 + \sqrt{-7}}{4}.$$

REFERENCES

1 H. Butts and G. Pall, Modules and binary quadratic forms. To appear.

2 H. Butts and D. Estes, Modules and binary quadratic forms over integral domains. *Linear Algebra and Its Applications* **1**(1968), 153–180.

3 J. H. M. Wedderburn, *Lectures on Matrices*, American Math. Soc., Providence, Rhode Island, 1934.

4 M. Marcus and S. Pierce, Elementary divisors of associated transformations. To appear.

5 H. Poincaré, Sur un mode nouveau de représentation géometrique des formes quadratiques définies ou indéfinies, *Oeuvres*, Tome V, Paris, 1956, pp. 117–183.

6 D. Hilbert, *Gesammelte Abhandlungen* I, Springer, Berlin, 1932.

Received January 9, 1968