ON A NORMAL FORM OF THE ORTHOGONAL TRANSFORMATION III

Hans Zassenhaus

§4. Equivalence invariants of matrix pairs.

Under the assumptions of case b) of theorem 1 we derive from (3.32) the matrix equation

$$(P(X)\,(XDP(X))^{-1})^{\mu-1} = (\delta_{i\mu}\,\delta_{k1}\,I_n)$$

so that there corresponds the matrix B to the bilinear form

$$(4.1)\quad f_{P,\mu}^{(X,A)}(u,v) = f_{P,\mu}(u,v) = (-1)^{[\mu/2]} f(u,(P(\sigma)(\sigma DP(\sigma))^{-1})^{\mu-1}v)$$

on the linear space

$$(4.2)\quad M_{P,\mu}^{(X,A)} = M_{P,\mu} = M_{P^\mu}/(M_{P^{\mu-1}} + (P(\sigma)M \cap M_{P^\mu})$$

and $f_{P,\mu}$ is symmetric if $\varepsilon = (-1)^{\mu+1}$, anti-symmetric if $\varepsilon = (-1)^\mu$.

The last statement remains true in the case a) if P is symmetric irreducible because in that case $f_{P,\mu}$ is 0.

If (X, A) is any matrix pair with regular second constituent satisfying (3.1) and with representation space M then we decompose M into the direct sum of mutually orthogonal and orthogonally indecomposable invariant subspaces for each of which the previous observations hold. Hence also on M the bilinear form defined by (4.1) on the linear space (4.2) is symmetric if $\varepsilon = (-1)^{\mu+1}$ and anti-symmetric if $\varepsilon = (-1)^\mu$. Moreover the number $a_{P,\mu}$ of orthogonally indecomposable components of M of the type a) of theorem 1, the number $b_{P,\mu}$ of orthogonally indecomposable components of M of the type b) of theorem 1, the degree dim $M_{P,\mu}$ of $f_{P,\mu}$ and the rank $\rho(f_{P,\mu})$ of $f_{P,\mu}$ are connected by the relations

$$(4.3)\quad \rho(f_{P,\mu}) = [P]\; b_{P,\mu}$$

Can. Math. Bull., vol. 1, no. 3, Sept. 1958

https://doi.org/10.4153/CMB-1958-021-6 Published online by Cambridge University Press

(4.4) $\dim M_{P,\mu} = \rho(f_{P,\mu}) + 2 \; [P] \; a_{P,\mu}$

(4.5) $a_{P,\mu} = 0$ if $P(x) \neq x - \delta$

where δ satisfies (3.2) and $[P]$ is the degree of the symmetric irreducible separable polynomial P.

The linear transformation σ of M induces a linear transformation $\sigma_{P,\mu}$ of $M_{P,\mu}$ under which $f_{P,\mu}$ is invariant such that $P(\sigma_{P,\mu}) = 0$. Therefore $M_{P,\mu}$ is a linear space over the finite extension $E_P = F[Y_P]$ according to

(4.6) $Q(Y_P)\,(u/M_{P,\mu}) = (Q(\sigma)u)/M_{P,\mu}$

for $Q(x)$ of $F[x]$ and u of $M_{P,\mu}$.

We study the case $[P] > 1$ in which E_P has the involutoric automorphism α_P over F that maps Y_P onto Y_P^{-1}. According to (3.4) one finds

(4.7) $f_{P,\mu}(\xi u, v) = f_{P,\mu}(u, \alpha_P(\xi)v)$

for ξ of E_P, u, v of $M_{P,\mu}$. This rule suggests the interpretation of $f_{P,\mu}$ as trace of an hermitian or anti-hermitian bilinear form $h_{P,\mu}$ on the E_P -module $M_{P,\mu}$ such that

(4.8) $f_{P,\mu}(u,v) = \operatorname{tr} h_{P,\mu}(u,v)$

for any two elements u, v of $M_{P,\mu}$ where we make use of the trace function of $E_{P,\mu}$ relative to F.

In order to solve (4.8) we observe that for any matrix Y the traces of $Y^0, Y^1, Y^2, \ldots$ vanish identically if and only if the multiplicity of each separable irreducible divisor of the characteristic polynomial of Y is divisible by the characteristic of F. Hence there is a power of the matrix Y_P that has non vanishing trace and hence the linear homogeneous system of equations

(4.9) $0 = \operatorname{tr}\left(\left(\sum \xi_k Y_P^{k-1}\right) Y_P^{i-1}\right) \quad (i = 1, 2, \ldots, P)$

for the coefficients $\xi_1, \xi_2, \ldots\ldots, \xi_{[P]}$ in F has only the trivial solution. It follows that for any pair of elements u, v of $M_{P,\mu}$ the system of linear equations

$$f_{P,\mu}(\sigma^{i-1}u, v) = \operatorname{tr}\left(\left(\sum_{k=1}^{P} \eta_k Y_P^{k-1}\right) Y_P^{i-1}\right) \quad (i = 1, 2, \ldots [P])$$

https://doi.org/10.4153/CMB-1958-021-6 Published online by Cambridge University Press

has precisely one solution set of elements $\eta_1, \eta_2, .\ \eta_P$ of F. Setting

$$(4.10) \qquad h^{(X,A)}_{P,\mu}(u,v) = h_{P,\mu}(u,v) = \sum_{k=1}^{[P]} \eta_\kappa Y_P^{\,k-1}$$

it follows that (4.8) is satisfied and that each of the elements

$$h_{P,\mu}(u_1+u_2,v) - h_{P,\mu}(u_1,v) - h_{P,\mu}(u_2,v),$$

$$h_{P,\mu}(u,v) - (-1)^{\mu+1}\varepsilon\, h_{P,\mu}(v,u)$$

$$h_{P,\mu}(\xi u,v) - \xi\, h_{P,\mu}(u,v)$$

$$h_{P,\mu}(u, \xi v) - \alpha_P(\xi)\, h_{P,\mu}(u,v)$$

satisfies (4.9) if it is substituted in place of $\sum \xi_k Y_P^{\,k-1}$ so that $h_{P,\mu}$ is indeed an hermitian or anti-hermitian bilinear form on $M_{P,\mu}$.

Observing that the normal form of theorem 1 case b) is uniquely determined by μ and by B i.e. the equivalence class of $h_{P,\mu}$ we derive from theorem 1 and the previous observations the following.

THEOREM 2. Let F be a field of reference of characteristic different from 2. Let (X, A) be a matrix pair over F with regular second constituent satisfying (3.1) such that every symmetric irreducible divisor P of the characteristic polynomial of X is separable. The matrix pairs (X, A) and (Y, B) are equivalent if and only if

1) X is similar to Y,

2) B is regular satisfying the condition $B^T = \varepsilon B$,

3) the bilinear forms $f^{(X,A)}_{P,\mu}$ and $f^{(Y,B)}_{P,\mu}$ are equivalent for $P(x)=x-\delta$ subject to (3.2),

4) for each symmetric irreducible divisor P of χ_X of degree greater than 1 the hermitian or anti-hermitian bilinear forms $h^{(X,A)}_{P,\mu}$, $h^{(Y,B)}_{P,\mu}$ are equivalent for $\mu = 1, 2, \ldots, \mu_P$ where μ_P denotes the multiplicity of P in the minimal polynomial of X.

5. Application to the real field and to Galois fields.

If $\varepsilon = (-1)^{\mu}$ and if $P(x)$ is a symmetric irreducible separable polynomial of degree greater than 1, then $h_{P,\mu}$ is an anti-hermitian bilinear form on $M_{P,\mu}$ over E_P. Hence the bilinear form $(Y_P - Y_P^{-1})h_{P,\mu}$ is hermitian. The equivalence of $h_{P,\mu}^{(X,A)}$, $h_{P,\mu}^{(Y,B)}$ implies the equivalence of $(Y_P - Y_P^{-1})h_{P,\mu}^{(X,A)}$, $(Y_P - Y_P^{-1})h_{P,\mu}^{(Y,B)}$ and conversely. By forming the trace bilinear form we obtain the symmetric bilinear form

$$(5.1)\quad g_{P,\mu}(u,v) = tr((Y_P - Y_P^{-1})h_{P,\mu}(u,v)) = tr(h_{P,\mu}((\sigma - \sigma^{-1})u,v))$$

so that the equivalence of the hermitian bilinear forms $(Y_P - Y_P^{-1})h_{P,\mu}^{(X,A)}$, $(Y_P - Y_P^{-1})h_{P,\mu}^{(Y,B)}$ on $M_{P,\mu}$ over E_P implies the equivalence of $g_{P,\mu}^{(X,A)}$, $g_{P,\mu}^{(Y,B)}$ on $M_{P,\mu}$ over F

If F is the real field then $h_{P,\mu}$ is obtained by solving the equations : $f_{P,\mu}(\sigma^i u,v) = h(\sigma^i u,v) + \overline{h(\sigma^i u,v)} = Y_P^i h(u,v) + Y_P^{-i}\overline{h(u,v)}$ where $\bar{\xi}$ is the complex conjugate of ξ such that $\overline{Y_P} = Y_P^{-1}$. Hence,

$$h_{P,\mu}(u,v) = \frac{f_{P,\mu}(\sigma u,v) - Y_P^{-1} f_{P,\mu}(u,v)}{Y_P - Y_P^{-1}}$$

$$(5.2)\quad g_{P,\mu}(u,v) = 2f_{P,\mu}(\sigma u,v) - (Y_P + Y_P^{-1})\, f_{P,\mu}(u,v)$$

Note that two hermitian bilinear forms over the complex field are equivalent if and only if the corresponding symmetric trace bilinear forms are equivalent over the real field.

A symmetric bilinear form on a linear space M is called a <u>splitting</u> bilinear form if M is the direct sum of two isotropic linear subspaces.

In case a) of theorem 1 for $\varepsilon = 1$ the bilinear form f is a splitting symmetric bilinear form. Similarly in case b) for $\varepsilon = 1$ the bilinear form f is splitting if μ is even whereas f is the sum of a splitting bilinear form and another bilinear form equivalent to $f_{P,\mu}$ with disjoint variables if μ is odd.

We note that because of (3.11) the number μ must be odd if $\varepsilon = 1$ and $P(x) = x - \delta$.

We denote the degree of a quadratic matrix by d(A) and its rank by r(A). If A is symmetric over the real field then the

https://doi.org/10.4153/CMB-1958-021-6 Published online by Cambridge University Press

number $p(A)$ of positive characteristic roots and the number $q(A)$ of negative characteristic roots are equivalence invariants so that $p(A) + q(A) = r(A)$. For a bilinear form f on a d-dimensional linear space with corresponding quadratic matrix A we set $d(f) = d(A) = d$, $r(f) = r(A)$ and if the field of reference is real and if f is symmetric then we set $p(f) = p(A)$, $q(f) = q(A)$. It is well known that two symmetric bilinear forms f, g are equivalent if and only if $d(f) = d(g)$, $p(f) = p(g)$, $q(f) = q(g)$. The symmetric bilinear form f splits if and only if $p(f) = q(f)$. Two anti-symmetric bilinear forms f, g are equivalent if $d(f) = d(g)$ and $r(f) = r(g)$.

After these preparations we can answer the following questions:

A) E. Wigner's question: What are the characteristic roots of a matrix X over the real field satisfying (0.1) for a given regular symmetric matrix A over the real field?

Answer:

$$(5.3)\quad \det(I_d - xX) = (x-1)^{\nu_1} (x+1)^{\nu_2} \prod_{j=3}^{s} ((x-e^{i\phi_j})(x-e^{-i\phi_j}))^{\nu_j}$$
$$\prod_{j=s+1}^{t} ((x-r_j)(x-r_j^{-1}))^{\nu_j} \prod_{j=t+1}^{\mu} ((x+r_j^{-1}))^{\nu_j}$$
$$\prod_{j=\mu+1}^{\nu} ((x-r_j e^{i\phi_j})(x-r_j e^{-i\phi_j})(x-r_j^{-1} e^{i\phi_j})(x-r_j^{-1} e^{i\phi_j}))^{\nu_j}$$

subject to the conditions

(5.4) $\quad 2 < s \leq t \leq u \leq v$

(5.5) $\quad \nu_i \geq 0, \quad \nu_i$ integral

(5.6) $\quad 1 > r_j$ if $s < j \leq v$

(5.7) $\quad r_j > r_{j+1}$ if $s < j < t$ or $t < j < u$,

(5.8) $\quad r_j \geq r_{j+1}$ if $u < j < v$, and $\phi_j > \phi_{j+1}$ if $r_j = r_{j+1}$

or if $3 \leq j < s$, $\quad 0 < \phi_j < \pi$ if $3 \leq j \leq s$ or $u < j \leq v$

(5.9) $\quad \nu_1 + \nu_2 + 2\sum_{j=3}^{u} \nu_j + 4\sum_{j=u+1}^{v} \nu_j = d(A)$

(5.10) $\quad \sum_{j=s+1}^{u} \nu_j + 2\sum_{j=u+1}^{v} \nu_j \leq \mathrm{Min}(p(A), q(A))$

(5.11) $\quad$ if $\nu_1 = \nu_2 = 0$ then $p(A) \equiv \sum_{j=s+1}^{u} \nu_j \equiv q(A) \pmod 2$.

https://doi.org/10.4153/CMB-1958-021-6 Published online by Cambridge University Press

B) What are the characteristic roots of a matrix X satisfying (0.1) for a regular skew symmetric matrix A?

Answer:

(5.3) subject to the conditions (5.4), (5.5), (5.6), (5.7), (5.8), (5.9) and the condition

$$(5.12) \qquad \nu_1 \equiv \nu_2 \equiv d(A) \equiv 0(2)$$

C) What are the condition that two matrix pairs (X, A), (Y, B) are equivalent over the real field if A, B are regular matrices satisfying $A = \varepsilon A$, $B = \bar{\varepsilon} B$ ($\varepsilon = +1$, $\bar{\varepsilon} = +1$) ?

Answer:

$$\varepsilon = \bar{\varepsilon} \quad , \dim M_{P^{\mu}} = \dim M_{P^{*\mu}} = \dim \overline{M}_{P^{\mu}} = \dim \overline{M}_{P^{*\mu}}$$

where M is a representation space of (X, A) and $\overline{M}$ is a representation space of (Y, B) and P(x) is any asymmetric irreducible polynomial dividing the minimal polynomial of X in a multiplicity $\nu(P)$ that is not smaller than μ. Furthermore for any symmetric irreducible polynomial P(x) dividing the minimal polynomial of X with a positive multiplicity $\nu(P)$ the bilinear forms $f_{P,\mu}^{(X,A)}$, $f_{P,\mu}^{(Y,B)}$ are equivalent if $0 < \mu \leq \nu(P)$ and if $\varepsilon = (-1)^{\mu+1}$, and the bilinear forms $g_{P,\mu}^{(X,A)}$, $g_{P,\mu}^{(Y,B)}$ are equivalent if $0 < \mu \leq \nu(P)$ and if $\varepsilon = (-1)^{\mu}$.

D) To give a complete system of independent numerical equivalence invariants for matrix pairs (X, A) with representation space M over the real field in the event that A is regular and symmetric or skew symmetric?

Answer:

$\varepsilon = \pm 1$ determined by $A^T = \varepsilon A$; moreover for irreducible polynomials P(x) with highest coefficient 1 the numbers $e_{\mu}(P)$ denoting the dimension of the factor space of $M_{P^{\mu}}$ over $P(\sigma)M \cap M_{P^{\mu}} + M_{P^{\mu-1}}$ and for symmetric irreducible polynomials P(x) the numbers

https://doi.org/10.4153/CMB-1958-021-6 Published online by Cambridge University Press

$$p_{P,\mu} = \begin{cases} p(f_{P,\mu}) & \text{if } \varepsilon = (-1)^{\mu+1} \\ p(g_{P,\mu}) & \text{if } \varepsilon = (-1)^{\mu} \end{cases}$$

$$q_{P,\mu} = \begin{cases} q(f_{P,\mu}) & \text{if } \varepsilon = (-1)^{\mu+1} \\ q(g_{P,\mu}) & \text{if } \varepsilon = (-1)^{\mu} \end{cases}$$

subject to the conditions

(5.13) $e_{P,\mu} \geqslant 0,\ e_{P,\mu} \equiv 0\ ([P]),\ \sum\sum_{\mu} e_{P,\mu} = d(X) = d(A)$

(5.14) $p_{P,\mu} \geqslant 0,\ q_{P,\mu} \geqslant 0,\ p_{P,\mu} \equiv q_{P,\mu} \equiv 0\ ([P]),\ p_{P,\mu} + q_{P,\mu} \leq e_{P,\mu}$

(5.15) $p(A) = \frac{1}{2}(d(A) + \sum(p_{P,\mu} - q_{P,\mu})),\ q(A) = \frac{1}{2}(d(A) + \sum(q_{P,\mu} - p_{P,\mu}))$

(5.16) If $P(x) = x - \delta$ and $\delta^{\mu} + \varepsilon = 0$ then $p_{P,\mu} + q_{P,\mu} = e_{P,\mu}$.

E) To give a normal form of the equivalence class of matrix pairs with invariants as delimited under D.

Answer:

$$\dot{\sum_{\substack{P = P^{**} \\ \text{irreducible}}}} C_{P,\mu} \left(\begin{pmatrix} Y_{P^{\mu}} & \\ & Y_{P^{\mu}}^{-T} \end{pmatrix} \begin{pmatrix} & I_{[P]\mu} \\ \varepsilon I_{[P]\mu} & \end{pmatrix} \right)$$

$$\dot{+} \dot{\sum_{\substack{P = P^{*} \\ \text{irreducible}}}} \left[\dot{p}_{P,\mu} (X_{P,\mu}, A_{P,\mu}) + \dot{q}_{P,\mu} (X_{P,\mu} - A_{P,\mu-1}) \right]$$

where $[P]\,C_{P,\mu} = e_{P,\mu}$ if $P \neq P^*$, $[P]\,C_{P,\mu} = e_{P,\mu} - p_{P,\mu} - q_{P,\mu}$

if $P = P^*$, $X_{P,\mu} = ((\delta_{ik} + \delta_{i,k+1}) Y_P)$ $(i, k = 1, 2, \ldots, \mu)$,

$$A_{P,\mu} = \begin{cases} A_{\mu}(I_{[P]}) & \text{if } \varepsilon = (-1)^{\mu+1} \\ (-1)^{[\mu/2]} A_{\mu}\left(\begin{pmatrix} & 1 \\ -1 & \end{pmatrix}\right) & \text{if } \varepsilon = (-1)^{\mu} \end{cases}$$

F) Another question of E. Wigner: Let A be a regular symmetric matrix over the real field. What are the connected components of the multiplicative group G(A) formed by all matrices X satisfying (0.1) after all elements of G(A) with multiple characteristic roots i.e. the elements X of G(A) that

https://doi.org/10.4153/CMB-1958-021-6 Published online by Cambridge University Press

satisfy $R(\chi_x, D\chi_x) = 0$ (R the resultant polynomial), have been removed?

Answer:

The connected components of the residual space are in 1-1-correspondence to the system of all sequences

$$\alpha_1, \alpha_2, \ldots, \alpha_s, t, u, v$$

subject to the conditions (5.4) and

(5.17) $\alpha_i = 0, \pm 1$ if $i = 1, 2$; $\alpha_i = \pm 1$ if $2 < i \leq s$

(5.18) $\sum_{i=1}^{2} |\alpha_i| + 2\sum_{i=3}^{s} |\alpha_i| + 2(2v-u-S) = d(A)$

(5.19) $\frac{1}{2} \sum_{i=1}^{2} (\alpha_i + |\alpha_i|) + \sum_{i=3}^{s} (\alpha_i + |\alpha_i|) + 2v-u-s = p(A)$

and they are formed by all matrix pairs (X, A) subject to (5.3) where

(5.20) $\nu_i = |\alpha_i|$ if $1 \leq i \leq s$, $\nu_i = 1$ if $s < i \leq v$

G) To give a complete system of independent numerical equivalence invariants for matrix pairs (X, A) with representation space M over a Galois field of characteristic $\neq 2$ in the case that A is regular and either symmetric or skew symmetric.

Answer: $\varepsilon = \pm 1$ determined by $A^T = \varepsilon A$, moreover for irreducible polynomials P(x) with highest coefficient 1 the numbers $e_\mu(P)$ denoting the dimension of the factor space of M_{P^μ} over $P(\sigma)M \cap M_{P^\mu} + M_{P^{\mu-1}}$ and for symmetric irreducible polynomials P(x) the numbers

$$r_{P,\mu} = \begin{cases} r(f_{P,\mu}) \text{ if } \varepsilon = (-1)^{\mu+1} \\ r(g_{P,\mu}) \text{ if } \varepsilon = (-1)^{\mu} \end{cases}$$

and the numbers

$$\varepsilon_{P,\mu} = \begin{cases} \varepsilon(h_{P,\mu}) \text{ if } \varepsilon = (-1)^{\mu+1} \\ \varepsilon((Y_P - Y_P^{-1})h_{P,\mu}) \text{ if } \varepsilon = (-1)^{\mu} \end{cases}$$

where for any hermitian form h over a Galois field of q^2 elements (q odd)

$$\varepsilon(h) = \begin{cases} 1 \text{ if } h \sim \sum_{i=1}^{r} \xi_i \eta_i^q \\ -1 \text{ if } h \sim \nu \xi_1 \eta_1^q + \sum_{i=2}^{r} \xi_i \eta_i^q \\ 0 \text{ if } h = o \end{cases}$$

and ν is a non square of GF(q).

Conclusion: Our methods fail in the case of inseparable irreducible divisors of the characteristic polynomial of X. The case of characteristic 2 leads to more complicated equations the solution of which is not given here. Due to the results of this paper the problem of the classification of the hermitian forms over arbitrary fields with an involutoric automorphism has taken on still another aspect.

BIBLIOGRAPHY

L.E. Dickson, *Linear groups*, (Leipzig,1901).
J. Dieudonné, *Sur les groupes classiques*, (Paris, 1948).

McGill University

https://doi.org/10.4153/CMB-1958-021-6 Published online by Cambridge University Press