ON THE CONJUGACY CLASSES IN THE UNITARY, SYMPLECTIC AND ORTHOGONAL GROUPS

G. E. WALL

(received 1 May 1962)

Introduction

This paper is concerned with two main problems:

(a) *the determination of the conjugacy classes in the finite-dimensional unitary, symplectic and orthogonal groups over division rings or fields*;

(b) *the determination of the equivalence classes of non-degenerate sesquilinear forms on finite-dimensional vector spaces.*

To "solve" these problems means to reduce them to standard problems of linear algebra such as the similarity of matrices and the equivalence of Hermitian or quadratic forms.

The matrix formulation of (b) is as follows:

(b)′ *to find the conditions that two non-singular matrices A, B be "congruent", i.e. that $B = T^*AT$ for some matrix T, where $*$ is an operation of the conjugate transpose type.*

Now, if the characteristic of the coefficient domain is not 2, we may split A, B into their "Hermitian" and "skew-Hermitian" components:

$$A' = \tfrac{1}{2}(A + A^*),\ A'' = \tfrac{1}{2}(A - A^*),\ B' = \cdots;$$

and then the single equation $B = T^*AT$ is replaced by the pair of equations $B' = T^*A'T$, $B'' = T^*A''T$. Thus problem (b) is substantially equivalent to the classification problem for non-singular pairs of forms (f, g), f Hermitian, g skew-Hermitian. For earlier work on problem (b) in this form, see Trott ([10]), Ingraham and Wegner ([4]), Turnbull ([11]).

The decisive contributions to the present subject were made by J. Williamson, who solved the conjugacy problem over perfect fields of characteristic not 2 ([17], [18]) and the equivalence problem over arbitrary fields of characteristic not 2 ([14], [15], [19]). T. A. Springer ([9]) later determined the conjugacy classes, and the centralizers of the elements, in symplectic groups over arbitrary fields of characteristic not 2. See also Venkatachaliengar ([12]), Klingenberg ([7]), Zassenhaus ([20]) and a forthcoming paper by V. Ennola ([21]). A general survey is given in Pickert's encyclopedia article [8].

In the present paper, Williamson's work is coordinated and extended. Firstly, I show that there is a quite simple reduction of the conjugacy problem to the equivalence problem (§ 1). With each element of a classical group is associated a non-degenerate sesquilinear form (§ 1.1) and two elements are conjugate if, and only if, their sesquilinear forms are equivalent (thm. 1.3.1). This association of a sesquilinear form with a group element generalizes the classical Cayley parametrization in orthogonal groups.

Secondly, I clarify and extend Williamson's matrix methods for the equivalence problem by putting them in a structural setting (§ 2). Let D be a division ring with an involutory anti-automorphism J, $f(x, y)$ a non-degenerate sesquilinear form over D. The *multiplier* of f is the non-singular linear transformation P defined by $f(y, x)^J = f(x, yP)$. Let $\mathscr{C}$ denote the ring of linear transformations which commute with P. The adjoint $X^\dagger$ of a linear transformation X with respect to f is defined by $f(xX, y) = f(x, yX^\dagger)$. The adjoint mapping $X \to X^\dagger$ $(X \in \mathscr{C})$ is an involutory anti-automorphism of $\mathscr{C}$. In the equivalence problem we may, without loss of generality, suppose that all forms have the one fixed multiplier P. Let $f_i(x, y) = f(x, yQ_i)$ $(i = 1, 2)$ be two such forms, represented with respect to the fixed standard form f by the linear transformations Q_i. Then Q_1, Q_2 are non-singular, $\dagger$-symmetric $(Q_i = Q_i^\dagger)$ elements of $\mathscr{C}$ and f_1, f_2 are equivalent if, and only if, Q_1, Q_2 are $\dagger$-congruent in $\mathscr{C}$ ($Q_2 = XQ_1X^\dagger$ for some $X \in \mathscr{C}$). This preliminary reduction of the equivalence problem is perfectly elementary and depends only on simple calculations with forms or matrices (§ 2.1).

Now let $\mathscr{N}$ be the radical of $\mathscr{C}$. The factor ring $\mathscr{C}/\mathscr{N}$ is a direct sum of total matrix algebras over certain division rings. The canonical mapping $X \to X + \mathscr{N}$ carries the anti-automorphism $\dagger$ over to $\mathscr{C}/\mathscr{N}$ and the problem of $\dagger$-congruence in $\mathscr{C}/\mathscr{N}$ in fact reduces to the problem of congruence of Hermitian matrices over division rings. In our formulation, Williamson's central result is that the following *Approximation theorem* holds whenever D is a field of characteristic not 2:

two non-singular, $\dagger$-symmetric elements of $\mathscr{C}$ are $\dagger$-congruent in $\mathscr{C}$ if, and only if, their canonical images in $\mathscr{C}/\mathscr{N}$ are $\dagger$-congruent in $\mathscr{C}/\mathscr{N}$ (*thm.* 2.2.1). Clearly, when the Approximation theorem holds, the equivalence class of a non-degenerate sesquilinear form is determined by the similarity class of its multiplier and the equivalence classes of the Hermitian matrices which arise from the $\dagger$-congruence problem in $\mathscr{C}/\mathscr{N}$.

The Approximation theorem certainly holds when either (a) the characteristic of D is not 2 or (b) the restriction of J to the centre of D is not the identity or (c) $I - P$ is non-singular (lemma 2.2.1, corr.). Thus, e.g., the essential features of Williamson's theory hold in the finite unitary groups $U(n, 2^{2a})$, though not in the finite symplectic or orthogonal groups $Sp(2m, 2^a)$, $O(n, 2^a)$.

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

In § 3, I examine the exceptional case where D is a field of characteristic 2, J is the identity and P has characteristic polynomial $(1-t)^k$. A weak form of the Approximation theorem (in which $\mathcal{N}$ is replaced by a smaller ideal) gives some reduction of the problem, though the resulting system of equations is still formidable. A complete solution is obtained when D satisfies the hypothesis:

every ternary quadratic form $x^2 + xy + \lambda y^2 + \mu z^2$ $(\lambda, \mu \in D)$ *represents zero non-trivially.*

This hypothesis certainly holds when D is perfect or when every quadratic extension of D is inseparable. Thus, e.g., the theory can be applied to the symplectic and orthogonal groups over $GF(2^a)$.

In §§ 2.6, 3.7, I determine the total number of conjugacy classes, and the individual class orders, in the finite groups $U(n, q^2)$, $Sp(2m, q)$, $O(n, q)$. This is a necessary first step in the more difficult problem of their matrix representations.

0.1 *Notation.* D denotes a division ring with a fixed involutory anti-automorphism J. Thus $(\alpha+\beta)^J = \alpha^J + \beta^J$, $(\alpha\beta)^J = \beta^J\alpha^J$, $\alpha^{J^2} = \alpha$ for all $\alpha, \beta \in D$. We remark that if J is the identity D must be a field, for then $\alpha\beta = (\alpha\beta)^J = \beta^J\alpha^J = \beta\alpha$ whenever $\alpha, \beta \in D$. The element $\alpha \in D$ is called symmetric if $\alpha^J = \alpha$, skew if $\alpha^J = -\alpha$. The elements of D are usually called scalars and denoted by lower case Greek letters.

p (= 0 or a prime) is the characteristic of D. Z is the centre of D. $D[t]$, $Z[t]$ denote the rings of polynomials in an indeterminate t over D, Z respectively. $\phi(t) \in D[t]$ is called monic if the coefficient of the highest power of t is 1. If $\phi(t) = \sum \alpha_i t^i$, we define $\phi^J(t) = \sum \alpha_i^J t^i$.

All vector spaces considered are left vector spaces over D. Linear transformations are regarded as right multipliers. Thus, if λ is a scalar, v a vector and T a linear transformation, we write λv, vT. Composition of linear transformations is defined by $v(T_1T_2) = (vT_1)T_2$. If T is a linear transformation and $\phi(t) = \sum \alpha_i t^i \in Z[t]$, then the linear transformation $\phi(T)$ is defined, as usual, by $v\phi(T) = \sum \alpha_i(vT^i)$. The commutator ring, $\mathscr{C}(T)$, of T is the ring formed by the linear transformations which commute with T.

Suppose that V is a vector space over D (of finite or infinite dimension). A sesquilinear form on V (strictly: J-sesquilinear form on V) is a mapping $f : V \times V \to D$ such that $f(u, v)$ is linear in u for each fixed v and anti-linear in v for each fixed u, i.e.,

$$f(\lambda_1 u_1 + \lambda_2 u_2, v) = \lambda_1 f(u_1, v) + \lambda_2 f(u_2, v),$$
$$f(u, \mu_1 v + \mu_2 v) = f(u, v_1)\mu_1^J + f(u, v_2)\mu_2^J,$$

for all $u, v, u_i, v_i \in V$ and $\lambda_i, \mu_i \in D$. f is called non-degenerate if $f(u, v) = 0$ for all v implies $u = 0$ and $f(u, v) = 0$ for all u implies $v = 0$; otherwise f is degenerate. f is called Hermitian (strictly: J-Hermitian) if $f(u, v) = f(v, u)^J$

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

for all u, v, skew-Hermitian if $f(u, v) = -f(v, u)^J$ for all u, v. Let f_1, f_2 be sesquilinear forms on vector spaces V_1, V_2. Then f_1, f_2 are said to be equivalent if there exists a linear isomorphism $T : V_1 \to V_2$ (i.e., 1 — 1 linear mapping of V_1 onto V_2) such that $f_2(uT, vT) = f_1(u, v)$ for all u, $v \in V$. If f_1, f_2 are equivalent, we write $f_1 \approx f_2$.

Let $\Omega = (\omega_{ij})$ be a matrix over D. The transpose and conjugate transpose of Ω are denoted by Ω^T, Ω^* respectively: $\Omega^T = (\omega_{ji})$, $\Omega^* = (\omega_{ji}^J)$. If square matrices Ω_1, Ω_2 are similar we write $\Omega_1 \sim \Omega_2$; the same notation is used for linear transformations. Two square matrices Ω_1, Ω_2 are called J-congruent if $\Omega_2 = T\Omega_1 T^*$ for some non-singular T.

Matrix notation is introduced as follows. Suppose that the vector space V is finite-dimensional with basis $e_1, \cdots, e_n$. Then $x = \sum \xi_i e_i \in V$ is represented by the row vector $x = (\xi_i)$, a linear transformation T on V by the square matrix $T = (\tau_{ij})$, where $e_i T = \sum \tau_{ij} e_j$ $(i = 1, \cdots n)$, and a sesquilinear form f on V by the square matrix $\Phi = (\phi_{ij})$, where $\phi_{ij} = f(e_i, e_j)$ $(i, j = 1, \cdots n)$. With these conventions, xT is represented by xT, $T_1 T_2$ by $T_1 T_2$ and $f(x, y) = x\Phi y^*$. Forms f_1, f_2 are equivalent if, and only if, their matrices Φ_1, Φ_2 are J-congruent.

0.2. *Direct decompositions.* Let V be a vector space over D, T a linear transformation on V. Let

$$V = V_1 \oplus \cdots \oplus V_k$$

be a direct sum decomposition of V. If each V_i is invariant under T, we write

$$T = T_1 \oplus \cdots \oplus T_k, \tag{0.2.1}$$

where T_i is the restriction of T to V_i $(i = 1, \cdots k)$. (0.2.1) is called a direct decomposition of T.

Suppose now that V is finite-dimensional. Let $\phi(t)$ be a non-constant element of $Z[t]$. Then there exists a unique Fitting decomposition

$$V = V_0 \oplus V_1, \quad T = T_0 \oplus T_1,$$

where $\phi(T_1)$ is nilpotent and $\phi(T_0)$ non-singular. V_0, V_1 are respectively the image- and null-spaces of $\phi(T)^r$ for sufficiently large r.

More generally, let $\phi_1(t), \cdots, \phi_s(t)$ be non-constant elements of $Z[t]$ such that $\phi_i(t)$, $\phi_j(t)$ are relatively prime whenever $i \neq j$. Then there is a unique Fitting decomposition:

$$V = V_0 \oplus \cdots \oplus V_s, \quad T = T_0 \oplus \cdots \oplus T_s,$$

where, for $i = 1, \cdots s$, $\phi_i(T_i)$ is nilpotent and $\phi_i(T_j)$ $(j \neq i)$ non-singular.

Let $S \in \mathscr{C}(T)$. Because of the uniqueness of the Fitting decomposition, $S = S_0 \oplus \cdots \oplus S_s$, where $S_i \in \mathscr{C}(T_i)$ $(i = 0, \cdots s)$.

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

0.3. *Linear transformations of commutative type.* Let V be a finite-dimensional space over D, T a linear transformation on V. Let $\phi(t)$ be an element of $Z[t]$ *which is irreducible qua element of* $D[t]$ (e.g., $\phi(t) = t - \alpha$, $\alpha \in Z$). Then T is called ϕ-nul if $\phi(T)$ is nilpotent. The term ϕ-nul will only be used when $\phi(t)$ satisfies the conditions above. T is said to be of commutative type if it is ϕ-nul for some ϕ.

For a ϕ-nul linear transformation T, the theory of elementary divisors holds in its customary form. The matrix $\boldsymbol{T}$ of T with respect to a suitable basis has coefficients in Z and the reduction of $\boldsymbol{T}$ to the direct sum of the companion matrices of its elementary divisors gives rise to a splitting of T into indecomposable parts (see Jacobson [6], Ch. 3, §§ 9—10). In view of the Fitting decomposition, we may speak of the multiplicity of ϕ^e as elementary divisor of an *arbitrary* linear transformation S on V.

Suppose now that T is *indecomposable* and ϕ-nul, say with minimum polynomial ϕ^e. Then there is a vector u such that $u, uT, \cdots, uT^{n-1}$ form a basis of V. If $g(t) = \sum \alpha_i t^i \in D[t]$, we may define the linear transformation $g(T)$ by

$$(\sum \lambda_i u T^i) g(T) = \sum \lambda_i \alpha_j u T^{i+j}. \tag{0.3.1}$$

(The definition of $g(T)$ depends on the choice of u unless $g(t) \in Z[t]$.) The $g(T)$ form a subring $D[T]$ of the commutator ring $\mathscr{C}(T)$. On the other hand, an element S of $\mathscr{C}(T)$ is uniquely determined by uS, which has the form $ug(T)$ for some g. Hence $D[T] = \mathscr{C}(T)$. It is easy to see that $Z[T]$ is the centre of $\mathscr{C}(T)$. We have $\mathscr{C}(T) \cong D[t]/\phi(t)^e D[t]$, $Z[T] \cong Z[t]/\phi(t)^e Z[t]$. The only ideals (left or right) of $\mathscr{C}(T)$ are the 2-sided ideals $\mathscr{C}(T)\phi(T)^i$ $(i = 0, 1, \cdots e)$. The only subspaces of V invariant under T are the subspaces $V\phi(T)^i (i = 0, 1, \cdots e)$.

We remark that the division ring $D[t]/\phi(t)D[t]$ can be identified with the ring of polynomials over D in a quantity τ which commutes with the elements of D and satisfies $\phi(\tau) = 0$. We say that this ring is obtained by adjoining a root of $\phi(t)$ to D.

1. Conjugacy

1.1 *Parametrization.* Let V be a left vector space over D, of finite or infinite dimension, and $F(x, y) = x \cdot y$ a non-degenerate Hermitian or skew-Hermitian form on V. Thus $(y \cdot x)^J \equiv \varepsilon(x \cdot y)$, where $\varepsilon = \pm 1$.

Let $M, N \subset V$. We say that M is *F-perpendicular* to N (notation: $M \perp N$) if $x \cdot y = 0$ whenever $x \in M$, $y \in N$. The subspace $M^\perp = \{x \mid M \perp x, x \in V\}$ is the *F-perpendicular space* of M. If W is a subspace of V, the subspace $W^\rho = W \cap W^\perp$ is called the *radical* of W.

Let W_1, W_2 be subspaces of V. A linear isomorphism $X : W_1 \to W_2$ is

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

called an *isometry* of W_1 onto W_2 if

$$xX \cdot yX = x \cdot y \quad \text{for all} \quad x, y \in W_1.$$

The isometries of V onto itself form the *unitary* group $U(F)$ of F. This formulation includes the symplectic and unitary groups as usually defined and the orthogonal groups over fields of characteristic not 2. The orthogonal groups over fields of characteristic 2 are treated separately in § 1.5. In this section we obtain a parametrization for the elements of $U(F)$ (cf. [13]).

Let $X \in U = U(F)$. Write $X = I - T$, where I is the identity mapping. Then the subspace VT is called the *space of* X and denoted by V_X. The *dimension of* X is defined to be the dimension of V_X. It is easily verified that the finite-dimensional elements of $U(F)$ form a normal subgroup $U_\phi(F)$.

In terms of T, the equation of invariance $xX \cdot yX = x \cdot y$ becomes

$$x \cdot yT + xT \cdot y = xT \cdot yT. \tag{1.1.1}$$

(1.1.1) shows that the scalar $x \cdot yT$ is determined by the vectors xT, yT alone; therefore the equation

$$(xT, yT) = x \cdot yT \tag{1.1.2}$$

uniquely defines a function (u, v) of the variables u, v in V_X. (u, v) is called the *form of* X and denoted by F_X. By (1.1.1),

$$(u, v) + \varepsilon(v, u)^J = u \cdot v. \tag{1.1.3}$$

Since $X^{-1} = I + X^{-1}T = I + TX^{-1}$, we have $V_{X^{-1}} = V_X$. Let $[u, v]$ denote the form of X^{-1}. Then, by definition,

$$x \cdot yT = -[xX^{-1}T, yT], \quad y \cdot xX^{-1}T = (yT, xX^{-1}T),$$

and, by (1.1.1),

$$-(x \cdot yT) = xT \cdot yX = xX^{-1}T \cdot y = \varepsilon(y \cdot xX^{-1}T)^J,$$

so that

$$[u, v] = \varepsilon(v, u)^J. \tag{1.1.4}$$

LEMMA 1.1.1. *F_X is a non-degenerate sesquilinear form.*

PROOF. This means that

(a) the mappings $u \to (u, v)$ and $u \to (v, u)^J = \varepsilon[u, v]$ of V_X into D are linear for each $v \in V_X$, and

(b) if $(u, v) = 0$ for all $u \in V_X$, or if $(v, u) = \varepsilon[u, v]^J = 0$ for all $u \in V_X$, then $v = 0$.

(a) is obvious from (1.1.2). (b) follows from (1.1.2) and the non-degeneracy of F.

COROLLARY. *$V_X^\perp$ is the null-space of $I - X$.*

This follows from the lemma and (1.1.2).

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

LEMMA 1.1.2. *X is uniquely determined by V_X and F_X.*

PROOF. Suppose that $X = I - T$ and $Y = I - S$ both have the same space W and form (u, v) Then

$$x \cdot w = (xT, w) = (xS, w)$$

for all $x \in V$, $w \in W$. Since (u, v) is non-degenerate, $xT = xS$ for all $x \in V$, i.e. $T = S$. Hence $X = Y$, Q.E.D

We consider the families

$$L = \{l_v | v \in V_X\}, \quad R = \{r_v | v \in V_X\}, \quad K = \{k_x | x \in V\},$$

where l_v, r_v, k_x are the following linear functions on V_X:

$$l_v(u) = (u, v), \quad r_v(u) = [u, v], \; k_x(u) = u \cdot x.$$

LEMMA 1.1.3. $L = R = K$.

PROOF. The equation $[v, xT] = v \cdot x$ is an easy consequence of (1.1.2), (1.1.4). Hence $R = K$ and similarly $L = K$.

THEOREM 1.1.1. *Let W be a subspace of V, $(u, v)'$ a non-degenerate sesquilinear form on W satisfying*

$$(u, v)' + \varepsilon(v, u)'^J = u \cdot v \tag{1.1.3}'$$

for all $u, v \in W$. Suppose further that $L' = R' = K'$, where

$$L' = \{l'_v | v \in W\}, \quad R' = \{r'_v | v \in W\}, \quad K' = \{k'_x | x \in V\},$$
$$l'_v(w) = (w, v)', \quad r'_v(w) = \varepsilon(v, w)'^J, \quad k'_x(w) = w \cdot x \qquad (w \in W).$$

Then $V_X = W$, $F_X = (u, v)'$ for a unique $X \in U(F)$.

PROOF. Since $R' = K'$, and since F and $(u, v)'$ are non-degenerate, there is a unique linear mapping $T : V \to V$ with $VT = W$ such that

$$(xT, v)' = x \cdot v \quad \text{for all} \quad x \in V, v \in W.$$

Similarly, there is a unique linear mapping $S : V \to V$ with $VS = W$ such that

$$(u, yS)' = u \cdot y \quad \text{for all} \quad u \in W, y \in V.$$

Let $X = I - T$, $Y = I - S$. We prove the theorem by showing that X and Y are mutually inverse elements of U.

Using (1.1.3)', we get

$$-\varepsilon(v, xT)'^J = xX \cdot v = \varepsilon(v \cdot xX)^J = \varepsilon(v, xXS)'^J.$$

Since $(u, v)'$ is non-degenerate, it follows that $-xT = xXS$ and so $-T = XS$. Thus $XY = I$ and similarly $YX = I$. Putting $u = xT$, $v = yT$ in (1.1.3)', we get $xX \cdot yX = x \cdot y$ so that $X \in U$. This completes the proof.

Theorem 1.1.1 simplifies when the dimension of W is finite.

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

THEOREM 1.1.2. *Let W be a finite-dimensional subspace of V, $(u, v)'$ a non-degenerate sesquilinear form on W satisfying* (1.1.3)'. *Let $\Omega = (\omega_{ij})$ be the matrix of $(u, v)'$ with respect to a basis $e_1, \cdots, e_r$ of W and let $\Theta = (\theta_{ij})$ be the inverse of Ω. Then the linear mapping*

$$xX = x - \sum_{i,j=1}^{r} (x \cdot e_i)\theta_{ij} e_j \tag{1.1.5}$$

is an element of U such that $V_X = W$, $F_X = (u, v)'$.

PROOF. Since W is finite-dimensional and $(u, v)'$ non-degenerate, each of the families L', R', K' coincides with the dual space of W. Hence there is a unique $X \in U$ such that $V_X = W$, $F_X = (u, v)'$. Let $X = I - T$, $xT = \sum \lambda_i e_i$. Then

$$x \cdot e_j = (xT, e_j)' = \sum \lambda_j \omega_{ij} \quad (j = 1, \cdots r)$$

and so

$$\lambda_j = \sum (x \cdot e_i)\theta_{ij} \quad (j = 1, \cdots r)$$

as required.

We note that the matrix formulation of (1.1.3)' is

$$\Omega + \varepsilon \Omega^* = \Phi, \tag{1.1.6}$$

where Ω^* is the conjugate transpose $(\omega_{ij}^J)^T$ of Ω and Φ the matrix of the restriction of F to W.

EXAMPLE. Let U be the n-dimensional orthogonal group in the classical Euclidean sense. Let $W = V$ and let $e_1, \cdots e_n$ be an orthonormal basis of V. (1.1.6) becomes $\Omega + \Omega^T = I$. The general non-singular solution is $\Omega = \frac{1}{2}(I + S)$, where S is skew-symmetric. By (1.1.5), the matrix of X is $X = (S - I)(S + I)^{-1}$. This is the Cayley paramatrization for the "non-exceptional" elements of U.

1.2 *Witt's theorem.* This deals with the extension of isometries $X : W_1 \to W_2$ to isometries $Y : V \to V$, i.e. to elements of U. The present account is rather more general than that of Dieudonné ([2]).

If $x \in V$ then $(x \cdot x)^J = \varepsilon(x \cdot x)$. We call x *trace-valued* if $x \cdot x = \lambda + \varepsilon\lambda^J$ for some $\lambda \in D$.

LEMMA 1.2.1. *The trace-valued vectors form a subspace V^τ. If $p \neq 2$, $V^\tau = V$.*

PROOF. Let $x, y \in V^\tau$, so that $x \cdot x = \lambda + \varepsilon\lambda^J$, $y \cdot y = \mu + \varepsilon\mu^J$. Then $(\alpha x + \beta y) \cdot (\alpha x + \beta y) = \rho + \varepsilon\rho^J$, where $\rho = \alpha\lambda\alpha^J + \beta\mu\beta^J + \alpha(x \cdot y)\beta^J$. Hence V^τ is a subspace. If $x \in V$ and $p \neq 2$, then $x \cdot x = \frac{1}{2}(x \cdot x) + \varepsilon(\frac{1}{2}(x \cdot x))^J$; hence $V = V^\tau$.

LEMMA 1.2.2. *If $X \in U$, $V_X \subset V^\tau$.*

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

This follows from (1.1.3).

COROLLARY. *If $X \in U$, X leaves the spaces V/V^τ and $(V^\tau)^\perp$ pointwise invariant.*

This follows from the lemma and lemma 1.1.1, cor.

THEOREM 1.2.1. (Witt). *Let W_1, W_2 be finite-dimensional subspaces of V^τ such that $W_1 \cap (V^\tau)^\perp = W_2 \cap (V^\tau)^\perp = \{0\}$. Then every isometry X of W_1 onto W_2 can be extended to an element of $U_\phi(F)$.*

PROOF. The theorem is trivial when the dimension, r, of W_1, W_2 is 0. Suppose that $r > 0$ and that the theorem holds for lower dimensions. Let $e_1, \cdots e_r$ form a basis of W_1. Then $f_1, \cdots, f_r$, where $f_i = e_i X$, form a basis of W_2. By the induction hypothesis, and since each element of U leaves V^τ and $(V^\tau)^\perp$ invariant, we may assume that $e_i = f_i$ for $i = 1, \cdots, r-1$. Then, since X is an isometry, we have

$$(1.2.1) \qquad e \cdot c + \varepsilon(e \cdot c)^J = c \cdot c, \quad e_i \cdot c = 0 \quad (i = 1, \cdots, r-1)$$

where

$$e = e_r, \quad c = e_r - f_r.$$

We may of course assume that $c \neq 0$. Since $W_1 \cap (V^\tau)^\perp = W_2 \cap (V^\tau)^\perp = \{0\}$, each of the sets $e_1, \cdots e_{r-1}$, e and $e_1, \cdots, e_{r-1}$, $e - c$ is linearly independent modulo $(V^\tau)^\perp$.

If $e \cdot c \neq 0$, the 1-dimensional element

$$xY = x - (x \cdot c)(e \cdot c)^{-1} c$$

of U extends X.

Suppose next that $e \cdot c = 0$. We seek a 2-dimensional element Z of U which extends X. Choose a vector $d \in V^\tau$ such that

$$(1.2.2) \qquad e_i \cdot d = 0 \ (i = 1, \cdots, r-1), \quad e \cdot d = 1, \quad c \cdot d \neq 1.$$

Since $e_1, \cdots e_{r-1}$, e and c $(= e - f_r)$ are in V^τ by hypothesis, such a choice is possible when $e_1, \cdots e_{r-1}$, e, c are linearly independent modulo $(V^\tau)^\perp$. In the contrary case, since $e_1, \cdots e_{r-1}$, e and $e_1, \cdots e_{r-1}$, $e - c$ are linearly independent sets modulo $(V^\tau)^\perp$, we have a relation

$$c \equiv \sum_1^{r-1} \lambda_i e_i + \lambda e \ (\mathrm{mod}(V^\tau)^\perp),$$

where $\lambda \neq 1$. Then, choosing $d \in V^\tau$ such that the first r equations in (1.2.2) hold, we have $c \cdot d = \lambda \neq 1$ as required.

Since $d \in V^\tau$, $d \cdot d = \alpha + \varepsilon\alpha^J$ for some $\alpha \in D$. Since $e \cdot c = 0$ and $e \cdot d = 1$, c and d are linearly independent. Let W be the subspace with basis c, d. Then it is easily verified that the following conditions define an element Z of U which extends X:

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

$$V_Z = W,$$
$$F_Z(c, c) = 0, \qquad F_Z(c, d) = 1,$$
$$F_Z(d, c) = \varepsilon((c \cdot d)^J - 1), \qquad F_Z(d, d) = \alpha.$$

This proves the theorem.

DEFINITION. *F is called trace-valued if $V^\tau = V$.*

COROLLARY. *If F is trace-valued, every isometry of finite-dimensional subspaces of V can be extended to an element of $U_\phi(F)$.*

It is easily seen that F is trace-valued if, and only if, its matrix has the form $\Omega + \varepsilon\Omega^*$ for some matrix Ω. If $p \neq 2$, every F is trace-valued by lemma 1.2.1. If $p = 2$ and J is the identity, F is trace-valued if, and only if, it is an alternate form, i.e. $x \cdot x = 0$ for all x.

1.3 *Conjugacy.*

LEMMA 1.3.1. *Let X, Y be conjugate elements of $U(F) : Y = Z^{-1}XZ$, where $Z \in U(F)$. Then $V_Y = V_X Z$ and $F_Y(uZ, vZ) = F_X(u, v)$ for all u, $v \in V_X$.*

PROOF. Let $X = I - T$. Then $Y = I - Z^{-1}TZ$ and so $V_Y = VZ^{-1}TZ = V_X Z$. Also

$$F_Y(xTZ, yTZ) = xZ \cdot yTZ = x \cdot yT = F_X(xT, yT),$$

which gives the second part of the lemma.

COROLLARY 1. *Conjugate elements of U have equivalent forms.*

COROLLARY 2. *Let X, $Z \in U$. Then $XZ = ZX$ if, and only if, Z leaves V_X and F_X invariant. In particular, $XZ = ZX$ if V_X, V_Z are F-perpendicular.*

The following theorem reduces the conjugacy problem in finite-dimensional, trace-valued classical groups to the equivalence problem for finite-dimensional, non-degenerate sesquilinear forms.

THEOREM 1.3.1. *Let X, $Y \in U_\phi(F)$ and suppose that $V_X \cap (V^\tau)^\perp = V_Y \cap (V^\tau)^\perp = \{0\}$. Then X, Y are conjugate in $U_\phi(F)$ (or $U(F)$) if (and only if) F_X, F_Y are equivalent forms.*

PROOF. By hypothesis, there is a linear isomorphism $\tilde{Z}$ of V_X onto V_Y such that $F_Y(u\tilde{Z}, v\tilde{Z}) = F_X(u, v)$ for all $u, v \in V_X$. By (1.1.3), $\tilde{Z}$ is an isometry. Hence, by Witt's theorem, $\tilde{Z}$ can be extended to an element Z of U_ϕ. Then Y, $Z^{-1}XZ$ have the same space and form, so that $Y = Z^{-1}XZ$. This proves the theorem.

COROLLARY. *Let F be trace-valued. Then two elements of U_ϕ are conjugate in U_ϕ (or U) if, and only if, their forms are equivalent.*

1.4 *Direct Sums and Products.* We consider the direct and semi-direct

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

decompositions of F_X and their relation to the group-theoretical properties of X. We write $V_X = W$, $F_X = f$ and assume throughout that the dimension, m, of W is finite.

Let $M, N \subset W$. We say that M is *f-perpendicular* to N (notation: $M\omega N$) if $(x, y) = 0$ whenever $x \in M, y \in N$. The relation $M\omega N$ is in general not symmetric. The subspaces

$$M^\omega = \{x | M\omega x, x \in W\}, \quad {}^\omega M = \{x | x\omega M, x \in W\},$$

are called the *right* and *left f-perpendicular* spaces of M respectively.

Let

$$W = W_1 \oplus \cdots \oplus W_k \tag{1.4.1}$$

be a direct decomposition of the vector space W. If $W_j \omega W_i$ whenever $1 \leqq i < j \leqq k$, we write

$$f = f_1 + f_2 + \cdots + f_k, \tag{1.4.2}$$

where f_i is the restriction f_{W_i} of f to W_i. We call (1.4.2) a *semi-direct sum* and say that each f_i is a semi-direct summand of f. Notice that the addition in (1.4.2) is not in general commutative. If $W_i \omega W_j$ whenever $i \neq j$, we write

$$f = f_1 \oplus f_2 + \cdots \oplus f_k. \tag{1.4.3}$$

We call (1.4.3) a *direct sum* and say that each f_i is a direct summand of f. In this case

$$W_i^\omega = {}^\omega W_i = W_1 \oplus \cdots \oplus W_{i-1} \oplus W_{i+1} \oplus \cdots \oplus W_k$$

for each i, and it follows from (1.1.3) that (1.4.1) *is an F-perpendicular decomposition*, i.e. $W_i \perp W_j$ whenever $i \neq j$.

Choose a basis of W adapted to the direct decomposition (1.4.1). If (1.4.2) or (1.4.3) holds, the matrix of f has the form

$$\begin{pmatrix} \Omega_{11} & \Omega_{12} & \cdots & \Omega_{1k} \\ 0 & \Omega_{22} & \cdots & \Omega_{2k} \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & \Omega_{kk} \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} \Omega_{11} & 0 & \cdots & 0 \\ 0 & \Omega_{22} & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & \Omega_{kk} \end{pmatrix},$$

and conversely. Clearly, Ω_{ii} is the matrix of f_i and f_i is non-degenerate.

LEMMA 1.4.1. *Let M be a subspace of W. Then f_M (the restriction of f to M) is a semi-direct summand of f if (and only if) it is non-degenerate.*

PROOF. Since f_M is non-degenerate, $M \cap {}^\omega M = \{0\}$ and therefore $W = M \oplus {}^\omega M$. Hence $f = f_M + f_{({}^\omega M)}$, Q.E.D.

In order to have an analogous criterion for direct summands, we introduce the *multiplier* of f. This is defined to be the (non-singular) linear transforma-

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

tion P on W such that

$$(y, x)^J = (x, yP) \quad \text{for all} \quad x, y \in W. \tag{1.4.4}$$

In terms of the matrices $\boldsymbol{\Omega}$, $\boldsymbol{P}$ of f, P:

$$\boldsymbol{P} = \boldsymbol{\Omega}\boldsymbol{\Omega}^{*-1}. \tag{1.4.4}'$$

We shall also refer to P as the multiplier of X and write $P = P_X$.

It follows easily from (1.4.4) that

$$(x\phi(P), y) = (x, y\phi^J(P^{-1})) \tag{1.4.5}$$

when $\phi(t) \in Z[t]$ (Z the centre of D).

LEMMA 1.4.2. *Let M be a subspace of W. Then ${}^\omega M = M^\omega$ if, and only if, M is invariant under P.*

PROOF. By (1.4.5), $M^\omega = {}^\omega(MP)$. On the other hand, since M is a subspace the relation ${}^\omega M = {}^\omega(MP)$ is equivalent to the relation $M = MP$.

LEMMA 1.4.3. *Let M be a subspace of W. Then f_M is a direct summand of f if, and only if, f_M is non-degenerate and M is invariant under P.*

PROOF. If f_M is a direct summand it is a semi-direct summand and ${}^\omega M = M^\omega$. By the previous lemmas, f_M is non-degenerate and M is invariant under P. Conversely, if these conditions hold then ${}^\omega M = M^\omega$ by lemma 1.4.2, so that the sum $f = f_M + f_{({}^\omega M)}$ in lemma 1.4.1 is direct.

THEOREM 1.4.1. *Let* (1.4.1), (1.4.2) *hold. Then*

$$X = X_1 X_2 \cdots X_k, \tag{1.4.6}$$

where X_i is the element of U such that $V_{X_i} = W_i$, $F_{X_i} = f_i$ $(i = 1, \cdots k)$. (1.4.3) *holds if, and only if, $X_i X_j = X_j X_i$ for all i, j.*

PROOF. The first statement was proved in [13]. The second follows immediately from the fact that (1.4.3) holds if, and only if, $f = f_{i_1} + \cdots + f_{i_k}$ for every permutation $i_1, \cdots i_k$ of $1, \cdots k$.

We call (1.4.6) a *semi-direct factorization*. If (1.4.3) holds, we call it a *direct factorization*, We mention, without proof, the

COROLLARY *. *If F_X is not an alternate form, the m-dimensional element X is a product of m 1-dimensional elements.*

There is a simple direct relation between X and its multiplier P.

LEMMA 1.4.4. *The restriction of X to V_X is $-\varepsilon P^{-1}$.*

PROOF. Let $u, v \in W$. By (1.1.3) and (1.4.5),

$$u \cdot v = (u(I + \varepsilon P^{+1}), v)$$

* The well known theorem that every orthogonal transformation is a product of symmetries follows almost at once from this corollary.

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

and by (1.1.2),

$$u \cdot v = (u(I - X), v).$$

Comparing these formulae we get the lemma.

COROLLARY. *W^ρ is the null-space of $I + \epsilon P^{-1}$.*

This follows from the lemma and lemma 1.1.1, cor.

It is clear that a direct decomposition of f gives rise to one of P. We now consider two cases where the converse holds.

Notation: if $\phi(t) = \alpha_0 + \alpha_1 t + \cdots + t^r$ is a monic element of $Z[t]$ such that $\phi(0) = \alpha_0 \neq 0$, we define the monic polynomial

$$\tilde{\phi}(t) = (\alpha_0^{-1})^J t^r \phi^J(t^{-1}) = (\alpha_0^{-1})^J + \cdots + (\alpha_0^{-1}\alpha_1)^J t^{r-1} + t^r.$$

(1) Suppose now that

(1.4.7)
- (a) $\phi_1, \cdots, \phi_k$ are monic elements of $Z[t]$, none of which is divisible by t;
- (b) $\tilde{\phi}_i = \phi_i$ $(i = 1, \cdots, k)$;
- (c) ϕ_i, ϕ_j are relatively prime whenever $i \neq j$.

Let

$$W = W_0 \oplus \cdots \oplus W_k, \quad P = P_0 \oplus \cdots \oplus P_k \tag{1.4.8}$$

be the corresponding Fitting decomposition of P, so that

(i) $\phi_i(P_i)$ is nilpotent $(i = 1, \cdots, k)$;

(ii) $\phi_i(P_j)$ is non-singular $(i = 1, \cdots k; j = 0, \cdots k; i \neq j)$.

LEMMA 1.4.5. *The decomposition* (1.4.8) *of P gives rise to a direct decomposition $f = f_0 \oplus \cdots \oplus f_k$ of f.*

PROOF. By lemma 1.4.2, it is sufficient to prove that $W_j \omega W_i$ whenever $i < j$. Let $x \in W_j$, $y \in W_i$. By (i), there is a power ϕ of ϕ_j such that $x\phi(P) = 0$. By (ii) and (1.4.7) (b), there is a $z \in W_i$ such that $y = z\phi^J(P^{-1})$. Then, by (1.4.5), $(x, y) = (x, z\phi^J(P^{-1})) = (x\phi(P), y) = (0, y) = 0$, so that $W_j \omega W_i$ as required.

We now apply the lemma to the element X, taking (for convenience of exposition) $\phi_1(t) = t + \varepsilon$. Let $u_1, \cdots, u_r$ be a basis of W^ρ and choose $v_1, \cdots, v_r \in (W_0 \oplus W_2 \oplus \cdots \oplus W_k)^\perp$ so that $u_i \cdot v_j = \delta_{ij} (i, j = 1, \cdots, r)$. Then $V' = W + \{v_1, \cdots, v_r\}$ is a finite-dimensional subspace of V and we have corresponding direct decompositions

$$V = V' \oplus V'', \quad F = F' \oplus F'', \quad X = X' \oplus X'',$$

where $V'' = (V')^\perp$, X'' is the identity on V'' and X' the element of $U(F')$ with the same form and space as X. We define

$$\begin{aligned} V_1' &= W_1 + \{v_1, \cdots, v_r\}, \\ V_i' &= W_i \qquad\qquad (2 \leqq i \leqq k), \\ V_0' &= (V_1' + \cdots + V_k')^\perp \cap V' \end{aligned}$$

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

Then it is easy to check that

$$V' = V'_0 \oplus V'_1 \oplus \cdots \oplus V'_k$$

is an F-perpendicular direct decomposition which gives rise to the Fitting decomposition $-\varepsilon X_0'^{-1} \oplus \cdots$ of $-\epsilon X'^{-1}$ corresponding to (1.4.7). Notice that $X'_i = -\varepsilon P_i^{-1}$ when $i \geqq 2$, and that $-\varepsilon P_1^{-1}$ is the restriction of X'_1 to $W_1 = V'_1(I - X'_1)$. These considerations yield the following immeidate corollaries.

COROLLARY 1. *The Fitting decomposition of* $-\varepsilon X^{-1}$ *corresponding to* (1.4.7) *exists and is F-perpendicular (even when the dimension of V is infinite).*

COROLLARY 2. *If* μ'_i, ν_i *denote the multiplicities of* $(t-1)^i$, $(t+\varepsilon)^i$ *as elementary divisors of* X', P *respectively, then*

$$\mu'_i = \nu_{i-1} (i = 2, 3, \cdots), \sum_{i \geqq 1} \mu'_i = \dim V' - \dim W.$$

If the dimension of V *is finite and* μ_i *denotes the multiplicity of* $(t-1)^i$ *as elementary divisor of* X, *then*

$$\mu_i = \nu_{i-1} (i = 2, 3, \cdots), \sum_{i \geqq 1} \mu_i = \dim V - \dim W.$$

COROLLARY 3. *Two elements of* $U_\phi(F)$ *are similar if, and only if, their multipliers are similar.*

(2) The second type of decomposition of P which gives rise to a decomposition of f is more special. It applies when P is ϕ-nul (cf. § 0.3). We remark that in this case $\phi = \tilde{\phi}$, by (1.4.5). Let the elementary divisors of P by $\phi^{e_1}, \cdots, \phi^{e_r} (e_1 > e_2 > \cdots > e_r > 0)$ with respective multiplicities $\nu_1, \cdots \nu_r$. Then there is at least one decomposition

$$W = W_1 \oplus \cdots \oplus W_r, \quad P = P_1 \oplus \cdots \oplus P_r, \tag{1.4.9}$$

where P_i has the single elementary divisor ϕ^{e_i} with multiplicity ν_i $(i = 1, \cdots r)$. Let f_i denote the restriction of f to W_i. Then we have

LEMMA 1.4.6. f_1 *is a direct summand of* f.

PROOF. By lemma 1.4.3, it is sufficient to prove that f_1 is non-degenerate. Suppose f_1 is degenerate. Then W_1 contains a non-zero vector u such that $u \omega W_1$. Since $W_1 \cap {}^\omega W_1$ is invariant under P (by lemma 1.4.2), we may suppose that $u\phi(P) = 0$. From the theory of elementary divisors it follows that $u = v\phi(P)^{e_1 - 1}$, where $v \in W_1$. Now let $w \in W_i$, where $i > 1$. Then $(u, w) = (v, w\phi^J (P^{-1})^{e_1 - 1}) = (v, 0) = 0$, since $\tilde{\phi} = \phi$ and $e_1 > e_i$. Hence u is f-perpendicular to $W_1 + W_2 + \cdots + W_r = W$, contrary to the non-degeneracy of f. This contradiction proves the lemma.

COROLLARY. *At least one decomposition* (1.4.9) *gives rise to a direct decomposition of* f.

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

PROOF. By the lemma, there exist corresponding direct decompositions $f = f_1 \oplus f'$, $P = P_1 \oplus P'$. The corollary now follows easily by induction.

1.5 *Orthogonal groups for $p = 2$.* Let D be a field of characteristic 2 and J the identity. Let $Q(x) = |x|$ be a non-degenerate quadratic form on V with polar form $F(x, y) = x \cdot y$. Thus

$$|\lambda x + \mu y| = \lambda^2|x| + \mu^2|y| + \lambda\mu(x \cdot y) \quad (\lambda, \mu \in D, x, y \in V) \tag{1.5.1}$$

and

$$|x + y| = 0 \text{ (all } y \in V) \text{ implies that } x = 0. \tag{1.5.2}$$

(1.5.1) shows that F is alternate ($x \cdot x = 0$ for all $x \in V$). Q is called *defective* or *non-defective* according as F is degenerate or non-degenerate.

The symbol $\perp$ refers to F-perpendicularity. If W is a subspace of V, the F-perpendicular space $W^\perp$ and radical $W^\rho = W \cap W^\perp$ are defined as before. The *singular radical* W^σ is defined as the set of $x \in W$ such that $|x + y| = |y|$ for all $y \in W$. By (1.5.1), $W^\sigma \subset W^\rho$. The restriction of Q to W is non-degenerate if, and only if, $W^\sigma = \{0\}$, non-defective if, and only if, $W^\rho = \{0\}$.

Let W_1, W_2 be subspaces of V. A linear isomorphism $X : W_1 \to W_2$ is called an isometry of W_1 onto W_2 if

$$|xX| = |x| \quad \text{for all} \quad x \in W_1. \tag{1.5.3}$$

By (1.5.1), this implies that

$$xX \cdot yX = x \cdot y \quad \text{for all} \quad x, y \in W_1. \tag{1.5.4}$$

The isometries of V onto itself form the *orthogonal group* $O(Q)$ of Q. By (1.5.4), $O(Q)$ is a subgroup of the unitary (= symplectic) group $U(F)$ of F.

The space, V_X, and form, F_X, of an element $X = I - T$ of $O = O(Q)$ are defined as before. In place of (1.1.1), the equation of invariance $|xX| = |x|$ yields the stronger relation

$$x \cdot xT = |xT| \quad \text{for all} \quad x \in V. \tag{1.5.5}$$

This implies that

$$(u, u) = |u| \quad \text{for all} \quad u \in V_X. \tag{1.5.6}$$

LEMMA 1.5.1. $V_X \cap V^\perp = \{0\}$.

PROOF. Let $xT \in V_X \cap V^\perp$. By (1.5.5), $|xT| = 0$ and so $xT \in V^\sigma$. Since $V^\sigma = \{0\}$ (by the non-degeneracy of Q), we have $xT = 0$ as required.

With the help of this lemma, it is easy to show that lemmas 1.1.1 to 1.1.3 and corollary are still valid. Theorems 1.1.1 and 1.1.2 are valid if we replace (1.1.3)′ by

$$(u, u)' = |u| \quad \text{for all} \quad u \in W. \tag{1.5.6}'$$

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

Trace-valued vectors are defined as before with reference to F. But since F is alternate, every x is trace-valued and so $V^\tau = V$. Witt's theorem becomes

THEOREM 1.5.1. *Let W_1, W_2 be finite-dimensional subspaces of V such that $W_1 \cap V^\perp = W_2 \cap V^\perp = \{0\}$. Then every isometry of W_1 onto W_2 can be extended to an element of $O_\phi(Q)$.*

From theorem 1.5.1 and lemma 1.5.1, we now deduce

THEOREM 1.5.2. *Two elements of O_ϕ are conjugate in O_ϕ (or O) if, and only if, their forms are equivalent.*

The results of § 1.4 carry over with the obvious minor alterations.

We add one simple result about the singular radical of the space of X.

LEMMA 1.5.2. *Let W be the space, P the mulitplier, of an element X of $O(Q)$. Then $W^\rho \cap W(I + P) \subset W^\sigma$.*

PROOF. Let $v = w(I + P) \in W^\rho$, where $w \in W$; we have to prove that $|v| = 0$. By (1.5.6), $|wP| = (wP, wP) = (w, w) = |w|$. On the other hand, $|v| = |w| + |wP|$ since $v \in W^\rho$. Hence $|v| = 0$ as required.

2. Equivalence

2.1 *Preliminary transformation of problem.* Let W be a left vector space over D of finite dimension m. Let $f_1(x, y) = (x, y)_1$ and $f_2(x, y) = (x, y)_2$ be non-degenerate sesquilinear forms on W. Our problem is to determine the conditions that $f_1 \approx f_2$.

LEMMA 2.1.1. *Equivalent forms have similar multipliers.*

PROOF. Let $f_1 \approx f_2$, so that $(x, y)_2 = (xY, yY)_1$ for some non-singular linear transformation Y. Let P_1, P_2 be the multipliers of f_1, f_2. Then $(y, x)_2^J = (yY, xY)_1^J = (xY, yYP_1)_1 = (x, yYP_1Y^{-1})_2$, so that $P_2 = YP_1Y^{-1}$.

By lemma 2.1.1, we may confine attention to forms with the one fixed multiplier P. We choose one such form $f(x, y) = (x, y)$ as a fixed reference form and write

$$(x, y)_i = (x, yQ_i) \quad (i = 1, 2). \tag{2.1.1}$$

Q_i is called the *representative of f_i with respect to f.*

Let Y be a linear transformation on W. The *f-adjoint* $Y^\dagger$ of Y is defined by $(xY, y) \equiv (x, yY^\dagger)$. We say that Y is *$\dagger$-symmetric* (or f-symmetric) if $Y^\dagger = Y$. Since

$$(y, xYP)^J = (xY, y) = (x, yY^\dagger) = (yY^\dagger, xP)^J = (y, xPY^{\dagger\dagger}),$$

we have

$$Y^{\dagger\dagger} = P^{-1}YP. \tag{2.1.2}$$

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

Also, since

$$(y, x)^J = (x, yP) = (x, yP)^{J^2} = (yP, xP)^J = (y, xPP^\dagger)^J,$$

we have

$$P^\dagger = P^{-1}. \tag{2.1.3}$$

Finally, since

$$(y, x)_i = (y, xQ_i)^J = (xQ_i, yP) = (x, yPQ_i^\dagger) = (x, yPQ_i^\dagger Q_i^{-1})_i,$$

we have

$$P_i = PQ_i^\dagger Q_i^{-1}, \tag{2.1.4}$$

where P_i is the multiplier of f_i.

Let $\mathscr{C} = \mathscr{C}(P)$ denote the commutator ring of P. The next two lemmas follow immediately from (2.1.2)—(2.1.4).

LEMMA 2.1.2. *The mapping $Y \to Y^\dagger$ induces an involutory antiautomorphism of $\mathscr{C}$. Every $\dagger$-symmetric linear transformation lies in $\mathscr{C}$.*

LEMMA 2.1.3. *f_i has multiplier P if, and only if, its representative Q_i with respect to f is $\dagger$-symmetric.*

Linear transformations Y_1, Y_2 on W are called $\dagger$-*congruent* if $Y_2 = YY_1Y^\dagger$ for some non-singular linear transformation Y *belonging to* $\mathscr{C}$. The following theorem reduces the original equivalence problem to a congruence problem in $\mathscr{C}$.

THEOREM 2.1.1. *Let f_1, f_2 be non-degenerate forms with multiplier P and let Q_1, Q_2 be their (non-singular, $\dagger$-symmetric) representatives with respect to the reference form f with multiplier P. Then $f_1 \approx f_2$ if, and only if, Q_1, Q_2 are $\dagger$-congruent.*

PROOF. If Y is a linear transformation on W,

$$(xY, yY)_1 = (xY, yYQ_1) = (x, yYQ_1Y^\dagger) = (x, yYQ_1Y^\dagger Q_2^{-1})_2.$$

Therefore $f_1 \approx f_2$ if, and only if, $Q_2 = YQ_1Y^\dagger$ for some non-singular Y. But this equation implies that $Y \in \mathscr{C} : YQ_1Y^\dagger = (YQ_1Y^\dagger)^\dagger = Y^{\dagger\dagger}Q_1Y^\dagger$, hence $Y^{\dagger\dagger} = Y$, hence, by (2.1.2), $Y \in \mathscr{C}$. This proves the theorem.

2.2 *Approximation Theorem.* Let $\mathscr{N} = \mathscr{N}(P)$ denote the radical of $\mathscr{C} = \mathscr{C}(P)$. The main theorem of this section is proved under the assumption that $\mathscr{C}$ satisfies the following *trace condition*:

(2.2.1) *If $\mathscr{I} \subset \mathscr{N}$ is an ideal of $\mathscr{C}$ and N an element of $\mathscr{I}$ such that $N^\dagger = \varepsilon N (\varepsilon = \pm 1)$, then $N = M + \varepsilon M^\dagger$ for some $M \in \mathscr{I}$.*

Apart from this, only the following properties of $\mathscr{C}$ are used:

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

(i) $\mathscr{C}$ is a ring with unit element I;
(ii) $\mathscr{C}$ has an involutory anti-automorphism $\dagger$;
(iii) $\mathscr{N}$ is nilpotent (Jacobson [6], Ch. 4, § 7).

The following simple lemma shows that the trace condition "usually" hold in $\mathscr{C}(P)$.

LEMMA 2.2.1. *If $\mathscr{C}$ contains a central element Y such that $Y + Y^\dagger = I$, $\mathscr{C}$ satisfies the trace condition.*

PROOF. If $N^\dagger = \varepsilon N$, $N = (YN) + \varepsilon(YN)^\dagger$.

COROLLARY. *If any one of the following conditions holds, $\mathscr{C}(P)$ satisfies the trace condition*:
(a) $p \neq 2$;
(b) *the restriction of J to the centre Z of D is not the identity;*
(c) *$I + P$ is non-singular.*

PROOF. (a) $Y = \frac{1}{2}I$. (b) We may assume that $p = 2$. By hypothesis, there is an $\alpha \in Z$ such that $\alpha + \alpha^J = \beta \neq 0$. Let $Y = \alpha\beta^{-1}I$. (c) $Y = P(I + P)^{-1}$.

(2.2.2) NOTATION. The following conventions are observed both here and in later sections. Suppose R is a ring with radical S. Then $\bar{R}$ denotes the factor ring R/S and $\bar{Y} = Y + S$ the canonical image of an element Y of R in $\bar{R}$. If $\dagger$ is an anti-automorphism of R, the same symbol $\dagger$ will be used to denote the induced anti-automorphism of $\bar{R}$: $(\bar{Y})^\dagger = (\overline{Y^\dagger})$. An element of a ring with I is called non-singular if it has a two-sided inverse.

THEOREM 2.2.1. (*Approximation theorem*) *Suppose that $\mathscr{C}$ satisfies the trace condition* (2.2.1). *Then*
(a) *every non-singular, $\dagger$-symmetric (or $\dagger$-skewsymmetric) element of $\overline{\mathscr{C}} = \mathscr{C}/\mathscr{N}$ is the canonical image of a non-singular, $\dagger$-symmetric (or $\dagger$-skewsymmetric) element of $\mathscr{C}$;*
(b) *two non-singular, $\dagger$-symmetric (or $\dagger$-skewsymmetric) elements of $\mathscr{C}$ are $\dagger$-congruent if, and only if, their canonical images in $\overline{\mathscr{C}}$ are $\dagger$-congruent.*

PROOF. Let $\bar{Q}$ be a non-singular element of $\overline{\mathscr{C}}$ such that $\bar{Q}^\dagger = \varepsilon\bar{Q}$, i.e. $Q^\dagger - \varepsilon Q = N \in \mathscr{N}$ ($\varepsilon = \pm 1$). Then $N^\dagger = -\varepsilon N$ and so, by the trace condition, $N = M^\dagger - \varepsilon M$, where $M \in \mathscr{N}$. Let $Q_1 = Q - M$. Then $\bar{Q} = \bar{Q}_1$ and $Q_1^\dagger = \varepsilon Q_1$. Also Q_1 is non-singular. For, since $\bar{Q}$ is non-singular, there exists an $R \in \mathscr{C}$ such that $Q_1 R = I - N_1$, where $N_1 \in \mathscr{N}$. Then $R(I + N_1 + N_1^2 + \cdots)$ is a right inverse of Q_1. Similarly Q_1 has a left inverse.

Suppose now that S_1, S_2 are non-singular elements of $\mathscr{C}$ such that $S_i^\dagger = \varepsilon S_i$ ($\varepsilon = \pm 1$; $i = 1, 2$). Obviously $\bar{S}_1$, $\bar{S}_2$ are $\dagger$-congruent if S_1, S_2 are. Conversely, let $\bar{S}_1$, $\bar{S}_2$ be $\dagger$-congruent, so that $S_1 - YS_2Y^\dagger = N \in \mathscr{N}$, where $\bar{Y}$ is

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

non-singular. By the argument of the first paragraph, Y is non-singular. Since $\mathscr{N}$ is nilpotent, it is sufficient to prove that if $S_1 - Y_1 S_2 Y_1^\dagger = N_1 \in \mathscr{N}^i$ $(i \geqq 1)$, where Y_1 is non-singular, then $S_1 - Y_2 S_2 Y_2^\dagger = N_2 \in \mathscr{N}^{i+1}$, where Y_2 is non-singular. Now $N_1^\dagger = \varepsilon N_1$ and so, by the trace condition, $N_1 = M_1 + \varepsilon M_1^\dagger$, where $M_1 \in \mathscr{N}^i$. It is now easily verified that $Y_2 = (I + M_1(Y_1 S_2 Y_1^\dagger)^{-1})Y_1$ meets the requirements. This proves the theorem.

The following reduction theorem holds without any restriction on the commutator rings involved. It is an immediate consequence of the uniqueness of the Fitting decomposition.

THEOREM 2.2.2. *Let f, g be non-degenerate sesquilinear forms on W and*

$$f = f_0 \oplus \cdots \oplus f_k, \quad g = g_0 \oplus \cdots \oplus g_k$$

their Fitting decompositions corresponding to (1.4.7). *Then $f \approx g$ if and only if, $f_i \approx g_i$ for $i = 0, 1, \cdots k$.*

The theorem shows, in particular, that the general case of equivalence reduces to the two subcases: (a) $P + I$ is non-singular and (b) $P + I$ is nilpotent. By lemma 2.2.1, cor., the approximation theorem is valid in case (a), even when $p = 2$. Thus the essential difficulties are concentrated into case (b).

Definition. Let R be a ring with 1 having an involutory anti-automorphism α. The group

$$N(\alpha, R) = \{x \in R | x x^\alpha = 1\}$$

is called the *norm group* of R with respect to α.

In the theory of the centralizers of the elements of a classical group, the norm groups $N = N(\dagger, \mathscr{C})$ play an important part. Let $N_i = N_i(\dagger, \mathscr{C})$ $(i = 0, 1, \cdots)$ denote the normal subgroup of N formed by the $Y \in N$ such that $Y \equiv I \pmod{\mathscr{N}^{2^i}}$. Let $S_i^\pm = S_i^\pm(\dagger, \mathscr{C})$ $(i = 0, 1, \cdots)$ denote the additive group formed by the $Y \in \mathscr{N}^{2^i}$ such that $Y^\dagger = \pm Y$.

THEOREM 2.2.3. *If $\mathscr{C}$ satisfies the trace condition,*

$$N/N_0 \cong N(\dagger, \overline{\mathscr{C}}),$$
$$N_i/N_{i+1} \cong S_i^-/S_{i+1}^- \quad (i = 0, 1, \cdots).$$

PROOF. Consider the group homomorphism $\eta(Y) = \bar{Y}$ of N into $N(\dagger, \overline{\mathscr{C}})$. Let $\bar{Y} \in N(\dagger, \overline{\mathscr{C}})$, where $Y \in \mathscr{C}$. Then $YY^\dagger = I \pmod{\mathscr{N}}$ and the proof of the approximation theorem shows that $\bar{Y} = \bar{Y}_1$, where $Y_1 \in N$. Thus $\eta(N) = N(\dagger, \overline{\mathscr{C}})$. Since the kernel of η is clearly N_0, we have $N/N_0 \cong N(\dagger, \overline{\mathscr{C}})$.

Consider next the mapping $\zeta(I - M) = M + S_{i+1}$ of N_i into S_i/S_{i+1}, where S_i denotes the additive group of $\mathscr{N}^{2^i}$. If $I - M$ and $I - M' \in N_i$, we have

$$(I - M)(I - M') = I - M - M' (\mathrm{mod}\, S_{i+1}),$$

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

whence ζ is a homomorphism with kernel N_{i+1}. Also

$$I = (I - M)(I - M^\dagger) \equiv I - M - M^\dagger (\text{mod } S_{i+1}),$$

so that $M + M^\dagger \in S_{i+1}$. By the trace condition, $M + M^\dagger = M_1 + M_1^\dagger$, where $M_1 \in S_{i+1}$. Thus

$$\zeta(I - M) = (M - M_1) + S_{i+1} \in (S_i^- + S_{i+1})/S_{i+1}$$

and so $\zeta(N_i) \subset (S_i^- + S_{i+1})/S_{i+1}$.

Conversely, let $R \in S_i^-$. We prove that $R + S_{i+1} \in \zeta(N_i)$ by showing that if $R + S_{i+1} = R_1 + S_{i+1}$, where $R_1 + R_1^\dagger - R_1 R_1^\dagger \in S_{i+j}$ $(j \geqq 1)$, then $R + S_{i+1} = R_2 + S_{i+1}$, where $R_2 + R_2^\dagger - R_2 R_2^\dagger \in S_{i+j+1}$. In fact, by the trace condition, $R_1 + R_1^\dagger - R_1 R_1^\dagger = R_3 + R_3^\dagger$, where $R_3 \in S_{i+j}$. Then $R_2 = R_1 - R_3(I - R_1^\dagger)^{-1}$ meets the requirements. This proves that $\zeta(N_i) = (S_i^- + S_{i+1})/S_{i+1}$.

We now have

$$N_i/N_{i+1} \cong (S_i^- + S_{i+1})/S_{i+1} \cong S_i^-/S_{i+1}^-,$$

which proves the theorem.

2.3. *Multipliers.* In the present section, we determine the conditions that a given linear transformation Π be the multiplier of some form.

Let $\boldsymbol{\Pi}$ be the matrix of Π. Then (1.4.5)′ shows that, if Π is a multiplier, $\boldsymbol{\Pi}$ is similar to $\boldsymbol{\Pi}^{*-1}$. We indicate this by the symbolical notation *

$$\Pi \sim \Pi^{*-1}. \tag{2.3.1}$$

Let $\Pi = \Pi_1 \oplus \cdots$ be a splitting of Π into indecomposable parts and let R be any indecomposable linear transformation. If k of the summands Π_i are similar to R, we say that R has multiplicity k in Π. Then (2.3.1) holds if, and only if, R and R^{*-1} have the same multiplicity in Π for every R.

LEMMA 2.3.1. *Every linear transformation $S \oplus S^{*-1}$ is a multiplier.*

This follows at once from the matrix identity

$$\begin{pmatrix} \mathbf{0} & \boldsymbol{I} \\ \boldsymbol{S}^{*-1} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{0} & \boldsymbol{I} \\ \boldsymbol{S}^{*-1} & \mathbf{0} \end{pmatrix}^{*-1} = \begin{pmatrix} \boldsymbol{S} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{S}^{*-1} \end{pmatrix}. \tag{2.3.2}$$

LEMMA 2.3.2. *If Π satisfies* (2.3.1) *and $\Pi^2 - I$ is non-singular, Π is a multiplier.*

PROOF. By the previous lemma it is sufficient to prove this when Π is indecomposable. By (2.3.1), there exists a non-degenerate sesquilinear form $f(x, y)$ such that $\Pi^\dagger = \Pi^{-1}$, where $\dagger$ denotes the f-adjoint. Let P be the multiplier of f and write $Y = P^{-1}\Pi$. Since the multiplier of $g(x, y) = f(x, yQ)$

* $\sim$ indicates similarity of matrices or linear transformations.

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

is $PQ^{\dagger}Q^{-1}$, it is sufficient to prove that $Y = Q^{\dagger}Q^{-1}$ for some Q.

Let $\mathscr{K}$ denote the commutator ring of Π. Since $\Pi^{\dagger\dagger} = \Pi$, we have $Y \in \mathscr{K}$ by (2.1.2). Since Π is indecomposable, $\mathscr{K}$ is a completely primary ring and so every element of $\mathscr{K}$ not in the radical is non-singular. Since $\Pi - I$ is non-singular and $\Pi - I = \Pi(Y + I) - (\Pi Y + I)$, at least one of $Y + I$ and $\Pi Y + I$ is non-singular. In the first case, take $Q = (Y + I)^{-1}$ and in the second, $Q = (\Pi + I)(\Pi Y + I)^{-1}$ (this choice being possible because $\Pi + I$ is non-singular). This proves the lemma.

LEMMA 2.3.3. *Let Π be an indecomposable ϕ-nul * linear transformation on W, where $\phi = \tilde{\phi}$. Suppose that*

(i) *if J is the identity and $\phi(t) = t - 1$, then m is odd;*

(ii) *if J is the identity, $p \neq 2$ and $\phi(t) = t + 1$, then m is even.*

*Then Π is a multiplier **.*

PROOF. Let θ denote the involutory automorphism $g(\Pi) \to g^J(\Pi^{-1})$ of $\mathscr{Z} = Z[\Pi]$. Suppose that we have determined a mapping $\chi : \mathscr{Z} \to D$ such that

(a) χ is Z-linear;

(b) $\chi(Y^{\theta}) = \chi(Y\Pi)^J$ for all $Y \in \mathscr{Z}$;

(c) χ does not vanish identically on the minimal ideal $\mathscr{M} = \mathscr{Z}\phi(\Pi)^{e-1}$ of $\mathscr{Z}$.

Let u be a vector such that $u, u\Pi, \cdots, u\Pi^{m-1}$ are a basis of W. Then we prove that the equations

$$(2.3.3) \qquad (u\Pi^i, u\Pi^j) = \chi(\Pi^{i-j+1}) \qquad (i, j = 0, 1, \cdots m - 1)$$

define a non-degenerate sesquilinear form $f(x, y) = (x, y)$ on W, whose multiplier is Π.

We first remark that

$$(2.3.4) \qquad (x\Pi, y) = (x, y\Pi^{-1}) \qquad (x, y \in W),$$

$$(2.3.5) \qquad (uR, uS) = \chi(RS^{\theta}\Pi) \qquad (R, S \in \mathscr{Z}).$$

By (2.3.4), the "left radical" ${}^{\omega}W$ is invariant under Π. Hence, if ${}^{\omega}W \neq \{0\}$, $W\mathscr{M} \subset {}^{\omega}W$. By (2.3.5), this implies that χ vanishes on $\mathscr{M}$, contrary to (c). Thus ${}^{\omega}W = \{0\}$ and so f is nondegenerate. Finally, since

$$\begin{aligned}(uR, uS\Pi) = \chi(RS^{\theta}\Pi^{\theta}\Pi) = \chi(RS^{\theta}) = \chi(SR^{\theta}\Pi)^J \\ = (uS, uR)^J \qquad (R, S \in \mathscr{Z}),\end{aligned}$$

Π is the multiplier of f.

* Cf. § 0.3.

** Zassenhaus ([20], theorem 1(b)) gives an interesting construction which essentially reduces the present lemma to the case where $\phi(\Pi) = 0$.

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

It remains to construct χ. Let $\phi(t) = t^d + \cdots + \beta$, $\phi(t)^e = t^m + \cdots + \alpha$, so that $m = de$, $\alpha = \beta^e$. We suppose first that m is odd: $m = 2\mu + 1$. Then each $Y \in \mathscr{Z}$ has a unique representation

$$Y = \sum_{-\mu}^{\mu} \eta_i \Pi^i \qquad (\eta_i \in Z),$$

and we prove that $\chi(Y) = \lambda\eta_{-\mu}$ satisfies (a)—(c) for suitable $\lambda \in D$.

(a) obviously holds for any λ. Since $\Pi^{-\mu}\phi(\Pi)^{e-1} = \beta^{e-1}\Pi^{-\mu} + \cdots$, (c) holds provided that $\lambda \neq 0$. Finally, since

$$Y^\theta = \sum_{-\mu}^{\mu} \eta_{-i}^J \Pi^i, \quad Y\Pi = \sum_{1-\mu}^{\mu} \eta_{i-1}\Pi^i - \eta_\mu \Pi^{-\mu}(\alpha + \cdots),$$

(b) holds if, and only if,

$$\lambda^J + \lambda\alpha = 0. \tag{2.3.6}$$

Since $\phi = \tilde{\phi}$, $\alpha\alpha^J = 1$. Hence $\lambda = \tau\alpha^J - \tau^J$ satisfies (2.3.6) for any $\tau \in D$. Thus we get a non-zero solution λ unless J is the identity, $p \neq 2$ and $\alpha = 1$. But these conditions are excluded by hypothesis (ii), for $\alpha = 1$, m odd and $\phi = \tilde{\phi}$ imply that $(t+1) | \phi(t)^e$ and thus that $t + 1 = \phi(t)$.

Suppose now that m is even: $m = 2\mu$. Each $Y \in \mathscr{Z}$ has a unique representation

$$Y = \sum_{1-\mu}^{\mu} \eta_i \Pi^i \qquad (\eta_i \in Z),$$

and we prove that $\chi(Y) = \lambda\eta_{1-\mu} + \lambda^J\eta_\mu$ satisfies (a)—(c) for suitable $\lambda \in D$. (a) holds for any λ. An easy calculation shows that (b) also holds for all λ. If $d > 1$, the equation

$$\Pi^{1-\mu}\phi(\Pi)^{e-1} = \beta^{e-1}\Pi^{1-\mu} + \cdots + \Pi^{\mu-d+1}$$

shows that (c) holds for any $\lambda \neq 0$ If $d = 1$, so that $\phi(t) = t + \beta$, $\chi(\Pi^{1-\mu}\phi(\Pi)^{e-1}) = \lambda\beta^{m-1} + \lambda^J$ is non-zero for suitable λ, unless J is the identity and $\beta^{m-1} = -1$ But these conditions are excluded by hypothesis (i), for since $\beta\beta^J = \beta^2 = 1$, $\beta^{m-1} = -1$ implies that $t + \beta = t - 1$. This completes the proof.

THEOREM 2.3.1. *Let Π be a linear transformation on the finite-dimensional space W. Then Π is the multiplier of a non-degenerate sesquilinear form on W if, and only if,*

(i) $\Pi \sim \Pi^{*-1}$;

(ii) *if J is the identity, the multiplicity of $(t-1)^{2k}$ as elementary divisor of Π is even $(k = 1, 2, \cdots)$;*

(iii) *if J is the identity and $p \neq 2$, the multiplicity of $(t+1)^{2k-1}$ as elementary divisor of Π is even $(k = 1, 2, \cdots)$.*

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

PROOF. The sufficiency of the conditions is easily proved by splitting Π into indecomposable parts and applying the preceding lemmas. Suppose, conversely, that Π is the multiplier of the non-degenerate form $f(x, y) = (x, y)$. We have proved that (i) is necessary. In proving (ii), (iii) necessary, we may, by the Fitting decomposition and lemma 1.4.6, cor., assume that Π is of commutative type with a single elementary divisor $(t-1)^{2k}$ or $(t+1)^{2k-1}$, say of multiplicity μ. We may also assume that $p \neq 2$ in the second case and that J is the identity in both cases. Then the left hand side of the equation $(x, y) - (y, x) = (y, x(\Pi - I))$ is an alternate form, whence the rank of $\Pi - I$, viz. $\mu(2k-1)$, is even. Thus the multiplicity μ is even as required. This proves the theorem.

Supplementary Remarks. The following discussion of the construction in lemma 2.3.3. is required only in § 2.6. Since we shall be comparing the constructions corresponding to different powers ϕ^e of the one fixed polynomial ϕ, we shall write $\Pi_e, \mathscr{Z}_e, \cdots$ instead of $\Pi, \mathscr{Z}, \cdots$. We assume that J is the identity, $d > 1$, $p \neq 2$. Since $\phi = \tilde{\phi}$, d is even. A fortiori, the dimension de of the Z-space $\mathscr{Z}_e$ is even. Since $d > 1$, we may take $\lambda = 1$ in the definition $\chi_e(Y) = \lambda\eta_{1-\mu} + \lambda^J\eta_\mu$ of χ_e. Let us now consider the quadratic form $\rho_e(Y) = \chi_e(YY^\theta)$ on $\mathscr{Z}_e$. The matrix of ρ_e is $\frac{1}{2}(\Omega + \Omega^T) = \frac{1}{2}(I + \Pi^{-1})\Omega$, where Ω is the matrix of f. Therefore ρ_e is non-degenerate. *If e is even, ρ_e is a form of maximum Witt index $\frac{1}{2}de$.* In fact, $\mathscr{Z}_e\phi(\Pi_e)^{\frac{1}{2}e}$ is a totally isotropic subspace of dimension $\frac{1}{2}de$. *If e is odd, ρ_e is a form Witt index $\frac{1}{2}de$ or $\frac{1}{2}de - 1$; moreover, all the forms $\rho_1, \rho_3, \cdots$ have the same Witt type.** To see this, we remark that $\mathscr{U}_e = \mathscr{Z}_e\phi(\Pi_e)^{\frac{1}{2}(e+1)}$ is totally isotropic and that $\mathscr{V}_e = \mathscr{Z}_e\phi(\Pi_e)^{\frac{1}{2}(e-1)}$ is its perpendicular space. Hence if $\mathscr{W}_e$ is any complement of $\mathscr{U}_e$ in $\mathscr{V}_e$, the restriction σ_e of ρ_e to $\mathscr{W}_e$ has the same Witt type as ρ_e. We take $\mathscr{W}_e$ as the subspace with basis elements $\Pi_e^{\kappa-\frac{1}{2}d(e-1)}\phi(\Pi_e)^{\frac{1}{2}(e-1)}$ $(\kappa = 0, 1, \cdots, d-1)$. The first $\frac{1}{2}d - 1$ elements span a totally isotropic subspace, so that the Witt index of σ_e is $\frac{1}{2}d$ or $\frac{1}{2}d - 1$. It is easy to see that the matrix of σ_e with respect to the above basis is independent of e, whence $\rho_1, \rho_3, \cdots$ are all of the same Witt type. This proves our result.

2.4. *Equivalence invariants.* We consider †-congruence in $\bar{\mathscr{C}} = \mathscr{C}/\mathscr{N}$. This leads to explicit solutions of the equivalence and conjugacy problems when the trace condition holds in $\mathscr{C}$. We make constant use of the notational conventions in (2.2.2).

Let P be a multiplier on W. We choose a direct decomposition of P into multipliers, none of which is itself a proper direct sum of multipliers; say,

$$P = P_1 \oplus P_1' \oplus \cdots \oplus P_s \oplus \cdots \oplus P_s'^{\cdots'} \tag{2.4.1}'$$

where terms with like subscripts are similar. Then we choose a reference form

* Cf. Bourbaki [1], § 8.

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

(2.4.2)′ $$f = f_1 \oplus f_1' \oplus \cdots \oplus f_s \oplus \cdots \oplus f_s'^{\cdots'},$$

where terms with like subscripts are equivalent and where, if P_i is decomposable, f_i is constructed by the method of lemma 2.3.1. A reference form of this kind will be called *standard. When comparing forms with similar multipliers, we shall always assume that the corresponding reference forms are equivalent.*

Suppose that P_i is indecomposable for $i = 1, \cdots, r$, and decomposable for $i = r+1, \cdots s$. In a suitable coordinate system, the matrices of f, P have the forms

(2.4.1) $$\Omega = \operatorname{diag}(\Omega^1, \cdots, \Omega^s), \quad P = \operatorname{diag}(P^1, \cdots, P^s),$$

where, if $1 \leqq i \leqq r$,

(2.4.2) $$\Omega^i = \operatorname{diag}(\underbrace{\Omega_i, \cdots, \Omega_i}_{m_i}), \quad P^i = \operatorname{diag}(\underbrace{P_i, \cdots, P_i}_{m_i}),$$

$$P_i = \Omega_i \Omega_i^{*-1}, \ P_i \text{ indecomposable,}$$

and where, if $r+1 \leqq i \leqq s$,

(2.4.3) $$\Omega^i = \begin{pmatrix} 0 & I \\ (\Pi^i)^{*-1} & 0 \end{pmatrix}, \quad P^i = \begin{pmatrix} \Pi^i & 0 \\ 0 & (\Pi^i)^{*-1} \end{pmatrix}$$
$$\Pi^i = \operatorname{diag}(\underbrace{\Pi_i, \cdots, \Pi_i}_{m_i}), \quad \Pi_i \text{ indecomposable.}$$

Let $Q \in \mathscr{C}$. The matrix of Q has the form

(2.4.4) $$Q = (Q^{ij})_{i,j=1\cdots s},$$

where $P^i Q^{ij} = Q^{ij} P^j$. The matrix $Q^\dagger$ of $Q^\dagger$ is given by

$$(Q^\dagger)^{ij} = (\Omega^j)^* (Q^{ji})^* (\Omega^i)^{*-1}.$$

In particular, the diagonal block Q^{ii} belongs to the commutator ring $\mathscr{C}^i = \mathscr{C}(P^i)$ and $(Q^\dagger)^{ii}$ is the adjoint $(Q^{ii})^\dagger$ of Q^{ii} with respect to Ω^i. But the mapping

$$\bar{Q} \to \overline{Q^{11}} \oplus \cdots \oplus \overline{Q^{ss}}$$

is an isomorphism of $\bar{\mathscr{C}}$ onto the direct sum $\bar{\mathscr{C}}^1 \oplus \cdots \oplus \bar{\mathscr{C}}^s$ (Jacobson [6], Ch. 4, § 8). It follows that $\bar{Q}$ is non-singular and †-symmetric if, and only if, $\overline{Q^{ii}}$ is non-singular and †-symmetric $(i = 1, \cdots s)$ and that $\bar{Q}$, $\bar{R}$ $(R \in \mathscr{C})$ are †-congruent if, and only if, $\overline{Q^{ii}}$, $\overline{R^{ii}}$ are †-congruent $(i = 1, \cdots s)$.

Let $1 \leqq i \leqq r$. Then

(2.4.5) $$Q^{ii} = (Q^{ii}_{\lambda\mu})_{\lambda,\mu=1,\cdots m_i},$$

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

where $Q_{\lambda\mu}^{ii}$ belongs to the commutator ring $\mathscr{C}_i = \mathscr{C}(P_i)$. Let J_i denote the Ω_i-adjoint mapping in $\mathscr{C}_i$:

$$Y^{J_i} = \Omega_i^* Y^* \Omega_i^{*-1} \qquad (Y \in \mathscr{C}_i)$$

Then

$$(Q_{\lambda\mu}^{ii})^\dagger = (Q_{\mu\lambda}^{ii\,J_i}),$$

i.e. $\dagger$ is the conjugate transpose operation with respect to J_i. The mapping

$$\overline{Q^{ii}} \to (\overline{Q_{\lambda\mu}^{ii}})$$

is an isomorphism of $\overline{\mathscr{C}^i}$ onto the total matrix algebra $M_{m_i}(\overline{\mathscr{C}_i})$ over $\overline{\mathscr{C}_i}$ and, since P_i is indecomposable, $\overline{\mathscr{C}_i}$ is a division ring. Suppose now that $\bar{Q}$, $\bar{R}$ are non-singular and $\dagger$-symmetric. Then the matrices $(\overline{Q_{\lambda\mu}^{ii}})$, $(\overline{R_{\lambda\mu}^{ii}})$ are non-singular and J_i-Hermitian, and $\overline{Q^{ii}}$, $\overline{R^{ii}}$ are $\dagger$-congruent if, and only if, they are J_i-congruent. The J_i-Hermitian form with matrix $(\overline{Q_{\lambda\mu}^{ii}})$ is called the *i-th Hermitian invariant* of the $\dagger$-symmetric element $\bar{Q}$.

LEMMA 2.4.1. *Let $\bar{Q}$, $\bar{R}$ be non-singular, $\dagger$-symmetric elements of $\overline{\mathscr{C}}$. Except when $p = 2$, J is the identity and $I + P$ is singular, $\bar{Q}$, $\bar{R}$ are $\dagger$-congruent if (and only if) their corresponding Hermitian invariants are equivalent.*

PROOF. We have to prove that under the conditions stated $\overline{Q^{ii}}$, $\overline{R^{ii}}$ are $\dagger$-congruent for $i = r + 1, \cdots s$.

First Case: $\Pi_i \nsim \Pi_i^{*-1}$. The matrix Q^{ii} has the form

$$Q^{ii} = \begin{pmatrix} Y_1 & * \\ * & Y_2 \end{pmatrix},$$

where Y_1, Y_2 belong to the anti-isomorphic commutator rings $\Gamma^i = \mathscr{C}(\Pi^i)$, $\Gamma^{i*} = \mathscr{C}(\Pi^{i*})$. Since Π^i, $(\Pi^{i*})^{-1}$ have no common indecomposable part, the mapping $\overline{Q^{ii}} \to \overline{Y_1} \oplus \overline{Y_2}$ is an isomorphism of $\overline{\mathscr{C}^i}$ onto $\overline{\Gamma^i} \oplus \overline{\Gamma^{i*}}$. Since $\overline{Q^{ii}}$ is non-singular and $\dagger$-symmetric, and since

$$(Q^\dagger)^{ii} = \begin{pmatrix} Y_2^* & * \\ * & Y_1^* \end{pmatrix},$$

we have $\overline{Y_2} = \overline{Y_1^*}$ and $\overline{Y_1}$ is non-singular. Similarly, $\overline{R^{ii}} \to \overline{Z_1} \oplus \overline{Z_1^*}$, where $\overline{Z_1}$ is non-singular. Then

$$\overline{S^{ii}}\,\overline{Q^{ii}}(\overline{S^{ii}})^\dagger = \overline{R^{ii}}$$

where

$$S^{ii} = \begin{pmatrix} Z_1 Y_1^{-1} & 0 \\ 0 & I \end{pmatrix},$$

so that $\overline{Q^{ii}}$, $\overline{R^{ii}}$ are $\dagger$-congruent as required.

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

Second Case: $\Pi_i \sim \Pi_i^{*-1}$. By theorem 2.3.1, J is the identity and Π_i is of commutative type with a single elementary divisor $(t-1)^{2k}$ or $(t+1)^{2k-1}$. It is convenient to modify the previous notation by taking

$$P^i = \operatorname{diag}(\underbrace{\Pi_i, \cdots \Pi_i}_{2m_i}).$$

By hypothesis, $p \neq 2$. Hence, by theorem 2.3.1, $-\Pi_i$ is a multiplier, say of the form g_i with matrix Ψ_i. Then $-P^i$ is the multiplier of $g^i = g_i \oplus \cdots \oplus g_i$. Let $\ddagger$ denote the adjoint with respect to g^i. It is easily verified that the elements of $\overline{\mathscr{C}^i}$ corresponding to $\overline{Q^{ii}}$, $\overline{R^{ii}}$ are $\ddagger$-skewsymmetric and that they are $\dagger$-congruent if, and only if, they are $\ddagger$-congruent. On the other hand, these elements may be identified by our previous method with non-singular, $2m_i$-rowed, J_i'-skew-Hermitian matrices over the division ring $\overline{\mathscr{C}_i}$, where $\mathscr{C}_i = \mathscr{C}(\Pi_i)$ and J_i' denotes the g_i-adjoint. Now $\overline{\mathscr{C}_i} \cong D$ and J_i' is the identity, since Π_i is $(t \pm 1)$-null and J is the identity. Thus the representing matrices are essentially skew-symmetric matrices over the field D. Since any two non-singular, skew-symmetric matrices are congruent, our result follows. This proves the lemma.

We are now in a position to prove the main theorems about equivalence and conjugacy. Let g be a non-degenerate sesquilinear form on W with multiplier P and let f be a standard reference form with multiplier P. Let Q be the representative of g with respect to f. Then the Hermitian invariants of $\bar{Q}$ are called the *Hermitian invariants of g with respect to f*. If X is an element of a unitary or orthogonal group, the Hermitian invariants of X are defined to be those of its form F_X.

THEOREM 2.4.1. (*Equivalence theorem*) *Let D be a division ring with involutory anti-automorphism J. Let f_1, f_2 be non-degenerate J-sesquilinear forms on the finite-dimensional space W over D. Let P_1, P_2 be their multipliers. If the characteristic of D is 2 and J leaves invariant every element of the centre of D, let $P_1 + I$ be non-singular. Then f_1, f_2 are equivalent if, and only if,*

(a) *P_1, P_2 are similar;*

(b) *the corresponding Hermitian invariants of f_1, f_2 (with respect to equivalent standard forms) are equivalent.*

This follows from the theorem 2.2.1 and lemmas 2.2.1 (cor.), 2.4.1.

COROLLARY. *Under the conditions of the theorem, f_1 is directly decomposable if (and only if) P_1 has a proper direct decomposition $P_1 = P_1' \oplus P_1''$, where P_1', P_1'' are multipliers.*

PROOF. Write f', P instead of f_1, P_1 and suppose that P has a proper decomposition (2.4.1)'. Choose a standard form f as in (2.4.2)'. Let f' be represented with respect to f by the matrix Q in (2.4.4). We choose a form g

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

with multiplier P as follows. If $s > 1$, the matrix $\boldsymbol{R}$ representing g with respect to f is taken as diag $(\boldsymbol{Q}^{11}, \cdots, \boldsymbol{Q}^{ss})$. If $s = 1$, and $r = 0$, g is taken as f. If $s = r = 1$, $\boldsymbol{R}$ is taken as any diagonal matrix J_1-congruent to the J_1-Hermitian matrix $\boldsymbol{Q}^{11}$. Then g is decomposable and has the same Hermitian invariants as f. Hence f is decomposable.

THEOREM 2.4.2. (*Conjugacy theorem*). *Let D be a division ring with involutory anti-automorphism J, V a vector space over D. Let G be either the unitary group of a non-degenerate J-Hermitian or J-skew Hermitian form on V or (when D is a field and J the identity) the orthogonal group of a non-degenerate quadratic form on V. Let $X_1, X_2 \in G_\phi$, where G_ϕ is the subgroup formed by the finite-dimensional elements of G. If the characteristic of D is 2 and J leaves invariant every element of the centre of D, let $W_1 \cap W_1^\perp = W_2 \cap W_2^\perp = \{0\}$, where W_1, W_2 are the spaces of X_1, X_2. Then X_1, T_2 are conjugate in G_ϕ (or G) if, and only if,*

(a) *X_1, X_2 are similar;*

(b) *the corresponding Hermitian invariants of X_1, X_2 (with respect to equivalent standard forms) are equivalent.*

PROOF. The theorem follows from theorems 1.3.1. and 2.4.1. We have only to check that the hypotheses of these theorems hold, viz. that

(i) if $p = 2$ and the restriction of J to Z is the identity, then $P_i + I$ is non-singular;

(ii) $W_i \cap (V^\tau)^\perp = \{0\}$.

(i) follows from lemma 1.4.4 (cor.) and our hypothesis that $W_i^\rho = \{0\}$ in the case under consideration. If either $p \neq 2$ or the restriction of J to Z is not the identity, then (ii) is obvious because $V^\tau = V$. In the contrary case, (ii) follows from the hypothesis that $W_i^\rho = \{0\}$; for, since $W_i \subset V^\tau$, we have $W_i^\rho = W_i \cap W_i^\perp \supset W_i \cap (V^\tau)^\perp$. This completes the proof.

Supplementary Remarks. With the notation of the beginning of this section, let g be a form with multiplier P which is represented by Q with respect to f. Let $\ddagger$ denote the adjoint with respect to g. It is easy to see that the norm group $N(\ddagger, \mathscr{C})$ consists of the $Y \in \mathscr{C}$ such that $YQY^\dagger = Q$. It follows that $N(\ddagger, \overline{\mathscr{C}})$ is isomorphic to the direct product $H_1 \times H_2 \times \cdots \times H_s$, where H_i is the group formed by the $\overline{\boldsymbol{Y}}^i \in \overline{\mathscr{C}}^i$ such that $\overline{\boldsymbol{Y}}^i \overline{\boldsymbol{Q}}^{ii} \overline{\boldsymbol{Y}}^{i\dagger} = \overline{\boldsymbol{Q}}^{ii}$. Using the same methods as before, we see that

(i) if $i \leqq r$, H_i is isomorphic to the unitary group of the i-th Hermitian invariant of g;

(ii) if $i > r$ and $\boldsymbol{\Pi}_i \sim \boldsymbol{\Pi}_i^{*-1}$, H_i is isomorphic to the full linear group $GL(m_i, \overline{\Gamma}_i)$, where $\overline{\Gamma}_i$ is the division ring $\mathscr{C}(\overline{\boldsymbol{\Pi}}_i)$;

(iii) if $i > r$ and $\boldsymbol{\Pi}_i = \boldsymbol{\Pi}_i^{*-1}$, and if we assume that $p \neq 2$, then H_i is isomorphic to the symplectic group $Sp(2m_i, D)$.

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

In order to determine these groups in practice, we need to know the division rings and anti-automorphisms involved. The following lemma gives some information or these points.

LEMMA 2.4.2. *If P is indecomposable and ϕ-nul, $\bar{\mathscr{C}} \cong \Delta$, where Δ is the division ring obtained by adjoining * a root τ of $\phi(t)$ to D. For a suitable choice of the reference form f with multiplier P, $\bar{\mathscr{C}}$ can be identified with Δ in such a way that*

$$(\sum \alpha_i \tau^i)^\dagger = \sum (\lambda^{-1} \alpha_i^J \lambda) \tau^{-i} \qquad (\alpha_i \in D),$$

where $\lambda = \pm \lambda^J \in D$. λ can be chosen as 1 except when D is non-commutative, the restriction of J to Z is the identity and $\phi(t) = t \pm 1$.

PROOF. Using the definition (0.3.1), we may identify $\mathscr{C}$ with the ring $D[P]$. P commutes with the scalars $\alpha I \in D[P]$ and $P^\dagger = P^{-1}$. Also, since $\phi(P)$ is nilpotent and $\bar{\mathscr{C}}$ a division ring (because P is indecomposable), $\phi(\bar{P}) = 0$. Thus the mapping $\sum \alpha_i \tau^i \to \sum \bar{\alpha}_i \bar{P}^i$ is an isomorphism of Δ onto $\bar{\mathscr{C}}$. It remains to prove that, for a suitable choice of f,

$$(\bar{\alpha})^\dagger = \overline{(\lambda^{-1} \alpha^J \lambda)} \qquad (\alpha \in D), \tag{2.4.6}$$

where λ satisfies the stated conditions.

If the matrix of f with respect to the basis $u, uP, uP^2, \cdots$ has coefficients in Z, it is clear that (2.4.6) holds with $\lambda = 1$ (for $\alpha I \in D[P]$ has matrix $\alpha \boldsymbol{I}$ with respect to this basis). An examination of the proof of lemma 2.3.3 shows that such a choice of f is possible except when D is non-commutative, J is the identity on Z and $\phi(t) = t \pm 1$.

Let us now consider this exceptional case. Write $P = \pm I + M$, where M is nilpotent. Choose a form f with multiplier P and let $\boldsymbol{\Omega} = (\omega_{ij})$ be its matrix with respect to the basis $u, uM, uM^2, \cdots$. Let $(\alpha I)^\dagger = \sum_{i=0}^{m-1} \beta_i M^i$. It is clear that $\beta_0 = \alpha^\theta$, where θ is an involutory anti-automorphism of D. Comparing the last columns in the adjoint equation

$$\operatorname{diag}(\alpha, \alpha, \cdots \alpha)\, \boldsymbol{\Omega} = \boldsymbol{\Omega} \begin{pmatrix} \beta_0 & \beta_1 & \cdots & \cdot \\ 0 & \beta_0 & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & \beta_0 \end{pmatrix}^*,$$

we get $\alpha \omega_{im} = \omega_{im} (\alpha^\theta)^J$ $(i = 1, \cdots m)$. Since not all ω_{im} are zero, α^θ has the form $\lambda^{-1} \alpha^J \lambda$ for some $\lambda \in D$. Since θ is involutory, $\alpha = \lambda^{-1} \lambda^J \alpha (\lambda^{-1})^J \lambda$, whence $\lambda^J = \lambda \zeta$, $\zeta \in Z$. Since $\zeta \zeta^J = 1$ and J is the identity on Z, $\zeta = \pm 1$ and so $\lambda^J = \pm \lambda$ as required. This proves the lemma.

* Cf. § 0.3.

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

2.5 *Centralizers.* We investigate the centralizers of the finite-dimensional elements of a classical group G. We suppose first that G is the unitary group of a *trace-valued* form $F(x, y) = x \cdot y$ $(= \varepsilon(y \cdot x), \varepsilon = \pm 1)$ on V and later describe the modifications necessary in the orthogonal case. Let X be a finite-dimensional element of G with space $V_X = W$, form $f_X(u, v) = (u, v)$ and multiplier $P_X = P$. Let $C = C(X)$ denote the centralizer of X in G, $\mathscr{C} = \mathscr{C}(P)$ the commutator ring of P.

Each element Y of C leaves W and f_X invariant. Let $\eta(Y)$ denote the restriction of Y to W. Clearly, the mapping $Y \to \eta(Y)$ is a homomorphism of C into the group of invertible elements of $\mathscr{C}$. The kernel $C_1 = C_1(X)$ consists of the Y which leave every element of W invariant. Thus

$$C_1 = \{Y \in G | V_Y \subset W^\perp\}.$$

LEMMA 2.5.1. *C/C_1 is isomorphic to the norm group $N(\ddagger, \mathscr{C})$, where $\ddagger$ denotes the adjoint mapping with respect to f_X.*

PROOF. We have to show that $\eta(C) = N(\ddagger, \mathscr{C})$. Let $Y \in C$. Since Y leaves f_X invariant, we have

$$(u, v) = (u\eta(Y), v\eta(Y)) = (u, v\eta(Y)\eta(Y)^\ddagger)$$

when $u, v \in W$. Thus $\eta(Y) \in N(\ddagger, \mathscr{C})$ and so $\eta(C) \subset N(\ddagger, \mathscr{C})$.

Conversely, let $R \in N(\ddagger, \mathscr{C})$. Since R leaves f_X invariant it also leaves the restriction of F to W invariant, i.e. it is an isometry of W onto W. By Witt's theorem, R can be extended to an element Y of G. Clearly, $Y \in C$ and $R = \eta(Y)$, whence $N(\ddagger, \mathscr{C}) \subset \eta(C)$. This proves the lemma.

Notation. If L is a subspace of V, the equations

$$(x + L^\rho) \cdot (y + L^\rho) = x \cdot y \qquad (x, y \in L)$$

define a non-degenerate form on L/L^ρ called the *core* of F on L.

Each element Y of C_1 leaves $W^\perp$ and $(W^\perp)^\rho = W^\rho$ invariant. Let $\eta_1(Y)$ denote the linear transformation on $W^\perp/W^\rho$ induced by Y. Then η_1 is a homomorphism of C_1 into the group of non-singular linear transformations on $W^\perp/W^\rho$. The kernel $C_2 = C_2(X)$ of η_1 consists of the Y which map every coset $x + W^\rho (x \in W^\perp)$ into itself:

$$C_2 = \{Y \in G | V_Y \subset W^\perp \quad \text{and} \quad W^\perp(I - Y) \subset W^\rho\}.$$

LEMMA 2.5.2. *$C_1/C_2 \cong U(F')$, where F' is the core of F on $W^\perp$.*

PROOF. It is evident that $\eta_1(C_1) \subset U(F')$. It remains to prove that $U(F') \subset \eta_1(C_1)$, i.e. that each $S' \in U(F')$ can be "extended" to an element of C_1.

Choose a basis $u_1, \cdots, u_r$ of W^ρ and elements $v_1, \cdots, v_r$ of V such that $u_i \cdot v_j = \delta_{ij}$ $(i, j = 1, \cdots, r)$. Then the subspace $M = H^\perp$, where $H = W +$

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

$\{v_1, \cdots, v_r\}$, is a complement of W^ρ in $W^\perp$ and of H in V: $W^\perp = M \oplus W^\rho$, $V = M \oplus H$. We now define a linear transformation S on V as follows:

$$hS = h \quad (h \in H),$$
$$mS = m' \text{ when } (m + W^\rho)S' = (m' + W^\rho) \quad (m, m' \in M).$$

Then it is clear that $S \in C$, and $\eta_1(S) = S'$. This proves the lemma.

Let $Y \in C_2$. Since Y is a finite-dimensional element of G, its restriction $\eta_2(Y)$ to $W^\perp$ has the form

$$x \to x - \sum_{i=1}^{r} (x \cdot w_i)u_i \qquad (w_1, \cdots, w_r \in W^\perp), \tag{2.5.1}$$

where, as before $u_1, \cdots, u_r$ are basis a of W^ρ. The kernel $C_3 = C_3(X)$ of the homomorphic mapping η_2 consists of the Y which leave every element of $W^\perp$ invariant. Thus

$$C_3 = \{Y \in G | V_Y \subset (W + W^\perp)^\perp = W^\rho\}.$$

LEMMA 2.5.3. *$C_2/C_3 \cong A$, where A is the abelian group formed by all linear transformations on $W^\perp$ of the form* (2.5.1).

PROOF. Let Y' be the element (2.5.1) of A. It is required to prove that Y' can be extended to an element of C_2. Let $v_1, \cdots, v_r$ and M be defined as in the proof of the previous lemma. Without loss of generality we may suppose that the w_i are in M. We now define a linear transformation Y on V by:

$$xY = xY' \quad \text{when} \quad x \in W^\perp;$$
$$xY = x \qquad \text{when} \quad x \in W;$$
$$v_iY = v_i - \varepsilon w_i - \sum_{1}^{r} \lambda_{ij}u_j \qquad (i = 1, \cdots, r),$$

where the λ_{ij} are arbitrary elements of D. Y is of course an extension of Y'. It is easily verified that $Y \in C_2$ if, and only if, $\lambda_{ij} + \varepsilon\lambda_{ji}^J = (w_i \cdot w_j)$ $(i, j = 1, \cdots, r)$. Such a choice of the λ_{ij} is possible because F is assumed to be trace-valued. This completes the proof.

The structure of C_3 is easily determined. Each of its elements has the form

$$x \to x - \sum_{i,j=1}^{r} (x \cdot u_i)\omega_{ij}u_j, \tag{2.5.2}$$

where $u_1, \cdots, u_r$ are a basis of W^ρ. Conversely, the linear transformation (2.5.2) is an element of C_3 if, and only, if, $\omega_{ij} + \varepsilon\omega_{ji}^J = 0$ $(i, j = 1, \cdots, r)$. Thus we have

LEMMA 2.5.4. *C_3 is isomorphic to the additive group formed by all $r \times r$ matrices $\boldsymbol{K}$ over D such that $\boldsymbol{K} + \varepsilon\boldsymbol{K}^* = \boldsymbol{0}$ $(r = \dim W^\rho)$.*

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

The following important reduction theorem is an immediate consequence of the uniqueness of the Fitting decomposition. It allows us, in particular, to restrict attention to the two special cases: (a) $W = V$ and (b) $X - I$ is nilpotent.

THEOREM 2.5.1. *Let*

$$V = V_0 \oplus \cdots \oplus V_k, \quad -\epsilon X^{-1} = (-\epsilon X_0^{-1}) \oplus \cdots (-\varepsilon X_k^{-1}),$$

be the Fitting decomposition of $-\varepsilon X^{-1}$ *corresponding to* (1.4.7). *Let* $G_i = U(F_i)$, *where* F_i *is the restriction of* F *to* V_i, *and let* $C(X_i)$ *be the centralizer of* X_i *in* G_i $(i = 0, \cdots, k)$. *If* Y *is a linear transformation on* V, *then* $Y \epsilon C(X)$ *if, and only if, it has the form* $Y_0 \oplus \cdots \oplus Y_k$, *where* $Y_i \epsilon C(X_i)$ $(i = 0, \cdots, k)$.

Although lemmas 2.5.2—2.5.4 are perfectly explicit, lemma 2.5.1 requires further elucidation. When $\mathscr{C}$ satisfies the trace condition, the finer structure of $N(\ddagger, \mathscr{C})$ is given by theorem 2.2.3. The group $N(\ddagger, \overline{\mathscr{C}})$ was discussed in detail in § 2.4. With the groups S_i^- in mind, we now study the following situation. Let $\mathscr{I}$ be an ideal of $\mathscr{C}$ such that $\mathscr{I} = \mathscr{I}^\ddagger \subset \mathscr{N}$. Consider its additive subgroups

$$\mathscr{I}^\theta(\ddagger) = \mathscr{I}^\theta = \{Y \epsilon \mathscr{I} | Y^\ddagger = \theta Y\} \qquad (\theta = \pm 1).$$

The $\mathscr{I}^\theta$ are vector spaces over the subfield, Z_0, of the centre Z formed by its symmetric elements. We shall determine their dimensions $\dim_0 \mathscr{I}^\theta$ over Z_0. We assume that the dimension of D over Z is finite and that, when $p = 2$ and $Z = Z_0$, $I + P$ is non-singular. The first asumption ensures that $\dim_0 \mathscr{C}$ is finite, for the elements of $\mathscr{C}$ are finite-dimensional Z_0-linear transformations.

We begin with three simple observations. Firstly, our assumptions imply that $\mathscr{C}$ satisfies the trace condition. Secondly, $Y \to Y + Y^\ddagger (Y \epsilon \mathscr{I})$ is a Z_0-linear mapping of $\mathscr{I}$ onto $\mathscr{I}^+$ with kernel $\mathscr{I}^-$, so that

$$\dim_0 \mathscr{I}^+ + \dim_0 \mathscr{I}^- = \dim_0 \mathscr{I}. \tag{2.5.3}$$

Thirdly, if $\mathscr{C}$ contains a non-singular central element M such that $M^\ddagger = -M$, then $Y \leftrightarrow MY$ $(Y \epsilon \mathscr{I}^+)$ is a Z_0-linear isomorphism between $\mathscr{I}^+$ and $\mathscr{I}^-$. Hence, in this case,

$$\dim_0 \mathscr{I}^\theta = \tfrac{1}{2} \dim_0 \mathscr{I} \qquad (\theta = \pm 1). \tag{2.5.4}$$

(We shall regard $\dim_0 \mathscr{I}$ as known, so that (2.5.4), when it holds, is a complete answer to our problem.)

Suppose that we are given a direct decomposition

$$fX = f_1 \oplus \cdots \oplus f_k, \qquad P = P_1 \oplus \cdots \oplus P_k.$$

Let $I = E_1 + \cdots + E_k$ be the corresponding idempotent decomposition of the identity. If $\mathscr{J}$ is an ideal of $\mathscr{C}$, we write $\mathscr{J}_{ij} = E_i \mathscr{J} E_j$ $(i, j = 1, \cdots, k)$.

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

Then each $Y \in \mathscr{J}$ has a unique decomposition $Y = \sum_{i,j} Y_{ij}$, where Y_{ij} $(= E_i Y E_j) \in \mathscr{J}_{ij} \cdot \mathscr{C}_{ii}$ may be identified with the commutator ring of P_i and then $\ddagger$ (applied to the elements of $\mathscr{C}_{ii}$) becomes the adjoint mapping with respect to $f_i \cdot \mathscr{N}_{ii}$ is the radical of $\mathscr{C}_{ii}$, $\mathscr{J}_{ii}$ an ideal of $\mathscr{C}_{ii}$. We prove now that

$$(2.5.5) \qquad (\dim_0 \mathscr{J}^\theta - \tfrac{1}{2}\dim_0 \mathscr{J}) = \sum_{i=1}^{k} (\dim_0 \mathscr{J}^\theta_{ii} - \tfrac{1}{2}\dim_0 \mathscr{J}_{ii}).$$

Indeed let $Y \in \mathscr{J}$. Since $E_i^\ddagger = E_i$, we have $(Y^\ddagger)_{ij} = (Y_{ji})^\ddagger$. Therefore $Y \in \mathscr{J}^\theta$ if, and only if, $Y_{ii} \in \mathscr{J}^\theta_{ii} (1 \leqq i \leqq k)$, $Y_{ij} = \theta(Y_{ji})^\ddagger (1 \leqq i < j \leqq k)$. (2.5.5) now follows after a simple computation.

In view of (2.5.5), we may now assume that f is directly indecompsable. By theorem 2.4.1 (cor.), P has no proper direct decomposition into multipliers. By theorem 2.3.1 and our assumption that $I + P$ is non-singular when $p = 2$ and $Z = Z_0$, either P is directly indecomposable or P has a proper direct decomposition $P = (-P') \oplus (-P'')$, where P' and P'' are directly indecomposable multipliers. If g is a form with multiplier P and $\dagger$ the adjoint mapping with respect to g, it is easy to see that $\mathscr{J}^{-\theta}(\ddagger)$ and $\mathscr{J}^\theta(\dagger)$ are isomorphic Z_0-spaces. We may therefore now assume that P is directly indecomposable.

We distinguish two cases. *First case*: either $Z \neq Z_0$ or $P^2 - I$ is non-singular. Here $\mathscr{C}$ contains a non-singular central element M such that $M^\ddagger = -M$ and so $\dim_0 \mathscr{J}^\theta$ is given by (2.5.4). In fact, if $Z_0 \neq Z$, we may take $M = (\lambda - \lambda^J)I$, where λ is any element of Z not in Z_0, and if $P^2 - I$ is non-singular, we may take $M = P - P^{-1}$. *Second case*: $Z = Z_0$ and $P = \pm I - R$, where R is nilpotent and indecomposable. Let $R^e = 0$, $R^{e-1} \neq 0$; then $\mathscr{J}$ is one of the ideals $\mathscr{J}_i = \mathscr{C}R^i (1 \leqq i \leqq e)$. Let now $Y = \alpha R^i + \beta R^{i+1} + \cdots (i < e)$. By lemma 2.4.2, $Y^\ddagger = \alpha^{J'}(-1)^i R^i + \cdots$, where $\alpha^{J'}$ has the form $\lambda^{-1}\alpha^J\lambda$, $\lambda^J = \pm\lambda$. Let D^θ denote the Z_0-subspace of D formed by its elements μ such that $\mu^{J'} = \theta\,\mu$. Then the dimensions of D and D^θ have the forms

$$\dim_0 D = k^2, \quad \dim_0 D^\theta = \tfrac{1}{2}k(k + \theta\tau) \qquad (\tau = \pm 1)$$

(cf. Dieudonné [2]). Since $Y \pm Y^\ddagger = (\alpha \pm \alpha^{J'}(-1)^i)R^i + \cdots$ and $\mathscr{C}$ satisfies the trace condition, it follows that

$$\begin{aligned} \dim_0 \mathscr{J}_i^\theta - \dim_0 \mathscr{J}_{i+1}^\theta &= \dim_0 (\mathscr{J}_i^\theta + \mathscr{J}_{i+1})/\mathscr{J}_{i+1} \\ &= \tfrac{1}{2}k(k + (-1)^i\theta\tau). \end{aligned}$$

Thus

$$(2.5.6) \quad (\dim_0 \mathscr{J}_i^\theta - \tfrac{1}{2}\dim_0 \mathscr{J}_i) = (\dim_0 \mathscr{J}_{i+1}^\theta - \tfrac{1}{2}\dim_0 \mathscr{J}_{i+1}) + \tfrac{1}{2}(-1)^i k\theta\tau,$$

from which the value of $\dim_0 \mathscr{J}_i^\theta - \tfrac{1}{2}\dim_0 \mathscr{J}_i$ follows at once.

In particular, suppose that G is a finite-dimensional symplectic or ortho-

gonal group over a field of characteristic not 2. Let μ_i^ω be the multiplicity of $(t+\omega)^i$ as elementary divisor of X ($\omega = \pm 1$; $i = 1, 2, \cdots$). The above methods give, for the group S_0^- in theorem 2.2.3,

$$\dim_0 S_0^-(\dagger, \mathscr{C}) - \tfrac{1}{2}\dim_0 \mathscr{N} = \tfrac{1}{2}\varepsilon \sum_{k\geq 1} (\mu_{2k+1}^- - \mu_{2k}^+), \tag{2.5.7}$$

where $\varepsilon = 1$ or -1 according as G is symplectic or orthogonal.

Orthogonal groups. ($G = O(Q)$, $Q(x) = |x|$, $p = 2$). The principal modifications necessary (apart from the obvious replacements of F by Q, F_i by Q_i, etc.) is to substitute the singular radical for the radical throughout. Thus, the core of Q on L is the non-degenerate quadratic form on L/L^σ defined by

$$|x + L^\sigma| = |x| \quad (x \in L).$$

Also W^ρ is to be replaced by W^σ in the definitions of C_2, C_3, A and in lemma 2.5.4. All the unitary results carry over with the possible exception of lemma 2.5.2. I do not know whether C_1/C_2 is always isomorphic to the orthogonal group $O(Q')$ of the core Q' of Q on $W^\perp$. However, this is true if $W^\rho = W^\sigma$, and in any case C_1/C_2 is isomorphic to a subgroup of $O(Q')$ which contains all finite-dimensional elements.

2.6 *Example*: *finite classical groups.* We suppose now that D is a Galois field $GF(q)$, where $q = p^\alpha$, and that the dimension, n, of V is finite. The symplectic and orthogonal groups over fields of characteristic 2 are excluded from the discussion. Three cases arise.

(A) *Unitary.* Here q is a square r^2 and $\lambda^J = \lambda^r$. Any two non-degenerate skew-Hermitian forms on V are equivalent, so that the unitary group $U(n, r^2)$ is essentially unique. Its order is

$$|U(n, r^2)| = r^{\frac{1}{2}(n^2-n)} \prod_1^n (r^i - (-1)^i). \tag{2.6.1}$$

In this unitary case we allow p to be 2.

(B) *Symplectic.* Here n is even: $n = 2\nu$. We assume that $p \neq 2$. The 2ν-dimensional symplectic group $Sp(2\nu, q)$ is essentially unique and

$$|Sp(2\nu, q)| = q^{\nu^2} \prod_1^\nu (q^{2i} - 1). \tag{2.6.2}$$

(C) *Orthogonal.* We assume that $p \neq 2$. If $n = 2\nu$, a non-degenerate quadratic form $|x|$ on V is equivalent to

$$\sum_1^\nu x_{2i-1}x_{2i} \quad \text{(of Witt index } \nu)$$

or

$$(x_1^2 - \delta x_2^2) + \sum_2^\nu x_{2i-1}x_{2i} \quad \text{(of Witt index } \nu - 1),$$

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

where δ is a fixed non-square of D. The corresponding orthogonal groups are denoted by $O_+(2\nu, q)$, $O_-(2\nu, q)$ respectively. Their orders are

$$|O_\pm(2\nu, q)| = 2q^{\nu^2-\nu}(q^\nu \mp 1)\prod_1^{\nu-1}(q^{2i}-1). \tag{2.6.3}$$

If $n = 2\nu + 1$, $|x|$ is equivalent to

$$\sum_1^n x_i^2 \quad \text{or} \quad \delta\sum_1^n x_i^2.$$

The common orthogonal group is denoted by $O(n, q)$. Its order is

$$|O(2\nu+1, q)| = 2q^{\nu^2}\prod_1^{\nu}(q^{2i}-1). \tag{2.6.4}$$

For notational convenience, we observe the following conventions. If n is even, $O(n, q)$ stands for an unspecified one of $O_+(n, q)$ and $O_-(n, q)$. If n is odd, $O_+(n, q) = O_-(n, q) = O(n, q)$.

Notation. X stands for a non-singular linear transformation on V. $\phi = \phi(t)$ stands generically for an irreducible monic polynomial over $GF(q)$, distinct from t. $|\phi|$ denotes the degree of ϕ and $m(\phi^\kappa)$ the multiplicity of ϕ^κ as elementary divisor of X. $|K|$ denotes the number of elements in a finite set K.

Case (A). This is the simplest case. In a number of respects, $U(n, q)$ behaves like the full linear group $GL(n, q)$.

(i) *X is similar to an element of $U(n, q)$ if, and only if, $X \sim X^{*-1}$, i.e. $m(\phi^\kappa) = m(\tilde{\phi}^\kappa)$ for all ϕ, κ.*

(ii) *Two elements of $U(n, q)$ are conjugate in $U(n, q)$ if, and only if, they are similar.*

(iii) *The number of conjugacy classes in $U(n, q)$ is the coefficient of t^n in*

$$\prod_{\lambda=1}^{\infty}\left(\frac{1+t^\lambda}{1-q^{\frac{1}{2}}t^\lambda}\right).$$

(iv) *Let $X \in U(n, q)$. Write*

$$A(\phi^\mu) = \begin{cases} |U(m_\mu, Q)| & (\phi = \tilde{\phi}), \\ |GL(m_\mu, Q)|^{\frac{1}{2}} & (\phi \neq \tilde{\phi}), \end{cases}$$

$$B(\phi) = Q^{\sum_{\mu<\nu}\mu m_\mu m_\nu + \frac{1}{2}\sum_\mu(\mu-1)m^2{}_\mu}\prod_\mu A(\phi^\mu),$$

where $Q = q^{|\phi|}$, $m_\mu = m(\phi^\mu)$. Then the order of the conjugacy class of X in $U(n, q)$ is

$$|U(n, q)|/\prod_\phi B(\phi).$$

Proof of (i). Suppose that $X \sim X^{*-1}$. Let P be the restriction of X^{-1}

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

to $W = V(I - X)$. Since $P \sim P^{*-1}$, P is the multiplier of a certain form f, by theorem 2.3.1. Let

$$G(x, y) = f(x, y) - f(y, x)^J. \tag{2.6.5}$$

Now, since $W = V(I - X)$, $\dim V - \dim W = m_1 + m_2 + \cdots$, where $m_i = m((t-1)^i)$. Also, by (2.6.5), the G-radical W^ρ of W is the null-space of $I - P$. Hence $\dim W^\rho = m_2 + m_3 + \cdots$. Since $\dim W^\rho \leqq \dim V - \dim W$, G can be extended to a non-degenerate skew-Hermitian form F on V. Then X is similar to the element of $U(F)$ with space W and form f, which proves our result.

Proof of (ii). Suppose that $X \sim Y$, where $X, Y \in U(n, q)$. By theorem 2.4.2, it is sufficient to prove that corresponding Hermitian invariants of X, Y are equivalent. Let χ, ψ be the Hermitian invariants of X, Y corresponding to the elementary divisor ϕ^κ (where $\phi = \tilde{\phi}$). Let m' be the common multiplicity of ϕ^κ as elementary divisor of X, Y, $\Delta = D[\tau]$ the field obtained by adjoining a root τ of ϕ to D, J' the involutory automorphism of Δ which extends J and maps τ into τ^{-1}. Then χ, ψ are both m'-dimensional, non-degenerate, J'-Hermitian forms over Δ. Since J' is not the identity, $\chi \approx \psi$ as required.

Proof of (iii). We use a method of W. Feit and N. J. Fine ([3]). Let k_n denote the number of conjugacy classes in $U(n, q)$. By (i), k_n is the number of similarity classes of linear transformations X such that $X \sim X^{*-1}$. Let $f_1(t), f_2(t), \cdots$ be the invariant factors of X, where f_{i+1} is a divisor of f_i for each i. In terms of the quotients $g_i = f_i/f_{i+1}$, the condition $X \sim X^{*-1}$ means that

$$\begin{aligned} &g_i(0) \neq 0, \quad g_i = \tilde{g}_i \quad (i = 1, 2, \cdots), \\ &|g_1| + 2|g_2| + 3|g_3| + \cdots = n. \end{aligned} \tag{2.6.6}$$

Let $c(d)$ denote the number of monic polynomials $g(t)$ such that

$$g(0) \neq 0, \quad g = \tilde{g}, \quad |g| = d.$$

Then, by (2.6.6),

$$k_n = \sum_{d_1 + 2d_2 + \cdots = n} c(d_1)c(d_2)\cdots$$

and so

$$\sum_0^\infty k_n t^n = \prod_{\lambda=1}^\infty \sum_{i=0}^\infty c(i)t^{i\lambda} \quad (c(0) = k_0 = 1).$$

An easy enumeration gives $c(i) = q^{\frac{1}{2}i} + q^{\frac{1}{2}(i-1)}$, so that $\sum c(i)t^i = (1 + t)(1 - q^{\frac{1}{2}}t)^{-1}$. This proves our result.

Proof of (iv). It is required to prove that $|C(X)| = \prod_\phi B(\phi)$. We use the notation of § 2.5. By lemma 1.4.4 (cor.) and lemma 1.4.5 (cor. 2), we have

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

$$\dim W^{\perp} = \sum_{i \geq 1} m_i', \quad \dim W^{\rho} = \sum_{i \geq 2} m_i',$$

where $m_i' = m((t-1)^i)$. By § 2.5 and theorem 2.2.3,

$$\begin{aligned} |C/C_1| &= |N(\ddagger, \mathscr{C})| = |N(\ddagger, \bar{\mathscr{C}})||S_0^-(\ddagger, \mathscr{C})| \\ &= |N(\ddagger, \bar{\mathscr{C}})||\mathscr{N}|^{\frac{1}{2}}, \\ |C_1/C_2| &= |U(m_i', q)|, \\ |C_2| &= q^{\frac{1}{2}(\Sigma_{i \geq 1} m_i')^2 - \frac{1}{2}m_1'^2}. \end{aligned}$$

The value of $|N(\ddagger, \bar{\mathscr{C}})|$ follows from the discussion in § 2.4 and $|\mathscr{N}|$ is easily evaluated from elementary divisor theory. We find that

$$|C/C_1| = \Big(\prod_{\phi \neq t-1} B(\phi)\Big) q^{\frac{1}{2}\Sigma_{i \geq 1}(i-1)m_{i+1}'^2 + \Sigma_{1 \leq i < j} i m_{i+1}' m_{j+1}'},$$

whence $|C| = \prod_{\phi} B(\phi)$ as required.

Case (B). This case is similar to the unitary one, but with added complications due to the elementary divisors $(t \pm 1)^k$. *Notation.* If $X \in Sp(n, q)$, let $\psi_{2i}^+(X)$, $\psi_{2i}^-(X)$ denote the Hermitian invariants of X associated with the elementary divisors $(t+1)^{2i}$, $(t-1)^{2i-1}$ of P respectively, and thus, by lemma 1.4.5 and corollaries, with the elementary divisors $(t+1)^{2i}$, $(t-1)^{2i}$ of X respectively $(i = 1, 2, \cdots)$. The $\psi_{2i}^{\pm}(X)$ are symmetric bilinear forms over $GF(q)$.

(i) *X is similar to an element of $Sp(n, q)$ if, and only if,*
  (a) $X \sim X^{*-1}$,
  (b) *each elementary divisor $(t \pm 1)^{2k+1}$ of X has even multiplicity.*

(ii) *Two elements X, Y of $Sp(n, q)$ are conjugate in $Sp(n, q)$ if, and only if,*
  (a) $X \sim Y$,
  (b) $\psi_{2i}^+(X) \approx \psi_{2i}^+(Y)$ *and* $\psi_{2i}^-(X) \approx \psi_{2i}^-(Y)$ $(i = 1, 2, \cdots)$.

(iii) *The number of conjugacy classes in $Sp(n, q)$ is the coefficient of t^n in*

$$\prod_{\lambda=1}^{\infty} \left[\frac{(1+t^{2\lambda})^4}{1-qt^{2\lambda}}\right].$$

(iv) *Let $X \in Sp(n, q)$. Then the order of the conjugacy class of X in $Sp(n, q)$ is*

$$|Sp(n, q)| / \prod_{\phi} B(\phi),$$

where $B(\phi)$, $A(\phi^\mu)$ are defined as in the unitary case, except that, when $\phi(t) = t \pm 1$,

$$A(\phi^\mu) = \begin{cases} |Sp(m_\mu, q)| & (\mu \text{ odd}), \\ q^{\frac{1}{2}m_\mu}|O(m_\mu, q)| & (\mu \text{ even}). \end{cases}$$

Here $O(m_\mu, q)$ is the orthogonal group of the corresponding Hermitian invariant $\psi_\mu^{\pm}(X)$.

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

Proof of (i). Let $X \in Sp(n, q)$. (a) holds as in the unitary case. By lemma 1.4.5 and corollaries, the multiplicities of $(t+1)^{2k+1}$, $(t-1)^{2k+3}$ $(k \geqq 0)$ as elementary divisors of X are the multiplicities of $(t+1)^{2k+1}$, $(t-1)^{2k+2}$ as elementary divisors of P. By theorem 2.3.1, (b) holds for $(t+1)^{2k+1}$, $(t-1)^{2k+3} (k \geqq 0)$. It also holds for $(t-1)^1$, because the multiplicity of this elementary divisor is the dimension of the (non-degenerate, alternate) core of F on $W^\perp$. Thus the conditions are necessary. Their sufficiency is proved as in the unitary case.

Proof of (ii). As in the unitary case. The $\psi_{2i}^{\pm}$ are the only Hermitian invariants for which the corresponding automorphism J' is the identity (by lemma 2.4.2).

Proof of (iii). Let $F(t) = \sum k_n t^n$ be the generating function for the number of conjugacy classes in $Sp(n, q)$. By the Fitting decomposition,

$$F(t) = F_0(t) F_+(t) F_-(t),$$

where F_0, F_+, F_- are the generating functions for the numbers of conjugacy classes of $X \in Sp(n, q)$ such that $X^2 - I$ is non-singular, $X + I$ is nilpotent, $X - I$ is nilpotent, respectively. By the method of the unitary case,

$$F_0(t) = \prod_{\lambda=1}^{\infty} \sum_{i=0}^{\infty} c'(i) t^{i\lambda},$$

where $c'(d)$ is the number of monic polynomials $g(t)$ such that

$$g(0)g(1)g(-1) \neq 0, \quad g = \tilde{g}, \quad |g| = d.$$

Let $c(d)$ be the number of monic polynomials $g(t)$ such that

$$g(0) \neq 0, \quad g = \tilde{g}, \quad |g| = d.$$

Out of these $c(d)$ polynomials of degree d, $c(d-1)$ are divisible by $t+1$, $c(d-1)$ by $t-1$ and $c(d-2)$ by $t^2 - 1$. Hence

$$c'(d) = c(d) - 2c(d-1) + c(d-2)$$

and so

$$\sum_i c'(i) t^i = (1-t)^2 \sum_i c(i) t^i.$$

An easy enumeration gives $\sum c(i) t^i = (1+t)^2 (1 - qt^2)^{-1}$, whence

$$F_0(t) = \prod_{\lambda=1}^{\infty} \left[\frac{(1 - t^{2\lambda})^2}{1 - qt^{2\lambda}} \right].$$

Since $X - I$ is nilpotent if, and only if, $(-X) + I$ is nilpotent, we have

$$F_+(t) = F_-(t).$$

Let X be a linear transformation such that $X + I$ is nilpotent and let m_i

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

be the multiplicity of $(t+1)^i$ as elementary divisor of X. The conditions that X be similar to an element of $Sp(n, q)$ are:

$$m_{2i+1} \equiv 0 \pmod 2 \quad (i = 0, 1, \cdots)$$
$$m_1 + 2m_2 + \cdots = n.$$

Each such set of m_i gives rise to 2^s conjugacy classes of $Sp(n, q)$, where s is the number of non-zero multiplicities $m_2, m_4, \cdots$; for if $m_{2i} \neq 0$ there are two possible equivalence classes for the Hermitian invariant ψ_{2i}^+. It follows that

$$F_+(t) = (\sum_0^\infty t^{2i})(1 + 2\sum_1^\infty t^{2i})(\sum_0^\infty t^{6i})(1 + 2\sum_1^\infty t^{4i}) \cdots$$
$$= (1-t^2)^{-1}\left(\frac{1+t^2}{1-t^2}\right)(1-t^6)^{-1}\left(\frac{1+t^4}{1-t^4}\right)\cdots$$
$$= \prod_{\lambda=1}^\infty \frac{(1+t^{2\lambda})(1-t^{4\lambda})}{(1-t^{2\lambda})^2} = \prod_{\lambda=1}^\infty \left[\frac{(1+t^{2\lambda})^2}{1-t^{2\lambda}}\right].$$

Hence

$$F(t) = F_0(t)F_+(t)^2 = \prod_{\lambda=1}^\infty \left[\frac{(1+t^{2\lambda})^4}{1-qt^{2\lambda}}\right],$$

as required.

Proof of (iv). As in unitary case. We remark only that $|S_0^-|$ differs from $|\mathcal{N}(P)|^{\frac{1}{2}}$ in this case. The correction factor (2.5.7) is incorporated in the modified definition of $A((t \pm 1)^{2\mu})$.

Case (C) Two kinds of complication arise in this case, the first due to the elementary divisors $(t \pm 1)^k$ and the second to the fact that there are two orthogonal groups for a given even dimension. *Notation.* If $X \in O(n, q)$, let $\psi_{2i-1}^+(X)$, $\psi_{2i+1}^-(X)$ denote the Hermitian invariants of X associated with the elementary divisors $(t-1)^{2i-1}$, $(t+1)^{2i}$ of P *respectively,* and thus, by lemma 1.4.5 and corollaries, with the elementary divisors $(t+1)^{2i-1}$, $(t-1)^{2i+1}$ of X respectively $(i = 1, 2, \cdots)$. The ψ_{2i-1}^+, ψ_{2i+1}^- are symmetric bilinear forms over $GF(q)$. It is formally convenient to define ψ_1^- as the core of F on $W^\perp$. Then $\psi_{2i-1}^\pm$ is defined for $i = 1, 2, \cdots$.

(i) *X is similar to an element of some orthogonal group $O(n, q)$ if, and only if,*
  (a) $X \sim X^{*-1}$,
  (b) *each elementary divisor $(t \pm 1)^{2k}$ of X has even multiplicity.*

(i)′ *Let n be even and suppose that* (a), (b) *in* (i) *are satisfied. If any elementary divisor $(t \pm 1)^{2k+1}$ of X has positive multiplicity, X is similar to an element of $O_+(n, q)$ and also to an element of $O_-(n, q)$. If every such elementary divisor has multiplicity zero, X is similar to an element of $O_+(n, q)$ $(O_-(n, q))$ if, and only if, $\sum_{\phi,\mu} \mu m(\phi^\mu) \equiv 0 \pmod 2$ $(\sum_{\phi,\mu} \mu m(\phi^\mu) \equiv 1 \pmod 2)$.*

(ii) *Two elements X, Y of $O(n, q)$ are conjugate in $O(n, q)$ if, and only if,*

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

(a) $X \sim Y$,

(b) $\psi^+_{2i+1}(X) \approx \psi^+_{2i+1}(Y)$ *and* $\psi^-_{2i+1}(X) \approx \psi^-_{2i+1}(Y)$ $(i = 0, 1, \cdots)$.

(iii) *Let* k^+_n, k^-_n *denote the numbers of conjugacy classes in* $O_+(n, q)$, $O_-(n, q)$ *respectively. Then*

$$\sum_{n=0}^{\infty} (k^+_n + k^-_n)t^n = \prod_{\lambda=1}^{\infty} \left[\frac{(1 + t^{2\lambda-1})^4}{1 - qt^{2\lambda}}\right],$$

$$\sum_{n=0}^{\infty} (k^+_n - k^-_n)t^n = \prod_{\lambda=1}^{\infty} \left(\frac{1 - t^{4\lambda-2}}{1 - qt^{4\lambda}}\right).$$

(iv) *Let* $X \in O(n, q)$. *Then the order of the conjugacy class of* X *in* $O(n, q)$ *is*

$$|O(n, q)| / \prod_{\phi} B(\phi),$$

where $B(\phi)$, $A(\phi^\mu)$ *are defined as in the unitary and symplectic cases, expect that, when* $\phi(t) = t \pm 1$,

$$A(\phi^\mu) = \begin{matrix} |O(m_\mu, q)| & (\mu \text{ odd}), \\ q^{-\frac{1}{2}m_\mu}|Sp(m_\mu, q)| & (\mu \text{ even}). \end{matrix}$$

Here $O(m_\mu, q)$ *is the orthogonal group of the corresponding Hermitian invariant* $\psi^\pm_\mu(X)$.

Proof of (i), (ii), (iv). As in the symplectic case.

As a preliminary to the proof of (i)′, we derive a formula for the Witt type * of the fundamental quadratic form in terms of the conjugacy invariants of an element X of its orthogonal group $O(n, q)$. (We may regard $O(n, q)$ either as the unitary group $U(F)$ of the nondegenerate symmetric bilinear form $F(x, y) = x \cdot y$ or as the orthogonal group $O(|x|)$ of the quadratic form $|x| = \frac{1}{2}F(x, x)$.) There are 4 Witt types over $GF(q)$, viz. $\mathbf{0}, \mathbf{1}, \boldsymbol{\delta}, \boldsymbol{\omega} = \mathbf{1} - \boldsymbol{\delta}$, corresponding to the forms $0, x^2, \delta x^2, x^2 - \delta y^2$, where δ is a fixed non-square of $GF(q)$. The Witt type of a quadratic form χ is denoted by $\boldsymbol{\tau}(\chi)$ and we write $\boldsymbol{\tau}^\pm_{2i+1} = \boldsymbol{\tau}(\psi^\pm_{2i+1}(X))$.

Let $W, g(x, y), P$ be the space, form and multiplier of X, $f(x, y)$ a standard reference form with multiplier P and Q the representative of g with respect to f. The Hermitian invariants $\psi^\pm_{2i-1}(X)$ are calculated with respect to f. We suppose that the matrices of f, P, Q are as in § 2.4. By theorem 2.4.2, we may assume that, in (2.4.4), $\boldsymbol{Q} = \text{diag}\ (\boldsymbol{Q}^{11}, \cdots, \boldsymbol{Q}^{ss})$. An index i $(1 \leqq i \leqq s)$ will be called exceptional when it corresponds to one of the Hermitian invariants $\psi^+_1, \psi^\pm_3, \cdots$, i.e. when ($i \leqq r$ and) the minimum polynomial of $\boldsymbol{P}^i$ is $(t - 1)^{2k-1}$ or $(t + 1)^{2k}$. By (ii) above, we may assume that, in (2.4.5), $\boldsymbol{Q}^{ii}$ is the unit matrix when i is not exceptional, and a block-diagonal matrix diag $(\alpha\boldsymbol{I}, \beta\boldsymbol{I}, \cdots)$ when i is exceptional, where $\alpha x^2 + \beta y^2 + \cdots$

* Cf. Bourbaki [1], § 8.

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

is equivalent to the corresponding Hermitian invariant $\psi^{\pm}_{2k+1}$.

LEMMA 2.6.1. *For a suitable choice of f,*

$$\tau(|x|) = \Big(\sum_{\phi,\mu}{}' \mu m(\phi^\mu)\Big)\omega + \sum_{i,\pm} \tau^{\pm}_{2i-1}, \tag{2.6.7}$$

where summation $\sum'$ is over all powers ϕ^μ of irreducible monic polynomials $\phi(t)$ distinct from t, $t+1$, $t-1$.

PROOF. Let $|x|'$ denote the restriction of $|x|$ to W. By (1.1.3), $|x|' = g(x, x)$. Also it is easy to see that

$$\tau(|x|) = \tau^{-}_{1} + \tau(|x|'). \tag{2.6.8}$$

Now let f^i, g^i, $|x|^{i'}$ denote the direct summands of f, g, $|x|'$ corresponding to $\boldsymbol{P}^i$ $(1 \leqq i \leqq s)$ and f_i the direct summand of f corresponding * to $\boldsymbol{P}_i$ $(1 \leqq i \leqq r)$. Suppose first that $r+1 \leqq i \leqq s$. Then $f^i(x, x) = g^i(x, x) = |x|^{i'}$. By (2.4.3), the matrix of $|x|^{i'}$ is

$$\tfrac{1}{2}\begin{pmatrix} \boldsymbol{0} & \boldsymbol{I} + (\boldsymbol{\Pi}^i)^{*-1} \\ \boldsymbol{I} + (\boldsymbol{\Pi}^i)^{*-1} & \boldsymbol{0} \end{pmatrix},$$

whence $\tau(|x|^{i'}) = \boldsymbol{0}$. (It is easy to see that this is true whether $\boldsymbol{I} + (\boldsymbol{\Pi}^i)^{*-1}$ is singular or not.) The corresponding part $\big(\sum'' \mu m(\phi^\mu)\big)\omega$ in (2.6.7) is also zero. In fact, summation is over the powers $\phi^\mu = (t \pm 1)^{2k}$ (for which the multiplicities $m(\phi^\mu)$ are even) and over the powers ϕ^μ such that $\phi \neq \tilde{\phi}$ (which occur in pairs ϕ^μ, $\tilde{\phi}^\mu$ with $m(\phi^\mu) = m(\tilde{\phi}^\mu)$).

Suppose next that $1 \leqq i \leqq r$. First let i be exceptional. By the way in which g^i was chosen, we have

$$\tau(|x|^{i'}) = \tau\big(g^i(x, x)\big) = \tau(\psi) \times \tau\big(f_i(x, x)\big),$$

where ψ is the corresponding Hermitian invariant and where $\times$ denotes multiplication in the ring of Witt types. The elementary divisor of $\boldsymbol{P}_i$ is either $(t-1)^{2k-1}$ or $(t+1)^{2k}$. Since the matrix of $f_i(x, x)$ is $\frac{1}{2}(\boldsymbol{\Omega}_i + \boldsymbol{\Omega}_i^T) = \frac{1}{2}(\boldsymbol{I} + \boldsymbol{P}_i^{-1})\boldsymbol{\Omega}_i$, $f_i(x, x)$ is a form of *odd* rank $2k-1$. Hence, replacing f^i by δf^i if necessary, we may suppose that $\tau\big(f_i(x, x)\big) = \boldsymbol{1}$ and so

$$\tau(|x|^{i'}) = \tau(\psi).$$

These terms, with τ^{-}_{1} in (2.6.8), give the part $\sum \tau^{\pm}_{2i-1}$ in (2.6.7).

Finally, let i be non-exceptional (and $\leqq r$). Then the minimum polynomial of $\boldsymbol{P}_i$ is a power ϕ^μ, where $\phi = \tilde{\phi}$, $\phi(t) \neq (t \pm 1)$. Since $\boldsymbol{I} + \boldsymbol{P}_i$ is non-singular, $f_i(x, x)$ is non-degenerate. By lemma 2.4.2, the field obtained by adjoining a root of ϕ to $GF(q)$ has an involutory automorphism which

* There are actually m_i different summands of f corresponding to the m_i occurrences of $\boldsymbol{P}_i$ in (2.4.2). But all are equivalent because f is standard.

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

is the identity on $GF(q)$. Hence $|\phi|$ is even, and so the dimension of $f_i(x, x)$ is even. By the form of (2.6.7), and since $f^i(x, x) = g^i(x, x) = |x|^{i'}$, it remains only to prove that

$$\tau(f_i(x, x)) = \mu\omega, \tag{2.6.9}$$

i.e. that $f_i(x, x)$ has type $\mathbf{0}$, ω according as μ is even or odd.

To prove (2.6.9), it is convenient to change to the notation of the proof of lemma 2.3.3. Thus we identify $f_i(x, x)$ with the quadratic form $\rho(R) = \chi(RR^\theta)$ on the Z-space $\mathscr{Z}$. The Supplementary Remarks in § 2.3 show that it is sufficient to prove our result when $\mu = 1$. In this case, $\mathscr{Z} \cong GF(q^d)$, $R^\theta = R^{q^{\frac{1}{2}d}+1}$, where $d = |\phi|$. Let $\mathscr{Z}_0$ be the subfield of $\mathscr{Z}$ of index 2. There are $q^{\frac{1}{2}d-1} - 1$ non-zero solutions S of $\chi(S) = 0$ in $\mathscr{Z}_0$ and, for each such S, $q^{\frac{1}{2}d} + 1$ elements R of $\mathscr{Z}$ such that $RR^\theta = S$. Hence the number of non-zero solutions of $\rho(R) = 0$ in $\mathscr{Z}$ is $(q^{\frac{1}{2}d-1} - 1)(q^{\frac{1}{2}d} + 1)$. This shows that $\tau(\rho) = \omega$, as required.

Proof of (i)′. This follows immediately from (i) and lemma 2.6.1.

Proof of (iii). Let

$$F^\pm(t) = \sum (k_n^+ \pm k_n^-)t^n,$$

where $k_n^\pm$ is the number of conjugacy classes in $O_\pm(n, q)$. Let $F_0^\pm$, $F_+^\pm$, $F_-^\pm$, $F_\phi^\pm$ be the similarly defined functions for the numbers of conjugacy classes of elements X of $O_\pm(n, q)$ such that $X^2 - I$ is non-singular, $X + I$ is nilpotent, $X - I$ is nilpotent, $\phi(X)\tilde{\phi}(X)$ is nilpotent, respectively.

By lemma 2.6.1, and the methods of the symplectic case, we have

$$F^\pm = F_0^\pm F_+^\pm F_-^\pm = F_0^\pm (F_+^\pm)^2, \tag{2.6.10}$$

$$F_0^\pm = \Big(\prod_{\phi=\tilde{\phi}\neq(t\pm1)} F_\phi^\pm\Big)\Big(\prod_{\phi\neq\tilde{\phi}} F_\phi^\pm\Big)^{\frac{1}{2}}. \tag{2.6.11}$$

$$F_0^+(t) = \prod_{\lambda=1}^{\infty}\left[\frac{(1-t^{2\lambda})^2}{1-qt^{2\lambda}}\right]. \tag{2.6.12}$$

If $\phi = \tilde{\phi} \neq (t \pm 1)$, the coefficient of t^n in F_ϕ^+ is the number of similarity classes of linear transformations X on V such that $\phi(X)$ is nilpotent, i.e., the number of partitions of $n/|\phi|$. Hence $F_\phi^+(t) = P(t^{|\phi|})$, where $P(t) = \prod_{\lambda=1}^{\infty}(1-t^\lambda)^{-1}$. Similarly, if $\phi \neq \tilde{\phi}$, $F_\phi^+(t) = P(t^{2|\phi|})$. Hence

$$F_0^+(t) = \prod_{\lambda=1}^{\infty} P(t^\lambda)^{N_\lambda} P(t^{2\lambda})^{M_\lambda}, \tag{2.6.13}$$

where N_λ is the number of ϕ such that $\phi = \tilde{\phi} \neq (t \pm 1)$, $|\phi| = \lambda$, and M_λ the number of pairs of ϕ such that $\phi \neq \tilde{\phi}$, $|\phi| = \lambda$.

A similar argument, together with lemma 2.6.1, gives

$$F_0^-(t) = \prod_{\lambda=1}^{\infty} \big(P(-t^\lambda)^{N_\lambda} P(t^{2\lambda})^{M_\lambda}\big). \tag{2.6.14}$$

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

Let $\Phi_0(t)$ be the generating function for the number of similarity classes of linear transformations X on V such that $X(X^2 - I)$ is non-singular. Applying the two different methods used to evaluate $F_0^+(t)$, we get

$$\begin{aligned}\Phi_0(t) &= \prod_{\lambda=1}^{\infty} P(t^\lambda)^{N_\lambda + 2M_\lambda} \\ &= \prod_{\lambda=1}^{\infty} \left[\frac{(1-t^\lambda)^3}{1-qt^\lambda}\right].\end{aligned} \tag{2.6.15}$$

By (2.6.12)—(2.6.15), we have

$$\begin{aligned}F_0^-(t) &= F_0^+(t^2)^2 \Phi_0(t^2)/F_0^+(t)\Phi_0(t^4) \\ &= \prod_{\lambda=1}^{\infty} \left[\frac{(1-t^{2\lambda})(1-t^{4\lambda})}{(1-qt^{4\lambda})}\right].\end{aligned} \tag{2.6.16}$$

Let X be an element of $O(n, q) = O(|x|)$ such that $X + I$ is nilpotent and let m_i be the multiplicity of $(t+1)^i$ as elementary divisor of X. Then

$$\begin{gathered}m_{2i} \equiv 0 \pmod 2 \quad (i = 1, 2, \cdots), \\ m_1 + 2m_2 + \cdots = n, \\ \tau_1^+ + \tau_3^+ + \cdots = \tau(|x|).\end{gathered}$$

It follows, as in the symplectic case, that

$$\begin{aligned}F_+^+(t) &= (1 + 2\sum_1^\infty t^i)(\sum_0^\infty t^{4i})(1 + 2\sum_1^\infty t^{3i})(\sum_0^\infty t^{8i}) \cdots \\ &= \prod_{\lambda=1}^{\infty} \left[\frac{(1+t^{2\lambda-1})^2}{1-t^{2\lambda}}\right].\end{aligned} \tag{2.6.17}$$

If s of the multiplicities $m_1, m_3, \cdots$ are positive and $s > 0$, then one half of the 2^s choices of the types $\tau_1^+, \tau_3^+, \cdots$ give rise to an element of $O_+(n, q)$ and the other half to an element of $O_-(n, q)$. It follows that the coefficient of t^n in $F_+^-(t)$ is the number of solutions of

$$\begin{gathered}m_{2i} \equiv 0 \pmod 2 \quad (i = 1, \cdots) \\ 2m_{2i} + 4m_{4i} + \cdots = n.\end{gathered}$$

Thus

$$F_+^-(t) = P(t^4) = \prod_{\lambda=1}^{\infty} (1 - t^{4\lambda})^{-1}. \tag{2.6.18}$$

The values of $F^{\pm}(t)$ now follow from (2.6.10), (2.6.12), (2.6.16), (2.6.17), (2.6.18). (Note that $\prod_{\lambda=1}^{\infty}(1 - t^{4\lambda-2}) = \prod_{\lambda=1}^{\infty}(1 + t^{2\lambda})^{-1}$.)

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

3. Equivalence (Exceptional Case)

Throughout § 3, D is a field of characteristic 2, J is the identity, W a vector space over D of finite dimension M and P a linear transformation on W such that $N = I + P$ is nilpotent. We consider the equivalence of forms on W with multiplier P. This is the simplest case in which theorem 2.4.1 breaks down. It is essentially more complicated than the cases treated already and our results are complete only when D satisfies the additional hypothesis:

(3.0.1) *every ternary quadratic form* $x^2 + xy + \lambda y^2 + \mu z^2$ *over* D *is indefinite* (*i.e. represents* 0 *non-trivially*).

We remark that (3.0.1) holds when D is perfect or when every quadratic extension of D is inseparable. Hence we shall be able to solve the conjugacy problem for the finite classical groups $O(n, 2^\alpha)$ and $Sp(n, 2^\alpha)$.

3.1 *A matrix representation of* $\mathscr{C}$. For brevity, we omit most of the proofs in this section. They are direct, though somewhat long, verifications.

Consider a splitting of P into indecomposable parts, say

$$W = \bigoplus_{i=1}^{r} W_i, \quad W_i = \bigoplus_{j=1}^{m_i} W_{ij},$$

$$P = \bigoplus_{i=1}^{r} P_i, \quad P_i = \bigoplus_{j=1}^{m_i} P_{ij},$$

where P_{ij} is indecomposable and has minimum polynomial $(1 + t)^i$, and where some of the multiplicities m_i may be zero. By theorem 2.3.1, m_i is even when i is even. By the same theorem, we may suppose that the standard reference form f has the form

$$f = \bigoplus_{i=1}^{r} f_i, \quad f_i = \bigoplus_{j=1}^{m_i} f_{ij} \ (i \text{ odd}), \quad f_i = \bigoplus_{j=1}^{\frac{1}{2}m_i} \phi_{i,2j} \ (i \text{ even}),$$

where $\phi_{i,2j}$ is a form on $W_{i,2j-1} \oplus W_{i,2j}$, which vanishes identically on each of $W_{i,2j-1}$ and $W_{i,2j}$.

We define the f_{ij} and $\phi_{i,2j}$ explicitly as follows. Choose a basis of W_{ij} of the form $u_{ij}, u_{ij}N, \cdots, u_{ij}N^{i-1}$. We first postulate that

$$(u_{ij}P^{\frac{1}{2}(i-1)}, u_{ij}N^\lambda) = 1 \qquad (i \text{ odd}, 0 \leqq \lambda \leqq i - 1),$$

$$(u_{i,2j-1}P^{\frac{1}{2}i-1}, u_{i,2j}N^\lambda) = 1 \quad (i \text{ even}, 0 \leqq \lambda \leqq i - 1).$$

Then we extend the definition to the whole of W_{ij} or $W_{i,2j-1} \oplus W_{i,2j}$, by requiring that $(xP^{-1}, y) = (x, yP) = (y, x)$ for all x, y in these subspaces. It may be verified that this definition is possible and unambiguous, and that it yields a non-degenerate form f with multiplier P.

Let D_t denote the ring of formal power series $\sum_0^\infty a_\lambda t^\lambda$ over D and M_t the

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

ring of $m' \times m'$ matrices over D_t, where $m' \equiv \sum_1^r m_i$ (the number of indecomposable parts of P). We write the elements of M_t as $r \times r$ block matrices:

$$\Omega = (\Omega_{ij}),$$

where Ω_{ij} is an $m_i \times m_j$ matrix over D_t, and we denote the element in the k-th row and l-th column of Ω_{ij} by $\omega_{ik;jl}$. The notation $\Omega(t^a)$ indicates a matrix whose elements are power series in t^a.

We now show that the elements of $\mathfrak{C} \equiv \mathfrak{C}(P)$ can be represented by certain elements of M_t. Consider the set C of all matrices

$$\Phi \equiv \left(t^{|i-j|}\Phi_{ij}(t^2)\right) = \left(t^{|i-j|}\phi_{ik;jl}(t^2)\right)$$

and the subset K of all matrices

$$\Psi \equiv \left(t^{i+j}\Psi_{ij}(t^2)\right) = \left(t^{i+j}\psi_{ik;jl}(t^2)\right).$$

It is easily seen that C is a D-algebra and K a (D-algebra) ideal of C.

LEMMA 3.1.1. *If $\Phi \in C$, there is a unique element $[\Phi]$ of $\mathfrak{C}$ such that*

$$u_{ik}[\Phi] \equiv \sum_{j,l} u_{jl} N^{\frac{1}{2}(j-i+|j-i|)}\phi_{ik;jl}(N) \quad (1 \leqq i \leqq r;\ 1 \leqq k \leqq m_i).$$

The mapping $\Phi \to [\Phi]$ is a (D-algebra) homomorphism of C onto $\mathfrak{C}$ with kernel K, so that $\mathfrak{C} \cong C/K$.

Let $\omega = \sum a_\lambda t^\lambda \in D_t$, $\Omega \equiv (\omega_{ik;jl}) \in M_t$.
We define

$$\bar{\omega} = \sum a_\lambda \bar{t}^\lambda, \quad \text{where} \quad \bar{t} = \sum_1^\infty t^\lambda,$$

$$\Omega^* \equiv (\theta_{ik;jl}), \quad \text{where} \quad \theta_{ik;jl} \equiv \bar{\omega}_{jl;ik}.$$

Then the mapping $\Omega \to \Omega^*$ is an involutory anti-automorphism of M_t. Let T denote the 2×2 matrix $\begin{pmatrix} 0 & 1+t \\ 1+\bar{t} & 0 \end{pmatrix}$. We define the matrix

$$W \equiv \operatorname{diag}(W_1, \cdots, W_r),$$

where the $m_i \times m_i$ matrix W_i is the unit matrix for odd i and the block diagonal matrix diag $(\underbrace{T, T, \cdots, T}_{(\frac{1}{2}m_i)})$ for even i. Then $W = W^{-1} = W^*$, so that the mapping

$$\Omega \to \Omega^\dagger = W\Omega^* W \tag{3.1.1}$$

is also an involutory anti-automorphism of M_t. It is easily seen that $\dagger$ induces involutory anti-automorphisms of C and K, and thus of C/K.

LEMMA 3.1.2. *If $\Phi \in C$, $[\Phi^\dagger] = [\Phi]^\dagger$.*

In other words, if we identify C/K with $\mathfrak{C}$ according to the isomorphism $\Phi + K \leftrightarrow [\Phi]$, then the mapping (3.1.1) induces the f-adjoint mapping in $\mathfrak{C}$.

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

3.2 *Analogue of theorem* 2.2.1. Although the Approximation Theorem itself is false in the present case, we can prove a weaker result of the same kind. Consider, in $\mathscr{C}$, the set $\mathscr{E}$ of elements

$$X = \bigoplus_{i=1}^{r} X_i,$$

and the subset $\mathscr{F}$ of elements

$$Y = \bigoplus_{i=1}^{r} N^{\lambda_i} Y_i,$$

where

$$\lambda_i = \left.\begin{matrix} 3 & (i \text{ odd}) \\ 2 & (i \text{ even}) \end{matrix}\right\}.$$

Clearly $\mathscr{E}$ is a subalgebra of $\mathscr{C}$ and $\mathscr{F}$ an ideal of $\mathscr{E}$.

LEMMA 3.2.1. *Every non-singular, †-symmetric element Q of $\mathscr{C}$ is †-congruent to an element of $\mathscr{E}$.*

PROOF. Let E_i be the element of $\mathscr{C}$ which maps W_i identically onto itself and every other W_j onto $\{0\}$. Write

$$Q_1 = \sum_i E_i Q E_i, \quad Q_2 = \sum_{i<j} E_i Q E_j, \quad Q_3 = \sum_{i>j} E_i Q E_j.$$

Clearly, $Q_1 \in \mathscr{E}$. Since $Q_2, Q_3 \in \mathscr{N}$, Q_1 is non-singular. Since Q is †-symmetric, Q_1 is †-symmetric and $Q_3 = Q_2^\dagger$.

Assume now that $Q_2 \in \mathscr{N}^k$ $(k \geqq 1)$. We prove the lemma by showing that Q is †-congruent to S, where (in the same notation) $S_2 \in \mathscr{N}^{k+1}$. In fact, let

$$S = (I + Q^{-1}Q_2)^\dagger Q (I + Q^{-1}Q_2) = Q_1 + Q_2^\dagger Q^{-1} Q_2.$$

Then $S_2 = \sum_{i<j} E_i Q_2^\dagger Q^{-1} Q_2 E_j \in (\mathscr{N}^k)^2 \subseteq \mathscr{N}^{k+1}$, as required.

The above proof actually gives the slightly stronger result:

COROLLARY 3.2.1. *Let $\mathscr{G}$ be an ideal of $\mathscr{C}$, Q_1 and Q_2 non-singular, †-symmetric elements of $\mathscr{C}$ such that $Q_1 \equiv Q_2 \pmod{\mathscr{G}}$. Then Q_1, Q_2 are †-congruent to elements R_1, R_2 of $\mathscr{E}$ such that $R_1 \equiv R_2 \pmod{\mathscr{G}}$.*

LEMMA 3.2.2. *Let Q be a †-symmetric element of $N^{2k}\mathscr{F}$ $(k \geqq 0)$. Then Q has the form $R + R^\dagger$, where $R \in N^{2k}\mathscr{F}$.*

PROOF. By the method of proof of lemma 3.2.1 we may suppose f indecomposable. Thus, either (a) $f = f_{i1}$ (i odd) or (b) $f \equiv \phi_{i2}$ (i even).

Case (a) $\mathscr{C}$ is the polynomial algebra $D[P]$ and $N^{2k}\mathscr{F} = N^{2k+3}\mathscr{C}$. Let Σ be the subspace of $\mathscr{C}$ formed by the †-symmetric elements. Write $i = 2e + 1$. Since the matrices P^λ $(-e \leqq \lambda \leqq e)$ form a basis of $\mathscr{C}$, the elements I and $P^\lambda + P^{-\lambda}$ $(1 \leqq \lambda \leqq e)$ obviously span Σ. Hence the elements I and

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

$(P + P^{-1})^\lambda (1 \leqq \lambda \leqq e)$ also span Σ. Since $P + P^{-1} = N^2 + N^3 + \cdots$, the elements $(P + P^{-1})^\lambda$ $(k + 2 \leqq \lambda \leqq e)$ span $\Sigma \cap N^{2k+3}\mathscr{C}$. But $(P + P^{-1})^\lambda = X_\lambda + X_\lambda^\dagger$, where $X_\lambda = N(P + P^{-1})^\lambda$, and $X_\lambda \in N^{2k+3}\mathscr{C}$ when $k + 2 \leqq \lambda \leqq e$. This proves the lemma in case (a).

Case (b) Each $S \in \mathscr{C}$ is determined by the pair of equations

$$u_{i\lambda} S = \sum_{\mu=1}^{2} u_{i\mu} s_{\lambda\mu}(P) \quad (\lambda = 1, 2),$$

and therefore can be represented by the 2×2 matrix $(s_{\lambda\mu}(P))$ over $D[P]$. It can be verified that the lemma is equivalent to the following statement:

(3.2.1) *if $g(P)$ is an element of $N^{2k+2}D[P]$ such that $g(P) = Pg(P^{-1})$, then $g(P) = h(P) + Ph(P^{-1})$ for some $h(P) \in N^{2k+2}D[P]$.*

Let Σ be the subspace of $D[P]$ formed by the elements $g(P)$ such that $g(P) = Pg(P^{-1})$. Let $i = 2e$. Writing $g(P)$ in terms of the basis elements $P^\lambda (-e + 1 \leqq \lambda \leqq e)$, we see that the elements $P^\lambda + P^{1-\lambda}$ span Σ. Therefore Σ consists of the elements of $D[P]$ of the form $h(P) + Ph(P^{-1})$. Hence the elements $N^\lambda + P(N^\dagger)^\lambda = N(N^{\lambda-1} + (N^\dagger)^{\lambda-1})$ span Σ, and so the elements $N(N + N^\dagger)^\lambda$ also span Σ. Since $N + N^\dagger = NN^\dagger = N^2 + N^3 + \cdots$, the elements $N(N + N^\dagger)^\lambda$ $(k + 1 \leqq \lambda \leqq 2e - 1)$ span $\Sigma \cap N^{2k+2}D[P]$. But $Y_\lambda + PY_\lambda^\dagger = N(N + N^\dagger)^\lambda$, where $Y_\lambda = (N + N^\dagger)^\lambda$, and $Y_\lambda \in N^{2k+2}D[P]$ when $k + 1 \leqq \lambda \leqq 2e - 1$. This proves (3.2.1) and the lemma.

COROLLARY 3.2.2. *Let $Q \in \mathscr{E}$. Then $Q + \mathscr{F}$ is the canonical image of a $\dagger$-symmetric element of $\mathscr{E}$ if, and only if, $Q \equiv Q^\dagger \pmod{\mathscr{F}}$.*

PROOF. If $Q \equiv Q^\dagger \pmod{\mathscr{F}}$, then, by lemma 3.2.2, $Q + Q^\dagger = R + R^\dagger$ for some $R \in \mathscr{F}$. Thus $Q + \mathscr{F}$ is the canonical image of the $\dagger$-symmetric element $Q + R$. The converse is obvious.

THEOREM 3.2.1. *Let $\mathscr{G}$ be an ideal of $\mathscr{C}$ such that $\mathscr{G} \cap \mathscr{E} \subseteq \mathscr{F}$. Let Q_1, Q_2 be non-singular, $\dagger$-symmetric elements of $\mathscr{C}$ such that $Q_1 \equiv Q_2 \pmod{\mathscr{G}}$. Then Q_1 is $\dagger$-congruent to Q_2.*

PROOF. By corollary 3.2.1, we may suppose that $Q_1, Q_2 \in \mathscr{E}$ and thus that $Q_1 \equiv Q_2 \pmod{\mathscr{F}}$. Assuming that $Q_1 \equiv Q_2 \pmod{N^{2k}\mathscr{F}}$ $(k \geqq 0)$, we prove the theorem by showing that $X^\dagger Q_1 X \equiv Q_2 \pmod{N^{2k+2}\mathscr{F}}$ for some $X \in \mathscr{E}$.

By lemma 3.2.2, $Q_1 + Q_2 = L + L^\dagger$, where $L \in N^{2k}\mathscr{F}$. Let $X = I + Q_1^{-1}L$. Since * $L^\dagger Q_1^{-1} L \in N^{4k}\mathscr{F}^2 \subseteq N^{2k+2}\mathscr{F}$, we have $X^\dagger Q_1 X = Q_2 + L^\dagger Q_1^{-1} L \equiv Q_2 \pmod{N^{2k+2}\mathscr{F}}$, as required.

* Up to this point, we could have taken $\lambda_i = 1$ (i odd) in the definition of $\mathscr{F}$. But here we definitely require $\lambda_i \geqq 2$ (all i) in order that $N^{4k}\mathscr{F}^2 \subseteq N^{2k+2}\mathscr{F}$ when $k = 0$. Since lemma 3.2.2 is only valid when $\lambda_i \equiv i \pmod 2$ (all i), we are led to the values $\lambda_i = 3$ (i odd), $\lambda_i = 2$ (i even) adopted in the definition.

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

The results of this section reduce the problem of †-congruence to the following one:

(3.2.2) *given non-singular elements Q, R of $\mathscr{E}$ such that $Q \equiv Q^\dagger$ and $R \equiv R^\dagger$ (mod $\mathscr{F}$), to determine the conditions that $X^\dagger Q X \equiv R$ (mod $\mathscr{G}$) for some $X \in \mathscr{C}$.*

We may assume, where convenient, that Q, R are actually †-symmetric, so that the forms

$$g(x, y) = (x, yQ), \quad h(x, y) = (x, yR) \tag{3.2.3}$$

have multiplier P. Then, since Q, $R \in \mathscr{E}$, g, h have direct decompositions

$$g = g_1 \oplus \cdots \oplus g_r, \quad h = h_1 \oplus \cdots \oplus h_r. \tag{3.2.4}$$

If $H = (i_1, \cdots, i_s)$ is a subset of $(1, \cdots r)$ we write

$$g_H = g_{i_1} \oplus \cdots \oplus g_{i_s}, \quad h_H = h_{i_1} \oplus \cdots \oplus h_{i_s}. \tag{3.2.5}$$

3.3 *Detailed equations of congruence.* We consider the problem (3.2.2) in terms of the matrix algebra C. Let E, F, G denote the subrings of C which correspond to $\mathscr{E}$, $\mathscr{F}$, $\mathscr{G}$. Thus E consists of the matrices diag $(\boldsymbol{X}_1, \cdots \boldsymbol{X}_r) \in C$, and F of the matrices in E such that* $\boldsymbol{X}_1 \equiv \boldsymbol{0}$ (mod t^2), $\boldsymbol{X}_i \equiv \boldsymbol{0}$ (mod t^6) (i odd and $\geqq 3$), $\boldsymbol{X}_i \equiv \boldsymbol{0}$ (mod t^4) (i even). We define G as the set of matrices $(t^{|i-j|}\boldsymbol{\Phi}_{ij}(t^2))$ such that

$$\begin{aligned} &\text{diag } (\boldsymbol{\Phi}_{11}, \cdots \boldsymbol{\Phi}_{rr}) \in F, \\ &\boldsymbol{\Phi}_{i,j} \equiv \boldsymbol{0} \pmod{t^2} \quad \text{when} \quad |i - j| = 1 \text{ (all } i\text{)} \\ &\boldsymbol{\Phi}_{i,j} \equiv \boldsymbol{0} \pmod{t^2} \quad \text{when} \quad |i - j| = 2 \text{ (all odd } i\text{)}. \end{aligned}$$

It is easily verified that G is an ideal of C and that $K \subset G$, $G \cap E = F$. Thus, the corresponding ideal $\mathscr{G}$ in $\mathscr{C}$ satisfies $\mathscr{G} \cap \mathscr{E} \subseteq \mathscr{F}$ as required in theorem 3.2.1.

Choose matrices $\boldsymbol{Q}$, $\boldsymbol{R} \in C$ which represent the elements Q, R in (3.2.2). Write

$$\begin{aligned} \boldsymbol{W}\boldsymbol{Q} &= \text{diag } (\boldsymbol{Q}_1, \cdots, \boldsymbol{Q}_r) \\ \boldsymbol{W}\boldsymbol{R} &= \text{diag } (\boldsymbol{R}_1, \cdots, \boldsymbol{R}_r). \end{aligned} \tag{3.3.1}$$

Since Q is non-singular and $Q \equiv Q^\dagger$ (mod $\mathscr{F}$), we have

$$\begin{aligned} &\boldsymbol{Q}_1 \equiv \boldsymbol{S}_1 \pmod{t^2} \\ &\boldsymbol{Q}_i \equiv \boldsymbol{S}_i + (\boldsymbol{P}_i + \boldsymbol{P}_i^T)t^2 + \boldsymbol{P}_i t^4 \pmod{t^6} \quad (i \text{ odd and} \geqq 3), \end{aligned} \tag{3.3.2}$$

where the $\boldsymbol{S}_i$, $\boldsymbol{P}_i$ are matrices over D and the $\boldsymbol{S}_i$ non-singular and symmetric;

* A matrix congruence $(a_{ij}) \equiv (b_{ij})$ (mod t^k) means that $a_{ij} \equiv b_{ij}$ (mod t^k) for all i, j.

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

and

$$(3.3.3) \qquad (1+t)\boldsymbol{Q}_i \equiv \boldsymbol{P}_i + \boldsymbol{P}_i^T + \boldsymbol{P}_i t^2 \pmod{t^4} \quad (i \text{ even})$$

where the $\boldsymbol{P}_i$ are matrices over D such that $\boldsymbol{P}_i + \boldsymbol{P}_i^T$ is non-singular.

The form of $\boldsymbol{Q}_i$ can be simplified by a preliminary transformation $\boldsymbol{Q} \to \boldsymbol{X}^\dagger \boldsymbol{Q}\boldsymbol{X}$, where $\boldsymbol{X}$ is an element of E in which only the i-th diagonal block, $\boldsymbol{X}_i$, differs from the unit matrix. Let the i-th diagonal block of $\boldsymbol{W}\boldsymbol{X}^\dagger \boldsymbol{Q}\boldsymbol{X}$ be $\tilde{\boldsymbol{Q}}_i$. Taking $\boldsymbol{X}_i = \boldsymbol{M}_i$, $\boldsymbol{I} + \boldsymbol{Y}_i t^2$, $\boldsymbol{I} + \boldsymbol{Z}_i t^4$ in turn, where $\boldsymbol{M}_i$, $\boldsymbol{Y}_i$, $\boldsymbol{Z}_i$ are matrices over D, we get, when i is odd,

$$\tilde{\boldsymbol{Q}}_i = \left.\begin{array}{l} \boldsymbol{M}_i^T \boldsymbol{S}_i \boldsymbol{M}_i + \cdots \\ \boldsymbol{S}_i + ((\boldsymbol{Y}_i^T \boldsymbol{S}_i) + (\boldsymbol{Y}_i^T \boldsymbol{S}_i)^T + \boldsymbol{P}_i + \boldsymbol{P}_i^T)t^2 + \cdots \\ \boldsymbol{S}_i + (\boldsymbol{P}_i + \boldsymbol{P}_i^T)t^2 + ((\boldsymbol{Z}_i^T \boldsymbol{S}_i) + (\boldsymbol{Z}_i^T \boldsymbol{S}_i)^T + \boldsymbol{P}_i)t^4 + \cdots \end{array}\right\}$$

and, when i is even,

$$(1+t)\tilde{\boldsymbol{Q}}_i = \left.\begin{array}{l} \boldsymbol{M}_i^T (\boldsymbol{P}_i + \boldsymbol{P}_i^T)\boldsymbol{M}_i + \cdots \\ (\boldsymbol{P}_i + \boldsymbol{P}_i^T) + (\boldsymbol{Y}_i^T (\boldsymbol{P}_i + \boldsymbol{P}_i^T) + (\boldsymbol{Y}_i^T (\boldsymbol{P}_i + \boldsymbol{P}_i^T))^T + \boldsymbol{P}_i)t^2 + \cdots \end{array}\right\}.$$

When m_i is even, let $\boldsymbol{J}_i$ denote the "canonical" alternate * matrix

$$\boldsymbol{J}_i = \operatorname{diag}\Big(\underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}_{(\frac{1}{2}m_i)}\Big).$$

A non-singular symmetric matrix over a field of characteristic 2 is congruent to the canonical matrix if it is alternate and to a diagonal matrix otherwise (Jacobson [5]). We may therefore suppose that, in (3.3.2),

$$(3.3.4) \qquad \text{either } \boldsymbol{S}_i = \boldsymbol{J}_i \quad \text{or} \quad \boldsymbol{S}_i = \operatorname{diag}(s_1^i, \cdots s_{m_i}^i),$$

and that, in (3.3.3),

$$(3.3.5) \qquad \boldsymbol{P}_i + \boldsymbol{P}_i^T = \boldsymbol{J}_i.$$

By the second and third equations for $\tilde{\boldsymbol{Q}}_i$, we may further suppose that, in (3.3.2),

$$(3.3.6) \qquad \boldsymbol{P}_i = \operatorname{diag}\ (p_1^i, \cdots p_{m_i}^i),$$

and that, in (3.3.3),

$$(3.3.7) \qquad \boldsymbol{P}_i = \operatorname{diag}\left(\begin{pmatrix} p_1^i & 1 \\ 0 & p_2^i \end{pmatrix}, \cdots \begin{pmatrix} p_{m_i-1}^i & 1 \\ 0 & p_{m_i}^i \end{pmatrix}\right).$$

Thus

$$(3.3.2)' \qquad \begin{array}{l} \boldsymbol{Q}_1 \equiv \boldsymbol{S}_1 \pmod{t^2}, \\ \boldsymbol{Q}_i \equiv \boldsymbol{S}_i + \boldsymbol{P}_i t^4 \pmod{t^6} \quad (i \text{ odd and } \geqq 3), \end{array}$$

* An alternate matrix is one of the form $\boldsymbol{M} + \boldsymbol{M}^T$.

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

(3.3.3)′ $$(1+t)\boldsymbol{Q}_i \equiv \boldsymbol{J}_1 + \boldsymbol{P}_i t^2 \pmod{t^4} \quad (i \text{ even}),$$

where (3.3.4)—(3.3.7) hold. Similarly, we may suppose that

(3.3.2)″ $$\boldsymbol{R}_1 \equiv \boldsymbol{\Sigma}_1 \pmod{t^2},$$
$$\boldsymbol{R}_i \equiv \boldsymbol{\Sigma}_i + \boldsymbol{\Pi}_i t^4 \pmod{t^6} \quad (i \text{ odd and } \geqq 3);$$

(3.3.3)″ $$\boldsymbol{R}_i \equiv \boldsymbol{J}_i + \boldsymbol{\Pi}_i t^2 \pmod{t^4} \quad (i \text{ even}).$$

Let now $\boldsymbol{X} = (t^{|i-j|}\boldsymbol{U}_{ij}(t^2))$, where

(3.3.8) $$\boldsymbol{U}_{ii} \equiv \boldsymbol{X}_i + \boldsymbol{Y}_i t^2 + \boldsymbol{Z}_i t^4 \pmod{t^6},$$
$$\boldsymbol{U}_{ij} \equiv \boldsymbol{X}_{ij} \pmod{t^2} \ (i \neq j),$$

the $\boldsymbol{X}_i$, $\boldsymbol{Y}_i$, $\boldsymbol{Z}_i$, $\boldsymbol{X}_{ij}$ being matrices over D. Write

(3.3.9) $$\boldsymbol{K}_i = \boldsymbol{Y}_i^T \boldsymbol{S}_i \boldsymbol{X}_i \quad (i \text{ odd and } \geqq 3).$$

Then the equation $\boldsymbol{X}^\dagger \boldsymbol{Q}\boldsymbol{X} \equiv \boldsymbol{R} \pmod{G}$, i.e. $\boldsymbol{X}^* \boldsymbol{W}\boldsymbol{Q}\boldsymbol{X} \equiv \boldsymbol{W}\boldsymbol{R} \pmod{\boldsymbol{W}G}$, yields the following system of equations:

I (λ *odd*).

(3.3.10) $$\boldsymbol{\Sigma}_\lambda \equiv \boldsymbol{X}_\lambda^T \boldsymbol{S}_\lambda \boldsymbol{X}_\lambda$$

(3.3.11) $$\boldsymbol{K}_\lambda + \sum_{k=\pm 1} \boldsymbol{X}_{\lambda+k,\lambda}^T \boldsymbol{P}_{\lambda+k} \boldsymbol{X}_{\lambda+k,\lambda} \text{ is symmetric } (\lambda \geqq 3).$$

(3.3.12) $$\left.\begin{aligned} \boldsymbol{\Pi}_\lambda \approx{}& \boldsymbol{X}_\lambda^T \boldsymbol{P}_\lambda \boldsymbol{X}_\lambda + \boldsymbol{K}_\lambda + \boldsymbol{K}_\lambda \boldsymbol{\Sigma}_\lambda^{-1} \boldsymbol{K}_\lambda^T \\ &+ \sum_{k=\pm 2} \boldsymbol{X}_{\lambda+k,\lambda}^T \boldsymbol{S}_{\lambda+k} \boldsymbol{X}_{\lambda+k,\lambda} + \sum_{k=\pm 1} \boldsymbol{X}_{\lambda+k,\lambda}^T \boldsymbol{P}_{\lambda+k} \boldsymbol{X}_{\lambda+k,\lambda} \end{aligned}\right\} \ (\lambda \geqq 3).$$

(3.3.13) $$\boldsymbol{0} = \boldsymbol{X}_\lambda^T \boldsymbol{S}_\lambda \boldsymbol{X}_{\lambda,\lambda+1} + \boldsymbol{X}_{\lambda+1,\lambda}^T \boldsymbol{J}_{\lambda+1} \boldsymbol{X}_{\lambda+1} \quad (\lambda \geqq 3).$$

(3.3.14) $$\boldsymbol{0} = \boldsymbol{X}_\lambda^T \boldsymbol{S}_\lambda \boldsymbol{X}_{\lambda,\lambda+2} + \boldsymbol{X}_{\lambda+1,\lambda}^T \boldsymbol{J}_{\lambda+1} \boldsymbol{X}_{\lambda+1,\lambda+2} + \boldsymbol{X}_{\lambda+2,\lambda}^T \boldsymbol{S}_{\lambda+2} \boldsymbol{X}_{\lambda+2}. \quad (\lambda \geqq 3)$$

II (μ *even*).

(3.3.15) $$\boldsymbol{J}_\mu = \boldsymbol{X}_\mu^T \boldsymbol{J}_\mu \boldsymbol{X}_\mu.$$

(3.3.16) $$\boldsymbol{\Pi}_\mu \approx \boldsymbol{X}_\mu^T \boldsymbol{P}_\mu \boldsymbol{X}_\mu + \sum_{k=\pm 1} \boldsymbol{X}_{\mu+k,\mu}^T \boldsymbol{S}_{\mu+k} \boldsymbol{X}_{\mu+k,\mu}.$$

(3.3.17) $$\boldsymbol{0} = \boldsymbol{X}_\mu^T \boldsymbol{J}_\mu \boldsymbol{X}_{\mu,\mu+1} + \boldsymbol{X}_{\mu+1,\mu}^T \boldsymbol{S}_{\mu+1} \boldsymbol{X}_{\mu+1}.$$

Remarks. (1) The symbol $\approx$ in (3.3.12) and (3.3.16) means that the two sides of the equation are equal apart from terms of the form $\boldsymbol{M} + \boldsymbol{M}^T$. Taking into account the form of these equations and (3.3.15), this simply means that corresponding diagonal elements on both sides are to be equated.

(2) When $\lambda = 1$, (3.3.13) and (3.3.14) do actually form part of the equation system. But we can omit them because they serve only to define $\boldsymbol{X}_{21}$ and $\boldsymbol{X}_{31}$ which do not appear in any other equations.

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

(3) The complete set of equations (3.3.10)—(3.3.17) (for all λ, μ) will be denoted by Θ. Let H be a subset of the indices $1, \cdots r$. Then the system of equations (3.3.10)—(3.3.17) (for all $\lambda, \mu \in H$), with the omission of all terms involving indices not in H, will be denoted by Θ_H. Clearly, Θ_H expresses the equivalence $g_H \approx h_H$ of the forms in (3.2.5).

3.4 *Case of a single elementary divisor.* With g, h as in (3.2.3), we study the equivalence $g_i \approx h_i$ of individual summands in (3.2.4). (It should be emphasized, however, that the equivalence $g \approx h$ does not in general reduce to the system $g_i \approx h_i$ $(i = 1, \cdots r)$.) We always assume that (3.3.4)—(3.3.7) hold.

Consider first an even index μ. We associate with g_μ the non-defective quadratic form

$$p_\mu(x) = \sum_{i=1}^{\frac{1}{2}m_\mu} (p^\mu_{2i-1}x^2_{2i-1} + x_{2i-1}x_{2i} + p^\mu_{2i}x^2_{2i}), \tag{3.4.1}$$

in the vector $x = (x_1, \cdots, x_{m_\mu})$. Then we have

LEMMA 3.4.1. *Let μ be an even index and let $p_\mu(x)$, $\pi_\mu(x)$ be the quadratic forms associated with g_μ, h_μ. Then $g_\mu \approx h_\mu$ if, and only if, $p_\mu(x)$ is equivalent to $\pi_\mu(x)$.*

PROOF. The system Θ_μ consists of the two equations $\boldsymbol{J}_\mu = \boldsymbol{X}_\mu^T \boldsymbol{J}_\mu \boldsymbol{X}_\mu$ and $\boldsymbol{\Pi}_\mu \approx \boldsymbol{X}_\mu^T \boldsymbol{P}_\mu \boldsymbol{X}_\mu$. It is easily seen that these express the equivalence of $p_\mu(x)$ and $\pi_\mu(x)$.

The situation for an odd index λ is more complicated and we require some preliminaries on bilinear forms. Consider a non-degenerate symmetric bilinear form $\phi = [x, y]$ on the m-dimensional space W. The mapping $\theta : x \to [x, x]$ is an additive homomorphism of W into D. Let $\mathscr{W} = \theta(W)$, $W_0 = \theta^{-1}(0)$. Since $[\lambda x, \lambda x] = \lambda^2[x, x]$,

(A) *$\mathscr{W}$ is a D^2-subspace of D of dimension $k \leqq m$.*

Let U denote the perpendicular space $W_0^\perp$ of W_0 with respect to ϕ and write $\mathscr{U} = \theta(U)$. Let l be the dimension of $\mathscr{U}$ as D^2-space. Then

$$l = \dim U - \dim (U \cap W_0) = m - \dim W_0 - \dim (U \cap W_0)$$

and, since the restriction of ϕ to W_0 is alternate,

$$\dim W_0 - \dim (U \cap W_0) \equiv 0 \pmod 2,$$

whence

$$l \equiv m \pmod 2. \tag{3.4.2}$$

Thus,

(B) *$\mathscr{U}$ is a D^2-subspace of $\mathscr{W}$ whose dimension l satisfies* (3.4.2).

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

Let x_1, x_2, y_1, y_2 be elements of U such that $[x_1, x_1] = [x_2, x_2]$ and $[y_1, y_1] = [y_2, y_2]$. Then $x_1 - x_2, y_1 - y_2 \in U \cap W_0$ so that $[x_1, y_1] = [x_2, y_2]$. Therefore the equations

$$\langle [x, x], [y, y] \rangle = [x, y] \quad (x, y \in U)$$

uniquely define a function $\gamma = \langle \lambda, \mu \rangle$ of the variables $\lambda, \mu \in \mathscr{U}$. It is easily verified that

$$\text{(C)} \qquad \left.\begin{array}{l} \langle \lambda, \mu \rangle \in D \\ \langle \lambda, \mu \rangle = \langle \mu, \lambda \rangle \\ \langle \lambda + \mu, \nu \rangle = \langle \lambda, \nu \rangle + \langle \mu, \nu \rangle \\ \langle \lambda \alpha^2, \mu \rangle = \alpha \langle \lambda, \mu \rangle \\ \langle \lambda, \lambda \rangle = \lambda \end{array}\right\}$$

whenever $\lambda, \mu, \nu \in \mathscr{U}$, $\alpha \in D$. One constructs a function γ satisfying (C) by choosing a D^2-basis $s_1, \cdots, s_l$ of $\mathscr{U}$ and defining

$$\text{(3.4.3)} \qquad \langle \sum \alpha_i^2 s_i, \sum \beta_i^2 s_i \rangle = \sum \rho_{ij} \alpha_i \beta_j,$$

where $\rho_{ij} = \rho_{ji} \in D$ and $\rho_{ii} = s_i$ for all i, j. Conversely, every γ is obtained in this way. We call $\mathscr{W}$, $\mathscr{U}$, γ the *invariants* of ϕ.

LEMMA 3.4.2. *If $\mathscr{W}$, $\mathscr{U}$, γ satisfy* (A)—(C), *they are the invariants of a non-degenerate symmetric bilinear form on W. Two such forms are equivalent if, and only if, they have the same invariants.*

PROOF. Choose a fixed D^2-basis $s_1, \cdots, s_k$ of $\mathscr{W}$ such that $s_1, \cdots, s_l$ is a basis of $\mathscr{U}$. Let

$\boldsymbol{A}$ be the $l \times l$ matrix $(\langle s_i, s_j \rangle)$ $(i, j = 1, \cdots l)$,
$\boldsymbol{B}$ be the $(k - l) \times (k - l)$ matrix diag $(s_{l+1}, \cdots, s_k)$,
$\boldsymbol{J}$ be the $(m - 2k + l) \times (m - 2k + l)$ canonical alternate matrix,
$\boldsymbol{I}$ be the $(k - l) \times (k - l)$ unit matrix,
$\boldsymbol{T}$ be the $m \times m$ matrix

$$\text{(3.4.4)} \qquad \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{I} \\ \mathbf{0} & \boldsymbol{A} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{J} & \mathbf{0} \\ \boldsymbol{I} & \mathbf{0} & \mathbf{0} & \boldsymbol{B} \end{pmatrix}.$$

It is easily verified that a bilinear form with matrix $\boldsymbol{T}$ is symmetric and non-degenerate, and that it has the invariants $\mathscr{W}$, $\mathscr{U}$, γ.

Suppose now that ϕ is any form with these invariants. We prove the last part of the lemma by showing that ϕ has matrix $\boldsymbol{T}$ with respect to a suitable basis. Let $d_1', \cdots, d_{k-l}'$ be elements of W such that $\theta(d_i') = s_{l+i}$ $(1 \leqq i \leqq k-l)$. Clearly, the d_1^i are linearly independent and span a complement of $U + W_0$ in W. Since $U + W_0 = (U \cap W_0)^\perp$, there exist elements $a_1, \cdots, a_{k-l}$ of

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

$U \cap W_0$ such that $[d_i', a_j] = \delta_{ij}$ (Kronecker delta) and there also exist complements U', W_0' of $U \cap W_0$ in U, W_0 which are perpendicular to $d_1', \cdots, d_{k-l}'$. Let $d_i = d_i' + \sum_{j>i}[d_i', d_j']a_j$. Then $\{d_1, \cdots d_{k-l}\}$ is a complement of $U + W_0$ which is perpendicular to U', W_0' and $[d_i, d_j] = \delta_{ij}s_i$, $[d_i, a_j] = \delta_{ij}$. The (unique) elements $b_1, \cdots, b_l$ in U' such that $\theta(b_i) = s_i$ clearly form a basis of U'. Since the restriction of ϕ to W_0 is alternate and $U \cap W_0 = W_0^{\perp} \cap W_0$, the restriction ψ of ϕ to W_0' is non-degenerate and alternate. Hence there is a basis $c_1, \cdots, c_{m-2k+l}$ with respect to which ψ has matrix $\boldsymbol{J}$. It is now easily verified that ϕ has matrix $\boldsymbol{T}$ with respect to the basis $a_1, \cdots, b_1, \cdots, c_1, \cdots, d_1, \cdots$ of W. This completes the proof.

By § 3.3, $g_\lambda \approx h_\lambda$ (λ odd) if, and only if,

$$\boldsymbol{\Sigma}_\lambda = \boldsymbol{X}_\lambda^T \boldsymbol{S}_\lambda \boldsymbol{X}_\lambda \tag{3.4.5}$$

$$\boldsymbol{\Pi}_\lambda \approx \boldsymbol{X}_\lambda^T \boldsymbol{P}_\lambda \boldsymbol{X}_\lambda + \boldsymbol{K}_\lambda \boldsymbol{\Sigma}_\lambda^{-1} \boldsymbol{K}_\lambda + \boldsymbol{K}_\lambda \tag{3.4.6}$$

for a non-singular matrix $\boldsymbol{X}_\lambda$ and symmetric matrix $\boldsymbol{K}_\lambda$. (3.4.5) states that the non-degenerate symmetric bilinear forms with matrices $\boldsymbol{\Sigma}_\lambda$, $\boldsymbol{S}_\lambda$ are equivalent. Let us suppose that this condition holds. For convenience of notation, we take W to be equal to its summand W_λ. Then we may suppose that $\boldsymbol{\Sigma}_\lambda \equiv \boldsymbol{S}_\lambda \equiv \boldsymbol{T}$ (see (3.4.4)). Let ϕ be the form on W with matrix $\boldsymbol{T}$, X_λ the linear transformation on W with matrix $\boldsymbol{X}_\lambda$. By (3.4.5), X_λ leaves ϕ invariant. Therefore X_λ leaves U, W_0, $U \cap W_0$ invariant and so $\boldsymbol{X}_\lambda$ has the form

$$\begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \mathbf{0} & \cdot & \mathbf{0} & \cdot \\ \mathbf{0} & \mathbf{0} & \cdot & \cdot \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdot \end{pmatrix}.$$

It is also easy to see (e.g. by considering the corresponding quadratic forms) that the matrix equation

$$\boldsymbol{Y}^T \begin{pmatrix} \boldsymbol{A} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{B} \end{pmatrix} \boldsymbol{Y} \approx \begin{pmatrix} \boldsymbol{A} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{B} \end{pmatrix}$$

implies that $\boldsymbol{Y} = \boldsymbol{I}$. Using these facts, we get

$$\boldsymbol{X}_\lambda = \begin{pmatrix} \boldsymbol{I} & \mathbf{0} & \boldsymbol{X} & \boldsymbol{Y} \\ \mathbf{0} & \boldsymbol{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{Z} & \boldsymbol{W} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{I} \end{pmatrix}$$

where

$$\boldsymbol{Z}^T \boldsymbol{J} \boldsymbol{Z} = \boldsymbol{J}, \quad \boldsymbol{X} = \boldsymbol{W}^T \boldsymbol{J} \boldsymbol{Z}, \quad \boldsymbol{Y} + \boldsymbol{Y}^T = \boldsymbol{W}^T \boldsymbol{J} \boldsymbol{W}.$$

We may take

$$\boldsymbol{P}_\lambda = \operatorname{diag}(\boldsymbol{P}_1^\lambda, \boldsymbol{P}_2^\lambda, \boldsymbol{P}_3^\lambda, \boldsymbol{P}_4^\lambda), \quad \boldsymbol{\Pi}_\lambda = \operatorname{diag}(\boldsymbol{\Pi}_1^\lambda, \boldsymbol{\Pi}_2^\lambda, \boldsymbol{\Pi}_3^\lambda, \boldsymbol{\Pi}_4^\lambda),$$
$$\boldsymbol{K}_\lambda = (\boldsymbol{K}_{ij}) \quad (i, j = 1, \cdots, 4),$$

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

where the P_i^λ, Π_i^λ are diagonal and $K_{ij}^T = K_{ji}$. Then (3.4.6) gives

$$\begin{aligned} P_1^\lambda + \Pi_1^\lambda &\approx K_{11} + K_{11}BK_{11} + K_{12}A^{-1}K_{12}^T \\ P_2^\lambda + \Pi_2^\lambda &\approx K_{22} + K_{12}^T BK_{12} + K_{22}A^{-1}K_{22} \end{aligned} \tag{3.4.7}$$

and two other equations which may be omitted since they serve merely as "defining equations" for K_{33}, K_{44}. (3.4.7) can be written as a single matrix equation

$$\begin{pmatrix} \Pi_1^\lambda & 0 \\ 0 & \Pi_2^\lambda \end{pmatrix} \approx \begin{pmatrix} P_1^\lambda & 0 \\ 0 & P_2^\lambda \end{pmatrix} + K\begin{pmatrix} B & 0 \\ 0 & A^{-1} \end{pmatrix} K + K \tag{3.4.8}$$

where K is a symmetric matrix.

It follows that we can "normalize" P_λ, Π_λ by taking

$$P_\lambda = \operatorname{diag}(P_1^\lambda, P_2^\lambda, 0, 0), \quad \Pi_\lambda = \operatorname{diag}(\Pi_1^\lambda, \Pi_2^\lambda, 0, 0), \tag{3.4.9}$$

where the P_i^λ, Π_i^λ are diagonal. Then (with $\Sigma_\lambda = S_\lambda = T$) $g_\lambda \approx h_\lambda$ if, and only if, (3.4.8) has a symmetric solution K.

3.5 *A reduction theorem.* With g, h as before, we prove now that, under suitable conditions, the equivalence $g \approx h$ reduces to certain equivalences $g_H \approx h_H$ (cf. (3.2.5)). We also obtain an important necessary condition for equivalence in the general case.

THEOREM 3.5.1. *With g, h as in* (3.2.3), (3.3.4)—(3.3.7), *let i be an odd index such that either $m_i = 0$ or $S_i = \Sigma_i = J_i$. Then $g \approx h$ if, and only if, $g_L \approx h_L$ and $g_M \approx h_M$, where $L = (1, \cdots i-1)$, $M = (i+1, \cdots r)$.*

PROOF. The system of equations Θ_L forms part of the complete system Θ, except that the equations in Θ_L corresponding to (3.3.12) for $\lambda = i-2$ and (3.3.16) for $\mu = i-1$ lack the terms $X_{i,i-2}^T S_i X_{i,i-2}$ and $X_{i,i-1}^T S_i X_{i,i-1}$. But these terms are immaterial because, under our hypotheses, they have the form $M + M^T$. Hence Θ implies Θ_L and similarly Θ implies Θ_M.

Conversely, let $g_L \approx h_L$, $g_M \approx h_M$. It is obviously sufficient to prove that $g_i \approx h_i$ when $\Sigma_i = S_i = J_i$. But in this case $X_i = I$, $K_i = \Pi_i + P_i$ is a solution of the corresponding system Θ_i. This proves the theorem.

COROLLARY 3.5.1. *Suppose that, for each odd index λ, either $m_\lambda = 0$ or $S_\lambda = \Sigma_\lambda = J_\lambda$. Then $g \approx h$ if, and only if, the quadratic forms $p_\mu(x)$, $\pi_\mu(x)$ in* (3.4.2) *are equivalent for all even indices μ.*

Let

$$(1+t)^{e_1}, \cdots, (1+t)^{e_s} \quad (0 < e_1 < \cdots < e_s) \tag{3.5.2}$$

be the elementary divisors of P of positive multiplicity. With theorem 3.5.1 in mind, we now study the situation where

(3.5.3) *every odd integer k such that $e_1 \leqq k \leqq e_s$ is an e_i, and $S_k \not\approx 0$.*

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

Let us also suppose for the moment that $e_1 > 1$ (i.e. f has no symmetric bilinear form as direct summand). Suppose that $g \approx h$ and, as before, let $\boldsymbol{Q}, \boldsymbol{R}$ be matrices in C which represent Q, R. By slightly modifying the proof of theorem 3.2.1, it is easy to show that $\boldsymbol{Y}^\dagger \boldsymbol{Q} \boldsymbol{Y} \equiv \boldsymbol{R} \pmod{t^6 C}$ for some $\boldsymbol{Y} \in C$. We take determinants of both sides of the congruence, or more conveniently of the congruence $\boldsymbol{Y}^* \boldsymbol{W} \boldsymbol{Q} \boldsymbol{Y} \equiv \boldsymbol{W} \boldsymbol{R} \pmod{t^6 \boldsymbol{W} C}$. It is easy to see that $|\boldsymbol{Y}|$ has the form $\lambda(1 + \mu t^2 + \nu t^4) \pmod{t^6}$ and that therefore

$$\begin{aligned} |\boldsymbol{Y}||\boldsymbol{Y}^*| &\equiv \lambda^2(1 + \mu t^2 + \nu t^4)(1 + \mu t^2 + \nu t^4) \\ &\equiv \lambda^2(1 + (\mu^2 + \mu) t^4) \pmod{t^6}. \end{aligned}$$

Also

$$|\boldsymbol{W}\boldsymbol{Q}| \equiv \alpha(1 + \beta t^4) \pmod{t^6}, \quad |\boldsymbol{W}\boldsymbol{R}| \equiv \alpha_1(1 + \beta_1 t^4) \pmod{t^6},$$

where

$$\left.\begin{aligned} \alpha &= \prod_{\lambda \text{ odd}} \prod_{i=1}^{m_\lambda} s_i^\lambda \\ \beta &= \sum_{\lambda \text{ odd}} \sum_{i=1}^{m_\lambda} p_i^\lambda / s_i^\lambda + \sum_{\mu \text{ even}} \sum_{j=1}^{\frac{1}{2} m_\mu} p_{2j-1}^\mu p_{2j}^\mu \end{aligned}\right\} \tag{3.5.4}$$

and similarly for α_1, β_1.

Equating determinants, we get

$$\alpha_1 = \alpha\lambda^2, \beta_1 = \beta + \rho^2 + \rho. \tag{3.5.5}$$

We call α the *first*, β the *second, discriminant* of g. We say that the second discriminants of g, h are *essentially equal* when they are related as in (3.5.5), and we write $\beta_1 \sim \beta$. It is evident that the first discriminant is defined even when $e_1 = 1$. We assign the value 0 to the second discriminant in this case.

Let us now return to the general case where (3.5.3) is not assumed. We say that there is a *gap between the indices e_i and e_{i+1} with respect to g* when at least one of the following conditions holds:

(a) $e_i < k < e_{i+1}$ for some odd k;

(b) e_i is odd and $\boldsymbol{S}_{e_i} \approx \boldsymbol{0}$;

(c) e_{i+1} is odd and $\boldsymbol{S}_{e_{i+1}} \approx \boldsymbol{0}$.

The gaps divide the indices $e_1, \cdots e_r$ into *component sets with respect to* g:

$$L_1 = (e_1, \cdots e_{s_1}), \quad L_2 = (e_{s_1+1}, \cdots e_{s_1+s_2}), \cdots. \tag{3.5.5}'$$

The second discriminants of $g_{L_1}, g_{L_2}, \cdots$ are called the *partial second discriminants* of g. (If L_i consists of a single odd index λ and $\boldsymbol{S}_\lambda \approx \boldsymbol{0}$, we define the corresponding second partial discriminant to be 0.) By (3.5.5) and theorem 3.5.1, we have

THEOREM 3.5.2. *With g, h as in* (3.2.3), (3.3.4)—(3.3.7), *let* $g \approx h$.

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

Then g, h have the same component sets and their corresponding partial second discriminants are essentially equal.

3.6 *Complete solution in a special case.*

LEMMA 3.6.1. *Let a, $b(\neq 0) \in D$. If D satisfies* (3.0.1), *the equation*

$$x^2 + x + a + by^2 = 0 \tag{3.6.1}$$

has a solution (x, y).

PROOF. By (3.0.1), the equation $\xi^2 + \xi\eta + a\eta^2 + b\zeta^2 = 0$ has solutions $(\xi, \eta, \zeta) \neq (0, 0, 0)$. If $\eta \neq 0$, $(x, y) = (\xi/\eta, \zeta/\eta)$ satisfies (3.6.1). If $\eta = 0$, $(x, y) = (a, \zeta a/\xi)$ satisfies (3.6.1). Q.E.D.

Consider a non-defective quadratic form $q(\boldsymbol{x}) = \sum_{i,j=1}^{2u} a_{ij}x_ix_j$ over D. In terms of suitable new variables $\boldsymbol{y}$,

$$q(\boldsymbol{x}) = \sum_{i=1}^{u} (b_{2i-1}y_{2i-1}^2 + y_{2i-1}y_{2i} + b_{2i}y_{2i}^2).$$

The *pseudo-discriminant* of $q(\boldsymbol{x})$ is defined to be $\sum_{i=1}^{u} b_{2i-1}b_{2i}$. It is uniquely determined up to essential equality, i.e. apart from additive terms $\lambda^2 + \lambda$ ($\lambda \in D$). (Cf. Dieudonné [2]).

LEMMA 3.6.2. *Suppose that D satisfies* (3.0.1). *Then two non-defective quadratic forms $q_1(\boldsymbol{x})$, $q_2(\boldsymbol{x})$ over D are equivalent if, and only if, their pseudo-discriminants are essentially equal.*

PROOF. The result holds for binary forms over *any* field of characteristic 2. By (3.0.1), a non-defective form in $2u > 2$ variables is indefinite. By Cahit Arf's theorem *, it is equivalent to a form $\phi = x_1x_2 + \omega(x_3, \cdots x_{2u})$. If $\psi = x_1x_2 + \chi(x_3, \cdots x_{2u})$ is another such form, then (again by Cahit Arf's theorem) ϕ is equivalent to ψ if, and only if, ω is equivalent to χ. The lemma now follows obviously by induction.

THEOREM 3.6.1. *Let g, h be as in* (3.2.3), (3.3.4)—(3.3.7). *If D satisfies* (3.0.1), *$g \approx h$ if, and only if,*

(3.6.2) *S_λ is congruent to Σ_λ for each odd index λ;*

(3.6.3) *the corresponding partial second discriminants of g, h are essentially equal.*

REMARK. Necessary and sufficient conditions that (3.6.2) should hold are given by lemma 3.4.2.

PROOF. We have proved that the conditions are necessary. In proving that they are sufficient, we may suppose, by (3.6.2), that

$$\Sigma_\lambda = S_\lambda \quad \text{for all odd indices } \lambda. \tag{3.6.4}$$

* i.e. "Witt's theorem" in the case $p = 2$.

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

Let Ψ denote the system of equations (3.3.10), (3.3.13)—(3.3.17) (for all λ, μ). We now show that Ψ has a solution; in other words, we may make the further assumption that

(3.6.5) $$\boldsymbol{\Pi}_\mu = \boldsymbol{P}_\mu \quad \text{for all even indices } \mu.$$

Let $(\omega_{ij}) = \boldsymbol{\Pi}_\mu$, $(\rho_{ij}) = \boldsymbol{P}_\mu + \boldsymbol{Y}_1^T \boldsymbol{S}_{\mu+1} \boldsymbol{Y}_1 + \boldsymbol{Y}_{-1}^T \boldsymbol{S}_{\mu-1} \boldsymbol{Y}_{-1}$, where $\boldsymbol{Y}_1$, $\boldsymbol{Y}_{-1}$ are $m_{\mu+1} \times m_\mu$ and $m_{\mu-1} \times m_\mu$ matrices respectively. In order to show that Ψ has a solution, it is sufficient to prove that $u \sim v$ for suitable $\boldsymbol{Y}_{\pm 1}$, where u, v are the pseudo-discriminants of the non-defective quadratic forms $\sum \omega_{ij} x_i x_j$, $\sum \rho_{ij} x_i x_j$. For if this is so, (3.3.15), (3.3.16) have a solution $\boldsymbol{X}_\mu$, $\boldsymbol{X}_{\mu\pm1,\mu}$ by lemmas 3.4.1 and 3.6.2, and then the remaining equations in Ψ can obviously be satisfied by choosing the $\boldsymbol{X}_\lambda$, $\boldsymbol{X}_{\lambda\pm1,\lambda}$ and $\boldsymbol{X}_{\lambda+2,\lambda}$ appropriately.

If the index μ is itself a component set, then, by (3.6.3), we can take the $\boldsymbol{Y}_{\pm1}$ as zero matrices. If not, $\boldsymbol{S}_{\mu+\varepsilon}$ is a non-singular matrix diag $(s, \cdots)$ for $\varepsilon = 1$ or -1. We take $\boldsymbol{Y}_{-\varepsilon} = \boldsymbol{0}$, the (1,1), (1,2) elements of $\boldsymbol{Y}_\varepsilon$ as parameters ξ, η and all other elements in $\boldsymbol{Y}_\varepsilon$ zero. Then the equation $u \sim v$, i.e.

$$\sum_1^{\frac{1}{2}m_\mu} \pi_{2i-1}^\mu \pi_{2i}^\mu \sim (p_1^\mu + s\xi^2)(p_2^\mu + s\eta^2) + \sum_2^{\frac{1}{2}m_\mu} p_{2i-1}^\mu p_{2i}^\mu$$

has the form

$$\alpha^2 + \alpha + k + a\xi^2 + b\eta^2 + s^2\xi^2\eta^2 = 0,$$

and it is easily seen from lemma 3.6.1 that this has solutions.

We now complete the proof of the theorem by showing that the full system, Θ, of equations for equivalence has a solution. We take the $\boldsymbol{X}_\lambda$, $\boldsymbol{X}_\mu$ as unit matrices and the $\boldsymbol{X}_{\lambda\pm1,\lambda}$ $\boldsymbol{X}_{\mu\pm1,\mu}$ as zero matrices, so that all equations except (3.3.12), (3.3.14) are satisfied. Let Ψ_λ denote these two equations for a given λ (Ψ_λ being empty when $\lambda = 1$ or $m_\lambda = 0$). By theorem 3.5.1, we may assume, without loss of generality, that $e_1, \cdots, e_s$ (cf. (3.5.2)) form a single component set. By lemma 3.4.1, we may also assume that at least one e_i is odd. Then the odd e_i, arranged in descending order, form a consecutive sequence

$$\lambda_1 = 2\omega + 1, \lambda_2 = 2\omega - 1, \cdots, \lambda_t = 2\rho + 1 \ (t = \omega - \rho + 1 \geqq 1).$$

To prove the theorem, it obviously suffices to show that $\Psi_{\lambda_1} \cup \Psi_{\lambda_2} \cdots \cup \Psi_{\lambda_u}$ can be satisfied, assuming that $\Psi_{\lambda_1} \cup \cdots \cup \Psi_{\lambda_{u-1}}$ holds $(0 < u \leqq t)$. Thus, we may assume that $\boldsymbol{\Pi}_\lambda = \boldsymbol{P}_\lambda$ $(\lambda = \lambda_1, \cdots, \lambda_{u-1})$. We shall take the $\boldsymbol{X}_{\lambda+2,\lambda}$ and $\boldsymbol{X}_{\lambda,\lambda+2}$ as zero matrices for $\lambda = \lambda_1, \cdots, \lambda_u$.

Consider now the equations Ψ_λ, where $\lambda = \lambda_u$. Excluding the trivial case $u = t$, $\lambda = 1$, we have to consider the cases (a) $u = t$, $\lambda > 1$ and (b) $u < t$. In (a), Ψ_λ reduces to

(3.6.6)′ $$\boldsymbol{\Pi}_\lambda \approx \boldsymbol{P}_\lambda + \boldsymbol{K}_\lambda \boldsymbol{S}_\lambda^{-1} \boldsymbol{K}_\lambda + \boldsymbol{K}_\lambda \quad (\boldsymbol{K}_\lambda^T = \boldsymbol{K}_\lambda).$$

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

If $S_\lambda = J_\lambda$, take $K_\lambda = \Pi_\lambda + P_\lambda$. If not, (3.6.6)′ becomes

$$(3.6.6) \qquad (s_\alpha^\lambda)^{-1}(\pi_\alpha^\lambda + p_\alpha^\lambda) = \sum_{\substack{\kappa=1 \\ \kappa \neq \alpha}}^{m_\lambda} (k_{\kappa\alpha}^\lambda)^2/s_\kappa^\lambda s_\alpha^\lambda + (k_{\alpha\alpha}^\lambda/s_\alpha^\lambda)^2 + (k_{\alpha\alpha}^\lambda/s_\alpha^\lambda)$$
$$(\alpha = 1, \cdots m_\lambda;\ k_{\alpha\beta}^\lambda = k_{\beta\alpha}^\lambda).$$

We replace these equations by the first $m_\lambda - 1$ equations and the sum of all equations, viz.

$$(3.6.7) \qquad \sum_{\alpha=1}^{m_\lambda} (s_\alpha^\lambda)^{-1}(\pi_\alpha^\lambda + p_\alpha^\lambda) = k^2 + k,$$

where $k = \sum_{\alpha=1}^{m_\lambda} k_{\alpha\alpha}^\lambda/s_\alpha^\lambda$. We take all the elements of K_λ zero, except those on the main diagonal and in the last row and column. Then the first $m_\lambda - 1$ equations are satisfied for suitable $k_{\alpha\alpha}^\lambda$ and $k_{m_\lambda\alpha}^\lambda$ $(\alpha = 1, \cdots, m_\lambda - 1)$ by lemma 3.6.1. Finally, in view of (3.6.3), the sum equation has a solution $k_{m_\lambda m_\lambda}^\lambda$.

In case (b), Ψ_λ reduces to

$$(3.6.8) \qquad \Pi_\lambda \approx P_\lambda + K_\lambda S_\lambda^{-1} K_\lambda + K_\lambda + X_{\lambda-2,\,\lambda}^T S_{\lambda-2} X_{\lambda-2,\lambda} \quad (K_\lambda^T = K_\lambda).$$

This gives (3.6.6) with an added term $\sum_{\kappa=1}^{m_{\lambda-2}} s_\kappa^{\lambda-2}(x_{\kappa\alpha}/s_\alpha^\lambda)^2$ in the "α" equation $(\alpha = 1, \cdots m_\lambda)$.

The first $m_\lambda - 1$ equations can be solved as before. The sum of all m_λ equations, viz.

$$\sum_{\alpha=1}^{m_\lambda} (s_\alpha^\lambda)^{-1}(\pi_\alpha^\lambda + p_\alpha^\lambda) = k^2 + k + \sum_{\alpha=1}^{m_\lambda} \sum_{\kappa=1}^{m_{\lambda-2}} s_\kappa^{\lambda-2} \left(\frac{x_{\kappa\alpha}}{s_\alpha^\lambda}\right)^2$$

can be solved for $k_{m_\lambda, m_\lambda}^\lambda$ and the x_{ij} by lemma 3.6.1. This completes the proof.

3.7 *Example*: *finite classical groups*. We consider the finite-dimensional symplectic and orthogonal groups * over Galois fields $D = GF(q)$, $q = 2^\alpha$. For brevity, some of the proofs will be omitted.

Let $n = \dim V$. In the symplectic case, $n = 2\nu$ and $Sp(2\nu, q)$ is essentially unique. The order $|Sp(2\nu, q)|$ is given by (2.6.2). In the orthogonal case, if $n = 2\nu$ the fundamental form $|x|$ on V is equivalent to

$$(3.7.1) \qquad \left.\begin{array}{l} \sum_1^\nu x_{2i-1}x_{2i} \quad \text{(of Witt index } \nu\text{)} \\ \text{or} \quad (x_1^2 + x_1x_2 + \delta x_2^2) + \sum_2^\nu x_{2i-1}x_{2i} \quad \text{(of Witt index } \nu - 1\text{)} \end{array}\right\}$$

where δ is a fixed element of D such that $t^2 + t + \delta$ is irreducible (i.e. $\delta \nsim 0$, in the notation of § 3.5). The numerical invariant which distinguishes the two cases is the pseudo-discriminant ** $\Delta(|x|)$, defined up to essential

* The unitary groups $(U(n, 2^\alpha)$ were considered in § 2.6.
** Cf. § 3.6.

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

equality. The orthogonal groups of the forms (3.7.1) are denoted by $O_+(2\nu, q)$, $O_-(2\nu, q)$ and their orders are given by (2.6.3). If $n = 2\nu + 1$, $|x|$ is equivalent to $x_0^2 + \sum_1^\nu x_{2i-1}x_{2i}$ and the corresponding orthogonal group is isomorphic to $Sp(2\nu, q)$. *We shall therefore omit this case and assume throughout that* $n = 2\nu$.

In the orthogonal case, let L be a subspace of V. Then the equations

$$s(x + L^\sigma) = |x| \quad (x \in L)$$

define a non-degenerate form s on L/L^σ, called the *core* of $|x|$ on L. Notice that s is an odd- or even-dimensional form according to whether $L^\rho \neq L^\sigma$ or $L^\rho = L^\sigma$.

Let $X \in G = Sp(2\nu, q)$ or $O_\pm(2\nu, q)$. As before, $\phi(t)$ denotes an irreducible monic polynomial distinct from t and $m(\phi^i)$ is the multiplicity of ϕ^i as elementary divisor of X. Then we have

$$\left.\begin{aligned} &\sum_{\phi, i} i|\phi|m(\phi^i) = n; \\ &m(\phi^i) = m(\tilde{\phi}^i); \\ &m\big((t+1)^{2i-1}\big) \text{ is even } (i = 1, 2, \cdots). \end{aligned}\right\} \tag{3.7.2}$$

The last two statements follow from theorem 2.3.1, lemma 1.4.5 (cor. 2) and the evenness of n.

Let W, f, P denote the space, form and multiplier of X. We may suppose that the Fitting $(t+1)$-component of f is represented (with respect to a standard reference form chosen as in § 3.1) by the matrix $\boldsymbol{Q}$ in (3.3.1), where (3.3.2)′, (3.3.3)′, (3.3.4)—(3.3.7) hold.

A non-singular symmetric matrix over $GF(2^\alpha)$ is congruent to the canonical alternate matrix $\boldsymbol{J}$ or unit matrix $\boldsymbol{I}$. Thus we may suppose in (3.3.2)′ that $\boldsymbol{S}_\lambda = \boldsymbol{J}_\lambda$ or $\boldsymbol{I}_\lambda$ $(\lambda = 1, 3, \cdots)$. Accordingly, we define an invariant s_λ $(\lambda = 1, 3, \cdots)$ by:

$$s_\lambda = \left.\begin{aligned} &0 \quad \text{if} \quad \boldsymbol{S}_\lambda = \boldsymbol{J}_\lambda \text{ or the dimension of } \boldsymbol{S}_\lambda = 0, \\ &1 \quad \text{if} \quad \boldsymbol{S}_\lambda = \boldsymbol{I}_\lambda. \end{aligned}\right\}$$

Then clearly

$$\left.\begin{aligned} &s_\lambda = 0 \quad \text{when} \quad m\big((t+1)^{\lambda+1}\big) = 0, \\ &s_\lambda = 1 \quad \text{when} \quad m\big((t+1)^{\lambda+1}\big) \text{ is odd.} \end{aligned}\right\} \tag{3.7.3}$$

Further we have

LEMMA 3.7.1. *In the orthogonal case,* $W^\rho = W^\sigma$ *if, and only if,* $s_1 = 0$. (Proof omitted.)

The values of the s_λ determine the component sets $L_1, L_2, \cdots$ in (3.5.5)′. Let $\delta_1, \delta_2, \cdots$ be the corresponding partial second discriminants. We may suppose that $\delta_i = 0$ or δ since δ_i is determined only up to essential equality. It is convenient to define one further "partial second discriminant" δ_0 by:

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

$$\delta_0 = \begin{matrix} \Delta(s) & \text{if} \quad G = O, s_1 = 1, \\ 0 & \text{otherwise,} \end{matrix} \Bigg\}$$

where s is the core of $|x|$ on $W^{\perp}$. With this notation we have

LEMMA 3.7.2. *In the orthogonal case,*

$$\Delta(|x|) = \sum \delta_i + (\sum_{\phi, i} im(\phi^i))\delta. \tag{3.7.4}$$

(Proof omitted.)

The δ_i satisfy the conditions

$$\left.\begin{aligned} &\delta_0 = 0 \quad \text{if} \quad G = Sp \quad \text{or} \quad G = O, s_1 = 0; \\ &\delta_1 = 0 \quad \text{if} \quad s_1 = 1; \\ &\delta_i = 0 \text{ if the corresponding component set consists of a single odd index } \lambda \text{ such that } s_\lambda = 0. \end{aligned}\right\} \tag{3.7.5}$$

(Cf. § 3.5.)

With the methods of § 2.6 and theorem 3.6.1, the following two theorems are now easily deduced.

THEOREM 3.7.1. *Two elements of G are conjugate in G if, and only if, they have the same invariants $m(\phi^i)$, s_λ, δ_i.*

THEOREM 3.7.2. *Choose a non-negative integer $m'(\phi^i)$ for each power ϕ^i of a monic irreducible ϕ, an integer $s'_\lambda = 0$ or 1 for each index $\lambda = 1, 3, \cdots$ and an element $\delta'_i = 0$ or δ of D for each of the component sets formally defined by the $m'(\phi^i)$ and s'_λ. Then the $m'(\phi^i)$, s'_λ, δ'_i are the invariants of an element X of G if, and only if, they satisfy* (3.7.2)—(3.7.5).

It is now merely a combinatorial question to enumerate the conjugacy classes in G. We define a sequence of polynomials $\chi_i = \chi_i(a, b; t)$ $(i = 1, 2, \cdots)$ and a power series $\chi = \chi(a, b; t)$ as follows:

$$\left.\begin{aligned} \chi_{-1} = a, \quad \chi_0 = b, \\ \chi_{2\nu+1} - \chi_{2\nu} = t^{2\nu+1}\chi_{2\nu-1}, \\ \chi_{2\nu+2} - \chi_{2\nu+1} = t^{\nu+1}(1 + t^{\nu+1})(\chi_{2\nu+1} + (1 - t^{2\nu+1})\chi_{2\nu-1}) \\ \chi(t) \equiv \chi_{2r}(t) \pmod{t^r} \quad (r = 0, 1, 2, \cdots). \end{aligned}\right\} \quad (\nu \geqq 0),$$

THEOREM 3.7.3. *Let the numbers of conjugacy classes in $Sp(2\nu, q)$, $O_{\pm}(2\nu, q)$ be the coefficients of $t^{2\nu}$ in the power series $sp(t^2)$, $\omega_{\pm}(t^2)$. Then*

$$sp(t^2) = \chi(0, 1; t^2) \prod_{\lambda=1}^{\infty} (1 - qt^{2\lambda})^{-1},$$

$$\omega_+(t^2) + \omega_-(t^2) = \chi(1, 1; t^2) \prod_{\lambda=1}^{\infty} (1 - qt^{2\lambda})^{-1},$$

$$\omega_+(t^2) - \omega_-(t^2) = \prod_{\lambda=1}^{\infty} \left(\frac{1 - t^{4\lambda-2}}{1 - qt^{4\lambda}}\right).$$

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

(Proof omitted.)

We determine finally the order of the conjugacy class of X in G.

Notation. A component set is called *big* if it contains an odd index λ such that $s_\lambda = 1$ and if, when $G = Sp$, it does not contain the index 1. An odd index $\mu + 1$ is called *isolated* if *either* μ is positive and forms a component set by itself *or* $\mu = 0$, $G = O$ and $s_1 = 0$. With such an isolated index $\mu + 1$ is associated a partial second discriminant δ', viz. $\delta' = \delta_0$ when $\mu = 0$, $\delta' =$ partial discriminant determined by μ when $\mu > 0$. $O(m_{\mu+1}, q)$ denotes the orthogonal group of the form of pseudo-discriminant δ' and dimension $m_{\mu+1} = m\big((t+1)^{\mu+1}\big)$.

We write

$$B(\phi) = Q^{\sum_{i<j} i\, m_i m_j + \frac{1}{2}\sum_i (i-1) m_i^2} \prod_i A(\phi^i),$$

where $Q = q^{|\phi|}$, $m_i = m(\phi^i)$ and $A(\phi^i)$ is defined as follows:
if $\phi(t) \neq (t+1)$,

$$A(\phi^i) = \left.\begin{array}{ll} |U(m_i, Q)| & (\phi = \tilde{\phi}), \\ |GL(m_i, Q)|^{\frac{1}{2}} & (\phi \neq \tilde{\phi}); \end{array}\right\}$$

if $\phi(t) = (t+1)$,

$$A(\phi^i) = \left\{\begin{array}{ll} q^{\varepsilon \frac{1}{2} m_i} |Sp(m_i, q)| & (i \text{ even}, s_{i-1} = 0) \\ q^{\varepsilon \frac{1}{2} m_i} |Sp(m_i - 1, q)| & (i \text{ even}, s_{i-1} = 1, m_i \text{ odd}) \\ q^{(1+\frac{1}{2}\varepsilon) m_i - 1} |Sp(m_i - 2, q)| & \\ \qquad (i \text{ even}, s_{i-1} = 1, m_i \text{ even}) & \\ q^{\frac{1}{2}(1+\varepsilon) m_i} |O(m_i, q)| & (i \text{ odd, isolated}) \\ q^{-\frac{1}{2}(1-\varepsilon) m_i} |Sp(m_i, q)| & (i \text{ odd, non-isolated}), \end{array}\right.$$

where $\varepsilon = +1$ or -1 according as $G = Sp$ or O.

THEOREM 3.7.4. *The order of the conjugacy class of X in G is*

$$|G| / 2^k \prod_\phi B(\phi),$$

where k is the number of big component sets.

OUTLINE OF PROOF. By the methods of § 2.6 and the Fitting decomposition, it is sufficient to consider the case where $I + P$ is nilpotent. Let Q be the representative of $f = f_X$ with respect to a standard reference form f' and let $\dagger$ denote the adjoint with respect to f'. Let N denote the group of $T \in \mathscr{C} = \mathscr{C}(P)$ such that $T^\dagger Q T = Q$, $\bar{N}$ the group of $\bar{T} \in \bar{\mathscr{C}}$ such that $\bar{T}^\dagger \bar{Q} \bar{T} = \bar{Q}$, where the $-$ now denotes passage to the quotient $\mathscr{C}/\mathscr{G}$ (cf. § 3.3.) As in theorem 2.2.3,

$$|N| = |\bar{N}| |\mathscr{G}|^{\frac{1}{2}}.$$

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

The mapping $Y \to Y_W$ gives a homomorphism of $C(X)$ onto N, whose kernel H consists of the $Y \in G$ which leave every element of W fixed. Thus

$$|C(X)| = |N||H|.$$

Since the calculation of $|\mathscr{G}|$ is straightforward, it remains to consider $|\bar{N}|$ and $|H|$.

$|\bar{N}|$ is the number of ways of choosing the matrices $\boldsymbol{X}_i (i = 1, 2, \cdots)$, $\boldsymbol{Y}_i (i = 2, 3, \cdots)$, $\boldsymbol{Z}_\lambda (\lambda = 3, 5, \cdots)$, $\boldsymbol{X}_{i,i+1}$, $\boldsymbol{X}_{i+1,i}$, $\boldsymbol{X}_{i,i+2}$, $\boldsymbol{X}_{i+2,i}$ $(i = 1, 2, \cdots)$ so that (3.3.10)—(3.3.17) hold with $\boldsymbol{S}_\lambda$, $\boldsymbol{P}_\lambda$, $\boldsymbol{P}_\mu$ in place of $\boldsymbol{\Sigma}_\lambda$, $\boldsymbol{\Pi}_\lambda$, $\boldsymbol{\Pi}_\mu$ on the left hand sides. This number is conveniently calculated as the quotient u/v, where u is the number of ways of choosing the $\boldsymbol{X}_i$, $\boldsymbol{Y}_i$, etc. so that (3.3.10)—(3.3.17) hold with $\boldsymbol{\Sigma}_\lambda = \boldsymbol{S}_\lambda$ and with $\boldsymbol{\Pi}_\lambda$, $\boldsymbol{\Pi}_\mu$ having the forms

$$\boldsymbol{\Pi}_\lambda = \operatorname{diag}(\pi_1^\lambda, \cdots \pi_{m_\lambda}^\lambda)$$

$$\boldsymbol{\Pi}_\mu = \operatorname{diag}\left(\begin{pmatrix} \pi_1^\mu & 1 \\ 0 & \pi_2^\mu \end{pmatrix}, \cdots\right),$$

and where v is the number of ways in which such $\boldsymbol{\Sigma}_\lambda$, $\boldsymbol{\Pi}_\lambda$, $\boldsymbol{\Pi}_\mu$ can be chosen. The calculation of u, v is straightforward except for the determination of the number $\boldsymbol{X}_\lambda$ satisfying $\boldsymbol{X}_\lambda^T \boldsymbol{S}_\lambda \boldsymbol{X}_\lambda = \boldsymbol{S}_\lambda$. By the methods of § 3.4, this number is found to be

$$\begin{aligned} |Sp(m_\lambda, q)| &\quad \text{if} \quad \boldsymbol{S}_\lambda = \boldsymbol{J}_\lambda, \\ |Sp(m_\lambda - 1, q)| &\quad \text{if} \quad \boldsymbol{S}_\lambda = \boldsymbol{I}_\lambda, m_\lambda \text{ odd}, \\ q^{m_\lambda - 1}|Sp(m_\lambda - 2, q)| &\quad \text{if} \quad \boldsymbol{S}_\lambda = \boldsymbol{I}_\lambda, m_\lambda \text{ even}, \end{aligned}$$

where $m_\lambda = \dim \boldsymbol{S}_\lambda$.

It remains to calculate $|H|$. To fix ideas, consider the case $G = O$, $W^\rho \neq W^\sigma$. Write $L = W^\perp$, $m_i = m((t+1)^i)$. Then

$$\dim L = m_1 + m_2 + \cdots$$

$$1 + \dim L^\sigma = \dim L^\rho = m_2 + m_3 + \cdots$$

If $Y \in H$, let $\eta(Y)$ denote the induced linear transformation on L/L^σ. By Witt's theorem, $\eta(H) \cong O(m_1 + 1, q) \cong Sp(m_1, q)$. Let K be the kernel of η and let $\zeta(Y)$ denote the transformation on L induced by $Y \in K$. It is easy to see that $\zeta(K)$ consists of the linear transformations which leave every element of L/L^σ and L^ρ fixed, so that

$$|\zeta(K)| = q^{m_1(m_2 + m_3 + \cdots - 1)}.$$

The kernel T of ζ consists of the $Y \in G$ which leave every element of W and $W^\perp$ fixed, i.e. the Y with spaces in $(W + W^\perp)^\perp = W \cap W^\perp = W^\rho$. Let $e_1, e_2, \cdots$ be a basis of W^ρ, where $e_2, \cdots$ are a basis of W^σ. Each element of T has the form

https://doi.org/10.1017/S1446788700027622 Published online by Cambridge University Press

$$x \to x + \sum (x \cdot e_i)\omega_{ij} e_j.$$

Conversely, it is found that such a linear transformation belongs to G if, and only if,

$$\omega_{ij} = \omega_{ji} \quad (\text{all } i, j)$$
$$\omega_{ii} = \omega_{i1}^2 \quad (\text{all } i).$$

Hence

$$|T| = 2q^{\frac{1}{2}(m_2+\cdots)(m_2+\cdots-1)}.$$

Putting these values together we get the required order $|H|$. The calculation of $|H|$ in the other cases is similar.

References

[1] Bourbaki, N., *Formes sesquilinéaires et formes quadratiques* (Éléments de Mathématique I, livre II, chapitre 9), Hermann (Paris), 1959.

[2] Dieudonné, J., *La géometrie des groupes classiques*, Springer (Berlin), 1955.

[3] Feit, W. and Fine, N. J., *Pairs of commuting matrices over a finite field*, Duke Math. J. 27 (1960), 91–94.

[4] Ingraham, M. H. and Wegner, K. W., *The equivalence of pairs of Hermitian matrices*, Prans. Amer. Math. Soc. 38 (1935), 145–162.

[5] Jacobson, N., *Lectures in Abstract Algebra*, vol. II, Van Nostrand (New York), 1952.

[6] Jacobson, N., *The theory of rings*, Math. Surveys No. II, American Mathematical Society (New York), 1943.

[7] Klingenberg, W., *Paare symmetrischer und alternierender Formen zweiten Grades*, Abh. Math. Sem. Univ. Hamburg 19 (1954), 78–93.

[8] Pickert, G., *Normalformen von Matrizen*, Enz. der Math. Wiss. Bd. II, Heft 3, Teil 1, Teubner (Leipzig), 1953.

[9] Springer, T. A., *Over Symplectische Transformaties*, Thesis, University of Leiden, 1951.

[10] Trott, G. R., *On the canonical form of a non-singular pencil of Hermitian matrices*, Amer. J. Math. 56 (1934), 359–371.

[11] Turnbull, H. W., *On the equivalence of pencils of Hermitian forms*, Proc. London Math. Soc. (2) 39 (1935), 232–248.

[12] Venkatachalienger, K., *Pairs of symmetric and skew matrices in an arbitrary field I*, Proc. Indian Acad. Sci., Sect. A, 22 (1945), 243–264.

[13] Wall, G. E., *The structure of a unitary factor group*, Publ. math. de l'Inst. des hautes études scientifiques No. 1 (1959).

[14] Williamson, J., *The equivalence of non-singular pencils of Hermitian matrices in an arbitrary field*, Amer. J. Math. 57 (1935), 475–490.

[15] Williamson, J., *On the algebraic problem concerning the normal forms of linear dynamical systems*, Amer. J. Math. 58 (1936), 141–163.

[16] Williamson, J., *Quasi-unitary matrices*, Duke Math. J. 3 (1937) 715–725.

[17] Williamson, J., *On the normal forms of linear canonical transformations in dynamics*, Amer. J. Math. 59 (1937), 599–617.

[18] Williamson, J., *Normal matrices over an arbitrary field of characteristic zero*, Amer. J. Math. 61 (1939), 335–356.

[19] Williamson, J., *Note on the equivalence of non-singular pencils of Hermitian matrices*, Bull. Amer. Math. Soc. 51 (1945), 894–897.

[20] Zassenhaus, H., *On a normal form of the orthogonal transformation*, Canad. Math. Bull. 1 (1958), (I) 31–39, (II) 101–111, (III) 183–191.

[21] Ennola, V., *On the conjugacy classes of the finite unitary groups*, Ann. Acad. Sci. Fenn. AI, No. 313 (1962).

Department of Mathematics,
University of Sydney.