Classification of Hamiltonian Linear Systems

ANTONIO CIAMPI

Communicated by the Editors

1. Introduction. A *hamiltonian linear system* is a system of first order differential equations in $2n$ variables of the form:

$$(1.1) \quad \begin{cases} \dot{p}^k = -\frac{\partial H}{\partial q_k} \\ \dot{q}_k = \frac{\partial H}{\partial p^k} \end{cases} \quad k = 1, \dots, n$$

where H is a *quadratic form* in the p 's and the q 's, and therefore $\partial H/\partial q_k$, $\partial H/\partial p^k$ are *linear functions*. H is called the *hamiltonian function* of the system. Two such systems are *equivalent* iff there is a linear transformation of the p 's and the q 's which carries one into the other; we shall call such an equivalence *symplectic similarity*.

In 1936, J. Williamson [1] obtained a complete classification of hamiltonian linear systems up to symplectic similarity. Both the results and the method are difficult to penetrate and they have remained relatively unknown to applied mathematicians and physicists. On the other hand, the original problem has been stated more recently in the wider context of the theory of bilinear forms and in that of algebraic groups, leading to several interesting works by pure algebraists [2, 3, 4]. These works are very elegant, but still impenetrable to non-specialists.

In this paper we present a reformulation of Williamson's classification problem which should be accessible to physicists and applied mathematicians, since in the development little more is required than the classical results of linear algebra. Our method should be considered as a compromise between Williamson's matrix approach and the sophisticated algebraic one that can be extracted from recent works.

2. Formulation of the problem and results. A *symplectic space* is a pair (W, ω) consisting of a linear space W over a field K , and a *symplectic form* ω (*symplectic* = anti-symmetric, non-degenerated, bilinear). A *symplectic subspace*

U of (W, ω) is a linear subspace of W such that the restriction of ω to $U \times U$ is a *symplectic form* (in general, the restriction is degenerate, hence not symplectic). The definition of *orthogonality* in a symplectic space is the same as in ordinary inner product spaces (we interpret the number $\omega(u, v)$ as “inner product” of the vectors u and v). We shall use the symbol $U \perp V$ to denote orthogonal direct sum of two subspaces U and V , and the symbol $U^\perp$ to denote the space of all vectors of W orthogonal to the subspace U . Unlike the case of ordinary inner product spaces, a subspace in a symplectic space can be orthogonal to itself or to one of its subspaces. A subspace orthogonal to itself is called *isotropic*. The *radical* of a subspace U , denoted by $\text{rad } U$, is the subspace of U consisting of the vectors orthogonal to all the vectors in U . The following lemma, whose proof we leave to the reader, shall be useful in the development:

Lemma 2.1.

- (a) $U \cap U^\perp = \text{rad } U$, for any subspace U of W ;
- (b) $U \perp V$ is symplectic iff U and V are symplectic;
- (c) U is symplectic iff $U^\perp$ is; if this is the case, then $U \perp U^\perp = W$.

Definition 2.1. A system of first order differential equations:

$$(2.1) \quad \dot{x} = Bx,$$

where x is a vector in a symplectic space (W, ω) and B is a linear transformation of W such that:

$$(2.2) \quad \omega(Bu, v) = -\omega(u, Bv) \quad \forall u, v \in W,$$

is called a *linear hamiltonian system*. The symmetric bilinear form:

$$(2.3) \quad \alpha(u, v) = \omega(Bu, v)$$

is called the *hamiltonian* of the system, and the quadratic form:

$$(2.4) \quad H(u) = \alpha(u, u)$$

is called the *hamiltonian function*.

A *symplectic basis* of W is a basis $\{u_i, v_i\}$, $i, j = 1, \dots, n$, such that:

$$(2.5) \quad (\omega) = \begin{bmatrix} \omega(u_i, u_i) & \omega(u_i, v_i) \\ \omega(v_i, u_i) & \omega(v_i, v_i) \end{bmatrix} = \begin{bmatrix} 0 & 1_n \\ -1_n & 0 \end{bmatrix}$$

where 1_n denote the $n \times n$ identity matrix. It is a classical result that a symplectic space always has symplectic bases. It is also a standard result that a *skew-self-adjoint* operator (*i.e.* an operator satisfying (2.2)) has with respect to any symplectic basis, the matrix:

$$(2.6) \quad (B) = \begin{bmatrix} X_1 & X_2 \\ X_3 & -{}^t X_1 \end{bmatrix}$$

where X_2 and X_3 are $n \times n$ symmetric matrices and X_1 is an arbitrary $n \times n$ matrix.

It can be seen that, with respect to a symplectic basis, equations (2.1) take the form (1.1) of the original definition. A more detailed account of this can be found in our work [5] to which we also refer for more extensive proofs. For more information on symplectic spaces we refer the reader to the monograph by Artin [6].

The classification problem can then be stated as follows: *given a symplectic space (W, ω) describe the equivalence classes of skew-self-adjoint operators under the equivalence relation; $B \sim B'$ iff $B' = SBS^{-1}$, where S is a symplectic transformation of W , i.e. a linear transformation such that $\omega(u, v) = \omega(Su, Sv)$ for all $u, v \in W$.*

We recall from linear algebra [7] that the *elementary divisors* of a linear transformation are a complete set of invariants under *similarity*. Let A denote a linear transformation of a linear space V and let $A = \bigoplus_{i \leq i \leq r} A_i$ $V = \bigoplus_{i \leq i \leq r} V_i$ denote a direct sum decomposition of V into A -invariant, irreducible cyclic subspaces, where $A_i = A|_{V_i}$, $j = 1, \dots, r$. Then we recall that the elementary divisors are the characteristic polynomials of the A_i 's. We shall say that an elementary divisor $\pi(\lambda)$ occurs with multiplicity ν if there are ν A_i 's, each having $\pi(\lambda)$ as characteristic polynomial.

We are now ready to state our results:

Theorem 2.2. *Let B denote a skew-self-adjoint operator on a symplectic space (W, ω) over a field K of characteristic $p \neq 2$. Then:*

- (i) *(Properties of the elementary divisors) Its elementary divisors are of two types:*
 - (a) $p^k(\lambda)$, where $p(\lambda)$ is an irreducible polynomial and $p^k(\lambda) = p^k(-\lambda)$.
(Notice that this case includes the elementary divisors $p^k(\lambda) = \lambda^{2h}$.)
 - (b) $q^k(\lambda)$, where $q(\lambda)$ is an irreducible polynomial and $q^k(\lambda) \neq q^k(-\lambda)$.
(Notice that this case includes the elementary divisors: $q^k(\lambda) = \lambda^{2h+1}$.)

The elementary divisors of type (b) occur always in pairs: $q^k(\lambda)$, $q^k(-\lambda)$.

- (ii) *(Existence of a decomposition "adapted to the geometry") There exists a direct sum decomposition of W into symplectic orthogonal subspaces invariant under B :*

$$W = W_1 \perp W_2 \perp \dots W_i \perp \dots$$

$$B = B_1 \oplus B_2 \oplus \dots B_i \oplus \dots$$

such that each subspace corresponds either to an elementary divisor of type (a) or to a pair of divisors $q^k(\lambda)$, $\pm q^k(-\lambda)$, where $q^k(\lambda)$ is of type (b).

If W_i corresponds to a divisor of type (a), it is irreducible and $B_i = B|_{W_i}$ has minimal polynomial equal to the elementary divisor. If W_i corresponds to a pair of elementary divisors of type (b), it can be reduced further as the sum of two isotropic subspaces V^+ , V^- , each irreducible and such that $B|_{V^+}$ has minimal polynomial $q^k(\lambda)$, and $B|_{V^-}$ has minimal polynomial $\pm q^k(-\lambda)$.

$a > 0$, of multiplicity ν , and with hermitian invariant equivalent to $(-1, \nu)$, there is a term in the expansion of H which has the form of the hamiltonian function of a ν -dimensional harmonic oscillator.

3. Proof of Theorem 2.2. We break up the proof into several steps. Let B , W , ω be as in the previous section; we group the irreducible factors of the characteristic polynomial of B into two classes:

- (a) $p(\lambda)$, such that $p(\lambda) = \pm p(-\lambda)$;
- (b) $q(\lambda)$, such that $q(\lambda) \neq \pm q(\lambda)$.

We define: $q^*(\lambda) = q(-\lambda)$ for factors of class b. We shall also use the symbol:

$$W_\phi = \{v \in W; \phi^i(B)v = 0, \text{ for some } i \in \mathbf{Z}\},$$

where ϕ denotes any polynomial.

Lemma 3.1. *W admits the orthogonal B -invariant decomposition:*

$$W = \perp_p W_p \perp_q W_{q^*}.$$

Proof. The existence of the invariant decomposition: $W = \bigoplus_p W_p \bigoplus_q W_{q^*}$ is a consequence of the *primary decomposition* [7]. If W_ϕ and $W_{\phi'}$ denote any two distinct summands in the above decomposition, then we have to prove that: $\omega(W_\phi, W_{\phi'}) = 0$. Using the fact that $\phi'(B)$ is one-to-one on W_ϕ , we have:

$$\omega(W_\phi, W_{\phi'}) = \omega(\phi'(B)^k W_\phi, W_{\phi'}),$$

for every integer k . On the other hand: $\phi'(B)^\dagger = \pm \phi'(B)$ where $\dagger$ denotes adjunction with respect to ω ; so, for k large enough, we have:

$$\omega(\phi'(B)^k W_\phi, W_{\phi'}) = \pm \omega(W_\phi, \phi'(B)^k W_{\phi'}) = \pm \omega(W_\phi, 0) = 0.$$

Using Lemma 2.1, it is easy to see that each W_ϕ is symplectic. ■

The above lemma reduces our problem to that of classifying linear transformations having characteristic polynomials either of the type:

- (a) $\chi(\lambda) = p^k(\lambda)$, or of the type
- (b) $\chi(\lambda) = q^k(\lambda)q^{k'}(-\lambda)$.

We shall deal with type (b) first.

Lemma 3.2. *Let B have characteristic polynomial $\chi(\lambda) = q^k(\lambda)q^{k'}(-\lambda)$, for positive integers k, k' . Then in the primary decomposition:*

$$W = V_q \oplus V_{q^*},$$

each of the subspaces V_q, V_{q^} is isotropic.*

The proof of this lemma is left to the reader, as it is similar to that of Lemma 3.1.

Theorem 3.3. Let B have characteristic polynomial $\chi(\lambda) = q^k(\lambda)q^{k'}(-\lambda)$, and let $q^{k_1}(\lambda), q^{k_2}(\lambda), \dots, q^{k_s}(\lambda)$ denote the sequence of the elementary divisors (with possible repetitions) of $B|_{V_q}$. Then $B|_{V_{q^*}}$ has elementary divisors: $q^{k_1}(-\lambda), q^{k_2}(-\lambda), \dots, q^{k_s}(-\lambda)$. (Hence $k = k'$.) There exists at least one orthogonal direct sum invariant decomposition

$$W = (V_1^+ \oplus V_1^-) \perp (V_2^+ \oplus V_2^-) \perp \dots \perp (V_s^+ \oplus V_s^-)$$

$$B = (B_1^+ \oplus B_1^-) \oplus (B_2^+ \oplus B_2^-) \oplus \dots \oplus (B_s^+ \oplus B_s^-)$$

where each V_i^+, V_i^- is isotropic and irreducible and where $(V_i^+ \oplus V_i^-)$ is symplectic: $B_i^+ = B|_{V_i^+}$ has characteristic (minimal) polynomial equal to $q^{k_i}(\lambda)$, and $B_i^- = B|_{V_i^-}$ has characteristic (minimal) polynomial, equal to $q^{k_i}(-\lambda)$.

Also there exists a symplectic basis with respect to which B has the irreducible block-form

$$(B) = \begin{bmatrix} (B_1) & & & & & & \\ & (B_2) & & & & & \\ & & \ddots & & & & \\ & & & (B_s) & & & \\ & & & & -{}^t(B_1) & & \\ & & & & & -{}^t(B_2) & \\ & & & & & & \ddots & \\ & & & & & & & -{}^t(B_s) \end{bmatrix}.$$

Proof. The proof follows almost immediately from Lemma 3.2 and from the fact that a set of vectors which span an isotropic subspace can be completed to a symplectic basis [5] and from equation (2.6). ■

We now consider case (a).

Lemma 3.4. Let B denote a skew-self-adjoint transformation of (W, ω) . Suppose that its characteristic polynomial is:

$$\chi(\lambda) = p^k(\lambda)$$

with $p(\lambda)$ irreducible and such that $p(-\lambda) = \pm p(\lambda)$. Then there exists at least one decomposition of W into the direct sum of orthogonal symplectic spaces invariant under B :

$$W = W_1 \perp W_2 \perp \dots \perp W_s$$

$$B = B_1 \oplus B_2 \oplus \dots \oplus B_s$$

such that each B_i has a single distinct elementary divisor $p^{k_i}(\lambda)$, possibly occurring with a certain multiplicity $\nu_i > 1$.

Proof. Consider an arbitrary decomposition (not necessarily orthogonal) of the form:

$$W = V_1 \oplus V_2 \oplus \cdots \oplus V_s$$

$$B = B'_1 + B'_2 \oplus \cdots \oplus B'_s$$

where $p^{k_i}(B)V_i = 0$, $p^{k_i-1}(B)V_i \neq 0$; in other words, B_i has ν_i identical elementary divisors $p^{k_i}(\lambda)$. Suppose that the order of the term is such that $k_1 > k_2 > \cdots > k_s$. It suffices to show that V_1 is necessarily symplectic; for then we can choose $W_1 = V_1$ and select W_2 by the same procedure applied to $W_1^\perp$, which is also invariant under B (Lemma 4.3), and so on, until we exhaust the whole space W .

Suppose V_1 not symplectic. Then its radical $\text{rad } V_1$ contains a non-zero vector u ; that is, there exists a non-zero $u \in V_1$ such that $\omega(u, V_1) = 0$. Since $\text{rad } V_1$ is invariant under B (Lemma 4.4), we can assume $p(B)u = 0$, so that $u \in \text{kern } p(B)$, and therefore $u = p^{k_1-1}(B)v$ for some $v \in V_1$. But then for $i > 1$ we have, using $p(-\lambda) = \pm p(\lambda)$:

$$\omega(u, V_i) = \omega(p^{k_1-1}(B)v, V_i) = \pm \omega(v, p^{k_1-1}(B)V_i) = \pm \omega(v, 0) = 0.$$

Hence u is orthogonal to the whole W , which is absurd. ■

Corollary. For a skew-self-adjoint operator B on the symplectic space (W, ω) the elementary divisors of the form λ^{2h+1} appear in pairs.

Notice that with this corollary, we have completely proved statement (i) of Theorem 2.2. We are now ready to construct the hermitian invariant which appear in statement (iii).

Lemma 3.5. Let B have a single elementary divisor λ^k with multiplicity ν . Define:

$$\tilde{W} = W/BW, \quad \gamma(\bar{u}, \bar{v}) = \omega(u, B^{n-1}v),$$

where u, v are arbitrary representatives of the classes $\bar{u}, \bar{v} \in W/BW$. Then:

- (i) γ is a well defined, non-degenerate, bilinear form on $\tilde{W}$.
- (ii) $\gamma(\bar{u}, \bar{v}) = (-1)^k \gamma(\bar{v}, \bar{u})$; hence γ is symmetric if $k = 2h$, and anti-symmetric if $k = 2h + 1$.

Lemma 3.6. Let B have a single elementary divisor:

$$p^k(\lambda) = (\lambda^2 + a^2)^k, \quad a \neq 0$$

with multiplicity ν . Then:

- (i) $H = W/p(B)W$ is a ν -dimensional vector space over the field $\mathbf{R}[B]/p(B) \cong \mathbf{C}$, where $\mathbf{R}[\lambda]$ denotes, as usual, the algebra of the polynomials in one variable over $\mathbf{R}$, $\mathbf{C}$ is the field of the complex numbers.

- (ii) Let $\bar{B} \in \mathbf{R}[B]/p(B)$ denote a solution of $p(B) = 0$; i.e. $\bar{B}^2$ is the subset of $\mathbf{R}[B]$ whose elements are transformations congruent to $-\alpha^2 1$ modulo $p(B)$. Let $*$ denote the automorphism of $\mathbf{R}[\bar{B}] \cong \mathbf{R}[B]/p(B)$ defined by:

$$\bar{B}^* = -\bar{B}$$

and extended by linearity to the whole of $\mathbf{R}[\bar{B}]$. Then the mapping:

$$h : H \times H \rightarrow \mathbf{R}[\bar{B}]$$

$$h(\bar{u}, \bar{v}) = \omega(Bp^{k-1}(B)u, v) + \omega(p^{k-1}(B)u, v)\bar{B},$$

where u, v are arbitrary representatives of the classes $\bar{u}, \bar{v}$ modulo $p(B)W$, is a $\mathbf{R}[B]$ -sesquilinear, non-degenerate, hermitian (i.e. $h(\bar{u}, \bar{v}) = h(\bar{v}, \bar{u})^*$) form.

We leave to the reader the simple but lengthy verifications of these statements.

We recall that an inner product space with *antisymmetric* inner product is a “trivial invariant”, in the sense that any two such spaces of the same dimensions are equivalent. On the other hand, for *symmetric* inner product spaces we have non-trivial equivalence classes. In fact the “inertia theorem” (or Sylvester theorem) states that every symmetric non-degenerate inner product is equivalent to an inner product with matrix:

$$\begin{bmatrix} \eta_1 & & & & \mathbf{0} \\ & \eta_2 & & & \\ & & \ddots & & \\ & & & \ddots & \\ \mathbf{0} & & & & \eta_m \end{bmatrix}$$

and in particular, for spaces over the field of the reals $\mathbf{R} : \eta_i = \pm 1$. The number of $+1$'s, occurring in this matrix is called the *index* of the inner product space; two nondegenerate inner product spaces are equivalent if and only if they have the same dimension and the same index.

Theorem 3.7. Let B have a single elementary divisor: λ^{2k+1} , with multiplicity $\nu = 2\mu$ (ν is even, according to the Corollary of Lemma 3). Then:

- (i) there exists at least one direct sum B -invariant irreducible decomposition:

$$W = (V_1^+ \oplus V_1^-) \perp (V_2^+ \oplus V_2^-) \perp \cdots \perp (V_\mu^+ \oplus V_\mu^-)$$

$$B = (B_1^+ \oplus B_1^-) \oplus (B_2^+ \oplus B_2^-) \oplus \cdots \oplus (B_\mu^+ \oplus B_\mu^-)$$

with $V_i^+ \oplus V_i^-$ symplectic and V_i^+, V_i^- isotropic for each i .

- (ii) there exists a symplectic basis of W with respect to which B is represented by a matrix which is a sum of μ identical blocks of the form:

$$\left[\begin{array}{cccc|cccc} 0 & 1 & & & & & & \\ & 0 & 1 & & & & & \\ & & 0 & \ddots & & & & \\ & & & \ddots & \ddots & & & \\ & & & & \ddots & 1 & & \\ & 0 & & & & 0 & & \\ \hline & & & & & & 0 & \\ & & & & 0 & & & \\ & & & & -1 & 0 & & \\ & & & & & -1 & \ddots & \\ & & & & & & \ddots & 0 \\ & & & & & & & -1 \end{array} \right]$$

(iii) any two transformations B, B' of this type are symplectically similar if and only if they are similar.

Proof. In this case we have only one trivial hermitian invariant, namely an anti-symmetric non-degenerate inner product γ on $\tilde{W} = W/BW$.

There exists a basis of W , $\{\tilde{u}_i, \tilde{v}_i\}$ $i = 1, \dots, \mu$, such that γ takes the canonical form $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, or:

$$\gamma(\tilde{u}_i, \tilde{u}_i) = \gamma(\tilde{v}_i, \tilde{v}_i) = 0$$

$$\gamma(\tilde{u}_i, \tilde{v}_i) = \delta_{ii}.$$

If $\{u_i^{(0)}, v_i^{(0)}\}$ is any set of representatives of $\tilde{u}_i, \tilde{v}_i$ in W , each vector of the set generates a $2k + 1$ dimensional B -cyclic subspace, and:

$$\omega(u_i^{(0)}, B^{2k}v_i^{(0)}) = \delta_{ii},$$

$$\omega(u_i^{(0)} B^{2k}u_i^{(0)}) = \omega(v_i^{(0)} B^{2k}v_i^{(0)}) = 0.$$

Now we shall use the freedom we have in choosing representatives to construct a representative u_1 of $\tilde{u}$, and a representative v_1 of $\tilde{v}_1$ such that:

$$\omega(u_1, B^m v_1) = 0, \quad m = 0, 1, \dots, 2k - 1,$$

$$\omega(u_1, B^m u_1) = \omega(v_1, B^m v_1) = 0, \quad m = 1, \dots, 2k - 1.$$

Then, $\text{span}(u_1, Bu_1, \dots, B^{2k}u_1, v_1, Bv_1, \dots, B^{2k}v_1)$ is a *symplectic space* with $\text{span}(u_1, Bu_1, \dots, B^{2k}u_1)$ and $\text{span}(v_1, Bv_1, \dots, B^{2k}v_1)$ *isotropic*. So we can choose:

$$V^+_i = \text{span}(u_1, Bu_1, \dots, B^{2k}u_1)$$

$$V^-_i = \text{span}(v_1, Bv_1, \dots, B^{2k}v_1)$$

where $\text{span}(\dots)$ denotes the linear space spanned by the vectors in the brackets. Applying the same procedure to the symplectic B -invariant subspace $(V^+_1 V^-_1)^\perp$,

we construct $(V^+{}_2 V^-{}_2)$, and so on. So we have only to show that we can construct u_1, v_1 satisfying the above conditions.

We set:

$$\begin{aligned} u_1^{(1)} &= u_1^{(0)} + a^{(1)} B v_1^{(0)} \\ v_1^{(1)} &= v_1^{(0)} + b^{(1)} B u_1^{(0)} + c^{(1)} B v_1^{(0)} \end{aligned}$$

and determine $a^{(1)}, b^{(1)}, c^{(1)}$ from the conditions:

$$\begin{aligned} \omega(u_1^{(1)}, B^{2k-1} u_1^{(1)}) &= 0 \\ \omega(v_1^{(1)}, B^{2k-1} v_1^{(1)}) &= 0 \\ \omega(u_1^{(1)}, B^{2k-1} v_1^{(1)}) &= 0. \end{aligned}$$

A direct calculation shows that:

$$\begin{aligned} a^{(1)} &= -\frac{1}{2} \omega(u_1^{(0)}, B^{2k-1} u_1^{(0)}) \\ b^{(1)} &= \frac{1}{2} \omega(v_1^{(0)}, B^{2k-1} v_1^{(0)}) \\ c^{(1)} &= -\omega(u_1^{(0)}, B^{2k-1} v_1^{(0)}). \end{aligned}$$

Next, we construct a sequence: $\{u_1^{(p)}, v_1^{(p)}\}$

$$\begin{aligned} u_1^{(p)} &= u_1^{(p-1)} + a^{(p)} B^p v_1^{(p-1)}, \\ v_1^{(p)} &= v_1^{(p-1)} + b^{(p)} B^p u_1^{(p-1)} + c^{(p)} B^p v_1^{(p-1)} \end{aligned}$$

$p = 2, \dots, k-1$; the coefficients are determined from the conditions:

$$\omega(u_1^{(p)}, B^{2k-p} u_1^{(p)}) = \omega(v_1^{(p)}, B^{2k-p} v_1^{(p)}) = \omega(u_1^{(p)}, B^{2k-p} v_1^{(p)}) = 0.$$

These conditions give

$$\begin{aligned} a^{(p)} &= -\frac{1}{2} \omega(u_1^{(p-1)}, B^{2k-p} u_1^{(p-1)}) \\ b^{(p)} &= \frac{1}{2} \omega(v_1^{(p-1)}, B^{2k-p} v_1^{(p-1)}) \\ c^{(p)} &= -\omega(u_1^{(p-1)}, B^{2k-p} v_1^{(p-1)}). \end{aligned}$$

It is readily verified that:

$$\omega(u_1^{(p)}, B^{2k-q} u_1^{(p)}) = \omega(v_1^{(p)}, B^{2k-q} v_1^{(p)}) = \omega(u_1^{(p)}, B^{2k-q} v_1^{(p)}) = 0,$$

for $q = 1, \dots, p-1$; in other words, each element of the sequence we have constructed satisfies the relationships defining the elements which precede it in the sequence. It follows that we can set:

$$\begin{aligned} u_1 &= u_1^{(k-1)} \\ v_1 &= v_1^{(k-1)}. \end{aligned}$$

This proves the first part of the theorem.

The second statement follows immediately from taking as a basis for W the ordered set:

$$\{u_1, Bu_1, \dots, B^{2k}u_1, v_1, Bv_1, \dots, B^{2k}v_1; \dots; \\ u_\mu, Bu_\mu, \dots, B^{2k}u_\mu, v_\mu, Bv_\mu, \dots, B^{2k}v_\mu\}.$$

The last statement follows trivially from the first two. $\blacksquare$

Theorem 3.8. *Let B have a single elementary divisor λ^{2k} with multiplicity ν . Then:*

- (i) *there exists at least one orthogonal direct sum decomposition of W into B -invariant irreducible symplectic subspaces:*

$$W = W_1 \perp W_2 \perp \dots \perp W_\nu$$

$$B = B_1 \oplus B_2 \oplus \dots \oplus B_\nu.$$

- (ii) *Suppose $K = \mathbf{R}$. There is a symplectic basis for W on which the matrix of B is a sum irreducible blocks of the form:*

$$\eta_i \left[\begin{array}{ccc|ccc} & & & & & 0 \\ & & & \mathbf{0} & & 0 \\ & & & & 0 & 1 \\ & \mathbf{0} & & & \cdot & \\ & & & 0 & \cdot & 1 \\ & & & & \cdot & \\ & & & & \cdot & \\ & & & 1 & & \mathbf{0} \\ \hline & & & & & \\ & \mathbf{0} & & 1 & & \\ & & \cdot & & & \\ & 1 & \mathbf{0} & & & \\ & & & & \mathbf{0} & \end{array} \right] \quad i = 1, \dots, \nu$$

where $\eta_i = \pm 1$ and the number of $+1$'s equals the index of the hermitian invariant associated to λ^{2k} .

- (iii) *Any two transformations B, B' of this type are symplectically equivalent if and only if they are similar and have equivalent hermitian invariants.*

Proof. B has a symmetric hermitian invariant γ . There is a basis of $\tilde{W} = W/BW$, $\{\bar{e}_i\}$, $i = 1, \dots, \nu$, such that:

$$\gamma(\bar{e}_i, \bar{e}_i) = \eta_i \delta_{ii},$$

$\eta_i \in K^* = (K - \{0\})$. Moreover, if $K = \mathbf{R}$ then $\eta_i = \pm 1$.

If $\{e_i^{(0)}\}$ is an arbitrary set of representatives of the $\bar{e}_i$'s, each of them generates a cyclic subspace under the action of B , and:

$$\omega(e_i^{(0)}, B^{2k-1}e_i^{(0)}) = \eta_i \delta_{ii}$$

As in the preceding theorem, we select an appropriate representative e_1 of $\bar{e}_1$; e_1 shall satisfy the condition:

$$\omega(e_1, B^m e_1) = 0, \quad m = 1, 2, \dots, 2k - 2.$$

Then, $\text{span}(e_1, B e_1, \dots, B^{2k-1} e_1)$ is a symplectic space, obviously B -invariant, which we can choose as W_1 : repeating this process for $W_1^\perp$ we construct W_2 and so on, proving (i). Thus we have only to show how to select e_1 . The sequence:

$$e_1^{(p)} = e^{(p-1)} + a^{(p)} B^p e_1^{(p-1)}, \quad p = 1, \dots, 2k - q$$

with:

$$a^{(p)} = -\frac{1}{2} \omega(e_1^{(p-1)}, B^{2k-p-1} e_1^{(p-1)}),$$

has the property

$$\omega(e_1^{(p)}, B^{2k-q-1} e_1^{(p)}) = 0, \quad q = 1, \dots, p - 1.$$

Hence we can choose:

$$e_1 = e_1^{(2k-2)}.$$

To prove (ii), we observe that the set:

$$\begin{aligned} u_{1,i} &= \eta_i e_i, & u_{2,i} &= \eta_i B^2 e_i, \dots, u_{k,i} = \eta_i B^{2k-2} e_i \\ v_{1,i} &= B^{2k-1} e_i, & v_{2,i} &= B^{2k-3} e_i, \dots, v_{k,i} = B e_i \end{aligned}$$

$i = 1, \dots, \nu$, is a *symplectic basis*. It is easy to compute the action of B on this basis:

$$\begin{aligned} B u_{1,i} &= \eta_i v_{k,i}, & B u_{2,i} &= v_{k-1,i}, \dots, B u_{k,i} = \eta_i v_1 \\ B v_{1,i} &= 0, & B v_{2,i} &= \eta_i u_{k,i}, \dots, B v_{k,i} = \eta_i u_{2,i}. \end{aligned}$$

We obtain a canonical matrix as in (ii).

The third statement is an immediate consequence of the first two. ■

Theorem 3.9. *Let B have a single elementary divisor: $p^k(\lambda)$ such that: $p(\lambda) = p(-\lambda)$, occurring with multiplicity ν . Then:*

(i) *there exists at least one orthogonal decomposition of W into B -invariant irreducible symplectic subspaces:*

$$\begin{aligned} W &= W_1 \perp W_2 \perp \dots \perp W_\nu, \\ B &= B_1 \oplus B_2 \oplus \dots \oplus B_\nu. \end{aligned}$$

(ii) *Let $K = \mathbf{R}$. There is a symplectic basis with respect to which B has a matrix which is a direct sum of irreducible blocks of the form:*

$$\eta_i \left[\begin{array}{ccc|ccc} & & & & & -a^2 \\ & & & 0 & & \\ & & & & -a^2 & 1 \\ & 0 & & & \ddots & \\ & & & & \ddots & 0 \\ & & & -a^2 & 1 & \\ \hline & & 1 & & & \\ & 0 & & & & \\ & & 1 & & & \\ & & & \ddots & & \\ & & & \ddots & & \\ 1 & & & & 0 & \end{array} \right]$$

$$i = 1, \dots, \nu$$

with $\eta_i = \pm 1$, and the number of +1's equals the index of the hermitian invariant associated to $p^k(\lambda)$.

- (iii) Any two transformations B, B' of this type are symplectically equivalent if and only if they are similar and have equivalent hermitian invariants.

Proof. The only hermitian invariant of B is (H, h) ,

$$H = W/pW, \quad h(\bar{u}, \bar{v}) = \omega(Bp^{k-1}(B)u, v) + \omega(p^{k-1}(B)u, v)\bar{B}.$$

There is a basis of H , $\{\bar{x}_i\}$, such that:

$$h(\bar{x}_i, \bar{x}_i) = \eta_i \delta_{ii}, \quad n_i \in K^*,$$

with $\eta_i = \pm 1$ if $K = \mathbf{R}$. Let $x_1^{(0)}$ be an arbitrary representative of x_1 . Then $x_1^{(0)}$ generates a cyclic subspace and:

$$\begin{aligned} \omega(p^{k-1}(B)x_1^{(0)}, x_1^{(0)}) &= 0 \\ \omega(Bp^{k-1}(B)x_1^{(0)}, x_1^{(0)}) &= \eta_1. \end{aligned}$$

We construct a representative $x_1 \in W$ of $\bar{x}_1$ such that:

- (1) $\omega(p^l(B)x_1, x_1) = 0, \quad l = 1, \dots, k-2$
 (2) $\omega(Bp^m(B)x_1, x_1) = 0, \quad m = 0, 1, \dots, k-2.$

While conditions (1) are automatically satisfied for any x due to the particular form of p (in fact, $p(\lambda) = p(-\lambda)$, $p(\lambda) = p(\lambda^2)$), conditions (2) are non-trivial. As in the preceding theorems we obtain x_1 as the last term of a sequence. In this case we set:

$$x_1^{(r)} = x_1^{(r-1)} + a^{(r)} p^r(B)x_1^{(r-1)}, \quad r = 1, \dots, k-1,$$

with

$$a^{(k)} = -\frac{1}{2}\omega(p^{k-r-1}(B)Bx_1^{(0)}, x_1^{(0)}).$$

The sequence has the property that:

$$\omega(p^{k-s-1}(B)Bx_1^{(r)}, x_1^{(r)}) = 0$$

for $s = 1, \dots, r$. Thus:

$$x_1 = x_1^{(k-1)}$$

has the desired property.

It is easily verified that the cyclic subspace $\langle x_1 \rangle$ generated by x_1 under B is *symplectic* and is:

$$x_1 = \text{span}(p'(B)x_1, Bp'(B)x_1; l = 0, 1, \dots, k-1).$$

Hence we can take

$$W_1 = \langle x_1 \rangle$$

and the first statement follows easily as in the preceding theorems.

A canonical matrix for $B|W_1$, which has the form of (ii), is obtained by choosing the basis:

$$\begin{aligned} u_1 &= \eta_1 x_1, & v_1 &= p^{k-1}(B)x_1 \\ u_2 &= \eta_1 p(B)x_1, & v_2 &= p^{k-2}(B)x_1 \\ &\dots\dots\dots, & & \dots\dots\dots \\ u_k &= \eta_1 p^{k-1}(B)x_1, & v_k &= Bx_1, \end{aligned}$$

and computing the action of B on it.

The third statement follows from (i) and (ii). ■

With this theorem we have proved all the statements of Theorem 2.2.

ACKNOWLEDGMENTS

I wish to thank Professor Giles for his constant help at every stage of this work, Professor Coleman for a careful reading of the manuscript and stimulating comments, and Professor T. A. Springer for a private communication.

REFERENCES

1. J. WILLIAMSON, *On the algebraic problem concerning the normal forms of linear dynamical systems*, Amer. J. of Math. **58** (1936), 141-163.
2. J. MILNOR, *On isometries of inner product spaces*. Inventiones mat. **8** (1969), 83-97.
3. T. A. SPRINGER, R. STEINBERG, *Seminar on algebraic groups and related finite groups*, E. conjugacy classes, Ch. 4, §2.
4. G. E. WALL, *On the conjugacy classes in the unitary, symplectic and orthogonal groups*, J. Austr. Math. Soc. **3** (1963), 1-62.
5. A. CIAMPI, *Classical hamiltonian linear systems*, Queen's Paper in Pure and Applied Mathematics (to appear).
6. E. ARTIN, *Geometric algebra*, Interscience Publishers, New York, 1957.
7. N. JACOBSON, *Lectures in abstract algebra, Vol. II, Linear algebra*, Van Nostrand, New York, 1953.

Queen's University, Kingston, Ontario
Date communicated: DECEMBER 1, 1972