A FACTORIZATION IN $GSp(V)$

C. RYAN VINROOT

Abstract. Let V be a vector space over the field F such that $\mathrm{char}(F) \neq 2$, and let V have a skew-symmetric nondegenerate bilinear form. Wonenburger proved that any element g of $Sp(V)$ is the product of two skew-symplectic involutions. Let $GSp(V)$ be the group of general similitudes with similitude character μ. We give a generalization of Wonenburger's result in the following form. Let $g \in GSp(V)$ with $\mu(g) = \beta$. Then $g = t_1 t_2$ such that t_1 is a skew-symplectic involution, and t_2 is such that $t_2^2 = \beta I$ and $\mu(t_2) = -\beta$. One application that follows from this result is a necessary and sufficient condition for an element of $GL(V)$ to be conjugate to a scalar multiple of its inverse. Another result is that we find an extension of the group $Sp(2n, \mathbb{F}_q)$, for $q \equiv 3 (\mathrm{mod}\ 4)$, all of whose complex representations have real-valued characters.

2000 *AMS Subject Classification:* 15A23, 20G40
Key words and phrases: similitude group, symplectic group, factorization of matrices, conjugacy classes, real-valued characters.

1. Introduction

It was proven by Wonenburger [10] and Djoković [1] that an invertible matrix is conjugate to its inverse if and only if it is the product of two involutions. This brought about the question, if a matrix has determinant ± 1, can it be factorized as a product of involutions, and if so, what is the fewest number of involutions needed in a factorization? This question was answered by Gustafson, Halmos, and Radjavi [7], who showed that any matrix with determinant ± 1 can be written as a product of 4 involutions, and that not all matrices of determinant ± 1 are a product of 3 involutions.

Wonenburger proved that in the case of the orthogonal group, every element is the product of two orthogonal involutions. In the case of the symplectic group, however, the two involutions in a factorization are not necessarily symplectic, but rather the following result is obtained [10, Theorem 2].

Theorem 1. *Let $G = Sp(2n, F)$ where $\mathrm{char}(F) \neq 2$. Then every element of $g \in G$ may be written $g = h_1 h_2$, where h_1 and h_2 are skew-symplectic involutions.*

Now consider the group of similitudes of an F-vector space, $GSp(2n, F)$, with similitude character μ. Our main result is the following, which generalizes Theorem 1.

Theorem 2. *Let $g \in GSp(2n, F)$ and $\mu(g) = \beta$, and suppose* $\mathrm{char}(F) \neq 2$. *Then $g = t_1 t_2$, where t_1 is a skew-symplectic involution, and where t_2 is such that $\mu(t_2) = -\beta$ with $t_2^2 = \beta I$.*

We note that Theorem 2 follows from Theorem 1 directly if β is a square in F, but the nonsquare case is not immediate.

One application of Theorem 2 is a necessary and sufficient condition for a linear transformation to be conjugate to a scalar multiple of its inverse. Specifically, the result in Theorem 3 is that an element of $GL(n, F)$, $\mathrm{char}(F) \neq 2$, is conjugate to λ times its inverse for some $\lambda \in F^\times$, if and only if it is the product of an involution and an element whose square is λI.

Another application of Theorem 2 is motivated by a result of a factorization of matrices given by R. Gow [3]. Gow proved that any invertible matrix is the product of an involution with a symmetric matrix. He then observed [5, Theorem 1] that it follows that every element of the split extension of $GL(n, \mathbb{F}_q)$ by the transpose-inverse automorphism is conjugate to its inverse. That is, if $G = GL(n, \mathbb{F}_q)$, then all of the irreducible complex representations of the group

$$G^+ = \langle G, \tau \mid \tau^2 = I, \tau^{-1} g \tau = {}^t g^{-1} \text{ for every } g \in G \rangle$$

have real-valued characters.

In our situation, we consider the group $G = Sp(2n, \mathbb{F}_q)$ with $q \equiv 3 (\mathrm{mod}\ 4)$, and the order 2 automorphism ι of G defined by

$${}^\iota g = \begin{pmatrix} -I_n & \\ & I_n \end{pmatrix} g \begin{pmatrix} -I_n & \\ & I_n \end{pmatrix}.$$

The result we give in Theorem 4 is that every element of the group

$$G^{\iota, -I} = \langle G, \tau \mid \tau^2 = -I, \tau^{-1} g \tau = {}^\iota g \text{ for every } g \in G \rangle,$$

is conjugate to its inverse.

The author would like to thank Daniel Bump, David Ginzburg, and the referee for helpful comments and suggestions.

2. Initial Reduction

Throughout, we assume that V is a $2n$-dimensional F-vector space and $\mathrm{char}(F) \neq 2$. Suppose V has a fixed nondegenerate skew-symmetric form, $\langle \cdot, \cdot \rangle : V \times V \to F$. The *general symplectic group*, (or *group of similitudes of* $\langle \cdot, \cdot \rangle$) is defined as $GSp(2n, F) = \{g \in GL(2n, F) : \langle gv, gw \rangle = \mu(g) \langle v, w \rangle$ for some $\mu(g) \in F^\times$ for all $v, w \in V\}$. The function $\mu : GSp(2n, F) \to F^\times$ is a multiplicative character called the *similitude character*. Then the *symplectic group* $Sp(2n, F)$ is the subgroup of $GSp(2n, F)$ which is the kernel of μ, leaving the inner product invariant. We will also write $GSp(V) = GSp(2n, F)$ and $Sp(V) = Sp(2n, F)$.

Suppose $g \in GSp(V)$, $\mu(g) = \beta$, and β is a square in F, say $\gamma^2 = \beta$. Then $\mu(\gamma I) = \beta$, and so $\gamma^{-1} g \in Sp(2n, F)$. Then we can write $\gamma^{-1} g = h_1 h_2$, where h_1 and h_2 are both skew-symplectic involutions, applying Theorem

1. Multiplying both sides by γ, we have $g = h_1(\gamma h_2)$, where $t_1 = h_1$ is a skew-symplectic involution and $t_2 = \gamma h_2$ is such that $\mu(t_2) = -\beta$ and $t_2^2 = \beta I$. So in this case, Theorem 2 follows directly from Theorem 1. We may therefore assume that the g we start with in Theorem 2 is such that $\mu(g)$ is not a square in F. If $g \in GSp(2n, F)$ is such that $\mu(g) = \beta$, then we call g *β-symplectic*, and from now on we fix a β which is not a square in F.

Let V be a $2n$-dimensional F-vector space with nondegenerate skew-symmetric bilinear form $\langle \cdot, \cdot \rangle$. If J is the matrix representing this form, we have, for a β-symplectic g, ${}^t gJg = \beta J$. Then $JgJ^{-1} = \beta\, {}^t g^{-1}$, and since ${}^t g$ and g are conjugate in $GL(2n, F)$, we have g is conjugate to βg^{-1} over $GL(2n, F)$. Therefore, for a β-symplectic g, we have g and βg^{-1} have the same minimal polynomial. We thus introduce the *β-adjoint* of a polynomial in $F[x]$. The development in this section will follow parts of [10] and [8].

If $f(x) \in F[x]$ is such that $f(0) \neq 0$, and $\deg(f) = d$, define the *β-adjoint* of $f(x)$, written $\hat{f}(x)$, to be

$$\hat{f}(x) = f(0)^{-1} x^d f(\beta/x).$$

Then $\deg(\hat{f}) = d$, and the roots of $\hat{f}$ are β times the reciprocals of the roots of f. The relevance of the definition to our situation is made clear in the following proposition.

Proposition 1. *Let $g \in GL(V)$ and suppose $m(x)$ is the monic minimal polynomial of g. Then the element βg^{-1} has monic minimal polynomial $\hat{m}(x)$.*

Proof. For any monic $f(x)$, we have $\hat{\hat{f}}(x) = f(x)$. Also, if $f(x)$ factors as $f(x) = f_1(x) f_2(x)$ and $f(x)$ is monic, then we also have the factorization $\hat{f}(x) = \hat{f}_1(x)\hat{f}_2(x)$. So, a monic polynomial $f(x)$ is irreducible if and only if $\hat{f}(x)$ is irreducible. Now, for g with monic minimal polynomial $m(x)$, we have $\hat{m}(\beta g^{-1}) = m(0)^{-1}\beta^d g^{-d} m(g) = 0$, where $\deg(m) = d$. So βg^{-1} satisfies $\hat{m}(x)$, and $\hat{m}(x)$ is irreducible, so $\hat{m}(x)$ must be the monic minimal polynomial of βg^{-1}. $\square$

So when g is β-symplectic, since g and βg^{-1} have the same monic minimal polynomial $f(x)$, we have $f(x) = \hat{f}(x)$ by Proposition 1. We will call such a monic polynomial *self-β-adjoint*. Then if $f(x)$ is self-β-adjoint, for every root γ of $f(x)$ in an algebraic closure $\overline{F}$ of F, $f(x)$ also has $\beta\gamma^{-1}$ as a root. If $p(x)$ is the minimal polynomial of γ in $F[x]$, then by the same argument as in the proof of Proposition 1, $\hat{p}(x)$ is the minimal polynomial of $\beta\gamma^{-1}$ in $F[x]$. Since both are irreducible in $F[x]$, then either they are relatively prime and both are divisors of $f(x)$, or they are equal to the same polynomial, which is also a divisor of $f(x)$. We have the following factorization of a self-β-adjoint

polynomial $f(x)$ into irreducibles:

$$f(x) = \prod_{i=1}^{k} (p_i(x)\hat{p}_i(x))^{n_i} \prod_{j=1}^{l} q_j(x)^{m_j} = \prod_{i=1}^{k+l} r_i(x)^{s_i}, \tag{1}$$

where the $p_i(x)$ and $q_j(x)$ are irreducible, $q_j(x)$ are self-β-adjoint but $p_i(x)$ are not, and $r_i(x) = p_i(x)\hat{p}_i(x)$ for $i = 1, \ldots, k$, $r_{k+j}(x) = q_j(x)$ for $j = 1, \ldots, l$, and the $r_i(x)$ are all distinct.

For a β-symplectic g, we have for any $v, w \in V$,

$$\langle gv, gw \rangle = \beta \langle v, w \rangle$$

and so

$$\langle \beta g^{-1} v, w \rangle = \langle v, gw \rangle. \tag{2}$$

Furthermore, we have the following.

Lemma 1. *Let g be β-symplectic and let $r(x)$ be any polynomial in $F[x]$. Then for any $v, w \in V$, we have*

$$\langle v, r(g)w \rangle = \langle r(\beta g^{-1})v, w \rangle.$$

Proof. Let $r(x) = \sum_{i=1}^{d} a_i x^d$. Then

$$\langle v, r(g)w \rangle = \sum_{i=1}^{d} a_i \langle v, g^i w \rangle = \sum_{i=1}^{d} a_i \langle (\beta g^{-1})^i v, w \rangle = \langle r(\beta g^{-1})v, w \rangle,$$

the second equality coming from Equation 2. $\square$

Lemma 2. *Let g be β-symplectic and $r(x)$ any polynomial in $F[x]$ satisfying $r(0) \neq 0$. Then the subspaces $\mathrm{im}(r(g))$ and $\ker(\hat{r}(g))$ are mutually orthogonal.*

Proof. Let $u \in \mathrm{im}(r(g))$, where $u = r(g)w$, and let $v \in \ker(\hat{r}(g))$. Then, if $d = \deg(r)$,

$$\langle v, u \rangle = \langle v, r(g)w \rangle = \langle r(\beta g^{-1})v, w \rangle = \beta^{-d} \langle g^d r(\beta g^{-1})v,\, g^d w \rangle$$

$$= \beta^{-d} r(0) \langle \hat{r}(g)v, g^d w \rangle = \beta^{-d} r(0) \langle 0, g^d w \rangle = 0,$$

where we have applied, respectively, Lemma 1, the definition of β-symplectic, the definition of $\hat{r}(x)$, and the fact that $v \in \ker(\hat{r}(g))$. $\square$

Proposition 2. *Let g be a β-symplectic transformation of V with minimal polynomial $m(x)$. Let $m(x) = \prod_i r_i(x)^{s_i}$ be the factorization of $m(x)$ into self-β-adjoint polynomials as in Equation (1). Then the direct sum*

$$V = \bigoplus_i \ker(r_i(g)^{s_i})$$

is a decomposition into nondegenerate mutually orthogonal g-invariant subspaces.

Proof. Since the $r_i(x)^{s_i}$ are pairwise relatively prime, we may write V as the direct sum $V = \bigoplus_i \ker(r_i(g)^{s_i})$, and these subspaces are g-invariant. We have $\mathrm{im}(r_j(g)^{s_j}) \subseteq \bigoplus_{i\neq j} \ker(r_i(g)^{s_i})$, since if $v = (r_j(g)^{s_j})w$, we have $(\prod_{i\neq j} r_i(g)^{s_i})v = m(g)w = 0$. But the dimensions of $\mathrm{im}(r_j(g)^{s_j})$ and $\bigoplus_{i\neq j} \ker(r_i(g)^{s_i})$ are both equal to $\dim(V) - \dim(\ker(r_j(g)^{s_j})$, and so we have

$$(*) \qquad \mathrm{im}(r_j(g)^{s_j}) = \bigoplus_{i\neq j} \ker(r_i(g)^{s_i}).$$

Now $r_i(x) = \hat{r}_i(x)$, so $\ker(\hat{r}_i(x)^{s_i}) = \ker(r_i(x)^{s_i})$. From Lemma 2, we then have $\ker(r_i(x)^{s_i})$ and $\mathrm{im}(r_i(x)^{s_i})$ are mutually orthogonal. By (*), we have that the subspaces $V_i = \ker(r_i(g)^{s_i})$ are mutually orthogonal. Mutual orthogonality then implies that the inner product restricted to V_i is nondegenerate. □

So, by Proposition 2, given any β-symplectic transformation g of V, we may write $g = \bigoplus_i g_i$, where each g_i is g restricted to $V_i = \ker(r_i(g)^{s_i})$, and g_i is β-symplectic on the space V_i. Then g_i has minimal polynomial $r_i(x)^{s_i}$. If each g_i then satisfies Theorem 2, that is, if $g_i = t_{i1}t_{i2}$, where t_{i1} is skew-symplectic with $t_{i1}^2 = I$, and t_{i2} is such that $\mu(t_{i2}) = -\beta$ with $t_{i2}^2 = \beta I$, then g will satisfy Theorem 2 by taking $t_1 = \bigoplus_i t_{i1}$ and $t_2 = \bigoplus_i t_{i2}$. From the factorization in Equation (1), each $r_i(x)$ is either of the form $p(x)\hat{p}(x)$, where $p(x)$ is irreducible, or is an irreducible polynomial $q(x)$ satisfying $q(x) = \hat{q}(x)$.

3. Technical Lemmas

The proofs of the lemmas in this section are almost exactly the same as in Wonenburger, [10, Lemmas 1,3, and 4], respectively.

Recall that a vector space V is *cyclic* with respect to a transformation g if there exists a vector $v \in V$ such that V is spanned by vectors of the form $g^i v$, and call g a *cyclic transformation* if V is cyclic with respect to g. The first lemma reduces the case of a g with minimal polynomial $(p(x)\hat{p}(x))^s$ to cyclic transformations. We will say a vector v has *order* $r(g)^k$ when $r(g)^k v = 0$, but $r(g)^{k-1}v \neq 0$, where $r(x) \in F[x]$.

Lemma 3. *Let g be a β-symplectic transformation for V such that the minimal polynomial of g is $(p(x)\hat{p}(x))^s$, where $p(x)$ and $\hat{p}(x)$ are distinct irreducible polynomials. Then V can be decomposed as a direct sum of mutually orthogonal subspaces which are cyclic with respect to g.*

Proof. First, we have $V = \ker(p(g)^s) \oplus \ker(\hat{p}(g)^s)$, since $p(x)$ and $\hat{p}(x)$ are relatively prime. Then $\ker(p(g)^s) = \mathrm{im}(\hat{p}(g)^s)$, which is mutually orthogonal with $\ker(p(g)^s)$ by Lemma 2, so each subspace is totally isotropic. Consider an element $u \in \ker(p(g)^s)$ of order $p(g)^s$, which exists since the minimal polynomial of g restricted to $\ker(p(g)^s)$ is $p(x)^s$. Then $p(g)^{s-1}u \neq 0$, and

so there exists an element $v \in \ker(\hat{p}(g)^s)$ such that $\langle p(g)^{s-1}u, v\rangle \neq 0$, since $\ker(p(g)^s)$ is totally isotropic. Then, if $\deg(p) = d$, we have

$$(**) \qquad \begin{array}{rcl} 0 \neq \langle p(g)^{s-1}u, v\rangle & = & \langle u,\, p(\beta g^{-1})^{s-1}v\rangle \\ & = & \beta^{-d(s-1)}p(0)^{s-1}\langle g^{d(s-1)}u,\, \hat{p}(g)^{s-1}v\rangle. \end{array}$$

Therefore v has order $\hat{p}(g)^s$. This means $u+v$ has order $(p(g)\hat{p}(g))^s$.

Let W be the cyclic subspace generated by $u+v$, spanned by vectors of the form $g^i(u+v)$. We will show that W is a nondegenerate subspace of V. Let $w \in W$ be a nonzero element of W. Then there is some polynomial $f(x) \in F[x]$ such that $w = f(g)(u+v)$. We factor $f(x)$ in two different ways: factor out all powers of $p(x)$, and factor out all powers of $\hat{p}(x)$. Then we write

$$w = g^m q_1(g)p(g)^k u + g^m q_2(g)\hat{p}(g)^l v,$$

where $q_1(0) \neq 0$, $q_2(0) \neq 0$, $p(x)$ is relatively prime to $q_1(x)$, and $\hat{p}(x)$ is relatively prime to $q_2(x)$. One of these terms is nonzero, so suppose $g^m q_1(g)p(g)^k u \neq 0$. Since $q_1(x)$ and $p(x)$ are relatively prime, then we must have $\hat{q}_1(x)$ is relatively prime to $\hat{p}(x)$, and thus to $\hat{p}(x)^s$. We may then find polynomials $a(x)$ and $b(x)$ such that

$$a(x)\hat{q}_1(x) + b(x)\hat{p}(x)^s = 1.$$

Then we have $a(g)\hat{q}_1(g)v = v$, since $v \in \ker(\hat{p}(g)^s)$.

Now let $y = g^\rho a(g)\hat{p}(g)^{s-k-1}v$, where $\rho = m + \deg(q_1(x)) + d(k-s+1)$ (recall that $d = \deg(p(x))$). Since $v \in \ker(\hat{p}(g)^s)$, which is a g-invariant subspace, then we have $y \in \ker(\hat{p}(g)^s)$ and $g^m q_2(g)\hat{p}(g)^l v \in \ker(\hat{p}(g)^s)$. Then since $\ker(\hat{p}(g)^s)$ is totally isotropic, we have $\langle g^m q_2(g)\hat{p}(g)^l v, y\rangle = 0$. Now we compute $\langle w, y\rangle$:

$$\begin{array}{rcl} \langle w, y\rangle & = & \langle g^m q_1(g)p(g)^k u,\; g^\rho a(g)\hat{p}(g)^{s-k-1}v\rangle \\ & = & \langle u,\; \beta^m g^{\rho-m} q_1(\beta g^{-1})a(g)p(\beta g^{-1})^k \hat{p}(g)^{s-k-1}v\rangle. \end{array}$$

We now apply to both sides of the bilinear form a factor of $g^{dk+\deg(q_1(x))-\rho+m} = g^{d(s-1)}$, which will put a factor of $\beta^{-d(s-1)}$ in front. The goal of this is to get a $\hat{q}_1(g)$ and a $\hat{p}(g)^k$ on the right-hand side of the bilinear form, for which we also need a factor of $q_1(0)^{-1}$ and a $p(0)^{-k}$. We can then apply the fact that $a(g)\hat{q}_1(g)v = v$. Doing this, we have

$$\begin{array}{rcl} \langle w, y\rangle & = & q_1(0)p(0)^k\beta^{m-d(s-1)}\langle g^{d(s-1)}u,\; \hat{p}(g)^{s-1}a(g)\hat{q}_1(g)v\rangle \\ & = & q_1(0)p(0)^k\beta^{m-d(s-1)}\langle g^{d(s-1)}u,\; \hat{p}(g)^{s-1}v\rangle \neq 0. \end{array}$$

The last expression is not zero because of (**). We therefore have $\langle w, y\rangle \neq 0$. If instead we had $g^m q_1(g)p(g)^k u = 0$, and $w = g^m q_2(g)\hat{p}(g)^l v$, we go through a similar computation, and let $y' = g^{\rho'}a'(g)p(g)^{s-l-1}u$, where $\rho' = m + \deg(q_2(x)) + dl$, and $a'(g)q_2(g)u = u$. Then $\langle p(g)^{s-l}u, v\rangle \neq 0$ implies that $\langle w, y'\rangle \neq 0$.

In any case, we have that W is a nondegenerate cyclic subspace of V. Then we can write $V = W \oplus W^\perp$. Then g restricted to $W^\perp$ is again a β-symplectic transformation with minimal polynomial a power of $p(x)\hat{p}(x)$,

and we can continue with this process. Doing this, we have decomposed V as a direct sum of cyclic mutually orthogonal subspaces. $\square$

The next two lemmas deal with the case that the minimal polynomial of g is of the form $q(x)^s$.

Lemma 4. *Let g be a β-symplectic transformation for V such that the minimal polynomial of g is of the form $q(x)^s$, where $q(x)$ is an irreducible self-β-adjoint polynomial. Let $u \in V$ have order $q(g)^s$. Then the subspace U cyclically generated by u and g is nondegenerate if and only if there exists a $v \in U$ such that $\langle q(g)^{s-1}u, v\rangle \neq 0$.*

Proof. The "only if" statement follows directly from the definition of nondegenerate. So suppose there is a $v \in U$ such that $\langle q(g)^{s-1}u,\, v\rangle \neq 0$. Then $v = h(g)u$ for some polynomial $h(x) \in F[x]$. Let w be any nonzero vector in U, then $w = g^m f(g)q(g)^k u$, where $f(x)$ and $q(x)$ are relatively prime, $f(0) \neq 0$, and $k < s$, which is obtained by factoring the polynomial in g when expressing w in terms of u.

Since $q(x)$ and $f(x)$ are relatively prime, then $\hat{q}(x) = q(x)$ and $\hat{f}(x)$ are relatively prime. So we may find polynomials $a(x)$ and $b(x)$ such that

$$a(x)\hat{f}(x) + b(x)q(x)^s = 1.$$

This implies that $a(g)\hat{f}(g)u = u$.

Now let $y = g^\rho a(g)h(g)q(g)^{s-k-1}u$, where

$$\rho = \deg(f(x)) - (s-k-1)\deg(q(x)) + m,$$

and recall that $v = h(g)u$. We compute $\langle w, y\rangle$:

$$\begin{aligned}
\langle w, y\rangle &= \langle g^m f(g)q(g)^k u,\ g^\rho a(g)h(g)q(g)^{s-k-1}u\rangle \\
&= \langle g^m q(g)^k q(\beta g^{-1})^{s-k-1}u,\ g^\rho h(g)f(\beta g^{-1})a(g)u\rangle \\
&= \beta^{\rho-\deg(f(x))}q(0)^{s-k-1}f(0)\langle q(g)^{s-1}u,\ h(g)a(g)\hat{f}(g)u\rangle \\
&= \beta^{\rho-\deg(f(x))}q(0)^{s-k-1}f(0)\langle q(g)^{s-1}u,\ v\rangle \neq 0.
\end{aligned}$$

So for any $w \in U$, we have found a $y \in U$ such that $\langle w, y\rangle \neq 0$, and so U is nondegenerate. $\square$

Lemma 5. *Let g be a β-symplectic transformation for V such that the minimal polynomial for g is $q(x)^s$, where $q(x)$ is an irreducible self-β-adjoint polynomial. Then either*

(a) there exists a vector u of order $q(g)^s$ which generates a nondegenerate cyclic subspace U, or

(b) there exist vectors u and v of order $q(g)^s$ which generate cyclic subspaces U and U' respectively, such that $U \cap U' = \{0\}$ and such that $U \oplus U'$ is nondegenerate.

Proof. Let $u \in V$ have order $q(g)^s$, and suppose that the cyclic subspace U generated by u is degenerate. By applying Lemma 4, for a $v \in V$ such that

$$\langle q(g)^{s-1}u, v\rangle \neq 0,$$

we must have $v \notin U$. This also implies that v has order $q(g)^s$, since

$$\begin{aligned} \langle q(g)^{s-1}u, v\rangle &= \langle u,\, q(\beta g^{-1})^{s-1}v\rangle \\ &= q(0)^{s-1}\beta^{-(s-1)\deg(q(x))}\langle g^{(s-1)\deg(q(x))}u,\, q(g)^{s-1}v\rangle \neq 0. \end{aligned}$$

Let U' be the cyclic subspace generated by v, and suppose that $U \cap U' \neq \{0\}$. So let y be a nonzero element of $U \cap U'$. Then y may be written in terms of u and v:

$$y = f(g)q(g)^k u = f'(g)q(g)^{k'}v,$$

where $f(x)$ and $f'(x)$ are relatively prime to $q(x)$. Note that from these expressions for y, we have that y has order $q(g)^{s-k}$ and $q(g)^{s-k'}$, and so $k = k'$. Since $f(x)$ and $q(x)^s$ are relatively prime, we find polynomials $a(x)$ and $b(x)$ such that $a(x)f(x) + b(x)q(x)^s = 1$. This implies that $a(g)f(g)y = y$. Using the expression for y in terms of u, we compute

$$q(g)^{s-k-1}a(g)y = q(g)^{s-k-1}a(g)f(g)q(g)^k u = q(g)^{s-1}u.$$

Using the expression for y in terms of v, we have

$$q(g)^{s-k-1}a(g)y = q(g)^{s-k-1}a(g)f'(g)q(g)^k v = a(g)f'(g)q(g)^{s-1}v.$$

So now $q(g)^{s-1}u = a(g)f'(g)q(g)^{s-1}v$. We know $\langle q(g)^{s-1}u, v\rangle \neq 0$, and so

$$\begin{aligned} \langle q(g)^{s-1}u, v\rangle &= \langle a(g)f'(g)q(g)^{s-1}v,\ v\rangle \\ &= \langle q(g)^{s-1}v,\ a(\beta g^{-1})f(\beta g^{-1})v\rangle \neq 0 \end{aligned}$$

Since $w = a(\beta g^{-1})f(\beta g^{-1})v \in U'$, and $\langle q(g)^{s-1}v, w\rangle \neq 0$, we have by Lemma 4 that the cyclic subspace U' generated by v is nondegenerate.

Now suppose that $U \cap U' = \{0\}$. Take any $w \in U \oplus U'$, and write w in terms of u and v:

$$w = g^m f_1(g)q(g)^k u + g^{m'}f_2(g)q(g)^l v,$$

where, as usual, $f_1(0), f_2(0) \neq 0$, and $f_1(x)$, $f_2(x)$ are relatively prime to $q(x)$. Also suppose that $k \geq l$. Since $f_2(x)$ is relatively prime to $q(x)$, then $\hat{f}_2(x)$ is relatively prime to $q(x)^s$. Let $a(x), b(x) \in F[x]$ be such that $a(x)\hat{f}_2(x) + b(x)q(x)^s = 1$, which implies that $a(g)\hat{f}_2(g)u = u$. Now let

$$z = g^\rho a(g)q(g)^{s-l-1}u,$$

where $\rho = l \cdot \deg(q(x)) + \deg(f_2(x)) + m'$. Then we have

$$\begin{aligned} \langle z,\ g^{m'}f_2(g)q(g)^l v\rangle &= \langle g^\rho a(g)q(g)^{s-l-1}u,\ g^{m'}f_2(g)q(g)^l v\rangle \\ &= f_2(0)q(0)^l\beta^{m'}\langle q(g)^{s-1}a(g)\hat{f}_2(g)u,\ v\rangle \\ &= f_2(0)q(0)^l\beta^{m'}\langle q(g)^{s-1}u,\ v\rangle \neq 0, \end{aligned}$$

from our assumption on v at the beginning of the proof. On the other hand,

$$\langle z,\ g^m f_1(g)q(g)^k u\rangle = \langle g^\rho a(g)q(g)^{s-l-1}u,\ g^m f_1(g)q(g)^k u\rangle$$

$$= \langle q(g)^{s-1+k-l}u,\ q(0)^k(\beta g^{-1})^{\rho-k\cdot\deg(q(x))}g^m a(\beta g^{-1})f_1(g)u\rangle = 0,$$

which, when $k > l$, follows immediately since u has order $q(g)^s$, and when $k = l$, follows from Lemma 4, since U is assumed degenerate. So now we have

$$\langle z, w\rangle \neq 0,$$

which implies in this case that $U \oplus U'$ is nondegenerate.

In the case that $k < l$, first choose $a'(x), b'(x)$ such that $a'(x)\hat{f}_1(x) + b'(x)q(x)^s = 1$ and $a'(0), b'(0) \neq 0$. Then let

$$z = g^{\rho'} a'(g) q(q)^{s-k-1} v,$$

where $\rho' = m - (s-k-1)\deg(q(x)) - \deg(a'(x))$. A computation similar to the previous one again gives $\langle z, w\rangle \neq 0$. So now $U \oplus U'$ is nondegenerate. □

4. Proof of the Main Theorem

The remaining cases required to obtain Theorem 2 are given in the following proposition.

Proposition 3. *(i) Let g be a cyclic transformation for V (ignoring any inner product structure) such that the minimal polynomial for g is self-β-adjoint. Then $g = t_1 t_2$, where $t_1^2 = I$ and $t_2^2 = \beta I$.*

(ii) If g is taken to be β-symplectic in (i), then t_1 can be taken to be skew-symplectic and t_2 can be taken to satisfy $\mu(t_2) = -\beta$.

(iii) Let g be a β-symplectic transformation for V such that the minimal polynomial of g is $q(x)^s$, where $q(x)$ is an irreducible self-β-adjoint polynomial. Then $g = t_1 t_2$, where t_1 is a skew-symplectic involution, and t_2 satisfies $\mu(t_2) = -\beta$ and $t_2^2 = \beta I$.

Proof. (i): The minimal polynomial of g, $f(x)$, is self-β-adjoint, and we first factor $f(x) = r(x)(x^2 - \beta)^s$, where $r(x)$ is relatively prime to the irreducible $x^2 - \beta$, where we are assuming β is not a square in F. Then $r(x)$ is a self-β-adjoint polynomial of even degree $2m$. We may write $V = \ker(r(x)) \oplus \ker((x^2-\beta)^s)$, and consider the cases for the minimal polynomial of g being of the form $r(x)$ as above or $(x^2 - \beta)^s$ separately.

Case I. The transformation g is cyclic (for a space V), and has minimal polynomial $r(x)$ which is self-β-adjoint, where $r(x)$ is either relatively prime to $x^2 - \beta$ and has degree $2m$, or is of the form $(x^2 - \beta)^s$, where s is even. For simplicity, we will write $s = m$, although m is not necessarily even in the first case. Let $r(x) = \sum_{i=0}^{2m} a_i x^i$, where $a_{2m} = 1$ and $a_0 = \beta^m$, which also holds for $r(x) = (x^2 - \beta)^s$ since in this case $a_0 = (-\beta)^s = \beta^s$, since s is even. Then since $r(x) = \hat{r}(x) = r(0)^{-1} x^{2m} r(\beta/x)$, we have

$$\begin{aligned} \textstyle\sum_{i=0}^{2m} a_i x^i &= \beta^{-m} x^{2m} \textstyle\sum_{i=0}^{2m} a_i \beta^i x^{-i} \\ &= \textstyle\sum_{i=0}^{2m} a_i \beta^{i-m} x^{2m-i} \\ &= \textstyle\sum_{i=0}^{2m} a_{2m-i} \beta^{m-i} x^i \end{aligned}$$

and so we have

$$a_{2m-i} = \beta^{i-m} a_i.$$

Since g is cyclic, there is a vector $w \in V$ such that the vectors

$$w, gw, g^2w, \ldots, g^{2m-1}w$$

form a basis of V. Then also the vectors

$$g^{-m}v, g^{-(m-1)}v, g^{-(m-2)}v, \ldots, v, gv, \ldots, g^{m-1}v$$

form a basis for V, where $v = g^m w$. Consider the basis consisting of the vectors $v, (g+\beta g^{-1})v, \ldots, (g^{m-1}+\beta^{m-1}g^{-(m-1)})v, (g-\beta g^{-1})v, \ldots, (g^{m-1}-\beta^{m-1}g^{-(m-1)})v, (g^m-\beta^m g^{-m})v$, where linear independence boils down to the fact that $a_{2m} = 1$ and $a_0 = \beta^m$ in the minimal polynomial for g.

Let P be the subspace of V generated by vectors of the form $(g^i+\beta^i g^{-i})v$ for $0 \leq i < m$, and let Q be the subspace of V generated by vectors of the form $(g^i - \beta^i g^{-i})v$ for $0 < i \leq m$. Then $V = P \oplus Q$.

First we observe that the transformation $g - \beta g^{-1}$ maps the space P to Q. If i is such that $0 \leq i < m$, we see that

$$(g-\beta g^{-1})(g^i+\beta^i g^{-i})v = (g^{i+1} - \beta^{i+1}g^{-(i+1)})v - \beta(g^{i-1} - \beta^{i-1}g^{-(i-1)})v.$$

Since $i+1$ and $i-1$ are both no bigger than m, then the image above is always in Q. These image vectors in Q are linearly independent, and so we have $(g-\beta g^{-1})P = Q$.

Next we see that $g + \beta g^{-1}$ maps P to P. Let i be such that $0 \leq i < m$. Then

$$(g+\beta g^{-1})(g^i+\beta^i g^{-i})v = (g^{i+1} + \beta^{i+1}g^{-(i+1)})v + \beta(g^{i-1} + \beta^{i-1}g^{-(i-1)})v,$$

and the image is immediately seen to be in P, except for the case $i = m-1$, where we need for $(g^m+\beta^m g^{-m})v$ to be in P also. The minimal polynomial of g is $r(x) = \sum_{i=0}^{2m} a_i x^i$, and it was observed above that $a_{2m} = 1, a_0 = \beta^m$, and $a_{2m-i} = \beta^{i-m}a_i$. We have $\sum_{i=0}^{2m} a_i g^i v = 0$, and multiplying by g^{-m} gives $\sum_{i=0}^{2m} a_i g^{i-m}v = 0$. So now

$$(g^m+\beta^m g^{-m})v = -\sum_{i=1}^{2m-1} a_i g^{i-m}v = -a_m v - \sum_{i=1}^{m-1} a_i g^{i-m}v - \sum_{i=m+1}^{2m-1} a_i g^{i-m}v$$

$$= -a_m v - \sum_{i=1}^{m-1} a_{m-i}g^{-i}v - \sum_{i=1}^{m-1} a_{i+m}g^i v,$$

where the last equality is obtained by shifting the indices for each sum. In the first sum we now apply $a_{2m-j} = \beta^{j-m}a_j$ to $j = i+m$ to see that $a_{m-i} = \beta^i a_{i+m}$, and we finally obtain

$$(g^m+\beta^m g^{-m})v = -a_m v - \sum_{i=1}^{m-1} a_{m+i}(g^i+\beta^i g^{-i})v \in P,$$

as desired. So $(g+\beta g^{-1})P \subset P$.

We now have $(g-\beta g^{-1})P = Q$ and $(g+\beta g^{-1})P \subset P$, and so

$$(g-\beta g^{-1})Q = (g-\beta g^{-1})^2 P = [(g+\beta g^{-1})^2 - 4\beta I]P \subset P,$$

and

$$(g+\beta g^{-1})Q = (g+\beta g^{-1})(g-\beta g^{-1})P$$
$$= (g-\beta g^{-1})(g+\beta g^{-1})P \subset (g-\beta g^{-1})P \subset Q,$$

the last equality coming from the fact that $g-\beta g^{-1}$ and $g+\beta g^{-1}$ commute. Now let t_1 be the involution acting on V with P as its $+1$ eigenspace and Q as its -1 eigenspace. Since $V = P \oplus Q$, $(g-\beta g^{-1})P = Q$, and $(g-\beta g^{-1})Q \subset P$, we have $t_1(g-\beta g^{-1}) = -(g-\beta g^{-1})t_1$. Also, since $(g+\beta g^{-1})Q \subset Q$ and $(g+\beta g^{-1})P \subset P$, we have $t_1(g+\beta g^{-1}) = (g+\beta g^{-1})t_1$. Adding these two equations together, we obtain $2t_1 g = 2\beta g^{-1} t_1$, and since $\mathrm{char}(F) \neq 2$, we have $t_1 g = \beta g^{-1} t_1$, or $(t_1 g)^2 = \beta I$. So letting $t_2 = t_1 g$, we have $g = t_1 t_2$ with $t_1^2 = I$ and $t_2^2 = \beta I$.

Case II. The cyclic transformation g acting on V has minimal polynomial $(x^2-\beta)^s$, where $s = 2k+1$ is odd. Letting $(x^2-\beta)^s = \sum_{i=1}^{2s} a_i x^i$, we observe that $a_0 = -\beta^s$ and $a_{2s-i} = -\beta^{i-s} a_i$, and in particular $a_s = 0$.

We may take the following as a basis for V:

$$v, (g+\beta g^{-1})v, \ldots, (g^s + \beta^s g^{-s})v, (g-\beta g^{-1})v, \ldots, (g^{s-1} - \beta^{s-1} g^{-(s-1)})v,$$

where linear independence essentially comes from the observation above that $a_0 = -\beta^s$ and $a_{2s} = 1$ in the minimal polynomial for g. We must define bases for P and Q slightly differently from the previous case, because of the slight change in basis for V. Let P be the subspace of V with basis vectors of the form $(g^i + \beta^i g^{-i})v$ for $0 \leq i \leq s$, and let Q be the subspace of V with basis vectors of the form $(g^i - \beta^i g^{-i})v$ for $0 < i < s$.

We first show that $(g-\beta g^{-1})P = Q$. We calculate, as in the previous case, that for $0 \leq i \leq s$,

$$(g-\beta g^{-1})(g^i + \beta^i g^{-i})v = (g^{i+1} - \beta^{i+1} g^{-(i+1)})v - \beta(g^{i-1} - \beta^{i-1} g^{-(i-1)})v,$$

which can be seen to be in Q immediately unless $i = s-1$ or $i = s$. For $i = s-1$, we need to show that $(g^s - \beta^s g^{-s})v$ is an element of Q. By plugging g into its minimal polynomial and multiplying by g^{-s}, we have $\sum_{i=0}^{2s} a_i g^{i-s} = 0$. Now, $a_0 = 1$, $a_{2s} = -\beta^s$, $a_s = 0$, and $a_{s-i} = -\beta^i a_{s+i}$, which in the end gives

$$(g^s - \beta^s g^{-s})v = -\sum_{i=1}^{s-1} a_{s+i}(g^i - \beta^i g^{-i})v \in Q,$$

as desired. For the case $i = s$, it needs to be shown now that $(g^{s+1} - \beta^{s+1} g^{-(s+1)})v \in Q$. Observe that since $(g^2-\beta)^s v = (g^2-\beta)(g^2-\beta)^{2k} v = 0$, multiplying by $g^{-(2k+2)}(g^2+\beta)$ yields

$$g^{-(2k+2)}(g^4 - \beta^2)(g^2-\beta)^{2k} v$$
$$= g^{-2k}(g^2 - \beta^2 g^{-2})(g^4 - 2\beta g^2 + \beta^2)^k v$$
$$= (g^2 - \beta^2 g^{-2})(g^2 + \beta^2 g^{-2} - 2\beta)^k v = 0.$$

If we exchange the terms g^2 and $\beta^2 g^{-2}$ in the expression $(g^2 - \beta^2 g^{-2})(g^2 + \beta^2 g^{-2} - 2\beta)^k v$, it is negated. So in the expansion, the powers of g^2 and

$\beta^2 g^{-2}$ will have coefficients that are negatives of each other. So we have

$$(g^2 - \beta^2 g^{-2})(g^2 + \beta^2 g^{-2} - 2\beta)^k v = \sum_{i=1}^{k+1} c_i (g^{2i} - \beta^{2i} g^{-2i}) v = 0,$$

with $c_{k+1} = 1$, and so

$$(g^{s+1} - \beta^{s+1} g^{-(s+1)})v = -\sum_{i=1}^{k} c_i (g^{2i} - \beta^{2i} g^{-2i}) v \in Q.$$

So we finally have $(g - \beta g^{-1})P = Q$, equality coming from the fact that the images of $(g^i + \beta^i g^{-i})v$ for $0 \leq i \leq s-2$ are linearly independent, and $\dim(Q) = s - 1$.

We next show that $(g + \beta g^{-1})P \subset P$, and we calculate that for $0 \leq i \leq s$,

$$(g + \beta g^{-1})(g^i + \beta^i g^{-i})v = (g^{i+1} + \beta^{i+1} g^{-(i+1)})v + \beta(g^{i-1} + \beta^{i-1} g^{-(i-1)})v,$$

which is readily seen to be in P, except when $i = s$. For this case, we need to show that $(g^{s+1} + \beta^{s+1} g^{-(s+1)})v \in P$. Since $(g^2 - \beta)^{2k+1} v = 0$, multiplying by $g^{-(2k+2)}(g^2 - \beta)$ gives

$$g^{-(2k+2)}(g^2 - \beta)^{2k+2} v = (g^2 + \beta^2 g^{-2} - 2\beta)^{k+1} v = 0.$$

Since $(g^2 + \beta^2 g^{-2} - 2\beta)^{k+1} v$ is invariant if we exchange g^2 and $\beta^2 g^{-2}$, then powers of these will have the same coefficients in the expansion. So we have

$$(g^2 + \beta^2 g^{-2} - 2\beta)^{k+1} v = \sum_{i=0}^{k+1} d_i (g^{2i} + \beta^{2i} g^{-2i}) v = 0,$$

with $d_{k+1} = 1$, and so

$$(g^{s+1} + \beta^{s+1} g^{-(s+1)})v = -\sum_{i=0}^{k} d_i (g^{2i} + \beta^{2i} g^{-2i}) v \in P.$$

So now $(g + \beta g^{-1})P \subset P$.

As in the previous case, it follows that $(g - \beta g^{-1})Q \subset P$ and $(g + \beta g^{-1})Q \subset Q$. If we define, as before, t_1 to be the involution having $+1$ eigenspace P and -1 eigenspace Q, it follows from the earlier computation that $g = t_1 t_2$ with $t_1^2 = I$ and $t_2^2 = \beta I$.

We note that in both cases above, we may prove by induction, and making use of the fact that $(g - \beta g^{-1})P = Q$ and $(g + \beta g^{-1})P \subset P$, that we have for any i, $(g^i + \beta^i g^{-i})v \in P$ and $(g^i - \beta^i g^{-i})v \in Q$.

(ii): If g is β-symplectic, we must show that t_1 is skew-symplectic and t_2 satisfies $\mu(t_2) = -\beta$. As we have just shown, we may take t_1 to be an involution with $+1$ eigenspace P, spanned by vectors of the form $(g^i + \beta^i g^{-i})v$, and -1 eigenspace Q, spanned by vectors of the form $(g^i - \beta^i g^{-i})v$. We first show that P and Q are totally isotropic. For P, take $(g^i + \beta^i g^{-i})v$, $(g^j + \beta^j g^{-j})v \in$

P, and we have:

$$\begin{aligned}
&\langle (g^i + \beta^i g^{-i})v,\ (g^j + \beta^j g^{-j})v\rangle \\
&= \langle (g^i + \beta^i g^{-i})v,\ g^j v\rangle + \langle g^i v,\ \beta^j g^{-j} v\rangle + \langle \beta^i g^{-i} v,\ \beta^j g^{-j} v\rangle \\
&= \langle (g^i + \beta^i g^{-i})v,\ g^j v\rangle + \langle g^j v,\ \beta^i g^{-i} v\rangle + \langle g^j v,\ g^i v\rangle \\
&= \langle (g^i + \beta^i g^{-i})v,\ g^j v\rangle + \langle g^j v,\ (g^i + \beta^i g^{-i})v\rangle \\
&= 0.
\end{aligned}$$

And similarly for Q, we may show that for $(g^i - \beta^i g^{-i})v, (g^j - \beta^j g^{-j})v \in Q$, we have:

$$\langle (g^i - \beta^i g^{-i})v,\ (g^j - \beta^j g^{-j})v\rangle = 0.$$

So for every pair of vectors in a set that spans P or Q, the inner product is zero, and so $P \subset P^\perp$ and $Q \subset Q^\perp$. Now let u and u' be any two vectors in $V = P \oplus Q$. Write $u = w + y$, $u' = w' + y'$, where $w, w' \in P$ and $y, y' \in Q$. We compute $\langle t_1 u, t_1 u'\rangle$:

$$\begin{aligned}
\langle t_1 u, t_1 u'\rangle &= \langle t_1(w+y),\ t_1(w'+y')\rangle = \langle w - y,\ w' - y'\rangle \\
&= \langle w, w'\rangle + \langle y, y'\rangle - \langle y, w'\rangle - \langle w, y'\rangle \\
&= -\langle y, w'\rangle - \langle w, y'\rangle.
\end{aligned}$$

While computing $\langle u, u'\rangle$ gives us:

$$\begin{aligned}
\langle u, u'\rangle &= \langle w + y,\ w' + y'\rangle \\
&= \langle w, w'\rangle + \langle y, y'\rangle + \langle y, w'\rangle + \langle w, y'\rangle \\
&= \langle y, w'\rangle + \langle w, y'\rangle.
\end{aligned}$$

Therefore we have $\langle t_1 u,\ t_1 u'\rangle = -\langle u, u'\rangle$, and t_1 is skew-symplectic. Since g satisfies $\mu(g) = \beta$, then $t_2 = t_1 g$ satisfies $\mu(t_2) = -\beta$.

(iii): From Lemma 5, we may write V as an orthogonal sum of subspaces that are either cyclic, or the direct sum of two cyclic subspaces. The case for the nondegenerate cyclic pieces is covered in (i) and (ii), and so we may assume that V is the sum of two cyclic subspaces. We may further assume, from Lemmas 4 and 5, that for any $u_1 \in V$ of order $q(g)^s$, the cyclic space U_1 generated by u_1 is degenerate, and if u_2 is any vector satisfying $\langle q(g)^{s-1} u_1,\ u_2\rangle \neq 0$, then $V = U_1 \oplus U_2$, where U_2 is the cyclic space generated by u_2.

From part (i) above, we may write $U_1 = P_1 \oplus Q_1$, where P_1 is spanned by vectors of the form $(g^k + \beta^k g^{-k})u_1$, and Q_1 is spanned by vectors of the form $(g^k - \beta^k g^{-k})u_1$. If $q(x)$ is relatively prime to $x^2 - \beta$ and $\deg(q(x)) = 2m$, let $w = g^{-m(s-1)} q(g)^{s-1} u_1$. If $q(x) = x^2 - \beta$ and s is odd, let $w = g^{-(s-1)}(x^2 - \beta)^{s-1} u_1$. Then in either case we have $w \neq 0$ and $w \in P_1$, and in particular, $w \notin Q_1$. The case where $q(x) = x^2 - \beta$ and s is even is taken care of below. Since $(Q_1^\perp)^\perp = Q_1$, and $w \notin Q_1$, we can find a vector $u_2 \in Q_1^\perp$ such that $\langle w, u_2\rangle \neq 0$. Since U_1 is cyclically generated by u_1, and so also by $g^{-m(s-1)} u_1$, and U_1 is degenerate, we have by Lemma 4 that $u_2 \notin U_1$, and is of order $q(g)^s$. Then $V = U_1 \oplus U_2$, where U_2 is cyclically generated by $u_2 \in Q_1^\perp$.

Now let $U_2 = P_2 \oplus Q_2$, where P_2 is spanned by the vectors $(g^k + \beta^k g^{-k})u_2$, and Q_2 is spanned by the vectors $(g^k - \beta^k g^{-k})u_2$. Let $P = P_1 \oplus Q_2$ and $Q = P_2 \oplus Q_1$. From part (i), we know that for $i = 1, 2$, $(g - \beta g^{-1})P_i = Q_i$, $(g - \beta g^{-1})Q_i \subset P_i$, $(g + \beta g^{-1})P_i \subset P_i$, and $(g + \beta g^{-1})Q_i \subset Q_i$. These together give us $(g - \beta g^{-1})P \subset Q$, $(g - \beta g^{-1})Q \subset P$, $(g + \beta g^{-1})P \subset P$, and $(g + \beta g^{-1})Q \subset Q$. So if we define t_1 to be the involution on V with P as its $+1$ eigenspace and Q its -1 eigenspace, we duplicate the argument in (i) to obtain $(t_1 g)^2 = \beta I$, so letting $t_2 = t_1 g$ gives $t_2^2 = \beta I$.

We have shown in (ii) above that $P_i \subset P_i^{\perp}$ and $Q_i \subset Q_i^{\perp}$ for $i = 1, 2$. Now we show that $P_i \perp Q_j$ for $i \neq j$. We have

$$\begin{aligned}
&\langle (g^k \pm \beta^k g^{-k})u_1,\ (g^l \mp \beta^l g^{-l})u_2 \rangle \\
&= \langle (g^k \pm \beta^k g^{-k}u_1,\ g^l u_2 \rangle \mp \langle (g^k \pm \beta^k g^{-k})u_1,\ \beta^l g^{-l} u_2 \rangle \\
&= \langle \beta^l g^{-l}(g^k \pm \beta^k g^{-k})u_1,\ u_2 \rangle \mp \langle g^l (g^k \pm \beta^k g^{-k})u_1,\ u_2 \rangle \\
&= \mp \langle (g^l \mp \beta^l g^{-l})(g^k \pm \beta^k g^{-k})u_1,\ u_2 \rangle \\
&= 0,
\end{aligned}$$

since

$$\begin{aligned}
&(g^l \mp \beta^l g^{-l})(g^k \pm \beta^k g^{-k})u_1 \\
&= (g^{l+k} - \beta^{l+k} g^{-(l+k)})u_1 \pm \beta^k (g^{l-k} - \beta^{l-k} g^{-(l-k)})u_1 \in Q_1
\end{aligned}$$

and $u_2 \in Q_1^{\perp}$. So $P_i \perp Q_j$ for $i \neq j$. Now $P \subset P^{\perp}$ and $Q \subset Q^{\perp}$. From the argument in (ii), it follows that $\mu(t_1) = -1$ and $\mu(t_2) = -\beta$.

If $q(x) = x^2 - \beta$ and s is even, then we again let $w = g^{-(s-1)}(x^2 - \beta)^{s-1}u_1$, but now we have $w \neq 0$ and $w \in Q_1$, so $w \notin P_1$. Now we find a $u_2 \in P_1^{\perp}$ such that $\langle w, u_2 \rangle \neq 0$, and by Lemma 4, $u_2 \notin U_1$ and is of order $q(g)^s$. Now $V = U_1 \oplus U_2$, where U_2 is cyclically generated by $u_2 \in P_1^{\perp}$. We have $U_2 = P_2 \oplus Q_2$, where P_2 is spanned by the vectors $(g^k + \beta^k g^{-k})u_2$ and Q_2 is spanned by the vectors $(g^k - \beta^k g^{-k})u_2$.

In this case, we let $P = P_1 \oplus P_2$ and $Q = Q_1 \oplus Q_2$. Applying the method in part (i) to the P_i and Q_i, we obtain $(g - \beta g^{-1})P \subset Q$, $(g - \beta g^{-1})Q \subset P$, $(g + \beta g^{-1})P \subset P$, and $(g + \beta g^{-1})Q \subset Q$. Define t_1 to be the involution on $V = P \oplus Q$ to have $+1$ eigenspace P and -1 eigenspace Q, and by part (i), we have $(t_1 g)^2 = \beta I$, and $t_2 = t_1 g$ satisfies $t_2^2 = \beta I$.

As in the previous case, we may show that $P_i \perp P_j$ and $Q_i \perp Q_j$ for $i \neq j$. Since also $P_i \subset P_i^{\perp}$ and $Q_i \subset Q_i^{\perp}$ from part (ii), we finally have for this case that $P \subset P^{\perp}$ and $Q \subset Q^{\perp}$. From part (ii), it follows that $\mu(t_1) = -1$ and $\mu(t_2) = -\beta$. $\square$

Proof of Theorem 2. From Proposition 2, it is sufficient to prove the theorem for g whose minimal polynomial is a power of a polynomial of either the form $p(x)\hat{p}(x)$, where $p(x)$ is an irreducible polynomial which is not self-β-adjoint, or $q(x)$, an irreducible polynomial which is self-β-adjoint. For the first case, Lemma 3 reduces the task to proving the theorem for cyclic transformations, which is proven in Proposition 3 (i) and (ii). The second case is taken care of in Proposition 3 (iii). $\square$

5. A Type of Conjugacy Class in GSp(V)

In this section we prove a proposition about a specific type of conjugacy class of $GSp(V)$. The proof is adapted from unpublished notes of D. Bump and D. Ginzburg. The proposition is a generalization of a result of R. Gow [6, Lemma 1], and the proof of part (i) is essentially the same as Gow's proof.

Proposition 4. *Let V be an F-vector space such that $\mathrm{char}(F) \neq 2$, equipped with a nondegenerate skew-symmetric bilinear form $\langle \cdot, \cdot \rangle$.*

(i) If $-\beta \in F$ is a square in F, there exists a unique conjugacy class of $GSp(V)$ whose elements g satisfy $g^2 = -\beta I$, $\mu(g) = \beta$.

(ii) Suppose that $-\beta \in F$ is not a square in F, and let K be a quadratic extension of F containing the square roots of $-\beta$. Let $\varphi : \lambda \mapsto \overline{\lambda}$ be the nontrivial element of $Gal(K/F)$. If the norm map $N : K \to F$, $N(\lambda) = \lambda\overline{\lambda}$ is surjective, then there exists a unique conjugacy class of $GSp(V)$ whose elements g satisfy $g^2 = -\beta I$, $\mu(g) = \beta$.

Proof. Since $\mathrm{char}(F) \neq 2$, g has the two square roots of $-\mu(g) = -\beta$ as distinct eigenvalues, and so g is semisimple. Let γ and $-\gamma$ be the two square roots of $-\beta$, so if $\gamma \notin F$, $\overline{\gamma} = -\gamma$. Let $V(\gamma)$ and $V(-\gamma)$ be the eigenspaces for γ and $-\gamma$ respectively. In the case that $\gamma \notin F$, these eigenspaces are defined over K, and are subspaces of $V_K = V \otimes_F K$. We may extend the bilinear form $\langle \cdot, \cdot \rangle$ to $V_K \times V_K \to K$ by linearity. We define the action of $GSp(V)$ on V_K as $g(v \otimes \lambda) = gv \otimes \lambda$, where $g \in GSp(V)$ and $\lambda \in K$. Now note that $V(\gamma)$ and $V(-\gamma)$ are totally isotropic spaces, since for any $v, w \in V(\pm\gamma)$, we have, using $\mu(g) = \beta$,

$$\beta\langle v, w\rangle = \langle gv, gw\rangle = \langle \pm\gamma v, \pm\gamma w\rangle = (\pm\gamma)^2\langle v, w\rangle = -\beta\langle v, w\rangle.$$

So $\langle v, w\rangle = 0$. If the dimension of V is $2n$, then the maximum dimension of a totally isotropic subspace of V (or V_K) is n, and since $V(\gamma) \oplus V(-\gamma) = V$ (or V_K), then $V(\gamma)$ and $V(-\gamma)$ are each dimension n.

For part (i), we suppose that $-\beta$ is a square in F, so that $\gamma \in F$. Then $V(\gamma)$ and $V(-\gamma)$ are F-subspaces of V, each of dimension n. Take a basis $v_1, \ldots, v_n$ of $V(\gamma)$. Now, $V(-\gamma)$ is isomorphic to $V(\gamma)^*$, the dual of $V(\gamma)$, as it acts on $V(\gamma)$ through the inner product. So we may choose a basis $w_1, \ldots, w_n$ of $V(-\gamma)$ dual to $v_1, \ldots, v_n$, so that $\langle v_i, w_j\rangle = \delta_{ij}$. Now we have found a symplectic basis $v_1, \ldots, v_n, w_1, \ldots, w_n$, with respect to which the element g acts by the matrix

$$\begin{pmatrix} \gamma I_n & \\ & -\gamma I_n \end{pmatrix},$$

and so the conjugacy class of g is uniquely determined.

In (ii), if $-\beta$ is not a square in F, then $V(\gamma)$ and $V(-\gamma)$ are K-subspaces of V_K. We may extend the action of φ to V_K, by letting $\varphi(v \otimes \lambda) = v \otimes \overline{\lambda}$. Then for any $g \in GSp(V)$, the action of φ and g on V_K commute. It follows

we have the map $\varphi : V(\gamma) \to V(-\gamma)$. Now define $[\cdot,\cdot]$ on $V(\gamma) \times V(\gamma)$ by

$$[v,w] = -2\gamma\langle v, \overline{w}\rangle.$$

Then $\overline{[v,w]} = [w,v]$, that is $[\cdot,\cdot]$ is Hermitian, and is nondegenerate, since $\langle\cdot,\cdot\rangle$ is nondegenerate. It will follow from the surjectivity of the norm map that there is an orthonormal basis of $V(\gamma)$ with respect to $[\cdot,\cdot]$, call it $u_1, \ldots, u_n$. Then $\langle u_i, \overline{u_j}\rangle = -\delta_{ij}/2\gamma$. Now define

$$v_i = \gamma(u_i - \overline{u_i}), \quad w_i = u_i + \overline{u_i}.$$

Then $v_i, w_i \in V$, since they are invariant under φ, and also $\langle v_i, v_j\rangle = \langle w_i, w_j\rangle = 0$ and $\langle v_i, w_j\rangle = \delta_{ij}$. Now with respect to the symplectic basis $v_1, \ldots, v_n, w_1, \ldots, w_n$ of V, the element g acts by the matrix

$$\begin{pmatrix} & I_n \\ -\beta I_n & \end{pmatrix},$$

and so again the conjugacy class of g is uniquely determined.

We finish the proof of (ii) by showing that there is an orthonormal basis for $V(\gamma)$ with respect to the nondegenerate Hermitian form $[\cdot,\cdot]$. The proof is by induction on the dimension of $V(\gamma)$. For the one-dimensional case, we may first find a v such that $[v,v] \neq 0$ since the form is nondegenerate. Now for any $\lambda \in K$, we have $[\lambda v, \lambda v] = N(\lambda)[v,v]$, and we are assuming that the norm map is surjective. So we let λ be such that $N(\lambda) = [v,v]^{-1}$, and then $u_1 = \lambda v$ satisfies $[u_1, u_1] = 1$.

For the n-dimensional case, if we can find any vector such that $[v,v] \neq 0$, we may normalize as above, from the surjectivity of the norm map, to find a u_1 such that $[u_1, u_1] = 1$. The orthogonal complement of the span of u_1 is then of smaller dimension, and by an induction hypothesis we may find an orthonormal basis. If $[v,v] = 0$ for all v, then we would have $T([v,w]) = [v+w, v+w] - [v,v] - [w,w] = 0$ for all v, w, where $T : \lambda \mapsto \lambda + \overline{\lambda}$ is the trace map from K down to F. Now, a basis for K over F is $\{1, \gamma\}$, where γ is a square root of $-\beta$, and $\overline{\gamma} = -\gamma$. So an element of K with trace zero must be of the form $\lambda\gamma$, where $\lambda \in F$. Now $[v,w]$ must be of this form for all v, w, but then $[\gamma v, w] = \gamma[v,w]$ must also be of this form, which implies we must have $[v,w] = 0$. This contradicts the fact that $[\cdot,\cdot]$ is nondegenerate. We therefore have an orthonormal basis and the proof is complete. □

6. Applications

First, we have the following immediate consequence of Theorem 2.

Corollary 1. *Let $g \in GSp(V)$, where V is an F-vector space and* $\mathrm{char}(F) \neq 2$. *Then g is conjugate to $\mu(g)g^{-1}$ by a skew-symplectic involution.*

The statement that g is conjugate to $\mu(g)g^{-1}$ by some element of $GSp(V)$ is stated by Prasad [9, Proposition 1], but a very different proof is suggested there.

In [6, Theorem 1], Gow proves that if F is a field in which -1 is a square, then $Sp(2n, F)$ is the square of a conjugacy class, which is a special case of

Thompson's conjecture. Gow's proof uses Wonenburger's Theorem 1 and a version of Proposition 4 for the symplectic group. We obtain the following similar result for $GSp(V)$, which says that under certain conditions on the field, we may write $GSp(V)$ as the product of one conjugacy class with the union of conjugacy classes indexed by the multiplicative group of the field.

Corollary 2. *Let V be an F-vector space such that* $\mathrm{char}(F) \neq 2$*, and such that the norm map of any quadratic extension of F is surjective. Let C_β be the unique conjugacy class corresponding to $\beta \in F^\times$ as described in Proposition 4,*

$$C_\beta = \{g \in GSp(V) \mid g^2 = -\beta I, \mu(g) = \beta\},$$

and let C be the union of these conjugacy classes,

$$C = \bigcup_{\beta \in F^\times} C_\beta.$$

Then $GSp(V) = C_1 \cdot C$.

Proof. This follows directly from Theorem 2 and Proposition 4. □

One of the main results of Wonenburger [10, Theorem 1] is that an element of $GL(V)$ is conjugate to its inverse if and only if it is the product of two involutions. Wonenburger proves this under the assumption that V is an F-vector space such that $\mathrm{char}(F) \neq 2$, but Djoković [1] proved the result for general characteristic. We are able to prove the following generalization.

Theorem 3. *Let V be an F-vector space with* $\mathrm{char}(F) \neq 2$*, and $g \in GL(V)$. Then g is conjugate to βg^{-1} for some $\beta \in F^\times$ if and only if g has a factorization $g = t_1 t_2$ such that $t_1^2 = I$ and $t_2^2 = \beta I$.*

Proof. If g has such a factorization, then $\beta g^{-1} = t_2 t_1$, and so g is conjugate to βg^{-1} since $t_2^{-1} g t_2 = \beta g^{-1}$.

Now assume that g is conjugate to βg^{-1} for some $\beta \in F^\times$. Then g and βg^{-1} have the same invariant factors and minimal polynomial, and so from Proposition 1, the minimal polynomial of g is self-β-adjoint. In general, if g has invariant factors $\delta_1(x), \ldots, \delta_\nu(x)$, then βg^{-1} has invariant factors $\hat{\delta}_1(x), \ldots, \hat{\delta}_\nu(x)$, where $\hat{\delta}_i(x)$ is the β-adjoint of $\delta_i(x)$. Since g and βg^{-1} have the same invariant factors in this case, then each $\delta_i(x)$ is self-β-adjoint.

We may write V as the direct sum of subspaces, $V = \oplus_{i=1}^{\nu} V_i$, where g restricted to V_i, call it g_i, has minimal polynomial $\delta_i(x)$, and V_i is a cyclic space with respect to g_i. Now g_i is a cyclic transformation which is self-β-adjoint, and so it follows from Proposition 3(i) that g_i has the desired factorization. Since $g = \oplus_{i=1}^{\nu} g_i$, the theorem follows. □

We now consider the case when V is an $\mathbb{F}_q$-vector space. If $q \equiv 1 (\mathrm{mod}\ 4)$, then -1 is a square, and it follows immediately from Wonenburger's Theorem 1 that every element of $Sp(2n, \mathbb{F}_q)$ is conjugate to its inverse. This is not true, however, if $q \equiv 3 (\mathrm{mod}\ 4)$, as there are unipotent elements which are

not conjugate to their inverses in $Sp(2n,\mathbb{F}_q)$, as given by Feit and Zuckerman [2, Lemma 5.3]. It follows from the results of Wonenburger and of Feit and Zuckerman that if we throw a skew-symplectic involution into $Sp(2n,\mathbb{F}_q)$ for $q \equiv 3 (\text{mod } 4)$, then every element of $Sp(2n,\mathbb{F}_q)$ is conjugate to its inverse by an element of this larger group. However, there are elements in the other coset which are not conjugate to their inverses by elements of this group. The following result gives a group in which $Sp(2n,\mathbb{F}_q)$, $q \equiv 3 (\text{mod } 4)$, is an index 2 subgroup, such that every element is conjugate to its inverse.

Theorem 4. *Let $G = Sp(2n,\mathbb{F}_q)$, where $q \equiv 3 (\text{mod } 4)$. Let ι be the order 2 automorphism of G defined by the following conjugation by a skew-symplectic element:*

$$^{\iota}g = \begin{pmatrix} -I_n & \\ & I_n \end{pmatrix} g \begin{pmatrix} -I_n & \\ & I_n \end{pmatrix}.$$

Now define $G^{\iota,-I}$ as the following group containing G as an index 2 subgroup:

$$G^{\iota,-I} = \langle G, \tau \mid \tau^2 = -I, \tau^{-1} g \tau = {}^{\iota}g \text{ for every } g \in G \rangle.$$

Then every element of $G^{\iota,-I}$ is conjugate to its inverse, and so every complex character of $G^{\iota,-I}$ is real-valued.

Proof. Let $\alpha \in \mathbb{F}_{q^2}$ be a square root of -1. Note that conjugation by the element

$$\tau = \begin{pmatrix} \alpha I_n & \\ & -\alpha I_n \end{pmatrix}$$

also gives the order 2 automorphism ι of G, and $\tau^2 = -I$, so the group $G^{\iota,-I}$ is isomorphic to $\langle G, \tau \rangle$, which is a subgroup of $Sp(2n,\mathbb{F}_{q^2})$.

Let $g \in G$. From Wonenburger's Theorem 1, there is a skew-symplectic involution h such that $h^{-1}gh = g^{-1}$. The element

$$\kappa = \begin{pmatrix} -I_n & \\ & I_n \end{pmatrix}$$

is also a skew-symplectic involution, and so κh is symplectic, and then $(\kappa h)^{-1} g (\kappa h) = {}^{\iota}g^{-1}$. Conjugating both sides by τ, we have $(\kappa h \tau)^{-1} g (\kappa h \tau) = g^{-1}$, and $\kappa h \tau \in G^{\iota,-I}$.

Now let $k \in \tau G = G^{\iota,-I} \setminus G$. Then $\alpha k = \alpha \tau g$ for some $g \in G$. We have $\alpha \tau = \kappa$, and so $\alpha k = \kappa g$. So now $\alpha k \in GSp(2n,\mathbb{F}_q)$ with $\mu(\alpha k) = -1$. We now apply Corollary 1 to say that there exists a skew-symplectic involution $t \in GSp(2n,\mathbb{F}_q)$ such that $t^{-1}(\alpha k) t = -(\alpha k)^{-1}$. Multiplying both sides of this by $-\alpha$, we obtain $k^{-1} = (\alpha t)^{-1} k (\alpha t)$.

Now $\alpha t = -\tau \kappa t$. Since $-\kappa$ and t are skew-symplectic, their product is in G, so we have $\alpha t \in \tau G$. So k is conjugate to its inverse in $G^{\iota,-I}$. $\square$

Remarks. Gow proved [4] that if V is an F-vector space and $\text{char}(F) = 2$, then every element of $Sp(V)$ is the product of two symplectic involutions. If F is any algebraic extension of $\mathbb{F}_2$, then every element of F is a square, and so a result for $GSp(V)$ is immediate. We do not have results for $GSp(V)$,

however, when F is a characteristic 2 field which is not an algebraic extension of $\mathbb{F}_2$.

References

1. D.Ž. Djoković, Product of two involutions, *Arch. Math. (Basel)*, **18** (1967), 582–584.
2. W. Feit and G.J. Zuckerman, Reality properties of conjugacy classes in spin groups and symplectic groups, In *Algebraists' homage: papers in ring theory and related topics (New Haven, Conn., 1981)*, volume 13 of *Contemp. Math.*, pages 239–253, Amer. Math. Soc., Providence, R.I., 1982.
3. R. Gow, The equivalence of an invertible matrix to its transpose, *Linear and Multilinear Algebra*, **8** (1980), no. 4, 329–336.
4. R. Gow, Products of two involutions in classical groups of characteristic 2, *J. Algebra*, **71** (1981), no. 2, 583–591.
5. R. Gow, Properties of the characters of the finite general linear group related to the transpose-inverse involution, *Proc. London Math. Soc. (3)*, **47** (1983), no. 3, 493–506.
6. R. Gow, Commutators in the symplectic group, *Arch. Math. (Basel)*, **50** (1988), no. 3, 204–209.
7. W.H. Gustafson, P.R. Halmos, and H. Radjavi, Products of involutions, Collection of articles dedicated to Olga Taussky Todd, *Linear Algebra and Appl.*, **13** (1976), no. 1/2, 157–162.
8. B. Huppert, Isometrien von Vektorräumen I, *Arch. Math. (Basel)*, **35** (1980), no. 1-2, 424–431.
9. D. Prasad, On the self-dual representations of finite groups of Lie type, *J. Algebra* **210** (1998), no. 1, 298–310.
10. M. Wonenburger, Transformations which are products of two involutions, *J. Math. Mech.*, **16** (1966), 327–338.

Department of Mathematics, Stanford University, Stanford, CA 94305-2125 USA

E-mail address: vinroot@math.stanford.edu