Finite Matrix Groups

Mendel Keller

June 7, 2019

Abstract

We look at finite linear groups, in particular projective special linear and symplectic groups, as these provide a rich set of examples of finite simple groups. Using linear algebra, group theory and some simple combinatorics we can list these groups and determine their basic properties like size.

1 Introduction

In this paper wee look at projective special linear groups and projective symplectic groups over finite fields as providing a collection of finite simple groups. This serves then as an important introduction to finite group theory, as the methods and reasoning developed here is of the sort used whenever working with finite groups of lie type which form a large an important section of finite groups. For example the theory for orthogonal groups is fairly similar to although a bit more complicated than the treatment here for symplectic groups.

When working with these groups, we can work from a few different perspectives. On the one hand, the objects of study are groups, so we can rely on group theory and formulate results in that language. The group theory angle helps when counting things using orbit-stabilizer theorems, as well as when looking at conjugation and generators. We can also write down matrices explicitly and use some basic linear algebra to prove properties by sheer force, this allows us to work with underlying vector spaces themselves, which allows us to exploit linearity and dimension properties. The vector spaces will then also admit bilinear forms, which are a particularly neat tool for computation. We can often work explicitly with the form and with some basic algebra have it take on some desired value, or count the number of vectors which yield such a value. Having these various angles to work with simultaneously leads to a vibrant theory where many results can be obtained as well as understood and even visualized.

Our treatment here follow [1] chapters 0-3 pretty closely.

Section 2 will introduce the most basic matrix groups over finite fields. In section 3 we explore the projective special linear group and in particular show that it is in general simple. In section 4 we introduce bilinear forms as a computational tool for linear algebra. In section 5 we take a close look at symplectic groups and show that the projective symplectic groups are in general simple. In section 6 we address the exceptions - the groups of the right type that still fail to be simple.

2 The Main Players

The first group of matrices to consider is the *general linear group* $GL(n,q)$, the $n \times n$ invertible matrices over a finite field $\mathbb{F}_q$. All the other groups we consider are obtained from these groups.

The next thing to consider is the *special linear group* $SL(n,q)$. This consists of $n \times n$ matrices over $\mathbb{F}_q$ of determinant 1. We have that $SL(n,q) \triangleleft GL(n,q)$ is a normal subgroup, because it is the kernel of the determinant homomorphism $\det : GL(n,q) \to \mathbb{F}_q$.

We also consider the *projective general linear group* $PGL(n,q)$. This is what we obtain when we mod out $GL(n,q)$ by its center.

Theorem 1. *The center of* $GL(n,q)$ *is* $Z(GL(n,q)) = \{\lambda I : \lambda \in \mathbb{F}_q^\times\}$.

Proof. Clearly for any $M \in GL(n,q)$ we have that $(\lambda I)M = \lambda M = M(\lambda I)$. We then only need check that $AM = MA$ for all $M \in GL(n,q)$ implies that $A = \lambda I$ for some $\lambda \in \mathbb{F}_q^\times$.

In order to prove this fact, we consider the set of matrices M_{ij} with ones down the main diagonal and a single off-diagonal entry of 1 in the ij-th place. It is clear that these are in $GL(n.q)$. If the set of all of these M_{ij} commutes with a matrix A then we'll have $(M_{ij} - I)A = A(M_{ij} - I)$, since we already know that $-I$ commutes with everything and matrix multiplication is distributive. But $M_{ij} - I$ is just a matrix E_{ij} with a single entry in ij-th place, so we consider the set of these products instead.

To see that A must be diagonal, consider the matrices E_{ii}. Then we have that $E_{ii}A$ is the matrix with only the i-th row of A and AE_{ii} is the matrix having only the i-th column of A. So that $E_{ii}A = AE_{ii}$ will force that the only nonzero entry of the i-th row of A is in the i-th spot i.e. A must be diagonal.

Once we know that A is diagonal, we show that it is in fact scalar. Multiplying $E_{ij}A$ where A is diagonal just yields a_{jj} of A as the ij-th entry and zero everywhere else. On the other hand, AE_{ij} is a matrix with just the ii-th entry of A in its ij-th spot, and zeros otherwise. These two matrices are equal iff $a_{ii} = a_{jj}$, so that in fact all nonzero entries of A are equal and lie on the diagonal, and the desired result is proved. ■

Since $PGL(n, q)$ is equal to $GL(n, q)/\lambda I$, what we get is the group of $GL(n, q)$ actions on the lines in $\mathbb{F}_q^n$. This is because we have that $M \sim \lambda^{-1}M$ so that the $PGL(n, q)$ action doesn't distinguish between v and λv. This is where the group gets its name, as we call a space of with lines as points a projective space.

We similarly also have the *projective special linear group* $PSL(n, q) \triangleleft PGL(n, q)$, which can be given as $\pi(SL(n, q))$ where $\pi : GL(n, q) \to PGL(n, q)$ is the canonical projection map. The projective special linear group is then given by $SL(n, q)/Z(SL(n, q))$ where $Z(SL(n, q)) = (Z(GL(n, q)) \cap SL(n, q)) = \{\lambda I | \lambda^n = 1\}$ since $|\lambda I| = \lambda^n$. The fact that there isn't anything new in the center of $SL(n, q)$ can be seen by examining our proof of theorem 1 and noting that this proof does not need to utilize matrices with a determinant other than 1.

That $PSL(n, q)$ is a normal subgroup can be seen from the fact that it is the kernel of a map $PGL(n, q) \to \mathbb{F}_q^\times/(\mathbb{F}_q^\times)^n$ which we explain further in section 3.1.

2.1 Sizes

Having introduced the relevant groups, we begin to explore facts about them. One basic fact to start with is their sizes. We start with $GL(n, q)$.

Theorem 2. *The size of $GL(n, q)$ is given by*

$$|GL(n, q)| = \prod_{i=0}^{n-1} (q^n - q^i)$$

.

Proof. Once a basis is chosen for $\mathbb{F}_q^n$, an element of $GL(n, q)$ is given by a ordered choice of n linearly independent elements of $\mathbb{F}_q^n$ to send this basis to. For the first basis vector, any nonzero vector will do, leaving $q^n - 1$ choices. Having chosen i vectors, there is a subspace $\mathbb{F}_q^i \subset \mathbb{F}_q^n$ which they span, any vector outside of this subspace will be linearly independent of our set thus far, leaving $|\mathbb{F}_q^n \setminus \mathbb{F}_q^i| = q^n - q^i$ elements to choose from. This justifies our above equality. ■

We now consider the size of $SL(n, q)$. We have that $\mathbb{F}_q^\times \cong GL(n, q)/SL(n, q)$ because det is a surjective homomorphism $GL(n, q) \to \mathbb{F}_q^\times$ with kernel $SL(n, q)$. This then tells us that $q - 1 = |\mathbb{F}_q^\times| = |GL(n, q)|/|SL(n, q)|$. By cross-multiplying, we get that

$$|SL(n, q)| = |GL(n, q)|/(q - 1) = \frac{\prod_{i=0}^{n-1}(q^n - q^i)}{q - 1}$$

.

We similarly have that

$$|PGL(n, q)| = \frac{\prod_{i=0}^{n-1}(q^n - q^i)}{q - 1}$$

. Here the computation is more direct even, as $PGL(n, q)$ is obtained as the quotient of $GL(n, q)$ by the subgroup $\mathbb{F}_q^\times I$ which has size $q - 1$.

A bit more complicated is the size of $PSL(n,q)$ which is given by

$$\frac{\prod_{i=0}^{n-1}(q^n - q^i)}{\gcd(n, q-1)(q-1)}$$

. To show this we need to show that the number of $\lambda \in \mathbb{F}_q^\times$ such that $\lambda^n = 1$ is equal to $\gcd(n, q-1)$.

Theorem 3. *The subgroup* $\{\lambda \in \mathbb{F}_q^\times : \lambda^n = 1\} \subset \mathbb{F}_q^\times$ *has order* $\gcd(n, q-1)$.

Proof. We have that $\lambda^{q-1} = 1$ for all of $\mathbb{F}_q^\times$ so that $\lambda^n = 1 \Rightarrow \lambda^{\gcd(n,q-1)} = 1$. This tells us that at most $\gcd(n, q-1)$ many λ solve $\lambda^n = 1$, since this means they also solve $x^{\gcd(n,q-1)} = 1$ which has degree $\gcd(n, q-1)$. On the other hand, raising a generator to the $\frac{q-1}{\gcd(n,q-1)}$ power will provide an element of order $\gcd(n, q-1)$ which gives at least this many as well. ■

We now consider whether these groups are isomorphic when they are the same size.

2.2 Equivalences

Since $|SL(n,q)| = |PGL(n,q)|$, it is natural to ask whether these are isomorphic. It turns out that this depends on whether n and $q-1$ are coprime or not.

Theorem 4. *The groups* $PGL(n,q)$ *and* $SL(n,q)$ *are isomorphic iff* $\gcd(n, q-1) = 1$.

Proof. If $\gcd(n, q-1) = 1$ then $|PSL(n,q)| = |SL(n,q)|$ so that $PSL(n,q) \cong SL(n,q)$. But we also have then than $|PSL(n,q)| = |PGL(n,q)|$ so that this subgroup is the whole group and $SL(n,q) \cong PSL(n,q) \cong PGL(n,q)$. On the other hand, if $\gcd(n, q-1) \neq 1$ then $|Z(SL(n,q))| = \gcd(n, q-1) \neq 1$. But meanwhile we have that $|Z(PGL(n,q))| = 1$ has trivial center. This is true because if some action ϕ in $PGL(n,q)$ is nontrivial, then is sends a line $[v]$ to a line $[w]$, this then fails to commute with the action θ that fixes $[v]$ and sends $[w]$ to $[u]$, because $\phi\theta[v] = [w]$ while $\theta\phi[v] = [u]$. Thus the two cannot be isomorphic, as their centers have different degrees. ■

3 More About Projective Linear Groups

3.1 Cokernel of special projective

We have a natural map $SL(n,q) \to PGL(n,q)$ which is simply the restriction of the canonical projection $\pi : GL(n,q) \to PGL(n,q)$ to the subgroup $SL(n,q) \subset GL(n,q)$. We can then consider the cokernel of this map, which is simply $PGL(n,q)/PSL(n,q)$.

In order to do this, it would serve us to define the equivalent of a determinant for $PGL(n,q)$, since $SL(n,q)$ is defined in a way given by the determinant map. The issue however is that the determinant of an element of $PGL(n,q)$ can only be defined up to an element $\lambda^n \in \mathbb{F}_q^*$, since elements of $PGL(n,q)$ are only defined up to λI, which has determinant λ^n. What we want then is illustrated in the commutative diagram:

$$\begin{array}{ccc} GL(n,q) & \xrightarrow{\det} & \mathbb{F}_q^* \\ \downarrow{\scriptstyle \pi} & & \downarrow \\ PGL(n,q) & \xrightarrow{\widetilde{\det}} & \mathbb{F}_q^*/(\mathbb{F}_q^*)^n \end{array}$$

In order then to define a determinant analogue $\widetilde{\det}$ for $PGL(n,q)$ we must mod out $\mathbb{F}_q^*$ by $(\mathbb{F}_q^*)^n$. The idea here is that $PGL(n,q) \cong GL(n,q)/\mathbb{F}_q^* I$ and the determinant is a map $\det : GL(n,q) \to \mathbb{F}^*$. So in order to have something that works with $PGL(n,q)$ we need only mod out $\mathbb{F}_q^*$ by $\mathrm{Im}(\mathbb{F}_q^* I)$ since $\mathbb{F}_q^* I$ is exactly what is modded out by to get to $PGL(n,q)$. What we get then is $\mathbb{F}_q^*/(\mathbb{F}_q^*)^n$ which holds because $\det(\lambda I) = \lambda^n$. This will of course be a cyclic group, since $\mathbb{F}_q^*$ is cyclic. It will also have order $\frac{q-1}{\gcd(q-1,n)}$

We then of course will get that $PGL(n,q)/PSL(n,q)$ is exactly the group $\mathbb{F}_q^*/(\mathbb{F}_q^*)^n$, because $PSL(n,q) \triangleleft PGL(n,q)$ is exactly the kernel of $\widetilde{\det} : PGL(n,q) \to \mathbb{F}_q^*/(\mathbb{F}_q^*)^n$.

We now have a pretty clear picture of what the projective special linear group looks like inside of the projective general linear group. We now move to show that the projective special linear group is in general simple.

3.2 The simplicity of PSL

3.2.1 Simplicity in general

Before we can show that PSL is in general simple, we will need a more general result which will allow us to determine when a group is in fact simple. But before we get to the theorem, we need to define the notion of *primitive* group actions.

Definition 5. For a group G and a set S the action $G \to \mathrm{Aut}(S)$ is primitive if (i) it is transitive (ii) for every subset $B \subsetneq S$ with $|B| \geq 2$ there is $x \in G$ and $a, b \in B$ such that $xa \in B, xb \notin B$.

We are now ready to introduce an important condition for simplicity. The theorem is a bit of a mouthful, so it helps to read through it carefully a few times to get the picture.

Theorem 6 (Iwasawa)**.** *Suppose G is a group with a faithful, primitive action on a set S, where G is equal to its derived group. Fixing $s \in S$, and setting $H = Stab(s)$. Then if there is a solvable subgroup $K \triangleleft H$ such that $G = \langle \cup\{xKx^{-1} : x \in G\}\rangle$, then G is simple.*

Proof. Suppose there were a normal subgroup N of G containing a nonidentity element x. Since the action of G is faithful, we have that x takes some $a \in S$ to some b in S. Then $B = Na$ is a subset of S with $|B| \geq 2$. If $y \in G$ then $N \triangleleft G$ gives $yN = Ny$ so that $yB = N(ya)$. Since $Na, N(ya)$ are orbits of elements in S under the action of N, either $Na = N(ya)$ or $Na \cap N(ya) = \emptyset$. But G being primitive must mean that either $B = S$, or for some y we have $\alpha, \beta \in B$ with $y\alpha \in B, y\beta \notin B$. But since $yB = N(ya)$ is either B or disjoint from it, we cannot have the second option, so we must have $B = S$, so that in particular we have the action of N transitive.

Given that N is transitive on S, we find in fact that $G = NH$. This is because for our $s \in S$ that we chose in order to give us $H = \mathrm{Stab}(s)$, we have that for any $x \in G$ it will be true that $xs = ns$ for some $n \in N$ by N's transitivity. That tells us that $(n^{-1}x)s = s$ so that $n^{-1}x \in H$ so that $x \in nH \subset NH$. Thus any $x \in G$ can be written as $nh : n \in N, h \in H$, so that in fact we have the desired $G = NH$.

But since $K \triangleleft H$ we have $KN \triangleleft HN = G$. We then have that for any $x \in G$ that $xKx^{-1} \subset xKNx^{-1} = KN$ so that $\cup\{xKx^{-1}\} \subset KN$ so that $G \subset KN \Rightarrow G = KN$.

But we have that K is solvable, so that the series of taking its commutator subgroups terminates in the trivial group, let's say at the m-th step. Denote taking the m-th commutator subgroup as $K^{(m)}$. Then we have that

$$G = G^{(m)} = (KN)^{(m)} \subset K^{(m)}N = N$$

So that in particular $N = G$, thus any nontrivial normal subgroup is the whole group, so that G is simple. ∎

3.2.2 Transvections

Thus, in order to demonstrate that projective special linear groups are simple, we need only show that they satisfy the assumptions of theorem 6. Before we get to that however, it will do us well to introduce transvections, and show that they can generate the special linear group.

Definition 7. A linear operator $1 \neq \tau \in GL(V)$ is a *transvection* if it fixes a subspace $W \subset V$ of codimension 1, and $\tau v - v \in W$ for all $v \in V$.

Note that the transvections are all contained in the special linear group. This can be seen by inspection from the matrix representation obtained when choosing a basis with $n - 1$ elements in W. Note also that the inverse of a transvection is a transvection as well because $\tau^{-1}x - x = -\tau^{-1}(\tau x - x) \in \tau^{-1}W = W$.

Lemma 8. *The set of transvections generates $SL(V)$.*

Proof. We prove this by induction on dimension. Having been given an arbitrary $\rho \in SL(V)$, we choose a vector v such that $v, \rho v$ are linearly independent, and set $v_1 = \rho v$. If ρ gives rise to no such vector we then choose any vector v and then choose a transvection τ_0 such that $v, \tau v$ are linearly independent and set $v_1 = \tau_0 v$.

We then choose a codimension 1 subspace $W \subset V$ such that $v_1 - v \in W$ while $v_1, v \notin W$. We can then define a transvection τ by having it act as the identity on W, with $\tau_1 v_1 = v$. The defined τ is a transvection because we can write any $x \in V$ as $x = av_1 + w$ where $a \in \mathbb{F}_q$ and $w \in W$ so that $\tau(x) - x = \tau(av_1 + w) - (av_1 + w) = av + w - av_1 - w = a(v_1 - v) \in W$. Now define $\tau_1 = \tau\tau_0$ if there was no v with ρv linearly independent, and otherwise simply take $\tau_1 = \tau$. So we now have that $\tau_1 \rho v = v$. We then set $W_1 = \tau_1 \rho W$ and we wish to multiply W_1 by another transvection to get back W. If indeed it is already true that $W_1 = W$, then we just leave things as is.

Note that $W + W_1 = V$ since these are distinct subspaces of codimension 1. This then tells us, using dimension formula, that $W \cap W_1$ is of dimension $2(n-1) - n = n - 2$. We then have that $W_2 = W \cap W_1 + \langle v \rangle$ is of codimension 1, since $v \notin W$. Now, we can write $v = w - w_1$ with $w \in W$ and $w_1 \in W_1$. But since $v \notin W \cup W_1$ we must have $w \in W \setminus W_1$ and $w_1 \in W_1 \setminus W$, so that $V = W \cap W_1 + \langle w \rangle + \langle w_1 \rangle$. It then follows that $w, w_1 \notin W_2$ since $w - w_1 \in W_2$ and $W_2 \subsetneq V$. We can then define a transvection τ_2 as acting as the identity on W_2 and $\tau_2 w_1 = w$, this is clearly a transvection because $w - w_1 \in W_2$. Now we get that $v = \tau_2 v$ wince $v \in W_2$. We also have that $\tau_2 W_1 = \tau_2(W \cap W_1 + \langle w_1 \rangle) = W \cap W_1 + \langle w \rangle = W$. So then we get that $\tau_2 \tau_1 \rho$ fixes v and permutes W. So we find that ρ is the product of some transvections and an element of $SL(W)$, thus by induction down to $\dim V = 2$, in which case we have $\tau_2 \tau_1 \rho$ fixing a line and sending the other to itself but $\det \tau_2 \tau_1 \rho = 1$ so that $\tau_2 \tau_1 \rho$ is the identity and we are done. ■

Lemma 9. *If* $\dim V > 2$ *all transvections are conjugate in* $SL(V)$.

Proof. Choose any two transvections τ_1, τ_2 with respective fixed planes W_1, W_2. Choose points $x_1 \notin W_1$ and $x_2 \notin W_2$, and define $W_1 \ni w_1 = \tau_1 x_1 - x_1$ and similarly for w_2. Now we choose bases $\{w_1, v_3, \ldots, v_n\}$ of W_1 and $\{w_2, u_3, \ldots, u_n\}$ of W_2. Define σ so that $\sigma x_1 = x_2 . \sigma w_1 = w_2, \sigma v_i = u_i$. Now $\sigma\tau_1\sigma^{-1}x_2 = \sigma\tau_1 x_1 = \sigma(w_1 + x_1) = w_2 + x_2 = \tau_2 x_2$, similarly $\sigma\tau_1\sigma^{-1}w_2 = \sigma\tau_1 w_1 = \sigma w_1 = w_2 = \tau_2 w_2$ and similarly for the u_i since τ_1 fixes the v_i and τ_2 fixes the u_i. So we see by explicit computation that $\sigma\tau_1\sigma^{-1} = \tau_2$. All that's left is to make σ and element of $SL(V)$. This is achieved by sending v_i to $\det(\sigma)^{-1}u_i$. ■

Dimension 2 For the 2 dimensional case things are a bit more tricky, as $SL(V)$ isn't quite big enough to conjugate all transvections. Instead we have that all transvections are conjugate in $SL(V)$ to *some* transvection of a specified form.

Lemma 10. *If* $\dim V = 2$ *and* $\{v_1, v_2\}$ *is a basis for* V *then every transvection is* $SL(V)$*-conjugate to a transvection whose* $\{v_1, v_2\}$ *representation is given by*

$$\begin{bmatrix} 1 & 0 \\ b & 1 \end{bmatrix}$$

for some $b \in \mathbb{F}_q^*$.

Proof. For τ a transvection, and some v outside the line fixed by τ set $w = \tau v - v$ so that the matrix of τ relative to $\{v, w\}$ is

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

. This mean that the matrix M of τ relative to $\{v_1, v_2\}$ is $M = BAB^{-1}$ for some matrix $B \in GL(V)$. For $b^{-1} = \det(B)$ we can take

$$B' = B \begin{bmatrix} b & 0 \\ 0 & 1 \end{bmatrix}$$

so that $B' \in SL(V)$ and we then get

$$B'^{-1} M B = \begin{bmatrix} 1 & 0 \\ b & 1 \end{bmatrix}$$

as desired. ■

3.2.3 For PSL

Theorem 11. *If $n \geq 2$ then $PSL(n,q)$ is simple, except for when $PSL(2,2)$ and $PSL(2,3)$.*

We wish to use theorem 6, so we are looking for: a set S for PSL to act on faithfully and primitively, a natural such set is simply the lines in V. Once we found this we want to find a normal subgroup of some element's stabilizer, for which we do the work in SL. That this normal subgroup's conjugates generate PSL is something we will be able to derive from our transvections work in the previous subsection. Lastly, the reason PSL isn't simple when $n = 2$ and $q = 2, 3$ will be found to be in their failure to be equal to their commutator subgroups.

Proof. Take the stabilizer of some $v \neq 0 \in \mathbb{F}_q^n$, and then choose a codimension 1 subspace $W \not\ni v$. We can then take the W component of σw for any $\sigma \in \mathrm{Stab}_{SL(n,q)}(v), w \in W$. In this way we get a homomorphism $\phi : \mathrm{Stab}_{SL(n,q)}(v) \to GL(W)$, the kernel of which $\ker(\phi)$ is a normal subgroup of $\mathrm{Stab}_{SL(n,q)}(v)$.

We can define elements $\tau_b \in \ker(\phi)$ for $b \in \mathbb{F}_q^*$ so that τ_b is a transvection by choosing a basis $\{w_1, \ldots, w_{n-1}\}$ for W and setting $\tau_b w_1 = w_1 + bv$ and $\tau_b w_i = w_i$ for $i \geq 2$, and lastly $\tau_b v = v$. This then fixes the subspace spanned by $\{w_2, \ldots, v\}$ and we have $\tau_b w_1 - w_1 = w_1 + bv - w_1 = bv$ so that τ is indeed a transvection, and when restricting the action to the W component τ_b acts as the identity and so is indeed in $\ker(\phi)$. For $n = 2$ note that this includes all matrices of the form

$$\begin{bmatrix} 1 & 0 \\ b & 1 \end{bmatrix}$$

relative to the basis $\{w, v\}$, so we can apply theorem 10. We find then indeed using theorem 9 in general and 10 for dimension 2 that indeed the conjugates of $\ker(\phi)$ generate all of $SL(V)$ by theorem 8.

Note also that $\ker(\phi)$ is abelian, since everything in $\ker(\phi)$ is the identity when looking only at W, so that in block matrix form the elements of $\ker(\phi)$ look like:

$$\begin{bmatrix} I & 0 \\ u & 1 \end{bmatrix}$$

for some $1 \times (n-1)$ matrix u. So that multiplying two of these will give

$$\begin{bmatrix} I & 0 \\ u & 1 \end{bmatrix} \begin{bmatrix} I & 0 \\ u' & 1 \end{bmatrix} = \begin{bmatrix} I & 0 \\ u+u' & 1 \end{bmatrix} = \begin{bmatrix} I & 0 \\ u' & 1 \end{bmatrix} \begin{bmatrix} I & 0 \\ u & 1 \end{bmatrix}$$

. In particular we now have that $\ker(\phi)$ is solvable.

We can pass all this information down to PSL when we quotient by the center $Z(SL)$. Clearly the quotient $\ker(\phi)/(Z(SL) \cap \ker(\phi))$ remains abelian. Similarly, the generators of a group will generate the quotient group, and conjugates remain conjugate in the quotient as well. The only difference is that $\mathrm{Stab}_{SL(V)}(v)$ will become $\mathrm{Stab}_{PSL(V)}([v])$, which comes out to the same subgroup modded out by its center. Thus we need only that the derived group PSL' is equal to the whole group PSL, this too we compute in SL.

For $n > 2$ we need only demonstrate that SL' contains a transvection. Then we will have by theorem 9 and the fact that SL' is normal that it contains all transvections, and then by theorem 8 SL' must be all of SL. We can explicitly construct a commutator that is a transvection however by choosing a basis $\{v_i\}$ and defining σ_1, σ_2 as $\sigma_1 v_1 = v_1 - v_2$ and $\sigma_2 v_2 = v_2 - v_3$ and $\sigma_i v_j = v_j$ for $i \neq j$. We then set $\tau = \sigma_1^{-1}\sigma_2^{-1}\sigma_1\sigma_2$. Clearly $\tau v_i = v_i$ for $i \geq 2$, so we need only check for $i = 1, 2$. We then get

$$\tau v_1 = (\sigma_1^{-1}\sigma_2^{-1}\sigma_1\sigma_2)v_1 = (\sigma_1^{-1}\sigma_2^{-1}\sigma_1)v_1 = (\sigma_1^{-1}\sigma_2^{-1})(v_1 - v_2) = (\sigma_1^{-1})(v_1 - v_2 - v_3) = v_1 - v_3$$
$$\tau v_2 = (\sigma_1^{-1}\sigma_2^{-1}\sigma_1\sigma_2)v_2 = (\sigma_1^{-1}\sigma_2^{-1}\sigma_1)(v_2 - v_3) = (\sigma_1^{-1}\sigma_2^{-1})(v_2 - v_3) = \sigma_1^{-1}v_2 = v_2$$

This means that τ stabilizes $W = \{v_2, \ldots, v_n\}$ and also that $\tau v_1 - v_1 = (v_1 - v_3) - v_1 = -v_3 \in W$ so that τ is indeed a transvection and so $SL'(V) = SL(V)$ for $\dim V > 2$.

Now for $n = 2$ we need to have all matrices of the form

$$\begin{bmatrix} 1 & 0 \\ b & 1 \end{bmatrix}$$

for $b \in \mathbb{F}_q^*$ to be inside of $SL'(n,q)$. We can show that we have this by choosing a basis $\{v,w\}$ and some $a \neq \pm 1$ in $\mathbb{F}_q^*$ (which is possible whenever $q > 3$). We then multiply for each $b \in \mathbb{F}_q^*$

$$\left(\begin{bmatrix} a^{-1} & 0 \\ 0 & a \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -b & 1 \end{bmatrix}\right)\left(\begin{bmatrix} a & 0 \\ 0 & a^{-1} \end{bmatrix}\begin{bmatrix} 1 & 0 \\ b & 1 \end{bmatrix}\right) = \begin{bmatrix} a^{-1} & 0 \\ -ab & a \end{bmatrix}\begin{bmatrix} a & 0 \\ a^{-1}b & a^{-1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ b(1-a^2) & 1 \end{bmatrix}$$

since $a^2 \neq 1$ we have that $1 - a^2$ is just some element of $\mathbb{F}_q^*$ so that by varying b we get every value of $\mathbb{F}_q^*$ in the lower left corner, which is what we need for $SL(V)$ to be equal to its derived group, so that in particular $PSL'(V) = PSL(V)$. We can now apply theorem 6 to show that $PSL(V)$ is simple, completing the proof. ■

4 Bilinear forms

A useful tool for computation is a vector space is a bilinear form, in particular we will be using these to cut out the symplectic subgroups of $GL(n,q)$, in order to get to the sequence of simple groups given by $PSP(2n,q)$, the projective symplectic groups.

A bilinear form is a function $B : V \times V \to F$ on an ordered pair of entries in a vector space to the ground field such that when fixing an entry we get a linear function $B(v,-), B(-,w) : V \to F$. In other words, we require that for all $v, u, w \in V$ and $a \in F$

$$B(v+u,w) = B(v,w) + B(u,w) \text{ and } B(av,w) = aB(v,w)$$
$$B(v,u+w) = B(v,u) + B(v,w) \text{ and } B(v,aw) = aB(v,w)$$

A bilinear form $B : V \times V \to F$ is called *nondegenerate* if $\forall v \in V \setminus \{0\}, \exists u, w \in V : B(v,u) \neq 0$ and $B(w,v) \neq 0$. That is, if B doesn't send any nonzero vector to zero. Otherwise B is called *degenerate*. Note that whether B is degenerate depends on the vector space in question, in particular B may be nondegenerate on a vector space V yet degenerate on a subspace $W \subset V$. For this reason, once we have fixed a bilinear form B, we may call a subspace degenerate or nondegenerate.

4.1 Reflexive forms

A bilinear form is called *reflexive* if $B(w,v) = 0 \Rightarrow B(v,w) = 0$. It is these forms that we are most interested in.

For a reflexive form $B : V \times V \to F$ and subspace $W \subset V$ we define the *orthogonal complement* $W^\perp$ to be the set of $\{v \in V : B(v,w) = 0, \forall w \in W\}$. We then have the following important theorem:

Theorem 12. *If B is a reflexive form on V and W is a nondegenerate subspace of V then $V = W \oplus W^\perp$.*

Proof. First note that W nondegenerate means that $W \cap W^\perp = 0$ so that we have $W + W^\perp = W \oplus W^\perp$. What we need now is only to show that $\dim W + \dim W^\perp = \dim V$ and we will be done. We will have that $\dim W = k$ for some $k \leq n$, so we need only show that $\dim W^\perp \geq n - k$. In order to show this, choose a basis $\{v_1, \ldots, v_k\}$ for W and extend by $\{v_{k+1}, \ldots, v_n\}$ to basis of V. Then every $v \in V$ can be expressed as $v = \sum_1^n a_i v_i$, and $B(v_j, v) = \sum_1^n a_i B(v_j, v_i)$. Then

$$v \in W^\perp \Leftrightarrow \begin{bmatrix} B(v_1,v_1) & \cdots & B(v_1,v_n) \\ \vdots & \ddots & \\ B(v_k,v_1) & \cdots & B(v_k,v_n) \end{bmatrix}\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = 0$$

where $v = \sum_i^n a_i v_i$. But since this is a $k \times n$ matrix with $k < n$ it has rank $\leq k$. Since W is nondegenerate the matrix has rank k, and has nullspace of dimension $n - k$, which is the desired result. ■

4.2 Alternate forms

One class of reflexive forms is *alternate forms*, which are forms such that $B(v,w) = -B(w,v) \forall v, w$. It is immediately clear that alternate forms are reflexive, and so in particular theorem 12 can be applied to them. One immediate thing to note about alternating forms is that $B(v,v) = -B(v,v)$ so that $\forall v \in V : B(v,v) = 0$.

Theorem 13. *If B is a nondegenerate alternating form on a vector space V then* $\dim V$ *is even.*

Proof. We do this proof by induction on dimension of V.

Set $N = V^\perp$ so that the assumption that B is nondegenerate tells us that $N = 0$. Then choose some $u \in V$, by nondegeneracy we have some $v \in V$ such that $B(u,v) \neq 0$. Then consider the subspace W spanned by $\{u, v\}$ clearly W is nondegenerate because any $W \setminus \{0\} \ni x = au + bv$ has at least one of $a \neq 0$ or $b \neq 0$. If $a \neq 0$ then $B(x,v) = aB(u,v) \neq 0$, similarly if $b \neq 0$, so W. So we can apply theorem 12 and $V = W \oplus W^\perp$. We then get $N = V^\perp = (W \oplus W^\perp)^\perp = W^\perp \cap (W^\perp)^\perp = 0$ so that $W^\perp$ is nondegenerate and of dimension $\dim V - 2$. Thus by induction down to dimension ≤ 2 and nondegeneracy forces $\dim V = 2$ because otherwise V is the span of a single vector v, but B is alternating so $B(av, bv) = abB(v,v) = 0$. ■

As a result of theorem 13 when dealing with a nondegenerate alternating form we usually say that V has dimension $n = 2m$ and write the basis of V as $\{v_1, u_2, \ldots, v_m, u_m\}$ where $B(v_i, u_j) = \delta_{ij}$.

5 Symplectic groups

For a nondegenerate alternating form B on vector space V define the *symplectic group* $Sp(V)$ as the set of $\tau \in GL(V)$ such that $B(v,w) = B(\tau v, \tau w)$ for all $v, w \in V$, so that $Sp(V)$ is the set of linear transformations of V respecting the alternating form B. We will show that $Sp(V)$ mod its center is simple, and compute its size.

5.1 Dimension 2

Particularly simple is the case when $\dim V = 2$, in fact we will find that we have already done the work in this case, as $Sp(V) = SL(V)$ in this case.

Theorem 14. *If* $\dim V = 2$ *is nondegenerate with alternating form B then $Sp(V) = SL(V)$.*

Proof. First note that there is a one to one correspondence between matrices and bilinear forms via $B \leftrightarrow M \Leftrightarrow B(v,w) = v^T M w$. The idea is that a bilinear form is entirely determined by its action on a basis. Denote by $\hat{B}$ the matrix corresponding to B relative to some basis then $\widehat{B \circ \tau} = \tau^T B \tau$ so that τ fixes B iff $\tau^T \hat{B} \tau = \hat{B}$. For an alternating form B choose $v \in V$ then by nondegeneracy $B(v,w) \neq 0$ for some $w \in V$ linearly independent of v, thus by rescaling w we can write

$$\hat{B} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

relative to $\{v, w\}$. Thus the condition on τ becomes

$$\tau^T \hat{B} \tau = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 0 & ad - bc \\ bc - ad & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

which is exactly the condition that $\det \tau = 1$ so that $Sp(V) = SL(V)$. ■

So in studying $Sp(V)$ we need only concern ourselves with $\dim V > 2$, since we already covered the case when $\dim V = 2$.

5.2 Some transvection results

Like in the $SL(V)$ case, we turn to transvections in order to simplify things. Unlike in SL, here in Sp not every transvection is symplectic. A symplectic transvection is in fact τ such that $\tau v = v + aB(v,u)u$ for some choice of $a \in \mathbb{F}_q^*$ and $u \in V$. We can see that if B is nondegenerate then $u^\perp$ will be some codimension 1 subspace W so that τ will fix W. Note also that $u \in W$ since $B(u,u) = 0$ so that $\tau v - v = bu \in W$ for all v for some b. We can also see that $B \circ \tau = B$ since the cross terms will come out with opposite signs. We write these symplectic transvections as $\tau_{u,a}$. It isn't immediately clear that this consists of all symplectic transvections, but seeing as any transvections we use here will be constructed in this way, we won't show that fact here, even though it is the case.

Note that for there transvections, if $a = 0$ so that we have $\tau_{u,0}$ we simply get the identity, also multiplying $\tau_{u,a}\tau_{u,b}$ yields $\tau_{u,a+b}$ and conjugating $\sigma\tau_{u,a}\sigma^{-1}$ yields $\tau_{\sigma u,a}$. Lastly, note that $\tau_{bu,a} = \tau_{u,ab^2}$ because u will appear with this factor twice, once inside B and once outside. These are facts that we will use later.

Theorem 15. *The symplectic group $Sp(V)$ is generated by transvections.*

Proof. We use induction on $m = \dim V/2$, where $\dim V = 2$ is already covered by theorems 8 and 14.

For some nondegenerate v with $\dim V = 2m$ we choose $u, v \in V$ such that $B(u,v) \neq 0$ and set W to be the subspace spanned by u and v. Then W is nondegenerate and by theorem 12 we have $V = W \oplus W^\perp$. If we choose any $\sigma \in Sp(V)$ then we will have again that $B(\sigma u, \sigma v) \neq 0$. Then we can find a product of transvections τ to undo the action of σ on u, v i.e. such that $\tau\sigma u = u$ and $\tau\sigma v = v$. Then we will find that $\tau\sigma$ is really just an element of $Sp(W^\perp)$ and can proceed by induction. Thus we need only show that such τ exists.

To find τ taking σu to u if $B(u, \sigma u) \neq 0$ just choose the scalar $B(\sigma u, u)^{-1}$ and the vector $\sigma u - u$ to define our transvection. This will give

$$\tau\sigma u = \sigma u + \frac{B(\sigma u, \sigma u - u)}{B(\sigma u, u)}(\sigma u - u) = \sigma u - (\sigma u - u) = u$$

otherwise we can find some $w \in V$ with $B(w,u) \neq 0$ and $B(w, \sigma u) \neq 0$. So that by a similar method we can get a transvection to move σu to w and then w to u. This takes care of u, now we need transvections to get $\tau\sigma v$ back to v while not moving around u. But if $B(\tau\sigma v, v) \neq 0$ then when we apply that trick with $(\tau\sigma v - v)$ we will find that since $\tau\sigma u = u$ and $\tau\sigma \in Sp(V)$ we have $B(u, \tau\sigma v) = B(u, v)$ so that $B(u, \tau\sigma v - v) = 0$ and so u is indeed fixed by the transvection which puts v back. The last case is then $B(\tau\sigma v, v) = 0$. But here we have $B(v,u) \neq 0 \Rightarrow B(\tau\sigma v, \tau\sigma v + u) \neq 0$ so just as above we can fix u and by transvection move $\tau\sigma v$ to $\tau\sigma v + u$ but then we still have $B(\tau\sigma v + u, v) = B(u,v) \neq 0$ so that we can use a transvection fixing u to move $\tau\sigma v + u$ back to v and so we have a way to multiply by transvection and undo the action of σ on W which allows for our inductive step, and thus $Sp(V)$ is generated by symplectic transvections. ■

It follows that $Sp(V)$ is a subgroup of $SL(V)$.

Theorem 16. *If $|F| > 3$ then $Sp'(n,q) = Sp(n,q)$.*

Proof. Choose some $u \in V$ and some $a, b \in \mathbb{F}_q^*$ such that $b \neq \pm$ then set $c = a/(1-b^2)$ and $d = -b^2c$. Then we have that $c + d = c - b^2c = c(1-b^2) = a$ but the symplectic transvections when composed are additive in their scalar component so that $\tau_{u,c}\tau_{u,d} = \tau_{u,a}$. Choosing then some $\sigma \in Sp(V)$ such that $\sigma u = bu$ this will then give us that

$$\tau_{u,c}\sigma\tau_{u,c}^{-1}\sigma^{-1} = \tau_{u,c}\tau_{\sigma u,-c} = \tau_{u,c}\tau_{bu,-c} = \tau_{u,c}\tau_{u,-b^2c} = \tau_{u,c}\tau_{u,d} = \tau_{u,a}$$

since u and a were arbitrary, that means that we can have any symplectic transvection in $Sp'(v)$ so that $Sp(V)$ is equal to its derived group. ■

Theorem 17. *If $|F| = 3$ and $n \geq 4$ then $Sp'(n,q) = Sp(n,q)$.*

Proof. Choosing a basis $\{u_i, v_i\}$ define

$$\sigma u_1 = u_1 + u_2, \sigma v_1 = v_2, \sigma u_2 = u_1, \sigma v_2 = v_1 - v_2$$
$$\tau u_1 = u_1 - v_1 + v_2, \tau u_2 = u_2 + v_1$$

and both σ, τ are the identity on all the rest. We will find then that $\sigma\tau\sigma^{-1}\tau^{-1} = \tau_{v_1,1}$. By conjugating by some $\theta \in Sp(V)$ with $\theta v_1 = v$ we will get $\tau_{v,1}$ for any v, similarly we get that $\tau_{v,1}^{-1} = \tau_{v,-1}$ so that we in fact have all symplectic transvections and thus all of $Sp(v)$. ■

Theorem 18. *If $|F| = 2$ and $n \geq 6$ then $Sp'(n,q) = Sp(n,q)$.*

Proof. Similar to above, we compute explicitly.

Choosing a basis $\{u_i, v_i\}$ define

$$\sigma u_1 = u_1 + u_3, \sigma v_1 = v_3, \sigma u_2 = u_1, \sigma v_2 = v_1 + v_3, \sigma u_3 = u_2, \sigma v_3 = v_2$$
$$\tau u_1 = u_1 + u_2, \tau u_2 = v_1 + v_2 + u_2 + u_3, \tau u_3 = v_2 + v_3 + u_3$$

and both σ, τ are the identity on all the rest. We will find then that $\sigma\tau\sigma^{-1}\tau^{-1} = \tau_{v_1,1}$. By conjugating by some $\theta \in Sp(V)$ with $\theta v_1 = v$ we will get $\tau_{v,1}$ for any v, so that we in fact have all symplectic transvections and thus all of $Sp(v)$. ■

5.3 Simplicity

We now have almost all the ingredients we need to apply theorem 6, we just need to find a set to act on primitively and a normal subgroup that stabilizes things. Again here a natural place to look is the set of lines in V, so that we will be passing to the quotient of $Sp(V)$ by its center.

It turns out that the center of $Sp(V)$ is ± 1. To see this realize that something in the center needs to commute with transvections, so that $\sigma\tau_{u,a}\sigma^{-1} = \tau_{u,a}$. If $\sigma = \lambda I$ then $\sigma\tau_{u,a}\sigma^{-1} = \tau_{u,\lambda^2 a}$ so that we need $\lambda = \pm 1$. The worry might be however that something more complicated slipped into the center, but actually note that if $\sigma u = v$ then $\sigma\tau_{u,a}\sigma^{-1} = \tau_{v,a}$ so that σ has got to fix lines to be in the center, so the center is ± 1.

Theorem 19. *Sp(V) acts primitively on the lines in V.*

Proof. As above, we only need to worry about $n \geq 4$. Suppose to the contrary that for some set of lines S in V with $|S| \geq 2$ we have that $\sigma S = S$ or $\sigma S \cap S = \emptyset$ for all $\sigma \in Sp(V)$.

Suppose then that for every pair of lines $[u], [v] \in S$ we have that $B(u,v) = 0$. Then choose some pair $[u] \neq [v]$ in S we will have then by nondegeneracy of B that for some $x \in V$ $B(u,x) = 1$ and $B(v,x) = 0$. Setting W to be the subspace of V spanned by u and x (x isn't a multiple of u because in that case $B(u,x) = 0$) then let H be the subset of $Sp(V)$ acting as the identity on W. We have that $B(u,x)$ means that W is nondegenerate so that $V = W \oplus W^\perp$ so that every $\sigma \in Sp(W^\perp)$ yields an element of H, in fact in this way $Sp(W^\perp) \cong H$. Choosing some $w \in W^\perp \setminus \{0\}$ such that $B(v,w) \neq 0$ (such exists because B nondegenerate and $v \in W^\perp$) we get that for some $\tau \in H \cong Sp(W^\perp)$ we have $\tau v = w$ since we can choose $a = B(v,w)^{-1}$ and $y = v - w$ so that $\tau_{a,y} v = w$. But we also have $\tau u = u$ so that $[u] \in S \cap \tau S$ so that by assumption of S we have $\tau S = S$ so that $\tau[v] = [w] \in S$ but this contradicts that $B(y,z) = 0$ for all $[y], [z] \in S$,

We must then have some $[u], [v] \in S$ such that $B(u,v) \neq 0$. Choose any line $[w]$ in V. If $B(u,w) \neq 0$ we will be able to find as we did in theorem 15 a σ such that $\sigma u = u$ and $\sigma v = w$ then $u \in S \cap \sigma S$ so that $\sigma S = S$ by the assumption we used to define S, and $[w] \in \sigma S = S$. If $B(u,w) = 0$ we find some $x \in V$ with $B(u,x), B(w,x) \neq 0$ then applying the same notion twice we will have some sigma with $\sigma u = w$ and $\sigma x = x$ but as before with x playing the role of w we have that $[x] \in S$ and $\sigma x = x$ so that $\sigma S = S$, so that again we find $\sigma[u] = [w] \in S$.But $[w]$ was arbitrary, so S is all the lines in V, and $Sp(V)$ acts primitively. ■

Theorem 20. *With the exceptions $PSp(2,2), PSp(2,3), PSp(4,2)$ every projective symplectic group $PSp(V)$ is simple.*

Proof. This follow from the fact that $H = \text{Stab}_{Sp(V)}[v]$ will contain the subgroup of transvections $\tau_{a,v}, a \in \mathbb{F}_q^*$. This is an abelian subgroup of the stabilizer because it acts additively in the scalar entry, It is normal in the stabilizer because conjugation by σ just sends $\tau_{v,a}$ to $\tau_{\sigma v,a}$ but all σ in the stabilizer fix v as a line so only change the scalar entry, which is free to take on any value. This set of transvections also has its conjugates generating all of $Sp(V)$, since its conjugates include all transvection because it has all scalar values and for each pair of vectors v, u there;s a symplectic linear transformation τ with $\tau v = u$. We then need only to have $Sp(V)$ equal to its derived group and we can apply theorem 6 to $PSp(V)$ which acts faithfully on the lines in V. ■

5.4 Size

Theorem 21. *The size of $Sp(2m, q)$ is*

$$q^{m^2} \prod_{i=1}^{m} (q^{2i} - 1)$$

.

Proof. We prove by induction on m. The base case of $m = 1$ is covered by theorem 14 and section 2.1 so that we get $Sp(2, q) = (q^2 - 1)(q^2 - q)/(q - 1) = q(q^2 - 1)$.

In order to show the inductive step we use orbit-stabilizer on a nondegenerate plane $W \subset V$. If we give W an ordered basis v, w with $B(v, w) \neq 0$ then we can send v to any nonzero vector u in V via an element of $Sp(V)$ so long as we send w to a vector x such that $B(u, x) = B(v, w)$. The number of nonzero vectors u in V is just $q^n - 1$, and the vectors x with $B(u, x) = B(v, w)$ will be given by the choice of a line in $V \setminus u^\perp$ since every line in that space will have exactly one vector satisfying $B(u, x) = B(v, w)$ so this will be $(q^n - q^{n-1})/(q - 1) = q^{n-1}$ so that the orbit has size $(q^n - 1)q^{n-1}$, the stabilizer will then have size $|Sp(2(m - 1), q)|$ because it is isomorphic to the symplectic group of $W^\perp$. We have then

$$|Sp(2m, q)| = (q^{2m} - 1)q^{2m-1}|Sp(2(m - 1), q)| = q^{2m-1}q^{(m-1)^2}\left(\prod_{i=1}^{m-1}(q^{2i} - 1)\right)(q^{2m} - 1) = q^{m^2}\prod_{i=1}^{m}(q^{2i} - 1)$$

as desired. ■

We will also have that $|PSp| = |Sp|/2$ except when $q = 2^k$ for some k in which case $1 = -1$ so that $Sp/\{\pm 1\} = Sp$.

6 The groups we left behind

We found 3 cases for which $PSp(V)$ fails to be simple, and we look now at how they fail at this.

The first group is $PSp(2, 2) \cong Sp(2, 2) \cong SL(2, 2)$ in this case our vector space has 3 nonzero elements. We find that $SL(2, 2) = GL(2, 2)$ since 1 is the only nonzero element of $\mathbb{F}_2$ so that in fact $Sp(2, 2) = SL(2, 2) = GL(2, 2)$ acts as the symmetric group on three letters S_3 on this vector space. This group isn't simple as it has normal subgroup $A_3 \triangleleft S_3$.

The second group that we found not to be simple is $PSp(2, 3)$. Here we find that $Sp(2, 3) = SL(2, 3)$ with order $3(3^2 - 1) = 24$. This is suggestive, as it is the order of S_4, the symmetric group on 4 letters. We can see that it fact it is the same group by seeing $\mathbb{F}_3^2$ as a grid of nine dots with the zero vector in the middle. Then the four objects to permute are the four lines in the grid, two diagonals, one vertical and one horizontal. If some permutation seems not to be in SL we can make the determinant one by flipping one of the lines, which keeps the same action on lines but changes the determinant. We see then that $PSp(2, 3)$ must be A_4, the alternating group on 4 letters. This group isn't normal however, as it has a normal subgroup given by the double transpositions $\{1, (12)(34), (13)(24), (14)(23)\}$ this subgroup is normal since conjugation in S_n preserves the cycle presentation of permutations.

The last of these is $Sp(4, 2)$ which has size 720. This is suggestive again because it is the size of S_6 the group of permutations of six letters. It isn't too easy to see how these two groups are the same, and a more robust and elegant treatment can be found in [2]. The idea is to build $\mathbb{F}_2^4$ as a quotient of a subspace of $\mathbb{F}_2^6$

with the standard symplectic form by setting $u = (1,1,1,1,1,1)$ and then taking $u^\perp/u \cong \mathbb{F}_2^4$. The action of S_6 is of course then given by permuting the entries of a vector. We now take the symplectic form on this reduced space to be the normal inner product, or in other words the $\mathbb{F}_2$ sum of the number of overlaps between a pair of vectors (this is well defined mod u because being in $u^\perp$ means having an even number of nonzero entries) e.g. $\langle (0,1,1,0,1,1),(0,0,1,0,1,0)\rangle = 0\times 0+1\times 0+1\times 1+0\times 0+1\times 1+1\times 0 = 1+1 = 0$. This is stable under S_6 action since the number of overlaps remains unchanged when you permute two vectors in the same way. It isn't too hard to see that this S_6 action is faithful, the only worry would be something like that for some $\sigma \in S_6$ and $v \in \mathbb{F}_2^6$ we would have $\sigma v = v + u$ and that v is the unique vector that σ acts non-trivially on. For this we note that the number of nonzero entries of $u + v$ is different than the number of nonzero entries of v and so σ cannot map v to $u + v$. Thus we have that S_6 acts faithfully on $\mathbb{F}_2^4$ and preserves the symplectic form we defined, and we also have that $|PSp(4,2)| = |Sp(4,2)| = |S_6|$ so that we see that in fact $PSp(4,2) = Sp(4,2) \cong S_6$. Finally, we have that $PSp(4,2)$ isn't simple because it contains the normal subgroup $A_6 \lhd S_6 \cong PSp(4,2)$.

References

[1] Larry Grove, *Classical Groups and Geometric Algebra*. American Mathematical Society, 2002.

[2] Tim Silverman, *Re: A Wrinkle in the Mathematical Universe* August 18, 2015 https://golem.ph.utexas.edu/category/2015/08/a_wrinkle_in_the_mathematical.html#c049471