Diameters and Mixing Times of Classical Groups

Roshani Nanayakkara

July 27, 2007

ABSTRACT

This thesis studies the diameter of the Special Linear and Symplectic groups over finite fields with prime order.

We shall establish two bounds on its diameter of $SL_n(p)$ with respect to a small, well known generating set. This partially solves a problem by Lubotzky. Then we use one of these bounds to establish an upper bound on the mixing time of the Special Linear Group given by the uniform random walk on this generating set with the identity.

Finally, we establish similar bounds for $Sp_{2n}(p)$.

I would like to thank my supervisor, Martin Liebeck, for his help throughout my four years as a PhD student and EPSRC for providing me with funding for my research. I would also like to thank K. Magaard for allowing me to use his proof of 4.1.1.

CONTENTS

1. INTRODUCTION AND STATEMENT OF RESULTS

1.1 Group Diameters

Let G be a finite group with a symmetric generating set S (note that a generating set is said to be symmetric if $s \in S \Rightarrow s^{-1} \in S$). Each element of G can be written as a product of a finite number of elements of S. We define the length of g with respect to S, denoted $l_S(g)$, to be the smallest positive number k such that $g = s_1 s_2 ... s_k$ with the $s_i \in S$.

Let H be any subgroup of G. Then the diameter of H with respect to S, denoted $l_S(H)$ is defined as follows.

$$l_S(H) = \max\{l_S(h) : h \in H\}$$

We may also define the diameter of G with respect to a given generating set in terms of an associated graph.

A graph $\Gamma(V, E)$ is a set of vertices, V, with an associated set of edges, E, where E is a subset of $V \times V$. If there is an edge between two vertices of a graph, the vertices are said to be adjacent. We denote an edge between vertices v_1 and v_2 as $(v_1, v_2) \in E$.

Given two vertices $v_a, v_b \in V$, we say there is a path of length k between v_a and v_b if there are vertices $v_1, v_2, ...v_{k-1} \in V$ such that $(v_a, v_1), (v_1, v_2), ...(v_{k-1}, v_b) \in E$. The distance between v_a and v_b, denoted $d(v_a, v_b)$, is the smallest number k such that there exists a path of length k between v_a and v_b.

We define the diameter of a graph $\Gamma(V, E)$ to be the largest number m such that there exist vertices v_a and v_b with $d(v_a, v_b) = m$.

Given a group G and a symmetric generating set, S, we may define an associated graph $X(G, S)$, called the Cayley Graph of G with respect to S, as follows. We set the vertex set of $X(G, S)$ to be the set of elements in G and we define the edge set by setting $(g_i, g_j) \in E$ if and only if $g_i = sg_j$ for some $s \in S$.

Note that the diameter of the Cayley graph $X(G, S)$ is the same as $l_S(G)$.

Lemma 1.1.1: Let G be a group with a generating set S. Then we have $l_S(G) \geq \frac{\log |G|}{\log |S|} - 1$.

Proof.

For ease of notation I will use $d = l_S(G)$.

Note that each element of G can be written as the product of at most d elements of S. So for each $g \in G$ we have $g = s_1 s_2 s_3 ....s_k$ where each $s_i \in S$ and $0 \leq k \leq d$. There are at most $|S|^k$ expressions of this form. This gives us that

$$\begin{aligned} |G| &\leq \sum_{k=0}^{d} |S|^k \\ &= \frac{|S|^{d+1} - 1}{|S| - 1} \\ &\leq |S|^{d+1} \end{aligned}$$

Taking logs we now have

$$\log |G| \leq (d+1) \log |S|$$

and so $d \geq \frac{\log |G|}{\log |S|} - 1$.

■

In [1], Babai, Lubotzky and Kantor show that, if G is a simple group, then there exists a generating set S of G such that $|S| \leq 7$ and $l_S(G) < m \log |G|$.

Here m is a constant that is approximately 10^{10}. However the S_G are complicated and difficult to construct. There are several open questions surrounding the diameters of groups with respect to more natural generating sets.

In [8] (problem 8.13), Lubotzky conjectured that, if $G = SL_n(p)$, there is a natural generating set, S, such that $l_S(G)$ is of order $\log|G|$. More precisely, there exists a constant C such that

$$l_S(G) \leq C \log|G|.$$

Since $\log|G|$ is roughly $n^2 \log p$, Lubotzky's conjecture would mean that $l_S(G)$ would be $Cn^2 \log p$ for some constant C. By 1.1.1, if $|S|$ is bounded, this bound is tight up to the constant.

The generating set in Lubotzky's conjecture is defined as follows. Let the standard basis of $(\mathbb{F}_p)^n$ be denoted by $e_1, e_2, ...e_n$. Define y to be the matrix in G that sends e_i to e_{i+1} for $i = 1, 2, 3, ...n-1$ and sends e_n to $(-1)^{n+1}e_1$. That is

$$y = \begin{pmatrix} & & & & (-1)^{n+1} \\ 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \end{pmatrix}.$$

Define x to be the transvection $1 + e_{1,2}$, i.e

$$x = \begin{pmatrix} 1 & 1 & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & & 1 \end{pmatrix}.$$

We define S to be the set $\{x^{\pm 1}, y^{\pm 1}\}$.

In Chapter 3 we take Lubotzky's suggested generating set, S, and prove the following diameter results, which go some way towards proving Lubotzky's conjecture.

Theorem 1.1.2: The diameter of $SL_n(p)$ with respect to S is at most $50n^2p$.

Theorem 1.1.3: There exists a constant K, which is not dependent on n and p, such that the diameter of $SL_n(p)$ with respect to S is at most $Kn^3 \log p$.

Remark 1.1.4: We may bound K above very crudely by 36500.

In chapter 5 we generalize this to the Symplectic Groups, $Sp_{2n}(p)$, over fields with prime order. We construct a generating set, S' as follows.

Using the same notation as in our definition of S, define v to be the matrix

$$\begin{pmatrix} x & 0 \\ 0 & x^{-TJ} \end{pmatrix}$$

and w to be the product

$$\begin{pmatrix} & & & & -1 \\ & 1 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & 1 \\ 1 & & & & \end{pmatrix} \begin{pmatrix} \begin{pmatrix} & & 1 \\ 1 & & \\ & \ddots & \\ & 1 & \end{pmatrix} & 0 \\ 0 & \begin{pmatrix} & & 1 \\ 1 & & \\ & \ddots & \\ & 1 & \end{pmatrix}^{-TJ} \end{pmatrix}$$

where M^{-TJ} is the matrix obtained by taking the inverse of the transpose of a matrix M, and conjugating by a fixed matrix J which will be described later. Define S' to be the set $\{v^{\pm 1}, w^{\pm 1}\}$.

Note that we have chosen this generating set because, as in the case of our generating set of $SL_n(p)$, this set contains a short root element and an element of the Weyl Group of $Sp_{2n}(p)$.

We have the following two Theorems about the diameter of $Sp_{2n}(p)$ with respect to S'.

Theorem 1.1.5: Let p be an odd prime. Then the diameter of $Sp_{2n}(p)$ with respect to S' is at most $187n^2p$.

Theorem 1.1.6: Let p be an odd prime. Then the diameter of $Sp_{2n}(p)$ with respect to S' is at most $Kn^3 \log p$, where K is a constant that is not dependent on n and p.

Lubotzky's conjecture for $SL_n(p)$ was recently shown to be true in the paper by Kassabov and Riley, [7]. This used a different, less elementary method than mine to prove the result. My bound is better for small values of p. My results about the diameter of $Sp_{2n}(p)$ are completely new and have been published in [10].

1.2 Random Walks on Groups

Let G be a finite group with generating set $S = \{s_1, s_2, ...s_k\}$. Let P be a probability measure on G, so $P(g) \geq 0$ for each $g \in G$ and $\sum_{g \in G} P(g) = 1$.

Consider the process where at each step we choose an element of G with probability given by P and study the evolution of the product of all the

chosen elements. This process is called the random walk on G generated by the probability measure P.

For example, a random walk can be used to model the following scenario. Suppose there are n cards labeled 1 to n from left to right. A process is carried out where, at each stage, one card is chosen at random, and then another, with repetition allowed. The two cards are interchanged or, if the same card was chosen both times, the cards are left alone.

We are interested in the order of the cards after k steps of this process have been carried out.

This can be modeled as follows. The order of the cards can be thought of as an element of S_n. For example, if the only first two cards have been interchanged, we can view the new order of the cards as the permutation $(1, 2) \in S_n$.

At each step a permutation, is chosen using the probability measure P where

$$P(g) = \begin{cases} \frac{2}{n^2} & \text{if } g \text{ is a transposition} \\ \frac{1}{n} & \text{if } g \text{ is the identity} \\ 0 & \text{otherwise.} \end{cases}$$

This permutation is applied to the cards. The order of the cards after k steps

of this process is the product of the k permutations that have been chosen.

If, after k steps of a random walk, the product of all the chosen elements is g, we say that the random walk is at g after k steps. We define $P^{*k} : G \to [0, 1]$ by setting $P^{*k}(g)$ to be the probability the walk generated by P is at g after k steps.

We are usually interested in how large k has to be before the element produced by k steps of the walk is 'random.' By this we mean that P^{*k} is very similar to the uniform probability distribution on G.

In order to formalize this concept we need to introduce the notion of a metric on probability measures.

Definition 1.2.1: Suppose Q and R are probability measures on G. Then the total variation distance between Q and R is $\|Q - R\|$ where

$$\|Q - R\| = \frac{1}{|2|} \sum_{g \in G} |Q(g) - R(g)|.$$

To proceed, we will need to know when the probability distribution corresponding to a random walk converges to the uniform distribution. The following section will show that convergence depends only on the elements of G for which $P(g) \neq 0$.

1.3 Convergence of Random Walks

Throughout this section, G will be a finite group with a symmetric generating set S. We will study the random walk given by a probability distribution P, where $P(g) = \frac{1}{|S|}$ if and only if $g \in S$.

The following well known Theorem gives us a condition which guarantees that P^{*k} will converge to the uniform distribution, U on G.

Theorem 1.3.1: Let G, S and P be as described above. Then P^{*k} will converge to the uniform distribution on G as k tends to infinity if and only if S is not contained in a coset of a proper normal subgroup of G of index 2. It is not easy to find a published proof of this result so I provide one here.

We will need to make a few definitions before we prove 1.3.1.

Definition 1.3.2: Let Γ be a graph with vertex set V. Then Γ is bipartite if there exist two disjoint, non-empty subsets of the vertex set, V_1 and V_2, with $V_1 \cup V_2 = V$ so that no two vertices in V_1 are adjacent and no two vertices in V_2 are adjacent.

We may label the elements of G as $g_1, g_2, ..., g_n$. Now we can define a $|G| \times |G|$ matrix, $A = (a_{i,j})$ as follows. The entry $a_{i,j}$ of A is 1 if $g_i = sg_j$ for some

$s \in S$ and zero otherwise. We call A the adjacency matrix of G with respect to S. Note that A is also the adjacency matrix of the Cayley graph $\Gamma(G, S)$.

Suppose now that P is a probability distribution as described in 1.3.1. We may define a $|G|$- dimensional vector, p_k, as $p_k = (P^{*k}(g_1), P^{*k}(g_2), ...P^{*k}(g_n))^T$. The is called the probability vector of P^{*k}.

Note that $a_{i,j} = 1$ if and only if $g_i g_j^{-1} \in S$. So $a_{i,j} = 1$ if and only if $P(g_i g_j^{-1}) = \frac{1}{m}$, where $m = |S|$, and hence $\frac{1}{m} a_{i,j} = P(g_i g_j^{-1})$.

Now

$$P^{*k}(g_j) = \sum_{i=1}^{n} P(g_j g_i^{-1}) P^{*k-1}(g_i)$$

$$\sum_{i=1}^{n} \frac{1}{m} a_{j,i} P^{*k-1}(g_i).$$

So we have

$$\begin{pmatrix} P^{*k}(g_1) \\ P^{*k}(g_2) \\ \vdots \\ \vdots \\ P^{*k}(g_n) \end{pmatrix} = \frac{1}{m} \begin{pmatrix} \sum_{i=1}^{n} a_{1,i} P^{*k-1}(g_i) \\ \sum_{i=1}^{n} a_{2,i} P^{*k-1}(g_i) \\ \vdots \\ \vdots \\ \sum_{i=1}^{n} a_{n,i} P^{*k-1}(g_i) \end{pmatrix}$$

$$= \frac{1}{m} A p_{k-1}.$$

To prove 1.3.1 we will need the following well-known, elementary Lemma,

which I shall state without proof.

Lemma 1.3.3: Let Γ be a regular connected graph with valency m and adjacency matrix, A. Then

1) m is an eigenvalue of A with multiplicity 1. The eigenspace of m is generated by $(1, 1, ..., 1)$.

2) if λ is any eigenvalue of A, then $|\lambda| \leq m$.

3) $-m$ is an eigenvalue of A if and only if Γ is bipartite.

Theorem 1.3.4: Let G, S and P be as described at the beginning of this section. If the Cayley Graph of G with respect to S is not bipartite, we have that $\|P^{*k} - U\| \to 0$ as k tends to infinity.

Proof. Let A be the adjacency matrix of G with respect to S. Define M to be the matrix $\frac{1}{m}A$, where $m = |S|$. Since S is symmetric, M is symmetric so $M = TDT^{-1}$ for some matrix T and some diagonal matrix D. Suppose

$$D = \begin{pmatrix} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \ddots & & \\ & & & \lambda_{n-1} & \\ & & & & \lambda_n \end{pmatrix}$$

The λ_i are the eigenvalues of M. From 1.3.3 part 1), We know that precisely one of these is 1. By part 2), we know that for any other eigenvalue, λ, we have $|\lambda| \leq 1$. Since the Cayley graph of G with respect to S is bipartite, we know that -1 is not an eigenvalue of M.

So, we may assume that $\lambda_1 = 1$ and for $i > 1$, $|\lambda_i| < 1$. Now,

$$M^k p = T D^k T^{-1} p$$

tends to

$$T \begin{pmatrix} 1 & & & & \\ & 0 & & & \\ & & \ddots & & \\ & & & 0 & \\ & & & & 0 \end{pmatrix} T^{-1} p$$

as k tends to infinity. Define this limit as v.

Now $M^{k+1}p$ also tends to v and so $Mv = v$. From 1.3.3 part 1) we know that $v = \lambda(1, 1, ...1)^T$. Also, since v is a probability vector we have $\lambda = \frac{1}{n}$ and v represents the uniform probability distribution on G. ■

Lemma 1.3.5: If the Cayley graph $X(G, S)$ is bipartite, then S is contained

in a coset of a normal subgroup of G with index 2.

Proof. Since $X(G, S)$ is bipartite we may divide the vertex set, V, into two disjoint, non-empty sets V_1 and V_2.

Without loss of generality, we'll say that the vertex corresponding to the identity lies in V_1. Then, since there is an edge between the vertex corresponding to the identity and each vertex corresponding to elements of S, each element of S must lie in V_2. Repeating this argument, we see that, if $l_S(g)$ is odd, the vertex corresponding to g must lie in V_2 and, if $l_S(g)$ is even, it must lie in V_1. Hence the elements of g whose vertices lie in V_1 form a group.

This group is of index 2 and hence is normal. So, the elements of S lie in the coset of a normal subgroup of index 2. ■

Proof. (of 1.3.1) Suppose S is a symmetric generating set of G which is not contained in a coset of a normal subgroup of G. Then from 1.3.5 we see that the Cayley graph $X(G, S)$ is not bipartite and so by 1.3.3 P^{*k} tends to the uniform distribution as k tends to infinity.

Conversely, if S is contained in the coset of any normal subgroup, say Ng, where N is the normal subgroup and g is a coset representative, then the support of P^{*k} will be contained within Ng^k. Hence P^{*k} is always zero on all but $|N|$ elements of G and cannot tend to the uniform distribution on G. ■

1.4 Mixing Times

Definition 1.4.1: The mixing time of G with respect to the random walk generated by P is the smallest number, k, such that $\|P^{*k} - U\| < (2e)^{-1}$.

The reason we choose $(2e)^{-1}$ is that, if $\|P^{*k} - U\| < (2e)^{-1}$, then we can guarantee that $\|P^{*mk} - U\|$ decays exponentially as m increases linearly.

Many of the original problems on mixing times of finite groups were answered by Persi Diaconis. Several problems were posed to answer the question of how long it takes to shuffle a pack of n cards. The example given above illustrates the model used to answer this question when the pack of cards is shuffled by repeatedly interchanging two cards.

The mixing time of this 'Random Transposition' walk was determined by Diaconis and Shashahani in [5]. The method, which will be discussed in more detail in Chapter 2, relied on the fact that the probability measure, P, is constant on conjugacy classes.

In situations where the probability measure is not a class function, other methods are needed to determine mixing times. A technique known as 'Comparison' is used. As the name suggests, comparison involves comparing some properties of the random walk being studied with those of another walk which has already been studied. In order to use comparison techniques we

will need to know an upper bound for the diameter of the group in question. The techniques and results of Comparison will be discussed further in Chapter 2.

We shall be studying a random walk generated by the uniform probability on a small generating set. The generating set is not a union of conjugacy classes and so we will need to use comparison techniques to work out its mixing time.

In Chapter 4 we use comparison with the uniform random walk on the set of transvections to determine a bound for the mixing time of $SL_n(p)$ given by the following probability distribution, where x and y are as defined in section 1.1.

$$Q(g) = \begin{cases} \frac{1}{5} & g = x, y, x^{-1}, y^{-1} \text{ or } I \\ 0 & \text{otherwise.} \end{cases}$$

We shall show

Theorem 1.4.2: The mixing time of $SL_n(p)$ given by the probability distribution Q is of order at most $n^6(\log p)^3$.

In Chapters 6 we shall solve a similar problem for the Symplectic Group,

$Sp_{2n}(p)$. We use the probability distribution,

$$Q(g) = \left\{ \begin{array}{cc} \frac{1}{5} & g = v, w, v^{-1}, w^{-1} \text{ or } I \\ 0 & \text{otherwise.} \end{array} \right\}$$

where v and w are as defined in section 1.1.

Theorem 1.4.3: The mixing time of $Sp_{2n}(p)$ given by the probability measure Q described above is of order at most $n^8(\log p)^3$.

2. PRELIMINARIES

2.1 Classical Groups

Let p be a prime and define $\mathbb{F}_p$ to be the finite field with p elements. We use V to denote the vector space of dimension n over $\mathbb{F}_p$. Let $\mathbb{F}_p^*$ denote $\mathbb{F}_p \backslash \{0\}$.

We define the general linear group, $GL_n(p)$, to be the group of $n \times n$ matrices with entries in $\mathbb{F}_p$ that have non-zero determinant.

Note that the diagonal matrices in $GL_n(p)$ form a subgroup which is isomorphic to $(\mathbb{F}_p^*)^n$.

The special linear group, $SL_n(p)$, is the subgroup of $GL_n(p)$ which consists of all the matrices with determinant 1.

The classical groups are subgroups of $GL_n(p)$ and $SL_n(p)$ that stabilise certain forms on V. We shall only need to deal with the Symplectic Group which is defined as follows.

If the dimension of V is $2n$, there exists a skew-symmetric, non-degenerate, bilinear form on V. This is known as a symplectic form. Given a symplectic form (.), on V, there exists a basis, $e_1, e_2, ..., e_n, f_1, f_2, ..., f_n$ with the following properties. For all $i, j \in \{1, 2, ..., n\}$ we have

$$(e_i.e_j) = (f_i.f_j) = 0$$

and

$$(e_i.f_j) = \delta_{i,j} = -(f_j.e_i).$$

We define the symplectic group, denoted $Sp_{2n}(p)$, to be the subgroup of $SL_{2n}(p)$ that preserves this symplectic bilinear form.

Now we discuss the structure of $SL_n(p)$. Fix a basis, $e_1, e_2, ...e_n$ of V and define B to be the stabiliser in $SL_n(p)$ of $\{\langle e_1, ...e_i\rangle : i = 1, 2, ...n - 1.\}$ Then B is the group of upper triangular matrices in $SL_n(p)$. Define N to be the stabiliser of $\{\langle e_1\rangle, \langle e_2\rangle, ...\langle e_n\rangle\}$. Then N is the group of matrices with precisely one non-zero entry in each row and column and is known as the group of monomials.

The Weyl Group of $SL_n(p)$ is the quotient $\frac{N}{B\cap N}$. The following important result about the structure of $SL_n(p)$ has been taken from [2] (Theorem 8.2.2).

Theorem 2.1.1: Using the notation described above we have $SL_n(p) = BNB$.

We may also define subgroups B and N for $Sp_{2n}(p)$. In this case we let $e_1, e_2, e_3, ...e_n, f_1, f_2, ...f_n$ be a basis as descibed above and define B as the stabiliser of

$$\{\langle e_1, ...e_i\rangle : i = 1, 2, ...n\}$$

.If we order the standard symplectic basis $e_1, e_2, ...e_n, f_n, f_{n-1}, ..., f_1$ then B is the subgroup of upper triangular matrices in $Sp_{2n}(p)$. The subgroup N is defined as the stabiliser of $\{\langle e_1\rangle, \langle e_2\rangle, ...\langle e_n\rangle, \langle f_1\rangle, \langle f_2\rangle, ...\langle f_n\rangle\}$ in $Sp_{2n}(p)$. Again, N is just the subgroup of monomials in $Sp_{2n}(p)$.

In the case of $Sp_{2n}(p)$, the Weyl Group is isomorphic to the semi-direct product $2^n.S_n$. The subgroup 2^n is the normal subgroup generated by the set of matrices $\{t_i | i = 1, 2, ..., n\}$ where t_i sends e_i to f_i and f_i to $-e_i$.

Again [2] provides us with the following result about the structure of $Sp_{2n}(p)$.

Theorem 2.1.2: Using the notation described above we have $Sp_{2n}(p) = BNB$.

Note that in general each classical group can be written as a product of subgroups BNB. This is proved in [2].

For any matrix M denote M^R to be the matrix obtained by reversing the order of the columns of M. Now define J to be the matrix I^R.

Remark 2.1.3: A matrix M lies in $Sp_{2n}(p)$ if and only if

$$M \begin{pmatrix} 0 & J \\ -J & 0 \end{pmatrix} M^T = \begin{pmatrix} 0 & J \\ -J & 0 \end{pmatrix}.$$

This follows immediately from the definition of the symplectic form.

For ease of notation, given any matrix, M, I will denote $(M^{-1})^T$ by M^{-T} and $J^{-1}(M^{-1})^T J$ by M^{-TJ}.

We will use the following property of $Sp_{2n}(p)$ several times in Chapter 5.

Lemma 2.1.4: Let A and C be matrices in $GL_n(p)$. Then the $2n \times 2n$ matrix

$$\begin{pmatrix} A & B \\ 0 & C \end{pmatrix},$$

where B is an $n \times n$ matrix, lies in $Sp_{2n}(p)$ if and only if $C = A^{-TJ}$, and $AJB^T = BJA^T$.

Proof.

From the above remark we know that

$$\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

is in $Sp_{2n}(p)$ if and only if

$$\begin{pmatrix} A^T & 0 \\ B^T & C^T \end{pmatrix} = \begin{pmatrix} 0 & J \\ -J & 0 \end{pmatrix}^{-1} \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}^{-1} \begin{pmatrix} 0 & J \\ -J & 0 \end{pmatrix}$$

so

$$\begin{pmatrix} A & B \\ 0 & C \end{pmatrix} \begin{pmatrix} 0 & J \\ -J & 0 \end{pmatrix} \begin{pmatrix} A^T & 0 \\ B^T & C^T \end{pmatrix} = \begin{pmatrix} 0 & J \\ -J & 0 \end{pmatrix}$$

and hence

$$\begin{pmatrix} -BJA^T + AJB^T & AJC^T \\ -CJA^T & 0 \end{pmatrix} = \begin{pmatrix} 0 & J \\ -J & 0 \end{pmatrix}.$$

Now, looking at the bottom left hand quadrants of the matrices in the last equation we have $-CJA^T = -J$ so $C = JA^{-T}J^{-1} = A^{-TJ}$. Equating the top left hand quadrants we have $AJB^T = BJA^T$.

■

Corollary 2.1.5: Suppose M is a matrix in $GL_n(p)$. Then the map $\phi : SL_n(p) \rightarrow Sp_{2n}(p)$ given by

$$\phi(M) = \begin{pmatrix} M & 0 \\ 0 & M^{-TJ} \end{pmatrix}$$

is an embedding.

The transvections in $SL_n(p)$ are mapped by ϕ to the short root elements of $Sp_{2n}(p)$.

2.2 Two Useful Diameter Results

The following result about the diameter of $SL_2(p)$ from [1] will be needed to prove Theorems 1.1.3 and 1.1.6.

Theorem 2.2.1: (8.1 from [1]) Let $G = SL_2(p)$ and let T be the set $\left\{ \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\}$. Then T generates G and $l_T(G) \leq m \log p$ for some constant m.

Remark 2.2.2: As stated in [1], a crude upper bound for m is 500.

The following result about $SL_2(p)$ will be used in the proof of 1.1.6.

Theorem 2.2.3: (proposition 2.2.2 of [9]) Let p be an odd prime. Define Σ_p^2 to be the subset of $SL_2(p)$

$$\left\{ \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \right\}.$$

Then the set $\{X(SL_2(p), \Sigma_p^2)\}_{p \text{ prime}}$ is an expander family and hence $l_{\Sigma_p^2}(SL_2(p)) \leq c \log p$ for some constant c.

2.3 Mixing Times

Let G be a finite group and let P be a probability measure on G. Recall from Chapter 1, that the mixing time of the random walk on G generated by P is the lowest number k such that $\|P^{*k} - U\| < (2e)^{-1}$ where U is the uniform distribution on G.

In order to determine the mixing time we need to know about P^{*k}. Usually we only have information about about the probability measure P and know very little about P^{*k}. We can begin to relate these two measures using the idea of convolution.

Definition 2.3.1: Suppose Q and R are probability measures on a group G. The convolution of Q with R, denoted $Q * R$ is the probability measure on

G given by

$$Q * R(g) = \sum_{h \in G} Q(gh^{-1})R(h).$$

It is easy to see that $P^{*k}(g)$ is given by $P^{*(k-1)} * P(g)$. Hence we see that

$$P^{*k} = \underbrace{(...((P * P) * P)... * P)}_{\text{k times}}$$

To proceed further I will need to introduce the idea of a Fourier Transform of a probability measure on G.

Suppose that ρ is an irreducible representation of G. The Fourier Transform of P at ρ is given by

$$\hat{P}(\rho) = \sum_{g \in G} P(g)\rho(g).$$

These Fourier transforms share many properties of Fourier transforms of continuous functions. A property that will be particularly useful is that Fourier transforms turn convolutions into products. By this we mean

$$\widehat{Q * R}(\rho) = \hat{Q}(\rho)\hat{R}(\rho)$$

and so

$$\hat{P^{*k}}(\rho) = (\hat{P}(\rho))^k.$$

In [3], Diaconis has shown that a Fourier transform $\hat{P}(\rho)$ can be used to bound the mixing time of G given by P in the following Lemma.

Lemma 2.3.2: [The Upper Bound Lemma]

Let P be a probability on a group G. Then

$$\|P^{*k} - U\|^2 \leq \frac{1}{4} \sum_{\rho \neq 1} d_\rho tr(\hat{P}^{*k}(\rho)\hat{P}^{*k}(\rho)^*)$$

where $\hat{P}^{*k}(\rho)^*$ is the complex conjugate transpose of $\hat{P}^{*k}(\rho)$.

Here the sum is taken over all irreducible representations, ρ of G with $\rho \neq 1$. Now, if P is constant on conjugacy classes, then $P(g) = P(hgh^{-1})$ for all $g, h \in G$. So for a fixed irreducible representation ρ we have

$$\begin{aligned}\hat{P}(\rho) &= \sum_{g\in G} P(hgh^{-1})\rho(hgh^{-1}) \\ &= \rho(h)\left(\sum_{g\in G} P(g)\rho(g)\right)\rho(h^{-1}) \\ &= \rho(h)\left(\hat{P}(\rho)\right)\rho(h^{-1})\end{aligned}$$

for any element h of G. So $\rho(h)\hat{P}(\rho) = \hat{P}(\rho)\rho(h)$ for all $h \in S_n$. Schur's Lemma now gives us that $\hat{P}(\rho) = c_\rho I$ for some constant c_ρ. Now since

$$\hat{P^{*k}}(\rho) = (\hat{P}(\rho))^k,$$

we have $P^{*k}(\rho) = c_\rho^k I$ and so $Tr(P^{*k}(\rho)P^{*k}(\rho)^*) = d_\rho |c_\rho|^{2k}$ where d_ρ is the dimension of ρ.

Substituting into the Upper Bound Lemma we now have

$$\|P^{*k} - U\|^2 \leq \frac{1}{4}\sum_{\rho\neq 1} d_\rho^2 |c_\rho|^{2k}$$

Now if we bound c, we will be able to find an upper bound for the mixing time for the walk generated by P. To illustrate how this can be used to find bounds for mixing times, we return to the random transpositions example

that was introduced in Chapter 1.

Recall that

$$\hat{P}(\rho) = c_\rho I = \sum_{g \in S_n} P(g)\rho(g).$$

Taking traces we get

$$d_\rho c_\rho = \sum_{g \in S_n} P(g)\chi_\rho(g)$$

where χ_ρ is the character of ρ. We know that P takes value $\frac{2}{n^2}$ on the $\frac{n(n-1)}{2}$ transpositions, takes value $\frac{1}{n}$ on the identity and is zero everywhere else. Substituting in the values of $P(g)$ into the above equality, we get

$$d_\rho c_\rho = \left(\frac{n(n-1)}{2}\right)\left(\frac{2}{n^2}\right)\chi_\rho(\tau) + \frac{1}{n}\chi_\rho(id)$$

where τ is a transposition. Simplifying this gives

$$c_\rho = \left(\frac{n-1}{n}\right)\left(\frac{\chi_\rho(\tau)}{d_\rho}\right) + \frac{1}{n}$$

and the Upper Bound Lemma gives us

$$\|P^{*k} - U\|^2 \leq \frac{1}{4} \sum_{\rho \neq 1} d_\rho^2 \left(\left(\frac{n-1}{n} \right) \left(\frac{\chi_\rho(\tau)}{d_\rho} \right) + \frac{1}{n} \right)^{2k}$$

Bounding this expression requires some detailed knowledge of the Representation Theory of the Symmetric Group. Diaconis and Shashahani used a formula by Frobenius that relates the character ratio, $\frac{\chi_\rho(\tau)}{d_\rho}$, to the shape of the diagram of the corresponding λ-tableaux to establish that the largest term in the sum corresponds to the $n-1$-dimensional irreducible representation.

The largest term to bound is

$$(n-1)^2(1 - \frac{2}{n})^{2k}.$$

Using Stirling's formula and a MacLaurin expansion we have

$$(n-1)^2(1 - \frac{2}{n})^{2k} = e^{2\log(n-1)+2k(1-\frac{2}{n})} = e^{-\frac{4k}{n}+2\log n+o(\frac{k}{n^2})}$$

Now if we put $k = \frac{1}{2} n \log n + cn$, the above decreases exponentially with c. Hence the mixing time of S_n given by the random walk generated by P is of order $n \log n$. More detail about how this sum was bounded is given in [4].

2.4 Comparison

If we have a probability measure that is not constant on conjugacy classes, the techniques described above do not help us.

A probability measure P on a group G is defined as being symmetric if for each $g \in G$ we have $P(g) = P(g^{-1})$. If a probability measure on a group has this property, we may use a method known as Comparison to bound the mixing time of the group given by that measure. The Comparison results in this section are taken from [4].

Let Q be a symmetric probability on a finite group G. We may associate a $|G| \times |G|$ matrix, M, to Q with $M_{st} = Q(st^{-1})$. Since the matrix M has real value entries, is symmetric and is doubly stochastic, the Perron-Frobenius Theorem states that its eigenvalues, $\pi_0, \pi_, ...\pi_{|G|-1}$ have the property $1 = \pi_0 \geq \pi_1 ... \geq \pi_{|G|-1} \geq -1$.

Lemma 2.4.1:

$$4\|Q^{*k} - U\|^2 \leq |G|\pi_{max}^{2k}$$

where $\pi_{max} = \max\{\pi_1, |\pi_{|G|-1}|\}$.

From this we see that if we can establish bounds on the eigenvalues of the matrix, M, we can find upper bounds for the mixing time given by Q.

We may bound the smallest eigenvalue using the following Theorem from [4].

Theorem 2.4.2: If Q is a symmetric probability function on a symmetric set of generators of a finite group, the smallest eigenvalue of Q, $\pi_{|G|-1}$ is bounded below by $-1 + 2Q(id)$.

We can bound the other eigenvalues by using properties of $L^2(G)$, the linear space of all real functions on G. Usually the eigenvalues of Q can be bounded in terms of the eigenvalues of another walk, which we'll denote by $\tilde{Q}$, that we already know something about.

Define an inner product on real valued functions on G by

$$\langle f_1, f_2 \rangle = \sum_{s \in G} f_1(s) f_2(s).$$

If Q is symmetric, it defines a Laplace operator on the space of all linear functions on G. This is given by

$$(I - Q)f(s) = f(s) - \sum_{t \in G} f(t) Q(ts^{-1})$$

and has eigenvalues $1 - \pi_i$. There is a quadratic form associated with this operator. It is given by

$$\varepsilon(f,f) = \langle (I-Q)f, f\rangle$$

and is known as the Dirichlet form.

Theorem 2.4.3: Let Q and $\tilde{Q}$ be symmetric probabilities with eigenvalues $\{\pi_i\}_{i=0}^{|G|-1}$ and $\{\tilde{\pi}_i\}_{i=0}^{|G|-1}$ respectively and associated Dirichlet forms ε and $\tilde{\varepsilon}$. If there exists a constant $A > 0$ such that $\tilde{\varepsilon}(f,f) < A\varepsilon(f,f)$ for all f, then we have $\pi_i \leq 1 - \frac{1-\tilde{\pi}_i}{A}$ for all i.

A value for A is given by the following theorem from [4]. Recall that a generating set S of a group G is symmetric if for each $x \in S$, we have $x^{-1} \in S$.

Theorem 2.4.4: Let Q and $\tilde{Q}$ be symmetric probabilities on G. Let S be a symmetric generating set and let S be the support of Q. Then $\tilde{\varepsilon} \leq A\varepsilon$ for

$$A = \max_{s\in S} \frac{1}{Q(s)} \sum_{g\in G} |g| N(s,g) \tilde{Q}(g).$$

Here $N(s,g)$ is the minimum number times s appears when g is written as a product of elements in S, and $|g| = l_S(g)$.

To illustrate how these results may be used, we shall use comparison with

the Random Transpositions example to work out the mixing time of S_n given by the following walk.

Let $S = \{id, (1,2), (1,2,...,n), (1,2,...,n)^{-1}\}$ and define P to be the uniform distribution on S. It can be shown that any transposition in S_n can be written as a product of at most $3n$ elements of S. We'll use comparison with the random walk on transpositions given by the probability distribution

$$\tilde{P}(g) = \begin{cases} \frac{2}{n^2} & \text{if } g \text{ is a transposition} \\ \frac{1}{n} & \text{if } g \text{ is the identity} \\ 0 & \text{otherwise.} \end{cases}$$

Substituting this into the theorem, we get that $A \leq 36n^2$.

In order to apply any of the other results we need to know the eigenvalues of the matrix corresponding to $\tilde{P}$. In [5], Diaconis and Shahshahani have shown that the largest of these, $\tilde{\pi}_1$, is $1 - \frac{2}{n}$.

Now 2.4.3 gives us that $\pi_1 \leq 1 - \frac{1}{18n^3}$.

The smallest eigenvalue can be bounded using 2.4.2. This is greater than or equal to $-\frac{1}{2}$.

Recall $\pi_{max} = \max\{\pi_1, |\pi_{|G|-1}|\}$. From the above we have $\pi_{max} = |\pi_{|G|-1}| \leq 1 - \frac{1}{18n^2}$.

Using the bound for the smallest eigenvalue in 2.4.1 we have

$$4||Q^k - U|| \leq |G|\pi_1^{2k}$$

$$\leq n!(1 - \frac{1}{18n^3})^{2k}$$

$$\leq \exp\left(n \log n - \frac{2k}{18n^3}\right).$$

So the mixing time decreases exponentially as $\frac{2k}{18n^3}$ decreases by multiples of $n \log n$, so k will have to be of order $n^4 \log n$.

Note that in [4], Diaconis shows that this bound may be improved to $n^3 \log n$. It can also be shown that this bound is tight.

2.5 Mixing Times of Classical Groups

The random walks on the classical groups we will be studying are not generated by class probability functions. Hence we will need to use comparison techniques to bound their mixing times.

The following result comes from [9] (page 175).

Theorem 2.5.1: Let P be a probability function on G such that the support of P is a single conjugacy class, C, and P is constant on C. Then the eigenvalues

of the matrix, M associated with P are precisely the character ratios $\frac{\chi_\rho(C)}{d_\rho}$, where ρ is an irreducible representation of G, d_ρ is the dimension of ρ and $\chi_\rho(C)$ is the value the character of ρ takes on the elements of C.

In the case of $SL_n(p)$, we will use comparison with the set of transvections to find a bound for the mixing time given by the probability function Q on our generating set, where Q is given by

$$Q(g) = \begin{cases} \frac{1}{5} & g = x, y, x^{-1}, y^{-1} \text{ or } I \\ 0 & \text{otherwise.} \end{cases}$$

Here x and y are the elements of our generating set S, which was defined in Chapter 1.

Let T be the set of transvections in G and let $\tilde{Q}$ be the uniform distribution on T. The random walk on G generated by $\tilde{Q}$ has an associated matrix $\tilde{M}$. Since the support of $\tilde{Q}$ is the single conjugacy class, T, the eigenvalues of $\tilde{M}$ are the character ratios

$$\left\{ \frac{\chi_\rho(t)}{\chi_\rho(1)} : \chi_\rho \text{ is an irreducible representation of } G \right\}$$

where $t \in T$.

In his paper, [6], Gluck has shown that these character ratios are bounded

above by $\frac{19}{20}$ in the following theorem.

Theorem 2.5.2: Suppose G is a quasi-simple group of Lie type. Then, if χ is an irreducible representation of G, we have $|\frac{\chi(x)}{\chi(1)}| \leq \frac{19}{20}$ for all $x \in G \backslash Z(G)$.

Using this bound in 2.4.3 will give us a bound on the eigenvalues of the matrix associated with Q. Then, by applying 2.4.1 we will be able to get a bound for the mixing time of G given by the random walk generated by Q.

In the case of $Sp_{2n}(p)$ we set Q to be the probability function

$$Q(g) = \begin{cases} \frac{1}{5} & g = v, w, v^{-1}, w^{-1} \text{ or } I \\ 0 & \text{otherwise.} \end{cases}$$

where u and v are the elements of the generating set S' described in Chapter 5.

The set of short root elements in $Sp_{2n}(p)$, is the conjugacy class v^G, where v is the element of our generating set S of $Sp_{2n}(p)$, described in Chapter 5. We will use comparison with the set of short root elements to determine the mixing time of $Sp_{2n}(p)$ given by the following probability function.

We define $\tilde{Q}$ to be the uniform probability function on v^G. As before, we can bound the eigenvalues of $\tilde{M}$, the matrix associated with $\tilde{Q}$, using 2.5.2. Again the eigenvalues of the matrix associated with $\tilde{M}$ are bounded above

by $\frac{19}{20}$.

From this and 2.4.3 we obtain a bound for the eigenvalues of the matrix associated with Q and 2.4.1 will give us a bound on the mixing time of $Sp_{2n}(p)$ with respect to the probability function Q.

3. THE DIAMETER OF THE SPECIAL LINEAR GROUP

3.1 Proof of Theorem 1.1.2

Let G be $SL_n(p)$, the Special Linear Group of $n \times n$ matrices over $\mathbb{F}_p$. In this chapter we will establish two bounds on the diameter of G with respect to the generating set $\{x^{\pm 1}, y^{\pm 1}\}$ which was described in Chapter 1.

For convenience, we recall the definition of x and y. We use the standard notation $e_{i,j}$ to denote the matrix with 1 in position (i, j) and zeroes everywhere else. Define $E_{i,j}$ as the matrix $I + e_{i,j}$. To illustrate the proofs we will display $n \times n$ matrices with the zero entries omitted.

As in Chapter 1, we may define a generating set, S, of G as follows. Let the standard basis of $(\mathbb{F}_p)^n$ be denoted by $e_1, e_2, ...e_n$. Define y to be the matrix in G that sends e_i to e_{i+1} for $i = 1, 2, 3, ...n - 1$ and sends e_n to $(-1)^{n+1}e_1$. That is

$$y = \begin{pmatrix} & & & (-1)^{n+1} \\ 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}.$$

Define x to be the transvection $E_{1,2}$, i.e

$$x = \begin{pmatrix} 1 & 1 & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}.$$

Set $S = \{x^{\pm 1}, y^{\pm 1}\}$. In this chapter we shall show that S generates G and bound the diameter of G with respect to S. For $g \in G$ we will denote the length of g in S by $l_S(g)$. If H is a subset of G we define $l_S(H) = \max\{l_S(h) | h \in H\}$.

Most of this chapter will be devoted to proving Theorem 1.1.2,

The diameter of $SL_n(p)$ with respect to S is at most $50n^2p$.

Theorem 1.1.3, the statement of which was

There exists a constant K, which is not dependent on n and p, such that the diameter of $SL_n(p)$ with respect to S is at most $Kn^3 \log p$.

will then be deduced at the end of the chapter.

In Chapter 2 we saw how G can be written as the product BNB where N is the subgroup of monomial matrices and B is the group of upper triangular matrices. Since any matrix in B may be written as the product of an upper uni-triangular matrix and a diagonal matrix, we may write $G = UNU$ where U is the group of upper uni-triangular matrices.

We shall bound the diameter of G by bounding $l_S(U)$ and $l_S(N)$. To do this we shall make use of the following identities.

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} x = \begin{pmatrix} a_{1,1} & a_{1,1}+a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,1}+a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n,1} & a_{n,1}+a_{n,2} & \cdots & a_{n,n} \end{pmatrix}$$

$$x \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} = \begin{pmatrix} a_{1,1}+a_{2,1} & a_{1,2}+a_{2,2} & \cdots & a_{1,n}+a_{2,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix}$$

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} y = \begin{pmatrix} a_{1,2} & a_{1,3} & \cdots & a_{1,n} & -a_{1,1} \\ a_{2,2} & a_{2,3} & \cdots & a_{2,n} & -a_{2,1} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ & & & & \\ a_{n,2} & a_{n,3} & \cdots & a_{n,n} & -a_{n,1} \end{pmatrix}$$

$$y \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} = \begin{pmatrix} -a_{n,1} & -a_{n,2} & \cdots & -a_{n,n} \\ a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n-2,1} & a_{n-2,2} & \cdots & a_{n-2,n} \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n} \end{pmatrix}$$

We think of a post-multiplication of a matrix by x as adding the first column of the matrix to its second column. Similarly post-multiplying a matrix by x^{-1} subtracts the first column from the second column.

We may view pre-multiplication by x as adding the second row to the first row. Pre-multiplication by x^{-1} corresponds to subtracting the second row from the first.

Post-multiplication by y applies the permutation $(1, n, n-1..., 3, 2)$ to its columns and multiplies the nth column by $(-1)^{n+1}$. Pre-multiplication by y applies the permutation $(1, 2, 3, ...n)$ to the rows and multiplies the top

row by $(-1)^{n+1}$. Multiplication by y^{-1} reverses the operations described above, so, for example, post multiplying by y^{-1} multiplies the nth column by $((-1)^{n+1}$ then applies the permutation $(1, 2, ...n)$ to the columns.

Construction of N

The matrices in N have precisely one non-zero entry in each row and column. Fix a matrix $X \in N$ with entry x_i in row i. There exists a permutation $\pi \in S_n$ such that for each i, x_i appears in position $(i, \pi(i))$ of X.

We shall construct X (i.e., express X as a product of elements of S) by first constructing a matrix P, where P is monomial with entries of ± 1 in positions $(i, \pi(i))$. Our first aim is to construct the transposition matrices, $\{t_{i,j} : 1 \leq i, j \leq n\}$ which are defined as follows. Suppose $i \neq j$. Define $t_{i,j} \in G$ to be the matrix that sends e_i to $-e_j$, sends e_j to e_i and leaves all other basis elements fixed. To construct $t_{i,j}$ we will need to determine the length of a matrix of form $E_{i,j}$.

Lemma 3.1.1: $l_S(E_{1,k}) \leq 10k$.

Proof. Let Z_k be the $n \times n$ matrix with entries of 1 on the diagonal and in

positions (i, j) for $i < j \leq k$. All other entries are zero. So Z_k has this form.

$$\begin{pmatrix} 1 & 1 & \cdots & 1 & 1 & & & \\ & 1 & \cdots & 1 & 1 & & & \\ & & \ddots & \vdots & \vdots & & & \\ & & & 1 & 1 & & & \\ & & & & 1 & & & \\ & & & & & \ddots & & \\ & & & & & & 1 & \\ & & & & & & & 1 \end{pmatrix}$$

Note that $Z_2 = x$. We construct Z_{i+1} from Z_i as follows. Post multiplying by y performs the permutation $(1, n, n-1, ...2)$ on the columns of Z_k and multiplies the last column by δ where $\delta = (-1)^{n+1}$. So post multiplying Z_k by y^{k-1} gives us the matrix of form

$$\begin{pmatrix} 1 & & & & \delta & \delta & \cdots & \delta \\ 1 & & & & & \delta & \cdots & \delta \\ \vdots & & & & & & \ddots & \vdots \\ 1 & & & & & & & \delta \\ 1 & & & & & & & \\ & 1 & & & & & & \\ & & \ddots & & & & & \\ & & & 1 & & & & \end{pmatrix}.$$

Now multiplying by x adds the first column to the second column so

$$Z_k y^{k-1} x = \begin{pmatrix} 1 & 1 & & & \delta & \delta & \cdots & \delta \\ 1 & 1 & & & & \delta & \cdots & \delta \\ \vdots & \vdots & & & & & \ddots & \vdots \\ 1 & 1 & & & & & & \delta \\ 1 & 1 & & & & & & \\ & 1 & & & & & & \\ & & \ddots & & & & & \\ & & & 1 & & & & \end{pmatrix}$$

Now, if we post multiply by y^{-1}, we apply the permutation $(1, 2, ..., n)$ to the columns and multiply the first column by δ. So post multiplying by $y^{-(k-1)}$ gives

$$\begin{pmatrix} 1 & 1 & \cdots & 1 & 1 & 1 & & \\ & 1 & \cdots & 1 & 1 & 1 & & \\ & & \ddots & \vdots & \vdots & \vdots & & \\ & & & 1 & 1 & 1 & & \\ & & & & 1 & 1 & & \\ & & & & & 1 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix}$$

which is Z_{i+1}. So for any $i \in \{2, 3, ...n-1\}$ we have $Z_{i+1} = Z_i y^{i-1} x y^{-(i-1)}$. Hence for each $2 \leq k \leq n$, we have

$$Z_k = Z_{k-1} y^{k-2} x y^{-(k-2)}.$$

Substituting $Z_{k-1} = Z_{k-2} y^{k-3} x y^{-(k-3)}$ we get,

$$\begin{aligned} Z_k &= Z_{k-2} y^{k-3} x y^{-(k-3)} y^{k-2} x y^{-(k-2)} \\ &= Z_{k-2} y^{k-3} x y x y^{-(k-2)}. \end{aligned}$$

Now writing Z_{k-2} in terms of Z_{k-3} we have

$$\begin{aligned} Z_k &= Z_{k-3} y^{k-4} x y^{-(k-4)} y^{k-3} x y^{-1} x y^{-(k-2)} \\ &= Z_{k-3} y^{k-4} x y x y x y^{-(k-2)}. \end{aligned}$$

Continuing like this we eventually have

$$Z_k = Z_2 y (xy)^{k-3} x y^{-(k-2)}$$

and since $Z_2 = x$,

$$Z_k = xy(xy)^{k-3}xy^{-(k-2)}$$

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

From Z_k we construct the matrix Y_k which has form

$$\begin{pmatrix}
1 & & & & & & 1 & & & \\
& 1 & & & & & 1 & & & \\
& & 1 & & & & 1 & & & \\
& & & 1 & & & 1 & & & \\
& & & & \ddots & & \vdots & & & \\
& & & & & 1 & 1 & & & \\
& & & & & & 1 & 1 & & \\
& & & & & & 1 & & & \\
& & & & & & & \ddots & & \\
& & & & & & & & \ddots & \\
& & & & & & & & & 1
\end{pmatrix}.$$

By post multiplying the matrix Z_k by y^{k-3}, we repeatedly multiply the first column by δ then apply the permutation $(1, n, n-1, ..., 2)$.

Hence

$$Z_k y^{k-3} = \begin{pmatrix} 1 & 1 & 1 & & & \delta & \delta & \cdots & \delta \\ 1 & 1 & 1 & & & & \delta & \cdots & \delta \\ \vdots & \vdots & \vdots & & & & & \ddots & \vdots \\ 1 & 1 & 1 & & & & & & \delta \\ 1 & 1 & 1 & & & & & & \\ & 1 & 1 & & & & & & \\ & & 1 & & & & & & \\ & & & 1 & & & & & \\ & & & & 1 & & & & \end{pmatrix}.$$

Now post multiplying by x^{-1} subtracts the first column from the second so

$$Z_k y^{k-3} x^{-1} = \begin{pmatrix} 1 & & 1 & & & \delta & \delta & \cdots & \delta \\ 1 & & 1 & & & & \delta & \cdots & \delta \\ \vdots & & \vdots & & & & & \ddots & \vdots \\ 1 & & 1 & & & & & & \delta \\ 1 & & 1 & & & & & & \\ & 1 & 1 & & & & & & \\ & & 1 & & & & & & \\ & & & 1 & & & & & \\ & & & & 1 & & & & \end{pmatrix}.$$

Post multiplying by y^{-1} gives us

$$Z_k y^{k-3} x^{-1} y^{-1} = \begin{pmatrix} 1 & 1 & & 1 & & & \delta & \cdots & \delta \\ 1 & 1 & & 1 & & & & \ddots & \vdots \\ \vdots & \vdots & & \vdots & & & & & \delta \\ 1 & 1 & & 1 & & & & & \\ & 1 & & 1 & & & & & \\ & & 1 & 1 & & & & & \\ & & & 1 & & & & & \\ & & & & 1 & & & & \\ & & & & & 1 & & & \\ & & & & & & & & \end{pmatrix}.$$

Now post multiplication by x^{-1} gives us

$$Z_k y^{k-3} x^{-1} y^{-1} x^{-1} = \begin{pmatrix} 1 & & & 1 & & & \delta & \cdots & \delta \\ 1 & & & 1 & & & & \ddots & \vdots \\ \vdots & & & \vdots & & & & & \delta \\ 1 & & & 1 & & & & & \\ & 1 & & 1 & & & & & \\ & & 1 & 1 & & & & & \\ & & & 1 & & & & & \\ & & & & 1 & & & & \\ & & & & & 1 & & & \\ & & & & & & & & \end{pmatrix}.$$

Continuing to repeatedly post multiply by y^{-1} and x^{-1} we eventually have

$$Y_k = Z_k y^{k-3}(x^{-1}y^{-1})^{(k-3)}x^{-1}$$

$$= (xy)^{k-2}xy^{(-k-2)}y^{k-3}(x^{-1}y^{-1})^{(k-3)}x^{-1}$$

$$= (xy)^{k-2}xy^{-1}(x^{-1}y^{-1})^{(k-3)}x^{-1}.$$

Pre-multiplying Y_k by y performs the permutation $(1, n, n-1, n-2, ...2)$ on the rows of Y_k and multiplies the last row by δ. So

$$y^{-1}Y_k = \begin{pmatrix} 1 & & & & & & 1 & & \\ & 1 & & & & & 1 & & \\ & & 1 & & & & 1 & & \\ & & & \ddots & & & \vdots & & \\ & & & & 1 & & 1 & & \\ & & & & & 1 & 1 & & \\ & & & & & & 1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & \ddots \\ & & & & & & & & 1 \\ \delta & & & & & & \delta & & \end{pmatrix}$$

.

Now pre-multiplying this matrix by x^{-1} subtracts the second row from the first.

$$x^{-1}y^{-1}Y_k = \begin{pmatrix} 1 & -1 & & & & & 0 & & & \\ & 1 & & & & & 1 & & & \\ & & 1 & & & & 1 & & & \\ & & & \ddots & & & \vdots & & & \\ & & & & 1 & & 1 & & & \\ & & & & & 1 & 1 & & & \\ & & & & & & 1 & & & \\ & & & & & & & \ddots & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & 1 \\ \delta & & & & & & \delta & & & \end{pmatrix}.$$

Pre multiplying by y^{-1} now gives us

$$\begin{pmatrix}
 & 1 & & & & 1 & & \\
 & & 1 & & & 1 & & \\
 & & & \ddots & & \vdots & & \\
 & & & & 1 & 1 & & \\
 & & & & & 1 & 1 & \\
 & & & & & 1 & & \\
 & & & & & & \ddots & \\
 & & & & & & & \ddots \\
 & & & & & & & & 1 \\
\delta & & & & & \delta & & \\
 & \delta & -\delta & & & & &
\end{pmatrix}$$

and so

$$x^{-1}y^{-1}x^{-1}y^{-1}Y_k = \begin{pmatrix} & 1 & -1 & & & 0 & & \\ & & 1 & & & 1 & & \\ & & & \ddots & & \vdots & & \\ & & & & 1 & 1 & & \\ & & & & & 1 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \\ \delta & & & & & \delta & & \\ & \delta & -\delta & & & & & \end{pmatrix}.$$

Continuing to pre multiply by $x^{-1}y$ we eventually have

$$(x^{-1}y^{-1})^{k-2}Y_k = \begin{pmatrix} & & & & 1 & 0 & & \\ & & & & & 1 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \\ & & & & & & & \\ \delta & & & & & \delta & & \\ & \delta & -\delta & & & & & \\ & & \ddots & \ddots & & & & \\ & & & \delta & -\delta & & & \end{pmatrix}.$$

Pre multiplying by y performs the permutation $(1, 2, ...n)$ on the rows and multiplies row 1 by δ, so

$$y^{(k-2)}(x^{-1}y^{-1})^{k-2}Y_k = \begin{pmatrix} 1 & & & & & & & 1 & & \\ & 1 & -1 & & & & & & & \\ & & 1 & -1 & & & & & & \\ & & & 1 & -1 & & & & & \\ & & & & \ddots & \ddots & & & & \\ & & & & & 1 & -1 & & & \\ & & & & & & 1 & & & \\ & & & & & & & 1 & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & \ddots \\ & & & & & & & & & & 1 \end{pmatrix}$$

We call this matrix W_k. So

$$W_k = y^{(k-2)}(x^{-1}y^{-1})^{(k-2)}Y_k$$

$$= y^{(k-2)}(x^{-1}y^{-1})^{(k-2)}(xy)^{k-2}xy^{-1}(x^{-1}y^{-1})^{(k-3)}x^{-1}$$

To get $E_{1,k}$ from this matrix we just need to remove the -1 entries from above the diagonal. Using similar arguments as before, post-multiplying W_k by yxy^{-1} gives us

$$\begin{pmatrix}
1 & & & & & & 1 & & \\
& 1 & & & & & & & \\
& & 1 & -1 & & & & & \\
& & & 1 & -1 & & & & \\
& & & & \ddots & \ddots & \vdots & & \\
& & & & & 1 & -1 & & \\
& & & & & & 1 & & \\
& & & & & & & 1 & \\
& & & & & & & & \ddots \\
& & & & & & & & & 1
\end{pmatrix}.$$

Pre-multiplying this matrix by y^2xy^{-2} gives us

$$\begin{pmatrix} 1 & & & & & & & 1 & & \\ & 1 & & & & & & & & \\ & & 1 & & & & & & & \\ & & & 1 & -1 & & & & & \\ & & & & \ddots & \ddots & & \vdots & & \\ & & & & & 1 & -1 & & & \\ & & & & & & 1 & & & \\ & & & & & & & 1 & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & \ddots \\ & & & & & & & & & & 1 \end{pmatrix}$$

Carrying on like this we have

$$E_{1,k} = W_k yxy^{-1}y^2xy^{-2}...y^{k-2}xy^{-(k-2)}$$

and so

$$E_{1,k} = W_k(yx)^{(k-2)}y^{-(k-2)}$$

$$= y^{(k-2)}(x^{-1}y^{-1})^{(k-2)}(xy)^{k-2}xy^{-1}(x^{-1}y^{-1})^{(k-3)}x^{-1}(yx)^{(k-2)}y^{-(k-2)}$$

Hence $l_S(E_{1,k}) < 10k$. ■

Corollary 3.1.2: $l_S(E_{i,j}) \leq 12n$

Proof. Suppose $1 < k < n$ .Pre-multiplying $E_{1,k}$ by y^{-1} applies the permutation $(1, 2, ...n)$ to the rows and multiplies the top row by δ. So

$$yE_{1,k} = \begin{pmatrix} & & & & & & \delta \\ 1 & & & 1 & & & \\ & 1 & & & & & \\ & & 1 & & & & \\ & & & 1 & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}.$$

Now post multiplying by y^{-1} performs the permutation $(1, 2, 3, ...n)$ on the columns and multiplies the first column by δ. So We are left with $yE_{1,k}y^{-1} = E_{2,k+1}$. By repeating this argument we have $y^r E_{1,k} y^{-r} = E_{1+r,k+r}$ providing $k + r \leq n$. Applying the same argument to any $E_{a,b}$ shows that we have $yE_{a,b}y^{-1} = E_{a+1,b+1}$ providing $i, j < n$.

If, $i < j \leq n$ we have $E_{i,j} = y^{(i-1)} E_{1,j-(i-1)} y^{-(i-1)}$. Hence if $i < j \leq n$, we have

$$l_S(E_{i,j}) \leq E_{1,j-(i-1)} + 2(i-1)$$

$$\leq 10(j - (i-1)) + 2(i-1) < 10j - 8(i-1).$$

This shows that if $i < j \leq n$ then $l_S(E_{i,j}) < 10n$. Also it shows that a matrix $E_{b,n}$ has length of at most $10n - 8(b-1)$.

Now observe that

$$yE_{1,n}y^{-1} = y \begin{pmatrix} 1 & & & & & 1 \\ & 1 & & & & \\ & & \ddots & & & \\ & & & \ddots & & \\ & & & & 1 & \\ & & & & & 1 \end{pmatrix} y^{-1}$$

$$= \begin{pmatrix} 1 & & & & & \\ \delta & 1 & & & & \\ & & 1 & & & \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}$$

This is $E_{2,1}^{\pm 1}$. Similarly for all $1 \leq b \leq n-1$ we have $yE_{b,n}y^{-1} = E_{b+1,1}^{\pm 1}$.

Suppose $2 \leq a \leq n$. From the argument in the previous paragraph we have

$$l_S(E_{a,1}) = l_S(E_{a,1})^{-1} \leq 2 + l_S(E_{a-1,n}) \leq 10n - 8(a-2) + 2.$$

Now if $j < i \leq n$, we can write $E_{i,j} = y^{-(j-1)} E_{i-(j-1),1} y^{(j-1)}$. So

$$l_S(E_{i,j}) < l_S(E_{i-(j-1),1}) + 2(j-1)$$

$$< 10n - 8(i-(j-1)) + 2 + 2(j-1) < 10n + 10(j-1) - 8i + 2$$

Since $i > j$ this gives us

$$l_S(E_{i,j}) < 10n + 2(j-1) + 2 \leq 12n.$$

■

Lemma 3.1.3: $l_S(t_{i,j}) \leq 34n$

Proof.

Note that, since $t_{i,j}^{-1} = t_{j,i}$, we only need to show how to construct $t_{i,j}$ where $j > i$. The matrix $E_{1,k}$ has form

$$\begin{pmatrix} 1 & & & & 1 & & & \\ & 1 & & & & & & \\ & & \ddots & & & & & \\ & & & 1 & & & & \\ & & & & 1 & & & \\ & & & & & 1 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix}$$

As mentioned in the remark above, pre-multiplying by $E_{k,1}^{-1}$ will subtract the 1st row from the kth row. So

$$E_{k,1}^{-1}E_{1,k} = \begin{pmatrix} 1 & & & & 1 & & & \\ & 1 & & & & & & \\ & & \ddots & & & & & \\ & & & 1 & & & & \\ -1 & & & & 0 & & & \\ & & & & & 1 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix}.$$

Now pre-multiplying by $E_{1,k}$ adds the kth row to the 1st so

$$E_{1,k}E_{k,1}^{-1}E_{1,k} = \begin{pmatrix} 0 & & & & 1 & & \\ & 1 & & & & & \\ & & \ddots & & & & \\ & & & 1 & & & \\ -1 & & & & 0 & & \\ & & & & & 1 & \\ & & & & & & \ddots \\ & & & & & & & 1 \end{pmatrix} = t_{1,k}.$$

By applying 3.1.1 and 3.1.2 we see that $l_S(t_{1,k}) \leq 32n$. Also $y^m t_{1,k} y^{-m} = t_{1+m,k+m}$ for $m \leq n-k$. So we have $t_{i,j} = y^{(i-1)} t_{1,j-(i-1)} y^{(i-1)}$. and hence $l_S(t_{i,j}) \leq 34n$. ■

Recall that for each matrix $X \in N$, there is a permutation π such that the non-zero entries of X appear in the positions $(i, \pi(i))$. We denote the entry in the ith row of X as x_i. We will now construct P, a monomial with entries of ± 1 in positions $(i, \pi(i))$.

We may write π as a product of disjoint cycles $\pi = \pi_1\pi_2...\pi_r$. Suppose $\pi_i = (p_{i_1}, p_{i_2}...p_{i_k})$. Then the product $P_i = t_{p_{i_{k-1}},p_{i_k}} t_{p_{i_{k-2}},p_{i_{k-1}}}...t_{p_{i_1},p_{i_2}}$ is a monomial matrix with the non-zero entry of the jth row appearing in the $\pi_i(j)$th column. Also $(-1)^{i_k+1}$ is the non-zero entry in the row p_{i_k} and all other rows have a non-zero entry of 1.

If we define $P = P_1P_2...P_r$ then P is monomial with the non-zero entry of row j in the $\pi(j)$th column. We can write P as a product of at most n of the matrices $t_{i,j}$. Without loss of generality we can say P has a $(-1)^{i_k+1}$ in the row corresponding to the last number appearing in each π_i and 1 in every other row.

Lemma 3.1.4: $l_S(P) \leq 34n^2$

Proof. This follows immediately from 3.1.3 and the discussion above. ■

Define A to be the matrix

$$\begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix},$$

where $a_j = (-1)^{i_k+1}x_j$ if j is the last number appearing in one of the π_i and $a_j = x_j$ otherwise. It is easy to see that $AP = X$. In order to construct A

we will need to construct the matrices of form

$$M_a = \begin{pmatrix} a & & & & \\ & a^{-1} & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}$$

where $a \in \mathbb{F}_p^*$.

Lemma 3.1.5: $l_S(M_a) \leq 5np$

Proof. We have

$$x^{-a} = \begin{pmatrix} 1 & -a & & & \\ & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}.$$

Now, using the identities from the beginning of this section, post multiplying the above matrix by $(E_{2,1})^{a^{-1}}$ adds a^{-1} times its second column from its first column.

So

$$x^{-a}(E_{2,1})^{a^{-1}} = \begin{pmatrix} {\scriptstyle 1+(-a)a^{-1}} & -a & & & \\ {\scriptstyle a^{-1}} & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} = \begin{pmatrix} 0 & -a & & & \\ a^{-1} & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}.$$

By post multiplying by x^{-a} we add $-a$ times the first column to the second column. We get

$$x^{-a}E_{2,1}^{a^{-1}}x^{-a} = \begin{pmatrix} 0 & -a & & & \\ {\scriptstyle a^{-1}} & {\scriptstyle 1+(-a)a^{-1}} & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} = \begin{pmatrix} 0 & -a & & & \\ a^{-1} & 0 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}.$$

Now to get M_a we just post multiply by $t_{1,2}$.

We have $M_a = x^{-a}E_{2,1}x^{-a}t_{1,2}$. Now $E_{2,1} = y^{-n}xy^n$ so $l_S(E_{2,1}) \leq 2n-1$. Also, following the same proof as 3.1.3 we have $t_{1,2} = xE_{2,1}^{-1}x$. So $l_S(t_{1,2}) \leq 2n+1$.

Now

$$l_S(M_a) \leq 2p + (2n-1)p + (2n+1)p \leq 5np.$$

■

Lemma 3.1.6: $l_S(A) \leq 6n^2p$

Proof.

It is easy to check that

$$y^{-1}...y^{-1}y^{-1}M_{a_1}yM_{a_2a_1}yM_{a_3a_2a_1}...M_{a_na_{n-1}...a_1}y$$

gives the matrix A.

Each matrix M_a has length of at most $5np$. Hence A has length less than or equal to $5n^2p + 2n$, which is less than or equal to $6n^2p$. ■

Combining 3.1.6 with 3.1.4 we have the following corollary.

Corollary 3.1.7: If X is an arbitrary matrix in N then $l_S(X) \leq 40n^2p$.

Construction of U

Fix $W \in U$ and suppose

$$W = \begin{pmatrix} 1 & w_{1,2} & w_{1,3} & \cdots & w_{1,n} \\ & 1 & w_{2,3} & \cdots & w_{2,n} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & w_{n-1,n} \\ & & & & 1 \end{pmatrix}.$$

Let $2 \leq k \leq n$ and define W_k to be the matrix

$$\begin{pmatrix} 1 & w_{1,2} & w_{1,3} & \cdots & w_{1,k} & & & \\ & 1 & w_{2,3} & \cdots & w_{2,k} & & & \\ & & \ddots & \ddots & \vdots & & & \\ & & & \ddots & w_{k-1,k} & & & \\ & & & & 1 & & & \\ & & & & & 1 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix}.$$

We can construct W_{k-1} from W_k as follows.

Let

$$C_{k-1} = \begin{pmatrix} 1 & & & & w_{1,k} & & & \\ & 1 & & & w_{2,k} & & & \\ & & \ddots & & & & & \\ & & & \ddots & w_{k-1,k} & & & \\ & & & & 1 & & & \\ & & & & & 1 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix}.$$

It is not hard to check that $C_{k-1}W_{k-1} = W_k$. Since $W = W_n$ and $W_2 = C_1$ we have

$$W = C_{n-1}W_{n-1} = C_{n-1}C_{n-2}W_{n-2} = ... = \left(\prod_{i=n-1}^{1} C_i\right)$$

and so, to construct W, it is enough to construct the matrices C_i.

Lemma 3.1.8: $l_S(C_k) \leq 10kp$

Proof.

In the proof of 3.1.2 we saw that the matrix

$$Y_k = \begin{pmatrix} 1 & & & 1 & & & \\ & 1 & & 1 & & & \\ & & \ddots & \vdots & & & \\ & & & 1 & & & \\ & & & & \ddots & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

has length of at most $4k$ in S.

Now pre multiply this matrix by the product

$$\left(y^{(k-1)} \prod_{m=1}^{k-1} y^{-1} x^{w_{m,k}-1} \right).$$

This gives the matrix

$$\begin{pmatrix}
1 & w_{1,k}-1 & & & & & & & w_{1,k} & & \\
 & 1 & w_{2,k}-1 & & & & & & w_{2,k} & & \\
 & & 1 & \ddots & & & & & \vdots & & \\
 & & & \ddots & \ddots & & & & & & \\
 & & & & 1 & w_{k-2,k}-1 & & & & & \\
 & & & & & 1 & w_{k-1,k} & & & & \\
 & & & & & & 1 & & & & \\
 & & & & & & & 1 & & & \\
 & & & & & & & & \ddots & & \\
 & & & & & & & & & 1 &
\end{pmatrix}$$

We need to get rid of the unwanted entries in the $(i, i+1)$ positions. To do this we post multiply by

$$\left(\prod_{m=k-1}^{1}(x^{1-w_{m,k}}y)\right)y^{-(k-1)}$$

to get the matrix C_k.

Now we have $l_S(C_k) \leq 8k + 2kp \leq 10kp$.

■

Corollary 3.1.9: $l_S(W) < 5np^2$

Proof. We have

$$l_S(W) = \sum_{i=2}^{n-1} l_S(C_i) \leq \sum_{i=2}^{n-1} 10kp < 5n^2p.$$

■

Now Theorem 1.1.2 can be deduced immediately from 3.1.7 and 3.1.9.

3.2 Proof of Theorem 1.1.3

To prove Theorem 1.1.3 we need 2.2.1, which was stated in Section 2.2. and proved by Babai, Lubotzky and Kantor in [1].

Lemma 3.2.1: Let $H = \langle E_{i,j}, t_{j,i} \rangle$.

Then $l_S(H) \leq 34mn \log(p)$ where m is the constant from 2.2.1.

Proof. H is clearly isomorphic to $SL_2(p)$. So, from 2.2.1 we see that H has length $m \log(p)$ with respect to the set $\{E_{i,j}, t_{j,i}\}$. From 3.1.2 and 3.1.3 we know that $l_S(E_{j,i})$ and $l_S(t_{j,i})$ are both bounded above by $34n$. Hence $l_S(H) \leq 34nm \log(p)$. ■

By using this bound and following exactly the same methods as before we may deduce Theorem 1.1.3 as follows. We begin by bounding the diameter of N.

Lemma 3.2.2: $l_S(N) \leq 38n^2 \log p$

Proof. By 3.2.1, $l_S(E_{i,j}^q) \leq 34nm \log p$ for any $q \in \mathbb{F}_p$.

Let P be a monomial matrix as described in the paragraph before 3.1.1. By 3.1.4, $l_S(P) \leq 34n^2$.

Using the notation of 3.1.5, each matrix of form M_a lies in the group generated by $\{E_{1,2}, t_{2,1}.\}$ Observe that $t_{2,1} = t_{1,2}^{-1} = (E_{1,2}E_{2,1}^{-1}E_{1,2})^{-1}$. Also $E_{2,1}^{-1} = y^{-n}E_{1,2}y^n$. So $l_S(E_{1,2}) < 2n+1$ and $t_{2,1}$ has length of at most $2n+3$, which is at most $3n$. This gives us that each matrix M_a has length at most $3mn \log p$.

Now, the matrix A can be written as the product

$$y^{-1}...y^{-1}y^{-1}M_{a_1}yM_{a_2a_1}yM_{a_3a_2a_1}...M_{a_na_{n-1}...a_1}y.$$

Hence $l_S(A) \leq 2n + 3n^2m \log p \leq 4n^2m \log p$. Since each $X \in N$ can be written $X = AP$, we have $l_S(X) \leq 4n^2m \log p + 34n^2 \leq 38n^2m \log p$.

■

Lemma 3.2.3: Let W be any matrix in U. Then $l_S(W) \leq 17.5n^3 m \log p$.

Proof. In this proof we will use the same notation as we did for the proof of 3.1.8.

Each $x^{w_{m,k}-1}$ has length of at most $34nm \log p$. As in the proof of 3.1.8, we obtain the matrix C_k by pre-multiplying the matrix

$$\begin{pmatrix} 1 & & & 1 & & & \\ & 1 & & 1 & & & \\ & & \ddots & \vdots & & & \\ & & & 1 & & & \\ & & & & \ddots & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

by

$$\left(y^{(k-1)} \prod_{m=k-1}^{1} y^{-1} x^{w_{m,k}-1} \right).$$

and then post multiplying by

$$\left(\prod_{m=1}^{k-1}(x^{1-w_{m,k}}y)\right)y^{-(k-1)}.$$

Since each $x^{w_{m,k}-1}$ has length of at most $34nm\log p$ and the matrix

$$\begin{pmatrix} 1 & & & 1 & & & \\ & 1 & & 1 & & & \\ & & \ddots & \vdots & & & \\ & & & 1 & & & \\ & & & & \ddots & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

has length $4k$, we have $l_S(C_k) \leq 4k + 2(2k + \sum_{k=1}^{n} 34mnk\log p) \leq 8k + 34n^2m\log p \leq 35n^2m\log p$.

Finally, since $W = \Pi_{k=n-1}^{1} C_k$ we have $l_S(W) \leq 17.5mn^3\log p$.

■

Theorem 1.1.3 now follows immediately from these last two lemmas.

4. USING COMPARISON TO DETERMINE A MIXING TIME OF THE SPECIAL LINEAR GROUP

4.1

In this chapter we prove Theorem 1.4.2. Recall that the random walk we are studying is generated by the probability distribution, Q, where Q is given by

$$Q(g) = \begin{cases} \frac{1}{5} & g = x, y, x^{-1}, y^{-1} \text{ or } I \\ 0 & \text{otherwise.} \end{cases}$$

In Chapter 2 we saw the mixing time of a group with respect to a given random walk can be determined using the eigenvalues of the matrix associated with the random walk. In particular, we saw how to bound the Variational Distance between Q^k and U, the uniform distribution on $SL_n(p)$ using 2.4.1.

To use 2.4.1 we needed bounds on the eigenvalues of the matrix corresponding to Q. These bounds could be obtained using comparison with another ran-

dom walk. In particular2.4.3 and 2.4.4 bounded the eigenvalues of the matrix corresponding to Q in terms of the eigenvalues of the matrix corresponding to another random walk.

In order to establish bounds on the eigenvalues of the matrix corresponding to Q we shall use comparison with the uniform random walk on the set of transvections in $SL_n(p)$.

We shall use these comparison theorems with some results determined about the character ratios on the set of transvections in $SL_n(p)$ to determine the eigenvalues of Q.

Let $\tilde{Q}$ be the uniform distribution on the set of transvections in $SL_n(p)$. Then $\tilde{Q}$ has corresponding matrix with eigenvalues $\tilde{\pi}_i$.

We first determine a suitable value of A using 2.4.4. We need to determine a bound for the values $|t|$ and $N(s,t)$. Now, for any transvection t, $|t| = l_S(t)$ and $N(s,t) \leq l_S(t)$, so finding a bound for $l_S(T)$ where T is the set of transvections in S will provide us with a bound for all values of $|t|$ and $N(s,t)$.

We will use the following result, which K. Magaard has proved in personal correspondence with Martin Liebeck, and Lemma 3.1.2 to find $l_S(T)$.

Lemma 4.1.1: Let t be a transvection in $SL_n(p)$. Then t is the product of at

most $4n-5$ elementary matrices.

We need to introduce some notation before beginning the proof.

Recall that a transvection, t, in $SL_n(p)$ is an element that fixes an $n-1$ dimensional subspace of $\mathbb{F}_p^n$ which we shall denote $C_V(t)$. This subspace is called the centre of t. There is a vector $v_t \in C_V(t)$ such that, for each v in $V - C_V(t)$, $vt = v + \alpha v_t$ for some $\alpha \in \mathbb{F}_p^*$. The subspace $\langle v_t \rangle$ is called the axis of t and we denote it by $[V,t]$.

For ease of notation I will use $E_{i,j}(q)$ to denote $E_{i,j}^q$. We define $U_{i,j}$ to be the group $= \{E_{i,j}(q) | q \in \mathbb{F}_p\}$. The root groups of G are the conjugates of $U_{1,2}$.

In the proof of 4.1.1 we will need to consider the action of $SL_n(p)$ on the dual space, V^*, of V. Let $e_1, e_2, ...e_n$ be a standard basis of $SL_n(p)$ and let $f_1, f_2, ...f_n$ be the corresponding dual basis of V^*. We will say $SL_n(p)$ acts on V from the right and acts on V^* from the left as follows.

For any $f \in V^*$ and $g \in SL_n(p)$ we have $(g(f))(v) = f(vg^{-1})$. Note that each transvection $t \in SL_n(p)$ is also a transvection when it acts n V^*.

We denote the centre of t in V^* by $C_{V^*}(t)$ and the axis by $[V^*, t]$.

We will use the following three remarks in the proof of 4.1.1.

Remark 4.1.2: If t is a transvection in $SL_n(p)$, then the root group of t is

the set of transvections $\{t^q | q \in \mathbb{F}_p\}$. If two transvections lie in the same root group then they have the same axis and the same center. Conversely if two transvections have equal center and axis, then they must lie in the same root group.

Remark 4.1.3: $C_V(U_{1,2}) = \langle e_2, \ldots, e_n \rangle$, $[V, U_{1,2}] = \langle e_2 \rangle$.
$C_{V^*}(U_{1,2}) = \langle f_1, f_3, f_4 \ldots, f_n \rangle$, $[V^*, U_{1,2}] = \langle f_1 \rangle$. Note that $C_V(U_{1,2}) = Ker(f_1)$.

Remark 4.1.4: Let R be a root group and $t \in R$. If $s \in \mathrm{GL}(V)$, then $C_V(s^{-1}ts) = C_V(t)s$, $[V, s^{-1}Rs] = [V, R]s$, and $C_{V^*}(s^{-1}ts) = s^{-1}C_{V^*}(t)$, $[V^*, s^{-1}Rs] = s^{-1}[V^*, R]$.

Proof. (of 4.1.1) Let τ be a transvection and let $v^* \neq 0$ be a vector in $[V^*, \tau]$. With respect to the standard basis of V^*, $v^* = \sum_{i=1}^n \alpha_i f_i$. As $v^* \neq 0$ we have some j for which $\alpha_j \neq 0$. Now let $\sigma = \Pi_{i \neq j} E_{i,j}(\frac{-\alpha_i}{\alpha_j})$. Then $\sigma v^* = \alpha_j f_j$ and thus by Remark 4.1.4 we see that $[V^*, \sigma\tau\sigma^{-1}] = \langle f_j \rangle$.

As in Remark 4.1.3 we have $C_V(\sigma\tau\sigma^{-1}) = \langle e_1, \ldots, e_{j-1}, e_{j+1}, \ldots e_n \rangle$ and $[V, \sigma\tau\sigma^{-1}] \subset C_V(\sigma\tau\sigma^{-1})$.

Any element $\rho \in \mathrm{GL}(V)$ which centralizes f_j stabilizes $C_V(\sigma\tau\sigma^{-1})$. Let $v \neq 0$ be a vector in $[V, \sigma\tau\sigma^{-1}]$. Now $v = \sum_{k=1, k \neq j}^n \beta_k e_k$. For some i we must have

$\beta_i \neq 0$. Hence, if we let $\rho = \Pi_{k \neq i,j} E_{i,k}(\frac{-\beta_k}{\beta_i})$, we have $\rho f_j = f_j$ and again by Remark 4.1.4, $[V, \rho^{-1}\sigma\tau\sigma^{-1}\rho] = [V, \sigma\tau\sigma^{-1}]\rho = \langle e_i \rangle$.

We know that $\rho^{-1}\sigma\tau\sigma^{-1}\rho$ is a transvection with axis e_i and center $\langle e_1, \ldots, e_{j-1}, e_{j+1}, \ldots e_n \rangle$. Thus $\rho^{-1}\sigma\tau\sigma^{-1}\rho$ is in $U_{j,i}$ and so $\rho^{-1}\sigma\tau\sigma^{-1}\rho = E_{j,i}(\alpha)$ for some $\alpha \in \mathbb{F}$. As ρ and σ are products of elementary matrices we have now expressed τ as the following product of elementary matrices $\tau = \sigma^{-1}\rho E_{j,i}(\alpha)\rho^{-1}\sigma$. The number of factors in the product is at most $2(n-1) + 2(n-2) + 1 = 4n - 5$, as claimed. ■

Proposition 4.1.5: Let T be the set of transvections in $SL_n(p)$. Then $l_S(T) \leq 136mn^2 \log p$.

Proof. Each matrix $E_{i,j}(c)$ lies in the group generated by $S' = \langle E_{i,j}, t_{i,j} \rangle$ and from 2.2.1 we see that $l_{S'}(E_{i,j}(c)) \leq m \log p$ where m is the constant in 2.2.1. Also,from 3.1.2, we know that each matrix of the form $E_{i,j}$ and $t_{i,j}$ can be written as a product of at most $34n$ elements of S. Hence $l_S(E_{i,j}(c)) \leq 34mn \log p$.

Now, applying 4.1.1, we have $l_S(T) \leq (4n-5)34mn \log p \leq 136mn^2 \log p$. ■

Substituting this value into 2.4.4 we get

$$\begin{aligned} A = & \max_{s\in S} \tfrac{1}{Q(s)} \sum_{g\in G} |g| N(s,g)\tilde{Q}(g) \\ \leq & \quad 5\sum_{t\in T}(136mn^2\log p)^2\tilde{Q}(t). \\ = & \quad 5\sum_{t\in T}(136m)^2 n^4 (logp)^2\tilde{Q}(t) \\ = & \quad Kn^4\log^2 p \end{aligned}$$

where K is a constant.

Next we bound $\tilde{\pi}_1$. Since the support of $\tilde{Q}$ is a single conjugacy class, from 2.5.1 we see that the eigenvalues of $\tilde{M}$ are the character ratios $\frac{\chi_\rho(t)}{\chi_\rho(1)}$ where ρ is an irreducible representation of G and t is a transvection.

In 2.5.2 these ratios are bounded above by $\frac{19}{20}$.

Now we have $\frac{|\chi(x)|}{|\chi(1)|} \leq \frac{19}{20}$ and so $\tilde{\pi}_1 \leq \frac{19}{20}$.

Now

$$\pi_1 \leq 1 - \frac{1-\tilde{\pi}_1}{A} \leq 1 - \frac{1-\frac{19}{20}}{A} = 1 - \frac{1}{20A}$$

We can write $\frac{1}{20A} \geq \frac{1}{K'n^4(\log p)^2}$ where K' is a constant So we have $\pi_1 \leq 1 - \frac{1}{K'n^4(\log p)^2}$.

Now we can use 2.4.2 to bound the eigenvalue $\pi_{|G|-1}$.

$$-\pi_{|G|-1} \geq -1 + 2Q(id) = -1 + \frac{2}{5}$$

Hence $|\pi_{|G|-1}|$ is bounded above by $\frac{3}{5}$ and the largest eigenvalue π_1, is bounded above by $1 - \frac{1}{Kn^4p(\log p)^2}$.

We can now prove Theorem 1.4.2.

Proof. We have

$$4||Q^k - U||^2 \leq |G|\pi_1^{2k}$$

$$\leq p^{n^2}\left(1 - \frac{1}{K'n^4(\log p)^2}\right)^{2k}$$

$$\leq \exp\left(n^2 \log p - \frac{2k}{K'n^4(\log p)^2}\right)$$

$$= \exp\left(\frac{K'n^6(\log p)^3 - 2k}{K'n^4(\log p)^2}\right).$$

This shows that the mixing time of $SL_n(p)$ is of order $n^6(\log p)^3$.

■

5. THE DIAMETER OF THE SYMPLECTIC GROUP OVER FINITE FIELDS WITH PRIME ORDER

5.1 Definitions and Notation

In this chapter we will prove 1.1.3 and 1.1.5. Let G denote the Symplectic group of $2n \times 2n$ matrices over $\mathbb{F}_p$ where p is an odd prime. We order the symplectic basis of $\mathbb{F}_p^{2n}$ as $e_1, e_2, ...e_n, f_n, f_{n-1}, ...f_1$. Any matrix in G may be written

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

where A, B, C and D are $n \times n$ matrices.

Recall that, in Chapter 1, we defined a generating set of $Sp_{2n}(p)$ as follows.

Using the notation of Chapter 3, define v to be the matrix

$$\begin{pmatrix} x & 0 \\ 0 & x^{-TJ} \end{pmatrix}$$

where x is the element of our generating set of $SL_n(p)$ described in Chapter 1 and at the beginning of Chapter 3. Recall that x^{-TJ} means the inverse transpose of x conjugated by the matrix J, which is described just before the statement of 2.1.4. By 2.1.4 v lies in $Sp_{2n}(p)$.

Define w to be the product

$$\begin{pmatrix} & & & & -1 \\ & 1 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & 1 \\ 1 & & & & \end{pmatrix} \begin{pmatrix} \begin{pmatrix} & & 1 \\ 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} & 0 \\ 0 & \begin{pmatrix} & & 1 \\ 1 & & \\ & \ddots & \\ & 1 & \end{pmatrix}^{-TJ} \end{pmatrix}.$$

Then v is a short root element of G and w is a a member of the Weyl group of G. We shall show the set $S' = \{v^{\pm 1}, w^{\pm 1}\}$ generates G and determine a bound for $l_{S'}(G)$.

We may use similar methods to those used in Chapter 3 to express matrices as products of our generators. Since the generators we use to generate $Sp_{2n}(p)$ are similar to those we used to generate $SL_n(p)$ we may also use some results from Chapter 3 .

In particular, in early sections of the proof,we will often be dealing with matrices of form

$$\begin{pmatrix} A & 0 \\ 0 & A^{-TJ} \end{pmatrix}$$

where A^{-TJ} denotes the inverse transpose of A conjugated by the matrix J. By 2.1.4, these matrices lie in $Sp_{2n}(p)$.

Observe that left multiplying the matrix above by v would give us

$$\begin{pmatrix} xA & 0 \\ 0 & (xA)^{-TJ} \end{pmatrix}.$$

The identities from chapter 3 can now enable us to work out how matrices may be written as the product of elements of S'.

We deal with multiplication by w in a similar way to how we dealt with y. In particular, post-multiplication by w can be thought of as applying the permutation $(1, n+1, n+2, ..., 2n, n-1, n-2, ..., 2)$. to the columns of the matrix, followed by reversing the sign of all entries in the $n+1$th column.Pre-multiplication by w can be thought of as applying the permutation $(1, 2, ...n-$

$1, 2n, 2n-1, ...n)$ to the rows followed by reversing the sign of all entries in the 1st row.

From 2.1.2 we know that we may write $G = BNB$ where N is the subgroup of monomials in G and B is the group of upper triangular matrices. Since each matrix in B is the product of a diagonal matrix with an upper uni-triangular matrix we may write $G = UNU$ and so it is sufficient to determine the diameters of U and N.

In order to construct either of these subgroups we need to construct the root groups of G as products of our generators in S'. We begin by constructing the short root elements.

5.2 Construction of Short Root Elements

Define $F_{i,j}$ to be the matrix

$$\begin{pmatrix} E_{i,j} & 0 \\ 0 & E_{i,j}^{-TJ} \end{pmatrix}$$

where $E_{i,j}$ is the same as described in Chapter 3. The matrices $F_{i,j}$ are short root elements of G.

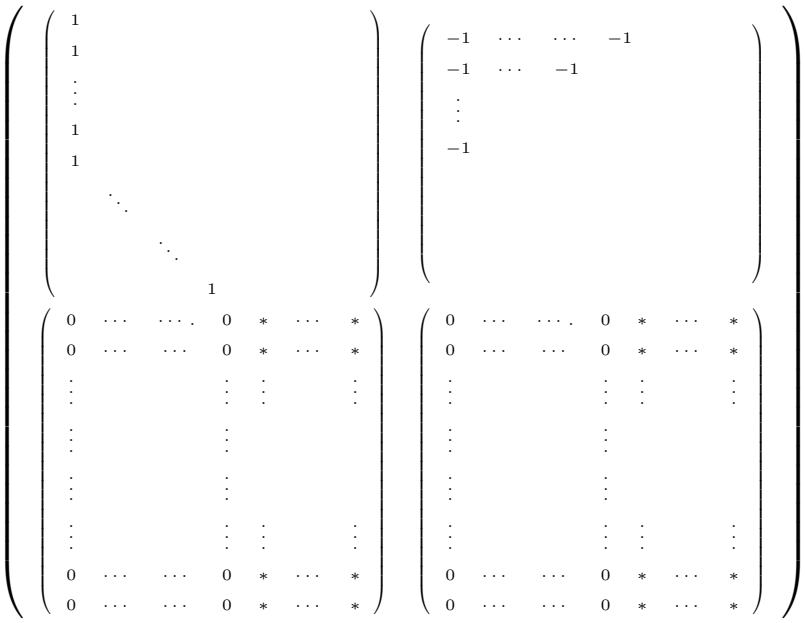
Lemma 5.2.1: Fix $k \leq n$. Then $l_{S'}(F_{1,k}) < 10k$.

Proof. Let Z'_k be the $2n \times 2n$ matrix of form

$$\begin{pmatrix} Z_k & 0 \\ 0 & Z_k^{-TJ} \end{pmatrix}$$

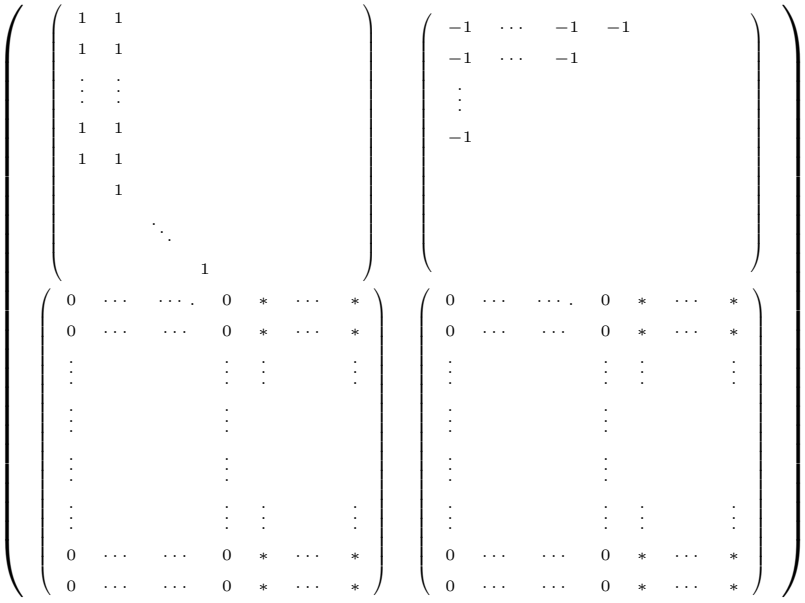
where Z_k is as described in Lemma 3.1.2.

Note that $Z'_2 = v$. We construct Z'_{k+1} from Z'_k as follows. Post multiplying by w performs the permutation $(1, n+1, n+2..., 2n, n-1, n-2, ...2)$ on the columns of Z'_k and multiplies the $(n-1)$th and $(n+1)$th column by -1.So $Z'_k w^{(k-1)}$ looks like this.



Here we keep track of possible non-zero matrix entries with stars.

Now, multiplying by v adds the first column to the second and subtracts the $2n-1$th column from the $2n$th. Hence $Z'_k w^{(k-1)} v$ looks like this.



Now post-multiplying by w^{-1} performs the permutation $(1, 2, ...n, 2n, 2n - 1, ...n{+}1)$ to the columns and multiplies the first column by -1 so $Z'_k w^{(k-1)} v w^{-(k-1)}$ looks like this.

$$\begin{pmatrix} \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 & & & & \\ & 1 & \cdots & 1 & 1 & & & & \\ & & \ddots & \vdots & \vdots & & & & \\ & & & 1 & 1 & & & & \\ & & & & 1 & & & & \\ & & & & & \ddots & & & \\ & & & & & & 1 & & \\ & & & & & & & 1 & \end{pmatrix} & & 0 \\ & & \\ 0 & & \begin{pmatrix} * & * & \cdots & & \cdots & * & * \\ * & * & \cdots & & \cdots & * & * \\ \vdots & \vdots & & & & \vdots & \vdots \\ & & & & & & \\ \vdots & \vdots & & & & \vdots & \vdots \\ * & * & \cdots & & \cdots & * & * \\ * & * & \cdots & & \cdots & * & * \end{pmatrix} \end{pmatrix}$$

Since we are certain that the bottom left and top right quadrants are zero, 2.1.4 tells us that the bottom right quadrant is Z_{k+1}^{-TJ} and so we have ended up with the matrix Z'_{k+1}. So for any $i \in \{2, 3, ...n-1\}$ we have $Z'_{i+1} = Z'_i w^{(i-1)} v w^{-(i-1)}$. Hence for each $2 \leq k \leq n$, we have

$$Z'_k = Z'_{k-1} w^{(k-2)} v w^{-(k-2)}.$$

Substituting $Z'_{k-1} = Z'_{k-2} w^{(k-3)} v w^{-(k-3)}$ we get,

$$\begin{aligned} Z'_k &= Z'_{k-2} w^{(k-3)} v w^{-(k-3)} w^{(k-2)} v w^{-(k-2)} \\ &= Z'_{k-2} w^{(k-3)} v w v w^{-(k-2)}. \end{aligned}$$

Now writing Z'_{k-2} in terms of Z'_{k-3} we have

$$Z'_k = Z'_{k-3}w^{(k-4)}vw^{-(k-4)}w^{(k-3)}vwvw^{-(k-2)}$$

$$= Z'_{k-3}w^{(k-4)}vwvwv^{-(k-2)}.$$

Continuing like this we eventually have

$$Z'_k = Z'_2w(vw)^{k-3}vw^{-(k-2)}$$

and since $Z'_2 = v$,

$$Z'_k = (vw)^{k-2}vw^{-(k-2)}.$$

From Z'_k we construct the matrix Y'_k which has form

$$\begin{pmatrix} Y_k & 0 \\ 0 & Y_k^{-TJ} \end{pmatrix}$$

where Y_k is as described in 3.1.2.

By post multiplying the matrix Z'_k by $w^{(k-3)}$, we repeatedly apply the permutation $(1, n+1, n+2, ..2n, n-1, n-2..., 2)$ and multiply the $n+1$th

column by -1.

Hence we get a matrix of form

$$\begin{pmatrix} \begin{pmatrix} 1 & 1 & 1 & & & & \\ 1 & 1 & 1 & & & & \\ \vdots & \vdots & \vdots & & & & \\ 1 & 1 & 1 & & & & \\ 1 & 1 & 1 & & & & \\ 1 & 1 & 1 & & & & \\ & 1 & 1 & & & & \\ & & 1 & & & & \end{pmatrix} & \begin{pmatrix} -1 & \cdots & \cdots & \cdots & -1 & & & \\ -1 & \cdots & \cdots & -1 & & & & \\ \vdots & & & & & & & \\ -1 & -1 & & & & & & \\ -1 & & & & & & & \end{pmatrix} \\ \begin{pmatrix} 0 & 0 & 0 & * & \cdots & \cdots & * \\ 0 & 0 & 0 & * & \cdots & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & & & \vdots \\ \vdots & \vdots & \vdots & \vdots & & & \vdots \\ \vdots & \vdots & \vdots & \vdots & & & \vdots \\ \vdots & \vdots & \vdots & \vdots & & & \vdots \\ 0 & 0 & 0 & * & \cdots & \cdots & * \\ 0 & 0 & 0 & * & \cdots & \cdots & * \end{pmatrix} & \begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 & * & \cdots & \cdots & * \\ 0 & \cdots & \cdots & \cdots & 0 & * & \cdots & \cdots & * \\ \vdots & & & & \vdots & \vdots & & & \vdots \\ \vdots & & & & \vdots & \vdots & & & \vdots \\ \vdots & & & & \vdots & \vdots & & & \vdots \\ \vdots & & & & \vdots & \vdots & & & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & * & \cdots & \cdots & * \\ 0 & \cdots & \cdots & \cdots & 0 & * & \cdots & \cdots & * \end{pmatrix} \end{pmatrix}$$

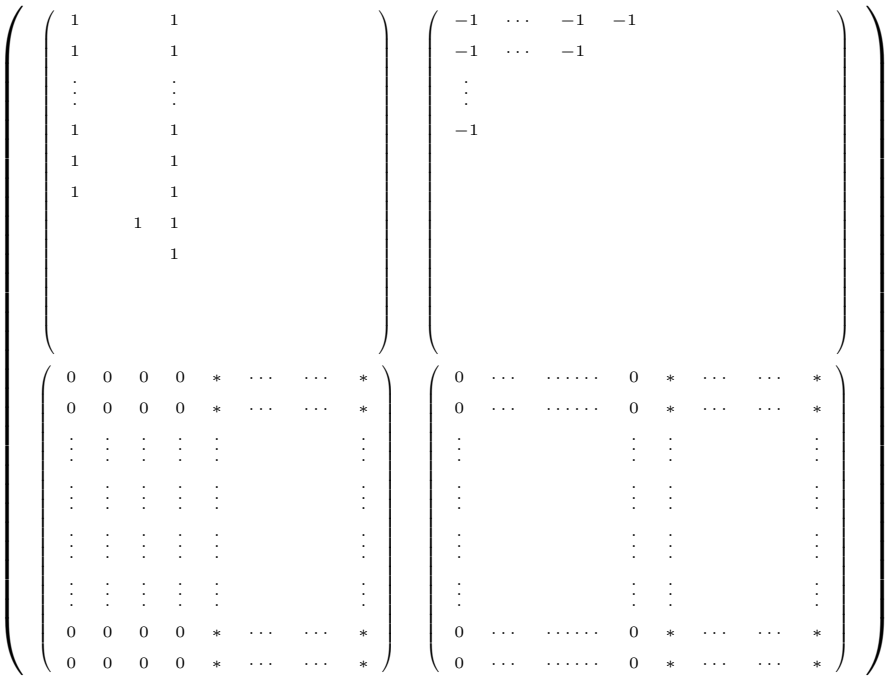
Now post multiplying by v^{-1} subtracts the first column from the second so the $Z_k w^{(k-3)} v^{-1}$ has form

$$\begin{pmatrix} \begin{pmatrix} 1 & & 1 & & & & \\ 1 & & 1 & & & & \\ \vdots & & \vdots & & & & \\ 1 & & 1 & & & & \\ 1 & & 1 & & & & \\ 1 & & 1 & & & & \\ & 1 & 1 & & & & \\ & & 1 & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \end{pmatrix} & \begin{pmatrix} -1 & \cdots & \cdots & -1 & -1 & & & \\ -1 & \cdots & \cdots & -1 & & & & \\ \vdots & & & & & & & \\ -1 & -1 & & & & & & \\ -1 & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \end{pmatrix} \\ \begin{pmatrix} 0 & 0 & 0 & * & \cdots & & \cdots & * \\ 0 & 0 & 0 & * & \cdots & & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & & & & \vdots \\ \vdots & \vdots & \vdots & \vdots & & & & \vdots \\ \vdots & \vdots & \vdots & \vdots & & & & \vdots \\ \vdots & \vdots & \vdots & \vdots & & & & \vdots \\ 0 & 0 & 0 & * & \cdots & & \cdots & * \\ 0 & 0 & 0 & * & \cdots & & \cdots & * \end{pmatrix} & \begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 & * & \cdots & \cdots & * \\ 0 & \cdots & \cdots & \cdots & 0 & * & \cdots & \cdots & * \\ \vdots & & & & \vdots & \vdots & & & \vdots \\ \vdots & & & & \vdots & \vdots & & & \vdots \\ \vdots & & & & \vdots & \vdots & & & \vdots \\ \vdots & & & & \vdots & \vdots & & & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & * & \cdots & \cdots & * \\ 0 & \cdots & \cdots & \cdots & 0 & * & \cdots & \cdots & * \end{pmatrix} \end{pmatrix}$$

Post multiplying by w^{-1} gives us

$$\begin{pmatrix} \begin{pmatrix} 1 & 1 & & 1 & & & & \\ 1 & 1 & & 1 & & & & \\ \vdots & \vdots & & \vdots & & & & \\ 1 & 1 & & 1 & & & & \\ 1 & 1 & & 1 & & & & \\ & 1 & & 1 & & & & \\ & & 1 & 1 & & & & \\ & & & 1 & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \end{pmatrix} & \begin{pmatrix} -1 & \cdots & -1 & -1 & & & & \\ -1 & \cdots & -1 & & & & & \\ \vdots & & & & & & & \\ -1 & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \end{pmatrix} \\ \begin{pmatrix} 0 & 0 & 0 & 0 & * & \cdots & \cdots & * \\ 0 & 0 & 0 & 0 & * & \cdots & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & & & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & & & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & & & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & & & \vdots \\ 0 & 0 & 0 & 0 & * & \cdots & \cdots & * \\ 0 & 0 & 0 & 0 & * & \cdots & \cdots & * \end{pmatrix} & \begin{pmatrix} 0 & \cdots & \cdots & 0 & * & \cdots & \cdots & * \\ 0 & \cdots & \cdots & 0 & * & \cdots & \cdots & * \\ \vdots & & & \vdots & \vdots & & & \vdots \\ \vdots & & & \vdots & \vdots & & & \vdots \\ \vdots & & & \vdots & \vdots & & & \vdots \\ \vdots & & & \vdots & \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 & * & \cdots & \cdots & * \\ 0 & \cdots & \cdots & 0 & * & \cdots & \cdots & * \end{pmatrix} \end{pmatrix}.$$

Now post multiplication by v^{-1} gives us



Continuing to repeatedly post multiply by w^{-1} and v^{-1} we eventually have

$$Y'_k = Z'_k w^{(k-3)}(v^{-1}w^{-1})^{(k-3)}v^{-1}$$

$$= (vw)^{k-2}vw^{-(k-2)}w^{(k-3)}(v^{-1}w^{-1})^{(k-3)}v^{-1}$$

$$= (vw)^{k-2}vw^{-1}(v^{-1}w^{-1})^{(k-3)}v^{-1}.$$

Hence Y'_k has length of at most $4k$.

Pre-multiplying Y'_k by w^{-1} performs the permutation $(1, n{+}1, n{+}2, ...2n, n, n{-}1, ...2)$ on the rows of Y'_k and multiplies $n+1$th row by -1. This gives us the

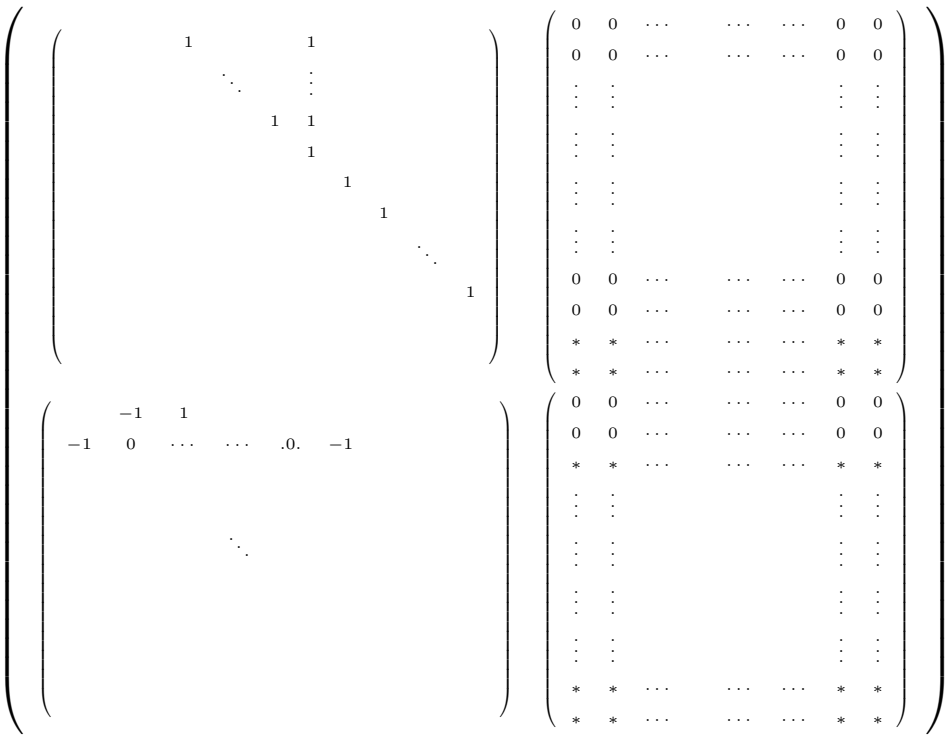
following matrix.

$$\begin{pmatrix} \begin{pmatrix} 1 & & & & & & & & & \\ & 1 & & & & & & 1 & & \\ & & \ddots & & & & & \vdots & & \\ & & & 1 & & & & 1 & & \\ & & & & 1 & & & 1 & & \\ & & & & & & & 1 & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & 1 \end{pmatrix} & \begin{pmatrix} 0 & 0 & \cdots & \cdots & \cdots & 0 & 0 \\ 0 & 0 & \cdots & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & & & & \vdots & \vdots \\ \vdots & \vdots & & & & \vdots & \vdots \\ \vdots & \vdots & & & & \vdots & \vdots \\ \vdots & \vdots & & & & \vdots & \vdots \\ \vdots & \vdots & & & & \vdots & \vdots \\ 0 & 0 & \cdots & \cdots & \cdots & 0 & 0 \\ 0 & 0 & \cdots & \cdots & \cdots & 0 & 0 \\ * & * & \cdots & \cdots & \cdots & * & * \end{pmatrix} \\ \begin{pmatrix} -1 & 0 & \cdots\cdots & 0 & -1 & & \\ & & & & & & \\ & & & & & & \\ & & & \ddots & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \end{pmatrix} & \begin{pmatrix} 0 & 0 & \cdots & \cdots & \cdots & 0 & 0 \\ * & * & \cdots & \cdots & \cdots & * & * \\ \vdots & \vdots & & & & \vdots & \vdots \\ \vdots & \vdots & & & & \vdots & \vdots \\ \vdots & \vdots & & & & \vdots & \vdots \\ \vdots & \vdots & & & & \vdots & \vdots \\ \vdots & \vdots & & & & \vdots & \vdots \\ * & * & \cdots & \cdots & \cdots & * & * \\ * & * & \cdots & \cdots & \cdots & * & * \\ * & * & \cdots & \cdots & \cdots & * & * \end{pmatrix} \end{pmatrix}$$

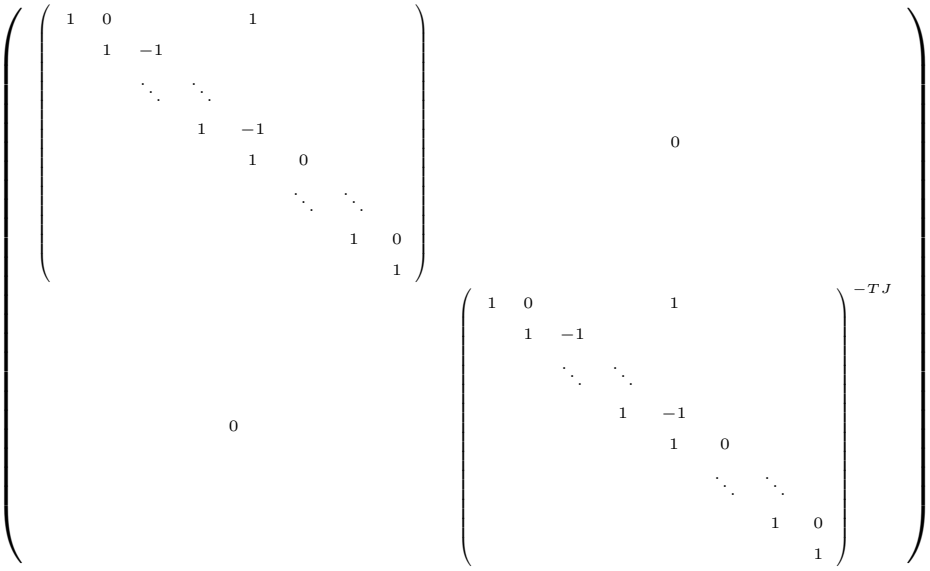
Now pre-multiplying by v^{-1} subtracts the second row from the first and adds the $2n-1$th row to the $2n$th. So we have

$$\begin{pmatrix} \begin{pmatrix} 1 & -1 & & & & 0 & & \\ & 1 & & & & 1 & & \\ & & \ddots & & & \vdots & & \\ & & & 1 & & 1 & & \\ & & & & 1 & 1 & & \\ & & & & & 1 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix} & \begin{pmatrix} 0 & 0 & \cdots & \cdots & \cdots & 0 & 0 \\ 0 & 0 & \cdots & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & & & & \vdots & \vdots \\ \vdots & \vdots & & & & \vdots & \vdots \\ \vdots & \vdots & & & & \vdots & \vdots \\ \vdots & \vdots & & & & \vdots & \vdots \\ 0 & 0 & \cdots & \cdots & \cdots & 0 & 0 \\ 0 & 0 & \cdots & \cdots & \cdots & 0 & 0 \\ * & * & \cdots & \cdots & \cdots & * & * \end{pmatrix} \\ \begin{pmatrix} -1 & 0 & \cdots & \cdots & 0 & -1 & & \\ & & & & & & & \\ & & & & & & & \\ & & & \ddots & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \\ & & & & & & & \end{pmatrix} & \begin{pmatrix} 0 & 0 & \cdots & \cdots & \cdots & 0 & 0 \\ * & * & \cdots & \cdots & \cdots & * & * \\ \vdots & \vdots & & & & \vdots & \vdots \\ \vdots & \vdots & & & & \vdots & \vdots \\ \vdots & \vdots & & & & \vdots & \vdots \\ \vdots & \vdots & & & & \vdots & \vdots \\ * & * & \cdots & \cdots & \cdots & * & * \\ * & * & \cdots & \cdots & \cdots & * & * \\ * & * & \cdots & \cdots & \cdots & * & * \end{pmatrix} \end{pmatrix}$$

Pre-multiplying by w^{-1} again we get



By continuing to pre-multiply by $v^{-1}w^{-1}$ and then post multiplying by w^{k-3} we end up with the matrix



We need to remove the -1 entries from the top-left matrix. To do this we first post multiply by w to move our first unwanted entry into the second column. Then we post multiply by v. This adds the first column to the second column, which eliminates the -1 entry. By multiplying by wv repeatedly, we end up eliminating the unwanted -1s. Now multiplying by $w^{-(k-2)}$ we have the matrix $F_{1,k}$.

So our matrix $F_{1,k}$ may be written as $w^{k-2}(v^{-1}w^{-1})^{k-2}Y'_k(wv)^{k-3}w^{k-3}$. So $l_{S'}(F_{1,k}) \leq l_{S'}(Y'_k) + 4k < 10k$. ■

Corollary 5.2.2: Suppose $i \neq j$ and $i, j \leq n$. Then $l_{S'}(F_{i,j}) < 12n$.

Proof. First suppose $i < j$. We have $wF_{a,b}w^{-1} = F_{a+1,b+1}$. Hence $F_{i,j}$ can be obtained by conjugating a matrix of form $F_{1,k}$ by w up to $n-k$ times and so $l_{S'}(F_{i,j}) < 10k + 2(n-k) = 8k + 2n \leq 10n$. Finally observe that $F_{j,i} = w^n F_{i,j} w^{-n}$. ■

5.3 Construction of Long Root Elements

Recall that the matrix $E_{i,j}$ is the matrix $I + e_{i,j}$. The matrices $E_{i,2n-i+1}$ for $1 \leq i \leq 2n$ are long root elements in G.

Lemma 5.3.1: $l_{S'}((E_{1,2n})^2) < 10n$

Proof. We have

$$(vw)^{n-1}v(w^{-1}v^{-1})^{n-1} = I + \sum_{i=n+1}^{2n-1}(e_{1,i}) + \sum_{i=2}^{n}(e_{i,2n}) + 2e_{1,2n} =$$

$$E_{1,2n}^{-2} + \sum_{i=n+1}^{2n-1}(e_{1,i}) + \sum_{i=2}^{n}(e_{i,2n}).$$

This matrix has the following form.

$$\begin{pmatrix} \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} & \begin{pmatrix} 1 & \cdots & 1 & 2 \\ & & & 1 \\ & & & \vdots \\ & & & 1 \end{pmatrix} \\ 0 & \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} \end{pmatrix}$$

We wish to remove the entries of 1 in the top right quadrant. Pre-multiplying by $w^{(n-1)}(v^{-1}w^{-1})^{n-1}$ leaves us with the matrix

$$\begin{pmatrix} \begin{pmatrix} 1 & & & & & \\ & 1 & 1 & & & \\ & & \ddots & \ddots & & \\ & & & 1 & 1 & \\ & & & & 1 & \\ & & & & & 1 \end{pmatrix} & \begin{pmatrix} 0 & \cdots & 0 & 2 \\ \vdots & & & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & & 0 \end{pmatrix} \\ 0 & \begin{pmatrix} 1 & & & & & \\ & 1 & 1 & & & \\ & & \ddots & \ddots & & \\ & & & 1 & 1 & \\ & & & & 1 & \\ & & & & & 1 \end{pmatrix}^{-TJ} \end{pmatrix}$$

Now we need to remove the entries above the diagonal. We achieve this by post multiplying by $w(vw)^{n-3}vw^{-(n-2)}$.

So we have $(E_{1,2n})^2 = w^{(n-1)}(v^{-1}w^{-1})^{n-1}(vw)^{n-1}v(w^{-1}v^{-1})^{n-1}w(vw)^{n-3}vw^{-(n-2)}$. Hence $l_{S'}((E_{1,2n}^2)) \leq 10n$. ■

Lemma 5.3.2: Suppose $q \in \mathbb{F}_p$. Then $l_S(E^q_{1,2n}) < 10pn$. For $1 < i \leq n$ we have $l_S(E^q_{i,2n-i+1}) < 12pn$.

Proof. Suppose $1 < i \leq n$. We have $w^{(i-1)}(E^2_{1,2n})w^{-(i-1)} = E^{-2}_{i,2n-i+1}$. Hence $l_S(E^{-2}_{i,2n-i+1}) < 12n$.

Now observe that, since p is odd, for $1 \leq i \leq n$, $E^{-2}_{i,2n}$ generates the group $\{E^q_{i,2n} | q \in \mathbb{F}^*_p\}$. ■

5.4 Construction of N

Define r_i to be the matrix in G sending e_i to f_i and $-f_i$ to e_i. So r_i has entries of 1 in positions (k,k) for $k \neq i, 2n-i+1$, an entry of 1 in position $(i, 2n-i+1)$ and an entry of -1 in position $(2n-i+1, i)$. Let $R = \langle r_1, r_2, ..., r_n \rangle$.

Define H to be the group

$$\left\{ \begin{pmatrix} M & 0 \\ 0 & M^{-TJ} \end{pmatrix} \mid \text{M is a monomial in } GL_n(p) \right\}.$$

An arbitrary matrix in N may be written as the product of some $h \in H$ and at most n of the r_i. Also, each element of H is the product of a diagonal matrix with a matrix of form

$$P' = \begin{pmatrix} P & 0 \\ 0 & (P^{-TJ}) \end{pmatrix}$$

where P is a permutation matrix in $GL_n(p)$. Hence, to construct N we need to work out the maximum possible lengths of each of the r_i, each diagonal matrix in H and each matrix of form P'.

We shall now construct each r_i from the long roots of G. Note that pre-multiplying a matrix by $E_{i,j}$ adds the ith column to the jth column. Post multiplying a matrix by $E_{i,j}$ adds the jth row to the ith row.

Lemma 5.4.1: $l_{S'}(r_i) < 32pn$

Proof. From the discussion above we see

$$E_{1,2n}E_{2n,1}^{-1} = E_{1,2n}\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ -1 & & & 1 \end{pmatrix} = \begin{pmatrix} 0 & & & 1 \\ & 1 & & \\ & & \ddots & \\ -1 & & & 1 \end{pmatrix}.$$

Post multiplying this matrix by $E_{1,2n}^{-1}$ subtracts the 1st row from the $2n$th row and so $r_1 = E_{1,2n}E_{2n,1}^{-1}E_{1,2n}^{-1}$.

By 5.3.2, $l_{S'}(E_{1,2n}) \leq 10np$. Also we have $E_{2n,1} = w^n E_{1,2n} w^{-n}$ and so

$$l_{S'}(E_{2n,1}) < 2n + 10np \leq 11np.$$

Hence the length of r_1 is less than $31pn$. Finally observe that $wr_1w^{-1} = r_n$ and if $2 < j \leq n$, then $wr_jw^{-1} = r_{j-1}$. ■

Now we construct the diagonal matrices in H. Let $a \in \mathbb{F}_p^*$. We'll first construct diagonal matrices of form D_a where D_a has entry a in rows 1 and $2n-1$, entry a^{-1} in rows 2 and $2n$ and ones along the rest of the diagonal. That is

$$D_a = \begin{pmatrix} a & & & & & & \\ & a^{-1} & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & & & 1 & & \\ & & & & & a & \\ & & & & & & a^{-1} \end{pmatrix}.$$

Define $s_{i,j}$ to be the matrix that sends e_j to e_i, e_i to $-e_j$, f_i to $-f_j$, and f_j to f_i. So, in the notation of the previous section,

$$s_{i,j} = \begin{pmatrix} t_{i,j} & 0 \\ 0 & t_{i,j}^{-TJ} \end{pmatrix}.$$

Lemma 5.4.2: $l_{S'}(s_{i,j}) < 32n$.

Proof. Observe that

$$F_{i,j}F_{j,i}^{-1}F_{i,j} =$$

$$\begin{pmatrix} E_{i,j} & 0 \\ 0 & E_{i,j}^{-TJ} \end{pmatrix} \begin{pmatrix} E_{j,i}^{-1} & 0 \\ 0 & E_{j,i}^{TJ} \end{pmatrix} \begin{pmatrix} E_{i,j} & 0 \\ 0 & E_{i,j}^{-TJ} \end{pmatrix} =$$

$$\begin{pmatrix} E_{i,j}E_{j,i}^{-1}E_{i,j} & 0 \\ 0 & (E_{i,j}E_{j,i}^{-1}E_{i,j})^{-TJ} \end{pmatrix}.$$

Looking back to the proof of 3.2.5 we see that $E_{i,j}E_{j,i}^{-1}E_{i,j} = t_{i,j}$ and so the above matrix is precisely $s_{i,j}$. Now apply 5.2.2. ■

Let $a \in \mathbb{F}_p^*$ and define q_a to be the matrix in G that sends e_1 to ae_1 and f_1 to $a^{-1}f_1$. So

$$q_a = \begin{pmatrix} a & & & & & \\ & 1 & & & & \\ & & \ddots & & & \\ & & & 1 & & \\ & & & & a^{-1} & \end{pmatrix}.$$

Lemma 5.4.3: $l_{S'}(q_a) \leq 64np$

Proof. Since $E_{2n,1}^{-2} = w^n E_{1,2n}^{-2} w^{-n}$, by applying 5.3.2 we have $l_{S'}(E_{2n,1}^{-2}) < 12n$. Hence $l_{S'}(E_{2n,1})^q < 12np$ for any $q \in \mathbb{F}_p$.

Let $a \in \mathbb{F}_p^*$ As discussed earlier, post-multiplication by $E_{1,2n}^{-a}$ adds the 1st column of a matrix to its $2n$th column $-a$ times. Hence

$$\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & \ddots & \\ a^{-1} & & & 1 \end{pmatrix} E_{1,2n}^{-a} = \begin{pmatrix} 1 & & & -a \\ & \ddots & & \\ & & \ddots & \\ a^{-1} & & & 1-a(a^{-1}) \end{pmatrix}$$

$$= \begin{pmatrix} 1 & & & & -a \\ & \ddots & & & \\ & & \ddots & & \\ & & & 1 & \\ a^{-1} & & & & 0 \end{pmatrix}.$$

Now pre-multiplying this by $E_{1,2n}^{-a}$ will add the $2n$th row to the 1st row $-a$ times. This leaves us with the matrix

$$\begin{pmatrix} 0 & & & & -a \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ a^{-1} & & & & 0 \end{pmatrix}.$$

To get q_a we now just multiply by r_1. Hence $q_a = (E_{1,2n}^{-a})(E_{2n,1}^{a^{-1}})(E_{1,2n}^{-a})r_1$. So by the previous paragraph, 5.3.2 and 5.4.1, $l_{S'}(q_a) < 64np$. ■

Lemma 5.4.4: $l_{S'}(D_a) < 5np$

Proof. We have

$$v^{-a}F_{2,1}^{a^{-1}}v^{-a}s_{1,2} =$$

$$\begin{pmatrix} x^{-a} & 0 \\ 0 & (x^{-a})^{-TJ} \end{pmatrix}\begin{pmatrix} E_{2,1}^{a^{-1}} & 0 \\ 0 & (E_{2,1}^{a^{-1}})^{-TJ} \end{pmatrix}\begin{pmatrix} x^{-a} & 0 \\ 0 & (x^{-a})^{-TJ} \end{pmatrix}\begin{pmatrix} t_{1,2} & 0 \\ 0 & t_{1,2}^{-TJ} \end{pmatrix} =$$

$$\begin{pmatrix} x^{-a}E_{2,1}^{a^{-1}}x^{-a}t_{1,2} & 0 \\ 0 & (x^{-a}E_{2,1}^{a^{-1}}x^{-a}t_{1,2})^{-TJ} \end{pmatrix}$$

From 3.1.5 we see this is equal to

$$\begin{pmatrix} M_a & 0 \\ 0 & (M_a)^{-TJ} \end{pmatrix}$$

which is precisely the matrix D_a. Since $s_{1,2} = vF_{2,1}^{-1}v$ we have $l_{S'}(s_{1,2}) < 2n+3$. Now applying 5.2.2 we have $l_{S'}(D_a) < 5np$. ■

Lemma 5.4.5: Suppose D is a diagonal matrix in H. Then $l_{S'}(D) < 32pn^2$.

Proof.

Let D be the diagonal matrix in $Sp_{2n}(p)$ with entry d_i in row i for $1 \leq i \leq n$. Now

$$wD_{d_n^{-1}}w^{-1} = \begin{pmatrix} 1 & & & & & & & & \\ & d_n^{-1} & & & & & & & \\ & & d_n & & & & & & \\ & & & 1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & 1 & & & \\ & & & & & & d_n^{-1} & & \\ & & & & & & & d_n & \\ & & & & & & & & 1 \end{pmatrix}.$$

Multiplying by $D_{(d_n d_{n-1})^{-1}}$ gives the matrix

$$\begin{pmatrix} (d_n d_{n-1}^{-1}) & & & & & & & & \\ & d_{n-1} & & & & & & & \\ & & d_n & & & & & & \\ & & & 1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & 1 & & & \\ & & & & & & d_n^{-1} & & \\ & & & & & & & d_n & \\ & & & & & & & & 1 \end{pmatrix}.$$

By continuing to conjugate by w and multiplying by matrices of form D_a we

see that

$$w^{-(n-2)}D_{d_n{}^{-1}}wD_{d_nd_{n-1}^{-1}}wD_{d_nd_{n-1}d_{n-2}^{-1}}...wD_{d_nd_{n-1}...d_2{}^{-1}}$$

$$= \begin{pmatrix} \begin{pmatrix} (d_2d_3d_4...d_n)^{-1} & & & & & \\ & d_2 & & & & \\ & & d_3 & & & \\ & & & \ddots & & \\ & & & & d_{n-1} & \\ & & & & & d_n \end{pmatrix} & 0 \\ 0 & \begin{pmatrix} d_n & & & & & \\ & d_{n_1} & & & & \\ & & \ddots & & & \\ & & & d_3 & & \\ & & & & d_2 & \\ & & & & & (d_2d_3d_4...d_n)^{-1} \end{pmatrix} \end{pmatrix}.$$

Now it is easy to see that multiplying by $q_{d_1d_2d_3...d_n}$ we are left with the matrix D.

Using 5.4.4, and 5.4.3 we see $l_{S'}(D) < 5n^2p + 64pn$. This is less than or equal to $5n^2p + 32n^2p$ and so $l_{S'}(D) < 37n^2p$.

■

Lemma 5.4.6: $l_{S'}(H) \leq 69n^2p$

Proof. Suppose

$$h' = \begin{pmatrix} h & 0 \\ 0 & h^{-TJ} \end{pmatrix} \in H.$$

Then there exists a permutation $\pi \in S_n$ such that the non-zero entry in row i of h' appears in the $\pi(i)$th column of h'. Now π^{-1} may be written as a product of transpositions, $\tau_1\tau_2...\tau_k$ where $k \leq n$. Suppose $\tau_i = (a_i, b_i)$ then

$$P' = \begin{pmatrix} P & 0 \\ 0 & P^{-TJ} \end{pmatrix} = \prod_{i=1}^{k} s_{a_i,b_i}$$

is a monomial matrix where the ith row entry of P appears in the $\pi^{-1}(i)$th column. Applying 5.4.2 we have the length of this matrix is $32n^2p$.

Also $h'P'$ is diagonal. So $h' = (P')^{-1}D$ for some diagonal matrix, D and by 5.4.5, $l_{S'}(h')$ is less than $69n^2p$. ■

Corollary 5.4.7: $l_{S'}(N) \leq 101n^2p$

Proof. This follows directly from 5.4.1, 5.4.6 and the discussion at the beginning of the section. ■

5.5 Construction of U

Each matrix in U is a product of a matrix of form

$$X = \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix}$$

where X_1 and X_2 are upper uni-triangular matrices in $SL_n(p)$, and a one of form

$$M' = \begin{pmatrix} I & M \\ 0 & I \end{pmatrix}.$$

Note that each entry $(m_{i,j})$ of M has the property $m_{i,j} = m_{n-j+1,n-i+1}$.

We construct the matrices of form X in a similar way to the upper uni-triangular matrices from Chapter 3.

Suppose X is as described above. Then as shown in 3.1.9,

$$X_1 = \left(\prod_{i=1}^{n-1} b_i \right)$$

where the b_i are as described in Chapter 3. Hence

$$X' = \left(\prod_{i=1}^{n-1} b_i' \right)$$

where

$$b_i' = \begin{pmatrix} b_i & 0 \\ 0 & b_i^{-TJ} \end{pmatrix}.$$

and we now need to work out the length of each of these matrices.

Lemma 5.5.1: $l_{S'}(X') < 5pn^2$.

Proof.

From the proof of 5.2.1 we know $(vw)^{k-2}v(w^{-1}v^{-1})^{k-2}$ has form

$$\begin{pmatrix} \begin{pmatrix} 1 & & & 1 & & \\ & 1 & & 1 & & \\ & & \ddots & \vdots & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix} & 0 \\ 0 & \begin{pmatrix} 1 & & & 1 & & \\ & 1 & & 1 & & \\ & & \ddots & \vdots & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}^{-TJ} \end{pmatrix}.$$

Pre-multiplying this matrix by $\left(w^{(k-1)}\prod_{m=1}^{k-1} w^{-1}v^{w_{m,k}-1}\right)$ gives the matrix

$$\begin{pmatrix} Q & 0 \\ 0 & Q^{-TJ} \end{pmatrix}$$

where

$$Q = \begin{pmatrix} 1 & w_{1,k}-1 & & & & & & w_{1,k} & & \\ & 1 & w_{2,k}-1 & & & & & w_{2,k} & & \\ & & 1 & \ddots & & & & \vdots & & \\ & & & \ddots & \ddots & & & & & \\ & & & & 1 & w_{k-2,k}-1 & & & & \\ & & & & & 1 & w_{k-1,k} & & & \\ & & & & & & 1 & & & \\ & & & & & & & 1 & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & 1 \end{pmatrix}.$$

Finally post multiplying by $\left(\prod_{m=1}^{k-2}(v^{1-w_{m,k}}w)\right)w^{(k-2)}$ eliminates the entries above the diagonal and leaves us with $b_k.'$ Hence $l_{S'}(b'_k) < 10pk$.

Now $l'_S(X') \leq \sum_{i=1}^{n} l_{S'}(b'_k) \leq 5n^2p$. ■

To construct M' we will need to use the long roots of G.

Define M_k to be the matrix

$$M_k = I + \sum_{i=1}^{k} m_{i,n-k+1} e_{i,n-k+1} + \sum_{j=n-k+1}^{n} m_{k,j} e_{k,j}.$$

So M_k is the matrix

$$\begin{pmatrix} \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & 1 \end{pmatrix} & \begin{pmatrix} m_{1,n-k+1} & & & \\ m_{2,n-k+1} & & & \\ \vdots & & & \\ \vdots & & & \\ m_{k,n-k+1} & \cdots & \cdots & m_{k,n} \end{pmatrix} \\ 0 & \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & 1 \end{pmatrix} \end{pmatrix}.$$

Now it is easy to see that

$$M_1 M_2 ... M_n = M'$$

We now need to construct the matrices M_k.

Lemma 5.5.2: Fix $2 \leq k \leq n$. Then $l_{S'}(M_k) \leq 32pn + 11pk$.

Proof. In this proof we only illustrate the top right quadrants of the matrices. In each case the top left and bottom right quadrants will be the identity and the bottom left will be the zero matrix.

Define L_2 to be the matrix with top right quadrant

$$\begin{pmatrix} & m_{k-1,2n-k+1} & \\ & & \\ & 1 & m_{k+1,2n-k+2} \\ & & \\ & & \\ & & \\ & & \end{pmatrix},$$

and L_3 to be the matrix

$$\begin{pmatrix} & m_{k-2,2n-k+1} & & & \\ & m_{k-1,2n-k+1} & & & \\ & 1 & m_{k+1,2n-k+2} & m_{k+1,2n-k+3} & \\ & & & & \\ & & & & \\ & & & & \end{pmatrix},$$

and so on. So $L_j = I + e_{k,2n-k+1} + \sum_{i=1}^{j-1} m_{k-(j-i),2n-k+1} e_{i,2n-j+1} + \sum_{i=n+j+1}^{2n} m_{k,2n-k+j-1} e_{j,n+i}$.

Define r to be the smallest number such that L_r has some non-zero $m_{i,j}$ entries. Now, for $j < r$, the matrices L_j are precisely the matrices $E_{j,2n-j+1}$. We now construct L_r from $E_{r,2n-r+1}$.

The matrix $F_{1,r}^{m_{k-r+1,2n-k+1}-1} E_{r,2n-r+1} F_{1,r}^{m_{k,2n}}$ is the matrix with top right quadrant

$$\begin{pmatrix} & m_{k-r+1,2n-k+1} & m^2_{k-r+1,2n-k+1} \\ & & \\ & & \\ & 1 & m_{k-r+1,2n-k+1} \\ & & \\ & & \\ & & \end{pmatrix}.$$

If we remove the entry in the top right hand corner we will be left with L_r. We achieve this by post-multiplying by $E_{1,2n}^{-m^2_{k-r+1,2n-k+1}}$.

Now suppose $j \geq r$. We can construct L_{j+1} from L_j.

Set $c_j = \min\{c \in \mathbb{N} | m_{k-c,n-k+1} \neq 0.\}$ So c_j is the first row in L_j to have a non zero entry in its top right quadrant.

We construct L_{j+1} from L_j as follows. We first conjugate L_j by w. The product $wL_jw^{-}1$ has top right quadrant

$$\begin{pmatrix} & 0 & & & & \\ & m_{k-(j-1),n-k+1} & & & & \\ & \vdots & & & & \\ & m_{k-2,n-k+1} & & & & \\ & m_{k-1,n-k+1} & & & & \\ & 1 & m_{k,n-k+2} & \cdots & m_{k,} & 0 \\ & & & & & \\ & & & & & \\ & & & & & \end{pmatrix}$$

We now aim to place an entry of $m_{k-j,n-k+1}$ in the top row. If $m_{k-j,n-k+1} = 0$ we are done.

If $m_{k-j,n-k+1} \neq 0$, set $c_j = \min\{c \in \mathbb{N} | m_{k-c,n-k+1} \neq 0.\}$. So c_j is the first row in L_j to have a non zero entry in its top right quadrant. By repeatedly adding the c_jth row to the first row we get the required entry in the top row. We do this by repeatedly pre-multiplying by E_{1,c_j}. The matrix

$$(F_{1,c_j})^{m^{-1}_{k-(j-c_j),n-k+1} m_{k-j+1,k}} w L_j w^- 1$$

has top right quadrant

$$\begin{pmatrix} & & & & \\ m_{k-j,n-k+1} & & & & \\ m_{k-(j-1),n-k+1} & & & & \\ \vdots & & & & \\ m_{k-2,n-k+1} & & & & \\ m_{k-1,n-k+1} & & & & \\ 1 & m_{k,n-k+2} & \cdots & m_{k,} & 0 \\ & & & & \\ & & & & \\ & & & & \end{pmatrix}$$

Then we add the $j - c_j$th column to the last column to place the right entry in the last column. We have

$$L_{j+1} = (F_{1,c_j})^{m^{-1}_{k-(j-c_j),n-k+1} m_{k-j+1,k}} w^{-1} L_j w (F_{1,j-c_j})^{-m^{-1}_{k-j,k} m_{k-j+1,k}}$$

We have shown that to get L_{k-1} from L_r we conjugate by w up to k times and multiply by powers of the F_{1,c_j} up to $2k$ times. Now the total length contributed by the F_{1,c_j} is at most

$$\sum_{j=2}^{k-1} l_{S'} (F_{1,c_j})^{p-1}$$

$$\leq (p-1)\sum_{j=2}^{k-1} 10c_j$$

Now the total sum of the c_j cannot be greater than k. Hence the contribution of the F_{1,c_j} is at most $10pk$. Also, L_r contained $E_{1,2n}$ and a power of $E_{1,2n}$. By 5.3.1 and 5.3.2 these together give a contribution of less than $20np$. Accounting for the number of times we conjugate by w, we have $l_{S'}(L_{k-1}) \leq 11pk + 20np$.

Now M_k only differs from L_{k-1} in the entry in position $(k, 2n-k+1)$. To obtain M_k from L_{k-1} we need to repeatedly add the kth column to the $(2n-k+1)$th column until we have the correct entry in this position. Hence $M_k = L_{k-1}F_{k,2n-k+1}^{m_{k,n-k+1}-1}$. Now we have $l_{S'}(M_k) \leq 11pk + 20np + 12np = 32np + 11pk$. ■

Corollary 5.5.3: $l_{S'}(M') < 38n^2p$

Proof. From the discussion earlier, $M' = M_1M_2...M_n$. Hence

$$l_{S'}(M') \leq \sum_{i=1}^{n} l_{S'}(M_i)$$

$$\leq \sum_{i=1}^{n}(32pn + 11pk) < 38n^2p.$$

■

Now from the discussion at the beginning of this section we have

Corollary 5.5.4: $l_{S'}(U) \leq 43n^2p$

Combining this with 5.4.7 we now have Theorem 1.1.5.

The diameter of G with respect to S' is at most $187n^2p$.

5.6 Another Bound for The Diameter

Finally, we prove Theorem 1.1.6.

Lemma 5.6.1: Let the matrix D_a be as described in 5.4.4. Then $l_{S'}(D_a) < 4mn \log p$, where m is the constant from 2.2.1.

Proof. Firstly observe that $s_{2,1} = F_{1,2}F_{2,1}^{-1}F_{1,2}$. Also $F_{2,1} = w^n v w^{-n}$. So $l_{S'}(F_{2,1}) < 2n + 1$ and $l_S(s_{2,1}) \leq 2n + 3 \leq 4n$. Now,the group generated by $s_{2,1}$ and $F_{1,2}$ is equal to the group of matrices of form

$$\begin{pmatrix} A & 0 \\ 0 & A^{-TJ} \end{pmatrix}$$

where A has non-zero determinant and has form

$$A = \begin{pmatrix} a & b & & & \\ c & d & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}.$$

From the result by Babai, Lubotzky and Kantor, 2.2.1, we also know that the diameter of the group with respect to the set $\{s_{2,1}, F_{1,2}\}$ is $m \log p$ where m is a constant that does not depend on n or p. Since each matrix D_a lies in this group, the length of D_a with respect to the set $\{F_{1,2}, s_{2,1}\}$ is at most $m \log p$ and so $l_S(D_a) \leq 4mn \log p$. ■

Lemma 5.6.2: Let D be a diagonal matrix in $Sp(2n, p)$. Then $l_{S'}(D) < 133m'n^2 \log p$ where $m' = \max\{m, c\}$. Here m denotes the constant from 2.2.1 and c is the constant from 2.2.3.

Proof. Recall that from 5.3.1 we have $l_{S'}((E_{1,2n})^2) < 10n$ and from 5.4.2 we have $l_{S'}(s_{2n,1}) \leq 32n$.

Now using 2.2.3 we see that the set $T = \{(E_{1,2n})^2, (E_{2n,1})^2\}$ generates the group of matrices with form

$$\begin{pmatrix} a & & & & b \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ c & & & & d \end{pmatrix}$$

in G and this group of matrices has diameter $c \log p$ with respect to T. Hence $l_T((E_{1,2n})^q) \leq m' \log p$, where $m' = \max\{m, c\}$, and so $l_{S'}((E_{1,2n})^q) \leq 32m'n \log p$. Similarly each $l_{S'}((E_{2n,1})^a)$ and $l_{S'}(r_1)^a$ (where the r_i are the matrices constructed in 5.4.1) are each bounded above by $32m'n \log p$ for each $a \in F_p$.

Since $q_a = E_{1,2n}E_{2n,1}^{-1}E_{1,2n}(r_1)^a$ we have $l_{S'}(q_a) < 128mn \log p$.

As in 5.4.5, we can write D as the product of $n-1$ of the D_a, one matrix of form q_a and $2n$ elements of S'. Hence we have $l_{S'}(D) < 4m'n^2 \log p + 128m'n \log p + 2n$. Since $n \geq 2$ we can bound this above by $133m'n^2 \log p$. ■

Lemma 5.6.3: $l_{S'}(N) < 288m'n^2 \log p$

Proof. Any matrix in H can be written as $P^{-1}D$ where D is diagonal and P is as described in 5.4.6 . Now P is a product of n of the $s_{i,j}$ and has length at most $32n^2$. So, applying 5.6.2 any matrix in H has length less than $165m'n^2 \log p$. Any matrix in N is the product of a matrix in H

with at most n of the r_i, which have length of at most $128'mn \log p$. Hence $l_{S'}(N) < 288m'n^2 \log p$. ■

Lemma 5.6.4: Suppose X' is as described in 5.5.1. Then $l_{S'}(X') < 32n^3 \log p$.

Proof. In the proof of 5.5.1 we saw that the matrix X' could be written as a product of less than n^2 of the matrices of form v^q and w. where $q \in \mathbb{F}_p$. Since each matrix v^q has length of at most $32m'n \log p$, the length of X' is bounded above by $32n^3 \log p$. ■

Lemma 5.6.5: In the notation of 5.5.2, $l_S(M_k) < 129m'nk \log p + 2n + 96m'n \log p$.

Proof. Let $a \in \mathbb{F}_p^*$. By copying the proof of 5.2.1, we can see that

$$(F_{1,k})^q = w^{(k-2)}((v^q)^{-1}w^{-1})^{k-2}(v^a w)^{k-1}w^{-1}(w(v^q)^{-1})^{k-2}(w^{-1}(v^q))^{k-3}w^{k-3}.$$

Since $l_{S'}(v^q) < 32m'n \log p$, we have $l_{S'}(F_{1,k}) < 129mnk \log p$.

Recall that the matrix L_{k-1} is the product of a number of matrices of form $F_{1,c}$, $2n$ of the matrix w, a power of $E_{1,2n}$ and a power of $E_{2n,1}$.

The total length contributed by the $F_{1,c}$ is $129mnk \log p$. Also $l_{S'}(E_{1,2n}^q)$ and $l_{S'}(E_{2n,1})^q$ have length at most $32m'n \log p$. Hence $l_{S'}(L_{k-1}) < 129m'nk \log p + 2n + 64m'n \log p$.

Now $M_k = L_{k-1}F_{k,2n-k+1}^{m_{k,n-k+1}-1}$. Now $l_{S'}(F_{i,j}) < 12n$ by 5.2.2. By a a similar argument to that in 5.6.1 we see that $F_{i,j}$ has length of at most $m' \log p$ with respect to the set $\{F_{i,j}, s_{j,i}\}$. Hence $l_{S'}(F_{k,2n-k+1})^{m'_{k,n-k+1}-1} < 32m'n \log p$. This gives us that $l_{S'}(M_k) < 129m'nk \log p + 2n + 96m'n \log p$. ■

Corollary 5.6.6: $l_S(M) < 163m'n^3 \log p$.

Proof. Set $M_1 = I + m_{1,2n}E_{1,2n}$. Then

$$l_S(M) = \sum_{k=1}^{n} l_S(M_k)$$

$$< \sum_{k=1}^{n} 129nck \log p + 2n + 96m'n \log p$$

$$< 65n^3m'k \log p + 2n^2 + 96m'n^2 \log p$$

$$< 163m'n^3 \log p$$

■

Now we may deduce 1.1.6 from the previous corollary, 5.6.4 and 5.6.3.

6. USING COMPARISON TO DETERMINE A MIXING TIME OF THE SYMPLECTIC GROUP

6.1

In this chapter we prove Theorem 1.1.6. We establish an upper bound on the mixing time of $G = Sp_{2n}(p)$ given by a random walk with respect to the uniform distribution on the set $S' = \{I, v^{\pm 1}, w^{\pm 1}\}$, described in Chapter 5. We call this probability distribution, Q, so Q is given by

$$Q(g) = \begin{cases} \frac{1}{5} & g = v, w, v^{-1}, w^{-1} \text{ or } I \\ 0 & \text{otherwise.} \end{cases}$$

As in Chapter 4, we will bound the mixing time of G with respect to this random walk by bounding the distance between Q^k and U, the uniform distribution on G. To do this we will need to use 2.4.1.

To use 2.4.1 we'll need to bound the eigenvalues of the matrix associated

with Q. To do this, we will use comparison with the random walk given by the uniform distribution on the set of short root elements in $Sp_{2n}(p)$.

Let $\tilde{Q}$ be the uniform distribution on the set of short root elements in G. Then $\tilde{Q}$ has corresponding matrix $\tilde{M}$ with eigenvalues $\tilde{\pi}_i$.

To bound the eigenvalues $\tilde{\pi}_i$ we'll use 2.4.3 and 2.4.4. We first determine a suitable value of A in 2.4.4. Let T be the set of short root elements in G. Note that, if $t \in T$, then $N(s,t)$ and $|t|$ are both bounded above by $l_{S'}(G)$, which was shown to be less than $Kn^3 \log p$ for some constant K in 1.1.6.

$$\begin{aligned} A = & \quad \max_{s\in S'} \tfrac{1}{Q(s)} \textstyle\sum_{g\in G} |g| N(s,g)\tilde{Q}(g) \\ \leq & \quad 5 \textstyle\sum_{t\in T} (Kn^3 \log p)^2 \tilde{Q}(t). \\ = & \quad 5 \textstyle\sum_{t\in T} (K^2) n^6 (logp)^2 \tilde{Q}(t) \\ = & \quad K^2 n^6 \log^2 p. \end{aligned}$$

So, using 2.4.4, π_1 is bounded above by $1 - \frac{\tilde{\pi}_1}{K^2 n^6 (\log p)^2}$.

As in Chapter 4, we can establish a bound on $\tilde{\pi}_1$ with the following result using 2.5.2. Again, since the support of $\tilde{Q}$ is a single conjugacy class we know, from 2.5.1, that the eigenvalues of $\tilde{M}$ are the character ratios $\frac{\chi_\rho(t)}{\chi_\rho(1)}$

where ρ is an irreducible representation of G and t is a short root element in G.

Theorem 2.5.2 bounds the absolute values of these above by $\frac{19}{20}$.

Now we have $\frac{|\chi(x)|}{|\chi(1)} \leq \frac{19}{20}$ and so $\tilde{\pi}_i \leq \frac{19}{20}$.

Now

$$\pi_i \leq 1 - \frac{1-\tilde{\pi}_i}{A} \leq 1 - \frac{1-\frac{19}{20}}{A} = 1 - \frac{1}{20A}$$

Since $\frac{1}{20A} \geq \frac{1}{20K^2n^6(\log p)^2}$, we have $\pi_1 \leq 1 - \frac{1}{20K^2n^6(\log p)^2}$.

By using 2.4.2 we now can bound $-\pi_{|G|-1}$ above by $\frac{3}{5}$ which gives us that the largest eigenvalue, π_{max}, is bounded above by $1 - \frac{1}{20K^2n^6(\log p)^2}$. We are now ready to prove 1.4.3.

Proof. Let $K' = 20K^2$. Using 2.4.1 we have

$$4||Q^k - U|| \leq |G|\pi_1^{2k}.$$

$$\leq p^{n^2}\left(1 - \frac{1}{K'n^6(\log p)^2}\right)^{2k}$$

$$\leq \exp\left(n^2 \log p - \frac{2k}{K' n^6 (\log p)^2}\right)$$

$$= \exp\left(\frac{K' n^8 (\log p)^3 - 2k}{K' n^6 (\log p)^2}\right).$$

Hence the mixing time of $Sp_{2n}(p)$ is of order $n^8(\log p)^3$.

■

REFERENCES

[1] L. Babai, W.M. Kantor, and A. Lubotzky. Small-diameter cayley graphs for finite simple groups. *European J. Combin.*, 10(6):507–522, 1989.

[2] R.W. Carter. *Simple Groups of Lie Type.* John Wiley and Sons, 1972.

[3] P. Diaconis. *Group Representations in Probability and Statistics.* Institute of Mathematical Statistics, 1983.

[4] P. Diaconis. Random walks on groups: Characters and geometry. *London Math. Soc. Lecture Note Series*, 1:120–142, 2003.

[5] P. Diaconis and M. Shahshahani. Generating a random permutation with random transpositions. *Z. Wahrsch. Verw. Gebiete*, (2):159–179, 1981.

[6] D. Gluck. Sharper character value estimates for groups of lie type. *Journal of Algebra*, (218):155–181, 1995.

[7] M. Kassabov and T.R. Riley. Diameters of cayley graphs of chevalley groups. *European J. Combin.*, (3):791–800, 2007.

[8] A. Lubotzky. *Discrete groups, expanding graphs and invariant measures.* Birkhuser Verlag, Basel,, 1994.

[9] A. Lubotzky. Cayley graphs: Eigenvalues, expanders and random walks. *Surveys in Combinatorics*, (218):155–181, 1995.

[10] R.M. Nanayakkara. The diameters of some classical groups with respect to a small generating set. *Communications in Algebra*, pages 4259–4267, 2006.