Chapters 6 and 7 of Steinberg's Yale Notes

B.Sury

ATM workshop on Chevalley Groups
IISER Pune
May 2013.

Introduction

Chevalley's construction of the Chevalley groups unified and explained the classical simplicity results. Steinberg (*Colloque théorie des groupes algébriques*, Bruxelles 1962) constructed using generators and relations, the universal central extension of Chevalley groups over any field - these are now called Steinberg groups. Later, these have been generalized to arbitrary commutative rings with unity by M.R.Stein. While studying the congruence subgroup problem for quasi-split groups, Vinay Deodhar also generalized Steinberg's construction to non-split groups. Over the years, Abe, Bak, Hurley, Suzuki, Vavilov and others have developed extensive details on Chevalley-Demazure group schemes over rings. The theory of (abstract and topological) central extensions is the analogue of covering group theory for abstract groups. Steinberg also obtained an explicit presentation for the Chevalley group itself. The Schur multiplier of a Chevalley group over a field can be described by these means. Over local and global fields, the description is in terms of the norm residue symbol and leads to rich mathematics with connections to several parts of mathematics. We report on chapters 6 and 7 of Steinberg's celebrated Yale notes.

Steinberg's construction of Chevalley groups

This section reviews preliminary information and can be skipped if that answer has already been recalled.

For any abstract root system Φ with a basis $\Delta = \{\alpha_1, \cdots, \alpha_l\}$, one considers the free Lie algebra generated by $3l$ symbols $x_i, y_i, h_i (1 \leq i \leq l)$. On this free Lie algebra, if we force the 'Serre' relations

$$[x_i, y_i] = h_i \ , \ [x_i, y_j] = 0 \ \forall \ i \neq j$$

$$[h_i, h_j] = 0 \ \forall \ i, j$$

$$[h_i, x_j] = < \alpha_j, \alpha_i > x_j \ \forall \ i, j$$

$$[h_i, y_j] = - < \alpha_j, \alpha_i > y_j \ \forall \ i, j$$

$$(ad \ x_i)^{-<\alpha_j, \alpha_i>+1}(x_j) = 0 \ \forall \ i \neq j$$

$$(ad \ y_i)^{-<\alpha_j, \alpha_i>+1}(y_j) = 0 \ \forall \ i \neq j$$

where $< \alpha_j, \alpha_i >$ are the corresponding Cartan integers, one obtains a finite-dimensional semisimple Lie algebra $\mathcal{G}$ as Serre proved. Then, the next step is to show that this Lie algebra has a certain basis $\{x_\alpha; \alpha \in \Phi\} \cup \{h_i; i \leq l\}$ with certain special properties like the structure constants being in $\mathbf{Z}$ (what is known as a Chevalley basis). All this depends only on the root system really. The $\mathbf{Z}$-span $\mathcal{G}(\mathbf{Z})$ of a Chevalley basis is then a lattice in $\mathcal{G}$. For any $m \geq 0$ and any $\alpha \in \Phi$, the operators $(ad \ x_\alpha)^m/m!$ leave $\mathcal{G}(\mathbf{Z})$ invariant; hence the operator $exp(ad \ x_\alpha)$ itself leaves this lattice invariant (as $ad \ x_\alpha$ is nilpotent). Then, the group A of inner derivations of $\mathcal{G}$ is a matrix group which has the subgroup G generated by all $exp(ad \ cx_\alpha)$ as c varies in $\mathbf{Z}$. This G gives the algebraic group sought for, as it leaves the lattice $\mathcal{G}(\mathbf{Z})$ invariant, and hence consists of some integral matrices of determinant 1. Indeed, if T is a general indeterminate, then the matrix group generated by all $exp(ad \ Tx_\alpha)$ has entries from $\mathbf{Z}[T]$ and determinant 1. Specializing T to elements of any field (including finite fields), we get an algebraic group G. This is the group of adjoint type. Similarly, using other representations of the Lie algebra (the adjoint representation was used in the above construction), one can construct other 'covers'. This has essentially been done in earlier lectures but we shall rephrase and repeat them in Steinberg's language.

0.1. **A Chevalley basis.** We will recall a number of results proved in the earlier lectures on Lie algebras. Let $\mathcal{G}$ be a complex semisimple Lie algebra, and $\mathcal{H}$, Cartan subalgebra. We have

$$\mathcal{G} = \mathcal{H} \oplus \bigoplus_{\alpha \neq 0} \mathcal{G}_\alpha$$

as usual. A crucial fact proved in earlier lectures is that the set Φ of roots spans $\mathcal{H}^*$ as a vector space over $\mathbf{C}$. The nondegeneracy of the Killing form implies the existence of an element $H'_\alpha \in \mathcal{H}$ such that $(H, H'_\alpha) = \alpha(H)$ for all $H \in \mathcal{H}$. Thus, the definition $(\alpha, \beta) = (H'_\alpha, H'_\beta)$ gives a positive-definite, symmetric bilinear form on the $\mathbf{Q}$-vector space $\mathcal{H}^*_{\mathbf{Q}}$ spanned by the roots. We know that the reflections

$$s_\alpha : v \mapsto v - 2\frac{(v, \alpha)}{(\alpha, \alpha)}\alpha$$

in the hyperplane orthogonal to the root α, takes roots to roots and generate a finite group - the Weyl group of the root system. It is also useful (as we have seen in earlier lectures) to write the Cartan integers $2\frac{(\alpha,\beta)}{(\beta,\beta)}$ as $< \alpha, \beta >$. Moreover, if $\{\alpha_1, \cdots, \alpha_l\}$ is a simple system of roots, then W is generated already by s_{α_i}'s and every root is a W-translate of a simple root. The first observation is :

Lemma 1.
For each root α, the element $H_\alpha = \frac{2}{(\alpha,\alpha)}H'_\alpha$ is a $\mathbf{Z}$-linear combination of H_{α_i}.

The proof of this follows from the simple calculation which establishes

$$s_{\alpha_i}(H_{\alpha_j}) = H_{s_{\alpha_i}(\alpha_j)} \ \forall \ i, j.$$

For roots α, β, we write $\beta - r\alpha, \cdots, \beta, \cdots, \beta + q\alpha$ for the α-string through β. We know that there is such a string of roots and also that if $\alpha + \beta$ is a root, then $r - q =< \beta, \alpha >$ (as s_α maps $\beta - r\alpha$ to $\beta + q\alpha$). We have :

Lemma 2.
If $\alpha + \beta$ is a root, then

$$q\frac{|\alpha + \beta|^2}{|\beta|^2} = r + 1.$$

Having chosen H_{α_i}'s as above, we would like to choose generators for the root spaces in a suitable manner as follows :

Lemma 3.
There exist $X_\alpha \in \mathcal{G}_\alpha$ satisfying :
(a) $[X_\alpha, X_{-\alpha}] = H_\alpha$, and
(b) if $\beta \neq \pm\alpha$, then $[X_\alpha, X_\beta] = \pm(r+1)X_{\alpha+\beta}$ or $= 0$ according as to whether $\alpha + \beta$ is a root or not.
One usually writes these as $[X_\alpha, X_\beta] = N_{\alpha,\beta}X_{\alpha+\beta}$.

One calls a basis of $\mathcal{G}$ made up of H_{α_i}'s and X_α's as above, a Chevalley basis. More precisely, we have a *Chevalley basis* for $\mathcal{G}$ to be a basis which is a union of a basis $\{H_i; i \leq l\}$ of $\mathcal{H}$ and a basis $\{X(\alpha) : \alpha \in \Phi\}$ of $\bigoplus_{\alpha\in\Phi} \mathcal{G}_\alpha$, which satisfy the properties :
(a) $[H_i, X(\alpha)] =< \alpha, \alpha_i > X(\alpha)$,
(b) $[X(\alpha), X(-\alpha)] \in \mathcal{H}$ is a $\mathbf{Z}$-linear combination of the H_i's,
(c) $[X(\alpha), X(\beta)] = \pm(r+1)X(\alpha+\beta)$ or 0 according as to whether $\alpha + \beta$ is a root or not.

§ A Z-basis for $U(\mathcal{G})$.

Recall that by the PBW-theorem, the universal enveloping algebra $U(\mathcal{G})$ of $\mathcal{G}$ has as a basis, all monomials $X_1^{k_1} \cdots X_n^{k_n}$ where $X_1, \cdots, X_n$ gives a basis of the Lie algebra $\mathcal{G}$. Given a Chevalley basis, it is possible to get a corresponding basis of the universal enveloping algebra over $\mathbf{Z}$ itself. Indeed :

Proposition 4.
Let $\{H_{\alpha_i}; 1 \leq i \leq l\}$ be a basis of $\mathcal{H}$ as in lemma 1 and X_α (for $\alpha \in \Phi$) be as in lemma 3. Fix an ordering of these elements of $\mathcal{G}$. Consider the $\mathbf{Z}$-algebra $U(\mathcal{G})_\mathbf{Z}$ which is generated by the elements $\frac{X_\alpha^m}{m!}$ as α varies over roots and m varies over non-negative integer. Then, for every choice of non-negative integers n_i, m_α, the collection of products of all $\binom{H_{\alpha_i}}{n_i}$ and $\frac{X_\alpha^{m_\alpha}}{m_\alpha!}$ in the fixed order, forms a basis for $U(\mathcal{G})_\mathbf{Z}$.

For the proof, we need the following steps.

Step 1 :

$$\frac{X_\alpha^m}{m!}\frac{X_{-\alpha}^n}{n!} = \sum_i \frac{X_{-\alpha}^{n-i}}{(n-i)!}\binom{H_{\alpha_i} - m - n + 2i}{i}\frac{X_\alpha^{m-i}}{(m-i)!}$$

for each root α.
This is proved by first showing by induction on n that

$$X_\alpha \frac{X_{-\alpha}^n}{n!} = (\frac{X_{-\alpha}^n}{n!}X + \frac{X_{-\alpha}^{n-1}}{(n-1)!})(H_\alpha - n + 1)$$

and then using induction on m.
Step 2 :

$$\binom{H_\alpha}{n} \in U(\mathcal{G})_{\mathbf{Z}}.$$

To see this, one puts $m = n$ in step 1, applies induction on n and uses the elementary fact that a complex polynomial of l variables $z_1, \cdots, z_l$ that takes integral values for integral z_i's must be an integral linear combination of the polynomials $\prod_{i=1}^{l} \binom{z_i}{n_i}$ with n_i's bounded by the degree of the starting polynomial.
Step 3 :
If $\mathcal{G}_{\mathbf{Z}}$ denotes the $\mathbf{Z}$-span of H_{α_i}'s and X_α's, then each $\frac{X_\alpha^m}{m!}$ preserves each tensor product $\mathcal{G}_{\mathbf{Z}} \otimes \mathcal{G}_{\mathbf{Z}} \otimes \cdots \otimes \mathcal{G}_{\mathbf{Z}}$ under the adjoint representation extended to $U(\mathcal{G})$.
This is checked by simply using the definitions.
Step 3 :
Let $S \subset \Phi$ be closed under addition and satisfy $S \cap -S = \emptyset$. Then, the product of all $\frac{X_\alpha^{m_\alpha}}{m_\alpha!}$ for $\alpha \in S$ and $m_\alpha \geq 0$ (taken in the fixed order) is a $\mathbf{Z}$-basis for the $\mathbf{Z}$-algebra $\mathcal{A}$ generated by all $\frac{X_\alpha^m}{m!}$ with $\alpha \in S$.
The PBW-theorem applied to the Lie algebra with basis $\{X_\alpha : \alpha \in S\}$ shows that each element of the above $\mathbf{Z}$-algebra is at least a complex combination of the given elements. That the coefficients are integral can be proved using the previous step.
Step 4 :
For any two roots α, β, and any $m, n \geq 0$, the element $\frac{X_\alpha^m}{m!}\frac{X_\beta^n}{n!}$ is an integral combination of $\frac{X_\beta^n}{n!}\frac{X_\alpha^m}{m!}$ and of monomials whose total X-degrees are smaller.
This follows for $\alpha = -\beta$ from step 1. In the contrary case, step 3 can be applied to the set S of roots of the form $i\alpha + j\beta$ arranged in the order $\alpha, \beta, \alpha + \beta, \cdots$

Sketch of proof of proposition 4.
Firstly, we keep in mind the observation that for any roots α and β, and any polynomial f, we have

$$X_\alpha^n f(H_\beta) = f(H_\beta - n\alpha(H_\beta))X_\alpha^n.$$

This simply follows from the case when f is a power, in which case, it is a consequence of the equality $[H_\beta, X_\alpha] = \alpha(H_\beta)X_\alpha$ and induction on the two exponents.
Now, by step 2, $\binom{H_{\alpha_i}}{n}$ are in the $\mathbf{Z}$-algebra under consideration and we need to show that the integral combinations of these elements give all elements. Of course, it suffices to show that all monomials arise as

integral combinations. One may apply induction on the total degree in the X's. Step 4 and the observation made in the beginning of the proof here, allows us to write every monomial as an integral combination of monomials where, for each α, the X_α-terms can be put together and in the order fixed. As

$$\frac{X_\alpha^m}{m!}\frac{X_\alpha^n}{n!} = \binom{m+n}{n}\frac{X_\alpha^{m+n}}{(m+n)!}$$

each X_α needs to be represented at most once. Also, (again by the observation in the beginning), the H-terms can be brought in front and, by the proof in step 2, can be written as integral combinations of the asserted elements in the fixed order.

0.2. **Lattices in representations of $\mathcal{G}$.** We know the highest weight theory for representations of $\mathcal{G}$ and we would like to get lattices invariant under the $\mathbf{Z}$-form of the universal enveloping algebra. We start with a consequence of the previous proposition.

Corollary 5 (of proposition 4).
If $U(\mathcal{G})^+_{\mathbf{Z}}$ (respectively, $U(\mathcal{G})^-_{\mathbf{Z}}$) denotes the $\mathbf{Z}$-subalgebra of $U(\mathcal{G})_{\mathbf{Z}}$ generated by the elements $\frac{X_\alpha^m}{m!}$ for $\alpha > 0$ (respectively, $\alpha < 0$) and m varies over non-negative integers, and if $U(\mathcal{G})^0_{\mathbf{Z}}$ denotes the $\mathbf{Z}$-subalgebra generated by all $\binom{H_{\alpha_i}}{n}$ (for $i \leq l$; $n \geq 0$), then we have

$$U(\mathcal{G})_{\mathbf{Z}} = U(\mathcal{G})^-_{\mathbf{Z}} U(\mathcal{G})^0_{\mathbf{Z}} U(\mathcal{G})^+_{\mathbf{Z}}.$$

This follows from step 3.

Proposition 6.
Every finite-dimensional representation V of $\mathcal{G}$ contains (under the induced action of $U(\mathcal{G})_{\mathbf{Z}}$) a lattice M which is left invariant. Moreover, every such lattice is the direct sum of its weight components.
Proof.
We need to get hold of a lattice M so that all the generators $\frac{X_\alpha^m}{m!}$ of $U(\mathcal{G})_{\mathbf{Z}}$ leave it stable.

Corollary 7.
If V is a faithful, finite-dimensional representation of $\mathcal{G}$ and M, a lattice in it invariant under $U(\mathcal{G})_{\mathbf{Z}}$), then one has

$$\{X \in \mathcal{G} : X(M) \subseteq M\} = \oplus_\alpha \mathbf{Z} X_\alpha \oplus \{H \in \mathcal{H} : \mu(H) \in \mathbf{Z}\}$$

where μ on the right runs over the weights of the representation. Thus, this set $\mathcal{G}_{\mathbf{Z}}$ is a lattice in $\mathcal{G}$ which depends on the representation but not on the lattice M.

Example 8.
We mentioned that the definition of the lattice $\mathcal{G}_{\mathbf{Z}}$ depends on the representation. For instance, look at the 3-dimensional Lie algebra generated by X, Y, H and such that $[H, X] = 2X, [H, Y] = -2Y, [X, Y] = H$. Under the adjoint representation V, the only weights are $\pm\alpha$ with $\alpha(H) = 2$. So, we have $\mathcal{G}_{\mathbf{Z}} = \mathbf{Z}X \oplus \mathbf{Z}Y \oplus \mathbf{Z}(\frac{H}{2})$. On the other hand, the Lie algebra is isomorphic to $\mathcal{G}' = sl_2$ with H corresponding to $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$; under the natural representation, the weights are $\pm\mu$ where $\mu(H) = 1$. Thus, $\mathcal{G}'_{\mathbf{Z}} = \mathbf{Z}X \oplus \mathbf{Z}Y \oplus \mathbf{Z}H$.

0.3. **Chevalley groups. Definition 9.**
Let k be *any* field. Then, for $\mathcal{G}, \mathcal{H}, V, M$ as above, one defines $V^k = M \otimes k, \mathcal{G}^k = \mathcal{G}_{\mathbf{Z}} \otimes k, \mathcal{H}^k = \mathcal{H}_{\mathbf{Z}} \otimes k V^k_\mu = M_\mu \otimes k, kX^k_\alpha = \mathbf{Z}X_\alpha \otimes k$, where the tensoring is over $\mathbf{Z}$.

Lemma 10.
(i) $V^k = \oplus_\mu V^k_\mu$,
(ii) $\mathcal{G}^k = kX^k_\alpha \oplus \mathcal{H}^k$, and every $X^k_\alpha \neq 0$, $\dim_k \mathcal{H}^k = \dim_{\mathbf{C}} \mathcal{G}$ and $\dim_{\mathbf{C}} \mathcal{G}$.
This is immediate from corollary 7.

As the action of $\frac{X^n_\alpha}{n!}$ on M is zero for large n, the action of $T^n \frac{X^n_\alpha}{n!}$ on $M \otimes \mathbf{Z}[T]$ is zero as well for large n, where T is a variable. Thus, there is an action of $exp(TX_\alpha) := \sum_{n \geq 0} T^n \frac{X^n_\alpha}{n!}$ on $M \otimes \mathbf{Z}[T]$ and hence, on $M \otimes \mathbf{Z}[T] \otimes k$. Following this up with a homomorphism $M \otimes \mathbf{Z}[T] \otimes k \to M \otimes k = V^k$ given by a specialization $T \mapsto t$ for some $t \in k$, we therefore have an action of $exp(tX_\alpha)$ on V^k. The element $exp(tX_\alpha)$ is also written as $X_\alpha(t)$ and is clearly an automorphism of V^k (the inverse corresponds to $-t$). Indeed, it is clear that for each root α, the automorphism $X_\alpha(t)$ is additive as a function of $t \in k$.
The *Chevalley group over k with respect to the representation V* is defined to be the subgroup of $Aut(V^k)$ generated by $X_\alpha(t)$ for $\alpha \in \Phi, t \in k$.

1. Some relations in Chevalley groups

Before starting with chapter 6 of Steinberg's notes, we recall some basic facts on Chevalley groups which will be relevant for our study of the Steinberg groups.

Let $\mathbf{g}$ be a complex, simple Lie algebra. Let Φ be its (irreducible, reduced) root system with respect to a Cartan subalgebra and let Δ be

the corresponding set of simple roots. Recall that a group of automorphisms of $\mathbf{g}$ - called a Chevalley group - was constructed. It was constructed with the help of a Chevalley basis of $\mathbf{g}$. Recall that this is a basis of the form $\{h_\alpha; \alpha \in \Delta\} \cup \{e_\beta; \beta \in \Phi\}$ where $h_\alpha = \frac{2\alpha}{(\alpha,\alpha)}$ is the co-root corresponding to α and e_β is a non-zero element in the root space $\mathbf{g}_\beta$ and they satisfy the following properties:

$$[e_\alpha, e_{-\alpha}] = h_\alpha \ \forall \ \alpha \in \Phi.$$

$$[h_\alpha, e_\beta] =< \beta, \alpha > e_\beta \ \forall \ \alpha \in \Delta, \beta \in \Phi$$

where $< \beta, \alpha >= 2\frac{(\beta,\alpha)}{(\alpha,\alpha)}$ are the Cartan integers.

$$[e_\alpha, e_\beta] = 0 \ if \ \alpha + \beta \notin \{0\} \cup \Phi.$$

$$[e_\alpha, e_\beta] = N_{\alpha,\beta} e_{\alpha+\beta} \ if \ \alpha + \beta \in \Phi$$

where the structure constant $N_{\alpha,\beta} = \pm(p+1)$ with p the largest integer with $\beta - p\alpha \in \Phi$.
The signs of the structure constants depend on the choice of the Chevalley basis.

Let K be any field. For each root α in Φ and each $t \in K$, $x_\alpha(t) := exp(t \ ade_\alpha)$ is an automorphism of $\mathbf{g}$. These automorphisms generate the Chevalley group $G(\Phi, K)$. The effect of these automorphisms on the Chevalley basis is described as follows.
$x_\alpha(t)$ leaves invariant, the 3-dimensional space spanned by $e_\alpha, h_\alpha, e_{-\alpha}$; on this ordered basis it acts by the matrix $\begin{pmatrix} 1 & -2t & -t^2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}$.
If α, β are linearly independent roots, then

$$x_\alpha(t)(h_\beta) = h_\beta - t \ < \alpha, \beta > e_\alpha$$

and

$$x_\alpha(t)(e_\beta) = \sum_{i \geq 0} \frac{t^i}{i!} N_{\alpha,\beta} N_{\alpha,\alpha+\beta} \cdots N_{\alpha,(i-1)\alpha+\beta} e_{i\alpha+\beta} = \sum_{i \geq 0} \pm t^i \binom{p+i}{i} e_{i\alpha+j\beta}.$$

Chevalley showed that the elements $x_\alpha(t)$ as α varies over Φ and t varies over K, satisfy the following commutator relations:

$$x_\alpha(s) x_\alpha(t) = x_\alpha(s+t)$$

$$[x_\alpha(s), x_\beta(t)] = \prod_{i\alpha+j\beta \in \Phi} x_{i\alpha+j\beta}(N_{\alpha,\beta,i,j} s^i t^j)$$

where the product is over roots of the form $i\alpha + j\beta$ with $i, j \geq 1$.
Here, $N_{\alpha,\beta,i,j}$ are integers depending on the order of terms in the product and the roots α and β but not on s, t.
The Chevalley commutator relations depend on the structure constants only, as follows:

$$N_{\alpha,\beta,i,1} = \frac{1}{i!} N_{\alpha,\beta} N_{\alpha,\alpha+\beta} \cdots N_{\alpha,(i-1)\alpha+\beta};$$

$$N_{\alpha,\beta,1,j} = \frac{-1}{j!} N_{\beta,\alpha} N_{\beta,\alpha+\beta} \cdots N_{\beta,(j-1)\beta+\alpha}.$$

In particular,

$$N_{\alpha,\beta,1,1} = N_{\alpha,\beta}.$$

So, if $\alpha + \beta$ is the only root of the form $i\alpha + j\beta$ (for instance, when Φ is simply-laced = types ADE), we have

$$[x_\alpha(s), x_\beta(t)] = x_{\alpha+\beta}(N_{\alpha,\beta} st).$$

$$N_{\alpha,\beta,3,2} = \frac{1}{3} N_{\alpha,\beta} N_{\alpha,\beta+\alpha} N_{\alpha,\beta+2\alpha} N_{\beta,\beta+3\alpha} \ if \ \alpha + \beta < 2\alpha + \beta;$$

$$N_{\alpha,\beta,3,2} = \frac{-1}{6} N_{\alpha,\beta} N_{\alpha,\beta+\alpha} N_{\alpha,\beta+2\alpha} N_{\beta,\beta+3\alpha} \ if \ \alpha + \beta > 2\alpha + \beta;$$

$$N_{\alpha,\beta,2,3} = \frac{-1}{3} N_{\beta,\alpha} N_{\beta,\alpha+\beta} N_{\beta,\alpha+2\beta} N_{\alpha,\alpha+3\beta} \ if \ \alpha + 2\beta < \alpha + \beta;$$

$$N_{\alpha,\beta,2,3} = \frac{1}{6} N_{\beta,\alpha} N_{\beta,\alpha+\beta} N_{\beta,\alpha+2\beta} N_{\alpha,\alpha+3\beta} \ if \ \alpha + 2\beta > \alpha + \beta.$$

Of course, if Φ has rank 1, the last commutator identity is vacuous; it has the following analogue in case Φ has rank 1.
For $t \in K^*$, define $w_\alpha(t) = x_\alpha(t) x_{-\alpha}(t^{-1}) x_\alpha(t)$ where $\Phi = \{\pm\alpha\}$. Then, we have the relations:

$$w_\alpha(s) x_\alpha(t) w_\alpha(-s) = x_{-\alpha}(-s^{-2} t)$$

We are going to discuss the beautiful work of Robert Steinberg in which he shows that an abstract group defined by the above relations is an analogue of a universal cover for the Chevalley group and the Chevalley group itself is determined by these relations along with the additional relations

$$h_\alpha(s) h_\alpha(t) = h_\alpha(st) \ \forall \ s, t \in K^*$$

where $h_\alpha(t) = w_\alpha(t) w_\alpha(1)^{-1}$ for each $\alpha \in \Phi$.
M.R.Stein has generalized Steinberg's results to commutative rings with unity. Also, Vinay Deodhar generalized Steinberg's results to quasi-split groups over K.

Let us discuss Steinberg's results now. The proofs will depend crucially on the Bruhat decomposition; so, we recall that first.

Consider a Chevalley group $G := G(\Phi, K)$ as above. Let N and H denote, respectively, the subgroups of G generated by $w_\alpha(t)$ and by $h_\alpha(t)$ as α varies in Φ and t varies in K^*. Further, if U denotes the subgroup of G generated by $x_\alpha(t)$ as varies over the positive roots and t over K, and if we write $B := UH$, then we have:
(i) *H normalizes U;*
(ii) *H is normal in N and the quotient N/H can be identified with the Weyl group W; denote by w_α the element of W corresponding to the coset of $w_\alpha(1)$;*
(iii) *$G = \sqcup_{w\in W} UHn_wU$ where, for each $w \in W$, n_w is an arbitrary lift in N;*
(iv) (Normal form of Bruhat decomposition)
Each element of UHn_wU has a unique expression as uhn_wv where $u \in U$ and v is in the group generated by $x_\alpha(t)$ as α varies over positive roots such that $w(\alpha) < 0$ and t varies in K;
(v) *UH is the normalizer in G of U and also of UH;*
(vi) *N is the normalizer in G of H when $|K| > 3$.*

The above properties are deduced from the following set of relations which are well-known **in the Chevalley group:**

$$x_\alpha(s)x_\alpha(t) = x_\alpha(s+t).$$

$$[x_\alpha(s), x_\beta(t)] = \prod_{i\alpha+j\beta\in\Phi} x_{i\alpha+j\beta}(N_{\alpha,\beta,i,j}s^it^j).$$

$$w_\alpha(s)x_\beta(t)w_\alpha(s)^{-1} = x_{w_\alpha(\beta)}(cs^{-<\beta,\alpha>}t)$$

where $c = \pm 1$ depends on α, β but is independent of s, t;

$$w_\alpha(s)w_\beta(t)w_\alpha(s)^{-1} = w_{w_\alpha(\beta)}(cs^{-<\beta,\alpha>}t)$$

with c same as above.

$$h_\alpha(s)x_\beta(t)h_\alpha(s)^{-1} = x_\beta(s^{<\beta,\alpha>}t).$$

$$h_\alpha(s)w_\beta(t)h_\alpha(s)^{-1} = w_\beta(s^{<\beta,\alpha>}t).$$

$$w_\alpha(s)h_\beta(t)w_\alpha(s)^{-1} = h_{w_\alpha(\beta)}(cs^{-<\beta,\alpha>}t)h_{w_\alpha(\beta)}(cs^{-<\beta,\alpha>})^{-1} = h_{w_\alpha(\beta)}(t).$$

Also, the signs $c(\alpha, \beta)$ above satisfy:

$$c(\alpha, \beta) = c(\alpha, -\beta);$$

$$c(\alpha, \alpha) = c(\alpha, -\alpha) = -1;$$

$$c(\alpha,\beta)c(\alpha,w_\alpha(\beta)) = (-1)^{<\beta,\alpha>};$$
$$c(\alpha,\beta) = 1 \; if \; \alpha+\beta \notin \Phi \cup \{0\};$$
$$c(\alpha,\beta) = c(\beta,\alpha) = -1 \; if \; <\alpha,\beta> = -1 = <\beta,\alpha>;$$
$$c(\alpha,\beta) = -1 \; if \; <\alpha,\beta> = 0 \; , \; \alpha \pm \beta \in \Phi.$$

2. The Steinberg groups

Let Φ, K, G be as above. Consider the set $\hat{X}$ of symbols $\hat{x}_\alpha(t)$ as α varies in Φ and t varies in K. Define, for $t \in K^*$, symbols $\hat{w}_\alpha(t), \hat{h}_\alpha(t)$ in the obvious manner in terms of the various $\hat{x}_\beta(s)$'s.
The *Steinberg group* corresponding to Φ is defined to be the abstract group $St(\Phi,K) = <\hat{X}; R>$ where R is the set of relations:

$$\hat{x}_\alpha(s)\hat{x}_\alpha(t) = \hat{x}_\alpha(s+t) \cdots (R1).$$

$$[\hat{x}_\alpha(s), \hat{x}_\beta(t)] = \prod_{i\alpha+j\beta \in \Phi} \hat{x}_{i\alpha+j\beta}(N_{\alpha,\beta,i,j}s^i t^j) \cdots (R2)$$

if Φ has rank > 1; or

$$\hat{w}_\alpha(s)\hat{x}_\alpha(t)\hat{w}_\alpha(-s) = \hat{x}_{-\alpha}(-s^{-2}t) \cdots (R2')$$

if Φ has rank 1.
The numbers $N_{\alpha,\beta,i,j}$ are as in the Chevalley commutator relations above and take values among $\pm 1, \pm 2, \pm 3$.
Clearly, there is a surjective homomorphism Π from $St(\Phi,K)$ to $G(\Phi,K)$ which maps $\hat{x}_\alpha(t)$ to $x_\alpha(t)$.
The first theorem of Steinberg we wish to prove is:

Theorem 1.
The following relations must also hold in $St(\Phi,K)$:

$$\hat{w}_\alpha(s)\hat{x}_\beta(t)\hat{w}_\alpha(-s) = \hat{x}_{w_\alpha(\beta)}(cs^{-<\beta,\alpha>}t) \cdots (R3)$$

where $c = \pm 1$ is the constant occurring in the analogous equality for the Chevalley group $G(\Phi,K)$;

$$\hat{h}_\alpha(s)\hat{x}_\beta(t)\hat{h}_\alpha(s)^{-1} = \hat{x}_\beta(s^{<\beta,\alpha>}t) \cdots (R4).$$

$$\hat{w}_\alpha(s)\hat{h}_\beta(t)\hat{w}_\alpha(-s) = \hat{h}_{w_\alpha(\beta)}(cs^{-<\beta,\alpha>}t)\hat{h}_{w_\alpha(\beta)}(cs^{-<\beta,\alpha>})^{-1} \cdots (R5).$$

$$\hat{h}_\alpha(s)\hat{h}_\beta(t)\hat{h}_\alpha(s)^{-1} = \hat{h}_\beta(s^{<\beta,\alpha>}t)\hat{h}_\beta(s^{<\beta,\alpha>})^{-1} \cdots (R6).$$

$$\hat{w}_\alpha(s)\hat{w}_\beta(t)\hat{w}_\alpha(-s) = \hat{w}_{w_\alpha(\beta)}(cs^{-<\beta,\alpha>}t) \cdots (R7)$$

$$\hat{h}_\alpha(s)\hat{w}_\beta(t)\hat{h}_\alpha(s)^{-1} = \hat{w}_\beta(s^{<\beta,\alpha>}t) \cdots (R8).$$

Before proving theorem 1, we recall two facts which have been proved earlier in this workshop:

Fact I (Lemma 17): If $S \subset \Phi$ is a closed subset such that $S \cap -S = \emptyset$, then every element of the subgroup X_S of $G(\Phi, K)$ generated by $x_\alpha(t)$ as α varies in S and t varies in K, can be expressed uniquely as a product $\prod_{\alpha \in S} x_\alpha(t_\alpha)$ where the product is taken in any fixed order.

Fact II (Corollary to Lemma 33): If Φ has rank at least 2, then any $\alpha \in \Phi$ can be written as $\beta + n\gamma$ for some β, γ in Φ and some positive integer $n > 0$ such that $N_{\beta,\gamma,1,n} \neq 0$.
In fact, most of the time $n = 1$ itself works (see page 51).

Consequence of Fact I:
For any closed $S \subset \Phi$ with $S \cap -S = \emptyset$, the homomorphism $\Pi : St(\Phi, K) \to G(\Phi, K)$ maps the subgroup $\hat{X}_S$ isomorphically to X_S.
This is so because any element of $St(\Phi, K)$ can be reduced to a product of the form $\prod_{\alpha \in S} \hat{x}_\alpha(t_\alpha)$ using the commutator relations and any element $\prod_{\alpha \in S} x_\alpha(t_\alpha)$ in X_S has a unique such expression.

Proof of theorem 1.
We notice that (R4) to (R8) are consequences of (R1),(R2) and (R3).
Let us deduce (R3) from (R1) and (R2).
Clearly, the case $\alpha = -\beta$ reduces to that of $\alpha = \beta$ because $\hat{w}_\alpha(-t)$ is the inverse of $\hat{w}_\alpha(t)$ and $w_\alpha(\alpha) = -\alpha$.
Therefore, assume $\alpha \neq -\beta$.
If $\alpha \neq \pm\beta$, look at the closed set

$$S = \{i\alpha + j\beta \in \Phi : j > 0\}.$$

Write X_α and $\hat{X}_\alpha$ for X_T and $\hat{X}_T$ with $T = \{\alpha\}$. Then, (R2) shows that $\hat{X}_S$ is normalized by the subgroups $\hat{X}_\alpha$ and $\hat{X}_{-\alpha}$. Thus, by definition, $\hat{w}_\alpha(s)$ being in the subgroup generated by $\hat{X}_\alpha$ and $\hat{X}_{-\alpha}$ normalizes $\hat{X}_S$ as well. As we observed (consequence of fact I), $\hat{X}_S$ maps isomorphically onto its image in $G(\Phi, K)$ and we need verify (R3) only in $G(\Phi, K)$ where, of course, it is already observed.
Next, look at $\alpha = \beta$.
If Φ has rank 1, then (R3) follows simply from (R2).
Assume Φ has rank > 1 and $\alpha = \beta$. Then fact II recalled above shows there exist roots β, γ such that

$$\alpha = \beta + n\gamma$$

with $n > 0$ and $N_{\beta,\gamma,1,n} \neq 0$.
Let us consider then the subset S of Φ consisting of $iw_\alpha(\beta) + jw_\alpha(\gamma)$ with $i, j > 0$. Now

$$[\hat{x}_\beta(s), \hat{x}_\gamma(t)] = \prod_{i\beta+j\gamma \in S} \hat{x}_{i\beta+j\gamma}(N_{\beta,\gamma,i,j} s^i t^j)$$

has $N_{\beta,\gamma,1,n} \neq 0$.
Conjugating both sides by $\hat{w}_\alpha(u)$, and applying the case already treated above, all terms on the right hand side (other than the one corresponding to α) are known to belong to $\hat{X}_S$. Hence, we must have

$$\hat{w}_\alpha(u)\hat{x}_\alpha(N_{\beta,\gamma,1,n} u t^n)\hat{w}_\alpha(-t) \in \hat{X}_S.$$

Once again, the proof of (R3) reduces therefore to the corresponding assertion in G where it is known/easily verified.
The proof is complete.

3. Presentation for Chevalley groups

In this section, we use the Steinberg group to obtain a presentation for the abstract group $G(\Phi, K)$. More precisely, we prove:

Theorem 2.
Consider the abstract group $\tilde{G}$ defined by the presentation $< \tilde{X}; S >$ where $\tilde{X}$ consists of symbols $\tilde{x}_\alpha(t)$ as α varies over Φ and t varies over K, and $S = \{R1, R2/R2', T\}$ where

$$\tilde{x}_\alpha(s)\tilde{x}_\alpha(t) = \tilde{x}_\alpha(s+t) \cdots (R1).$$

$$[\tilde{x}_\alpha(s), \tilde{x}_\beta(t)] = \prod_{i\alpha+j\beta \in \Phi} \tilde{x}_{i\alpha+j\beta}(N_{\alpha,\beta,i,j} s^i t^j) \cdots (R2)$$

if Φ has rank > 1; or

$$\tilde{w}_\alpha(s)\tilde{x}_\alpha(t)\tilde{w}_\alpha(-s) = \tilde{x}_{-\alpha}(-s^{-2}t) \cdots (R2')$$

if Φ has rank 1; and

$$\tilde{h}_\alpha(rs) = \tilde{h}_\alpha(r)\tilde{h}_\alpha(s) \cdots (T).$$

Here s, t vary in K^*. Also, $\tilde{w}_\alpha(t)$, $\tilde{h}_\alpha(t)$ etc. here are defined in the obvious manner in terms of the generators in $\tilde{X}$. Then, $\tilde{G}$ is isomorphic to the universal Chevalley group $G(\Phi, K)$.

As mentioned earlier, a crucial result to be used is a Bruhat decomposition for the Steinberg group. We recall now how the Bruhat decomposition for G carries over to $St(\Phi, K)$. We define $\hat{U}$ in the obvious way; that is, as the subgroup generated by $\hat{x}_\alpha(t)$ as α varies over the

positive roots and t varies in K.
For each α, consider the reflection w_α. We define $\hat{w}_\alpha := \hat{w}_\alpha(1) \in St(\Phi)$.
For each $w \in W$, the Weyl group, one may write

$$w = w_{\alpha_1} \cdots w_{\alpha_k}$$

Define analogous elements $\hat{n}_w \in St$ as

$$\hat{n}_w = \hat{w}_{\alpha_1} \cdots \hat{w}_{\alpha_k}.$$

Recall that we have a surjective homomorphism Π from St to G which takes $\hat{x}_\alpha(t)$ to $x_\alpha(t)$. We claim:

Lemma 1 (Proposition+ Corollary 1 on P.39-40).
In the following, any quotient group of $St(\Phi, K)$ can be used although we use the notations for the Steinberg group.
(i) *Every element of $\hat{U}$ can be written uniquely as $\prod_{\alpha>0} \hat{x}_\alpha(t_\alpha)$.*
(ii) *Every element of $St(\Phi, K)$ has a unique expression of the form $uh\hat{n}_w v$ for some $w \in W$ with $u \in \hat{U}, h \in \hat{H}$ and v belonging to the subgroup generated by $\hat{x}_\alpha(t)$ as α varies over positive roots which are carried to negative roots by w and t varies in K.*
(iii) *Ker $\Pi \leq Z(St) \leq \hat{H}$.*
Proof.
We have already seen that (i) holds (consequence of fact I).
To see that (ii) is true, it is firstly clear that each element of St has an expression of the form $uh\hat{n}_w v$ (we can use (R3),(R4),(R5) as they hold good in St by theorem 1).
Suppose $uh\hat{n}_w v = u'h'\hat{n}_{w'}v'$. Applying Π and noting the uniqueness of the normal form of the Bruhat decomposition for G, and the isomorphism of $\hat{U}$ with its image under Π, we have $w = w', u = u'$ and $v = v'$. Thus, we get $h = h'$ as well.
To prove (iii), take an element z of Ker Π and express it as in (ii). Taking the image under Π and using uniqueness of the Bruhat normal form, we obtain $z \in \hat{H}$. We may write $z = \prod_{i=1}^{k} \hat{h}_{\alpha_i}(t_i)$. As z is in the center, it commutes with any $\hat{x}_\beta(t)$ for any β and any t. But, the relation (R4) gives

$$z\hat{x}_\beta(t)z^{-1} = \hat{x}_\beta\left((\prod_{i=1}^{k} t_i^{<\beta,\alpha_i>})t\right).$$

Hence, applying Π, we have

$$x_\beta\left((\prod_{i=1}^{k} t_i^{<\beta,\alpha_i>})t\right) = x_\beta(t)$$

which gives

$$\prod_{i=1}^{k} t_i^{<\beta,\alpha_i>} = 1.$$

Therefore, z commutes with each $\hat{x}_\beta(t)$, which shows that z is in the center of $St(\Phi, K)$.
Finally, to show that $Z(St) \leq \hat{H}$, it suffices to show that $Z(G) \leq H$ (because $H = \Pi(\hat{H})$ and Ker $\Pi \leq \hat{H}$).
Let $1 \neq c \in Z(G)$; write $c = uhn_w v$ in G with v in the subgroup generated by all $x_\alpha(t)$ where α varies over positive roots which are carried to negative roots by w if $w \neq 1$. If $w \neq 1$, then choose $\alpha > 0$ with $w(\alpha) < 0$. Then $cx_\alpha(1) = x_\alpha(1)c$ contradicts the uniqueness of the normal form of the Bruhat decomposition.
Therefore, we must have $c = uh$. This is super-diagonal. But the longest element w_0 of W satisfies the property that $n_{w_0} c n_{w_0}^{-1}$ is sub-diagonal. These are equal (as $c \in Z(G)$); so $u = 1$ and thus $x \in H$.

Proof of theorem 2.
Clearly, we have a surjective homomorphism $\tilde{\Pi} : \tilde{G} \to G$.
Moreover, theorem 1 shows that (R3),(R4),(R5),(R6),(R7),(R8) are also satisfied by $\tilde{G}$ (as they are satisfied by St which surjects onto $\tilde{G}$). Hence, for each root α, the subgroup $\tilde{H}_\alpha$ generated by $\tilde{h}_\alpha(t)$ as t varies in K^*, is normalized by each $\tilde{H}_\beta$ for any root β (from (R5).

Claim: *The subgroup $\tilde{H}$ generated by $\tilde{h}_\alpha(t)$ as α varies over roots and t varies in K^* is expressible as $\prod_{i=1}^{k} \tilde{H}_{\alpha_i}$ where $\alpha_1, \cdots, \alpha_k$ are the simple roots.*
We need to show that for each root β, $\tilde{H}_\beta$ is contained in the right hand side. This follows by writing β as $w\alpha_i$ for some $w \in W$ and applying induction on the length of w. Let $1 \neq w$ and write w as a product of simple reflections, say $w = w_\alpha \cdots$ where α is a simple root. We put $\gamma = w_\alpha(\beta)$. Then (R5) implies

$$\tilde{h}_\beta(t) = \tilde{w}_\alpha(1)\tilde{h}_{w_\alpha(\beta)}(c(-1)^{-<\beta,\alpha>}t)\tilde{h}_{w_\alpha(\beta)}(c(-1)^{-<\beta,\alpha>})^{-1}\tilde{w}_\alpha(-1).$$

Using (R8), this can be rewritten as

$$\tilde{h}_\beta(t) = \tilde{h}_\gamma(c(-1)^{-<\beta,\alpha>}t)\tilde{h}_\gamma(c(-1)^{-<\beta,\alpha>})^{-1}\tilde{w}_\alpha(t^{-<\beta,\alpha>})\tilde{w}_\alpha(-1).$$

The right hand side above belongs to $\tilde{H}_\gamma \tilde{H}_\alpha$.
Applying induction hypothesis, the subgroup $\tilde{H}_\gamma$ is contained in the product $\prod_{i=1}^{k} \tilde{H}_{\alpha_i}$. The claim follows.
To continue with the proof of the theorem, look at an element z in

the kernel of the surjection $\tilde{\Pi} : \tilde{G} \to G$. By lemma 1 (applied to the quotient group $\tilde{G}$ of $St(\Phi, K)$), $z \in \tilde{H}$. Note also that the relation T implies that, for any root α, any element of $\tilde{H}_\alpha$ is of the form $\tilde{h}_\alpha(t)$ for some $t \in K^*$. Hence,

$$z = \prod_{i=1}^{k} \tilde{h}_{\alpha_i}(t_i)$$

for some $t_i \in K^*$. Applying $\tilde{\Pi}$, we have

$$1 = \prod_{i=1}^{k} h_{\alpha_i}(t_i)$$

in $G(\Phi, K)$.
As this is the universal Chevalley group, each $t_i = 1$ and so $z = \prod_{i=1}^{k} \tilde{h}_{\alpha_i}(1) = 1$.
This proves that $\tilde{G}$ is isomorphic to G and hence, completes the proof of theorem 2.

4. Over Finite fields

We saw in the previous section that the Steinberg group is a central extension of the universal Chevalley group for any irreducible root system Φ and any field K. We prove in this section that if K is algebraic over a finite field, the groups coincide! First, we begin with the following definition and lemma over any field:

Definition (Symbols).
Fix any $\alpha \in \Phi$. For $u, v \in K^*$, the element

$$f_\alpha(u, v) = \hat{h}_\alpha(u)\hat{h}_\alpha(v)\hat{h}_\alpha(uv)^{-1}$$

is called a symbol.
We will not discuss symbols in detail but they will be of significance when the field K is local or global. The symbol will turn out to be essentially independent of α.

In this section, we fix α and, for convenience of notation, do not write the subscript in $f_\alpha(u, v)$ (we write $f(u, v)$ etc.), and also write $h(t)$ in place of $\hat{h}_\alpha(t)$, $x(t)$ in place of $\hat{x}_\alpha(t)$ and $y(t)$ in place of $\hat{x}_{-\alpha}(t)$. There should be no confusion with the corresponding elements in G.

Note that $f(u,v) \in \text{Ker } \Pi \leq Z(St)$ since the image is trivial in G.
Further, note that by definition, for $t \in K^*$,

$$w(t) = x(t)y(-t^{-1})x(t).$$

So, $w(-t) = w(t)^{-1}$.
Further, by (R3),

$$w(t)x(u)w(t)^{-1} = y(-t^{-2}u)\ ,\ w(t)y(u)w(t)^{-1} = x(-t^2u).$$

Hence, with $t = u$, we have

$$w(t)x(t)w(-t) = y(-t^{-1})$$

which gives

$$w(t) = y(-t^{-1})w(t)x(-t) = y(-t^{-1})x(t)y(-t^{-1})$$

Lemma 2.
$f : K^* \times K^* \to Z(St)$ satisfies:
(i) $f(t,u^2) = f(t,u)f(u,t)^{-1} = [h(t),h(u)] := h(t)h(u)h(t)^{-1}h(u)^{-1}$;
(ii) $f(t,u^2v) = f(t,u^2)f(t,v)$;
(iii) *If* $f(u,v) = f(v,u)$, *then* $f(u,v^2) = f(v,u^2) = 1$;
(iv) *If* $u = t^m, v = t^n$ *for some* t *and integers* m,n, *then* $f(u,v) = f(v,u)$;
(v) $f(t,1-t) = 1$ *if* $t \neq 1$.
(vi) (skew-symmetry) $f(t,-t) = 1$ *for all* $t \in K^*$.
Proof.
We first note that (R6) reduces (as $\alpha = \beta$) to:

$$h(x)h(y)h(x)^{-1} = h(x^2y)h(x^2)^{-1}.$$

(i) The above statement implies with $x = t, y = u$ that

$$h(t)h(u)h(t)^{-1}h(u)^{-1} = h(t)(h(u^2t)h(u^2)^{-1})^{-1}$$
$$= h(t)h(u^2)h(tu^2)^{-1} = f(t,u^2).$$

Clearly, the left hand side is also equal to $f(t,u)f(u,t)^{-1}$.

(ii) $f(t,u^2v) = h(t)h(u^2v)h(tu^2v)^{-1}$
$= h(t)h(u^2v)\left(h(u)h(tv)h(u)^{-1}h(u^2)\right)^{-1}$ (with $x = u, y = tv$)
$= h(t)h(u^2v)h(u^2)^{-1}h(u)h(tv)^{-1}h(u)^{-1}$
$= h(t)(h(u)h(v)h(u)^{-1})h(u)h(tv)^{-1}h(u)^{-1}$
$= h(t)h(u)h(v)h(tv)^{-1}h(u)^{-1}$
$= h(t)h(u)h(t)^{-1}f(t,v)h(u)^{-1}$
$= h(t)h(u)h(t)^{-1}h(u)^{-1}f(t,v)$
(as we can push the central element $f(t,v)$ to the end)

$= f(t,u^2)f(t,v)$ by (i).

(iii) This is immediate from (i).

(iv) $h(u) = h(t^m) = h(t)^m x$ for some $x \in Ker\Pi$.
Similarly, $h(v) = h(t^n) = h(t)^n y$ for some $y \in Ker\Pi$.
This implies $[h(u), h(v)] = 1$; that is, $f(u,v) = f(v,u)$.

(v) Now $f(t, 1-t) = h(t)h(1-t)h(t-t^2)^{-1} = 1$ if and only if

$$h(t)h(1-t) = h(t-t^2).$$

Equivalently, we wish to prove

$$w(t)w(-1)w(1-t) = w(t-t^2).$$

The LHS equals

$$\begin{aligned}
&w(t)y(1)x(-1)y(1)w(1-t)\\
&= \Big(w(t)y(1)w(-t)\Big)w(t)x(-1)y(1)w(1-t)\\
&= \Big(w(t)y(1)w(-t)\Big)w(t)x(-1)w(1-t)\Big(w(1-t)^{-1}y(1)w(1-t)\Big)\\
&= x(-t^2)w(t)x(-1)w(1-t)\Big(w(t-1)y(1)w(1-t)\Big)\\
&= x(-t^2)x(t)y(-t^{-1})x(t)x(-1)w(1-t)x(-(t-1)^2)\\
&= x(t-t^2)y(-t^{-1})x(t-1)x(1-t)y(-(1-t)^{-1})x(-(t-1)^2)\\
&= x(t-t^2)y(-t^{-1}-(1-t)^{-1})x(-(t-1)^2) = w(t-t^2)
\end{aligned}$$

since $-t^{-1} - (1-t)^{-1} = -(t-t^2)^{-1}$.

(vi) To show $f(t,-t) = 1$, we need to show $h(t)h(-t) = h(-t^2)$. Equivalently, we need to check if

$$w(t)w(-1)w(t)^{-1} = w(-t^2).$$

Now, the LHS equals

$$\begin{aligned}
&w(t)x(-1)y(1)x(-1)w(-t)\\
&= (w(t)x(-1)w(-t))(w(t)y(1)w(-t))(w(t)x(-1)w(-t))\\
&= y(t^{-2})x(-t^2)y(t^{-2}) = w(-t^2).
\end{aligned}$$

This finishes the proof.

Theorem 3.
Let K be algebraic over a finite field. Then, Π is an isomorphism from $St(\Phi, K)$ onto $G(\Phi, K)$.

Proof.
We need only show that $f(t,u) = 1$ for all t, u.
Look at the finite subfield k generated by t, u. If t or u is a square in k, this follows from lemma 2 (iii),(iv). In particular, the result follows when K has characteristic 2.
Assume that both t, u are non-squares.
As k has odd characteristic (so, odd cardinality), there exist $r, s \in k^*$ such that $r + s = 1$. Since the non-squares form the unique non-trivial coset in $k^*/(k^*)^2$, we have

$$t = x^2 r, u = y^2 s$$

So, $f(t,u) = f(x^2 r, y^2 s) = f(r,s) = 1$ by lemma 2.

5. Central extensions - generalities

Definition. A *central extension* is an exact sequence of groups

$$1 \to A \to E \to G \to 1$$

where the image of A is contained in the center of E.
For instance, for an abelian group A, the direct product of G and A gives such a central extension.

A central extension as above is to be thought of as a way of extending G by A. With this point of view, it is natural to call another such central extension

$$1 \to A \to F \to G \to 1$$

equivalent to the first one if there is an isomorphism between E and F giving a commutative diagram as in the figure. This is clearly an equivalence relation. Also, any central extension is equivalent to one in which the homomorphism from A to E is simply inclusion (exercise).

A central extension

$$1 \to A \to E \to G \to 1$$

is said to be *split* if it is equivalent to the trivial extension

$$1 \to A \to A \times G \to G \to 1$$

The terminology comes because these are precisely the extensions for which there is a *splitting* homomorphism from G to E giving the identity on G on composing it with the given surjection from E to G.

Natural examples of central extensions:
$1 \to K^* \to GL_n(K) \to PGL_n(K) \to 1$ where K is any field and K^*

sits as the scalar matrices in $GL_n(K)$.
$1 \to Z_n(K) \to SL_n(K) \to PSL_n(K) \to 1$ where K is any field and $Z_n(K)$ denotes the scalar matrices in $SL_n(K)$.
It should be noted here that $Z_n(K)$ is a finite subgroup of the group of all n-th roots of unity in K.

Exercises.
(i) *Let $\rho : G \to PGL_n(\mathbf{C})$ be a homomorphism (also called a projective representation of G). Show that it lifts to an actual representation from G to $GL_n(\mathbf{C})$ if the central extension*

$$1 \to \mathbf{C}^* \to \pi^{-1}(\rho(G)) \to \rho(G) \to 1$$

induced by

$$1 \to \mathbf{C}^* \to GL_n(\mathbf{C}) \xrightarrow{\pi} PGL_n(\mathbf{C}) \to 1$$

is split.
(ii) *Show that any central extension is equivalent to one in which the homomorphism from A to E is inclusion.*
Hint: Given any central extension

$$1 \to A \xrightarrow{\alpha} E \xrightarrow{\pi} G \to 1$$

choose any section $s : G \to E$ so that $s(1) = 1$ and $\pi \circ s = Id_G$. Define $f : G \times G \to \alpha(A)$ by $f(x, y) = s(x)s(y)s(xy)^{-1}$. Consider the set $F = A \times G$ with multiplication defined by $(a_1, g_1)(a_2, g_2) = (a_1 a_2 \alpha^{-1}(f(g_1, g_2)), g_1 g_2)$ gives the central extension

$$1 \to A \to F \to G \to 1$$

where $A \to F$ is the inclusion $a \mapsto (a, 1)$.

Central extensions arise naturally in the context of projective representations as seen in the exercise (i).

Let us see what the obstruction is to the existence of a splitting for a given central extension

$$1 \to A \to E \xrightarrow{\pi} G \to 1$$

One can, of course, choose some section i.e., set-theoretic splitting $s : G \to E$. Then, s is a group-theoretic splitting if $f(x, y) := s(x)s(y)s(xy)^{-1}$ is the identity. Note that the values of f land in A, the kernel of π. The map $f : G \times G \to A$ is, in fact, a 2-cocycle where the action of G on A is trivial. Moreover, the element defined in $H^2(G, A)$ is independent of the choice of s (see exercise below). In other words, there is a group-theoretic splitting precisely when the corresponding f

gives the trivial element in $H^2(G, A)$. In particular, *if* $H^2(G, A)$ *itself is trivial, any central extension is trivial.*
Notice that if

$$1 \to A \to E \xrightarrow{\pi} G \to 1$$

is an exact sequence with A abelian, then G acts on A by means of the inner automorphisms of E. In this way, even for a nontrivial action of G, the cohomology group $H^2(G, A)$ characterises *all extensions of* G *by* A i.e., exact sequences as above. In this more general situation, the trivial element of H^2 corresponds to the semi-direct product of G and A.

Exercise.
If s *is a set-theoretic splitting of a central extension*

$$1 \to A \to E \xrightarrow{\pi} G \to 1$$

then show that $f_s : G \times G \to A \ ; \ (x, y) \mapsto s(x)s(y)s(xy)^{-1}$ *is an element of* $Z^2(G, A)$ *for the trivial action of* G *on* A.
Further, if t *is any other splitting, then* $f_s = f_t$ *in* $H^2(G, A)$.
Hint: The proof of the first part was already given as a part of a hint. For the second part, note that if s, t are two splittings, then for any $x \in G$, the element $s(x)^{-1}t(x) \in A$ i.e., is central. So, $f_s^{-1}f_t$ as an element of $Z^2(G, A)$ is given by $(x, y) \mapsto \alpha(x)\alpha(y)\alpha(xy)^{-1}$ where $\alpha : G \to A$ with $\alpha(x) = s(x)^{-1}t(x)$. Here, one has used the fact that $s(x)t(x)^{-1} = t(x)^{-1}s(x)$ which holds good because $s(x) = at(x)$ for some $a \in A$.

Calculating central extensions of finite groups.

Given a finite presentation $< X \mid R >$ for a group G there is a canonical central extension induced. This is

$$1 \to R/[F, R] \to F/[F, R] \to G \to 1$$

Here, we have used R to denote also the normal subgroup of $F = F(X)$ generated by the relations R. The context will make it clear whether one is talking about the normal subgroup R or the set of relations R. Moreover, if G is finite, it is easy to see that the finitely generated abelian group $R/[F, R]$ is isomorphic to the direct product of $\mathbf{Z}^n$ and the finite subgroup $([F, F] \cap R)/[F, R]$ where $n = rank(F)$.

We noted that the notion of central extensions is an algebraisation of the notion of covering spaces. In covering space theory, one has the universal covers which have no nontrivial covers themselves. The

corresponding notion here is that of *universal central extensions* (abbreviated u.c.e).
A central extension

$$1 \to A \to E \xrightarrow{\pi} G \to 1$$

is *universal* if for any other central extension

$$1 \to B \to E' \xrightarrow{\pi'} G \to 1$$

there is a *unique* homomorphism $\theta : E \to E'$ so that $\pi = \pi' \circ \theta$. By the requirement of a unique θ, it follows that if there is a u.c.e of G, then it is unique upto equivalence. Sometimes, one simply writes (π, E) for the u.c.e. and $\mathrm{Ker}(\pi)$ is called the *Schur muliplier* of G.

Lemma 3.
(a) *If (π, E) is a u.c.e of G, then $E = [E, E]$ and $[G, G] = G$.*
(b) *If $G = [G, G]$, there exists a u.c.e of G.*

Proof
Look at the extension

$$1 \to B \to E' \xrightarrow{\pi'} G \to 1$$

where $E' = E \times E/[E, E]$ and $\pi'(x, y) = \pi(x)$. This is clearly a central extension. Moreover, the two homomorphisms $\theta_1(x) = (x, 1)$ and $\theta_2(x) = (x, x[E, E])$ from E to E' satisfy $\pi = \pi' \circ \theta_i$. By the uniqueness of such a map, one has $\theta_1 = \theta_2$ i.e., $E = [E, E]$. The last assertion that $G = [G, G]$ then follows trivially. This proves (a).
We construct a u.c.e when $G = [G, G]$. Consider the group E defined by generators $s(g)$ for each $g \in G$ and the relations

$$[s(x)s(y)s(xy)^{-1}, s(z)] \ ; \ x, y, z \in G.$$

The map $s(g) \mapsto g$ extends to a homomorphism $\pi : E \to G$. Let us denote the central element $s(x)s(y)s(xy)^{-1}$ of E by $t(x, y)$ for simplicity. Suppose $w \in \mathrm{Ker}(\pi)$. Write $w = s(x_1)s(x_2)\cdots s(x_n)$. Then, $x_1 \cdots x_n = 1$. Moreover, $w = t(x_1, x_2)s(x_1x_2)s(x_3)\cdots s(x_n)$. By induction on n, it follows that $w = cs(x_1 \cdots x_n) = c$ where c is a central element. Thus, (π, E) is a central extension.
Next, we need to know that for any other central extension

$$1 \to A \to E' \xrightarrow{\pi'} G \to 1$$

there is a homomorphism $\theta : E \to E'$ such that $\pi = \pi' \circ \theta$. For this, let us pick any arbitrary lifts $s'(g) \in E'$ of $g \in G$. Since $t'(x, y) := s'(x)s'(y)s'(xy)^{-1} \in \mathrm{Ker}(\pi') \leq Center(E')$, for any $x, y \in G$, the relations $[t'(x, y), s'(z)] = 1$ hold good for all $x, y, z \in G$. This means

that the map $s(g) \mapsto s'(g)$ extends to a homomorphism $\theta : E \to E'$. Evidently, $\pi = \pi' \circ \theta$ from the very definition.
But, θ may not be the unique such homomorpism i.e., (π, E) may not be universal. To get around this, one considers $F = [E, E]$ and the restriction π_F of π to F. Since $\pi(F) = [\pi(E), \pi(E)] = [G, G] = G$ as given, one has a central extension (π_F, F) of G. If θ_1 and θ_2 are two homomorphisms from F to E' such that $\pi_F = \pi' \circ \theta_1 = \pi' \circ \theta_2$, then $\theta_1(x)\theta_2(x)^{-1} \in \mathrm{Ker}(\pi') \leq Center(E')$. Thus, $\alpha : x \mapsto \theta_1(x)\theta_2(x)^{-1}$ is a homomorphism from F to an abelian group. Now, $\pi(E) = G = [G, G] = \pi([E, E])$ shows that $E = [E, E]\mathrm{Ker}(\pi) = F\mathrm{Ker}(\pi)$. Hence, $[E, E] = [F, F]$ as $\mathrm{Ker}(\pi)$ is central in E. Thus, $F = [F, F]$ i.e., α is trivial and so $\theta_1 = \theta_2$. This completes the proof of the lemma.

Exercises.
If (π, E) is a u.c.e of G, then prove :
(i) *that (Id, E) is a u.c.e of E, and*
(ii) *that every projective representation of G can be lifted uniquely to an actual representation of E.*
(iii) *For any abelian group A, one has $H^2(G, A) \cong \mathrm{Hom}(SchG, A)$ where $SchG$ is the Schur multiplier of G.*
Hint : E is perfect and does not admit nontrivial central extensions.

Topological central extensions

Let us consider a topological group G which is locally compact and second countable. Then, another such topological group E is said to be *a topological central extension* of G by a group A if there is a central extension

$$1 \to A \to E \xrightarrow{\sigma} G \to 1$$

with A closed in E, σ continuous and inducing an isomorphism $E/A \to G$ of topological groups.

Mackey and Moore showed that on equivalence classes of topological central extensions there is a natural multiplication under which the group becomes isomorphic to $H^2_m(G, A)$ where the cocycles are defined in terms of Borel-measurable cochains. Thus, for a connected topological group, covering space theory can be seen in terms of central extensions. To develop a 'covering space theory' for other types of groups like the p-adic Lie groups, the notion of topological central extensions proves very useful.

The correct analogue of $G = [G, G]$ (which was the condition for the existence of a u.c.e) in covering space theory is the connectedness of G. A covering $E \to G$ of connected topological groups is a topological central extension.

6. Steinberg is universal

We prove that $St(\Phi, K)$ is the u.c.e. of $G(\Phi, K)$. We give a proof which is valid for fields with at least 5 elements (and also avoid the field of 9 elements when Φ has rank 1).

Theorem 4.
Let $|K| > 4$ (and $|K| \neq 9$ when Φ has rank 1. Then, $\Pi : St(\Phi, K) \to G(\Phi, K)$ defines a u.c.e.
Here, G is the universal Chevalley group.
Proof.
Noting that $[\hat{h}_\alpha(a), x_\alpha(t)] = x((a^2-1)t)$, it follows that we may choose $a \neq 0, 1, -1$ when $|K| > 4$, and so the groups St and G are perfect. Hence, we only need to prove that St has no nontrivial central extensions. As any central extension of St gives a central extension of G itself, the idea of the proof is to show that for a central extension E of G, the relations (R1), (R2/R2') can be lifted. Consider any central extension $\pi : E \to G$. The crucial fact to be used is:
The commutator of two arbitrary lifts in E of elements of G is a well-defined element.
Denote $C = Ker\pi$.
Choose $a \neq 0, 1, -1$ in K; we noted that

$$[h_\alpha(a), x_\alpha(t)] = x((a^2-1)t)$$

in G, for any α and t.
Define $\phi : G \to E$ such that:
$\phi(x_\alpha(t))$ is a lift of $x_\alpha(t)$ to E and $\phi(w_\alpha(t)), \phi(h_\alpha(t))$ etc. in terms of $\phi(x_\alpha(t))$ etc. so that the following property holds:

$$[\phi(h_\alpha(a)), \phi(x_\alpha(t))] = \phi(x_\alpha((a^2-1)t)) \cdots (\spadesuit)$$

Note that this is not a cyclic definition because of the observation on commutator lifts made above.
In this manner, we may define ϕ on all of G.

Take any $h \in H$; write $hx_\alpha(t)h^{-1} = x_\alpha(dt)$ for some $d \in K^*$ which depends on h and is independent of α, t etc.
Conjugating ($\spadesuit$) by $\phi(h)$, and remembering that commutator does not

depend on the choice of lifts, we get

$$[\phi(h_\alpha(a)), \phi(x_\alpha(dt)] = \phi(h)\phi(x_\alpha((a^2-1)t))\phi(h)^{-1}$$

Thus,

$$\phi(x_\alpha((a^2-1)dt)) = \phi(h)\phi(x_\alpha((a^2-1)t))\phi(h)^{-1}$$

The LHS also equals

$$\phi(hx_\alpha((a^2-1)t))h^{-1})$$

so that we have (as $(a^2-1)t$ runs through arbitrary elements in K^*),

$$\phi(h)\phi(x_\alpha(u)\phi(h)^{-1} = \phi(hx_\alpha(u)h^{-1}).$$

Similarly, we have

$$\phi(n)\phi(x_\alpha(u)\phi(n)^{-1} = \phi(nx_\alpha(u)n^{-1})$$

for all $n \in N$.
The last assertion also shows that if Φ has rank 1, then ϕ respects the relation (R2').

Claim: *If α, β are roots with $\alpha + \beta$ is not a root and not zero also, then the commutator*

$$[\phi(x_\alpha(t)), \phi(x_\beta(u))] = 1.$$

Call this commutator $g(t,u)$. Note that clearly g is bi-additive. The claim is equivalent to showing that g is the constant map 1.

To prove the claim, we shall repeatedly use the relation

$$h_\alpha(t)x_\beta(u)h_\alpha(t)^{-1} = x_\beta(t^{<\beta,\alpha>}u)\cdots(\heartsuit)$$

Consider firstly the case when $\alpha \neq \beta$.
If $(\alpha, \beta) = 0$, then by conjugating the central element

$$g(t,u) = [\phi(x_\alpha(t)), \phi(x_\beta(u))]$$

by $\phi(h_\alpha(v))$, we obtain by $(\heartsuit)$ that

$$g(t,u) = [\phi(h_\alpha(v))\phi(x_\alpha(t))(\phi(h_\alpha(v)))^{-1}, \phi(h_\alpha(v))\phi(x_\beta(u))(\phi(h_\alpha(v)))^{-1}]$$
$$= [\phi(x_\alpha(v^2t)), \phi(x_\beta(u))] = g(v^2t, u)$$

since $(\alpha, \beta) = 0$.
Thus, $g(t(1-v^2), u) = 1$ which gives on choosing some $v \neq \pm 1$ that

$$g(x,u) = 1 \ \forall \ x, u.$$

If $(\alpha, \beta) > 0$, then choosing $h = h_\alpha(v^2)h_\beta(v^{-<\beta,\alpha>})$, we obtain

$$g(tv^d, u) = g(t,u)$$

where $d = 4 - < \alpha, \beta >< \beta, \alpha >= 1, 2$ or 3.
Hence $g(t(1 - v^d), u) = 1$ for some $d \in \{1, 2, 3\}$. As $|K| > 4$, we can choose v with $1 \neq v^d$. Once again, we get $g \equiv 1$.

Now, look at the case $\alpha = \beta$.
Firstly, assume that the rank of Φ is > 1.
If Φ is not of type C_n, there is always a root γ so that $< \alpha, \gamma >= 1$.
Choosing $h = h_\gamma(v)$ and applying

$$\phi(h)\phi(x_\alpha(u)\phi(h)^{-1} = \phi(hx_\alpha(u)h^{-1})$$

we obtain

$$g(tv, uv) = g(t, u).$$

Therefore, for v satisfying $v - v^2 \neq 0, 1$, we get

$$g(t(v-v^2), u) = g(t, \frac{u}{v - v^2}) = g(t, \frac{1}{v})g(t, \frac{1}{1 - v}) = g(tv, u)g(t(1-v), u) = g(t, u)$$

Thus, $g(t(1 - v + v^2), u) = 1$ and hence by the choice of v, we again have

$$g \equiv 1.$$

If Φ is of type C_n, one may write $\alpha = \delta + 2\beta$.
By explicit computation in C_2 (lemma 33) - this was used also in the "consequence of fact I" - we have

$$[x_\delta(t), x_\gamma(u)] = x_{\delta+\gamma}(\pm tu)x_{\delta+2\gamma}(\pm tu^2).$$

Thus, we obtain

$$[\phi(x_\delta(t)), \phi(x_\gamma(u))] = cx_{\delta+\gamma}(\pm tu)x_{\delta+2\gamma}(\pm tu^2)$$

for some $c \in Ker\pi$.
By the cases of the claim already proved above, $\phi(x_\alpha(v))$ commutes with each of the factors other than possibly the last, it must commute with the last one also, which proves our claim in the case when Φ has type C_n and $n > 1$.

Finally, consider the case when the rank of Φ is 1 and $\alpha = \beta$.
If $|K|$ is a prime, then $g \equiv 1$ since $x_\alpha(t)$ and $x_\alpha(u)$ are both powers of the element $x_\alpha(1)$. Thus, we assume $|K|$ is not prime.
Using $h = h_\alpha(v)$, we have by ($\heartsuit$) that

$$g(tv^2, uv^2) = g(t, u).$$

The argument given above (for $\alpha = \beta$ and Φ of type other than C_n and if rank > 1) implies that it suffices to get hold of non-zero v such that $v - v^2 \neq 0, 1$ and, in addition, v and $1 - v$ are both squares (because there we had $g(t, u) = g(tv, uv)$ for all v whereas here we have it only

for squares).
This is clearly possible when K is finite of characteristic 2 (because every element is a square).
Otherwise, choose $v = (\frac{2x}{1+x^2})^2$. Then, since we have $|K| \geq 25$ (this is why we avoided $|K| = 9$), it can be checked that there are at least 12 values of x with the desired properties $v - v^2 \neq 0, 1$. This completes the proof of the claim.

Now, we show that the relations (R1),(R2/R2') are respected by ϕ.
By the claim above, the element

$$z = \phi(x_\alpha(t(a^2-1)^{-1})\phi(x_\alpha(u(a^2-1)^{-1})\phi(x_\alpha((t+u)(a^2-1)^{-1})^{-1}$$

is in Ker π.
Conjugating this central element by $\phi(h_\alpha(a))$, the above observations imply that

$$z = z\phi(x_\alpha(t)\phi(x_\alpha(u))\phi(x_\alpha(t+u))^{-1}$$

which means that (R1) holds in E.

Now, look at (R2). We have

$$\phi(x_\alpha(t))\phi(x_\beta(u))\phi(x_\alpha(t))^{-1} = g(t,u)\Big(\prod_{i\alpha+j\beta} \phi(x_{i\alpha+j\beta}(N_{ij}t^i u^j)\Big)\phi(x_\beta(u)).$$

We need to prove that $g \equiv 1$. Look at the number k of roots in the above product. If $k = 0$, then we have shown above that $g \equiv 1$. We apply induction on k. Now, the induction hypothesis and the fact that ϕ respects (R1) implies that g is bi-additive. Once again, we may apply the method used in proving the above claim.
The proof is complete.

7. Moore's description of Schur multiplier

We have seen that the Steinberg group is the u.c.e. of the universal Chevalley group G and the kernel of the surjection Π can be identified with the Schur multiplier of G. We have seen that this is trivial when K is a finite field. While proving that, we noticed that certain 2-cocycles $f(u,v)$ played a key role. In general, Moore and Mastumoto described the Schur multiplier in the following fashion:

Theorem 5 (Moore-Matsumoto).
Let α be a long root. Then, the function

$$f : K^* \times K^* \to Ker\Pi;$$

$$(u,v) \mapsto \hat{h}_\alpha(u)\hat{h}_\alpha(v)\hat{h}_\alpha(uv)^{-1}$$

defines an isomorphism between $Ker\Pi$ *and the abstract generated by all symbols* $f(u,v)$ *with the relations:*
(a) f *is a 2-cocycle;* $f(u,1) = 1 = f(1,v)$*;*
(b) $f(u,v) = f(v^{-1},u)$*;*
(c) $f(u,v) = f(u,-uv)$*;*
(d) $f(u,v)f(u,-v^{-1}) = f(u,-1)$*;*
(e) $f(u,v) = f(u,(1-u)v)$*;*
(f) *When* Φ *has type other than* C_n*;* f *is bi-multiplicative.*

In the proof, the fact that it suffices to stick to a single long root to describe $Ker\Pi$, and the fact that the mapping from the abstract group to Ker Π is a surjective homomorphism are fairly easy (see pages 87-88). The injectivity requires more work.

8. $K_2(\mathbf{Z})$ has order 2

As mentioned earlier, Steinberg's results have been generalized by M.R.Stein to all types of root systems over any commutative rings. We discuss only the case of type A_n and the ring $\mathbf{Z}$. However, we define the Steinberg group over a general ring - it will be useful to study this group over the finite rings $\mathbf{Z}/k\mathbf{Z}$ as it leads to some applications over $\mathbf{Z}$.

Definition.
Let R be any commutative ring (with unity $1 \neq 0$).
For $n \geq 3$, define the *Steinberg group* $St_n(R)$ to be generated by the symbols $x_{ij}(t)$ for $i \neq j$ and $t \in R$, subject to the following relations:

$t \mapsto x_{ij}(t)$ are homomorphisms such that

$$[x_{ij}(t), x_{jk}(u)] = x_{ik}(tu) \; if \; i \neq k$$

$$[x_{ij}(t), x_{kl}(u)] = 1 \; if \; j \neq k \; , i \neq l$$

For $n = 2$, $St_2(R)$ is defined by the generators $x_{12}(t), x_{21}(u)$ where x_{12}, x_{21} are homomorphisms subject to the relations

$$w_{12}(t)x_{12}(u)w_{12}(-t) = x_{21}(-t^{-2}u) \text{if } t \in R^*$$

$$w_{21}(t)x_{21}(u)w_{21}(-t) = x_{12}(-t^{-2}u) \text{if } t \in R^*$$

where for any $n \geq 2$ and $i \neq j$, $w_{ij}(t) = x_{ij}(t)x_{ji}(-t^{-1})x_{ij}(t)$.

As we saw over fields, a Bruhat decomposition for the Steinberg group was vital to the discussion. To carry this over to rings, we need rings with many units. We discuss rings like $\mathbf{Z}/p^r\mathbf{Z}$ which are actually generated by the units. The ring $\mathbf{Z}$ is much harder to study!

Theorem 6.
For $n \geq 3$ and $k \geq 1$, the group $SL_n(\mathbf{Z}/k\mathbf{Z})$ is generated by the $n(n-1)$ elementary matrices $X_{ij}; i \neq j$ subject to the relations

$$[X_{ij}, X_{jk}] = X_{ik} \; if \; i \neq k$$
$$[X_{ij}, X_{kl}] = I \; if \; j \neq k \; , i \neq l$$
$$X_{12}^k = I$$
$$(X_{12}X_{21}^{-1}X_{12})^4 = I$$

The last relation is redundant unless $4 \mid k$.

Proof.
First, suppose that the asserted presentation holds for $SL_n(\mathbf{Z}/p^r)$ for prime powers. Now, by the Chinese remainder theorem, if $k = p_1^{r_1} \cdots p_s^{r_s}$, then $SL_n(\mathbf{Z}/k) \cong \prod_{i=1}^s SL_n(\mathbf{Z}/p_i^{r_i})$. Each $SL_n(\mathbf{Z}/p_i^{r_i})$ has a presentation $< F_i | R_i >$ with F_i consisting of certain symbols $x_{jk}^{(i)}$ for $j \neq k$ such that $(x_{jk}^{(i)})^{p_i^{r_i}} = 1$. Then, a presentation for $\prod_{i=1}^s SL_n(\mathbf{Z}/p_i^{r_i})$ is obtained by the generators

$$(x_{jk}^{(1)}, 1, \cdots), (1, x_{jk}^{(2)}, 1, \cdots), \cdots, (1, \cdots, 1, x_{jk}^{(s)}).$$

Call their product f_{jk}. Now, $ap_1^{r_1} - bq_1 = 1$ for some a, b where $q_1 = k/p_1^{r_1}$. Then, $f_{jk}^{-bq_1} = (x_{jk}^{(1)}, 1, \cdots, 1)$. Thus, the elements f_{jk} for $j \neq k$ generate $\prod_{i=1}^s SL_n(\mathbf{Z}/p_i^{r_i})$. Mapping the f_{jk} to the elementary matrices x_{jk}, it follows that the presentation holds good for $SL_n(\mathbf{Z}/k)$ for a general k once it is known for prime powers.
From now on, we consider $SL_n(\mathbf{Z}/p^r)$ for a prime p and $n \geq 3$. Consider the corresponding Steinberg group $St_n(\mathbf{Z}/p^r)$. Write L and U respectively, for the subgroup generated by $x_{ij}, i > j$ and that generated by $x_{ij}, i < j$. Also, denote by W, the subgroup generated by $w_{ij}(t)$ for t a unit and $i \neq j$. Then, we have the following Bruhat-type decomposition in $St_n(\mathbf{Z}/p^r)$:

Claim.
For any $n \geq 2$ and any prime p, $St_n(\mathbf{Z}/p^r) = LUW$.

The following calculation will prove useful :

Exercise.
Let A be any commutative ring and u, v be units. Then, in $St_n(A)$, we have for any $i \neq j$,

$$x_{ij}(-u)x_{ji}(u^{-1}(1-v))x_{ij}(u) =$$
$$x_{ji}(u^{-1}(v^{-1}-1))h_{ij}(uv)h_{ij}(u)^{-1}x_{ij}(u(1-v^{-1})) =$$

$$x_{ji}(-u^{-1})x_{ij}(u(v-1))x_{ji}(u^{-1}).$$

Proof of claim
We first note that $St_n(\mathbf{Z}/p^r)$ is generated by

$$x_{i+1,i}\ ,\ w_{i+1,i}\ ,\ 1 \leq i \leq n-1.$$

This is clear from the commutator relations $[x_{ij}, x_{jk}] = x_{ik}$ and the relation $w_{ij}x_{ij}w_{ij}^{-1} = x_{ji}^{-1}$. So, we need only show that LUW is stable under left multiplication by any $x_{i+1,i}^{\pm}$ and by any $w_{i+1,i}^{\pm}$. The first is evident. Also, $w_{i+1,i}^{\pm}LUW = x_{i+1,i}^{\pm}x_{i,i+1}^{\mp}LUW$; thus it suffices to prove that $x_{i,i+1}^{\pm}LUW \subseteq LUW$.
Start with any $g \in x_{i,i+1}^{\pm}LUW$ We can write $g = x_{i,i+1}^{e}x_{i+1,i}^{a}luw$ with $e = \pm 1, a \in \mathbf{Z}, l \in L^{i+1,i}$ where $L^{i+1,i}$ consists of those elements of L which can be written as products of x_{jk} with $j > k$ other than $x_{i+1,i}$.
We shall use the first equality in the exercise to conclude that $x_{i,i+1}^{e}x_{i+1,i}^{a} = x_{i+1,i}^{*}hx_{i,i+1}^{*}$ for some h which is a product of $h_{i,i+1}(v)$ for some suitable unit u. This is valid if $1 + ea$ is a unit i.e., if $a \not\equiv -e^{-1} \bmod p$.
This is one place where we use the existence of sufficiently many units in the ring.
Therefore, for such a, we get $g \in LUW$ because h normalises each $< x_{jk} >$ and $x_{i,i+1}$ normalises $L^{i+1,i}$.
Now, we are left with the case $a \equiv -e^{-1} \bmod p$. Let us write $e = -a^{-1} + tp$ for some integer t. Then, $g = x_{i,i+1}^{-a^{-1}+tp}x_{i+1,i}^{a}luw = x_{i,i+1}^{tp}x_{i,i+1}^{-a^{-1}}x_{i+1,i}^{a}luw$.
Using $x_{i,i+1}^{-a^{-1}}x_{i+1,i}^{a} = x_{i+1,i}^{-a}w_{i+1,i}(a)$, we get $g = x_{i,i+1}^{tp}x_{i+1,i}^{-a}w_{i+1,i}(a)luw = x_{i,i+1}^{tp}x_{i+1,i}^{-a}l'w_{i+1,i}(a)uw$ for some $l' \in L^{i+1,i}$ since $w_{i+1,i}(.)$ normalises $L^{i+1,i}$. Putting $u = x_{i,i+1}^{*}u_1$ for some $u_1 \in U^{i,i+1}$, we get

$$g = x_{i,i+1}^{tp}x_{i+1,i}^{-a}l'x_{i+1,i}^{*}u'w$$

as $w_{i+1,i}(.)$ normalises $U^{i,i+1}$ and conjugates $x_{i,i+1}$ to $x_{i+1,i}^{-1}$.
But, we can rewrite $x_{i,i+1}^{tp}x_{i+1,i}^{-a}$ as $x_{i+1,i}^{*}hx_{i,i+1}^{rp}$ for some integer r. So, $g = x_{i+1,i}^{*}l_0x_{i,i+1}^{rp}x_{i+1,i}^{b}u_0w_0$ for some integer b, some l_0, u_0 in $L^{i+1,i}, U^{i,i+1}$ respectively, and $w_0 \in W$. Therefore, it suffices to show that $g_0 = x_{i,i+1}^{rp}x_{i+1,i}^{b}u_0w_0 \in LUW$.
This is immediate if b is a unit. If $b = mp$, then $x_{i,i+1}^{rp}x_{i+1,i}^{mp}u_0w_0 = x_{i,i+1}x_{i,i+1}^{rp-1}x_{i+1,i}^{mp}u_0w_0 = x_{i,i+1}x_{i+1,i}^{tp}hx_{i,i+1}^{*}u_0w_0$ by the first part of the lemma. Once again, one can apply the first equality to rewrite the first two terms above and thereby obtain $g_0 \in LUW$. This completes the proof of the Bruhat-like decomposition claimed.

Let us continue with the proof of the theorem.
Let $\theta : St_n(\mathbf{Z}/p^r) \to SL_n(\mathbf{Z}/p^r)$ be the natural homomorphism.

Note that the conjugate of a Steinberg generator $x_{ij}(u)$ by any $w \in W$ is again of the form $x_{kl}(v)^{\pm}$, and so, $\mathrm{Ker}(\theta) \cap W$ is central. We shall show, in fact, that $\mathrm{Ker}(\theta) \leq W$. Let $h \in \mathrm{Ker}(\theta)$. Writing $h = luw$, we get $\theta(l)\theta(u)\theta(w) = I$. This gives $\theta(l) = 1 = \theta(u)$ since an expression $x = ym$ in SL_n with x lower triangular unipotent, y upper triangular unipotent and w monomial necessarily implies that $x = y = m = I$. Thus $l = u = 1$ since θ is an isomorphism on L and on U. So $\mathrm{Ker}(\theta) \leq W$, and $\theta(W)$ is the group of monomial matrices. Thus, we have shown that $\mathrm{Ker}(\theta) \leq W$. Therefore, $\mathrm{Ker}(\theta)$ is central.

We prove now a result which is valid for a general commutative ring R. First, we introduce a notion and a notation in $St_n(R)$.
A *Steinberg symbol* is an element of $St_n(R)$ of the form

$$h_{ij}(uv)h_{ij}(u)^{-1}h_{ij}(v)^{-1} \ , \ i \neq j$$

where u, v are units. Note that if we take $u = v = -1$, then $w_{ij}(1)^4$ is a symbol for any $i \neq j$. The symbols have remarkable properties when $n \geq 3$.
For instance, $h_{ij}(uv)h_{ij}(u)^{-1}h_{ij}(v)^{-1} = [h_{ik}(u), h_{ij}(v)]$ for any k different from i, j.
This immediately makes it clear that since the symbol is a central element, it is fixed under conjugation and therefore, it is independent of the choice of the distinct indices i, j, k. One suppresses the h_{ij}'s and writes $\{u, v\}$ for the symbol. Thus, it is obvious that the symbol is skew-symmetric and bilinear.

Lemma 4.
For any commutative ring R, consider the kernel C of the homomorphism from $St_n(R)$ onto $E_n(R)$. Then the central subgroup $C \cap W$ of $St_n(R)$ is generated by Steinberg symbols.
Proof
The subgroup H generated by $h_{ij}(u)$ is normal in W. In W/H, one has relations $w_{ij}(u) = w_{ij}(1)$ for every unit u. One can call this common class simply as w_{ij}. If $x = w_{i_1,j_1}(u_1) \cdots w_{il jl}(u_l) \in C \cap W$, one has $x \equiv w_{i_1,j_1} \cdots w_{i_l j_l} \bmod H$. One can use the conjugation formulae to push all the terms of the form w_{1r} to the beginning. Moreover, $w_{1r}^2 = 1 \bmod H$ and $w_{1r}w_{1s}w_{1r} = w_{rs}$ for $r \neq s$. Thus, we can cancel off the w_{1r}'s one or two at a time. After this is done, if there is a single w_{1r} left, it cannot map to the identity in $SL_n(R)$. Similarly, we can do with the elements of the form w_{2s} and so on to get $c \in H$.
If D denotes the subgroup generated by the symbols, then clearly one has $h_{ij}(uv) \equiv h_{ij}(v)h_{ij}(u) \equiv h_{ij}(u)h_{ij}(v) \bmod D$. Let us write c as a product of elements of the form $h_{1r}(u)^{\pm}$ which we can do again by

the conjugation relations. Then, $c \equiv h_{12}(u_1)\cdots h_{1n}(u_{n-1})$ mod D for certain units u_i.
As $h_{12}(u_1)\cdots h_{1n}(u_{n-1})$ maps to the diagonal matrix

$$diag(u_1\cdots u_{n-1}, u_1^{-1}, \cdots, u_{n-1}^{-1})$$

while c maps to the identity element, it follows that $c \in D$. This proves the lemma.

Exercises.
Let R be any commutative ring and $n \geq 3$. Prove:
(i) For $i \neq j$, the Steinberg symbol $h_{ij}(uv)h_{ij}(u)^{-1}h_{ij}(v)^{-1}$ equals the commutator $[h_{ik}(u), h_{ij}(v)]$ for any k different from i, j.
(ii) The symbol $\{u, v\}$ is skew-symmetric and bilinear in u, v.
(iii) $\{u, 1-u\} = 1$ for all units u.

The theorem on the presentation of $SL_n(\mathbf{Z}/k)$ would follow if we could compute for each prime power p^r, the subgroup $D(n, p^r)$ of $St_n(\mathbf{Z}/p^r)$ generated by the symbols. This is the contention of the following:

Lemma 5.
Let $n \geq 3$. If p is odd, then $D(n, p^r)$ is trivial.
If $p = 2$ and $r \geq 2$, then $D(n, 2^r)$ is the cyclic group generated by the symbol $\{-1, -1\}$.
Proof
The main idea of the proof is the fact that the group of units of $\mathbf{Z}/p^r$ is cyclic if p is odd; for this reason the proof works for the finite fields also (see corollary below). As $\mathbf{Z}/2^r$ is not cyclic if $r \geq 3$, the proof is in this case is slightly more cumbersome. We follow a computation due to M.R. Stein in this case.
Consider first the odd prime case. Now, an integer a is a square mod p^r if, and only if, it is a square mod p. Look at the homomorphism $u \mapsto 1-u$ from $(\mathbf{Z}/p)^* \setminus 1$ to itself. As there are exactly $(p-1)/2$ nonsquares, one of them has to map to a nonsquare; otherwise the $(p-1)/2$ squares that the nonsquares map to will together with 1 give $(p+1)/2$ squares in $(\mathbf{Z}/p)^*$. Thus, there is a unit $u \in \mathbf{Z}/p^r$ such that both u and $1-u$ are nonsquares in $\mathbf{Z}/p^r$. If λ is a generator of the cyclic group $(\mathbf{Z}/p^r)^*$, then $u = \lambda^r, 1-u = \lambda^s$ for some odd r, s. As the symbol is bilinear, we have $\{u, 1-u\} = \{\lambda, \lambda\}^{rs}$. But, we know from the exercise that for any unit v, $\{v, 1-v\} = 1$. Therefore, we get $\{\lambda, \lambda\}^{rs} = 1$. But, by skew-symmetry, $\{\lambda, \lambda\}^2 = 1$. As rs is odd, we get $\{\lambda, \lambda\} = 1$ and so $D(n, p^r) = 1$ if p is odd.
If $p = 2$, we show that $D(n, 2^r) = D(n, 4)$ for all $r \geq 2$. This will complete the group because evidently, $D(n, 4)$ is generated by $\{-1, -1\}$.

We notice for further use that $(\mathbf{Z}/2^r)^*$ is generated by the two units 5 and -1. To show that $D(n, 2^{r+1}) = D(n, 2^r)$ for all $r \geq 2$, we notice that for any integer $u = 1 + 2u_1 + \cdots + 2^r u_r < 2^{r+1}$ which is a unit mod $\mathbf{Z}/2^{r+1}$, either $u_r = 0$ or $u_r = 1$. In the second case, $u \equiv (1 + 2^r)(1 + 2u_1 + \cdots + 2^{r-1}u_{r-1}) \bmod 2^{r+1}$. Thus, by the bilinearity of the symbol, it suffices to show that $\{1 + 2^r, u\} = 1$ for any unit in $\mathbf{Z}/2^{r+1}$. Now, $1 + 2^r \equiv 5^{2^{r-2}} \bmod 2^{r+1}$ for all $r \geq 2$. Now, $\{1 + 2^r, -1\} = \{1 + 2^r, 1 + 2^r\} = \{1 + 2^r, 5^s\} = \{1 + 2^r, 5\}^{2^{r-2}}$. Thus, it suffices to prove that $\{1 + 2^r, 5\} = 1$. But, this is just $\{5^{2^{r-2}}, 5\} = \{5, 5\}^{2^{r-2}} = 1$ if $r \geq 3$. Finally, if $r = 2$, we must show that $\{5, u\} = 1$ for any unit u of $\mathbf{Z}/8$. As the units are $\pm 5, \pm 1$, and as $\{5, 1\} = 1 = \{5, -5\}$, we need only show that $\{5, -1\} = \{5, 5\}^{-1} = 1$. We leave it as an exercise.

Recall that $St_n(\mathbf{Z})$ is generated by elements x_{ij} for $i \neq j$. We also denote by w_{ij} the element $w_{ij}(1)$.

Lemma 6.
For any $n \geq 2$, let W_n denote the group generated by $w_{ij}, i \neq j$. Then, $\mathrm{Ker}(\phi_n) \cap W_n$ is a central subgroup of $St_n(\mathbf{Z})$ if $n \geq 2$.
Proof
This is easy to see; also, every element $w \in W_n$ conjugates any Steinberg generator x_{ij} to some $x_{kl}^{\pm 1}$.

Theorem 7.
For $n \geq 3$, $SL_n(\mathbf{Z})$ is generated by the $n(n-1)$ elementary matrices X_{ij} for $i \neq j$ subject to the relations

$$[X_{ij}, X_{jk}] = X_{ik} \; if \; i \neq k$$

$$[X_{ij}, X_{kl}] = I \; if \; j \neq k \; , i \neq l$$

$$(X_{12}X_{21}^{-1}X_{12})^4 = I$$

For $SL_2(\mathbf{Z})$, one has an analogous presentation by two generators X_{12}, X_{21} and two relations

$$X_{12}X_{21}^{-1}X_{12} = X_{21}^{-1}X_{12}X_{21}^{-1}$$

$$(X_{12}X_{21}^{-1}X_{12})^4 = I$$

Idea of the proof.
Let us lead to the proof in easy steps.
We shall show that

$$1 \to C_n \to St_n(\mathbf{Z}) \stackrel{\phi_n}{\to} SL_n(\mathbf{Z}) \to 1$$

is a central extension and that C_n is a cyclic group, which is generated by the element $(x_{12}x_{21}^{-1}x_{12})^4$.

This will be done in two steps:
(i) $C_n \subseteq W_n$, and hence, central,
(ii) C_n is cyclic, generated by w_{12}^4 where $w_{12} = x_{12}x_{21}^{-1}x_{12}$.
For each $n \geq 2$, there is an action of $St_n(\mathbf{Z})$ on $\mathbf{Z}^n$ on the right by means of the homomorphism $\phi_n : St_n(\mathbf{Z}) \to SL_n(\mathbf{Z})$. Define a *norm* on $\mathbf{Z}^n$ by $\parallel (a_1, \cdots, a_n) \parallel = \mid a_1 \mid + \cdots + \mid a_n \mid$. The subgroup W_n of $St_n(\mathbf{Z})$ generated by the elements w_{ij} clearly preserves the norm. As we mentioned earlier, in the absence of a Bruhat-type of decomposition for $St_n(\mathbf{Z})$, one looks for some sort of normal form for the elements of $St_n(\mathbf{Z})$. This is provided by the following lemma due to Silvester:

Lemma 7.
For any $n \geq 2$, every element in $St_n(\mathbf{Z})$ has an expression as a product $x_1 \cdots x_r w$ with $w \in W_n$ and each x_k one of the $x_{ij}^{\pm 1}$ in such a way that

$$\parallel ex_1 \parallel \leq \parallel ex_1x_2 \parallel \leq \cdots \parallel ex_1x_2 \cdots x_r \parallel$$

where $e = (0, 0, \cdots, 1)$.

Proof
The proof uses an appropriate induction hypothesis although it is somewhat laborious to carry out. The deviation from monotonicity of the sequence $\theta_i = \parallel ex_1x_2 \cdots x_i \parallel$ is measured by a pair (λ, μ) of positive integers defined as follows. If $1, \theta_1, \cdots, \theta_r$ is monotonic i.e., if $1 = \theta_0 \leq \theta_1 \leq \cdots\cdots \leq \theta_r$, set $\lambda = \mu = 1$. If the sequence is not monotonic, look at those $i \geq 0$ for which $\theta_i > \theta_{i+1}$ and set λ to be the maximum value of θ_i. Of course, λ could equal θ_i for several i, and one sets μ to be the maximum i for which $\lambda = \theta_i$. One can order the pairs (λ, μ) lexicographically as though they were two-digit numbers. With this set-up, the proof of Silvester's lemma proceeds by showing that each word $x_1 \cdots x_r w$ with $(\lambda, \mu) > (1, 1)$ can be altered by the Steinberg relations so that (λ, μ) is decreased. This is done as follows. Now, $\lambda = \theta_\mu > \theta_{\mu+1}$ since $(\lambda, \mu) > (1, 1)$. Obviously, $\mu \neq 0$. We may assume, by renaming the Steinberg generators that $x_\mu = x_{12}^{\pm}$. Moreover, if $x_\mu = x_{12}^{-1}$, one could conjugate each x_l by w_{12}, replace e by the vector ew_{12}^{-1} and w by $w_{12}w$ so that $x_\mu = x_{21}$. So, we may assume that $x_\mu = x_{12}$.
Write $ex_1 \cdots x_\mu = (a, b, c, \cdots) \in \mathbf{Z}^n$. Hence $ex_1 \cdots x_{\mu-1} = (a, b - a, c, \cdots)$. Thus, $x_{\mu-1} \leq x_\mu$ can be rephrased as $\mid b - a \mid \leq \mid b \mid$.
Equivalently, $\mid a \mid \leq 2 \mid b \mid$, and $ab > 0$ unless $a = 0$.
Let $x_{\mu+1} = x_{ij}^{\pm}$. We shall argue depending on the various choices of i, j.

We outline the proof in some cases and leave the other cases which can be dealt with on the same lines.
First, if $x_{\mu+1} = x_\mu = x_{12}$, then $(a, b, c, \cdots)x_{\mu+1} = (a, b + a, c, \cdots)$ and thus $\mid b - a \mid \leq \mid b \mid > \mid b + a \mid$, an impossibility.
If $x_{\mu+1} = x_{12}^{-1}$, one can simply cancel $x_\mu x_{\mu+1}$ and this reduces (λ, μ).
If $x_{\mu+1} = x_{1i}^{\pm}$ with $i \geq 3$, then we may assume $i = 3$; and so

$$(a, b, c, \cdots)x_{\mu+1} = (a, b, \pm a + c, \cdots) \ , |c| > | \pm a + c|.$$

Replace $x_\mu x_{\mu+1} = x_{12}x_{13}^{\pm}$ by $x_{13}^{\pm}x_{12}$. Observe that the transformation

$$(a, b - a, c, \cdots) \overset{x_\mu}{\mapsto} (a, b, c, \cdots) \overset{x_{\mu+1}}{\mapsto} (a, b, \pm a + c, \cdots)$$

becomes

$$(a, b - a, c, \cdots) \mapsto (a, b - a, \pm a + c, \cdots) \mapsto (a, b, \pm a + c, \cdots).$$

This means that all θ_i are unchanged excepting

$$\theta_\mu = \parallel (a, b, c, \cdots) \parallel$$

which becomes

$$\theta'_\mu = \parallel (a, b - a, \pm a + c, \cdots) \parallel .$$

As $\mid c \mid > \mid \pm a + c \mid$, we have $\theta_{\mu-1} > \theta'_\mu$ and so the pair (λ', μ') associated with the new sequence is less than (λ, μ).
If $x_{\mu+1} = x_{ij}^{\pm}$ with $i, j > 2$, the proof is the same as the above case.
The other cases can be worked out on the same lines.

We continue with the proof of the main theorem.
Using the lemma, let us show by induction on n that $C_n \subseteq W_n$.
In this set-up, the inclusion $\theta_{n-1} : St_{n-1}(\mathbf{Z}) \subset St_n(\mathbf{Z})$ corresponds to the *left hand upper corner* inclusion; $SL_{n-1}(\mathbf{Z}) \subset SL_n(\mathbf{Z})$. If $c \in C_n$, let us write $c = x_1 \cdots x_r w$ as in the lemma. Then,

$$1 \leq \parallel ex_1 \parallel \leq \parallel ex_1x_2 \parallel \leq \cdots \leq \parallel ex_1x_2 \cdots x_r w \parallel = \parallel e \parallel = 1$$

and so, equality holds everywhere. Inductively, it follows that each x_i leaves e fixed, and since $\phi_n(x_1 \cdots x_r w) = 1$, w leaves e fixed too. Thus none of the x_k's can be $x_{nj}^{\pm 1}$ for some j. Using the Steinberg relations, one can push all the factors of the form x_{in} to the left and write $x_1 \cdots x_r = xy$ where x is a product of factors of the form $x_{in}^{\pm 1}$ for some i, and y is a product of the other types of Steinberg generators. Thus, $\phi_n(x)$ is of the form $\begin{pmatrix} I_{n-1} & * \\ 0 & 1 \end{pmatrix}$ while $\phi_n(yw)$ is of the form $\begin{pmatrix} * & 0 \\ 0 & 1 \end{pmatrix}$. Since $I_n = \phi_n(xyw)$, we must have separately $\phi_n(x) = I_n = \phi_n(yw)$. But, it is clear from the definition of ϕ_n that then $x = 1$. We note also that $y \in \theta_{n-1}(St_{n-1}(\mathbf{Z}))$. Further, $\phi_n(w) = \phi_n(z)$ for some

$z \in \theta_{n-1}(W_{n-1})$; so, we can write $w = zt$ for some $t \in W_n \cap C_n$.Thus, the element $yz \in \theta_{n-1}(W_{n-1})$ by the induction hypothesis. Therefore, $c = xyw = yw = yzt \in W_n$. This proves step I i.e., that $C_n \leq W_n$ and is central.
Finally, we have to show that C_n is cyclic, and generated by the element w_{12}^4. For $n = 2$, this is clear since $w_{12}w_{21} = Id$ and so W_n is generated by w_{12}; as $\phi_n(w_{12})$ has order 4, $C_n =< w_{12}^4 >$.
For $n \geq 3$, one considers the subgroup H of W_n generated by w_{ij}^2 for $i \neq j$. We first show that $C_n \subseteq H$. Let $c \in C_n \subseteq W_n$. We write $c = w_{i_1,j_1} \cdots w_{i_r,j_r} h$ where $h \in H$. Now $I = \phi_n(c)$, $\phi_n(h)$ is a diagonal matrix and $\phi_n(w_{ij})$ is a permutation matrix corresponding to the transposition (i, j), we must have $c = h$. But, each w_{ij} can be written in terms of $w_{12}, w_{13}, \cdots w_{1n}$. Hence c is conjugate in H to $w_{12}^{2u_2} \cdots w_{1n}^{2u_n}$ for some integers u_i. This gives

$$I_n = \phi_n(c) = \begin{pmatrix} (-1)^{\sum u_i} & 0 & \cdots & 0 \\ 0 & (-1)^{u_2} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & (-1)^{u_n} \end{pmatrix}$$

Hence u_i are all even. So, C_n is generated by the 4th powers of w_{ij}. Moreover, obviously $w_{ij}^4 \in C_n$ for all $i \neq j$. As C_n is central, and as $w_{1j}w_{1k}w_{1j}^{-1} = w_{jk}^{-1}$, we have $w_{ij}^4 = w_{kl}^4$ for all $i \neq j$, $k \neq l$. Thus, $C_n =< w_{12}^4 >$ where $w_{12} = x_{12}x_{21}^{-1}x_{12}$. The proof of the theorem is complete.

Remarks.
The above central extension

$$1 \to C_n \to St_n(\mathbf{Z}) \to SL_n(\mathbf{Z}) \to 1$$

can be shown to be the universal central extension.
Moreover, it can be shown that w_{12}^4 has order 2 or infinity in $St_n(\mathbf{Z})$ according as $n \geq 3$ or $n = 2$.

9. Application : CSP for $SL(n, \mathbf{Z}); n > 2$

We use the above presentations to prove the CSP for $SL_n(\mathbf{Z})$ for $n \geq 3$. In fact, we prove the sharper result that infinite normal subgroups are automatically of finite index and are, in fact, congruence subgroups. The proof uses some matrix computations due to J.L.Brenner.

Lemma.
Let $g \in GL_n(\mathbf{Z})$ with $n \geq 3$. Let m denote the greatest common divisor

of all g_{ij} and $g_{ii} - g_{jj}, i \neq j$. Then, the normal subgroup $N(g)$ of $GL_n(\mathbf{Z})$ generated by g contains X_{12}^m.

Proof.
If we show that $N(g)$ contains $X_{12}^{m_i}$ for all i where m_i is the GCD of all the nondiagonal entries of the i-th column, it follows that $N(g)$ contains X_{12}^h where h is the GCD of all the nondiagonal entries of g. Moreover, for a permutation $\sigma \in S_n$, if $w \in GL_n(\mathbf{Z})$ is the permutation matrix $w_{ij} \neq 0 \Leftrightarrow j = \sigma(i)$, then $(wgw^{-1})_{ij} = g_{\sigma(i),\sigma(j)}$. So, conjugation by a permutation matrix allows us to permute the nondiagonal entries of any two columns. Further, for $i \neq j$, the conjugate $X_{ij}gX_{ij}^{-1}$ has (i,j)-th entry $g_{ij} - g_{ji} + g_{jj} - g_{ii}$. Therefore, to prove the lemma, it suffices to show that $N(g)$ contains X_{12}^d where d is the GCD of all the nondiagonal entries of the first column.

Observation I *Each $g \in GL_n(\mathbf{Z})$ has a conjugate whose first column is $(g_{11}, s, 0, \cdots, 0)$ for some s dividing d.*
To see this, write $d = \sum_{i=2}^n t_i g_{i1}$. If r is the GCD of the t_i's, the vector $(\frac{t_2}{r}, \cdots, \frac{t_n}{r})$ is the first row of some $x \in GL_{n-1}(\mathbf{Z})$. Then $h = \begin{pmatrix} 1 & 0 \\ 0 & x \end{pmatrix}$ is in $GL_n(\mathbf{Z})$ and h^{-1} has a similar form. So, hgh^{-1} has the first column $(g_{11}, d/r, d_3, \cdots, d_n)$ where d_i are integral combinations of $g_{i1}, i \geq 2$. Hence a further conjugation by the matrix $X_{32}^{-rd_3/d} X_{42}^{-rd_4/d} \cdots X_{n2}^{-rd_n/d}$ gives a conjugate ygy^{-1} whose first column is $(g_{11}, d/r, 0, \cdots, 0)$.

Observation II *If $g_{i1} = 0$ for $i > 2$, then the commutator $u = [X_{12}, g] = X_{12}gX_{12}^{-1}g^{-1}$ has the last $n-2$ rows as that of the identity matrix.*

Observation III *If $u \in SL_n(\mathbf{Z})$ has the last $n-2$ rows as that of the identity matrix, then $X_{12}^{(u_{12}, 1-u_{11})} \in N(u)$.*
This is due to the fact that u^{-1} also has the last $n-2$ rows as that of the identity matrix and so $u^{-1}X_{23}^{-1}uX_{23}$ differs from the identity matrix only in the $(1,3)$-th and $(2,3)$-th entries. These are, respectively, u_{12} and $1 - u_{11}$. Arguing as in the proof of observation I, one shows that this is a conjugate of $X_{13}^{(u_{12}, 1-u_{11})}$ and, hence, of $X_{12}^{(u_{12}, 1-u_{11})}$.

To complete the proof of the lemma, we compute the values u_{12} and $1 - u_{11}$ where $u = [X_{12}, g]$ and g has its last $n-2$ rows the same as the identity matrix. Then, $u_{12} = 1 \mp (g_{11} + g_{21})g_{11}$ and $u_{11} = 1 \pm (g_{11} + g_{21})g_{21}$ where the signs are according as $\det g = \pm 1$. As $X_{12}^{u_{12}}, X_{12}^{1-u_{11}} \in N(u) \subset N(g)$, we get $X_{12}^{g_{21}} \in N(g)$. This completes

the proof of the lemma since g_{21} is d/r where d is the GCD of the nondiagonal entries of the first column of the original g.

Theorem.
Let $n \geq 3$ and let N be a normal subgroup of $SL_n(\mathbf{Z})$, not contained in $\{\pm I\}$. Then, there exists a unique integer $k \geq 0$ such that

$$\Gamma(k) \leq N \leq \Gamma(k)^*$$

where $\Gamma(k)^*$ denote the normal subgroup of $SL_n(\mathbf{Z})$ consisting of all the matrices congruent to a scalar matrix modulo k. Here, $\Gamma(0)$ stands for the trivial group and $\Gamma(0)^*$ stands for the scalars in $SL(n, \mathbf{Z})$.
In particular, any normal subgroup of $SL_n(\mathbf{Z})$ with $n \geq 3$ is either a finite, central subgroup, or is a congruence subgroup.
Proof. Let k be the G.C.D of $g_{ij}, g_{ii} - g_{jj}; i \neq j$ as g runs through the elements of N. Then, $N \leq \Gamma(k)^*$ by definition. Now, $X_{12}^k \in N$. Therefore, if $E(k)$ denotes the normal subgroup of $SL_n(\mathbf{Z})$ generated by X_{12}^k, then $E(k) \leq N \leq \Gamma(k)^*$. It suffices to prove that $E(k) = \Gamma(k)$. If $k = 0$, this is clear. So, let us assume $k > 0$. Now, the inclusion $E(k) \leq \Gamma(k)$ induces a homomorphism $f : SL_n(\mathbf{Z})/E(k) \to SL_n(\mathbf{Z})/\Gamma(k) = SL_n(\mathbf{Z}/k\mathbf{Z})$. Since $X_{12}^k \in E(k)$, we have a homomorphism in the opposite direction from the above-mentioned presentation for $SL_n(\mathbf{Z}/k\mathbf{Z})$. Thus, f is an isomorphism which means that $E(k) = \Gamma(k)$. This completes the proof.