Matrices Elementary, My Dear Homs

B.Sury
Stat-Math Unit
Indian Statistical Institute
Bangalore Centre.

ATM Workshop on Classical groups and K-theory
IISER Pune
December 2013

(Ravi) Rao and (van der) Kallen transformations have taken the place of Row and column transformations

... Yrus ...

Introduction

The definition of the K-groups over rings, led to interpretations of geometric and arithmetic questions and to their eventual solutions also at times. A typical example is the famous result that algebraic vector bundles over affine spaces are trivial. For the rings of integers in algebraic number fields, these K-groups carry information like the unit group, the class group etc. Quillen introduced higher K-groups which match Milnor's K groups (only) for $n \leq 2$. These gave a beautiful bridge between topology and group theory. We shall not discuss the fascinating subject of Quillen's K groups in these lectures.

Apart from Hahn-O'Meara's book, the two other basic references we use are Bass's and Milnor's classic texts on algebraic K-Theory. Without further ado, we introduce the most relevant object of study.

Definition 0.1 Let $e_{ij}(\lambda)$, where $i \neq j$, and $\lambda \in R$, denote the matrix $\mathrm{I}_n + \lambda \mathrm{E}_{ij}$, where E_{ij} is the matrix with 1 in the (i, j)-th position and zeros elsewhere. The subgroup of $\mathrm{GL}_n(R)$ generated by $e_{ij}(\lambda)$, $\lambda \in R$, is denoted by $\mathrm{E}_n(R)$. It is called the *elementary subgroup* of $\mathrm{GL}_n(R)$, and e_{ij}'s are called its *elementary generators.*

The elementary subgroup $\mathrm{E}_n(R)$ plays a crucial role for the development of classical algebraic K-theory. It turns out that it is not always equal to the special linear group $\mathrm{SL}_n(R)$. But, if R is a field, then the groups coincide. The matrices $e_{ij}(\lambda)$'s are linear operators on row and column vectors. Indeed, observe:
Multiplication of a matrix by $e_{ij}(\lambda)$ on the right, is the elementary column operation of adding λ times the i'th column to the j'th column.
The multiplication by $e_{ij}(\lambda)$ on the left can be similarly described - it adds to the i-th row, λ times the j-th row. The set $\mathrm{GL}_n(R)/E_n(R)$ measures the obstruction to reducing an invertible matrix to the identity matrix by applying these linear operators. Thus, the question of normality of $\mathrm{E}_n(R)$ in $\mathrm{SL}_n(R)$ or $\mathrm{GL}_n(R)$ is of interest. The Russian mathematician Andrei Suslin proved that for a commutative ring R, $\mathrm{E}_n(R)$ is a normal subgroup of $\mathrm{GL}_n(R)$ for $n \geq 3$. Interestingly, the result is not true for the case $n = 2$. We shall discuss some counter examples. We shall discuss the proof of Suslin's theorem. For $n \geq 3$, Anthony Bak proved that the group $\mathrm{GL}_n(R)/E_n(R)(n \geq 3)$ is a solvable group for any commutative ring R.

1 Transvections and Dilations

In this section, we consider R to be a general **possibly noncommutative** *ring with identity and, we look at right R-modules. Ring homomorphisms for us will always mean that the identities are preserved.*
For a (right) R-module M, the abelian group

$$M^* := Hom(M, R)$$

of all (right) R-module homomorphisms, is naturally a left R-module. It is called the *dual module* to M.
If we look at the right R-module M^{**}, there is a natural right R-module homomorphism

$$\theta : M \to M^{**} \ ;$$

$$x \mapsto (\phi : M^* \to R; \rho \mapsto \rho(x)).$$

Let $Z(R)$ denote the center of R and R^* denote the unit group of R (unfortunate, but standard notation!), then one may look at any $\lambda \in Z(R)^* := Z(R) \cap R^*$. One denotes by $End(M)$ is a ring under composition and $GL(M)$ is the group of R-linear automorphism. Any central unit λ defines an element of the general linear group $GL(M)$ of all R-linear automorphisms of M by:

$$x \mapsto x\lambda.$$

Such λ's define the "scalar" subgroup of $GL(M)$, an abelian central subgroup and the quotient group is called the projective linear group $PGL(M)$. If M is a faithful R-module, then the scalar subgroup is isomorphic to $Z(R)^*$.

When R is commutative and M is free:
One may then define the notion of a determinant; it is a homomorphism from $GL(M)$ to R^* and, the kernel is the special linear group $SL(M)$.
When M is R^n, $GL(M)$ can be identified with invertible $n \times n$ matrices and $SL(M)$ can be identified with those matrices which have determinant 1.

If we consider an arbitrary R module M, then there is a way to define a subgroup of $GL(M)$ generalizing the elementary subgroup - and coinciding with $E_n(R)$ when M is the free module R^n. Here is the definition.

Definition. Let $v \in M$, $\rho \in M^*$. Define

$$\tau_{v,\rho} : M \to M \ ; \ x \mapsto x + v\rho(x).$$

Clearly, this is a (right) R-module homomorphism.
If $\rho(v) = 0$, then $\tau_{v,\rho} \in GL(M)$; its inverse is $\tau_{v,-\rho}$.

Such an element $\tau_{v,\rho}$ (if nontrivial) is called a *transvection.*
If $\rho(v) \neq 0$ and if $\tau_{v,\rho}$ is invertible, then it is called a *dilation.*
For instance, let v be a *unimodular* vector; that is, there is a submodule N of M for which $M = vR \oplus N$. Then, choosing some $s \in R^*; s \neq 1$ and taking

$$\rho : M \to R\ ;\ N \to 0\ ;\ v \mapsto s - 1$$

defines a dilation $\tau_{v,\rho}$.

If R is commutative, $M = R^n$ (considered as column vectors), look at the following two examples:
(a) $v = e_i$ and $\rho = te_j^*$ for some $t \in R$ and some $i \neq j$, then with respect to the canonical ordered basis, the transvection $\tau_{v,\rho}$ has the matrix representation $I_n + tE_{ij}$.
(b) $v = e_i$, $\rho'(e_i) = t - 1$ for some $t(\neq 1) \in R^*$ and $\rho'(e_j) = 0$ for all $j \neq i$, then the dilation $\tau_{v,\rho'}$ has the matrix diag $(1, 1, \cdots, t, 1, \cdots, 1)$ where t is the i-th diagonal entry.

Properties of transvections and dilations.

(i) $\tau_{v,\rho_1}\tau_{v,\rho_2} = \tau_{v,\rho_1+\rho_2}$ if $\rho_1(v) = 0$;
(ii) $\tau_{v_1,\rho}\tau_{v_2,\rho} = \tau_{v_1+v_2,\rho}$ if $\rho(v_2) = 0$;
(iii) $\tau_{vr,\rho} = \tau_{v,r\rho}$;
(iv) $g\tau_{v,\rho}g^{-1} = \tau_{gv,\rho g^{-1}} \forall g \in GL(M)$;
(v) τ_{v_1,ρ_1} and τ_{v_2,ρ_2} commute if $\rho_1(v_2) = 0 = \rho_2(v_1)$.
Proof. Tutorial sessions!

Note, by (iv), that conjugates of transvections are transvections and conjugates of dilations are dilations.

1.1 "Watson" Elementary group?

If M is a free R-module with a basis $B = \{v_1, \cdots, v_n\}$, and $\{\rho_1, \cdots, \rho_n\}$ is the dual basis of M^*, then consider the transvections of the form $\tau_{v_ir,\rho_j} (= \tau_{v_i,r\rho_j})$ where $i \neq j$ and $r \in R$. Such transformations are called *elementary transvections.* The subgroup of $GL(M) \cong GL_n(R)$ generated by the elementary transvections is said to be the elementary subgroup and is denoted by $E_B(M)$. In the matrix avataar, this subgroup coincides $E_n(R)$, the subgroup of elementary matrices in $GL_n(R)$ as defined above. Namely, it is the subgroup of $GL_n(R)$ generated by $e_{ij}(r) = I_n + rE_{ij}$ as r varies over R and $i \neq j$ vary in $\{1, 2, \cdots, n\}$. Simply identify $\tau_{v_i,r\rho_j}$ with $e_{ij}(r)$.

We will study the elementary subgroup in detail. But, we begin by observing some "elementary" properties:

$$e_{ij}(r)e_{ij}(s) = e_{ij}(r+s).$$

$$[e_{ij}(r), e_{jk}(s)] = e_{ik}(rs) \ if \ i,j,k \ distinct.$$

$$[e_{ij}(r), e_{kl}(s)] = 1 \ if \ j \neq k \ and \ i \neq l.$$

$$diag(t_1, \cdots, t_n)e_{ij}(s)diag(t_1, \cdots, t_n)^{-1} = e_{ij}(t_i s t_j^{-1}) \ if \ t_i \in R^*.$$

We also define for each $t \in R^*$, and each $i \neq j$, the element $w_{ij}(t) \in E_n(R)$ as :

$$w_{ij}(t) = e_{ij}(t)e_{ji}(-t^{-1})e_{ij}(t).$$

Observe that if a matrix h is multiplied *on the left* by $w_{ij}(1)$, the i-th and j-th rows of h get interchanged with the j-th row also changing in sign.
Similarly, right multiplication by $w_{ij}(1)$ has an analogous effect on the columns.

2 Properties of $\mathrm{E}_n(R)$ for a Euclidean ring R

In this section, we shall observe that the usual Euclidean algorithm (done in a noncommutative set-up!) gives the basic properties of the elementary subgroup.
Firstly, note that a noncommutative ring R is said to be *Euclidean*, if there exists a size function $\delta : R - \{0\} \to \mathbf{Z}^{\geq 0}$ such that for each $a, b \in R$ with $b \neq 0$, there exist $q_1, r_1, q_2, r_2 \in R$ such that

$$a = q_1 b + r_1 = b q_2 + r_2$$

with $r_i = 0$ or $\delta(r_i) < \delta(b)$ for $i = 1, 2$.
In particular, $\delta(b) \neq 0$ for each $b \in R^*$.

Before stating the main result, we introduce a natural embedding of $GL_n(R)$ as a subgroup of $GL_{n+1}(R)$; the former group sits as the upper left block of an $(n+1) \times (n+1)$ matrix whose last row and last column are the unit vector $(0, 0, \cdots, 0, 1)$.

Proposition 1.
If R is a Euclidean ring, then $GL_n(R) = GL_1(R)E_n(R)$ for any $n \geq 1$. In particular, $E_n(R)$ is a normal subgroup. Further, if R is also commutative, then ($SL_n(R)$ makes sense and) $E_n(R) = SL_n(R)$.
Proof.
Notice that $GL_1(R) \cong R^*$, when thought of as a subgroup of $GL_n(R)$, simply consists of the $n \times n$ diagonal matrices with first entry a unit and other entries equal to 1.

This subgroup $GL_1(R)$ conjugates each $e_{ij}(r)$ into another elementary matrix as observed above. For the rest of the proof, we just inductively show that for any $g \in GL_n(R)$, there are elements $x, y \in E_n(R)$ so that $xgy \in GL_{n-1}(R)$ (where the latter is regarded as a subgroup of $GL_n(R)$ under the upper-left corner embedding described above.

The proof is merely carrying out the Euclidean algorithm on the left and on the right. Indeed, corresponding to the given element g, consider the set

$$\Omega := \{xgy : x, y \in E_n(R)\}$$

Choose an element of Ω for which a non-zero entry b has the smallest size under δ among all non-zero entries of all elements of Ω. Suppose xgy has b at the (i,j)-th place. From this element xgy of Ω, we will get an element of Ω for which this entry s is at the (n,n)-th place, in the following manner.

Recall the matrices $w_{ij}(1)$ in $E_n(R)$ which were defined above; we write w_{ij} instead of $w_{ij}(1)$.

Multiplying xgy on the left by w_{in} (if $i \neq n$) and by w_{jn} on the right (if $j \neq n$) gives a matrix $h \in \Omega$ such that $h_{nn} = b$.

We claim that h can further be changed to an element of Ω which has no non-zero entries in the last row and the last column excepting the (n,n)-th entry. Now, if $0 \neq a \in R$ and $h_{in} = a$, then write

$$a = qb + r$$

with $r = 0$ or $\delta(r) < \delta(b)$.

Therefore, $e_{in}(-q)h$ has i-th row equal to $R_i - qR_n$ where R_i is the i-th row of h. Hence (i,n)-th entry of $e_{in}(-q)h$ is $a - qb = r$. This means by the choice of b that $r = 0$. In this manner, the last column can be reduced to 0 excepting the (n,n)-th entry. Note that this new matrix z continues to have the (n,n)-th entry equal to b.

Similarly, if $0 \neq c \in R$ is so that $c = z_{nj}$, then writing

$$c = bq_2 + r_2$$

with $r_2 = 0$ or $\delta(r_2) < \delta(b)$, and noting that the matrix $ze_{nj}(-q_2) \in \Omega$ has (n,j)-th entry r_2, it follows that $r_2 = 0$. In this manner, we arrive at an element $w \in \Omega$ such that $w_{kn} = w_{nl} = 0$ for all $k, l < n$ and $w_{nn} = b$. Since w is

invertible, $b \in R^*$.
Finally, $h_{1n}(b)w$ is in $GL_{n-1}(R)$.
This completes the proof of the first two assertions.
To deduce the last assertion, note that evidently $E_n(R) \leq SL_n(R)$ for any commutative ring R. In case of such a ring being Euclidean also, we have $SL_n(R) \leq GL_1(R)E_n(R)$ which gives on comparing determinants that $SL_n(R) \leq E_n(R)$. The proof is complete.

3 Properties of $\mathrm{E}_n(R)$ and $E_n(R, I)$

When R is an arbitrary ring with unity, we discuss some useful properties of the elementary subgroup, and use these to prove the normality of $\mathrm{E}_n(R)$ in $GL_n(R)$ for $n \geq 3$ when R is commutative. We start with some examples of elements of the elementary subgroup. Before that, we define a certain type of normal subgroup $E_n(R, I)$ of $E_n(R)$ which will be studied alongside $E_n(R)$.

Definition. For a two-sided ideal I of R, the group $E_n(R, I)$ is defined to be the *normal subgroup of* $E_n(R)$ generated by all $e_{ij}(t)$ as t varies in I and $i \neq j$.

Note that $E_n(R, I) \leq GL_n(R, I) := Ker(GL_n(R) \to GL_n(R/I)$. The latter kernel is known as a principal congruence subgroup and will be studied in a later section.

Example 3.1

(*i*) The matrix

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \in \mathrm{E}_2(R)$$

Note that it has finite order.

(*ii*) For each $a \in R$, the matrix
$\mathrm{E}(a) = \begin{pmatrix} a & 1 \\ -1 & 0 \end{pmatrix} = e_{12}(1-a)e_{21}(-1)e_{12}(1) \in \mathrm{E}_2(R)$.

(*iii*) Upper and lower triangular $n \times n$ matrices with 1's on the diagonal, are in $\mathrm{E}_n(R)$.

(*iv*) For $g \in M_n(R)$, the matrices $\begin{pmatrix} \mathrm{I}_n & g \\ 0 & \mathrm{I}_n \end{pmatrix}$ and $\begin{pmatrix} \mathrm{I}_n & 0 \\ g & \mathrm{I}_n \end{pmatrix}$ are in $\mathrm{E}_{2n}(R)$.

(iv) For any unit $u \in R^*$ and any $i \neq j$, the matrix $w_{ij}(u)$ which has u in the (i,j)-th place, u^{-1} in the (j,i)-th place, 1 in the diagonal places (k,k) for $k \neq i,j$ and 0's at all other places, is in $E_n(R)$. Indeed

$$w_{ij}(u) = e_{ij}(u)e_{ji}(-u^{-1})e_{ij}(u).$$

For any unit u and any $i \neq j$, the diagonal matrix $h_{ij}(u)$ which has u and u^{-1} as the i-th and the j-th diagonal entries and 1's as other diagonal entries, is in $E_n(R)$. Indeed

$$h_{ij}(u) = w_{ij}(u)w_{ij}(-1).$$

Now, we list a few properties of $\mathrm{E}_n(R)$ and $E_n(R,I)$ for a two-sided ideal I.

(i) $\{I_n\} < \mathrm{E}_n(R) \leq \mathrm{GL}_n(R)$ for $n \geq 2$.
If R is *commutative*, we also have $E_n(R) \leq SL_n(R)$.
For $m \geq 1, n > 1$, the embedding $\mathrm{GL}_n(R) \to \mathrm{GL}_{n+m}(R)$, given by $\alpha \mapsto \begin{pmatrix} \alpha & 0 \\ 0 & \mathrm{I}_m \end{pmatrix}$, induces embeddings $\mathrm{E}_n(R) \to \mathrm{E}_{n+m}(R)$ and, if R is commutative, $\mathrm{SL}_n(R) \to \mathrm{SL}_{n+m}(R)$.
This allows us to *define* the groups

$$\mathrm{GL}(R) = \bigcup_{n=1}^{\infty} \mathrm{GL}_n(R), \quad \mathrm{SL}(R) = \bigcup_{n=1}^{\infty} \mathrm{SL}_n(R), \quad \mathrm{E}(R) = \bigcup_{n=1}^{\infty} \mathrm{E}_n(R)$$

where we talk of SL in the case when R is commutative.

(iii) $e_{ij}(x+y) = e_{ij}(x)e_{ij}(y) \;\; \forall\, x, y \in R$ when $i \neq j$.

(iv) $\forall\, x, y \in R$,
(a) $[e_{ij}(x), e_{kl}(y)] = 1$ if $j \neq k$, $i \neq l$.
(b) $[e_{ij}(x), e_{jk}(y)] = e_{ik}(xy)$ if $i \neq j$, $j \neq k$, $i \neq k$.
For $n \geq 3$, by using commutator formula one can deduce that $\mathrm{E}_n(R)$ is generated by the set $\{e_{1j}(\lambda), e_{i1}(\mu) \mid 1 \leq i.j \leq n,\ \lambda, \mu \in R\}$.
Note also the identities involving the elements $w_{ij}(u), h_{ij}(u)$ in $E_n(R)$ where u is a unit:

$$w_{ij}(u)e_{ji}(t)w_{ij}(u)^{-1} = e_{ij}(-utu);$$

$$h_{ij}(u)e_{ij}(t)h_{ij}(u)^{-1} = e_{ij}(utu).$$

(v) The subgroup $\mathrm{E}_2(R)$ is generated by the set $\{\mathrm{E}(a) \mid a \in R\}$, where

$$\mathrm{E}(a) = \begin{pmatrix} a & 1 \\ -1 & 0 \end{pmatrix}.$$

Indeed, as mentioned above, one can check that $\mathrm{E}(a) = e_{12}(1-a)e_{21}(-1)e_{12}(1)$.
Moreover, $e_{12}(a) = \mathrm{E}(-a)\mathrm{E}(0)^{-1}$ and $e_{21}(a) = \mathrm{E}(0)^{-1}\mathrm{E}(a)$.

(vi) (Perfectness): $\mathrm{E}_n(R)$ is a perfect group if $n > 2$.
In particular, $[\mathrm{GL}(R), \mathrm{GL}(R)] = \mathrm{E}(R)$.
On the other hand, $\mathrm{E}_2(\mathbf{F}_2)$ and $\mathrm{E}_2(\mathbf{F}_3)$ are not perfect.
Note that if R has a maximal ideal $\mathbf{m}$ such that $R/\mathbf{m} \cong \mathbf{F}_2$ or $\mathbf{F}_3$, then $E_2(R)$ is not perfect for the above reason.
It is easy to deduce the perfectness of $\mathrm{E}_n(R)$ for $n \geq 3$ using the commutator formulas above. It follows then that $\mathrm{E}(R)$ is perfect, i.e., $[\mathrm{E}(R), \mathrm{E}(R)] = \mathrm{E}(R)$. The last assertion is now a consequence of (vi).

(viii) Let I be a two-sided ideal in the ring R. Then, the homomorphism

$$\mathrm{E}_n(R) \to \mathrm{E}_n(R/I)$$

is surjective, for all $n \geq 2$.
Indeed, the generators $e_{ij}(\overline{\lambda})$ of $\mathrm{E}_n(R/I)$ can be lifted to the generators $e_{ij}(\lambda)$ of $\mathrm{E}_n(R)$.

The next property is a beautiful one which proves that the product operation in GL coincides with that in the direct sum; it is also the key to defining K_1. The original Whitehead's Lemma is a topological assertion; the matrix version below is due to A. Suslin.

Whitehead's lemma.
Let I be a two-sided ideal. For $g \in GL_n(R)$ and $h \in GL_n(R, I)$, we have the congruences (both as right cosets and as left cosets):

$$\begin{pmatrix} g & 0 \\ 0 & h \end{pmatrix} \equiv \begin{pmatrix} gh & 0 \\ 0 & I_n \end{pmatrix} \equiv \begin{pmatrix} hg & 0 \\ 0 & I_n \end{pmatrix} \equiv (mod\ E_{2n}(R, I).$$

Further, we have the congruence

$$\begin{pmatrix} g & 0 \\ 0 & h \end{pmatrix} \equiv \begin{pmatrix} 0 & g \\ -h & 0 \end{pmatrix} (mod\ E_{2n}(R).$$

Proof.
We prove it for left cosets; the proof for right cosets is similar.
The last congruence is easy to see because

$$\begin{pmatrix} 0 & g \\ -h & 0 \end{pmatrix} = \begin{pmatrix} g & 0 \\ 0 & h \end{pmatrix} \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$$

and the last matrix is in $E_{2n}(R)$.
Let us prove the first congruences now.

We have

$$\begin{pmatrix} h & 0 \\ 0 & h^{-1} \end{pmatrix} = \begin{pmatrix} I_n & h - I_n \\ 0 & I_n \end{pmatrix} \begin{pmatrix} I_n & 0 \\ I_n & I_n \end{pmatrix} \begin{pmatrix} I_n & h^{-1} - I_n \\ 0 & I_n \end{pmatrix} \begin{pmatrix} I_n & 0 \\ -h & I_n \end{pmatrix}$$

$$= \begin{pmatrix} I_n & h - I_n \\ 0 & I_n \end{pmatrix} \left(\begin{pmatrix} I_n & 0 \\ I_n & I_n \end{pmatrix} \begin{pmatrix} I_n & h^{-1} - I_n \\ 0 & I_n \end{pmatrix} \begin{pmatrix} I_n & 0 \\ I_n & I_n \end{pmatrix}^{-1} \right) \begin{pmatrix} I_n & 0 \\ I_n - h & I_n \end{pmatrix}.$$

This is in $E_{2n}(R, I)$ as the entries of $I_n - h$ and of $h^{-1} - I_n$ are in I. Therefore,

$$\begin{pmatrix} gh & 0 \\ 0 & I_n \end{pmatrix} = \begin{pmatrix} g & 0 \\ 0 & h \end{pmatrix} \begin{pmatrix} h & 0 \\ 0 & h^{-1} \end{pmatrix} \equiv \begin{pmatrix} g & 0 \\ 0 & h \end{pmatrix} (mod\ E_{2n}(R, I)).$$

We are left with showing that

$$\begin{pmatrix} hg & 0 \\ 0 & I_n \end{pmatrix} \equiv \begin{pmatrix} g & 0 \\ 0 & h \end{pmatrix} (mod\ E_{2n}(R, I)).$$

We observe that

$$\begin{pmatrix} g^{-1}h^{-1} & 0 \\ 0 & I_n \end{pmatrix} \begin{pmatrix} g & 0 \\ 0 & h \end{pmatrix}$$

$$= \begin{pmatrix} I_n & g^{-1}(I_n - h^{-1}) \\ 0 & I_n \end{pmatrix} \left(\begin{pmatrix} I_n & 0 \\ -g & I_n \end{pmatrix} \begin{pmatrix} I_n & -g^{-1}(h - I_n) \\ 0 & I_n \end{pmatrix} \begin{pmatrix} I_n & 0 \\ -g & I_n \end{pmatrix}^{-1} \right)$$

$$\begin{pmatrix} I_n & 0 \\ (h^{-1} - I_n)g & I_n \end{pmatrix}$$

$$\in E_{2n}(R, I).$$

Corollary. *For any two-sided ideal I, we have*

$$[GL_n(R), GL_n(R, I)] \leq E_{2n}(R, I).$$

In particular, $[G(R), G(R, I)] \leq E(R, I)$. Hence $E(R, I)$ is normal in $G(R)$.

Proof.

Let $g \in GL_n(R)$ and $h \in GL_n(R, I)$. From the Whitehead lemma, we have the equivalence of $\begin{pmatrix} gh & 0 \\ 0 & I_n \end{pmatrix}$ and $\begin{pmatrix} hg & 0 \\ 0 & I_n \end{pmatrix}$ modulo $E_{2n}(R, I)$. It follows that

$$\begin{pmatrix} gh(hg)^{-1} & 0 \\ 0 & I_n \end{pmatrix} \in E_{2n}(R, I).$$

Remarks.

The finite analogue that $E_n(R,I)$ is normal in $GL_n(R)$ will be proved later for commutative R when $n \geq 3$.

The analogue of the surjectivity in (viii) above is not true for $\mathrm{SL}_n(R)$ in general. Here is an example.
Let $\langle XY - ZT - 1\rangle$ denote the ideal generated by the element $XY - ZT - 1$ in the ring $\mathbb{R}[X,Y,Z,T]$. Then the homomorphism

$$\mathrm{SL}_2(\mathbb{R}[X,Y,Z,T]) \to \mathrm{SL}_2\left(\frac{\mathbb{R}[X,Y,Z,T]}{\langle XY - ZT - 1\rangle}\right)$$

is not surjective. In fact, if bar denotes the reduction modulo the above ideal, then there is no lift of the matrix

$$\begin{pmatrix} \overline{X} & \overline{Z} \\ \overline{T} & \overline{Y} \end{pmatrix} \in \mathrm{SL}_2\left(\frac{\mathbb{R}[X,Y,Z,T]}{\langle XY - ZT - 1\rangle}\right)$$

to a matrix in $\mathrm{SL}_2(\mathbb{R}[X,Y,Z,T])$.

3.1 When $n = 2$

In this section, we describe examples of P.M.Cohn which show:
(i) $\mathrm{E}_2(R)$ may not be a normal subgroup of $\mathrm{SL}_2(R)$;
(ii) $\mathrm{E}_2(\mathbb{Z}[Y]) \neq \mathrm{SL}_2(\mathbb{Z}[Y])$.
In fact, let us look at $R = k[X,Y]$ where k is afield. If we consider the matrix

$$\alpha = \begin{pmatrix} 1+XY & X^2 \\ -Y^2 & 1-XY \end{pmatrix}$$

then, it was shown by P.M.Cohn that $\alpha \in \mathrm{SL}_2(R)$ but not in $\mathrm{E}_2(R)$.
This will be discussed in the tutorials sessions.
P.M. Cohn has also proved that $\mathrm{E}_2(\mathbb{Z}[Y]) \neq \mathrm{SL}_2(\mathbb{Z}[Y])$ by proving the matrix

$$\begin{pmatrix} 1+2Y & 4 \\ -Y^2 & 1-2Y \end{pmatrix} \notin \mathrm{E}_2(\mathbb{Z}[Y]).$$

Consider the matrix $\beta = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. In the tutorials, we will deduce that $\alpha\beta\alpha^{-1} \notin \mathrm{E}_2(R)$.

Remark.
For the ring of integers O_d of an imaginary quadratic field $\mathbf{Q}(\sqrt{-d})$, P.M.Cohn also proved that $E_2(O_d) \neq SL_2(O_d)$ unless O_d is a Euclidean domain (only if $d = 1, 2, 3, 7$ or 11).

4 Normality of $\mathrm{E}_n(R)$ for $n > 2$ when R is commutative

In this section, we describe the proof of Suslin's theorem, using a variant of Whitehead's Lemma due to L.N. Vaserstein. The theorem of Suslin alluded to above is:

Theorem 4.1 (A. Suslin) *Let R be a commutative ring with identity. The elementary subgroup $\mathrm{E}_n(R)$ is normal in $\mathrm{GL}_n(R)$, for $n \geq 3$.*

Lemma 4.2 (L.N. Vaserstein) *Let $\mathrm{M}_{r,s}(R)$ denote the set of $r \times s$ matrices over R. Let $\alpha \in \mathrm{M}_{r,s}(R)$ and $\beta \in \mathrm{M}_{s,r}(R)$. If $\mathrm{I}_r + \alpha\beta \in \mathrm{GL}_r(R)$, then $\mathrm{I}_s + \beta\alpha \in \mathrm{GL}_s(R)$ and*

$$\begin{pmatrix} \mathrm{I}_r + \alpha\beta & 0 \\ 0 & (\mathrm{I}_s + \beta\alpha)^{-1} \end{pmatrix} \in \mathrm{E}_{r+s}(R).$$

Proof Note that

$$I_s - \beta(I_r + \alpha\beta)^{-1}\alpha$$

is easily verified to be the inverse of $(I_s + \beta\alpha)$; a nice way to arrive at this expression is to view the sought-for inverse in analogy with a geometric series. Hence, the invertibility of $I_r + \alpha\beta$ implies that of $I_s + \beta\alpha$ and vice versa. Moreover,

$$\begin{pmatrix} \mathrm{I}_r + \alpha\beta & 0 \\ 0 & (\mathrm{I}_s + \beta\alpha)^{-1} \end{pmatrix}$$
$$= \begin{pmatrix} \mathrm{I}_r & 0 \\ (\mathrm{I}_s + \beta\alpha)^{-1}\beta & \mathrm{I}_s \end{pmatrix} \begin{pmatrix} \mathrm{I}_r & -\alpha \\ 0 & \mathrm{I}_s \end{pmatrix} \begin{pmatrix} \mathrm{I}_r & 0 \\ -\beta & \mathrm{I}_s \end{pmatrix} \begin{pmatrix} \mathrm{I}_r & (\mathrm{I}_r + \alpha\beta)^{-1}\alpha \\ 0 & \mathrm{I}_s \end{pmatrix}.$$

The lemma follows now from the fact (which we already noted and used) that a triangular matrix with 1 in the diagonal is a product of elementary matrices.

Observation.
As a consequence of the above lemma, if $v = (v_1, \ldots, v_n)^t$ and $w = (w_1, \ldots, w_n)^t$ are two column vectors with the property that the dot product $w^t v = 0$, then the matrix $\mathrm{I}_n + vw^t$ is invertible, and 1-stably elementary, i.e.,

$$\begin{pmatrix} \mathrm{I}_n + vw^t & 0 \\ 0 & 1 \end{pmatrix} \in \mathrm{E}_{n+1}(R).$$

As a matter of fact, the above observation holds in a stronger form under a certain condition as follows:

Lemma 4.3 *Suppose $v = (v_1, \ldots, v_n)^t$ and $w = (w_1, \ldots, w_n)^t$ are two column vectors with the property that the dot product $w^t v = 0$, and suppose also that $w_i = 0$ for* **some** *$i \leq n$. Then, $I_n + vw^t \in E_n(R)$.*

Proof.

We may assume without loss of generality that $w_n = 0$; indeed, if $w_i = 0$ and $i < n$, then for

$$g = e_{in}(-1)e_{ni}(1)e_{in}(-1),$$

we have

$$g(I_n + vw^t)g^{-1} = I_n + v(0)w(0)^t$$

where $v_0 = gv$, $w(0)^t = w^t g^{-1}$.

Note that $w(0)^t v(0) = 0$ and $w(0)_n = 0$.

Therefore, let us assume that the given w has $w_n = 0$.

If we consider $w' = (w_1, \cdots, w_{n-1})^t \in R^n, v' = (v_1, \cdots, v_{n-1})^t \in R^n$, then $I_{n-1} + v'w'^t \in E_{n-1}(R)$ by the above observation since $w'^t v' = w^t v = 0$.

We note that $I_n + vw^t = \begin{pmatrix} I_{n-1} + v'w'^t & 0 \\ * & 1 \end{pmatrix}$ where the last column has 0's above the last entry 1.

Adding appropriate multiples of the last column to the other columns (which is equivalent to multiplication by elements in $E_n(R)$, we may make the last row to consist of 0's excepting the last entry. In other words, $I_n + vw^t$ is equivalent to the matrix $\begin{pmatrix} I_{n-1} + v'w'^t & 0 \\ 0 & 1 \end{pmatrix}$ which itself is in $E_n(R)$.

In the proof of Suslin's theorem, we will use the following lemma. We view elements of R^n as column vectors.

Recall that we defined a vector v in an R=module to be unimodular if the cyclic module vR has a complement; this means that v can be completed to a basis of R^n. If v is unimodular, clearly the ideal generated by the v_i's is the unit ideal.

Lemma 4.4 *If v is a unimodular vector and $f : R^n \to R$ is the R-linear map given by $e_i \mapsto v_i$ (e_i being the canonical basis of R^n for $1 \leq i \leq n$), then*

$$\ker(f) = \{w = (w_1, \ldots, w_n)^t \in R^n \mid \sum_{i=1}^{n} w_i v_i = 0\}$$

is generated by the elements

$$\{v_j e_i - v_i e_j \mid 1 \leq i \leq n\}.$$

Proof.
Basically, the proof is to construct a splitting of the short exact sequence. There are elements $r_1, \cdots, r_n \in R$ such that $\sum_{i=1}^n r_i v_i = 1$. Consider the R-module homomorphism

$$g : R \to R^n \ ; \ 1 \mapsto (r_1, \cdots, r_n)^t.$$

This is a splitting on the right of the above exact sequence
(because $\sum_i r_i v_i = 1$).
Then, the map

$$\theta : x \mapsto x - g(f(x))$$

is a splitting on the left side of the exact sequence (from A^n to $Ker(f)$); that is, $\theta|_{Ker(f)} = 1_{Ker(f)}$.
So, θ is surjective and thus, the elements $\theta(e_i)$ generate $Ker(f)$. Note that

$$\theta(e_i) = e_i - g(f(e_i)) = e_i - g(v_i) = e_i - v_i \sum_j r_j e_j$$

$$= (\sum_j r_j v_j)e_i - \sum_j r_j v_i e_j = \sum_j r_j(v_j e_i - v_i e_j).$$

A further important generalization of the observation above is:

Lemma 4.5 *(i) Let $n \geq 3$. If $v \in R^n$ is a unimodular vector, and $w \in R^n$ such that $w^t v = 0$, then $\mathrm{I}_n + vw^t \in \mathrm{E}_n(R)$.*
(ii) Let $n \geq 3$. If w is unimodular and v is arbitrary satisfying $w^t v = 0$, then $I_n + vw^t \in E_n(R)$.

Proof.
We prove (i); then (ii) follows by taking transposes.
Define the R-linear map $f : R^n \to R$ by $e_i \mapsto v_i$. The hypothesis $w^t v = 0$ implies that $w^t \in Ker(f)$. By the above lemma, we have some elements $r_{ij} \in R$ such that

$$w^t = \sum_{i<j} r_{ij}(v_i e_j - v_j e_i).$$

Writing $w_{ij}^t = v_i e_j - v_j e_i$ for all $i < j$, we observe

$$I_n + vw^t = I_n + v \sum_{i<j} w_{ij}^t = \prod_{i<j} (I_n + vw_{ij}^t).$$

As $n \geq 3$, we have $I_n + vw_{ij}^t \in E_n(R)$ for all $i < j$. This completes the proof.

Now we can deduce the proof of Theorem.

Proof of Suslin's Theorem.
As $\mathrm{E}_n(R)$ is generated by $e_{ij}(\lambda)$; $\lambda \in R$, it is sufficient to show that for $\alpha \in \mathrm{GL}_n(R)$ we must have $\alpha e_{ij}(\lambda)\alpha^{-1} \in \mathrm{E}_n(R)$. Let α_i and β_i $(1 \leq i \leq n)$ be the i-th column of α and i-th row of α^{-1} respectively. Then

$$\alpha e_{ij}(\lambda)\alpha^{-1} = \alpha(\mathrm{I}_n + \lambda \mathrm{E}_{ij})\alpha^{-1} = \mathrm{I}_n + \lambda\alpha_i\beta_j.$$

Since $\alpha, \alpha^{-1} \in \mathrm{GL}_n(R)$, we observe that the ideal generated by the entries of α_i for any i is the unit ideal; likewise, the entries of β_j generate the unit ideal for each j.
Also, $\alpha^{-1}\alpha = I_n$ implies that $\beta_j\alpha_i = 0$, for $j \neq i$. Hence from the above mentioned remark, it follows that $\mathrm{I}_n + \lambda\alpha_i\beta_j \in \mathrm{E}_n(R)$. □

Remarks.
(i) If we take $v = (X, -Y)^t$, and $w = (Y, X)^t$. Then $w^t v = 0$. Hence, the observation above gives P.M.Cohn's example

$$\begin{pmatrix} \mathrm{I}_2 + vw^t & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 + XY & X^2 \\ -Y^2 & 1 - XY \end{pmatrix} \in \mathrm{E}_3(A).$$

Indeed, it is $[e_{31}(Y)e_{32}(X), e_{13}(-X)e_{23}(y)]$ as can be worked out from Vaserstein's lemma.

Cohn's matrix is stably elementary as seen above even though it is not in the elementary subgroup as we will prove in the tutorials.

(ii) Suslin's theorem on normality of $E_n(R)$ holds for a noncommutative ring R (for $n > 2$) if R is finitely generated as a $Z(R)$-module. This is due to Tulenbaev.
More generally, it holds for R under the condition that n is bigger than the stable range of R.

(iii) For a general ring R and for $n \geq 2$, the centralizer of $E_n(R)$ in $GL_n(R)$ is the 'scalar'subgroup

$$RL_n(R) := \{rI_n : r \in Z(R)^*\}.$$

5 Congruence subgroups

Let R be an arbitrary ring with unity. Recall that we have fixed embeddings of $M_n(R)$ in $M_m(R)$ for each $m > n$. This enabled us to define the *stable* groups

$$GL(R) = \bigcup_{n \geq 1} GL_n(R)$$

and

$$E(R) = \bigcup_{n \geq 1} E_n(R).$$

For any two-sided ideal I of R, we define the *principal congruence subgroup of level* I to be the normal subgroup

$$GL_n(R, I) := \{g \in GL_n(R) : g_{ij} - \delta_{ij} \in I \ \forall \ i, j\}.$$

Notice that this is the kernel of the homomorphism from $GL_n(R)$ to $GL_n(R/I)$ induced by the quotient map $R \to R/I$.
Recall also that we defined $E_n(R, I)$ to be the *normal subgroup of* $E_n(R)$ *generated by* $e_{ij}(t)$ *as* t *varies in* I *and* $i \neq j$ *vary.* Clearly, $E_n(R, I) \leq GL_n(R, I)$ as generators are all in the congruence subgroup. We may define the stable groups $E(R, I)$ and $GL(R, I)$ similarly.
In order to study some natural questions (like the question of normality of $E_n(R, I)$ in $GL_n(R)$), it is convenient to introduce the notion of Cartesian squares. Specifically, one problem is to determine if $E_n(R, I)$ defined as above coincides with the kernel of $E_n(R) \to E_n(R/I)$. We will use special Cartesian squares called "doubles" to prove that this is so.

5.1 Doubles & Cartesian squares

Roughly speaking, if I is a two-sided ideal of a ring R and G is a functor from rings to groups, then we would like to define a natural group $G(R, I)$ such that we may get an exact sequence of the form $G(R, I) \to G(R) \to G(R/I)$.
The tool which helps us to do this is the following.

Let R be a ring with unity and I, a two-sided ideal. The *double of* R *along* I is defined to be the subring

$$R \times_I R := \{(a, b) \in R \times R : a - b \in I\}$$

of $R \times R$. Consider the diagram below where p_1, p_2 are the two projections and i is the natural quotient map.

$$\begin{tikzcd}
R \times_I R \arrow[r, "p_1"] \arrow[d, "p_2"'] & R \arrow[d, "i"] \\
R \arrow[r, "i"'] & R/I.
\end{tikzcd}$$

Note that

$$i(r_1) = i(r_2) \Leftrightarrow (r_1, r_2) \in R \times_I R.$$

Moreover, the tuple (r_1, r_2) above is unique with the above property and the diagonal map $\delta : R \to R \times R$ splits both p_1 and p_2 simultaneously.
In fact, there is another way to view the double of R along I. Consider the semidirect product

$$R \propto I := \{(a, b) : a \in R, b \in I\}$$

where the product is defined as

$$(a, b)(c, d) := (ac, ad + bc - bd).$$

Observe that

$$(a, b) \mapsto (a, a - b)$$

yields an isomorphism from the double to the semidirect product.
We make a useful observation now.

Observation. *Let G be an exact functor from rings to groups. Suppose R is a ring and I, a two-sided ideal. Then,*
(i) $G(p_2)$:*ker* $(G(p_1) : G(R \times_I R) \to G(R)) \to$ *ker* $(G(i) : G(R) \to G(R/I))$ *is an isomorphism.*
(ii) *Denote the above kernel by $G(R, I)$. Then, $G(R \times_I R) \cong G(R) \propto G(R, I)$.*
Proof.
(i) follows by definition.
For (ii), just observe that the diagonal map δ simultaneously splits both the projections p_1 and p_2 from the double.
Details to be worked out in the tutorials!

Proposition.
Let R be any ring and I, any two-sided ideal in it.
(i) *The diagonal map δ from $GL_n(R)$ to $GL_n(R \times_I R)$ maps $E_n(R)$ into $E_n(R \times_I R)$.*
(ii) $p_2($ *ker* $p_1 | E_n(R \times_I R)) = E_n(R, I)$.
(iii) *If R is commutative and $n \geq 3$, $E_n(R, I)$ is normal in $GL_n(R)$.*
Proof.
Part (i) is clear since $\delta(e_{ij}(t)) = e_{ij}(t, t)$.
Part (ii) is really just (ii) of the observation. More precisely, first note that

$\delta(E_n(R))$ker $(p_1|_{E_n(R\times_I R)})$ is a subgroup of $E_n(R\times_I R)$. We observe that they are equal. Indeed, let $e_{ij}(s,t)\in E_n(R\times_I R)$. Then,

$$e_{ij}(s,t)=e_{ij}(s,s)e_{ij}(0,t-s)\in\delta(E_n(R))ker(p_1|_{E_n(R\times_I R)}).$$

It is clear that

$$\delta(E_n(R))\cap ker(p_1|_{E_n(R\times_I R)})=\{1\}.$$

Thus, we have a unique decomposition of each element of $E_n(R\times_I R)$ is a unique product of the form xy where $x\in\delta(E_n(R))$ and $y\in kerp_1|_{E_n(R\times_I R)}$. So, we have shown that

$$1\to p_2(kerp_1|_{E_n(R\times_I R)}\overset{p_2^{-1}}{\to}E_n(R\times_I R)\overset{p_1}{\to}E_n(R)\to 1$$

is a split extension with a splitting $\delta: E_n(R)\to E_n(R\times_I R)$.
If $S=\{e_{ij}(s,s): s\in R\}$ and $T=\{e_{ij}(0,t): t\in I\}$, then these are subsets of $\delta(E_n(R))$ and ker $(p_1|_{E_n(R\times_I R)})$ respectively, and together generate the whole of $E_n(R\times_I R)$. Thus,

$$E_n(R\times_I R)=<S><T>_{<S>}$$

Since $<T>_{<S>}\leq ker(p_1|_{E_n(R\times_I R)})$, they must be equal since $<S>\leq\delta(E_n(R))$. Taking the images under p_2, we get

$$p_2(kerp_1|_{E_n(R\times_I R)})=p_2<T>_{p_2<S>}=<\{e_{ij}(t): t\in I\}>_{E_n(R)}=E_n(R,I).$$

So, (ii) is proved.
To prove (iii), let R be commutative and $n\geq 3$. Then, as we proved, $E_n(R\times_I R)$ is normal in $GL_n(R\times_I R)$. Therefore, $(kerp_1)\cap E_n(R\times_I R)$ is normal in $GL_n(R\times_I R)$. Since p_2 is an epimorphism from $GL_n(R\times_I R)$ onto $GL_n(R)$, we have that $p_2(kerp_1\cap E_n(R\times_I R)$ is normal in $GL_n(R)$. The left hand side is $E_n(R,I)$ by (ii). The proof is complete.

5.2 Normality results for elementary congruence subgroups

Theorem. Let $n\geq 3$ and R be any ring with unity.
(i) If $H\leq GL_n(R)$ is normalized by $E_n(R)$, and F is a *subset* of $E_n(R)\cap H$, then $H\geq E_n(R,I)$ where I is the two-sided ideal generated by all the entries of I_n-e as e varies in F.
(ii) If I,J are two-sided ideals, then

$$[E_n(R,I),E_n(R,J)]\geq E_n(R,IJ).$$

In particular,

$$[E_n(R), E_n(R,I)] = E_n(R,I).$$

(iii) $[GL_n(R), GL_n(R,I)] \leq E_{2n}(R,I)$.

Proof.

(i) Let $e_{ij}(t) \in H$. Then, since H is normalized by $E_n(R)$, each elementary matrix of the form $e_{kl}(atb) \in H$ where $k \neq l$ and $a, b \in R$. Therefore, $H \geq E_n(I)$ where $I = RtR$. As t varies, we get (i).

(ii) As $[e_{ij}(a), e_{jk}(b)] = e_{ik}(ab)$ for i, j, k distinct, we have that $[E_n(R,I), E_n(R,J)]$ contains all $e_{ij}(c)$ for $i \neq j$ and $c \in IJ$. Since $[E_n(R,I), E_n(R,J)]$ is normalized by $E_n(R)$, (i) implies the result. The particular case follows from the $J = R$ case.

Finally, (iii) is a consequence of Whitehead's lemma. Indeed, let $g \in GL_n(R)$ and $h \in GL_n(R,I)$. From the Whitehead lemma, we have the equivalence of $\begin{pmatrix} gh & 0 \\ 0 & I_n \end{pmatrix}$ and $\begin{pmatrix} hg & 0 \\ 0 & I_n \end{pmatrix}$ modulo $E_{2n}(R,I)$. It follows that

$$\begin{pmatrix} gh(hg)^{-1} & 0 \\ 0 & I_n \end{pmatrix} \in E_{2n}(R,I).$$

6 Stable linear groups and K_1

Before defining the group K_1 etc., we define the notion of a level.

Definition. A subgroup $H \leq GL_n(R)$ is said to be of *level I* for a two-sided ideal I if

$$E_n(R,I) \leq H \leq GL_n(R,I).$$

This is meaningful because of the following observation:

Lemma. *If $n > 1$, then a subgroup H of $GL_n(R)$ which has a level, has a unique level.*

Proof.

Suppose H has levels I, J for two two-sided ideals. The epimorphism of rings

$$\theta : R \to R/J,$$

implies an epimorphism of groups

$$E_n(R,I) \to E_n(R/J, \theta(I)).$$

But, the inclusion $E_n(R,I) \leq GL_n(R,J)$ shows that $E_n(R/J, \theta(I))$ is trivial. Hence, the ideal $\theta(I) = 0$ (as $n > 1$). That is, $I \subseteq J$. By symmetry, they are equal.

Define $K_1(R) := GL(R)/E(R)$. More generally, for any ideal I, we define the groups $K_1(R,I) = GL(R,I)/E(R,I)$ and $SK_1(R,I) = SL(R,I)/E(R,I)$ (the latter when R is commutative).
At the finite level, we define the coset spaces $K_{1,n}(R,I) = GL_n(R,I)/E_n(I)$ and there are obvious maps from $K_{1,n}(R,I)$ to $K_{1,m}(R,I)$ for $m > n$ and hence to the whole of $K_1(R,I)$.
Some natural interesting questions are to determine the possible injectivity of the maps $K_{1,n}(R,I) \to K_{1,m}(R,I)$ for $m > n$ and the possible surjectivity of the maps $K_{1,n}(I) \to K_1(I)$.
As we proved earlier, for Euclidean rings R, we have $SK_1(R) = \{1\}$ and so, $K_1(R) \cong R^*$.
Thus, the study of K_1 of a ring generalizes the unit group.
Finally, we prove the main theorem of this section which enables us to classify all normal subgroups of the stable group $GL(R)$.

Theorem. Let R be any ring with identity.
(i) If $H \leq GL(R)$ is normalized by $E(R)$, then there is a unique two-sided ideal I in R such that H has level I; that is,

$$E(R,I) \leq H \leq GL(R,I).$$

(ii) If H is any subgroup of $GL(R)$ whose level is a two-sided ideal I, then

$$[GL(R),H] = [E(R),H] = E(R,I) \leq H.$$

In particular, H is *normal* in $GL(R)$. A still particular case shows that $E(R,I)$ is normal in $GL(R)$.
(iii) Let $\theta : R \to S$ be an epimorphism of rings. If $H \leq GL(R)$ has level I for some two-sided ideal I in R, then

$$E(R,I) \to E(S,\theta(I))$$

is an epimorphism of groups and $\theta(H)$ is a normal subgroup of $GL(S)$ with level $\theta(I)$.
Proof.
Part (iii) follows from (i) and (ii).
Part (ii) follows from assertion proved earlier that:

$$[E_n(R), E_n(R,I)] = E_n(R,I) \quad if \quad n \geq 3.$$

and from Whitehead's lemma.
To prove (i), we first observe that the uniqueness of I follows from (ii).

If H is nontrivial, we first prove the existence of some non-zero ideal I such that $E(R, I) \leq H$. Indeed, for each n, look at the subgroup

$$H_n := H \cap GL_n(R).$$

Look at H_n in the affine group $\begin{pmatrix} GL_n(R) & R^n \\ 0 & 1 \end{pmatrix} \cong GL_n(R) \propto R^n$ of $GL_{n+1}(R)$. The subgroup H_n is nontrivial for large enough n.

Note that since H is normalized by $E(R)$, the subgroup $[H_n, R^n]$ of R^n (in the affine group above) is contained in H.

But, $[H_n, R^n]$ is clearly the additive subgroup $\sum_{g \in H_n} Im(g - I_n)$ of R^n in the affine group and it is invariant under $E_n(R)$. But, an additive subgroup of R^n which is invariant under $E_n(R)$ is also invariant under the additive group generated by $E_n(R)$ which is the whole of $M_n(R)$. Thus, such a nontrivial subgroup must be $R^n L$ for some non-zero *left ideal* L of R. That is, we have proved that H contains all matrices of the form $\begin{pmatrix} I_n & a \\ 0 & 1 \end{pmatrix}$ with $a \in L^n$. As shown earlier, this implies that H contains $E(R, LR)$ for the two-sided ideal LR.

Thus, we have shown that a nontrivial H in (i) must contain $E(R, I)$ for some non-zero two-sided ideal I.

Let I_0 be the biggest such two-sided ideal. We claim that $H \leq GL(R, I_0)$. Suppose this is not true.

Then, look at the homomorphism $GL(R) \to GL(R/I_0)$ and induced by $R \to R/I_0$. We know that $E(R) \to E(R/I_0)$ is an epimorphism. The image $\bar{H}$ of H in $GL(R/I_0)$ is nontrivial (by assumption that $H \not\leq GL(R, I_0)$) and is normalized by $E(R/I_0)$. Applying the earlier argument to R/I_0 instead of R, there is a two-sided ideal $I_1 \supset I_0, I \neq I_0$ in R so that $E(R, I_1/I_0) \leq \bar{H}$. Looking at the preimage in $GL(R)$, we get

$$E(R, I_1) \leq GL(R, I_0)H.$$

So $E(R, I_1) = [E(R), E(R, I_1)] \leq H$ because $E(R)$ normalizes H and

$$[E(R), GL(R, I_0)] \leq E(R, I_0) \leq H.$$

This contradicts the choice of I_0. Therefore, we have a two-sided ideal as in (i) of the theorem.

The theorem enables us to define, for each two-sided ideal, the group

$$K_1(R, I) := GL(R, I)/E(R, I).$$

Moreover, the theorem implies that:

The determination of all normal subgroups of $GL(R)$ is equivalent to determining $K_1(R, I)$ for every two-sided ideal I.

7 Steinberg group

Definition.
Let R be any associative ring with unity.
For $n \geq 3$, recall we defined 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

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

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

Further, for any unit u in R, the elements

$$w_{ij}(u) = x_{ij}(u)x_{ji}(-u^{-1})x_{ij}(u)$$

and

$$h_{ij}(u) = w_{ij}(u)w_{ij}(-1)$$

will be useful while discussing properties of $St_n(R)$.
If R is commutative, it is easy to check (exercise for tutorials!) that

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

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

$$x_{ij}(t)x_{ij}(u) = x_{ij}(t+u)$$

and the relations

$$w_{ij}(t)x_{ji}(u)w_{ij}(-t) = x_{ij}(-tut)\text{if } t \in R^*$$

for $(i,j) = (1,2)$ or $(2,1)$ where,

$$w_{ij}(t) = x_{ij}(t)x_{ji}(-t^{-1})x_{ij}(t).$$

Remark.
The last relation for $(i,j) = (1,2)$ implies the relation for $(i,j) = (2,1)$; that is,

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

Definition.
There is an obvious epimorphism from $St_n(R)$ to $E_n(R)$ given by

$$\phi_n : x_{ij}(t) \mapsto e_{ij}(t).$$

Passing to the direct limit, we thus have an epimorphism

$$\phi : St(R) \to E(R) = [GL(R), GL(R)] = [E(R), E(R)].$$

We define $K_2(R) = Ker(\phi)$.
$K_2(R)$ can be thought of as the set of all nontrivial relations between elementary matrices over R.
More generally, we define $K_2(n, R) = Ker(\phi_n)$.

Remarks.
We will prove that $K_2(R) =$ Center of $St(R)$. In fact, we will also show that $St(R)$ is the so-called universal central extension of $E(R)$. Thus, one may identify $K_2(R)$ with the Schur multiplier of $E(R)$ (which is the homology group $H_2(E(R), \mathbf{Z})$ by definition). In other words, $K_2(R)$ is like the (dual of the) "fundamental group".

Proposition. *$K_2(R)$ is the center of $St(R)$.*
Proof.
Now, an $n \times n$ matrix over R which commutes with all $e_{ij}(t)$ as t varies in R and $i \neq j$ vary, if and only if, it is a scalar matrix diag $(a, \cdots, a)$ with $a \in R^*$ (exercise for the tutorials). Thus, no nontrivial element of $E_{n-1}(R)$ centralizes the whole of $E_n(R)$. Hence, the center of $E(R)$ is trivial.
Suppose $c \in Z(St(R))$. Then, $\phi(c)$ is in the center of $E(R)$ (which is trivial by the above observation) and, hence $c \in Ker(\phi)$.
Conversely, suppose $c \in Ker(\phi)$.
We will show that c commutes with each Steinberg generator $x_{ij}(t)$.
Firstly, choose n so large that c is a word in the generators $x_{ij}(t)$ for $i, j < n$ and $t \in R$.
Now, we have a simple observation (proof in tutorials):
Fact. *The subgroup R_n of $St(R)$ generated by $x_{in}(t)$ as i varies from 1 to $n-1$ and t varies in R, is abelian. Moreover, each of its elements is a unique word of the form*

$$x_{1n}(t_1)x_{2n}(t_2)\cdots x_{n-1,n}(t_{n-1}).$$

From this fact, it follows that R_n maps isomorphically onto its image in $E(R)$.
Now, evidently the generators

$$x_{ij}(t)R_n x_{ij}(-t) \leq R_n$$

if both $i, j < n$. Hence, the element $cR_nc^{-1} \leq R_n$. As $\phi(c) = 1$, the injectivity of ϕ restricted to R_n shows that c centralizes R_n. In particular, c commutes with each $x_{in}(t)$ for $i < n$. Similarly, it follows that c commutes with each $x_{nj}(t)$ for $j < n$. Therefore, c commutes with

$$x_{ij}(t) = [x_{in}(t), x_{nj}(1)]$$

for all $i, j < n$. Since n can be arbitrarily large, this means that c is in the center of $St(R)$.

Remarks.
Now, we digress a bit to recall some generalities on central extensions which will be useful while proving the universality of $St(R)$ over $E(R)$.

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

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

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.

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 $H^2(G, A)$ characterizes *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.

Calculating central extensions of 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 $n = rank(F)$.

The notion of central extensions is an algebraization 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 up to equivalence. Sometimes, one simply writes (π, E) for the u.c.e. and $\mathrm{Ker}(\pi)$ is called the *Schur muliplier* of G.

Lemma.
(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.*

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.*

We can prove now:

Theorem. *If $n \geq 5$ and, if $St_n(R) \to E_n(R)$ is a central extension, then it is the universal central extension.*
Therefore, for any ring R, the stable Steinberg $St(R)$ is the universal central extension of $E(R)$.
In particular, $K_2(R) = H_2(E(R), \mathbf{Z})$.
Proof.
The second and third statements follow from the first one, which we now prove. Consider any central extension

$$1 \to C \to X \xrightarrow{\phi} St_n(R) \to 1$$

where $n \geq 5$.
Because of centrality, we know that for any pair of elements $x, y \in St_n(R)$, the commmutator $[x', y']$ is uniquely defined in X for arbitrary lifts x' and y' of x, y respectively. Therefore, we denote the above commutator in X as $[\phi^{-1}(x), \phi^{-1}(y)]$ for simplicity.
Step I: *If $j \neq k, l \neq i$, then for all $s, t \in R$,*

$$[\phi^{-1}(x_{ij}(s)), \phi^{-1}(x_{kl}(t))] = 1.$$

Proof of step I:
Indeed, choose h different from i, j, k, l (as $n \geq 5$) and write

$$x_{ij}(s) = [x_{ih}(s), x_{hj}(1)].$$

Let $x \in \phi^{-1}(x_{ih}(s)), y \in \phi^{-1}(x_{hj}(1)), z \in \phi^{-1}(x_{kl}(t))$.
Now, $[x,y] \in \phi^{-1}(x_{ij}(s))$. Write $[x,z] = c_1 \in C, [y,z] = c_2 \in C$.
So, $xzx^{-1} = c_1 z, yzy^{-1} = c_2 z$.
Hence $x^{-1}zx = c_1^{-1}z, y^{-1}zy = c_2^{-1}z$.
Then,

$$xyx^{-1}y^{-1}z(yxy^{-1}x^{-1} = z$$

which shows step I.

Step II: *Let i,j,k,l be distinct. Then, for $s,t \in R$, we have*

$$[\phi^{-1}(x_{ij}(s)), \phi^{-1}(x_{jl}(t))] = [\phi^{-1}(x_{ik}(1)), \phi^{-1}(x_{kl}(st))].$$

Proof of step II:
Consider the subgroup Y of X generated by the elements of

$$\phi^{-1}(x_{ik}(1)), \phi^{-1}(x_{kj}(s)), \phi^{-1}(x_{jl}(t)).$$

Then, the commutator subgroup $DY := [Y,Y]$ is generated by the elements of $\phi^{-1}(x_{ij}(s)), \phi^{-1}(x_{kl}(st))$ and $\phi^{-1}(x_{il}(st))$. BY step I, these generating elements commute among themselves and, thus $D_2Y := [DY, DY]$ is trivial.
Now, for any group G, and elements x,y,z, we have (attend tutorial to work out the proof!)

$$[x,[y,z]] = [[x,y],z] \quad mod \quad D_2G.$$

Hence, in $Y \leq X$, taking x,y,z to be elements of $\phi^{-1}(x_{ik}(1)), \phi^{-1}(x_{kj}(s))$ and $\phi^{-1}(x_{jl}(t))$ respectively, we have that

$$[x,[y,z]] = [[x,y],z]$$

which is exactly step II.

Step III: *For distinct i,j,k and elements $t \in R$, the element*

$$s_{ij}(t) := [\phi^{-1}(x_{ik}(1)), \phi^{-1}(x_{kj}(t))]$$

is independent of the choice of k.
Proof of step III:
Take $s = 1$ in step II.

Step IV: *The elements $s_{ij}(t)$ give rise to a splitting of the central extension; that is, they satisfy the Steinberg relations.*
Proof of step IV:
Firstly, note the following commutator identity is valid in an arbitrary group:

$$[u,v][u,w] = [u,vw][v,[w,u]].$$

Now, we consider in X, elements u in $\phi^{-1}(x_{ik}(1))$, v in $\phi^{-1}(x_{kj}(s))$ and w in $\phi^{-1}(x_{kj}(t))$ where i, j, k are distinct and $s, t \in R$. The commutator relation above shows that $s_{ij}(s)s_{ij}(t) = s_{ij}(s+t)$ because

$$[u, v] = s_{ij}(s), [u, w] = s_{ij}(t), [u, vw] = s_{ij}(s+t), [v, [u, w]] = 1.$$

Finally, step II gives immediately that for i, j, k distinct and $s, t \in R$, we have

$$[s_{ik}(s), s_{kj}(t)] = s_{ij}(st).$$

Hence, the central extension splits; that is, $St_n(R)$ is a u.c.e. of $E_n(R)$ if it is a c.e. and if $n \geq 4$.

8 Steinberg symbols and $K_2(\mathbf{Z})$

In this section, we shall show that $K_2(\mathbf{Z})$ is of order 2. This follows once we establish that the $St_n(\mathbf{Z})$ is a central extension of $E_n(\mathbf{Z}) = SL_n(\mathbf{Z})$ whose kernel is of order 2 when $n \geq 3$ (the kernel is infinite cyclic when $n = 2$). That will also thus give a presentation for $SL_n(\mathbf{Z})$. Steinberg's results have been generalized by M.R.Stein to all types of root systems over any commutative rings. Thus, the case we discuss here corresponds to the type A_n and the ring $\mathbf{Z}$. Over fields, a Bruhat decomposition for the Steinberg group is vital to the study. This can be carried over to rings which are actually generated by the units. The ring $\mathbf{Z}$ is harder to study!

We first prove 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

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

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.
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

$$x = w_{i_1,j_1}(u_1) \cdots w_{il j_l}(u_l) \in C \cap W,$$

one has

$$x \equiv w_{i_1,j_1} \cdots w_{i_l j_l} \ mod \ H.$$

One can use the conjugation formulae in the lemma below to push all the terms of the form w_{1r} to the beginning. Moreover, $w_{1r}^2 = 1$ mod 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 $h_{ij}(uv) \equiv h_{ij}(v)h_{ij}(u) \equiv h_{ij}(u)h_{ij}(v)$ mod 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

$$h_{jk}(u) = h_{1k}(u)h_{1j}(u)^{-1}$$

which follows from the lemma below.

Further, $h_{1l}(uv) \equiv h_{1l}(u)h_{1l}(v)$ mod D; so,

$$h_{1l}(u)^{-1} \equiv h_{1l}(u^{-1}) \ mod \ D.$$

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

$$u_1 = u_2 = \cdots = u_{n-1} = 1;$$

so, $c \in D$. This proves the lemma.

Exercises (tutorials!).

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.

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.
Let R be any commutative ring. For any $n \geq 2$, let W_n denote the group generated by $w_{ij}(t), i \neq j$ and $t \in R^*$. 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 for $n \geq 3$ from the following lemma which shows that every element $w \in W_n$ conjugates any Steinberg generator $x_{ij}(s)$ to some $x_{kl}^{\pm 1}(t)$. For $n = 2$, it follows because of the very definition of $St_2(R)$.

Lemma.
Let R be a commutative ring. If $n \geq 3$, i, j, k, l are distinct and $t, u \in R^*$, then we have :
(a) $[w_{ij}(u), x_{kl}(t)] = 1$.
(b) $w_{ij}(u)x_{ik}(t)w_{ij}(u)^{-1} = x_{jk}(-tu^{-1})$.
(c) $w_{ij}(u)x_{kj}(t)w_{ij}(u)^{-1} = x_{ki}(tu^{-1})$.
(d) $w_{ij}(u)x_{ki}(t)w_{ij}(u)^{-1} = x_{kj}(-tu)$.
(e) $w_{ij}(u)x_{jk}(t)w_{ij}(u)^{-1} = x_{ik}(tu)$.
(f) $w_{ij}(u)x_{ij}(t)w_{ij}(u)^{-1} = x_{ji}(-tu^{-2})$.
(g) $w_{ij}(u)x_{ji}(t)w_{ij}(u)^{-1} = x_{ij}(-tu^{2})$.
(h) $w_{ij}(t)^{-1} = w_{ij}(-t)$.
(i) $w_{ij}(u)w_{ik}(t)w_{ij}(u)^{-1} = w_{jk}(-tu^{-1})$.
(j) $w_{ij}(u)w_{jk}(t)w_{ij}(u)^{-1} = w_{ik}(tu)$.
(k) $w_{ij}(u)w_{kj}(t)w_{ij}(u)^{-1} = w_{ki}(tu^{-1})$.
(l) $w_{ij}(u)w_{ki}(t)w_{ij}(u)^{-1} = w_{kj}(-tu)$.

Proof
(a) is obvious.
It is fun to prove the other parts (tutorials!).
We indicate some of them. For instance, let us prove (e):

$$w_{ij}(u)x_{jk}(t)w_{ij}(u)^{-1} = x_{ik}(tu)$$

The left hand side equals

$$x_{ij}(u)x_{ji}(-u^{-1})x_{ij}(u)x_{jk}(t)x_{ij}(-u)x_{ji}(u^{-1})x_{ij}(-u)$$

$$= x_{ij}(u)x_{ji}(-u^{-1})x_{ik}(ut)x_{jk}(t)x_{ji}(u^{-1})x_{ij}(-u)$$

on using $[x_{ij}(u), x_{jk}(t)] = x_{ik}(tu)$. This is further equal to

$$x_{ij}(u)x_{jk}(t)x_{ji}(-u^{-1})x_{ik}(ut)x_{ji}(u^{-1})x_{ij}(-u)$$

$$= x_{ij}(u)x_{jk}(t)x_{jk}(-t)x_{ik}(ut)x_{ij}(-u)$$

$$= x_{ij}(u)x_{ik}(ut)x_{ij}(-u) = x_{ik}(ut).$$

One can similarly prove (b),(c), and (d).
To prove (f), one only needs to write

$$x_{ij}(t) = [x_{ik}(t), x_{kj}(1)]$$

for some k different from i, j. This is possible because $n \geq 3$.
Finally, (h) follows by applying (f) and (g).

Theorem.
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

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

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

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

Further, $K_2(\mathbf{Z})$ has order 2.
For $SL_2(\mathbf{Z})$, one has an analogous presentation by two generators X_{12}, X_{21} and two relations

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

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

Proof.
Let us lead to the proof in easy steps.
We shall show that, for all $n \geq 2$,

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

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.
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 is a standard unit vector, say $(0,0,\cdots,1)$.
Proof - Tutorials!

Using the lemma, let us show by induction on n that $C_n \subseteq W_n$ for all $n \geq 2$.

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,

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

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 $x_1 \cdots x_r = x\theta_{n-1}(y)$ where x is a product of factors of the form $x_{in}^{\pm 1}$ for the i's, and $\theta_{n-1}(y) \in \theta_{n-1}(St_{n-1}(\mathbf{Z}))$ 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 $x = 1$. 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 where $w_{12} := w_{12}(-1)$.
For $n = 2$, this is clear since $w_{12}(-1)w_{21}(-1) = Id$ 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 4-th 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, $C_n =< w_{12}^4 >$.
Finally, it can be seen that w_{12}^4 has order 2 in $St_n(\mathbf{Z})$. Indeed, the symbol $\{-1,-1\}$ is

$$h_{12}(1) h_{12}(-1)^{-1} h_{12}(-1)^{-1} = w_{12}^4$$

is bilinear since

$$\{uv, w\} = [h_{12}(uv), h_{13}(w)] = [\{u,v\} h_{12}(v) h_{12}(u), h_{13}(w)]$$

$$= [h_{12}(v) h_{12}(u), h_{13}(w)] = [h_{12}(u), h_{13}(w)][h_{12}(v), h_{13}(w)] = \{u,w\}\{v,w\}.$$

That is, $K_2(\mathbf{Z})$ has order ≤ 2. It can be shown (proof to be given!) that the symbol $\{-1,-1\}$ is not trivial in $St_n(\mathbf{Z})$ for $n > 2$.
The proof of the theorem is complete.

Remarks.
It can be shown that w_{12}^4 has infinite order in $St_2(\mathbf{Z})$.

9 Matsumoto's theorem and $K_2(\mathbf{Q})$

In this section, we do two things:
(a) we state a deep theorem of Matsumoto on K_2 of any field which was applied by him to solve the congruence subgroup problem for Chevalley groups;
(b) state and indicate a proof due to Tate of the computation of $K_2(\mathbf{Q})$.

Theorem (Matsumoto). *For any field F, the group $K_2(F)$ has a presentation where the generators are symbols $\{x,y\}$ (for $x, y \in F^*$) and the relations are:*

(i) the symbol is bilinear and
(ii) $\{x, 1-x\} = 1$ *for* $x \neq 0 \neq 1-x$.

Theorem (Tate).

$$K_2(\mathbf{Q}) \cong \{1, -1\} \oplus \bigoplus_{p>2} (\mathbf{Z}/p\mathbf{Z})^*.$$

The proof - according to Tate - is a rewriting of Gauss's first proof of the quadratic reciprocity law! Indeed, the norm-residue symbol

$$(x, y)_p := (-1)^{v_p(x)v_p(y)} \frac{x^{v_p(y)}}{y^{v_p(x)}}$$

is a symbol from $\mathbf{Q}^* \times \mathbf{Q}^*$ to $(\mathbf{Z}/p\mathbf{Z})^*$ for prime $p > 2$.
Define $(x, y)_2 = -1$ if both $x, y < 0$ and 1 otherwise. The asserted isomorphism is then given by

$$(x, y) \mapsto ((x, y)_p)_p.$$

10 Congruence subgroup problem

In a naive form, the congruence subgroup problem for $SL_n(\mathbf{Z})$ asks whether every subgroup of finite index in $SL_n(\mathbf{Z})$ contains a congruence subgroup; viz., a subgroup of the form Ker $(SL_n(\mathbf{Z}) \to SL_n(\mathbf{Z}/m\mathbf{Z}))$, for some integer $m > 1$.

Note that the question is meaningful because of the existence of plenty of congruence subgroups in the sense that their intersection consists just of the identity element.

Already in the late 19th century, Fricke and Klein showed that the answer to this question is negative if $n = 2$.
Indeed, since the free group of rank 2 is the principal congruence subgroup $\Gamma(2)$ of level 2, any 2-generated finite group is a quotient of this group.

It turns out that the finite, simple groups which can occur as quotients by congruence subgroups are the groups $PSL_2(\mathbf{F}_p)$ for primes p.
But there are many finite, simple 2-generated groups (like A_n $(n > 5)$, $PSL_3(\mathbf{F}_q)$ for odd prime q) which are not isomorphic to $PSL_2(\mathbf{F}_p)$ (the latter has abelian q-Sylow subgroups). So, the corresponding kernel cannot be a congruence subgroup of $SL_2(\mathbf{Z})$.

An explicit example of a noncongruence subgroup is the following:

Let k be a positive integer which is not a power of 2.

In the (free) group generated by the matrices $X = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ and $Y = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$, the group of all matrices which are words in X, Y such that the exponents of X, Y add up to multiples of k, is a subgroup of finite index in $SL_2(\mathbf{Z})$ but is not a congruence subgroup.

In fact, it can be shown that there are many more noncongruence subgroups of finite index in $SL_2(\mathbf{Z})$ than there are congruence subgroups!
It was only in 1962 that Bass-Lazard-Serre - and, independently, Mennicke - showed that the answer to the question is affirmative when $n \geq 3$.

Later, in 1965, Bass-Milnor-Serre generalised this to the special linear and the symplectic groups over number fields.

The question can be generally asked for more general groups than SL_n and for the rings of S-integers in number fields. In this more general form, the answer can be 'no' but a certain group called the *congruence subgroup kernel* defined by Serre, measures the deviation from the property holding good. The general problem is to compute this kernel. In particular, a conjecture of Serre predicts when this is a finite group and this has been proved in many cases. For $SL_n(O_S)$, the congruence kernel is simply $SK_1(O_S, I)$ when $n > 2$.

We indicate here the proof of CSP for $SL_3(\mathbf{Z})$. In the tutorials, we can complete the details.

Look at $G = SL_3(\mathbf{Z})$ and the embedded $SL_2(\mathbf{Z})$ in its left hand top corner.

For any r, consider the normal subgroup $E_3(r)$ generated by $U_{12}(r)$ in G.

For any element g of $\Gamma_3(r)$ (for SL_3), one may apply elementary row and column operations and get x, y in $E_3(r)$, with xgy in the principal congruence subgroup $\Gamma_2(r)$ of level r for the embedded SL_2.

This implies easily that the action of G on $\Gamma_3(r)/E_3(r)$ by conjugation is trivial.

Now, the first rows (a, b) of $\Gamma_2(r)$ for $SL_2(\mathbf{Z})$ give rise to corresponding elements $\begin{pmatrix} a & b & 0 \\ * & * & 0 \\ 0 & 0 & 1 \end{pmatrix}$ of $\Gamma_3(r)/E_3(r)$; these are the so-called Mennicke symbols $M(a, b)$ which will be shown to be trivial.

The Mennicke symbol has nice properties like:

(i) $M(a + tb, b) = M(a, b)$ for all $t \in \mathbf{Z}$,

(ii) $M(a, rta + b) = M(a, b)$ for all $t \in \mathbf{Z}$,

(iii) $M(a, b)$ is multiplicative in b,

(iv) $M(a, b) = 1$ if $b \equiv \pm 1 \bmod a$.

Using these, one can show that it is actually trivial (tutorials!).

Therefore, $\Gamma_3(r) = E_3(r)$; so, a normal subgroup of finite index which contains $E_3(r)$ (as it must, for some r), also contains $\Gamma_3(r)$.

THANK YOU!