Bounded generation of some linear groups

Aleksander Viktorovich Morgan
Bristol, VA

Masters of Science, University of Virginia, 2011

A Dissertation presented to
the Graduate Faculty of the University of Virginia,
in Candidacy for the Degree of Doctor of Philosophy

Department of Mathematics
University of Virginia
Charlottesville, VA 22904-4137, USA
December, 2019

Advisor: Andrei Rapinchuk

Abstract. Let R be a commutative ring with identity. For $n \geq 2$, we let $\mathrm{E}_n(R)$ denote the subgroup of the special linear group $\mathrm{SL}_n(R)$ generated by all elementary matrices, and for an ideal J of R we let $\mathrm{E}_n(R, J)$ denote the normal subgroup of $\mathrm{E}_n(R)$ generated by the elementaries that are congruent to the identity matrix modulo J. The goal of this dissertation is to establish several results asserting that in certain situations the groups $\mathrm{E}_n(R)$ and $\mathrm{E}_n(R, J)$ have finite width with respect to their natural generating sets. In particular, it is shown in Chapter II that if R is a ring of algebraic S-integers with infinitely many units, then $\mathrm{E}_2(R)$ (which in this case coincides with $\mathrm{SL}_2(R)$) has width ≤ 9 with repect to elementary matrices. This result yields bounded generation of $\mathrm{SL}_2(R)$ in the case at hand, as an abstract group. The results of Chapter III apply to the ideals J in an arbitrary Noetherian ring R for which the quotient R/J is finite. We show that if $\mathrm{E}_n(R)$ has finite width with respect to elementaries, then every element in $\mathrm{E}_n(R, J^2)$ is a product of a bounded number of elementary matrices congruent to I_n modulo J. Assuming further that R satisfies Bass's stable range condition, we derive from this result that the normal subgroup of $\mathrm{E}_n(R)$ generated by a given matrix subject to some conditions, has finite index and finite width with respect to the union of the conjugacy classes of the matrix and its inverse.

Contents

CHAPTER I

Overview of Groups with Bounded Generation, and Facts from Alg. Number Theory

I.1. Statements of the main results

The focus of this work is the property of *bounded generation* of special linear groups over commutative rings, as well as of their natural subgroups associated to the ideals of the ring, with respect to their natural generating sets.

Definition. Let Γ be an abstract group with a symmetric generating set T. (Recall that T is *symmetric* if $T = T^{-1}$.) Then

- Γ has *finite width* with respect to T if there exists an integer m so that every element of Γ is a product of no more than m elements from T. The smallest such m is called the *width* of Γ with respect to T.
- Γ has *bounded generation* if furthermore we can take T to be a finite union of cyclic subgroups of G. This is equivalent to the existence of finitely many elements $g_1, \ldots, g_d \in \Gamma$ so that Γ can be decomposed as a product of the cyclic subgroups:

  $$\Gamma = \langle g_1 \rangle \cdots \langle g_d \rangle, \text{ where } \langle g \rangle \text{ is the cyclic subgroup generated by } g.$$

  In this case, the smallest such integer d is called the *degree* of bounded generation.

We refer the reader to Subsection I.3.1 for examples of groups with bounded generation.

Let k be a number field, S be a finite set of places of k containing all the archimedean ones. Let $\mathcal{O}_k$ be the ring of integers of k, and $\mathcal{O} = \mathcal{O}_{k,S}$ be its ring of S-integers (see Subsection I.4.1 for precise definitions). It is known (cf. [**Va**]) that if the group of units $\mathcal{O}^\times$ is infinite, then the group $\mathrm{SL}_2(\mathcal{O})$ of unimodular matrices over $\mathcal{O}$ is generated by the set of elementary matrices. Our first main result asserts that it actually has finite width with respect to this set.

Theorem II.3.6. *(cf. [**MRS**]) If $\mathcal{O}^\times$ is infinite, then $\mathrm{SL}_2(\mathcal{O})$ has width ≤ 9 with respect to the set of elementary matrices.*

This theorem implies bounded generation of $\mathrm{SL}_2(\mathcal{O})$, and completes a long line of research in this direction – see Subsection I.3.2 for a historical survey. Furthermore, it yields a bound on the width of $\mathrm{SL}_n(\mathcal{O})$ for all $n \geq 2$, which is independent of the discriminant of k. More precisely, we have

Corollary II.3.7. *Assume that the group $\mathcal{O}^\times$ is infinite. Then for $n \geq 2$, the width of $\mathrm{E}_n(\mathcal{O})$ with respect to the set of elementary matrices is $\leq \frac{1}{2}(3n^2 - n) + 4$.*

The above theorem is established in Chapter II. The goal of Chapter III is to use the finiteness of the width obtained in Theorem II.3.6, in order to establish the finiteness of the width of $F_n(\mathcal{O}, J)$ with respect to $\mathcal{F}_n(\mathcal{O}, J)$. It turns out that the methods of Chapter III apply not only to rings of

S-integeres of number fields, but also to all commutative Noetherian rings R over which $\mathrm{E}_n(R)$ has finite width with respect to the set of elementary matrices – but, for $n \geq 3$. Let $\mathcal{F}_n(R,J)$ be the set of elementary matrices whose entry different from that of the identity matrix is in J, and $\mathcal{E}_n(R,J)$ the set of $\mathrm{E}_n(R)$-conjugates of matrices in $\mathcal{F}_n(R,J)$. Let $\mathrm{F}_n(R,J)$ and $\mathrm{E}_n(R,J)$ denote the groups generated by the sets $\mathcal{F}_n(R,J)$ and $\mathcal{E}_n(R,J)$, respectively. Then the first section of Chapter III proves:

Theorem III.1.6. *Let R be a Noetherian ring, and let $n \geq 3$. Assume that $\mathrm{E}_n(R)$ has finite width with respect to $\mathcal{E}_n(R)$. Then for any ideal $J \subset R$ satisfying $[R:J] < \infty$:*

(a) *$\mathrm{E}_n(R,J)$ has finite width with respect to $\mathcal{E}_n(R,J)$;*

(b) *there exists an integer N, so that $\mathrm{E}_n(R,J^2)$ is contained in the product of N copies of $\mathcal{F}_n(R,J)$, namely $\mathcal{F}_n(R,J)^N$;*

(c) *$[\mathrm{E}_n(R):\mathrm{F}_n(R,J)] < \infty$;*

Theorem III.1.13. *With the above setup:*

$\mathrm{F}_n(R,J)$ has finite width with respect to $\mathcal{F}_n(R,J)$.

When R is the ring of S-integers of a number field, then R is Noetherian, and every (non-zero) ideal J satisfies $[R:J] < \infty$. Therefore, the above theorem applies.

A ring is said to satisfy the stable range condition $\mathtt{SR}_n$ (due to Bass, cf. [**B**, P. 14]) for $n \geq 1$, if for any n elements $a_1, \ldots, a_n \in R$ which are coprime (i.e., $a_1R + \ldots + a_nR = R$), there exist elements $b_1, \ldots, b_{n-1} \in R$ so that the elements $\alpha_i := a_i + b_i a_n$ for $i = 1, \ldots, n-1$ are also coprime. The second section of Chapter III applies to rings which also satisfy this stable range condition, and shows that in this case, the normal subgroups $\mathrm{E}_n(R,J)$ of $\mathrm{SL}_n(R)$ have finite width with respect to conjugacy classes of a given matrix and its inverse.

Proposition III.2.7. *Suppose R is a commutative ring which satisfies the stable range condition $\mathtt{SR}_{n-1}$. Fix $B \in \mathrm{SL}_n(R)$, and for $i = 1, \ldots, n$, let J_i be the ideal generated by the off-diagonal entries of $\mathcal{C}_i(B)$; let $J = J_1 + \ldots + J_n$. Then:*

(a) *$\mathrm{E}_{1n}(J_1) \subset \mathrm{C}_B^{32}$.*

(b) *$\mathrm{E}_{1n}(J) \subset \mathrm{C}_B^{32n}$.*

Combining this result with Theorem III.1.6, we obtain the following:

Theorem III.2.1. *Suppose R is a Noetherian ring which satisfies the stable range condition $\mathtt{SR}_{n-1}$, and for which $\mathrm{E}_n(R)$ has finite width with respect to $\mathcal{E}_n(R)$. Suppose we are given a (non-diagonal) matrix $A \in \mathrm{SL}_n(R)$. Let C_A be the union of the conjugacy class of A and the conjugacy class of A^{-1}. Let J be the ideal generated by the off-diagonal entries of A. If $[R:J] < \infty$, then the normal group generated by A has finite width with respect to C_A.*

The structure of the current chapter is the following. The next section introduces notations, which will be used throughout this work. Section I.3 deals with groups having bounded generation. In Subsection I.3.1, we give examples and non-examples of bounded generation. In Subsection I.3.2 we outline a brief history of research on bounded generation, and in Subsection 1.3.3 we list some applications of bounded generation. Section I.4 is devoted to results from Algebraic Number Theory, which will be used in Chapter II. Subsection I.4.1 discusses splitting of primes in finite extensions of number fields, Subsection I.4.2 focuses on abelian extensions, introducing the Artin map, and Subsection I.4.3 recalls a few results related to the proof of the m-th power reciprocity law.

I.2. Notations

Let R be an arbitrary ring. A (two-sided) ideal J of R is said to be of *finite index* if $[R : J] < \infty$. For the sake of consistency, for an arbitrary $g, h \in R$, we let $[g, h]$ denote $ghg^{-1}h^{-1}$.

A ring R is said to satisfy Bass's stable range condition $\mathtt{SR}_n$ (cf. [**B**, P. 14]) for $n \geq 1$, if for any n elements $a_1, \ldots, a_n \in R$ which are coprime (i.e., $a_1R + \ldots + a_nR = R$), there exist elements $b_1, \ldots, b_{n-1} \in R$ so that the elements $\alpha_i := a_i + b_ia_n$ for $i = 1, \ldots, n-1$ are also coprime. For example, whenever R is a commutative ring, and n is the Krull dimension of Spec R, then for any finite R-algebra A, the ring $\mathrm{GL}(A)$ satisfies the stable range condition $\mathtt{SR}_{n+1}$.

2.1. Important subsets and subgroups of $\mathrm{SL}_n(R)$. For given ideals $J, J' \subseteq R$, a subset $T \subset R$, and integers $1 \leq i \neq j \leq n$, we define the following:

- $\mathrm{E}_{ij}(T)$ is the set of elementary matrices $\{\mathrm{E}_{ij}(\alpha) \mid \alpha \in T\}$.
- $\mathcal{F}_n(R, J)$ is the union $\bigcup\limits_{1\leq i\neq j\leq n} \mathrm{E}_{ij}(J)$; $\mathcal{E}_n(R) = \mathcal{F}_n(R, R)$.
- $\mathrm{F}_n(R, J)$ is the subgroup of $\mathrm{SL}_n(R)$ generated by $\mathcal{F}_n(R, J)$; $\mathrm{E}_n(R) = \mathrm{F}_n(R, R)$.
- $\mathcal{E}_n(J, J') = \{BAB^{-1} \mid B \in \mathrm{E}_n(J), A \in \mathcal{F}_n(R, J')\}$;
- $\mathrm{E}_n(J, J')$ is the subgroup of $\mathrm{SL}_n(R)$ generated by $\mathcal{E}_n(J, J')$;

In particular, $\mathrm{E}_n(R, J)$ is the normal subgroup of $\mathrm{E}_n(R)$ generated by $\mathcal{F}_n(R, J)$.

2.2 Field Theory For a field k, we let k^{ab} denote the maximal abelian extension of k (in a fixed algebraic closure of k). Furthermore, $\mu(k)$ will denote the group of all roots of unity in k; if $\mu(k)$ is finite, as is the case when k is a number field, then we let μ denote its order. For n a positive integer prime to char k, we let ζ_n denote a primitive n-th root of unity.

If we are given a prime number p, we can write any integer n in the form $n = p^e \cdot m$, for some $e \geq 0$, where $p \nmid m$. We then call p^e the *p-primary component* of n.

2.3 Algebraic Number Theory

Let k be a number field, i.e. a finite extension of $\mathbb{Q}$. Then the set V^k of all valuations on k is a union of the set of archimedean valuations, denoted V^k_∞, and the set of non-archimedean valuations, denoted V^k_f (cf. Subsection I.4.1). The ring of integers of k is denoted $\mathcal{O}_k$, and more generally for a finite subset S of V^k containing V^k_∞, $\mathcal{O} = \mathcal{O}_{k,S}$ will denote the ring of S-integers of k – namely, the set

$$\mathcal{O}_{k,S} = \{a \in k \mid v(a) \leq 1 \ \text{ for all } v \in V^k \setminus S\}.$$

Note that $\mathcal{O}_k = \mathcal{O}_{k,V^k_\infty}$. Then the nonzero prime ideals of $\mathcal{O}_{k,S}$ are in a natural bijective correspondence with the valuations in $V^k \setminus S$. For any $v \in V^k$, we let k_v denote the corresponding completion; if $v \in V^k_f$ then $\mathcal{O}_v$ will denote the valuation ring in k_v with the valuation ideal $\hat{\mathfrak{p}}_v$ and the group of units $U_v = \mathcal{O}_v^\times$; if $v \notin S$, then $\mathfrak{p}_v$ will denote the corresponding prime ideal of $\mathcal{O}_{k,S}$ (note that $\mathfrak{p}_v = \mathcal{O} \cap \hat{\mathfrak{p}}_v$). Conversely, for a nonzero prime ideal $\mathfrak{p}$ of $\mathcal{O}$ or of $\mathcal{O}_k$, we let $v_{\mathfrak{p}} \in V^k \setminus S$ denote the corresponding valuation.

Next, suppose we are given a nonzero ideal $\mathfrak{a}$ of $\mathcal{O}$. Generalizing Euler's φ-function, we set

$$\phi(\mathfrak{a}) = |(\mathcal{O}/\mathfrak{a})^\times|.$$

Also, since $\mathcal{O}$ is Dedekind, we can uniquely write any ideal $\mathfrak{a}$ as $\mathfrak{a} = \mathfrak{p}_1^{\alpha_1} \cdots \mathfrak{p}_r^{\alpha_r}$; and then, we set $V(\mathfrak{a}) = \{\mathfrak{p}_1, \ldots, \mathfrak{p}_r\}$. For a given $a \in k^\times$, we define $V(a) = \{\, v \in V_f^k \mid v(a) \neq 0\,\}$; for example, if $\mathcal{O} = \mathcal{O}_{k,S}$, then $V(a\mathcal{O}) = V(a) \setminus S$.

Finally, we would like to define idèles. The adèle ring V_k is defined as the restricted product of the k_v with respect to their rings of integers $\mathcal{O}_v$. Namely,

$$V_k = \{\mathbf{i} = (\mathbf{i}_v)_{i \in V^k} \mid \mathbf{i}_v \in k_v \text{ for all } v, \mathbf{i}_v \in \mathcal{O}_v \text{ for all but finitely many } v\}$$

The multiplication and addition on V_k are defined componentwise. The basis for the topology on V_k consists of

$$\left\{ \prod_{v \in V^k} \Gamma_v \mid \Gamma_v \text{ is open in } k_v \text{ for all } v, \text{ and } \Gamma_v = \mathcal{O}_v \text{ for all but finitely many } v \right\}.$$

The set of units $V_k^\times$ is not necessarily a topological group when endowed with the subset topology, since inversion is not necessarily continuous. Instead, we give to $V_k^\times$ the subset topology of the injection of $V_k^\times$ into $V_k \times V_k$ sending $\mathbf{i}_v \mapsto (\mathbf{i}_v, \mathbf{i}_v^{-1})$. The idèles J_k are defined to be $V_k^\times$ with this topology, which makes J_k a topological group.

There is an embedding of $k^\times$ into V_k sending $x \mapsto \mathbf{i}(x)$, where $\mathbf{i}(x)_v = x$ for all x. If x is written as $x = a/b$ with $a, b \in \mathcal{O}$, then the only places v where $\mathbf{i}(x)_v \notin \mathcal{O}_v$ are those in $v \in V(b)$; so the adèle $\mathbf{i}(x) \in J_k$ for all $x \in k^\times$. Therefore, we identify $k^\times$ as a subgroup of J_k.

I.3. Background on groups with bounded generation

3.1 Examples and non-examples of groups with bounded generation.

It is trivial to see that every group Γ with bounded generation is finitely generated; it will follow from the following lemma that the converse holds if Γ is abelian.

LEMMA I.3.1. *Let G be a group, and H a subgroup. If H is of finite index in G and has (BG), or if H is normal and both H and G/H have (BG), then G also has (BG).*

PROOF. Since in both cases we assume H has (BG), we can write $H = \langle h_1 \rangle \cdots \langle h_m \rangle$, for some $h_i \in H$. Similarly, if $[G : H] < \infty$, we can write $G = g_1H \cup g_2H \cup \ldots \cup g_nH$ with $g_i \in G$; if G/H has (BG), we can write $G/H = \langle \bar{g}_1 \rangle \cdots \langle \bar{g}_n \rangle$, with $\bar{g}_i \in G/H$. In the latter case, for each i let us fix $g_i \in G$ so that $g_iH = \bar{g}_iH$. Therefore, in both cases we have $G = \langle g_1 \rangle \cdots \langle g_n \rangle \cdot H$, which implies

$$G = \langle g_1 \rangle \cdots \langle g_n \rangle \cdot \langle h_m \rangle \cdots \langle h_1 \rangle,$$

as desired. □

Furthermore, non-abelian groups can also have bounded generation.

LEMMA I.3.2. *Every finitely generated nilpotent group has bounded generation.*

PROOF. If Γ is nilpotent, then there is a filtration $\Gamma = \Gamma_0 \geq \Gamma_1 \geq \ldots \geq \Gamma_r = \{e\}$ for some m, where Γ_{i+1} is defined to be the commutator $[\Gamma_i, \Gamma]$. Since $\Gamma_r = \{e\}$ obviously has (BG), the result will follow by induction from Lemma I.3.1, once we show that each factor Γ_{i-1}/Γ_i has (BG). Since each quotient Γ_{i-1}/Γ_i is Abelian, it suffices to show it is finitely generated. We do this by induction on i.

So suppose Γ is generated by $\{g_1, \ldots, g_M\}$; then Γ_0/Γ_1 is also finitely generated by the images of the g_i's, showing the claim for $i = 1$. Now, suppose $i > 1$. By the induction hypothesis, suppose Γ_{i-1}/Γ_i is generated by $\{\bar{\gamma}_1, \ldots, \bar{\gamma}_N\}$; for each $j \leq N$ fix a preimage $\gamma_j \in \Gamma_{i-1}$ of $\bar{\gamma}_j$. We claim that Γ_i/Γ_{i+1} is generated by the images of $T := \{[g_j, \gamma_k] \mid j \leq M; k \leq N\}$ in Γ_i/Γ_{i+1}. Now we know a priori that Γ_i is generated by $\{[g, \gamma] \mid g \in \Gamma_{i-1}, \gamma \in \Gamma\}$. It remains to show how to write an arbitrary $[g, \gamma]$ as an product of the $[g_j, \gamma_k]$'s. First, suppose $\gamma, \delta \in \Gamma_{i-1}$ and $\epsilon \in \Gamma$, or vice versa. Then using the identity $[\gamma, g] = [g, \gamma]^{-1}$, we have:

- $[\gamma\delta, \epsilon] = \gamma[\delta, \epsilon]\gamma^{-1}[\gamma, \epsilon] = [\gamma, [\delta, \epsilon]] \cdot [\delta, \epsilon][\gamma, \epsilon] \equiv [\gamma, \epsilon][\delta, \epsilon] = [\gamma, \epsilon][\delta, \epsilon] (\text{mod } \Gamma_{i+1})$,
- $[\gamma^{-1}, \epsilon] = \gamma^{-1}[\epsilon, \gamma]\gamma = \left[\gamma^{-1}, [\epsilon, \gamma]\right] \cdot [\epsilon, \gamma]^{-1} \equiv [\epsilon, \gamma] = [\gamma, \epsilon]^{-1} (\text{mod } \Gamma_{i+1})$,

Now if we are given an arbitrary $[g, \gamma] \in [\Gamma_{i-1}, \Gamma] = \Gamma_i$, we can write

$$g = \prod_{k=1}^{r} g_{j_k}^{n_k} \quad \text{and} \quad \gamma = \prod_{l=1}^{s} \gamma_{h_l}^{m_l}.$$

Then, the above identities imply that

$$[g, \gamma] \equiv \prod_{k=1}^{r} \prod_{l=1}^{s} [g_{j_k}, \gamma_{h_l}]^{n_k \cdot m_l} (\text{mod } \Gamma_{i+1}).$$

Therefore, Γ_i/Γ_{i+1} is finitely generated. As explained above, this implies that

□

While it is fairly easy to come up with examples of groups with bounded generation, proving that a group does not have bounded generation is more difficult. We start with the standard (non-)example.

Non-example. *The free group $\mathbb{F}_r$ does not have bounded generation for $r > 1$.*

Proof. If Γ has bounded generation, then we can write $\Gamma = \langle g_1 \rangle \cdots \langle g_n \rangle$ for some n. Then $|\Gamma/\Gamma^m| \leq n^m$ for all $m \geq 0$: in particular, Γ/Γ^m is finite for all m. On the other hand, when Γ is the free group $\mathbb{F}_r$ of rank r, the quotient Γ/Γ^m is called the Burnside group $B(r, m)$ (cf. definition of $\Gamma(r, m)$ in [**NA3**, P. 669]). Furthermore, the solution to Burnside's problem was shown to be negative by P. Novikov and S. Adjan (cf. [**NA1**], [**NA2**], and [**NA3**]), by showing that the Burnside group $B(r, m)$ is infinite for all $r > 1$, and odd $m \geq 4381$. This implies that $\mathbb{F}_r$ cannot have bounded generation for $r > 1$. See [**T**] for a direct argument. □

The above example can be used to show that some other groups do not have bounded generation, via the following partial converse of Lemma I.3.1.

Lemma I.3.3. *Suppose Γ has bounded generation, and $H < \Gamma$ is a subgroup of finite index. Then H also has bounded generation.*

Proof. Suppose that $\gamma_1, \ldots, \gamma_m \in \Gamma$ are the elements such that

$$\Gamma = \langle \gamma_1 \rangle \cdots \langle \gamma_m \rangle.$$

First, we note that it suffices to assume that H is normal. Since otherwise, we can let H' be the kernel of the left multiplication action of Γ on the coset space Γ/H, which, as a kernel, is a normal subgroup. Since left-multiplying the identity coset by any element outside H will result in a non-trivial coset, then H' is contained in H. Furthermore, since G/H is finite, and the set of elements fixing each coset is of finite index in Γ, then $H' = \bigcap_{g \in \Gamma/H} gHg^{-1}$ is also of finite index in Γ, and therefore in H. Since $H' \subseteq H$, the (BG) property of H' will by Lemma I.3.1 imply the (BG) property of H.

So assume that H is normal, and let $r = [\Gamma : H]$. Then the set

$$\Lambda_i := \{\gamma_1^{e_1} \dots \gamma_i^{e_i} \mid 0 \leq e_j \leq r \text{ for } j = 1, \dots, i\}$$

is finite for $i = 1, \dots, m$; and therefore, each of the sets

$$M_{j,i} := \bigcup_{g \in \Lambda_i} \langle (g^{-1}\gamma_j g)^r \rangle \quad \text{for } 1 \leq i, j \leq m$$

is a union of finitely many cyclic groups. Since $[\Gamma : H] = r$, and each of the cyclic groups comprising $M_{j,i}$ is generated by r-th powers, then $M_{j,i}$ is contained in H for all i, j. Set $M_i = \prod_{j=1}^{i} M_{j,i}$.

Furthermore, let us define the set $L_i := \Lambda_i \cdot M_i$ for all $i = 1, \dots, m$. The strategy of the proof is to show that if $h \in L_{i-1}$, then $h \cdot \gamma_i^n \in L_i$ for all n. Since every element $g \in \Gamma$ can be written as $g = \gamma_1^{n_1} \cdots \gamma_m^{n_m}$, this will imply that $L_m = \Gamma$. But, $H \subseteq L_m = \Lambda_m \cdot M_m$, and since $M_m \subseteq H$, then $H \subseteq (\Lambda_m \cap H) \cdot M_m \subseteq H$. Now $\Lambda_m \cap H$ is a finite set, and M_m is a finite product of finite unions of cyclic subgroups of H. This will imply that $H = (\Lambda_m \cap H) \cdot M_m$ has bounded generation.

First we show that if we are given $h \in L_{i-1}$ and an integer $e_i = e'_i + e''_i$, where $0 \leq e'_i < r$ and $r \mid e''_i$, then we have $h \cdot \gamma_i^{e_i} \in L_i$. This is because writing $h = h' \cdot h'' \in \Lambda_{i-1} \cdot M_{i-1}$, we obtain

$$h \cdot \gamma_i^{e_i} = h' \cdot h'' \cdot \gamma_i^{e'_i} \gamma_i^{e''_i} = h' \cdot \gamma_i^{e'_i} \cdot (\gamma_i^{-e'_i} h'' \gamma_i^{e'_i}) \gamma_i^{e''_i}.$$

Since $h' \in \Lambda_{i-1}$ and $0 \leq e_i < r$, then $h' \cdot \gamma_i^{e'_i} \in \Lambda_i$. Since $h'' \in M_{i-1} = \prod_{j=1}^{i-1} M_{j,i-1}$ and $0 \leq e'_i < r$, then $\gamma_i^{-e'_i} h'' \gamma_i^{e'_i} \in \prod_{j=1}^{i-1} M_{j,i}$. Furthermore, since $r \mid e''_i$, we have $\gamma_i^{e''_i} \in \langle \gamma_i^r \rangle \subseteq M_{i,i}$. This implies that $h \cdot \gamma_i^{e_i} \in \Lambda_i \cdot \prod_{j=1}^{i-1} M_{j,i} \cdot M_{i,i} = L_i$. □

Corollary I.3.4. *If H is a subgroup of G of finite index, then H has (BG) if and only if G has (BG).*

Now, we will apply Lemma I.3.3 to show

Corollary I.3.5. *The group* $\mathrm{SL}_2(\mathbb{Z})$ *does not have bounded generation.*

Proof. Consider the homomorphism ϕ of the free group $\mathbb{F}_2 = \langle x, y \rangle$ to $\mathrm{SL}_2(\mathbb{Z})$ sending $x \mapsto \mathrm{E}_{12}(2)$ and $y \mapsto \mathrm{E}_{21}(2)$. By Lemma I.3.3, it suffices to show that $\mathrm{Im}(\phi)$ is free, and that $\mathrm{Im}(\phi)$ is of finite index in $\mathrm{SL}_2(\mathbb{Z})$. That it is free is usually proved by showing that its action on the hyperbolic plane is free; however, there is a more elementary argument. For $i \geq 0$, let H_i be set of images of all words of length i which start with x, and H'_i be set of images of all words of length i which start with y.

Clearly, $|(I_2)_{11}| > |(I_2)_{12}|$; a simple induction argument shows that for any $A \in H_n$, if n is even then $|(A)_{11}| > |(A)_{12}|$; otherwise $|(A)_{12}| > |(A)_{11}|$. Consequently, one shows by induction that for each $A \in H_n$, $\max(|(A)_{11}|, |(A)_{12}|) \geq n$, implying $A \neq I_2$ when $n > 0$; similarly, $I_2 \notin H'_n$ for $n > 0$. This means that the only way to map an element of $\mathbb{F}_2$ to the identity matrix is, if it is the trivial word. This implies that $\mathrm{Im}(\phi)$ is free.

On the other hand, the elements $\alpha := \mathrm{E}_{12}(1)$ and $\beta := \mathrm{E}_{21}(-1)$ generate $\mathrm{SL}_2(\mathbb{Z})$. Let $\bar{\alpha}$ and $\bar{\beta}$ be the images of α and β, respectively, in $\mathrm{PSL}_2(\mathbb{Z})$. Then they satisfy the identities

$$\bar{\alpha}\bar{\beta}^2\bar{\alpha}^{-1} = \bar{\beta}^{-2} \quad \text{and} \quad \bar{\beta}\bar{\alpha}^2\bar{\beta}^{-1} = \bar{\alpha}^{-2}.$$

Let H be the image of ϕ in $\mathrm{PSL}_2(\mathbb{Z})$; the above implies that H is normal. One readily checks that

$$\{I_2, \bar{\alpha}, \bar{\beta}, \bar{\alpha}\bar{\beta}, \bar{\beta}\bar{\alpha}, \bar{\alpha}\bar{\beta}\bar{\alpha} = \bar{\beta}\bar{\alpha}\bar{\beta}\}$$

forms the full set of representatives of $\mathrm{PSL}_2(\mathbb{Z})/H$, so $[\mathrm{PSL}_2(\mathbb{Z}) : H] = 6$. (In fact, H is precisely the kernel of the map $\mathrm{PSL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/2\mathbb{Z})$.) Since $[\mathrm{SL}_2(\mathbb{Z}) : \mathrm{PSL}_2(\mathbb{Z})] = 2$, then the index of $\mathrm{Im}(\phi)$ in $\mathrm{SL}_2(\mathbb{Z})$ is at most 12. Since $\mathrm{Im}(\phi)$ is free, and the order of $-I_2$ is 2, then $-I_2 \notin \mathrm{Im}(\phi)$, and the index of $\mathrm{Im}(\phi)$ in $\mathrm{SL}_2(\mathbb{Z})$ is equal to 12. □

While it was straightforward to show that finitely generated nilpotent groups have bounded generation, in fact bounded generation is not restricted to nilpotent or solvable groups. As outlined next, many Chevalley groups, which are non-solvable, also have bounded generation.

3.2 Previous results on bounded generation of linear groups. The quest to validate bounded generation of non-solvable Chevalley groups has a considerable history. It has been known that if the ring of S-integers $\mathcal{O} = \mathcal{O}_{k,S}$ has infinitely many units, then the group $\Gamma = \mathrm{SL}_n(\mathcal{O})$ is generated by elementary matrices for any n (see [**Va**] for $n = 2$, and [**BMS**] for $n \geq 3$). However, the question of whether the width with respect to elementary matrices is actually bounded has yielded many papers giving a partial answer. The first result in this direction was given by G. Cooke and P. J. Weinberger [**CW**], who established that the width equals 5, assuming the truth of a suitable form of the Generalized Riemann Hypothesis, which still remains unproven.

The subsequent research initially focused on bounded generation of SL_n for $n \geq 3$. The first unconditional proof of bounded generation of special linear groups was given by D. Carter and G. Keller [**CK1**], who showed that $\mathrm{SL}_n(\mathcal{O})$ for $n \geq 3$ is boundedly generated by elementary matrices for any ring $\mathcal{O}$ of algebraic integers. O. I. Tavgen [**T**] was able to generalize their proof to most (regular and twisted) Chevalley groups of rank > 1; note that these results again do not apply to SL_2. The first unconditional proof which proved the bounded generation of $\mathrm{SL}_2(\mathcal{O}_{k,S})$ in all cases where it holds was given by D. Carter, G. Keller and E. Paige in an unpublished preprint; their argument was streamlined and made available to the public by D. W. Morris [**MCKP**]. Around this time, also I. V. Erovenko and A. S. Rapinchuk [**ER**] also were able to establish bounded generation of S-arithmetic subgroups of some isotropic, but not necessarily split or quasi-split, orthogonal groups. Furthermore, A. Heald [**He**] was able to establish bounded generation of special linear groups over certain quaternion algebras, and also of special unitary groups over some rings.

Soon after Carter and Keller's paper [**CK1**], B. Liehl [**L2**] applied some of their techniques to show (BG) of $\mathrm{SL}_2(\mathcal{O}_{k,S})$, under heavy restrictions on k. After this result, no new results appeared, until [**LM**] (see also [**M**]) showed using analytic tools, that $\mathrm{SL}_2(\mathcal{O}_{k,S})$ has a width of at most 5 for an arbitrary k, essentially by proving that Artin's Primitive Root Conjecture holds in this case; however, this result required that $|S| \geq \max(5, 2[k : \mathbb{Q}] - 3)$. The argument in [**MCKP**] finally proved (BG) of $\mathrm{SL}_2(\mathcal{O}_{k,S})$ for an arbitrary k and S. However, it was based on model theory, and provides no explicit bound on the number of elementary matrices required; besides, it uses difficult results from additive number theory. In [**Vs**], M. Vsemirnov proved Theorem II.3.6 for $\mathcal{O} = \mathbb{Z}[1/p]$ using the results of D. R. Heath-Brown [**Hb**] on Artin's Primitive Root Conjecture (in a broad sense, this proof develops the initial approach of Cooke and Weinberger [**CW**]); his bound on the number of elementaries required is ≤ 5. Subsequently, B. Sury re-worked the argument from [**Vs**] to avoid the use of [**Hb**] in an unpublished note. These notes were the beginning of the work which led to our proof of Theorem II.3.6 in the general case (cf. [**MRS**]). It should be noted that our proof used only standard results from number theory, and is relatively short and constructive with an explicit bound which is independent of the field k and the set S. This, in particular, implies that Theorem II.3.6 remains valid for any *infinite* S.

Next, it should be pointed out that the assumption that the unit group $\mathcal{O}^\times$ is infinite is *necessary* for the bounded generation of $\mathrm{SL}_2(\mathcal{O})$, hence cannot be omitted. Indeed, it follows from Dirichlet's Unit Theorem [**CF**, §2.18] that $\mathcal{O}^\times$ is finite only when $|S| = 1$ which happens precisely when S is the set of archimedian valuations in the following two cases:

1) $k = \mathbb{Q}$ and $\mathcal{O} = \mathbb{Z}$. In this case, as shown in Corollary I.3.5, the group $\mathrm{SL}_2(\mathbb{Z})$ is generated by the elementaries, but has a nonabelian free subgroup of finite index, which prevents it from having bounded generation.

2) $k = \mathbb{Q}(\sqrt{-d})$ for some square-free integer $d \geq 1$, and $\mathcal{O}_d$ is the ring of algebraic integers integers in k. According to [**GS**], the group $\Gamma = \mathrm{SL}_2(\mathcal{O}_d)$ has a finite index subgroup that admits an epimorphism onto a nonabelian free group, hence again cannot possibly be boundedly generated. Moreover, P. M. Cohn [**Co**] showed that if $d \notin \{1, 2, 3, 7, 11\}$ then Γ is not even generated by elementary matrices.

The upper bound on the number of factors required to write every matrix in $\mathrm{SL}_n(\mathcal{O})$ as a product of elementaries given in [**CK1**] is $\frac{1}{2}(3n^2 - n) + 68\Delta - 1$, where Δ is the number of prime divisors of the discriminant of k. In particular, this estimate depends on the field k. Using our Theorem II.3.6, one shows in all cases where the group of units $\mathcal{O}^\times$ is infinite, this estimate can be improved to $\frac{1}{2}(3n^2 - n) + 4$, hence made independent of k – see Corollary II.3.7. The situation not covered by this result are when $\mathcal{O}$ is either $\mathbb{Z}$ or the ring of integers in an imaginary quadratic field, as outlined above. The case $\mathcal{O} = \mathbb{Z}$ for $n \geq 3$ was treated in [**CK2**] with an estimate $\frac{1}{2}(3n^2 - n) + 36$, so only in the case of imaginary quadratic fields the question of the existence of a bound on the number of elementaries independent of the field k remains open.

From a more general perspective, Theorem II.3.6 should be viewed as a contribution to the sustained effort aimed at proving that all higher rank lattices are boundedly generated as abstract groups. Given the implications of (BG) described in Subsection I.3.3, we would like to point out the following consequence of Theorem II.3.6.

Corollary I.3.6. *Let $\mathcal{O} = \mathcal{O}_{k,S}$ be the ring of S-integers, in a number field k. If the group of units $\mathcal{O}^\times$ is infinite, then the group $\Gamma = \mathrm{SL}_2(\mathcal{O})$ has bounded generation.*

3.3 Properties and applications of groups with bounded generation. The interest in bounded generation of a group stems from the fact that while being purely combinatorial in nature, it is known to have a number of far-reaching consequences for the structure and representations of a group, particularly if the latter is S-arithmetic. See §I.4 for definitions of terms used in this section.

One application is to representation rigidity. Suppose Γ is an arbitrary finitely generated group, and fix an arbitrary set of generators $\Gamma = \langle \gamma_1, \ldots, \gamma_r \rangle$. Then we have a surjection from the free group $\mathbb{F}_r = \langle x_1, \ldots, x_r \rangle$ to Γ sending $x_1 \mapsto \gamma_r$; let N be the kernel of this map. Then for a fixed integer $d \geq 1$, any d-dimensional complex representation ρ of Γ can be defined by specifying the image of the generators $(\rho(\gamma_1), \ldots, \rho(\gamma_r)) \in \mathrm{GL}_d(\mathbb{C})^r$; conversely, any r-tuple $(M_1, \ldots, M_r) \in \mathrm{GL}_d(\mathbb{C})^r$ corresponds to a representation of Γ if and only if these elements satisfy the relations $w(M_1, \ldots, M_r) = 1 \in \Gamma$ for all $w \in N$. Therefore, the set

$$R_d(\Gamma) = \{(M_1, \ldots, M_r) \in \mathrm{GL}_d(\mathbb{C})^r \mid w(M_1, \ldots, M_r) = 1 \text{ for all } w \in N\}$$

will parametrize all the d-dimensional complex representations of Γ. It is an affine algebraic variety called the *representation variety* of d-dimensional representations of Γ. Up to a biregular isomorphism, this variety is independent of the generators $\{\gamma_1, \ldots, \gamma_r\}$ of Γ (cf. [**PR2**, 2.4.7]).

Furthermore, there is a natural conjugation action of $\mathrm{GL}_d(\mathbb{C})$ on $R_d(\Gamma)$. The orbits under this action correspond to equivalence classes of representations. Then the categorical quotient of $R_d(\Gamma)$ under this action is an algebraic variety, called the *character variety* [**PR2**, 2.4.7] of Γ, denoted $X_d(\Gamma)$. The elements of $X_d(\Gamma)$ are in a bijective correspondence with the Zariski closures of the $\mathrm{GL}_d(\mathbb{C})$-orbits under the conjugation action; furthermore, the closure of each orbit contains a unique fully reducible representation. Therefore, $X_d(\Gamma)$ parametrizes the set of d-dimensional complex completely reducible representations of Γ. Then, a group Γ is defined to be *semisiply rigid* (a.k.a. SS-rigid) when the dimension of $X_d(\Gamma)$ is 0 for all $d \geq 1$. This is equivalent to the existence of at most finitely many equivalence classes of complex completely reducible representations in each dimension.

There are several obstructions to a group being SS-rigid. First, if the abelianization Γ^{ab} is infinite, then the fact that Γ is finitely generated implies that Γ^{ab} has a free factor, say corresponding to $\gamma \in \Gamma$. This means for each $d \geq 1$ and each $M \in \mathrm{GL}_d(\mathbb{C})$, the map $\gamma \mapsto M$ can be extended to a representation of Γ, thereby resulting in infinitely many equivalence classes of complex completely reducible representations of Γ in each dimension – meaning that Γ is not SS-rigid.

The group $\mathrm{SL}_2(\mathbb{Z})$ is an example of a group whose abelianization is finite, of size 12. However, it has a finite-index subgroup H, namely the one generated by the elementary matrices $\mathrm{E}_{12}(2)$ and $\mathrm{E}_{21}(2)$ (cf. Corollary I.3.5), and therefore $X_n(H)$ has a positive-dimensional subvariety in each dimension. One shows that after inducing these representations to $\mathrm{SL}_2(\mathbb{Z})$, at most finitely many will be in the closure of any given orbit, and therefore there will be infinitely many classes of representations in $X_n(\mathrm{SL}_2(\mathbb{Z}))$ for a large enough n. Again, we see that $\mathrm{SL}_2(\mathbb{Z})$ is not SS-rigid. Applying analogous reasoning using induced representations, one shows that if $\Delta \subset \Gamma$ is a finite-index subgroup with an infinite abelianization, then again Γ is not SS-rigid.

Therefore, the finiteness of the abelianization of Γ, and of abelianizations of all of its finite-index subgroups, are necessary for Γ to be SS-rigid. However, it is not sufficient. The group $\mathrm{SL}_3(\mathbb{Z}[x])$ is a Kazhdan group [**EJ**], and therefore the abelianizations of it and of all its finite-index subgroups are finite. On the other hand, when we are considering the three-dimensional representations $\mathrm{SL}_3(\mathbb{Z}[x]) \to \mathrm{GL}_3(\mathbb{C})$, the element x can be sent to any element of $\mathbb{C}$, which will yield infinitely many closed orbits in dimension 3, and therefore imply that the group $\mathrm{SL}_3(\mathbb{Z}[x])$ is not SS-rigid. Rapinchuk ([**R**]; see also Appendix A in [**PR2**]) showed that there is a number of additional conditions, so that any group Γ satisifying one of them in addition to the above necessary conditions, will be SS-rigid. If Γ is boundedly generated, then it satisfies one of those additional conditions; so as long as the abelianizations of Γ and of all its finite-index subgroups are finite, Γ will be SS-rigid.

Another application of bounded generation is to the Congruence Subgroup Problem. Let $\mathbf{G}$ be an algebraic group over a number field k, S be a finite subset of V^k containing V^k_∞, and let $\mathcal{O} = \mathcal{O}_{k,S}$. Fix a k-defined embedding $\mathbf{G} \to \mathrm{GL}_n$. Then, we define $\mathbf{G}_k$ to be the set of k-rational points (i.e., $\mathbf{G} \cap \mathrm{GL}_n(k)$; cf. [**PR2**, §2.1]), and $\mathbf{G}_\mathcal{O}$ to be $\mathbf{G}_k \cap \mathrm{GL}_n(\mathcal{O})$.

For any (non-zero) ideal $\mathfrak{a}$ of $\mathcal{O}$, we define the congruence subgroup of modulus $\mathfrak{a}$ to be

$$\mathbf{G}(\mathcal{O}, \mathfrak{a}) := \{g \in \mathbf{G}_\mathcal{O} \mid g \equiv I_n (\mathrm{mod}\ \mathfrak{a})\}.$$

Since all non-zero ideals of $\mathcal{O}$ are of finite index, it is straightforward to see that the index $[\mathbf{G}_{\mathcal{O}} : \mathbf{G}(\mathcal{O}, \mathfrak{a})]$ is finite. The Congruence Subgroup Problem asks whether every finite-index normal subgroup contains some congruence subgroup.

Serre proposed a way of looking at this problem, in terms of topologies on $\mathbf{G}_k$. Define the S-arithmetic topology τ_a on $\mathbf{G}_k$ by letting the set of finite-index normal subgroups of $\mathbf{G}_{\mathcal{O}}$ comprise the basis of neighbourhoods of the identity. Similarly, the S-congruence topology τ_c on $\mathbf{G}_k$ is defined by letting the set of all congruence subgroups of $\mathbf{G}_{\mathcal{O}}$ comprise the basis of neighbourhoods of the identity. Choosing a different k-defined embedding of $\mathbf{G}$ yields a different $\mathbf{G}_{\mathcal{O}}$; however, one can show that the topologies τ_a and τ_c are independent of the embedding. One can then complete $\mathbf{G}$ with respect to both τ_a and τ_c; these completions are denoted $\hat{\mathbf{G}}^S$ and $\bar{\mathbf{G}}^S$, respectively. Then, since τ_a is finer than τ_c, we have a natural map $\hat{\mathbf{G}}^S \to \bar{\mathbf{G}}^S$, which can be shown to be a surjection. The kernel of this surjection is called the S-congruence kernel of $\mathbf{G}$.

Section III.2 contains a proof of Bass's result that, if $\mathbf{G} = \mathrm{SL}_n(\mathcal{O})$, where $\mathcal{O}$ is a ring satisfying Bass's stable range condition $\mathtt{SR}_{n-1}$, then every normal subgroup contains the group $\mathrm{E}_n(\mathcal{O}, \mathfrak{a})$, for some ideal $\mathfrak{a}$ of $\mathcal{O}$ – which is also of finite index in $\mathbf{G}(\mathcal{O}, \mathfrak{a})$ in case $\mathfrak{a}$ is of finite index in $\mathcal{O}$. In this case, the S-congruence kernel is the quotient $\varprojlim\,(\mathbf{G}(\mathcal{O}, \mathfrak{a})/\mathrm{E}_n(\mathcal{O}, \mathfrak{a}))$, where the limit is taken over all (non-zero) ideals of $\mathcal{O}$.

A positive answer to the Congruence Subgroup Problem is equivalent to the topologies τ_a and τ_c being equivalent, meaning that the S-congruence kernel is trivial. For many applications, the finiteness of the S-congruence kernel is just as good as triviality. Lubotzky, Platonov, and Rapinchuk ([**Lu**], [**PR1**]) were able to show that if the group $\mathbf{G}_{\mathcal{O}_{k,S}}$ has bounded generation, then under certain easier-to-check assumptions, the S-congruence kernel is finite. The exact statement is:

Theorem I.3.7. *Let k be an algebraic number field, $\mathbf{G}$ an absolutely simple simply connected algebraic group over a number field k, and T be the set of (non-archimedean) places where $\mathbf{G}$ is k_v-anisotropic. Let S be a finite set of places, containing the infinite ones. Suppose $\mathbf{G} = \mathbf{G}_{\mathcal{O}_{k,S}}$ has bounded generation, and $S \cap T = \emptyset$. If the Margulis-Platonov Conjecture holds for $\mathbf{G}_k$, then the S-congruence kernel is finite.*

The Margulis-Platonov conjecture says that for all non-central subgroups $N \leq \mathbf{G}_k$, there exists a normal open subgroup $W \trianglelefteq \prod_{v \in T} \mathbf{G}_{k_v}$ so that $N = \mathbf{G}_k \cap W$. This conjecture has been shown to hold in many cases(cf. [**PR2**]).

We note that combining bounded generation of $\mathrm{SL}_2(\mathcal{O})$ with the results of [**Lu**], [**PR1**], one obtains an alternative proof of the centrality of the congruence kernel for $\mathrm{SL}_2(\mathcal{O})$ (provided that $\mathcal{O}^\times$ is infinite), originally established by J.-P. Serre [**S1**].

A third application of bounded generation is to the Commensurator-Normalizer Property. If Γ is a higher-rank S-arithmetic group, such as $\mathrm{SL}_3(\mathbb{Z})$ or $\mathrm{SL}_2(\mathbb{Z}[1/p])$ where p is a prime, then Margulis's Normal Subgroup Theorem states that any non-central normal subgroup $\Lambda \subset \Gamma$ is of finite index. Zimmer was interested in studying the generalization of this phenomenon to commensurated groups. Two subgroups Λ_1 and Λ_2 of an abstract group G are commensurable with each other (denoted $\Lambda_1 \sim_c \Lambda_2$), if the index $[\Lambda_i : \Lambda_1 \cap \Lambda_2]$ is finite for $i = 1, 2$. Then for an abstract group G containing Γ, we say that $\Lambda \subseteq G$ is commensurated by Γ if $\gamma\Lambda\gamma^{-1} \sim_c \Lambda$ for all $\gamma \in \Gamma$.

It is clear that the exact statement of Margulis's N.S.T. is no longer true for commensurated subgroups. For example, already over $\mathbb{Z}$, we have $\Gamma = \mathrm{SL}_3(\mathbb{Z})$ is commensurated by $\mathrm{SL}_3(\mathbb{Z}[1/p])$;

nevertheless, it is neither central nor of finite index. However, in this case, $\mathbb{Z}[1/p]$ is a localization of $\mathbb{Z}$.

This phenomenon generalizes to other algebraic groups. Let $\mathbf{G}$ be a simple algebraic group, k be a global field, and S be a finite set of places containing the infinite ones. As long as $\Gamma \sim_c \mathbf{G}_{\mathcal{O}_{k,S}}$, then any group $\Gamma' \sim_c \mathbf{G}_{\mathcal{O}_{k,S'}}$ for $V_\infty^k \subseteq S' \subseteq S$ already is commensurated by Γ – even though it is of infinite index in Γ when $S' \neq S$. Zimmer conjectured that these are the only (infinite) ones. Shalom and Willis (cf. [**SW**]) used bounded generation to show that Zimmer's conjecture holds for the Chevalley groups $\Gamma = \mathrm{SL}_2(\mathcal{O})$ having (BG).

One other application of bounded generation is to calculating the Kazhdan's constant of special linear groups. Suppose G is a topological group, and $(\pi, \mathcal{H})$ is a continuous unitary representation of G. For a subset $K \subseteq G$ and $\varepsilon > 0$, we say that a vector $v \in \mathcal{H}$ is (K, ε)-invariant if

$$||\pi(g)v - v|| < \varepsilon \cdot ||v|| \text{ for all } g \in K.$$

Then G is said to have Property (T) is there exists a *compact* K, and $\varepsilon > 0$, so that every representation with (K, ε)-invariant vectors contains some non-zero G-invariant vector. In this case, this ε is called a *Kazhdan constant* for G. For any integer n, number field k, and a finite set of places S containing the infinite ones, Burger [**Bu**] was able to compute a Kazhdan constant for $\mathrm{SL}_n(\mathcal{O}_{k,S})$ when $n \geq 3$. Shalom [**Sh**] followed Burger's proof by computing a Kazhdan constant when K is the set of elementary matrices, using the bound on the width of $\mathrm{SL}_n(\mathcal{O}_{k,S})$ with respect to elementary matrices in his proof.

I.4. Brief review of algebraic number theory

4.1. Splitting of primes in Galois extensions. Let L be a number field, i.e. a finite extension of $\mathbb{Q}$, and R a (non-zero) subring. Recall that an element $\alpha \in L$ is said to be *integral* over a ring R if it is a root of a minimal polynomial over R. An equivalent condition is, that $R[\alpha]$ is contained in some finitely generated R-module. Using this equivalent condition, one shows that the set of all elements of L integral over R is a ring. This ring is called the *integral closure* of R in L. When the ring R coincides with its integral closure, then we say that R is *integrally closed* in L; a ring is said to be integrally closed if it is integrally closed in its fraction field. For example, $\mathbb{Z}$ is integrally closed.

In each number field L, the integral closure of $\mathbb{Z}$ is the smallest (non-zero) integrally closed subring, and is called the *ring of integers* of L, denoted $\mathcal{O}_L$. It is a free $\mathbb{Z}$-module of rank equal to the index $[L : \mathbb{Q}]$.

One distinguishing property of $\mathbb{Z}$ is that each of its elements has a unique factorization as a product of prime elements. Arbitrary rings of integers $\mathcal{O}_L$ no longer have this property, since they can have irreducible elements which do not generate a prime ideal; however, they satisfy a weaker condition – every ideal of $\mathcal{O}_L$ has a unique decomposition as a product of prime ideals. Noetherian integrally closed rings satisfying this condition are called *Dedekind* (cf. [**Hu**, P. 405]). Thus for a Dedekind ring, unique factorization is equivalent to being a principal ideal domain.

Theorem I.4.1. [**ZS**, Ch. V, §1] *If L is a number field, then the ring of integers $\mathcal{O}_L$ is a Dedekind ring.*

The fact that each ideal has a unique factorization into prime ideals implies that every prime ideal is maximal. This would mean that any two prime ideals are comaximal, and the Chinese Remainder Theorem applies:

THEOREM I.4.2. *Suppose $\mathfrak{n}_1, \ldots, \mathfrak{n}_r$ are pairwise comaximal ideals of a ring R, and let $\mathfrak{a}$ be their product. If the index $[R : \mathfrak{a}]$ is finite, then there is a ring isomorphism*

$$R/\mathfrak{a} \to \bigoplus_{i=1}^{r} R/\mathfrak{n}_i.$$

PROOF. First, we note this ring homomorphism is well-defined, since if $a \equiv b (\mathrm{mod}\ \mathfrak{a})$, then $a \equiv b (\mathrm{mod}\ \mathfrak{n}_i)$ for all i. Since the ideals $\mathfrak{n}_i$ are pairwise coprime, then the converse is also true, so this ring homomorphism is injective. Finally, it is surjective, since the index of $\mathfrak{a}$ in R is finite and equal to the product of the indices in R of the $\mathfrak{n}_i$'s. □

When this theorem applies, it allows us to replace congruences modulo powers of prime ideals with a single congruence modulo their product. And conversely, if we are given a congruence modulo an arbitrary ideal $\mathfrak{a}$ of $\mathcal{O}_L$, then since $\mathcal{O}_L$ is a Dedekind ring, the ideal $\mathfrak{a}$ has a unique factorization as a product of prime ideals, say $\mathfrak{a} = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_n^{e_n}$. The Chinese Remainder Theorem then implies that solving an equation modulo $\mathfrak{a}$ is equivalent to solving it modulo $\mathfrak{p}_i^{e_i}$ for each $i = 1, \ldots, n$.

In order to determine this factorization, however, we need to know the set of prime ideals of $\mathcal{O}_L$. It is known that for each (non-zero) prime ideal $\mathfrak{p}$ of $\mathcal{O}_L$, the ideal $\mathfrak{p} \cap \mathbb{Z}$ is always a (non-zero) prime ideal of $\mathbb{Z}$. Since the set of (non-zero) prime ideals of $\mathbb{Z}$ is the set of $p\mathbb{Z}$ for each prime number p, the problem of finding the set of all prime ideals of $\mathcal{O}_L$ reduces to the problem of determining for each $\mathbb{Z}$-prime p how the ideal $p\mathcal{O}_L$ splits as a product of prime ideals.

More generally, suppose $k \subset L$ are number fields; in this case we have $\mathcal{O}_k \subset \mathcal{O}_L$. For every prime ideal $\mathfrak{p}$ of $\mathcal{O}_k$, the ideal $\mathfrak{p}\mathcal{O}_L$ will have a unique decomposition as a product of prime ideals of $\mathcal{O}_L$, say

$$\mathfrak{p}\mathcal{O}_L = \mathfrak{P}_1^{e(\mathfrak{P}_1/\mathfrak{p})} \mathfrak{P}_2^{e(\mathfrak{P}_2/\mathfrak{p})} \ldots \mathfrak{P}_r^{e(\mathfrak{P}_r/\mathfrak{p})}.$$

The exponents $e(\mathfrak{P}_i/\mathfrak{p})$ are called *ramification indices*.

Suppose $L/K/k$ is a tower of field extensions, and let us fix one such $\mathfrak{P}_i$. Let $\tilde{\mathfrak{p}}$ be the prime of $\mathcal{O}_K$ lying below $\mathfrak{P}_i$. We observe that we can write

$$\mathfrak{p}\mathcal{O}_K = \tilde{\mathfrak{p}}^{e(\tilde{\mathfrak{p}}/\mathfrak{p})} \cdot \mathfrak{Q} \text{ with } \mathfrak{Q} \text{ coprime to } \tilde{\mathfrak{p}}, \text{ and}$$

$$\tilde{\mathfrak{p}}\mathcal{O}_L = \mathfrak{P}_i^{e(\mathfrak{P}_i/\tilde{\mathfrak{p}})} \cdot \mathfrak{Q}' \text{ with } \mathfrak{Q}' \text{ coprime to } \mathfrak{P}_i.$$

Then $e(\tilde{\mathfrak{p}}/\mathfrak{p})$ and $e(\mathfrak{P}_i/\tilde{\mathfrak{p}})$ are the intermediate ramification indices. Since also we have

$$\mathfrak{p}\mathcal{O}_L = \mathfrak{P}_i^{e(\mathfrak{P}_i/\mathfrak{p})} \cdot \mathfrak{q} \text{ with } \mathfrak{q} \text{ coprime to } \mathfrak{P}_i,$$

then the ramification indices are multiplicative, meaning they satisfy

$$e(\mathfrak{P}_i/\tilde{\mathfrak{p}}) \cdot e(\tilde{\mathfrak{p}}/\mathfrak{p}) = e(\mathfrak{P}_i/\mathfrak{p})$$

A prime $\mathfrak{p}$ is said to be unramified in L if all of the indices $e(\mathfrak{P}_i/\mathfrak{p})$ for $i = 1, \ldots, r$ are 1. It is known [**FT**, Thm. 22] that a prime number p is ramified in L if and only if it divides the discriminant of $L/\mathbb{Q}$, and so there are only finitely many such primes. Because the ramification indices are multiplicative, the only primes of $\mathcal{O}_k$ which can be ramified in L, are those lying above a $\mathbb{Z}$-prime ramified in L, and so again all but finitely many $\mathcal{O}_k$-primes are unramified in L.

Since $\mathfrak{P}_i$ is a prime ideal of $\mathcal{O}_L$, the quotient $\mathcal{O}_L/\mathfrak{P}_i$ is a field, called the residue field. Since L is a number field, $\mathcal{O}_L/\mathfrak{P}_i$ is finite. Similarly, $\mathcal{O}_k/\mathfrak{p}$ is a finite field. Then the index $[\mathcal{O}_L/\mathfrak{P}_i : \mathcal{O}_k/\mathfrak{p}]$ is also finite; it is called the *residual degree*, and is denoted by $f(\mathfrak{P}_i/\mathfrak{p})$. For a tower of field extensions, the residual degree is by definition also multiplicative. The ramification indices $e(\mathfrak{P}_i/\mathfrak{p})$

and the residual degrees $f(\mathfrak{P}_i/\mathfrak{p})$ do not need to equal each other for $i = 1, \dots, r$; however, they satisfy the formula

$$\sum_{i=1}^{r} e(\mathfrak{P}_i/\mathfrak{p}) f(\mathfrak{P}_i/\mathfrak{p}) = [L : k]. \tag{1}$$

However, when the extension L/k is Galois, then the Galois group $\mathrm{Gal}(L/k)$ acts transitively on the set of $\mathfrak{P}_i$'s [**CF**, Prop. 2 of §I.10]. Consequently, for a given $\mathfrak{p}$, all the $e(\mathfrak{P}_i/\mathfrak{p})$'s must be equal; their value will be denoted simply e; similarly, all the $f(\mathfrak{P}_i/\mathfrak{p})$'s equal the same value, which is denoted f. In this case, the formula (1) simplifies to become $e \cdot f \cdot r = [L : k] = |\mathrm{Gal}(L/k)|$.

For the remainder of this section, we will assume that the extension L/k is Galois. Let $\mathfrak{P}$ denote the fixed prime $\mathfrak{P}_i$; then the set of all elements of $\mathrm{Gal}(L/k)$ which fix $\mathfrak{P}$ forms a subgroup of $\mathrm{Gal}(L/k)$, called the *decomposition group* $G(\mathfrak{P})$. Since there are r ideals $\{\mathfrak{P}_1, \dots, \mathfrak{P}_r\}$ above $\mathfrak{p}$, and the Galois group acts transitively on this set, then $|G(\mathfrak{P})| = |\mathrm{Gal}(L/k)|/r = ef$. Let $\bar{L} = \mathcal{O}_L/\mathfrak{P}$ be the residue field of L, and $\bar{k} = \mathcal{O}_k/\mathfrak{p}$ be the residue field of k. Suppose that $\bar{k}$ is isomorphic to the finite field $\mathbb{F}_q$, where q is a power of a prime. Since L is a finite extension of K, then $\bar{L}$ is isomorphic to $\mathbb{F}_{q^d}$, for some $d \geq 1$, so we have a canonical surjection

$$G(\mathfrak{P}) = \mathrm{Gal}(L/k) \xrightarrow{\phi} \mathrm{Gal}(\bar{L}/\bar{k}) \cong \mathrm{Gal}(\mathbb{F}_{q^d}/\mathbb{F}_q).$$

The kernel of the above map ϕ consists of the elements that fix $\bar{L} = \mathcal{O}_L/\mathfrak{P}^1$, and is called the *inertia group* of $\mathfrak{P}$, denoted $I(\mathfrak{P})$. Since the map is onto, and the degree of the extension $\bar{L}/\bar{K}$ equals to f, then $d = f$, and the size of $I(\mathfrak{P})$ is $|\mathrm{Gal}(L/K)|/f = e$, which is the ramification index. Therefore, the ideal $\mathfrak{p}$ is unramified in L precisely when the inertia group is trivial.

The inertia group has further subgroups, called the higher ramification groups, which are obtained by looking at elements which act trivially on $\mathcal{O}_L/\mathfrak{P}^h$ for various $h > 1$, and which control the type of ramification that occurs; however, since we always avoid ramified primes in our arguments, we will not discuss the inertia group here.

So let us assume that $\mathfrak{p}$ is unramified in L; then the inertia group is trivial, and so the map ϕ is an isomorphism. From field theory, we know that the extension of finite fields $\mathbb{F}_{q^d}/\mathbb{F}_q$ is cyclic, so $G(\mathfrak{P})$ is generated by a single element – namely, the pre-image of $(x \mapsto x^q) \in \mathrm{Gal}(\mathbb{F}_{q^d}/\mathbb{F}_q)$. When restricted to the residue fields, this generator sends $x + \mathfrak{P}$ to $x^q + \mathfrak{P}$; its image in $\mathrm{Gal}(L/k)$ is called the *Frobenius automorphism* of $\mathfrak{P}$ over $\mathfrak{p}$, denoted $\mathrm{Fr}_{L/k}(\mathfrak{P}|\mathfrak{p})$. So for example, when the prime $\mathfrak{p}$ splits completely in L, then $G(\mathfrak{P})$ is trivial; therefore, the Frobenius $\mathrm{Fr}_{L/k}(\mathfrak{P}|\mathfrak{p})$ acts on L trivially, so equals the identity element of $\mathrm{Gal}(L/k)$. The converse is also true, since if $G(\mathfrak{P})$ is non-trivial, then so must be its generator.

The Frobenius automorphism possesses functorial properties. For example, if $L/K/k$ is a tower of field extensions and $\tilde{\mathfrak{p}}$ is the prime of $\mathcal{O}_K$ lying below $\mathfrak{P}$, then $\mathrm{Fr}_{K/k}(\tilde{\mathfrak{p}}|\mathfrak{p}) = \mathrm{Fr}_{L/k}(\mathfrak{P}|\mathfrak{p})_{|K}$, since the generator of a Galois group over a bigger field still is a generator when restricted to a smaller field.

For any other prime $\mathfrak{P}'$ above $\mathfrak{p}$, there exists an element $\tilde{\sigma} \in G/G(\mathfrak{P})$, so that for any $\sigma \in G$ in the left coset of $\tilde{\sigma}$ in $G/G(\mathfrak{P})$, we have $\sigma(\mathfrak{P}) = \mathfrak{P}'$. Then σ^{-1} takes $x + \mathfrak{P}'$ to $\sigma^{-1}(x) + \mathfrak{P}$, $\mathrm{Fr}_{L/k}(\mathfrak{P}|\mathfrak{p})$ sends $\sigma^{-1}(x) + \mathfrak{P}$ to $\sigma^{-1}(x^q) + \mathfrak{P}$, and σ takes $\sigma^{-1}(x^q) + \mathfrak{P}$ back to $x^q + \mathfrak{P}'$; so the composition $\sigma \mathrm{Fr}_{L/k}(\mathfrak{P}|\tilde{\mathfrak{p}})\sigma^{-1}$ is the Frobenius automorphism $\mathrm{Fr}_{L/k}(\mathfrak{P}'|\tilde{\mathfrak{p}})$. Conversely, if we are given a conjugate $\rho := \tau \mathrm{Fr}_{L/k}(\mathfrak{P}|\tilde{\mathfrak{p}})\tau^{-1}$ for $\tau \in \mathrm{Gal}(L/k)$, then we know that $\tau(\mathfrak{P})$ is an ideal of $\mathcal{O}_k$;

then similarly $\rho = \mathrm{Fr}_{L/k}((\tau\mathfrak{P})|\mathfrak{p})$. Therefore, the set of the Frobenius automorphisms $\mathrm{Fr}_{L/k}(\mathfrak{P}_j/\mathfrak{p})$ comprises a conjugate class of $\mathrm{Gal}(L/k)$.

For a given rational prime p, the factorization of $p\mathcal{O}_L$ is determined by the action of the Frobenius automorphism $\mathrm{Fr}_{L/\mathbb{Q}}(p)$ on L. More generally than just over $\mathbb{Q}$, the following lemma would allow us, given an extension of number fields L/k and a prime $\mathfrak{P}$ of $\mathcal{O}_L$ lying above a prime $\mathfrak{p}$ of $\mathcal{O}_k$ unramified in L, to determine the decomposition field $L^{G(\mathfrak{P})}$ of L – i.e., the maximal subfield of L in which $\mathfrak{p}$ splits completely. Then for an Abelian extension, the number of extensions of $v_\mathfrak{p}$ to V^L will match the degree $[L^{G(\mathfrak{P})} : k]$.

Lemma I.4.3. *Suppose F/k is a Galois extension of number fields. Let $\mathfrak{p}$ be an ideal of $\mathcal{O}_k$ unramified in F, and let $\tilde{\mathfrak{P}}$ be an ideal of $\mathcal{O}_F$ lying above $\mathfrak{p}$. Then $\mathfrak{p}$ splits completely in F if and only if $\mathrm{Fr}_{F/k}(\tilde{\mathfrak{P}}|\mathfrak{p})$ is trivial.*

In particular, suppose $\mathfrak{p}$ is a prime of $\mathcal{O}_k$ unramified in $L \supseteq F$, and $\mathfrak{P}$ is a prime of $\mathcal{O}_L$ lying above $\mathfrak{p}$. Since $\mathrm{Fr}_{L/k}$ restricted to F is $\mathrm{Fr}_{F/k}$, then the decomposition field $F = L^{G(\mathfrak{P})}$ will be the largest subfield of L fixed by $\mathrm{Fr}_{L/k}(\mathfrak{P}|\mathfrak{p})$.

Now, let us return to the problem of solving a congruence modulo an ideal $\mathfrak{a}$ of $\mathcal{O}_k$. As stated above, if the unique prime decomposition of $\mathfrak{a}$ is $\mathfrak{a} = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_r^{e_r}$, then by the Chinese Remainder Theorem, solving an equation modulo $\mathfrak{a}$ is equivalent to solving it modulo $\mathfrak{p}_i^{e_i}$ for each $i = 1, \ldots, r$. Because the exponents e_i can be quite large, it is often helpful to use Hensel's Lemma to speed up solving an equation (mod $\mathfrak{p}_i^{e_i}$). Hensel's Lemma states that:

Lemma I.4.4. *Let k be a number field, $\mathcal{O}_K$ its ring of integers, and $\mathfrak{p}$ a prime ideal of $\mathcal{O}_K$. If we are given $f(x) \in \mathcal{O}_k[x]$, let $f'(x)$ be its (formal) derivative. If there is an integer $n > 0$ so that $\alpha \in \mathcal{O}_k$ is a solution to $f(x) \equiv 0(\mathrm{mod}\ \mathfrak{p}^{2n-1})$, but $f'(\alpha) \not\equiv 0(\mathrm{mod}\ \mathfrak{p}^n)$, then for each integer $m > 0$ the equation $f(x) \equiv 0(\mathrm{mod}\ \mathfrak{p}^m)$ has a solution $\alpha_m \in \mathcal{O}_k$ so that $\alpha_m \equiv \alpha(\mathrm{mod}\ \mathfrak{p}^n)$.*

Therefore, once an equation has a solution modulo a congruence, with its derivative not satisfying another congruence, Hensel's Lemma provides an algorithm to find a solution (mod $\mathfrak{p}_i^m$) for each $m \geq 1$. To allow topological tools to apply to produce a solution modulo a congruence, the existence of such a solution can be recast as a distance in a metric induced by a valuation, on a certain local field, which is the completion of k with respect to the topology induced by this valuation. We will now introduce valuations.

Definition. A valuation over a field k is a multiplicative homomorphism $v : k \to \mathbb{R}_{\geq 0}$ satisfying
(a) the triangle inequality (i.e., $v(a + b) \leq v(a) + v(b)$), and
(b) $v^{-1}(0) = \{0\}$.

Two valuations are said to be equivalent if the topology induced by them is the same. A valuation is said to be trivial if the topology induced by it is the discrete topology. Within each equivalence class, normalizing any valuation yields a canonical representative of that class; for instance, the representative of the class with the trivial valuation satisfies $v(x) = 1$ for all $x \neq 0$. Since all valuations will be equivalent to a normalized valuation, the question of finding all valuations becomes the question of finding the set of all the non-trivial normalized valuations. For a field k, this set is denoted V^k. A valuation is called non-archimedian if it satisfies

$$v(a + b) \leq \max(v(a), v(b)) \text{ for all } a, b \in k;$$

otherwise, it is called archimedian. The set of non-archimedian valuations is denoted V_f^k, and is in one-to-one correspondence with the set of prime ideals of $\mathcal{O}_k$; the set of archimedian valuations is denoted V_∞^k.

For each valuation, we can complete k with respect to that valuation. When $k = \mathbb{Q}$, the set $V_f^{\mathbb{Q}}$ consists of a valuation v_p for each $\mathbb{Z}$-prime p, When $k = \mathbb{Q}$, the set $V_f^{\mathbb{Q}}$ consists of a valuation v_p for each $\mathbb{Z}$-prime p, which satisfies $v_p(p) = 1/p$ and $v_p(q) = 1$ for all (non-zero) integers q coprime to p; the completion of $\mathbb{Q}$ with respect to v_p is $\mathbb{Q}_{v_p} = \mathbb{Q}_p$. The set $V_\infty^{\mathbb{Q}}$ consists of the single valuation v_∞, defined by $v_\infty(q) = |q|$; the completion of $\mathbb{Q}$ with respect to v_∞ is $\mathbb{Q}_{v_\infty} = \mathbb{R}$. Then over the integers, the equation $a \equiv b \pmod{p^\alpha}$ is equivalent to the condition $v_p(a-b) \leq p^{-\alpha}$, in the p-adic topology on $\mathbb{Q}_p$.

Before discussing how to complete a general number field L with respect to its valuations, we would like to show how to find the set V^L for a general number field L. When restricted to $\mathbb{Q}$, each valuation $w : L \to \mathbb{R}$ becomes a valuation v of $\mathbb{Q}$; then the image $w(L)$ is contained in $v(\mathbb{Q}_v)$, and we can factor w as $L \xrightarrow{\tilde{w}} \mathbb{Q}_v \xrightarrow{v} \mathbb{R}$. So in order to determine the set of all valuations of L, it suffices to show the set of all homomorphisms $L \to \mathbb{Q}_v$ for all $v \in V^{\mathbb{Q}}$. By the Primitive Element Theorem [**FT**, 1.16], we know that L can be written as $\mathbb{Q}(\alpha)$ for some element α. Let $f(x)$ be the minimal polynomial of α over $\mathbb{Q}$. This is a polynomial of degree $[L : \mathbb{Q}]$, and splits as a product of irreducible factors $f(x) = f_1(x) \cdots f_n(x)$ over $\mathbb{Q}_v$. Then each of the irreducible factors $f_i(x)$ yields a distinct valuation, as follows. If we let $\bar{\mathbb{Q}_v}$ be the algebraic closure of $\mathbb{Q}_v$ and $q(x)$ be one of the linear factors of $f_i(x)$ over $\bar{\mathbb{Q}_v}$, then the map $\alpha \mapsto N_{\bar{\mathbb{Q}_v}/\mathbb{Q}_v}(q(\alpha))$ can be extended to a homomorphism $L \to \mathbb{Q}_v$, which as above leads to a valuation $L \to \mathbb{R}$.

As a special case, to find all archimedian valuations of a number field L, we will apply the above process to $v = v_\infty$. Then each of the roots of $f(x)$ in $\mathbb{C}$ leads to an embedding $L \to \mathbb{C}$, meaning that there are $[L : \mathbb{Q}]$ distinct homomorphisms from L into $\mathbb{C}$. An homomorphism is called a real valuation when the corresponding root of $f(x)$ lies in $\mathbb{R}$; if $\varphi : L \to \mathbb{C}$ is a homomorphism which is not a real valuation, then φ composed with complex conjugation yields another homomorphism $L \to \mathbb{C}$, meaning that homomorphisms which are not real valuations will come in pairs. In this case, the valuation obtained from this homomorphism will be called a complex valuation, obtained by composing it with the norm map $\mathrm{N}_{\mathbb{C}/\mathbb{R}}$; in particular, both homomorphisms in a given the pair will yield the same valuation. Therefore, setting r to be the number of real valuations and $2s$ to be the number of other homomorphisms $L \to \mathbb{C}$, we obtain $r + 2s = [L : \mathbb{Q}]$, but $|V_\infty^L| = r + s$.

Now, let us return to the question of completing a number field L with respect to a valuation $v \in V^L$. If v is a real valuation, then $L_v = \mathbb{R}$; if v is a complex valuation, then $L_v = \mathbb{C}$. Otherwise, v is a non-archimedian valuation, of the form $v_{\mathfrak{P}_i}$ for some ideal $\mathfrak{P}_i$ lying above a prime number p. Then the completion L_v of L with respect to this valuation will be a finite extension of $\mathbb{Q}_p$. Since $\mathbb{Q}_p$ is already complete, its valuation v_p has a unique extension to L_v; this will determine the values of v on L. More precisely, for each $a \in \mathfrak{P}_i \setminus \mathfrak{P}_i^2$ we will have $v(a) = p^{-1/e(\mathfrak{P}_i/p)}$. Since each integer $b \in p\mathcal{O}_L \setminus p^2\mathcal{O}_L$ is in $\mathfrak{P}_i^{e(\mathfrak{P}_i/p)}$, then we will have $v(b) = p^{-1}$ for each such b, and therefore $v_{\mathfrak{P}_i}$ is indeed an extension of v_p to L.

Using the above tools, we would like now to end this section with a proof of Lemma I.4.3, and also show how to rephrase Hensel's Lemma in the language of valuations, so that its proof is given in [**CF**, II.C].

Proof of Lemma I.4.3. Let L be a finite extension of F, $\mathfrak{P}$ be a prime of L lying above $\tilde{\mathfrak{P}}$, and $\sigma = \mathrm{Fr}_{L/k}(\mathfrak{P}|\mathfrak{p})$. Fix an extension $\tilde{\sigma}$ of σ to an element of $\mathrm{Gal}(L_{\mathfrak{P}}/k_{\mathfrak{p}})$. Since L/k is Galois, we know that $\mathfrak{p}$ splits completely in F if and only if the residual degree of $\tilde{\mathfrak{P}}/\mathfrak{p}$ is one. This means that the residue fields of k and F are the same, and is equivalent to the completion $k_{\mathfrak{p}}$ coinciding with the completion $F_{\tilde{\mathfrak{P}}}$. So it suffices to show that σ acts on F trivially if and only if $F_{\tilde{\mathfrak{P}}} = k_{\mathfrak{p}}$.

Suppose that σ acts on F trivially. Then in particular, σ fixes the residue field of F. But, we know that σ acts on the residue fields by sending $x \mapsto x^q$, where q is the size of the residue field of k; so every element of the residue field is a root of the polynomial $x^q - x$, and therefore $|\mathcal{O}_F/\tilde{\mathfrak{P}}| = q$. This implies $F_{\tilde{\mathfrak{P}}} \cong \mathbb{F}_q$, and so $F_{\tilde{\mathfrak{P}}} = k_{\mathfrak{p}}$. Conversely, if $F_{\tilde{\mathfrak{P}}} = k_{\mathfrak{p}}$; then by definition that the Frobenius σ fixes k – so $\tilde{\sigma}$ fixes $k_{\mathfrak{p}} = F_{\tilde{\mathfrak{P}}}$, and therefore also F. Since σ is a restriction of $\tilde{\sigma}$, then σ also fixes F. □

Proof of Hensel's Lemma I.4.4. The lemma in [**CF**, II.C] states that

Lemma I.4.5. *Let k be a field complete with respect to the non-archimedian valuation $|\ |$, and let*

$$f(X) \in \mathcal{O}[X],$$

where $\mathcal{O} \subset k$ is the ring of integers for $|\ |$. Let $\alpha_0 \in \mathcal{O}$ be such that

$$|f(\alpha_0)| < |f'(\alpha_0)|^2,$$

where $f'(X)$ is the (formal) derivative of $f(X)$. Then there is a solution of

$$f(\alpha) = 0, \qquad |\alpha - \alpha_0| \leq |f(\alpha)|/|f'(\alpha)|.$$

Its proof gives an algorithm for how to find the α_m in a more general setup; we will only show that with our setup, their conditions there are satisfied.

The k in [**CF**, II.C] should be the completion $K_{v_{\mathfrak{p}}}$, and the α_0 should be our α. Let w be the value of $v_{\mathfrak{p}}(s)$ for $s \in \mathfrak{p} \setminus \mathfrak{p}^2$. Then the condition $f'(\alpha) \not\equiv 0 (\mathrm{mod}\ \mathfrak{p}^n)$ but $f(\alpha) \equiv 0 (\mathrm{mod}\ \mathfrak{p}^{2n-1})$ means that $v_{\mathfrak{p}}(f'(\alpha)) \geq w^{n-1}$, but $v_{\mathfrak{p}}(f(\alpha)) \leq w^{2n-1} < w^{2n-2} = (w^{n-1})^2$. Then the Lemma of [**CF**, II.C] guarantees a solution in $K_{v_{\mathfrak{p}}}$, which implies a solution (mod $\mathfrak{p}^m$) for all m. The last statement follows since $v_{\mathfrak{p}}(f(\alpha))/v_{\mathfrak{p}}(f'(\alpha)) \leq w^{2n-1}/w^{n-1} = w^n$. □

4.2. Artin map for abelian extensions. Now, let us look closer at what happens for *abelian* extensions of k, for a fixed ideal $\mathfrak{p}$. Let L/k be an abelian (Galois) extension, and $S \subset V^k$ a finite set containing V^k_∞ and the valuations that ramify in L/k. Fix a prime ideal $\mathfrak{p}$ of $\mathcal{O}_k$ so that $v_{\mathfrak{p}} \notin S$. Then the set of Frobenius automorphisms $\mathrm{Fr}(\mathfrak{P}|\mathfrak{p})$, as $\mathfrak{P}$ ranges over the set of ideals of $\mathcal{O}_L$ lying above $\mathfrak{p}$, is a conjugate class of $\mathrm{Gal}(L/k)$. However, since we assumed that $\mathrm{Gal}(L/k)$ is *abelian*, then the Frobenius is actually unique, independent of the choice of $\mathfrak{P}$; we refer to it as $\mathrm{Fr}_{L/k}(\mathfrak{p})$ for short.

The Frobenius possesses some functorial properties as L varies. As explained above, if K is an intermediate field between k and L, then the restriction of $\mathrm{Fr}_{L/k}(\mathfrak{p})$ to K is the Frobenius $\mathrm{Fr}_{K/k}(\mathfrak{p})$. Also, if L can be written as a product of two subfields L_1 and L_2 with $L_1 \cap L_2 = k$, then the isomorphism map

$$\mathrm{Gal}(L/k) \overset{\sim}{\to} \mathrm{Gal}(L_1/k) \times \mathrm{Gal}(L_2/k)$$

sends

$$\mathrm{Fr}_{L/k}(\mathfrak{p}) \mapsto (\mathrm{Fr}_{L_1/k}(\mathfrak{p}), \mathrm{Fr}_{L_2/k}(\mathfrak{p})),$$

since an element is a generator of $\mathrm{Gal}(L/k)$ if and only if it generates both $\mathrm{Gal}(L_1/k)$ and $\mathrm{Gal}(L_2/k)$.

Now, let L be the maximal abelian extension of k (in a given algebraic closure). The resultant inertia field $L^{I(\mathfrak{p})}$ will be the maximal abelian extension of k unramified at $\mathfrak{p}$, and will be denoted $k^{nr}(\mathfrak{p})$. In this case, the fuctorial properties of the Frobenius automorphism allow us to define a map from the set of all finite extensions of k which are a subfield of $k^{nr}(\mathfrak{p})$, to their Galois groups, sending a field L where $\mathfrak{p}$ is unramified, to the Frobenius $\mathrm{Fr}_{L/k}(\mathfrak{p})$. This map is called an Artin symbol; the image of a field L under this map is denoted $(\mathfrak{p}, L/k)$.

Above we listed properties of $(\mathfrak{p}, L/k)$ for a single prime ideal $\mathfrak{p}$ of $\mathcal{O}_k$, as L varies over the finite sub-extensions of $k^{\mathrm{nr}}(\mathfrak{p})$. Let us now reverse the procedure: fix an abelian extension L/k, and consider all finite primes $\mathfrak{p}$ of $\mathcal{O}_k$, unramified in $\mathcal{O}_L$. For each $v \in V_f^k$, define the element $\mathbf{i}(v)$ to be the idèle with the components

$$\mathbf{i}(v)_{v'} = \begin{cases} 1 & , \; v' \neq v \\ \pi_v & , \; v' = v \end{cases} .$$

For an arbitrary $S \subset V^k$, define J_k^S to be the set of idèles $\mathbf{j} \in J_k$ satisfying $\mathbf{j}_v = 1$ for all $v \in S$. When S consists of V_∞^k and the set of finite valuations whose primes are ramified in L/k, we can define the Artin map on J_k^S via

$$\alpha_S : J_k^S \to \mathrm{Gal}(L/k)$$

$$\mathbf{i} = (\mathbf{i}_v) \mapsto \prod_{v \in V^k \setminus S} (\mathfrak{p}_v, L/k) \cdot v(\mathbf{i}_v)$$

We would like to extend the Artin map to a map $\alpha_{L/k} : J_k \to \mathrm{Gal}(L/k)$, defined over all the idèles. This extension will follow by linearity if we can define it on the set of $\mathbf{i}(v)$, where $\mathfrak{p}_v$ is ramified in L/k, and also on all idèles $\mathbf{i}$ whose $\mathbf{i}_v$-component is 1 for all v except some $v \in V_\infty^k$.

It turns out that there is only one way to extend $\alpha_{L/k}$ if we wish to make the Artin map continuous when the Galois group is endowed with the discrete topology. More generally, let S be a finite set of places, outside which definition of the Artin map is known. For a given $\mathbf{x} = (\mathbf{x})_v \in J_k$, let $\mathbf{x}^S$ denote the idèle satisfying

$$\mathbf{x}_v^S = \begin{cases} 1 & , \; v \in S \\ \mathbf{x}_v & , \; v \notin S \end{cases} .$$

Then we can write $\mathbf{x} = \mathbf{x}^S \cdot \mathbf{x}_1$, where $\mathbf{x}^S \in J_k^S$, and $(\mathbf{x}_1)_v = 1$ for all $v \in J_k \setminus S$. By assumption, $\alpha_{L/k}(\mathbf{x}^S)$ is already defined; therefore, to define $\alpha_{L/k}(\mathbf{x})$, it suffices to define $\alpha_{L/k}(\mathbf{x}_1)$. Now, $\mathbf{x}_1^{-1} \in J_k$, and by the Weak Approximation Theorem, we can find a sequence of $a_n \in k^\times$, so that $v(a_n - \mathbf{x}_1^{-1}) \to 0$ for all $v \in S$. We define

$$\alpha_{L/k}(\mathbf{x}_1) := \lim_{n \to \infty} \alpha_{L/k}((a_n \mathbf{x}_1)^S). \tag{2}$$

To show that this map is independent of the choice of the sequence a_n, we use a theorem from class field theory.

Definition. Suppose G is a topological group. A homomorphism $\phi : J_k^S \to G$ is *admissible* if for each neighborhood N of the identity $1 \in G$, there exists $\epsilon > 0$ such that $\phi((a)^S) \in N$ whenever $a \in K^\times$ and $v(a - 1) < \epsilon$ for all $v \in S$.

It is a highly non-trivial theorem, which is equivalent to Artin's Reciprocity Law, that the Artin map is admissible (cf. [CF], §§VII.4,10). Suppose that also $v(b_n - \mathbf{x}_1^{-1}) \to 0$ for all $v \in S$; write

$c_n = b_n^{-1}$; then since inversion of idèles is continuous, we have $v(c_n - \mathbf{x}_1) \to 0$; thus,

$$\begin{gathered}
v(a_n/b_n - 1) = v\left((a_n - \mathbf{x}_1^{-1} + \mathbf{x}_1^{-1})(c_n - \mathbf{x}_1 + \mathbf{x}_1) - 1\right) = \\
v\left((a_n - \mathbf{x}_1^{-1})(c_n - \mathbf{x}_1) + (a_n - \mathbf{x}_1^{-1})\mathbf{x}_1 + (c_n - \mathbf{x}_1)\mathbf{x}_1^{-1} + 1 - 1\right) \leq \\
\max\left(v(a_n - \mathbf{x}_1^{-1}) \cdot v(c_n - \mathbf{x}_1), v(a_n - \mathbf{x}_1^{-1}) \cdot v(\mathbf{x}_1), v(c_n - \mathbf{x}_1) \cdot v(\mathbf{x}_1^{-1})\right)
\end{gathered}$$

Since the former factor in each product goes to 0, then $v(a_n/b_n - 1) \to 0$ for all $v \in S$. Then, the fact that the Artin map is admissible implies that $\alpha_{L/k}(a_n/b_n) = e$ for n high enough. This means that the choice of $\alpha_{L/k}(\mathbf{x}_1)$ is independent of the sequence a_n. Then, the formula

$$\alpha_{L/k}(\mathbf{x}) = \alpha_{L/k}(\mathbf{x}^S) \cdot \alpha_{L/k}(\mathbf{x}_1),$$

allows us to extend the Artin map to all of J_k. Computationally, we have

$$\alpha_{L/k}(\mathbf{x}) := \lim_{n\to\infty} \alpha_{L/k}((a_n\mathbf{x})^S).$$

Next, we would like to point out that the image of a norm of a field under the Artin map corresponds to the restriction onto its Galois group. But first, an accessory lemma.

Lemma I.4.6. *Suppose L/k is an abelian extension. Then $k^\times \subseteq \text{Ker}\ \alpha_{L/k}$.*

Proof. When $\mathbf{x} \in k$, then $\mathbf{x}^{-1} \in k$; take $a_n = \mathbf{x}^{-1}$ for all n in the definition in (2). Then,

$$\alpha_{L/k}(\mathbf{x}) := \lim_{n\to\infty} \alpha_{L/k}((a_n\mathbf{x})^S) = \lim_{n\to\infty} \alpha_{L/k}((\mathbf{1})^S) = \lim_{n\to\infty} e = e$$

□

Lemma I.4.7. *Suppose $K \subseteq L \subseteq M$ are finite abelian extensions of k. Then the following diagram commutes:*

$$\begin{tikzcd}
J_K \arrow[r, "\alpha_{M/K}"] \arrow[d, "\mathrm{N}_{K/k}"'] & \text{Gal}(M/K) \arrow[d, "\iota"] \\
J_k \arrow[r, "\alpha_{L/k}"] & \text{Gal}(L/k)
\end{tikzcd},$$

where ι is the restriction to L.

The proof consists of verifying details we have not introduced, and is therefore only sketched here. The above diagram commutes when $\mathbf{x} \in J_K^S$ almost by definition (cf. Proposition VII.4.3 in [CF]); it commutes when $x \in K^\times$ because $\alpha_{M/K}(K)$ and $\alpha_{L/k}(k)$ are both the identity element by Lemma I.4.6. Consequently, it commutes on $J_K^S \cdot K^\times$, which is a dense subset of J_K. Since $\alpha_{M/K}$ and $\alpha_{L/k}$ are continuous, the above diagram is then commutative on all of J_K.

Taking $K = L$ and $M = L$, Lemma I.4.7 implies that $\alpha_{L/k}(\mathrm{N}_{L/k}(J_L)) = e$. Combining this fact with Lemma I.4.6, we obtain that:

Corollary I.4.8. $\mathrm{N}_{L/k}(J_L) \cdot k^\times \subseteq \text{Ker}\ \alpha_{L/k}$.

It is a deep result from Class Field Theory [**CF**, VII.5.1(B)], that the above containment is actually an equality. Then, identifying $k^\times$ with the (discrete) subgroup of principal idèles in J_k, define C_k to be the quotient $J_k/k^\times$. Then, we obtain an injective homomorphism

$$J_k/\text{Ker}\ \alpha_{L/k} = C_k/N_{L/k}C_L \overset{\psi_{L/k}}{\to} \text{Gal}(L/k),$$

where $\psi_{L/k}$ is the homomorphism obtained by projecting $\alpha_{L/k}$ onto the quotient of J_k by Ker $\alpha_{L/k}$. Taking $K = k$ in Lemma I.4.7 and taking the quotient by the kernel, the map $N_{K/k}$ becomes the projection map, which will be called j. Then, we obtain that:

Corollary I.4.9. *The following diagram commutes:*

$$\begin{tikzcd} C_k/N_{M/k}C_M \arrow[r, "\psi_{M/k}"] \arrow[d, "j"'] & \mathrm{Gal}(M/k) \arrow[d, "\iota"] \\ C_k/N_{L/k}C_L \arrow[r, "\psi_{L/k}"] & \mathrm{Gal}(L/k) \end{tikzcd},$$

Proof. The only new thing to verify is, that j is well-defined. This follows since $\mathrm{N}_{M/k}(C_M) = \mathrm{N}_{L/k}(\mathrm{N}_{M/L}(C_M)) \subseteq N_{L/k}(C_L)$. □

4.3. m-th power reciprocity law.

Suppose one needs to check whether an integer a is a square modulo a prime number p. Rather than computing the list of all squares of residue classes, we can simply compute the power residue symbol

$$\left(\frac{a}{p}\right) := a^{\frac{\varphi(p\mathbb{Z})}{2}} (\text{mod } p).$$

Lemma I.4.10. *a is a square modulo p if and only if* $\left(\frac{a}{p}\right) = 1$.

Proof. If $a \equiv c^2 \pmod p$ for some c, then $\left(\frac{a}{p}\right)$ will equal $c^{2\cdot(\frac{\varphi(p\mathbb{Z})}{2})} \equiv 1$. Conversely, the group $(\mathbb{Z}/p)^\times$ is the multiplicative group of a field, and therefore cyclic; let c be a primitive root modulo p. Then, if $a^{\varphi(p\mathbb{Z})/2} \equiv 1(\text{mod } p)$, then a must be an even power of c, say $c^{2s} \equiv a$. Then $a \equiv (c^s)^2 \pmod p$, proving the converse. □

A key reason the converse is true is, because the group $(\mathbb{Z}/p)^\times$ is cyclic. More generally, we will prove in Chapter II that

Lemma I.4.11. *For all $n \geq 1$ and odd primes p, the group $G = (\mathbb{Z}/p^n)^\times$ is cyclic.*

In particular, we can use the power residue symbol to compute not only whether an element is a square modulo p – but in fact for any m dividing $\phi(p)$, a will be an m-th power (mod p) if and only if the power residue symbol

$$\left(\frac{a}{p}\right)_m := a^{\frac{\phi(p)}{m}} (\text{mod } p)$$

is congruent to 1. More generally, for any integer m, we can determine whether a is an m-th power (mod p^n) as follows:

• a is an m-th power (mod p^n) if and only if whenever a power q^e of a prime q divides m, a is a q^e-th power (mod p^n).

• If $q \neq p$ and q^d is the highest power of q dividing $\phi(p)$, then a is a q^e-th power (mod p^n) if and only if

$$\left(\frac{a}{p}\right)_{q^d}^{\lceil q^{d-e} \rceil} = 1,$$

where $\lceil\ \rceil$ means we round up if q^{d-e} is not an integer (i.e., if $d < e$).

- The case $q = p$ is more difficult, and is covered in §I.3 of [**BMS**]; since there are only finitely many primes p so that $p = q$ for some $q^e \mid m$, we avoid such primes in our argument.

This procedure can be generalized to rings of S-integers of arbitrary number fields. Let k be a number field, and $\mathcal{O} = \mathcal{O}_{k,S}$ be the ring of S-integers of k. Recall that for an arbitrary ideal $\mathfrak{p}$ of $\mathcal{O}$, we defined $\phi(\mathfrak{p})$ to be the size of $(\mathcal{O}/\mathfrak{p})^\times$; then, once an ideal $\mathfrak{p}$ satisfies $\mathcal{O}/\mathfrak{p} = \mathbb{Z}/p$ and p is an odd prime, Lemma I.4.10 applies the above proof works to show that for any prime $\mathfrak{p}$ satisfying $\mathcal{O}/\mathfrak{p} = \mathbb{Z}/p$ and integer m dividing $\phi(\mathfrak{p})$, if $\mathfrak{p} \nmid 2\mathcal{O}$ then a given element $a \in \mathcal{O}$ is an m-th power (mod $\mathfrak{p}$) if and only if $a^{\phi(\mathfrak{p})/m} \equiv 1 (\bmod\ \mathfrak{p})$. Using Hensel's Lemma I.4.4, this happens if and only if a is an m-th power (mod $\mathfrak{p}^n$) for all n. We will call a prime $\mathfrak{p}$ of $\mathcal{O}$ $\mathbb{Q}$-split if it satisfies the conditions to make this argument work: namely, if $\mathfrak{p}$ is a non-dyadic prime satisfying $\mathcal{O}/\mathfrak{p} = \mathbb{Z}/p$ (cf. Chapter II).

Thus far, our definition only allows us to view the power residue symbol as a residue class modulo $\mathfrak{p}$. However, in order to do meaningful computation using the power residue symbol, we need to be able to view it as a root of unity in k. In order to do so, we now introduce the Hilbert symbol.

Let m be an integer dividing $\mu(k)$, and $L = k(\sqrt[m]{a})$. We will now use the Artin map we introduced in subection I.4.2 in order to introduce local Artin maps for each $v \in V^k$, as follows. If we let i_v be the composition of the canonical injections $k^\times \to k_v^\times \to J_k$, then we get a diagram of the form

$$k^\times \xrightarrow{i_v} J_k \xrightarrow{\alpha_{L/k}} \mathrm{Gal}(L/k) \tag{3}$$

The composition $\alpha_{L/k} \circ i_v$ is the local Artin map ψ_v. This is the map which comes from the local class field theory for L_w/k_v, where w is the valuation of L above v. Since the product of all of the i_v's corresponds to the injection $k^\times \to J_k$, then for all $b \in k^\times$ we can decompose $\alpha_{L/k}$ as

$$\alpha_{L/k}(b) = \prod_{v \in V^k} \psi_v(b), \tag{4}$$

a product of the local maps $\psi_v(b)$.

Suppose we are given a prime ideal $\mathfrak{p}$ of $\mathcal{O}_k$. Then for each $b \in k^\times$, $\psi_{v_\mathfrak{p}}(b)$ is an element of $\mathrm{Gal}(k(\sqrt[m]{a})/k)$. The extension $k(\sqrt[m]{a})/k$ is a cyclic extension of degree dividing m, and so each element of the Galois group has order dividing m. In particular, $\psi_{v_\mathfrak{p}}(b)$ acts on $\sqrt[m]{a}$ by multiplying with some m-th root of unity. The value of the Hilbert symbol is defined to be equal to that root. More precisely,

$$\left(\frac{a, b}{\mathfrak{p}}\right)_m = \frac{\psi_{v_\mathfrak{p}}(b)(\sqrt[m]{a})}{\sqrt[m]{a}}.$$

To show that the Hilbert symbol is well-defined, we need to show that it does not depend on the choice of $\sqrt[m]{a}$. This is because two such roots differ by ζ_m^n for some $n \geq 1$; and, by assumption, $\zeta_m \in k$, so $\psi_{v_\mathfrak{p}}$, being an automorphism of $k(\sqrt[m]{a})/k$, must fix ζ_m^n.

For an archimedian prime v, if v is complex, then k_v is algebraically closed, so for all $b \in k_v^\times$ we have $k_v(\sqrt[m]{b}) = k_v$, and therefore ψ_v is the trivial automorphism, and the value of the Hilbert symbol is 1. On the other hand, if v is real, then $m = 2$. In this case, if $v(a) > 0$, then $\sqrt{a} \in k_v$, and so $k(\sqrt{a})_v = k_v$, meaning that again the Hilbert symbol must be trivial. Otherwise, $k(\sqrt{a})_v/k_v$ is an extension of degree 2; in this case, for $v(b) > 0$, we have $b = (\sqrt{b})^2$, so the above argument shows that $\psi_v(b)$ must act trivially on $k(\sqrt{a})$, and the Hilbert symbol is again trivial. Finally, $\mathrm{Gal}(k(\sqrt{a})_v/k_v) \neq \{e\}$, and the image under ψ_v of the open subgroup of k_v consisting of the set of b such that $v(b) < 0$ acts non-trivially on $\sqrt{a}$; this means that whenever $b \in k$ satisfies $v(b) < 0$ and $v(a) < 0$, the Hilbert symbol is equal to $\zeta_m = -1$.

For non-archimedian primes, a tool which is frequently used to identify the value of the Hilbert symbol is, the fact that the Hilbert symbol is bilinear, and satisfies the Steinberg relation. But first, we need a technical lemma.

Lemma I.4.12. *Suppose $m \mid \mu(k)$, and we are given $a, b \in k^\times$ so that $a + b \in k^m$. Then for all primes $\mathfrak{p} \in V^k$, $\left(\dfrac{a, b}{\mathfrak{p}}\right)_m = 1$.*

Proof. Let ϵ be a fixed m-th root of a, and $\beta \in \mathcal{O}_k$ be an element so that $a + b = \beta^m$. To show that ϵ is fixed by $\psi_{v_\mathfrak{p}}(b)$, we will show that $b \in \mathrm{N}_{k(\epsilon)/k}(k(\epsilon)) \subset \mathrm{Ker}(\alpha_{k(\epsilon)/k})$. Since $\psi_{v_\mathfrak{p}} = \iota_v \circ \alpha_{k(\epsilon)/k}$ (see (3)), this will imply that $\psi_{v_\mathfrak{p}}(b)(\sqrt[m]{a}) = \sqrt[m]{a}$, so the Hilbert symbol is trivial.

Now the map $\sigma \mapsto \sigma(\epsilon)/\epsilon$ is an isomorphism of $\mathrm{Gal}(k(\epsilon)/k)$ with μ_d, for some $d = |\mathrm{Gal}(k(\epsilon)/k)|$ dividing m; as shown when defining the Hilbert symbol, this map is independent of the choice of $\sqrt[m]{a}$. Then

$$\mathrm{N}_{k(\epsilon)/k}(\beta - \epsilon) = \prod_{\zeta \in \mu_d} (\beta - \zeta\epsilon).$$

Setting $\{z_1, \ldots, z_{m/d}\}$ to be a set of coset representatives of μ_d in μ_m, we know all z_i are in k, and so $\prod_{i=1}^{m/d}(\beta - z_i\epsilon) \in k(\epsilon)$. Therefore,

$$\mathrm{N}_{k(\epsilon)/k}\left(\prod_{i=1}^{m/d}(\beta - z_i\epsilon)\right) = \prod_{z \in \mu_m}(\beta - z\epsilon) = \beta^m - \epsilon^m = (a+b) - a = b,$$

completing the proof. □

Lemma I.4.13. *For a valuation $v \in V^k$, let $\mathfrak{p} = \mathfrak{p}_v$. Then the Hilbert symbol is bilinear, and satisfies the Steinberg relation*

$$\left(\frac{a, 1-a}{\mathfrak{p}}\right)_m = 1.$$

Proof. To show bilinearity, we have to show that for arbitrary $a, a', b, b' \in k^\times$, we have

$$\left(\frac{a, b}{\mathfrak{p}}\right)_m \left(\frac{a', b}{\mathfrak{p}}\right)_m = \left(\frac{aa', b}{\mathfrak{p}}\right)_m \quad \text{and} \quad \left(\frac{a, b}{\mathfrak{p}}\right)_m \left(\frac{a, b'}{\mathfrak{p}}\right)_m = \left(\frac{a, bb'}{\mathfrak{p}}\right)_m.$$

So let $a, a', b \in k^\times$ be fixed, $L = k(\sqrt[m]{a}, \sqrt[m]{a'})$, and ψ be the $\psi_\mathfrak{p}$ map for the extension L/k. Then by Lemma I.4.7, for each $b \in k(\sqrt[m]{a})$, we know $\psi(b)|_{k(\sqrt[m]{a})}$ is the $\psi_{v_\mathfrak{p}}(b)$ map in $\mathrm{Gal}(k(\sqrt[m]{a})/k)$, and for each $b \in k(\sqrt[m]{a'})$, we know $\psi(b)|_{k(\sqrt[m]{a'})}$ is the $\psi_{v_\mathfrak{p}}(b)$ map in $\mathrm{Gal}(k(\sqrt[m]{a'})/k)$. With these notations, bilinearity in the first component follows since

$$\left(\frac{aa', b}{\mathfrak{p}}\right)_m \sqrt[m]{aa'} = (\sqrt[m]{aa'})^{\psi(b)} = (\sqrt[m]{a})^{\psi(b)}(\sqrt[m]{a'})^{\psi(b)} = \left(\frac{a, b}{\mathfrak{p}}\right)_m \left(\frac{a', b}{\mathfrak{p}}\right)_m \sqrt[m]{aa'}.$$

Bilinearity in the second component follows since the local Artin map is a homomorphism. Namely,

$$\left(\frac{a, bb'}{\mathfrak{p}}\right)_m \sqrt[m]{a} = (\sqrt[m]{a})^{\psi(bb')} = (\sqrt[m]{a})^{\psi(b)\psi(b')} = (\left(\frac{a, b}{\mathfrak{p}}\right)_m \sqrt[m]{a})^{\psi(b')} = \left(\frac{a, b}{\mathfrak{p}}\right)_m (\sqrt[m]{a})^{\psi(b')} = \left(\frac{a, b}{\mathfrak{p}}\right)_m \left(\frac{a, b'}{\mathfrak{p}}\right)_m \sqrt[m]{a}.$$

The third equality follows since Hilbert symbols are in μ_m, which is fixed by $\mathrm{Gal}(k(\sqrt[m]{a})/k) \ni \psi(b')$. Finally, the fact that

$$\left(\frac{a, 1-a}{\mathfrak{p}}\right)_m = 1$$

follows from Lemma I.4.12, since $a + (1-a) = 1 \in k^m$. □

The group generated by elements of k^2 subject to the relations of bilinearity and the Steinberg relation is called the Milnor's K-group $K_2(k)$ (cf., [**M**, P. 40]). As a consequence, all the identities that hold in $K_2(k)$ are also satisfied by Hilbert Symbols. One of these identities we will use is,

$$\left(\frac{a,b}{\mathfrak{p}}\right)_m = \left(\frac{b,a}{\mathfrak{p}}\right)_m^{-1}. \tag{5}$$

For a direct proof of this identity, see [**M**].

Corollary I.4.14. *If $v \in V^k$, and $\mathfrak{p} = \mathfrak{p}_v$, then there is a homomorphism from the Milnor's K-group $K_2(k)$ into $\mu(k)$, sending $(a,b) \mapsto \left(\frac{a,b}{\mathfrak{p}}\right)_m$.*

Now, in order to utilize the m-th power reciprocity formula, we would like to connect the Hilbert symbol with the power residue symbol introduced earlier. Suppose that $\mathfrak{p}$ is a prime ideal of $\mathcal{O}$ so that $v_\mathfrak{p} \notin V(\mu) \cup S$, and we are given $a, b \in k^\times$ satisfying $v_\mathfrak{p}(a) = 0$ and $v_\mathfrak{p}(b) = 1$. Then $\psi_{v_\mathfrak{p}}(b)$ will be the Frobenius automorphism $\mathrm{Fr}_{k(\sqrt[m]{a})/k}(\mathfrak{p})$, and therefore we will have

$$\left(\frac{a,b}{\mathfrak{p}}\right)_m \sqrt[m]{a} = \sqrt[m]{a}^{\psi_{v_\mathfrak{p}}(b)} = \mathrm{Fr}_{k(\sqrt[m]{a})/k}(\mathfrak{p})(\sqrt[m]{a})$$

This definition is independent of the exact choice of b. Since the minimal power n so $\left(\frac{a,b}{\mathfrak{p}}\right)_m^n = 1$ matches the minimal power n' so $\left(\frac{a}{\mathfrak{p}}\right)^{n'} \equiv 1$, then we can identify

$$\left(\frac{a,b}{\mathfrak{p}}\right)_m = \left(\frac{a}{\mathfrak{p}}\right)_m$$

in this case. Using linearity in the second factor, we obtain that for arbitrary $b \in k^\times$, we have

$$\left(\frac{a,b}{\mathfrak{p}}\right)_m = \left(\frac{a}{\mathfrak{p}}\right)_m^{v_\mathfrak{p}(b)} = \left(\frac{a^{v_\mathfrak{p}(b)}}{\mathfrak{p}}\right)_m. \tag{6}$$

Next, for the case $v_\mathfrak{p}(a) = v_\mathfrak{p}(b) = 1$, we have $v_\mathfrak{p}(-ab^{-1}) = 0$, so the above argument, in conjunction with bilinearity of the Hilbert symbol and Lemma I.4.12, shows

$$\left(\frac{a,b}{\mathfrak{p}}\right)_m = \left(\frac{-ab^{-1},b}{\mathfrak{p}}\right)_m \left(\frac{-b,b}{\mathfrak{p}}\right)_m = \left(\frac{-ab^{-1},b}{\mathfrak{p}}\right)_m = \left(\frac{-ab^{-1}}{\mathfrak{p}}\right)_m.$$

Finally, using bilinearity, it is straightforward to check that whenever $v_\mathfrak{p} \notin V(\mu) \cup S$, the following formula holds for the Hilbert symbols:

$$\left(\frac{a,b}{\mathfrak{p}}\right)_m = \left(\frac{(-1)^{v_\mathfrak{p}(a)v_\mathfrak{p}(b)} a^{v_\mathfrak{p}(b)} b^{-v_\mathfrak{p}(a)}}{\mathfrak{p}}\right)_m.$$

The definition of the Artin map for the $\mathfrak{p}$'s satisfying $v_\mathfrak{p} \in V(\mu) \cup S$ was given using the density of $J_k^S \cdot k^\times$ in J_k; and therefore for those primes, the Hilbert symbol does not have a simple definition in terms of the power residue symbol. Even though the exact value is hard to determine, the following product formula still holds:

Lemma I.4.15. *For any* $a, b \in k^\times$, $\displaystyle\prod_{v \in V^k} \left(\frac{a,b}{\mathfrak{p}_v}\right)_m = 1$.

Proof.

$$\left(\prod_{v \in V^k} \left(\frac{a,b}{\mathfrak{p}_v}\right)_m\right)(\sqrt[m]{a}) = \left(\prod_{v \in V^k} \psi_v(b)\right)(\sqrt[m]{a}) = \alpha_{k(\sqrt[m]{a})/k}(b)(\sqrt[m]{a}) = e(\sqrt[m]{a}),$$

implying the result. The last equality follows from Lemma I.4.6 since $\alpha_{L/k}(k^\times) = e$ for all finite abelian extensions L/k. □

CHAPTER II

Algebraic Number Fields

Introduction

The goal of this chapter is to prove the following theorem, and a few of its refinements:

THEOREM II.3.6. *Let $\mathcal{O} = \mathcal{O}_{k,S}$ be the ring of S-integers in a number field k, and assume that the group of units $\mathcal{O}^\times$ is infinite. Then every matrix in $\mathrm{SL}_2(\mathcal{O})$ is a product of at most 9 elementary matrices.*

The proof of these results requires the use of an algebraic result about abelian subextensions of radical extensions of a general field – namely:

PROPOSITION II.1.1. *Let $n \geq 1$ be an integer prime to* char k, *let $\mu(k)_p$ denote the set of $x \in k$ such that $x^{p^\alpha} = 1$ for some integer $\alpha \neq 0$, and let $u \in k^\times$ be such that $u \notin \mu(k)_p k^{\times p}$ for all primes $p \mid n$. Then the polynomial $x^n - u$ is irreducible over k, and for $t = \sqrt[n]{u}$ we have*

$$k(t) \cap k^{\mathrm{ab}} = k(t^m) \quad \textit{where} \quad m = \frac{n}{\displaystyle\prod_{p \mid n} \gcd(n, |\mu(k)_p|)},$$

with the convention that if $|\mu(k)_p| = \infty$, then $\gcd(n, |\mu(k)_p|)$ *is simply the p-primary component of n.*

We start our chapter with the proof of this result.

One of the key notions in this chapter is that of a $\mathbb{Q}$-split prime: we say that a prime $\mathfrak{p}$ of a number field k is $\mathbb{Q}$-split if it is non-dyadic and its local degree over the corresponding rational prime is 1. In §II.2, we establish some relevant for us properties of such primes (see Subsection II.2.1) and prove for them in §II.2.2 the following refinement of Dirichlet's Theorem from [**BMS**].

THEOREM II.2.3. *Let $\mathcal{O}$ be the ring of S-integers in a number field k for some finite $S \subset V^k$ containing V^k_∞. If nonzero $a, b \in \mathcal{O}$ are relatively prime (i.e., $a\mathcal{O} + b\mathcal{O} = \mathcal{O}$) then there exist infinitely many principal $\mathbb{Q}$-split prime ideals $\mathfrak{p}$ of $\mathcal{O}$ with a generator π such that $\pi \equiv a \pmod{b\mathcal{O}}$ and $\pi > 0$ in all real completions of k.*

Subsection §II.2.3 is devoted to the statement and proof of the Theorem II.2.7, which is another key number-theoretic result needed in the proof of Theorem II.3.6. In §II.3, we prove Theorem II.3.6 and Corollary I.3.6 (from the introduction). Next, in §II.4 we correct the faulty example from [Vs] of a matrix in $\mathrm{SL}_2(\mathbb{Z}[1/p])$, where p is a prime $\equiv 1 (\mathrm{mod}\ 29)$, that is not a product of four elementary matrices – see Proposition II.4.1, confirming thereby that the bound of 5 in [Vs] is optimal.

II.1. Abelian subextensions of radical extensions.

In this section, k is an arbitrary field. For a prime $p \neq \operatorname{char} k$, we let $\mu(k)_p$ denote the subgroup of $\mu(k)$, consisting of elements satisfying $x^{p^d} = 1$ for any $d \geq 0$. If this subgroup is finite, we set $\lambda(k)_p$ to be the non-negative integer satisfying $|\mu(k)_p| = p^{\lambda(k)_p}$; otherwise, set $\lambda(k)_p = \infty$. Clearly if $\mu(k)$ is finite, then $\mu = \prod_p p^{\lambda(k)_p}$. For $a \in k^\times$, we write $\sqrt[n]{a}$ to denote an arbitrary root of the polynomial $x^n - a$.

The goal of this section is to prove the following.

Proposition II.1.1. *Let $n \geq 1$ be an integer prime to* char k*, and let $u \in k^\times$ be such that $u \notin \mu(k)_p (k^\times)^p$ for all $p \mid n$. Then the polynomial $x^n - u$ is irreducible over k, and for $t = \sqrt[n]{u}$ we have*

$$k(t) \cap k^{\mathrm{ab}} = k(t^m) \quad \textit{where} \quad m = \frac{n}{\prod\limits_{p \mid n} \gcd(n, p^{\lambda(k)_p})},$$

with the convention that $\gcd(n, p^\infty)$ *is simply the p-primary component of n.*

We first treat the case $n = p^d$ where p is a prime.

Proposition II.1.2. *Let p be a prime number $\neq$* char k*, and let $u \in k^\times \setminus \mu(k)_p (k^\times)^p$. Fix an integer $d \geq 1$, set $t = \sqrt[p^d]{u}$. Then*

$$k(t) \cap k^{\mathrm{ab}} = k(t^{p^\gamma}) \quad \textit{where} \quad \gamma = \max(0, d - \lambda(k)_p).$$

We begin with the following lemma.

Lemma II.1.3. *Let p be a prime number $\neq$* char k*, and let $u \in k^\times \setminus \mu(k)_p (k^\times)^p$. Set $k_1 = k(\sqrt[p]{u})$. Then*

(i) *$[k_1 : k] = p$;*

(ii) *$\mu(k_1)_p = \mu(k)_p$*

(iii) *None of the $\sqrt[p]{u}$ are in $\mu(k_1)_p (k_1^\times)^p$.*

Proof. (i) follows from [**La**, Ch. VI, §9], as $u \notin (k^\times)^p$.

(ii): If $\lambda(k)_p = \infty$, then there is nothing to prove. Otherwise, we need to show that for $\lambda = \lambda(k)_p$, we have $\zeta_{p^{\lambda+1}} \notin k_1$. Assume the contrary. Then, first, $\lambda > 0$. Indeed, we have a tower of inclusions $k \subseteq k(\zeta_p) \subseteq k_1$. Since $[k_1 : k] = p$ by (i), and $[k(\zeta_p) : k] \leq p - 1$, we conclude that $[k(\zeta_p) : k] = 1$, i.e. $\zeta_p \in k$.

Now, since $\zeta_{p^{\lambda+1}} \notin k$, we have

$$k_1 = k(\zeta_{p^{\lambda+1}}) = k(\sqrt[p]{\zeta_{p^\lambda}}). \tag{7}$$

But according to Kummer's theory (which applies because $\zeta_p \in k$), the fact that $k(\sqrt[p]{a}) = k(\sqrt[p]{b})$ for $a, b \in k^\times$ implies that the images of a and b in $k^\times/(k^\times)^p$ generate the same subgroup. So, it follows from (7) that $u\zeta_p^i \in (k^\times)^p$ for some i, and therefore $u \in \mu(k)_p (k^\times)^p$, contradicting our choice of u.

(iii): Assume the contrary, i.e. some p-th root $\sqrt[p]{u}$ can be written in the form $\sqrt[p]{u} = \zeta a^p$ for some $a \in k_1^\times$ and $\zeta \in \mu(k_1)_p$. Let $N = N_{k_1/k}\colon k_1^\times \to k^\times$ be the norm map. Then

$$N(\sqrt[p]{u}) = N(\zeta)N(a)^p.$$

Clearly, $N(\zeta) \in \mu(k)_p$, so $N(\sqrt[p]{u}) \in \mu(k)_p(k^\times)^p$. On the other hand, $N(\sqrt[p]{u}) = u$ for p odd, and $-u$ for $p = 2$. In all cases, we obtain that $u \in \mu(k)_p(k^\times)^p$. A contradiction. □

A simple induction now yields the following:

Corollary II.1.4. *Let p be a prime number $\neq$ char k, and let $u \in k^\times \setminus \mu(k)_p(k^\times)^p$. For a fixed integer $d \geq 1$, set $k_d = k(\sqrt[p^d]{u})$. Then:*

(i) *$[k_d : k] = p^d$;*

(ii) *$\mu(k_d)_p = \mu(k)_p$, hence $\lambda(k_d)_p = \lambda(k)_p$.*

Of course, assertion (i) is well-known and follows, for example, from [**La**, Ch. VI, §9].

Lemma II.1.5. *Let p be a prime number $\neq$ char k, and let $u \in k^\times \setminus \mu(k)_p(k^\times)^p$. Fix an integer $d \geq 1$, and set $t = \sqrt[p^d]{u}$ and $k_d = k(t)$. Furthermore, for an integer j between 0 and d define $\ell_j = k(t^{p^{d-j}}) \simeq k(\sqrt[p^j]{u})$. Then any intermediate subfield $k \subseteq \ell \subseteq k_d$ is of the form $\ell = \ell_j$ for some $j \in \{0, \ldots, d\}$.*

Proof. Suppose we are given such an ℓ; it follows from Corollary II.1.4(i) that $[k_d : \ell] = p^j$ for some $0 \leq j \leq d$. Since any conjugate of t is of the form $\zeta \cdot t$ where $\zeta^{p^d} = 1$, we see that the norm $N_{k_d/\ell}(t)$ is of the form $\zeta_0 t^{p^j}$, where again $\zeta_0^{p^d} = 1$. Then $\zeta_0 \in \mu(k_d)_p$, and using Corollary II.1.4(ii), we conclude that $\zeta_0 \in k \subseteq \ell$. So, $t^{p^j} \in \ell$, implying the inclusion $\ell_{d-j} \subseteq \ell$. Now, the fact that $[k_d : \ell_{d-j}] = p^j$ implies that $\ell = \ell_{d-j}$, yielding our claim. □

Proof of Proposition II.1.2. Set $\lambda = \lambda(k)_p$. Then for any $d \leq \lambda$ the extension $k(\sqrt[p^d]{u})/k$ is abelian, and our assertion is trivial. So, we may assume that $\lambda < \infty$ and $d > \lambda$. It follows from Lemma II.1.5 that $\ell := k(t) \cap k^{\text{ab}}$ is of the form $\ell_{d-j} = k(t^{p^j})$ for some $j \in \{0, \ldots, d\}$. On the other hand, ℓ_{d-j}/k is a Galois extension of degree p^{d-j}, so must contain the Galois conjugate $\zeta_{p^{d-j}} t^{p^{d-j}}$ of $t^{p^{d-j}}$, implying that $\zeta_{p^{d-j}} \in \ell_{d-j}$. Since $\ell_{d-j} \simeq k(\sqrt[p^{d-j}]{u})$, we conclude from Corollary II.1.4(ii) that $d-j \leq \lambda$, i.e. $j \geq d - \lambda$. This proves the inclusion $\ell \subseteq k(t^{p^\gamma})$; the opposite inclusion is obvious.

Proof of Proposition II.1.1. Let $n = p_1^{\alpha_1} \cdots p_s^{\alpha_s}$ be the prime factorization of n, and for $i = 1, \ldots, s$ set $n_i = n/p_i^{\alpha_i}$. Let $t = \sqrt[n]{u}$ and $t_i = t^{n_i}$ (so, t_i is a $p_i^{\alpha_i}$-th root of u). Using again [**La**, Ch. VI, §9] we conclude that $[k(t) : k] = n$, which implies that

$$[k(t) : k(t_i)] = n_i \quad \text{for all} \quad i = 1, \ldots, r. \tag{8}$$

Since for $K := k(t) \cap k^{\text{ab}}$ the degree $[K : k]$ divides n, we can write $K = K_1 \cdots K_s$ where K_i is an abelian extension of k of degree $p_i^{\beta_i}$ for some $\beta_i \leq \alpha_i$. Then the degree $[K_i(t_i) : k(t_i)]$ must be a power of p_i. Comparing with (8), we conclude that $K_i \subseteq k(t_i)$. Applying Proposition II.1.2 with $d = \alpha_i$, we obtain the inclusion

$$K_i \subseteq k(t_i^{p_i^{\gamma_i}}) = k(t^{n_i p_i^{\gamma_i}}) \quad \text{where} \quad \gamma_i = \max(0\,,\, \alpha_i - \lambda(k)_{p_i}). \tag{9}$$

It is easy to see that the g.c.d. of the numbers $n_i p_i^{\gamma_i}$ for $i = 1, \dots, s$ is

$$m = \frac{n}{\displaystyle\prod_{p|n} \gcd(n, p^{\lambda(k)_p})}.$$

Furthermore, the subgroup of $k(t)^\times$ generated by $t^{n_1 p_1^{\gamma_1}}, \dots, t^{n_s p_s^{\gamma_s}}$ coincides with the cyclic subgroup with generator t^m. Then (9) yields the following inclusion

$$K = K_1 \cdots K_s \subseteq k(t^m).$$

Since the opposite inclusion is obvious, our claim follows. □

Corollary II.1.6. *Assume that $\mu = |\mu(k)| < \infty$. Let P be a finite set of rational primes, and define*

$$\mu' = \mu \cdot \prod_{p \in P} p.$$

If we are given $u \in k^\times$ such that

$$u \notin \mu(k)_p (k^\times)^p \ \ \textit{for all} \ \ p \in P,$$

then for any abelian extension F of k the intersection

$$E := F \cap k\left(\sqrt[\mu']{u}, \zeta_{\mu'}\right)$$

is contained in $k\left(\sqrt[\mu]{u}, \zeta_{\mu'}\right)$.

Proof. Without loss of generality we may assume that $\zeta_{\mu'} \in F$, and then we have the following tower of field extensions

$$k\left(\sqrt[\mu]{u}\,,\, \zeta_{\mu'}\right) \subset E\left(\sqrt[\mu]{u}\right) \subset k\left(\sqrt[\mu']{u}\,,\, \zeta_{\mu'}\right).$$

We note that the degree $\left[k\left(\sqrt[\mu']{u}\,,\, \zeta_{\mu'}\right) : k\left(\sqrt[\mu]{u}\,,\, \zeta_{\mu'}\right)\right]$ divides $\prod_{p \in P} p$. So, if we assume that the assertion of the lemma is false, then we should be able to find to find a prime $p \in P$ that divides the degree $\left[E\left(\sqrt[\mu]{u}\right) : k\left(\sqrt[\mu]{u}\,,\, \zeta_{\mu'}\right)\right]$, and therefore does *not* divide the degree $\left[k\left(\sqrt[\mu']{u}\,,\, \zeta_{\mu'}\right) : E\left(\sqrt[\mu]{u}\right)\right]$. The latter implies that $\sqrt[p\mu]{u} \in E\left(\sqrt[\mu]{u}\right)$. But this contradicts Proposition II.1.1 since $E\left(\sqrt[\mu]{u}\right) = E \cdot k\left(\sqrt[\mu]{u}\right)$ is an abelian extension of k. □

II.2. Results from Algebraic Number Theory

1. $\mathbb{Q}$-split primes. Our proof of Theorem II.3.6 heavily relies on properties of so-called $\mathbb{Q}$-split primes in $\mathcal{O}$.

Definition. Let $\mathfrak{p}$ be a nonzero prime ideal of $\mathcal{O}$, and let p be the corresponding rational prime. We say that $\mathfrak{p}$ is $\mathbb{Q}$-*split* if $p > 2$, and for the valuation $v = v_{\mathfrak{p}}$ we have $k_v = \mathbb{Q}_p$.

For the convenience of further references, we list some simple properties of $\mathbb{Q}$-split primes.

Lemma II.2.1. *Let $\mathfrak{p}$ be a $\mathbb{Q}$-split prime in $\mathcal{O}$, and for $n \geq 1$ let $\rho_n \colon \mathcal{O} \to \mathcal{O}/\mathfrak{p}^n$ be the corresponding quotient map. Then:*

(a) *the group of invertible elements $(\mathcal{O}/\mathfrak{p}^n)^\times$ is cyclic for any n;*

(b) *if $c \in \mathcal{O}$ is such that $\rho_2(c)$ generates $(\mathcal{O}/\mathfrak{p}^2)^\times$ then $\rho_n(c)$ generates $(\mathcal{O}/\mathfrak{p}^n)^\times$ for any $n \geq 2$.*

Proof. Let $p > 2$ be the rational prime corresponding to $\mathfrak{p}$, and $v = v_{\mathfrak{p}}$ be the associated valuation of k. By definition, $k_v = \mathbb{Q}_p$, hence $\mathcal{O}_v = \mathbb{Z}_p$. So, for any $n \geq 1$ we will have canonical ring isomorphisms

$$\mathcal{O}/\mathfrak{p}^n \simeq \mathcal{O}_v/\hat{\mathfrak{p}}_v^n = \mathbb{Z}_p/p^n\mathbb{Z}_p \simeq \mathbb{Z}/p^n\mathbb{Z}. \tag{10}$$

Then the isomorphisms in (10) are compatible for different n's. Since the kernel of the group homomorphism $(\mathbb{Z}/p^n\mathbb{Z})^\times \to (\mathbb{Z}/p^2\mathbb{Z})^\times$ is contained in the Frattini subgroup of $(\mathbb{Z}/p^n\mathbb{Z})^\times$ for $n \geq 2$, the same is true for the homomorphism $(\mathcal{O}/\mathfrak{p}^n)^\times \to (\mathcal{O}/\mathfrak{p}^2)^\times$. This easily implies (b). Finally, (a) follow from the fact that:

Lemma I.4.11. *$(\mathbb{Z}/p^n)^\times$ is cyclic for all prime numbers p and positive integers n.*

Proof. When $n = 1$, the group $(\mathbb{Z}/p)^\times$ is cyclic, since it's the multiplicative group of a field. Namely, if the exponent of $(\mathbb{Z}^p)^\times$, call it e, is less than $p-1$, then $x^e - 1$ has $p - 1 > e$ solutions in the Unique Factorization Domain $(\mathbb{Z}/p)[x]$, contradiction.

To show $\mathbb{Z}/p^{n+1}$ is cyclic for an arbitrary n, we must show that $(\mathbb{Z}/p^{n+1})^\times$ has an element of order $p^n(p-1)$. Consider the map of reduction modulo p

$$\begin{array}{rcl} \pi : (\mathbb{Z}/p^{n+1})^\times & \to & (\mathbb{Z}/p)^\times \\ g & \mapsto & g(\mathrm{mod}\ p) \end{array}.$$

Then $(\mathbb{Z}/p)^\times$ has a primitive root of order $p-1$; the order of all its pre-images will be divisible by $p-1$. Since $p-1$ is coprime to p^n, in order to find a primitive root of $(\mathbb{Z}/p^{n+1})^\times$ it suffices to find an element of order p^n.

Now the kernel of π consists of elements that are congruent to 1 modulo p, therefore $1 + p\mathbb{Z}$. The order of Ker π is p^n, so it is a p-group. Let $h \in$ Ker π be so $h \not\equiv 1(\mathrm{mod}\ p^2)$. To show the order of h is p^n, it suffices to show $h^{p^{n-1}} \not\equiv 1$. Now we can write $h = 1 + p\alpha + p^2\beta$ with $\alpha, \beta \in \mathbb{Z}$, where $p \nmid \alpha$. Then we will have

$$h^{p^{n-1}} = 1 + p\alpha \cdot p^{n-1} + p^n \cdot \ldots \equiv 1 + \alpha p^n (\mathrm{mod}\ p^{n+1}).$$

Since $p \nmid \alpha$, then $h^{p^{n-1}} \not\equiv 1(\mathrm{mod}\ p^{n+1})$. Therefore, the order of h is p^n, which implies that $(\mathbb{Z}/p^{n+1})^\times$ has a primitive root.

□

Let $\mathfrak{p}$ be a $\mathbb{Q}$-split prime, let $v = v_{\mathfrak{p}}$ be the corresponding valuation. We will now define the *level* $\ell_{\mathfrak{p}}(u)$ of an element $u \in \mathcal{O}_v^\times$ and establish some properties of this notion that we will need later.

Let $p > 2$ be the corresponding rational prime. The group of p-adic units $\mathbb{U}_p = \mathbb{Z}_p^\times$ has the natural filtration by the congruence subgroups

$$\mathbb{U}_p^{(i)} = 1 + p^i\mathbb{Z}_p \quad \text{for} \quad i \in \mathbb{N}.$$

It is well-known [**PR2**, 1.1.3] that

$$\mathbb{U}_p = C \times \mathbb{U}_p^{(1)}$$

where C is the cyclic group of order $(p-1)$ consisting of all roots of unity in $\mathbb{Q}_p$. Furthermore, the logarithmic map yields a continuous isomorphism $\mathbb{U}_p^{(i)} \to p^i\mathbb{Z}_p$, which implies that for any $u \in \mathbb{U}_p \setminus C$, the closure of the cyclic group generated by u has a decomposition of the form

$$\overline{\langle u \rangle} = C' \times \mathbb{U}_p^{(\ell)}$$

for some subgroup $C' \subset C$ and some integer $\ell = \ell_p(u) \geq 1$ which we will refer to as the *p-level* of u. We also set $\ell_p(u) = \infty$ for $u \in C$.

Returning now to a $\mathbb{Q}$-split prime $\mathfrak{p}$ of $\mathcal{O}$ and keeping the above notations, we define the $\mathfrak{p}$-*level* $\ell_{\mathfrak{p}}(u)$ of $u \in \mathcal{O}_v^{\times}$ as the the p-level of the element in $\mathbb{U}_p$ that corresponds to u under the natural identification $\mathcal{O}_v = \mathbb{Z}_p$. We will need the following.

Lemma II.2.2. *Let $\mathfrak{p}$ be a $\mathbb{Q}$-split prime in $\mathcal{O}$, let p be the corresponding rational prime, and $v = v_{\mathfrak{p}}$ the corresponding valuation. Suppose we are given an integer $d \geq 1$ not divisible by p, an element $u \in \mathcal{O}_v^{\times}$ of infinite order having $\mathfrak{p}$-level $s = \ell_{\mathfrak{p}}(u)$, an integer n_s, and an element $c \in \mathcal{O}_v$ such that $u^{n_s} \equiv c \pmod{\mathfrak{p}^s}$. Then for any $t \geq s$ there exists an integer $n_t \equiv n_s \pmod{d}$ for which $u^{n_t} \equiv c \pmod{\mathfrak{p}^t}$.*

Proof. In view of the identification $\mathcal{O}_v = \mathbb{Z}_p$, it is enough to prove the corresponding statement for $\mathbb{Z}_p$. More precisely, we need to show the following: *Let $u \in \mathbb{U}_p$ be a unit of infinite order and p-level $s = \ell_p(u)$. If $c \in \mathbb{U}_p$ and $n_s \in \mathbb{Z}$ are such that $u^{n_s} \equiv c \pmod{p^s}$, then for any $t \geq s$ there exists $n_t \equiv n_s (\mathrm{mod}\, d)$ such that $u^{n_t} \equiv c \pmod{p^t}$.* Thus, we have that $u^{n_s} \in c\mathbb{U}_p^{(s)}$, and we wish to show that

$$u^{n_s} \cdot \langle u^d \rangle \bigcap c\mathbb{U}_p^{(t)} \neq \emptyset.$$

Since $c\mathbb{U}_p^{(t)}$ is open, it is enough to show that

$$u^{n_s} \cdot \overline{\langle u^d \rangle} \bigcap c\mathbb{U}_p^{(t)} \neq \emptyset. \tag{11}$$

But since $\ell_p(u) = s$ and d is prime to p, we have the inclusion $\overline{\langle u^d \rangle} \supset \mathbb{U}_p^{(s)}$, and (11) is obvious. □

2. Dirichlet's Theorem for $\mathbb{Q}$-split primes. We will now establish the existence of $\mathbb{Q}$-split primes in arithmetic progressions.

Theorem II.2.3. *Let $\mathcal{O}$ be the ring of S-integers in a number field k for some finite $S \subset V^k$ containing V_{∞}^k. If nonzero $a, b \in \mathcal{O}$ are relatively prime (i.e., $a\mathcal{O} + b\mathcal{O} = \mathcal{O}$) then there exist infinitely many principal $\mathbb{Q}$-split prime ideals $\mathfrak{p}$ of $\mathcal{O}$ with a generator π such that $\pi \equiv a \pmod{b\mathcal{O}}$ and $\pi > 0$ in all real completions of k.*

The proof follows the same general strategy as the proof of Dirichlet's Theorem in [BMS] – see Theorem A.10 in the Appendix on Number Theory. First, we will quickly recall some basic facts from global class field theory (cf. §I.4 or [**CF**, Ch. VII]) and fix some notations. Let J_k denote the *group of idèles* of k with the natural topology; as usual, we identify $k^{\times}$ with the (discrete) *subgroup of principal idèles* in J_k. Then for every open subgroup $\mathcal{U} \subset J_k$ of finite index containing $k^{\times}$ there exists a finite abelian Galois extension L/k and a continuous surjective homomorphism $\alpha_{L/k} \colon J_k \to \mathrm{Gal}(L/k)$ (known as the *Artin map*; see subsection I.4.2 for details) such that

- $\mathcal{U} = \mathrm{Ker}\, \alpha_{L/k} = N_{L/k}(J_L)k^{\times}$;
- for every nonarchimedean $v \in V^k$ which is unramified in L we let $\mathrm{Fr}_{L/k}(v)$ denote the Frobenius automorphism of L/k at v (i.e., the Frobenius automorphism $\mathrm{Fr}_{L/k}(w|v)$ associated to some (equivalently, any) extension $w|v$) and let $\mathbf{i}(v) \in J_k$ be an idèle with the components
$$\mathbf{i}(v)_{v'} = \begin{cases} 1 & , \quad v' \neq v \\ \pi_v & , \quad v' = v \end{cases},$$
where $\pi_v \in k_v$ is a uniformizer; then $\alpha_{L/k}(\mathbf{i}(v)) = \mathrm{Fr}_{L/k}(v)$.

For our fixed finite subset $S \subset V^k$ containing V^k_∞, we define the following open subgroup of J_k:

$$U_S := \prod_{v \in S} k_v^\times \times \prod_{v \in V^k \setminus S} U_v.$$

Then the abelian extension of k corresponding to the subgroup $\mathcal{U}_S := U_S k^\times$ will be called the *Hilbert S-class field* of k and denoted K throughout the rest of the paper.

Next, we will introduce the idelic S-analogs of *ray groups*Let $\mathfrak{b}$ be a nonzero ideal of $\mathcal{O} = \mathcal{O}_{k,S}$ with the prime factorization

$$\mathfrak{b} = \mathfrak{p}_1^{n_1} \cdots \mathfrak{p}_t^{n_t}, \tag{12}$$

let $v_i = v_{\mathfrak{p}_i}$ be the valuation in $V^k \setminus S$ associated with $\mathfrak{p}_i$, and let $V(\mathfrak{b}) = \{v_1, \ldots, v_t\}$. We then define an open subgroup

$$R_S(\mathfrak{b}) := \prod_{v \in S \setminus V_r^k} k_v^\times \times \prod_{v \in V_r^k} (k_v^+)^\times \times \prod_{v \in V^k \setminus S} U_v^{(v(\mathfrak{b}))},$$

where V_r^k is the set of real valuations, k_v^+ is the subgroup of positive elements, and $U_v^{(v(\mathfrak{b}))}$ is the congruence subgroup of U_{v_i} modulo $\hat{\mathfrak{b}} U_v$. More precisely, if $v \notin V(\mathfrak{b})$ then $U_v^{(v(\mathfrak{b}))} = U_v$, and for $v = v_i \in V(\mathfrak{b})$, $U_v^{(v(\mathfrak{b}))} = U_v^{(n_i)}$, the congruence subgroup of U_{v_i} modulo $\hat{\mathfrak{p}}_{v_i}^{n_i}$.

We then let $K(\mathfrak{b})$ denote the abelian extension of k corresponding to $\mathbf{R}_S(\mathfrak{b}) := R_S(\mathfrak{b}) k^\times$ ("ray class field"). (Obviously, $K(\mathfrak{b})$ contains K for any (nonzero) ideal $\mathfrak{b}$ of $\mathcal{O}$.) Furthermore, for a given $c \in k^\times$, we let $\mathbf{j}_\mathfrak{b}(c)$ denote an idèle with the following components:

$$\mathbf{j}_\mathfrak{b}(c)_v = \begin{cases} c & , \quad v \in V(\mathfrak{b}), \\ 1 & , \quad v \notin V(\mathfrak{b}). \end{cases}$$

Then $\theta_\mathfrak{b} \colon k^\times \to \mathrm{Gal}(K(\mathfrak{b})/k)$ defined by $c \mapsto \alpha_{K(\mathfrak{b})/k}(\mathbf{j}_\mathfrak{b}(c))^{-1}$ is a group homomorphism.

The following lemma summarizes some simple properties of these definitions.

Lemma II.2.4. *Let $\mathfrak{b} \subset \mathcal{O}$ be a nonzero ideal.*

(a) *If a nonzero $c \in \mathcal{O}$ is relatively prime to $\mathfrak{b}$ (i.e. $c\mathcal{O} + \mathfrak{b} = \mathcal{O}$) then $\theta_\mathfrak{b}(c)$ restricts to the Hilbert S-class field K trivially.*

(b) *If nonzero $c_1, c_2 \in \mathcal{O}$ are both relatively prime to $\mathfrak{b}$ then $c_1 \equiv c_2 \pmod{\mathfrak{b}}$ is equivalent to*

$$\mathrm{pr}_\mathfrak{b}\left(\mathbf{j}_\mathfrak{b}(c_1) R_S(\mathfrak{b})\right) = \mathrm{pr}_\mathfrak{b}\left(\mathbf{j}_\mathfrak{b}(c_2) R_S(\mathfrak{b})\right) \tag{13}$$

where $\mathrm{pr}_\mathfrak{b} \colon J_k \to \prod_{v \in V(\mathfrak{b})} k_v^\times$ is the natural projection.

Proof. (a): Since c is relatively prime to $\mathfrak{b}$, we have $\mathbf{j}_\mathfrak{b}(c) \in U_S$. So, using the functoriality properties of the norm residue map, we obtain

$$\theta_\mathfrak{b}(c)|K = \alpha_{K(\mathfrak{b})/k}(\mathbf{j}_\mathfrak{b}(c))^{-1}|K = \alpha_{K/k}(\mathbf{j}_\mathfrak{b}(c))^{-1} = \mathrm{id}_K$$

because $\mathbf{j}_\mathfrak{b}(c) \in U_S \subset \mathcal{U}_S = \mathrm{Ker}\, \alpha_{K/k}$, as required.

(b): As above, let (12) be the prime factorization of $\mathfrak{b}$, let $v_i = v_{\mathfrak{p}_i} \in V^k \setminus S$ be the valuation associated with $\mathfrak{p}_i$. Then for any $c_1, c_2 \in \mathcal{O}$, the congruence $c_1 \equiv c_2 \pmod{\mathfrak{b}}$ is equivalent to

$$c_1 \equiv c_2 \pmod{\hat{\mathfrak{p}}_{v_i}^{n_i}} \quad \text{for all} \quad i = 1, \ldots, t. \tag{14}$$

On the other hand, for any $v \in V_f^k$ and any $u_1, u_2 \in U_v$, the congruence $u_1 \equiv u_2 \pmod{\hat{\mathfrak{p}}_v^n}$ for $n \geq 1$ is equivalent to

$$u_1 U_v^{(n)} = u_2 U_v^{(n)},$$

where $U_v^{(n)}$ is the congruence subgroup of U_v modulo $\hat{\mathfrak{p}}_v^n$. Thus, for (nonzero) $c_1, c_2 \in \mathcal{O}$ prime to $\mathfrak{b}$, the conditions (13) and (14) are equivalent, and our assertion follows. □

We will now establish a result needed for the proof of Theorem II.2.3 and its refinements.

Proposition II.2.5. *Let $\mathfrak{b}$ be a nonzero ideal of $\mathcal{O}$, let $a \in \mathcal{O}$ be relatively prime to $\mathfrak{b}$, and let F be a finite Galois extension of $\mathbb{Q}$ that contains $K(\mathfrak{b})$. Assume that a rational prime p is unramified in F and there exists an extension w of the p-adic valuation v_p to F such that $\mathrm{Fr}_{F/\mathbb{Q}}(w|v_p)|K(\mathfrak{b}) = \theta_{\mathfrak{b}}(a)$. If the restriction v of w to k is not in $S \cup V(\mathfrak{b})$ then:*

(a) $k_v = \mathbb{Q}_p$;

(b) *the prime ideal $\mathfrak{p} = \mathfrak{p}_v$ of $\mathcal{O}$ corresponding to v is principal with a generator π satisfying $\pi \equiv a \pmod{\mathfrak{b}}$ and $\pi > 0$ in every real completion of k.*

(We note since v is unramified in F which contains $K(\mathfrak{b})$, we in fact *automatically* have that $v \notin V(\mathfrak{b})$.)

Proof. (a): Since the Frobenius $\mathrm{Fr}(w|v_p)$ generates $\mathrm{Gal}(F_w/\mathbb{Q}_p)$, our claim immediately follows from the fact that it acts trivially on k.

(b): According to (a), the local degree $[k_v : \mathbb{Q}_p]$ is 1, hence the residual degree $f(v|v_p)$ is also 1, and therefore

$$\mathrm{Fr}(w|v) = \mathrm{Fr}(w|v_p)^{f(v|v_p)} = \mathrm{Fr}(w|v_p).$$

Thus,

$$\alpha_{K(\mathfrak{b})/k}(\mathbf{i}(v)) = \mathrm{Fr}(w|v)|K(\mathfrak{b}) = \theta_{\mathfrak{b}}(a) = \alpha_{K(\mathfrak{b})/k}(\mathbf{j}_{\mathfrak{b}}(a))^{-1},$$

and therefore

$$\mathbf{i}(v)\mathbf{j}_{\mathfrak{b}}(a) \in \mathrm{Ker}\, \alpha_{K(\mathfrak{b})/K} = \mathbf{R}_S(\mathfrak{b}) = R_S(\mathfrak{b})k^{\times}.$$

So, we can write

$$\mathbf{i}(v)\mathbf{j}_{\mathfrak{b}}(a) = \mathbf{r}\pi \quad \text{with} \quad \mathbf{r} \in R_S(\mathfrak{b}),\ \pi \in k^{\times}. \tag{15}$$

Then

$$\pi = \mathbf{i}(v)(\mathbf{j}_{\mathfrak{b}}(a)\mathbf{r}^{-1}).$$

Since a is prime to $\mathfrak{b}$, the idèle $\mathbf{j}_{\mathfrak{b}}(a) \in U_S$, and then $\mathbf{j}_{\mathfrak{b}}(a)\mathbf{r}^{-1} \in U_S$. For any $v' \in V^k \setminus (S \cup \{v\})$, the v'-component of $\mathbf{i}(v)$ is trivial, so we obtain that $\pi \in U_{v'}$. On the other hand, the v-component of $\mathbf{i}(v)$ is a uniformizer π_v of k_v implying that π is also a uniformizer. Thus, $\mathfrak{p} = \pi\mathcal{O}$ is precisely the prime ideal associated with v. For any real v', the v'-components of $\mathbf{i}(v)$ and $\mathbf{j}_{\mathfrak{b}}(a)$ are trivial, so π equals the inverse of the v'-component of $\mathbf{r}$, hence positive in $k_{v'}$. Finally, it follows from (15) that

$$\mathrm{pr}_{\mathfrak{b}}(\mathbf{j}_{\mathfrak{b}}(a)) = \mathrm{pr}_{\mathfrak{b}}(\mathbf{j}_{\mathfrak{b}}(\pi)\mathbf{r}),$$

so $\pi \equiv\ a \pmod{\mathfrak{b}}$ by Lemma II.2.4(b), as required. □

Proof of Theorem II.2.3. Set $\mathfrak{b} = b\mathcal{O}$ and $\sigma = \theta_{\mathfrak{b}}(a) \in \mathrm{Gal}(K(\mathfrak{b})/k)$. Let F be the Galois closure of $K(\mathfrak{b})$ over $\mathbb{Q}$, and let $\tau \in \mathrm{Gal}(F/\mathbb{Q})$ be such that $\tau|K(\mathfrak{b}) = \sigma$. Applying Chebotarev's Density Theorem (see [**CF**, Ch. VII, 2.4] or [**BMS**, A.6]), we find infinitely many rational primes $p > 2$ for which the p-adic valuation v_p is unramified in F, does not lie below any valuations in $S \cup V(\mathfrak{b})$, and has an extension w to F such that $\mathrm{Fr}_{F/\mathbb{Q}}(w|v_p) = \tau$. Let $v = w|k$, and let $\mathfrak{p} = \mathfrak{p}_v$ be the corresponding prime ideal of $\mathcal{O}$. Since $p > 2$, part (a) of Proposition II.2.5 implies that $\mathfrak{p}$ is $\mathbb{Q}$-split. Furthermore, part (b) of it asserts that $\mathfrak{p}$ has a generator π such that $\pi \equiv a \pmod{\mathfrak{b}}$ and $\pi > 0$ in every real completion of k, as required. □

We will now prove a statement from Galois theory that we will need in the next subsection.

Lemma II.2.6. *Let $F/\mathbb{Q}$ be a finite Galois extension, and let κ be an integer for which $F \cap \mathbb{Q}^{\mathrm{ab}} \subseteq \mathbb{Q}(\zeta_\kappa)$. Then $F(\zeta_\kappa) \cap \mathbb{Q}^{\mathrm{ab}} = \mathbb{Q}(\zeta_\kappa)$.*

Proof. We need to show that

$$[F(\zeta_\kappa) : F(\zeta_\kappa) \cap \mathbb{Q}^{\mathrm{ab}}] = [F(\zeta_\kappa) : \mathbb{Q}(\zeta_\kappa)]. \tag{16}$$

Let

$$G = \mathrm{Gal}(F(\zeta_\kappa)/\mathbb{Q}) \quad \text{and} \quad H = \mathrm{Gal}(F/\mathbb{Q}).$$

Then the left-hand side of (16) is equal to the order of the commutator subgroup $[G, G]$, while the right-hand side equals

$$[F : F \cap \mathbb{Q}(\zeta_\kappa)] = [F : F \cap \mathbb{Q}^{\mathrm{ab}}] = |[H, H]|.$$

Now, the restriction gives an *injective* group homomorphism

$$\psi \colon G \to H \times \mathrm{Gal}(\mathbb{Q}(\zeta_\kappa)/\mathbb{Q}).$$

Since the restriction $G \to H$ is surjective, we obtain that ψ implements an isomorphism between $[G, G]$ and $[H, H] \times \{1\}$. Thus, $[G, G]$ and $[H, H]$ have the same order, and (16) follows.

□

3. Key statement. In this subsection we will establish another number-theoretic statement which plays a crucial role in the proof of Theorem II.3.6. To formulate it, we need to introduce some additional notations. As above, let $\mu = |\mu(k)|$ be the number of roots of unity in k, let K be the Hilbert S-class field of k, and let $\tilde{K}$ be the Galois closure of K over $\mathbb{Q}$. Suppose we are given two finite sets P and Q of rational primes. Let

$$\mu' = \mu \cdot \prod_{p \in P} p,$$

pick an integer $\lambda \geq 1$ which is divisible by μ and for which $\tilde{K} \cap \mathbb{Q}^{\mathrm{ab}} \subseteq \mathbb{Q}(\zeta_\lambda)$, and set

$$\lambda' = \lambda \cdot \prod_{q \in Q} q.$$

Theorem II.2.7. *Let $u \in \mathcal{O}^\times$ be a unit of infinite order such that $u \notin \mu(k)_p (k^\times)^p$ for every prime $p \in P$, and let $\mathfrak{q}$ be a $\mathbb{Q}$-split prime of $\mathcal{O}$ which is relatively prime to $\lambda'\mathcal{O}$. Then there exist infinitely many principal $\mathbb{Q}$-split primes $\mathfrak{p} = \pi\mathcal{O}$ of $\mathcal{O}$ with a generator π such that*

(1) *for each $p \in P$, the p-primary component of $\phi(\mathfrak{p})/\mu$ divides the p-primary component of the order of u* $\pmod{\mathfrak{p}}$*;*

(2) $\pi \pmod{\mathfrak{q}^2}$ *generates* $(\mathcal{O}/\mathfrak{q}^2)^\times$*;*

(3) $\gcd(\phi(\mathfrak{p}), \lambda') = \lambda$.

Proof. As in the proof of Theorem II.2.3, we will derive the required assertion by applying Chebotarev's Density Theorem (see [**CF**, Ch. VII, 2.4] or [**BMS**, A.6]) to a specific automorphism of an appropriate finite Galois extension.

Let $K(\mathfrak{q}^2)$ be the abelian extension $K(\mathfrak{b})$ of k introduced in Subsection II.2.2 for the ideal $\mathfrak{b} = \mathfrak{q}^2$. Set

$$L_1 = K(\mathfrak{q}^2)(\zeta_{\lambda'}), \quad L_2 = k\left(\zeta_{\mu'}, \sqrt[\mu']{u}\right), \quad L = L_1L_2 \quad \text{and} \quad \ell = L_1 \cap L_2.$$

Then

$$\mathrm{Gal}(L/k) = \{\, \sigma = (\sigma_1, \sigma_2) \in \mathrm{Gal}(L_1/k) \times \mathrm{Gal}(L_2/k) \ \mid\ \sigma_1|\ell = \sigma_2|\ell \,\}. \tag{17}$$

So, to construct $\sigma \in \mathrm{Gal}(L/k)$ that we will need in the argument it is enough to construct appropriate $\sigma_i \in \mathrm{Gal}(L_i/k)$ for $i = 1, 2$ that have the same restriction to ℓ.

Lemma II.2.8. *The restriction maps define the following isomorphisms:*

(1) $\mathrm{Gal}(L_1/K) \simeq \mathrm{Gal}(K(\mathfrak{q}^2)/K) \times \mathrm{Gal}(K(\zeta_{\lambda'})/K)$*;*

(2) $\mathrm{Gal}(K(\zeta_{\lambda'})/K(\zeta_\lambda)) \simeq \mathrm{Gal}(\mathbb{Q}(\zeta_{\lambda'})/\mathbb{Q}(\zeta_\lambda)) \simeq \displaystyle\prod_{q\in Q} \mathrm{Gal}(\mathbb{Q}(\zeta_{q\lambda})/\mathbb{Q}(\zeta_\lambda))$.

Proof. (1): We need to show that $K(\mathfrak{q}^2) \cap K(\zeta_{\lambda'}) = K$. But the Galois extensions $K(\mathfrak{q}^2)/K$ and $K(\zeta_{\lambda'})/K$ are respectively totally and un-ramified at the extensions of $v_{\mathfrak{q}}$ to K (since $\mathfrak{q}$ is prime to λ'), so the required fact is immediate.

(2): Since $K(\zeta_{\lambda'}) = K(\zeta_\lambda) \cdot \mathbb{Q}(\zeta_{\lambda'})$, we only need to show that

$$K(\zeta_\lambda) \cap \mathbb{Q}(\zeta_{\lambda'}) = \mathbb{Q}(\zeta_\lambda). \tag{18}$$

We have

$$K(\zeta_\lambda) \cap \mathbb{Q}(\zeta_{\lambda'}) \subseteq \tilde{K}(\zeta_\lambda) \cap \mathbb{Q}^{\mathrm{ab}} = \mathbb{Q}(\zeta_\lambda)$$

by Lemma II.2.6. This proves one inclusion in (18); the other inclusion is obvious. □

Since $\mathfrak{q}$ is $\mathbb{Q}$-split, the group $(\mathcal{O}/\mathfrak{q}^2)^\times$ is cyclic (by Lemma II.2.1(a)), and we pick $c \in \mathcal{O}$ so that $c \pmod{\mathfrak{q}^2}$ is a generator of this group. We then set

$$\sigma_1' = \theta_{\mathfrak{q}^2}(c) \in \mathrm{Gal}(K(\mathfrak{q}^2)/K)$$

in the notations of Subsection II.2.2 (cf. Lemma II.2.4(a)). Next, for $q \in Q$, we let $q^{e(q)}$ be the q-primary component of λ. Then using the isomorphism from Lemma II.2.8(2), we can find $\sigma_1'' \in \mathrm{Gal}(K(\zeta_{\lambda'})/K)$ such that

$$\sigma_1''(\zeta_\lambda) = \zeta_\lambda \ \text{ but } \ \sigma_1''(\zeta_{q^{e(q)+1}}) \neq \zeta_{q^{e(q)+1}} \ \text{ for all } \ q \in Q. \tag{19}$$

We then define $\sigma_1 \in \mathrm{Gal}(L_1/K)$ to be the automorphism corresponding to the pair (σ_1', σ_1'') in terms of the isomorphism from Lemma II.2.8(1) (in other words, the restrictions of σ_1 to $K(\mathfrak{q}^2)$ and $K(\zeta_{\lambda'})$ are σ_1' and σ_1'', respectively).

We fix a μ'-th root $\sqrt[\mu']{u}$, and for $\nu|\mu'$ set $\sqrt[\nu]{u} = \left(\sqrt[\mu']{u}\right)^{\mu'/\nu}$ (also denoted $u^{\nu^{-1}}$). To construct $\sigma_2 \in \mathrm{Gal}(L_2/k)$, we need the following.

Lemma II.2.9. *Let $\sigma_0 \in \mathrm{Gal}(\ell/k)$. Then there exists $\sigma_2 \in \mathrm{Gal}(L_2/k)$ such that*

(1) $\sigma_2|\ell = \sigma_0$;

(2) *for any $p \in P$, if $p^{d(p)}$ is the p-primary component of μ then*

$$\sigma_2\left(u^{p^{-(d(p)+1)}}\right) \neq u^{p^{-(d(p)+1)}},$$

and consequently either $\sigma_2(\zeta_{p^{d(p)+1}}) \neq \zeta_{p^{d(p)+1}}$ or σ_2 acts nontrivially on <u>all</u> $p^{d(p)+1}$-th roots of u.

Proof. Since L_1/k is an abelian extension, we conclude from Corollary II.1.6 that

$$\ell \subseteq k\left(\sqrt[\mu]{u}, \zeta_{\mu'}\right) \subseteq k^{\mathrm{ab}}. \tag{20}$$

On the other hand, according to Proposition II.1.1, none of the roots $\sqrt[p\mu]{u}$ for $p \in P$ lies in k^{ab}, and the restriction maps yield an isomorphism

$$\mathrm{Gal}\left(k\left(\sqrt[\mu']{u}, \zeta_{\mu'}\right)/k\left(\sqrt[\mu]{u}, \zeta_{\mu'}\right)\right) \to \prod_{p \in P} \mathrm{Gal}\Big(k\left(\sqrt[p\mu]{u}, \zeta_{\mu'}\right)/k\left(\sqrt[\mu]{u}, \zeta_{\mu'}\right)\Big).$$

It follows that for each $p \in P$ we can find $\tau_p \in \mathrm{Gal}\Big(k\left(\sqrt[\mu']{u}, \zeta_{\mu'}\right)/k\left(\sqrt[\mu]{u}, \zeta_{\mu'}\right)\Big)$ such that

$$\tau_p\left(u^{p^{-(d(p)+1)}}\right) = \zeta_p \cdot u^{p^{-(d(p)+1)}} \quad \text{and} \quad \tau_p\left(u^{q^{-(d(q)+1)}}\right) = u^{q^{-(d(q)+1)}} \quad \text{for all} \quad q \in P \setminus \{p\}.$$

Now, let $\tilde{\sigma}_0$ be any extension of σ_0 to L_2. For $p \in P$, define

$$\chi(p) = \begin{cases} 1 & , \quad \tilde{\sigma}_0\left(u^{p^{-(d(p)+1)}}\right) = u^{p^{-(d(p)+1)}} \\ 0 & , \quad \tilde{\sigma}_0\left(u^{p^{-(d(p)+1)}}\right) \neq u^{p^{-(d(p)+1)}} \end{cases}$$

Set

$$\sigma_2 = \tilde{\sigma}_0 \cdot \prod_{p \in P} \tau_p^{\chi(p)}.$$

In view of (20), all τ_p's act trivially on ℓ, so $\sigma_2|\ell = \tilde{\sigma}_0|\ell = \sigma_0$ and (1) holds. Furthermore, the choice of the τ_p's and the $\chi(p)$'s implies that (2) also holds. □

Continuing the proof of Theorem II.2.7, we now use $\sigma_1 \in \mathrm{Gal}(L_1/k)$ constructed above, set $\sigma_0 = \sigma_1|\ell$, and using Lemma II.2.9 construct $\sigma_2 \in \mathrm{Gal}(L_2/k)$ with the properties described therein. In particular, part (1) of this lemma in conjunction with (17) implies that the pair (σ_1, σ_2) corresponds to an automorphism $\sigma \in \mathrm{Gal}(L/k)$. As in the proof of Theorem II.2.3, we let F denote the Galois closure of L over $\mathbb{Q}$, and let $\tilde{\sigma} \in \mathrm{Gal}(F/Q)$ be such that $\tilde{\sigma}|L = \sigma$. By Chebotarev's Density Theorem, there exist infinitely many rational primes $p > 2$ that are relatively prime to $\lambda' \cdot \mu'$ and for which the p-adic valuation v_p is unramified in F, does not lie below any valuation in $S \cup \{v_{\mathfrak{q}}\}$, and has an extension w to F such that $\mathrm{Fr}_{F/\mathbb{Q}}(w|v_p) = \tilde{\sigma}$. Let $v = w|k$, and let $\mathfrak{p} = \mathfrak{p}_v$ be the corresponding prime ideal of $\mathcal{O}$. As in the proof of Theorem II.2.3, we see that $\mathfrak{p}$ is $\mathbb{Q}$-split. Furthermore, since $\sigma|K(\mathfrak{q}^2) = \theta_{\mathfrak{q}^2}(c)$, we conclude that $\mathfrak{p}$ has a generator π such that $\pi \equiv c \pmod{\mathfrak{q}^2}$ (cf. Proposition II.2.5(b)). Then by construction $\pi \pmod{\mathfrak{q}^2}$ generates $(\mathcal{O}/\mathfrak{q}^2)^{\times}$, verifying condition (2) of Theorem II.2.7.

To verify condition (1), we fix $p \in P$ and consider two cases. First, suppose $\sigma(\zeta_{p^{d(p)+1}}) \neq \zeta_{p^{d(p)+1}}$. Since p is prime to $\mathfrak{p}$, this means that the residue field $\mathcal{O}/\mathfrak{p}$ does not contain an element of order

$p^{d(p)+1}$ (although, since μ is prime to $\mathfrak{p}$, it does contain an element of order μ, hence of order $p^{d(p)}$). So, in this case $\phi(\mathfrak{p})/\mu$ is prime to p, and there is nothing to prove. Now, suppose that $\sigma(\zeta_{p^{d(p)+1}}) = \zeta_{p^{d(p)+1}}$. Then by construction σ acts nontrivially on every $p^{d(p)+1}$-th root of u, and therefore the polynomial $X^{p^{d(p)+1}} - u$ has no roots in $k_{v_{\mathfrak{p}}}$. Again, since p is prime to $\mathfrak{p}$, we see from Hensel's Lemma I.4.4 that $u \pmod{\mathfrak{p}}$ is not a $p^{d(p)+1}$-th power in the residue field. It follows that the p-primary component of the order of $u \pmod{\mathfrak{p}}$ is not less than the p-primary component of $\phi(\mathfrak{p})/p^{d(p)}$, and (1) follows.

Finally, by construction σ acts trivially on ζ_λ but nontrivially on $\zeta_{q\lambda}$ for any $q \in Q$. Since $\mathfrak{p}$ is prime to λ', we see that the residue field $\mathcal{O}/\mathfrak{p}$ contains an element of order λ, but does not contain an element of order $q\lambda$ for any $q \in Q$. This means that $\lambda \mid \phi(\mathfrak{p})$ but $\phi(\mathfrak{p})/\lambda$ is relatively prime to each $q \in Q$, which is equivalent to condition (3) of Theorem II.2.7. □

II.3. Proof of Theorem II.3.6

First, we will introduce some additional notations needed to convert the task of factoring a given matrix $A \in \mathrm{SL}_2(\mathcal{O})$ as a product of elementary matrices into the task of reducing the first row of A to $(1, 0)$. Let

$$\mathcal{R}(\mathcal{O}) = \{(a, b) \in \mathcal{O}^2 \mid a\mathcal{O} + b\mathcal{O} = \mathcal{O}\}$$

(note that $\mathcal{R}(\mathcal{O})$ is precisely the set of all first rows of matrices $A \in \mathrm{SL}_2(\mathcal{O})$). For $\lambda \in \mathcal{O}$, one defines two permutations, $e_+(\lambda)$ and $e_-(\lambda)$, of $\mathcal{R}(\mathcal{O})$ given respectively by

$$(a, b) \mapsto (a, b + \lambda a) \quad \text{and} \quad (a, b) \mapsto (a + \lambda b, b).$$

These permutations will be called *elementary transformations* of $\mathcal{R}(\mathcal{O})$. For (a, b), $(c, d) \in \mathcal{R}(\mathcal{O})$ we will write $(a, b) \overset{n}{\Rightarrow} (c, d)$ to indicate the fact that (c, d) can be obtained from (a, b) by a sequence of n (equivalently, $\leq n$) elementary transformations. For the convenience of further reference, we will record some simple properties of this relation.

Lemma II.3.1. *Let $(a, b) \in \mathcal{R}(\mathcal{O})$.*

(1a) *If $(c, d) \in \mathcal{R}(\mathcal{O})$ and $(a, b) \overset{n}{\Rightarrow} (c, d)$, then $(c, d) \overset{n}{\Rightarrow} (a, b)$.*

(1b) *If $(c, d), (e, f) \in \mathcal{R}(\mathcal{O})$ are such that $(a, b) \overset{m}{\Rightarrow} (c, d)$ and $(c, d) \overset{n}{\Rightarrow} (e, f)$, then $(a, b) \overset{m+n}{\Rightarrow} (e, f)$.*

(2a) *If $c \in \mathcal{O}$ such that $c \equiv a (\mathrm{mod}\, b\mathcal{O})$, then $(c, b) \in \mathcal{R}(\mathcal{O})$, and $(a, b) \overset{1}{\Rightarrow} (c, b)$.*

(2b) *If $d \in \mathcal{O}$ such that $d \equiv b (\mathrm{mod}\, a\mathcal{O})$, then $(a, d) \in \mathcal{R}(\mathcal{O})$, and $(a, b) \overset{1}{\Rightarrow} (a, d)$.*

(3a) *If $(a, b) \overset{n}{\Rightarrow} (1, 0)$ then any matrix $A \in \mathrm{SL}_2(\mathcal{O})$ with the first row (a, b) is a product of $\leq n + 1$ elementary matrices.*

(3b) *If $(a, b) \overset{n}{\Rightarrow} (0, 1)$ then any matrix $A \in \mathrm{SL}_2(\mathcal{O})$ with the second row (a, b) is a product of $\leq n + 1$ elementary matrices.*

(4a) *If $a \in \mathcal{O}^\times$ then $(a, b) \overset{2}{\Rightarrow} (0, 1)$.*

(4b) *If $b \in \mathcal{O}^\times$ then $(a, b) \overset{2}{\Rightarrow} (1, 0)$.*

Proof. For (1a), we observe that the inverse of an elementary transformation is again an elementary transformation given by $[e_\pm(\lambda)]^{-1} = e_\pm(-\lambda)$, so the required fact follows. Part (1b) is obvious.

(Note that (1) implies that the relation between (a,b) and $(c,d) \in \mathcal{R}(\mathcal{O})$ defined by $(a,b) \overset{n}{\Rightarrow} (c,d)$ for *some* $n \in \mathbb{N}$ is an equivalence relation.)

In (2a), we have $c = a + \lambda b$ with $\lambda \in \mathcal{O}$. Then

$$c\mathcal{O} + b\mathcal{O} = a\mathcal{O} + b\mathcal{O} = \mathcal{O},$$

so $(c,b) \in \mathcal{R}(\mathcal{O})$, and $e_+(\lambda)$ takes (a,b) to (c,b). The argument for (2b) is similar.

(3a) Suppose $A \in \mathrm{SL}_2(\mathcal{O})$ has the first row (a,b). Then for $\lambda \in \mathcal{O}$, the first row of the product $A\mathrm{E}_{12}(\lambda)$ is $(a, b + \lambda a) = e_+(\lambda)(a,b)$, and similarly the first row of $A\mathrm{E}_{21}(\lambda)$ is $e_-(\lambda)(a,b)$. So, the fact that $(a,b) \overset{n}{\Rightarrow} (1,0)$ implies that there exists a matrix $U \in \mathrm{SL}_2(\mathcal{O})$ which is a product of n elementary matrices and is such that AU has the first row $(1,0)$. This means that $AU = \mathrm{E}_{21}(z)$ for some $z \in \mathcal{O}$, and then $A = \mathrm{E}_{21}(z)U^{-1}$ is a product of $\leq n+1$ elementary matrices. The argument for (3b) is similar.

Part (4a) follows since $e_-\big(-a\big)e_+\big(a^{-1}(1-b)\big)(a,b) = (0,1)$. The proof of (4b) is similar. □

Remark. All assertions of Lemma II.3.1 are valid over any commutative ring $\mathcal{O}$.

COROLLARY II.3.2. *Let $\mathfrak{q}$ be a principal $\mathbb{Q}$-split prime ideal of $\mathcal{O}$ with generator q, and let $z \in \mathcal{O}$ be such that $z(\mathrm{mod}\ \mathfrak{q}^2)$ generates $(\mathcal{O}/\mathfrak{q}^2)^\times$. If we are given an integer t_0 and an element of $\mathcal{R}(\mathcal{O})$ of the form (b, q^n) with $n \geq 2$, then there exists an integer $t \geq t_0$ such that $(b, q^n) \overset{1}{\Rightarrow} (z^t, q^n)$.*

PROOF. By Lemma II.2.1(b), the element $z(\mathrm{mod}\ \mathfrak{q}^n)$ generates $(\mathcal{O}/\mathfrak{q}^n)^\times$. Since b is prime to $\mathfrak{q}$, one can find $t \in \mathbb{Z}$ such that $b \equiv z^t(\mathrm{mod}\ \mathfrak{q}^n)$. Adding to t a suitable multiple of $\phi(\mathfrak{q}^n)$ if necessary, we can assume that $t \geq t_0$. Our assertion then follows from Lemma II.3.1(2a). □

LEMMA II.3.3. *Suppose we are given $(a,b) \in \mathcal{R}(\mathcal{O})$, a finite subset $T \subseteq V_f^k$, and an integer $n \neq 0$. Then there exists $\alpha \in \mathcal{O}_k$ and $r \in \mathcal{O}^\times$ such that $V(\alpha) \cap T = \emptyset$, and $(a,b) \overset{1}{\Rightarrow} (\alpha r^n, b)$.*

PROOF. Let h_k be the class number of k. If for each $v \in S \setminus V_\infty^k$ we let $\mathfrak{m}_v$ denote the maximal ideal of $\mathcal{O}_k$ corresponding to v, then the ideal $(\mathfrak{m}_v)^{h_k}$ is principal, and its generator π_v satisfies $v(\pi_v) = h_k$ and $w(\pi_v) = 0$ for all $w \in V_f^k \setminus \{v\}$. Let R be the subgroup of $k^\times$ generated by π_v for $v \in S \setminus V_\infty^k$; note that $R \subset \mathcal{O}^\times$. We can pick $r \in R$ so that $a' := ar^{-n} \in \mathcal{O}_k$. We note that since a and b are relatively prime in $\mathcal{O}$, we have $V(a') \cap V(b) \subset S$.

Now, it follows from the Strong Approximation Theorem that there exists $\gamma \in \mathcal{O}_k$ such that

$$\begin{aligned} v(\gamma b) \geq 0 \text{ and } v(\gamma b) \equiv 0(\mathrm{mod}\ nh_k) \quad & \text{for all } v \in S \setminus V_\infty^k, \\ \text{and } v(\gamma b) = 0 \quad & \text{for all } v \in V(a') \setminus S. \end{aligned}$$

Then, in particular, we can find $s \in R$ so that $v(\gamma b s^{-1}) = 0$ for all $v \in S \setminus V_\infty^k$. Set

$$\gamma' := \gamma s^{-1} \in \mathcal{O} \text{ and } b' := \gamma' b \in \mathcal{O}_k.$$

By construction,

$$v(b') = 0 \quad \text{for all} \quad v \in V(a') \cup (S \setminus V_\infty^k), \tag{21}$$

implying that $b' \in \mathcal{O}_k$ and $V(a') \cap V(b') = \emptyset$, which means that a' and b' are relatively prime in $\mathcal{O}_k$.

Again, by the Strong Approximation Theorem we can find $t \in \mathcal{O}_k$ such that

$$v(t) = 0 \text{ for } v \in T \cap V(a') \text{ and } v(t) > 0 \text{ for } v \in T \setminus V(a').$$

Set $\alpha = a' + tb' \in \mathcal{O}_k$. Then for $v \in T \cap V(a')$ we have $v(a') > 0$ and $v(tb') = 0$ (in view of (21)), while for $v \in T \setminus V(a')$ we have $v(a') = 0$ and $v(tb') > 0$. In either case,

$$v(\alpha) = v(a' + tb') = 0 \text{ for all } v \in T,$$

i.e. $V(\alpha) \cap T = \emptyset$. On the other hand,

$$a + r^n t\gamma' b = r^n a' + r^n tb') = r^n \alpha,$$

which means that $(a, b) \overset{1}{\Rightarrow} (\alpha r^n, b)$, as required.

□

Recall that we let μ denote the number of roots of unity in k.

Lemma II.3.4. *Let $(a, b) \in \mathcal{R}(\mathcal{O})$ be such that $a = \alpha \cdot r^\mu$ for some $\alpha \in \mathcal{O}_k$ and $r \in \mathcal{O}^\times$ where $V(\alpha)$ is disjoint from $S \cup V(\mu)$. Then there exist $a' \in \mathcal{O}$ and infinitely many $\mathbb{Q}$-split prime principal ideals $\mathfrak{q}$ of $\mathcal{O}$ with a generator q such that for any $m \equiv 1(\mathrm{mod}\ \phi(a'\mathcal{O}))$ we have $(a, b) \overset{3}{\Rightarrow} (a', q^{\mu m})$.*

Proof. The argument below is adapted from the proof of Lemma 3 in [CK1]. It relies on the properties of the power residue symbol (in particular, the power reciprocity law) described in [**BMS**, Appendix on Number Theory]. We will work with all $v \in V^k$ (and not only $v \in V^k \setminus S$), so to each such v we associate a symbol ("modulus") $\mathfrak{m}_v$. For $v \in V_f^k$ we will identify $\mathfrak{m}_v$ with the corresponding maximal ideal of $\mathcal{O}_k$ (obviously, $\mathfrak{p}_v = \mathfrak{m}_v\mathcal{O}$ for $v \in V^k \setminus S$); the valuation ideal and the group of units in the valuation ring $\mathcal{O}_v$ (or $\mathcal{O}_{\mathfrak{m}_v}$) in the completion k_v will be denoted $\hat{\mathfrak{m}}_v$ and U_v respectively. For any divisor $\nu|\mu$, we let

$$\left(\frac{*,*}{\mathfrak{m}_v}\right)_\nu$$

be the (bi-multiplicative, skew-symmetric) power residue symbol of degree ν on $k_v^\times$ (cf. [BMS, p. 85]; see Subsection I.4.3 for details). We recall that $\left(\frac{x,y}{\mathfrak{m}_v}\right)_\nu = 1$ if one of the elements x, y is a ν-th power in $k_v^\times$ (in particular, if v is either complex, or is real and one of the elements x, y is positive in k_v) or if v is nonarchimedean $\notin V(\nu)$ and $x, y \in U_v$. It follows that for any $x, y \in k^\times$, we have $\left(\frac{x,y}{\mathfrak{m}_v}\right)_\nu = 1$ for almost all $v \in V^k$. Furthermore, we have the *reciprocity law* (cf. Lemma I.4.15):

$$\prod_{v \in V^k} \left(\frac{x,y}{\mathfrak{m}_v}\right)_\nu = 1. \tag{22}$$

Now, let $\mu = p_1^{e_1} \cdots p_n^{e_n}$ be a prime factorization of μ. For each $i = 1, \ldots, n$, pick $v_i \in V(p_i)$. According to [BMS, (A.17)], the values

$$\left(\frac{x,y}{\mathfrak{m}_{v_i}}\right)_{p_i^{e_i}} \quad \text{for} \quad x, y \in U_{v_i}$$

cover all $p_i^{e_i}$-th roots of unity. Thus, we can pick units $u_i, u_i' \in U_{v_i}$ for $i = 1, \ldots, n$ so that

$$\left(\frac{u_i, u_i'}{\mathfrak{m}_{v_i}}\right)_{p_i^{e_i}} = \zeta_{p_i^{e_i}},$$

a primitive $p_i^{e_i}$-th root of unity. Then

$$\zeta_\mu := \prod_{i=1}^{n} \left(\frac{u_i, u_i'}{\mathfrak{m}_{v_i}}\right)_{p_i^{e_i}} \tag{23}$$

is a primitive μ-th root of unity. Furthermore, it follows from Hensel's Lemma (cf. Lemma I.4.4) (or, from the Inverse Function Theorem) that we can find an integer $N > 0$ such that

$$1 + \hat{\mathfrak{m}}_v^N \subset k_v^{\times\mu} \quad \text{for all} \quad v \in V(\mu). \tag{24}$$

We now write $b = \beta t^\mu$ with $\beta \in \mathcal{O}_k$ and $t \in \mathcal{O}^\times$. Since a, b are relatively prime in $\mathcal{O}$, then so are α, β, hence $V(\alpha) \cap V(\beta) \subset S$. On the other hand, by our assumption, $V(\alpha)$ is disjoint from $S \cup V(\mu)$, so we conclude that $V(\alpha)$ is disjoint from $V(\beta) \cup V(\mu)$. Applying Theorem II.2.3 to the ring $\mathcal{O}_k$ we obtain that there exists $\beta' \in \mathcal{O}_k$ having the following properties:

$(1)_1$ $\mathfrak{b} := \beta'\mathcal{O}_k$ is a prime ideal of $\mathcal{O}_k$ and the corresponding valuation $v_\mathfrak{b} \notin S \cup V(\mu)$;

$(2)_1$ $\beta' > 0$ in every real completion of k;

$(3)_1$ $\beta' \equiv \beta (\mathrm{mod}\, \alpha\mathcal{O}_k)$;

$(4)_1$ for each $i = 1, \ldots, n$, we have
$\beta' \equiv u_i' (\mathrm{mod}\, \hat{\mathfrak{m}}_{v_i})$, and
$\beta' \equiv 1 (\mathrm{mod}\, \hat{\mathfrak{m}}_v)$ for all $v \in V(p_i) \setminus \{v_i\}$.

Set $b' = \beta' t^\mu$. It is a consequence of $(3)_1$ that $b \equiv b' (\mathrm{mod}\, a\mathcal{O})$, so by Lemma II.3.1(2) we have $(a, b) \overset{1}{\Rightarrow} (a, b')$. Furthermore, it follows from $(4)_1$ and (24) that $\beta'/u_i' \in (k_{v_i}^\times)^\mu$, so

$$\left(\frac{u_i, \beta'}{\mathfrak{m}_{v_i}}\right)_\mu = \left(\frac{u_i, u_i'}{\mathfrak{m}_{v_i}}\right)_\mu = \zeta_{p_i^{e_i}}.$$

Since ζ_μ defined by (23) is a primitive μ-th root of unity, we can find an integer $d > 0$ such that

$$1 = \left(\frac{\alpha, \beta'}{\mathfrak{b}}\right)_\mu \cdot \zeta_\mu^d = \left(\frac{\alpha, \beta'}{\mathfrak{b}}\right)_\mu \cdot \prod_{i=1}^{n} \left(\frac{u_i^d, \beta'}{\mathfrak{m}_{v_i}}\right)_{p_i^{e_i}}. \tag{25}$$

By construction, $v_\mathfrak{b} \notin V(\alpha) \cup V(\mu)$, so applying Theorem II.2.3 one more time, we find $\alpha' \in \mathcal{O}_k$ such that

$(1)_2$ $\mathfrak{a} := \alpha'\mathcal{O}_k$ is a prime ideal of $\mathcal{O}_k$ and the corresponding valuation $v_\mathfrak{a} \notin S \cup V(\mu)$;

$(2)_2$ $\alpha' \equiv \alpha (\mathrm{mod}\, \mathfrak{b})$;

$(3)_2$ $\alpha' \equiv u_i^d (\mathrm{mod}\, \hat{\mathfrak{m}}_{v_i}^N)$ for $i = 1, \ldots, n$.

Set $a' = \alpha' r^\mu$. Then $a'\mathcal{O} = \alpha'\mathcal{O}$ is a prime ideal of $\mathcal{O}$ and $a' \equiv a (\mathrm{mod}\, b'\mathcal{O})$, so $(a, b') \overset{1}{\Rightarrow} (a', b')$.

Now, we note that $\left(\frac{\alpha', \beta'}{\mathfrak{m}_v}\right)_\mu = 1$ if either $v \in V_\infty^k$ (since $\beta' > 0$ in all real completions of k) or $v \in V_f^k \setminus (V(\alpha') \cup V(\beta') \cup V(\mu))$. Since the ideals $\mathfrak{a} = \alpha'\mathcal{O}_k$ and $\mathfrak{b} = \beta'\mathcal{O}_k$ are prime by construction, we have $V(\alpha') = \{v_\mathfrak{a}\}$ and $V(\beta') = \{v_\mathfrak{b}\}$. Besides, it follows from (24) and $(4)_1$ that for $v \in V(p_i) \setminus \{v_i\}$ we have $\beta' \in k_v^{\times\mu}$, and therefore again $\left(\frac{\alpha', \beta'}{\mathfrak{m}_v}\right)_\mu = 1$. Thus, the reciprocity law

(22) for α', β' reduces to the relation

$$(26) \qquad \left(\frac{\alpha', \beta'}{\mathfrak{a}}\right)_{\mu} \cdot \left(\frac{\alpha', \beta'}{\mathfrak{b}}\right)_{\mu} \cdot \prod_{i=1}^{n} \left(\frac{\alpha', \beta'}{\mathfrak{m}_{v_i}}\right)_{\mu} = 1.$$

It follows from $(2)_2$ and $(3)_2$ that

$$\left(\frac{\alpha', \beta'}{\mathfrak{b}}\right)_{\mu} = \left(\frac{\alpha, \beta'}{\mathfrak{b}}\right)_{\mu} \quad \text{and} \quad \left(\frac{\alpha', \beta'}{\mathfrak{m}_{v_i}}\right)_{\mu} = \left(\frac{u_i^d, \beta'}{\mathfrak{b}}\right)_{\mu} \quad \text{for all} \quad i = 1, \ldots, n.$$

Comparing now (25) with (26), we find that

$$\left(\frac{\beta', \alpha'}{\mathfrak{a}}\right)_{\mu} = \left(\frac{\alpha', \beta'}{\mathfrak{a}}\right)_{\mu}^{-1} = 1.$$

This implies (cf. [BMS, (A.16)]; or, (5) and (6)) that β' is a μ-th power modulo $\mathfrak{a}$, i.e. $\beta' \equiv \gamma^{\mu} (\mathrm{mod} \mathfrak{a})$ for some $\gamma \in \mathcal{O}_k$. Clearly, the elements $a' = \alpha' r^{\mu}$ and γt are relatively prime in $\mathcal{O}$, so applying Theorem II.2.3 to this ring, we find infinitely many $\mathbb{Q}$-split principal prime ideals $\mathfrak{q}$ of $\mathcal{O}$ having a generator $q \equiv \gamma t (\mathrm{mod}\, a'\mathcal{O})$. Then for any $m \equiv 1 (\mathrm{mod}\, \phi(a'\mathcal{O}))$ we have

$$q^{\mu m} \equiv q^{\mu} \equiv \beta' t^{\mu} = b' (\mathrm{mod}\, a'\mathcal{O}),$$

so $(a', b') \overset{1}{\Rightarrow} (a', q^{\mu m})$. Then by Lemma II.3.1(1b), we have $(a, b) \overset{3}{\Rightarrow} (a', q^{\mu m})$, as required. □

The final ingredient that we need for the proof of Theorem II.3.6 is the following lemma which uses the notion of the *level* $\ell_{\mathfrak{p}}(u)$ of a unit u of infinite order with respect to a $\mathbb{Q}$-split ideal $\mathfrak{p}$ introduced in Subsection II.2.1.

Lemma II.3.5. *Let $\mathfrak{p}$ be a principal $\mathbb{Q}$-split ideal of $\mathcal{O}$ with a generator π, and let $u \in \mathcal{O}^{\times}$ be a unit of infinite order. Set $s = \ell_{\mathfrak{p}}(u)$, and let δ, λ, and m be positive integers satisfying $\delta \mid \lambda \mid \phi(\mathfrak{p})$, and $m \equiv 0 (\mathrm{mod}\, \phi(\mathfrak{p}^s)/\lambda)$. If $b \in \mathcal{O}$ is a δ-th power (*mod $\mathfrak{p}$*) and is prime to π, while $\nu := \lambda/\delta$ divides the order of u(*mod $\mathfrak{p}$*), then for any integer $t \geq s$ there exists an integer n_t for which*

$$(\pi^t, b^m) \overset{1}{\Rightarrow} (\pi^t, u^{n_t}).$$

Proof. Let p be the rational prime corresponding to $\mathfrak{p}$. Being a divisor of λ, the integer δ is relatively prime to p. So, the fact that b is a δ-th power (mod $\mathfrak{p}$) implies that it is also a δ-th power (mod $\mathfrak{p}^s$). On the other hand, it follows from our assumptions that $\lambda m = \delta \nu m$ is divisible by $\phi(\mathfrak{p}^s)$, and therefore $(b^m)^{\nu} \equiv 1 (\mathrm{mod}\, \mathfrak{p}^s)$. But since ν is prime to p, the subgroup of elements in $(\mathcal{O}/\mathfrak{p}^s)^{\times}$ of order dividing ν is isomorphic to a subgroup of $(\mathcal{O}/\mathfrak{p})^{\times}$, hence cyclic. So, the fact that the order of $u(\mathrm{mod}\, \mathfrak{p})$, and consequently the order of $u(\mathrm{mod}\, \mathfrak{p}^s)$, is divisible by ν implies that every element in $(\mathcal{O}/\mathfrak{p}^s)^{\times}$ whose order divides ν lies in the subgroup generated by $u(\mathrm{mod} \mathfrak{p}^s)$. Thus, $b^m \equiv u^{n_s} (\mathrm{mod}\, \mathfrak{p}^s)$ for some integer n_s. Since $\mathfrak{p}$ is $\mathbb{Q}$-split, we can apply Lemma II.2.2 to conclude that for any $t \geq s$ there exists an integer n_t such that $b^m \equiv u^{n_t} (\mathrm{mod}\, \mathfrak{p}^t)$. Then $(\pi^t, b^m) \overset{1}{\Rightarrow} (\pi^t, u^{n_t})$ by Lemma II.3.1(2). □

We will call a unit $u \in \mathcal{O}^{\times}$ *fundamental* if it has infinite order and the cyclic group $\langle u \rangle$ is a direct factor of $\mathcal{O}^{\times}$. Since the group $\mathcal{O}^{\times}$ is finitely generated (Dirichlet's Unit Theorem, cf. [CF,

§2.18]) it always contains a fundamental unit when it is infinite. We note that any fundamental unit has the following property:

$$u \notin \mu(k)_p (k^\times)^p \text{ for any prime } p.$$

We are now in a position to give the proof of the main theorem of this section.

Theorem II.3.6. *Let $\mathcal{O} = \mathcal{O}_{k,S}$ be the ring of S-integers in a number field k, and assume that the group of units $\mathcal{O}^\times$ is infinite. Then every matrix in $\mathrm{SL}_2(\mathcal{O})$ is a product of at most 9 elementary matrices.*

Proof. We return to the notations of Subsection II.2.3: we let K denote the Hilbert S-class field of k, let $\tilde{K}$ be its normal closure over $\mathbb{Q}$, and pick an integer $\lambda \geq 1$ which is divisible by μ and for which $\tilde{K} \cap \mathbb{Q}^{\mathrm{ab}} \subseteq \mathbb{Q}(\zeta_\lambda)$. Furthermore, since $\mathcal{O}^\times$ is infinite by assumption, we can find a fundamental unit $u \in \mathcal{O}^\times$. By Lemma II.3.1(3), it suffices to show that for any $(a, b) \in \mathcal{R}(\mathcal{O})$, we have

$$(a, b) \overset{8}{\Rightarrow} (1, 0). \tag{27}$$

First, applying Lemma II.3.3 with $T = (S \setminus V_\infty^k) \cup V(\mu)$ and $n = \mu$, we see that there exist $\alpha \in \mathcal{O}_k$ and $r \in \mathcal{O}^\times$ such that

$$V(\alpha) \cap (S \cup V(\mu)) = \emptyset \text{ and } (a, b) \overset{1}{\Rightarrow} (\alpha r^\mu, b).$$

Next, applying Lemma II.3.4 to the last pair, we find $a' \in \mathcal{O}$ and a $\mathbb{Q}$-split principal prime ideal $\mathfrak{q} = q\mathcal{O}$ such that $v_\mathfrak{q} \notin S \cup V(\lambda) \cup V(\phi(a'\mathcal{O}))$ and $(\alpha r^\mu, b) \overset{3}{\Rightarrow} (a', q^{\mu m})$ for any $m \equiv 1 (\mathrm{mod}\, \phi(a'\mathcal{O})$. Then

$$(a, b) \overset{4}{\Rightarrow} (a', q^{\mu m}) \text{ for any } m \equiv 1 (\mathrm{mod}\, \phi(a'\mathcal{O})). \tag{28}$$

To proceed with the argument we will now specify m. We let P and Q denote the sets of prime divisors of λ/μ and $\phi(a'\mathcal{O})$, respectively, and define λ' and μ' as in Subsection II.2.3; we note that by construction $\mathfrak{q}$ is relatively prime to λ'. So, we can apply Theorem II.2.7 which yields a $\mathbb{Q}$-split principal prime ideal $\mathfrak{p} = \pi\mathcal{O}$ so that $v_\mathfrak{p} \notin V(\phi(a'\mathcal{O}))$ and conditions (1) - (3) are satisfied. Let $s = \ell_\mathfrak{p}(u)$ be the $\mathfrak{p}$-level of u. Condition (3) implies that

$$1 = \gcd(\phi(\mathfrak{p})/\lambda, \lambda'/\lambda) = \gcd(\phi(\mathfrak{p})/\lambda, \phi(a'\mathcal{O})),$$

the last equality holding because λ'/λ is the product of all prime divisors of $\phi(a'\mathcal{O})$. It follows that the numbers $\phi(\mathfrak{p}^s)/\lambda$ and $\phi(a'\mathcal{O})$ are relatively prime, and therefore one can pick a positive integer m so that

$$m \equiv 0 (\mathrm{mod}\, \phi(\mathfrak{p}^s)/\lambda) \text{ and } m \equiv 1 (\mathrm{mod}\, \phi(a'\mathcal{O})).$$

Fix this m for the rest of the proof.

Condition (2) of Theorem II.2.7 enables us to apply Corollary II.3.2 with $z = \pi$ and $t_0 = s$ to find $t \geq s$ so that $(a', q^{\mu m}) \overset{1}{\Rightarrow} (\pi^t, q^{\mu m})$. Since P consists of all prime divisors of λ/μ, condition (1) of Theorem II.2.7 implies that λ/μ divides the order of $u (\mathrm{mod}\, \mathfrak{p})$. Now, applying Lemma II.3.5 with $\delta = \mu$ and $b = q^\mu$, we see that $(\pi^t, q^{\mu m}) \overset{1}{\Rightarrow} (\pi^t, u^{n_t})$ for some integer n_t. Finally, since u is a unit, we have $(\pi^t, u^{n_t}) \overset{2}{\Rightarrow} (1, 0)$. Combining these computations with (28), we obtain (27), completing the proof. □

Corollary II.3.7. *Assume that the group $\mathcal{O}^\times$ is infinite. Then for $n \geq 3$, any matrix $A \in \mathrm{SL}_n(\mathcal{O})$ is a product of $\leq \frac{1}{2}(3n^2 - n) + 4$ elementary matrices.*

Proof. Suppose we are given $A \in \mathrm{SL}_n(\mathcal{O})$. The case $n = 2$ is a direct consequence of Theorem II.3.6. The case of general n will follow by induction, if we can show that $3n - 2$ operations can reduce A to a matrix of the form

$$\left(\begin{array}{c|c} \ast & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline 0 \cdots 0 & 1 \end{array}\right)$$

These $3n - 2$ elementary operations consist of the following:

• 1 **elementary operation to ensure that the first** $n-1$ **entries of the last row are coprime.** This elementary operation will be of the form $B = A\mathrm{E}_{n1}(r)$, where r is as follows. Let $\mathfrak{a}$ be the ideal generated by $a_{n2}, \dots, a_{n,n-1}$. Then the $r \in \mathcal{O}$ can be any element satisfying $V(r) \cap V(a_{n1}) = \emptyset$ and $V(\mathfrak{a}) \setminus V(a_{n1}) \subseteq V(r)$. To show that the first $n-1$ entries of the last row of B are coprime, it suffices to show $V(a_{n1} + ra_{nn}) = V((B)_{n1})$ does not intersect $V(\mathfrak{a})$.

Now, for a fixed $v \in V(\mathfrak{a})$, if $v \notin V(a_{n1})$, then by construction $v \in V(r)$, which implies that $v \notin V(a_{n1} + ra_{nn})$. On the other hand, since $A \in \mathrm{SL}_n(\mathcal{O})$, we know that

$$\big(\, V(a_{n1}) \cap V(\mathfrak{a})\,\big) \cap V(a_{nn}) = \bigcap_{i=1}^{n} V(a_{ni}) = \emptyset.$$

Therefore, if we are given $v \in V(a_{n1}) \cap V(\mathfrak{a})$, then $v \notin V(a_{nn})$; by construction, $v \notin V(r)$, and so the fact that $v \in V(a_{n1})$ implies that $v \notin V(a_{n1}+ra_{nn})$. Thus, in either case we have $v \notin V((B)_{n1})$, and therefore $V(\mathfrak{a}) \cap V((B)_{n1}) = \emptyset$, as desired.

• $n-1$ **elementary operations to ensure that the last entry is** 1. Let $b_{i,j}$ denote $(B)_{ij}$. Since the above construction implies that $b_{n,1}, \dots, b_{n,n-1}$ are coprime, we can find $\lambda_1, \dots, \lambda_{n-1} \in \mathcal{O}$ so that

$$\lambda_1 b_{n,1} + \ldots + \lambda_{n-1} b_{n,n-1} = 1 - b_{n,n}.$$

This means $C := B \cdot \prod_{i=1}^{n-1} \mathrm{E}_{in}(\lambda_i)$ has the (n,n)-th entry equal to $b_{n,n} + (1 - b_{n,n}) = 1$.

• $n-1$ **elementary operations to ensure that all off-diagonal entries of the last row are** 0**'s.** With the above notation, $D := C \cdot \prod_{i=1}^{n-1}(-b_{i,n})$ will satisfy this condition.

• $n-1$ **elementary operations to ensure that all off-diagonal entries of the last column are** 0's. Let $d_{i,j} = (D)_{ij}$. Then $E : \prod_{i=1}^{n-1}(-d_{i,n}) \cdot D$ will satisfy this condition.

Then clearly, the matrix E is of the desired form, and $3n - 2$ elementary operations were used to transform A into E. □

Proof of Corollary I.3.6. Let

$$e_+ : \alpha \mapsto \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad e_- : \alpha \mapsto \begin{pmatrix} 1 & 0 \\ \alpha & 1 \end{pmatrix}$$

be the standard 1-parameter subgroups. Set $U^{\pm} = e_{\pm}(\mathcal{O})$. In view of Theorem II.3.6, it is enough to show that each of the subgroups U^+ and U^- is contained in a product of finitely many cyclic subgroups of $\mathrm{SL}_2(\mathcal{O})$. Let h_k be the class number of k. Then there exists $t \in \mathcal{O}^{\times}$ such that $v(t) = h_k$ for all $v \in S \setminus V_{\infty}^{k}$ and $v(t) = 0$ for all $v \in V^k \setminus S$. Then $\mathcal{O} = \mathcal{O}_k[1/t]$. So, letting $U_0^{\pm} = e_{\pm}(\mathcal{O}_k)$ and $h = \begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix}$, we will have the inclusion

$$U^{\pm} \subset \langle h \rangle \, U_0^{\pm} \, \langle h \rangle.$$

On the other hand, if $w_1, \dots, w_n$ (where $n = [k : \mathbb{Q}]$) is a $\mathbb{Z}$-basis of $\mathcal{O}_k$ then

$$U_0^{\pm} = \langle e_{\pm}(w_1)\rangle \cdots \langle e_{\pm}(w_n)\rangle,$$

hence

$$U^{\pm} \subset \langle h\rangle\langle e_{\pm}(w_1)\rangle \cdots \langle e_{\pm}(w_n)\rangle\langle h\rangle, \tag{29}$$

as required. (Quantitatively, it follows from the proof of Theorem II.3.6 that $\mathrm{SL}_2(\mathcal{O}) = U^-U^+ \cdots U^-$ (nine factors), so since the right-hand side of (23) involves $n+2$ cyclic subgroups, with $\langle h\rangle$ at both ends, we obtain that $\mathrm{SL}_2(\mathcal{O})$ is a product of $9[k : \mathbb{Q}] + 10$ cyclic subgroups.) □

II.4. Example

For a ring of S-integers $\mathcal{O}$ in a number field k such that the group of units $\mathcal{O}^{\times}$ is infinite, we let $\nu(\mathcal{O})$ denote the smallest positive integer with the property that every matrix in $\mathrm{SL}_2(\mathcal{O})$ is a product of $\leq \nu(\mathcal{O})$ elementary matrices. So, the result of [Vs] implies that $\nu(\mathbb{Z}[1/p]) \leq 5$ for any prime p, and our Theorem II.3.6 yields that $\nu(\mathcal{O}) \leq 9$ for any $\mathcal{O}$ as above. It may be of some interest to determine the exact value of $\nu(\mathcal{O})$ in some situations. In Example 2.1 on p. 289, Vsemirnov claims that the matrix

$$M = \begin{pmatrix} 5 & 12 \\ 12 & 29 \end{pmatrix}$$

is not a product of four elementary matrices in $\mathrm{SL}_2(\mathbb{Z}[1/p])$ for any $p \equiv 1(\operatorname{mod} 29)$, and therefore $\nu(\mathbb{Z}[1/p]) = 5$ in this case. However this example is faulty because for any prime p, in $\mathrm{SL}_2(\mathbb{Z}[1/p])$ we have

$$M = \begin{pmatrix} 5 & 12 \\ 12 & 29 \end{pmatrix} = \left(\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \right)^2$$

However, it turns out that the assertion that $\nu(\mathbb{Z}[1/p]) = 5$ is valid not only for $p \equiv 1(\operatorname{mod} 29)$ but in fact for all $p > 7$. More precisely, we have the following.

Proposition II.4.1. *Let $\mathcal{O} = \mathbb{Z}[1/p]$, where p is prime > 7. Then not every matrix in $\mathrm{SL}_2(\mathcal{O})$ is a product of four elementary matrices.*

In the remainder of this section, unless stated otherwise, we will work with congruences over the ring $\mathcal{O}$ rather than $\mathbb{Z}$, so the notation $a \equiv b(\operatorname{mod} n)$ means that elements $a, b \in \mathcal{O}$ are congruent modulo the ideal $n\mathcal{O}$. We begin the proof of the proposition with the following lemma.

Lemma II.4.2. *Let $\mathcal{O} = \mathbb{Z}[1/p]$, where p is any prime, and let r be a positive integer satisfying $p \equiv 1(\operatorname{mod}\ r)$. Then any matrix $A \in \mathrm{SL}_2(\mathcal{O})$ of the form*

$$A = \begin{pmatrix} 1 - p^{\alpha} & * \\ * & 1 - p^{\beta} \end{pmatrix}, \ \alpha, \beta \in \mathbb{Z} \tag{30}$$

which is a product of four elementary matrices, satisfies the congruence

$$A \equiv \pm \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} (\operatorname{mod} r).$$

Proof. We note right away that the required congruence is obvious for the diagonal entries, so we only need to establish it for the off-diagonal ones. Since A is a product of four elementary matrices, it admits one of the following presentations:

$$A = \mathrm{E}_{12}(a)\mathrm{E}_{21}(b)\mathrm{E}_{12}(c)\mathrm{E}_{21}(d), \tag{31}$$

or

$$A = \mathrm{E}_{21}(a)\mathrm{E}_{12}(b)\mathrm{E}_{21}(c)\mathrm{E}_{12}(d), \tag{32}$$

with $a, b, c, d \in \mathcal{O}$.

First, suppose we have (31). Then reducing it modulo $b\mathcal{O}$, we obtain the following congruence:

$$A \equiv \mathrm{E}_{12}(a+c)\mathrm{E}_{21}(d) = \begin{pmatrix} 1+d(a+c) & a+c \\ d & 1 \end{pmatrix} (\operatorname{mod} b).$$

Looking back at (30) and comparing the $(2,2)$ entries, we obtain that $1 - p^{\beta} \equiv 1 (\operatorname{mod} b)$. It follows that $b \in \mathcal{O}^{\times}$, i.e. $b = \pm p^{\gamma}$ for some integer γ. Similarly, one shows that $c = \pm p^{\delta}$ for some integer δ.

Next, we observe that the signs involved in b and c must be different. Indeed, otherwise we would have

$$A = \mathrm{E}_{12}(a)\mathrm{E}_{21}(\pm p^{\gamma})\mathrm{E}_{12}(\pm p^{\delta})\mathrm{E}_{21}(d) = \begin{pmatrix} * & * \\ * & 1+p^{\gamma+\delta} \end{pmatrix},$$

which is inconsistent with (30). Thus, A looks as follows:

$$A = \mathrm{E}_{12}(a)\mathrm{E}_{21}(\pm p^{\gamma})\mathrm{E}_{12}(\mp p^{\delta})\mathrm{E}_{21}(d) = \begin{pmatrix} * & a(1-p^{\gamma+\delta}) \mp p^{\delta} \\ d(1-p^{\gamma+\delta}) \pm p^{\delta} & * \end{pmatrix}.$$

Consequently, the required congruences for the off-diagonal entries immediately follow from the fact that $p \equiv 1 (\operatorname{mod} r)$, proving the lemma in this case.

Now, suppose we have (32). Then

$$A^{-1} = \mathrm{E}_{12}(-d)\mathrm{E}_{21}(-c)\mathrm{E}_{12}(-b)\mathrm{E}_{21}(-a),$$

which means that A^{-1} has a presentation of the form (31). Since the required congruence in this case has already been established, we conclude that

$$A^{-1} \equiv \pm \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} (\operatorname{mod} r).$$

But then we have

$$A \equiv \pm \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} (\operatorname{mod} r),$$

as required. □

To prove the proposition, we will consider two cases.

Case 1. *$p-2$ is composite.* Write $p - 2 = r_1 \cdot r_2$, where r_1, r_2 are positive integers > 1, and set $r = p - 1$. Then

$$r_i \not\equiv \pm 1 (\operatorname{mod} r) \quad \text{for} \quad i = 1, 2. \tag{33}$$

Indeed, we can assume that $r_2 \leq \sqrt{p-2}$. If $r_2 \equiv \pm 1 (\operatorname{mod} r)$ then $r_2 \mp 1$ would be a nonzero integral multiple of r, and therefore $r \leq r_2 + 1$, hence $p - 2 \leq \sqrt{p-2}$. But this is impossible since $p > 3$. Thus, $r_2 \not\equiv \pm 1 (\operatorname{mod} r)$. Since $r_1 \cdot r_2 \equiv -1 (\operatorname{mod} r)$, condition (33) follows.

Now, consider the matrix

$$A = \begin{pmatrix} 1-p & r_1 \cdot p \\ r_2 & 1-p \end{pmatrix}$$

One immediately checks that $A \in \mathrm{SL}_2(\mathcal{O})$. At the same time, A is of the form (30). Then Lemma II.4.2 in conjunction with (33) implies that A is not a product of four elementary matrices.

Case 2. *p and $p-2$ are both primes.* In this paragraph we will use congruences in $\mathbb{Z}$. Clearly, a prime > 3 can only be congruent to $\pm 1 (\bmod 6)$. Since $p > 5$ and $p-2$ is also prime, in our situation we must have $p \equiv 1 (\bmod 6)$. Furthermore, since $p > 7$, the congruence $p \equiv 0$ or $2 (\bmod 5)$ is impossible. Thus, in the case at hand we have

$$p \equiv 1, 13, \text{ or } 19 (\bmod 30).$$

If $p \equiv 13 (\bmod 30)$ in $\mathbb{Z}$, then $p^3 \equiv 7$ (mod 30), and therefore $p^3 - 2$ is an integral multiple of 5. Set $r = p - 1$ and $s = (p^3 - 2)/5$, and consider the matrix

$$A = \begin{pmatrix} 1 - p^3 & 5p^3 \\ s & 1 - p^3 \end{pmatrix}$$

Then A is a matrix in $\mathrm{SL}_2(\mathcal{O})$ having form (30). Note that $5p^3 \equiv 5 (\bmod r)$, which is different from $\pm 1 (\bmod r)$ since $r > 6$. Now, it follows from Lemma II.4.2 that A is not a product of four elementary matrices.

It remains to consider the case where $p \equiv 1$ or $19 (\bmod 30)$. Consider the following matrix:

$$A = \begin{pmatrix} 900 & 53 \cdot 899 \\ 17 & 900 \end{pmatrix},$$

and note that $A \in \mathrm{SL}_2(\mathbb{Z})$ and

$$A^{-1} = \begin{pmatrix} 900 & -53 \cdot 899 \\ -17 & 900 \end{pmatrix}.$$

It suffices to show that neither A nor A^{-1} can be written in the form

$$\mathrm{E}_{12}(a)\mathrm{E}_{21}(b)\mathrm{E}_{12}(c)\mathrm{E}_{21}(d) = \begin{pmatrix} * & c + a(1 + bc) \\ b + d(1 + bc) & (1 + bc) \end{pmatrix}, \text{ with } a, b, c, d \in \mathcal{O} \tag{34}$$

Assume that either A or A^{-1} is written in the form (34). Then $1 + bc = 900$, so

$$b, c \in \{\pm p^n, \pm 29p^n, \pm 31p^n, \pm 899p^n \mid n \in \mathbb{N}\}.$$

On the other hand, setting $t = b + d(1+bc)$ and $u = c + a(1+bc)$, we have the following congruences

$$t \equiv b (\bmod 30) \quad \text{and} \quad u \equiv c (\bmod 30).$$

Analyzing the above list of possibilities for b and c, we conclude that both $t, u \equiv \pm p^n (\bmod 30)$ for some integer n. Thus, if $p \equiv 1 (\bmod 30)$ then $t, u \equiv \pm 1 (\bmod 30)$, and if $p \equiv 19 (\bmod 30)$ then $t, u \equiv \pm 1, \pm 19 (\bmod 30)$. Since $17 \not\equiv \pm 1, \pm 19 (\bmod 30)$, we obtain a contradiction in either case. (We observe that the argument in this last case is inspired by Vsemirnov's argument in his Example 2.1.)

CHAPTER III

Boundedness results for special linear groups over commutative rings

In this chapter, we present two results, which establish bounded width of certain subgroups of special linear groups $\mathrm{SL}_n(R)$ over an arbitrary commutative ring R, for $n \geq 3$. In §III.1 we consider the subgroups generated by elementary matrices with entries in a finite index ideal of the ring and generalize in this situation the result of Tavgen [**T**], which was originally established over the rings of algebraic S-integers – see Theorem III.1.13. In §III.2 we focus our attention on the normal subgroups of $\mathrm{SL}_n(R)$ generated by a given non-central matrix $A \in \mathrm{SL}_n(R)$, when R is a ring satisfying Bass's stable range condition $\mathtt{SR}_{n-1}$, and $\mathrm{SL}_n(R)$ has finite width with respect to the set of elementary matrices. We will show that if the ideal generated by the off-diagonal entries of A has finite index in R, then the normal subgroup generated by A has finite width with respect to the union of the conjugacy classes of A and A^{-1} – see Theorem III.2.1.

III.1. Subgroups of the special linear group defined by a finite-index ideal.

Let R be a (commutative) Noetherian ring. We say that an ideal J of R has *finite index* in R if the quotient (ring) R/J is finite. The goal of this section is to show that the fact that $\mathrm{E}_n(R)$ $(n \geq 3)$ has finite width with respect to $\mathcal{E}_n(R)$ implies that for all ideals $J \subset R$ of finite index, the groups $\mathrm{F}_n(R,J)$ and $\mathrm{E}_n(R,J)$ have finite width with respect to their natural generating sets $\mathcal{F}_n(R,J)$ and $\mathcal{E}_n(R,J)$, respectively. For any ideal $J \subseteq R$ of finite index, let us fix a set $S_J \subset R$ of coset representatives of R/J, and define

$$\mathcal{R}_J = \bigcup_{1 \leq i \neq j \leq n} \mathrm{E}_{ij}(S_J)$$

Since S_J is a finite set, so is $\mathcal{R}_J$. We start with a preliminary result:

Lemma III.1.1. *Let R be a Noetherian ring, and $J \subseteq R$ be an ideal of finite index. Then for all positive integers m, the m-th power J^m is also of finite index.*

Proof. We argue by induction on m. For $m = 1$, the result is true by assumption. So suppose the result is true for all $m \leq k$; so in particular, R/J^k is finite. Let $m = k+1$. Since R is Noetherian, then the ideal J^k is also finitely generated; hence J^k/J^{k+1} is a finitely generated module over the finite ring R/J, and therefore finite. Since

$$(R/J^{k+1})/(J^k/J^{k+1}) \cong R/J^k$$

is finite, this implies R/J^{k+1} is also finite. □

Lemma III.1.2. *If $\mathrm{E}_n(R)$ has finite width with respect to $\mathcal{E}_n(R)$, then for any ideal $J \subset R$ of finite index, $[\mathrm{E}_n(R) : \mathrm{E}_n(R,J)] < \infty$.*

Proof. Let w be the width of $\mathrm{E}_n(R)$ with respect to the set of elementary matrices $\mathcal{E}_n(R)$; then

$$\mathrm{E}_n(R) = \mathcal{E}_n(R)^w = (\mathrm{E}_n(R,J)\cdot\mathcal{R}_J)^w.$$

It follows that the size of the quotient group $\mathrm{E}_n(R)/\mathrm{E}_n(R,J)$ is at most that of the finite set $\mathcal{R}_J^w$. □

Proposition III.1.3. *If $\mathrm{E}_n(R)$ has finite width with respect to $\mathcal{E}_n(R)$, then for any ideal $J \subset R$ of finite index, $\mathrm{E}_n(R,J)$ has finite width with respect to $\mathcal{E}_n(R,J)$.*

Before proving Proposition III.1.3, we would like to introduce a result which will be used in the proof.

Lemma III.1.4. *Let m be an integer, and let $A \in \mathcal{E}_n(R)^m$. For any ideal $J \subset R$ of finite index, we can rewrite A in the form*

$$A = K\cdot L, \quad \text{with} \quad K \in \mathcal{R}_J^m, \quad L \in \mathcal{E}_n(R,J)^m, \tag{35}$$

Proof. We prove the lemma by induction on m. When $m = 1$, then $A = \mathrm{E}_{ij}(\alpha)$ for some $\alpha \in R$ and integers i, j. Then (35) is just the standard decomposition

$$\mathrm{E}_{ij}(\alpha) = \mathrm{E}_{ij}(\beta)\cdot\mathrm{E}_{ij}(\gamma), \quad \text{with } \beta \in S_J,\ \gamma \in J, \text{ and } \beta+\gamma = \alpha.$$

Next, suppose $m > 1$, and the lemma holds for all matrices in $\mathcal{E}_n(R)^{m-1}$. Suppose we are given an arbitrary $A \in \mathcal{E}_n(R)^m$. Write $A = A_1\cdot A_2$, with $A_1 \in \mathcal{E}_n(R)^{m-1}$, and $A_2 = \mathrm{E}_{ij}(\alpha) \in \mathcal{E}_n(R)$. Then we can write and write $\alpha = \beta+\gamma$, with $\beta \in S_J$ and $\gamma \in J$. By the induction hypothesis, we know

$$A_1 = \tilde{K}\cdot\tilde{L} \text{ with } \tilde{K} \in \mathcal{R}_J^{m-1} \text{ and } \tilde{L} \in \mathcal{E}_n(R,J)^{m-1}.$$

Then

$$\begin{aligned} A = A_1\cdot\mathrm{E}_{ij}(\beta+\gamma) &= \tilde{K}\cdot\tilde{L}\cdot\mathrm{E}_{ij}(\beta)\cdot\mathrm{E}_{ij}(\gamma) \\ &= \tilde{K}\cdot\mathrm{E}_{ij}(\beta)\cdot(\mathrm{E}_{ij}(\beta)^{-1}\tilde{L}\,\mathrm{E}_{ij}(\beta))\cdot\mathrm{E}_{ij}(\gamma). \end{aligned}$$

Clearly, $\tilde{K}\cdot\mathrm{E}_{ij}(\beta) \in \mathcal{R}_J^m$, and

$$(\mathrm{E}_{ij}(\beta)^{-1}\tilde{L}\mathrm{E}_{ij}(\beta))\cdot\mathrm{E}_{ij}(\gamma) \in \mathcal{E}_n(R,J)^m,$$

(since the first factor remains in $\mathcal{E}_n(R,J)^{m-1}$), so our assertion follows. □

Proof of Proposition III.1.3. Let w be the width of $\mathrm{E}_n(R)$ with respect to the elementary matrices; then $\mathrm{E}_n(R) = \mathcal{E}_n(R)^w$. Now, since $\mathcal{R}_J$ is a finite set, then there exists an integer M so that every element of $\mathcal{R}_J^w \cap \mathrm{E}_n(R,J)$ is in $\mathcal{E}_n(R,J)^M$. Then for an arbitrary element $A \in \mathrm{E}_n(R,J)$, applying Lemma III.1.4 with $m = w$ we obtain that A can be written in the form (35). Observe that A and L are in $\mathrm{E}_n(R,J)$, hence also K is in $\mathrm{E}_n(R,J)$; since also $K \in \mathcal{R}_J^w$, then $K \in \mathcal{E}_n(R,J)^M$. This implies that the arbitrarily chosen A is in $\mathcal{E}_n(R,J)^{w+M}$; hence, the width of $\mathrm{E}_n(R,J)$ with respect to $\mathcal{E}_n(R,J)$ is finite. □

To put things into perspective, a result which is to be published implies that if k is an algebraic number field, S is a finite set of places of k containing the infinite ones, and $R = \mathcal{O}_{k,S}$ is the ring of S-integers of k, where the group of units $R^\times$ is infinite, then for any ideal $J \subset R$, the width of $\mathrm{E}_2(R,J)$ with respect to $\mathcal{E}_2(R,J)$ is 6. Using this bound, it is easy to obtain a bound on the width of $\mathrm{E}_n(R,J)$ with respect to $\mathcal{E}_n(R,J)$, via a technique similar to that used to prove Corollary II.3.7. However, for some applications, the set $\mathcal{E}_n(R,J)$ may be difficult to identify and utilize. We would like to use this bound to show that $\mathcal{E}_n(R,J^2)$ is contained in a product of finitely many copies of the (very explicit) set $\mathcal{F}_n(R, J)$. Before stating the theorem that provides a bound on this width, we would like to introduce some notation.

For the remainder of this section, we will assume that $n \geq 3$. For fixed integers $1 \leq i \neq j \leq n$, let $\mathcal{S}_{ij}(J) := \mathrm{E}_{ij}(J) \cup \mathrm{E}_{ji}(J)$, and let $\mathrm{S}_{ij}(J)$ be the group generated by $\mathcal{S}_{ij}(J)$. Let $\mathrm{T}_{ij}(R, J)$ be the set

$$\{BAB^{-1} \mid A \in \mathrm{E}_{ij}(J) \text{ and } B \in \mathrm{S}_{ij}(R)\},$$

$\mathcal{T}_{ij}(R, J) = \mathrm{T}_{ij}(R, J) \cup \mathrm{T}_{ji}(R, J)$, and $\mathrm{T}_n(R, J) = \bigcup_{1 \leq i < j \leq n} \mathcal{T}_{ij}(R, J)$.

Theorem III.1.5. *Assume that $n \geq 3$, and J is an ideal of R. Then:*

(a) $\mathrm{E}_n(R, J^2) \subseteq \mathrm{F}_n(R, J)$*;*

(b) *If $\mathrm{E}_n(R)$ has finite width w with respect to $\mathcal{E}_n(R)$ and J is generated by r elements, then*

$$\mathcal{E}_n(R, J^2) \subseteq \mathcal{F}_n(R, J)^d, \quad \text{where} \quad d = 4r \cdot 6^w.$$

Here is an immediate corollary.

Theorem III.1.6. *Let R be a Noetherian ring, and let $n \geq 3$. Assume that $\mathrm{E}_n(R)$ has finite width with respect to $\mathcal{E}_n(R)$. Then for any ideal $J \subset R$ of finite index, we have:*

(a) $\mathrm{E}_n(R, J)$ *has finite width with respect to $\mathcal{E}_n(R, J)$;*

(b) $\mathrm{E}_n(R, J^2)$ *is contained in a product of finitely many copies of $\mathcal{F}_n(R, J)$;*

(c) $[\mathrm{E}_n(R) : \mathrm{F}_n(R, J)] < \infty$.

Proof. Part (a) has already been established in Proposition III.1.3. To prove part (b), we note that since J is of finite index in R, Lemma III.1.1 implies that J^2 is also of finite index in R. So, part (a) implies that the width of $\mathrm{E}_n(R, J^2)$ with respect to $\mathcal{E}_n(R, J^2)$ is finite. Now according to Theorem III.1.5(b), $\mathcal{E}_n(R, J^2)$ is contained in the product $\mathcal{F}_n(R, J)^d$ for some integer $d > 0$. Combining these two facts, we obtain that $\mathrm{E}_n(R, J^2)$ is contained in the product of finitely many copies of $\mathcal{F}_n(R, J)$, implying (b).

Finally, since J^2 is of finite index in R, then Lemma III.1.2 implies that $[\mathrm{E}_n(R) : \mathrm{E}_n(R, J^2)] < \infty$. Combining this with the inclusion $\mathrm{E}_n(R, J^2) \subseteq \mathrm{F}_n(R, J)$ in Theorem III.1.5(a), we obtain (c). □

Before proving Theorem III.1.5, we first give a preliminary result.

Lemma III.1.7. *Let J be an ideal of R, and suppose that $A \in \mathrm{E}_{ij}(J)$, $C \in \mathcal{S}_{kl}(R)^m$, and $\{i, j\} \neq \{k, l\}$. Then*

$$CAC^{-1} \in \mathcal{F}_n(R, J)^{2^m}.$$

Proof. We argue by induction on m. When $m = 1$, then we have $A = \mathrm{E}_{ij}(\gamma)$ for some $\gamma \in J$, and without loss of generality we can assume $C = \mathrm{E}_{kl}(\beta)$ for some $\beta \in R$. We will use the following well-known commutator relations for elementary matrices (cf. [**M**, P. 39]):

$$[\mathrm{E}_{kl}(\beta), \mathrm{E}_{ij}(\gamma)] = \begin{cases} \mathrm{E}_{kj}(\beta\gamma) & \text{if} \quad l = i, k \neq j \\ \mathrm{E}_{il}(-\gamma\beta) & \text{if} \quad k = j, l \neq i \\ I_n & \text{if} \quad l \neq i, k \neq j \end{cases} \tag{36}$$

Using the fact that $CAC^{-1} = [C, A]A$, we obtain

$$CAC^{-1} = \begin{cases} \mathrm{E}_{kj}(\beta\gamma)A & \text{if } \ l = i, k \neq j \\ \mathrm{E}_{il}(-\gamma\beta)A & \text{if } \ k = j, l \neq i \\ A & \text{if } \ l \neq i, j \neq k \end{cases}$$

Clearly, in all cases, $CAC^{-1} \in \mathcal{F}_n(R, J)^2$. This proves the lemma for $m = 1$.

Next, suppose that our assertion is true for all $C' \in \mathcal{S}_{kl}(R)^m$, and prove it for $C \in \mathcal{S}_{kl}(R)^{m+1}$. In order to do this, we will write

$$C = C_1C_2 \text{ with } C_1 \in \mathcal{S}_{kl}(R) \ \text{ and } \ C_2 \in \mathcal{S}_{kl}(R)^m.$$

Then

$$CAC^{-1} = C_1(C_2AC_2^{-1})C_1^{-1}.$$

By the induction hypothesis, $C_2AC_2^{-1} \in \mathcal{F}_n(R, J)^{2^m}$. On the other hand, by the case $m = 1$, we have

$$C_1\mathcal{F}_n(R, J)C_1^{-1} \subseteq \mathcal{F}_n(R, J)^2.$$

This implies that

$$CAC^{-1} \in C_1\mathcal{F}_n(R, J)^{2^m}C_1^{-1} = (C_1\mathcal{F}_n(R, J)C_1^{-1})^{2^m} \subseteq (\mathcal{F}_n(R, J)^2)^{2^m} = \mathcal{F}_n(R, J)^{2^{m+1}},$$

completing the argument. □

Since there is no simple formula for the commutator $[\mathrm{E}_{ij}(\alpha), \mathrm{E}_{ji}(\beta)]$, the proof of Lemma III.1.7 does not apply to the conjugates CAC^{-1} with $A \in \mathrm{E}_{ij}(J)$ and $C \in \mathcal{S}_{kl}(R)^m$ when $\{i, j\} = \{k, l\}$. Nevertheless, we can prove the following partial analog of the lemma in this case.

Lemma III.1.8. *Suppose that $n \geq 3$, $1 \leq i \neq j \leq n$, and J is an ideal of R. Let $A = \mathrm{E}_{ij}(\alpha)$, where $\alpha \in J^2$, and $C \in \mathcal{S}_{ij}(R)^m$; choose an integer r so that α can be expressed as*

$$\alpha = \sum_{i=1}^{r} \beta_i\gamma_i \ \text{ with } \beta_i, \gamma_i \in J. \tag{37}$$

Then $CAC^{-1} \in \mathcal{F}_n(R, J)^{2^m \cdot 4r}$.

Proof. Pick $k \in \{1, \dots, n\}$ different from i and j, and set $T = \mathcal{S}_{ik}(J) \cup \mathcal{S}_{kj}(J)$. Then Lemma III.1.7 applies to each element of T, implying that

$$CDC^{-1} \in \mathcal{F}_n(R, J)^{2^m} \ \text{ for any } \ D \in T.$$

On the other hand, in the above notations, we have

$$A = \mathrm{E}_{ij}(\alpha) = \prod_{i=1}^{r} \mathrm{E}_{ij}(\beta_i\gamma_i) = \prod_{i=1}^{r} [\mathrm{E}_{ik}(\beta_i), \mathrm{E}_{kj}(\gamma_i)] \in T^{4r},$$

which implies that $CAC^{-1} \in \mathcal{F}_n(R, J)^{4r \cdot 2^m}$. □

Corollary III.1.9. *In the notations of Lemma III.1.8, assume that J can be generated by r elements. Then for any $A \in \mathrm{E}_{ij}(J^2)$ and $C \in \mathcal{S}_{ij}(R)^m$, we have $CAC^{-1} \in \mathcal{F}_n(R, J)^d$, where $d = 4r \cdot 2^m$.*

Proof. Since $A \in \mathrm{E}_{ij}(J^2)$, then we can write $A = \mathrm{E}_{ij}(\alpha)$ with $\alpha \in J^2$. Our assertion will follow from Lemma III.1.8 if we show that α has a presentation (37) of length r. But, if J is generated by $\gamma_1, \dots, \gamma_r$, then $J^2 = J\gamma_1 + \ldots + J\gamma_r$, implying the result. □

In order to prove part (a) of Theorem III.1.5, we need to show that $\mathcal{E}_n(R,J^2) \subseteq \mathrm{F}_n(R,J)$. According to Lemma III.1.8, for any $i \neq j$, an arbitrary element of $\mathrm{T}_{ij}(R,J^2)$ can indeed be written as a product of a bounded number of elements from $\mathcal{F}_n(R,J)$. So, to prove Theorem III.1.5, it suffices to show that every element of $\mathcal{E}_n(R,J^2)$ is a bounded product of elements from $\{\mathrm{T}_{ij}(R,J^2) \mid 1 \leq i \neq j \leq n\}$. This is done in the following statement, which we will state for an arbitrary ideal J and later use it for J^2.

Proposition III.1.10. *Let $A \in \mathcal{F}_n(R,J)$ and $B = B_1 \cdots B_m$ with $B_t \in \mathcal{E}_n(R)$ for all t. Then*
$$BAB^{-1} = (C_1A_1C_1^{-1})\cdots(C_{3^m}A_{3^m}C_{3^m}^{-1}) \text{ with } A_t \in \mathrm{E}_{i_tj_j}(J) \text{ and } C_t \in \mathcal{S}_{i_tj_t}(R)^m \text{ for } t = 1,\ldots,3^m,$$
hence belongs to $\mathrm{T}_n(J)^{3^m}$.

Proof. We begin the proof with the following computation.

Proposition III.1.11. *Let $m \geq 1$, and suppose that $A \in \mathcal{T}_{ij}(R,J)^m, B \in \mathrm{E}_{kl}(R)$, and $\{i,j\} \neq \{k,l\}$. Then*
$$[B,A] \in W, \quad \text{where } W = \begin{cases} \{I_n\} & \text{if } \{i,j\} \cap \{k,l\} = \emptyset, \\ \mathrm{E}_{ki}(J)\mathrm{E}_{kj}(J) & \text{if } l \in \{i,j\}, \\ \mathrm{E}_{il}(J)\mathrm{E}_{jl}(J) & \text{if } k \in \{i,j\}. \end{cases}$$

Proof. The first case follows immediately from the commutator formula (36). Furthermore, the second and the third cases are similar, so we will treat only the second case. Our argument is based on the following commutator identity, which is true in any group:

$$[x_1x_2,y] = x_1[x_2,y]x_1^{-1}[x_1,y]. \tag{38}$$

We also observe that it easily follows from (36) that W is a subgroup normalized by $\mathrm{S}_{ij}(R) \supset \mathcal{T}_{ij}(R,J)^m$. We will first consider the case where $m = 1$, so $A \in \mathcal{T}_{ij}(R,J)$. Switching i and j if necessary, we may assume that

$$A = C\tilde{A}C^{-1} \text{ with } A \in \mathrm{E}_{ij}(J) \text{ and } C \in \mathcal{S}_{ij}(R)^\ell \text{ for some } \ell \geq 0.$$

We will prove the inclusion $[B,A] \in W$ by induction on ℓ. If $\ell = 0$, then $[B,A] \in \mathrm{E}_{kj}(J) \subseteq W$ (in fact, $[B,A] = I_n$ if $l = j$) by (36). Suppose $\ell > 0$. We can then write

$$C = C_1C_2 \text{ with } C_1 \in \mathcal{S}_{ij}(R) \text{ and } C_2 \in \mathcal{S}_{ij}(R)^{\ell-1}.$$

We have

$$[B,A] = [B, C\tilde{A}C^{-1}] = C_1[C_1^{-1}BC_1, C_2\tilde{A}C_2^{-1}]C_1^{-1},$$

and it is enough to prove that $X := [C_1^{-1}BC_1, C_2\tilde{A}C_2^{-1}] \in W$. We have

$$C_1^{-1}BC_1 = YB \text{ where } Y = [C_1^{-1}, B] \in \mathrm{E}_{kj}(R) by (36).$$

Then using (38), we obtain

$$X = Y[B, C_2\tilde{A}C_2^{-1}]Y^{-1} \cdot [Y, C_2\tilde{A}C_2^{-1}].$$

By the induction hypothesis, $[B, C_2\tilde{A}C_2^{-1}] \in W$, hence commutes with Y. Similarly, by the induction hypothesis (applied with $l = j$), we have $[Y, C_2\tilde{A}C_2^{-1}] \in W$. Thus, $X \in W$, as required, completing the proof when $m = 1$.

We will now prove the proposition by induction on m. So, let $A \in \mathcal{T}_{ij}(R,J)^m$ with $m > 1$. We write

$$A = A_1 \cdot A_2 \text{ with } A_1 \in \mathcal{T}_{ij}(R,J) \text{ and } A_2 \in \mathcal{T}_{ij}(R,J)^{m-1},$$

noting that $A_1, A_2 \in \mathrm{S}_{ij}(R)$. Using the commutator relation (38), we obtain

$$[B, A] = [A, B]^{-1} = (A_1[A_2, B]A_1^{-1}[A_1, B])^{-1} = [B, A_1]A_1[B, A_2]A_1^{-1}$$

Since $[B, A_1]$ and $[B, A_2] \in W$ by the induction hypothesis, and W is normalized by $\mathrm{S}_{ij}(R)$, we conclude that $[B, A] \in W$, as required. □

Corollary III.1.12. *Let $A = CDC^{-1}$ with $D \in \mathrm{E}_{ij}(J)$ and $C \in \mathcal{S}_{ij}(R)^m$. If we are given $B \in \mathrm{E}_{kl}(R)$, then the conjugate BAB^{-1} belongs to $\mathcal{F}_n(R, J)^2 A$ if $\{i, j\} \neq \{k, l\}$, and is of the form $C'DC'^{-1}$ with $C' \in \mathcal{S}_{ij}(R)^{m+1}$ otherwise.*

Proof. In the first case, it is a consequence of the identity $BAB^{-1} = [B, A] \cdot A$ and Proposition III.1.11. In the second case, $BAB^{-1} = (BC)D(BC)^{-1}$ and $BC \in \mathcal{S}_{ij}(R)^{m+1}$, so the claim follows immediately. □

Proof of Proposition III.1.10. We will argue by induction on m. First, let $m = 1$, so $A \in \mathcal{S}_{ij}(J)$ and $B \in \mathrm{E}_{kl}(R)$. If $\{k, l\} = \{i, j\}$ then BAB^{-1} itself can serve as one factor in the required presentation with the other two factors being trivial. If $\{k, l\} \neq \{i, j\}$ then it follows from the identity $BAB^{-1} = [B, A] \cdot A$ and Proposition III.1.11 that $BAB^{-1} \in \mathcal{F}_n(R, J)^3$, meaning that it has a required presentation with all the C_t's equal to the identity.

Let now $m > 1$, and set $B' = B_2 \cdots B_m$. By the induction hypothesis,

$$B'AB'^{-1} = (C'_1 A'_1 {C'_1}^{-1}) \cdots (C'_{3^{m-1}} A'_{3^{m-1}} {C'}^{-1}_{3^{m-1}})$$

where each $A'_t \in \mathrm{E}_{i_t j_t}(J)$ and $C'_t \in \mathcal{S}_{i_t j_t}(R)^{m-1}$. Then

$$BAB^{-1} = \prod_{t=1}^{3^{m-1}} B_1(C'_t A'_t {C'_t}^{-1})B_1^{-1}.$$

We will now analyze each factor in this product. Suppose $B_1 \in \mathrm{E}_{ij}(R)$. If $\{i, j\} = \{i_t, j_t\}$, then

$$B_1(C'_t A'_t {C'_t}^{-1})B_1^{-1} = (B_1 C'_t)A'_t(B_1 C'_t)^{-1} \quad \text{and} \quad B_1 C'_t \in \mathcal{S}_{ij}(R)^m.$$

Otherwise, according to Corollary III.1.12,

$$B_1(C'_t A'_t {C'_t}^{-1})B_1^{-1} \in \mathcal{F}_n(R, J)^2 (C'_t A'_t {C'_t}^{-1}).$$

In either case, the term is a product of three factors, each of the form CAC^{-1} with $A \in \mathrm{E}_{kl}(J)$ and $C \in \mathcal{S}_{kl}(R)^m$, and the lemma follows.

□

Proof of Theorem III.1.5. (a) Since $\mathrm{E}_n(R, J^2)$ is generated by elements BAB^{-1} for $A \in \mathcal{F}_n(R, J^2)$ and $B \in \mathcal{E}_n(R)^m$, it suffices to show each such element is in $\mathrm{F}_n(R, J)$. By Proposition III.1.10, we can write

(39)

$$BAB^{-1} = (C_1 A_1 C_1^{-1}) \cdots (C_{3^m} A_{3^m} C_{3^m}^{-1}) \text{ with } A_t \in \mathrm{E}_{i_t j_j}(J^2) \text{ and } C_t \in \mathcal{S}_{i_t j_t}(R)^m \text{ for } t = 1, \ldots, 3^m,$$

Then Lemma III.1.8 shows that each $(C_t A_t C_t^{-1}) \in \mathrm{F}_n(R, J)$. This proves that $BAB^{-1} \in \mathrm{F}_n(R, J)$, as required.

(b) Suppose we are given an element of $\mathcal{E}_n(R, J^2)$; by definition, it is of the form BAB^{-1} for some $A \in \mathcal{F}_n(R, J^2)$ and $B \in \mathrm{E}_n(R)$. Since we assume that $\mathrm{E}_n(R) = \mathcal{E}_n(R)^w$, then we can use Proposition III.1.10 with $m = w$ to pick a presentation of BAB^{-1} of the form (39). Then it follows from Corollary III.1.9 that each factor $C_tA_tC_t^{-1}$ lies in $\mathcal{F}_n(R,J)^{2^w\cdot 4r}$. Then (39) yields that $BAB^{-1}\in\mathcal{F}_n(R,J)^{6^w\cdot 4r}$, as claimed. □

The following theorem is the main result of this section.

Theorem III.1.13. *Let R be a Noetherian ring, and let $n\geq 3$. If $\mathrm{E}_n(R)$ has bounded width with respect to $\mathcal{E}_n(R)$, then for every ideal $J\subset R$ of finite index, the subgroup $\mathrm{F}_n(R,J)$ has bounded width with respect to $\mathcal{F}_n(R,J)$.*

Proof. According to Theorem III.1.5(a), the subgroup $\mathrm{F}_n(R,J)$ contains $\mathrm{E}_n(R,J^2)$, which is a subgroup of finite index in $\mathrm{E}_n(R)$. By Theorem III.1.6(b), there exists $d\geq 1$ such that $\mathrm{E}_n(R,J^2)\subseteq\mathcal{F}_n(R,J)^d$. On the other hand, considering a finite system of coset representatives of $\mathrm{F}_n(R,J)$ modulo $\mathrm{E}_n(R,J^2)$, we see that there exists $m\geq 1$ such that

$$\mathrm{F}_n(R,J)=\mathrm{E}_n(R,J^2)\mathcal{F}_n(R,J)^m.$$

Then $\mathrm{F}_n(R,J)=\mathcal{F}_n(R,J)^{d+m}$, as required. □

III.2. Finite width with respect to a union of conjugacy classes

Let G be an arbitrary group. For an element $g\in G$, we let C_g denote the union of the conjugacy classes of g and of g^{-1}, and $\mathcal{N}_g$ denote the minimal normal subgroup of G containing g. Then

$$\mathcal{N}_g=\bigcup_{n\geq 0}\mathrm{C}_g^n.$$

Therefore, for every $h\in\mathcal{N}_g$ there exists a minimal number $n_g(h)$ so that $h\in\mathrm{C}_g^{n_g(h)}$. One can ask the question of whether the numbers $n_g(h)$ are uniformly bounded, as h ranges over $\mathcal{N}_g$. An affirmative answer to this question is equivalent to $\mathcal{N}_g$ having bounded width with respect to its natural generating set C_g. Using non-standard models, Morris [**MCKP**] established an affirmative answer for $G=\mathrm{SL}_n(\mathcal{O})$, when $\mathcal{O}$ is a ring of S-integers of a number field, and $n\geq 3$. These considerations also played a role, for example, in [**ALM**].

The goal of this section is to examine the property of having finite width with respect to a conjugacy class for some normal subgroups of the group $G=\mathrm{E}_n(R)$, where $n\geq 3$ and R is a ring satisfying some conditions. Before stating the main result of this section more precisely, we recall the definition of the stable range condition, due to Bass [**B**, P. 14].

Definition. Let $m\geq 1$. A ring R is said to satisfy the *stable range* condition $\mathtt{SR}_m$ if for any $a_1,\dots,a_m\in R$ which are relatively prime (i.e., $a_1R+\dots+a_mR=R$), there exist $\lambda_1,\dots,\lambda_{m-1}$ so that $a_1-\lambda_1a_m,\dots,a_{m-1}-\lambda_{m-1}a_m$ are relatively prime.

Theorem III.2.1. *Suppose R is a commutative Noetherian ring which satisfies the stable range condition $\mathtt{SR}_{n-1}$, and for which $\mathrm{E}_n(R)$ has finite width with respect to $\mathcal{E}_n(R)$. Fix a (non-central) matrix $A\in\mathrm{E}_n(R)$, and let J be the ideal generated by the off-diagonal entries of A. If the index of J in R is finite, then $\mathcal{N}_A$ has finite width with respect to C_A.*

We will prove this theorem using the results of §III.1, and the following statement:

Proposition III.2.8. *Suppose R is a commutative Noetherian ring which satisfies the stable range condition $\mathtt{SR}_{n-1}$, and for which $\mathrm{E}_n(R)$ has finite width with respect to $\mathcal{E}_n(R)$. Fix a (non-central) matrix $A\in\mathrm{E}_n(R)$, and let J be the ideal generated by the off-diagonal entries of A. If the index of J in R is finite, then there exists a constant N, so that C_A^N contains the group $\mathrm{F}_n(R,J)$.*

Proof of Theorem III.2.1 (assuming Proposition III.2.8). Proposition III.2.8 asserts that $\mathrm{F}_n(R,J) \subseteq \mathcal{N}_A$; since $\mathcal{N}_A$ is closed under conjugation, then also $\mathrm{E}_n(R,J) \subseteq \mathcal{N}_A$. But, Lemma III.1.1 asserts that the index $[\mathrm{E}_n(R) : \mathrm{E}_n(R,J)]$ is finite; hence, $[\mathcal{N}_A : \mathrm{E}_n(R,J)]$ is also finite. Let S_A be a set of coset representatives of $\mathcal{N}_A/\mathrm{E}_n(R,J)$; since it is a finite set and each element $s \in S_A$ is contained in $\mathrm{C}_A^{N_s}$ for some integer N_s, then there is an integer $N = \max_{s \in S_A} N_s$ so that each element of S_A is in C_A^N. Also, Proposition III.2.8 provides an integer M so that each element of $\mathrm{F}_n(R,J)$ is in C_A^M; by conjugating we obtain that each element of $\mathrm{E}_n(R,J)$ is in C_A^M as well. Combining these facts, we obtain that $\mathcal{N}_A = \mathrm{C}_A^{M+N}$. □

The remainder of the section is devoted to the proof of Proposition III.2.8. But first, we would like to give a survey of the previous work on this topic.

Let $n \geq 3$ be an integer, and suppose R is a commutative ring. Fix a non-central matrix $A = (a_{ij}) \in \mathrm{E}_n(R)$, and let $J = \langle a_{21}, \dots, a_{n-1\,1} \rangle$ be the ideal generated by the off-diagonal entries of the first column of A. Bass [**B**, Lem. I.2.5] showed that if $a_{n1} = 0$, then $\mathrm{E}_{1n}(J) \subseteq \mathrm{C}_A^8$ (with an analogous result holding when any other off-diagonal a_{ij} is zero). Furthermore, Bass [**B**, Thm. I.4.2(e)] has also showed that if a ring R satisfies the stable range condition $\mathtt{SR}_{n-1}$, then we can find a non-central matrix $B \in \mathrm{C}_A^2$, which has an off-diagonal zero entry in the first column. If we let J' be the ideal generated by $\{(B)_{i,1} \mid i > 1\}$, then Bass's argument implies that $\mathrm{E}_{1n}(J') \subseteq \mathrm{C}_A^{16}$. However, in general the ideals J' and J are different (more precisely, $J' = r \cdot J$ for some $r \in R$, and so $[J : J'] = \infty$); therefore, the fact that $\mathrm{E}_{1n}(J') \subseteq \mathrm{C}_A^{16}$ does not imply that $\mathrm{E}_{1n}(J) \subseteq \mathrm{C}_A^M$ for any fixed integer M.

When R is a number ring, then we will always have $[J : J'] < \infty$; in this case, Brenner [**Br**] showed that $\mathrm{E}_{1n}(J) \subseteq \mathrm{C}_A^{16n}$. His argument inherently uses the Euclidean algorithm, and therefore only works for Euclidean domains.

Let $n \geq 3$, and R be an arbitrary (unital) ring. We present a variation of Bass's argument, which proves an analogue of Brenner's result for a more general class of rings R.

Notation. Let us fix $n \geq 3$. For a fixed matrix $A \in \mathrm{M}_n(R)$ and $1 \leq i,j \leq n$, the element a_{ij} is defined as $(A)_{ij}$. For $m = 1, \dots, n$, $\mathcal{C}_m(A)$ is the m-th column of A; similarly, $\mathcal{R}_m(A)$ is defined to be the m-th row of A. Define $\mathcal{R}= (0 \quad \dots \quad 0 \quad 1)$; then $\mathcal{R}(A) = \mathcal{R} \cdot A = \mathcal{R}_n(A)$. Let $\{\overrightarrow{e_1}, \dots, \overrightarrow{e_n}\}$ be the standard basis of $\mathbb{Z}^n$ (as column vectors), and $\{e_{ij} \mid 1 \leq i,j \leq n\}$ be the standard basis of $\mathrm{M}_n(\mathbb{Z})$.

To start, we will assume that R is a ring satisfying the Stable Range condition $\mathtt{SR}_{n-1}$, and repeat Bass's argument for how to obtain a zero entry for some matrix $B \in \mathrm{C}_A^2$. Incidentally, this procedure also makes $(B)_{nn} = 1$, which is necessary for our argument. The next two lemmas are extracted from [**B**].

Lemma III.2.2. *Suppose we are given a matrix $A \in \mathrm{M}_n(R)$, integers $c \neq d$ and m between 1 and n, and $\lambda \in R$.*

(a) *If $d \neq m$ then $\mathcal{C}_m(\mathrm{E}_{cd}(\lambda) A \mathrm{E}_{cd}^{-1}(\lambda)) = \mathcal{C}_m(A) + \lambda a_{dm} \overrightarrow{e_c}$.*

(b) *If $c \neq m$ then $\mathcal{R}_m(\mathrm{E}_{cd}(\lambda) A \mathrm{E}_{cd}^{-1}(\lambda)) = \mathcal{R}(A) - a_{mc} \lambda \overrightarrow{e_d}^t$.*

Proof. When $A = e_{ij}$ for some $1 \leq i,j \leq n$, then this follows from [**B**, Lemma I.2]; for a general matrix A, this follows by writing A as a linear combination of the e_{ij}'s. □

Lemma III.2.3. *Suppose R satisfies the stable range condition $\mathtt{SR}_{n-1}$. If $A \in \mathrm{SL}_n(R)$ then there exists $\Lambda \in \mathrm{SL}_n(R)$ so that, writing $\mathcal{C}_1(\Lambda A \Lambda^{-1}) = (f_1 \ \cdots \ f_n)^t$, we have $f_1 R + \ldots + f_{n-1} R = R$, $f_n = a_{n1}$, and $f_2 R + \ldots + f_n R$ equals $a_{21} R + \ldots + a_{n1} R$.*

Proof. From the definition of stable range, there exist $\lambda_1, \ldots, \lambda_{n-1} \in R$ so that

$$(a_{11} + \lambda_1 a_{n1})R + \cdots + (a_{n-1\,1} + \lambda_{n-1} a_{n1})R = R.$$

Let $\Lambda = \prod_{i=1}^{n-1} \mathrm{E}_{in}(\lambda_i)$. Then by Lemma III.2.2(a), we see that when $i \neq n$, then

$$\big(E_{jn}(\lambda_j) A E_{jn}(\lambda_j)^{-1}\big)_{i1} = (A)_{i1} \quad \text{for any } j \neq i, n,$$

and therefore

$$f_i := (\Lambda A \Lambda^{-1})_{i1} = \big(E_{in}(\lambda_i) A E_{in}(\lambda_i^{-1})\big)_{i1} = a_{i1} + \lambda_i a_{ni} \quad \text{for } i = 1, \ldots, n-1,$$

and $f_n := (\Lambda A \Lambda^{-1})_{n1} = a_{n1}$. Therefore, by the choice of λ_i, we have $f_1 R + \ldots + f_{n-1} R = R$; furthermore, it's clear that $f_2 R + \ldots + f_n R$ contains $a_{21} R + \ldots + a_{n1} R$ because each f_i can be written as a linear combination $f_i = a_{i1} + \lambda_i a_{n1}$ for $i = 2, \ldots, n-1$, and $f_n = a_{n1}$. The reverse inclusion follows for the same reason, since $a_{i1} = f_i - \lambda_i f_n$ for $i = 2, \ldots, n = 1$, and $a_{n1} = f_n$. Therefore, the two ideals are equal. □

Lemma III.2.4. *Assume R satisfies the stable range condition $\mathtt{SR}_{n-1}$, and let $A \in \mathrm{SL}_n(R)$. Then for each solution $(\lambda_1, \ldots, \lambda_{n-1})$, all $\lambda_i \in R$, to the equation*

$$\lambda_1 a_{11} + \ldots + \lambda_{n-1} a_{n-1\,1} = a_{n1},$$

the set of commutators $[\mathrm{C}_A, \mathrm{E}_n(R)] \subset \mathrm{C}_A^2$ contains a matrix B satisfying $\mathcal{R}(B) = (0, \lambda_1, 0, \ldots, 0, 1)$.

Proof. Let $D = \prod_{h=1}^{n-1} E_{nh}(-\lambda_h)$, and define a'_{ij} to be $(A^{-1})_{ij}$. Then, $\mathcal{R}(D) = (-\lambda_1, \cdots, -\lambda_{n-1}, 1)$, and $D \cdot (a_{11} \ \ldots \ a_{n1})^t$ is a column vector, whose last entry is $a_{n1} - \sum_{\ell=1}^{n-1} \lambda_\ell a_{h1}$, which equals 0 by choice of $\{\lambda_i\}$. Now

$$ATA^{-1} = A(I_n + e_{12})A^{-1} = I_n + (\vec{0} \ \ \mathcal{C}_1(A) \ \ \vec{0} \ \ \ldots \ \ \vec{0}) \cdot A^{-1} = I_n + (a_{11} \ \ldots \ a_{n1})^t \cdot (a'_{21} \ \ldots \ a'_{2n}).$$

Then using the above equation, we see that

$$\begin{aligned} \mathcal{R}(DATA^{-1}) &= \mathcal{R}(D) + \mathcal{R}(D \cdot (a_{11} \ \ldots \ a_{n1})^t \cdot (a'_{21} \ \ldots \ a'_{2n})) = \\ &= \mathcal{R}(D) + (0) \cdot (a'_{21} \ \ldots \ a'_{2n}) = \mathcal{R}(D). \end{aligned} \tag{40}$$

Now clearly, the element $B := D[A,T]D^{-1} \in [\mathrm{C}_A, \mathrm{E}_n(R)]$. Let $b_{ij} = (B)_{ij}$. Since $T^{-1} = I_n - e_{12}$, we observe that

$$\mathcal{R}(D[A,T]) = \mathcal{R}(DATA^{-1})T^{-1} \overset{(40)}{=} \mathcal{R}(DT^{-1}) = \mathcal{R}(D) - \mathcal{R}((\vec{0} \ \ \mathcal{C}_1(D) \ \ \vec{0} \ \ \ldots \ \ \vec{0})) = \mathcal{R}(D) + \lambda_1 \vec{e_2}^{\,t}.$$

Since the second row of $D^{-1} = \prod_{i=1}^{n-1} E_{ni}(\lambda_i)$ is $\vec{e_2}^{\,t}$, then $\lambda_1 \vec{e_2}^{\,t} D^{-1} = \lambda_1 \mathcal{R}_2(D^{-1}) = \lambda_1 \vec{e_2}^{\,t}$, and

$$\mathcal{R}(D[A,T]D^{-1}) = \mathcal{R}(DD^{-1}) + \lambda_1 \vec{e_2}^{\,t} D^{-1} = (1 \ \ \lambda_1 \ \ 0 \ \ \ldots \ \ 0).$$

This is the statement of the lemma. □

The next result will make use of writing matrices $F \in \mathrm{SL}_n(R)$ satisfying $\mathcal{R}(F) = (0, \dots, 0, 1)$ as an affine transformation – namely, in the form $F = (B|\vec{v})$, where the notation $(B|\vec{v})$ denotes a block matrix of the form

$$\begin{pmatrix} B & | & \vec{v} \\ \vec{0}^{\,t} & | & 1 \end{pmatrix}$$

where B is an $(n-1) \times (n-1)$-matrix, and $\vec{v}$ is a $(n-1) \times 1$-matrix. It is straightforward to verify that matrices in this form have the following formulae for multiplication and inversion:

$$(B|\vec{v})(C|\vec{w}) = (BC|B\vec{w} + \vec{v})$$

$$(B|\vec{v})^{-1} = (B^{-1}| - B^{-1}\vec{v})$$

Lemma III.2.5. *Suppose $A \in \mathrm{SL}_n(R)$ satisfies $a_{nn} = 1$, and let $\lambda = a_{n1}$. Then the set of commutators $[\mathrm{E}_n(R), A]$ contains a matrix of the form*

$$I_n + \left(\begin{array}{c|r} -\mathcal{R}(A) & \kappa_1 \\ \hline \text{O} & \mathcal{C}_1(A) - \lambda\mathcal{C}_n(A) \\ \hline \lambda\mathcal{R}(A) & \kappa_2 \end{array}\right) \quad \text{with } \kappa_1 \in R,\ \kappa_2 \in \lambda R.$$

Proof. Let $D = \prod_{i=1}^{n-1} E_{ni}(a_{ni}) \in \mathrm{E}_n(R)$; then $\mathcal{R}(AD^{-1}) = (0\ \dots\ 0\ 1)$, so in the above notations, $AD^{-1} = (B|\vec{v})$ for some matrices B and $\vec{v}$.

Set $E = D^{-1}(I_{n-1}|-\vec{e_1})D$, and let us compute the matrix $[E, A] \in [\mathrm{E}_n(R), A]$:

$$\begin{aligned} [E,A] = [D^{-1}(I_{n-1}|-\vec{e_1})D, (B|\vec{v})D] &= \\ \big(D^{-1}(I_n - e_{1n})\cdot D\big)(B|\vec{v})(I_{n-1}|\vec{e_1}) \quad (B|\vec{v})^{-1} &= \\ \big(D^{-1}D + \big(\text{O}|-\mathcal{C}_1(D^{-1})\big)D\ \big)(B|B\vec{e_1} + \vec{v})(B^{-1}| - B^{-1}\vec{v}) &= \\ \big(I_n - \vec{e_1} \cdot \mathcal{R}(D) + \lambda\vec{e_n} \cdot \mathcal{R}(D)\big)(I_{n-1}|B\vec{e_1}). \end{aligned}$$

The last equality holds since $\mathcal{C}_1(D^{-1}) = \vec{e_1} - \lambda\vec{e_n}$. Letting $b_i = (B)_{i1} = a_{i1} - \lambda a_{in}$ and $d_j = (D)_{nj} = a_{nj}$, we see that $[E, A]$ is a matrix of the form

$$[E,A] = I_n + \left(\begin{array}{c|c} -\mathcal{R}(D) & \kappa_1 \\ \hline \text{O} & B\vec{e_1} \\ \hline \lambda\mathcal{R}(D) & \kappa_2 \end{array}\right) \quad \begin{array}{rl} \text{with} & \kappa_1 = b_1(1-d_1) - \left(\sum_{i=2}^{n-1} b_i \cdot d_i\right) - d_n \\ \text{and} & \kappa_2 = \lambda\left(\sum_{i=1}^{n-1} b_i \cdot d_i\right) + \lambda d_n \end{array}$$

The statement of the lemma now follows since $\mathcal{R}(D) = \mathcal{R}(A)$, and $B\vec{e_1}$ is the first column of B, which in the first $n-1$ entries is $\mathcal{C}_1(A) - \lambda\mathcal{C}_n(A)$. □

Lemma III.2.6. *Let R be a commutative ring. Suppose $A \in \mathrm{SL}_n(R)$ satisfies $a_{nn} = 1$ and $a_{n1} = 0$. Then the subset C_A^4 contains a matrix of the form $\mathrm{E}_{1n}(r)$ for any r in the ideal generated by the off-diagonal entries of the first column and the last row of A.*

Proof. By Lemma III.2.5, we know the subset $[\mathrm{E}_n(R), A]$ contains a matrix of the form

$$\Lambda = I_n + \left(\begin{array}{c|r} -\mathcal{R}(A) & \kappa_1 \\ \hline \text{O} & \mathcal{C}_1(A) - \lambda\mathcal{C}_n(A) \\ \hline \lambda\mathcal{R}(A) & \kappa_2 \end{array}\right) \quad \begin{array}{rl} \text{with} & \kappa_1 \in R \text{ and } \kappa_2 \in \lambda R, \\ \text{where} & (B)_{i1} = a_{i1} - \lambda a_{in} \text{ and } \lambda = a_{n1}. \end{array}$$

Now by assumption, $\lambda = a_{n1} = 0$. In this case, the last row of Λ is just $(0 \quad \dots \quad 0 \quad 1)$, and $\mathcal{C}_1(A) - \lambda\mathcal{C}_n(A) = \mathcal{C}_1(A)$; furthermore, $(\Lambda)_{11} = 1 - \lambda = 1$. Thus, Λ is of the form

$$I_n + \left(\begin{array}{c|c|r} 0 & -\mathcal{R}(A) & \kappa_1 \\ \hline \vec{0} & \text{O} & \mathcal{C}_1(A) \\ \hline 0 & 0 \ \cdots \ 0 & 0 \end{array}\right) \quad \text{with } \kappa_1 \in R. \tag{41}$$

Then Lemma III.2.2 implies that for any $\lambda \in R$, the following conjugation identities hold for $i = 2, \dots, n-1$:

$$\mathrm{E}_{in}(\lambda)^{-1} \cdot \Lambda \cdot \mathrm{E}_{in}(\lambda) = \Lambda + a_{ni}\lambda e_{1n}$$
$$\mathrm{E}_{1i}(\lambda) \cdot \Lambda \cdot \mathrm{E}_{1i}(\lambda)^{-1} = \Lambda + \lambda a_{i1} e_{1n}.$$

Moreover, these conjugates are in the form (41), which means that Lemma III.2.2 applies to them, and by further conjugating them we can continue adding multiples of a_{ni} and of a_{i1} to the $(1,n)$-entry of the resulting matrix.

Therefore, if we are given r in the ideal generated by the off-diagonal entries of $\mathcal{R}(A)$ and $\mathcal{C}_1(A)$, then we can express r as

$$r = \sum_{i=2}^{n-1} a_{ni}\lambda_i + \sum_{i=2}^{n-1} \lambda_i' a_{i1},$$

and setting

$$F = \prod_{i=2}^{n-1} E_{in}(\lambda_i) \cdot \prod_{i=2}^{n-1} E_{1i}(\lambda_i') \in \mathrm{E}_n(R),$$

we see that $F\Lambda F^{-1} = \Lambda + re_{1n} \in [\mathrm{E}_n(R), \mathrm{C}_A]$. Then

$$\Lambda^{-1}F\Lambda F^{-1} = \Lambda^{-1}(\Lambda + re_{1n}) = (I_{n-1} \mid \mathcal{C}_1(\Lambda^{-1}) \cdot r) = I_n + re_{1n} \in \mathrm{C}_A^4.$$

□

Remark. The above lemma also holds for non-commutative rings, with the same proof, if we take r to be in the left ideal generated by the off-diagonal entries of the first column of A, or in the right ideal generated by the off-diagonal entries of the last row of A.

In order to prove the main results of this section, we will need to define for $1 \leq i, j \leq n$ $\sigma_{ij} := \mathrm{E}_{ij}(1)\mathrm{E}_{ji}(-1)$, which is equivalent to the identity matrix which has the i-th and the j-th columns switched, and the i-th column multiplied by -1.

PROPOSITION III.2.7. *Suppose R is a commutative ring which satisfies the stable range condition $\mathtt{SR}_{n-1}$. Fix $B \in \mathrm{SL}_n(R)$, and for $i = 1, \dots, n$, let J_i be the ideal generated by the off-diagonal entries of $\mathcal{C}_i(B)$; let $J = J_1 + \dots + J_n$. Then:*

(a) $\mathrm{E}_{1n}(J_1) \subset \mathrm{C}_B^{32}$.

(b) $\mathrm{E}_{1n}(J) \subset \mathrm{C}_B^{32n}$.

PROOF. (a) If $(B)_{n1} = 0$, let $A = B$; otherwise, apply Lemma III.2.3 to pick a conjugate A of B so that $a_{11}R + \dots + a_{n-1\,1}R = R$. In either case, $a_{n1} = (B)_{n1}$, the ideal generated by $\{a_{21}, \dots, a_{n1}\}$ equals J_1, and there is a solution to the equation

$$\lambda_1 a_{11} + \dots + \lambda_{n-1} a_{n-1\,1} = a_{n1}. \tag{42}$$

Let us fix one such solution $(\lambda_1, \ldots, \lambda_{n-1})$, with each $\lambda_i \in R$. Then Lemma III.2.4 yields a matrix $C \in \mathrm{C}_A^2$ satisfying $\mathcal{R}(B) = (0 \ \ \lambda_1 \ \ 0 \ \ \ldots \ \ 0 \ \ 1)$. Since $A \in \mathrm{C}_B$, then $\mathrm{C}_A = \mathrm{C}_B$, so $C \in \mathrm{C}_B^2$. Lemma III.2.6 then guarantees that $\mathrm{E}_{1n}(\lambda_1) \in (\mathrm{C}_B^2)^4 = \mathrm{C}_B^8$.

Our goal is to show that for every element

$$r = \sum_{i=2}^{n} \alpha_i a_{i1} \in J_1, \tag{43}$$

we have $\mathrm{E}_{1n}(r) \in \mathrm{C}_B^{32}$. First, we treat the case when $\alpha_n = 0$. Observe that

$$(\lambda_1 + r)a_{11} - \lambda_1 a_{11} = \sum_{i=2}^{n-1} \alpha_i a_{11} \cdot a_{i1},$$

so adding the left hand side and subtracting the right hand side from (42), we see that

$$(\lambda_1 + r)a_{11} + \sum_{i=2}^{n-1} (\lambda_i - \alpha_i a_{11}) a_{i1} = a_{n1}.$$

This means $(\lambda_1 + r, \lambda_2 - \alpha_2 a_{11}, \ldots, \lambda_{n-1} - \alpha_{n-1} a_{11})$ is another solution to (42). Just as above, $\mathrm{E}_{1n}(\lambda_1 + r) \in \mathrm{C}_B^8$, and so $\mathrm{E}_{1n}(r) = \mathrm{E}_{1n}(\lambda_1 + r)\mathrm{E}_{1n}(\lambda_1)^{-1} \in \mathrm{C}_B^{16}$.

And now, we treat the general case. Suppose we are given an element r in the form (43). Let $r' = \sum_{i=2}^{n-1} \alpha_i a_{i1}$. Then r' can be expressed in the form (43) with $\alpha_n = 0$. The above procedure then applies to show that $\mathrm{E}_{1n}(r') \in \mathrm{C}_B^{16}$. Furthermore, we can write $r = r' + \alpha_n a_{n1}$; setting $A' := \sigma_{n-1,n} A \sigma_{n-1,n} \in \mathrm{C}_B$, we see that $(A')_{n-1,1} = -a_{n1}$. This means the above procedure applies to A' to show that $\mathrm{E}_{1n}(\alpha_n a_{n1}) \in \mathrm{C}_B^{16}$. Multiplying the above two matrices, we obtain $\mathrm{E}_{1n}(s) \in \mathrm{C}_B^{32}$.

(b) Suppose we are given $\alpha \in J$; write $\alpha = \alpha_1 + \ldots + \alpha_n$, with $\alpha_i \in J_i$ for each i. Let $B_1 = A$, and for each $i = 2, \ldots, n$, let $B_i = \sigma_{1i} A \sigma_{1i}^{-1} \in \mathrm{C}_A$. Then the ideal generated by the off-diagonal entries of $\mathcal{C}_1(B_i)$ is equal to the ideal generated by the off-diagonal entries $\mathcal{C}_i(A)$. Applying Proposition III.2.7(b) to the B_i's, we see that $E_{1n}(\alpha_i) \in \mathrm{C}_A^{32}$ for $i = 1, \ldots, n$. Multiplying these elements together, we obtain that

$$\mathrm{E}_{1n}(\alpha) = \mathrm{E}_{1n}(\alpha_1) \cdots \mathrm{E}_{1n}(\alpha_n) \in \mathrm{C}_A^{32n}.$$

□

Proposition III.2.8. *Suppose R is a commutative Noetherian ring which satisfies the stable range condition $\mathtt{SR}_{n-1}$, and for which $\mathrm{E}_n(R)$ has finite width with respect to $\mathcal{E}_n(R)$. Fix a non-central matrix $A \in \mathrm{SL}_n(R)$, and let J be the ideal generated by the off-diagonal entries of A. If the index of J in R is finite, then there exists a constant N, so that C_A^N contains the group $\mathrm{F}_n(R, J)$.*

Proof. By Proposition III.2.7(b), we know that $\mathrm{E}_{1n}(J) \subset \mathrm{C}_A^{32n}$. Since for a fixed $\alpha \in J$ all the elementary matrices $\{E_{ij}(\alpha) \mid 1 \leq i \neq j \leq n\}$ are conjugate one to another, it follows that $\mathcal{F}_n(R, J) \subset \mathrm{C}_A^{32n}$. Finally, since the index of J in R is finite, Theorem III.1.13 implies that the set of elementary matrices $\mathcal{F}_n(R, J)$ boundedly generates $\mathrm{F}_n(R, J)$, say $\mathrm{F}_n(R, J) = \mathcal{F}_n(R, J)^2$. Then $\mathrm{F}_n(R, J) \subseteq \mathrm{C}_A^{32nw}$. □

We end this work with a result which shows that Proposition III.2.7 applies to each non-central subset of $\mathrm{SL}_n(R)$ which is closed under conjugation.

LEMMA III.2.9. *Let R be a (commutative) ring, and $\mathcal{A} \subseteq \mathrm{SL}_n(R)$ be a non-central subset closed under conjugation. Then $\mathcal{A}$ contains a matrix with at least one non-zero off-diagonal entry in the first column.*

PROOF. Suppose we are given $B \in \mathcal{A} \backslash Z(\mathrm{SL}_n(R))$. Let us consider two cases: B is diagonal, and B is not. If B is diagonal, then since it is non-central, there must be some $h \in 2, \ldots, n$ so $(B)_{hh} \neq (B)_{11}$. Then $\mathrm{E}_{h1}(-1)B\mathrm{E}_{h1}(1) \in \mathcal{A}$ has the $(h, 1)$-th entry equal to $(B)_{hh} - (B)_{11} \neq 0$, implying the result. Otherwise, it has some non-zero off-diagonal entry, say $(B)_{ij} \neq 0$ for some $1 \leq i \neq j \leq n$. If $j = 1$ then we are done; otherwise, $\sigma_{1j} B \sigma_{1j} \in \mathcal{A}$ has the $(i, 1)$-th entry equal to $-(B)_{ij} \neq 0$, implying the result. □

Bibliography

[ALM] N. Avni, A. Lubotzky, C. Meiri, *First order rigidity*, Invent. Math., **217**(2019), No. 1, 219-240.

[AT] E. Artin, J. Tate, *Class Field Theory*, AMS, 1967.

[B] H. Bass, *K-theory and stable algebra*, Publications mathématiques de l'I.H.É.S., **22**(1964), 5-60.

[BMS] H. Bass, J. Milnor, J.-P. Serre, *Solution of the congruence subgroup problem for* $\mathrm{SL}_n(n \geq 3)$ *and* $Sp_{2n}(n \geq 2)$, Publications mathématiques de l'I.H.É.S., **33**(1967), 59-137.

[Br] J. L. Brenner, *The Linear Homogeneous Group, III*, Annals of Mathematics, **71**(1960), No. 2, 210-223.

[Bu] M. Burger, *Kazhdan Constants for* $\mathrm{SL}(3, \mathbb{Z})$, J. reine angew. Math., **413**(1991), 36-67.

[CF] J. W. S. Cassels, A. Fröhlich, *Algebraic Number Theory*, Thompson Book Company Inc., 1967.

[CK1] D. Carter, G. Keller, *Bounded Elementary Generation of* $\mathrm{SL}_n(O)$, American Journal of Mathematics, **105**(1983), No. 3, 673-687.

[CK2] D. Carter, G. Keller, *Elementary expressions for unimodular matrices*, Communications in Algebra, **12**(1984), No. 4, 379-389.

[Co] P. M. Cohn, *On the structure of the* GL_2 *of a ring*, Institut des Hautes Études Scientifiques, **30**(1966), 5-53.

[CW] G. Cooke, P.J. Weinberger, *On the construction of division chains in algebraic number rings, with applications to* SL_2, Communications in Algebra **3**(1975), No. 6, 481-524.

[EJ] M. Ershov, A. Jalkin-Zapirain, *Property (T) for noncommutative universal lattices*, Invent. Math., **179**(2010), No. 2, 303-347.

[ER] I. V. Erovenko, A. S. Rapinchuk, *Bounded generation of* S*-arithmetic subgroups of isotropic orthogonal groups over number fields*, J. Number Theory **119**(2006), No. 1, 28-48.

[FT] A. Fröhlich, M.J. Taylor *Algebraic Number Theory*, Cambridge University Press, 1991.

[GS] F.J. Grunewald, J. Schwermer, *Free Non-abelian Quotients of* SL_2 *Over Orders of Imaginary Quadratic Numberfields*, Journal of Algebra **69**(1981), 298-304.

[He] A. Heald, *Bounded Generation of Two Families of* S*-arithmetic Groups*, Ph.D. Thesis, University of Virginia, 2013.

[Hu] T.W. Hungerford, *Algebra*, Springer-Verlag, 1974.

[Hb] D.R. Heath-Brown, *Artin's conjecture for primitive roots*, Quart. J. Math. Oxford (2), **37**(1986), no. 145, 27-38.

[La] S. Lang, *Algebra*, Springer, 2002.

[LM] D. Loukanidis, V.K. Murty, *Bounded generation for* SL_n $(n \geq 2)$ *and* Sp_{2n} $(n \geq 1)$, preprint (1994).

[L1] B. Liehl, *On the group* SL_2 *over orders of arithmetic type.*, Journal für die reine und angewandte Mathematik **323**(1981), 153-171.

[L2] B. Liehl, *Beschränkte Wortlänge in* SL_2, Mathematische Zeitschrift **186**(1984), 509-524.

[Lu] A. Lubotzky, *Subgroup growth and congruence subgroups*, Invent. Math., **119**(1995), 267-295.

[M] J. Milnor, *Introduction to Algebraic K-Theory*, Princeton University Press, 1971.

[MRS] A.V. Morgan, A.S. Rapinchuk, B. Sury, *Bounded Generation of* SL_2 *over rings of* S*-integers with infinitely many units*, Algebra and Number Theory **8**(2018), no. 12, 1949-1974.

[MCKP] D.W. Morris, *Bounded generation of* $\mathrm{SL}(n, A)$ *(after D. Carter, G. Keller and E. Paige)*, New York Journal of Mathematics **13**(2007), 383-421.

[M] V.K. Murty, *Bounded and finite generation of arithmetic groups. (English summary)*, Number theory (Halifax, NS, 1994), CMS Conf. Proc., **15**, Amer. Math. Soc., Providence, RI, (1995), 249-261.

[MP] M.R. Murty, K.L. Petersen, *The generalized Artin conjecture and arithmetic orbifolds*, Groups and symmetries, CRM Proceeding & Lecture Notes **47**, Amer. Math. Soc., Providence, RI, (2009), 259-263.

[NA1] P.S. Novikov, S.I. Adjan, *Infinite periodic groups I*, Izv. Akad. Nauk SSSR, Ser. Mat. **32**, no. 1, 209-236.

[NA2] P.S. Novikov, S.I. Adjan, *Infinite periodic groups I*, Izv. Akad. Nauk SSSR, Ser. Mat. **32**, no. 2, 241-479.

[NA3] P.S. Novikov, S.I. Adjan, *Infinite periodic groups I*, Izv. Akad. Nauk SSSR, Ser. Mat. **32**, no. 3, 665-685.

[PR1] V.P. Platonov, A.S. Rapinchuk, *Abstract properties of S-arithmetic groups and the congruence problem*, Russian Academy of Sciences. Izvestiya Mathematics **40**(1993), No. 3, 455-476.

[PR2] V.P. Platonov, A.S. Rapinchuk, *Algebraic Groups and Number Theory*, Academy of Sciences, 1991.

[R] A.S. Rapinchuk, *Representations of groups of finite width*, Soviet Math. Dokl. **42**(1991), 816-820.

[S1] J.-P. Serre, *Le problème des groupes de congruence pour* SL_2, Annals of Mathematics, Second Series, **92**(1970), No. 3, 489-527.

[S2] J.-P. Serre, *Local Fields*, Springer-Verlag, 1979.

[Sh] Y. Shalom, *Bounded Generation and Kazhdan's property (T).*, Inst. Hautes Études Sci. Publ. Math., **90**(1999), No. 2001, 145-168.

[SW] Y. Shalom, G. Willis, *Commensurated subgroups of arithmetic groups, totally disconnected groups and adelic rigidity*, Geometric and Functional Analysis, **23**(2013), No. 5, 1631-1683.

[T] O.I. Tavgen, *Bounded generation of Chevalley groups over rings of algebraic S-integers*, Izv. Akad. Nauk SSSR Ser. Mat., **54**(1990), No. 1, 97-122.

[Va] L.N. Vaserštein, *The group* SL_2 *over Dedekind rings of arithmetic type*, Mathematics of the USSR – Sbornik, **18**(1972), No. 2, 321-332.

[Vs] M. Vsemirnov, *Short unitriangular factorizations of* $\mathrm{SL}_2(\mathbb{Z}[1/p])$, Quart. J. Math., **65** (2014), 279-290.

[ZS] O. Zariski, P. Samuel, *Commutative Algebra*, Vol. 1, Van Nostrand, 1958.