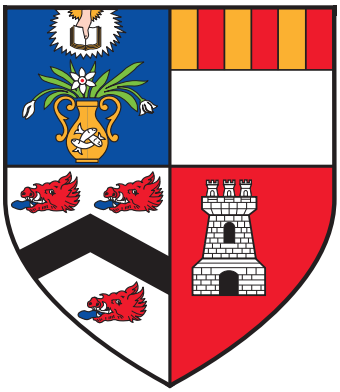
Quantitative aspects of normal generation of split Chevalley Groups

Alexander Alois Trost

BSc Mathematics, Technical University of Munich, 2014

MSc Mathematics, Ludwig Maximilian University of Munich, 2017

1 4 9 5



A thesis presented for the degree of

Doctor of Philosophy

at the

University of Aberdeen

supervised by Benjamin Martin and Jaroslaw Kędra

School of Natural and Computing Sciences

Department of Mathematics

2020

Declaration

I declare that this thesis has been composed entirely by myself and that the work contained within is my own. No portion of the work contained in this document has been submitted in support of an application for a degree or qualification at this or any other university or other institution of learning. All verbatim extracts have been distinguished by quotation marks and all sources of information have been specifically acknowledged.

Signed:

Abstract

This thesis is concerned with the diameter of certain word norms on S-arithmetic split Chevalley groups of higher ranks. Such groups are well known to be boundedly generated by root elements. We prove that word metrics given by finitely many conjugacy classes on S-arithmetic split Chevalley groups have an upper bound only depending on the number of those conjugacy classes. This property, called *strong boundedness*, was introduced by Kędra, Libman and Martin in [24] and proven for $\mathrm{SL}_n(R)$, assuming R is a principal ideal domain and $n \geq 3$. We give two methods to generalize such results to other Chevalley groups, one general and the other for specific examples: First, we give a general argument for all Chevalley groups using older results about normal subgroups of Chevalley groups and model-theory. Second, for the specific examples of $\mathrm{Sp}_{2n}(R)$ for $n \geq 2$, $E_6(R)$ and $G_2(R)$, we show such results by way of explicit calculations. We also provide examples of normally generating sets for S-arithmetic split Chevalley groups proving our upper bounds on the afore-mentioned word norms are sharp in an appropriate sense and give a complete account of the existence of small normally generating sets of $\mathrm{Sp}_4(R)$ and $G_2(R)$. For instance, we prove that $\mathrm{Sp}_4(\mathbb{Z}[\frac{1+\sqrt{-7}}{2}])$ cannot be generated by a single conjugacy class.

Acknowledgements

No achievements happen in a vacuum and there are a lot of people without whom this thesis would not exist. First, I want to thank my PhD supervisors Benjamin Martin and Jarek Kędra whose support and input during the research for this thesis and encouragement while writing it up were invaluable. Second, I want to thank my colleague Bastien Karlhofer who was always willing to talk with me about mathematics, poke holes in bad proofs and go drinking in the evening. This brings me to third group who I owe a debt of gratitude: The math department in Aberdeen, which is as sociable and pleasant as one could possibly imagine and with whose members like Dejan Govc, Ehud Meir, Ellen Henke and Irakli Patchkovia I shared countless fun evenings and discussions and who made me feel at home in Aberdeen and Scotland. Fourth, I want to thank my office mate Julian Kaspczyk who valiantly held out against my disorganization taking over our office without complaining once. Last, I want to thank my high school math teacher, Josef Fertl, whose one time reading of Simon Singh's 'Fermat's Last Theorem' sparked a lifelong obsession with mathematics, that he worked hard on fueling further, my close friends Michelle Vogl-Huff and Florian Kaspar whose friendships I all too often take for granted and my family and parents, who always supported me in my studies. My PhD studies and this work was funded by Leverhulme Trust Research Project Grant RPG-2017-159 and I want to express my gratitude to the Leverhulme Trust for the one-month extension of my studies that they funded swiftly and unbureaucratically.

Contents

Chapter 1

Main results and historical context

1.1 Historical context

Conjugation-invariant norms on groups are an old, if slightly non-standard topic in mathematics. They are defined as follows:

Definition 1.1.1. Let G be a group. Then a *conjugation invariant norm on* G is a map $\nu : G \to [0, +\infty)$ such that

1. for $a \in G : \nu(a) = 0$ holds precisely if $a = 1$.
2. for $a, b \in G : \nu(ab) \leq \nu(a) + \nu(b),\ \nu(aba^{-1}) = \nu(b)$ and $\nu(a) = \nu(a^{-1})$.

The group G is called *bounded*, iff for all conjugation invariant norms ν on G the diameter $\nu(G)$ is finite and *unbounded* if this is not the case.

A well-known example for such norms are bi-invariant Riemannian metrics on a compact Lie-group G, which naturally have finite diameter, because Riemannian metrics are continuous and G is compact. However, more often than not in the study of conjugation invariant norms, the topology induced by those norms is not the 'natural' topology of the corresponding group in question. A striking example of this is the Hofer-metric on the hamiltonian diffeomorphism groups:

If (M, ω) is a symplectic manifold, then let $H : M \times [0, 1] \to \mathbb{R}$ be a compactly supported smooth map and define vector fields $\{X_t\}_{t \in [0,1]}$ of the form

$$\iota_{X_t}\omega = dH(\cdot, t)$$

for $t \in [0, 1]$ and consider the corresponding flow $f_t : M \to M$ for $t \in [0, 1]$. Such diffeomorphisms f_t are called *hamiltonian diffeomorphisms* and the function H is said to generate the diffeomorphism f_1. The hamiltonian diffeomorphisms form a group $\mathrm{Ham}(M, \omega)$ and

this group admits the so-called *Hofer metric*:

$$\|f\|_H := \inf_{H \text{ generates} f} \int_0^1 (\max_{x\in M} H(x,t)) - (\min_{y\in M} H(y,t))dt.$$

This norm is conjugation invariant, but the natural C^1-topology on $\mathrm{Ham}(M,\omega)$ is finer than the topology induced by the Hofer-metric. The Hofer metric has been widely studied by Polterovich [36], Wagner and Ostrover [35] and many others and it is connected to various problems in symplectic topology and hamiltonian dynamics. Also beyond a couple of examples, like closed surfaces for which the diameter is unbounded [36, Corollary 7.2.D], it is not even known whether the Hofer metric has finite or infinite diameter in general. There are also other conjugation-invariant norms on Hamiltonian diffeomorphism groups as discovered by Viterbo [46] and conjugation invariant norms on contactomorphism groups found by Sandon [38],[39].

Less geometrically, there exist the so called fragmentation norms on compactly supported diffeomorphism or homeomorphism groups of manifolds. For example, the fragmentation norm on the homeomorphism group of a manifold, measures how many homeomorphisms supported in an open ball on this manifold, one needs to write a given compactly supported homeomorphism. So there are quite a lot of examples of such norms in geometry and topology. More algebraically, one can use these norms to study group theoretic properties of a group. For example the boundedness of a group is connected to quasi-morphisms:

Lemma 1.1.2. *[8] Let G be a group that admits an unbounded quasi-morphism $q : G \to \mathbb{R}$, meaning there exists a $D \geq 0$ called the defect of q such that for all $a, b \in G$*

$$|q(ab) - q(a) - q(b)| \leq D$$

holds and the set $q(G)$ is unbounded. Assume further that there is a finite subset $S \subset G$ with $\langle\langle S\rangle\rangle = G$. Then G is unbounded.

Proof. Define a conjugation invariant norm $\|\cdot\|_S : G \to \mathbb{N}_0$ by

$$\|a\|_S := \min\{n \in \mathbb{N}_0 |\ g_1, \ldots, g_n \text{ are up to conjugation elements of } S \cup S^{-1} \text{ and } a = g_1 \cdots g_n\}$$

for $a \neq 1$ and by $\|1\|_S := 0$. Further, for $a \in G$ the limit

$$q_\infty(a) := \lim_{n\to\infty} \frac{q(a^n)}{n}$$

is well-defined and defines a conjugation-invariant quasi-morphism $q_\infty : G \to \mathbb{R}$, that is

$$q_\infty(aba^{-1}) = q_\infty(b)$$

holds for all $a, b \in G$. Further, q_∞ is invariant under inverses and still unbounded. We leave these claims as an excercise to the reader. So q itself can be assumed to be conjugation-invariant and with defect D. Also choose $K := \max(\{|q(s)| \mid s \in S\} \cup \{D\})$. The number K is greater than 0 : This is the case because if D can be chosen as 0, then $q : G \to \mathbb{R}$ is a homomorphism. Hence if $\|q(s)\| = 0$ and so $q(s) = 0$ were to hold for all $s \in S$ as well, then the fact that the conjugacy classes of S generate G and $\mathbb{R}$ is abelian would imply that $q(G) = \{0\}$ and so q would not be unbounded. Next, let $a \in G - \{1\}$ be given with $\|a\|_S = n$ and choose $s_1, \ldots, s_n \in S \cup S^{-1}$ and $g_1, \ldots, g_n \in G$ with

$$a = \prod_{i=1}^{n} g_i s_i g_i^{-1}.$$

Then observe

$$\begin{aligned} |q(a)| = |q(\prod_{i=1}^{n} g_i s_i g_i^{-1})| &\leq (n-1)D + \sum_{i=1}^{n} |q(g_i s_i g_i^{-1})| \\ &= (n-1)D + \sum_{i=1}^{n} |q(s_i)| \leq nD + nK - D \leq 2nK = 2K\|a\|_S. \end{aligned}$$

But q is unbounded and $K > 0$, so $\|\cdot\|_S$ is unbounded as well. □

Further, one can show the following using a similar argument:

Lemma 1.1.3. *Let a group G with a finite subset S with $\langle\langle S \rangle\rangle = G$ and a group H with a conjugation invariant norm ν be given and let $\psi : G \to H$ be a homomorphism. Further, set $K := \max\{\nu(\psi(s))| \ s \in S\}$. Then for all $g \in G$:*

$$\nu(\psi(g)) \leq K\|g\|_S.$$

This implies for example that a bounded group generated by finitely many conjugacy classes, when acting by a hamiltonian group action on a symplectic manifold (M, ω) must have bounded image in $\mathrm{Ham}(M, \omega)$ with respect to the Hofer-metric. These lemmas also indicate the importance of the so called *conjugation generated word norms* $\|\cdot\|_S$ for S a finite set. One can ask more generally how such conjugation generated word norms $\|\cdot\|_S$ on a group G behave. For example, does the diameter $\|G\|_S$ of $\|\cdot\|_S$ being finite depend on the S in question or not? This is not the case:

Lemma 1.1.4. *[24, Corollary 2.5] Let a group G with a finite subset S with $\langle\langle S \rangle\rangle = G$. Then G is bounded precisely if the diameter of the conjugation generated word norm $\|\cdot\|_S$ is finite.*

However, boundedness is still not a property that is well behaved from a geometric group theory viewpoint. For example boundedness of a group does not pass to finite index

sub or supergroups in general.

Lemma 1.1.5. *[24, Example 2.8] The group $D_\infty = \mathbb{F}_2 \star \mathbb{F}_2$ is bounded and its index 2-subgroup $\mathbb{Z}$ is not.*

It is well-known that the kernel $\ker(j_G^2)$ of the comparison map $j_G^2 : H_b^2(G, \mathbb{R}) \to H^2(G, \mathbb{R})$ between bounded cohomology in degree 2 and ordinary group cohomology is isomorphic to the space of homogeneous quasi-morphisms $q : G \to \mathbb{R}$ modulo the space of actual homomorphisms $q : G \to \mathbb{R}$. Thus a group being bounded implies that j_G^2 is injective. However, these two properties are not equivalent: Each homogeneous quasi-morphism $\mathbb{Z} \to \mathbb{R}$ is a homomorphism and so $j_{\mathbb{Z}}^2$ is injective, but the group $\mathbb{Z}$ is obviously unbounded.

On the other hand, there is a classical result due to Burger and Monod [9, Corollary 1.3] stating the injectivity of j_G^2 for cocompact, irreducible lattices G in higher rank Lie groups and in a lot of cases one can also show that a lattice is bounded using straightforward calculations in the lattice itself as done by Gal, Kedra and myself [19]. To my knowledge, it is an open problem whether all lattices in groups of higher rank are bounded.

Assuming that a group is bounded, one can further ask how the diameter of G with respect to $\|\cdot\|_S$ depends on the chosen finite set S normally generating G. If G is a finite simple group, this is related to the classical question of covering numbers: For such a group G, its *covering number* $\mathrm{cn}(G)$ is the smallest natural number $n \in \mathbb{N}$ such that for every non-trivial conjugacy class C, one has $G = C^n$. If $S \neq \{1\}$ is a subset of the finite, simple group G, then $\|G\|_S \leq \mathrm{cn}(G)$ holds. Covering numbers have been extensively studied for various kind of finite simple groups for example Brenners FINASIG-papers [7] contain various different covering results for different finite simple groups, Lev's paper [26] contains covering results for $\mathrm{PSL}_n(K)$ for sufficiently large fields K and [4] is a survey about covering numbers containing various standard arguments and a table of covering numbers for all finite, simple groups with less than 1.000.000 elements. More recently, there have been a couple of remarkable papers about the asymptotics of covering numbers for finite simple groups by Liebeck and his collaborators [27],[25].

However, the first paper that studied conjugation invariant norms on groups named as such is presumably the Burago, Ivanov and Polterovich paper [8]. The paper establishes some general facts about conjugation invariant norms, their connection to quasi-morphisms and boundedness and uses these norms to determine commutator lengths on various diffeomorphism groups of spheres and 3-manifolds. More importantly, the line of ideas that lead to this thesis started with [8, Example 1.6]. There it was observed that the group $\mathrm{SL}_n(\mathbb{Z})$ is bounded for $n \geq 3$. This follows from two facts: First, the following delightful commutator formula for elementary matrices:

$$E_{13}(x) = (E_{12}(1), E_{23}(x))$$

for all $x \in \mathbb{Z}$. This implies $\nu(E_{13}(x)) \leq 2\nu(E_{12}(1))$ for a conjugation invariant norm ν on $\mathrm{SL}_n(\mathbb{Z})$. Second, each element A of $\mathrm{SL}_n(\mathbb{Z})$ can be written as a product of conjugates of matrices of the form $E_{13}(x)$ for $x \in \mathbb{Z}$ with a number of factors K independent of A according to a result by Carter and Keller [10]. These two facts, then imply $\nu(\mathrm{SL}_n(\mathbb{Z})) \leq 2K\nu(E_{12}(1))$ and hence boundedness of $\mathrm{SL}_n(\mathbb{Z})$. Later, it was observed by Gal, Kedra and myself in [19] that this idea generalizes immediately to other higher rank S-arithmetic Chevalley groups and finite index subgroups of them. The example of $\mathrm{SL}_n(\mathbb{Z})$ also indicates the connection to bounded generation by root elements, when studying the norms $\|\cdot\|_S$ on Chevalley groups.

But even knowing that $\mathrm{SL}_n(\mathbb{Z})$ is bounded for $n \geq 3$, it is still unclear how the diameter $\|\mathrm{SL}_n(\mathbb{Z})\|_S$ depends on the finite set S in question. Morris [30] has shown for any localization R of an order in a ring of algebraic integers (think a localization of $\mathbb{Z}[2i]$), that for $n \geq 3$ the diameter $\|\mathrm{SL}_n(R)\|_S$ has an upper bound only depending on the cardinality of S, as well as the ring R and n. This fascinating paper really highlights the connection to bounded generation result and I speak extensively about it in Chapter 6. However, it still does not describe the asymptotic of $\|\mathrm{SL}_n(R)\|_S$ in $|S|$. This question was answered by Kedra, Libman and Martin at least in the special case of rings of S-algebraic integers with class number 1. Class number 1 merely means that the ring is a principal ideal domain, but rings of S-algebraic integers are defined as follows:

Definition 1.1.6. [32, Chapter I, §11] Let K be a finite field extension of $\mathbb{Q}$. Then let T be a finite subset of the set V of all valuations of K such that T contains all archimedean valuations. Then the ring $\mathcal{O}_T$ is defined as

$$\mathcal{O}_T := \{a \in K |\ \forall v \in V - T : v(a) \geq 0\}$$

and $\mathcal{O}_T$ is called the *ring of T-algebraic integers in K* and rings $\mathcal{O}_T$ of this form are called *rings of S-algebraic integers.*

The answer given by Kędra, Libman and Martin is:

Corollary 1.1.7. *[24, Corollary 6.2] Let R be a ring of S-algebraic integers with class number 1 and let $n \geq 3$ be given. Then $\mathrm{SL}_n(R)$ is normally generated by the single element $E_{1,n}(1)$ and*

1. *for all finite, normally generating subsets S of G, it holds*

$$\|\mathrm{SL}_n(R)\|_S \leq (4n + 51)(4n + 4)|S|.$$

2. *for each $k \in \mathbb{N}$ there is a finite normally generating set S_k of $\mathrm{SL}_n(R)$ with $|S_k| = k$ and $\|\mathrm{SL}_n(R)\|_{S_k} \geq k$.*

This finally leads to the very questions, I set out to answer in this thesis: First, can the explicit bounds on the asymptotic of $\|\mathrm{SL}_n(R)\|_S$ also be shown for more general rings of S-algebraic integers R that are not principal ideal domains? And second, do these results hold for more general simply-connected Chevalley groups $G(\Phi, R)$ besides SL_n for $n \geq 3$? The groups $G(\Phi, R)$ are defined in Section 2.1, but for the sake of this introduction, one should think of classical matrix groups like $\mathrm{SL}_n, \mathrm{Sp}_{2n}$ defined by way of an irreducible root system Φ.

1.2 Main results and methods

I give positive answers to both of these questions and my main result is the following:

Theorem 1.2.1. *Let Φ be an irreducible root system of rank at least 2 and let R be a commutative ring with 1. Additionally, let $G(\Phi, R)$ be boundedly generated by root elements and if $\Phi = C_2$ or G_2, then further assume $(R : 2R) < \infty$. Then there is a constant $C(\Phi, R) \in \mathbb{N}$ such that for all finite, normally generating subsets S of G, it holds*

$$\|G(\Phi, R)\|_S \leq C(\Phi, R)|S|.$$

Remark 1.2.2. Root elements are natural generalizations of the elementary matrices $E_{ij}(x)$ in SL_n. Such root elements are usually denoted by $\varepsilon_\phi(x)$ with varying $\phi \in \Phi$ and $x \in R$. Most notably

$$\varepsilon_\phi(x_1 + x_2) = \varepsilon_\phi(x_1)\varepsilon_\phi(x_2)$$

holds for all $x_1, x_2 \in R$. I define them and bounded generation by root elements in Section 2.2.

The assumptions in Theorem 1.2.1 are quite general so the theorem can in principle be applied to a lot of different rings. The two main examples, I talk about in this thesis are rings of S-algebraic integers and semi-local rings. First regarding rings of S-algebraic integers, I show in Chapter 6, that the following generalization of the first part in Corollary 1.1.7 holds:

Theorem 6.1.4. *Let R be a ring of S-algebraic integers in a number field and Φ an irreducible root system of rank at least 2. Then there is a constant $C(\Phi, R) \geq 1$ such that for each finite, normally generating set S of $G(\Phi, R)$ the inequality $\|G(\Phi, R)\|_S \leq C(\Phi, R)|S|$ holds.*

Second, I state some results about boundedness of $G(\Phi, R)$ for R a semi-local ring:

Theorem 5.2.4. *Let R be a commutative, semilocal ring with 1 and let Φ an irreducible root system of rank at least 2. Furthermore, assume that $(R : 2R) < \infty$, if $\Phi = C_2$ or*

G_2. *Then there is a constant* $K(\Phi, R)$ *such that for each finite, normally generating set* S, *the inequality* $\|G(\Phi, R)\|_S \leq K(\Phi, R)$ *holds.*

Theorem 5.2.4 is essentially a corollary of Theorem 1.2.1 and the fact that for R a semi-local ring $G(\Phi, R)$ is boundedly generated by root elements.

The proofs of Theorem 1.2.1 and Corollary 1.1.7 are similar, so let us describe them briefly in the case of $\Phi \neq C_2$ or G_2: First, one obtains arbitrary root elements $\varepsilon_\phi(x)$ as products of conjugates of the finite normally generating set S with the number of factors proportional to $|S|$. Then secondly, one uses bounded generation of the group by root elements, as shown by Tavgen [42] and Carter, Keller [10], to finish the proof. The difference of the two strategies lies in how one accomplishes the first step. To prove Corollary 1.1.7, Kedra, Libman and Martin use extensive matrix calculations and rely heavily on the underlying ring being a principal ideal domain to construct root elements (or elementary matrices) in this case. Rather then using such explicit calculations, my strategy to show the first step considers for an element A of S the normal subgroup of $G(\Phi, R)$ generated by A. But the structure of such normal subgroups are well-understood for general commutative rings. For example, there is the following general theorem by Abe:

Theorem 3.3.1. *[2, Theorem 1,2,3,4] Let* Φ *be an irreducible root system that is not* A_1, C_2 *or* G_2 *and let* R *be a commutative ring with* 1. *Then for each subgroup* $H \subset G(\Phi, R)$ *normalized by the group* $E(\Phi, R)$, *there is an ideal* $J \subset R$ *and an additive subgroup of* L *of* J *such that* $\bar{E}(J, L) \subset H \subset E^*(J, L)$.

Remark 1.2.3. I define the subgroups $E(\Phi, R), \bar{E}(J, L)$ and $E^*(J, L)$ in Section 2.2, but crucially $\bar{E}(J, L)$ contain many root elements as long J is not trivial.

Thus considering H as the normal subgroup generated by A, certain root elements are contained in $\bar{E}(J, L)$. Hence one can describe an inconsistent theory in first order logic that describes that these root elements cannot be written as a product of conjugates of A or A^{-1} with any finite number u of factors. But then Gödel's Compactness Theorem [37, Theorem 3.2] implies that already a finite sub-theory is inconsistent. This in turn implies the possibility of writing these root elements as products with a certain bounded number of conjugates of A or A^{-1} as factors with the bound on the number of factors independent of the actual $A \in G(\Phi, R)$ in question. Then combining the various root elements obtained for varying $A \in S$, one can finish the proof of the first step.

This at least is the strategy for $\Phi \neq C_2$ or G_2. If Φ is either of those, then the first step, using results different from Theorem 3.3.1 instead, does not quite yield all root elements. Instead, it yields that the set

$$Q_{C_2} := \{A\varepsilon_\phi(2x)A^{-1} | \phi \in C_2, x \in R, A \in \mathrm{Sp}_4(R)\}$$

in case of $\Phi = C_2$ and

$$\begin{aligned} Q_{G_2} :=& \{A\varepsilon_\phi(2x)A^{-1} | \phi \in G_2 \text{ short}, x \in R, A \in G_2(R)\} \\ & \cup \{A\varepsilon_\phi(x)A^{-1} | \phi \in G_2 \text{ long}, x \in R, A \in G_2(R)\} \end{aligned}$$

in case of $\Phi = G_2$ are bounded with respect to $\| \cdot \|_S$ with a bound proportional to $|S|$. Then I use the remaining assumption of $(R : 2R)$ being finite to finish the proof of Theorem 1.2.1 in this case. In any case, I want to emphasize the difference between $\mathrm{Sp}_4(R)$, $G_2(R)$ and all other cases, because it will appear again and again in this thesis.

Ultimately, what this proof strategy shows is that rather than Corollary 1.1.7 and Theorem 1.2.1 being results about geometric group theory as initially suspected, they are much more directly connected to classical algebraic K-theory. A natural interpretation of Theorem 1.2.1 is to consider it as a quantitative version of results about normal subgroups in Chevalley groups.

Regarding the question of explicit upper bounds on $\|G(\Phi, R)\|_S$ however as they are stated in Corollary 1.1.7, I want to note that my argument to prove Theorem 1.2.1 cannot yield them. The reason for this is that the proof is non-constructive and does not provide an explicit algorithm to construct root elements instead merely showing that it is possible to do so. Instead, I give two different methods to calculate upper bounds explicitly. Both methods, the first quite similar to the matrix calculations done to prove Corollary 1.1.7 and the other dependent on a variant of the Bruhat decomposition [41, Chapter 8, p. 68, Corollary 1], rely on the ring R being a principal ideal domain. This is due to the fact that both require at different places, the possibility to represent the greatest common divisor of two elements of R as a principal ideal. In any case, I prove for example the following explicit result using matrix calculations:

Corollary 6.1.6. *Let R be a ring of S-algebraic integers with class number one and $n \geq 3$. Further set*

$$\Delta(R) := \begin{cases} 135, \text{ if } R \text{ is a quadratic imaginary ring of integers or } \mathbb{Z} \\ 12, \text{ if } R \text{ is neither of the above} \end{cases}$$

Then $\|\mathrm{Sp}_{2n}(R)\|_S \leq 192(1+5n)(12n+\Delta(R))|S|$ holds for each finite, normally generating S of $\mathrm{Sp}_{2n}(R)$.

There are three remarks, I want to make regarding this result: First, the proof uses R being of stable range at most 2: Namely, I use that according to Proposition 5.1.4, the decomposition $\mathrm{Sp}_{2n}(R) = (U^+(C_n, R)U^-(C_n, R))^2\mathrm{Sp}_4(R)$ holds for R a ring of stable range at most 2 to improve the asymptotic of bounded generation by root elements from the usual one that is quadratic in n appearing for example in Tavgens paper [42] to one that is under conjugation linear in n. Second, if one were interested in providing explicit

results for more general rings of S-algebraic integers and not just principal ideal domains, than the clearest strategy would be to use these stable range conditions for R by adapting the proof of Bass result [5, Theorem 4.2(e)] about normal subgroups of $\mathrm{SL}_n(R)$.

Third, as indicated by the decomposition $\mathrm{Sp}_{2n}(R) = (U^+(C_n, R)U^-(C_n, R))^2\mathrm{Sp}_4(R)$ and the previous remarks on the proof of Theorem 1.2.1, one would expect a difference between the cases $n = 2$ and $n \geq 3$ for $\mathrm{Sp}_{2n}(R)$.

In fact, the situation for $n = 2$ is more subtle and depends on number theoretic properties of the ring R in question, more precisely on the way the ideal $2R$ factors into prime ideals in R. In Chapter 6, I provide upper bounds on $\|\mathrm{Sp}_4(R)\|_S$ for rings of algebraic integers R with the property that each element of R can be written as a sum of an element of $2R$ and a unit by imitating the proof of the Bruhat-decomposition for fields in Steinberg's lecture notes [41]. For example, I show the following statement:

Proposition 6.2.14. *1. Let S be a finite, normally generating set of $\mathrm{Sp}_4(\mathbb{Z})$. Then $\|\mathrm{Sp}_4(\mathbb{Z})\|_S \leq 5 + 248064|S|$.*

2. Let S be a finite, normally generating set of $\mathrm{Sp}_4(\mathbb{Z}[\frac{1+\sqrt{-3}}{2}])$. Then

$$\|\mathrm{Sp}_4\left(\mathbb{Z}\left[\frac{1+\sqrt{-3}}{2}\right]\right)\|_S \leq 4 + 248064|S|.$$

Using the principal ideal domain version of the Bruhat decomposition [41, Chapter 8, p. 68, Corollary 1] mentioned before, I also prove the following two results:

Proposition 6.3.6. *1. Let S be a finite normally generating set of $G_2(\mathbb{Z})$. Then*

$$\|G_2(\mathbb{Z})\|_S \leq 82098906624|S| + 1.$$

2. Let S be a finite normally generating set of $G_2(\mathbb{Z}[\frac{1+\sqrt{-3}}{2}])$. Then

$$\|G_2\left(\mathbb{Z}\left[\frac{1+\sqrt{-3}}{2}\right]\right)\|_S \leq 61510897152|S|.$$

As mentioned, there is also an auxiliary step involved in proving Theorem 1.2.1 for $G_2(R)$, but this additional step is actually easier and less dependent on number theoretic properties for $G_2(R)$ than for $\mathrm{Sp}_4(R)$. This is the case, because in some sense the set Q_{G_2} is much larger than Q_{C_2}.

I also prove the following quantitative result for $E_6(R)$:

Theorem 6.1.7. *Let R be a ring of S-algebraic integers with class number one and let S*

be a finite normal generating set of $E_6(R)$*. Further set*

$$\Delta(R) := \begin{cases} 154, \text{ if } R \text{ is a quadratic imaginary ring of integers or } \mathbb{Z} \\ 117, \text{ if } R \text{ is neither of the above} \end{cases}$$

Then $\|E_6(R)\|_S \leq 120 \cdot 60^{2^{11}} \Delta(R)|S|$ *holds for all* $k \in \mathbb{N}$.

Comparing the bounds obtained from the decomposition [41, Chapter 8, p. 68, Corollary 1], like Theorem 6.1.7 and the ones in Corollary 1.1.7 or Corollary 6.1.6, one immediately notes that Theorem 6.1.7 is worse by several orders of magnitude. However, using the Bruhat decomposition [41, Chapter 8, p. 68, Corollary 1] has the advantage of being independent of any representation of the underlying groups, which makes them easier to work with for exceptional root systems Φ like E_6 and G_2.

But Corollary 1.1.7 does not only provide an upper bound on $\|\mathrm{SL}_n(R)\|_S$, it also states the existence of normally generating sets S_k of any given cardinality $k \in \mathbb{N}$ with diameter of $\|\cdot\|_{S_k}$ being at least k. In Chapter 7, I prove the following generalization of this statement:

Theorem 7.1.5. *Let* R *a Dedekind domain with finite class number and at least* k *distinct maximal ideals be given. Further, let* Φ *be one of the following root systems:*

1. A_n *for* $n \geq 2$,
2. B_n *for* $n \geq 3$,
3. C_n *for* $n \geq 3$,
4. D_n *for* $n \geq 4$,
5. E_6, E_7, E_8 *or* F_4

such that $G(\Phi, R)$ *is boundedly generated by root elements. Then there is a normally generating set* S_k *of* $G(\Phi, R)$ *with* $|S_k| = k$ *and*

1. $\|G(A_n, R)\|_{S_k} \geq k(n+1)$ *for* $n \geq 2$,
2. $\|G(B_n, R)\|_{S_k} \geq k(n+1)$ *for* $n \geq 3$,
3. $\|G(C_n, R)\|_{S_k} \geq 2nk$ *for* $n \geq 3$,
4. $\|G(D_n, R)\|_{S_k} \geq kn$ *for* $n \geq 4$,
5. $\|G(\Phi, R)\|_{S_k} \geq 2k$ *for* $\Phi = E_6, E_7, E_8, F_4$.

Notably, instead of just providing an asymptotic in k, these lower bounds also provide an asymptotic proportional to the rank of the root system. The way, these lower bounds are shown, not only here but also in Corollary 1.1.7 is by constructing root elements $\{\varepsilon_\phi(r_i)|i=1,\dots,k\}=S_k$ whose arguments $r_1,\dots,r_k\in R$ have a large set of common prime divisors T. Then one reduces modulo these prime divisors and considers the induced conjugation generated word norm on $\prod_{\mathcal{P}\in T}G(\Phi,R/\mathcal{P})$.

In the proof of Corollary 1.1.7 it was then observed that the induced conjugation generated word norms on the factors of $\prod_{\mathcal{P}\in T}G(\Phi,R/\mathcal{P})$ are not trivial. I on the other hand count the dimension of the 1-eigenspaces of the matrices induced on the factors $\prod_{\mathcal{P}\in T}G(\Phi,R/\mathcal{P})$ by S_k to obtain the lower bound proportional to the rank of the root system.

The situation for $\mathrm{Sp}_4(R)$ and $G_2(R)$ is again more complex. Instead of there being normally generating set of each size, the situation is restricted by number theoretic properties of R. Namely, I show the following:

Theorem 7.2.1. *Let Φ be C_2 or G_2 and let R be a ring of S-algebraic integers in a number field. Further define*

$$r:=r(R):=|\{\mathcal{P}|\ \mathcal{P}\text{ divides }2R\text{ and }R/\mathcal{P}=\mathbb{F}_2\}|.$$

1. *Then for all $k\in\mathbb{N}$ with $k\geq r(R)$, there is a normally generating set S_k of $G(\Phi,R)$ such that $\|\mathrm{Sp}_4(R)\|_{S_k}\geq 4k+r(R)$ and $\|G_2(R)\|_{S_k}\geq 2k$ holds.*

2. *Then there is no normally generating set of $G(\Phi,R)$ with less than $r(R)$ elements*

Showing that there are no normally generating sets of, say $\mathrm{Sp}_4(R)$, with less than $r(R)$ elements is straight forward. It amounts to proving that there is an epimorphism $\mathrm{Sp}_4(R)\to\mathbb{F}_2^{r(R)}$ and so any normally generating set of $\mathrm{Sp}_4(R)$ with less than $r(R)$ elements would give rise to a generating set of $\mathbb{F}_2^{r(R)}$ with less than $r(R)$ elements. But such generating sets clearly cannot exist. For example, there is the following corollary:

Corollary 7.2.9. *Let D be a square-free integer and R the ring of algebraic integers in $\mathbb{Q}[\sqrt{D}]$ and let $G:=\mathrm{Sp}_4(R)$ or $G_2(R)$ be given.*

1. *If $D\equiv 2,3,5,6,7$ mod 8, then G can be generated by a single conjugacy class.*

2. *If $D\equiv 1$ mod 8, then G cannot be generated by a single conjugacy class, but there are two conjugacy classes C_1,C_2 in G generating G.*

Somewhat surprising to me is the first claim of Theorem 7.2.1, that the size of $r(R)$ is the sole obstruction to the existence of normally generating sets in $\mathrm{Sp}_4(R)$ and $G_2(R)$. The proof of this is slightly involved. Namely, I analyze the additive subgroup generated by units in the ring $R/2R$ and conclude that the prime factors of $2R$ with residue field

bigger than $\mathbb{F}_2$ do not matter when constructing normally generating sets.

Last, I want to talk about possible generalizations of the main results and some obvious problems raised by the results. For example, both Theorem 1.2.1 and Theorem 5.2.4 contain the assumption that $(R : 2R)$ is finite if $\Phi = C_2$ or G_2. While this is not a problem for the main application of rings of S-algebraic integers, it raises the question, how the situation looks for, say rings of S-algebraic integers in global fields of characteristic 2 and whether the assumption is really necessary. In this context, I prove the following result indicating that the assumption of $(R : 2R)$ being finite is not necessary in all cases:

Theorem 5.3.2. *Let $\mathbb{F}$ be a finite field, $\mathcal{P}$ a prime ideal in $\mathbb{F}[T]$ and R the localization of $\mathbb{F}[T]$ at $\mathcal{P}$. Then there is a constant $K(R)$ such that for each finite normally generating set S of $\mathrm{Sp}_4(R)$, the inequality $\|\mathrm{Sp}_4(R)\|_S \leq K(R)$ holds.*

The proof is slightly different than the one of say Theorem 5.2.4 and uses the assumption $\mathrm{char}(R) = 2$ extensively in that case. Ultimately though, the proof of Theorem 5.3.2 shows that even if in the case that the ring R in question has $\mathrm{char}(R) = 2$, one still obtains strong boundedness for $\mathrm{Sp}_4(R)$ under the sole assumption of bounded generation by root elements for $\mathrm{Sp}_4(R)$.

Also there is the question whether Theorem 6.1.4 can be generalized to other arithmetic lattices. While I do not give a general statement about this, I do show the following theorem indicating that this might very well be the case:

Theorem 8.1.1. *Let R be a ring of S-algebraic integers and let H be a subgroup of finite index in $\mathrm{SL}_n(R)$ for $n \geq 3$. Then there is a constant $C(H) \in \mathbb{N}$ such that for each finite, normally generating set S of H, the inequality $\|H\|_S \leq C(H)|S|$ holds.*

The proof is quite similar to the proof of Theorem 1.2.1 in the case of $\mathrm{Sp}_4(R)$ or $G_2(R)$.

1.3 Structure of the thesis

As a document, the thesis is structured as follows: In Chapter 2, I define all needed notions like split Chevalley groups, their congruence subgroups and root elements, level ideals and the word norms.

In Chapter 3, I give a more precise formulation of Theorem 1.2.1 and prove it for general split Chevalley groups of higher rank using model theory and Gödel's Compactness Theorem. I also lay the ground work to talk about explicit bounds on the diameters $\|G(\Phi, R)\|_S$ by defining the main constants $L(\Phi), Q(\Phi, R)$ and in case of $\Phi = C_2$ or G_2 additionally the constants $K(\Phi, R)$ and $\Delta_\infty(G(\Phi, R)/N_\Phi)$ needed to provide explicit upper bounds.

In Chapter 4, I present two different methods to obtain an upper bound on $L(\Phi)$ in case that R is a principal ideal domain. The first method is by way of matrix calculations and is shown for the example $\mathrm{Sp}_{2n}(R)$ for $n \geq 3$ and $\mathrm{Sp}_4(R)$. The second method uses a version of the Bruhat decomposition in $G(\Phi, R)$ for R a principal ideal domain and is shown for the examples $G_2(R)$ and $E_6(R)$.

In Chapter 5, I talk about stable range conditions for commutative rings, how they can be used to obtain decompositions for Chevalley groups and to give a first step to obtain bounds on $Q(\Phi, R)$ in certain cases. Further, this chapter treats the case of semi-local rings and talks about boundedness properties in positive characteristic to an extent.

Chapter 6 talks about bounded generation results for rings of S-algebraic integers and uses them to prove Theorem 6.1.4 and to provide explicit upper bounds on $\|\mathrm{Sp}_{2n}(R)\|_S$ for $n \geq 3$ by using the discussion of the previous two chapters. Furthermore, Chapter 6 provides upper bounds for $K(C_2, R)$ and $K(G_2, R)$ as well as for $\Delta_\infty(G(C_2, R)/N_{C_2})$ and $\Delta_\infty(G(G_2, R)/N_{G_2})$ to provide upper bounds on $\|\mathrm{Sp}_4(R)\|_S$ and $\|G_2(R)\|_S$ in special cases of rings of S-algebraic integers, called $2R$-pseudo-good rings. For example, this chapter contains my proofs of Proposition 6.2.14 and Proposition 6.3.6. Lastly, I talk in this chapter about Morris' paper [30] and his results to some extent.

In Chapter 7, I explain how to construct the normally generating sets required to prove Theorem 7.1.5 and Theorem 7.2.1.

In Chapter 8, I prove Theorem 8.1.1 and talk about possible issues and ideas to generalize my results.

In Appendix A, I have for the convenience of the reader collected various statements and definitions regarding root systems used throughout the thesis. Appendix B describes the exceptional isomorphism of $\mathrm{Sp}_4(\mathbb{F}_2)$ and S_6 and Appendix C contains proofs of results in the thesis that for various reasons did not fit into the thesis proper.

Chapter 2

Definitions and basic properties

This chapter is divided into three sections. In the first section, we define the split Chevalley groups studied in this thesis. In the second section, we define special element of split Chevalley groups, called root elements and describe some of their properties. In the third section, we define the level ideals of elements of split Chevalley groups and explain their connection to central elements.

But first, we introduce the notions of boundedness and conjugation invariant word norms we study in this thesis anew:

Definition 2.0.1. Let G be a group.

1. The notation $A \sim B$ for $A, B \in G$ denotes that A, B are conjugate in G. Furthermore, we define $A^B := BAB^{-1}$ for $A, B \in G$.

2. For $S \subset G$, we define $\langle\langle S \rangle\rangle$ as the smallest normal subgroup of G containing S.

3. A subset $S \subset G$ is called a *normally generating set* of G, if $\langle\langle S \rangle\rangle = G$.

4. The group G is called *finitely normally generated*, if a finite normally generating set S exists.

5. For $k \in \mathbb{N}$ and $S \subset G$ define the following set

$$B_S(k) := \bigcup_{1 \leq i \leq k} \{x_1 \cdots x_i | \forall j \in \{1, \ldots, i\} : \ x_j \text{ or } x_j^{-1} \text{ are conjugate to elements of } S\} \cup \{1\}.$$

   Further set $B_S(0) := \{1\}$. If S only contains the single element A, then we write $B_A(k)$ instead of $B_{\{A\}}(k)$.

6. Define for a set $S \subset G$ the conjugation invariant word norm $\|\cdot\|_S : G \to \mathbb{N}_0 \cup \{+\infty\}$ by

$$\|A\|_S := \min\{k \in \mathbb{N}_0 | A \in B_S(k)\}$$

for $A \in \langle\langle S \rangle\rangle$ and by $\|A\|_S := +\infty$ for $A \notin \langle\langle S \rangle\rangle$. The diameter $\|G\|_S = \operatorname{diam}(\|\cdot\|_S)$ of G is defined as the minimal $N \in \mathbb{N}$, such that $\|A\|_S \leq N$ for all $A \in G$ holds, or as $+\infty$ if there is no such N.

7. Define for $k \in \mathbb{N}$ the invariant

$$\Delta_k(G) := \sup\{\operatorname{diam}(\|\cdot\|_S) |\ S \subset G \text{ with } |S| \leq k, \langle\langle S \rangle\rangle = G\} \in \mathbb{N}_0 \cup \{\infty\}$$

with $\Delta_k(G)$ defined as $-\infty$, if there is no normally generating set $S \subset G$ with $|S| \leq k$.

8. Define the invariant

$$\Delta_\infty(G) := \sup\{\operatorname{diam}(\|\cdot\|_S) |\ S \subset G \text{ finite with } \langle\langle S \rangle\rangle = G\} \in \mathbb{N}_0 \cup \{\infty\}$$

with $\Delta_\infty(G)$ defined as $-\infty$, if there is no finite, normally generating set $S \subset G$.

9. The group G is called *strongly bounded*, if $\Delta_k(G)$ is finite for all $k \in \mathbb{N}$. It is called *uniformly bounded*, if $\Delta_\infty(G)$ is finite.

Remark 2.0.2. Note $\Delta_k(G) \leq \Delta_{k+1}(G) \leq \Delta_\infty(G)$ holds for all $k \in \mathbb{N}$.

We will also use the following lemma throughout the thesis, usually without explicit reference:

Lemma 2.0.3. *Let G be a group and let $a, b, x \in G$ be given. Then*

1. $(ab, x) = (b, x)^a \cdot (a, x)$ *and*
2. $(ab, x) \sim (b, x) \cdot (x, a^{-1})$

hold.

Proof. Observe that

$$\begin{aligned}(ab, x) = abx(ab)^{-1}x^{-1} = a\left(bxb^{-1}\right) \cdot a^{-1}x^{-1} = a\left(bxb^{-1}x^{-1}\right) \cdot \left(xa^{-1}x^{-1}\right)\\ = a(b, x)a^{-1} \cdot \left(axa^{-1}x^{-1}\right) = (b, x)^a \cdot (a, x).\end{aligned}$$

This yields the first claim. But $(b, x)^a \cdot (a, x)$ is conjugate to

$$(b, x) \cdot a^{-1}(a, x)a = (b, x)a^{-1}axa^{-1}x^{-1}a = (b, x) \cdot \left(x, a^{-1}\right)$$

and this yields the second claim. □

2.1 Simply connected split Chevalley groups

To define split Chevalley groups, we will first define the Chevalley-Demazure group scheme. We do not prove various statements in the course of this definition. For a more complete description please consider [11] and [41, Theorem 1, Chapter 1, p.7; Theorem 6(e), Chapter 5, p.38; Lemma 27, Chapter 3, p. 29]. Also, we use various claims about root systems and Weyl group and have collected some of these statements in Appendix A for the convenience of the reader.

Let G be a simply-connected, semi-simple complex Lie group and T a maximal torus in G with associated irreducible root system Φ. Further, denote by Π a system of positive, simple roots of Φ, by $\mathfrak{g}$ the corresponding complex semi-simple Lie-algebra of G. The Cartan-subalgebra corresponding to T will be denoted by $\mathfrak{h}$ and the corresponding root spaces in $\mathfrak{g}$ by $\mathfrak{g}_\phi$ for $\phi \in \Phi$. These choices of Cartan-subalgebra and (simple, positive) roots will be fixed throughout the thesis. In particular, it will always be clear which roots in Φ are called positive and simple. We will usually denote the positive roots in Φ by Φ^+ and the negative ones by Φ^-. The Lie-algebra $\mathfrak{g}$ has a so-called *Chevalley basis*

$$\{X_\phi \in \mathfrak{g}_\phi\}_{\{\phi\in\Phi\}} \cup \{H_\alpha \in \mathfrak{h}\}_{\{\alpha\in\Pi\}}$$

such that for all $\phi, \psi \in \Phi$ and $\alpha \in \Pi$ the following conditions hold:

(a) $[H_\alpha, X_\phi] = \phi(H_\alpha)X_\phi$ and $\phi(H_\alpha) \in \mathbb{Z}$

(b) $[X_\phi, X_{-\phi}] \in \bigoplus_{\alpha\in\Pi} \mathbb{Z}H_\alpha$,

(c) $[X_\phi, X_\psi] = \pm(r+1)X_{\phi+\psi}$, if $\phi + \psi \in \Phi$ and $r := \max\{i \in \mathbb{N}_0 | \phi - i\psi \in \Phi\}$

(d) $[X_\phi, X_\psi] = 0$, if $\phi + \psi \neq 0$ and $\phi + \psi \notin \Phi$

Chevalley-basis are unique up to signs and automorphisms of $\mathfrak{g}$. Furthermore, one sets $\langle \lambda, \alpha \rangle := \lambda(H_\alpha)$ for any linear map $\lambda : \mathfrak{h} \to \mathbb{C}$ and $\alpha \in \Pi$.

For each faithful, smooth representation $\rho : G \to GL(V)$ for a complex vector space V, there is a lattice $V_\mathbb{Z}$ in V with the property:

$$\frac{d\rho(X_\phi)^k}{k!}(V_\mathbb{Z}) \subset V_\mathbb{Z} \text{ for all } \phi \in \Phi \text{ and } k \geq 0.$$

Fixing a minimal generating set $\{v_1, \ldots, v_n\}$ of $V_\mathbb{Z}$, then defines functions $t_{ij} : G \to \mathbb{C}$ for all $1 \leq i, j \leq n$ by:

$$\rho(g)(v_j) = \sum_{i=1}^{n} t_{ij}(g)v_i,$$

because the set $\{v_1, \ldots, v_n\}$ also defines a $\mathbb{C}$-basis of V. The functions t_{ij} generate a $\mathbb{Z}$-Hopf algebra called $\mathbb{Z}[G]$ by way of the multiplication in G and $\mathbb{Z}[G]$ defines the Chevalley-Demazure group scheme by

$$G(\Phi, \cdot) : R \mapsto G(\Phi, R) := \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}[G], R)$$

with the group structure on $G(\Phi, R)$ given by the Hopf-algebra structure on $\mathbb{Z}[G]$. For a ring homomorphism $R \to S$, the corresponding group homomorphism

$$G(\Phi, R) = \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}[G], R) \to \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}[G], S) = G(\Phi, S)$$

is obtained by postcomposing with the ring homomorphism $R \to S$. The group scheme $G(\Phi, \cdot)$ does not depend up to isomorphism on the choices of Chevalley basis, faithful representation ρ and lattice $V_{\mathbb{Z}}$.

Further note, that the ring $\mathbb{Z}[y_{ij}]$ is a finitely generated $\mathbb{Z}$-algebra and $\mathbb{Z}$ is noetherian. Hence the polynomial ring in several unknown $\mathbb{Z}[y_{ij}]$ is noetherian and hence there is a finite collection of polynomial functions $P \subset \mathbb{Z}[y_{ij}]$ such that $\mathbb{Z}[y_{ij}]/(P) \cong \mathbb{Z}[G]$ with the isomorphism given by $y_{ij} \mapsto t_{ij}$ and the Hopf-algebra structure on $\mathbb{Z}[y_{ij}]/(P)$ given by

$$y_{i,j} + (P) \mapsto (\sum_{k=1}^{n} y_{ik} \otimes y_{kj}) + (P).$$

Using this, one can equivalently define $G(\Phi, R)$ as a subgroup of $GL_n(R)$ by setting:

$$G(\Phi, R) := \{A \in R^{n \times n} | \forall p \in P : p(A) = 0\}.$$

In this notation, for another ring S with a ring homomorphism $R \to S$, the induced maps $G(\Phi, R) \to G(\Phi, S)$ are obtained by entry-wise application of the ring homomorphism $R \to S$. We will use mostly this interpretation of $G(\Phi, R)$ in the course of this thesis and also we will use the notation $P(A) = 0$ to denote that $p(A) = 0$ holds for all $p \in P$.

Remark 2.1.1. In terms of algebraic groups, the group $G(\Phi, R)$ is the group of R-points of the $\mathbb{Z}$-defined group scheme $G(\Phi, \cdot)$.

2.2 Root elements

Next, we will define the root elements of Chevalley groups. Again, we do not give all details. Fix a root $\phi \in \Phi$ and observe that for $Z \in \mathbb{C}$ arbitrary the following function is

an element of $\rho(G) \subset GL(V)$:

$$\varepsilon_\phi(Z) := \sum_{k=0}^{\infty} \frac{(Zd\rho(X_\phi))^k}{k!}$$

Further, $\rho(\varepsilon_\phi(Z))$ in $GL(V)$ has coordinates with respect to the basis $\{v_1, \ldots, v_n\}$ that are polynomial functions in Z with coefficients in $\mathbb{Z}$. This yields a ring homomorphism

$$\varepsilon_\phi : \mathbb{Z}[G] \to \mathbb{Z}[Z].$$

By precomposing, this defines another map as follows:

$$\varepsilon_\phi : \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}[Z], R) \to \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}[G], R) = G(\Phi, R)$$

Lastly, the root elements $\varepsilon_\phi(x) \in G(\Phi, R)$ for $x \in R$ are defined as the image of the map $x : \mathbb{Z}[Z] \to R, Z \mapsto x$ under the map ε_ϕ. The element $x \in R$ in question is often refered to as the *argument of* $\varepsilon_\phi(x)$. We also denote the subgroup

$$\langle \varepsilon_\phi(x) |\ x \in R \rangle$$

of $G(\Phi, R)$ by ε_ϕ or $\varepsilon_\phi(R)$. We refer the reader to [41] for further details regarding root elements.

Also note the following property:

Definition 2.2.1. Let R be a commutative ring with 1. Then $G(\Phi, R)$ is *boundedly generated by root elements*, if there is a natural number $N := N(\Phi, R) \in \mathbb{N}$ and roots $\phi_1, \ldots, \phi_N \in \Phi$ such that for all $A \in G(\Phi, R)$, there are $a_1, \ldots, a_N \in R$ (depending on A) such that:

$$A = \prod_{i=1}^{N} \varepsilon_{\phi_i}(a_i).$$

Further, we define the following two word norms:

Definition 2.2.2. Let R be a commutative ring with 1 and Φ an irreducible root system such that $G(\Phi, R)$ is generated by root elements. Then define the two sets

$$\mathrm{EL} := \{\varepsilon_\phi(t) |\ t \in R, \phi \in \Phi\} \text{ and } \mathrm{EL}_Q := \{A\varepsilon_\phi(t)A^{-1} |\ t \in R, \phi \in \Phi, A \in G(\Phi, R)\}.$$

Then

1. define the word norm $\| \cdot \|_{\mathrm{EL}} : G(\Phi, R) \to \mathbb{N}_0$ as $\|1\|_{\mathrm{EL}} := 0$ and as

$$\|X\|_{\mathrm{EL}} := \min\{n \in \mathbb{N} | \exists A_1, \ldots, A_n \in \mathrm{EL} : X = A_1 \cdots A_n\}$$

for $X \neq 1$.

2. define the word norm $\|\cdot\|_{\mathrm{EL}_Q} : G(\Phi, R) \to \mathbb{N}_0$ as $\|1\|_{\mathrm{EL}_Q} := 0$ and as

$$\|X\|_{\mathrm{EL}_Q} := \min\{n \in \mathbb{N} | \exists A_1, \ldots, A_n \in \mathrm{EL}_Q : X = A_1 \cdots A_n\}$$

for $X \neq 1$.

3. If $G(\Phi, R)$ is additionally boundedly generated by root elements, than we set $Q(\Phi, R) := \|G(\Phi, R)\|_{\mathrm{EL}_Q}$.

Remark 2.2.3. If $G(\Phi, R)$ is boundedly generated by root elements than both the diameters $\|G(\Phi, R)\|_{\mathrm{EL}}$ and $Q(\Phi, R) := \|G(\Phi, R)\|_{\mathrm{EL}_Q}$ are always finite, because

$$\|G(\Phi, R)\|_{\mathrm{EL}_Q} \leq \|G(\Phi, R)\|_{\mathrm{EL}} \leq N(\Phi, R)$$

holds. However, $\|G(\Phi, R)\|_{\mathrm{EL}_Q}$ might be smaller than $\|G(\Phi, R)\|_{\mathrm{EL}}$.

The group elements $\varepsilon_\phi(t)$ are *additive in* $t \in R$, that is $\varepsilon_\phi(t+s) = \varepsilon_\phi(t)\varepsilon_\phi(s)$ holds for all $t, s \in R$. Further, a couple of commutator formulas, expressed in the next lemma, hold. We will use the additivity and the commutator formulas implicitly throughout the thesis usually without reference.

Lemma 2.2.4. *[21, Proposition 33.2-33.5] Let R be a commutative ring with 1 and let Φ be an irreducible root system of rank at least 2. Let $\alpha, \beta \in \Phi$ be roots with $\alpha + \beta \neq 0$ and let $a, b \in R$ be given.*

1. *If $\alpha + \beta \notin \Phi$, then $(\varepsilon_\alpha(a), \varepsilon_\beta(b)) = 1$.*

2. *If α, β are positive, simple roots in a root subsystem of Φ isomorphic to A_2, then $(\varepsilon_\beta(b), \varepsilon_\alpha(a)) = \varepsilon_{\alpha+\beta}(\pm ab)$.*

3. *If α, β are positive, simple roots in a root subsystem of Φ isomorphic to C_2 with α short and β long, then*

$$(\varepsilon_{\alpha+\beta}(b), \varepsilon_\alpha(a)) = \varepsilon_{2\alpha+\beta}(\pm 2ab) \text{ and}$$
$$(\varepsilon_\beta(b), \varepsilon_\alpha(a)) = \varepsilon_{\alpha+\beta}(\pm ab)\varepsilon_{2\alpha+\beta}(\pm a^2 b).$$

4. *If α, β are positive, simple roots in a root system Φ isomorphic to G_2 with α short*

and β long, then

$$
\begin{aligned}
&(\varepsilon_\beta(b), \varepsilon_\alpha(a)) = \varepsilon_{\alpha+\beta}(\pm ab)\varepsilon_{2\alpha+\beta}(\pm a^2 b)\varepsilon_{3\alpha+\beta}(\pm a^3 b)\varepsilon_{3\alpha+2\beta}(\pm a^3 b^2),\\
&(\varepsilon_{\alpha+\beta}(b), \varepsilon_\alpha(a)) = \varepsilon_{2\alpha+\beta}(\pm 2ab)\varepsilon_{3\alpha+\beta}(\pm 3a^2 b)\varepsilon_{3\alpha+2\beta}(\pm 3ab^2),\\
&(\varepsilon_{2\alpha+\beta}(b), \varepsilon_\alpha(a)) = \varepsilon_{3\alpha+\beta}(\pm 3ab),\\
&(\varepsilon_{3\alpha+\beta}(b), \varepsilon_\beta(a)) = \varepsilon_{3\alpha+\beta}(\pm ab) \text{ and}\\
&(\varepsilon_{2\alpha+\beta}(b), \varepsilon_{\alpha+\beta}(a)) = \varepsilon_{3\alpha+2\beta}(\pm 3ab).
\end{aligned}
$$

Remark 2.2.5. The signs before the arguments on the right hand side of the above commutator formulas might vary depending on the choice of the Chevalley basis. These issues are commonly referred to as *pinning*. This problem will not be resolved in this thesis, due to the fact that our norms are invariant under taking inverses anyway.

Before continuing, we will define the Weyl group elements and diagonal elements in $G(\Phi, R)$:

Definition 2.2.6. Let R be a commutative ring with 1 and let Φ be a root system. Define for $t \in R^*$ and $\phi \in \Phi$ the elements:

$$w_\phi(t) := \varepsilon_\phi(t)\varepsilon_{-\phi}(-t^{-1})\varepsilon_\phi(t).$$

We will often write $w_\phi := w_\phi(1)$. We also define $h_\phi(t) := w_\phi(t)w_\phi(1)^{-1}$ for $t \in R^*$ and $\phi \in \Phi$.

Remark 2.2.7. Let $\Pi = \{\alpha_1, \ldots, \alpha_u\}$ be a system of simple, positive roots in the root system Φ. If $w = w_{\alpha_{i_1}} \cdots w_{\alpha_{i_k}}$ is an element of the Weyl group $W(\Phi)$ as defined in Appendix A, then there is an element $\widetilde{w} \in G(\Phi, R)$ defined by $\widetilde{w} := w_{\alpha_{i_1}}(1) \cdots w_{\alpha_{i_k}}(1)$. We will often denote this element $\widetilde{w}$ of $G(\Phi, R)$ by w as well.

Using these Weyl group elements, we can obtain the following lemma:

Lemma 2.2.8. *Let R be a commutative ring with 1 and Φ an irreducible root system with Π its system of simple roots. Let $\phi \in \Phi, \alpha \in \Pi$ and $x \in R, t \in R^*$ be given. Then $\varepsilon_\phi(x)^{w_\alpha} = \varepsilon_{w_\alpha(\phi)}(\pm x)$ and $\varepsilon_\phi(x)^{h_\alpha(t)} = \varepsilon_\phi(t^{\langle\phi,\alpha\rangle}x)$ hold and so for each $S \subset G(\Phi, R)$, one has*

$$\|\varepsilon_\phi(x)\|_S = \|\varepsilon_{w_\alpha(\phi)}(x)\|_S.$$

Here the element $w_\alpha(\phi)$ is defined by the action of $W(\Phi)$ on Φ from Appendix A.

Proof. This is a direct consequence of [41, Chapter 3, p. 23, Lemma 20(b)]. □

However according to Proposition A.0.8, for Φ an irreducible root system and $\phi_1, \phi_2 \in \Phi$ two roots of the same length, there is an element $w \in W(\Phi)$ such that $w(\phi_1) = \phi_2$. Hence Lemma 2.2.8 implies:

Lemma 2.2.9. *Let R be a commutative ring with 1 and Φ an irreducible root system. Further, let $\phi_1, \phi_2 \in \Phi$ be two roots of the same length and let $x \in R$ be given. Then for each $S \subset G(\Phi, R)$, one has*

$$\|\varepsilon_{\phi_1}(x)\|_S = \|\varepsilon_{\phi_2}(x)\|_S.$$

We will use the previous two lemmas throughout the thesis usually without explicit reference.

In particular, Lemma 2.2.9 implies for Φ an irreducible root system, $\phi \in \Phi, k \in \mathbb{N}$ and $S \subset G(\Phi, R)$, that the set

$$\{x \in R |\ \varepsilon_\phi(x) \in B_S(k)\}$$

only depends on the length of ϕ and not on the particular ϕ in question. But as seen in Proposition A.0.5, there are at most two root lengths in any irreducible root system Φ and hence the following sets are well-defined:

Definition 2.2.10. Let R be a commutative ring with 1 and Φ an irreducible root system and let $S \subset G(\Phi, R)$ be given. Then for $k \in \mathbb{N}_0$

1. define the subset $\varepsilon_s(S, k)$ of R as $\{x \in R |\ \varepsilon_\phi(x) \in B_S(k)\}$ for any short root $\phi \in \Phi$,
2. define the subset $\varepsilon_l(S, k)$ of R as $\{x \in R |\ \varepsilon_\phi(x) \in B_S(k)\}$ of R for any long root $\phi \in \Phi$,
3. for $\psi \in \Phi$ define the subset $\varepsilon(S, \psi, k)$ of R as $\{x \in R |\ \varepsilon_\psi(x) \in B_S(k)\}$.

Further, if Φ is simply-laced, that is if there is only one root length, then we denote $\varepsilon_s(S, k) = \varepsilon_l(S, k)$ by $\varepsilon(S, k)$.

Next, we will define subgroups of $G(\Phi, R)$ and some other notions that we need later on:

Definition 2.2.11. Let Φ be an irreducible root system and let R be a commutative ring with 1 in the following.

1. The *elementary subgroup* $E(\Phi, R)$ (or $E(R)$ if Φ is clear from the context) is defined as the subgroup of $G(\Phi, R)$ generated by the elements $\varepsilon_\phi(x)$ for $\phi \in \Phi$ and $x \in R$.
2. The subgroup $U^+(\Phi, R)$, called *the subgroup of upper unipotent elements of* $G(\Phi, R)$, is the subgroup of $G(\Phi, R)$ generated by the root elements $\varepsilon_\phi(x)$ for $x \in R$ and $\phi \in \Phi$ a positive root. Similarly, one can define $U^-(\Phi, R)$, *the subgroup of lower unipotent elements of* $G(\Phi, R)$.
3. The *upper Borel subgroup* $B^+(\Phi, R) = B^+(R) = B(R)$ *of* $G(\Phi, R)$ is the subgroup of $G(\Phi, R)$ generated by $U^+(\Phi, R)$ and all elements $h_\phi(t)$ for $\phi \in \Phi$ and $t \in R^*$. The *lower Borel subgroup* $B^-(\Phi, R) = B^-(R)$ *of* $G(\Phi, R)$ is the subgroup of $G(\Phi, R)$ generated by $U^-(\Phi, R)$ and all elements $h_\phi(t)$ for $\phi \in \Phi$ and $t \in R^*$.

4. For each pair (J, L), where J is an ideal in R and L an additive subgroup of J, we define the subgroup $E(J, L)$ of $G(\Phi, R)$ as the group generated by all elements of the form $\varepsilon_\alpha(x)$ for $\alpha \in \Phi$ short, $x \in J$ and $\varepsilon_\beta(y)$ for $\beta \in \Phi$ long, $y \in L$.

5. For each such pair (J, L), we define the subgroup $\bar{E}(J, L)$ as the normal closure of $E(J, L)$ in $E(R)$.

6. For each such pair (J, L), we define the subgroup $E^*(J, L)$ as follows:

$$E^*(J, L) := \{A \in G(R, \Phi) | (A, E(R)) \subset \bar{E}(J, L)\}.$$

7. For a proper ideal J in R the map $\pi_J : G(\Phi, R) \to G(\Phi, R/J)$ is the group homomorphism induced by the quotient map $R \to R/J$.

Remark 2.2.12. If there is a natural number $N := N(\Phi, R) \in \mathbb{N}$ and roots $\phi_1, \ldots, \phi_N \in \Phi$ such that for all $A \in E(\Phi, R)$, there are $a_1, \ldots, a_N \in R$ (depending on A) such that $A = \prod_{i=1}^{N} \varepsilon_{\phi_i}(a_i)$, then we call the group $E(\Phi, R)$ *boundedly generated by root elements.*

Another technical result that we use quite often in later chapters is:

Proposition 2.2.13. *[41, Chapter 3, p. 21, Lemma 16,17; Chapter 8, p. 68, Lemma 49] Let Φ be an irreducible root system and let S, T be two sets of roots in Φ such that $S \subset T$ and the following conditions hold:*

1. $\forall \alpha, \beta \in T : (\alpha + \beta \in \Phi) \Rightarrow (\alpha + \beta \in T)$,
2. $\forall \alpha \in S \forall \beta \in T : (\alpha + \beta \in \Phi) \Rightarrow (\alpha + \beta \in S)$,
3. $\forall \alpha \in T : -\alpha \notin T$

Further, let R be a commutative ring with 1. *Then*

1. $\prod_{\phi \in T} \varepsilon_\phi(R)$ *is a subgroup of* $G(\Phi, R)$.
2. $\prod_{\phi \in S} \varepsilon_\phi(R)$ *is a subgroup of* $G(\Phi, R)$ *normalized by* $\prod_{\phi \in T} \varepsilon_\phi(R)$ *and* $\{h_\phi(t) | \phi \in \Phi, t \in R^*\}$.
3. *every element* A *of* $\prod_{\phi \in T} \varepsilon_\phi(R)$ *can be written uniquely as* $\prod_{\phi \in T} \varepsilon_\phi(x_\phi)$ *for* $x_\phi \in R$.
4. *if* R *is additionally a principal ideal domain and* K *its fraction field, then* $\prod_{\phi \in T} \varepsilon_\phi(R) = G(\Phi, R) \cap (\prod_{\phi \in T} \varepsilon_\phi(K))$.

Remark 2.2.14. 1. A subset T of a root system Φ with

$$\forall \alpha, \beta \in T : (\alpha + \beta \in \Phi) \Rightarrow (\alpha + \beta \in T)$$

is called *closed* and a subset S of a closed set T with

$$\forall \alpha \in S \beta \in T : (\alpha + \beta \in \Phi) \Rightarrow (\alpha + \beta \in S)$$

is called an *ideal in* T. Note that the proposition implies in particular that for $T = \Phi^+$ and S an ideal in Φ^+, that $\prod_{\phi \in S} \varepsilon_\phi(R)$ is normalized by $B^+(\Phi, R)$.

2. If the indexing set I of a product $\prod_{i \in I} u_i$ is ordered in some manner it is always understood that elements u_i appear further to the left in the product $\prod_{i \in I} u_i$ than elements u_j, if i is bigger with respect to the ordering of I than j.

2.3 Central elements of Chevalley groups and level ideals

Let G be a complex, simply-connected, simple Lie-group with root system Φ which is not C_2 or G_2 and $\Pi = \{\alpha_1, \ldots, \alpha_u\}$ be a system of positive, simple roots and $\{H_{\alpha_1}, \ldots, H_{\alpha_u}\}$ the associated elements of the Chevalley basis of $\mathfrak{g}$ contained in $\mathfrak{h}$. For each linear function $\lambda : \mathfrak{h} \to \mathbb{C}$ with $\langle \lambda, \alpha_i \rangle \in \mathbb{Z}_{>0}$ for all $i = 1, \ldots, u$ there is a unique representation $\rho_\lambda : G \to V_\lambda$ with highest weight λ by [41, Chapter 2, p. 14, Theorem 2]. The so-called *fundamental weights* $\lambda_1, \ldots, \lambda_u$ are defined by $\langle \lambda_j, \alpha_i \rangle = \delta_{ij}$ for $1 \leq i, j \leq u$. These consequently define fundamental representations $\rho_i : G \to GL(V_{\lambda_i}) =: GL(V_i)$ for $1 \leq i \leq u$. Then define $V := V_1 \oplus \cdots \oplus V_u$. The induced direct sum representation $\rho : G \to GL(V)$ is faithful. The faithfulness of this ρ can be obtained by noting that $G(\Phi, \mathbb{C}) = G$ and using properties of maps between different Chevalley groups as described in [41, Chapter 3, p. 29].

As mentioned in the first section, the Chevalley group $G(\Phi, R)$ does not depend on the chosen faithful representation up to isomorphism. In case of $\Phi \neq C_2$ or G_2, the Chevalley group arising from the representation ρ, is what we refer to as the split Chevalley group $G(\Phi, R)$ in general. Setting further $n_i := dim_{\mathbb{C}}(V_i)$ for $1 \leq i \leq u$, note that $G(\Phi, R)$ is a subgroup of $\mathrm{GL}_{n_1}(R) \times \cdots \times \mathrm{GL}_{n_u}(R) \subset \mathrm{GL}_{n_1+n_2+\cdots+n_u}(R)$ such that there is a collection of polynomials $P \subset \mathbb{Z}[y_{ij}]$ with

$$G(\Phi, R) = \{A \in \mathrm{GL}_{n_1+n_2+\cdots+n_u}(R) | P(A) = 0\}.$$

One could also use representations of the form of ρ for $\Phi = C_2$ and G_2, but due to the representations used in the formulations of certain theorems in Chapter 3, we will use a different representation in those cases. Furthermore, we use the standard presentation as matrix groups for $G(C_n, R) = \mathrm{Sp}_{2n}(R)$ and $G(A_n, R) = \mathrm{SL}_n(R)$ quite often, but mostly in Chapter 4, Chapter 5 and Chapter 7.

Also there is the following description of central elements in $G(\Phi, R)$ with respect to the representation ρ:

Lemma 2.3.1. *Let R be a reduced, commutative ring with 1 and Φ an irreducible root system with $\Phi \neq C_2$ or G_2. Further, let $A \in G(\Phi, R)$ centralize $E(\Phi, R)$. Then there are $t_1, \ldots, t_u \in R^*$ such that $A = (t_1 I_{n_1}) \oplus \cdots \oplus (t_u I_{n_u}) \in GL_{n_1+n_2+\cdots+n_u}(R)$. Furthermore, elements of this form are central in $G(\Phi, R)$.*

Proof. We split the proof of the first claim of the lemma into three parts. First, we are going to show the statement for fields, then for integral domains and finally for general reduced rings. So let K be a field and $A = (a_{kl}) \in G(\Phi, K)$ be given. For fields, one has $G(\Phi, K) = E(\Phi, K)$ by Proposition 5.2.2 and hence A is central in $G(\Phi, K)$. Then by [41, Chapter 3, p. 29, Lemma 28] there are $t_1, \ldots, t_u \in K - \{0\}$ such that $A = \prod_{i=1}^{u} h_{\alpha_i}(t_i)$ with $\{\alpha_1, \ldots, \alpha_u\} = \Pi$ the system of simple, positive roots in Φ chosen. Further, [41, Chapter 3, p. 29, Lemma 28] implies

$$1 = \prod_{i=1}^{u} t_i^{\langle \phi, \alpha_i \rangle} \text{ for all } \phi \in \Phi. \tag{2.1}$$

Furthermore, due to the construction of $G(\Phi, K)$, we know that $G(\Phi, K)$ is a subgroup of $\mathrm{GL}_{n_1}(K) \times \cdots \times \mathrm{GL}_{n_u}(K)$ and according to the remark after [41, Chapter 3, p. 29, Corollary 4] for $1 \leq j \leq u$, the element A acts on the λ_j-weight component of $K^{n_j} \subset K^{n_1+n_2+\cdots+n_u}$ by multiplication with

$$\prod_{i=1}^{u} t_i^{\langle \lambda_j, \alpha_i \rangle} = \prod_{i=1}^{u} t_i^{\delta_{ij}} = t_j. \tag{2.2}$$

This follows from the fact that λ_j is defined as the fundamental weight corresponding to α_j, that is $\langle \lambda_j, \alpha_i \rangle = \delta_{ij}$ holds for all $1 \leq i, j \leq u$. Each other weight of the action of $G(\Phi, K)$ on K^{n_j} has the form $\lambda_j - \sum \phi$, where the ϕ are positive roots in Φ. Then (2.1) and (2.2) imply that A acts on K^{n_j} by multiplication with $t_j I_{n_j}$. This holds for all j and so yields the claim for fields.

For integral domains R, we distinguish two cases:

Case 1. R is finite.

Finite integral domains are fields and hence we are done.

Case 2. R is infinite.

Let $\alpha \in \Phi$ be given and observe that for K the algebraic closure of the field of fractions of R, we have the map

$$\phi_\alpha : \mathbb{G}_a(K) = K \to G(\Phi, K), \lambda \mapsto (A, \varepsilon_\alpha(\lambda)).$$

This is a morphism of algebraic varieties and note that as A commutes with elements in $\varepsilon_\alpha(R)$ by assumption, $\phi_\alpha|_R$ is equal to the identity. But R is infinite and so Zariski-dense in $\mathbb{G}_a(K)$. So $\phi_\alpha|_R$ being the identity implies that ϕ_α is constant. Hence A centralizes the entire subgroup $\varepsilon_\alpha(K)$ in $G(\Phi, K)$. However $G(\Phi, K)$ is generated by the elements $\{\varepsilon_\alpha(\lambda)|\lambda \in K, \alpha \in \Phi\}$. Hence A is central in $G(\Phi, K)$ and hence it has the form $A = (t_1 I_{n_1}) \oplus \cdots \oplus (t_u I_{n_u})$ for $t_1, \ldots, t_u \in K - \{0\}$. However, the element A is contained in the subgroup $G(\Phi, R)$ of $G(\Phi, K)$ and thus $t_1, \ldots, t_u \in R$ holds. But $A^{-1} = (t_1^{-1} I_{n_1}) \oplus \cdots \oplus (t_u^{-1} I_{n_u})$ is also an element of $G(\Phi, R)$ and thus $t_1^{-1}, \ldots, t_u^{-1}$ are also elements of R and hence the $t_1, \ldots, t_u$ are units in R.

Lastly, let R be a reduced ring. Further, let $\mathcal{P}$ be a prime ideal in R. So $\pi_{\mathcal{P}}(A) \in G(\Phi, R/\mathcal{P})$ centralizes all elements of $E(\Phi, R/\mathcal{P})$ and $R/\mathcal{P}$ is an integral domain. Thus we obtain

$$A \equiv (a_{11} I_{n_1}) \oplus \cdots \oplus (a_{n_1+\cdots+n_{u-1}+1, n_1+\cdots+n_{u-1}+1} I_{n_u}) \bmod \mathcal{P} \tag{2.3}$$

for all prime ideals $\mathcal{P}$. Recall here that $A = (a_{kl})$ and hence the congruence (2.3) does not depend on $\mathcal{P}$. This implies

$$A \equiv (a_{11} I_{n_1}) \oplus \cdots \oplus (a_{n_1+\cdots+n_{u-1}+1, n_1+\cdots+n_{u-1}+1} I_{n_u}) \bmod \bigcap_{\mathcal{P} \text{ prime in } R} \mathcal{P} = \sqrt{(0)}.$$

However R is reduced and so $\sqrt{(0)} = (0)$ holds. Thus

$$A = (a_{11} I_{n_1}) \oplus \cdots \oplus (a_{n_1+\cdots+n_{u-1}+1, n_1+\cdots+n_{u-1}+1} I_{n_u})$$

holds and as in the the integral domain case, one obtains that $a_{11}, \ldots, a_{n_1+\cdots+n_{u-1}+1}$ are units in R. This finishes the proof of the first claim.

For the second claim, note that elements of $G(\Phi, R)$ are block matrices in

$$\mathrm{GL}_{n_1}(R) \times \cdots \times \mathrm{GL}_{n_u}(R) \subset \mathrm{GL}_{n_1+\cdots+n_u}(R)$$

and so matrices of the form $A = (a_{11} I_{n_1}) \oplus \cdots \oplus (a_{n_1+\cdots+n_{u-1}+1, n_1+\cdots+n_{u-1}+1} I_{n_u})$ are obviously centralizing $G(\Phi, R)$. □

Presumably this statement holds for general rings R, but we were not able to find a reference. Next, we give the definitions of $G(C_2, R) = \mathrm{Sp}_4(R)$ and $G_2(R)$. Both are still instances of our general definition of $G(\Phi, R)$ in Section 2.1, but we will not describe the representations and choices involved explicitly for G_2.

Definition 2.3.2. Let R be a commutative ring with 1 and let

$$Sp_4(R) := \{A \in R^{4\times 4} | A^T J A = J\}$$

be given with

$$J = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$$

The root system C_2 has four different positive roots namely, $C_2^+ = \{\alpha, \beta, \alpha+\beta, 2\alpha+\beta\}$ with α short and β long and both simple. The corresponding root elements in $\mathrm{Sp}_4(R)$ have (subject to the choice of maximal torus as diagonal matrices in $\mathrm{Sp}_4(\mathbb{C})$) the following form for $t \in R$:

$$\varepsilon_\alpha(t) = I_4 + t(e_{12} - e_{43}), \varepsilon_{\alpha+\beta}(t) = I_4 + t(e_{14} + e_{23})$$
$$\varepsilon_\beta(t) = I_4 + te_{24}, \varepsilon_{2\alpha+\beta}(t) = I_4 + te_{13}$$

and $\varepsilon_\phi(t) = (\varepsilon_{-\phi}(t))^T$ for negative roots $\phi \in C_2$.

We could specify an explicit matrix description for G_2 with explicit root elements as well, but this would be rather lengthy and instead we refer to the description in the appendix of [15]. This appendix gives G_2 as a subgroup-scheme of GL_8 defined over $\mathbb{Z}$. We will not specify which elements of $G_2 \subset \mathrm{GL}_8$ correspond to root elements in particular, but note the positive roots in the root system G_2. They are $G_2^+ = \{\alpha, \beta, \alpha+\beta, 2\alpha+\beta, 3\alpha+\beta, 3\alpha+2\beta\}$ with α short and β long and both simple. Further, note that the root subsystem generated by β and $3\alpha+\beta$ is isomorphic to the root system A_2.

Next, we will define various variants of level ideals:

Definition 2.3.3. Let R be a commutative ring with 1, Φ an irreducible root system and let $A \in G(\Phi, R)$ be given. The *level ideal* $l(A)$ is defined

1. in case $\Phi \neq C_2$ or G_2 as the ideal in R generated by the elements $a_{i,j}$ and $a_{i,i} - a_{j,j}$ for all $1 \leq i \neq j \leq n_1 + \cdots + n_u$ such that there is a $k \in \{1, \ldots, u-1\}$ with $n_1 + \cdots n_k + 1 \leq i \neq j \leq n_1 + \cdots + n_{k+1}$ or $1 \leq i \neq j \leq n_1$.

2. in case $\Phi = C_2$ as $l(A) := (a_{i,j}, (a_{i,i} - a_{j,j}) | 1 \leq i \neq j \leq 4)$.

3. in case $\Phi = G_2$ as $l(A) := (a_{i,j}, (a_{i,i} - a_{j,j}) | 1 \leq i \neq j \leq 8)$.

Furthermore, define the following ideals: If $\Phi = C_2$ define

$$l(A)_2 := (a_{i,j}^2, (a_{i,i} - a_{j,j})^2 | 1 \leq i \neq j \leq 4)$$

and if $\Phi = G_2$ define

$$l(A)_3 := (a_{i,j}^3, (a_{i,i} - a_{j,j})^3 | 1 \leq i \neq j \leq 8).$$

Remark 2.3.4.

1. In case $\Phi = C_2$ or G_2, note $l(A) \subset \sqrt{l(A)_2}$ or $l(A) \subset \sqrt{l(A)_3}$.

2. All of the ideals $l(A), l(A)_2$ and $l(A)_3$ are finitely generated and independently of the irreducible Φ in question, any element $A \in G(\Phi, R)$ maps to a central element in $G(\Phi, R/l(A))$, if $l(A) \neq R$. This is obvious in case of $\Phi = C_2$ or G_2 and a consequence of Lemma 2.3.1 otherwise.

3. Using the definition of the ideal $l(A)$ for $\Phi \neq C_2$ or G_2, the first claim of Lemma 2.3.1 is equivalent to the fact that central elements $A \in G(\Phi, R)$ satisfy $l(A) = \{0\}$.

Chapter 3

Strong boundedness of Chevalley groups

In this chapter, we show how to obtain strong boundedness for split Chevalley groups of higher ranks as a consequence of applying Gödel's Compactness Theorem to Sandwich Theorems describing normal subgroups of split Chevalley groups. Instead of proving Theorem 1.2.1 directly, we prove two other theorems for different root systems Φ, Theorem 3.1.2 and Theorem 3.2.5, that together are equivalent to Theorem 1.2.1. We handle the three cases of Φ being not C_2 or G_2 and Φ being either of them separately, because the lower rank examples are more complicated then the higher rank ones. This distinction will persist throughout this thesis.

In the first section, we state and prove the strong boundedness Theorem 3.1.2 for higher rank Chevalley groups $G(\Phi, R)$ for Φ not C_2 and G_2. Further, we state the main technical statement, Theorem 3.1.1, needed to prove Theorem 3.1.2. In the second section, we state similar technical statements, Theorem 3.2.1 and Theorem 3.2.2, for $\mathrm{Sp}_4(R)$ and $G_2(R)$ and explain how to use them to prove the strong boundedness Theorem 3.2.5, for them. In the third section, we prove the main technical statement, Theorem 3.1.1, for higher rank Chevalley groups for $\Phi \neq C_2$ or G_2 using Gödel's compactness theorem and in the fourth and fifth section respectively, we prove the main technical statements, Theorem 3.2.1 and Theorem 3.2.2, for $\mathrm{Sp}_4(R)$ and $G_2(R)$.

The two main definitions used in this chapter are:

Definition 3.0.1. Let R be a commutative ring with 1, I an ideal in R, Φ an irrducible root system and S a subset of $G(\Phi, R)$. Then define the following two subsets of maximal ideals in R :

1. $V(I) := \{m \text{ maximal ideal in } R | I \subset m\}$ and
2. $\Pi(S) := \{m \text{ maximal ideal of } R |\ \forall A \in S : \pi_m(A) \text{ central in } G(\Phi, R/m)\}$

We also note the following observation:

Lemma 3.0.2. *Let R be a commutative ring with 1, I_1, I_2 two ideals in R, Φ an irreducible root system in R and S, T two subsets of $G(\Phi, R)$. Then $V(I_1 + I_2) = V(I_1) \cap V(I_2)$ and $\Pi(S \cup T) = \Pi(S) \cap \Pi(T)$ holds.*

The following lemma is elementary yet crucial for the later analysis:

Lemma 3.0.3. *Let Φ be an irreducible root system of rank at least 2 and R a commutative ring with 1 and $G := G(\Phi, R)$ the corresponding split Chevalley group. Further, let S be a normally generating set of G. Then $\Pi(S) = \emptyset$ holds.*

Proof. Assume for contradiction that m is an element of $\Pi(S)$ and set $K := R/m$. Then the set S maps to a set of central elements $\bar{S}$ in $G(\Phi, K)$. However, the set S normally generates $G(\Phi, R)$ and so in particular, the subgroup $E(\Phi, R)$ is contained in the normal subgroup of $G(\Phi, R)$ generated by S. But $E(\Phi, R)$ maps onto $E(\Phi, K)$ and $G(\Phi, R/m) = E(\Phi, R/m)$ holds according to Proposition 5.2.2. This implies in particular that $\bar{S}$ normally generates $G(\Phi, K)$. So $\bar{S}$ is a subset of the center of $G(\Phi, K)$ and normally generates $G(\Phi, K)$. But this is only possible if $G(\Phi, K)$ is an abelian group to begin with. But this is impossible, as can be seen in a number of different ways: The center of $G(\Phi, K)$ is a subset of

$$H := \langle h_\phi(t) | t \in K - \{0\}, \phi \in \Phi \rangle.$$

according to [41, Chapter 3, p. 29, Lemma 28(d)], but on the other hand $H \cap U(\Phi, K) = \{1\}$ holds by [41, Chapter 3, p. 24, Lemma 21]. So $G(\Phi, K)$ being abelian and hence $H = G(\Phi, K)$ would imply $U(\Phi, K) = \{1\}$. But according to [41, Chapter 3, p. 27, Lemma 17] for $\phi \in \Phi$ arbitrary, the subgroup $\varepsilon_\phi(K)$ of $U(\Phi, K) = \{1\}$ is isomorphic to $(K, +)$. So the field K is trivial, which implies $R = m$ and contradicts the assumption that m is a maximal ideal in R. □

3.1 Strong boundedness of higher rank Chevalley groups

The main technical theorem used in this section is the following:

Theorem 3.1.1. *Let Φ be an irreducible root system of rank at least 2, which is not C_2 or G_2 and let R be a commutative ring with 1. Then there are constants $L(\Phi) \in \mathbb{N}$ (depending only on Φ) such that for $A \in G(\Phi, R)$ it holds that, there is an ideal $I(A)$ contained in $\varepsilon_s(A, L(\Phi))$ and with the property $V(I(A)) \subset \Pi(\{A\})$.*

The main theorem in this section is the following version of Theorem 1.2.1:

Theorem 3.1.2. *Let Φ be an irreducible root system that is not C_2, G_2 or A_1 and let R be a commutative ring with 1 such that $G(\Phi, R)$ is boundedly generated by root elements with $\|G(\Phi, R)\|_{\mathrm{EL}_Q} \leq Q(\Phi, R)$ for $\| \cdot \|_{\mathrm{EL}_Q}$ and $Q(\Phi, R)$ defined as in Definition 2.2.2. Further, let $L(\Phi) \in \mathbb{N}_0$ be the constant given by Theorem 3.1.1.*

1. *If Φ is simply-laced, then $\Delta_k(G(\Phi,R)) \leq Q(\Phi,R)L(\Phi)k$ holds for all $k \in \mathbb{N}$.*

2. *If Φ is not simply-laced, then $\Delta_k(G(\Phi,R)) \leq 3Q(\Phi,R)L(\Phi)k$ holds for all $k \in \mathbb{N}$.*

We show next:

Proposition 3.1.3. *Let Φ be any irreducible root system that is not G_2, C_2 or A_1, R a commutative ring with 1 and let S be a finite subset of $G := G(\Phi,R)$ with $\Pi(S) = \emptyset$ and let $L(\Phi)$ be as in Theorem 3.1.1.*

1. *If Φ is simply-laced, then $\|\varepsilon_\phi(a)\|_S \leq |S|L(\Phi)$ holds for all $a \in R$ and for all $\phi \in \Phi$.*

2. *If Φ is not simply-laced, then $\|\varepsilon_\phi(a)\|_S \leq 3|S|L(\Phi)$ holds for all $a \in R$ and for all $\phi \in \Phi$.*

Proof. Let $S = \{A_1, \ldots, A_k\}$ be given and for $l = 1, \ldots, k$, let $I(A_l)$ be the ideal from Theorem 3.1.1. Next, consider the ideal $I := I(A_1) + \cdots + I(A_k)$. As $I(A_l) \subset \varepsilon_s(A_l, L(\Phi))$ holds for all l, it is immediately clear that $\|\varepsilon_\phi(a)\|_S \leq |S|L(\Phi)$ holds for all $a \in I$ and $\phi \in \Phi$ short. But if I were not R, then there would be a maximal ideal m containing I. So according to Lemma 3.0.2, the ideal m would be contained in

$$V(I) = V(I(A_1)) \cap \cdots \cap V(I(A_n)) \subset \Pi(\{A_1\}) \cap \cdots \cap \Pi(\{A_n\}) = \Pi(S) = \emptyset.$$

Hence $I = R$ holds.

This proves the claim of the proposition for the simply-laced case and shows in the not simply-laced case that

$$\|\varepsilon_\phi(x)\|_S \leq L(\Phi)|S|$$

holds for any $x \in R$ and $\phi \in \Phi$ short. If there are long and short roots in Φ, then each long root is conjugate to a positive, simple long root ϕ in Φ and there is a short, positive, simple root ψ such that the set $\{\psi, \phi\}$ spans a root subsystem of Φ isomorphic to C_2. Further, according to the short root case, we know $\|\varepsilon_\psi(a)\|_S \leq |S|L(\Phi)$ for all $a \in R$ already. But $\psi + \phi$ is a short root as well and so we obtain $\|\varepsilon_\psi(1)\|_S, \|\varepsilon_{\psi+\phi}(a)\|_S \leq |S|L(\Phi)$ for all $a \in R$. Hence as

$$(\varepsilon_\psi(1), \varepsilon_\phi(a)) = \varepsilon_{\psi+\phi}(\pm a)\varepsilon_{2\psi+\phi}(\pm a),$$

holds, we obtain $\|\varepsilon_{2\psi+\phi}(a)\|_S \leq 3|S|L(\Phi)$ for all $a \in R$. The root $2\psi + \phi$ is long however and so we obtain the claim for Φ not simply-laced. □

Having this proposition, we obtain Theorem 3.1.2:

Proof. Let $S = \{A_1, \ldots, A_k\}$ be a normally generating set of $G(\Phi, R)$. Then $\Pi(S) = \emptyset$ holds according to Lemma 3.0.3. Hence according to Proposition 3.1.3, one has

$$\|\varepsilon_\phi(a)\|_S \leq |S|L(\Phi) \text{ if } \Phi \text{ is simply-laced and}$$
$$\|\varepsilon_\phi(a)\|_S \leq 3|S|L(\Phi) \text{ if } \Phi \text{ is not simply-laced}$$

for all $a \in R$ and all $\phi \in \Phi$. But according to assumption $G(\Phi, R)$ is boundedly generated by root elements with

$$\|G(\Phi, R)\|_{\mathrm{EL}_Q} \leq Q(\Phi, R).$$

This implies

$$\|G(\Phi, R)\|_S \leq \|\mathrm{EL}_Q\|_S \cdot \|G(\Phi, R)\|_{\mathrm{EL}_Q} \leq |S|L(\Phi) \cdot Q(\Phi, R)$$

if Φ is simply laced and

$$\|G(\Phi, R)\|_S \leq \|\mathrm{EL}_Q\|_S \cdot \|G(\Phi, R)\|_{\mathrm{EL}_Q} \leq 3|S|L(\Phi) \cdot Q(\Phi, R)$$

if not. This finishes the proof. □

Chapter 4 is mainly concerned with determining possible values of $L(\Phi)$ for $\Phi = C_n$ for $n \geq 3$ and $\Phi = E_6$ and Chapter 5 and Chapter 6 are concerned among other things with determining possible values for $Q(\Phi, R)$ in case of specific rings.

3.2 Strong boundedness for Sp_4 and G_2

The two main tools in this section are the following two technical theorems:

Theorem 3.2.1. *Let R be a commutative ring with 1 and let $A \in \mathrm{Sp}_4(R)$ be given. Then there is a constant $L(C_2)$ (not depending on A or R) such that, there is an ideal $I(A)$ with $V(I(A)) \subset \Pi(\{A\})$ and $2I(A) \subset \varepsilon(A, \phi, L(C_2))$ for all $\phi \in C_2$. More precisely, $2l(A)_2 \subset \varepsilon(A, \phi, L(C_2))$ holds for all $\phi \in C_2$.*

and

Theorem 3.2.2. *Let R be a commutative ring with 1 and let $A \in G_2(R)$ be given. Then there is a constant $L(G_2)$ (not depending on A or R) such that, there is an ideal $I(A)$ with $V(I(A)) \subset \Pi(\{A\})$ and $I(A) \subset \varepsilon_l(A, L(G_2))$. More precisely $l(A)_3 \subset \varepsilon_l(A, L(G_2))$ holds.*

We further need the following:

Lemma 3.2.3. *Let R be a commutative ring with $1, (R:2R) < \infty$ and $\Phi = C_2$ or G_2 such that $G := G(\Phi, R)$ is boundedly generated by root elements. Further, define*

$$\begin{aligned} Q_{C_2} &:= \{A\varepsilon_\phi(2x)A^{-1} |\ x \in R, \phi \in C_2, A \in \mathrm{Sp}_4(R)\} \text{ and} \\ Q_{G_2} &:= \{A\varepsilon_\phi(2x)A^{-1} |\ x \in R, \phi \in G_2 \text{ short}, A \in G_2(R)\} \\ &\quad \cup \{A\varepsilon_\phi(x)A^{-1} |\ x \in R, \phi \in G_2 \text{ long}, A \in G_2(R)\} \end{aligned}$$

and $N_\Phi := \langle Q_\Phi \rangle$ and let $\|\cdot\|_{Q_\Phi} : N_\Phi \to \mathbb{N}_0$ be the word norm on N_Φ defined by the set Q_Φ, that is $\|X\|_{Q_\Phi} := \min\{k \in \mathbb{N} | \exists g_1, \ldots, g_k \in Q_\Phi : X = g_1 \cdots g_k\}$ for $X \in N_\Phi - \{1\}$ and $\|1\|_{Q_\Phi} := 0$.

1. *Then the group G/N_Φ is finite.*
2. *Then there is a $K(\Phi, R) \in \mathbb{N}$ such that $\|N_\Phi\|_{Q_\Phi} \leq K(\Phi, R)$.*

Proof. First, we show that N_Φ has finite index in G to show the first claim of the lemma. The ideal $2R$ has finite index in R so let $X \subset R$ be a finite set of representatives of $2R$ in R. The group G is boundedly generated by root elements and so there is an $n := n(R)$ and roots $\alpha_1, \ldots, \alpha_n \in \Phi$ such that for all $A \in G$ there are $r_1, \ldots, r_n$ with

$$A = \prod_{i=1}^{n} \varepsilon_{\alpha_i}(r_i). \tag{3.1}$$

Next, choose for each i an element $a_i \in R$ and an $x_i \in X$ such that $r_i = 2a_i + x_i$. Note:

$$A = \prod_{i=1}^{n} \varepsilon_{\alpha_i}(r_i) = \varepsilon_{\alpha_1}(2a_1) \left[\prod_{i=2}^{n} \varepsilon_{\alpha_i}(2a_i)^{(\varepsilon_{\alpha_1}(x_1)\cdots\varepsilon_{\alpha_{i-1}}(x_{i-1}))} \right] \cdot \left[\prod_{i=1}^{n} \varepsilon_{\alpha_i}(x_i) \right] \tag{3.2}$$

Yet the first two factors at the right are elements of N_Φ and there are only finitely many possibilities for the third factor, so the first claim of the lemma follows. For the second claim, observe that (3.2) implies for $A \in N_\Phi$:

$$\|A\|_{Q_\Phi} \leq \sum_{i=1}^{n} \|\varepsilon_{\alpha_i}(2a_i)\|_{Q_\Phi} + \|\prod_{i=1}^{n} \varepsilon_{\alpha_i}(x_i)\|_{Q_\Phi} \leq n + \|\prod_{i=1}^{n} \varepsilon_{\alpha_i}(x_i)\|_{Q_\Phi}.$$

But again, there are only finitely many possibilities for $\|\prod_{i=1}^{n} \varepsilon_{\alpha_i}(x_i)\|_{Q_\Phi}$ and this proves the second claim. □

Remark 3.2.4. As the group $G/N_\Phi = G(\Phi, R)/N_\Phi$ is finite, it is uniformly bounded, that is the constant $\Delta_\infty(G(\Phi, R)/N_\Phi)$ is finite as well.

In this section, we will prove strong boundedness for Sp_4 and G_2:

Theorem 3.2.5. *Let R be a commutative ring with $1, (R:2R) < \infty$ and $\Phi = C_2$ or G_2 such that $G(\Phi, R)$ is boundedly generated by root elements. Additionally, define N_Φ, Q_Φ*

and $K(\Phi, R)$ as in Lemma 3.2.3 and let $L(C_2)$ be the constant given in Theorem 3.2.1 and $L(G_2)$ be the constant given in Theorem 3.2.2. Further, define the constants $j(C_2) := 1$ and $j(G_2) := 6$. Then

$$\Delta_k(G(\Phi, R)) \leq j(\Phi)L(\Phi)K(\Phi, R)k + \Delta_\infty(G(\Phi, R)/N_\Phi)$$

holds for all $k \in \mathbb{N}$.

3.2.1 The Sp_4-case

First, we obtain the version of Proposition 3.1.3 for $\mathrm{Sp}_4(R)$:

Proposition 3.2.6. *Let R be a commutative ring with 1 and let $S \subset \mathrm{Sp}_4(R)$ be a finite set with $\Pi(S) = \emptyset$. Let $L(C_2)$ be as given in Theorem 3.2.1. Then $\|\varepsilon_\phi(2a)\|_S \leq |S|L(C_2)$ holds for all $a \in R$ and all $\phi \in C_2$. In particular, $\|q\|_S \leq |S|L(C_2)$ holds for all $q \in Q_{C_2}$.*

Proof. Let $S = \{A_1, \ldots, A_k\}$ be given and let $2I(A_l)$ be the ideal and $L(C_2)$ the constant from Theorem 3.2.1 for all $l = 1, \ldots, k$. Consider the ideal $I := I(A_1) + \cdots + I(A_k)$. As $2I(A_l) \subset \varepsilon(A_l, \phi, L(C_2))$ holds for all l and all $\phi \in C_2$, it is immediately clear that $\|\varepsilon_\phi(2a)\|_S \leq |S|L(C_2)$ holds for all $a \in I$. Thus to conclude $\|\varepsilon_\phi(2a)\|_S \leq |S|L(C_2)$ for all $a \in R$ and all $\phi \in C_2$, it suffices to show that $I = R$. But this follows from $\Pi(S) = \emptyset$ in the same manner as in the proof of Proposition 3.1.3.

The claim about elements of Q_{C_2} follows immediately now: Any $q \in Q_{C_2}$ has the form $q = A\varepsilon_\phi(2x)A^{-1}$ for $x \in R, \phi \in C_2$ and $A \in \mathrm{Sp}_4(R)$. Thus $\|q\|_S = \|A\varepsilon_\phi(2x)A^{-1}\|_S = \|\varepsilon_\phi(2x)\|_S \leq |S|L(C_2)$ and this finishes the proof. □

We can prove Theorem 3.2.5 for $\mathrm{Sp}_4(R)$ now:

Proof. Let S be a finite normally generating set of $\mathrm{Sp}_4(R)$. Set $G := \mathrm{Sp}_4(R)$ and recall

$$Q_{C_2} := \{A\varepsilon_\phi(2x)A^{-1} |\ x \in R, \phi \in C_2, A \in \mathrm{Sp}_4(R)\}.$$

and $N_{C_2} := \langle Q_{C_2} \rangle$. Next, set

$$E(G/N_{C_2}) := \{T \subset G/N_{C_2} |\ T \text{ normally generates } G/N_{C_2}\}$$

Let $\pi : G \to G/N_{C_2}$ be the quotient map. The quotient G/N_{C_2} is finite according to Lemma 3.2.3(1) and so we can define $M(R, C_2) := M := \Delta_\infty(G/N_{C_2}) \in \mathbb{N}_0$. So for all $T \subset G$ finite with $\pi(T) \in E(G/N_{C_2})$:

$$\forall g \in G : \exists t_1, \ldots, t_M \in T \cup T^{-1} \cup \{1\} : \exists A_1, \ldots, A_M \in G :$$
$$\pi(g) = \pi(\prod_{i=1}^{M} A_i t_i A_i^{-1}).$$

Hence $g(\prod_{i=1}^{M} A_i t_i A_i^{-1})^{-1} \in N_{C_2}$ holds and so

$$\|g\|_S \leq \|\prod_{i=1}^{M} A_i t_i A_i^{-1}\|_S + \|N_{C_2}\|_S \leq M \max\{\|t\|_S |\ t \in T\} + \|N_{C_2}\|_S.$$

This implies

$$\|\mathrm{Sp}_4(R)\|_S \leq M \max\{\|t\|_S |\ t \in T\} + \|N_{C_2}\|_S$$

for all $T \subset G$ finite with $\pi(T) \in E(G/N_{C_2})$.

Next, note that S itself satisfies $\pi(S) \in E(G/N_{C_2})$ and one clearly has $\|A\|_S \leq 1$ for $A \in S$. Thus

$$\|\mathrm{Sp}_4(R)\|_S \leq M + \|N_{C_2}\|_S = M(C_2, R) + \|N_{C_2}\|_S. \tag{3.3}$$

But according to Lemma 3.0.3, one has $\Pi(S) = \emptyset$. Thus according to Proposition 3.2.6, for all $q \in Q_{C_2}$, one has $\|q\|_S \leq L(C_2)|S|$ for the $L(C_2)$ from Theorem 3.2.1. Further, by Lemma 3.2.3(2), there is a $K(C_2, R) \in \mathbb{N}$ such that $\|N_{C_2}\|_{Q_{C_2}} \leq K(C_2, R)$. This implies $\|N_{C_2}\|_S \leq K(C_2, R)L(C_2)|S|$ and this finishes the proof together with (3.3). □

3.2.2 The G_2-case

Remember that the positive roots of G_2 are $\alpha, \beta, \alpha+\beta, 2\alpha+\beta, 3\alpha+\beta, 3\alpha+2\beta$ with α short and simple and β long and simple. Further, recall that the root subsystem spanned by β and $3\alpha+\beta$ is isomorphic to A_2. First, we give the version of Proposition 3.1.3 for G_2 :

Proposition 3.2.7. *Let R be a commutative ring with 1 and let S be a finite subset of $G_2(R)$ with $\Pi(S) = \emptyset$ and let $L(G_2)$ be chosen as in Theorem 3.2.2. Then for all $a \in R$:*

1. *$\|\varepsilon_\phi(a)\|_S \leq L(G_2)|S|$ holds for all $\phi \in G_2$ long.*
2. *$\|\varepsilon_\phi(2a)\|_S \leq 6L(G_2)|S|$ holds for all $\phi \in G_2$ short.*

In particular, $\|q\|_S \leq 6L(G_2)|S|$ holds for any $q \in Q_{G_2}$.

Proof. Note that according to Theorem 3.2.2, one has $I \subset \varepsilon_l(S, L(G_2)|S|)$ for the ideal $I := \sum_{A \in S} I(A)$. As in the proof of Proposition 3.1.3, $\Pi(S) = \emptyset$ implies $I = R$. This yields the claim of the proposition for long roots. To get the claim for short roots, observe first that $\varepsilon_\beta(a) \in B_S(L(G_2)|S|)$ holds for all $a \in R$, because β is long. This implies

$$B_S(2L(G_2)|S|) \ni (\varepsilon_\beta(a), \varepsilon_\alpha(1)) = \varepsilon_{\alpha+\beta}(\pm a)\varepsilon_{2\alpha+\beta}(\pm a)\varepsilon_{3\alpha+\beta}(\pm a)\varepsilon_{3\alpha+2\beta}(\pm a^2). \tag{3.4}$$

However, $\varepsilon_{3\alpha+\beta}(\pm a)\varepsilon_{3\alpha+2\beta}(\pm a^2)$ commutes with $\varepsilon_\alpha(1)$ and hence we obtain from equation

(3.4) that:

$$\begin{aligned}
&B_S(4L(G_2)|S|) \ni \big(\varepsilon_{\alpha+\beta}(\pm a)\varepsilon_{2\alpha+\beta}(\pm a)\varepsilon_{3\alpha+\beta}(\pm a)\varepsilon_{3\alpha+2\beta}(\pm a^2), \varepsilon_\alpha(1)\big)\\
&= (\varepsilon_{\alpha+\beta}(\pm a)\varepsilon_{2\alpha+\beta}(\pm a), \varepsilon_\alpha(1))\\
&\sim (\varepsilon_{2\alpha+\beta}(\pm a), \varepsilon_\alpha(1)) \cdot (\varepsilon_\alpha(1), \varepsilon_{\alpha+\beta}(\pm a))\\
&= \varepsilon_{3\alpha+\beta}(\pm 3a) \cdot \varepsilon_{2\alpha+\beta}(\pm 2a)\varepsilon_{3\alpha+\beta}(\pm 3a)\varepsilon_{3\alpha+2\beta}(\pm 3a^2)\\
&= \varepsilon_{2\alpha+\beta}(\pm 2a)\varepsilon_{3\alpha+\beta}(\pm 3a \pm 3a)\varepsilon_{3\alpha+2\beta}(\pm 3a^2).
\end{aligned}$$

Yet $\varepsilon_{3\alpha+\beta}(\pm 3a \pm 3a)$ and $\varepsilon_{3\alpha+2\beta}(\pm 3a^2)$ are both elements of $B_S(L(G_2)|S|)$ by assumption and hence

$$\varepsilon_{2\alpha+\beta}(\pm 2a) \in B_S(6L(G_2)|S|)$$

holds. But $2\alpha+\beta$ is a short root and hence as $a \in R$ is arbitrary, the claim for short roots follows as well. The claim for elements of Q_{G_2} follows, as in the proof of Proposition 3.2.6, from the conjugation-invariance of $\|\cdot\|_S$. □

Using this proposition, one can prove Theorem 3.2.5 for $G_2(R)$, but the proof is essentially the same as the one for $\mathrm{Sp}_4(R)$, so we are going to omit it.

Determining the value of $\Delta_\infty(G/N_\Phi)$ for $\Phi = C_2$ or G_2 is very similar to determining the so called covering number of the group G/N_Φ. This is a classical problem in the theory of finite groups and we talk about this to some extent in Chapter 6. Determining $K(\Phi, R)$ on the other hand is more difficult and we show how to do it for a special case of R in Chapter 6 as well. It is a problem related to the congruence subgroup property.

Lastly note the following corollary of the previous proofs:

Corollary 3.2.8. *Let R be a commutative ring with 1, Φ irreducible and of rank at least 2 and assume $G(\Phi, R) = E(\Phi, R) =: G$. Then a subset S of G normally generates G precisely if*

1. *one has $\Pi(S) = \emptyset$ in case $\Phi \neq C_2, G_2$.*
2. *one has $\Pi(S) = \emptyset$ and S maps to a normally generating set of G/N_Φ for N_Φ as in Lemma 3.2.3 in case $\Phi = C_2$ or G_2.*

Proof. First, if S normally generates $G(\Phi, R)$, then Lemma 3.0.3 implies $\Pi(S) = \emptyset$, so this condition is always necessary independent of Φ. However, if $\Phi = C_2$ or G_2, then it is obvious that if S normally generates $G(\Phi, R)$, it must also normally generate its quotient $G(\Phi, R)/N_\Phi$. This proves that the conditions named in the corollary are necessary.

On the other hand, assume first that $\Phi \neq C_2$ or G_2 and that $\Pi(S) = \emptyset$ holds. As $G(\Phi, R) = E(\Phi, R)$ holds by assumption, it suffices to prove that $G(\Phi, R)$ contains all root elements. This however is a direct consequence of Proposition 3.1.3. In case $\Phi = C_2$ or G_2, the crucial point is that $\Pi(S) = \emptyset$ implies $N_\Phi \subset \langle\langle S \rangle\rangle$ as shown by Proposition 3.2.6 and

Proposition 3.2.7. Hence if S maps to a normally generating set of G/N_Φ for $G = G_2(R)$ or $Sp_4(R)$, then S must normally generate G. □

3.3 Boundedness of root elements in higher rank Chevalley-groups

In this section, we prove Theorem 3.1.1. The main tool is the following theorem by Abe:

Theorem 3.3.1. *[2, Theorem 1,2,3] Let Φ be an irreducible root system that is not A_1, C_2 or G_2 and let R be a commutative ring with 1. Then for each subgroup $H \subset G(\Phi, R)$ normalized by the group $E(\Phi, R)$, there is an ideal $J \subset R$ and an additive subgroup L of J such that $\bar{E}(J, L) \subset H \subset E^*(J, L)$.*

Remark 3.3.2.

1. The paper [43] by Vaserstein deals with similar statements in the simply laced case and with the multiple laced case under some assumptions. The papers Abe, Suzuki [3] and Abe [1] deal with local rings.

2. The proof of Theorem 3.3.1 is quite complicated and requires a careful reduction to the case of R being local. However, under the assumption that R satisfies a stable range condition (see Chapter 5), results of the form of Theorem 3.3.1 are much more readily provable. For example, Bass earlier result [5, Theorem 4.2(e)] shows a similar description of normal subgroups of $\mathrm{SL}_n(R)$ by using much more elementary methods in case R satisfies a stable range condition.

3. Theorem 3.3.1 is enough to prove strong boundedness of $G(\Phi, R)$ for commutative rings with 1 and $\Phi \neq A_1, C_2, G_2$ with $G(\Phi, R)$ boundedly generated by root elements. However, this would not yield any linear bounds on Δ_k and is very similar to our argument, so we do not give more details.

We further need the following lemma about root elements:

Lemma 3.3.3. *Let Φ an irreducible root system that is not A_1, C_2 or G_2, R a commutative ring with 1 and $A \in G(\Phi, R)$ be given and assume that $\lambda \in \varepsilon_s(A, N)$ for some $N \in \mathbb{N}$. Then*

$$\lambda R \subset \varepsilon_s(A, 8N)$$

holds.

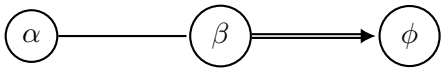
Proof. First note that $\lambda \in \varepsilon_s(A, N)$ is equivalent to $\varepsilon_\phi(\lambda) \in B_A(N)$ for any short root $\phi \in \Phi$. We distinguish two cases:

1. $\Phi \neq B_n$. Note that ϕ is a short root in Φ and that all of these root systems contain a root subsystem isomorphic to A_2 consisting of short roots. Hence after conjugating with a suitable Weyl group element, we can assume that $\Phi = A_2$ with simple positive roots α, β and $\phi = \alpha + \beta$. But α is also short and this implies $\varepsilon_\alpha(\lambda) \in B_A(N)$. For $x \in R$ arbitrary, we obtain further

$$\varepsilon_\phi(\pm x\lambda) = (\varepsilon_\alpha(\pm\lambda), \varepsilon_\beta(\pm x)) \in B_A(2N).$$

This yields the claim for $\Phi \neq B_n$.

2. $\Phi = B_n$ for $n \geq 3$. After conjugation with Weyl group elements, we assume that $n = 3$ and so there are positive, simple roots α, β, ϕ with the Dynkin-diagram corresponding to the system of positive, simple roots $\{\alpha, \beta, \phi\}$ looking as follows:



Then

$$B_A(2N) \ni (\varepsilon_\phi(\lambda), \varepsilon_\beta(x)) = \varepsilon_{\beta+\phi}(\pm x\lambda)\varepsilon_{\beta+2\phi}(\pm x\lambda^2) \tag{3.5}$$

holds for $x \in R$ arbitrary. The root $\beta+\phi$ is short however and so we have $\varepsilon_{\beta+\phi}(\lambda) \in B_A(N)$ as well. Thus for $x = 1$ we obtain $\varepsilon_{\beta+2\phi}(\lambda^2) \in B_A(3N)$ from (3.5). The root $\beta + 2\phi$ is long and hence $\varepsilon_{\beta+2\phi}(\lambda^2)$ is (up to sign) conjugate to $\varepsilon_\beta(\lambda^2)$ and so $\varepsilon_\beta(\lambda^2) \in B_A(3N)$. Yet α, β are simple roots in a root subsystem of B_3 isomorphic to A_2 and hence we obtain from the first case, that $\varepsilon_\beta(x\lambda^2) \in B_A(6N)$ holds for all $x \in R$. Summarizing this with equation (3.5) we get $\varepsilon_{\beta+\phi}(x\lambda) \in B_A(8N)$ for all $x \in R$. Hence after conjugation we are done with the case $\Phi = B_n$ as well.

□

Remark 3.3.4. This Lemma is a more quantitative version of Vasersteins [43, Theorem 4(a)].

We can prove Theorem 3.1.1 now:

Proof. First, let $n_1, \ldots, n_u$ be given as the dimensions of the representations involved in defining ρ as in Section 2.3 and set $n := n_1 + \cdots + n_u$. Also choose the polynomials P in $\mathbb{Z}[y_{ij}]$ characterizing elements of $G(\Phi, \cdot)$ that is

$$G(\Phi, R) = \{X \in R^{n\times n} | P(X) = 0\}$$

holds for any commutative ring R with 1. Further, let $1 \leq k, l \leq n$ be given with not both k and l equal to n and let $\phi \in \Phi$ be a short root.

Next, let a language $\mathcal{L}$ with the relation symbols, constants and function symbols

$$(\mathcal{R}, 0, 1, +, \times, (a_{i,j})_{1\leq i,j\leq n}, (e(k,l,v))_{v\in\mathbb{N}}, \cdot^{-1})$$

be given. Here $0, 1$ are constant symbols, $(a_{i,j})_{1\leq i,j\leq n}$ is a matrix of constant symbols and $(e(k,l,v))_{v\in\mathbb{N}}$ is an infinite sequence of constant symbols. Further, $\cdot^{-1} : \mathcal{R}^{n\times n} \to \mathcal{R}^{n\times n}$ is a function symbol and we often use the notation $X^{-1} := \cdot^{-1}(X)$ for X an $n \times n$-matrix of variable and constant symbols. Also the symbol $\mathcal{A}$ denotes the $n \times n$-matrix of constants $(a_{i,j})$ and X commonly refers to $n\times n$-matrices of variable symbols. Next, we define a first order theory $\mathcal{T}_{kl}$ in the language $\mathcal{L}$. This first order-theory $\mathcal{T}_{kl}$ contains formulas regarding $n \times n$-matrices with entries being variables or constant symbols in $\mathcal{L}$, but these formulas can always be rephrased into a conjunction of formulas about variables or constants in $\mathcal{L}$. Further, the formulas in $\mathcal{T}_{kl}$ also involve multiplication and conjugation of $n \times n$-matrices of constants and variables in the language $\mathcal{L}$. However, matrix multiplication can be phrased in terms of $\mathcal{L}$ as it is defined entry-wise: For example the (i, j)-entry of $Z := (z_{ij}) := X \cdot Y$ is defined as $z_{ij} := \sum_{t=1}^{n} x_{it}y_{tj}$ for $1 \leq i, j \leq n$ and $X = (x_{ij})$ and $Y = (y_{ij})$.

The theory $\mathcal{T}_{kl}$ is defined to contain the following sentences:

1. Sentences forcing the universe $R := \mathcal{R}^{\mathcal{M}}$ of each model $\mathcal{M}$ of $\mathcal{T}_{kl}$ to be a commutative ring with respect to the functions $+^{\mathcal{M}}, \times^{\mathcal{M}}$ and with $0^{\mathcal{M}}, 1^{\mathcal{M}}$ being 0 and 1 in the ring R.

2. For all $v \in \mathbb{N}$: If $k \neq l$ the sentence $e(k, l, v) = a_{k,l}^{v}$ should be included in $\mathcal{T}_{kl}$. If on the other hand $k = l$, then choose the smallest $w \in \{1, \ldots, u\}$ with $k \leq n_1 + \cdots + n_w$ and include the sentence $e(k, l, v) = (a_{k,k} - a_{n_1+\cdots+n_w})^{v}$.

3. The sentence $P(\mathcal{A}) = 0$.

4. The sentence $\forall X : (P(X) = 0) \to (X \cdot X^{-1} = I_n)$, where I_n denotes the unit matrix in $\mathcal{R}^{n\times n}$ with entries the constant symbols $0, 1$ as appropriate.

5. A family of sentences $(\theta_r)_{r\in\mathbb{N}}$ as follows:

$$\theta_r : \bigwedge_{1\leq v\leq r} \forall X_1^{(v)}, \ldots, X_r^{(v)}, \forall e_1^{(v)}, \ldots, e_r^{(v)} \in \{0, 1, -1\} :$$
$$\left[\left(P\left(X_1^{(v)}\right) = \cdots = P\left(X_r^{(v)}\right) = 0\right) \to \left(\varepsilon_\phi(e(k,l,v)) \neq (\mathcal{A}^{e_1^{(v)}})^{X_1^{(v)}} \cdots (\mathcal{A}^{e_r^{(v)}})^{X_r^{(v)}}\right)\right]$$

Here $\mathcal{A}^1 := \mathcal{A}, \mathcal{A}^{-1} := \mathcal{A}^{-1}$ and $\mathcal{A}^0 := I_n$. Also remember as mentioned in Section 2.2, that $\varepsilon_\phi(T)$ for T a variable, is a $n\times n$-matrix, whose entries are polynomials in $\mathbb{Z}[T]$, so in particular the last collection of sentences (θ_r) are in fact first-order sentences in the language $\mathcal{L}$.

We first show that the theory $\mathcal{T}_{kl}$ is inconsistent. To this end, let $\mathcal{M}$ be a model for the sentences in (1) through (4) and let $R := R^{\mathcal{M}}$ be the universe of $\mathcal{M}$. The sentences in (1) enforce that R is a commutative ring with $1 = 1^{\mathcal{M}}$ and $0 = 0^{\mathcal{M}}$ and (3) enforces that the matrix $A := (a_{i,j}^{\mathcal{M}}) \in R^{n\times n}$ is an element of the Chevalley group $G(\Phi, R)$. Let H be the subgroup of $G(\Phi, R)$ normally generated by A. According to Theorem 3.3.1 there is a pair (J, L) such that

$$\bar{E}(J, L) \subset H \subset E^*(J, L).$$

As $L \subset J$ holds, $A \in E^*(J, L)$ implies that $\pi_J(A)$ centralizes $E(R/J)$ and consequently that $\pi_{\sqrt{J}}(A)$ centralizes $E(R/\sqrt{J})$. The ring $R/\sqrt{J}$ is reduced and so $\pi_{\sqrt{J}}(A)$ has the form described in Lemma 2.3.1. According to Remark 2.3.4(3), this implies that $l(A) \subset \sqrt{J}$. Hence as $\bar{E}(J, L) \subset H$, there is a constant $r' \in \mathbb{N}$ such that $\varepsilon_\phi(e(k, l, v)^{\mathcal{M}}) \in B_A(r')$ holds for some $v \leq r'$. But this contradicts the statement $\theta_{r'}^{\mathcal{M}}$.

So summarizing: a model of the sentences in (1) through (4) cannot be a model of all of the sentences θ_r. Hence there is in fact no model of all of the above sentences and hence $\mathcal{T}_{kl}$ is inconsistent. Gödel's Compactness Theorem [37, Theorem 3.2] implies then, that a certain finite subset $\mathcal{T}_{kl}^0 \subset \mathcal{T}_{kl}$ is already inconsistent. But then only a finite collection of the θ_r is contained in $\mathcal{T}_{kl}^0$. So let $L_{kl}(\Phi) \in \mathbb{N}$ be the largest $r \in \mathbb{N}$ with $\theta_r \in \mathcal{T}_{kl}^0$. Observe further, that for all $r \in \mathbb{N}$, we have $\{(1) - (4), \theta_{r+1}\} \vdash \theta_r$. Hence the subset $\mathcal{T}_{kl}^1 \subset \mathcal{T}_{kl}$ that contains all sentences in (1) through (4) and the *single* sentence $\theta_{L_{kl}(\Phi)}$, must be inconsistent as well.

Let R be an arbitrary commutative ring with 1 and let $A \in G(\Phi, R)$ be given. This gives us a model $\mathcal{M}$ of the sentences in (1) through (4) and hence as $\mathcal{T}_{kl}^1$ is inconsistent, this model must violate the statement $\theta_{L_{kl}(\Phi)}^{\mathcal{M}}$. Thus there are elements $g_1, \ldots, g_{L_{kl}(\Phi)} \in G(\Phi, R)$ and $e_1, \ldots, e_{L_{kl}(\Phi)} \in \{0, 1, -1\}$ as well as a natural number $v \leq L_{kl}(\Phi)$ such that

$$\varepsilon_\phi(e(k, l, v)^{\mathcal{M}}) = (A^{e_1})^{g_1} \cdots (A^{e_{L_{kl}(\Phi)}})^{g_{L_{kl}(\Phi)}}.$$

Hence we obtain that either a power of a_{kl} (in case $k \neq l$) or a power of $a_{kk} - a_{n_1+\cdots+n_w, n_1+\cdots+n_w}$ (in case $k = l$) is an element of $\varepsilon(A, \phi, L_{kl}(\Phi))$. So setting

$$L(\Phi) := \sum_{1\leq k,l\leq n \text{ not both } k,l=n} 8L_{kl}(\Phi),$$

we get together with Lemma 3.3.3 an ideal $I(A)$ in R such that $I(A) \subset \varepsilon(A, \phi, L(\Phi))$ and $l(A) \subset \sqrt{I(A)}$ holds. But if m is a maximal ideal containing $I(A)$, then it contains $l(A)$ and so $\pi_m(A)$ is central in $G(\Phi, R/m)$ according to Lemma 2.3.1. Hence $V(I(A)) \subset \Pi(\{A\})$ holds. So the ideal $I(A)$ has the desired properties for the single root ϕ. But $\varepsilon(A, \phi, L(\Phi)) = \varepsilon_s(A, L(\Phi))$ holds and so the theorem is proven. □

3.4 Boundedness of root elements in $\mathrm{Sp}_4(R)$

In this section, we prove Theorem 3.2.1. Recall that the positive roots in C_2 are $\alpha, \beta, \alpha+\beta$ and $2\alpha+\beta$ with α short, positive and simple and β long, positive and simple. The main ingredient is the following observation due to Costa and Keller instead of Theorem 3.3.1:

Theorem 3.4.1. *[14, Theorem 2.6, 4.2, 5.1, 5.2] Let R be a commutative ring with 1 and let $A \in \mathrm{Sp}_4(R)$ be given. Then for all $x \in l(A)$, one has $\varepsilon_{2\alpha+\beta}(2x+x^2)\varepsilon_{\alpha+\beta}(x^2) \in \langle\langle A\rangle\rangle_{E(C_2,R)}$ with $\langle\langle A\rangle\rangle_{E(C_2,R)}$ denoting the subgroup of $\mathrm{Sp}_4(R)$ generated by the $E(C_2,R)$-conjugates of A.*

Root elements in Sp_4 are more complicated than in higher rank groups:

Lemma 3.4.2. *Let R be a commutative ring with 1 and $S \subset \mathrm{Sp}_4(R)$. Let $\lambda \in R$ and $N \in \mathbb{N}$ be given. Then*

1. *$\varepsilon_{2\alpha+\beta}(2\lambda+\lambda^2)\varepsilon_{\alpha+\beta}(\lambda^2) \in B_S(N)$ implies $\{\varepsilon_{2\alpha+\beta}(2x\lambda^2) | x \in R\} \subset B_S(2N)$.*
2. *$\varepsilon_{2\alpha+\beta}(\lambda) \in B_S(N)$ implies $\varepsilon_\phi(\lambda) \in B_S(3N)$ for all ϕ short in C_2.*
3. *$\varepsilon_\alpha(x\lambda) \in B_S(N)$ for all $x \in R$ implies $\varepsilon_\phi(x\lambda^2) \in B_S(3N)$ for all $\phi \in C_2$ long and all $x \in R$.*
4. *$\varepsilon_{2\alpha+\beta}(\lambda) \in B_S(N)$ implies $\{\varepsilon_{2\alpha+\beta}(2x\lambda) | x \in R\} \subset B_S(6N)$.*
5. *$\varepsilon_{2\alpha+\beta}(2\lambda+\lambda^2)\varepsilon_{\alpha+\beta}(\lambda^2) \in B_S(N)$ implies $\{\varepsilon_\phi(2x\lambda^2) | x \in R, \phi \in C_2\} \subset B_S(6N)$.*

All of the above implications stay true, if the balls B_S are replaced by a normal subgroup of $\mathrm{Sp}_4(R)$.

Proof. For the first claim inspect the commutator

$$\varepsilon_{2\alpha+\beta}(\pm 2x\lambda^2) = (\varepsilon_\alpha(x), \varepsilon_{2\alpha+\beta}(2\lambda+\lambda^2)\varepsilon_{\alpha+\beta}(\lambda^2))$$

for $x \in R$ arbitrary. For the second claim, note that $\varepsilon_{2\alpha+\beta}(\lambda)$ is conjugate to $\varepsilon_\beta(\lambda)$ and so $\varepsilon_\beta(\lambda) \in B_S(N)$. Note further

$$B_A(2N) \ni (\varepsilon_\beta(\lambda), \varepsilon_\alpha(1)) = \varepsilon_{\alpha+\beta}(\pm\lambda)\varepsilon_{2\alpha+\beta}(\pm\lambda).$$

These two facts imply $\varepsilon_{\alpha+\beta}(\lambda) \in B_S(3N)$. The element $\varepsilon_{\alpha+\beta}(\lambda)$ is conjugate to $\varepsilon_\phi(\lambda)$ for every short root $\phi \in C_2$. This proves the second claim of the lemma and the third claim follows by considering for $x \in R$ the commutator

$$B_A(2N) \ni (\varepsilon_\beta(x), \varepsilon_\alpha(\lambda)) = \varepsilon_{\alpha+\beta}(\pm x\lambda)\varepsilon_{2\alpha+\beta}(\pm x\lambda^2).$$

and noting $\varepsilon_{\alpha+\beta}(\pm x\lambda) \in B_A(N)$. For the fourth claim, note that we have by the second claim, that $\varepsilon_\alpha(\lambda) \in B_S(3N)$. Next inspect for $x \in R$ the commutator:

$$B_S(6N) \ni (\varepsilon_\alpha(\lambda), \varepsilon_{\alpha+\beta}(x)) = \varepsilon_{2\alpha+\beta}(\pm 2x\lambda).$$

This proves the fourth claim. The last claim follows from part (1) and (2). □

This enables us to prove Theorem 3.2.1:

Proof. The proof is very similar to the one of Theorem 3.1.1. Recall that Sp_4 is a subgroup scheme of GL_4 as seen in Section 2.3. First, let natural numbers k, l be given with $1 \leq k, l \leq 4$. Also if $k = l$, then we assume that $k = l < 4$. Further, let $P \subset \mathbb{Z}[y_{ij}]$ be a finite collection of polynomials describing membership in Sp_4. The language $\mathcal{L}$ and the theory $\mathcal{T}_{kl}$ is defined the same way as in the proof of Theorem 3.1.1 except for four differences: First, we include a constant symbol $e(k, l)$ instead of $e(k, l, v)$. Secondly, the sentence in (2) has the form

$$e(k,l) = \begin{cases} a_{kl}, & \text{if } k \neq l \\ a_{kk} - a_{k+1,k+1}, & \text{if not} \end{cases}$$

Third, the sentence in item (3) describes that for each model $\mathcal{M}$ the matrix $\mathcal{A}^{\mathcal{M}}$ is an element of $\mathrm{Sp}_4(\mathcal{R}^{\mathcal{M}})$. Fourth and most importantly, the sentences in (5) are a family of sentences $(\theta_r)_{r\in\mathbb{N}}$ such that

$$\begin{aligned} \theta_r :&\forall X_1, \ldots, X_r, \forall e_1, \ldots e_r \in \{0, 1, -1\} : (P(X_1) \wedge \cdots \wedge P(X_r)) \to \\ &\big(\varepsilon_{2\alpha+\beta}(2e(k,l) + e(k,l)^2)\varepsilon_{\alpha+\beta}(e(k,l)^2) \neq (\mathcal{A}^{e_1})^{X_1} \cdots (\mathcal{A}^{e_r})^{X_r}\big) \end{aligned}$$

Invoking Theorem 3.4.1 instead of Theorem 3.3.1 yields that a model of (1) through (4) cannot be a model of all sentences in (5). Hence $\mathcal{T}_{kl}$ is inconsistent. Using Gödel's compactness theorem [37, Theorem 3.2], we obtain, as in the proof of Theorem 3.1.1, that there is an $L_{k,l}(C_2) \in \mathbb{N}$ such that the subset $\mathcal{T}^1_{kl} \subset \mathcal{T}_{kl}$ that contains all sentences in (1) through (4) and the *single* sentence $\theta_{L_{k,l}(C_2)}$ is already inconsistent.

Let R be an arbitrary commutative ring with 1 and let $A \in \mathrm{Sp}_4(R)$ be given. This gives us a model $\mathcal{M}$ of the sentences in (1) through (4) and hence as $\mathcal{T}^1_{kl}$ is inconsistent this model must violate the statement $\theta^{\mathcal{M}}_{L_{k,l}(C_2)}$. Thus there are elements $g_1, \ldots, g_{L_{k,l}(C_2)} \in \mathrm{Sp}_4(R)$ and $e_1, \ldots e_{L_{k,l}(C_2)} \in \{0, 1, -1\}$ such that (abusing the notation slightly)

$$\varepsilon_{2\alpha+\beta}(2e(k,l) + e(k,l)^2)\varepsilon_{\alpha+\beta}(e(k,l)^2) = (A^{e_1})^{g_1} \cdots (A^{e_{L_{k,l}(C_2)}})^{g_{L_{k,l}(C_2)}}$$

Next, Lemma 3.4.2(5) implies $2(e(k,l)^2) \in \varepsilon(A, \phi, 6L_{k,l}(C_2))$ for all $\phi \in C_2$. If we sum

over all admissible k, l, this implies for all $\phi \in C_2$ that

$$2l(A)_2 = \sum_{k,l}(2e(k,l)^2) \subset \varepsilon(A, \phi, \sum_{k,l} 6L_{k,l}(C_2)).$$

So define $L(C_2) := \sum_{k,l} 6L_{k,l}(C_2)$ and then $l(A)_2$ has the desired property of $V(l(A)_2) \subset \Pi(\{A\})$: If m is a maximal ideal containing $l(A)_2$, then it contains $l(A)$ and so A maps to a scalar matrix in $\mathrm{Sp}_4(R/m)$, which is central in $\mathrm{Sp}_4(R/m)$. □

3.5 Boundedness of root elements in $G_2(R)$

In this section, we prove Theorem 3.2.2. This will be shown by using:

Theorem 3.5.1. *[15, (3.6) Main Theorem] Let R be a commutative ring with 1 and let H be an $E(G_2, R)$-normalized subgroup of $G_2(R)$. Then there is a pair of ideals J, J' in R with*

$$(x^3, 3x | x \in J) \subset J' \subset J$$

such that

$$[E(R), E(J, J')] \subset H \subset G(J, J').$$

Remark 3.5.2. We are not defining $G(J, J')$, but note that $H \subset G(J, J')$ implies that $\pi_J(H) = \{1\}$, if $J \neq R$ holds.

This implies:

Corollary 3.5.3. *Let R be a commutative ring with 1, $A \in G_2(R)$ and H the smallest subgroup of $G_2(R)$ normalized by $E(G_2, R)$ and containing A. Then $\varepsilon_{3\alpha+2\beta}(a^3), \varepsilon_{3\alpha+2\beta}(3a)$ are elements of H for all $a \in l(A)$.*

Proof. This follows directly from Theorem 3.5.1. Assume first, that the first ideal J from Theorem 3.5.1 is not R. Then J must contain $l(A)$. This is the case because $A \in H$ becomes scalar after reducing modulo J. Hence for $a \in l(A)$, we get that $3a, a^3$ are elements of the second ideal J' from Theorem 3.5.1. Lastly, $\{\varepsilon_\beta(b) |\ b \in J'\} \subset H$ holds, because β is a root in the long A_2 in G_2 and this finishes the proof if $J \neq R$. If J however is R, then $J' = R$ holds as well and so $\varepsilon_\beta(x^3), \varepsilon_\beta(3x) \in H$ holds for all $x \in R$, not just for $x \in l(A)$. □

Next, note the following:

Proposition 3.5.4. *Let R be a commutative ring with 1 and let $S \subset G_2(R)$ be given. Then*

1. *if for $N \in \mathbb{N}, \lambda \in R$ one has $\varepsilon_{3\alpha+2\beta}(\lambda) \in B_S(N)$, then*

(a) $\{\varepsilon_\phi(x\lambda)|x \in R\} \subset B_S(2N)$ *for* ϕ *long and*

(b) $\{\varepsilon_\phi(2x\lambda)|x \in R\} \subset B_S(8N)$ *for* ϕ *short hold.*

2. *if* $\varepsilon_\alpha(\lambda) \in B_S(N)$, *then* $\{\varepsilon_{3\alpha+2\beta}(x\lambda^3)|x \in R\} \subset B_S(4N)$ *holds.*

The implications are still true, if the balls B_S *are replaced by a normal subgroup of* $G_2(R)$.

Proof. Part (1a) can be obtained by arguing as in the proof of Lemma 3.3.3 using that the root subsystem of G_2 formed by long roots is isomorphic to A_2. For part (1b) inspect the following commutator formula for all $x \in R$:

$$\varepsilon_{\alpha+\beta}(\pm x\lambda)\varepsilon_{2\alpha+\beta}(\pm x^2\lambda)\varepsilon_{3\alpha+\beta}(\pm x^3\lambda)\varepsilon_{3\alpha+2\beta}(\pm x^3\lambda^2) = (\varepsilon_\beta(\lambda), \varepsilon_\alpha(x)) \in B_S(2N) \tag{3.6}$$

Note that $\varepsilon_{3\alpha+\beta}(x^3\lambda), \varepsilon_{3\alpha+2\beta}(x^3\lambda^2)$ both commute with $\varepsilon_\alpha(1)$. Hence we obtain from equation (3.6):

$$\begin{aligned}
&B_S(4N) \ni \big(\varepsilon_{\alpha+\beta}(\pm x\lambda)\varepsilon_{2\alpha+\beta}(\pm x^2\lambda)\varepsilon_{3\alpha+\beta}(\pm x^3\lambda)\varepsilon_{3\alpha+2\beta}(\pm x^3\lambda^2), \varepsilon_\alpha(1)\big)\\
&= \big(\varepsilon_{\alpha+\beta}(\pm x\lambda)\varepsilon_{2\alpha+\beta}(\pm x^2\lambda), \varepsilon_\alpha(1)\big)\\
&\sim (\varepsilon_{2\alpha+\beta}(\pm x^2\lambda), \varepsilon_\alpha(1)) \cdot (\varepsilon_\alpha(1), \varepsilon_{\alpha+\beta}(\pm x\lambda))\\
&= \varepsilon_{3\alpha+\beta}(\pm 3x^2\lambda) \cdot \varepsilon_{2\alpha+\beta}(\pm 2x\lambda)\varepsilon_{3\alpha+\beta}(\pm 3x\lambda)\varepsilon_{3\alpha+2\beta}(\pm 3x^2\lambda^2)\\
&= \varepsilon_{2\alpha+\beta}(\pm 2x\lambda)\varepsilon_{3\alpha+\beta}(\pm 3x\lambda \pm 3x^2\lambda)\varepsilon_{3\alpha+2\beta}(\pm 3x^2\lambda^2)
\end{aligned}$$

Yet $\varepsilon_{3\alpha+\beta}(\pm 3x\lambda \pm 3x^2\lambda), \varepsilon_{3\alpha+2\beta}(\pm 3x^2\lambda^2) \in B_S(2N)$ holds by claim (1a) and hence $\varepsilon_{2\alpha+\beta}(2x\lambda) \in B_S(8N)$ holds as well. This finishes the proof of the first claim of the lemma. For the second claim inspect first the commutator

$$B_S(2N) \ni (\varepsilon_\beta(x), \varepsilon_\alpha(\lambda)) = \varepsilon_{\alpha+\beta}(\pm x\lambda)\varepsilon_{2\alpha+\beta}(\pm x\lambda^2)\varepsilon_{3\alpha+\beta}(\pm x\lambda^3)\varepsilon_{3\alpha+2\beta}(\pm x^2\lambda^2).$$

However, all of the factors besides $\varepsilon_{3\alpha+\beta}(x\lambda^3)$ in this product commute with $\varepsilon_\beta(1)$. Thus taking the commutator with $\varepsilon_\beta(1)$, we obtain the second claim after conjugation. □

With this in hand, Theorem 3.2.2 follows:

Proof. The proof is very similar to the proof of Theorem 3.2.1. As mentioned in Section 2.3, G_2 is a subgroup-scheme of GL_8 defined over $\mathbb{Z}$. Let natural numbers k, l be given with $1 \leq k, l \leq 8$ and let $P \subset \mathbb{Z}[y_{ij}]$ be a collection of polynomials describing membership in G_2. Also if $k = l$, we further assume that $k = l < 8$.

The language $\mathcal{L}$ and the theory $\mathcal{T}_{kl}$ is defined the same way as in the proof of Theorem 3.2.1 except for two differences: First, the sentence in item (3) describes that for each model $\mathcal{M}$ the matrix $\mathcal{A}^{\mathcal{M}}$ is an element of $G_2(\mathcal{R}^{\mathcal{M}})$ instead of $\mathrm{Sp}_4(\mathcal{R}^{\mathcal{M}})$.

Second, the family of sentences $(\theta_r)_{r\in\mathbb{N}}$ in (5) has the form:

$$\begin{aligned}\theta_r :&\forall X_1,\dots,X_r,\forall e_1,\dots e_r\in\{0,1,-1\}:\\ &(P(X_1)\wedge\cdots\wedge P(X_r))\to\big((\varepsilon_\beta(e(k,l)^3)\neq(\mathcal{A}^{e_1})^{X_1}\cdots(\mathcal{A}^{e_r})^{X_r}\big)\end{aligned}$$

In particular, the sentence in (2) still has the form

$$e(k,l)=\begin{cases}a_{kl}, & \text{if } k\neq l\\ a_{kk}-a_{k+1,k+1}, & \text{if not}\end{cases}$$

One then obtains using Corollary 3.5.3, that a model of the sentences in (1) through (4) cannot be a model of all sentences θ_r in (5). Hence $\mathcal{T}_{kl}$ is inconsistent. As in previous proofs, we can by invoking Gödel's compactness Theorem, find an $L_{k,l}(G_2)\in\mathbb{N}$ such that the subset $\mathcal{T}^1_{kl}\subset\mathcal{T}_{kl}$ that contains all sentences in (1) through (4) and the *single* sentence $\theta_{L_{k,l}(G_2)}$, is already inconsistent.

Next, let R be an arbitrary commutative ring with 1 and let $A\in G_2(R)$ be given. This gives us a model $\mathcal{M}$ of the sentences in (1) through (4) of $\mathcal{T}_{kl}$ and hence as $\mathcal{T}^1_{kl}$ is inconsistent this model must violate the statement $\theta^{\mathcal{M}}_{L_{k,l}(G_2)}$. Thus there are elements $g_1,\dots,g_{L_{k,l}(G_2)}\in G_2(R)$ and $e_1,\dots e_{L_{k,l}(G_2)}\in\{0,1,-1\}$ such that (abusing the notation slightly)

$$\varepsilon_\beta(e(k,l)^3)=(A^{e_1})^{g_1}\cdots(A^{e_{L_{k,l}(G_2)}})^{g_{L_{k,l}(G_2)}}$$

Proposition 3.5.4(1a) implies $(e(k,l)^3)\subset\varepsilon_l(A,2L_{k,l}(G_2))$. Summing further over all admissible k,l implies

$$l(A)_3=\sum_{k,l}(e(k,l)^3)\subset\varepsilon_l(A,\sum_{k,l}2L_{k,l}(G_2)).$$

Define next $L(G_2):=\sum_{k,l}2L_{k,l}(G_2)$ and we are done, similar as in the proof of Theorem 3.2.1. □

Chapter 4

Quantitative bounds on root elements for principal ideal domains

In Chapter 3, we give a model theoretic argument for the existence of $L(\Phi)$ as in Theorem 3.1.1, Theorem 3.2.1 and Theorem 3.2.2. In this chapter in contrast, we give explicit values for $L(\Phi)$ for different Φ in case the underlying ring R used to define $G(\Phi, R)$ is a principal ideal domain.

In the first section, we give values for $L(C_n)$ for $n \geq 3$ by way of matrix calculations and generalizations of so-called Hessenberg-matrices. In the second section, we determine $L(C_2)$. In the third section, we introduce a particular version of the Bruhat decomposition for $G(\Phi, R)$ in case of R being a principal ideal domain and study some of the properties of this decomposition. In the fourth and fifth section, we use this Bruhat decomposition to give a value for $L(E_6)$ and $L(G_2)$ respectively.

4.1 Explicit bounds for root elements of $\mathrm{Sp}_{2n}(R)$

For this section, we use a representation of the complex, simply-connected Lie group $\mathrm{Sp}_{2n}(\mathbb{C})$ that gives the following, classical definition of $G(C_n, R) = \mathrm{Sp}_{2n}(R)$ instead of the representation ρ introduced in Section 2.3. However, remember that both of these representations still define the same group $G(C_n, R)$.

Definition 4.1.1. Let R be a commutative ring with 1 and let

$$\mathrm{Sp}_{2n}(R) := \{A \in R^{2n \times 2n} | A^T J A = J\}$$

be given with

$$J = \left(\begin{array}{c|c} 0_n & I_n \\ \hline -I_n & 0_n \end{array}\right)$$

This implies the following:

Lemma 4.1.2. *Let R be a commutative ring with 1 and let $A \in Sp_{2n}(R)$ be given with*

$$A = \left(\begin{array}{c|c} A_1 & A_2 \\ \hline A_3 & A_4 \end{array}\right)$$

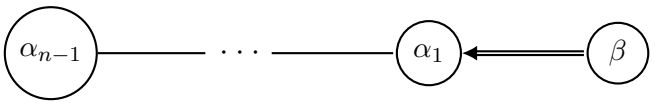
for $A_1, A_2, A_3, A_4 \in R^{n\times n}$. Then the equation

$$A^{-1} = -JA^TJ = \left(\begin{array}{c|c} A_1^T & -A_2^T \\ \hline -A_3^T & A_4^T \end{array}\right)$$

holds.

We use this identity frequently in the following matrix calculations usually without reference. Every symplectic matrix can be writen as a 4×4-block matrix of $n \times n$-matrices and this decomposition shows up naturally in the calculation. Therefore we will often signify this decomposition in blocks using vertical and horizontal lines in the following matrices as done in the above lemma for example. These lines serve merely as an optical help to read the calculations and have no mathematical meaning.

Let $n \geq 2$ be given. We can choose a system of positive simple roots $\{\alpha_1, \ldots, \alpha_{n-1}, \beta\}$ in C_n such that the Dynkin-diagram of this system of positive simple roots has the following form

C_n : 

Then subject to the choice of the maximal torus in $\mathrm{Sp}_{2n}(\mathbb{C})$ as diagonal matrices in $\mathrm{Sp}_{2n}(\mathbb{C})$, the root elements for simple roots in $G(C_n, R) = \mathrm{Sp}_{2n}(R)$ can be chosen as: $\varepsilon_{\alpha_i}(t) = I_{2n} + t(e_{n-i,n-i+1} - e_{2n-i+1,2n-i})$ for $1 \leq i \leq n-1$ and $\varepsilon_\beta(t) = I_{2n} + te_{n,2n}$ for all $t \in R$.

More generally, the root elements $\varepsilon_\phi(x)$ for short, positive roots in $\phi \in C_n$ and $x \in R$ are then either $I_{2n} + t(e_{ij} - e_{n+j,n+i})$ for $1 \leq i < j \leq n$ or $I_{2n} + t(e_{i,n+j} + e_{j,n+i})$ for $1 \leq i < j \leq n$. The root elements $\varepsilon_\psi(x)$ for long, positive roots in $\psi \in C_n$ and $t \in R$ are then $I_{2n} + xe_{i,n+i}$ for $1 \leq i \leq n$. Root elements for negative roots $\phi \in C_n$ and $x \in R$ are then $\varepsilon_\phi(x) = \varepsilon_{-\phi}(x)^T$.

The goal of this section is to prove the following:

Theorem 4.1.3. *Let R be a principal ideal domain, $n \geq 3$ and let $A \in \mathrm{Sp}_{2n}(R)$ be given. Then there is an ideal $I(A)$ in R such that*

1. *$V(I(A)) \subset \Pi(\{A\})$ and*

2. $I(A) \subset \varepsilon_s(A, 64(1+5n))$ *hold.*

Phrased differently, for R a principal ideal domain and $n \geq 3$, one can can pick $L(C_n)$ in Theorem 3.1.1 as $L(C_n) = 64(1+5n)$.

The first Hessenberg form

We start with a Lemma that gives us a Hessenberg form similar to the one used in [24]:

Lemma 4.1.4. *Let R be a principal ideal domain, $n \geq 3$ and $A \in Sp_{2n}(R)$ be given. Then there is an element $B \in Sp_{2n}(R)$ such that $A' := B^{-1}AB$ has the following form*

$$A' = \left(\begin{array}{cccccc|c} a'_{1,1} & a'_{1,2} & a'_{1,3} & \cdot & a'_{1,n-2} & a'_{1,n-1} & a'_{1,n} & \\ a'_{2,1} & a'_{2,2} & a'_{2,3} & \cdot & a'_{2,n-2} & a'_{2,n-1} & a'_{2,n} & \\ 0 & a'_{3,2} & a'_{3,3} & \cdot & a'_{3,n-2} & a'_{3,n-1} & a'_{3,n} & A'_2 \\ 0 & 0 & a'_{4,3} & \cdot & a'_{4,n-2} & a'_{4,n-1} & a'_{4,n} & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \\ 0 & 0 & 0 & \cdot & 0 & a'_{n,n-1} & a'_{n,n} & \\ \hline & & & & A'_3 & & & A'_4 \end{array}\right)$$

with $a'_{11} = a_{11}$ and $a'_{21} = gcd(a_{21}, a_{31}, \ldots, a_{n1})$ up to multiplication with a unit in R and $A'_2, A'_3, A'_4 \in R^{n\times n}$. We call a matrix of the form of A' in $Sp_{2n}(R)$ a matrix in first Hessenberg form.

Proof. If $a_{3,1} = 0$, then define $A^{(3)} := A$. Otherwise choose $t_3 := \gcd(a_{2,1}, a_{3,1})$. Observe that $x_3 := -\frac{a_{3,1}}{t_3}$ and $y_3 := \frac{a_{2,1}}{t_3}$ are coprime elements of R and hence, we can find elements $u_3, v_3 \in R$ with $u_3y_3 - x_3v_3 = 1$. This implies that the matrix

$$T_3 := \left(\begin{array}{ccc|ccc} 1 & & & & & \\ & u_3 & v_3 & & 0_n & \\ & x_3 & y_3 & & & \\ & & & I_{n-3} & & \\ \hline & & & 1 & & \\ & 0_n & & & y_3 & -x_3 \\ & & & & -v_3 & u_3 \\ & & & & & & I_{n-3} \end{array}\right)$$

is an element of $\mathrm{Sp}_{2n}(R)$. The matrix $A^{(3)} := T_3AT_3^{-1}$ has the $(1,1)$-entry $a_{1,1}$ and the $(3,1)$-entry

$$x_3a_{2,1} + y_3a_{3,1} = -\frac{a_{3,1}}{t_3}a_{2,1} + \frac{a_{2,1}}{t_3}a_{3,1} = 0.$$

The entries of $A^{(3)}$ are denoted by $a^{(3)}_{k,l}$.

Next, if $a^{(3)}_{4,1} = 0$, then define $A^{(4)} := A^{(3)}$. Otherwise choose $t_4 := \gcd(a^{(3)}_{2,1}, a^{(3)}_{4,1})$. Observe that $x_4 := -\frac{a^{(3)}_{4,1}}{t_4}$ and $y_4 := \frac{a^{(3)}_{2,1}}{t_4}$ are coprime elements of R and hence, we can find elements $u_4, v_4 \in R$ with $u_4 y_4 - x_4 v_4 = 1$. This implies that the matrix

$$T_4 := \left(\begin{array}{cccc|ccc} 1 & & & & & & \\ & u_4 & 0 & v_4 & & & \\ & 0 & 1 & 0 & & 0_n & \\ & x_4 & 0 & y_4 & & & \\ & & & & I_{n-4} & & \\ \hline & & & & 1 & & \\ & & & & & y_4 & 0 & -x_4 & \\ & & 0_n & & & 0 & 1 & 0 & \\ & & & & & -v_4 & 0 & u_4 & \\ & & & & & & & & I_{n-4} \end{array}\right)$$

is an element of $\mathrm{Sp}_{2n}(R)$. The matrix $A^{(4)} := T_4 A^{(3)} T_4^{-1}$ has the $(1,1)$-entry $a^{(3)}_{1,1} = a_{1,1}$, the $(3,1)$-entry 0 and the $(4,1)$-entry

$$x_4 a^{(3)}_{2,1} + y_4 a^{(3)}_{4,1} = -\frac{a^{(3)}_{4,1}}{t_4} a^{(3)}_{2,1} + \frac{a^{(3)}_{2,1}}{t_4} a^{(3)}_{4,1} = 0.$$

The entries of $A^{(4)}$ are denoted by $a^{(4)}_{k,l}$.

Carrying on this way, we find that the matrix $A^{(n)}$ is conjugate to A in $\mathrm{Sp}_{2n}(R)$ and has the $(1,1)$-entry $a_{1,1}$ and $a^{(n)}_{3,1} = a^{(n)}_{4,1} = \cdots = a^{(n)}_{n,1} = 0$. Further, the construction implies the existence of a matrix $D \in \mathrm{SL}_{n-1}(R)$ with

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \cdot & & & \\ \cdot & & D & \\ \cdot & & & \\ 0 & & & \end{pmatrix} \cdot \begin{pmatrix} a_{1,1} \\ a_{2,1} \\ a_{3,1} \\ \cdot \\ \cdot \\ a_{n,1} \end{pmatrix} = \begin{pmatrix} a_{1,1} \\ a^{(n)}_{2,1} \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}$$

But this implies that $a^{(n)}_{2,1}$ is a multiple of $\gcd(a_{2,1}, \ldots, a_{n,1})$. Further, note $D^{-1} \in \mathrm{SL}_{n-1}(R)$

and hence

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \cdot & & & \\ \cdot & & D^{-1} & \\ \cdot & & & \\ 0 & & & \end{pmatrix} \cdot \begin{pmatrix} a_{1,1} \\ a_{2,1}^{(n)} \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1,1} \\ a_{2,1} \\ a_{3,1} \\ \cdot \\ \cdot \\ a_{n,1} \end{pmatrix}$$

implies that all of the elements of $a_{2,1}, \ldots, a_{n,1}$ are multiples of $a_{2,1}^{(n)}$ and hence $\gcd(a_{2,1}, \ldots, a_{n,1})$ is also a multiple of $a_{2,1}^{(n)}$. So, up to multiplication with a unit $a_{2,1}^{(n)} = \gcd(a_{2,1}, \ldots, a_{n,1})$.

Hence the first column of the matrix $A^{(n)}$ has the form described in the Lemma. The remaining columns of $A^{(n)}$ can be brought to the desired form in a similar way, by conjugating with a matrix of the form

$$\left(\begin{array}{cc|cc} I_2 & & & \\ & D & \multicolumn{2}{c}{0_n} \\ \hline & & I_2 & \\ \multicolumn{2}{c|}{0_n} & & D^{-T} \end{array}\right)$$

for $D \in \mathrm{SL}_{n-2}(R)$. Note, that under conjugation with such a matrix, the first column of $A^{(n)}$ stays fixed and hence this yields the lemma. □

Remark 4.1.5.

1. Upper Hessenberg matrices in $R^{n\times n}$ are matrices $A = (a_{ij})$ with $a_{ij} = 0$ for $i > j + 1$. They are commonly used tools in numerical mathematics [20] and define subvarieties of flag varieties which have been extensively studied [16] as well.

2. The proof strategy for Lemma 4.1.4 is an adaption of [33, Theorem III.1] to the group $\mathrm{Sp}_{2n}(R)$. Lemma 4.1.4 (and Lemma 4.1.10 describing the second Hessenberg form) are actually the only steps in the proof of Theorem 4.1.3 requiring R to be a principal ideal domain.

The strategy to prove Theorem 4.1.3 is to calculate carefully chosen nested commuators of matrices in first (and second) Hessenberg-form with increasingly less entries until one arrives at root elements.

Lemma 4.1.6. *Let R be a commutative ring with 1 and $n \geq 3$ and let A be a matrix in first Hessenberg form in $Sp_{2n}(R)$ and $B := A^{-1}$. Then $X := (A, I_{2n} + e_{1,n+1})$ has the*

following form:

$$X = \left(\begin{array}{cccc|cccc} x_{1,1} & x_{1,2} & \cdot & x_{1,n} & x_{1,n+1} & x_{1,n+2} & 0 & \cdot & 0 \\ x_{2,1} & x_{2,2} & \cdot & x_{2,n} & x_{2,n+1} & x_{2,n+2} & 0 & \cdot & 0 \\ 0 & 0 & \cdot & 0 & 0 & 0 & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & 1 & 0 & 0 & 0 & \cdot & 0 \\ \hline x_{n+1,1} & x_{n+1,2} & \cdot & x_{n+1,n} & x_{n+1,n+1} & x_{n+1,n+2} & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{2n,1} & x_{2n,2} & \cdot & x_{2n,n} & x_{2n,n+1} & x_{2n,n+2} & 0 & \cdot & 1 \end{array}\right)$$

with $x_{1,n+1} = a_{11}(b_{n+1,n+1} - b_{n+1,1}) - 1$ *and* $x_{2,n+1} = a_{21}(b_{n+1,n+1} - b_{n+1,1})$.

Proof. Now let $A_2, A_3, A_4 \in R^{n\times n}$ be given such that A has the following form:

$$A = \left(\begin{array}{cccccc|c} a_{11} & a_{12} & a_{13} & \cdot & a_{1,n-2} & a_{1,n-1} & a_{1n} & \\ a_{21} & a_{22} & a_{23} & \cdot & a_{2,n-2} & a_{2,n-1} & a_{2n} & \\ 0 & a_{32} & a_{33} & \cdot & a_{3,n-2} & a_{3,n-1} & a_{3n} & A_2 \\ 0 & 0 & a_{43} & \cdot & a_{4,n-2} & a_{4,n-1} & a_{4n} & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \\ 0 & 0 & 0 & \cdot & 0 & a_{n,n-1} & a_{n,n} & \\ \hline \multicolumn{6}{c|}{A_3} & A_4 \end{array}\right)$$

Then for the matrices $B_2 := -A_2^T, B_3 := -A_3^T, B_1 := A_4^T \in R^{n\times n}$ one has:

$$B = \left(\begin{array}{c|ccccccc} B_1 & \multicolumn{7}{c}{B_2} \\ \hline B_3 & b_{n+1,n+1} & b_{n+1,n+2} & 0 & \cdot & 0 & 0 & 0 \\ & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & b_{2n-3,n+1} & b_{2n-3,n+2} & b_{2n-3,n+3} & \cdot & b_{2n-3,2n-2} & 0 & 0 \\ & b_{2n-2,n+1} & b_{2n-2,n+2} & b_{2n-2,n+3} & \cdot & b_{2n-2,2n-2} & b_{2n-2,2n-1} & 0 \\ & b_{2n-1,n+1} & b_{2n-1,n+2} & b_{2n-1,n+3} & \cdot & b_{2n-1,2n-2} & b_{2n-1,2n-1} & b_{2n-1,2n} \\ & b_{2n,n+1} & b_{2n,n+2} & b_{2n,n+3} & \cdot & b_{2n,2n-2} & b_{2n,2n-1} & b_{2n,2n} \end{array}\right)$$

Observe first:

$$Ae_{1,n+1}A^{-1} = Ae_{1,n+1}B$$

$$= \left(\begin{array}{ccc|ccccccc} 0 & \cdot & 0 & a_{1,1} & 0 & 0 & \cdot & 0 & 0 & 0 \\ 0 & \cdot & 0 & a_{2,1} & 0 & 0 & \cdot & 0 & 0 & 0 \\ 0 & \cdot & 0 & 0 & 0 & 0 & \cdot & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & 0 & 0 & 0 & 0 & \cdot & 0 & 0 & 0 \\ \hline 0 & \cdot & 0 & a_{n+1,1} & 0 & 0 & \cdot & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & 0 & a_{2n,1} & 0 & 0 & \cdot & 0 & 0 & 0 \end{array}\right) B$$

$$= \left(\begin{array}{ccc|ccc|ccc} a_{11}b_{n+1,1} & a_{11}b_{n+1,2} & \cdot & a_{11}b_{n+1,n} & a_{11}b_{n+1,n+1} & a_{11}b_{n+1,n+2} & 0 & \cdot & 0 \\ a_{21}b_{n+1,1} & a_{21}b_{n+1,2} & \cdot & a_{21}b_{n+1,n} & a_{21}b_{n+1,n+1} & a_{21}b_{n+1,n+2} & 0 & \cdot & 0 \\ 0 & 0 & \cdot & 0 & 0 & 0 & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & 0 & 0 & 0 & 0 & \cdot & 0 \\ \hline a_{n+1,1}b_{n+1,1} & a_{n+1,1}b_{n+1,2} & \cdot & a_{n+1,1}b_{n+1,n} & a_{n+1,1}b_{n+1,n+1} & a_{n+1,1}b_{n+1,n+2} & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{2n,1}b_{n+1,1} & a_{2n,1}b_{n+1,2} & \cdot & a_{2n,1}b_{n+1,n} & a_{2n,1}b_{n+1,n+1} & a_{2n,1}b_{n+1,n+2} & 0 & \cdot & 0 \end{array}\right)$$

This implies that

$$(A, I_{2n} + e_{1,n+1}) = A(I_{2n} + e_{1,n+1})A^{-1}(I_{2n} - e_{1,n+1}) = (I_{2n} + Ae_{1,n+1}A^{-1})(I_{2n} - e_{1,n+1})$$
$$= I_{2n} + Ae_{1,n+1}A^{-1} - e_{1,n+1} - Ae_{1,n+1}A^{-1}e_{1,n+1}$$

$$= \left(\begin{array}{ccc|ccc|ccc} 1 + a_{11}b_{n+1,1} & a_{11}b_{n+1,2} & \cdot & a_{11}b_{n+1,n} & a_{11}(b_{n+1,n+1} - b_{n+1,1}) - 1 & a_{11}b_{n+1,n+2} & 0 & \cdot & 0 \\ a_{21}b_{n+1,1} & 1 + a_{21}b_{n+1,2} & \cdot & a_{21}b_{n+1,n} & a_{21}(b_{n+1,n+1} - b_{n+1,1}) & a_{21}b_{n+1,n+2} & 0 & \cdot & 0 \\ 0 & 0 & \cdot & 0 & 0 & 0 & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & 1 & 0 & 0 & 0 & \cdot & 0 \\ \hline a_{n+1,1}b_{n+1,1} & a_{n+1,1}b_{n+1,2} & \cdot & a_{n+1,1}b_{n+1,n} & a_{n+1,1}(b_{n+1,n+1} - b_{n+1,1}) + 1 & a_{n+1,1}b_{n+1,n+2} & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{2n,1}b_{n+1,1} & a_{2n,1}b_{n+1,2} & \cdot & a_{2n,1}b_{n,n} & a_{2n,1}(b_{n+1,n+1} - b_{n+1,1}) & a_{2n,1}b_{n+1,n+2} & 0 & \cdot & 1 \end{array}\right)$$

This is precisely the form claimed in the lemma. □

Next, we use the commutator from the previous Lemma to obtain a double commutator with a low number of non-zero entries:

Lemma 4.1.7. *Let R be a commutative ring with 1 and $n \geq 3$ and let $X \in Sp_{2n}(R)$ be of the same form as the commutator X described in Lemma 4.1.6. Then the commutator*

$Z := (X, I_{2n} + e_{2n,1} + e_{n+1,n})$ *has the form*

$$Z = \left(\begin{array}{ccccc|ccc} 1 & \cdot & 0 & z_{1,n} & 0 & 0 & 0 \cdot 0 \\ 0 & \cdot & 0 & z_{2,n} & 0 & 0 & 0 \cdot 0 \\ 0 & \cdot & 0 & 0 & 0 & 0 & 0 \cdot 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \cdot \cdot \\ 0 & \cdot & 0 & 1 & 0 & 0 & 0 \cdot 0 \\ \hline 0 & \cdot & 0 & z_{n+1,n} & 1 & 0 & 0 \cdot 0 \\ 0 & \cdot & 0 & z_{n+2,n} & 0 & 1 & 0 \cdot 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \cdot \cdot \\ 0 & \cdot & 0 & z_{2n-1,n} & 0 & 0 & 0 \cdot 0 \\ z_{2n,1} & \cdot & z_{2n,n-1} & z_{2n,n} & z_{2n,n+1} & z_{2n,n+2} & 0 \cdot 1 \end{array}\right)$$

with $z_{1,n} = x_{1,n+1}$ *and* $z_{2,n} = x_{2,n+1}$.

Proof. Let $Y = (y_{ij})_{1 \le i,j \le 2n}$ be the inverse of X. We must study the following term:

$$X(e_{2n,1} + e_{n+1,n})X^{-1} = (Xe_{2n,1})X^{-1} + X(e_{n+1,n}X^{-1}) = e_{2n,1}Y + Xe_{n+1,n}$$

$$= \left(\begin{array}{ccccc|ccc} 0 & 0 & \cdot & 0 & 0 & 0 & 0 \cdot 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \cdot \cdot \\ 0 & 0 & \cdot & 0 & 0 & 0 & 0 \cdot 0 \\ 0 & 0 & \cdot & 0 & 0 & 0 & 0 \cdot 0 \\ \hline 0 & 0 & \cdot & 0 & 0 & 0 & 0 \cdot 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \cdot \cdot \\ 0 & 0 & \cdot & 0 & 0 & 0 & 0 \cdot 0 \\ y_{11} & y_{12} & \cdot & y_{1n} & y_{1,n+1} & y_{1,n+2} & 0 \cdot 0 \end{array}\right) + \left(\begin{array}{ccccc|ccc} 0 & \cdot & 0 & x_{1,n+1} & 0 & \cdot & 0 \\ 0 & \cdot & 0 & x_{2,n+1} & 0 & \cdot & 0 \\ 0 & \cdot & 0 & 0 & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & 0 & 0 & 0 & \cdot & 0 \\ \hline 0 & \cdot & 0 & x_{n+1,n+1} & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & 0 & x_{2n,n+1} & 0 & \cdot & 0 \end{array}\right)$$

$$= \left(\begin{array}{ccccc|ccc} 0 & \cdot & 0 & x_{1,n+1} & 0 & 0 & 0 \cdot 0 \\ 0 & \cdot & 0 & x_{2,n+1} & 0 & 0 & 0 \cdot 0 \\ 0 & \cdot & 0 & 0 & 0 & 0 & 0 \cdot 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \cdot \cdot \\ 0 & \cdot & 0 & 0 & 0 & 0 & 0 \cdot 0 \\ \hline 0 & \cdot & 0 & x_{n+1,n+1} & 0 & 0 & 0 \cdot 0 \\ 0 & \cdot & 0 & x_{n+2,n+1} & 0 & 0 & 0 \cdot 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \cdot \cdot \\ 0 & \cdot & 0 & x_{2n-1,n+1} & 0 & 0 & 0 \cdot 0 \\ y_{11} & \cdot & y_{1,n-1} & y_{1n} + x_{2n,n+1} & y_{1,n+1} & y_{1,n+2} & 0 \cdot 0 \end{array}\right)$$

Next, observe that

$$X(e_{2n,1}+e_{n+1,n})X^{-1}(e_{2n,1}+e_{n+1,n}) = y_{1,n+1}e_{2n,n}.$$

Hence the matrix

$$\begin{aligned}(X, I_{2n}+e_{2n,1}+e_{n+1,n}) &= (I_{2n}+X(e_{2n,1}+e_{n+1,n})X^{-1})(I_{2n}-e_{2n,1}-e_{n+1,n})\\ &= I_{2n}+X(e_{2n,1}+e_{n+1,n})X^{-1}-(e_{2n,1}+e_{n+1,n})\\ &\quad -X(e_{2n,1}+e_{n+1,n})X^{-1}(e_{2n,1}+e_{n+1,n})\\ &= I_{2n}+e_{2n,1}X^{-1}+Xe_{n+1,n}-e_{2n,1}-e_{n+1,n}-y_{1,n+1}e_{2n,n}\end{aligned}$$

has the desired form. □

Lemma 4.1.8. *Let R be a commutative ring with 1 and $n \geq 3$ and let $Z \in \mathrm{Sp}_{2n}(R)$ be of the same form as the commutator Z in Lemma 4.1.7.*

1. *Then the matrix $(Z, I_{2n}+e_{n+1,1})$ has the form $I_{2n}+a(e_{n+1,n}+e_{2n,1})+be_{2n,n}$ for $a=-z_{1,n}$ and $b=z_{1,n}^2$.*
2. *Then the matrix $(Z, I_{2n}+e_{n+2,2})$ has the form $I_{2n}+a(e_{n+2,n}+e_{2n,2})+be_{2n,n}$ for $a=-z_{2,n}$ and $b=z_{2,n}^2$*

Proof. Set $U := Z^{-1}$. Then U also has the form

$$\left(\begin{array}{cccc|cccc} 1 & \cdot & 0 & u_{1,n} & 0 & 0 & 0 & \cdot & 0\\ 0 & \cdot & 0 & u_{2,n} & 0 & 0 & 0 & \cdot & 0\\ 0 & \cdot & 0 & 0 & 0 & 0 & 0 & \cdot & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & \cdot & 0 & 1 & 0 & 0 & 0 & \cdot & 0\\ \hline 0 & \cdot & 0 & u_{n+1,n} & 1 & 0 & 0 & \cdot & 0\\ 0 & \cdot & 0 & u_{n+2,n} & 0 & 1 & 0 & \cdot & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & \cdot & 0 & u_{2n-1,n} & 0 & 0 & 0 & \cdot & 0\\ u_{2n,1} & \cdot & u_{2n,n-1} & u_{2n,n} & u_{2n,n+1} & u_{2n,n+2} & 0 & \cdot & 1 \end{array}\right)$$

First, observe

$$Ze_{n+1,1}Z^{-1} = (e_{n+1,1}+z_{2n,n+1}e_{2n,1})U = e_{n+1,1}+u_{1,n}e_{n+1,n}+z_{2n,n+1}(e_{2n,1}+u_{1,n}e_{2n,n}).$$

This implies

$$\begin{aligned}(Z, I_{2n} + e_{n+1,1}) &= (I_{2n} + Ze_{n+1,1}Z^{-1})(I_{2n} - e_{n+1,1}) \\ &= (I_{2n} + e_{n+1,1} + u_{1,n}e_{n+1,n} + z_{2n,n+1}(e_{2n,1} + u_{1,n}e_{2n,n}))(I_{2n} - e_{n+1,1}) \\ &= I_{2n} + u_{1,n}e_{n+1,n} + z_{2n,n+1}(e_{2n,1} + u_{1,n}e_{2n,n}).\end{aligned}$$

Next, observe $u_{1,n} = z_{2n,n+1} = -z_{1,n}$ and this gives the first claim of the lemma. The second claim follows the same way. □

Last, observe the following commutator formulas:

Lemma 4.1.9. *Let R be a commutative ring with 1 and $n \geq 3$ and let $a, b, x \in R$ be given.*

1. *Let $S = I_{2n} + a(e_{n+1,n} + e_{2n,1}) + be_{2n,n}$ be given. Then*

$$(S, I_{2n} + x(e_{12} - e_{n+2,n+1})) = I_{2n} + ax(e_{2n,2} + e_{n+2,n})$$

   holds.

2. *Let $S = I_{2n} + a(e_{n+2,n} + e_{2n,2}) + be_{2n,n} \in \mathrm{Sp}_{2n}(R)$ be given. Then*

$$(S, I_{2n} + x(e_{2,1} - e_{n+1,n+2})) = I_{2n} + ax(e_{2n,1} + e_{n+1,n})$$

   holds.

Proof. For the first commutator formula note:

$$S = (I_{2n} + a(e_{n+1,n} + e_{2n,1})) \cdot (I_{2n} + be_{2n,n}).$$

Further, $I_{2n} + be_{2n,n}$ commutes with $I_{2n} + x(e_{12} - e_{n+2,n+1})$. Hence

$$\begin{aligned}&(S, I_{2n} + x(e_{12} - e_{n+2,n+1})) = (I_{2n} + a(e_{n+1,n} + e_{2n,1}), I_{2n} + x(e_{1,2} - e_{n+2,n+1})) \\ &= [I_{2n} + x(I_{2n} + a(e_{n+1,n} + e_{2n,1})) \cdot (e_{1,2} - e_{n+2,n+1}) \cdot (I_{2n} - a(e_{n+1,n} + e_{2n,1}))] \\ &\quad \cdot (I_{2n} - x(e_{12} - e_{n+2,n+1})) \\ &= [I_{2n} + x(e_{12} - e_{n+2,n+1} + ae_{2n,2}) \cdot (I_{2n} - a(e_{n+1,n} + e_{2n,1}))] \cdot (I_{2n} - x(e_{1,2} - e_{n+2,n+1})) \\ &= [I_{2n} + x(e_{12} - e_{n+2,n+1} + ae_{2n,2} + ae_{n+2,n})] \cdot (I_{2n} - x(e_{1,2} - e_{n+2,n+1})) \\ &= I_{2n} + ax(e_{2n,2} + e_{n+2,n}).\end{aligned}$$

follows.

For the second commutator formula note:

$$S = (I_{2n} + a(e_{n+2,n} + e_{2n,2})) \cdot (I_{2n} + be_{2n,n}).$$

Further, $I_{2n} + be_{2n,n}$ commutes with $I_{2n} + x(e_{2,1} - e_{n+1,n+2})$. Hence

$$\begin{aligned}
&(S, I_{2n} + x(e_{2,1} - e_{n+1,n+2})) = (I_{2n} + a(e_{n+2,n} + e_{2n,2}), I_{2n} + x(e_{2,1} - e_{n+1,n+2})) \\
&= [I_{2n} + x(I_{2n} + a(e_{n+2,n} + e_{2n,2})) \cdot (e_{2,1} - e_{n+1,n+2}) \cdot (I_{2n} - a(e_{n+2,n} + e_{2n,2}))] \\
&\quad \cdot (I_{2n} - x(e_{2,1} - e_{n+1,n+2})) \\
&= [I_{2n} + x(e_{21} - e_{n+1,n+2} + ae_{2n,1}) \cdot (I_{2n} - a(e_{n+2,n} + e_{2n,2}))] \cdot (I_{2n} - x(e_{2,1} - e_{n+1,n+2})) \\
&= [I_{2n} + x(e_{21} - e_{n+1,n+2} + ae_{2n,1} + ae_{n+1,n})] \cdot (I_{2n} - x(e_{21} - e_{n+1,n+2})) \\
&= I_{2n} + ax(e_{2n,1} + e_{n+1,n}).
\end{aligned}$$

follows. □

The second Hessenberg Form

Lemma 4.1.10. *Let R be a principal ideal domain and let $n \geq 3$ be given. Then for each $A \in \mathrm{Sp}_{2n}(R)$ there is a matrix $B \in \mathrm{Sp}_{2n}(R)$ such that $A' := BAB^{-1}$ has the form:*

$$A' = \left(\begin{array}{cccccc|c}
\multicolumn{6}{c|}{A'_1} & A'_2 \\ \hline
a'_{n+1,1} & a'_{n+1,2} & a'_{n+1,3} & \cdot & a'_{n+1,n-2} & a'_{n+1,n-1} & a'_{n+1,n} & \\
a'_{n+2,1} & a'_{n+2,2} & a'_{n+2,3} & \cdot & a'_{n+2,n-2} & a'_{n+2,n-1} & a'_{n+2,n} & \\
0 & a'_{n+3,2} & a'_{n+3,3} & \cdot & a'_{n+3,n-2} & a'_{n+3,n-1} & a'_{n+3,n} & A'_4 \\
0 & 0 & a'_{n+4,3} & \cdot & a'_{n+4,n-2} & a'_{n+4,n-1} & a'_{n+4,n} & \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \\
0 & 0 & 0 & \cdot & 0 & a'_{2n,n-1} & a'_{2n,n} &
\end{array}\right)$$

with $a'_{n+2,1} = \gcd(a_{n+2,1}, a_{n+3,1}, \ldots, a_{2n,1})$ up to a multiplication by a unit in R. We call a matrix of the form of A' in $Sp_{2n}(R)$ a matrix in second Hessenberg form.

We omit the proof, as it is very similar to the one of Lemma 4.1.4.

Lemma 4.1.11. *Let R be a commutative ring with 1 and $n \geq 3$ and let A be a matrix in second Hessenberg form in $Sp_{2n}(R)$ and $B = A^{-1}$. Then $X := (A, I_{2n} + e_{1,n+1})$ has the*

following form:

$$X = \left(\begin{array}{ccccc|ccc} x_{1,1} & x_{1,2} & 0 & \cdot & 0 & x_{1,n+1} & \cdot & x_{1,2n} \\ x_{2,1} & x_{2,2} & 0 & \cdot & 0 & x_{2,n+1} & \cdot & x_{2,2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{n,1} & x_{n,2} & 0 & \cdot & 1 & x_{n,n+1} & \cdot & x_{n,2n} \\ \hline x_{n+1,1} & x_{n+1,2} & 0 & \cdot & 0 & x_{n+1,n+1} & \cdot & x_{n+1,2n} \\ x_{n+2,1} & x_{n+2,2} & 0 & \cdot & 0 & x_{n+2,n+1} & \cdot & x_{n+2,2n} \\ 0 & 0 & 0 & \cdot & 0 & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & 0 & 0 & \cdot & 1 \end{array}\right)$$

with $x_{n+2,n+1} = a_{n+2,1}(b_{n+1,n+1} - b_{n+1,1})$ *and* $x_{1,n+1} = a_{11}(b_{n+1,n+1} - b_{n+1,1}) - 1$.

Proof. Let $A_2, A_3, A_4 \in R^{n\times n}$ be given such that A has the following form:

$$A = \left(\begin{array}{cccccc|c} \multicolumn{6}{c|}{A_1} & A_2 \\ \hline a_{n+1,1} & a_{n+1,2} & a_{n+1,3} & \cdot & a_{n+1,n-2} & a_{n+1,n-1} & a_{n+1,n} \\ a_{n+2,1} & a_{n+2,2} & a_{n+2,3} & \cdot & a_{n+2,n-2} & a_{n+2,n-1} & a_{n+2,n} \\ 0 & a_{n+3,2} & a_{n+3,3} & \cdot & a_{n+3,n-2} & a_{n+3,n-1} & a_{n+3,n} \\ 0 & 0 & a_{n+4,3} & \cdot & a_{n+4,n-2} & a_{n+4,n-1} & a_{n+4,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & 0 & a_{2n,n-1} & a_{2n,n} \end{array}\right)$$

Then for the matrices $B_2 := -A_2^T, B_1 := A_4^T, B_4 := A_1^T \in R^{n\times n}$, one has:

$$B = \left(\begin{array}{cccccc|c} \multicolumn{6}{c|}{B_1} & B_2 \\ \hline b_{n+1,1} & b_{n+1,2} & 0 & \cdot & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ b_{2n-3,1} & b_{2n-3,2} & b_{2n-3,3} & \cdot & b_{2n-3,n-2} & 0 & 0 \\ b_{2n-2,1} & b_{2n-2,2} & b_{2n-2,3} & \cdot & b_{2n-2,n-2} & b_{2n-2,n-1} & 0 \\ b_{2n-1,1} & b_{2n-1,2} & b_{2n-1,3} & \cdot & b_{2n-1,n-2} & b_{2n-1,n-1} & b_{2n-1,n} \\ b_{2n,1} & b_{2n,2} & b_{2n,3} & \cdot & b_{2n,n-2} & b_{2n,n-1} & b_{2n,n} \end{array}\right) \begin{array}{c} \\ \\ \\ \\ \\ \\ B_4 \end{array}$$

Observe first:

$$Ae_{1,n+1}A^{-1} = Ae_{1,n+1}B$$

$$= \left(\begin{array}{ccc|ccccccc} 0 & \cdot & 0 & a_{1,1} & 0 & 0 & \cdot & 0 & 0 & 0 \\ 0 & \cdot & 0 & a_{2,1} & 0 & 0 & \cdot & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & 0 & a_{n,1} & 0 & 0 & \cdot & 0 & 0 & 0 \\ \hline 0 & \cdot & 0 & a_{n+1,1} & 0 & 0 & \cdot & 0 & 0 & 0 \\ 0 & \cdot & 0 & a_{n+2,1} & 0 & 0 & \cdot & 0 & 0 & 0 \\ 0 & \cdot & 0 & 0 & 0 & 0 & \cdot & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & 0 & 0 & 0 & 0 & \cdot & 0 & 0 & 0 \end{array}\right) B$$

$$= \left(\begin{array}{ccccc|ccc} a_{11}b_{n+1,1} & a_{11}b_{n+1,2} & 0 & \cdot & 0 & a_{11}b_{n+1,n+1} & \cdot & a_{11}b_{n+1,2n} \\ a_{21}b_{n+1,1} & a_{21}b_{n+1,2} & 0 & \cdot & 0 & a_{21}b_{n+1,n+1} & \cdot & a_{21}b_{n+1,2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \\ a_{n,1}b_{n+1,1} & a_{n,1}b_{n+1,2} & 0 & \cdot & 0 & a_{n,1}b_{n+1,n+1} & \cdot & a_{n,1}b_{n+1,2n} \\ \hline a_{n+1,1}b_{n+1,1} & a_{n+1,1}b_{n+1,2} & 0 & \cdot & 0 & a_{n+1,1}b_{n+1,n+1} & \cdot & a_{n+1,1}b_{n+1,2n} \\ a_{n+2,1}b_{n+1,1} & a_{n+2,1}b_{n+1,2} & 0 & \cdot & 0 & a_{n+2,1}b_{n+1,n+1} & \cdot & a_{n+2,1}b_{n+1,2n} \\ 0 & 0 & 0 & \cdot & 0 & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & 0 & 0 & \cdot & 0 \end{array}\right)$$

This implies that

$$(A, I_{2n} + e_{1,n+1}) = A(I_{2n} + e_{1,n+1})A^{-1}(I_{2n} - e_{1,n+1}) = (I_{2n} + Ae_{1,n+1}A^{-1})(I_{2n} - e_{1,n+1})$$
$$= I_{2n} + Ae_{1,n+1}B - e_{1,n+1} - Ae_{1,n+1}Be_{1,n+1}$$

$$= \left(\begin{array}{ccccc|ccc} 1+a_{11}b_{n+1,1} & a_{11}b_{n+1,2} & 0 & \cdot & 0 & a_{11}(b_{n+1,n+1}-b_{n+1,1})-1 & a_{11}b_{n+1,n+2} & \cdot & a_{11}b_{n+1,2n} \\ a_{21}b_{n+1,1} & 1+a_{21}b_{n+1,2} & 0 & \cdot & 0 & a_{21}(b_{n+1,n+1}-b_{n+1,1}) & a_{21}b_{n+1,n+2} & \cdot & a_{21}b_{n+1,2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n,1}b_{n+1,1} & a_{n,1}b_{n+1,2} & 0 & \cdot & 1 & a_{n,1}(b_{n+1,n+1}-b_{n+1,1}) & a_{n,1}b_{n+1,n+2} & \cdot & a_{n,1}b_{n+1,2n} \\ \hline a_{n+1,1}b_{n+1,1} & a_{n+1,1}b_{n+1,2} & 0 & \cdot & 0 & 1+a_{n+1,1}(b_{n+1,n+1}-b_{n+1,1}) & a_{n+1,1}b_{n+1,n+2} & \cdot & a_{n+1,1}b_{n+1,2n} \\ a_{n+2,1}b_{n+1,1} & a_{n+2,1}b_{n+1,2} & 0 & \cdot & 0 & a_{n+2,1}(b_{n+1,n+1}-b_{n+1,1}) & 1+a_{n+2,1}b_{n+1,n+2} & \cdot & a_{n+2,1}b_{n+1,2n} \\ 0 & 0 & 0 & \cdot & 0 & 0 & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & 0 & 0 & 0 & \cdot & 1 \end{array}\right)$$

This is precisely the form claimed in the lemma. □

Lemma 4.1.12. *Let R be a commutative ring with 1 and $n \geq 3$ and let $X \in Sp_{2n}(R)$ be of the same form as the commutator X described in Lemma 4.1.11. Then for $Y := X^{-1}$*

the commutator $Z := (X, I_{2n} + e_{n,1} - e_{n+1,2n})$ *has the form*

$$Z = \left(\begin{array}{ccccc|ccc} 1 & 0 & 0 & \cdot & 0 & 0 & \cdot & 0 & z_{1,2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ z_{n,1} & z_{n,2} & 0 & \cdot & 1 & z_{n,n+1} & \cdot & z_{n,2n-1} & z_{n,2n} \\ \hline 0 & 0 & 0 & \cdot & 0 & 1 & \cdot & 0 & z_{n+1,2n} \\ 0 & 0 & 0 & \cdot & 0 & 0 & \cdot & 0 & z_{n+2,2n} \\ 0 & 0 & 0 & \cdot & 0 & 0 & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & 0 & 0 & \cdot & 0 & 1 \end{array}\right)$$

with $z_{n+2,2n} = -x_{n+2,n+1}$ *and* $z_{1,2n} = -x_{1,n+1}$.

Proof. We must study the following term:

$$X(e_{n,1} - e_{n+1,2n})X^{-1} = (Xe_{n,1})X^{-1} - X(e_{n+1,2n}X^{-1}) = e_{n,1}Y - Xe_{n+1,2n}$$

$$= \left(\begin{array}{ccccc|ccc} 0 & 0 & 0 & \cdot & 0 & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ y_{11} & y_{12} & 0 & \cdot & 0 & y_{1,n+1} & \cdot & y_{1,2n} \\ \hline 0 & 0 & 0 & \cdot & 0 & 0 & \cdot & 0 \\ 0 & 0 & 0 & \cdot & 0 & 0 & \cdot & 0 \\ 0 & 0 & 0 & \cdot & 0 & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & 0 & 0 & \cdot & 0 \end{array}\right) + \left(\begin{array}{ccc|ccc} 0 & \cdot & 0 & 0 & \cdot & -x_{1,n+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & 0 & 0 & \cdot & -x_{n,n+1} \\ \hline 0 & \cdot & 0 & 0 & \cdot & -x_{n+1,n+1} \\ 0 & \cdot & 0 & 0 & \cdot & -x_{n+2,n+1} \\ 0 & \cdot & 0 & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & 0 & 0 & \cdot & 0 \end{array}\right)$$

$$= \left(\begin{array}{ccccc|ccc|c} 0 & 0 & 0 & \cdot & 0 & 0 & \cdot & 0 & -x_{1,n+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ y_{11} & y_{12} & 0 & \cdot & 0 & y_{1,n+1} & \cdot & y_{1,2n-1} & y_{1,2n} - x_{n,n+1} \\ \hline 0 & 0 & 0 & \cdot & 0 & 0 & \cdot & 0 & -x_{n+1,n+1} \\ 0 & 0 & 0 & \cdot & 0 & 0 & \cdot & 0 & -x_{n+2,n+1} \\ 0 & 0 & 0 & \cdot & 0 & 0 & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & 0 & 0 & \cdot & 0 & 0 \end{array}\right)$$

Next, observe

$$X(e_{n,1} - e_{n+1,2n})X^{-1}(e_{n,1} - e_{n+1,2n}) = -y_{1,n+1}e_{n,2n}.$$

Hence the matrix

$$\begin{aligned}(X, I_{2n} + e_{n,1} - e_{n+1,2n}) &= (I_{2n} + X(e_{n,1} - e_{n+1,2n})X^{-1})(I_{2n} - e_{n,1} + e_{n+1,2n}) \\ &= I_{2n} + X(e_{n,1} - e_{n+1,2n})X^{-1} - e_{n,1} + e_{n+1,2n} \\ &\quad - X(e_{n,1} - e_{n+1,2n})X^{-1}(e_{n,1} - e_{n+1,2n}) \\ &= I_{2n} + X(e_{n,1} - e_{n+1,2n})X^{-1} - e_{n,1} + e_{n+1,2n} + y_{1,n+1}e_{n,2n}\end{aligned}$$

has the desired form. □

Next, we have the following:

Lemma 4.1.13. *Let R be a commutative ring with 1 and $n \geq 3$ and let $Z \in \mathrm{Sp}_{2n}(R)$ be of the same form as the commutator Z in Lemma 4.1.12. Then*

1. *Then the matrix $(Z^{-1}, I_{2n}+e_{n+1,1})$ has the form $I_{2n} - z_{1,2n}(e_{n,1} - e_{n+1,2n}) + z_{1,2n}^2 e_{n,2n}$.*
2. *Then the matrix $(Z, I_{2n}+e_{2,n+2})$ has the form $I_{2n} - z_{n+2,2n}(e_{n,n+2}+e_{2,2n}) + z_{n+2,2n}^2 e_{n,2n}$.*

The proof is straightforward so we will omit it.

Lemma 4.1.14. *Let R be a commutative ring with 1 and $n \geq 3$ and let $a, b, x \in R$ be given.*

1. *Let $S = I_{2n} + a(e_{n,1} - e_{n+1,2n}) + be_{n,2n}$ be given. Then*

$$(S, I_{2n} + x(e_{1,n-1} - e_{2n-1,n+1})) = I_{2n} + ax(e_{n,n-1} - e_{2n-1,2n})$$

   holds.

2. *Let $S = I_{2n} + a(e_{n,n+2} + e_{2,2n}) + be_{n,2n}$ be given. Then*

$$(S, I_{2n} + x(e_{n+2,1} + e_{n+1,2})) = I_{2n} + ax(e_{n,1} - e_{n+1,2n})$$

   holds.

Again, the calculations are straightforward, so we are going to omit them.

Constructing the level ideal

We will apply the previous calculations to various matrices. First, note the following proposition:

Proposition 4.1.15. *Let R be a principal ideal domain, $n \geq 3$ and $A = (a_{ij})_{1 \leq i,j \leq 2n} \in \mathrm{Sp}_{2n}(R)$ be given. Then there are ideals*

1. *$I_1^{(1)}(A) \subset \varepsilon_s(A, 32)$ with $a_{2,1}, \ldots, a_{n,1} \in I_1^{(1)}(A)$ and*

2. $I_1^{(2)}(A) \subset \varepsilon_s(A, 32)$ *with* $a_{n+2,1}, \ldots, a_{2n,1} \in I_1^{(2)}(A)$.

We denote the ideal $I_1^{(1)}(A) + I_1^{(2)}(A) \subset \varepsilon_s(A, 64)$ *by* $I_1(A)$.

Proof. The proof will be split in two parts. First we are going to construct the ideal $I_1^{(1)}(A)$ containing $a_{2,1}, \ldots, a_{n,1}$ and then the second ideal $I_1^{(2)}(A)$ containing $a_{n+2,1}, \ldots, a_{2n,1}$.

For the first ideal put A in first Hessenberg form and call the resulting matrix $A' = (a'_{ij})_{1 \leq i,j \leq 2n}$ with inverse $B' = (b'_{ij})_{1 \leq i,j \leq 2n}$. Then apply Lemma 4.1.6 to A' to obtain a matrix $X = (x_{ij})_{1 \leq i,j \leq 2n}$ as in the lemma with entries $x_{1,n+1} = a'_{11}(b'_{n+1,n+1} - b'_{n+1,1}) - 1$ and $x_{2,n+1} = a'_{21}(b'_{n+1,n+1} - b'_{n+1,1})$. Note $X \in B_A(2)$. Next, apply Lemma 4.1.7 to obtain a matrix $Z = (z_{ij})_{1 \leq i,j \leq 2n}$ with $z_{1,n} = x_{1,n+1} = a'_{11}(b'_{n+1,n+1} - b'_{n+1,1}) - 1$ and $z_{2,n} = x_{2,n+1} = a'_{21}(b'_{n+1,n+1} - b'_{n+1,1})$. Note $Z \in B_A(4)$. Next, we apply Lemma 4.1.8 to obtain two matrices S_1 and S_2 and then Lemma 4.1.9 to obtain for all $x \in R$ matrices $T_1(x), T_2(x) \in B_A(16)$ with

$$T_1(x) = I_{2n} + xz_{1,n}(e_{2n,2} + e_{n+2,n}) = I_{2n} + xa'_{11}\left((b'_{n+1,n+1} - b'_{n+1,1}) - 1\right)(e_{2n,2} + e_{n+2,n})$$
$$T_2(x) = I_{2n} + xz_{2,n}(e_{2n,1} + e_{n+1,n}) = I_{2n} + xa'_{21}(b'_{n+1,n+1} - b'_{n+1,1})(e_{2n,1} + e_{n+1,n}).$$

Both of these matrices are root elements associated to short roots and hence as $x \in R$ is arbitrary, we obtain

$$I_1^{(1)}(A) := (a'_{11}(b'_{n+1,n+1} - b'_{n+1,1}) - 1, a'_{21}(b'_{n+1,n+1} - b'_{n+1,1})) \subset \varepsilon_s(A, 32).$$

Note

$$a'_{11}(b'_{n+1,n+1} - b'_{n+1,1}) \equiv 1 \bmod I_1^{(1)}(A).$$

Hence it follows

$$0 = 0 \cdot a'_{11} \equiv a'_{21}(b'_{n+1,n+1} - b'_{n+1,1})a'_{11} \equiv a'_{21} \cdot 1 = a'_{21} \bmod I_1^{(1)}(A).$$

and hence $a'_{21} \in I_1^{(1)}(A)$ holds. But according to Lemma 4.1.4, the entry a'_{21} of the matrix A' is up to multiplication with a unit $\gcd(a_{21}, \ldots, a_{n,1})$ for the entries $a_{21}, \ldots, a_{n,1}$ of the initial matrix A. So in particular, we obtain for an arbitrary matrix $A \in \mathrm{Sp}_{2n}(R)$ that $(a_{21}, \ldots, a_{n,1})$ is a subset of $I_1^{(1)}(A)$.

Running through the same line of calculations again, but using the second Hessenberg form and the Lemmas 4.1.11 through Lemma 4.1.14 instead, yields the ideal $I_1^{(2)}(A) \subset \varepsilon_s(A, 32)$ with $a_{n+2,1}, \ldots, a_{2n,1} \in I_1^{(2)}(A)$. □

The proposition yields all of-diagonal entries of the first column save for the single entry $a_{n+1,1}$ as arguments x for root elements $\varepsilon_\phi(x)$ for $x \in R$ and $\phi \in C_n$ short. In the

next proposition, we will explain how to obtain all off-diagonal entries in $\varepsilon_s(A, K)$ for an appropriate $K \in \mathbb{N}$.

Proposition 4.1.16. *Let R be a principal ideal domain, $n \geq 3$ and let $A = (a_{ij})_{1 \leq i,j \leq 2n} \in \mathrm{Sp}_{2n}(R)$ be given. Then there is an ideal $I'(A)$ such that the following two properties hold:*

1. *$I'(A) \subset \varepsilon_s(A, 320n)$ and*
2. *$(a_{i,j} | 1 \leq i \neq j \leq 2n) \cup (a_{i,i} - a_{i+1,i+1}, a_{n+i,n+i} - a_{n+i+1,n+i+1} | 1 \leq i < n) \subset I'(A)$.*

Proof. First, define for $2 \leq k \leq n$ the elements:

$$w_k := e_{1,k} - e_{k,1} + e_{n+1,n+k} - e_{n+k,n+1} + \sum_{1 \leq j \leq 2n, j \neq 1,k,n+1,n+k} e_{j,j} \in \mathrm{Sp}_{2n}(R).$$

The first column of the matrix $A_k := w_k A w_k^{-1}$ is

$$(a_{k,k}, a_{2,k}, \ldots, a_{k-1,k}, -a_{1,k}, a_{k+1,k}, \ldots, a_{n,k}, a_{n+k,k}, a_{n+2,k} \ldots, a_{n+k-1,k}, -a_{n+1,k}, a_{n+k+2,k}, \ldots, a_{2n,k})^T$$

Hence applying Proposition 4.1.15 to all of the matrices $A_2, \ldots, A_n$ and the matrix $A_1 := A$, there are ideals $I_1(A_1), \ldots, I_1(A_n)$ all of them contained in $\varepsilon_s(A, 64)$ with

$$a_{1,k}, \ldots, a_{n,k}, a_{n+1,k}, \ldots, a_{n+k-1,k}, a_{n+k+2,k}, \ldots, a_{2n,k} \in I_1(A_k)$$

for $k \geq 2$ and

$$a_{2,1}, \ldots, a_{n,1}, a_{n+2,1}, \ldots, a_{2n,1} \in I_1(A_1).$$

So, the ideal $I_2(A) := I_1(A_1) + \cdots + I_1(A_n)$ is contained in $\varepsilon_s(A, 64n)$ and contains all off-diagonal entries of the first n columns of A except possibly the entries $a_{n+1,1}, a_{n+2,2}, \ldots, a_{2n,n}$. Next, observe that J itself is an element of $\mathrm{Sp}_{2n}(R)$ and choose $M_1, M_2, M_3, M_4 \in R^{n \times n}$ with

$$A = \left(\begin{array}{c|c} M_1 & M_2 \\ \hline M_3 & M_4 \end{array}\right).$$

Then we obtain

$$A' := J^{-1} A J = \left(\begin{array}{c|c} 0_n & -I_n \\ \hline I_n & 0_n \end{array}\right) \cdot \left(\begin{array}{c|c} M_1 & M_2 \\ \hline M_3 & M_4 \end{array}\right) \cdot J = \left(\begin{array}{c|c} -M_3 & -M_4 \\ \hline M_1 & M_2 \end{array}\right) \cdot \left(\begin{array}{c|c} 0_n & I_n \\ \hline -I_n & 0_n \end{array}\right)$$

$$= \left(\begin{array}{c|c} M_4 & -M_3 \\ \hline -M_2 & M_1 \end{array}\right)$$

This implies, that if we apply the previous construction of $I_2(A)$ to the matrix A', then we obtain an ideal $I_2(A') \subset \varepsilon_s(A', 64n) = \varepsilon_s(A, 64n)$ that contains all off-diagonal entries of

the last n columns of A, except possibly the entries $a_{1,n+1}, \ldots, a_{n,2n}$. Thus if we consider the ideal $I_3'(A) := I_2(A) + I_2(A') \subset \varepsilon_s(A, 128n)$, it follows:

$$A \equiv \left(\begin{array}{cccc|cccc} a_{11} & 0 & 0 & 0 & a_{1,n+1} & 0 & 0 & 0 \\ 0 & a_{22} & 0 & 0 & 0 & a_{2,n+2} & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & a_{nn} & 0 & 0 & 0 & a_{n,2n} \\ \hline a_{n+1,1} & 0 & 0 & 0 & a_{n+1,n+1} & 0 & 0 & 0 \\ 0 & a_{n+2,2} & 0 & 0 & 0 & a_{n+2,n+2} & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & a_{2n,n} & 0 & 0 & 0 & a_{2n,2n} \end{array}\right) \bmod I_3'(A).$$

Thus the ideal $I_3(A) := (a_{ij}, a_{i,n+j}, a_{n+i,j}, a_{n+i,n+j} | 1 \leq i \neq j \leq n)$ is contained in $I_3'(A) \subset \varepsilon_s(A, 128n)$. Consequently, one also has

$$A^{-1} \equiv \left(\begin{array}{cccc|cccc} a_{n+1,n+1} & 0 & 0 & 0 & -a_{1,n+1} & 0 & 0 & 0 \\ 0 & a_{n+2,n+2} & 0 & 0 & 0 & -a_{2,n+2} & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & a_{2n,2n} & 0 & 0 & 0 & -a_{n,2n} \\ \hline -a_{n+1,1} & 0 & 0 & 0 & a_{1,1} & 0 & 0 & 0 \\ 0 & -a_{n+2,2} & 0 & 0 & 0 & a_{2,2} & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & -a_{2n,n} & 0 & 0 & 0 & a_{n,n} \end{array}\right) \bmod I_3(A).$$

These congruences for A and A^{-1} imply

$$\begin{aligned}
A'' &:= (A, I_{2n} + e_{1,2} - e_{n+2,n+1}) \\
&= \left(I_{2n} + A(e_{1,2} - e_{n+2,n+1})A^{-1}\right) \cdot (I_{2n} - e_{1,2} + e_{n+2,n+1}) \\
&\equiv \left[I_{2n} + (a_{11}e_{12} + a_{n+1,1}e_{n+1,2} - a_{2,n+2}e_{2,n+1} - a_{n+2,n+2}e_{n+2,n+1})A^{-1}\right] \\
&\quad \cdot (I_{2n} - e_{1,2} + e_{n+2,n+1}) \\
&\equiv [I_{2n} + a_{11}(a_{n+2,n+2}e_{12} - a_{2,n+2}e_{1,n+2}) + a_{n+1,1}(a_{n+2,n+2}e_{n+1,2} - a_{2,n+2}e_{n+1,n+2}) \\
&\quad - a_{2,n+2}(-a_{n+1,1}e_{2,1} + a_{11}e_{2,n+1}) - a_{n+2,n+2}(-a_{n+1,1}e_{n+2,1} + a_{11}e_{n+2,n+1})] \\
&\quad \cdot (I_{2n} - e_{1,2} + e_{n+2,n+1}) \\
&= I_{2n} + a_{11}(a_{n+2,n+2}e_{12} - a_{2,n+2}e_{1,n+2}) + a_{n+1,1}(a_{n+2,n+2}e_{n+1,2} - a_{2,n+2}e_{n+1,n+2}) \\
&\quad - a_{2,n+2}(-a_{n+1,1}e_{2,1} + a_{11}e_{2,n+1}) - a_{n+2,n+2}(-a_{n+1,1}e_{n+2,1} + a_{11}e_{n+2,n+1}) \\
&\quad - e_{1,2} + e_{n+2,n+1} - a_{n+1,1}a_{2,n+2}e_{22} - a_{n+1,1}a_{2,n+2}e_{n+1,n+1} \\
&\quad - a_{n+2,n+2}a_{n+1,1}e_{n+2,2} - a_{11}a_{2,n+2}e_{1,n+1} \\
&= \left(\begin{array}{cccc|cccc} 1 & a_{11}a_{n+2,n+2} - 1 & 0 & 0 & -a_{11}a_{2,n+2} & -a_{2,n+2}a_{11} & 0 & 0 \\ a_{2,n+2}a_{n+1,1} & 1 - a_{n+1,1}a_{2,n+2} & 0 & 0 & -a_{2,n+2}a_{11} & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & a_{n+1,1}a_{n+2,n+2} & 0 & 0 & 1 - a_{n+1,1}a_{2,n+2} & -a_{n+1,1}a_{2,n+2} & 0 & 0 \\ a_{n+2,n+2}a_{n+1,1} & -a_{n+2,n+2}a_{n+1,1} & 0 & 0 & 1 - a_{n+2,n+2}a_{11} & 1 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right) \bmod I_3(A).
\end{aligned}$$

Note that the $(n+2, 1)$-entry $a''_{n+2,1}$ of A'' is congruent to $a_{n+2,n+2}a_{n+1,1}$ modulo $I_3(A)$ and the $(1, 2)$-entry of A'' is congruent to $a_{n+2,n+2}a_{11} - 1$ modulo $I_3(A)$. Further, note that $A'' \in B_A(2)$.

Next, apply Proposition 4.1.15(2) to the matrix A'' to obtain an ideal

$$I_4^{(1)}(A) := I_1^{(2)}(A'') \subset \varepsilon_s(A'', 32) \subset \varepsilon_s(A, 64)$$

that contains $a''_{n+2,1}$, an element, which is congruent to $a_{n+2,n+2}a_{n+1,1}$ modulo $I_3(A)$.

So for each element $X = (x_{ij})$ of $\mathrm{Sp}_{2n}(R)$, there is an ideal $I_4^{(1)}(X) \subset \varepsilon_s(X, 64)$ which contains modulo $I_3(X)$ the element $x_{n+2,n+2}x_{n+1,1}$.

Consider next, the matrix $B_A(2) \ni A_2'' := w_2 A'' w_2^{-1}$ and note that its $(2,1)$-entry is congruent modulo $I_3(A)$ to $a_{11}a_{n+2,n+2} - 1$. Apply Proposition 4.1.15(1) to A_2'' to obtain an ideal

$$I_4^{(2)}(A) := I_1^{(1)}(A_2'') \subset \varepsilon_s(A_2'', 32) \subset \varepsilon_s(A, 64)$$

that contains the $(2,1)$-entry of A_2'', which is congruent to $a_{11}a_{n+2,n+2} - 1$ modulo $I_3(A)$. The properties of these ideals imply that the ideal

$$I_4^{(3)}(A) := I_1^{(2)}(A'') + I_1^{(1)}(A_2'') \subset \varepsilon_s(A, 64 + 64) = \varepsilon_s(A, 128)$$

contains modulo $I_3(A)$, the elements $a_{n+2,n+2}a_{11} - 1$ and $a_{n+2,n+2}a_{n+1,1}$ and consequently the element $a_{n+1,1}$ modulo $I_3(A)$.

Phrased differently, for each matrix $X \in \mathrm{Sp}_{2n}(R)$, there is an ideal $I_4^{(3)}(X) \subset \varepsilon_s(X, 128)$, which contains modulo the ideal $I_3(X)$ the element $x_{n+1,1}$ and $x_{n+2,n+2}x_{11} - 1$.

Observe that for $l = 3, \ldots, n$, the conjugate A_l of A defined before, has

1. $(n+1, 1)$-entry equal to $a_{n+l,l}$,
2. $(n+2, n+2)$-entry equal to $a_{n+2,n+2}$ and
3. $(1,1)$-entry equal to $a_{l,l}$.

Further, the conjugate A_2 of A defined before has

1. $(n+1, 1)$-entry equal to $a_{n+2,2}$,
2. $(n+2, n+2)$-entry equal to $a_{n+1,n+1}$ and
3. $(1,1)$-entry equal to $a_{2,2}$.

Hence applying the previous construction of the ideal $I_4^{(3)}(X)$ to the conjugates $A_2, A_3, \ldots, A_n$ then yields ideals $I_4^{(3)}(A_2), \ldots, I_4^{(3)}(A_n) \subset \varepsilon_s(A, 128)$ with the properties that

1. for $l = 2, 3 \ldots, n$ the ideal $I_4^{(3)}(A_l)$ contains the elements $a_{n+l,l}$ modulo the ideal $I_3(A_l) = I_3(A)$,
2. for $l = 3, \ldots, n$ the ideal $I_4^{(3)}(A_l)$ contains the element $a_{n+2,n+2}a_{l,l} - 1$ modulo the ideal $I_3(A_l) = I_3(A)$ and

To summarize, the ideal

$$I_4(A) := I_3(A) + I_4^{(3)}(A) + I_4^{(3)}(A_2) + \cdots + I_4^{(3)}(A_n) \subset \varepsilon_s(A, 256n)$$

contains all the entries $a_{n+1,1}, \ldots, a_{2n,n}$ and $a_{n+2,n+2}a_{1,1}-1, a_{n+2,n+2}a_{3,3}-1, \ldots, a_{n+2,n+2}a_{n,n}-1$. This implies:

$$A \equiv \left(\begin{array}{cccc|cccc} a_{11} & 0 & 0 & 0 & a_{1,n+1} & 0 & 0 & 0 \\ 0 & a_{22} & 0 & 0 & 0 & a_{2,n+2} & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & a_{nn} & 0 & 0 & 0 & a_{n,2n} \\ \hline 0 & 0 & 0 & 0 & a_{n+1,n+1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a_{n+2,n+2} & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_{2n,2n} \end{array}\right) \bmod I_4(A).$$

But A is an element of $\mathrm{Sp}_{2n}(R)$ and hence

$$a_{ll}a_{n+l,n+l} \equiv 1 \bmod I_4(A)$$

holds for all $l = 1, \ldots, n$. Thus $(a_{n+l,n+l} + I_4(A))^{-1} = a_{l,l} + I_4(A)$ holds in $R/I_4(A)$. On the other hand $a_{n+2,n+2}a_{1,1} - 1, a_{n+2,n+2}a_{3,3} - 1, \ldots, a_{n+2,n+2}a_{n,n} - 1$ are all elements of $I_4(A)$ and hence

$$a_{1,1} + I_4(A) = a_{3,3} + I_4(A) = \cdots = a_{n,n} + I_4(A) = (a_{n+2,n+2} + I_4(A))^{-1} = a_{2,2} + I_4(A)$$

holds in the ring $R/I_4(A)$ as well. Thus we obtain

$$A \equiv \left(\begin{array}{cccc|cccc} a_{22} & 0 & 0 & 0 & a_{1,n+1} & 0 & 0 & 0 \\ 0 & a_{22} & 0 & 0 & 0 & a_{2,n+2} & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & a_{22} & 0 & 0 & 0 & a_{n,2n} \\ \hline 0 & 0 & 0 & 0 & a_{n+2,n+2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a_{n+2,n+2} & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_{n+2,n+2} \end{array}\right) \bmod I_4(A).$$

Note in particular, that all diagonal entries of A reduce to units in $R/I_4(A)$.

Similarly, for $A' = J^{-1}AJ$ consider the conjugates $A'_l := w_l A' w_l^{-1}$ for $l = 2, \ldots, n$. Observe that for $l = 3, \ldots, n$ the $(n+1, 1)$-entry of A'_l is $-a_{l,n+l}$ and the $(n+2, n+2)$-entry is $a_{2,2}$. For A'_2 the $(n+1, 1)$-entry is $-a_{2,n+2}$ and the $(n+2, n+2)$-entry is $a_{1,1}$. Further, for A' the $(n+1, 1)$-entry is $-a_{1,n+1}$ and the $(n+2, n+2)$-entry is $a_{2,2}$.

Next, consider the ideals $I_4^{(1)}(A'), I_4^{(1)}(A'_2), \ldots, I_4^{(1)}(A'_n) \subset \varepsilon_s(A, 64)$ and observe that according to the construction of these ideals, one has that

1. the ideal $I_4^{(1)}(A')$ contains the element $-a_{1,n+1}a_{2,2}$ modulo $I_3(A') = I_3(A)$,
2. for $l = 3, \ldots, n$, the ideal $I_4^{(1)}(A'_l)$ contains the element $-a_{l,n+l}a_{2,2}$ modulo $I_3(A'_l) = I_3(A)$ and

3. the ideal $I_4^{(1)}(A'_2)$ contains the element $-a_{2,n+2}a_{1,1}$ modulo $I_3(A'_2) = I_3(A)$.

Next, consider the ideal:

$$\begin{aligned} I'(A) &:= I_4(A) + I_4^{(1)}(A') + I_4^{(1)}(A'_2) + \cdots + I_4^{(1)}(A'_n) \subset \varepsilon_s(A, 256n + 64n) \\ &= \varepsilon_s(A, 320n). \end{aligned}$$

As $I_3(A) \subset I'(A)$, one concludes that

1. $-a_{1,n+1}a_{2,2}$ is an element of $I'(A)$,
2. for $l = 3, \ldots, n$, the element $-a_{l,n+l}a_{2,2}$ is contained in $I'(A)$ and
3. the element $-a_{2,n+2}a_{1,1}$ is contained in $I'(A)$.

But remember that all diagonal entries of A reduce to units in $R/I_4(A)$ and consequently also reduce to units in $R/I'(A)$. Hence as $a_{1,n+1}a_{2,2}, a_{3,n+3}a_{2,2}, \ldots, a_{n,2n}a_{2,2}$ and $a_{2,n+2}a_{1,1}$ are all elements of $I'(A)$, we obtain that $a_{1,n+1}, a_{3,n+3}, \ldots, a_{n,2n}, a_{2,n+2}$ are also elements of $I'(A)$. Hence we obtain

$$A \equiv \left(\begin{array}{cccc|cccc} a_{22} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_{22} & 0 & 0 & 0 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & a_{22} & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & a_{n+2,n+2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a_{n+2,n+2} & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_{n+2,n+2} \end{array}\right) \bmod I'(A).$$

This finishes the proof. □

Remark 4.1.17. For a given element $A \in \mathrm{Sp}_{2n}(R)$, it is possible that any one of the many intermediate ideals I making up $I'(A)$ in the previous proof is already the entire ring R. In this case, it is problematic to speak about units in the quotient R/I or $R/I'(A)$, as we do in the proof. However, if any of the intermediate ideals I is already the entire ring R, then the claim of Proposition 4.1.16 is obvious anyway. This is an unstated caveat in many of the proofs to follow: If the construction of the sought after ideal yields an intermediate ideal I which is already the entire ring R (or the appropriate analogue, say $2R$ in case of $\mathrm{Sp}_4(R)$), the claim of the corresponding statement is then usually true for this intermediate ideal $I = R$ already. All following proofs should be read with this caveat in mind.

Modulo the ideal $I'(A)$ of the previous proposition, we have now reduced A to a diagonal matrix of the form $(a_{11} + I'(A))I_n \oplus (a_{11} + I'(A))^{-1}I_n$. Next, we are going to prove Theorem 4.1.3:

Proof. Consider the matrix $\widetilde{A} := (A^{-1}, I_{2n} + e_{2n,1} + e_{n+1,n}) \in B_A(2)$ and apply Proposition 4.1.15(2) to the matrix $\widetilde{A}$ to obtain an ideal

$$I_1^{(2)}\left(\widetilde{A}\right) \subset \varepsilon_s\left(\widetilde{A}, 32\right) \subset \varepsilon_s\left(A, 64\right)$$

that contains in particular the $(2n,1)$-entry of $\widetilde{A}$. Define next

$$I(A) := I'(A) + I_1^{(2)}\left(\widetilde{A}\right) \subset \varepsilon_s(A, 64 + 320n) = \varepsilon_s(A, 64(1+5n))$$

for $I'(A)$ the ideal of Proposition 4.1.16. Slightly abusing notation, we obtain next:

$$\begin{aligned}
\widetilde{A} &= (A^{-1}, I_{2n} + e_{2n,1} + e_{n+1,n}) = [I_{2n} + A^{-1}(e_{2n,1} + e_{n+1,n})A] \cdot (I_{2n} - e_{n+1,n} - e_{2n,1}) \\
&\equiv \left[I_{2n} + \left(\begin{array}{c|c} a_{1,1}^{-1} I_n & 0_n \\ \hline 0_n & a_{1,1} I_n \end{array}\right)(e_{2n,1} + e_{n+1,n})\left(\begin{array}{c|c} a_{1,1} I_n & 0_n \\ \hline 0_n & a_{1,1}^{-1} I_n \end{array}\right)\right] \cdot (I_{2n} - e_{n+1,n} - e_{2n,1}) \\
&= \left[I_{2n} + a_{11}(e_{2n,1} + e_{n+1,n})\left(\begin{array}{c|c} a_{1,1} I_n & 0_n \\ \hline 0_n & a_{1,1}^{-1} I_n \end{array}\right)\right] \cdot (I_{2n} - e_{n+1,n} - e_{2n,1}) \\
&= [I_{2n} + a_{1,1}^2(e_{2n,1} + e_{n+1,n})] \cdot (I_{2n} - e_{n+1,n} - e_{2n,1}) \\
&= I_{2n} + (a_{1,1}^2 - 1)(e_{2n,1} + e_{n+1,n}) \bmod I'(A).
\end{aligned}$$

This implies that the $(2n,1)$-entry of $\widetilde{A}$ is congruent to $a_{1,1}^2 - 1$ modulo $I'(A)$. This implies $a_{1,1}^2 - 1 \in I(A)$. Let m be a maximal ideal in $V(I(A))$. Note $(a_{1,1} - 1)(a_{1,1} + 1) = a_{1,1}^2 - 1 \in m$ and so either $a_{1,1} - 1$ or $a_{1,1} + 1$ is an element of m. But in either case, one has $(a_{1,1} + m)^{-1} = a_{1,1} + m$ in the ring R/m and so $\pi_m(A)$ is necessarily scalar and thus central in $\mathrm{Sp}_{2n}(R/m)$. Hence m is also an element of $\Pi(\{A\})$ and this finishes the proof. □

We also note the following corollary of the proof:

Corollary 4.1.18. *Let R be a principal ideal domain, $n \geq 3, A \in \mathrm{Sp}_{2n}(R)$. Then the ideal $I(A)$ of Theorem 4.1.3 is a sum of ideals $J_1(A), \ldots, J_{7n+1}(A)$ such that $J_i(A) \subset \varepsilon_s(A, 64)$ holds for all $1 \leq i \leq 7n+1$.*

Proof. Recall the Weyl group elements

$$w_k := e_{1,k} - e_{k,1} + e_{n+1,n+k} - e_{n+k,n+1} + \sum_{1 \leq j \leq 2n, j \neq 1,k,n+1,n+k} e_{j,j} \in \mathrm{Sp}_{2n}(R).$$

for $k = 2, \ldots, n$. Then X_k shall denote the conjugates $w_k X w_k^{-1}$ for $k = 2, \ldots, n$ and an arbitrary $X \in \mathrm{Sp}_{2n}(R)$. Going through the proofs, one can see that $I(A)$ is (contained in) the sum of the following ideals:

1. $I_1^{(1)}(A), I_1^{(1)}(A_2), \ldots, I_1^{(1)}(A_n), I_1^{(2)}(A), I_1^{(2)}(A_2), \ldots, I_1^{(2)}(A_n), I_1^{(1)}(A'), I_1^{(1)}(A'_2), \ldots, I_1^{(1)}(A'_n)$

and $I_1^{(2)}(A'), I_1^{(2)}(A'_2), \ldots, I_1^{(2)}(A'_n)$ for $A' := J^{-1}AJ$. These $4n$ ideals are all individually contained in $\varepsilon_s(A, 32)$.

2. $I_4^{(1)}(A), I_4^{(1)}(A_2), \ldots, I_4^{(1)}(A_n)$ and $I_4^{(2)}(A), I_4^{(2)}(A_2), \ldots, I_4^{(2)}(A_n)$. These $2n$ ideals are all individually contained in $\varepsilon_s(A, 64)$.

3. $I_4^{(1)}(A'), I_4^{(1)}(A'_2), \ldots, I_4^{(1)}(A'_n)$. These n ideals are all individually contained in $\varepsilon_s(A, 64)$.

4. $I_1^{(2)}(\widetilde{A})$ for $\widetilde{A} := (A^{-1}, I_{2n} + e_{2n,1} + e_{n+1,n})$. This ideal is contained in $\varepsilon_s(A, 64)$.

So to summarize: $I(A)$ is (contained in) the sum of $7n + 1$ ideals that are all individually contained in $\varepsilon_s(A, 64)$. □

Remark 4.1.19. We have used $n \geq 3$ at various places in the course of this section. First and foremost, one cannot even put matrices in Hessenberg forms if $n = 2$, because the constructions of Hessenberg forms rely on the block matrices used to conjugate A to have at least one trivial column and row, which cannot be done in the same way for $n = 2$. Second, for various commutator formulas, the root elements or products of root elements obtained would take on a 'degenerate form', where instead of having a root element, with two off-diagonal entries, we would only get one off-diagonal entry and this one would admit an additional factor of 2. This might happen for example in the first part of Lemma 4.1.9, where $n = 2$ implies:

$$I_{2n} + ax(e_{2n,2} + e_{n+2,n}) = I_4 + ax(e_{2*2,2} + e_{2+2,2}) = I_4 + 2axe_{4,2}.$$

The possibility to avoid the use of these degenerate commutator formulas is due to the presence of a root subsystem of C_n spanned by simple roots and isomorphic to A_{n-1} for $n \geq 3$.

4.2 Explicit bounds for root elements of $\mathrm{Sp}_4(R)$

In this section, we determine $L(C_2)$ for principal ideal domains:

Theorem 4.2.1. *Let R be a principal ideal domain and let $A \in \mathrm{Sp}_4(R)$ be given. Then there is an ideal $I(A)$ in R such that*

1. *$V(I(A)) \subset \Pi(\{A\})$ and*

2. *$2I(A) \subset \varepsilon(A, \phi, 384)$ holds for all $\phi \in C_2$.*

Phrased differently, for R a principal ideal domain, one can can pick $L(C_2)$ in Theorem 3.2.1 as $L(C_2) = 384$.

As a first step, we establish a form of Hessenberg matrices in $\mathrm{Sp}_4(R)$:

Lemma 4.2.2. *Let R be a principal ideal domain and $A = (a_{ij})_{1\leq i,j\leq 4} \in \mathrm{Sp}_4(R)$. Then there is a matrix $B \in \mathrm{Sp}_4(R)$ such that BAB^{-1} has $(3,2)$-entry 0 and the same $(4,2)$-entry as A.*

Proof. If $a_{3,2} = 0$, then we are done. Otherwise choose $t := \gcd(a_{3,2}, a_{1,2}), x := \frac{a_{3,2}}{t}$ and $y := -\frac{a_{1,2}}{t}$. Observe that $x, y \in R$ are coprime and hence we can find $u, v \in R$ such that $xv - yu = 1$. Then the matrix

$$B := \left(\begin{array}{cc|cc} u & 0 & v & 0 \\ 0 & 1 & 0 & 0 \\ \hline x & 0 & y & 0 \\ 0 & 0 & 0 & 1 \end{array}\right)$$

is an element of $\mathrm{Sp}_4(R)$ and has the desired property. □

From this one can obtain:

Lemma 4.2.3. *Let R be a commutative ring with 1 and let $A = (a_{ij})_{1\leq i,j\leq 4} \in \mathrm{Sp}_4(R)$ be given with $a_{3,2} = 0$. Then $X := (A, I_4 + e_{2,4})$ has the form*

$$X = \left(\begin{array}{cc|cc} 1 & x_{1,2} & x_{1,3} & x_{1,4} \\ 0 & x_{2,2} & x_{2,3} & x_{2,4} \\ \hline 0 & 0 & 1 & 0 \\ 0 & x_{42} & x_{4,3} & x_{4,4} \end{array}\right)$$

with $x_{4,2} = -a_{4,2}^2$.

We will omit the proof, as it is straight forward. Next:

Lemma 4.2.4. *Let R be a commutative ring with 1 and let $X \in \mathrm{Sp}_4(R)$ be of the same form as the commutator X in the Lemma 4.2.3. Then the commutator $Z := (X, I_4 + e_{2,3} + e_{1,4})$ has the form*

$$Z = I_4 + a(e_{12} - e_{43}) + b(e_{14} + e_{23}) + ce_{1,3}$$

with $a = -x_{4,2}, b = x_{2,2} - 1$ and $c = 2x_{1,2} + x_{4,2}$.

Again, we omit the proof. Next, we obtain:

Lemma 4.2.5. *Let R be a commutative ring with 1 and let $a, b, c, x \in R$ be given. Further, let $Z = I_4 + a(e_{12} - e_{43}) + b(e_{14} + e_{23}) + ce_{1,3}$ be given. Then $(Z, I_4 + x(e_{1,4} + e_{2,3})) = I_4 + 2axe_{1,3}$ holds.*

Again, we omit the proof. From this we can obtain:

Proposition 4.2.6. *Let R be a principal ideal domain and $A \in \mathrm{Sp}_4(R)$ be given. Then $2a_{42}^2 R \subset \varepsilon(A, \phi, 8)$ holds for all $\phi \in C_2$ long.*

Proof. Observe that according to Lemma 4.2.2 the matrix A is conjugate to a matrix A' with $a'_{3,2} = 0$ and $a'_{4,2} = a_{4,2}$. So assume $a_{3,2} = 0$. Then applying first Lemma 4.2.3 to obtain a matrix X, then Lemma 4.2.4 to X to obtain a matrix Z and then lastly applying Lemma 4.2.5 to Z yields that

$$2a_{4,2}^2 R = 2a_{4,2}'^2 R \subset \varepsilon(A, \phi, 8)$$

for $\phi \in C_2$ long. □

Using Proposition 4.2.6, we can prove Theorem 4.2.1:

Proof. The basic idea is to construct different matrices from $A = (a_{ij})$ that have (powers of) entries of A in their respective $(4,2)$-entry and then to apply Proposition 4.2.6.

First, note that applying Lemma 4.2.2, we may assume that $a_{32} = 0$ and let $\phi \in C_2$ be an arbitrary long root. First, we define the matrix:

$$w_2 := e_{1,2} - e_{2,1} + e_{3,4} - e_{4,3} \in \mathrm{Sp}_4(R).$$

Observe that the $(4,2)$-entry of $w_2 A w_2^{-1}$ is $a_{3,1}$ and hence Proposition 4.2.6 yields

$$2a_{3,1}^2 R \subset \varepsilon(A, \phi, 8).$$

Similarly $w_2 J^{-1} A J w_2^{-1}$ has the $(4,2)$-entry $-a_{1,3}$ and $J^{-1}AJ$ has the $(4,2)$-entry $-a_{2,4}$. Thus Proposition 4.2.6 yields

$$2(a_{4,2}^2, a_{3,1}^2, a_{1,3}^2, a_{2,4}^2) \subset \varepsilon(A, \phi, 32).$$

Next, consider the matrix

$$B_A(2) \ni A' = (A, I_4 + e_{4,2}) = I_4 + \left(\begin{array}{cc|cc} a_{14}a_{34} & a_{14}(a_{44}+a_{24}) & -a_{14}^2 & -a_{14}a_{24} \\ a_{24}a_{34} & a_{24}(a_{44}+a_{24}) & -a_{14}a_{24} & -a_{24}^2 \\ \hline a_{34}^2 & a_{34}(a_{44}+a_{24}) & -a_{34}a_{14} & -a_{34}a_{24} \\ a_{44}a_{34} & a_{44}(a_{44}+a_{24})-1 & -a_{14}a_{44} & -a_{44}a_{24} \end{array}\right)$$

Observe that $w_2 A' w_2^{-1}$ has the $(4,2)$-entry $a_{3,4}^2$ and hence Proposition 4.2.6 implies

$$2a_{3,4}^4 R \subset \varepsilon(w_2 A' w_2^{-1}, \phi, 8) = \varepsilon(A', \phi, 8) = \varepsilon(A, \phi, 16).$$

Thus for each matrix $X = (x_{ij}) \in \mathrm{Sp}_4(R)$ one has

$$I_1(X) := 2x_{3,4}^4 R \subset \varepsilon(X, \phi, 16). \tag{4.1}$$

Next, observe that $J^{-1}A'J$ has the form

$$J^{-1}A'J = I_4 + \left(\begin{array}{cc|cc} -a_{34}a_{14} & -a_{34}a_{24} & -a_{34}^2 & -a_{34}(a_{44}+a_{24}) \\ -a_{14}a_{44} & -a_{44}a_{24} & -a_{44}a_{34} & -a_{44}(a_{44}+a_{24})+1 \\ \hline a_{14}^2 & a_{14}a_{24} & a_{14}a_{34} & a_{14}(a_{44}+a_{24}) \\ a_{14}a_{24} & a_{24}^2 & a_{24}a_{34} & a_{24}(a_{44}+a_{24}) \end{array}\right)$$

The $(4,2)$-entry of $w_2 J^{-1}A'Jw_2^{-1}$ is $a_{1,4}^2$ and hence Proposition 4.2.6 implies

$$2a_{1,4}^4 R \subset \varepsilon(w_2 J^{-1}A'Jw_2^{-1}, \phi, 8) = \varepsilon(A', \phi, 8) = \varepsilon(A, \phi, 16).$$

and hence

$$2a_{1,4}^4 R + 2a_{3,4}^4 R \subset \varepsilon(A, \phi, 32)$$

holds.

Further, the $(4,2)$-entry of A' is $a_{44}(a_{44}+a_{24})-1$ and hence Proposition 4.2.6 implies

$$2(a_{44}(a_{44}+a_{24})-1)^2 R \subset \varepsilon(A', \phi, 8) = \varepsilon(A, \phi, 16).$$

Summarizing, we obtain

$$2(a_{1,4}^4, a_{3,4}^4, (a_{44}(a_{44}+a_{24})-1)^2) \subset \varepsilon(A, \phi, 3*16) = \varepsilon(A, \phi, 48).$$

Note, that the fourth column of $J^{-1}AJ$ is

$$(-a_{32}, -a_{42}, a_{12}, a_{22})^T.$$

Thus one obtains from equation (4.1) that

$$I_1(J^{-1}AJ) = 2a_{1,2}^4 R \subset \varepsilon(J^{-1}AJ, \phi, 16) = \varepsilon(A, \phi, 16).$$

Next, consider

$$T := J^{-1}(A, I_4 + e_{12} - e_{4,3})J \in B_A(2)$$

and use equation (4.1) to see:

$$I_1(T) = 2t_{34}^4 R \subset \varepsilon(T, \phi, 16) \subset \varepsilon(A, \phi, 32).$$

To summarize, we obtain for

$$I(A) := (a_{4,2}^2, a_{3,1}^2, a_{1,3}^2, a_{2,4}^2, a_{1,4}^4, a_{3,4}^4, (a_{44}(a_{44}+a_{24})-1)^2, a_{1,2}^4, t_{34}^4)$$

that $2I(A) \subset \varepsilon(A, \phi, 32+48+16+32) = \varepsilon(A, \phi, 128)$ holds for all $\phi \in C_2$ long. So according to Lemma 3.4.2(2), one has $2I(A) \subset \varepsilon(A, \phi, 384)$ for all $\phi \in C_2$. Remember that $l(A)$ is the ideal in R defined as $(a_{ij}, a_{ii} - a_{jj} | 1 \leq i \neq j \leq 4)$. We claim further that $\Pi(I(A)) \subset \Pi(\{A\})$, which if true finishes the proof. To this end, it suffices to show that each maximal m containing $I(A)$ must also contain $l(A)$.

So let m be a maximal ideal containing $I(A)$. Then clearly $a_{4,2}, a_{3,1}, a_{1,3}, a_{2,4}, a_{1,4}, a_{3,4}, a_{12}$ as well as t_{34} and $a_{44}(a_{44}+a_{24})-1$ are all elements of m. Summarizing, we obtain

$$A \equiv \left(\begin{array}{cc|cc} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 \\ \hline 0 & 0 & a_{33} & 0 \\ a_{41} & 0 & a_{43} & a_{44} \end{array}\right) \bmod m.$$

However, A is an element of $\mathrm{Sp}_4(R)$ and thus a_{21}, a_{23}, a_{41} and a_{43} must also be elements of m. Thus A is congruent to a diagonal matrix modulo m. For the same reason,

$$a_{11} + m = (a_{33} + m)^{-1} \text{ and } a_{22} + m = (a_{44} + m)^{-1}$$

must hold in the ring R/m. Let $\pi_m : \mathrm{Sp}_4(R) \to \mathrm{Sp}_4(R/m)$ be the reduction homomorphism induced by the quotient map $R \to R/m$ and set $u := a_{11} + m$ and $v := a_{22} + m$. Then we obtain:

$$
\begin{aligned}
\pi_m(T) &= J^{-1}(\pi_m(A), I_4 + e_{12} - e_{4,3})J \\
&= J^{-1}\left(I_4 + \pi_m(A)(e_{12} - e_{4,3})\pi_m(A^{-1})\right) \cdot (I_4 - e_{12} + e_{4,3})J \\
&= J^{-1}\left(I_4 + \left(\begin{array}{cc|c} u & 0 & \\ 0 & v & 0_2 \\ \hline & 0_2 & \begin{array}{cc} u^{-1} & 0 \\ 0 & v^{-1} \end{array} \end{array}\right)(e_{12} - e_{4,3})\pi_m(A^{-1})\right) \cdot (I_4 - e_{12} + e_{4,3})J \\
&= J^{-1}\left(I_4 + (ue_{12} - v^{-1}e_{43})\left(\begin{array}{cc|c} u^{-1} & 0 & \\ 0 & v^{-1} & 0_2 \\ \hline & 0_2 & \begin{array}{cc} u & 0 \\ 0 & v \end{array} \end{array}\right)\right) \cdot (I_4 - e_{12} + e_{4,3})J \\
&= J^{-1}(I_4 + uv^{-1}(e_{12} - e_{43})) \cdot (I_4 - e_{12} + e_{4,3})J \\
&= J^{-1}(I_4 + (uv^{-1} - 1)(e_{12} - e_{43}))J = I_4 + (uv^{-1} - 1)(e_{34} - e_{21}).
\end{aligned}
$$

But this implies that $t_{34} + m$ agrees with $uv^{-1} - 1$ in R/m. But t_{34} is an element of m and hence

$$a_{11} + m = u = v = a_{22} + m$$

as well as

$$a_{33} + m = u^{-1} = v^{-1} = a_{44} + m$$

follows. Further, $a_{44}(a_{44} + a_{24}) - 1$ and a_{24} are both elements of m and hence

$$a_{44}^2 - 1 = (a_{44} - 1) \cdot (a_{44} + 1)$$

must also be an element of m. But then either $a_{44} - 1$ or $a_{44} + 1$ must be an element of m. Hence $a_{44} + m$ agrees with either $1 + m$ or $-1 + m$ and thus

$$a_{44} + m = (a_{44} + m)^{-1}$$

holds in R/m. However, recall that $(a_{44} + m)^{-1} = a_{22} + m$ and hence we obtain

$$a_{11} + m = a_{22} + m = (a_{44} + m)^{-1} = a_{44} + m = a_{33} + m$$

and consequently $\pi_m(A)$ is a scalar matrix and so $l(A) \subset m$ holds, which finishes the proof. □

4.3 Bruhat decomposition for principal ideal domains

In this section, we will describe another method to provide the values $L(\Phi)$ for Chevalley groups $G(\Phi, R)$ defined over principal ideal domains using a variant of the Bruhat decomposition. It yields noticeably worse bounds than the one for $\mathrm{Sp}_{2n}(R)$ in this thesis and the one for $\mathrm{SL}_n(R)$ in [24]. However, it has the advantage of being more conceptual and so it is easier to apply it to the exceptional root systems. First recall the word norm l_S on groups:

Definition 4.3.1. Let G be a group and $S \subset G$ be given with $S = S^{-1}$ a generating set of G. Then define the function $l_S : G \to \mathbb{N}_0$ by $l_S(1) := 0$ and by

$$l_S(x) := \min\{n \in \mathbb{N} | \exists s_1, \ldots, s_n \in S : x = s_1 \cdots s_n\}$$

for $x \neq 1$.

Next, we need the following definition:

Definition 4.3.2. Let G be a group and $S \subset G$ be given with $S = S^{-1}$ a generating set of G. Further, let $w = s_1 \cdots s_n$ be given with all $s_i \in S$.

1. The tuple (or string) $(s_1, \ldots, s_n) \in S^n$ is called an *expression for w in terms of S of length n*. If $n = l_S(w)$ holds, then the tuple $(s_1, \ldots, s_n)$ is called *a minimal expression for w (with respect to S).*

2. Let a sequence of integers $1 \leq i_1 < i_2 < \cdots < i_k \leq n$ be given. Then

   $$((s_{i_1}, i_1), (s_{i_2}, i_2) \ldots, (s_{i_k}, i_k))$$

   is called a *subexpression* of $(s_1, \ldots, s_n)$.

3. An element $w' \in G$ is called a *subword of* $(s_1, \ldots s_n)$ if there is a sequence of integers $1 \leq i_1 < i_2 < \cdots < i_k \leq n$ such that $w' = s_{i_1} \cdots s_{i_k}$ and $l_S(w') = k$. Further, we denote the set of subwords of $(s_1, \ldots, s_n)$ by $S(s_1, \ldots, s_n)$.

Remark 4.3.3.

1. We will usually omit writing down the positions when denoting subexpressions to simplify notations. So for example, we will write $(s_{i_1}, s_{i_2} \ldots, s_{i_k})$ instead of $((s_{i_1}, i_1), (s_{i_2}, i_2) \ldots, (s_{i_k}, i_k))$.

2. The set $S(s_1, \ldots, s_n)$ depends on the *string* $(s_1, \ldots, s_n)$ and not on the group element $w = s_1 \cdots s_n$ represented by the string. Yet $S(s_1, \ldots, s_n)$ is a subset of G itself and not of the set $S^{<+\infty}$ of strings in S.

If G is the Weyl group $W(\Phi)$ of an irreducible root system, then the generating set S is usually chosen as the set $F = \{w_{\alpha_1}, \ldots, w_{\alpha_u}\}$ of fundamental reflections associated to a system of positive, simple roots $\Pi = \{\alpha_1, \ldots, \alpha_u\}$. However, according to [41, Chapter 8, p. 74, Lemma 53], if $(w_{\alpha_{j_1}}, \ldots, w_{\alpha_{j_k}})$ is a minimal expression with respect to F for an element $w \in W(\Phi)$, then the set $S(w_{\alpha_{j_1}}, \ldots, w_{\alpha_{j_k}})$ is actually independent of the minimal expression $(w_{\alpha_{j_1}}, \ldots, w_{\alpha_{j_k}})$ and only depends on the element w itself. Consequently, we will write $S(w)$ in this case.

Next, for Φ an irreducible root system, we describe certain subgroups of $G(\Phi, R)$ needed for the next proposition.

A principal ideal domain is an integral domain by definition and let Q be the field of fractions of R. Then according to [41, Chapter 3, p. 31, Corollary 6], for $\alpha \in \Phi$ there exists a group homomorphism $\psi_\alpha : \mathrm{SL}_2(Q) \to G(\Phi, Q)$ uniquely determined through

$$\psi_\alpha(I_2 + xe_{12}) = \varepsilon_\alpha(x) \text{ and } \psi_\alpha(I_2 + xe_{21}) = \varepsilon_{-\alpha}(x)$$

for $x \in Q$. Note that $\mathrm{SL}_2(R)$ is a subgroup of $\mathrm{SL}_2(Q)$ and define the subgroup $G_\alpha(R)$ of $G(\Phi, Q)$ as $G_\alpha(R) := \psi_\alpha(\mathrm{SL}_2(R))$. However the group $G_\alpha(R)$ is actually a subgroup of $G(\Phi, R)$ according to [41, Chapter 8, p. 67, Lemma 48].

Further, recall the Weyl group $W(\Phi)$ from Appendix A and the fact that according to Remark 2.2.7 each element w of $W(\Phi)$ has an associated element w of $G(\Phi, R)$. We can state the following proposition now:

Proposition 4.3.4. *[41, Chapter 8, p. 68, Corollary 1] Let R a principal ideal domain with fraction field Q, Φ an irreducible root system of rank at least 2, $F = \{w_{\alpha_1}, \ldots, w_{\alpha_u}\}$ the set of fundamental reflections associated with the system of positive, simple roots $\Pi = \{\alpha_1, \ldots, \alpha_u\}$ of Φ and $W(\Phi)$ the corresponding Weyl group be given.*

1. *Then for each $\alpha_i \in \Pi$, there is a subset $Y_{\alpha_i} \subset G_{\alpha_i}(R)$ such that*

$$G_{\alpha_i}(R) - \{\varepsilon_{\alpha_i}(x)h_{\alpha_i}(t) |\ t \in R^*, x \in R\} = \{\varepsilon_{\alpha_i}(x)h_{\alpha_i}(t) |\ t \in R^*, x \in R\} \cdot Y_{\alpha_i}$$

   holds with uniqueness of decomposition into factors on the right.

2. *Further, let $w \in W(\Phi)$ and $j_1, \ldots, j_k \in \{1, \ldots, u\}$ be given with $k = l_F(w)$ and $w = w_{\alpha_{j_1}} \cdots w_{\alpha_{j_k}}$. Then $(B(Q)wB(Q)) \cap G(\Phi, R) = B(R) \cdot Y_{\alpha_{j_1}} \cdots Y_{\alpha_{j_k}}$ holds for the Y_{α_i} from the first part of the proposition, with uniqueness of decomposition on the right.*

Note, further:

Lemma 4.3.5. *[41, Chapter 8, p. 65, Lemma 46] Let Φ an irreducible root system of rank at least 2, $F = \{w_{\alpha_1}, \ldots, w_{\alpha_u}\}$ the set of fundamental reflections associated with the*

system of positive, simple roots $\{\alpha_1, \ldots, \alpha_u\}$ *of* Φ, $W(\Phi)$ *the corresponding Weyl group and* w_0 *the longest word of* $W(\Phi)$ *with respect to* l_F. *Then* $S(w_0) = W(\Phi)$ *holds.*

Also note:

Definition 4.3.6. Let Φ an irreducible root system, $\Pi = \{\alpha_1, \ldots, \alpha_u\}$ a system of positive simple roots, $\chi \in \Phi$ the positive root of highest weight with respect to Π and $\psi \in \Phi$ a positive root be given. Then let $T(\psi)$ be defined as follows

$$T(\psi) := \{\psi + \sum_{i=1}^{u} k_i\alpha_i | \forall i \in \{1, \ldots, u\} : k_i \geq 0\} \cap \Phi.$$

Remark 4.3.7. Note that, $T(\psi)$ is an ideal in the set of positive roots of Φ : If $\alpha = \psi + \sum_{i=1}^{u} k_i\alpha_i \in T(\psi)$ is given and $\beta = \sum_{i=1}^{u} m_i\alpha_i$ is another positive root such that $\alpha + \beta$ is also a root, then obviously $\alpha + \beta = \psi + \sum_{i=1}^{u}(k_i + m_i)\alpha_i$ is also an element of $T(\psi)$. Hence according to Proposition 2.2.13, the set $\prod_{\mu \in T(\psi)} \varepsilon_\mu(R)$ is a normal subgroup of $B^+(\Phi, R)$ for R a commutative ring with 1.

Next, note the following:

Lemma 4.3.8. *Let* R *a principal ideal domain with fraction field* Q, Φ *an irreducible root system and* $F = \{w_{\alpha_1}, \ldots, w_{\alpha_u}\}$ *the set of fundamental reflections associated with the system of positive, simple roots* $\{\alpha_1, \ldots, \alpha_u\}$ *of* Φ *be given. Further, let* $w \in W(\Phi)$ *and* $w' \in S(w)$ *be given. Assume further that for* χ *the positive root of highest weight in* Φ, *the root* $\psi := w(\chi)$ *is a positive root. Then for* $A \in (B(\Phi, Q)w'B(\Phi, Q)) \cap G(\Phi, R)$, *the commutator* $(A, \varepsilon_\chi(1))$ *is an element of* $\prod_{\mu \in T(\psi)} \varepsilon_\mu(R)$.

Proof. We will prove this lemma in three steps. First, we will assume that R is an algebraically closed field K and $w' = w$. Second, we assume $w' \in S(w)$ is arbitrary and show the lemma in this case and last, we deduce the claim in the case that R is a principal ideal domain. For the first step, note that according to [41, Chapter 3, p. 26, Theorem 4'], we may assume that there is a $b \in B(\Phi, K)$ and an element $u \in U^+(\Phi, K)$ such that wuw^{-1} is an element of $U^-(\Phi, K)$ and $A = bwu$. Then

$$\begin{aligned}(A, \varepsilon_\chi(1)) = (bwu, \varepsilon_\chi(1)) = (u, \varepsilon_\chi(1))^{bw} \cdot (bw, \varepsilon_\chi(1)) = 1^{bw} \cdot (w, \varepsilon_\chi(1))^b \cdot (b, \varepsilon_\chi(1)) \\ = (\varepsilon_{w(\chi)}(\pm 1)\varepsilon_\chi(-1))^b \cdot (b, \varepsilon_\chi(1)) = \varepsilon_\psi(\pm 1)^b \varepsilon_\chi(-1))^b \cdot (b, \varepsilon_\chi(1))\end{aligned}$$

But according to Remark 4.3.7, the subgroup $\prod_{\mu \in T(\psi)} \varepsilon_\mu(K)$ is normalized by $B(\Phi, K)$. Hence $\psi \in T(\psi)$ implies that $\varepsilon_\psi(\pm 1)^b \in \prod_{\mu \in T(\psi)} \varepsilon_\mu(K)$. On the other hand, $\{\chi\}$ is an ideal in the set of positive roots of Φ and so $\varepsilon_\chi(-1))^b \cdot (b, \varepsilon_\chi(1))$ is an element of $\varepsilon_\chi(K)$ according to Proposition 2.2.13. But χ as the positive root of highest weight contains ψ as a summand and so $\varepsilon_\chi(K) \subset \prod_{\mu \in T(\psi)} \varepsilon_\mu(K)$ holds. So summarizing, the entire commutator $(A, \varepsilon_\chi(1))$ is an element of the subgroup $\prod_{\mu \in T(\psi)} \varepsilon_\mu(K)$. This finishes the first step.

For the second step, let $w' \in S(w)$ be given and let X_w denote the Zariski-closure of $B(\Phi,K)wB(\Phi,K)$ in $G(\Phi,K)$. Then the map

$$m : X_w \to G(\Phi,K), B \mapsto (B, \varepsilon_\chi(1))$$

is a morphism of algebraic varieties. Furthermore, $B(\Phi,K)wB(\Phi,K)$ maps into the set $\prod_{\mu\in T(\psi)} \varepsilon_\mu(K)$ under m according to the first step. However, the set $\prod_{\mu\in T(\psi)} \varepsilon_\mu(K)$ is Zariski-closed and hence the Zariski-closure X_w of $B(\Phi,K)wB(\Phi,K)$ must also map into $\prod_{\mu\in T(\psi)} \varepsilon_\mu(K)$. But [41, Chapter 8, p. 74, Theorem 23] implies that $B(\Phi,K)w'B(\Phi,K)$ is a subset of X_w and hence $m(B(\Phi,K)w'B(\Phi,K)) \subset \prod_{\mu\in T(\psi)} \varepsilon_\mu(K)$ holds. This finishes the second step.

For the third step, note first that if R is a principal ideal domain, Q its fraction field and K the algebraic closure of Q, then the second step implies for $A \in G(\Phi,R)$ that $(A, \varepsilon_\chi(1))$ is an element of $\prod_{\mu\in T(\psi)} \varepsilon_\mu(K)$. However, both $\varepsilon_\chi(1)$ and A are elements of $G(\Phi,Q)$ and the group operations in $G(\Phi,K)$ are defined over its prime field K_0 according to [41, Chapter 5, p. 39, Existence Theorem]. Thus $(A, \varepsilon_\chi(1))$ is actually an element of $\prod_{\mu\in T(\psi)} \varepsilon_\mu(Q)$ and thus Proposition 2.2.13 implies that $(A, \varepsilon_\chi(1))$ is actually an element of $\prod_{\mu\in T(\psi)} \varepsilon_\mu(R)$. This finishes the proof. □

Further, we need the following technical Lemma occasionally:

Lemma 4.3.9. *Let R be a principal ideal domain and let Φ be an irreducible root system and $\phi_1, \phi_2 \in \Phi$ with $\phi_1 + \phi_2 \neq 0$. Further, let $t \in R$ and $A \in G_{\phi_2}(R)$ be given and set*

$$Y(\phi_1,\phi_2) = \{k\phi_1 + l\phi_2 |\ k \in \mathbb{N}, l \in \mathbb{Z}\} \cap \Phi.$$

1. *If both $\phi_1 + \phi_2$ and $\phi_1 - \phi_2$ are not elements of Φ, then*

$$(A, \varepsilon_{\phi_1}(t)) = 1$$

follows.

2. *If $\phi_1 + \phi_2$ or $\phi_1 - \phi_2$ is an element of Φ, then*

$$(A, \varepsilon_{\phi_1}(t)) \in \prod_{\psi\in Y(\phi_1,\phi_2)} \varepsilon_\psi(R).$$

3. *If $\Phi \cong A_2$ with positive simple roots ϕ_1, ϕ_2,*

$$A = \psi_{\phi_2}\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

then

$$(\varepsilon_{\phi_1}(t), A) = \varepsilon_{\phi_1}(t(1-d))\varepsilon_{\phi_1+\phi_2}(-tb)$$

and

$$(\varepsilon_{\phi_1+\phi_2}(t), A) = \varepsilon_{\phi_1+\phi_2}(t(1-a))\varepsilon_{\phi_1}(tc).$$

Proof. Let Q be the field of fractions of R. In this proof, we consider $G(\Phi, R)$ as a subgroup of $G(\Phi, Q)$.

If neither $\phi_1+\phi_2$ nor $\phi_1-\phi_2$ are elements of Φ, then $\varepsilon_{\phi_1}(t)$ commutes with all elements in $\varepsilon_{\phi_2}(Q)$ and in $\varepsilon_{-\phi_2}(Q)$. However, we can find $a, b, c, d \in R$ such that

$$A = \psi_{\phi_2}\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

We assume first that $c \neq 0$ and so we can write A in $G(\Phi, Q)$ as

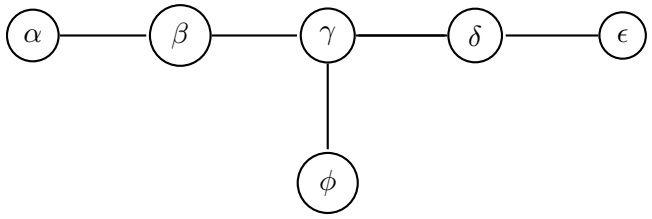
$$A = \varepsilon_{\phi_2}(c^{-1}(a-1)) \cdot \varepsilon_{-\phi_2}(c) \cdot \varepsilon_{\phi_2}(c^{-1}(d-1)).$$

But $\varepsilon_{\phi_1}(t)$ commutes with all these factors of A in $G(\Phi, Q)$ and hence the first claim of the lemma holds in case of $c \neq 0$. The cases of $b \neq 0$ and both b and c equal to 0 are dealt with similarly. This yields the first claim of the lemma. The second claim follows by a similar density argument as Lemma 4.3.8. The third claim of the lemma can simply be checked by matrix calculation in $\mathrm{SL}_3(R) = G(A_2, R)$. □

4.4 Explicit bounds for root elements of $E_6(R)$

Choose a system of positive simple roots $\Pi = \{\alpha, \beta, \gamma, \delta, \epsilon, \phi\} \subset E_6$ such that their corresponding Dynkin-diagram looks as follows

E_6 :



We will show the following:

Proposition 4.4.1. *Let R be a principal ideal domain and $A \in E_6(R)$. Then there is an ideal $I_0(A)$ in R with*

$$I_0(A) \subset \varepsilon(A, 10 \cdot 60^{2^{11}})$$

such that for each maximal ideal m with $I_0(A) \subset m$ the following equation holds

$$\pi_m((A, \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+2\phi}(1))) = 1.$$

Remark 4.4.2. The root $\chi = \alpha + 2\beta + 3\gamma + 2\delta + \epsilon + 2\phi$ is the positive root of highest weight in E_6. This can for example be seen from [18, Appendix, Table B, p. 528].

We will prove this proposition using Lemma 4.3.8. As a first step, note:

Lemma 4.4.3. *The longest element in the Weyl group $W(E_6)$ with respect to the fundamental reflections $F := \{w_\alpha, w_\beta, w_\gamma, w_\delta, w_\epsilon, w_\phi\}$ is*

$$w_0 = (w_\phi w_\beta w_\delta w_\gamma w_\alpha w_\epsilon)^6$$

and this equation gives a minimal expression for w_0 in terms of F.

Proof. According to [22, Exerc. 3.19], the longest element w_0 in $W(E_6)$ is equal to $w^{h(E_6)/2}$ for $w = w_\phi w_\beta w_\delta w_\gamma w_\alpha w_\epsilon$, if $h(E_6) := \mathrm{ord}(w)$ is even. Thus to prove the lemma, it suffices to show that $h(E_6) = 12$ and that $N(E_6) := l_F(w_0) = 36$. But according to [41, Appendix, p. 151, (2)Corollary], the length $l_F(w')$ of an element $w' \in W(E_6)$ is the number of positive roots $\psi \in E_6^+$ with $w'(\psi)$ a negative root and according to [41, Appendix, p. 151, (24)Theorem], the longest element w_0 of $W(E_6)$ maps each root in E_6^+ to a negative root. This implies that $N(E_6)$ is the total number of positive roots in E_6, which is 36 as can be seen from [18, Appendix, Table B, p. 528]. Further, one has the equation $72 = 2N(E_6) = h(E_6)\mathrm{rank}(E_6) = 6h(E_6)$ according to [22, 3.18]. But this implies $h(E_6) = 12$ and finishes the proof. □

Lemma 4.4.4. *The sequences*

1. $s_1 := (w_\alpha, w_\epsilon, w_\phi, w_\beta, w_\delta, w_\gamma, w_\alpha, w_\epsilon, w_\phi, w_\beta, w_\delta, w_\gamma, w_\alpha, w_\epsilon, w_\phi, w_\beta, w_\delta, w_\gamma, w_\alpha, w_\epsilon)$ *and*
2. $s_2 := (w_\delta, w_\beta, w_\phi, w_\epsilon, w_\alpha, w_\gamma, w_\delta, w_\beta, w_\phi, w_\epsilon, w_\alpha, w_\gamma, w_\delta, w_\beta, w_\phi)$

of fundamental reflections in $W(E_6)$ give minimal expressions with respect to the fundamental reflections for the corresponding Weyl group elements $w_1, w_2 \in W(E_6)$. Further, for χ the positive root of highest weight in E_6 :

$$w_1(\chi) = w_2(\chi) = \gamma \text{ and } T(\gamma) = E_6^+ - \{\alpha, \beta, \delta, \epsilon, \phi, \alpha + \beta, \delta + \epsilon\}.$$

The proof of this lemma can be found in Appendix C. The last preparatory lemma is the following:

Lemma 4.4.5. *Let R be a commutative ring with 1 and Φ a simply-laced irreducible root system of rank at least 2 and for each $\phi \in \Phi^+$ let $t_\phi \in R$ be given. Further, for each $k \in \mathbb{N}$, let v_k be the number of roots of weight k in Φ^+ and let l be the weight of the root of highest weight in Φ^+. Then define the sequence $(x_k)_{k=1,\ldots,l}$ of integers by reverse recursion as follows:*

$$x_l := 2, x_{l-k} := (2v_{l-k}x_{l-k+1} + 2v_{l-k} + 1)x_{l-k+1}.$$

Further assume that Φ^+ is ordered in some fixed way by decreasing weight of roots and set

$$u := \prod_{\phi \in \Phi^+} \varepsilon_\phi(t_\phi) \in U^+(\Phi, R).$$

Then

$$(t_\phi | \phi \in \Phi^+) \subset \varepsilon(u, x_1)$$

holds.

Proof. We will show by backwards induction on $k \in \{1, \ldots, l\}$, that the following claim holds:

Claim 4.4.5.1. *Let*

$$u := \prod_{\phi \in \Phi^+} \varepsilon_\phi(t_\phi) \in U^+(\Phi, R).$$

be given such that $t_\phi = 0$ holds for all $\phi \in \Phi^+$ with $\mathrm{wt}(\phi) < k$. Then

$$(t_\phi | \phi \in \Phi^+) \subset \varepsilon(u, x_k)$$

holds.

Clearly, this claim is equivalent to the claim of the lemma. First, for the case $k = l$ let $\chi \in \Phi$ be the root of positive highest weight in Φ. Then $u = \varepsilon_\chi(t_\chi)$ holds and thus the claim follows from the non B_n-case of the proof of Lemma 3.3.3.

So assume now that $k \geq 1$ is smaller than l. For each root $\phi \in \Phi^+$ with $t_\phi \neq 0$ and the weight of ϕ being k, there is a positive, simple root α_ϕ such that $\alpha_\phi + \phi$ is also a root in Φ. As k is the smallest possible weight for the roots in u, we can reorder the terms of u in such a way that

$$u = \left(\prod_{\psi \in \Phi^+ - \{\phi\}, \mathrm{wt}(\psi) \geq k} \varepsilon_\psi(t'_\psi) \right) \cdot \varepsilon_\phi(t_\phi) = u' \cdot \varepsilon_\phi(t_\phi).$$

holds for other $t'_\psi \in R$ for $\psi \in \Phi^+ - \{\phi\}$ with $\mathrm{wt}(\psi) \geq k$. Then, we obtain

$$\begin{aligned}(u, \varepsilon_{\alpha_\phi}(1)) &\sim (\varepsilon_\phi(t_\phi), \varepsilon_{\alpha_\phi}(1)) \cdot (\varepsilon_{\alpha_\phi}(1), u'^{-1}) \\ &= \varepsilon_{\phi+\alpha_\phi}(t_\phi) \cdot (\varepsilon_{\alpha_\phi}(1), u'^{-1}) := u_\phi.\end{aligned}$$

First, note that the weight of all roots appearing in the root elements in u_ϕ can be assumed to be greater or equal to $k+1$ and that we may assume that none of the roots appearing in the root elements in $(\varepsilon_{\alpha_\phi}(1), u'^{-1})$ are $\phi + \alpha_\phi$. Thus by applying the induction hypotheses,

we obtain in particular that

$$(t_\phi) \subset \varepsilon(u_\phi, x_{k+1}) \subset \varepsilon(u, 2x_{k+1}).$$

This implies further that $\varepsilon_\phi(t_\phi)$ is an element of $B_u(2x_{k+1})$. As we can do this for all roots ϕ of weight k appearing in u, we obtain that after multiplication from the right with the elements $\varepsilon_\phi(t_\phi)$ in appropriate order that

$$u'' := \prod_{\psi \in \Phi^+, \mathrm{wt}(\psi) \geq k+1} \varepsilon_\psi(t_\psi)$$

is an element of $B_u(2v_k x_{k+1} + 1)$. But then we can also apply the induction hypothesis to u'' and thus we obtain that

$$(t_\psi | \psi \in \Phi^+, \mathrm{wt}(\psi) \geq k+1) \subset \varepsilon(u'', x_{k+1}) \subset \varepsilon(u, (2v_k x_{k+1} + 1)x_{k+1}).$$

Thus in combination with the fact that $(t_\phi) \subset \varepsilon(u, 2x_{k+1})$ holds for all roots $\phi \in \Phi^+$ of weight k, we obtain

$$(t_\phi | \phi \in \Phi^+) \subset \varepsilon(u, (2v_k x_{k+1} + 2v_k + 1)x_{k+1})$$

However, the integer x_k is defined as $x_k = (2v_k x_{k+1} + 2v_k + 1)x_{k+1}$ so this finishes the induction step and the proof. □

Next, observe that:

Lemma 4.4.6. *Let R be a commutative ring with 1 and let*

$$u := \prod_{\psi \in E_6^+} \varepsilon_\psi(t_\psi) \in U^+(E_6, R).$$

be given. Then $(t_\psi | \psi \in E_6^+) \subset \varepsilon(u, 60^{2^{10}})$ holds.

Proof. According to Lemma 4.4.5 it suffices to show $x_1 \leq 60^{2^{10}}$ to show the claim of the lemma. To this end, observe that $l = 11$ and $v_k \leq 6$ holds for all $k \in \mathbb{N}$ for $\Phi = E_6$, as can be seen for example from the Hasse-diagram of E_6 in the proof of Lemma 4.4.5 in Appendix C. But the recursion $x_k = (2v_k x_{k+1} + 2v_k + 1)x_{k+1}$ implies then further that

$$x_k = (2v_k x_{k+1} + 2v_k + 1)x_{k+1} \leq 5v_k x_{k+1} \cdot x_{k+1} = 30x_{k+1}^2.$$

Then considering that $x_l = 2$, one can show by induction that $x_k \leq 30^{2^{l-k}-1} \cdot x_l^{2^{(l-k)}} = 60^{2^{l-k}}$ and hence $x_1 \leq 60^{2^{10}}$. □

We will prove Proposition 4.4.1 now:

Proof. Let Q be the fraction field of R. For ψ a positive, simple root in E_6, set $T_\psi(R) = Y_\psi \cup \{1\}$ for the Y_ψ as in Proposition 4.3.4. Then according to Lemma 4.4.3, Proposition 4.3.4 and Lemma 4.3.5, there are elements $X_\psi^{(i)} \in T_\psi(R)$ for the positive, simple roots $\psi \in E_6$ and $i = 1, \ldots, 6$ as well as $b \in B^+(E_6, R)$ such that

$$A = b \cdot \prod_{i=1}^{6} X_\phi^{(i)} X_\beta^{(i)} X_\delta^{(i)} X_\gamma^{(i)} X_\alpha^{(i)} X_\epsilon^{(i)}$$

holds. Setting further $Y_\psi^{(i)} := (X_\psi^{(i)})^{-1}$ for ψ positive and simple and $i = 1, \ldots, 6$, this implies

$$\begin{aligned}
&(A, \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+2\phi}(1)) \\
&= (b \prod_{i=1}^{6} X_\phi^{(i)} X_\beta^{(i)} X_\delta^{(i)} X_\gamma^{(i)} X_\alpha^{(i)} X_\epsilon^{(i)}, \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+2\phi}(1)) \\
&\sim (X_\alpha^{(3)} X_\epsilon^{(3)} \prod_{i=4}^{6} X_\phi^{(i)} X_\beta^{(i)} X_\delta^{(i)} X_\gamma^{(i)} X_\alpha^{(i)} X_\epsilon^{(i)}, \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+2\phi}(1)) \\
&\quad \cdot (\varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+2\phi}(1), (((b \prod_{i=1}^{2} X_\phi^{(i)} X_\beta^{(i)} X_\delta^{(i)} X_\gamma^{(i)} X_\alpha^{(i)} X_\epsilon^{(i)}) \cdot X_\phi^{(3)} X_\beta^{(3)} X_\delta^{(3)} X_\gamma^{(3)})^{-1}) \\
&\sim (X_\alpha^{(3)} X_\epsilon^{(3)} \prod_{i=4}^{6} X_\phi^{(i)} X_\beta^{(i)} X_\delta^{(i)} X_\gamma^{(i)} X_\alpha^{(i)} X_\epsilon^{(i)}, \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+2\phi}(1)) \\
&\quad \cdot \left[((Y_\gamma^{(3)} Y_\delta^{(3)} Y_\beta^{(3)} Y_\phi^{(3)} \prod_{i=2}^{3} Y_\epsilon^{(4-i)} Y_\alpha^{(4-i)} Y_\gamma^{(4-i)} Y_\delta^{(4-i)} Y_\beta^{(4-i)} Y_\phi^{(4-i)}) b^{-1}, \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+2\phi}(1)) \right]^{-1}.
\end{aligned}$$

The first part of Lemma 4.4.4 states that

$$s_1 = (w_\alpha, w_\epsilon, w_\phi, w_\beta, w_\delta, w_\gamma, w_\alpha, w_\epsilon, w_\phi, w_\beta, w_\delta, w_\gamma, w_\alpha, w_\epsilon, w_\phi, w_\beta, w_\delta, w_\gamma, w_\alpha, w_\epsilon)$$

is a minimal expression for its corresponding Weyl group element w_1. But note that, all $X_\psi^{(i)}$ are chosen as elements of $T_\psi(R)$. So according to Proposition 4.3.4, the element

$$X_\alpha^{(3)} X_\epsilon^{(3)} \prod_{i=4}^{6} X_\phi^{(i)} X_\beta^{(i)} X_\delta^{(i)} X_\gamma^{(i)} X_\alpha^{(i)} X_\epsilon^{(i)}$$

is an element of $(B(\Phi, Q) w' B(\Phi, Q)) \cap G(\Phi, R)$ for some $w' \in S(w_1)$. According to the first part of Lemma 4.4.4, the Weyl group element w_1 given by the sequence s_1 satisfies $w_1(\chi) = \gamma$ and $T(\gamma) = E_6^+ - \{\alpha, \beta, \delta, \epsilon, \phi, \alpha+\beta, \delta+\epsilon\} =: S$. Thus as $\chi = \alpha + 2\beta + 3\gamma + 2\delta + \epsilon + 2\phi$

is the positive root of highest weight in E_6, by Lemma 4.3.8, the commutator

$$B_1 := (X_\alpha^{(3)} X_\epsilon^{(3)} \prod_{i=4}^{6} X_\phi^{(i)} X_\beta^{(i)} X_\delta^{(i)} X_\gamma^{(i)} X_\alpha^{(i)} X_\epsilon^{(i)}, \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+2\phi}(1))$$

is an element of $\prod_{\psi\in S} \varepsilon_\psi(R)$. But reordering the terms, we may assume that

$$B_1 \in \varepsilon_\gamma(R) \prod_{\psi\in S-\{\gamma\}} \varepsilon_\psi(R).$$

Next, consider the following chain of equations

$$\begin{aligned}
&((Y_\gamma^{(3)} Y_\delta^{(3)} Y_\beta^{(3)} Y_\phi^{(3)} \prod_{i=2}^{3} Y_\epsilon^{(4-i)} Y_\alpha^{(4-i)} Y_\gamma^{(4-i)} Y_\delta^{(4-i)} Y_\beta^{(4-i)} Y_\phi^{(4-i)}) \cdot b^{-1}, \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+2\phi}(1)) \\
&= ((Y_\delta^{(3)} Y_\beta^{(3)} Y_\phi^{(3)} \prod_{i=2}^{3} Y_\epsilon^{(4-i)} Y_\alpha^{(4-i)} Y_\gamma^{(4-i)} Y_\delta^{(4-i)} Y_\beta^{(4-i)} Y_\phi^{(4-i)}) \cdot b^{-1}, \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+2\phi}(1)))^{Y_\gamma^{(3)}} \\
&\quad \cdot (Y_\gamma^{(3)}, \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+2\phi}(1)) \\
&= ((Y_\delta^{(3)} Y_\beta^{(3)} Y_\phi^{(3)} \prod_{i=2}^{3} Y_\epsilon^{(4-i)} Y_\alpha^{(4-i)} Y_\gamma^{(4-i)} Y_\delta^{(4-i)} Y_\beta^{(4-i)} Y_\phi^{(4-i)}) \cdot b^{-1}, \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+2\phi}(1)))^{Y_\gamma^{(3)}}
\end{aligned}$$

Then similarly to the previous discussion, one obtains from Lemma 4.4.4 and Proposition 4.3.4 that

$$(Y_\delta^{(3)} Y_\beta^{(3)} Y_\phi^{(3)} \prod_{i=2}^{3} Y_\epsilon^{(4-i)} Y_\alpha^{(4-i)} Y_\gamma^{(4-i)} Y_\delta^{(4-i)} Y_\beta^{(4-i)} Y_\phi^{(4-i)}) \cdot b^{-1}$$

is an element of $(B(\Phi, Q)w'B(\Phi, Q)) \cap G(\Phi, R)$ for some

$$w' \in S(w_\delta, w_\beta, w_\phi, w_\epsilon, w_\alpha, w_\gamma, w_\delta, w_\beta, w_\phi, w_\epsilon, w_\alpha, w_\gamma, w_\delta, w_\beta, w_\phi) = S(s_2).$$

According to the second part of Lemma 4.4.4, the Weyl group element w_2 given by the sequence s_2 satisfies $w_2(\chi) = \gamma$. Thus by Lemma 4.3.8, the commutator

$$B_2 := ((Y_\delta^{(3)} Y_\beta^{(3)} Y_\phi^{(3)} \prod_{i=2}^{3} Y_\epsilon^{(4-i)} Y_\alpha^{(4-i)} Y_\gamma^{(4-i)} Y_\delta^{(4-i)} Y_\beta^{(4-i)} Y_\phi^{(4-i)}) \cdot b^{-1}, \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+2\phi}(1))$$

is an element of $\prod_{\psi\in S} \varepsilon_\psi(R)$. But then reordering the terms yields

$$B_2 \in \varepsilon_\gamma(R) \prod_{\psi\in S-\{\gamma\}} \varepsilon_\psi(R).$$

Hence, we obtain

$$(B_2^{Y_\gamma^{(3)}})^{-1} \in (\prod_{\psi \in S-\{\gamma\}} \varepsilon_\psi(R)^{Y_\gamma^{(3)}}) \varepsilon_\gamma(R)^{Y_\gamma^{(3)}} \tag{4.2}$$

But for each $\psi \in S - \{\gamma\}$ one of three cases occurs:

1. Neither $\psi + \gamma$ nor $\psi - \gamma$ are elements of E_6. But as $\psi \neq \gamma$, Lemma 4.3.9 implies $\varepsilon_\psi(R)^{Y_\gamma^{(3)}} = \varepsilon_\psi(R)$.

2. The term $\psi + \gamma$ is an element of E_6. Then Lemma 4.3.9 implies

$$\varepsilon_\psi(R)^{Y_\gamma^{(3)}} = \varepsilon_\psi(R)\varepsilon_{\psi+\gamma}(R).$$

3. The term $\psi - \gamma$ is an element of E_6. Then Lemma 4.3.9 implies

$$\varepsilon_\psi(R)^{Y_\gamma^{(3)}} = \varepsilon_\psi(R)\varepsilon_{\psi-\gamma}(R).$$

Note, that if $\psi \in S - \{\gamma\}$ and $\psi + \gamma$ or $\psi - \gamma$ is an element of E_6, then this implies that either $\psi+\gamma \in (S-\{\gamma\})$ or $\psi-\gamma \in (S-\{\gamma\}) \cup \{\alpha+\beta, \beta, \delta, \phi, \delta+\epsilon\} = E_6^+ - \{\alpha, \gamma, \epsilon\} =: S'$. So in any case, one obtains fixing some order of the elements in S' that

$$\varepsilon_\psi(R)^{Y_\gamma^{(3)}} \in \prod_{\nu \in S'} \varepsilon_\nu(R).$$

But S' is a set of positive roots closed under addition and hence according to Proposition 2.2.13, the set $\prod_{\nu \in S'} \varepsilon_\nu(R)$ is a subgroup. Hence together with (4.2), we obtain

$$(B_2^{Y_\gamma^{(3)}})^{-1} \in (\prod_{\psi \in S'} \varepsilon_\psi(R)) \varepsilon_\gamma(R)^{Y_\gamma^{(3)}} \tag{4.3}$$

Remembering that

$$B_1 \in \varepsilon_\gamma(R) \prod_{\psi \in S-\{\gamma\}} \varepsilon_\psi(R)$$

and that $S - \{\gamma\}$ is a subset of S', we obtain that

$$B_1 \in \varepsilon_\gamma(R) \prod_{\psi \in S'} \varepsilon_\psi(R). \tag{4.4}$$

Hence, (4.3) and (4.4) together imply

$$\begin{aligned}(A, \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+2\phi}(1)) &\sim B_1 \cdot (B_2^{Y_\gamma^{(3)}})^{-1} \\ &\in (\varepsilon_\gamma(R) \prod_{\psi \in S'} \varepsilon_\psi(R)) \cdot (\prod_{\psi \in S'} \varepsilon_\psi(R)) \varepsilon_\gamma(R)^{Y_\gamma^{(3)}}\end{aligned}$$

So conjugating with an element of $\varepsilon_\gamma(R)$ yields that $(A, \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+2\phi}(1))$ is conjugate to an element B of

$$(\prod_{\psi\in S'} \varepsilon_\psi(R)) \cdot G_\gamma(R).$$

We assume that the set of roots S' is ordered by decreasing weight; the order of root elements associated to roots of the same weight should be fixed, but does not matter. So, we can pick elements $t_\psi \in R$ for each $\psi \in S'$ and a $Z_\gamma \in G_\gamma(R)$ with

$$B = (\prod_{\psi\in S'} \varepsilon_\psi(t_\psi)) \cdot Z_\gamma.$$

Further, choose $a, b, c, d \in R$ such that

$$Z_\gamma = \psi_\gamma \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Next, we will construct various ideals using Lemma 4.4.6, whose sum has the property desired of $I_0(A)$ in the proposition. First, pick a positive, simple root θ in E_6, which is not γ and consider the following commutator:

$$(B, \varepsilon_\theta(1)) \sim (Z_\gamma, \varepsilon_\theta(1)) \cdot (\varepsilon_\theta(1), (\prod_{\psi\in S'} \varepsilon_\psi(t_\psi))^{-1}).$$

According to Lemma 4.3.9, one obtains that

$$(Z_\gamma, \varepsilon_\theta(1))$$

is an element of $U^+(E_6, R)$. But $(\varepsilon_\theta(1), (\prod_{\psi\in S'} \varepsilon_\psi(t_\psi))^{-1})$ is clearly an element of $U^+(E_6, R)$ as well. Thus the commutator $(B, \varepsilon_\theta(1))$ is conjugate to an element u_θ of the subgroup $U^+(E_6, R)$. Hence using Lemma 4.4.6, we can find an ideal

$$I_\theta(A) \subset \varepsilon(u_\theta, 60^{2^{10}}) \subset \varepsilon(A, 4 \cdot 60^{2^{10}})$$

such that $\pi_{I_\theta(A)}(u_\theta) = 1$. Next, consider the commutator

$$(B, \varepsilon_{\gamma+\delta}(1)) \sim (Z_\gamma, \varepsilon_{\gamma+\delta}(1)) \cdot (\varepsilon_{\gamma+\delta}(1), (\prod_{\psi\in S'} \varepsilon_\psi(t_\psi))^{-1})$$

According to Lemma 4.3.9, one obtains that

$$(Z_\gamma, \varepsilon_{\gamma+\delta}(1)) = \varepsilon_{\gamma+\delta}(\pm(a-1))\varepsilon_\delta(\pm c)$$

and this is clearly an element of $U^+(E_6, R)$. But according to Proposition 2.2.13, the

commutator $(\varepsilon_{\gamma+\delta}(1), (\prod_{\psi\in S'}\varepsilon_\psi(t_\psi))^{-1})$ is an element of

$$\prod_{\psi\in E_6^+,\mathrm{wt}(\psi)\geq 3}\varepsilon_\psi(R).$$

So

$$(Z_\gamma, \varepsilon_{\gamma+\delta}(1))\cdot(\varepsilon_{\gamma+\delta}(1), (\prod_{\psi\in S'}\varepsilon_\psi(t_\psi))^{-1})$$

is an element of $U^+(E_6,R)$ and its terms can be ordered in such a way that the only root elements $\varepsilon_\psi(y_\psi)$ with $\mathrm{wt}(\psi)\leq 2$ appearing in it, are $\varepsilon_{\gamma+\delta}(\pm(a-1))$ and $\varepsilon_\delta(\pm c)$. Thus the commutator $(B, \varepsilon_{\gamma+\delta}(1))$ is conjugate to an element $u_{\gamma+\delta}$ of the subgroup $U^+(E_6,R)$. Hence using Lemma 4.4.6, we can find an ideal

$$I_{\gamma+\delta}(A)\subset\varepsilon(u_{\gamma+\delta}, 60^{2^{10}})\subset\varepsilon(A, 4\cdot 60^{2^{10}})$$

such that $\pi_{I_{\gamma+\delta}(A)}(u_{\gamma+\delta})=1$ and additionally $a-1, c$ are elements of $I_{\gamma+\delta}(A)$ as well. Similarly, one can find an ideal

$$I_{\gamma+\phi}(A)\subset\varepsilon(A, 4\cdot 60^{2^{10}})$$

such that $\pi_{I_{\gamma+\phi}(A)}(u_{\gamma+\phi})=1$ for an element $u_{\gamma+\phi}$ conjugate to the commutator $(B, \varepsilon_{\gamma+\phi}(1))$.

Further, we may reorder the terms in B in such a way that

$$B=(\prod_{\psi\in S'-\{\phi\}}\varepsilon_\psi(t'_\psi))\cdot\varepsilon_\phi(t_\phi)\cdot Z_\gamma$$

holds for other $t'_\psi\in R$ for $\psi\in S'-\{\phi\}$. Then for $x\in R$, consider the commutator

$$\begin{aligned}(B, \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+\phi}(x)) &= (\varepsilon_\phi(t_\phi)\cdot Z_\gamma, \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+\phi}(x))^{(\prod_{\psi\in S'-\{\phi\}}\varepsilon_\psi(t'_\psi))}\\ &\quad\cdot((\prod_{\psi\in S'-\{\phi\}}\varepsilon_\psi(t'_\psi)), \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+\phi}(x))\\ &= (\varepsilon_\phi(t_\phi)\cdot Z_\gamma, \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+\phi}(x))^{(\prod_{\psi\in S'-\{\phi\}}\varepsilon_\psi(t'_\psi))}\cdot 1\\ &\sim (\varepsilon_\phi(t_\phi)\cdot Z_\gamma, \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+\phi}(x))\\ &= (Z_\gamma, \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+\phi}(x))^{\varepsilon_\phi(t_\phi)}\cdot(\varepsilon_\phi(t_\phi), \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+\phi}(x))\end{aligned}$$

But note that according to Lemma 4.3.9, we obtain

$$(Z_\gamma, \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+\phi}(x))^{\varepsilon_\phi(t_\phi)}=(\varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+\phi}(\pm x(a-1))\varepsilon_{\alpha+2\beta+2\gamma+2\delta+\epsilon+\phi}(\pm xc))^{\varepsilon_\phi(t_\phi)}$$

However, we already know that

$$x(a-1), xc \in I_{\gamma+\delta}(A) \subset \varepsilon(A, 4 \cdot 60^{2^{10}})$$

and thus $(Z_\gamma, \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+\phi}(x))^{\varepsilon_\phi(t_\phi)}$ is an element of $B_A(8 \cdot 60^{2^{10}})$. Hence

$$(\varepsilon_\phi(t_\phi), \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+\phi}(x)) = \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+2\phi}(\pm x t_\phi)$$

and so in particular $\varepsilon_\phi(xt_\phi)$ is an element of $B_A(4 + 8 \cdot 60^{2^{10}})$.

In particular, this implies that

$$\begin{aligned} B' &:= (\prod_{\psi \in S' - \{\phi\}} \varepsilon_{w_\phi(\psi)}(\pm t_\psi)) \cdot w_\phi Z_\gamma w_\phi^{-1} \\ &= (\prod_{\psi \in S' - \{\phi\}} \varepsilon_{w_\phi(\psi)}(\pm t'_\psi)) \cdot Z_{\gamma+\phi} \end{aligned}$$

is an element of $B_A(6 + 8 \cdot 60^{2^{10}})$ for an appropriately chosen $Z_{\gamma+\phi} \in G_{\gamma+\phi}(R)$. Note that none of the roots $w_\phi(\psi)$ are negative for $\psi \in S' - \{\phi\}$. Thus $(\prod_{\psi \in S' - \{\phi\}} \varepsilon_{w_\phi(\psi)}(\pm t'_\psi))$ is in fact an element of $U^+(E_6, R)$. Hence according to [41, Chapter 3, p. 22, Corollary 4], the commutator

$$(\varepsilon_\phi(1), (\prod_{\psi \in S' - \{\phi\}} \varepsilon_{w_\phi(\psi)}(\pm t'_\psi))^{-1})$$

is an element of the subgroup $\prod_{\psi \in \Phi^+, \mathrm{wt}(\psi) \geq 2} \varepsilon_\psi(R)$ of $U^+(\Phi, R)$. So in particular,

$$w_\gamma \left(\varepsilon_\phi(1), (\prod_{\psi \in S' - \{\phi\}} \varepsilon_{w_\phi(\psi)}(\pm t'_\psi))^{-1} \right) w_\gamma^{-1}$$

is also an element of $U^+(E_6, R)$. Also Lemma 4.3.9 implies that there are $u, v \in R$ with

$$\begin{aligned} (B', \varepsilon_\phi(1)) &\sim (Z_{\gamma+\phi}, \varepsilon_\phi(1)) \cdot (\varepsilon_\phi(1), (\prod_{\psi \in S' - \{\phi\}} \varepsilon_{w_\phi(\psi)}(\pm t'_\psi))^{-1}) \\ &= \varepsilon_\phi(u)\varepsilon_{-\gamma}(v) \cdot (\varepsilon_\phi(1), (\prod_{\psi \in S' - \{\phi\}} \varepsilon_{w_\phi(\psi)}(\pm t'_\psi))^{-1}) \\ &\sim w_\gamma \varepsilon_\phi(u)\varepsilon_{-\gamma}(v) \cdot (\varepsilon_\phi(1), (\prod_{\psi \in S' - \{\phi\}} \varepsilon_{w_\phi(\psi)}(\pm t'_\psi))^{-1}) w_\gamma^{-1} \\ &= \varepsilon_{\gamma+\phi}(u)\varepsilon_\gamma(v) \cdot w_\gamma \left(\varepsilon_\phi(1), (\prod_{\psi \in S' - \{\phi\}} \varepsilon_{w_\phi(\psi)}(\pm t'_\psi))^{-1} \right) w_\gamma^{-1} := B'' \end{aligned}$$

So summarizing, the term B'' is an element of $U^+(E_6, R)$ and hence according to Lemma 4.4.6,

we can find an ideal

$$J(A) \subset \varepsilon(B'', 60^{2^{10}}) \subset \varepsilon(A, (6 + 8 \cdot 60^{2^{10}}) \cdot 60^{2^{10}}) \subset \varepsilon(A, 9 \cdot 60^{2^{11}})$$

such that $\pi_{J(A)}((B', \varepsilon_\phi(1))) = 1$. Next, we define the ideal

$$\begin{aligned} I_0(A) &:= I_\alpha(A) + I_\beta(A) + I_\delta(A) + I_\epsilon(A) + I_\phi(A) + I_{\gamma+\delta}(A) + I_{\gamma+\phi}(A) + J(A) \\ &\subset \varepsilon(A, 5 \cdot 4 \cdot 60^{2^{10}} + 2 \cdot 4 \cdot 60^{2^{10}} + 9 \cdot 60^{2^{11}}) \subset \varepsilon(A, 10 \cdot 60^{2^{11}}). \end{aligned}$$

To finish the proof, we now have to show that $I_0(A)$ has the desired property, that is for any maximal ideal m containing $I_0(A)$, one has

$$\pi_m(A, \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+2\phi}(1)) = 1$$

or equivalently that $\bar{B} := \pi_m(B) = 1$. To this end, consider the field $K := R/m$. First, we will show that $\pi_m(Z_\gamma)$ is trivial. Recall that:

$$Z_\gamma = \psi_\gamma \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

and observe that according to Lemma 4.3.9, one has

$$\begin{aligned} (B, \varepsilon_\beta(1)) &= ((\prod_{\psi \in S'} \varepsilon_\psi(t_\psi)) \cdot Z_\gamma, \varepsilon_\beta(1)) \\ &\sim (Z_\gamma, \varepsilon_\beta(1)) \cdot (\varepsilon_\beta(1), (\prod_{\psi \in S'} \varepsilon_\psi(t_\psi))^{-1}) \\ &= \varepsilon_{\beta+\gamma}(\pm b)\varepsilon_\beta(\pm(d-1)) \cdot (\varepsilon_\beta(1), (\prod_{\psi \in S'} \varepsilon_\psi(t_\psi))^{-1}) \\ &\sim \varepsilon_\beta(\pm(d-1)) \cdot (\varepsilon_\beta(1), (\prod_{\psi \in S'} \varepsilon_\psi(t_\psi))^{-1})\varepsilon_{\beta+\gamma}(\pm b) \end{aligned}$$

holds. Consider the set

$$T := S' - \{\beta + \gamma\} = E_6^+ - \{\alpha, \gamma, \epsilon, \beta + \gamma\}$$

of roots. Observe, that if ψ_1 is an element of T and ψ_2 is an element of S' such that $\psi_1 + \psi_2$ is an element of E_6, then $\psi_1 + \psi_2$ cannot be $\beta + \gamma$, because if it were than either ψ_1 or ψ_2 would have to be γ, which contradicts ψ_1, ψ_2 being elements of S'. Hence T is an *ideal* in the set S' and hence according to Proposition 2.2.13, the subgroup $\prod_{\psi \in T} \varepsilon_\psi(R)$ is a normal subgroup of $\prod_{\psi \in S'} \varepsilon_\psi(R)$. Thus as $\varepsilon_\beta(1)$ is an element of $\prod_{\psi \in T} \varepsilon_\psi(R)$, we obtain

that the commutator $(\varepsilon_\beta(1), (\prod_{\psi\in S'}\varepsilon_\psi(t_\psi))^{-1})$ is an element of the subgroup

$$\prod_{\psi\in T}\varepsilon_\psi(R)$$

of $U^+(E_6,R)$ as well. But consequently, the factors of $\varepsilon_\beta(\pm(d-1))(\varepsilon_\beta(1), (\prod_{\psi\in S'}\varepsilon_\psi(t_\psi))^{-1})$ can be rearranged in such a way that none of the roots appearing are $\beta+\gamma$. But remember that

$$\varepsilon_\beta(\pm(d-1))(\varepsilon_\beta(1), (\prod_{\psi\in S'}\varepsilon_\psi(t_\psi))^{-1})\varepsilon_{\beta+\gamma}(\pm b) \tag{4.5}$$

is conjugate to $(B, \varepsilon_\beta(1))$ and hence as m contains $I_\beta(A)$, the map π_m must map

$$\varepsilon_\beta(\pm(d-1))(\varepsilon_\beta(1), (\prod_{\psi\in S'}\varepsilon_\psi(t_\psi))^{-1})\varepsilon_{\beta+\gamma}(\pm b)$$

to the identity. This then implies that $\pi_m(\varepsilon_{\beta+\gamma}(\pm b))$ must be trivial as well and hence b must be an element of m. We have seen before already, that the ideal $I_{\gamma+\delta}(A)$ contains $a-1$ and c and so $a-1$ and c are elements of m. But

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

is an element of $\mathrm{SL}_2(R)$ and hence $d-1$ must be an element of m as well. Thus $\pi_m(Z_\gamma) = 1$.

Phrased differently, we obtain setting $\bar{t}_\psi := t_\psi + m \in K$ for $\psi \in S'$ that

$$\bar{B} = \prod_{\psi\in S'}\varepsilon_\psi(\bar{t}_\psi).$$

Next, we show by induction on $k \in \mathbb{N}_0$ that $\bar{t}_\psi = 0$ holds for all $\psi \in S'$ with $\mathrm{wt}(\psi) < k$. This implies $\bar{B} = 1$, because there is a maximal weight roots can have. So assume that $\bar{B}$ is given such that $\bar{t}_\psi = 0$ holds for all $\psi \in S'$ with $\mathrm{wt}(\psi) < k$. We have to show that $\bar{B}$ further has the property that $\bar{t}_\psi = 0$ holds for all $\psi \in S'$ with $\mathrm{wt}(\psi) = k$. To this end observe that for each root $\theta \in S'$ with weight k such that θ is not the root of highest weight in E_6^+, there must be a simple root θ_0 such that $\theta + \theta_0$ is a root in E_6^+. We distinguish two cases. First, let us assume that θ_0 can be chosen to not be γ. Then reordering the terms in $\bar{B}$, we can assume that

$$\bar{B} = \left(\prod_{\psi\in S'-\{\theta\},\mathrm{wt}(\psi)\geq k}\varepsilon_\psi(\bar{t}'_\psi)\right)\cdot\varepsilon_\theta(\bar{t}_\theta)$$

for certain other $\bar{t}'_\psi \in K$. But we know that $\bar{B}$ commutes with $\varepsilon_{\theta_0}(1)$ as $I_{\theta_0}(A)$ is a subset

of m. Hence

$$\begin{aligned} 1 = (\bar{B}, \varepsilon_{\theta_0}(1)) &\sim \left(\left(\prod_{\psi \in S' - \{\theta\}, \mathrm{wt}(\psi) \geq k} \varepsilon_\psi(\bar{t}'_\psi) \right) \cdot \varepsilon_\theta(\bar{t}_\theta), \varepsilon_{\theta_0}(1) \right) \\ &= (\varepsilon_\theta(\bar{t}_\theta), \varepsilon_{\theta_0}(1)) \cdot (\varepsilon_{\theta_0}(1), \left(\prod_{\psi \in S' - \{\theta\}, \mathrm{wt}(\psi) \geq k} \varepsilon_\psi(\bar{t}'_\psi) \right)^{-1}) \\ &= \varepsilon_{\theta+\theta_0}(\bar{t}_\theta) \cdot (\varepsilon_{\theta_0}(1), \left(\prod_{\psi \in S' - \{\theta\}, \mathrm{wt}(\psi) \geq k} \varepsilon_\psi(\bar{t}'_\psi) \right)^{-1}) \end{aligned}$$

This is an element of $U^+(E_6, K)$ again. But similar to how we showed that $\pi_m(Z_\gamma) = 1$, one can also show using Proposition 2.2.13 that the factors of

$$(\varepsilon_{\theta_0}(1), \left(\prod_{\psi \in S' - \{\theta\}, \mathrm{wt}(\psi) \geq k} \varepsilon_\psi(\bar{t}'_\psi) \right)^{-1})$$

can be rearranged as to not involve the root $\theta + \theta_0$ and hence one obtains that $\bar{t}_\theta = 0$ holds. To summarize, if there is a positive, simple root $\theta_0 \neq \gamma$ such that $\theta + \theta_0$ is a root, then $\bar{t}_\theta = 0$ holds. This settles the first case. The second case is that such a simple root θ_0 cannot be found. But looking again at the Hasse-diagram from the proof of Lemma 4.4.4 in Appendix C, the only positive roots θ in E_6 of this form are the roots

$$\phi, \alpha + \beta, \delta + \epsilon, \alpha + \beta + \gamma + \delta + \epsilon + \phi, \alpha + 2\beta + 2\gamma + 2\delta + \epsilon + \phi$$

and the root of highest weight χ. Disregarding χ for the moment, all the other ones of those roots θ, have one of two properties: either $\theta + \gamma + \delta$ or $\theta + \gamma + \phi$ is a root in E_6^+. But we know by construction of $I_0(A)$, that $\bar{B}$ commutes with both $\varepsilon_{\gamma+\delta}(1)$ and $\varepsilon_{\gamma+\phi}(1)$. Further, we can reorder the terms of $\bar{B}$ again such that

$$\bar{B} = \left(\prod_{\psi \in S' - \{\theta\}, \mathrm{wt}(\psi) \geq k} \varepsilon_\psi(\bar{t}'_\psi) \right) \cdot \varepsilon_\theta(\bar{t}_\theta)$$

and hence using Proposition 2.2.13 as before we obtain $\bar{t}_\theta = 0$ using either that $\bar{B}$ centralizes $\varepsilon_{\gamma+\delta}(1)$ or centralizes $\varepsilon_{\gamma+\phi}(1)$ depending on the θ in question. So proceeding by induction, we can assume finally that

$$\pi_m(B) = \bar{B} = \varepsilon_\chi(\bar{t}_\chi).$$

Recall the aforementioned element

$$\begin{aligned} B' &:= (\prod_{\psi\in S'-\{\phi\}} \varepsilon_{w_\phi(\psi)}(\pm t_\psi)) \cdot w_\phi Z_\gamma w_\phi^{-1} \\ &= (\prod_{\psi\in S'-\{\phi\}} \varepsilon_{w_\phi(\psi)}(\pm t'_\psi)) \cdot Z_{\gamma+\phi} \end{aligned}$$

and observe that $\bar{t}_\psi = 0$ for all $\psi \neq \chi$ and $\pi_m(Z_\gamma) = 1$, implies

$$\begin{aligned} \pi_m(B') &= \pi_m((\prod_{\psi\in S'-\{\phi\}} \varepsilon_{w_\phi(\psi)}(\pm t_\psi)) \cdot w_\phi Z_\gamma w_\phi^{-1}) \\ &= (\prod_{\psi\in S'-\{\phi\}} \varepsilon_{w_\phi(\psi)}(\pm \bar{t}_\psi)) \cdot w_\phi \cdot 1 \cdot w_\phi^{-1} \\ &= \varepsilon_{w_\phi(\chi)}(\pm \bar{t}_\chi)) = \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+\phi}(\pm \bar{t}_\chi). \end{aligned}$$

However, the ideal $I_0(A)$ contains $J(A)$ and hence $\pi_m(B')$ commutes with $\varepsilon_\phi(1)$ by construction of $J(A)$. Thus

$$1 = (\pi_m(B'), \varepsilon_\phi(1)) = (\varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+\phi}(\pm \bar{t}_\chi), \varepsilon_\phi(1)) = \varepsilon_\chi(\pm \bar{t}_\chi)$$

holds and consequently $\bar{t}_\chi = 0$ follows. So finally, we are done with the induction and the proof of the lemma. □

This proposition implies:

Proposition 4.4.7. *Let R be a principal ideal domain and $A \in E_6(R)$. Then there is an ideal $I_0(A)$ in R with*

$$I(A) \subset \varepsilon(A, 120 \cdot 60^{2^{11}})$$

such that $V(I(A)) \subset \Pi(\{A\})$. Phrased differently, for R a principal ideal domain, one can can pick $L(E_6)$ in Theorem 3.1.1 as $L(E_6) = 120 \cdot 60^{2^{11}}$.

Proof. To define the ideal $I(A)$, note first that according to Proposition A.0.8, for each simple root $\theta \in E_6$, positive and negative, there is an element $w^{(\theta)} \in W(E_6)$ such that $(w^{(\theta)})^{-1}(\chi) = \theta$ holds for χ the positive root of highest weight in E_6. Then consider the ideals

$$\begin{aligned} I_0\left(w^{(\alpha)}A(w^{(\alpha)})^{-1}\right), \ldots, I_0\left(w^{(\phi)}A(w^{(\phi)})^{-1}\right), I_0\left(w^{(-\alpha)}A(w^{(-\alpha)})^{-1}\right), \ldots, I_0\left(w^{(-\phi)}A(w^{(-\phi)})^{-1}\right) \\ \subset \varepsilon(A, 10 \cdot 60^{2^{11}}) \end{aligned}$$

from Proposition 4.4.1 and set

$$I(A) := \sum_{\theta\in E_6 \text{ simple}} I_0\left(w^{(\theta)}A(w^{(\theta)})^{-1}\right) \subset \varepsilon(A, 120 \cdot 60^{2^{11}}).$$

Let m be a maximal ideal, that contains $I(A)$ and define $K := R/m$. Further let $\pi_m : E_6(R) \to E_6(K)$ be the reduction homomorphism and set $\bar{A} := \pi_m(A)$. We will prove that $\bar{A}$ is a central element of $E_6(K)$. The key observation is that $\bar{A}$ commutes with all elements $\varepsilon_\theta(1)$ for θ a simple root in E_6: To see this note first that $w^{(\theta)}\bar{A}(w^{(\theta)})^{-1}$ commutes with $\varepsilon_\chi(1)$, because m contains $I_0\left(w^{(\theta)}A(w^{(\theta)})^{-1}\right)$. But conjugating, this implies that $\bar{A}$ commutes with

$$(w^{(\theta)})^{-1}\varepsilon_\chi(1)w^{(\theta)} = \varepsilon_{(w^{(\theta)})^{-1}(\chi)}(\pm 1) = \varepsilon_\theta(\pm 1).$$

Observe next that K is a field. Thus using the Bruhat-decomposition of $E_6(K)$ [41, Chapter 3, p. 26, Theorem 4'], there are $w \in W(E_6)$, an element $b \in B(K)$ and an element $u \in U^+(E_6, K)$ such that wuw^{-1} is an element of $U^-(E_6, K)$ and $A = bwu$. Let us first assume that $w \neq 1$. Then according to [41, Appendix, p. 151, (2)Corollary], there must be a positive simple root θ such that $w(\theta)$ is a negative root. However $\bar{A}$ commutes with $\varepsilon_\theta(1)$ and hence

$$\begin{aligned} 1 = (\bar{A}, \varepsilon_\theta(1)) &= (bwu, \varepsilon_\theta(1)) \sim (wu, \varepsilon_\theta(1)) \cdot (\varepsilon_\theta(1), b^{-1}) \\ &= (u, \varepsilon_\theta(1))^w \cdot (w, \varepsilon_\theta(1)) \cdot (\varepsilon_\theta(1), b^{-1}) \\ &= (u, \varepsilon_\theta(1))^w \cdot w\varepsilon_\theta(1)w^{-1}\varepsilon_\theta(-1) \cdot (\varepsilon_\theta(1), b^{-1}) \\ &= [(u, \varepsilon_\theta(1)) \cdot \varepsilon_\theta(1)]^w \cdot \left[\varepsilon_\theta(-1) \cdot (\varepsilon_\theta(1), b^{-1})\right] \end{aligned}$$

But consider the first factor

$$[(u, \varepsilon_\theta(1)) \cdot \varepsilon_\theta(1)]^w = \left[u\varepsilon_\theta(1)u^{-1}\varepsilon_\theta(-1)\varepsilon_\theta(1)\right]^w = u^w \cdot \varepsilon_\theta(1)^w(u^{-1})^w = u^w\varepsilon_{w(\theta)}(\pm 1)(u^{-1})^w.$$

and note that by assumption u^w is an element of $U^-(E_6, K)$ and hence the entire first factor is an element of $U^-(E_6, K)$. But the second factor

$$\left[\varepsilon_\theta(-1) \cdot (\varepsilon_\theta(1), b^{-1})\right]$$

is an element of $B(K)$ as all of its factors are. So

$$[(u, \varepsilon_\theta(1)) \cdot \varepsilon_\theta(1)]^w \cdot \left[\varepsilon_\theta(-1) \cdot (\varepsilon_\theta(1), b^{-1})\right]$$

is the trivial element in $U^-(E_6, K) \cdot B(K)$ and hence the uniqueness of the Big-cell-decomposition [41, Chapter 5, p. 40, Theorem 7] implies

$$1 = [(u, \varepsilon_\theta(1)) \cdot \varepsilon_\theta(1)] = \left[\varepsilon_\theta(-1) \cdot (\varepsilon_\theta(1), b^{-1})\right].$$

But this implies

$$1 = \varepsilon_\theta(-1) \cdot (\varepsilon_\theta(1), b^{-1}) = \varepsilon_\theta(-1)\varepsilon_\theta(1)b^{-1}\varepsilon_\theta(-1)b = b^{-1}\varepsilon_\theta(-1)b \sim \varepsilon_\theta(-1)$$

and hence $\varepsilon_\theta(-1)$ is the trivial element, which is obviously impossible. This contradiction yields $w = 1$ and hence $\bar{A}$ is an element of the upper Borel subgroup $B^+(E_6, K)$. But using the Bruhat-decomposition for the lower Borel-subgroup $B^-(E_6, K)$, one can show in the same fashion that $\bar{A}$ is also an element of $B^-(E_6, K)$. However

$$B^+(E_6, K) \cap B^-(E_6, K) = \{h_\alpha(s_\alpha)h_\beta(s_\beta)h_\gamma(s_\gamma)h_\delta(s_\delta)h_\epsilon(s_\epsilon)h_\phi(s_\phi) |\ s_\alpha, s_\beta, s_\gamma, s_\delta, s_\epsilon, s_\phi \in K - \{0\}\}.$$

holds and so we can pick $s_\alpha, s_\beta, s_\gamma, s_\delta, s_\epsilon, s_\phi \in K - \{0\}$ with

$$\bar{A} = h_\alpha(s_\alpha)h_\beta(s_\beta)h_\gamma(s_\gamma)h_\delta(s_\delta)h_\epsilon(s_\epsilon)h_\phi(s_\phi).$$

Next, observe that $\bar{A}$ commutes with $\varepsilon_\theta(1)$ for all positive, simple roots θ. Then using induction on the height of a root $\theta \in E_6^+$ and the commutator formulas in Lemma 2.2.4, one can show that $\bar{A}$ also commutes with $\varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+2\phi}(1)$. But $\varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+2\phi}(1)$ commutes with all factors in $\bar{A}$ except possibly $h_\phi(s_\phi)$ and hence the following holds:

$$\begin{aligned} 1 &= (\bar{A}, \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+2\phi}(1)) \sim (h_\phi(s_\phi), \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+2\phi}(1)) \\ &= \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+2\phi}(s_\phi^{\langle \alpha+2\beta+3\gamma+2\delta+\epsilon+2\phi, \phi \rangle} - 1) \\ &= \varepsilon_{\alpha+2\beta+3\gamma+2\delta+\epsilon+2\phi}(s_\phi - 1). \end{aligned}$$

But this implies that $s_\phi = 1$. Further, $\bar{A}$ commutes with $\varepsilon_\phi(1)$. But all factors of $\bar{A}$ commute with $\varepsilon_\phi(1)$ except possibly $h_\gamma(s_\gamma)$. This implies

$$1 = (\bar{A}, \varepsilon_\phi(1)) = (h_\gamma(s_\gamma), \varepsilon_\phi(1)) = \varepsilon_\phi(s_\gamma^{\langle \phi, \gamma \rangle} - 1) = \varepsilon_\phi(s_\gamma^{-1} - 1).$$

But this in turn implies that $s_\gamma = 1$. But if both s_ϕ and s_γ are 1, then observe that all factors in

$$\bar{A} = h_\alpha(s_\alpha)h_\beta(s_\beta)h_\delta(s_\delta)h_\epsilon(s_\epsilon)$$

commute with the root subgroups $\varepsilon_\phi(K)$ and $\varepsilon_{-\phi}(K)$ and so $\bar{A}$ centralizes those root subgroups.

Next, observe that the only factors of $\bar{A}$ that might not commute with $\varepsilon_\gamma(K)$ are

$h_\beta(s_\beta)$ and $h_\delta(s_\delta)$. This yields for $x \in K$ that

$$\begin{aligned}(\bar{A}, \varepsilon_\gamma(x)) &= (h_\beta(s_\beta)h_\delta(s_\delta), \varepsilon_\gamma(x)) \\ &= \varepsilon_\gamma(x s_\delta^{\langle\gamma,\delta\rangle} s_\beta^{-\langle\gamma,\beta\rangle} - 1) \\ &= \varepsilon_\gamma(x(s_\delta^{-1}s_\beta - 1))\end{aligned}$$

But observe that $\bar{A}$ commutes with $\varepsilon_\gamma(1)$ and hence setting $x = 1$, yields $s_\delta^{-1}s_\beta = 1$. This yields $(\bar{A}, \varepsilon_\gamma(x)) = 1$ for all $x \in K$. Hence $\bar{A}$ centralizes the root subgroup $\varepsilon_\gamma(K)$ and similarly $\bar{A}$ centralizes the root subgroup $\varepsilon_{-\gamma}(K)$. But analyzing the commutators $(\bar{A}, \varepsilon_\theta(x))$ for $\theta \in \{\pm\alpha, \pm\beta, \pm\delta, \pm\epsilon\}$ and $x \in R$ in the same manner yields that $\bar{A}$ also centralizes the root groups $\varepsilon_\theta(K)$.

To summarize: For each simple root $\theta \in E_6$ the element $\bar{A}$ centralizes the corresponding root subgroup $\varepsilon_\theta(K)$. But the group $E_6(K)$ is generated by

$$\{\varepsilon_\theta(x) |\ x \in K, \theta \in E_6 \text{ simple}\}$$

according to [41, Chapter 3, p. 24, Lemma 26] and hence $\bar{A}$ is a central element of $E_6(K)$. This finishes the proof. □

4.5 Explicit bounds for root elements of $G_2(R)$

Remember that the positive roots in G_2 are $\alpha, \beta, \alpha+\beta, 2\alpha+\beta, 3\alpha+\beta$ and $3\alpha+2\beta$ for α, β simple, positive roots in G_2 with α short and β long. Also note that the roots $3\alpha+\beta$ and β span a root subsystem of G_2 isomorphic to A_2. We will use the Bruhat decomposition Proposition 4.3.4 from Section 4.3. First, we need the longest element of the Weyl group:

Lemma 4.5.1. *The longest element in the Weyl group $W(G_2)$ with respect to the fundamental reflections $F := \{w_\alpha, w_\beta\}$ is $(w_\alpha w_\beta)^3$ and this product gives a minimal expression for w_0 with respect to F.*

Proof. As mentioned in the proof of Lemma 4.4.3, it suffices to show that all roots in G_2^+ get mapped by $(w_\alpha w_\beta)^3$ to negative roots to show that $(w_\alpha w_\beta)^3$ is the longest element in $W(G_2)$. Hence it suffices to show that $(w_\alpha w_\beta)^3(\alpha)$ and $(w_\alpha w_\beta)^3(\beta)$ are negative roots. But observe that

$$w_\alpha w_\beta(\alpha) = w_\alpha(\alpha+\beta) = 2\alpha+\beta \text{ and } w_\alpha w_\beta(\beta) = -w_\alpha(\beta) = -(3\alpha+\beta).$$

This then implies

$$\begin{aligned}(w_\alpha w_\beta)^3(\alpha) = (w_\alpha w_\beta)^2(2\alpha+\beta) &= w_\alpha w_\beta(2\cdot(2\alpha+\beta) - 3\alpha - \beta) \\ &= w_\alpha w_\beta(\alpha+\beta) = 2\alpha+\beta-3\alpha-\beta = -\alpha\end{aligned}$$

and

$$\begin{aligned}(w_\alpha w_\beta)^3(\beta) &= -(w_\alpha w_\beta)^2(3\alpha+\beta) = -w_\alpha w_\beta(3\cdot(2\alpha+\beta)-3\alpha-\beta)\\ &= -w_\alpha w_\beta(3\alpha+2\beta) = -(3\cdot(2\alpha+\beta)-2\cdot(3\alpha-\beta)) = -\beta\end{aligned}$$

and so $(w_\alpha w_\beta)^3$ is the longest element w_0 of $W(G_2)$. Further, $(w_\alpha w_\beta)^3$ gives a minimal expression for w_0, because $l_F(w_0) = |G_2^+| = 6$ holds as mentioned in the proof of Lemma 4.4.3. □

Next, we prove:

Proposition 4.5.2. *Let R be a principal ideal domain and $A \in G_2(R)$. Then there is an ideal $I_0(A) \subset \varepsilon_l(A, 2343808)$ such that for each maximal ideal m in R with $I_0(A) \subset m$, the equation $\pi_m((A, \varepsilon_{3\alpha+2\beta}(1))) = 1$ holds.*

The proof is very long, so we put parts of it in technical lemmas.

Lemma 4.5.3. *Let R be a commutative ring with 1, $A \in G_2(R)$ and let $a, b \in R$ be given such that there is an $M \in \mathbb{N}$ with*

$$\|\varepsilon_\alpha(\pm a)\varepsilon_{2\alpha+\beta}(\pm a^2)\varepsilon_{3\alpha+\beta}(\pm b)\|_A \leq M.$$

Then

$$(2b) \subset \varepsilon_l(A, 18M)$$

holds.

Proof. First, observe for $x \in R$ that

$$\begin{aligned}B_A(2M) \ni &(\varepsilon_\alpha(\pm a)\varepsilon_{2\alpha+\beta}(\pm a^2)\varepsilon_{3\alpha+\beta}(\pm b), \varepsilon_{2\alpha+\beta}(x))\\ &= (\varepsilon_\alpha(\pm a), \varepsilon_{2\alpha+\beta}(x)) = \varepsilon_{3\alpha+\beta}(\pm 3ax)\end{aligned}$$

and hence $(3a) \subset \varepsilon_l(A, 2M)$.

Second, observe for $x \in R$ that

$$\begin{aligned}&B_A(2M) \ni (\varepsilon_\alpha(\pm a)\varepsilon_{2\alpha+\beta}(\pm a^2)\varepsilon_{3\alpha+\beta}(\pm b), \varepsilon_{\alpha+\beta}(x))\\ &= (\varepsilon_\alpha(\pm a)\varepsilon_{2\alpha+\beta}(\pm a^2), \varepsilon_{\alpha+\beta}(x))\\ &\sim (\varepsilon_{2\alpha+\beta}(\pm a^2), \varepsilon_{\alpha+\beta}(x)) \cdot (\varepsilon_{\alpha+\beta}(x), \varepsilon_\alpha(\pm a))\\ &= \varepsilon_{3\alpha+2\beta}(\pm 3a^2x)\varepsilon_{2\alpha+\beta}(\pm 2ax)\varepsilon_{3\alpha+\beta}(\pm 3a^2x)\varepsilon_{3\alpha+2\beta}(\pm 3ax^2)\\ &= \varepsilon_{3\alpha+2\beta}(\pm 3a^2x \pm 3ax^2)\varepsilon_{3\alpha+\beta}(\pm 3a^2x)\varepsilon_{2\alpha+\beta}(\pm 2ax)\end{aligned}$$

But from $(3a) \subset \varepsilon_l(A, 2M)$, we obtain

$$\|\varepsilon_{3\alpha+2\beta}(\pm 3a^2x \pm 3ax^2)\varepsilon_{3\alpha+\beta}(\pm 3a^2x)\|_A \leq 4M$$

and hence $\|\varepsilon_{2\alpha+\beta}(\pm 2ax)\|_A \leq 6M$ holds for all $x \in R$. This implies $(2a) \subset \varepsilon(A, \alpha, 6M)$.

Third, consider

$$(\varepsilon_\alpha(\pm a)\varepsilon_{2\alpha+\beta}(\pm a^2)\varepsilon_{3\alpha+\beta}(\pm b))^2 = \varepsilon_\alpha(\pm 2a)\varepsilon_{2\alpha+\beta}(\pm 2a^2)\varepsilon_{3\alpha+\beta}(\pm 2b \pm 3a^3)$$

Observe that $\|\varepsilon_\alpha(2a)\|_A \leq 6M$. Hence $\|\varepsilon_{2\alpha+\beta}(\pm 2a^2)\varepsilon_{3\alpha+\beta}(\pm 2b \pm 3a^3)\|_A \leq 8M$ holds. So considering the commutator

$$B_A(16M) \ni (\varepsilon_{2\alpha+\beta}(\pm 2a^2)\varepsilon_{3\alpha+\beta}(\pm 2b \pm 3a^3), \varepsilon_\beta(x)) = \varepsilon_{3\alpha+2\beta}(x(\pm 2b \pm 3a^3))$$

and remembering $\|\varepsilon_{3\alpha+2\beta}(3xa^3)\|_A \leq 2M$ implies $(2b) \subset \varepsilon_l(A, 18M)$. This finishes the proof of the lemma. □

Next, we will prove the following lemma:

Lemma 4.5.4. *Let R be a commutative ring with 1, $A \in G_2(R)$ and let $a, b \in R$ be given such that there is an $M \in \mathbb{N}$ with*

$$\|\varepsilon_\alpha(\pm a)\varepsilon_{2\alpha+\beta}(\pm a^2)\varepsilon_{3\alpha+\beta}(\pm b)\|_A \leq M.$$

Then

$$(3b^2) \subset \varepsilon_l(A, 192M).$$

Proof. First, consider the commutator

$$\begin{aligned}
&(\varepsilon_\alpha(\pm a)\varepsilon_{2\alpha+\beta}(\pm a^2)\varepsilon_{3\alpha+\beta}(\pm b), \varepsilon_{-\alpha}(1)) \\
&\sim (\varepsilon_{2\alpha+\beta}(\pm a^2)\varepsilon_{3\alpha+\beta}(\pm b), \varepsilon_{-\alpha}(1)) \cdot (\varepsilon_{-\alpha}(1), \varepsilon_\alpha(\pm a)) \\
&= (\varepsilon_{3\alpha+\beta}(\pm b), \varepsilon_{-\alpha}(1))^{\varepsilon_{2\alpha+\beta}(\pm a^2)} \cdot (\varepsilon_{2\alpha+\beta}(\pm a^2), \varepsilon_{-\alpha}(1)) \cdot (\varepsilon_{-\alpha}(1), \varepsilon_\alpha(\pm a)) \\
&= (\varepsilon_{2\alpha+\beta}(\pm b)\varepsilon_{\alpha+\beta}(\pm b)\varepsilon_\beta(\pm b)\varepsilon_{3\alpha+2\beta}(\pm b^2))^{\varepsilon_{2\alpha+\beta}(\pm a^2)} \\
&\quad \cdot \varepsilon_{\alpha+\beta}(\pm 2a^2)\varepsilon_\beta(\pm 3a^2)\varepsilon_{3\alpha+2\beta}(\pm 3a^4) \cdot (\varepsilon_{-\alpha}(1), \varepsilon_\alpha(\pm a)) \\
&= \big(\varepsilon_{2\alpha+\beta}(\pm b)\varepsilon_{\alpha+\beta}(\pm b)\varepsilon_{3\alpha+2\beta}(\pm 3a^2 b)\varepsilon_\beta(\pm b)\varepsilon_{3\alpha+2\beta}(\pm b^2)\big) \\
&\quad \cdot \varepsilon_{\alpha+\beta}(\pm 2a^2)\varepsilon_\beta(\pm 3a^2)\varepsilon_{3\alpha+2\beta}(\pm 3a^4) \cdot (\varepsilon_{-\alpha}(1), \varepsilon_\alpha(\pm a)) \\
&= \varepsilon_{2\alpha+\beta}(\pm b)\varepsilon_{\alpha+\beta}(\pm b \pm 2a^2)\varepsilon_{3\alpha+2\beta}(\pm 3a^2 b \pm 3a^4 \pm b^2)\varepsilon_\beta(\pm b \pm 3a^2) \cdot (\varepsilon_{-\alpha}(1), \varepsilon_\alpha(\pm a)) \\
&\sim (\varepsilon_{-\alpha}(1), \varepsilon_\alpha(\pm a)) \cdot \varepsilon_{2\alpha+\beta}(\pm b)\varepsilon_{\alpha+\beta}(\pm b \pm 2a^2)\varepsilon_{3\alpha+2\beta}(\pm 3a^2 b \pm 3a^4 \pm b^2)\varepsilon_\beta(\pm b \pm 3a^2) =: B
\end{aligned}$$

But note that all factors of B besides $\varepsilon_{2\alpha+\beta}(\pm b)\varepsilon_{\alpha+\beta}(\pm b \pm 2a^2)$ and $(\varepsilon_{-\alpha}(1), \varepsilon_\alpha(\pm a))$

commute with $\varepsilon_{-\alpha}(1)$ and thus we obtain

$$\begin{aligned}
(B,\varepsilon_{-\alpha}(1)) &= \big((\varepsilon_{-\alpha}(1),\varepsilon_\alpha(\pm a))\cdot\varepsilon_{2\alpha+\beta}(\pm b)\varepsilon_{\alpha+\beta}(\pm b\pm 2a^2),\varepsilon_{-\alpha}(1)\big)\\
&\sim (\varepsilon_{2\alpha+\beta}(\pm b)\varepsilon_{\alpha+\beta}(\pm b),\varepsilon_{-\alpha}(1))\cdot\big(\varepsilon_{-\alpha}(1),(\varepsilon_{-\alpha}(1),\varepsilon_\alpha(\pm a))^{-1}\big)\\
&= (\varepsilon_{\alpha+\beta}(\pm b\pm 2a^2),\varepsilon_{-\alpha}(1))^{\varepsilon_{2\alpha+\beta}(\pm b)}(\varepsilon_{2\alpha+\beta}(\pm b),\varepsilon_{-\alpha}(1))\\
&\quad\cdot\big(\varepsilon_{-\alpha}(1),(\varepsilon_{-\alpha}(1),\varepsilon_\alpha(\pm a))^{-1}\big)\\
&= \varepsilon_\beta(3(\pm b\pm 2a^2))^{\varepsilon_{2\alpha+\beta}(\pm b)}\varepsilon_{\alpha+\beta}(\pm 2b)\varepsilon_\beta(\pm 3b)\varepsilon_{3\alpha+2\beta}(\pm 3b^2)\\
&\quad\cdot\big(\varepsilon_{-\alpha}(1),(\varepsilon_{-\alpha}(1),\varepsilon_\alpha(\pm a))^{-1}\big)\\
&= \varepsilon_\beta(3(\pm b\pm 2a^2))\varepsilon_{\alpha+\beta}(\pm 2b)\varepsilon_\beta(\pm 3b)\varepsilon_{3\alpha+2\beta}(\pm 3b^2)\\
&\quad\cdot\big(\varepsilon_{-\alpha}(1),(\varepsilon_{-\alpha}(1),\varepsilon_\alpha(\pm a))^{-1}\big)\\
&= \varepsilon_\beta\left(3(\pm b\pm 2a^2)\pm 3b\right)\varepsilon_{\alpha+\beta}(\pm 2b)\varepsilon_{3\alpha+2\beta}(\pm 3b^2)\cdot\big(\varepsilon_{-\alpha}(1),(\varepsilon_{-\alpha}(1),\varepsilon_\alpha(\pm a))^{-1}\big)
\end{aligned}$$

This is an element of $B_A(4M)$. Further, dependent on the respective signs the term $\varepsilon_\beta(3(\pm b\pm 2a^2)\pm 3b)$ is $\varepsilon_\beta(\pm 6b\pm 6a^2)$ or $\varepsilon_\beta(\pm 6a^2)$. But $\|\varepsilon_\beta(\pm 6b)\|_A\leq 18M$ holds according to Lemma 4.5.3 and we have seen $(3a)\subset\varepsilon_l(A,2M)$ in the proof of Lemma 4.5.3. Thus $\varepsilon_\beta(3(\pm b\pm 2a^2)\pm 3b)$ is an element of $B_A(20M)$. Hence

$$C:=\varepsilon_{\alpha+\beta}(\pm 2b)\varepsilon_{3\alpha+2\beta}(\pm 3b^2)\cdot(\varepsilon_{-\alpha}(1),(\varepsilon_{-\alpha}(1),\varepsilon_\alpha(\pm a))^{-1})\in B_A(20M+4M)=B_A(24M).$$

Note, that $(\varepsilon_{-\alpha}(1),(\varepsilon_{-\alpha}(1),\varepsilon_\alpha(\pm a))^{-1})$ commutes with $\varepsilon_{-3\alpha-2\beta}(1)$ according to Lemma 4.3.9(1) and hence

$$\begin{aligned}
(C,\varepsilon_{-3\alpha-2\beta}(1)) &= \big(\varepsilon_{\alpha+\beta}(\pm 2b)\varepsilon_{3\alpha+2\beta}(\pm 3b^2),\varepsilon_{-3\alpha-2\beta}(1)\big)\\
&\sim \big(\varepsilon_{3\alpha+2\beta}(\pm 3b^2),\varepsilon_{-3\alpha-2\beta}(1)\big)\cdot(\varepsilon_{-3\alpha-2\beta}(1),\varepsilon_{\alpha+\beta}(\pm 2b))\\
&= \big(\varepsilon_{3\alpha+2\beta}(\pm 3b^2),\varepsilon_{-3\alpha-2\beta}(1)\big)\cdot\varepsilon_{-2\alpha-\beta}(\pm 2b)\varepsilon_{-\alpha}(\pm 4b^2)\varepsilon_\beta(\pm 8b^3)\varepsilon_{-3\alpha-\beta}(\pm 8b^3)\\
&=: D.
\end{aligned}$$

Observe that $\|D\|_A\leq 48M$. Note next, that all factors in this product besides

$$\big(\varepsilon_{3\alpha+2\beta}(\pm 3b^2),\varepsilon_{-3\alpha-2\beta}(1)\big)$$

commute with $\varepsilon_\beta(1)$ and hence

$$\begin{aligned}
B_A(96M)\ni(D,\varepsilon_\beta(1)) &= ((\varepsilon_{3\alpha+2\beta}(\pm 3b^2),\varepsilon_{-3\alpha-2\beta}(1)),\varepsilon_\beta(1))\\
&= \varepsilon_{-3\alpha-\beta}(\pm 3b^2)\varepsilon_\beta\left((\pm 3b^2)\pm(\pm 3b^2)^2\right).
\end{aligned}$$

The last equation follows from Lemma 4.3.9(3). Lastly, consider for $x \in R$ the commutator

$$B_A(192M) \ni \left(\varepsilon_{-3\alpha-\beta}(\pm 3b^2)\varepsilon_\beta\left((\pm 3b^2) + (\pm 3b^2)^2\right), \varepsilon_{3\alpha+2\beta}(x)\right) = \varepsilon_\beta(\pm 3b^2 x).$$

This implies as $x \in R$ is arbitrary that

$$(3b^2) \subset \varepsilon(A, 3\alpha + 2\beta, 4 \cdot 48M) = \varepsilon_l(A, 192M)$$

and finishes the proof of the lemma. □

These two lemmata imply the following proposition:

Proposition 4.5.5. *Let R be a commutative ring with 1, $A \in G_2(R)$ and let $a, b \in R$ be given such that there is an $M \in \mathbb{N}$ with*

$$\|\varepsilon_\alpha(\pm a)\varepsilon_{2\alpha+\beta}(\pm a^2)\varepsilon_{3\alpha+\beta}(\pm b)\|_A \leq M.$$

Then

$$(b^2) \subset \varepsilon_l(A, 210M)$$

Proof. Lemma 4.5.3 implies

$$(2b^2) \subset (2b) \subset \varepsilon_l(A, 18M)$$

and Lemma 4.5.4 implies

$$(3b^2) \subset \varepsilon_l(A, 192M).$$

So we obtain as $b^2 = 3b^2 - 2b^2$ that

$$(b^2) \subset \varepsilon_l(A, 18M + 192M) = \varepsilon_l(A, 210M).$$

□

We further need the following lemma:

Lemma 4.5.6. *Let R be a principal ideal domain, $x \in R$ and let*

$$A = \psi_\beta \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G_\beta(R)$$

be given. Then

$$\begin{aligned}(A, \varepsilon_{\alpha+\beta}(x)) = {} & \varepsilon_{\alpha+\beta}(\pm x(a-1)) \cdot \varepsilon_{2\alpha+\beta}(\pm x^2 c(a-1)) \\ & \cdot \varepsilon_\alpha(\pm xc) \cdot \varepsilon_{2\alpha+\beta}(\pm cx^2) \cdot \varepsilon_{3\alpha+\beta}(\pm x^3 c^2(a-1) \pm c^2 x^3) \\ & \cdot \varepsilon_{3\alpha+2\beta}(\pm x^3 c(a-1)^2 \pm cx^3 \pm cx^3(a-1))\end{aligned}$$

and

$$\begin{aligned}(A, \varepsilon_{-\alpha}(x)) = \varepsilon_{-\alpha}(\pm x(a-1)) \cdot \varepsilon_{-2\alpha-\beta}(\pm cx^2(a-1)) \\ \cdot \varepsilon_{-\alpha-\beta}(\pm cx) \cdot \varepsilon_{-2\alpha-\beta}(\pm x^2 c) \cdot \varepsilon_{-3\alpha-\beta}(\pm cx^3(a-1)^2 \pm x^3 c \pm x^3 c(a-1)) \\ \cdot \varepsilon_{-3\alpha-2\beta}(\pm x^3 c^2 \pm c^2 x^3(a-1))\end{aligned}$$

hold.

Proof. For the first claim, we distinguish the cases $c \neq 0$ and $c = 0$. For the first case, note that we can write

$$A = \varepsilon_\beta(c^{-1}(a-1))\varepsilon_{-\beta}(c)\varepsilon_\beta(c^{-1}(d-1))$$

in $G_2(Q)$ with Q the fraction field of R. Then the following chain of equations holds:

$$\begin{aligned}(A, \varepsilon_{\alpha+\beta}(x)) &= \big(\varepsilon_\beta(c^{-1}(a-1))\varepsilon_{-\beta}(c)\varepsilon_\beta(c^{-1}(d-1)), \varepsilon_{\alpha+\beta}(x)\big) \\ &= \big(\varepsilon_{-\beta}(c)\varepsilon_\beta(c^{-1}(d-1)), \varepsilon_{\alpha+\beta}(x)\big)^{\varepsilon_\beta(c^{-1}(a-1))} \\ &= (\varepsilon_{-\beta}(c), \varepsilon_{\alpha+\beta}(x))^{\varepsilon_\beta(c^{-1}(a-1))} \\ &= \big(\varepsilon_\alpha(\pm xc)\varepsilon_{2\alpha+\beta}(\pm cx^2)\varepsilon_{3\alpha+2\beta}(\pm cx^3)\varepsilon_{3\alpha+\beta}(\pm c^2x^3)\big)^{\varepsilon_\beta(c^{-1}(a-1))} \\ &= \Big(\varepsilon_\alpha(\pm xc)^{\varepsilon_\beta(c^{-1}(a-1))}\Big)\, \varepsilon_{2\alpha+\beta}(\pm cx^2)\varepsilon_{3\alpha+2\beta}(\pm cx^3)\, \Big(\varepsilon_{3\alpha+\beta}(\pm c^2x^3)^{\varepsilon_\beta(c^{-1}(a-1))}\Big) \\ &= \varepsilon_{\alpha+\beta}(\pm x(a-1))\varepsilon_{2\alpha+\beta}(\pm x^2c(a-1))\varepsilon_{3\alpha+\beta}(\pm x^3c^2(a-1)) \\ &\quad \cdot \varepsilon_{3\alpha+2\beta}(\pm x^3c(a-1)^2)\varepsilon_\alpha(\pm xc) \\ &\quad \cdot \varepsilon_{2\alpha+\beta}(\pm cx^2)\varepsilon_{3\alpha+2\beta}(\pm cx^3)\varepsilon_{3\alpha+\beta}(\pm c^2x^3)\varepsilon_{3\alpha+2\beta}(\pm cx^3(a-1)) \\ &= \varepsilon_{\alpha+\beta}(\pm x(a-1)) \cdot \varepsilon_{2\alpha+\beta}(\pm x^2c(a-1)) \\ &\quad \cdot \varepsilon_\alpha(\pm xc) \cdot \varepsilon_{2\alpha+\beta}(\pm cx^2) \cdot \varepsilon_{3\alpha+\beta}(\pm x^3c^2(a-1) \pm c^2x^3) \\ &\quad \cdot \varepsilon_{3\alpha+2\beta}(\pm x^3c(a-1)^2 \pm cx^3 \pm cx^3(a-1))\end{aligned}$$

This finishes the case $c \neq 0$. If $c = 0$ hols, then $A = h_\beta(a)\varepsilon_\beta(a^{-1}b)$ and hence

$$\begin{aligned}(A, \varepsilon_{\alpha+\beta}(x)) &= \big(h_\beta(a)\varepsilon_\beta(a^{-1}b), \varepsilon_{\alpha+\beta}(x)\big) \\ &= (h_\beta(a), \varepsilon_{\alpha+\beta}(x)) = \varepsilon_{\alpha+\beta}(x(a-1))\end{aligned}$$

holds. This finishes the proof of the first claim. The proof for the second claim works the same way, so we omit it. □

This enables us to prove Proposition 4.5.2:

Proof. Let Q be the fraction field of R. Set $T_\alpha(R) = Y_\alpha \cup \{1\}$ and $T_\beta(R) = Y_\beta \cup \{1\}$ for Y_α, Y_β as in Proposition 4.3.4. The first step in proving this is to note that according to

Proposition 4.3.4 and Lemma 4.5.1, we can find $X_\alpha^{(1)}, X_\alpha^{(2)}, X_\alpha^{(3)} \in T_\alpha(R)$, $Y_\beta^{(1)}, Y_\beta^{(2)}, Y_\beta^{(3)} \in T_\beta(R)$ as well as an element $b \in B^+(G_2, R)$ such that

$$A = bX_\alpha^{(1)}Y_\beta^{(1)}X_\alpha^{(2)}Y_\beta^{(2)}X_\alpha^{(3)}Y_\beta^{(3)}.$$

Then observe:

$$\begin{aligned}
(A, \varepsilon_{3\alpha+2\beta}(1)) &= \left(bX_\alpha^{(1)}Y_\beta^{(1)}X_\alpha^{(2)}Y_\beta^{(2)}X_\alpha^{(3)}Y_\beta^{(3)}, \varepsilon_{3\alpha+2\beta}(1)\right) \\
&\sim \left(Y_\beta^{(2)}X_\alpha^{(3)}Y_\beta^{(3)}, \varepsilon_{3\alpha+2\beta}(1)\right) \cdot \left(\varepsilon_{3\alpha+2\beta}(1), (X_\alpha^{(2)})^{-1}(Y_\beta^{(1)})^{-1}(X_\alpha^{(1)})^{-1}b^{-1}\right) \\
&= \left[(X_\alpha^{(3)}Y_\beta^{(3)}, \varepsilon_{3\alpha+2\beta}(t))^{Y_\beta^{(2)}} \left(Y_\beta^{(2)}, \varepsilon_{3\alpha+2\beta}(1)\right)\right] \\
&\quad \cdot \left[((X_\alpha^{(2)})^{-1}(Y_\beta^{(1)})^{-1}(X_\alpha^{(1)})^{-1}b^{-1}, \varepsilon_{3\alpha+2\beta}(1))\right]^{-1}
\end{aligned}$$

But examine the various factors: First, $X_\alpha^{(i)} \in T_\alpha(R)$, $Y_\beta^{(i)} \in T_\beta(R)$ holds for all $i = 1, 2, 3$ and (w_α, w_β) is a minimal expression for $w_\alpha w_\beta$ with respect to the fundamental reflections. Thus according to Proposition 4.3.4, $X_\alpha^{(3)}Y_\beta^{(3)}$ is an element of $(B(G_2, Q)wB(G_2, Q) \cap G_2(R)$ for $w \in S(w_\alpha w_\beta)$. But note $w_\alpha w_\beta(3\alpha + 2\beta) = \beta$ is a positive root and so Lemma 4.3.8 implies together with $T(\beta) = \{\beta, \alpha+\beta, 2\alpha+\beta, 3\alpha+\beta, 3\alpha+2\beta\}$ that

$$\begin{aligned}
(X_\alpha^{(3)}Y_\beta^{(3)}, \varepsilon_{3\alpha+2\beta}(t))^{Y_\beta^{(2)}} &\in (\prod_{\psi \in T(\beta)} (\varepsilon_\psi(R))^{Y_\beta^{(2)}} \\
&= (\varepsilon_{3\alpha+2\beta}(R)\varepsilon_{3\alpha+\beta}(R)\varepsilon_{2\alpha+\beta}(R)\varepsilon_{\alpha+\beta}(R))^{Y_\beta^{(2)}}\varepsilon_\beta(R)^{Y_\beta^{(2)}} \\
&\subset (\varepsilon_{3\alpha+2\beta}(R)\varepsilon_{3\alpha+\beta}(R)\varepsilon_{2\alpha+\beta}(R)\varepsilon_{\alpha+\beta}(R)\varepsilon_\alpha(R))G_\beta(R).
\end{aligned}$$

The last inclusion follows from the second part of Lemma 4.3.9. Second, Lemma 4.3.9(3) also implies $(Y_\beta^{(2)}, \varepsilon_{3\alpha+2\beta}(1)) \in \varepsilon_{3\alpha+2\beta}(R) \cdot \varepsilon_{3\alpha+\beta}(R)$.

Next, $(w_\alpha, w_\beta, w_\alpha)$ is a minimal expression for $w_\alpha w_\beta w_\alpha$ with respect to fundamental reflections. Thus Proposition 4.3.4 implies that $(X_\alpha^{(2)})^{-1}(Y_\beta^{(1)})^{-1}(X_\alpha^{(1)})^{-1}b^{-1}$ is an element of $(B(G_2, Q)wB(G_2, Q)) \cap G_2(R)$ for some $w \in S(w_\alpha, w_\beta, w_\alpha)$. Further, $w_\alpha w_\beta w_\alpha(3\alpha + 2\beta) = \beta$ is a positive root. Thus Lemma 4.3.8 implies that

$$((X_\alpha^{(2)})^{-1}(Y_\beta^{(1)})^{-1}(X_\alpha^{(1)})^{-1}b^{-1}, \varepsilon_{3\alpha+2\beta}(1)) \in \prod_{\psi \in T(\beta)} \varepsilon_\psi(R) = \varepsilon_{3\alpha+2\beta}(R)\varepsilon_{3\alpha+\beta}(R)\varepsilon_{2\alpha+\beta}(R)\varepsilon_{\alpha+\beta}(R).$$

All of this is to say, that reordering the terms of

$$\left[(X_\alpha^{(3)}Y_\beta^{(3)}, \varepsilon_{3\alpha+2\beta}(t))^{Y_\beta^{(2)}} \left(Y_\beta^{(2)}, \varepsilon_{3\alpha+2\beta}(1)\right)\right] \cdot \left[((X_\alpha^{(2)})^{-1}(Y_\beta^{(1)})^{-1}(X_\alpha^{(1)})^{-1}b^{-1}, \varepsilon_{3\alpha+2\beta}(1))\right]^{-1},$$

we obtain that $(A, \varepsilon_{3\alpha+2\beta}(1))$ is conjugate to an element B of

$$\varepsilon_{3\alpha+2\beta}(R)\varepsilon_{3\alpha+\beta}(R)\varepsilon_{2\alpha+\beta}(R)\varepsilon_{\alpha+\beta}(R)\varepsilon_{\alpha}(R)G_{\beta}(R).$$

Hence, we can find $a, b, c, d, e \in R$ and an element

$$Z_{\beta} = \psi_{\beta}\begin{pmatrix} q & r \\ s & t \end{pmatrix} \in G_{\beta}(R)$$

such that

$$B = \varepsilon_{3\alpha+2\beta}(a)\varepsilon_{3\alpha+\beta}(b)\varepsilon_{2\alpha+\beta}(c)\varepsilon_{\alpha+\beta}(d)\varepsilon_{\alpha}(e)Z_{\beta}.$$

The remaining proof of the proposition proceeds in four steps: First, we will show that

$$(e^3) \subset \varepsilon_l(A, 32) \text{ and } (b^2) \subset \varepsilon_l(A, 45360).$$

Then second, we show that applying the calculations of the first step to $B^{w_{\beta}}$ implies

$$(d^3) \subset \varepsilon_l(A, 32) \text{ and } ((\pm a \pm 3d^2e)^2) \subset \varepsilon_l(A, 45360).$$

Third, we prove that

$$(\left[\pm s(q-1)^2 \pm s \pm (q-1)s \pm 3c \pm 3e\right]^2) \subset \varepsilon_l(A, 6720) \text{ and } (64c^2) \subset \varepsilon_l(A, 2246272).$$

Then fourth, we construct an ideal $I_0(A)$ with the desired properties.

For the first step, note that $\|B\|_A \leq 2$. Next, observe that for $x \in R$, we obtain from the third item of Lemma 4.3.9(3) that

$$(B, \varepsilon_{3\alpha+\beta}(x)) \sim (Z_{\beta}, \varepsilon_{3\alpha+\beta}(x)) = \varepsilon_{3\alpha+\beta}(x(t-1))\varepsilon_{3\alpha+2\beta}(xr).$$

Note that $x \in R$ is arbitrary, so we obtain by commuting $\varepsilon_{3\alpha+\beta}(x(t-1))\varepsilon_{3\alpha+2\beta}(xr)$ with $\varepsilon_{\beta}(1)$ that

$$(1-t) \subset \varepsilon_l(A, 8).$$

Similarly, we obtain by commuting $\varepsilon_{3\alpha+\beta}(x(t-1))\varepsilon_{3\alpha+2\beta}(xr)$ with $\varepsilon_{-\beta}(1)$ and conjugating the resulting term $\varepsilon_{3\alpha+\beta}(-xr)$ that

$$(r) \subset \varepsilon_l(A, 8).$$

Similarly, one obtains using for $x \in R$ the commutator $(B, \varepsilon_{3\alpha+2\beta}(x))$ that $(1-q), (s) \subset \varepsilon_l(A, 8)$.

Next, consider the commutator

$$\begin{aligned}(B,\varepsilon_\beta(1)) &\sim (\varepsilon_{3\alpha+\beta}(b)\varepsilon_\alpha(e)Z_\beta,\varepsilon_\beta(1))\\ &\sim (Z_\beta,\varepsilon_\beta(1))\cdot[(\varepsilon_{3\alpha+\beta}(-b)\varepsilon_\alpha(-e),\varepsilon_\beta(1))]^{-1}\\ &= (Z_\beta,\varepsilon_\beta(1))\cdot[(\varepsilon_\alpha(-e),\varepsilon_\beta(1))^{\varepsilon_{3\alpha+\beta}(-b)}(\varepsilon_{3\alpha+\beta}(-b),\varepsilon_\beta(1))]^{-1}\\ &= (Z_\beta,\varepsilon_\beta(1))\cdot[\varepsilon_{3\alpha+2\beta}(\pm e^3)\varepsilon_{3\alpha+\beta}(\pm e^3)\varepsilon_{2\alpha+\beta}(\pm e^2)\varepsilon_{\alpha+\beta}(\pm e)\varepsilon_{3\alpha+2\beta}(-b)]^{-1}\\ &\sim \varepsilon_{\alpha+\beta}(\pm e)\varepsilon_{2\alpha+\beta}(\pm e^2)\varepsilon_{3\alpha+\beta}(\pm e^3)\varepsilon_{3\alpha+2\beta}(b\pm e^3)(Z_\beta,\varepsilon_\beta(1)) =: C_1\in B_A(4).\end{aligned}$$

Next, we consider the commutator

$$\begin{aligned}(C_1,\varepsilon_\beta(1)) &= \big(\varepsilon_{\alpha+\beta}(\pm e)\varepsilon_{2\alpha+\beta}(\pm e^2)\varepsilon_{3\alpha+\beta}(\pm e^3)\varepsilon_{3\alpha+2\beta}(b\pm e^3)(Z_\beta,\varepsilon_\beta(1)),\varepsilon_\beta(1)\big)\\ &\sim \big(\varepsilon_{3\alpha+\beta}(\pm e^3)(Z_\beta,\varepsilon_\beta(1)),\varepsilon_\beta(1)\big)\\ &\sim ((Z_\beta,\varepsilon_\beta(1)),\varepsilon_\beta(1))\cdot(\varepsilon_\beta(1),\varepsilon_{3\alpha+\beta}(\pm e^3))\\ &= ((Z_\beta,\varepsilon_\beta(1)),\varepsilon_\beta(1))\cdot\varepsilon_{3\alpha+2\beta}(\pm e^3) =: C_2\in B_A(8).\end{aligned}$$

Next, for $x\in R$ consider the commutator

$$\begin{aligned}(C_2,\varepsilon_{-3\alpha-\beta}(x)) &\sim (\varepsilon_{3\alpha+2\beta}(\pm e^3),\varepsilon_{-3\alpha-\beta}(x))\cdot(\varepsilon_{-3\alpha-\beta}(x),(Z_\beta,\varepsilon_\beta(1)),\varepsilon_\beta(1))^{-1})\\ &= \varepsilon_\beta(\pm xe^3)\cdot(\varepsilon_{-3\alpha-\beta}(x),((Z_\beta,\varepsilon_\beta(1)),\varepsilon_\beta(1))^{-1})\end{aligned}$$

Note that

$$\begin{aligned}((Z_\beta,\varepsilon_\beta(1)),\varepsilon_\beta(1))^{-1} &= (\psi_\beta\begin{pmatrix}1-qs & (q-1)(q+1)+qs\\ -s^2 & 1+s(q+s)\end{pmatrix},\varepsilon_\beta(1))^{-1}\\ &= \psi_\beta\begin{pmatrix}1+(1-qs)s^2 & (1-(1-qs))(1+(1-qs))-s^2(1-qs)\\ -s^4 & 1-s^2(1-qs-s^2)\end{pmatrix}^{-1}\\ &= \psi_\beta\begin{pmatrix}1-s^2(1-qs-s^2) & -s(q(2-qs)+s(1-qs))\\ s^4 & 1+(1-qs)s^2\end{pmatrix}\end{aligned}$$

and consequently, we obtain from Lemma 4.3.9(3) that

$$(\varepsilon_{-3\alpha-\beta}(x),((Z_\beta,\varepsilon_\beta(1)),\varepsilon_\beta(1))^{-1}) = \varepsilon_{-3\alpha-\beta}(-xs^2(1-qs-s^2))\varepsilon_{-3\alpha-2\beta}(-xs^4).$$

But we already know that

$$(s)\subset\varepsilon_l(A,8)$$

and hence we obtain that

$$\|(\varepsilon_{-3\alpha-\beta}(x),((Z_\beta,\varepsilon_\beta(1)),\varepsilon_\beta(1))^{-1})\|_A = \|\varepsilon_{-3\alpha-\beta}(-xs^2(1-qs-s^2))\varepsilon_{-3\alpha-2\beta}(-xs^4)\|_A \\ \leq 2*8 = 16.$$

This in turn implies that

$$\|\varepsilon_\beta(xe^3)\|_A \leq \|(C_2,\varepsilon_{-3\alpha-\beta}(x))\|_A + \|(\varepsilon_{-3\alpha-\beta}(x),((Z_\beta,\varepsilon_\beta(1)),\varepsilon_\beta(1))^{-1})\|_A \leq 16+16 = 32.$$

So, we obtain

$$(e^3) \subset \varepsilon_l(A,32).$$

This proves the first statement of the first step.

For the second statement of the first step, observe first that $\|C_1\|_A \leq 4$ and $\|\varepsilon_{3\alpha+\beta}(\pm e^3)\varepsilon_{3\alpha+2\beta}(\pm e^3)\|_A \leq 64$. This implies

$$C_3 := \varepsilon_{\alpha+\beta}(\pm e)\varepsilon_{2\alpha+\beta}(\pm e^2)\varepsilon_{3\alpha+2\beta}(b)(Z_\beta,\varepsilon_\beta(1)) \in B_A(68).$$

Hence we obtain

$$B_A(136) \ni (C_3,\varepsilon_{-3\alpha-\beta}(1)) \sim ((Z_\beta,\varepsilon_\beta(1)),\varepsilon_{-3\alpha-\beta}(1)) \\ \cdot (\varepsilon_{-3\alpha-\beta}(1),\varepsilon_{\alpha+\beta}(\pm e)\varepsilon_{2\alpha+\beta}(\pm e^2)\varepsilon_{3\alpha+2\beta}(-b)).$$

But similar to the argument showing $\|(\varepsilon_{-3\alpha-\beta}(x),((Z_\beta,\varepsilon_\beta(1)),\varepsilon_\beta(1)))\|_A \leq 16$, we obtain $\|((Z_\beta,\varepsilon_\beta(1)),\varepsilon_{-3\alpha-\beta}(1))\|_A \leq 16$. Thus

$$\begin{aligned} B_A(152) = B_A(16+136) &\ni (\varepsilon_{\alpha+\beta}(\pm e)\varepsilon_{2\alpha+\beta}(\pm e^2)\varepsilon_{3\alpha+2\beta}(-b),\varepsilon_{-3\alpha-\beta}(1)) \\ &\sim (\varepsilon_{3\alpha+2\beta}(-b),\varepsilon_{-3\alpha-\beta}(1))\cdot(\varepsilon_{-3\alpha-\beta}(1),\varepsilon_{\alpha+\beta}(\pm e)\varepsilon_{2\alpha+\beta}(\pm e^2)) \\ &= \varepsilon_\beta(-b)[(\varepsilon_{2\alpha+\beta}(\pm e^2),\varepsilon_{-3\alpha-\beta}(1))^{\varepsilon_{\alpha+\beta}(\pm e)}\cdot(\varepsilon_{\alpha+\beta}(\pm e),\varepsilon_{-3\alpha-\beta}(1))]^{-1} \\ &= \varepsilon_\beta(-b)[(\varepsilon_{-\alpha}(\pm e^2)\varepsilon_{\alpha+\beta}(\pm e^4)\varepsilon_{3\alpha+2\beta}(\pm e^8)\varepsilon_\beta(\pm e^8))]^{\varepsilon_{\alpha+\beta}(\pm e)} \\ &= \varepsilon_\beta(-b)\varepsilon_{-\alpha}(\pm e^2)^{\varepsilon_{\alpha+\beta}(\pm e)}\varepsilon_{\alpha+\beta}(\pm e^4)\varepsilon_{3\alpha+2\beta}(\pm e^8)\varepsilon_\beta(\pm e^8) \\ &= \varepsilon_\beta(-b)\varepsilon_{-\alpha}(\pm e^2)\varepsilon_\beta(\pm 3e^3)\varepsilon_{\alpha+\beta}(\pm e^4)\varepsilon_{3\alpha+2\beta}(\pm e^8)\varepsilon_\beta(\pm e^8) \\ &= \varepsilon_\beta(-b)\varepsilon_{-\alpha}(\pm e^2)\varepsilon_{\alpha+\beta}(\pm e^4)\varepsilon_{3\alpha+2\beta}(\pm e^8)\varepsilon_\beta((\pm e^5\pm 3)e^3) \end{aligned}$$

Note further that $(e^3) \subset \varepsilon_l(A,32)$ implies

$$\|\varepsilon_{3\alpha+2\beta}(\pm e^8)\varepsilon_\beta((\pm e^5\pm 3)e^3)\|_A \leq 64.$$

This implies that

$$\begin{aligned} B_A(216) = B_A(64+152) \ni &(\varepsilon_\beta(-b)\varepsilon_{-\alpha}(\pm e^2)\varepsilon_{\alpha+\beta}(\pm e^4))^{w_\alpha} \\ &= \varepsilon_{3\alpha+\beta}(\pm b)\varepsilon_\alpha(\pm e^2)\varepsilon_{2\alpha+\beta}(\pm e^4) =: C_4. \end{aligned}$$

But applying Proposition 4.5.5 to C_4 implies

$$(b^2) \subset \varepsilon_l(A, 210 \cdot 216) = \varepsilon_l(A, 45360).$$

This finishes the first step.

Also note that applying Lemma 4.5.3 to C_4 implies

$$(2b) \subset \varepsilon_l(A, 18 * 216) = \varepsilon_l(A, 3888). \tag{4.6}$$

For the second step, observe first that

$$\begin{aligned} B^{w_\beta} &= \varepsilon_{3\alpha+\beta}(\pm a)\varepsilon_{3\alpha+2\beta}(\pm b)\varepsilon_{2\alpha+\beta}(\pm c)\varepsilon_\alpha(\pm d)\varepsilon_{\alpha+\beta}(\pm e)Z_\beta^{w_\beta} \\ &= \varepsilon_{3\alpha+2\beta}(\pm b \pm 3de^2)\varepsilon_{3\alpha+\beta}(\pm a \pm 3d^2e)\varepsilon_{2\alpha+\beta}(\pm c \pm 2de)\varepsilon_{\alpha+\beta}(\pm e)\varepsilon_\alpha(\pm d)Z_\beta^{w_\beta}. \end{aligned}$$

Note, that the first step does not use any particular properties arising from the definition of B beyond the fact that it is an element of

$$\varepsilon_{3\alpha+2\beta}(R)\varepsilon_{3\alpha+\beta}(R)\varepsilon_{2\alpha+\beta}(R)\varepsilon_{\alpha+\beta}(R)\varepsilon_\alpha(R)G_\beta(R).$$

But this is also the case for B^{w_β} and hence we obtain using the same calculations that

$$(d^3) \subset \varepsilon_l(A, 32)$$

and

$$((\pm a \pm 3d^2e)^2) \subset \varepsilon_l(A, 45360).$$

This finishes the second step.

For the third step, we split the argument into two parts again: First, we will show

$$(\left[\pm s(q-1)^2 \pm s \pm (q-1)s \pm 3c \pm 3e\right]^2) \subset \varepsilon_l(A, 6720).$$

and secondly, we will show

$$(64c^2) \subset \varepsilon_l(A, 2246272).$$

For the first part, consider the commutator

$$\begin{aligned}(B, \varepsilon_{\alpha+\beta}(1)) &\sim (\varepsilon_{2\alpha+\beta}(c)\varepsilon_\alpha(e)Z_\beta, \varepsilon_{\alpha+\beta}(1)) \\ &\sim (Z_\beta, \varepsilon_{\alpha+\beta}(1)) \cdot [(\varepsilon_\alpha(-e)\varepsilon_{2\alpha+\beta}(-c), \varepsilon_{\alpha+\beta}(1))]^{-1} \\ &= (Z_\beta, \varepsilon_{\alpha+\beta}(1)) \cdot [(\varepsilon_{2\alpha+\beta}(-c), \varepsilon_{\alpha+\beta}(1))^{\varepsilon_\alpha(-e)}(\varepsilon_\alpha(-e), \varepsilon_{\alpha+\beta}(1))]^{-1} \\ &= (Z_\beta, \varepsilon_{\alpha+\beta}(1)) \cdot [\varepsilon_{3\alpha+2\beta}(\pm 3c)^{\varepsilon_\alpha(-e)}\varepsilon_{2\alpha+\beta}(\pm 2e)\varepsilon_{3\alpha+\beta}(\pm 3e^2)\varepsilon_{3\alpha+2\beta}(\pm 3e)]^{-1} \\ &= (Z_\beta, \varepsilon_{\alpha+\beta}(1)) \cdot \varepsilon_{3\alpha+2\beta}(\pm 3c \pm 3e)\varepsilon_{3\alpha+\beta}(\pm 3e^2)\varepsilon_{2\alpha+\beta}(\pm 2e) =: D_1\end{aligned}$$

Next, note that, Lemma 4.5.6 implies

$$\begin{aligned}(Z_\beta, \varepsilon_{\alpha+\beta}(1)) = \varepsilon_{\alpha+\beta}(\pm(q-1)) &\cdot \varepsilon_{2\alpha+\beta}(\pm s(q-1)) \\ &\cdot \varepsilon_\alpha(\pm s) \cdot \varepsilon_{2\alpha+\beta}(\pm s) \cdot \varepsilon_{3\alpha+\beta}(\pm s^2(q-1) \pm s^2) \\ &\cdot \varepsilon_{3\alpha+2\beta}(\pm s(q-1)^2 \pm s \pm s(q-1))\end{aligned}$$

Hence D_1 looks as follows

$$\begin{aligned}D_1 &= (Z_\beta, \varepsilon_{\alpha+\beta}(1)) \cdot \varepsilon_{3\alpha+2\beta}(\pm 3c \pm 3e)\varepsilon_{3\alpha+\beta}(\pm 3e^2)\varepsilon_{2\alpha+\beta}(\pm 2e) \\ &= \varepsilon_{\alpha+\beta}(\pm(q-1)) \cdot \varepsilon_{2\alpha+\beta}(\pm s(q-1)) \\ &\quad \cdot \varepsilon_\alpha(\pm s) \cdot \varepsilon_{2\alpha+\beta}(\pm s) \cdot \varepsilon_{3\alpha+\beta}(\pm s^2(q-1) \pm s^2) \\ &\quad \cdot \varepsilon_{3\alpha+2\beta}(\pm s(q-1)^2 \pm s \pm s(q-1)) \\ &\quad \cdot \varepsilon_{3\alpha+2\beta}(\pm 3c \pm 3e)\varepsilon_{3\alpha+\beta}(\pm 3e^2)\varepsilon_{2\alpha+\beta}(\pm 2e) \\ &= \varepsilon_{\alpha+\beta}(\pm(q-1))\varepsilon_{3\alpha+2\beta}(\pm s(q-1)^2 \pm s \pm s(q-1) \pm 3c \pm 3e) \\ &\quad \cdot \varepsilon_{2\alpha+\beta}(\pm s(q-1))\varepsilon_\alpha(\pm s) \\ &\quad \cdot \varepsilon_{2\alpha+\beta}(\pm s) \cdot \varepsilon_{3\alpha+\beta}(\pm s^2(q-1) \pm s^2) \\ &\quad \cdot \varepsilon_{3\alpha+\beta}(\pm 3e^2)\varepsilon_{2\alpha+\beta}(\pm 2e)\end{aligned}$$

Note that all factors of D_1 besides $\varepsilon_{\alpha+\beta}(\pm(q-1))$ and

$$\varepsilon_{3\alpha+2\beta}\left(\pm s(q-1)^2 \pm s \pm s(q-1) \pm 3c \pm 3e\right)$$

commute with $\varepsilon_{-\beta}(1)$ and that $\|D_1\|_A \leq 4$. Thus

$$\begin{aligned}
&(D_1, \varepsilon_{-\beta}(1)) \\
&= \left(\varepsilon_{\alpha+\beta}(\pm(q-1))\varepsilon_{3\alpha+2\beta}\left(\pm s(q-1)^2 \pm s \pm s(q-1) \pm 3c \pm 3e\right), \varepsilon_{-\beta}(1)\right) \\
&\sim \left(\varepsilon_{3\alpha+2\beta}\left(\pm s(q-1)^2 \pm s \pm s(q-1) \pm 3c \pm 3e\right), \varepsilon_{-\beta}(1)\right) \\
&\quad \cdot (\varepsilon_{-\beta}(1), \varepsilon_{\alpha+\beta}(\pm(q-1))) \\
&= \varepsilon_{3\alpha+\beta}\left(\pm s(q-1)^2 \pm s \pm s(q-1) \pm 3c \pm 3e\right) \\
&\quad \cdot \varepsilon_{\alpha}(\pm(q-1))\varepsilon_{2\alpha+\beta}(\pm(q-1)^2)\varepsilon_{3\alpha+2\beta}(\pm(q-1)^3)\varepsilon_{3\alpha+\beta}(\pm(q-1)^3) \\
&=: D_2.
\end{aligned}$$

However, similarly to the calculations of the first step showing $(1-t), (r) \subset \varepsilon_l(A, 8)$, the factors $\varepsilon_{3\alpha+2\beta}(\pm(q-1)^3)$ and $\varepsilon_{3\alpha+\beta}(\pm(q-1)^3)$ are both elements of $B_A(8)$. Also $D_2 \in B_A(8)$ holds. Thus

$$\begin{aligned}
D_3 :=& \varepsilon_{\alpha}(\pm(q-1))\varepsilon_{2\alpha+\beta}(\pm(q-1)^2) \\
& \varepsilon_{3\alpha+\beta}\left(\pm s(q-1)^2 \pm s \pm (q-1)s \pm 3c \pm 3e\right)
\end{aligned}$$

is an element of $B_A(16 + 2 * 8) = B_A(32)$. But applying Proposition 4.5.5 to D_3 yields that

$$(\left[\pm s(q-1)^2 \pm s \pm (q-1)s \pm 3c \pm 3e\right]^2) \subset \varepsilon_l(A, 32 \cdot 210) = \varepsilon_l(A, 6720).$$

This finishes the proof of the the first statement of the third step.

For the second part of the third step, consider the commutator

$$\begin{aligned}
(B, \varepsilon_{-\alpha}(4)) &\sim (\varepsilon_{3\alpha+\beta}(b)\varepsilon_{2\alpha+\beta}(c)\varepsilon_{\alpha+\beta}(d)\varepsilon_{\alpha}(e)Z_{\beta}, \varepsilon_{-\alpha}(4)) \\
&\sim (\varepsilon_{2\alpha+\beta}(c)\varepsilon_{\alpha+\beta}(d)\varepsilon_{\alpha}(e)Z_{\beta}, \varepsilon_{-\alpha}(4))(\varepsilon_{-\alpha}(4), \varepsilon_{3\alpha+\beta}(-b)) \\
&= (Z_{\beta}, \varepsilon_{-\alpha}(4))^{\varepsilon_{2\alpha+\beta}(c)\varepsilon_{\alpha+\beta}(d)\varepsilon_{\alpha}(e)} \\
&\quad \cdot (\varepsilon_{2\alpha+\beta}(c)\varepsilon_{\alpha+\beta}(d)\varepsilon_{\alpha}(e), \varepsilon_{-\alpha}(4)) \cdot (\varepsilon_{-\alpha}(4), \varepsilon_{3\alpha+\beta}(-b)).
\end{aligned}$$

To proceed, we note two facts: First, observe that

$$(\varepsilon_{-\alpha}(4), \varepsilon_{3\alpha+\beta}(-b)) \sim \varepsilon_{3\alpha+2\beta}(\pm 64b^2)\varepsilon_{\beta}(\pm 64b)\varepsilon_{\alpha+\beta}(\pm 16b)\varepsilon_{2\alpha+\beta}(\pm 4b) \tag{4.7}$$

But

$$(2b) \subset \varepsilon_l(A, 3888)$$

holds according to (4.6) and hence Lemma 3.5.4 implies

$$(4b) \subset \varepsilon_s(A, 3888 * 8) = \varepsilon_s(A, 31104).$$

These two facts imply together with (4.7) that

$$\|(\varepsilon_{-\alpha}(4), \varepsilon_{3\alpha+\beta}(-b))\|_A \leq 2 * 3888 + 2 * 31104 = 69984. \tag{4.8}$$

Second, observe that Lemma 4.5.6 implies

$$\begin{aligned}(Z_\beta, \varepsilon_{-\alpha}(4)) &= \varepsilon_{-\alpha}(\pm 4(q-1)) \cdot \varepsilon_{-2\alpha-\beta}(\pm 16s(q-1)) \\ &\quad \cdot \varepsilon_{-\alpha-\beta}(\pm 4s) \cdot \varepsilon_{-2\alpha-\beta}(\pm 16s) \cdot \varepsilon_{-3\alpha-\beta}(\pm 64s(q-1)^2 \pm 64s \pm 64s(q-1)) \\ &\quad \cdot \varepsilon_{-3\alpha-2\beta}(\pm 64s^2 \pm 64s^2(q-1)) \\ &= \varepsilon_{-\alpha}(\pm 4(q-1)) \cdot \varepsilon_{-2\alpha-\beta}(\pm 16s(q-1) \pm 16s) \\ &\quad \cdot \varepsilon_{-\alpha-\beta}(\pm 4s) \cdot \varepsilon_{-3\alpha-\beta}(\pm 64s(q-1)^2 \pm 64s \pm 64s(q-1)) \\ &\quad \cdot \varepsilon_{-3\alpha-2\beta}(\pm 64s^2 \pm 64s^2(q-1) \pm 192s^2)\end{aligned}$$

But note that

$$(s) \subset \varepsilon_l(A, 8)$$

and hence Lemma 3.5.4 implies that

$$(2s) \subset \varepsilon_s(A, 64).$$

Hence $\varepsilon_{-2\alpha-\beta}(\pm 16s(q-1) \pm 16s), \varepsilon_{-\alpha-\beta}(\pm 4s) \in B_A(64)$ and $\varepsilon_{-3\alpha-\beta}(\pm 64s(q-1)^2 \pm 64s \pm 64s(q-1)), \varepsilon_{-3\alpha-2\beta}(\pm 64s^2 \pm 64s^2(q-1) \pm 192s^2) \in B_A(8)$ hold. Further, we have

$$(1-q) \subset \varepsilon(A, 3\alpha + 2\beta, 8)$$

and thus Lemma 3.5.4(2) implies

$$(2(1-q)) \subset \varepsilon(A, \alpha, 64).$$

Hence $\|\varepsilon_{-\alpha}(\pm 4(q-1))\|_A \leq 64$. It follows that

$$\|(Z_\beta, \varepsilon_{-\alpha}(4))^{\varepsilon_{2\alpha+\beta}(c)\varepsilon_{\alpha+\beta}(d)\varepsilon_\alpha(e)}\|_A = \|(Z_\beta, \varepsilon_{-\alpha}(4))\|_A \leq 3 * 64 + 2 * 8 = 208. \tag{4.9}$$

Summarizing (4.8) and (4.9) with

$$\begin{aligned}B_A(4) \ni (B, \varepsilon_{-\alpha}(4)) \sim &(Z_\beta, \varepsilon_{-\alpha}(4))^{\varepsilon_{2\alpha+\beta}(c)\varepsilon_{\alpha+\beta}(d)\varepsilon_\alpha(e)} \\ &\cdot (\varepsilon_{2\alpha+\beta}(c)\varepsilon_{\alpha+\beta}(d)\varepsilon_\alpha(e), \varepsilon_{-\alpha}(4))(\varepsilon_{-\alpha}(4), \varepsilon_{3\alpha+\beta}(-b))\end{aligned}$$

we obtain $\|(\varepsilon_{2\alpha+\beta}(c)\varepsilon_{\alpha+\beta}(d)\varepsilon_\alpha(e), \varepsilon_{-\alpha}(4))\|_A \leq 4 + 69984 + 208 = 70196$.

Next, observe

$$\begin{aligned}
&(\varepsilon_{2\alpha+\beta}(c)\varepsilon_{\alpha+\beta}(d)\varepsilon_\alpha(e), \varepsilon_{-\alpha}(4))\\
&\sim (\varepsilon_\alpha(e), \varepsilon_{-\alpha}(4)) \cdot (\varepsilon_{2\alpha+\beta}(-c)\varepsilon_{\alpha+\beta}(-d), \varepsilon_{-\alpha}(4))^{-1}\\
&= (\varepsilon_\alpha(e), \varepsilon_{-\alpha}(4)) \cdot [(\varepsilon_{\alpha+\beta}(-d), \varepsilon_{-\alpha}(4))^{\varepsilon_{2\alpha+\beta}(c)}(\varepsilon_{2\alpha+\beta}(-c), \varepsilon_{-\alpha}(4))]^{-1}\\
&= (\varepsilon_\alpha(e), \varepsilon_{-\alpha}(4)) \cdot [\varepsilon_\beta(\pm 12d)^{\varepsilon_{2\alpha+\beta}(c)}\varepsilon_{\alpha+\beta}(\pm 8c)\varepsilon_\beta(\pm 48c)\varepsilon_{3\alpha+2\beta}(\pm 12c^2)]^{-1}\\
&= (\varepsilon_\alpha(e), \varepsilon_{-\alpha}(4)) \cdot [\varepsilon_\beta(\pm 12d)\varepsilon_{\alpha+\beta}(\pm 8c)\varepsilon_\beta(\pm 48c)\varepsilon_{3\alpha+2\beta}(\pm 12c^2)]^{-1}\\
&= (\varepsilon_\alpha(e), \varepsilon_{-\alpha}(4)) \cdot [\varepsilon_\beta(\pm 12d \pm 48c)\varepsilon_{\alpha+\beta}(\pm 8c)\varepsilon_{3\alpha+2\beta}(\pm 12c^2)]^{-1}\\
&= (\varepsilon_\alpha(e), \varepsilon_{-\alpha}(4)) \cdot \varepsilon_\beta(\pm 12d \pm 48c)\varepsilon_{\alpha+\beta}(\pm 8c)\varepsilon_{3\alpha+2\beta}(\pm 12c^2) =: E_1.
\end{aligned}$$

Observe $\|E_1\|_A \leq 2 * 70196 = 140392$. To simplify notation, we set

$$u := \pm 12d \pm 48c, v := \pm 8c \text{ and } w := \pm 12c^2.$$

Note that $(\varepsilon_\alpha(e), \varepsilon_{-\alpha}(4))$ commutes with $\varepsilon_{-3\alpha-2\beta}(1)$ and hence

$$\begin{aligned}
&(E_1, \varepsilon_{-3\alpha-2\beta}(1))\\
&\sim (\varepsilon_\beta(u)\varepsilon_{\alpha+\beta}(v)\varepsilon_{3\alpha+2\beta}(w), \varepsilon_{-3\alpha-2\beta}(1))\\
&\sim (\varepsilon_{3\alpha+2\beta}(w), \varepsilon_{-3\alpha-2\beta}(1)) \cdot (\varepsilon_\beta(u)\varepsilon_{\alpha+\beta}(v), \varepsilon_{-3\alpha-2\beta}(1))^{-1}\\
&= (\varepsilon_{3\alpha+2\beta}(w), \varepsilon_{-3\alpha-2\beta}(1)) \cdot [(\varepsilon_{\alpha+\beta}(v), \varepsilon_{-3\alpha-2\beta}(1))^{\varepsilon_\beta(u)}(\varepsilon_\beta(u), \varepsilon_{-3\alpha-2\beta}(1))]^{-1}\\
&= (\varepsilon_{3\alpha+2\beta}(w), \varepsilon_{-3\alpha-2\beta}(1)) \cdot [(\varepsilon_{-3\alpha-\beta}(\pm v^2)\varepsilon_\beta(\pm v^2)\varepsilon_{-\alpha}(\pm v^2)\varepsilon_{-2\alpha-\beta}(\pm v))^{\varepsilon_\beta(u)}\varepsilon_{-3\alpha-\beta}(\pm u)]^{-1}\\
&= (\varepsilon_{3\alpha+2\beta}(w), \varepsilon_{-3\alpha-2\beta}(1)) \cdot [\varepsilon_{-3\alpha-\beta}(\pm v^2 \pm u)\varepsilon_\beta(\pm v^2)\varepsilon_{-\alpha}(\pm v^2)\varepsilon_{-2\alpha-\beta}(\pm v)]^{-1}\\
&= (\varepsilon_{3\alpha+2\beta}(w), \varepsilon_{-3\alpha-2\beta}(1)) \cdot \varepsilon_{-2\alpha-\beta}(\pm v)\varepsilon_{-\alpha}(\pm v^2)\varepsilon_\beta(\pm v^2)\varepsilon_{-3\alpha-\beta}(\pm v^2 \pm u) =: E_2
\end{aligned}$$

Observe that $\|E_2\|_A \leq 2 * 140392 = 280784$. But all factors of E_2 besides $(\varepsilon_{3\alpha+2\beta}(w), \varepsilon_{-3\alpha-2\beta}(1))$ and $\varepsilon_\beta(\pm v^2)$ commute with $\varepsilon_{-3\alpha-2\beta}(1)$ and hence

$$\begin{aligned}
(E_2, \varepsilon_{-3\alpha-2\beta}(1)) &= \big((\varepsilon_{3\alpha+2\beta}(w), \varepsilon_{-3\alpha-2\beta}(1))\, \varepsilon_\beta(\pm v^2), \varepsilon_{-3\alpha-2\beta}(1)\big)\\
&\sim \big(\varepsilon_\beta(\pm v^2), \varepsilon_{-3\alpha-2\beta}(1)\big) \cdot \big(\varepsilon_{-3\alpha-2\beta}(1), (\varepsilon_{3\alpha+2\beta}(w), \varepsilon_{-3\alpha-2\beta}(1))^{-1}\big)\\
&= \varepsilon_{-3\alpha-\beta}(\pm v^2) \cdot \big(\varepsilon_{-3\alpha-2\beta}(1), (\varepsilon_{3\alpha+2\beta}(w), \varepsilon_{-3\alpha-2\beta}(1))^{-1}\big) =: E_3.
\end{aligned}$$

Observe that $\|E_3\|_A \leq 2 * 280784 = 561568$. Define further

$$Z_{3\alpha+2\beta} := (\varepsilon_{-3\alpha-2\beta}(1), (\varepsilon_{3\alpha+2\beta}(w), \varepsilon_{-3\alpha-2\beta}(1))^{-1}).$$

Observe next that according to Lemma 4.3.9 there are $i, j \in R$ with

$$(Z_{3\alpha+2\beta}, \varepsilon_{-\beta}(1)) = \varepsilon_{-\beta}(i)\varepsilon_{3\alpha+\beta}(j).$$

Thus, we obtain

$$(E_3, \varepsilon_{-\beta}(1)) \sim (Z_{3\alpha+2\beta}, \varepsilon_{-\beta}(1)) \cdot (\varepsilon_{-\beta}(1), \varepsilon_{-3\alpha-\beta}(\pm v^2))$$
$$= \varepsilon_{-\beta}(i)\varepsilon_{3\alpha+\beta}(j)\varepsilon_{-3\alpha-2\beta}(\pm v^2).$$

Next, consider for $x \in R$ the commutator

$$(\varepsilon_{-\beta}(i)\varepsilon_{3\alpha+\beta}(j)\varepsilon_{-3\alpha-2\beta}(\pm v^2), \varepsilon_{3\alpha+\beta}(x)) \sim (\varepsilon_{-3\alpha-2\beta}(\pm v^2), \varepsilon_{3\alpha+\beta}(x)) = \varepsilon_{-\beta}(xv^2).$$

But $v = \pm 8c$ and hence we obtain that

$$(64c^2) = (v^2) \subset \varepsilon_l(A, 4 * 561568) = \varepsilon_l(A, 2246272).$$

This finishes the third step.

For the fourth step, define the ideal $I_0(A)$ as follows:

$$I_0(A) := (1-t, 1-q, s, r, e^3, b^2, d^3, (a \pm 3d^2e)^2, [\pm s(q-1)^2 \pm s \pm (q-1)s \pm 3c \pm 3e]^2, 64c^2)$$
$$\subset \varepsilon_l(A, 4 * 8 + 32 + 45360 + 32 + 45360 + 6720 + 2246272) = \varepsilon_l(A, 2343808).$$

To finish the proof of the proposition, we prove that each maximal ideal m with $I_0(A) \subset m$ contains $a, b, c, d, e, 1-t, 1-q, r, s$. Clearly $1-t, 1-q, s, r \in m$. Observe further that e, b, d are elements of m, because $e^3, b^2, d^3 \in m$. Also $(a \pm 3d^2e)$ is an element of m, because $(a \pm 3d^2e)^2$ is an element of m. But e is an element of m and hence so is $a = (a \pm 3d^2e) \mp 3d^2e$. Next, observe that

$$[\pm s(q-1)^2 \pm s \pm (q-1)s \pm 3c \pm 3e]^2 \in m$$

holds and hence

$$[\pm s(q-1)^2 \pm s \pm (q-1)s \pm 3c \pm 3e] \in m$$

holds as well. But all the elements $s(q-1)^2, s, (q-1)s, 3e$ are already known to be elements of m. Hence $3c$ is also an element of m. Next, observe that $64c^2 = (8c)^2$ is an element of m and hence $8c$ is an element of m as well. But if both $3c$ and $8c$ are elements of m, then so is

$$c = 3 * 3c - 8c.$$

But $a, b, c, d, e, 1-t, 1-q, r, s \in m$ implies $\pi_m(B) = 1$ and thus we obtain as $(A, \varepsilon_{3\alpha+2\beta}(1)) \sim B$ that

$$\pi_m(A, \varepsilon_{3\alpha+2\beta}(1)) = 1.$$

This finishes the proof. □

Next, we show the following:

Proposition 4.5.7. *Let R be a principal ideal domain and $A \in G_2(R)$. Then there is an ideal*

$$I(A) \subset \varepsilon_l(A, 14062848)$$

such that $V(I(A)) \subset \Pi(\{A\})$. Phrased differently, for R a principal ideal domain, one can pick $L(G_2)$ in Theorem 3.2.2 as $L(G_2) = 14062848$.

Proof. First, we will construct the ideal $I(A)$. For an element $T \in G_2(R)$ let $I_0(T)$ be the ideal constructed in Proposition 4.5.2.

Define

$$\begin{aligned} I(A) :=& I_0(A) + I_0(w_\beta A w_\beta^{-1}) + I_0((w_\alpha w_\beta)^3 A (w_\alpha w_\beta)^{-3}) \\ &+ I_0((w_\alpha w_\beta)^3 w_\beta A w_\beta^{-1} (w_\alpha w_\beta)^{-3}) + I_0(w_\beta w_\alpha (A, \varepsilon_\alpha(1)) w_\alpha^{-1} w_\beta^{-1}) \\ &\subset \varepsilon_l(A, 4 * 2343808 + 2 * 2343808) = \varepsilon_l(A, 14062848). \end{aligned}$$

We will prove that $I(A)$ has the desired properties. Let m be a maximal ideal containing $I(A)$ and set $\bar{A} := \pi_m(A)$ and $K := R/m$. We will show that $\bar{A}$ is central in $G_2(K)$. To see this, observe first that K is a field. Thus the Bruhat-decomposition [41, Chapter 3, p. 26, Theorem 4'] of $G_2(K)$ implies that, one can find elements $u_1 \in U^+(G_2, K), w \in W(G_2), t, s \in K - \{0\}$ as well as $u_2 \in U^+(G_2, K)$ with the property that $w u_2 w^{-1} \in U^-(G_2, K)$, such that

$$\bar{A} = u_1 h_\alpha(s) h_\beta(t) w u_2.$$

By definition of $I(A)$, the maximal ideal m contains $I_0(A)$ and hence as $\varepsilon_{3\alpha+2\beta}(1)$ commutes with all elements in $U^+(G_2, K)$, we obtain

$$\begin{aligned} 1 &= (\bar{A}, \varepsilon_{3\alpha+2\beta}(1)) \\ &= (u_1 h_\alpha(s) h_\beta(t) w u_2, \varepsilon_{3\alpha+2\beta}(1)) = (u_2, \varepsilon_{3\alpha+2\beta}(1))^{u_1 h_\alpha(s) h_\beta(t) w} \cdot (u_1 h_\alpha(s) h_\beta(t) w, \varepsilon_{3\alpha+2\beta}(1)) \\ &= 1^{u_1 h_\alpha(s) h_\beta(t) w} \cdot (u_1 h_\alpha(s) h_\beta(t) w, \varepsilon_{3\alpha+2\beta}(1)) = (u_1 h_\alpha(s) h_\beta(t) w, \varepsilon_{3\alpha+2\beta}(1)) \\ &\sim (h_\alpha(s) h_\beta(t) w, \varepsilon_{3\alpha+2\beta}(1)) = h_\alpha(s) h_\beta(t) w \varepsilon_{3\alpha+2\beta}(1) w^{-1} h_\beta(t^{-1}) h_\alpha(s^{-1}) \varepsilon_{3\alpha+2\beta}(-1) \\ &= h_\alpha(s) h_\beta \varepsilon_{w(3\alpha+2\beta)}(\pm 1) h_\beta(t^{-1}) h_\alpha(s^{-1}) \varepsilon_{3\alpha+2\beta}(-1) \\ &= \varepsilon_{w(3\alpha+2\beta)}(\pm s^{\langle w(3\alpha+2\beta), \alpha \rangle} t^{\langle w(3\alpha+2\beta), \beta \rangle}) \varepsilon_{3\alpha+2\beta}(-1). \end{aligned}$$

But according to [41, Chapter 3, p. 21, Corollary 2], the only way that the last term can possibly be conjugate to 1 and hence be trivial is if $w(3\alpha + 2\beta) = 3\alpha + 2\beta$. One easily, checks that this restricts the possible values for $w \in W(G_2)$ to $w = w_\alpha$ and $w = 1$.

Let us assume for contradiction that $w = w_\alpha$. But $u_2 \in U^+(G_2, K)$ has the property $w u_2 w^{-1} \in U^-(G_2, K)$ and hence, one obtains that there is a $y \in K$ with $u_2 = \varepsilon_\alpha(y)$.

Furthermore, we can find $a, b, c, d, e, f \in K$ with

$$u_1 = \varepsilon_{3\alpha+2\beta}(f)\varepsilon_{3\alpha+\beta}(e)\varepsilon_{2\alpha+\beta}(d)\varepsilon_{\alpha+\beta}(c)\varepsilon_\alpha(b)\varepsilon_\beta(a).$$

Recall that $I_0(w_\beta A w_\beta^{-1})$ is a subset of m and hence $w_\beta \bar{A} w_\beta^{-1}$ commutes with $\varepsilon_{3\alpha+2\beta}(1)$ in $G_2(K)$. This implies

$$\begin{aligned}
1 &= (\bar{A}, w_\beta^{-1}\varepsilon_{3\alpha+2\beta}(1)w_\beta) = (\bar{A}, \varepsilon_{3\alpha+\beta}(\pm 1)) \\
&= (\varepsilon_{3\alpha+2\beta}(f)\varepsilon_{3\alpha+\beta}(e)\varepsilon_{2\alpha+\beta}(d)\varepsilon_{\alpha+\beta}(c)\varepsilon_\alpha(b)\varepsilon_\beta(a)h_\alpha(s)h_\beta(t)w_\alpha\varepsilon_\alpha(y), \varepsilon_{3\alpha+\beta}(\pm 1)) \\
&= (\varepsilon_{3\alpha+2\beta}(f)\varepsilon_{3\alpha+\beta}(e)\varepsilon_{2\alpha+\beta}(d)\varepsilon_{\alpha+\beta}(c)\varepsilon_\alpha(b)\varepsilon_\beta(a)h_\alpha(s)h_\beta(t)w_\alpha, \varepsilon_{3\alpha+\beta}(\pm 1)) \\
&\sim (\varepsilon_\beta(a)h_\alpha(s)h_\beta(t)w_\alpha, \varepsilon_{3\alpha+\beta}(\pm 1)) \\
&\sim (h_\alpha(s)h_\beta(t)w_\alpha, \varepsilon_{3\alpha+\beta}(\pm 1)) \cdot (\varepsilon_{3\alpha+\beta}(\pm 1), \varepsilon_\beta(-a)) \\
&= h_\alpha(s)h_\beta(t)w_\alpha\varepsilon_{3\alpha+\beta}(\pm 1)w_\alpha^{-1}h_\beta(t^{-1})h_\alpha(s^{-1})\varepsilon_{3\alpha+\beta}(\mp 1) \cdot \varepsilon_{3\alpha+2\beta}(\pm a) \\
&= h_\alpha(s)h_\beta(t)\varepsilon_\beta(\pm 1)h_\beta(t^{-1})h_\alpha(s^{-1})\varepsilon_{3\alpha+\beta}(\mp 1) \cdot \varepsilon_{3\alpha+2\beta}(\pm a) \\
&= \varepsilon_\beta(\pm s^{\langle\beta,\alpha\rangle}t^{\langle\beta,\beta\rangle})\varepsilon_{3\alpha+\beta}(\mp 1)\varepsilon_{3\alpha+2\beta}(\pm a) \\
&= \varepsilon_\beta(\pm s^{-3}t^2)\varepsilon_{3\alpha+\beta}(\mp 1)\varepsilon_{3\alpha+2\beta}(\pm a)
\end{aligned}$$

But according to [41, Chapter 3, p. 21, Corollary 2], independently of the values of s, t and a, the last line is never trivial and this contradiction implies $w = 1$. Hence, $\bar{A}$ is an element of $B(G_2, K) = B^+(G_2, K)$. Recall further that

$$I_0((w_\alpha w_\beta)^3 A (w_\alpha w_\beta)^{-3}) \subset m$$

and

$$I_0((w_\alpha w_\beta)^3 w_\beta A w_\beta^{-1}(w_\alpha w_\beta)^{-3}) \subset m$$

This implies that $\bar{A}$ commutes with the two elements

$$(w_\alpha w_\beta)^{-3}\varepsilon_{3\alpha+2\beta}(1)(w_\alpha w_\beta)^3 = \varepsilon_{-3\alpha-2\beta}(\pm 1)$$

and

$$w_\beta^{-1}(w_\alpha w_\beta)^{-3}\varepsilon_{3\alpha+2\beta}(1)(w_\alpha w_\beta)^3 w_\beta = \varepsilon_{-3\alpha-\beta}(\pm 1)$$

and so we obtain similar to the previous calculations, that $\bar{A} \in B^-(G_2, K)$ holds as well. But this implies

$$\bar{A} \in B(G_2, K) \cap B^-(G_2, K) = \{h_\alpha(a)h_\beta(b) |\ a, b \in K - \{0\}\}.$$

Hence, we obtain $\bar{A} = h_\alpha(s)h_\beta(t)$ for certain $s, t \in K - \{0\}$. But this implies that

$$\begin{aligned} 1 = (\bar{A}, \varepsilon_{3\alpha+2\beta}(1)) &= h_\alpha(s)h_\beta(t)\varepsilon_{3\alpha+2\beta}(1)h_\beta(t^{-1})h_\alpha(s^{-1})\varepsilon_{3\alpha+2\beta}(-1) \\ &= \varepsilon_{3\alpha+2\beta}(s^{\langle 3\alpha+2\beta,\alpha\rangle}t^{\langle 3\alpha+2\beta,\beta\rangle})\varepsilon_{3\alpha+2\beta}(-1) \\ &= \varepsilon_{3\alpha+2\beta}(s^0t^1 - 1) = \varepsilon_{3\alpha+2\beta}(t-1). \end{aligned}$$

Hence $t = 1$ holds and thus $\bar{A} = h_\alpha(s)$. Further, $\bar{A}$ commutes with $\varepsilon_{3\alpha+\beta}(1)$ and hence

$$1 = (\bar{A}, \varepsilon_{3\alpha+\beta}(1)) = h_\alpha(s)\varepsilon_{3\alpha+\beta}(1)h_\alpha(s^{-1})\varepsilon_{3\alpha+\beta}(-1) = \varepsilon_{3\alpha+\beta}(s^3 - 1).$$

This implies $s^3 = 1$. Last, observe that

$$I_0(w_\beta w_\alpha(A, \varepsilon_\alpha(1))w_\alpha^{-1}w_\beta^{-1}) \subset m.$$

This implies that $w_\beta w_\alpha(\bar{A}, \varepsilon_\alpha(1))w_\alpha^{-1}w_\beta^{-1}$ commutes with $\varepsilon_{3\alpha+2\beta}(1)$ and hence that $(\bar{A}, \varepsilon_\alpha(1))$ commutes with

$$w_\alpha^{-1}w_\beta^{-1}\varepsilon_{3\alpha+2\beta}(1)w_\beta w_\alpha = w_\alpha^{-1}\varepsilon_{3\alpha+\beta}(\pm 1)w_\alpha = \varepsilon_\beta(\pm 1).$$

But observe

$$(\bar{A}, \varepsilon_\alpha(1)) = h_\alpha(s)\varepsilon_\alpha(1)h_\alpha(s^{-1})\varepsilon_\alpha(-1) = \varepsilon_\alpha(s^2 - 1).$$

Hence

$$\begin{aligned} 1 = ((\bar{A}, \varepsilon_\alpha(1)), \varepsilon_\beta(1)) &= (\varepsilon_\alpha(s^2-1), \varepsilon_\beta(1)) \\ &= \varepsilon_{3\alpha+2\beta}(\pm(s^2-1)^3)\varepsilon_{3\alpha+\beta}(\pm(s^2-1)^3)\varepsilon_{2\alpha+\beta}(\pm(s^2-1)^2)\varepsilon_{\alpha+\beta}(\pm(s^2-1)) \end{aligned}$$

follows. But this implies in turn that $s^2 - 1 = 0$ and hence $s^2 = 1$. But if both $s^2 = 1$ and $s^3 = 1$ hold, then $s = 1$ follows and hence $\bar{A} = 1$. This finishes the proof. □

Chapter 5

Strong and uniform boundedness and stable range conditions

The strong boundedness theorems, Theorem 3.1.2 and Theorem 3.2.5 of Chapter 3 naturally raise the question, which rings satisfy the main assumption, that is bounded generation by root elements and in case of $\Phi \neq C_2$ or G_2 what the value of $Q(\Phi, R)$ is. The main examples we talk about in this and the next chapter are rings of algebraic integers and semi-local rings. Both of these classes of rings have stable range at most 2, so we talk in the first section of this chapter about stable range conditions and their connection to bounded generation. In the second section, we analyze the boundedness properties of Chevalley groups defined over semi-local rings in greater depth. In the third section, we show for a specific infinite ring R that the assumption in Theorem 3.2.5 that $R/2R$ is finite is not necessary to show that $\mathrm{Sp}_4(R)$ is strongly bounded.

5.1 Stable range conditions and matrix decompositions

We first define the stable range of rings:

Definition 5.1.1. [5, Ch. 1,§4] The *(Bass) stable range* of a commutative ring R with 1 is the smallest $n \in \mathbb{N}$ with the following property: If any $v_0, \ldots, v_m \in R$ generate the unit ideal R for $m \geq n$, then there are $t_1, \ldots, t_m \in R$ such that the elements $v_1' := v_1 + t_1 v_0, \ldots, v_m' := v_m + t_m v_0$ also generate the unit ideal. If no such n exists, R has stable range $+\infty$.

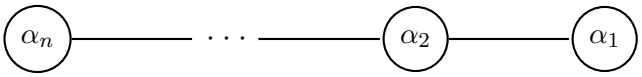
Remark 5.1.2. If for each $a \in R - \{0\}$ the ring R/aR has stable range 1, then R is said to have *stable range at most* 3/2. A ring R with stable range at most 3/2 has stable range at most 2 as well. Further, having stable range at most m for $m \in \mathbb{N}$ or at most 3/2 are first order properties.

Next, let $n \geq 1$ be given. Then picking the standard representation of $\mathrm{SL}_{n+1}(\mathbb{C})$ with

maximal torus the diagonal matrices, one obtains

$$G(A_n, R) = \mathrm{SL}_{n+1}(R) = \{A \in R^{(n+1)\times(n+1)} | \det(A) = 1\}$$

We can choose a system of positive simple roots $\{\alpha_1, \ldots, \alpha_n\}$ in the root system A_n such that the corresponding Dynkin diagram looks as follows



Then the root elements for positive, simple roots can be chosen as $\varepsilon_{\alpha_i}(t) = I_{n+1} + te_{n+1-i,n-i+2}$ for $t \in R$ and $1 \leq i \leq n$. More generally, the root elements for positive roots are then the elementary matrices $I_{n+1} + xe_{i,j}$ for $1 \leq i < j \leq n+1$ and the root elements for negative roots are $I_{n+1} + xe_{i,j}$ for $1 \leq j < i \leq n+1$. This further yields $U^+(A_n, R)$ as the group of upper unitriangular matrices and $U^-(A_n, R)$ as the group of lower unitriangular matrices in $R^{(n+1)\times(n+1)}$. Modulo these choices, note the following result:

Proposition 5.1.3. *[17, Lemma 9] Let $m \in \mathbb{N}$ and $n \geq m$ be given and let R be a commutative ring with 1 of stable range at most m.*

1. *If $m \geq 2$, let $\mathrm{SL}_m(R)$ be identified with a subgroup of $\mathrm{SL}_n(R)$ as follows*

$$\mathrm{SL}_m(R) = \{\begin{pmatrix} I_{n-m} & \\ & A \end{pmatrix} \mid A \in \mathrm{SL}_m(R)\}.$$

   Then $\mathrm{SL}_n(R) = (U^+(A_{n-1}, R) \cdot U^-(A_{n-1}, R))^2 \cdot \mathrm{SL}_m(R)$ holds.

2. *If $m = 1$, then $\mathrm{SL}_n(R) = (U^+(A_{n-1}, R) \cdot U^-(A_{n-1}, R))^2$ holds.*

Recalling the choices made for the symplectic group in Section 4.1, we obtain the following decomposition for symplectic groups:

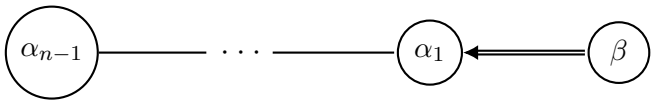
Proposition 5.1.4. *Let R be a ring of stable range at most 2 such that the group $\mathrm{Sp}_4(R)$ is generated by its root elements and let $n \geq 2$ be given. Then identifying $\mathrm{Sp}_4(R)$ with the subgroup*

$$\mathrm{Sp}_4(R) = \left\{ \left(\begin{array}{cc|cc} I_{n-2} & & 0_{n-2} & \\ & A & & B \\ \hline 0_{n-2} & & I_{n-2} & \\ & C & & D \end{array} \right) \mid \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right) \in \mathrm{Sp}_4(R) \right\}$$

of $\mathrm{Sp}_{2n}(R)$, the following decomposition holds for the elementary subgroup $E(C_n, R)$ of $\mathrm{Sp}_{2n}(R)$:

$$E(C_n, R) = (U^+(C_n, R) \cdot U^-(C_n, R))^2 \cdot \mathrm{Sp}_4(R)$$

Proof. In Section 4.1, we choose a system of positive simple roots $\{\alpha_1, \ldots, \alpha_{n-1}, \beta\}$ in C_n such that the Dynkin-diagram of this system of positive simple roots has the following form

C_n : 

and such that $\varepsilon_{\alpha_i}(t) = I_{2n} + t(e_{n-i,n-i+1} - e_{2n-i+1,2n-i})$ for $1 \leq i \leq n-1$ and $\varepsilon_\beta(t) = I_{2n} + te_{n,2n}$ hold for all $t \in R$.

We prove this proposition by adapting the strategy of the proof of [42, Proposition 1] and proceeding by induction on $n \in \mathbb{N}$. First, the statement is obvious for $n = 2$. Next, set

$$X := U^+(C_n, R)U^-(C_n, R)U^+(C_n, R)U^-(C_n, R)\mathrm{Sp}_4(R).$$

To prove $X = E(C_n, R)$ it suffices to show $A \cdot X \subset X$ for all $A \in E(C_n, R)$, because X contains the neutral element I_{2n}. However $E(C_n, R)$ is generated by root elements and hence it suffices to show $\varepsilon_\phi(t)X \subset X$ for all $\phi \in C_n$ and $t \in R$. However, each root element is conjugate modulo elements of the Weyl group to an element of the form $\varepsilon_\phi(t)$ for ϕ a simple positive root. Thus it suffices to show $w_\phi X \subset X$ and $\varepsilon_\phi(t)X \subset X$ for all positive, simple roots ϕ and all $t \in R$. But

$$w_\phi = \varepsilon_\phi(1)\varepsilon_{-\phi}(-1)\varepsilon_\phi(1)$$

holds. Hence it suffices to show $\varepsilon_\phi(t)X \subset X$ and $\varepsilon_{-\phi}(t)X \subset X$ for all positive, simple roots ϕ and all $t \in R$.

However due to the definition of X, it is clear that $\varepsilon_\phi(t)X \subset X$ holds for all positive, simple roots ϕ and all $t \in R$. Thus it suffices to show $\varepsilon_{-\phi}(t)X \subset X$ for all positive, simple roots ϕ and all $t \in R$.

We distinguish two cases for ϕ. First, assume that ϕ is not α_{n-1}. Then we separate C_n into two subsets: The subset Φ_1 of roots whose expression in terms of simple roots does not involve α_{n-1} and its complement Φ_2 in C_n. Clearly Φ_1 is as a root subsystem of C_n isomorphic to C_{n-1}. Next, observe that according to [41, Chapter 11, p. 104, Lemma 62] and slightly abusing notation by writing $U^\pm(\Phi_2, R)$ for the subgroup of $E(C_n, R)$ generated by the elements $\{\varepsilon_\phi(x) |\ x \in R, \phi \in \Phi_2 \cap \Phi^\pm\}$, we have

$$U^+(\Phi_1, R) \cdot U^-(\Phi_2, R) = U^-(\Phi_2, R) \cdot U^+(\Phi_1, R) \text{ and} \tag{5.1}$$
$$U^-(\Phi_1, R) \cdot U^+(\Phi_2, R) = U^+(\Phi_2, R) \cdot U^-(\Phi_1, R). \tag{5.2}$$

Reordering the terms in $U^+(C_n, R)$, we have further

$$U^+(C_n, R) = U^+(\Phi_2, R) \cdot U^+(\Phi_1, R) = U^+(\Phi_1, R) \cdot U^+(\Phi_2, R)$$

and similarly

$$U^-(C_n, R) = U^-(\Phi_2, R) \cdot U^-(\Phi_1, R) = U^-(\Phi_1, R) \cdot U^-(\Phi_2, R)$$

and applying these equation for $U^-(C_n, R)$ and $U^+(C_n, R)$ together with (5.1) and (5.2) repeatedly to the definition of X, we obtain

$$X = (U^+(\Phi_2, R)U^-(\Phi_2, R))^2(U^+(\Phi_1, R)U^-(\Phi_1, R))^2\mathrm{Sp}_4(R).$$

But note that $\varepsilon_{-\phi}(t)$ is an element of $U^-(\Phi_1, R)$ and hence applying (5.1) and ((5.2)) repeatedly, we obtain

$$\begin{aligned}\varepsilon_{-\phi}(t)X &= \varepsilon_{-\phi}(t) \cdot (U^+(\Phi_2, R)U^-(\Phi_2, R))^2 \cdot (U^+(\Phi_1, R)U^-(\Phi_1, R))^2 \cdot \mathrm{Sp}_4(R) \\ &\subset U^-(\Phi_1, R) \cdot (U^+(\Phi_2, R)U^-(\Phi_2, R))^2 \cdot (U^+(\Phi_1, R)U^-(\Phi_1, R))^2 \cdot \mathrm{Sp}_4(R) \\ &= (U^+(\Phi_2, R)U^-(\Phi_2, R))^2 \cdot [U^-(\Phi_1, R) \cdot (U^+(\Phi_1, R) \cdot U^-(\Phi_1, R))^2 \cdot \mathrm{Sp}_4(R)].\end{aligned}$$

However, note that Φ_1 is equal to the root subsystem generated by $\alpha_1, \ldots, \alpha_{n-2}, \beta$ and $U^-(\Phi_1, R) \cdot (U^+(\Phi_1, R)U^-(\Phi_1, R))^2\mathrm{Sp}_4(R)$ is a subset of the subgroup

$$\left\{ \left(\begin{array}{cc|cc} 1 & & 0 & \\ & A & & B \\ \hline 0 & & 1 & \\ & C & & D \end{array} \right) \mid \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right) \in E(C_{n-1}, R) \right\}$$

of $\mathrm{Sp}_{2n}(R)$. But this subgroup is isomorphic to $E(C_{n-1}, R)$ and so we know by induction that

$$U^-(\Phi_1, R) \cdot (U^+(\Phi_1, R)U^-(\Phi_1, R))^2\mathrm{Sp}_4(R) \subset (U^+(\Phi_1, R)U^-(\Phi_1, R))^2\mathrm{Sp}_4(R)$$

holds and hence

$$\varepsilon_{-\phi}(t)X \subset (U^+(\Phi_2, R)U^-(\Phi_2, R))^2(U^+(\Phi_1, R)U^-(\Phi_1, R))^2\mathrm{Sp}_4(R) = X$$

holds. This finishes the case $\phi \neq \alpha_{n-1}$.

If $\phi = \alpha_{n-1}$ holds, then we decompose C_n into the subset Φ_1 of roots whose expression in terms of simple roots does not involve β and its complement Φ_2 in C_n. Then as before, we can observe

$$\varepsilon_{-\phi}(t)X \subset (U^+(\Phi_2, R)U^-(\Phi_2, R))^2 \cdot [U^-(\Phi_1, R) \cdot (U^+(\Phi_1, R)U^-(\Phi_1, R))^2]\mathrm{Sp}_4(R).$$

However, the root system Φ_1 is equal to the root subsystem generated by $\alpha_1, \ldots, \alpha_{n-1}$

and $[U^-(\Phi_1, R) \cdot (U^+(\Phi_1, R)U^-(\Phi_1, R))^2]$ is a subset of the subgroup

$$\left\{ \left(\begin{array}{c|c} A & 0_n \\ \hline 0_n & A^{-T} \end{array} \right) \mid A \in E(A_{n-1}, R) \right\}$$

of $\mathrm{Sp}_{2n}(R)$. But this subgroup is isomorphic to the subgroup $E(A_{n-1}, R)$ of $\mathrm{SL}_n(R)$ Thus by applying Proposition 5.1.3 and the fact that R has stable range 2, we obtain

$$U^-(\Phi_1, R) \cdot (U^+(\Phi_1, R)U^-(\Phi_1, R))^2 \subset (U^+(\Phi_1, R)U^-(\Phi_1, R))^2 H$$

for

$$H := \left\{ \left(\begin{array}{cc|cc} I_{n-2} & & & \\ & A & & 0_n \\ \hline & & I_{n-2} & \\ 0_n & & & A^{-T} \end{array} \right) \mid A \in \mathrm{SL}_2(R) \right\}$$

But H is a subgroup of $\mathrm{Sp}_4(R)$ and hence we obtain

$$\begin{aligned} \varepsilon_\phi(t)X &\in (U^+(\Phi_2, R)U^-(\Phi_2, R))^2 \cdot [U^-(\Phi_1, R) \cdot (U^+(\Phi_1, R)U^-(\Phi_1, R))^2] \cdot \mathrm{Sp}_4(R) \\ &\subset (U^+(\Phi_2, R)U^-(\Phi_2, R))^2 \cdot (U^+(\Phi_1, R)U^-(\Phi_1, R))^2 \cdot H \cdot \mathrm{Sp}_4(R) = X \\ &\subset (U^+(\Phi_2, R)U^-(\Phi_2, R))^2 \cdot (U^+(\Phi_1, R)U^-(\Phi_1, R))^2 \cdot \mathrm{Sp}_4(R) = X. \end{aligned}$$

This finishes the proof. □

Remark 5.1.5. This is a result with quite a similar proof as the classical result [40, Theorem 2.5] and we suspect it is well-known or obvious to experts in algebraic K-theory.

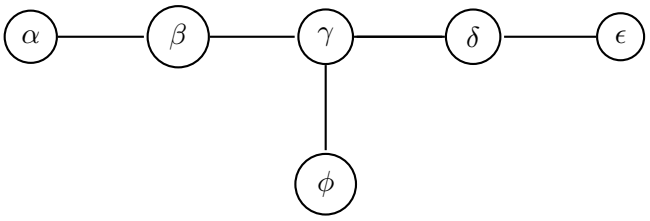
Note the following observation:

Lemma 5.1.6. *Let R be a principal ideal domain. Then R has stable range at most 2.*

The proof can be found in Appendix C. In a similar fashion to Proposition 5.1.4 one can prove the following two propositions:

Proposition 5.1.7. *Let R be a principal ideal domain such that $\mathrm{SL}_2(R) = G(A_1, R)$ is generated by root elements. Further, let $\{\alpha, \beta, \gamma, \delta, \epsilon, \phi\}$ be a system of simple, positive roots such that the corresponding Dynkin diagram has the following form*

E_6 :



Then $E(E_6, R) = (U^+(E_6, R) \cdot U^-(E_6, R))^2 \cdot G_\epsilon(R)$ holds.

and:

Proposition 5.1.8. *Let R be a commutative ring with 1 and $N \in \mathbb{N}$ such that*

$$G(A_1, R) = E(A_1, R) = (U^+(A_1, R)U^-(A_1, R))^N,$$
$$G(A_1, R) = E(A_1, R) = U^-(A_1, R)(U^+(A_1, R)U^-(A_1, R))^N$$

or

$$G(A_1, R) = E(A_1, R) = (U^+(A_1, R)U^-(A_1, R))^N U^+(A_1, R)$$

holds. Then

$$E(\Phi, R) = (U^+(\Phi, R)U^-(\Phi, R))^N,$$
$$E(\Phi, R) = U^-(\Phi, R)(U^+(\Phi, R)U^-(\Phi, R))^N$$

or

$$E(\Phi, R) = (U^+(\Phi, R)U^-(\Phi, R))^N U^+(\Phi, R)$$

respectively holds for all irreducible root systems Φ. In particular,

$$E(\Phi, R) = (U^+(\Phi, R)U^-(\Phi, R))^2$$

holds for R a ring of stable range 1.

This corollary is mainly useful for semi-local rings as seen in the next section. Next, we give a more detailed analysis of the asymptotics of bounded generation for SL_n and Sp_{2n}. First, recall the following word norm from Definition 2.2.2:

Definition 5.1.9. Let R be a commutative ring with 1 and Φ an irreducible root system such that $G(\Phi, R)$ is generated by root elements. Then define the set

$$\mathrm{EL}_Q := \{A\varepsilon_\phi(t)A^{-1} |\ t \in R, \phi \in \Phi, A \in G(\Phi, R)\}$$

and define the word norm $\|\cdot\|_{\mathrm{EL}_Q} : G(\Phi, R) \to \mathbb{N}_0$ as $\|1\|_{\mathrm{EL}_Q} := 0$ and as

$$\|X\|_{\mathrm{EL}_Q} := \min\{n \in \mathbb{N} | \exists A_1, \ldots, A_n \in \mathrm{EL}_Q : X = A_1 \cdots A_n\}$$

for $X \neq 1$.

Having this norm, we can show:

Proposition 5.1.10. *Let R be a principal ideal domain and let $n \geq 3$.*

1. *If $\mathrm{Sp}_4(R)$ and $\mathrm{Sp}_{2n}(R)$ are generated by its root elements and there is a $K \in \mathbb{N}$ with*

$$\|\mathrm{Sp}_4(R)\|_{\mathrm{EL}_Q} \leq K,$$

then

$$\|\mathrm{Sp}_{2n}(R)\|_{\mathrm{EL}_Q} \leq 12(n-2) + K$$

2. *If* $\mathrm{SL}_3(R)$ *and* $\mathrm{SL}_n(R)$ *are generated by its root elements and there is a* $K \in \mathbb{N}$ *with*

$$\|\mathrm{SL}_3(R)\|_{\mathrm{EL}_Q} \leq K,$$

then

$$\|\mathrm{SL}_n(R)\|_{\mathrm{EL}_Q} \leq 4(n-3) + K.$$

Proof. We only deal with the case of $\mathrm{Sp}_{2n}(R)$, because the statement for $\mathrm{SL}_n(R)$ is the content of [24, Proposition 6.20].

Considering $\mathrm{Sp}_4(R)$ as a subgroup of $\mathrm{Sp}_{2n}(R)$ as done in Proposition 5.1.4, we first prove by induction that:

Claim 5.1.10.1. *For each* $A \in U^+(C_n, R)$ *there is an* $A' \in U^+(C_2, R)$ *with* $\|A'^{-1}A\|_{\mathrm{EL}_Q} \leq 3(n-2)$ *for* $n \geq 2$.

First, the claim is clear for $n = 2$.

Let $A \in U^+(C_n, R)$ be given. Then it has the form

$$A = \left(\begin{array}{ccccc|cccc} 1 & a_{1,2} & \cdot & \cdot & a_{1,n} & a_{1,n+1} & \cdot & \cdot & \cdot & a_{1,2n} \\ & 1 & a_{2,3} & \cdot & a_{2,n} & a_{2,n+1} & \cdot & \cdot & \cdot & a_{2,2n} \\ & & 1 & \cdot & \cdot & a_{3,n+1} & \cdot & \cdot & \cdot & a_{3,2n} \\ & & & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & & 1 & a_{n,n+1} & \cdot & \cdot & \cdot & a_{n,2n} \\ \hline & & & & & 1 & & & & \\ & & & & & -a_{1,2} & 1 & & & \\ & & O_n & & & \cdot & -a_{2,3} & 1 & & \\ & & & & & \cdot & \cdot & \cdot & \cdot & \\ & & & & & -a_{1,n} & -a_{2,n} & \cdot & \cdot & 1 \end{array}\right)$$

Multiplying A with the matrix

$$T := (I_{2n} - a_{1,2}(e_{1,2} - e_{n+2,n+1})) \cdot (I_{2n} - a_{1,3}(e_{1,3} - e_{n+3,n+1})) \cdots (I_{2n} - a_{1,n}(e_{1,n} - e_{2n,n+1}))$$

$$= \left(\begin{array}{ccccc|c} 1 & -a_{1,2} & -a_{1,3} & \cdots & -a_{1,n} & \\ & 1 & & & & \\ & & 1 & & & 0_n \\ & & & . & & \\ & & & & 1 & \\ \hline & & & & & 1 \\ & & & & & a_{1,2} \quad 1 \\ & & 0_n & & & a_{1,3} \qquad 1 \\ & & & & & . \qquad\quad . \\ & & & & & a_{1,n} \qquad\quad 1 \end{array}\right)$$

from the right yields an element B of $U^+(C_n, R)$ with the first n entries of the first row of B being 0, except for the $(1,1)$-entry, which is 1. However, according to the proof of Lemma 4.1.4, there is a matrix $D \in \mathrm{Sp}_{2n}(R)$ of the form

$$D = \left(\begin{array}{cc|cc} 1 & & & \\ & D' & & 0_n \\ \hline & & 1 & \\ & 0_n & & D'^{-T} \end{array}\right)$$

for $D' \in \mathrm{SL}_{n-1}(R)$ such that the first column of DT^TD^{-1} has the form

$$(1, t, 0, \ldots, 0)^T$$

for $t = \gcd(-a_{1,2}, -a_{1,3}, \ldots, -a_{1,n})$. However, due to the form of T^T and D, this implies that $DT^TD^{-1} = I_{2n} + t(e_{21} - e_{n+1,n+2})$ and hence

$$D^TTD^{-T} = I_{2n} + t(e_{12} - e_{n+2,n+1})$$

holds. This implies $\|T\|_{\mathrm{EL}_Q} \leq 1$.

Then $B = A \cdot T$ has the form

$$B = \left(\begin{array}{ccccc|cccc} 1 & 0 & \cdot & \cdot & 0 & b_{1,n+1} & \cdot & \cdot & \cdot & b_{1,2n} \\ & 1 & b_{2,3} & \cdot & b_{2,n} & b_{2,n+1} & \cdot & \cdot & \cdot & b_{2,2n} \\ & & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & \cdot & \cdot & b_{n,n+1} & \cdot & \cdot & \cdot & b_{n,2n} \\ & & & & 1 & b_{n,n+1} & \cdot & \cdot & \cdot & b_{n,2n} \\ \hline & & & & & 1 & & & & \\ & & & & & 0 & 1 & & & \\ & & O_n & & & 0 & -b_{2,3} & 1 & & \\ & & & & & \cdot & \cdot & \cdot & \cdot & \\ & & & & & 0 & -b_{2,n} & \cdot & \cdot & 1 \end{array}\right)$$

Next, multiplying B with

$$\begin{aligned} S :=& (I_{2n} - b_{1,n+1}e_{1,n+1}) \cdot (I_{2n} - b_{1,n+2}(e_{1,n+2} + e_{2,n+1})) \cdot (I_{2n} - b_{1,n+3}(e_{1,n+3} + e_{3,n+1})) \\ & \cdots (I_{2n} - b_{1,2n}(e_{1,2n} + e_{n,n+1})) \\ =& \left(\begin{array}{c|ccccc} & -b_{1,n+1} & -b_{1,n+2} & -b_{1,n+3} & \cdots & -b_{1,2n} \\ & -b_{1,n+2} & & & & \\ & -b_{1,n+3} & & & & \\ I_n & \cdot & & 0_{n-1} & & \\ & \cdot & & & & \\ & -b_{1,2n} & & & & \\ \hline 0_n & & & I_n & & \end{array}\right) \end{aligned}$$

from the right yields an element $C \in U^+(C_n, R)$ whose first row is

$$(1, 0, \ldots, 0).$$

But applying the proof of Lemma 4.1.10, we can find a matrix of the form

$$E = \left(\begin{array}{cc|cc} 1 & & & \\ & E' & 0_n & \\ \hline & & 1 & \\ 0_n & & & E'^{-T} \end{array}\right)$$

for $E' \in \mathrm{SL}_{n-1}(R)$ such that the first column of ES^TE^{-1} has the form

$$(1, 0, \ldots, 0, -b_{1,n+1}, s, \ldots, 0)^T$$

for $s = \gcd(b_{1,n+2}, b_{1,n+3}, \ldots, b_{1,2n})$. However, due to the form of S^T and E, this implies that $ES^TE^{-1} = (I_{2n} - b_{1,n+1}e_{n+1,1}) \cdot (I_{2n} + s(e_{n+1,2} + e_{n+2,1}))$ and hence

$$E^TSE^{-T} = (I_{2n} - b_{1,n+1}e_{1,n+1}) \cdot (I_{2n} + s(e_{1,n+2} + e_{2,n+1}))$$

holds. This implies that $\|T\|_{\mathrm{EL}_Q} \leq 2$.

But note that C must be an element of the subgroup $U^+(C_{n-1}, R)$ of $U^+(C_n, R)$, if its first row is

$$(1, 0, \ldots, 0)^T.$$

This yields by induction that there is a $C' \in U^+(C_2, R)$ with

$$\|C'^{-1}C\|_{\mathrm{EL}_Q} \leq 3(n - 1 - 2) = 3(n - 3)$$

holds. Hence setting A' as C', one obtains from $C = ATS$ that

$$\begin{aligned}\|A'^{-1}A\|_{\mathrm{EL}_Q} &= \|C'^{-1}CS^{-1}T^{-1}\|_{\mathrm{EL}_Q} \leq \|C'^{-1}C\|_{\mathrm{EL}_Q} + \|T\|_{\mathrm{EL}_Q} + \|S\|_{\mathrm{EL}_Q} \\ &\leq 3(n-3) + 3 = 3(n-2).\end{aligned}$$

Thus the claim holds for all $n \geq 2$. Further, Proposition 5.1.4 yields that

$$\mathrm{Sp}_{2n}(R) = (U^+(C_n, R)U^-(C_n, R))^2\mathrm{Sp}_4(R)$$

for all $n \geq 2$. Let $A \in \mathrm{Sp}_{2n}(R)$ be given. Hence there are $u_1^+, u_2^+ \in U^+(C_n, R), u_1^-, u_2^- \in U^-(C_n, R)$ as well as $Z \in \mathrm{Sp}_4(R)$ with

$$A = u_1^+u_1^-u_2^+u_2^-Z$$

But $U^+(C_n, R)$ and $U^-(C_n, R)$ are conjugate in $\mathrm{Sp}_{2n}(R)$. Hence applying the claim of the first part of the proof to the $u_1^+, u_1^-, u_2^+, u_2^-$ yields $X_1, X_2, Y_1, Y_2 \in \mathrm{Sp}_{2n}(R)$ with

$$\|X_1\|_{\mathrm{EL}_Q}, \|X_2\|_{\mathrm{EL}_Q}, \|Y_1\|_{\mathrm{EL}_Q}, \|Y_2\|_{\mathrm{EL}_Q} \leq 3(n-2)$$

and $v_1^+, v_2^+ \in U^+(C_2, R)$ and $v_1^-, v_2^- \in U^-(C_2, R)$ such that

$$u_1^+ = v_1^+X_1, u_2^+ = v_2^+X_2, u_1^- = v_1^-Y_1, u_2^- = v_2^-Y_2.$$

But this implies

$$\begin{aligned} A &= u_1^+ u_1^- u_2^+ u_2^- Z = (v_1^+ X_1) \cdot (v_1^- Y_1) \cdot (v_2^+ X_2) \cdot (v_2^- Y_2) Z \\ &= (v_1^+ X_1 (v_1^+)^{-1}) \cdot (v_1^+ v_2^- Y_1 (v_1^+ v_2^-)^{-1}) \cdot (v_1^+ v_2^- v_2^+ X_2 (v_1^+ v_2^- v_2^+)^{-1}) \\ &\quad \cdot (v_1^+ v_2^- v_2^+ v_2^- X_2 (v_1^+ v_2^- v_2^+ v_2^-)^{-1}) \cdot (v_1^+ v_2^- v_2^+ v_2^-) \cdot Z \\ &= (X_1^{v_1^+}) \cdot (Y_1^{v_1^+ v_2^-}) \cdot (X_2^{v_1^+ v_2^- v_2^+}) \cdot (Y_2^{v_1^+ v_2^- v_2^+ v_2^-}) \cdot (v_1^+ v_2^- v_2^+ v_2^-) \cdot Z. \end{aligned}$$

But $(v_1^+ v_2^- v_2^+ v_2^-) \cdot Z$ is an element of $\mathrm{Sp}_4(R)$ and hence

$$\|(v_1^+ v_2^- v_2^+ v_2^-) \cdot Z\|_{\mathrm{EL}_Q} \leq K$$

holds. This implies

$$\begin{aligned} \|A\|_{\mathrm{EL}_Q} &= \|(X_1^{v_1^+}) \cdot (Y_1^{v_1^+ v_2^-}) \cdot (X_2^{v_1^+ v_2^- v_2^+}) \cdot (Y_2^{v_1^+ v_2^- v_2^+ v_2^-}) \cdot (v_1^+ v_2^- v_2^+ v_2^-) \cdot Z\|_{\mathrm{EL}_Q} \\ &\leq \|X_1\|_{\mathrm{EL}_Q} + \|Y_1\|_{\mathrm{EL}_Q} + \|X_2\|_{\mathrm{EL}_Q} + \|Y_2\|_{\mathrm{EL}_Q} + \|(v_1^+ v_2^- v_2^+ v_2^-) \cdot Z\|_{\mathrm{EL}_Q} \\ &\leq 4 * 3 * (n-2) + K = 12(n-2) + K. \end{aligned}$$

This yields the statement of the proposition for $\mathrm{Sp}_{2n}(R)$. □

5.2 Rings of stable range 1, semi-local rings and uniform boundedness

Proposition 5.1.8 states that $E(\Phi, R)$ is boundedly generated by root elements for all irreducible root systems Φ and all rings of stable range at most 1 and that that each element in $E(\Phi, R)$ can be written as a product of at most four upper and lower unipotent elements. This was observed by Vavilov, Smolenski, Sury in [45, Theorem 1]. The main example of rings of stable range 1 are semi-local rings, that is rings with only finitely many maximal ideals:

Lemma 5.2.1. *[5, Lemma 6.4, Corollary 6.5] Every semilocal ring, that is each ring with only finitely many maximal ideals has stable range* 1. *So also each field has stable range* 1.

Note:

Proposition 5.2.2. *[3, Corollary 2.4] Let R be a semi-local ring. Then for all irreducible root systems Φ of rank greater than one, the group $G(\Phi, R)$ is generated by root elements.*

This implies together with Proposition 5.1.8:

Proposition 5.2.3. *Let R be a semilocal ring and Φ a root system. Then the group $G(\Phi, R)$ is boundedly generated by root elements.*

Hence the strong boundedness theorems can be applied to $G(\Phi, R)$. For semi-local rings R, the group $G(\Phi, R)$ is in fact uniformly bounded. First, this can be seen quite abstractly:

Theorem 5.2.4. *Let R be a commutative, semilocal ring with 1 and let Φ an irreducible root system of rank at least 2. Furthermore, assume if $\Phi = C_2$ or G_2 that $(R : 2R) < \infty$ holds. Then $G(\Phi, R)$ is uniformly bounded.*

Proof. The strategy is to find a constant $K \in \mathbb{N}$ such that each finite normally generating subset S of $G := G(\Phi, R)$ has a subset $\bar{S}$ with $|\bar{S}| \leq K$ such that $\bar{S}$ is also a normally generating subset of $G(\Phi, R)$. Then Proposition 5.2.3 with Theorem 3.1.2 and Theorem 3.2.5 respectively yield the claim, because

$$\|G(\Phi, R)\|_S \leq \|G(\Phi, R)\|_{\bar{S}} \leq C(\Phi, R)|\bar{S}| \leq C(\Phi, R)K$$

and so uniform boundedness for $G(\Phi, R)$ holds.

Assume R has precisely q maximal ideals. Let S normally generate $G(\Phi, R)$. Corollary 3.2.8 implies $\Pi(S) = \emptyset$. But according to Lemma 3.0.2, we have $\Pi(T_1 \cup T_2) = \Pi(T_1) \cap \Pi(T_2)$ for all $T_1, T_2 \subset G(\Phi, R)$. This implies that if there are only q maximal ideals in R, then already some subset S' of S with $|S'| \leq q$ has the property $\bigcap_{A \in S'} \Pi(A) = \emptyset$. Hence in case $\Phi \neq C_2$ or G_2, Corollary 3.2.8 tells us that S' is already a normally generating subset of $G(\Phi, R)$. This finishes the case $\Phi \neq C_2, G_2$.

Next, we do the case $\Phi = C_2$ or G_2. We have $(R : 2R) < \infty$ by assumption and hence Lemma 3.2.3 implies for N_Φ, that the group G/N_Φ is finite. The set S normally generates the group G and hence the image of S in G/N_Φ normally generates G/N_Φ and so we can pick a subset $S'' \subset S$ with at most $M := |G/N_\Phi|$ elements such that the image of S'' in G/N normally generates G/N_Φ. Hence considering the set $\bar{S} := S' \cup S''$, we have

$$|\bar{S}| \leq |S'| + |S''| \leq q + M$$

and the upper bound $q + M$ clearly does not depend on S. Corollary 3.2.8 implies that $\bar{S}$ is a normally generating set of $G(\Phi, R)$. Thus we are done. □

We did give explicit values for $L(\Phi)$ for some Φ in case of principal ideal domains, so we obtain:

Corollary 5.2.5. *Let R be a semi-local ring and a principal ideal domain with at most q distinct maximal ideals. Further, let $n \geq 3$ and $k \in \mathbb{N}$ be given. Then*

1. $\Delta_k(\mathrm{SL}_n(R)) \leq 12(n-1)\min\{q, k(n+1)\}$ *and*
2. $\Delta_k(\mathrm{Sp}_{2n}(R)) \leq 768(3n-2)\min\{q, (5n+1)k\}$ *holds.*

Proof. The statement for $\mathrm{SL}_n(R)$ is the content of [24, Theorem 6.3]. For $\mathrm{Sp}_{2n}(R)$, let $S = \{A_1, \ldots, A_k\}$ be a subset normally generating $\mathrm{Sp}_{2n}(R)$. For $1 \leq i \leq k$, let $I(A_i) \subset \varepsilon_s(A_i, 64(1+5n))$ be the ideal given by Theorem 4.1.3 with $V(I(A_i)) \subset \Pi(\{A_i\})$. However Corollary 3.2.8 yields $V(I(A_1) + \cdots + I(A_k)) \subset \Pi(S) = \emptyset$ and so no maximal ideal can contain the ideal $I(A_1) + \cdots + I(A_k)$. Thus $\sum_{i=1}^{k} I(A_i) = R$ and so

$$R = \varepsilon_s(A, 64k(5n+1)). \tag{5.3}$$

According to Corollary 4.1.18, each of the $I(A_i)$ is a sum of $7n+1$ ideals $J_1(A_i), \ldots, J_{7n+1}(A_i)$, each of which is contained in $\varepsilon_s(A_i, 64)$. Hence

$$\sum_{i=1}^{k} \sum_{j=1}^{7n+1} J_j(A_i) = R$$

holds. Next, let m be one of the maximal ideals of R. Clearly not all of the ideals $J_j(A_i)$ can be contained in m. Hence there are $i(m) \in \{1, \ldots, k\}$ and $j(m) \in \{1, \ldots, 7n+1\}$ with

$$J_{j(m)}(A_{i(m)}) \not\subset m.$$

But this implies that

$$\sum_{m \text{ maximal ideal in R}} J_{j(m)}(A_{i(m)})$$

cannot be contained in any maximal ideal and thus must be the entire ring R. But this implies

$$R = \varepsilon_s(A, 64q) \tag{5.4}$$

Summarizing (5.3) and (5.4) yields

$$R = \varepsilon_s(A, 64 \min\{q, k(5n+1)\}) \tag{5.5}$$

holds. But then similar to the proof of Proposition 3.1.3, one can show that all root elements in $\mathrm{Sp}_{2n}(R)$ are contained in $B_S(3*64 \min\{q, k(5n+1)\}) = B_S(192 \min\{q, k(5n+1)\})$ and so are all elements in EL_Q. Thus

$$\|\mathrm{EL}_Q\|_S \leq 192 \min\{q, k(5n+1)\} \tag{5.6}$$

On the other hand, as R is semi-local and hence of stable range 1, one has

$$\mathrm{Sp}_4(R) = U^+(C_2, R)U^-(C_2, R)U^+(C_2, R)U^-(C_2, R)$$

according to Proposition 5.1.8 and hence

$$\|\mathrm{Sp}_4(R)\|_{\mathrm{EL}_Q} = 4 * \|U^+(C_2, R)\|_{\mathrm{EL}_Q} = 16$$

holds. Thus Proposition 5.1.10 yields

$$\|\mathrm{Sp}_{2n}(R)\|_{\mathrm{EL}_Q} \leq 12(n-2) + 16 = 12n - 8.$$

This bound together with (5.6) implies

$$\begin{aligned}\|\mathrm{Sp}_{2n}(R)\|_S &\leq \|\mathrm{Sp}_{2n}(R)\|_{\mathrm{EL}_Q} \cdot \|\mathrm{EL}_Q\|_S \leq (12n-8) \cdot 192 \min\{q, (5n+1)k\} \\ &= 768(3n-2)\min\{q, (5n+1)k\}.\end{aligned}$$

This finishes the proof for $\mathrm{Sp}_{2n}(R)$. □

Remark 5.2.6. One could improve the bound on $\|\mathrm{Sp}_{2n}(R)\|_{\mathrm{EL}_Q}$ further by observing that each element of the form $u_1^+ u_1^- u_2^+ u_2^-$ can be rewritten as the product $(u_1^-)^{u_1^+}(u_1^+ u_2^+)u_2^-$, and hence $\|\mathrm{Sp}_{2n}(R)\|_{\mathrm{EL}_Q} \leq 3\|U^+(C_n, R)\|_{\mathrm{EL}_Q}$ holds and thus it suffices to give bounds on $\|U^+(C_n, R)\|_{\mathrm{EL}_Q}$, which can be done in a similar way as in the proof of Proposition 5.1.10.

In this context, we also prove the following more explicit version of Theorem 5.2.4 in a special case where we can drop the assumption of R being a principal ideal domain:

Theorem 5.2.7. *Let R be a local ring and $n \geq 3$. Then $\Delta_\infty(\mathrm{SL}_n(R)) \leq 24(2n-3)$ holds.*

Proof. Let m be the unique maximal ideal of R. Further, let S be a normally generating subset of $\mathrm{SL}_n(R)$ and hence $\Pi(S) = \emptyset$ holds according to Corollary 3.2.8. But R has only one maximal ideal and so there must be at least one element $A(S) \in S$ such that already $\Pi(\{A(S)\}) = \emptyset$ holds. Hence using Corollary 3.2.8, we may assume that $S = \{A(S)\}$. For brevity write $A := (a_{ij}) := A(S)$ and so we only have to give an upper bound on the diameter $\|\mathrm{SL}_n(R)\|_A$. To this end, we distinguish two cases, first that there is no off-diagonal entry of A which is not an element of m and second that there is one. In the first case, there must be two $1 \leq k, l \leq n$ with $a_{kk} \not\equiv a_{ll} \bmod m$, because otherwise $\pi_m(A)$ would be scalar, which would imply $\Pi(\{A\}) = \{m\}$, which is not possible. After conjugation with possible Weyl group elements, we may assume $k = 1$ and $l = 2$. Then setting $\bar{a_{ij}} := a_{ij} + m \in R/m$, we obtain

$$\pi_m((A, I_n + e_{12})) = I_n + (\bar{a_{kk}}\bar{a_{ll}}^{-1} - 1)e_{12}.$$

This implies that the $(1,2)$-entry of $(A, I_n + e_{12})$ is not an element of m. Hence in both cases, we may assume that there is an element $B = (b_{ij}) \in B_A(2)$ with an off-diagonal entry that is not an element of m. Again after conjugation, we may assume that this entry

of B, which is not an element of m is the $(2,1)$-entry of B. In particular, the $(2,1)$-entry of B is a unit t in the ring R. But this in turn implies that

$$C := (I_n - t^{-1}b_{n1}e_{n,2})B(I_n + t^{-1}b_{1n}e_{n,2})$$

still has the $(2,1)$-entry equal to t, but also has the $(n,1)$-entry equal to 0.

Next, let $D = (c_{ij})$ be the inverse of C. Again, there must be an off-diagonal entry d_{uv} of C such that d_{uv} is not an element of m, because otherwise $\pi_m(C)$ and $\pi_m(D)$ would both be diagonal, which we know is not the case. Then [24, Lemma 6.7] implies that

$$E := ((C, I_n + e_{1u}), I_n + e_{v,n}) = I_n + d_{uv}Ce_{1n}.$$

Note, that E is an element of $B_A(4)$ and further observe that the $(2,n)$-entry of E is $d_{uv}t$. This finally implies for $x \in R$ arbitrary that

$$(I_n + xt^{-1}d_{uv}^{-1}e_{12}, E) = I_n + xe_{1n}$$

is an element of $B_A(8)$. In particular, this implies that $R = \varepsilon(A, 8)$ and so $\|\mathrm{EL}_Q\|_A \leq 8$. But R is local and hence has stable range 1, so Proposition 5.1.3 implies

$$\mathrm{SL}_n(R) = (U^+(A_{n-1}, R)U^-(A_{n-1}, R))^2.$$

So for $X \in \mathrm{SL}_n(R)$ arbitrary, there are $u_1^+, u_2^+ \in U^+(A_{n-1}, R)$ and $u_1^-, u_2^- \in U^-(A_{n-1}, R)$ such that $X = u_1^+u_1^-u_2^+u_1^-$. Hence we obtain

$$X = u_1^+u_1^-u_2^+u_1^- = (u_2^-)^{u_1^+}(u_1^+u_2^+)u_2^-$$

and so as the upper and lower unipotent groups are conjugate to each other, we obtain from $\|\mathrm{EL}_Q\|_A \leq 8$ that

$$\|X\|_A \leq 3\|U^+(A_{n-1}, R)\|_A \leq 3\|\mathrm{EL}_Q\|_A \cdot \|U^+(A_{n-1}, R)\|_{\mathrm{EL}_Q} \leq 24\|U^+(A_{n-1}, R)\|_{\mathrm{EL}_Q}.$$

Thus it suffices to give an upper bound on $\|U^+(A_{n-1}, R)\|_{\mathrm{EL}_Q}$ in order to give an upper bound on $\|\mathrm{SL}_n(R)\|_A$.

To this end we prove by induction on $n \geq 2$ that

$$\|U^+(A_{n-1}, R)\|_{\mathrm{EL}_Q} \leq 2n - 3.$$

First, observe that the case of $n = 2$ is clear. For the induction step observe first that for

$U = (u_{ij}) \in U^+(A_{n-1}, R)$, one obtains that

$$U' := U \cdot (I_n - u_{12}e_{12}) \cdot (I_n - u_{13}e_{13}) \cdots (I_n - u_{1n}e_{1n})$$

is an element of a subgroup of $U^+(A_{n-1}, R)$ isomorphic to $U^+(A_{n-2}, R)$. Thus by induction, we obtain $\|U'\|_{\mathrm{EL}_Q} \leq 2(n-1) - 3 = 2n - 5$. Next, consider the element

$$T := (I_n - u_{12}e_{12}) \cdot (I_n - u_{13}e_{13}) \cdots (I_n - u_{1n}e_{1n}) = \begin{pmatrix} 1 & -u_{12} & -u_{13} & \cdot & -u_{1n} \\ & 1 & 0 & \cdot & 0 \\ & & \cdot & & \\ & & & \cdot & \\ & & & & 1 \end{pmatrix}$$

We distinguish two cases now: The first case is that one of the $u_{12}, \ldots, u_{1n}$ is not an element of m and the second one is that they all are elements of m. In the first case, we may by conjugating with elements of the Weyl group assume that u_{12} is this element. In the second case, observe that

$$T' := (I_n + e_{12})T = (I_n + (1 - u_{12})e_{12}) \cdot (I_n - u_{13}e_{13}) \cdots (I_n - u_{1n}e_{1n})$$

has $(1,2)$-entry equal to $1 - u_{12}$. However as u_{12} is assumed to be an element of m, the element $1 - u_{12}$ cannot be an element of m. Thus up to multiplication with a single root element, we may assume in either case that the $(1,2)$-entry of T is not an element of m and we call this unit $s \in R$. But then observe further that

$$(I_n - s^{-1}u_{13}e_{23}) \cdots (I_n - s^{-1}u_{1n}e_{2n})T(I_n + s^{-1}u_{1n}e_{2n}) \cdots (I_n + s^{-1}u_{13}e_{23}) = I_n + se_{12}.$$

But this implies that in either of the two cases $\|T\|_{\mathrm{EL}_Q} \leq 2$ and so we obtain

$$\|U\|_{\mathrm{EL}_Q} = \|U'T^{-1}\|_{\mathrm{EL}_Q} \leq \|U'\|_{\mathrm{EL}_Q} + \|T\|_{\mathrm{EL}_Q} \leq 2n - 5 + 2 = 2n - 3.$$

But this finishes the induction and so we obtain $\|U^+(A_n, R)\|_{\mathrm{EL}_Q} \leq 2n-3$. This inequality together with the already seen inequality

$$\|\mathrm{SL}_n(R)\|_A \leq 24\|U^+(A_n, R)\|_{\mathrm{EL}_Q}$$

implies the claim of the theorem. □

Remark 5.2.8. One can generalize this theorem to the case of semi-local rings R as well, but this would be more involved.

We also obtain the following:

Theorem 5.2.9. *Let R be a commutative ring with 1 of stable range 1 and Φ an irreducible root system of rank at least 2. If $\Phi = C_2$ or G_2 assume further that $(R : 2R) < \infty$. Then for the elementary subgroup $E(\Phi, R)$ of $G(\Phi, R)$, there is a constant $C(\Phi, R)$ such that*

$$\Delta_k(E(\Phi, R)) \leq C(\Phi, R)k$$

for all $k \in \mathbb{N}$.

Proof. We want to show a version of Theorem 3.1.2 and Theorem 3.2.5 that speaks about $E(\Phi, R)$ instead of $G(\Phi, R)$. This can be done by following the same arguments essentially. The only differences in the proofs are found in the proofs of Theorem 3.1.1, Theorem 3.2.1 and Theorem 3.2.2. These differences are that in the sentences of the theories $\mathcal{T}_{kl}$, one must quantify over all elements of $E(\Phi, R)$ instead of elements of $G(\Phi, R)$. However, this issue can be resolved by breaking up the sentences called θ_r involved in the proofs further in such a way that the conjugating elements $X_1, \ldots, X_r$ appearing in the θ_r are only allowed to be products of at most r root elements. The rest of the proof then goes through in essentially the same way using that R being of stable range 1 implies $(U^+(\Phi, R)U^-(\Phi, R))^2 = E(\Phi, R)$ for such rings by Proposition 5.1.8 and hence $E(\Phi, R)$ is boundedly generated by root elements. □

5.3 Bounded generation and strong boundedness in positive characteristic

The main problem in positive characteristic is that there are no results concerning bounded generation by root elements known to us for these class of rings except for a result by Nica [34] stating bounded generation of $\mathrm{SL}_n(\mathbb{F}[T])$ for $\mathbb{F}$ a finite field and $n \geq 3$. We suspect that the following holds though:

Conjecture 5.3.1. *Let Φ be an irreducible root system of rank at least 2, K a global field and R a ring of S-algebraic integers in K. Then $G(\Phi, R)$ is boundedly generated by root elements.*

Another problem occurs in lower ranks, namely for $\Phi = C_2$ or G_2 : We often assume for those Φ that $R/2R$ is finite. But this is not generally the case for $\mathrm{char}(R) = 2$, say $R = \mathbb{F}_2[T]$. However, we still believe that strong boundedness is the norm for these rings and to illustrate this point we prove the following version of Theorem 5.2.4 in positive characteristic.

Theorem 5.3.2. *Let $\mathbb{F}$ be a finite field, $\mathcal{P}$ a prime ideal in $\mathbb{F}[T]$ and R the localization of $\mathbb{F}[T]$ at $\mathcal{P}$. Then $\mathrm{Sp}_4(R)$ is uniformly bounded.*

To show this, we need the following proposition:

Proposition 5.3.3. *Let $\mathcal{P}$ be a prime ideal of $\mathbb{F}[T]$ with $\mathbb{F}$ a finite field of characteristic 2 and let R be the localization of $\mathbb{F}[T]$ at $\mathcal{P}$. Further, let N be the normal subgroup of $\mathrm{Sp}_4(R)$ generated by*

$$A := \varepsilon_{\alpha+\beta}(1)\varepsilon_{2\alpha+\beta}(1).$$

Further, let $\|\cdot\|_A : \mathrm{Sp}_4(R) \to \mathbb{N}_0 \cup \{+\infty\}$ be the conjugation generated word norm on $\mathrm{Sp}_4(R)$ defined as in Defintion 2.0.1. Then

1. *N is a finite index subgroup of $\mathrm{Sp}_4(R)$ and*
2. *the norm $\|\cdot\|_A$ has finite diameter on N.*

Proof. First, we will show that

$$I := (x - x^2 | x \in R) \subset \varepsilon(A, \phi, 24)$$

for all $\phi \in C_2$ and second, we will deduce the two statements of the proposition from this. First, observe for any $x \in R$ that

$$B_A(2) \ni (\varepsilon_{\alpha+\beta}(1)\varepsilon_{2\alpha+\beta}(1), \varepsilon_{-\beta}(x)) = \varepsilon_\alpha(x)\varepsilon_{2\alpha+\beta}(x). \tag{5.7}$$

This yields that

$$\begin{aligned} B_A(2) \ni & w_\beta w_\alpha w_\beta \varepsilon_\alpha(x)\varepsilon_{2\alpha+\beta}(x) w_\beta^{-1} w_\alpha^{-1} w_\beta^{-1} \\ &= w_\beta w_\alpha \varepsilon_{\alpha+\beta}(x)\varepsilon_{2\alpha+\beta}(x) w_\alpha^{-1} w_\beta^{-1} \\ &= w_\beta \varepsilon_{\alpha+\beta}(x)\varepsilon_\beta(x) w_\beta^{-1} = \varepsilon_\alpha(x)\varepsilon_{-\beta}(x). \end{aligned}$$

On the other hand (5.7) implies for $x, y \in R$ that

$$\begin{aligned} B_A(4) \ni (\varepsilon_\alpha(y)\varepsilon_{2\alpha+\beta}(y), \varepsilon_{-(\alpha+\beta)}(x)) &= (\varepsilon_{2\alpha+\beta}(y), \varepsilon_{-(\alpha+\beta)}(x))^{\varepsilon_\alpha(y)} \varepsilon_{-\beta}(2xy) \\ &\sim \varepsilon_\alpha(xy)\varepsilon_{-\beta}(x^2 y). \end{aligned}$$

But this implies together with (5.7) that

$$\varepsilon_{-\beta}(x^2y - xy) = \varepsilon_{-\beta}(x^2y)\varepsilon_{-\beta}(xy) = \big(\varepsilon_{-\beta}(x^2y)\varepsilon_\alpha(xy)\big)\cdot(\varepsilon_\alpha(xy)\varepsilon_{-\beta}(xy)) \in B_A(4+2) = B_A(6).$$

This implies in particular that the ideal $(x^2 - x | x \in R)$ is contained in $\varepsilon(A, -\beta, 6)$ and hence according to Lemma 3.4.2, the inclusion $I = (x^2 - x | x \in R) \subset \varepsilon(A, \phi, 24)$ holds for all $\phi \in C_2$. Next, we will show that I has finite index in R. There are two possible cases: Either $\mathcal{P}$ has the property $\mathbb{F}[T]/\mathcal{P} = \mathbb{F}_2$ or not. If $\mathbb{F}[T]/\mathcal{P} \neq \mathbb{F}_2$, then there is an element $x \in \mathbb{F}[T]$ such that $x(1-x)$ is not an element of $\mathcal{P} \subset \mathbb{F}[T]$ and hence $x(1-x)$ is invertible in

R. Thus in this case I is the entire ring R. So assume the other case, that is $\mathbb{F}[T]/\mathcal{P} = \mathbb{F}_2$. This immediately implies that $\mathbb{F} = \mathbb{F}_2$ and $\mathcal{P} = T \cdot \mathbb{F}_2[T]$ or $\mathcal{P} = (T-1) \cdot \mathbb{F}_2[T]$. Wlog we assume that $\mathcal{P} = T \cdot \mathbb{F}_2[T]$. But observe, that $T^2 - T = T(T-1)$ is an element of I and hence as $T-1$ is a unit in R, this implies that T is an element of I. This yields that $R/I = \mathbb{F}_2$. So in either case I has finite index in R.

Next, pick a set $X \subset R$ of coset representatives of I in R. According to Proposition 5.2.3, $\mathrm{Sp}_4(R)$ is boundedly generated by root elements, because R is local. So let $K \in \mathbb{N}$ be given such that each element in $\mathrm{Sp}_4(R)$ can be written as a product of at most K root elements. Let $X \in \mathrm{Sp}_4(R)$ be given and choose $a_1, \ldots, a_K \in R$ as well as $\phi_1, \ldots, \phi_K \in C_2$ such that

$$X = \prod_{i=1}^{K} \varepsilon_{\phi_i}(a_i).$$

Each element $a \in R$ can be written as $a = b + x$ for $b \in I$ and $x \in X$. So choose $x_1, \ldots, x_K \in X$ and $b_1, \ldots, b_K \in I$ with $a_i = x_i + b_i$ for all $i = 1, \ldots, K$. Then, we obtain

$$X = \prod_{i=1}^{K} \varepsilon_{\phi_i}(b_i)\varepsilon_{\phi_i}(x_i) = \varepsilon_{\phi_1}(b_1) \cdot [\prod_{i=2}^{K} \varepsilon_{\phi_i}(b_i)^{\varepsilon_{\phi_1}(x_1)\cdots\varepsilon_{\phi_{i-1}}(x_{i-1})}] \cdot [\varepsilon_{\phi_1}(x_1)\cdots\varepsilon_{\phi_K}(x_K)] \quad (5.8)$$

But all the elements

$$\varepsilon_{\phi_1}(b_1), \{\varepsilon_{\phi_i}(b_i)^{\varepsilon_{\phi_1}(x_1)\cdots\varepsilon_{\phi_{i-1}}(x_{i-1})}\}_{2\leq i\leq K}$$

are elements of N and there are only finitely many possibilities for the product

$$\varepsilon_{\phi_1}(x_1)\cdots\varepsilon_{\phi_K}(x_K).$$

Hence N has finite index in $\mathrm{Sp}_4(R)$. On the other hand, if X is in N, then $\varepsilon_{\phi_1}(x_1)\cdots\varepsilon_{\phi_K}(x_K)$ is also an element of N. But there are only finitely many possibilities for $\varepsilon_{\phi_1}(x_1)\cdots\varepsilon_{\phi_K}(x_K)$, so there is an $M \in \mathbb{N}$ such that $\|\varepsilon_{\phi_1}(x_1)\cdots\varepsilon_{\phi_K}(x_K)\|_A \leq M$ holds for all the possible products $\varepsilon_{\phi_1}(x_1)\cdots\varepsilon_{\phi_K}(x_K) \in N$. But we already know that all the elements

$$\varepsilon_{\phi_1}(b_1), \{\varepsilon_{\phi_i}(b_i)^{\varepsilon_{\phi_1}(x_1)\cdots\varepsilon_{\phi_{i-1}}(x_{i-1})}\}_{2\leq i\leq K}$$

are elements of $B_A(18)$. Hence (5.8) implies

$$\|X\|_A = \|\varepsilon_{\phi_1}(b_1) \cdot [\prod_{i=2}^{K} \varepsilon_{\phi_i}(b_i)^{\varepsilon_{\phi_1}(x_1)\cdots\varepsilon_{\phi_{i-1}}(x_{i-1})}] \cdot [\varepsilon_{\phi_1}(x_1)\cdots\varepsilon_{\phi_K}(x_K)]\|_A \leq 18K + M$$

and this finishes the proof. □

Having this proposition, we can now prove Theorem 5.3.2:

Proof. There are two possible cases to consider here. First, the case that $\mathrm{char}(\mathbb{F}) \geq 3$

holds and second that $\mathrm{char}(\mathbb{F}) = 2$. The first case is a direct consequence of Theorem 5.2.4: The ring R is local and $|R/2R| = 1$ holds, because 2 is a unit in R. So let us assume that $\mathrm{char}(\mathbb{F}) = 2$ and let $S \subset \mathrm{Sp}_4(R)$ normally generate $\mathrm{Sp}_4(R)$. To prove this theorem, we will proceed in four steps similar to the proof of Theorem 3.2.5. First, we will show that there is a natural number K independent of R such that

$$\|\varepsilon_{\alpha+\beta}(1)\varepsilon_{2\alpha+\beta}(1)\|_S \leq K|S|$$

holds. Second, we will use the second part of Proposition 5.3.3 to bound the normal subgroup N generated by $\varepsilon_{\alpha+\beta}(1)\varepsilon_{2\alpha+\beta}(1)$ with respect to the norm $\|\cdot\|_S$. Third, we will show that $\mathrm{Sp}_4(R)$ is boundedly generated by root elements. Fourth, we will conclude from the first part of Proposition 5.3.3 and an argument similar to the proof of Theorem 3.2.5 that $\mathrm{Sp}_4(R)$ is strongly bounded. Then last, we will conclude the uniform boundedness of $\mathrm{Sp}_4(R)$ from the fact that R is a local ring.

For the first step, let $X \in \mathrm{Sp}_4(R), x \in R$ and i, j distinct elements of $\{1, \ldots, 4\}$ be given. Then observe that Theorem 3.4.1 and $\mathrm{char}(R) = 2$, implies that

$$\varepsilon_{\alpha+\beta}(y^2x_{i,j}^2)\varepsilon_{2\alpha+\beta}(y^2x_{i,j}^2) \in \langle\langle X \rangle\rangle.$$

Using a first-order compactness argument similar to the one in the proof of Theorem 3.2.1, one then proves the existence of a natural number K such that for all $y \in R$ and $X \in \mathrm{Sp}_4(R)$, one has

$$\|\varepsilon_{\alpha+\beta}(y^2x_{i,j}^2)\varepsilon_{2\alpha+\beta}(y^2x_{i,j}^2)\|_X \leq K$$

for all i, j distinct elements of $\{1, \ldots, 4\}$. Possibly enlarging K, one can also show for all $y \in R$ and $X \in \mathrm{Sp}_4(R)$, that

$$\|\varepsilon_{\alpha+\beta}(y^2(x_{i,i} - x_{j,j})^2)\varepsilon_{2\alpha+\beta}(y^2(x_{i,i} - x_{j,j})^2)\|_X \leq K$$

for all i, j distinct elements of $\{1, \ldots, 4\}$.

Next, observe that S is a normally generating set and hence Lemma 3.0.3 implies $\Pi(S) = \emptyset$. If there were a maximal ideal m containing $\sum_{X \in S} l(X)$, then π_m would map all elements of X to scalar elements in $\mathrm{Sp}_4(R/m)$, which are obviously central in $\mathrm{Sp}_4(R/m)$. But this would imply $m \in \Pi(S) = \emptyset$. This contradiction shows

$$\sum_{X \in S} l(X) = R.$$

Thus there are elements

$$\{y_{ij}^{(X)}, z_{ij}^{(X)}\}_{X \in S, 1 \leq i \neq j \leq 4} \subset R$$

with

$$1 = (\sum_{1\leq i\neq j\leq 4, X\in S} y_{ij}^{(X)} x_{ij}) + (\sum_{1\leq i\neq j\leq 4, X\in S} z_{ij}^{(X)}(x_{i,i} - x_{j,j})).$$

But remember that $\mathrm{char}(R) = 2$ and hence Freshman's dream implies

$$1 = (\sum_{1\leq i\neq j\leq 4, X\in S} (y_{ij}^{(X)})^2 x_{ij}^2) + (\sum_{1\leq i\neq j\leq 4, X\in S} (z_{ij}^{(X)})^2 (x_{i,i} - x_{j,j})^2).$$

Thus, enlarging K, we obtain

$$\begin{aligned} B_S(K|S|)) \ni [&\prod_{1\leq i\neq j, X\in S} \varepsilon_{\alpha+\beta}((y_{ij}^{(X)})^2 x_{ij}^2)\varepsilon_{2\alpha+\beta}((y_{ij}^{(X)})^2 x_{ij}^2)] \\ &\cdot [\prod_{1\leq i\neq j, X\in S} \varepsilon_{\alpha+\beta}((z_{ij}^{(X)})^2 (x_{ii} - x_{jj})^2)\varepsilon_{2\alpha+\beta}((z_{ij}^{(X)})^2 (x_{ii} - x_{jj})^2)] \\ &= \varepsilon_{\alpha+\beta}((\sum_{1\leq i\neq j, X\in S} (y_{ij}^{(X)})^2 x_{ij}^2) + (\sum_{1\leq i\neq j, X\in S} (z_{ij}^{(X)})^2 (x_{ii} - x_{jj})^2)) \\ &\cdot \varepsilon_{2\alpha+\beta}((\sum_{1\leq i\neq j, X\in S} (y_{ij}^{(X)})^2 x_{ij}^2) + (\sum_{1\leq i\neq j, X\in S} (z_{ij}^{(X)})^2 (x_{ii} - x_{jj})^2)) \\ &= \varepsilon_{\alpha+\beta}(1)\varepsilon_{2\alpha+\beta}(1). \end{aligned}$$

This finishes the first step. So we obtain

$$\|\varepsilon_{\alpha+\beta}(1)\varepsilon_{2\alpha+\beta}(1)\|_S \leq K|S|.$$

For the second step, note that the normal subgroup N generated by

$$A := \varepsilon_{\alpha+\beta}(1)\varepsilon_{2\alpha+\beta}(1)$$

is bounded with respect to the norm $\|\cdot\|_A$ according to the second statement of Proposition 5.3.3. Setting $L(R) := \|N\|_A$, this implies

$$\|N\|_S \leq \|N\|_A \cdot \|A\|_S \leq L(R)K|S|.$$

For the third step, remember that R is local and hence the bounded generation of $\mathrm{Sp}_4(R)$ by root elements follows from Proposition 5.2.3.

For the fourth step, observe that according to the first statement of Proposition 5.3.3, the normal subgroup N has finite index in $\mathrm{Sp}_4(R)$ and $\mathrm{Sp}_4(R)$ is boundedly generated by root elements. Hence, if we replace in the proof of Theorem 3.2.5 the subgroup N_{C_2} by N, we can find an $M(R) \in \mathbb{N}$ such that

$$\|\mathrm{Sp}_4(R)\|_S \leq M(R) + \|N\|_S$$

and hence

$$\|\mathrm{Sp}_4(R)\|_S \leq M(R) + L(R)K|S|.$$

Hence $\mathrm{Sp}_4(R)$ is strongly bounded. To finish the proof of uniform boundedness, it suffices to show that there is a natural number $C(R)$ such that each normally generating set S of R contains a normally generating subset S' with $|S'| \leq C(R)$. Note, that a set S normally generates $\mathrm{Sp}_4(R)$ precisely if it satisfies the following two conditions:

1. The group N is contained in the normal subgroup $\langle\langle S\rangle\rangle$ generated by S and
2. the set S maps to a normally generating set of $\mathrm{Sp}_4(R)/N$.

However, the first condition is equivalent to $\varepsilon_{\alpha+\beta}(1)\varepsilon_{2\alpha+\beta}(1)$ being an element of $\langle\langle S\rangle\rangle$. In the first step of the proof, we derived that $\varepsilon_{\alpha+\beta}(1)\varepsilon_{2\alpha+\beta}(1)$ is an element of $\langle\langle S\rangle\rangle$ solely from $\Pi(S) = \emptyset$. But R is local and hence has only one maximal ideal. Hence $\Pi(S) = \emptyset$ can only hold, if there is at least one element $X(S) \in S$ with $\Pi(\{X(S)\}) = \emptyset$. Hence the normal subgroup generated by $X(S)$ already contains $\varepsilon_{\alpha+\beta}(1)\varepsilon_{2\alpha+\beta}(1)$ and hence N.

For the second condition, observe that $\mathrm{Sp}_4(R)/N$ is a finite group according to Proposition 5.3.3. Thus there are only finitely many normally generating sets of $\mathrm{Sp}_4(R)/N$ and consequently there is a natural number $C_1(R)$ such that each normally generating set S of $\mathrm{Sp}_4(R)$ has a finite subset S_1 with at most $C_1(R)$ elements such that $S_1 \cup N$ normally generates $\mathrm{Sp}_4(R)$.

But this implies that already $S' := S_1 \cup \{X(S)\}$ normally generates $\mathrm{Sp}_4(R)$ and that $|S'| \leq C_1(R) + 1$. Hence the proof is finished. □

Remark 5.3.4.

1. The proof is slightly more general than necessary for $\mathrm{char}(\mathbb{F}) = 2$, considering the fact that R is a principal ideal domain and hence one could also use matrix calculations as in Subsection 4.2 to show the first step of the proof of Theorem 5.3.2. However, we wanted to demonstrate that the first step is possible for more general rings of characteristic 2.
2. The ideal I in the proof of Proposition 5.3.3 is commonly called the *booleanizing ideal* of R, often denoted by $\nu_2(R)$, and it can be shown to always have finite index in any global ring of S-algebraic integers R in a global field. Further, the proof of Theorem 5.3.2 shows that the main problem in proving strong boundedness for $\mathrm{Sp}_4(R)$ is not that R might have characteristic 2, but rather whether $\mathrm{Sp}_4(R)$ is boundedly generated by root elements or not.

Chapter 6

Rings of S-algebraic integers and orders

In this chapter, we talk about applications of our results Theorem 3.1.2 and Theorem 3.2.5 to rings of S-algebraic integers. In the first section, we prove strong boundedness for rings R of S-algebraic integers by way of known bounded generation results for $G(\Phi, R)$ for $\Phi \neq C_2$ and G_2. In the second section, we provide explicit upper bounds for $\Delta_k(G)$ in case of $G = \mathrm{Sp}_4(R)$ in a special case and in the third section, we provide similar bounds for $\Delta_k(G_2(R))$. In the fourth and fifth section, we speak about similar results by Morris [30] and how to interpret them in terms of (strong) boundedness.

6.1 Bounded generation results for rings of S-algebraic integers

First, recall the definition of S-algebraic integers:

Definition 6.1.1. [32, Chapter I, §11] Let K be a finite field extension of $\mathbb{Q}$. Then let S be a finite subset of the set V of all valuations of K such that S contains all archimedean valuations. Then the ring $\mathcal{O}_S$ is defined as

$$\mathcal{O}_S := \{a \in K | \ \forall v \in V - S : v(a) \geq 0\}$$

and $\mathcal{O}_S$ is called *the ring of S-algebraic integers in K*. Rings of the form $\mathcal{O}_S$ are called *rings of S-algebraic integers.*

Remember the word norm $\|\cdot\|_{\mathrm{EL}}$ from Definition 2.2.2. S-arithmetic Chevalley groups are boundedly generated by root elements:

Theorem 6.1.2. *[42] Let Φ be an irreducible root system of rank at least 2 and R a ring of S-algebraic integers in a number field K. Then $G(\Phi, R)$ is boundedly generated by root elements. More precisely, let K be a number field and*

$$\Delta := \max\{|\{p| \ p \text{ a prime divisor of } \mathrm{discr}_{K|\mathbb{Q}}\}|, 1\}$$

be given. Then the following inequalities hold:

1. $\|G(\Phi, R)\|_{\mathrm{EL}} \leq (68\Delta + 14)\frac{|\Phi|}{6}$ *for all* Φ *simply-laced,*
2. $\|G(\Phi, R)\|_{\mathrm{EL}} \leq (180\Delta + 27)\frac{|\Phi|}{8}$ *for all* Φ *non-simply-laced not equal to* G_2 *and*
3. $\|G_2(R)\|_{\mathrm{EL}} \leq 68\Delta + 25$.

Furthermore, if R *is a principal ideal domain or* $\Delta = 1$*, then the bounds can be improved to*

1. $\|G(\Phi, R)\|_{\mathrm{EL}} \leq 63\frac{|\Phi|}{6}$ *for all* Φ *simply-laced*
2. $\|G(\Phi, R)\|_{\mathrm{EL}} \leq 159\frac{|\Phi|}{8}$ *for all* Φ *non-simply-laced not equal to* G_2 *and*
3. $\|G_2(R)\|_{\mathrm{EL}} \leq 81$.

Remark 6.1.3. This is not the bounded generation result as found in [42]. Instead it is a summary of the result [42, Corollary 4] and [42, Proposition 1] for the first batch of inequalities. The second batch of inequalities comes from applying possible improvements as appearing in [10] in the principal ideal domain-case and the $\Delta = 1$-case.

Furthermore, all non-zero ideals I in a ring R of S-algebraic integers have finite index. So, rings R of S-algebraic integers in number fields have the property that $G(\Phi, R)$ is boundedly generated by root elements for all irreducible Φ of rank at least 2 and the ideal $2R$ (and all other non-zero ideals) have finite index in R. Hence Theorem 3.1.2 and Theorem 3.2.5 can be applied to the groups $G(\Phi, R)$. This gives us the following Theorem:

Theorem 6.1.4. *Let* R *be a ring of S-algebraic integers in a number field and* Φ *an irreducible root system of rank at least* 2*. Then there is a constant* $C(\Phi, R) \geq 1$ *such that*

$$\Delta_k(G(\Phi, R)) \leq C(\Phi, R)k$$

holds for all $k \in \mathbb{N}$.

Furthermore, we can give some explicit bounds for strong boundedness as well. To this end note the following bounded generation result for $\mathrm{SL}_2(R)$ by Rapinchuk, Morgan and Sury:

Theorem 6.1.5. *[29, Theorem 1.1] Let* R *be a ring of S-algebraic integers with infinitely many units. Then* $\|\mathrm{SL}_2(R)\|_{\mathrm{EL}} \leq 9$.

Using this result and Theorem 6.1.2, we can provide some explicit values:

Corollary 6.1.6. *Let R be a ring of S-algebraic integers with class number one and $n \geq 3$. Further set*

$$\Delta(R) := \begin{cases} 135, \text{ if } R \text{ is a quadratic imaginary ring of integers or } \mathbb{Z} \\ 12, \text{ if } R \text{ is neither of the above} \end{cases}$$

Then $\Delta_k(\mathrm{Sp}_{2n}(R)) \leq 192(1+5n)(12n+\Delta(R))k$ holds for all $k \in \mathbb{N}$.

Proof. Let $k \in \mathbb{N}$ be given. Then according to Theorem 3.1.2, the inequality

$$\Delta_k(\mathrm{Sp}_{2n}(R)) \leq 3Q(C_n, R) \cdot L(C_n)k \tag{6.1}$$

holds for $L(C_n)$ given as in Theorem 3.1.1 and $\|\mathrm{Sp}_{2n}(R)\|_{\mathrm{EL}_Q} \leq Q(C_n, R)$. However, according to Theorem 4.1.3, we can choose $L(C_n) \leq 64(1+5n)$.

Next, we give upper bounds on $\|\mathrm{Sp}_4(R)\|_{\mathrm{EL}_Q}$ depending on R. First, if R is a quadratic imaginary ring of integers or $\mathbb{Z}$, we have

$$\|\mathrm{Sp}_4(R)\|_{\mathrm{EL}_Q} \leq \|\mathrm{Sp}_4(R)\|_{EL} \leq 159.$$

according to Theorem 6.1.2.

On the other hand, if R is not a ring of quadratic imaginary integers or $\mathbb{Z}$, then R has infinitely many units according to [32, Corollary 11.7]. This implies $\|\mathrm{SL}_2(R)\|_{EL} \leq 9$ for those rings by Theorem 6.1.5. According to Proposition 5.1.8, this implies

$$\mathrm{Sp}_4(R) = (U^+(C_2,R)U^-(C_2,R))^4U^+(C_2,R) \text{ or } \mathrm{Sp}_4(R) = U^-(C_2,R)(U^+(C_2,R)U^-(C_2,R))^4.$$

But C_2 has four positive roots and hence

$$\|\mathrm{Sp}_4(R)\|_{\mathrm{EL}_Q} \leq \|\mathrm{Sp}_4(R)\|_{EL} \leq 4 * 9 = 36.$$

holds. Hence setting

$$\Delta'(R) := \begin{cases} 159, \text{ if } R \text{ is a quadratic imaginary ring of integers or } \mathbb{Z} \\ 36, \text{ if } R \text{ is neither of the above} \end{cases}$$

implies $\|\mathrm{Sp}_4(R)\|_{\mathrm{EL}_Q} \leq \Delta'(R)$ for all rings of S-algebraic integers with class number 1. Proposition 5.1.10 then implies

$$\|\mathrm{Sp}_{2n}(R)\|_{\mathrm{EL}_Q} \leq 12(n-2) + \|\mathrm{Sp}_4(R)\|_{\mathrm{EL}_Q} \leq 12(n-2) + \Delta'(R).$$

But this together with (6.1) and $L(C_n) \leq 64(1+5n)$ implies

$$\Delta_k(\mathrm{Sp}_{2n}(R)) \leq 3Q(C_n, R) \cdot L(C_n)k \leq (12n - 24 + \Delta'(R)) \cdot 192(5n+1)k$$

and finishes the proof. □

Further, we can show the following:

Theorem 6.1.7. *Let R be a ring of S-algebraic integers with class number one. Further set*

$$\Delta(R) := \begin{cases} 154, \text{ if } R \text{ is a quadratic imaginary ring of integers or } \mathbb{Z} \\ 117, \text{ if } R \text{ is neither of the above} \end{cases}$$

Then $\Delta_k(E_6(R)) \leq 120 \cdot 60^{2^{11}} \Delta(R)k$ holds for all $k \in \mathbb{N}$.

Proof. Let $k \in \mathbb{N}$ be given. Then according to Theorem 3.1.2, the inequality

$$\Delta_k(E_6(R)) \leq Q(E_6, R) \cdot L(E_6)k \tag{6.2}$$

holds for $L(E_6)$ given as in Theorem 3.1.1 and $\|E_6(R)\|_{\mathrm{EL}_Q} \leq Q(E_6, R)$. However, according to Proposition 4.4.7, we can choose $L(E_6) \leq 120 \cdot 60^{2^{11}}$.

Next, we give upper bounds on $\|E_6(R)\|_{\mathrm{EL}_Q}$ depending on R. To this end note that Proposition 5.1.7 implies

$$E_6(R) = (U^+(E_6, R)U^-(E_6, R))^2 G_\epsilon(R).$$

Thus for each $A \in E_6(R)$, there are $u_1^+, u_2^+ \in U^+(E_6, R)$ and $u_1^-, u_2^- \in U^-(E_6, R)$ as well as $Z \in G_\epsilon(R)$ such that $A = u_1^+ u_1^- u_2^+ u_2^- Z$. This implies

$$A = u_1^+ u_1^- u_2^+ u_2^- Z = \left(u_1^+ u_1^- (u_1^+)^{-1}\right) \cdot (u_1^+ u_2^+) u_2^- Z = (u_1^-)^{u_1^+} \cdot (u_1^+ u_2^+) u_2^- Z$$

and hence

$$\begin{aligned} \|A\|_{\mathrm{EL}_Q} &= \|(u_1^-)^{u_1^+} \cdot (u_1^+ u_2^+) u_2^- Z\|_{\mathrm{EL}_Q} \\ &\leq \|(u_1^-)^{u_1^+}\|_{\mathrm{EL}_Q} + \|u_1^+ u_2^+\|_{\mathrm{EL}_Q} + \|u_2^-\|_{\mathrm{EL}_Q} + \|Z\|_{\mathrm{EL}_Q} \\ &= \|u_1^-\|_{\mathrm{EL}_Q} + \|u_1^+ u_2^+\|_{\mathrm{EL}_Q} + \|u_2^-\|_{\mathrm{EL}_Q} + \|Z\|_{\mathrm{EL}_Q}. \end{aligned}$$

But $U^+(E_6, R)$ and $U^-(E_6, R)$ are conjugate to one another and A was arbitrary, so this implies

$$\|E_6(R)\|_{\mathrm{EL}_Q} \leq 3\|U^+(E_6, R)\|_{\mathrm{EL}_Q} + \|G_\epsilon(R)\|_{\mathrm{EL}_Q}.$$

The root system E_6 has 36 positive roots as can be for example seen in the proof of Lemma 4.4.4 in Appendix C and hence $\|U^+(E_6, R)\|_{\mathrm{EL}_Q} \leq 36$ holds. On the other hand,

$G_\epsilon(R)$ is a subgroup of the group

$$H := \langle G_\epsilon(R), G_\delta(R) \rangle$$

The root subsystem of E_6 spanned by ϵ and δ is isomorphic to A_2 and $\mathrm{SL}_3(R)$ is generated by root elements according to [42, Lemma 4]. Thus one can use Lemma 2.2.4 and [41, Chapter 8, p. 68, Lemma 49] to show that there is an epimorphism of $\mathrm{SL}_3(R)$ onto H with the property

$$E_{12}(x) \mapsto \varepsilon_\epsilon(x), E_{21}(x) \mapsto \varepsilon_{-\epsilon}(x)$$
$$E_{23}(x) \mapsto \varepsilon_\delta(x), E_{32}(x) \mapsto \varepsilon_{-\delta}(x)$$

for all $x \in R$. Using this epimorphism, the subgroup

$$\left\{ \begin{pmatrix} A & \\ & 1 \end{pmatrix} | A \in \mathrm{SL}_2(R) \right\}$$

of $\mathrm{SL}_3(R)$ maps onto $G_\epsilon(R)$. This is to say, that one can give an upper bound on $\|G_\epsilon(R)\|_{\mathrm{EL}_Q}$ by way of giving an upper bound on $\|\mathrm{SL}_2(R)\|_{\mathrm{EL}_Q}$ when considering $\mathrm{SL}_2(R)$ as a subgroup of $\mathrm{SL}_3(R)$. Now, we distinguish two cases. First, if R is a quadratic imaginary ring of integers or $\mathbb{Z}$, then one can see reading the proof of [10, Main Theorem] that

$$\|\mathrm{SL}_2(R)\|_{\mathrm{EL}_Q} \leq \|\mathrm{SL}_2(R)\|_{\mathrm{EL}} \leq 56.$$

On the other hand, if R is not a ring of quadratic imaginary integers or $\mathbb{Z}$, then R has infinitely many units according to [32, Corollary 11.7]. This implies $\|\mathrm{SL}_2(R)\|_{\mathrm{EL}_Q} \leq 9$ for those rings by Theorem 6.1.5. Thus setting

$$\Delta'(R) := \begin{cases} 56, \text{ if } R \text{ is a quadratic imaginary ring of integers or } \mathbb{Z} \\ 9, \text{ if } R \text{ is neither of the above} \end{cases}$$

implies $\|G_\epsilon(R)\|_{\mathrm{EL}_Q} \leq \Delta'(R)$ and hence

$$\|E_6(R)\|_{\mathrm{EL}_Q} \leq 3 * 36 + \Delta'(R) = \Delta(R)$$

holds. This together with (6.2) and $L(E_6) \leq 120 \cdot 60^{211}$ finishes the proof. □

Remark 6.1.8. It is obviously possible to give upper bounds on $\|U^+(E_6, R)\|_{\mathrm{EL}_Q}$, that are better than the naive $\|U^+(E_6, R)\|_{\mathrm{EL}_Q} \leq 36$. For example, one could consider the roots that do not involve the simple root ϕ and the ones that do separately. Then one uses an argument similar to the proof of Proposition 5.1.10 to give better upper bounds on the terms not involving ϕ. However, as the explicit bounds for the E_6-case are quite bad

anyway, we decided against it.

Invoking Theorem 6.1.5 also allows a slight improvement of the upper bound in the older result in [24, Corollary 6.2]:

Corollary 6.1.9. *Let R be a ring of S-algebraic integers with infinitely many units and class number one, $n \geq 3$ and $k \in \mathbb{N}$. Then*

$$\Delta_k(\mathrm{SL}_n(R)) \leq (4n+1)(4n+4)k$$

holds.

Remark 6.1.10. The only rings R of S-algebraic integers that have finitely many units and are principal ideal domains are $\mathbb{Z}$ and the rings of algebraic integers in the quadratic number fields $\mathbb{Q}[\sqrt{D}]$ for $D = -1, -2, -3, -7, -11, -19, -43, -67, -163$. This is a famous result by Heegner.

6.2 Explicit bounds for Sp_4

Let R be a commutative ring with 1 such that $(R : 2R)$ is finite. In this section G denotes the group $\mathrm{Sp}_4(R)$. Recall, that the set of positive roots in C_2 is $\alpha, \beta, \alpha+\beta$ and $2\alpha+\beta$ with α short and simple and β long and simple. Further, recall the set:

$$Q_{C_2} := \{A\varepsilon_\phi(2x)A^{-1} | x \in R, \phi \in C_2, A \in \mathrm{Sp}_4(R)\}$$

as well as the group $N_{C_2} := \langle Q_{C_2} \rangle$. Further, let $\pi : G \to G/N_{C_2}$ be the quotient map and let k be a natural number. Then recall that Theorem 3.2.5 implies:

$$\Delta_k(\mathrm{Sp}_4(R)) \leq L(C_2)K(C_2, R)k + \Delta_\infty(G/N_{C_2})$$

where

1. the constant $L(C_2)$ is given as in Theorem 4.2.1 or Theorem 3.2.1,
2. the constant $K(C_2, R)$ is defined to be $\|N_{C_2}\|_{Q_{C_2}}$

For principal ideal domains, an upper bound on the constant $L(C_2)$ is already known to us by Theorem 4.2.1. Namely, $L(C_2) \leq 384$ holds. So to give explicit upper bounds on $\Delta_k(\mathrm{Sp}_4(R))$, one must give upper bounds on $\Delta_\infty(G/N_{C_2})$ and $K(C_2, R)$. Determining $\Delta_\infty(G/N_{C_2})$ is relatively easy, because we only have to determine the maximal possible diameter of a conjugation generated word norm on some finite group of Lie-type or direct products of such. On the other hand, giving an upper bound on $K(C_2, R)$, is harder and we will in fact only do it in a special case.

6.2.1 Boundedness of the 2-congruence subgroup of $\mathrm{Sp}_4(R)$

Definition 6.2.1. Let R be a commutative ring with 1 such that the set of coset representatives X of $2R$ in R can be chosen in such a way that each $x \in X - 2R$ is a unit in $R, 0 \in X$ and if $R \neq 2R$, then $1 \in X$. Then R is called a $2R$-*pseudo-good* ring.

Remark 6.2.2. If R is $2R$-pseudo-good, then either $R/2R$ is a field or 2 is a unit in R. This is the case, because each element $\bar{x}$ in $R/2R - \{0\}$ can be written as $\bar{x} = x + 2R$ for some $x \in X$ a unit. But then $\bar{x}$ is itself a unit and hence each non-zero element of $R/2R$ is a unit and so R is a field. On the other hand, if $R/2R$ does not have non-zero elements, than this implies that 1 is an element of $2R$. We should also mention that $2R$-pseudo-goodness is our own concept named so as an homage to good rings.

We note the following characterization for rings of S-algebraic integers in quadratic number fields that are $2R$-pseudo-good:

Proposition 6.2.3. *Let D be a square-free integer, R' the ring of algebraic integers in the number field $\mathbb{Q}[\sqrt{D}]$ and S a finite set of non zero prime ideals in R'. Define*

$$R := \{a/b |\ a \in R', b \in R' - \{0\}, \{\text{ prime divisors of } bR'\} \subset S\}.$$

Then R is $2R-$pseudo-good if and only if at least one of the following conditions hold

1. *The set S contains a prime-divisor of $2R'$ or*
2. *$D \equiv 5 \bmod 8$ and $D > 0$ or*
3. *$D \equiv 5 \bmod 8$ and $S \neq \emptyset$ or*
4. *$D = -3$.*

Remark 6.2.4. After handing in this thesis and the thesis being approved by the examiners in the viva, I found a gap in the proof of this proposition and a counterexample to its claim. In a future paper I salvage part of this proposition though and give another class of examples of $2R$-pseudo-good rings.

Further, define for a $2R$-pseudo-good ring R with the corresponding set of coset representatives X, the set

$$B_R := \{\varepsilon_{2\alpha+\beta}(x_1)\varepsilon_{\alpha+\beta}(x_2)\varepsilon_\beta(x_3)\varepsilon_\alpha(x_4)h_\alpha(t)h_\beta(s) |\ t, s \in X \cap R^*, x_1, x_2, x_3, x_4 \in X\}.$$

Also recall that the Weyl group $W(C_2)$ is generated by the set $F := \{w_\alpha, w_\beta\}$. To determine $K(C_2, R)$, we prove the following proposition.

Proposition 6.2.5. *Let R be a $2R$-pseudo-good ring, $w_1 = s_1^{(1)} \cdots s_{k_1}^{(1)}$ and $w_2 = s_1^{(2)} \cdots s_{k_2}^{(2)}$ elements of $W(C_2)$ with $s_1^{(1)}, \ldots, s_{k_1}^{(1)}, s_1^{(2)}, \ldots, s_{k_2}^{(2)}$ elements of F and $l_F(w_1) = k_1$ and*

$l_F(w_2) = k_2$. *Then up to multiplication by* $l_F(w_2)$ *elements of* Q_{C_2}, *each element of* $(B(C_2,R)w_1B(C_2,R)) \cdot (B(C_2,R)w_2B(C_2,R))$ *is an element of* $B(C_2,R)wB(C_2,R)$ *for* w *some subword of the (possibly non-minimal) expression* $(s_1^{(1)},\ldots,s_{k_1}^{(1)},s_1^{(2)},\ldots,s_{k_2}^{(2)})$.

This proposition gives an upper bound on $K(C_2,R)$ as follows in case 2 is not a unit in R. Note that according to bounded generation by root elements, we can by grouping elements of $U^+(C_2,R)$ and $U^-(C_2,R)$ together and potentially conjugate by $w_0 := (w_\alpha w_\beta)^2$, the longest element in $W(C_2)$, find a $J \in \mathbb{N}$, such that each $A \in \mathrm{Sp}_4(R)$ can be written as

$$A = \prod_{i=1}^{J} u_i^+ u_i^-$$

for all u_i^+ elements of $U^+(C_2,R)$ and all u_i^- elements of $U^-(C_2,R)$. But each element $w_0^{-1}u_i^-w_0$ is a product of root elements of positive roots in C_2 and hence an element of $B(C_2,R)$. But then

$$u_i^+u_i^- = (u_i^+w_0)(w_0^{-1}u_i^-w_0)w_0^{-1} \in (B(C_2,R)w_0)\cdot(B(C_2,R)w_0) \subset (B(C_2,R)w_0B(C_2,R))^2.$$

holds for all i. This implies $A \in (B(C_2,R)w_0B(C_2,R))^{2J}$.

But $l_F(w_0) = 4$ holds, so according to Proposition 6.2.5, the matrix A can be written as a product $b_1'wb_2'$ for $b \in B(C_2,R)$ and $w \in W(C_2)$ after multiplication by $(2J-1)l_F(w_0) \leq 4(2J-1)$ elements of Q_{C_2}. But each element of $B(C_2,R)wB(C_2,R)$ is conjugate to an element of $B(C_2,R)w$. Observe that each element A of $B(C_2,R)$ has the form

$$\varepsilon_{2\alpha+\beta}(t_{2\alpha+\beta})\varepsilon_{\alpha+\beta}(t_{\alpha+\beta})\varepsilon_\beta(t_\beta)\varepsilon_\alpha(t_\alpha)h_\alpha(s_\alpha)h_\beta(s_\beta)$$

for $t_{2\alpha+\beta},t_{\alpha+\beta},t_\beta,t_\alpha \in R$ and $s_\alpha,s_\beta \in R^*$. Hence after multiplication with 4 elements of Q_{C_2}, we may assume that $t_{2\alpha+\beta},t_{\alpha+\beta},t_\beta,t_\alpha \in R$ are elements of the set X of coset representatives of $2R$ in R instead. Furthermore,

$$h_\alpha(s_\alpha) = w_\alpha(s_\alpha)w_\alpha^{-1}$$

holds and $w_\alpha(s_\alpha) = \varepsilon_\alpha(s_\alpha)\varepsilon_{-\alpha}(-s_\alpha^{-1})\varepsilon_\alpha(s_\alpha)$. Note, that all elements of $X - \{0\}$ are units in R and so we can consider the set

$$Y := \{-x^{-1} | x \in X - \{0\}\} \cup \{0\}.$$

One easily checks that this set Y is also a set of coset representatives of $2R$ in R. Thus after multiplication with 3 elements of Q_{C_2}, we may assume that s_α is an element of $X - \{0\}$. Similarly, we may assume after multiplication by 3 elements of Q_{C_2} that s_β is an element of $X - \{0\}$. So each element of $B(C_2,R)w$ agrees with an element of B_Rw

after multiplication by $4+3+3=10$ elements of Q_{C_2}.

To summarize: Up to multiplication by up to $4(2J-1)+10=8J+6$ elements of Q_{C_2}, each element A of $\mathrm{Sp}_4(R)$ can be rewritten as an element of $B_R w$ for some $w \in W(C_2)$. Next, observe that N_{C_2} is contained in $\ker(\pi_{2R} : \mathrm{Sp}_4(R) \to \mathrm{Sp}_4(R/2R)) := H_{C_2}$. We are going to show that $B_R w \cap H_{C_2} \neq \emptyset$ implies $w = I_4$ and $B_R w \cap H_{C_2} = \{I_4\}$. Together, this implies

$$K(C_2, R) = \|N_{C_2}\|_{Q_{C_2}} \leq 8J+6.$$

To show that $B_R w \cap H_{C_2} \neq \emptyset$ implies $w = I_4$ and $B_R w \cap H_{C_2} = \{I_4\}$, assume there is an $A = bw \in B_R w \cap N_{C_2}$ for some $w \in W(C_2)$. Observe that $\pi_{2R}(A) = I_4$. But $\pi_{2R}(b)$ is an element of $B(R/2R, C_2)$ of $\mathrm{Sp}_4(R/2R)$. Further, slightly abusing notation, we obtain $\pi_{2R}(w) = w$ and hence $\pi_{2R}(A)$ is an element of $B(R/2R, C_2)w$. But 2 is assumed to not be a unit in R and so the ring $R/2R$ is a field. Hence by the uniqueness of the Bruhat-decomposition for $\mathrm{Sp}_4(R/2R)$ [41, Chapter 3, p. 26, Theorem 4'], we obtain $\pi_{2R}(b) = w = I_4$. But according to the definition of B_R and remembering that X is a set of coset-representatives of $2R$ in R, this implies $b \in \{h_\alpha(t)h_\beta(s) |\ t, s \in X \cap R^*\}$. So there are $t, s \in X \cap R^*$ with

$$A = h_\alpha(t)h_\beta(s) = \left(\begin{array}{cc|cc} t & 0 & 0 & 0 \\ 0 & st^{-1} & 0 & 0 \\ \hline 0 & 0 & t^{-1} & 0 \\ 0 & 0 & 0 & s^{-1}t \end{array}\right)$$

But $\pi_{2R}(A) = I_4$ and hence $t \equiv 1 \bmod 2R$. But $1 \in X$ and so $t = 1$. Then $s = 1$ follows the same way. Hence $A = I_4$. This implies:

Proposition 6.2.6. *Let R be a $2R$-pseudo-good ring such that 2 is not a unit and let $J \in \mathbb{N}$ be given such that each $A \in \mathrm{Sp}_4(R)$ can be written as an element of $(U^+(C_2, R)U^-(C_2, R))^J$ or $(U^-(C_2, R)U^+(C_2, R))^J$. Then $K(C_2, R) \leq 8J+6$ holds.*

Further, the proof implies the following:

Corollary 6.2.7. *Let R be a $2R$-pseudo-good ring of S-algebraic integers with $R \neq 2R$. Then $N_{C_2} = \ker(\pi_{2R} : \mathrm{Sp}_4(R) \to \mathrm{Sp}_4(R/2R))$ holds.*

Remark 6.2.8. Milnor's, Serre's and Bass' solution for the Congruence subgroup problem [6, Theorem 3.6, Corollary 12.5] yields

$$N_{C_2} = \ker\left(\pi_{2R} : \mathrm{Sp}_4(R) \to \mathrm{Sp}_4(R/2R)\right).$$

more generally for all rings of S-algebraic integers.

To prove Proposition 6.2.5, we need:

Lemma 6.2.9. *Let R be a $2R$-pseudo-good ring. Then up to multiplication by an element of Q_{C_2}, we have*

$$(B(C_2,R)w_\alpha B(C_2,R))\cdot(B(C_2,R)w_\alpha B(C_2,R))\subset B(C_2,R)\cup(B(C_2,R)w_\alpha B(C_2,R))\,.$$

The same holds for β instead of α.

Proof. Let $b_1,b_2,b_1',b_2'\in B(C_2,R)$ be given. Note that we may write b_2b_1' as

$$b_2b_1'=\varepsilon_\alpha(a)u_{P-\alpha}h$$

for $a\in R$, $\varepsilon_{2\alpha+\beta}(b)\varepsilon_{\alpha+\beta}(c)\varepsilon_\beta(d)=u_{P-\{\alpha\}}$ for $b,c,d\in R$ and $h\in\{h_\alpha(t)h_\beta(s)|\ t,s\in R^*\}$. This implies:

$$b_1w_\alpha b_2b_1'w_\alpha b_2'=b_1w_\alpha\varepsilon_\alpha(a)u_{P-\alpha}hw_\alpha b_2'=b_1\varepsilon_{-\alpha}(\pm a)w_\alpha[u_{P-\alpha}(-h)]w_\alpha^{-1}b_2'.$$

Next, $w_\alpha[u_{P-\alpha}(-h)]w_\alpha^{-1}$ is an element of $B(C_2,R)$, because $w_\alpha u_{P-\alpha}w_\alpha^{-1}$ is a product of root elements associated to positive roots in C_2 and $w_\alpha(-h)w_\alpha^{-1}$ is an element of $\{h_\alpha(t)h_\beta(s)|\ t,s\in R^*\}$ as required. Thus $b_1w_\alpha b_2b_1'w_\alpha b_2'\in B(C_2,R)\varepsilon_{-\alpha}(\pm a)B(C_2,R)$ holds.

There are two possible cases now. Either a is an element of $2R$, then we are done after multiplying with one element of Q_{C_2}. On the other hand, if $a\notin 2R$ holds, then as R is $2R$-pseudo-good, there is a unit $x\in R$ such that $a\equiv -x^{-1}\bmod 2R$. Hence after multiplying with one element of Q_{C_2}, we may assume $a=-x^{-1}$ and so we obtain

$$\begin{aligned}\varepsilon_{-\alpha}(a)&=\varepsilon_\alpha(-x)(\varepsilon_\alpha(x)\varepsilon_{-\alpha}(-x^{-1})\varepsilon_\alpha(x))\varepsilon_\alpha(-x)\\&=\varepsilon_\alpha(-x)w_\alpha(x)\varepsilon_\alpha(-x)=\varepsilon_\alpha(-x)h_\alpha(x)w_\alpha\varepsilon_\alpha(-x).\end{aligned}$$

But $\varepsilon_\alpha(-x)h_\alpha(x)$ and $\varepsilon_\alpha(-x)$ are elements of $B(C_2,R)$, so $\varepsilon_{-\alpha}(a)$ is an element of $B(C_2,R)w_\alpha B(C_2,R)$. Hence

$$\begin{aligned}b_1w_\alpha b_2b_1'w_\alpha b_2'&\in B(C_2,R)\varepsilon_{-\alpha}(\pm a)B(C_2,R)\subset B(C_2,R)\cdot(B(C_2,R)w_\alpha B(C_2,R))\cdot B(C_2,R)\\&=B(C_2,R)w_\alpha B(C_2,R)\end{aligned}$$

holds after multiplication with up to one element of Q_{C_2}. □

Next, we are going to prove the Proposition 6.2.5:

Proof. Slightly abusing notation we set $T(w_1,w_2):=(s_1^{(1)},\dots,s_{k_1}^{(1)},s_1^{(2)},\dots,s_{k_2}^{(2)})$. We will

show first that

$$(B(C_2,R)w_1B(C_2,R))\cdot(B(C_2,R)w_2B(C_2,R))\subset \bigcup_{w \text{ subword of } T(w_1,w_2)} B(C_2,R)wB(C_2,R)$$

holds up to multiplication by $l_F(w_2)$ elements of Q_{C_2} by induction on $l_F(w_2)$.

For $l_F(w_2)=0$, we obtain $B(C_2,R)w_2B(C_2,R)=B(C_2,R)$ and hence the claim is obvious. So let $w_2\in W(C_2)$ be given and assume by induction that the claim holds for each subword of $(s_1^{(2)},\dots,s_{k_2}^{(2)})$. Further, assume that without loss of generality $w_2=w_2'w_\alpha$ and $l_F(w_2)=l_F(w_2')+1$. Then by induction hypothesis

$$\begin{aligned}
&(B(C_2,R)w_1B(C_2,R))\cdot(B(C_2,R)w_2B(C_2,R))\\
&=(B(C_2,R)w_1B(C_2,R))\cdot(B(C_2,R)w_2')\cdot(w_\alpha B(C_2,R))\\
&\subset\left[\bigcup_{w \text{ subword of } T(w_1,w_2')} B(C_2,R)wB(C_2,R)\right]\cdot w_\alpha B(C_2,R)\\
&=\bigcup_{w \text{ subword of } T(w_1,w_2')}(B(C_2,R)wB(C_2,R)\cdot w_\alpha B(C_2,R))
\end{aligned}$$

holds up to multiplication by $l_F(w_2')$ elements of Q_{C_2}. Hence it suffices to show the claim in the special case $w_2=w_\alpha$. We distinguish two cases: First $l_F(w_1w_\alpha)>l_F(w_1)$ and second $l_F(w_1w_\alpha)<l_F(w_1)$.

In the first case, it suffices to show that $w_1B(C_2,R)w_\alpha\subset B(C_2,R)w_1w_\alpha B(C_2,R)$. To see this let

$$b=\varepsilon_\alpha(a)u_{P-\{\alpha\}}h\in B(C_2,R)$$

be given with $a\in R$, $\varepsilon_{2\alpha+\beta}(b)\varepsilon_{\alpha+\beta}(c)\varepsilon_\beta(d)=u_{P-\{\alpha\}}$ for $b,c,d\in R$ and $h\in\{h_\alpha(t)h_\beta(s)|\,t,s\in R^*\}$. Note that

$$w_1\varepsilon_\alpha(a)w_1^{-1}=\varepsilon_{w_1(\alpha)}(\pm a).$$

Yet according to [41, Appendix, p. 151, (19)Lemma], the inequality $l_F(w_1w_\alpha)>l_F(w_1)$ implies that the root $w_1(\alpha)$ is positive root. Thus $w_1\varepsilon_\alpha(a)w_1^{-1}\in B(C_2,R)$. On the other hand, similar to the proof of the previous lemma, $w_\alpha^{-1}u_{P-\{\alpha\}}hw_\alpha$ is also an element of $B(C_2,R)$. Hence we obtain

$$w_1bw_\alpha=w_1\varepsilon_\alpha(a)u_{P-\{\alpha\}}hw_\alpha=(w_1\varepsilon_\alpha(a)w_1^{-1})w_1w_\alpha(w_\alpha^{-1}u_{P-\{\alpha\}}hw_\alpha)\in B(C_2,R)w_1w_\alpha B(C_2,R).$$

This finishes the proof of the first case. Note in particular that in the first case we need not multiply by an element of Q_{C_2}.

In the second case, we can write $w_1 = w_1' w_\alpha$ for $l_F(w_1) = l_F(w_1') + 1$ and so

$$\begin{aligned}&(B(C_2,R)w_1B(C_2,R)) \cdot (B(C_2,R)w_\alpha B(C_2,R))\\ &= (B(C_2,R)w_1') \cdot (w_\alpha B(C_2,R)) \cdot (B(C_2,R)w_\alpha B(C_2,R)).\end{aligned}$$

But according to Lemma 6.2.9, we know that up to multiplication by an element of Q_{C_2}, we have

$$(w_\alpha B(C_2,R)) \cdot (B(C_2,R)w_\alpha B(C_2,R)) \subset B(C_2,R) \cup (B(C_2,R)w_\alpha B(C_2,R)).$$

Thus up to multiplication by an element of Q_{C_2}, we have

$$\begin{aligned}&(B(C_2,R)w_1B(C_2,R)) \cdot (B(C_2,R)w_\alpha B(C_2,R))\\ &\subset (B(C_2,R)w_1') \cdot [B(C_2,R) \cup (B(C_2,R)w_\alpha B(C_2,R))]\\ &= (B(C_2,R)w_1'B(C_2,R)) \cup (B(C_2,R)w_1'B(C_2,R)w_\alpha B(C_2,R)).\end{aligned}$$

But according to the first case $B(C_2,R)w_1'B(C_2,R)w_\alpha B(C_2,R) \subset B(C_2,R)w_1'w_\alpha B(C_2,R)$. This finishes the second case and the proof of the proposition. □

6.2.2 Conjugation generated word norms on $\mathrm{Sp}_4(R/2R)$

To determine $\Delta_\infty(G/N_{C_2})$, we will first prove:

Lemma 6.2.10. *Let K be a field of characteristic 2. Then each subset S that normally generates $\mathrm{Sp}_4(K)$ contains an element $A \in S$ such that A alone normally generates $\mathrm{Sp}_4(K)$. This implies $\Delta_\infty(\mathrm{Sp}_4(K)) = \Delta_1(\mathrm{Sp}_4(K))$.*

Proof. Observe that $\mathrm{Sp}_4(K) = \mathrm{PSp}_4(K)$, because each scalar matrix in $\mathrm{Sp}_4(K)$ must be I_4 as $\mathrm{char}(K) = 2$. Next, assume that $K \neq \mathbb{F}_2$ and pick a non-trivial element A in the normally generating subset S of $\mathrm{Sp}_4(K)$. But as $K \neq \mathbb{F}_2$, the group $\mathrm{Sp}_4(K) = \mathrm{PSp}_4(K)$ is simple by [41, Chapter 4, p. 33, Theorem 5] and hence A normally generates $\mathrm{Sp}_4(K)$.

If $K = \mathbb{F}_2$, then $\mathrm{Sp}_4(K)$ is isomorphic to the permutation group S_6 according to Proposition B.0.1. However S_6 only has three normal subgroups, namely S_6, A_6 and the trivial subgroup. So for a normally generating set S of $\mathrm{Sp}_4(K)$ pick an element $A \in S$, not lying in A_6. Then clearly the normal subgroup generated by A must be the entire group $S_6 = \mathrm{Sp}_4(K)$.

So for each normally generating set S of $\mathrm{Sp}_4(K)$ there is an $A_S \in S$ that normally

generates $\mathrm{Sp}_4(K)$. This implies:

$$\begin{aligned}\Delta_\infty(\mathrm{Sp}_4(K)) \geq \Delta_1(\mathrm{Sp}_4(K)) &\geq \sup\{\|\mathrm{Sp}_4(K)\|_{A_S} |\ S \text{ normally generates } \mathrm{Sp}_4(K)\} \\ &\geq \sup\{\|\mathrm{Sp}_4(K)\|_S |\ S \text{ normally generates } \mathrm{Sp}_4(K)\} \\ &= \Delta_\infty(\mathrm{Sp}_4(K)).\end{aligned}$$

□

However $\Delta_1(G)$ is an invariant of a group closely related to the classical notion of *covering numbers of finite groups*. The invariant

$$\mathrm{cn}(G) := \min\{n \in \mathbb{N} | \forall \text{ conjugacy classes } C \text{ of } G : C^n = G\}.$$

is the covering number of the group G. Now clearly, $\Delta_1(G) \leq \mathrm{cn}(G)$ holds. In most cases we are interested in, one actually has equality between the two numbers.

Proposition 6.2.11. *Let K be a field of characteristic 2. Then*

1. $\Delta_\infty(\mathrm{Sp}_4(\mathbb{F}_2)) = 5$,
2. $\Delta_\infty(\mathrm{Sp}_4(\mathbb{F}_4)) = 4$ *and*
3. $\Delta_\infty(\mathrm{Sp}_4(K)) \leq 104$ *holds for* $|K| \geq 8$.

Proof. As mentioned in the proof of Lemma 6.2.10 the group $\mathrm{Sp}_4(\mathbb{F}_2)$ is isomorphic to S_6. However for S_6 the covering number can be determined to be 5 from the main result in [7]. Further, Proposition B.0.3 implies that the conjugation generated word metric induced by the transposition (12) in S_6 has diameter 5. Hence $\Delta_\infty(\mathrm{Sp}_4(\mathbb{F}_2)) = 5$. This proves the first claim of the proposition. The paper [23] contains a list of covering numbers calculated using a computer algebra system and states on page 61 that $\mathrm{cn}(\mathrm{Sp}_4(\mathbb{F}_4)) = 4$. This yields $\Delta_\infty(\mathrm{Sp}_4(\mathbb{F}_4)) \leq 4$. The lower bound $\Delta_\infty(\mathrm{Sp}_4(K)) \geq 4$ is a consequence of Proposition 7.1.3. This proves the second claim of the proposition. The third and last statement is a consequence of Liebeck's and Lawther's [25, Theorem 1], which implies in our terminology for q a power of 2, that $\Delta_\infty(\mathrm{Sp}_4(\mathbb{F}_q)) = \Delta_1(\mathrm{Sp}_4(\mathbb{F}_q)) \leq 8(5*2+3) = 104$. Then Lemma 6.2.10 yields the last claim of the lemma. □

Remark 6.2.12. To our knowledge, nobody calculated the covering numbers of $\mathrm{Sp}_4(K)$ for general finite fields. However, we suspect that $\mathrm{cn}(\mathrm{Sp}_4(K)) = 4$ holds for all finite fields with at least 4 elements and as mentioned, Proposition 7.1.3 implies $\Delta_\infty(\mathrm{Sp}_4(K)) \geq 4$.

Next, we can give an explicit upper bound for $\Delta_k(\mathrm{Sp}_4(R))$ in some cases:

Theorem 6.2.13. *Let R be a ring of S-algebraic integers such that R is a principal ideal domain and $2R$-pseudo-good with $R \neq \mathbb{Z}[\frac{1+\sqrt{-3}}{2}], \mathbb{Z}$. Then for all $k \in \mathbb{N}$ one has:*

1. $\Delta_k(\mathrm{Sp}_4(R)) \leq 5 + 17644k$, *if* $(R : 2R) = 2$.

2. $\Delta_k(\mathrm{Sp}_4(R)) \leq 4 + 17644k$, *if* $(R : 2R) = 4$.

3. $\Delta_k(\mathrm{Sp}_4(R)) \leq 104 + 17644k$, *if* $(R : 2R) \geq 8$.

4. $\Delta_k(\mathrm{Sp}_4(R)) \leq 13824k$, *if* $(R : 2R) = 1$.

Proof. Using Dirichlet's Unit Theorem [32, Corollary 11.7], one can see that every ring of S-algebraic integers, except rings of imaginary quadratic integers and $\mathbb{Z}$, has infinitely many units. Consequently, according to Proposition 6.2.3, all $2R$-pseudo-good rings R as in the theorem have infinitely many units. Thus Theorem 6.1.5, Proposition 5.1.8 and the fact that $\mathrm{Sp}_4(R) = E(C_2, R)$ holds, implies that

$$\mathrm{Sp}_4(R) = (U^+(C_2,R)U^-(C_2,R))^4U^+(C_2,R) \text{ or } \mathrm{Sp}_4(R) = (U^-(C_2,R)U^+(C_2,R))^4U^-(C_2,R)$$

Consequently, we can assume $J = 5$ in Proposition 6.2.6 and so

$$K(C_2, R) \leq 6 + 8 * 5 = 46$$

holds, if $R \neq 2R$. However, if $R = 2R$, then clearly $N_{C_2} = \mathrm{Sp}_4(R)$ and $K(C_2, R) \leq 36$ holds. Next, we know from Theorem 4.2.1, that $L(C_2) \leq 384$, as R is a principal ideal domain. Furthermore, Corollary 6.2.7 implies if $R \neq 2R$ that

$$N_{C_2} = \ker\left(\pi_{2R} : \mathrm{Sp}_4(R) \to \mathrm{Sp}_4(R/2R)\right).$$

This implies that $G/N_{C_2} = \mathrm{Sp}_4(R/2R)$ and so Proposition 6.2.11 implies

1. $\Delta_\infty(G/N_{C_2}) = 5$, if $(R : 2R) = 2$.

2. $\Delta_\infty(G/N_{C_2}) = 4$, if $(R : 2R) = 4$.

3. $\Delta_\infty(G/N_{C_2}) \leq 104$, if $(R : 2R) \geq 8$.

Further, $\Delta_\infty(G/N_{C_2}) = 0$, clearly holds in case of $(R : 2R) = 1$. Combining these facts with the following inequality from Theorem 3.2.1:

$$\Delta_k(\mathrm{Sp}_4(R)) \leq L(C_2)K(C_2, R)k + \Delta_\infty(G/N_{C_2})$$

for $k \in \mathbb{N}$ yields the claim of the theorem. □

We finish this subsection by giving bounds in the two omitted cases:

Proposition 6.2.14. *For all* $k \in \mathbb{N}$, *one has*

1. $\Delta_k(\mathrm{Sp}_4(\mathbb{Z})) \leq 5 + 248064k$.

2. $\Delta_k\left(\mathrm{Sp}_4(\mathbb{Z}[\frac{1+\sqrt{-3}}{2}])\right) \leq 4 + 248064k$.

Proof. Both of these rings are $2R$-pseudo-good. This is obvious for $\mathbb{Z}$ by considering the set of representatives $\{0, 1\}$ and follows for $\mathbb{Z}[\frac{1+\sqrt{-3}}{2}]$ from Proposition 6.2.3. Furthermore, both $\mathbb{Z}$ and $\mathbb{Z}[\frac{1+\sqrt{-3}}{2}]$ are principal ideal domains. Again, this is well-known for $\mathbb{Z}$ and can be seen for $\mathbb{Z}[\frac{1+\sqrt{-3}}{2}]$ by way of using the norm map

$$N_{\mathbb{Q}[\sqrt{-3}]|\mathbb{Q}} : \mathbb{Z}\left[\frac{1+\sqrt{-3}}{2}\right] \to \mathbb{Z}$$

to show that $\mathbb{Z}[\frac{1+\sqrt{-3}}{2}]$ is an euclidean domain. Also Theorem 6.1.2 implies for both rings R that J in Proposition 6.2.6 can be chosen as 80. So Proposition 6.2.6 implies $K(C_2, R) \leq 6 + 8 * 80 = 646$. Lastly, Proposition 6.2.11 and Corollary 6.2.7 implies

1. $\Delta_\infty\left(\mathrm{Sp}_4(\mathbb{Z})/N_{C_2}\right) = \Delta_\infty(\mathrm{Sp}_4(\mathbb{F}_2)) = 5$.
2. $\Delta_\infty\left(\mathrm{Sp}_4(\mathbb{Z}[\frac{1+\sqrt{-3}}{2}])/N_{C_2}\right) = \Delta_\infty\left(\mathrm{Sp}_4(\mathbb{F}_4)\right) = 4$.

This yields the proposition in the same way as in the proof of Theorem 6.2.13. □

6.3 Explicit bounds for G_2

In this section, we will explain how to give explicit values for $G := G_2(R)$ for R certain rings of algebraic integers. The strategy is similar to the one for $\mathrm{Sp}_4(R)$. Recall that the positive roots ing G_2 are $\alpha, \beta, \alpha+\beta, 2\alpha+\beta, 3\alpha+\beta$ and $3\alpha+2\beta$ for α short and simple and β long and simple. Further, recall:

$$\begin{aligned} Q_{G_2} :=& \{A\varepsilon_\phi(2x)A^{-1} | x \in R, \phi \in G_2 \text{ short}, A \in G_2(R)\} \\ &\cup \{A\varepsilon_\phi(x)A^{-1} | x \in R, \phi \in G_2 \text{ long}, A \in G_2(R)\} \end{aligned}$$

as well as the group $N_{G_2} := \langle Q_{G_2} \rangle$. Further, let $\pi : G \to G/N_{G_2}$ be the quotient map. Then from Theorem 3.2.5 for $G_2(R)$, one notes for $k \in \mathbb{N}$ that:

$$\Delta_k(G_2(R)) \leq 6K(G_2, R)L(G_2)k + \Delta_\infty(G/N_{G_2})$$

where

1. the constant $L(G_2)$ is given as in Theorem 3.2.2 and Proposition 4.5.7,
2. the constant $K(R, G_2)$ is defined to be $\|N_{G_2}\|_{Q_{G_2}}$.

However, in contrast to the situation for $\mathrm{Sp}_4(R)$, the group N_{G_2} is already the entire group $G_2(R)$ in a lot of cases and this implies $\Delta_\infty(G/N_{G_2}) = 0$ and $K(R, G_2) = \|G_2(R)\|_{Q_{G_2}}$ then. First, we will show the following useful lemma:

Lemma 6.3.1. *Let R be a ring of S-algebraic integers with $R/2R = \mathbb{F}_2$. Then there is an epimorphism $q : G_2(R)/N_{G_2} \to \mathbb{F}_2$ with*

$$q(\varepsilon_\phi(a)N_{G_2}) = \begin{cases} a + 2R, & \text{if } \phi \in G_2 \text{ short} \\ 0, & \text{if } \phi \in G_2 \text{ long} \end{cases}$$

Proof. First, observe that there is an epimorphism

$$p' : G_2(\mathbb{F}_2) \to \mathbb{F}_2$$

with

$$p'(\varepsilon_\phi(a)) = \begin{cases} a, & \text{if } \phi \in G_2 \text{ short} \\ 0, & \text{if } \phi \in G_2 \text{ long} \end{cases}$$

for $\phi \in G_2$ and $a \in \mathbb{F}_2$. This can be seen as follows: The group $G_2(\mathbb{F}_2)$ is generated by the root elements $\varepsilon_\phi(1)$ of order 2 for $\phi \in G_2$ and according to [41, Chapter 6, p. 43, Theorem 8(b)], the fact that root elements have order 2 and the following relations on the root elements of $G_2(\mathbb{F}_2)$ already give a finite presentation for the group $G_2(\mathbb{F}_2)$:

$(\varepsilon_\phi(1), \varepsilon_\psi(1)) = \varepsilon_{\psi+\phi}(1)\varepsilon_{2\psi+\phi}(1)\varepsilon_{3\psi+\phi}(1)\varepsilon_{3\psi+2\phi}(1)$, if $\phi + \psi \in G_2$ and ϕ long and ψ short,
$(\varepsilon_\phi(1), \varepsilon_\psi(1)) = \varepsilon_{\phi+\psi}(1)$, if $\phi + \psi \in G_2$ is long
$(\varepsilon_\phi(1), \varepsilon_\psi(1)) = \varepsilon_{2\phi+\psi}(1)\varepsilon_{\phi+2\psi}(1)$, if $\phi + \psi \in G_2$ and $\phi + \psi, \phi, \psi$ short.

However, the map p' defined as above on the root elements $\varepsilon_\phi(a)$ respects these relations and so p' extends to a group homomorphism $p' : G_2(\mathbb{F}_2) \to \mathbb{F}_2$ as required and this p' is obviously surjective. Hence, there is an epimorphism

$$p : G_2(R) \to G_2(R/2R) = G_2(\mathbb{F}_2) \to \mathbb{F}_2$$

with

$$p(\varepsilon_\phi(a)) = \begin{cases} a + 2R, & \text{if } \phi \in G_2 \text{ short} \\ 0, & \text{if } \phi \in G_2 \text{ long} \end{cases} \tag{6.3}$$

for $\phi \in G_2$ and $a \in R$. Thus to obtain an epimorphism

$$q : G_2(R)/N_{G_2} \to \mathbb{F}_2$$

it suffices to show that $p(N_{G_2}) = 0$ and so it suffices to show $p(Q_{G_2}) = 0$. However, this is obvious due to (6.3). □

Then we obtain:

Proposition 6.3.2. *Let R be a $2R-$pseudo-good ring and let $n \in \mathbb{N}$ be given such that $\|G_2(R)\|_{\mathrm{EL}} \leq n$ holds.*

1. *If $|R/2R| \geq 4$, then $K(G_2, R) \leq 9n$ and $G_2(R) = N_{G_2}$.*
2. *If $|R/2R| = 0$, then $K(G_2, R) \leq n$ and $G_2(R) = N_{G_2}$.*
3. *If $R/2R = \mathbb{F}_2$, then $K(G_2, R) \leq 12n + 1$.*

Proof. For the first claim of the lemma, it suffices to show that

$$\varepsilon_\phi(x) \in N_{G_2} \text{ and } \|\varepsilon_\phi(x)\|_{Q_{G_2}} \leq 9$$

for all $\phi \in G_2$ and $x \in R$. First, observe that this is obvious for $\phi \in G_2$ long, so we may assume that $\phi \in G_2$ is short. Furthermore, we can assume after conjugation by appropriate Weyl group elements that $\phi = \alpha$. Observe that $K := R/2R$ has at least four elements. So K contains an element $\bar{t}$ such that neither $\bar{t}$ nor $\bar{t} - 1$ are trivial, because otherwise K would only have two elements. Then as R is $2R-$pseudo-good, we can pick a unit $x \in R$ with $x + 2R = \bar{t} + 1$ and an element $y \in R$ with $y + 2R = \bar{t}^{-1}$. Further, let $z \in R$ be arbitrary. This implies

$$(h_\beta(x^{-1}), \varepsilon_\alpha(yz)) = h_\beta(x^{-1})\varepsilon_\alpha(yz)h_\beta(x^{-1})^{-1}\varepsilon_\alpha(-yz) = \varepsilon_\alpha(xyz)\varepsilon_\alpha(-yz) = \varepsilon_\alpha((x-1)yz).$$

But by definition of x and y, we obtain $(x-1)y - 1 \in 2R$. Hence there is a $u \in R$ such that $(h_\beta(x^{-1}), \varepsilon_\alpha(yz)) = \varepsilon_\alpha(z)\varepsilon_\alpha(2u)$. However, observe that $h_\beta(x^{-1})$ is an element of N_{G_2} and

$$\begin{aligned}\|h_\beta(x^{-1})\|_{Q_{G_2}} &= \|\varepsilon_\beta(x^{-1})\varepsilon_{-\beta}(-x)\varepsilon_\beta(x^{-1})\varepsilon_\beta(-1)\varepsilon_{-\beta}(1)\varepsilon_\beta(-1)\|_{Q_{G_2}} \\ &= \|\varepsilon_\beta(x^{-1} - 1)\varepsilon_{-\beta}(-x)\varepsilon_\beta(x^{-1} - 1)\varepsilon_{-\beta}(1)\|_{Q_{G_2}} \leq 4.\end{aligned}$$

Thus, we can conclude that

$$\|\varepsilon_\alpha(z)\|_{Q_{G_2}} = \|(h_\beta(x^{-1}), \varepsilon_\alpha(yz))\varepsilon_\alpha(-2u)\|_{Q_{G_2}} \leq 2 * 4 + 1 = 9.$$

The second claim of the lemma is obvious, because if $R/2R$ is trivial, then $2 \in R$ is a unit and hence not only are $\varepsilon_\phi(x)$ for $x \in R$ and $\phi \in G_2$ long elements of Q_{G_2}, but also $\varepsilon_\phi(x) = \varepsilon_\phi(2(x/2))$ for $x \in R$ and ϕ short.

For the third and last claim of the lemma, we will first show that each element of $\varepsilon_\phi(R)$ for $\phi \in G_2$ agrees with an element of $\varepsilon_\alpha(R)$ after multiplication with at most 12 elements of Q_{G_2}. This is obvious if ϕ is long, because then $\varepsilon_\phi(R) \subset Q_{G_2}$ holds. So let

$\phi \in G_2$ be short and $a \in R$ be given and consider $\varepsilon_\phi(a)$. Assume first, that ϕ is positive and consider the case $\phi = \alpha + \beta$. Then observe that $w_\beta \varepsilon_{\alpha+\beta}(a) w_\beta^{-1} = \varepsilon_\alpha(\pm a)$. But

$$w_\beta = \varepsilon_\beta(1)\varepsilon_{-\beta}(-1)\varepsilon_\beta(1) \sim \varepsilon_\beta(2)\varepsilon_{-\beta}(-1)$$

and hence $\|w_\beta\|_{Q_{G_2}} \leq 2$ and $\|w_\beta^{-1}\|_{Q_{G_2}} \leq 2$ hold. Hence $\varepsilon_{\alpha+\beta}(a)$ agrees with an element of $\varepsilon_\alpha(R)$ after multiplication with up to 4 elements of Q_{G_2}. Second, consider the case $\phi = 2\alpha + \beta$. Observe that

$$(\varepsilon_\beta(a), \varepsilon_\alpha(1)) = \varepsilon_{\alpha+\beta}(\pm a)\varepsilon_{2\alpha+\beta}(\pm a)\varepsilon_{3\alpha+\beta}(\pm a)\varepsilon_{3\alpha+2\beta}(\pm a^2).$$

Observe that $\|(\varepsilon_\beta(a), \varepsilon_\alpha(1))\|_{Q_{G_2}} \leq 2$ as well as $\|\varepsilon_{3\alpha+\beta}(\pm a)\varepsilon_{\alpha+\beta}(\pm a^2)\|_{Q_{G_2}} \leq 2$. This implies that

$$\varepsilon_{2\alpha+\beta}(\pm a) \cdot \left(\varepsilon_{3\alpha+\beta}(\pm a)\varepsilon_{3\alpha+2\beta}(\pm a^2)\,(\varepsilon_\beta(a), \varepsilon_\alpha(1))^{-1}\right) = \varepsilon_{\alpha+\beta}(\pm a)$$

and so $\varepsilon_{2\alpha+\beta}(\pm a)$ agrees with an element of $\varepsilon_{\alpha+\beta}(R)$ after multiplication of up to 4 elements of Q_{G_2} and hence, using the first case, with an element of $\varepsilon_\alpha(R)$ after multiplication with up to $8 = 4 + 4$ elements of Q_{G_2}.

Next, assume that ϕ is negative and short. First, assume that $\phi = -\alpha - \beta$. Similar to the case of $\phi = 2\alpha + \beta$, we need 8 elements of Q_{G_2} to turn

$$\varepsilon_\alpha(\pm a) = w_\alpha \varepsilon_{-\alpha}(a) w_\alpha^{-1}$$

into the element

$$\varepsilon_{-\alpha-\beta}(\pm a) = w_\alpha \varepsilon_{-2\alpha-\beta}(\pm a) w_\alpha^{-1}.$$

In the case of $\phi = -\alpha$ one needs another 4 elements of Q_{G_2} to turn $\varepsilon_{-\alpha-\beta}(\pm a)$ into $\varepsilon_{-\alpha}(\pm a)$. So in total one needs $12 = 8 + 4$ elements of Q_{G_2} to turn $\varepsilon_{-\alpha}(a)$ into an element of $\varepsilon_\alpha(R)$.

For $\phi = -2\alpha - \beta$, one needs 4 elements of Q_{G_2} to turn an element of $\varepsilon_{-2\alpha-\beta}(R)$ into an element of $\varepsilon_{-\alpha-\beta}(R)$ and one needs another 8 elements of Q_{G_2} to turn an element of $\varepsilon_{-\alpha-\beta}(R)$ into an element of $\varepsilon_\alpha(R)$. Hence in total, one needs 12 element of Q_{G_2} to turn an element of $\varepsilon_{-2\alpha-\beta}(R)$ into an element of $\varepsilon_\alpha(R)$. To summarize, one needs at most 12 elements of Q_{G_2} to turn an element of $\varepsilon_\phi(R)$ for $\phi \in G_2$ into an element of $\varepsilon_\alpha(R)$.

To finish the proof of the third claim let $A \in N_{G_2}$ be given and choose $\phi_1, \ldots, \phi_n \in G_2$ and $a_1, \ldots, a_n \in R$ with

$$A = \prod_{i=1}^{n} \varepsilon_{\phi_i}(a_i).$$

This implies that up to multiplication with $12n$ elements of Q_{G_2}, the element A is an

element of $\varepsilon_\alpha(R)$. Hence there is a $b \in R$ with

$$\|A\varepsilon_\alpha(-b)\|_{Q_{G_2}} \leq 12n.$$

However, A is an element of N_{G_2} and so $\varepsilon_\alpha(b)$ is an element of N_{G_2} as well. Yet, according to Lemma 6.3.1, this implies that $b \in 2R$ and hence $\varepsilon_\alpha(b)$ is an element of Q_{G_2}. Thus

$$\|A\|_{Q_{G_2}} \leq 12n + 1.$$

□

Remark 6.3.3. Further, $K(R, G_2) \leq 12n + r(R)$ holds, if the ideal $2R$ in R factorizes as follows:

$$2R = \mathcal{P}_1 \cdots \mathcal{P}_{r(R)}$$

with $R/\mathcal{P}_i = \mathbb{F}_2$ for all $i \in \{1, \ldots, r(R)\}$. This can be shown by following the proof strategy of the last part in Proposition 6.3.2 and finishing with the existence of an epimorphism $q : G_2(R)/N_{G_2} \to \mathbb{F}_2^{r(R)}$ with

$$p(\varepsilon_\phi(a)N_{G_2}) = \begin{cases} (a + \mathcal{P}_1, \ldots, a + \mathcal{P}_r), & \text{if } \phi \in G_2 \text{ short} \\ 0, & \text{if } \phi \in G_2 \text{ long} \end{cases} \tag{6.4}$$

Next, observe the following:

Corollary 6.3.4. *Let R be a $2R$-pseudo-good ring.*

1. *If $R/2R = \mathbb{F}_2$, then $G_2(R)/N_{G_2} = \mathbb{F}_2$ and so $\Delta_\infty(G_2(R)/N_{G_2}) = 1$.*
2. *If $|R/2R| \geq 4$ or $|R/2R| = 0$ then $G_2(R)/N_{G_2}$ is trivial and so $\Delta_\infty(G_2(R)/N_{G_2}) = 0$.*

We are in place now to give an explicit upper bound for $\Delta_k(G_2(R))$ in some cases:

Theorem 6.3.5. *Let R be a ring of S-algebraic integers such that R is a principal ideal domain and $2R$-pseudo-good with $R \neq \mathbb{Z}[\frac{1+\sqrt{-3}}{2}], \mathbb{Z}$. Then for all $k \in \mathbb{N}$, one has:*

1. *$\Delta_k(G_2(R)) \leq 41007264768k$, if $|R/2R| \geq 4$.*
2. *$\Delta_k(G_2(R)) \leq 4556362752k$, if $|R/2R| = 0$.*
3. *$\Delta_k(G_2(R)) \leq 54760730112k + 1$, if $R/2R = \mathbb{F}_2$.*

Proof. Similar to the proof of Theorem 6.2.13 each element of $G_2(R)$ can be written as a product of $54 = 9 * 6 = 9 * |G_2^+|$ root elements. Thus Proposition 6.3.2 implies

1. $K(G_2, R) \leq 9 * 54 = 486$, if $|R/2R| \geq 4$.

2. $K(G_2, R) \leq 54$, if $|R/2R| = 0$.

3. $K(G_2, R) \leq 12 * 54 + 1 = 649$, if $R/2R = \mathbb{F}_2$.

Further, Corollary 6.3.4 implies:

1. If $R/2R = \mathbb{F}_2$, then $\Delta_\infty(G_2(R)/N_{G_2}) = 1$.

2. If $|R/2R| \geq 4$ or $|R/2R| = 1$ then $\Delta_\infty(G_2(R)/N_{G_2}) = 0$.

Together with the result

$$L(G_2) \leq 14062848$$

from Proposition 4.5.7, the theorem follows from the inequality

$$\Delta_k(G_2(R)) \leq 6K(G_2, R)L(G_2)k + \Delta_\infty(G/N_{G_2})$$

from Theorem 3.2.5. □

We finish this subsection by giving bounds in the two cases omitted above:

Proposition 6.3.6. *For all $k \in \mathbb{N}$, one has*

1. $\Delta_k(G_2(\mathbb{Z})) \leq 82098906624k + 1$ *and*

2. $\Delta_k(G_2(\mathbb{Z}[\frac{1+\sqrt{-3}}{2}])) \leq 61510897152k$.

Proof. We already know that both of the rings are $2R$-pseudo-good and principal ideal domains. Hence Theorem 6.1.2 implies for both rings R that each element of $G_2(R)$ can be written as a product of 81 root elements. So Proposition 6.3.2 implies, as $\mathbb{Z}/2\mathbb{Z} = \mathbb{F}_2$ and $\mathbb{Z}[\frac{1+\sqrt{-3}}{2}]/2\mathbb{Z}[\frac{1+\sqrt{-3}}{2}] = \mathbb{F}_4$ hold, that

1. $K(G_2, \mathbb{Z}) \leq 12 * 81 + 1 = 973$ and

2. $K(G_2, \mathbb{Z}[\frac{1+\sqrt{-3}}{2}]) \leq 9 * 81 = 729$.

Lastly, Proposition 6.2.11 implies

1. $\Delta_\infty(G_2(\mathbb{Z})/N_{G_2}) = 1$.

2. $\Delta_\infty(G_2(\mathbb{Z}[\frac{1+\sqrt{-3}}{2}])/N_{G_2}) = 0$.

This yields the proposition together with $L(G_2) \leq 14062848$ from Proposition 4.5.7 in the same way as in the proof of Theorem 6.3.5. □

6.4 Orders in rings of algebraic integers

In this section, we talk about orders in rings of algebraic integers and Morris' results in [30] and how to use them to get strong boundedness results. We do not define orders precisely, but they are subrings of rings of algebraic integers that are also sublattices of the same ring of algebraic integers. The classical non-trivial example is $\mathbb{Z}[2i] = \{a + 2bi | a, b \in \mathbb{Z}\}$, a subring of the ring of Gaußian integers. Further, define for a subset X of $\mathrm{SL}_n(R)$ the level ideal $l(X)$ as the sum of all level ideals $l(A)$ for $A \in X$.

First, there is the following result by Morris that is very similar to our results, however wrong as stated.

Theorem 6.4.1. *[30, Theorem 6.1(1), Remark 6.2] Let B be an order in a ring of algebraic integers and S a multiplicative set in $B - \{0\}$. Further, assume either that $n \geq 3$ or that $S^{-1}B$ has infinitely many units. Also let X be a subset of $G := \mathrm{SL}_n(S^{-1}B)$, that is normalized by root elements and that does not consist entirely of scalar matrices. Then X boundedly generates a finite index subgroup N of $\mathrm{SL}_n(S^{-1}B)$ with a bound on the maximal length of a word in elements of X that depends on n, the degree $[K : \mathbb{Q}]$, the minimal numbers of generators of the level ideal $l(N)$ and the cardinality of $S^{-1}B/l(N)$. If $X := \{gsg^{-1} | s \in S, g \in \mathrm{SL}(S^{-1}B)\}$ for a finite set S with at least one non-scalar element, then the minimal number of generators of $l(N)$ is smaller than $n^2|S|$.*

As mentioned, this theorem is false as stated and we want to talk briefly about the error and how to interpret this theorem in the context of boundedness-considerations:

1. A dependence that Morris does not mention is that the bounded generation of N also depends on a first order description of the set X. The existence of this dependence can be seen as follows: If X is a finite collection of, say k, conjugacy classes in $\mathrm{SL}_n(R)$ generating $\mathrm{SL}_n(R)$, then none of the numbers mentioned in the theorem depend on k. Hence, $\mathrm{SL}_n(R)$ would be uniformly bounded. However, Corollary 7.1.7 shows that this is not the case. But adding this dependence on a first order description of X, Theorem 6.4.1 is correct.

2. We do not want to explain in detail how the dependence on a first order characterization of X arises in Morris' proof. Crucial to our investigation however is the fact that X being a collection of at most k many conjugacy classes generating $\mathrm{SL}_n(R)$ is a first order property. This holds because $A_1, \ldots, A_k \in \mathrm{SL}_n(R)$ normally generating $\mathrm{SL}_n(R)$ is equivalent to the first-order condition $\Pi(\{A_1, \ldots, A_k\}) = \emptyset$ according to Corollary 3.2.8.

3. In particular, the bounded generation result for $\mathrm{SL}_n(R)$ by k conjugacy classes obtained from the corrected Theorem 6.4.1 depends on k, but not on the particular classes themselves. Phrased in this way, this establishes that $\mathrm{SL}_n(S^{-1}B)$ is strongly

bounded. The main difference to our result, is that Morris has no control on the actual value of $\Delta_k(\mathrm{SL}_n(R))$, whereas we can establish that the dependence is at least linear in k. Structurally, the main reason for this difference is that Morris applies a first order compactness result to an entire *set of generators* to establish bounded generation. We, on the other hand, study the normal subgroup generated by a particular given element A of the group $G(\Phi, R)$ to obtain root elements with arguments lying in its level ideal $l(A)$ and only later consider the full generating set.

Morris [30, Theorem 5.26] proves bounded generation by root elements for the subgroup $E(A_1, R)$ of $\mathrm{SL}_2(R)$ also in the case that R is only a localization of an order, if said localization has infinitely many units. He further demonstrates that the elementary subgroup $E(A_2, R)$ of $\mathrm{SL}_3(R)$ is boundedly generated by root elements for R a localization of an order [30, Corollary 3.13]. Thus using Corollary 5.1.8 and modifying the proof of Theorem 3.1.2 and Theorem 3.2.5 in the same manner as we did to prove Theorem 5.2.9, one shows:

Proposition 6.4.2. *Let R be a localization of an order in a ring of algebraic integers and Φ an irreducible root system of rank at least 2. Assume further that R has infinitely many units in case Φ is not simply-laced. There is a constant $C(\Phi, R)$ such that*

$$\Delta_k(E(\Phi, R)) \leq C(\Phi, R)k$$

holds for all $k \in \mathbb{N}$.

6.5 Boundedness of $\mathrm{SL}_2(R)$ for rings with infinitely many units

We talk shortly about $\mathrm{SL}_2(R)$ in this section. Reading Morris' paper, especially the aforementioned Theorem 6.4.1, seems to imply that the boundedness properties for $\mathrm{SL}_2(R)$ for R a localization of an order with infinitely many units might be the same as for $\mathrm{SL}_n(R)$ for $n \geq 3$. So one could believe that Morris proved:

Conjecture 6.5.1. *Let R be a ring of S-algebraic integers with infinitely many units. Then $\mathrm{SL}_2(R)$ is strongly bounded.*

We believe this to be true, but Morris did not prove it. The problem is the requirement mentioned in Section 6.4 to give a first-order description of the property of a collection of k conjugacy classes to normally generate $\mathrm{SL}_2(R)$. In contrast to the case $n \geq 3$, no such characterization is known to us. For example, $\Pi(S) = \emptyset$ does not suffice to prove that $S \subset \mathrm{SL}_2(R)$ normally generates $\mathrm{SL}_2(R)$: The ring $R = \mathbb{Z}[\frac{1+\sqrt{17}}{2}]$ is a counter-example. It has infinitely many units according to [32, Corollary 11.7] and the element

$A = I_2 + e_{12} \in \mathrm{SL}_2(R)$ satisfies $\Pi(\{A\}) = \emptyset$. However, the ideal $2R$ factors in R as $2R = \mathcal{P}_1 \cdot \mathcal{P}_2$ for

$$\mathcal{P}_1 = \left(\frac{3+\sqrt{17}}{2}\right) \text{ and } \mathcal{P}_2 = \left(\frac{3-\sqrt{17}}{2}\right)$$

with $\mathcal{P}_1 \neq \mathcal{P}_2$ and $R/\mathcal{P}_1 = R/\mathcal{P}_2 = \mathbb{F}_2$, which implies that there is an epimorphism

$$\mathrm{SL}_2(R) \to \mathrm{SL}_2(R/\mathcal{P}_1) \times \mathrm{SL}_2(R/\mathcal{P}_2) = \mathrm{SL}_2(\mathbb{F}_2)^2.$$

But the group $\mathrm{SL}_2(\mathbb{F}_2)$ is isomorphic to the permutation group S_3, which can be seen from the fact that $\mathrm{SL}_2(\mathbb{F}_2)$ operates on the three non-zero vectors of $\mathbb{F}_2^2$. But then composing with the sign epimorphism $S_3 \to \mathbb{F}_2$ yields an epimorphism $q : \mathrm{SL}_2(R) \to \mathbb{F}_2^2$ and as $\mathbb{F}_2^2$ is abelian, the element A could only normally generate $\mathrm{SL}_2(R)$, if $q(A)$ would generate $\mathbb{F}_2^2$, which is obviously impossible.

However, a finite collection of conjugacy classes X generating $\mathrm{SL}_2(R)$ cannot be entirely scalar and this is a first order property. Further, $G(A_1, R) = E(A_1, R)$ holds according to Theorem 6.1.5. So a suitably adjusted version of Theorem 6.4.1 yields that X boundedly generates $\mathrm{SL}_2(R)$ with the bound depending on X. Hence one obtains:

Proposition 6.5.2. *Let R be a ring of S-algebraic integers with infinitely many units. Then $\mathrm{SL}_2(R)$ is bounded.*

Chapter 7

Finite, normally generating sets in Chevalley groups

In the previous chapters and theorems, we proved upper bounds on the diameter of conjugation generated word norms on split Chevalley groups $G(\Phi, R)$ for R a ring of S-algebraic integers. These bounds are linear in the number of conjugacy classes in the corresponding generating set. In this chapter, we show that this linearity is sharp in the sense, that for almost all $k \in \mathbb{N}$, there are finite normally generating sets S_k with

$$\|G(\Phi, R)\|_{S_k} \geq 2k$$

and $|S_k| = k$. So in general, linear bounds in the cardinality of the normally generating sets are the best possible. The dichotomy between G_2, C_2 and the other Φ persists here. Namely, for $\Phi = C_2$ or G_2 such lower bounds depend strongly on the ring R.

In the first section, we speak about root systems Φ of high rank and demonstrate the existence of normally generating sets with better lower bounds in some cases. In the second section, we speak about normally generating sets of $\mathrm{Sp}_4(R)$ and $G_2(R)$ for R rings of algebraic integers. It will turn out that the existence of these sets is restricted by number theoretic properties of R.

7.1 Conjugacy classes in groups of Lie type and lower bounds in the higher rank cases

In this section, we give lower bounds on $\Delta_k(G(\Phi, R))$ for Φ an irreducible root system of rank at least 2 not equal to C_2 or G_2. In order to do this, we will first give lower bounds on conjugation-generated norms on words norms defined over fields and then apply those lower bounds to obtain lower bounds for $G(\Phi, R)$.

First, we need the following statement:

Lemma 7.1.1. *Let K be a field, Φ an irreducible root system of rank at least 2, ϕ a root in Φ and $t \in K$ not-zero. If $\Phi = G_2$, then assume further that ϕ is short. Then $E := \varepsilon_\phi(t)$ normally generates $G(\Phi, K)$ and $\|G(\Phi, K)\|_E \geq 2$ holds.*

Proof. We will show first that E normally generates $G(\Phi, R)$. Every field is semi-local and hence according to Proposition 5.2.2, the group $G(\Phi, K)$ is boundedly generated by root elements. Hence it suffices to show the conditions of Corollary 3.2.8 for $\{E\}$ to prove that E normally generates $G(\Phi, K)$. If $\Phi \neq C_2$ or G_2, then, we are reduced to satisfying the condition $\Pi(\{E\}) = \emptyset$ and this condition is satisfied, because $t \neq 0$. Hence the only remaining case is $\Phi = C_2$ or G_2. Regarding $\mathrm{Sp}_4(K)$: If ϕ is long, then using Lemma 3.4.2(2), one obtains that the normal subgroup N generated by E, also contains $\varepsilon_\alpha(t)$ for α the simple, positive, short root in C_2. So we may assume $\phi = \alpha$. Then the normal subgroup N generated by $E = \varepsilon_\alpha(t)$ also contains

$$\varepsilon_\alpha(s) = \varepsilon_\alpha(st^{-1}t) = h_\beta(s^{-1}t)\varepsilon_\alpha(t)h_\beta(s^{-1}t)^{-1}$$

for all $s \in K - \{0\}$. Hence N contains all root elements for $\varepsilon_\psi(x)$ for $\psi \in C_2$ short and $x \in K$. Thus according to Lemma 3.4.2(3), the normal subgroup N also contains all root elements $\varepsilon_\psi(x)$ for $\psi \in C_2$ long and $x \in K$. Hence N contains all root elements of $\mathrm{Sp}_4(K)$ and hence $N = \mathrm{Sp}_4(K)$ holds according to Proposition 5.2.2. Similarly, for $G_2(K)$, one can show that the normal subgroup generated by E contains all root elements $\varepsilon_\psi(x)$ for $\psi \in G_2$ short and $x \in K$. Then Lemma 3.5.4(2) implies that N contains all root elements $\varepsilon_\psi(x)$ for $\psi \in G_2$ long and $x \in K$. Thus N contains all root elements of $G_2(K)$ and hence $N = G_2(K)$ holds according to Proposition 5.2.2. Thus E generates $G(\Phi, R)$ in all cases.

To finish the proof, it suffices to show that $\|G(\Phi, K)\|_{\varepsilon_\phi(t)} \leq 1$ is not true. If it were, then each element in $G(\Phi, K)$ would be conjugate to either $\varepsilon_\phi(t)$ or $\varepsilon_\phi(-t)$. But if $\bar{K}$ is the algebraic closure of K, then this would imply that each element of the subgroup $G(\Phi, K)$ of $G(\Phi, \bar{K})$ would be unipotent in the linear algebraic group $G(\Phi, \bar{K})$. However, if K is not the field $\mathbb{F}_2$, then for some simple root $\phi \in \Phi$ and $s \neq 0, 1$, the element $h_\phi(s)$ is not unipotent. This resolves the case $K \neq \mathbb{F}_2$.

If $K = \mathbb{F}_2$ holds and if each element in $G(\Phi, K)$ is conjugate to $\varepsilon_\phi(1) = \varepsilon_\phi(t) = \varepsilon_\phi(-t)$, then each element in $G(\Phi, K)$ would have the same order as $\varepsilon_\phi(1)$, that is $\mathrm{char}(K) = 2$. However, as $\mathbb{F}_2 = K$ holds, the Weyl group $W(\Phi)$ is actually a subgroup of $G(\Phi, R)$ according to [41, Chapter 3, p.24, Lemma 22]. But each Weyl group $W(\Phi)$ has one of the following groups as subgroups:

$$W(A_2) \cong S_3, W(C_2) \cong D_4 \text{ or } W(G_2) \cong D_6.$$

But clearly all three groups S_3, D_4 and D_6 contain elements of orders different than 2 and

this contradiction finishes the proof. □

Remark 7.1.2. The condition that ϕ is short is necessary in the case of $\Phi = G_2$, because for ϕ a long root in G_2 the element $\varepsilon_\phi(1) \in G_2(\mathbb{F}_2)$ is contained in the kernel of the epimorphism $q : G_2(\mathbb{F}_2) \to \mathbb{F}_2$ from Lemma 7.2.4 and hence cannot possibly normally generate a larger group than this kernel.

Second, we need the following:

Proposition 7.1.3. *Let K be a field, $t \in K - \{0\}$.*

1. *For $n \geq 2$ and $\phi \in A_n$, the element $E := \varepsilon_\phi(t)$ normally generates $G(A_n, K)$ and $\|G(A_n, K)\|_E \geq n + 1$.*

2. *For $n \geq 3$ and $\phi \in B_n$, the element $E := \varepsilon_\phi(t)$ normally generates $G(B_n, K)$ and $\|G(B_n, K)\|_E \geq n + 1$.*

3. *For $n \geq 2$ and $\phi \in C_n$ long, the element $E := \varepsilon_\phi(t)$ normally generates $G(C_n, K)$ and $\|G(C_n, K)\|_E \geq 2n$.*

4. *For $n \geq 4$ and $\phi \in D_n$, the element $E := \varepsilon_\phi(t)$ normally generates $G(D_n, K)$ and $\|G(D_n, K)\|_E \geq n$.*

Proof. That E normally generates $G(\Phi, K)$ is clear in all cases from Lemma 7.1.1. We only do the rest of the proof for $G(C_n, K) = \mathrm{Sp}_{2n}(K)$, because the proofs are very similar in all cases. Note that using the conventions from Section 4.1, we can (possibly after conjugation with Weyl group elements) assume $E = I_{2n} + te_{1,n+1}$. We define the subspace

$$I(l) := \{v \in K^{2n} | l(v) = v\}.$$

for a linear map $l : K^{2n} \to K^{2n}$. We prove next that for $l_1, l_2 : K^{2n} \to K^{2n}$, one has

$$\dim_k(I(l_1 l_2)) \geq \dim_k(I(l_1)) + \dim_k(I(l_2)) - 2n. \tag{7.1}$$

To see this, observe first that $I(l_1) \cap I(l_2) \subset I(l_1 l_2)$ and hence

$$\begin{aligned} \dim_K(I(l_1 l_2)) &\geq \dim_K(I(l_1) \cap I(l_2)) = \dim_K(I(l_1)) + \dim_K(I(l_2)) - \dim_K(\langle I(l_1), I(l_2)\rangle) \\ &\geq \dim_K(I(l_1)) + \dim_K(I(l_2)) - 2n. \end{aligned}$$

Observe that the linear map $E : K^{2n} \to K^{2n}$ induced by E has

$$I(E^{-1}) = I(E) = Ke_1 \oplus \cdots \oplus Ke_n \oplus Ke_{n+2} \oplus \cdots \oplus Ke_{2n}.$$

Hence $\dim_K I(E) = 2n-1 = \dim_K I(E^{-1})$ holds. Note further for $X \in K^{2n\times 2n}$, $A \in \mathrm{GL}_{2n}(K)$ and $v \in K^{2n}$, that the following holds:

$$v \in I(X) \text{ precisely if } Av \in I(AXA^{-1}).$$

Hence $I(AXA^{-1}) = AI(X)$ holds and thus $\dim_K I(X) = \dim_K I(AXA^{-1})$. Hence for each conjugate X of E or E^{-1} in $\mathrm{Sp}_{2n}(K)$, one has $\dim_K(I(X)) = 2n-1$. Next, let $X_1, \ldots, X_k$ be either conjugates of E or E^{-1} in $\mathrm{Sp}_{2n}(K)$ or I_{2n}. Then one can show by induction on $k \in \mathbb{N}$ that $\dim_K(I(X_1 \cdots X_k)) \geq 2n-k$.

First, this claim is clear for $k = 1$. For the induction step, observe for $k > 1$ that applying (7.1) implies:

$$\begin{aligned}\dim_K(I(X_1 \cdots X_{k-1}X_k)) &\geq \dim_K(I(X_1 \cdots X_{k-1})) + \dim_K(I(X_k)) - 2n \\ &\geq \dim_K(I(X_1 \cdots X_{k-1})) + 2n - 1 - 2n \\ &= \dim_K(I(X_1 \cdots X_{k-1})) - 1 \geq 2n - (k-1) - 1 = 2n-k.\end{aligned}$$

This implies in particular that for each $A \in B_E(2n-1)$ there is a non-trivial vector $v(A) \in K^{2n}$ fixed by A. Hence each element of $B_E(2n-1)$ has eigenvalue 1. So if $\|\mathrm{Sp}_{2n}(K)\|_E \leq 2n-1$ or equivalently $B_E(2n-1) = \mathrm{Sp}_{2n}(K)$ were to hold, then each element $A \in \mathrm{Sp}_{2n}(K)$ would have eigenvalue 1. Thus it suffices to give an element $A \in \mathrm{Sp}_{2n}(K)$ without the eigenvalue 1 to finish the proof. To this end, observe that for $B \in \mathrm{SL}_n(K)$, the matrix

$$A = \left(\begin{array}{c|c} B & 0_n \\ \hline 0_n & B^{-T} \end{array}\right)$$

is an element of $\mathrm{Sp}_{2n}(R)$ with characteristic polynomial

$$\chi_A(x) = \chi_B(x)\chi_{B^{-T}}(x) = \chi_B(x)\chi_{B^{-1}}(x).$$

But this implies that A has eigenvalue 1 precisely if either B or B^{-1} has eigenvalue 1. Yet B^{-1} has eigenvalue 1 precisely if B does. Thus it suffices to provide an element $B \in \mathrm{SL}_n(K)$ without eigenvalue 1 to finish the proof. If $n = 2m$ is even for $m \geq 1$, then consider for B a block-diagonal matrix of the form

$$B = \begin{pmatrix} C & & & \\ & C & & \\ & & \ddots & \\ & & & C \end{pmatrix}$$

with the block C equal to the 2×2-matrix

$$C := \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$$

Observe that this implies for the characteristic polynomial

$$\chi_B(x) = \chi_C(x)^m = [(1-x)(-x)+1]^m = [x^2 - x + 1]^m.$$

Obviously, 1 is not a root of this polynomial, so 1 is not an eigenvalue of B. This finishes the case n even. If $n = 3 + 2m$ is odd for $m \geq 0$, consider the block-diagonal matrix

$$B = \begin{pmatrix} D & & & & \\ & C & & & \\ & & C & & \\ & & & . & \\ & & & & C \end{pmatrix}$$

for D the 3×3-matrix

$$D = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

Observe that this implies for the characteristic polynomial

$$\chi_B(x) = \chi_D(x)\chi_C(x)^m = [-x^3 + x + 1] \cdot [x^2 - x + 1]^m.$$

Obviously 1 is not a root of this polynomial either, so 1 is not an eigenvalue of this B. This finishes the case n odd and the proof. □

Remark 7.1.4.

1. In the case of B_n or D_n, the linear action used is the one induced by the map

   $$G(B_n, K) = \mathrm{Spin}_{2n+1}(K) \to \mathrm{SO}_{2n+1}(K) \subset \mathrm{GL}_{2n+1}(K)$$

   and

   $$G(D_n, K) = \mathrm{Spin}_{2n}(K) \to \mathrm{SO}_{2n}(K) \subset \mathrm{GL}_{2n}(K)$$

   respectively and note that the covering map $\mathrm{Spin}_i(K) \to \mathrm{SO}_i(K)$ is always surjective.

2. This 'dimension counting'-strategy is quite well-known and was mentioned to me by B. Karlhofer in a different context, but it is also alluded to in Lawther's and Liebeck's paper [25, p. 120].

3. There are a couple of other ways one could show that $\|G(\Phi, K)\|_E$ is bounded below by constants linear in the rank of Φ. For example, one could compare the dimension of the centralizer of E in $G(\Phi, K)$ with the dimension of $G(\Phi, K)$.

The main theorem in this sections is a generalization of [24, Theorem 6.1] with better lower bounds:

Theorem 7.1.5. *Let R a Dedekind domain with finite class number and at least k distinct maximal ideals. Further, let Φ be one of the following root systems:*

1. *A_n for $n \geq 2$,*
2. *B_n for $n \geq 3$,*
3. *C_n for $n \geq 3$,*
4. *D_n for $n \geq 4$,*
5. *E_6, E_7, E_8 or F_4*

such that $G(\Phi, R)$ is boundedly generated by root elements. Then the following hold

1. *$\Delta_k(G(A_n, R)) \geq k(n+1)$ for $n \geq 2$,*
2. *$\Delta_k(G(B_n, R)) \geq k(n+1)$ for $n \geq 3$,*
3. *$\Delta_k(G(C_n, R)) \geq 2nk$ for $n \geq 3$,*
4. *$\Delta_k(G(D_n, R)) \geq kn$ for $n \geq 4$,*
5. *$\Delta_k(G(\Phi, R)) \geq 2k$ for $\Phi = E_6, E_7, E_8, F_4$.*

Proof. To prove the theorem we have to construct a normally generating set S of $G(\Phi, R)$ such that $|S| = k$ and the diameter $\|G(\Phi, R)\|_S$ is bounded below by constants as claimed in the theorem. To this end, let k distinct maximal ideals $\mathcal{P}_1, \ldots, \mathcal{P}_k$ be given and let c be the class number of R. All the ideals $\mathcal{P}_i^c$ are principal for $i = 1, \ldots, k$ so choose $t_1, \ldots, t_k$ as generators of $\mathcal{P}_1^c, \ldots, \mathcal{P}_k^c$ respectively and set

$$r_i := \prod_{1 \leq j \neq i \leq k} t_j$$

for all $i = 1, \ldots, k$. Fix a long root $\phi \in \Phi$ next and consider for $i = 1, \ldots, k$ the elements $A_i := \varepsilon_\phi(r_i)$ and the set $S := \{A_1, \ldots, A_k\}$. Then $\Pi(A_i) = \bigcup_{j \neq i} \{P_j\}$ holds for $i = 1, \ldots, k$ and thus $\Pi(S) = \emptyset$ follows. Hence Corollary 3.2.8 implies that S is a normally generating set of $G(\Phi, R)$. Next, set $K_i := R/\mathcal{P}_i$ for $i = 1, \ldots, k$ and consider the map

$$\pi : G(\Phi, R) \to \prod_{i=1}^{k} G(\Phi, K_i), A \mapsto (\pi_{\mathcal{P}_1}(A), \ldots, \pi_{\mathcal{P}_k}(A)).$$

Further observe that r_j is an element of $\mathcal{P}_i$ for all $1 \leq i \neq j \leq k$ and r_j is not an element of $\mathcal{P}_j$. Thus the only non-trivial component of $\pi(A_j)$ is the $G(\Phi, K_j)$-component equal to $\varepsilon_\phi(r_j + \mathcal{P}_j)$ and $G(\Phi, K_j)$ is normally generated by $\varepsilon_\phi(r_j + \mathcal{P}_j)$ according to Lemma 7.1.1. Also this implies that the only non-trivial component of any conjugate of $\pi(A_j)$ is the $G(\Phi, K_j)$-component. Together this implies that $\pi(S)$ normally generates $\prod_{i=1}^{k} G(\Phi, K_i)$ and

$$\|G(\Phi, R)\|_S \geq \|\prod_{i=1}^{k} G(\Phi, K_i)\|_{\pi(S)} = \sum_{i=1}^{k} \|G(\Phi, K_i)\|_{\varepsilon_\phi(r_i + \mathcal{P}_i)}.$$

First, observe that $\|G(\Phi, K_i)\|_{\varepsilon_\phi(r_i+\mathcal{P}_i)} \geq 2$ holds according to Lemma 7.1.1 for all Φ and all $i = 1, \ldots, k$. This implies

$$\|G(\Phi, R)\|_S \geq \sum_{i=1}^{k} \|G(\Phi, K_i)\|_{\varepsilon_\phi(r_i+\mathcal{P}_i)} \geq 2k.$$

This finishes the proof in the cases $\Phi = E_6, E_7, E_8$ and F_4.

For the other cases of Φ it suffices to apply Proposition 7.1.3 to obtain

1. $\|G(A_n, K_i)\|_{\varepsilon_\phi(r_i+\mathcal{P}_i)} \geq n + 1$,
2. $\|G(B_n, K_i)\|_{\varepsilon_\phi(r_i+\mathcal{P}_i)} \geq n + 1$,
3. $\|G(C_n, K_i)\|_{\varepsilon_\phi(r_i+\mathcal{P}_i)} \geq 2n$ and
4. $\|G(D_n, K_i)\|_{\varepsilon_\phi(r_i+\mathcal{P}_i)} \geq n$

for all $i = 1, \ldots, k$. Hence we obtain

$$\|G(\Phi, R)\|_S \geq \sum_{i=1}^{k} \|G(\Phi, K_i)\|_{\varepsilon_\phi(r_i+\mathcal{P}_i)} \geq \begin{cases} k(n+1), \text{ if } \Phi = A_n \text{ for } n \geq 2 \\ k(n+1), \text{ if } \Phi = B_n \text{ for } n \geq 3 \\ 2nk, \text{ if } \Phi = C_n \text{ for } n \geq 3 \\ kn, \text{ if } \Phi = D_n \text{ for } n \geq 4 \end{cases}$$

This finishes the proof. □

We want to point out that for $T \in \{A, B, C, D\}$ and $n, k \in \mathbb{N}$ this theorem provides a lower bound linear in n and k. Namely, the equation

$$\frac{\Delta_k(G(T_n, R))}{kn} > C_T$$

holds for some constant $C_T > 0$. Yet it is not possible to find lower bounds of $\Delta_k(G(A_n, R))$ for example with a better asymptotic behaviour in n using arguments as in the proof of Theorem 7.1.5. This is the case, because the covering numbers of finite groups of Lie type

are known to be linear in the rank of the corresponding root system, due to Liebeck's and Lawther's paper [25]. However, Theorem 7.1.5 gives bounds for semi-local rings, which are also principal ideal domains that have the 'correct' asymptotic in the rank of Φ and the number of maximal ideals:

Corollary 7.1.6. *Let R be a principal ideal domain with precisely q distinct maximal ideals. Further, let $n >> 0$ be a natural number. Then*

1. $q(n+1) \leq \Delta_\infty(\mathrm{SL}_n(R)) \leq 12(n-1)q$ *and*
2. $2nq \leq \Delta_\infty(\mathrm{Sp}_{2n}(R)) \leq 768(3n-2)q$

hold.

Proof. The upper bounds are consequences of Corollary 5.2.5 and the lower bounds consequences of Theorem 7.1.5. □

For rings of S-algebraic integers the situation is less well understood. There is a discrepancy between the asymptotic of the upper and lower bounds:

Corollary 7.1.7. *Let R be a ring of S-algebraic integers with class number 1 and infinitely many units. Further, let $n \geq 3$ be given. Then*

1. $k(n+1) \leq \Delta_k(\mathrm{SL}_n(R)) \leq (4n+1)(4n+4)k$ *and*
2. $2nk \leq \Delta_k(\mathrm{Sp}_{2n}(R)) \leq 192(1+5n)(12n+12)k$

hold for all $k \in \mathbb{N}$.

Proof. The lower bounds follow from Theorem 7.1.5 again and the upper bounds are a consequence of Corollary 6.1.6 and Corollary 6.1.9 respectively. □

Remark 7.1.8. There are similar statements for rings of algebraic integers with only finitely many units.

7.2 Finite normally generating sets of Sp_4 and G_2

Next, we are going to describe lower bounds on $\Delta_k(\mathrm{Sp}_4(R))$ and $\Delta_k(G_2(R))$ in the case of S-algebraic integers. It turns out that in this case the (existence of) lower bounds depends on the way 2 splits into primes in the ring R.

Theorem 7.2.1. *Let Φ be C_2 or G_2 and let R be a ring of S-algebraic integers in a number field. Further define*

$$r := r(R) := |\{\mathcal{P} |\ \mathcal{P} \text{ is a prime ideal with } R/\mathcal{P} = \mathbb{F}_2\}|.$$

Then

1. *the inequalities* $\Delta_k(G(C_2, R)) \geq 4k + r(R)$ *and* $\Delta_k(G_2(R)) \geq 2k$ *hold for all* $k \in \mathbb{N}$ *with* $k \geq r(R)$ *and*

2. *the equality* $\Delta_k(G(\Phi, R)) = -\infty$ *holds for* $k < r(R)$.

We show both parts of the theorem separately. For the first part, the main difficulty, compared to Theorem 7.1.5 comes, from the more complex conditions a set S has to fulfill to be a normally generating set. To address this, we first describe the way certain quotients of rings of S-algebraic integers are generated by their units:

Lemma 7.2.2. *Let* R *be a ring of S-algebraic integers,* $\mathcal{P}_1, \ldots, \mathcal{P}_s$ *distinct non-zero, prime ideals in* R *and* $l_1, \ldots, l_s$ *natural numbers and set* $\bar{R} := R/(\mathcal{P}_1^{l_1} \cdots \mathcal{P}_s^{l_s})$.

1. *If* $|R/\mathcal{P}_i| \geq 3$ *holds for all* $i = 1, \ldots, s$, *then each element in* $\bar{R}$ *is the sum of two units.*

2. *If at most one of the* $\mathcal{P}_i$ *has the property* $|R/\mathcal{P}_i| = 2$, *then each element in* $\bar{R}$ *is the sum of at most three units.*

Proof. We will show the first claim by induction on s. Let $s = 1$ and $x \in R$ be given. Then assume for all $a \in R$ that either $a \in \mathcal{P}_1$ or $x - a \in \mathcal{P}_1$ holds. Thus $a(x - a) \in \mathcal{P}_1$ would hold for all $a \in R$. Thus each element in the integral domain $R/\mathcal{P}_1$ is either $x + \mathcal{P}_1$ or trivial and hence $|R/\mathcal{P}_1| = 2$. This contradiction yields the existence of an $a \in R$ with neither a nor $x - a$ elements of $\mathcal{P}_1$. But then both $a + \mathcal{P}_1^{l_1}$ and $x - a + \mathcal{P}_1^{l_1}$ are units in $R/\mathcal{P}_1^{l_1}$. This solves the case $s = 1$.

For the induction step let prime ideals $\mathcal{P}_1, \ldots, \mathcal{P}_{s+1}$ be given and assume by induction that every element $y_1 \in R/(\mathcal{P}_1^{l_1} \cdots \mathcal{P}_s^{l_s}) := R_1$ is the sum of two units $u_1, u_2 \in R_1$. Also by the beginning of the induction, each element $y_2 \in R/\mathcal{P}_{s+1}^{l_{s+1}} := R_2$ is the sum of two units $u_3, u_4 \in R_2$. This implies that

$$(y_1, y_2) = (u_1 + u_2, u_3 + u_4) = (u_1, u_3) + (u_2, u_4) \in R_1 \times R_2 = R/(\mathcal{P}_1^{l_1} \cdots \mathcal{P}_s^{l_{s+1}}) = \bar{R}$$

is also the sum of two units. This proves the first claim of the lemma.

For the second claim of the lemma, assume wlog that $|R/\mathcal{P}_1| = 2$ and let $y_1 \in R/\mathcal{P}_1^{l_1} =: R_1, y_2 \in R/(\mathcal{P}_2^{l_2} \cdots \mathcal{P}_s^{l_s}) =: R_2$ be given. Then by the first claim of the lemma, there are units $u_1, u_2 \in R_2$ with $y_2 = u_1 + u_2$. We distinguish two cases: First, assume y_1 is not a unit in R_1. But then both $1 + y_1$ and -1 are units in R_1. Hence $(y_1, y_2) = (1 + y_1, u_1) + (-1, u_2)$ is a sum of two units in $\bar{R}$. On the other hand, assume y_1 is a unit in R_1. By the first claim of the lemma, the element $u_2 \in R_2$ can be written as the sum $u_3 + u_4$ for two units u_3, u_4 in R_2. This implies that

$$(y_1, y_2) = (y_1, u_1) + (0, u_2) = (y_1, u_1) + (1, u_3) + (-1, u_4)$$

is the sum of three units. □

This implies the following technical proposition:

Proposition 7.2.3. *Let R be a ring of S-algebraic integers such that the ideal $2R$ factorizes in prime ideals as follows*

$$2R = \left(\prod_{i=1}^{r} \mathcal{P}_i^{l_i}\right) \cdot \left(\prod_{j=1}^{s} \mathcal{Q}_j^{k_j}\right)$$

with $[R/\mathcal{P}_i : \mathbb{F}_2] = 1$ for $1 \leq i \leq r$ and $[R/\mathcal{Q}_j : \mathbb{F}_2] > 1$ for $1 \leq j \leq s$. Further, let $r_1, \ldots, r_r \in R$ be given such that for each $i = 1, \ldots, r$, the element r_i is contained in each $\mathcal{P}_k^{l_k}$ for $k \neq i$ and maps to a unit in $R_i := R/(\mathcal{P}_i^{l_i} \cdot \left(\prod_{j=1}^{s} \mathcal{Q}_j^{k_j}\right))$. Also let α be the simple, positive, short root in C_2 or G_2. Then for $\Phi = C_2$ or G_2, any normal subgroup N containing

$$\{\varepsilon_\phi(2x) | x \in R, \phi \in \Phi\} \cup \{\varepsilon_\alpha(r_i) | i = 1, \ldots, r\}$$

agrees with $\mathrm{Sp}_4(R)$ or $G_2(R)$ respectively.

Proof. According to Theorem 6.1.2, the groups $\mathrm{Sp}_4(R)$ and $G_2(R)$ are generated by root elements. Thus it suffices to show that N contains all root elements. Hence according to Lemma 3.4.2(3) and Lemma 3.5.4(2) it suffices to show that N contains the set $\{\varepsilon_\alpha(x) | x \in R\}$.

Let $\bar{R} := R/2R$ and define $\bar{R}_0 := \{a + 2R | \exists b \in a + 2R : \varepsilon_\alpha(b) \in N\}$. So to prove the proposition, it suffices to show that $\bar{R}_0 = \bar{R}$. We prove $\bar{R}_0 = \bar{R}$ in three steps. First, we show that $\bar{R}_0$ is closed under addition. Second, we show that $\bar{R}_0$ is closed under multiplication with units of $\bar{R}$. Then, we deduce that these two steps imply $\bar{R}_0 = \bar{R}$. For simplicity, we will restrict ourselves to the case of $\mathrm{Sp}_4(R)$ to show the three steps.

The first step is clear, because N is a subgroup of $\mathrm{Sp}_4(R)$. For the second step, let $\varepsilon_\alpha(b)$ be an element of N and let $u \in R$ be given such that $u + 2R$ is a unit in $\bar{R}$. Then we can pick a $v \in R$ such that $v + 2R$ is the inverse of $u + 2R$. Then observe that the following is an element of the normal subgroup N :

$$\begin{aligned}
&\varepsilon_\beta(v)\varepsilon_{-\beta}(-u)w_\beta(1)^{-1}\varepsilon_\alpha(b)w_\beta(1)\varepsilon_{-\beta}(u)\varepsilon_\beta(-v) \\
&= \varepsilon_\beta(v)\varepsilon_{-\beta}(-u)\varepsilon_{\alpha+\beta}(\pm b)\varepsilon_{-\beta}(u)\varepsilon_\beta(-v) \\
&= \varepsilon_\beta(v)\varepsilon_{\alpha+\beta}(\pm b)\varepsilon_\alpha(\pm bu)\varepsilon_{2\alpha+\beta}(\pm b^2 u)\varepsilon_\beta(-v) \\
&= \varepsilon_{\alpha+\beta}(\pm b)\varepsilon_\alpha(\pm bu)\varepsilon_{\alpha+\beta}(\pm bvu)\varepsilon_{2\alpha+\beta}(\pm b^2 u \pm b^2 vu^2) \\
&= \varepsilon_\alpha(\pm bu)\varepsilon_{\alpha+\beta}(\pm bvu \pm b)\varepsilon_{2\alpha+\beta}(\pm b^2 u \pm b^2 vu^2 \pm 2b^2 u).
\end{aligned}$$

Hence to finish the second step it suffices to show that $\pm bvu \pm b$ and $\pm b^2 u \pm b^2 vu^2 \pm 2b^2 u$

are elements of $2R$. To this end, observe first that

$$[\pm bvu \pm b] + 2R = [b \cdot (vu - 1)] + 2R = [b \cdot (1 - 1)] + 2R = 0 + 2R$$

and hence $\pm bvu \pm b$ is an element of $2R$. Second, note

$$\begin{aligned}[\pm b^2u \pm b^2vu^2 \pm 2b^2u] + 2R &= [b^2u - b^2vu^2] + 2R = [b^2u \cdot (1 - vu)] + 2R \\ &= [b^2u \cdot (1 - 1)] + 2R = 0 + 2R\end{aligned}$$

and hence $\pm b^2u \pm b^2vu^2 \pm 2b^2u$ is an element of $2R$ as well. This finishes the second step.

To show $\bar{R}_0 = \bar{R}$ and thus to finish the proof, observe first that by assumption on r_1,

$$\begin{aligned}r_1 + 2R &= (0 + \mathcal{P}_2^{l_2}, 0 + \mathcal{P}_3^{l_3}, \ldots, 0 + \mathcal{P}_r^{l_r}, r_1 + \mathcal{P}_1^{l_1} \cdot \mathcal{Q}_1^{k_1} \cdots \mathcal{Q}_s^{k_s}) \in \{0\} \times R_1 \\ &\subset (R/[\mathcal{P}_2^{l_2} \cdots \mathcal{P}_r^{l_r}]) \times R_1 = R/2R\end{aligned}$$

holds under the Chinese Remainder Theorem. Further, r_1 maps to a unit r_1' in the ring R_1 by assumption. But by definition of N and $\bar{R}_0$, the element $r_1 + 2R$ is an element of $\bar{R}_0$. Next, let u' be a unit in R_1 and choose a $u \in R$ such that

$$u + 2R = (1 + \mathcal{P}_2^{l_2}, 1 + \mathcal{P}_3^{l_3}, \ldots, 1 + \mathcal{P}_r^{l_r}, u'(r_1')^{-1}) \in (R/[\mathcal{P}_2^{l_2} \cdots \mathcal{P}_r^{l_r}]) \times R_1 = R/2R.$$

holds under the Chinese Remainder Theorem. Obviously $u + 2R$ is a unit in $\bar{R}$ and hence according to the second step

$$r_1u + 2R = (0 + \mathcal{P}_2^{l_2} \cdots \mathcal{P}_r^{l_r}, u'(r_1')^{-1}r_1') = (0 + \mathcal{P}_2^{l_2} \cdots \mathcal{P}_r^{l_r}, u') \in (R/[\mathcal{P}_2^{l_2} \cdots \mathcal{P}_r^{l_r}]) \times R_1$$

is an element of $\bar{R}_0$. But as $u' \in R_1$ is an arbitrary unit, this implies according to the first step, that $\bar{R}_0$ contains the subgroup of R_1 generated by the units of R_1. Yet Lemma 7.2.2 implies that this subgroup is already the entire subgroup R_1. So $\bar{R}_0$ contains the entire subgroup $R_1 = \{0\} \times R_1$ of $R/2R$. Similarly, $\bar{R}_0$ contains all the subgroups $R_2, \ldots, R_r$ and hence the entire ring $R/2R$. □

We can show the first part of Theorem 7.2.1 now.

Proof. For each $k \geq r(R) =: r$ and $\Phi = C_2$ or G_2, we have to construct a normally generating set S of $G(\Phi, R)$ such that $|S| = k$ and $\|G(\Phi, R)\|_S$ is bounded below by $4k + r(R)$ for $\Phi = C_2$ or $2k$ for $\Phi = G_2$. First, assume that 2 is a unit in R. In this case, the necessary condition in Corollary 3.2.8 on a set S to normally generate $\mathrm{Sp}_4(R)$ or $G_2(R)$ reduces to $\Pi(S) = \emptyset$. But this implies that the lower bounds on $\Delta_k(\mathrm{Sp}_4(R))$ and $\Delta_k(G_2(R))$ can be shown in the same manner as in Theorem 7.1.5 in this case. So we may assume that 2 in R is not a unit. Then let the ideal $2R$ in R split into distinct

prime ideals as follows:

$$2R = \left(\prod_{i=1}^{r} \mathcal{P}_i^{l_i}\right) \cdot \left(\prod_{j=1}^{s} \mathcal{Q}_j^{k_j}\right)$$

with $[R/\mathcal{P}_i : \mathbb{F}_2] = 1$ for $1 \leq i \leq r$ and $[R/\mathcal{Q}_j : \mathbb{F}_2] > 1$ for $1 \leq j \leq s$. Next, let c be a multiple of the class number of R greater than all $l_1, \ldots, l_r$. Pick elements $x_1, \ldots, x_r \in R$ such that $\mathcal{P}_i^c = (x_i)$ for all i. Also choose $r+1$ distinct primes $V_{r+1}, \ldots, V_k$ in R which do not agree with any of the $\mathcal{P}_1, \ldots, \mathcal{P}_r, \mathcal{Q}_1, \ldots, \mathcal{Q}_s$. Passing to the powers $V_{r+1}^c, \ldots, V_k^c$ we can find elements $v_{r+1}, \ldots, v_k \in R$ with $V_{r+1}^c = (v_{r+1}), \ldots, V_k^c = (v_k)$. Further, define the following elements for $1 \leq u \leq r$

$$r_u := \left(\prod_{1 \leq i \neq u \leq r} x_i\right) \cdot v_{r+1} \cdots v_k.$$

For $k \geq u \geq r+1$ set

$$r_u := x_1 \cdots x_r \cdot \left(\prod_{r+1 \leq u \neq q \leq k} v_q\right).$$

We consider the set $S := \{\varepsilon_\beta(r_1), \ldots, \varepsilon_\beta(r_k)\}$ in $\mathrm{Sp}_4(R)$ or $S := \{\varepsilon_\alpha(r_1), \ldots, \varepsilon_\alpha(r_k)\}$ in $G_2(K)$. Both cases are quite similar, so we will only write down the case of $\mathrm{Sp}_4(R)$.

Claim 7.2.3.1. *S is a normally generating set of* $\mathrm{Sp}_4(R)$.

Let N be the normal subgroup generated by S. First, note that

$$\Pi(\varepsilon_\beta(r_u)) = \begin{cases} \{\mathcal{P}_1, \ldots, \hat{\mathcal{P}}_u, \ldots, \mathcal{P}_r, V_{r+1}, \ldots, V_k\} & \text{, if } 1 \leq u \leq r \\ \{\mathcal{P}_1, \ldots, \mathcal{P}_r, V_{r+1}, \ldots, \hat{V}_u, \ldots, V_k\} & \text{, if } r+1 \leq u \leq k, \end{cases}$$

where the hat denotes the omission of the corresponding prime. This implies $\Pi(S) = \emptyset$. But then Proposition 3.2.6 implies that $\{\varepsilon_\phi(2x) | x \in R, \phi \in C_2\}$ is contained in N. Further, Lemma 3.4.2(2) implies that N also contains the elements $\varepsilon_\alpha(r_1), \ldots, \varepsilon_\alpha(r_r)$. Next, note that by definition of the $r_1, \ldots, r_r$, each r_i is contained in each $\mathcal{P}_j^{l_j}$ for $j \neq i$ and maps to a unit in $R_i := R/(\mathcal{P}_i^{l_i} \cdot \left(\prod_{j=1}^{s} \mathcal{Q}_j^{k_j}\right))$. Thus N and the $r_1, \ldots, r_r$ satisfy the assumptions of Proposition 7.2.3. Hence Proposition 7.2.3 implies that $N = \mathrm{Sp}_4(R)$. This proves the claim.

Claim 7.2.3.2. *The diameter of $\| \cdot \|_S$ is at least $4k + r(R)$. As $|S| = k$ this proves the first part of the theorem for* $\mathrm{Sp}_4(R)$.

This follows as in the proof of Theorem 7.1.5. Namely, one obtains again that for

$$\pi : \mathrm{Sp}_4(R) \to \left(\prod_{i=1}^{r} \mathrm{Sp}_4(R/\mathcal{P}_i)\right) \times \left(\prod_{j=r+1}^{k} \mathrm{Sp}_4(R/V_j)\right)$$

the image $\pi(S)$ is a normally generating set such that the only non-trivial component of $\pi(\varepsilon_\beta(r_u))$ is the $\mathrm{Sp}_4(R/\mathcal{P}_u)$-component if $u \leq r$ and the $\mathrm{Sp}_4(R/V_u)$-component if $u \geq r+1$. As in the proof of Theorem 7.1.5 this implies

$$\begin{aligned}\|\mathrm{Sp}_4(R)\|_S &\geq \| \left(\prod_{i=1}^{r} \mathrm{Sp}_4(R/\mathcal{P}_i)\right) \times \left(\prod_{j=r+1}^{k} \mathrm{Sp}_4(R/V_j)\right) \|_{\pi(S)} \\ &= \left(\sum_{i=1}^{r} \|\mathrm{Sp}_4(R/\mathcal{P}_i)\|_{\varepsilon_\beta(r_i+\mathcal{P}_i)}\right) + \left(\sum_{j=r+1}^{k} \|\mathrm{Sp}_4(R/V_j)\|_{\varepsilon_\beta(r_j+V_j)}\right).\end{aligned}$$

But Proposition 7.1.3(3) implies for $j = r+1, \ldots, k$ that $\|\mathrm{Sp}_4(R/V_j)\|_{\varepsilon_\beta(r_j+V_j)} \geq 4$ holds. Further, $R/\mathcal{P}_i$ is isomorphic to $\mathbb{F}_2$ for all $i = 1, \ldots, r$ and so Proposition B.0.3 implies $\|\mathrm{Sp}_4(R/\mathcal{P}_i)\|_{\varepsilon_\beta(r_i+\mathcal{P}_i)} = \|\mathrm{Sp}_4(\mathbb{F}_2)\|_{\varepsilon_\beta(1)} = 5$ for all $i = 1, \ldots, r$. Hence we obtain

$$\begin{aligned}\|\mathrm{Sp}_4(R)\|_S &\geq \left(\sum_{i=1}^{r} \|\mathrm{Sp}_4(R/\mathcal{P}_i)\|_{\varepsilon_\beta(r_i+\mathcal{P}_i)}\right) + \left(\sum_{j=r+1}^{k} \|\mathrm{Sp}_4(R/V_j)\|_{\varepsilon_\beta(r_j+V_j)}\right) \\ &\geq 5r + 4(k-r) = 4k + r\end{aligned}$$

and so the the first part of Theorem 7.2.1 follows. □

For the second part of Theorem 7.2.1, note the following:

Lemma 7.2.4. *There is an epimorphism $\mathrm{Sp}_4(\mathbb{F}_2) \to \mathbb{F}_2$ with $\varepsilon_\phi(a) \mapsto a$ for all $a \in \mathbb{F}_2$ and $\phi \in C_2$. Similarly, there is an epimorphism $G_2(\mathbb{F}_2) \to \mathbb{F}_2$ with*

$$\varepsilon_\phi(a) \mapsto \begin{cases} a, & \text{if } \phi \in G_2 \text{ short} \\ 0, & \text{if } \phi \in G_2 \text{ long} \end{cases}$$

Proof. The epimorphism $G_2(\mathbb{F}_2) \to \mathbb{F}_2$ with the required properties is constructed in the proof of Lemma 6.3.1. An epimorphism $\mathrm{Sp}_4(\mathbb{F}_2) \to \mathbb{F}_2$ as required is obtained from Proposition B.0.1: The isomorphism $\theta : \mathrm{Sp}_4(\mathbb{F}_2) \to S_6$ maps all root elements in $\mathrm{Sp}_4(\mathbb{F}_2)$ to an odd number of transpositions in S_6. Hence composing with the sign homomorphism $S_6 \to \mathbb{F}_2$, we find an epimorphism $\mathrm{Sp}_4(\mathbb{F}_2) \to \mathbb{F}_2$ with the required properties. □

Remark 7.2.5. The group $G_2(\mathbb{F}_2)$ has a simple subgroup U with $[G_2(\mathbb{F}_2) : U] = 2$. The subgroup U is isomorphic to the twisted finite group of Lie type ${}^2A_2(\mathbb{F}_9)$.

Using this lemma, the second part of Theorem 7.2.1 follows:

Proof. We restrict ourselves to the case $\mathrm{Sp}_4(R)$ again. Let $2R = (\prod_{i=1}^{r} \mathcal{P}_i^{l_i})(\prod_{j=1}^{s} \mathcal{Q}_j^{k_j})$ be given as in the proof of the first part of Theorem 7.2.1. Using the Chinese Remainder

Theorem, we know that the map

$$\mathrm{Sp}_4(R) \twoheadrightarrow \mathrm{Sp}_4(R/2R) = \prod_{i=1}^{r} \mathrm{Sp}_4(R/(\mathcal{P}_i^{l_i})) \times \prod_{i=1}^{r} \mathrm{Sp}_4(R/(\mathcal{Q}_j^{k_j})) \twoheadrightarrow \prod_{i=1}^{r} \mathrm{Sp}_4(R/\mathcal{P}_i) = \mathrm{Sp}_4(\mathbb{F}_2)^r$$

is an epimorphism. So composing with the epimorphism $\mathrm{Sp}_4(\mathbb{F}_2) \to \mathbb{F}_2$, we obtain an epimorphism $g : \mathrm{Sp}_4(R) \to \mathbb{F}_2^r$. This suffices to prove the second part of the theorem, because a given normally generating set S of $\mathrm{Sp}_4(R)$ with $|S| \leq r-1$ would map to a generating set of the abelian group $\mathbb{F}_2^r$ with less than r elements. The group $\mathbb{F}_2^r$ cannot be generated by less than r elements however. □

Remark 7.2.6. If R is a ring of S-algebraic integers with $r(R) = 0$, then the boundedness properties of $\mathrm{Sp}_4(R)$ and $G_2(R)$ are the same as for other $G(\Phi, R)$. For example, Theorem 3.3.1 is in fact also valid for $\Phi = C_2$ and G_2, if the ring R in question does not admit $\mathbb{F}_2$ as a quotient ring, which for rings of S-algebraic integers R can be easily seen to be equivalent to $r(R) = 0$. Consequently, it might seem possible to apply Theorem 3.3.1 instead of Theorem 3.4.1 directly in the proof of Theorem 3.2.1 for such rings R.

However, the condition on a ring R to not admit an ideal of index 2 is hard to formulate in first-order terms and we believe it is not a first-order property at all. However, for each ring R with an ideal I with $(R : I) = 2$, the ideal $2R$ must be contained in I. Hence, one could instead add additional properties that make it impossible for $R/2R$ to have the field $\mathbb{F}_2 = R/I$ as a quotient. For example, if $R/2R$ is finite, then one can describe the ring structure of $R/2R$ explicitly as a given direct product of finite, local non-reduced rings with residue fields bigger than $\mathbb{F}_2$. This is a first-order property.

This finishes the proof of Theorem 7.2.1. We note the following corollary:

Corollary 7.2.7. *Let R be a ring of S-algebraic integers and $r = r(R)$ defined as in Theorem 7.2.1. Then both $\mathrm{Sp}_4(R)$ and $G_2(R)$ have abelianization $\mathbb{F}_2^r$.*

Proof. We only do the case $\mathrm{Sp}_4(R)$. Note that $\langle\langle \varepsilon_\phi(2x) | x \in R, \phi \in C_2 \rangle\rangle \subset (\mathrm{Sp}_4(R), \mathrm{Sp}_4(R))$ by Lemma 3.4.2(4) and (2) and further that $\mathrm{Sp}_4(R)$ is boundedly generated by root elements by Theorem 6.1.2. Thus the abelianization $A(R)$ of $\mathrm{Sp}_4(R)$ is a finitely generated, 2-torsion group. Let $r' := dim_{\mathbb{F}_2}(A(R))$. The proof of Theorem 7.2.1 implies that $A(R)$ has the quotient $\mathbb{F}_2^r$ and hence $r' \geq r$. On the other hand, $r' > r$ is impossible, because it would imply, as in the proof of the second part of Theorem 7.2.1, that there are no normally generating sets of $\mathrm{Sp}_4(R)$ with precisely r elements, which we have seen to not be the case when proving the first part of Theorem 7.2.1. □

We call the minimal number of conjugacy classes of a group G, that can generate said group its *weight* $w(G)$. Then obviously $w(G/[G,G]) \leq w(G)$ holds for all groups G that can be generated by finitely many conjugacy classes.

One notes that for $r(R) \geq 1$, according to Corollary 7.2.7 the minimal number of group elements needed to generate the abelianization of $\mathrm{Sp}_4(R)$ or $G_2(R)$ is $r(R)$ and according to Theorem 7.2.1 the minimal number of conjugacy classes that can generate $\mathrm{Sp}_4(R)$ or $G_2(R)$ is $r(R)$ as well. We note the following problem:

Conjecture 7.2.8. *Let G be a non-perfect group such that it can be generated by a finite set of conjugacy classes. Then $w(G) = w(G/[G,G])$ holds.*

This conjecture is related to an old problem posed by Wiegold, we were told about by Alexander Lubotzky, asking whether there are perfect groups which cannot be generated by a single conjugacy class. Relatively little seems to be known about this problem and Conjecture 7.2.8 in general. Chiodo [12] gives a rather complete account of the groups G for which Conjecture 7.2.8 is known to hold, most prominently solvable and finite groups. In light of Corollary 7.2.7 and Theorem 7.2.1, we propose the more specialized conjecture that Conjecture 7.2.8 also holds for general non-perfect arithmetic lattices.

For rings of quadratic integers it is known how 2 splits into primes and hence we can give the following complete description of $r(R)$:

Corollary 7.2.9. *Let D be a square-free integer and R the ring of algebraic integers in $\mathbb{Q}[\sqrt{D}]$. Then*

1. *$r(R) = 1$ holds precisely if $D \equiv 2, 3, 5, 6, 7 \bmod 8$, so $\Delta_1(\mathrm{Sp}_4(R)), \Delta_1(G_2(R)) \neq -\infty$.*

2. *$r(R) = 2$ holds precisely if $D \equiv 1 \bmod 8$, so $\Delta_1(\mathrm{Sp}_4(R)) = \Delta_1(G_2(R)) = -\infty$ and $\Delta_2(\mathrm{Sp}_4(R)) = \Delta_2(G_2(R)) > -\infty$.*

Proof. We obtain from [28, Theorem 25] that the ideal $2R$ splits and ramifies in R as follows:

1. $2R$ is inert precisely if $D \equiv 5 \bmod 8$.

2. $2R$ ramifies precisely if $D \equiv 2, 3, 6, 7 \bmod 8$.

3. $2R$ splits precisely if $D \equiv 1 \bmod 8$.

In the first two cases, this implies $r(R) = 1$ and in the third case $r(R) = 2$. □

Chapter 8

Straightforward generalizations, open questions and closing remarks

In this chapter, we talk about a possible generalization of our results in the first section and talk a little more about improving and generalizing our results in the second section.

8.1 A straightforward generalization of Theorem 3.1.2

In this section, we show the following:

Theorem 8.1.1. *Let R be a ring of S-algebraic integers and let H be a subgroup of finite index in $\mathrm{SL}_n(R)$ for $n \geq 3$. Then there is a constant $C(H) \in \mathbb{N}$ such that $\Delta_k(H) \leq C(H)k$ holds for all $k \in \mathbb{N}$.*

The strategy is quite similar to the strategy to prove strong boundedness for $\mathrm{Sp}_4(R)$. First, one shows that a certain fixed subgroup of finite index in H is bounded and then one shows how to get from this group to the entire group H. First, we need the following definition:

Definition 8.1.2. Let R be a commutative ring with 1, I an ideal in R, $n \geq 2$. Then define the following subgroups of $\mathrm{GL}_n(R)$:

1. $E(n,R,I) := \langle A(I_n + te_{i,j})A^{-1} |\ 1 \leq i \neq j \leq n, A \in \mathrm{GL}_n(R), t \in I \rangle$ and
2. $C(n,R,I) := \langle A \in \mathrm{GL}_n(R) |\ \pi_I(A) = I_n \rangle$.

Further, for $A \in \mathrm{SL}_n(R)$ define the word norm $\| \cdot \|_{A,I} : \mathrm{SL}_n(R) \to \mathbb{N}_0 \cup \{+\infty\}$ by $\|I_n\|_{A,I} := 0$ and

$$\|X\|_{A,I} := \min\{m \in \mathbb{N}_0 \cup \{+\infty\} |\ \exists Y_1, \ldots, Y_m \in E(n,R,I), e_1, \ldots, e_m \in \{1,-1\} : X = \prod_{i=1}^{m} Y_i A^{e_i} Y_i^{-1}\}.$$

for $X \in \mathrm{SL}_n(R) - \{I_n\}$.

Then the following holds:

Theorem 8.1.3. *[44, Theorem 2] Let R be a commutative ring with 1, I an ideal in R, $n \geq 3$ and H a subgroup of $\mathrm{GL}_n(R)$ normalized by the subgroup $E(n, R, I)$. Then there is an ideal J in R such that*

$$E(n, R, I^5 J) \subset H \subset C(n, R, J).$$

Using this, one can prove the following preliminary proposition:

Proposition 8.1.4. *Let R be a commutative ring with 1, I a finitely generated ideal in R and $n \geq 3$. Then there is a constant $C_n(I)$ such that for all $A \in \mathrm{SL}_n(R)$, one has*

$$\|I_{2n} + xe_{1n}\|_{A,I} \leq C_n(I)$$

for all $x \in I^5 l(A)$ with $C_n(I)$ depending on the number of generators of I and n.

Proof. This will be proven by a compactness argument similar to the ones in Chapter 3 but employing Theorem 8.1.3. First, let $1 \leq u \neq v \leq n$ be given and let a language $\mathcal{L}$ with the relation symbols, constants and function symbols

$$(\mathcal{R}, 0, 1, +, \times, (a_{i,j})_{1 \leq i,j \leq n}, c, s_1, \ldots, s_k, (t_M)_{\{M \subset \{1,\ldots,k\} \text{ and } |M|=5\}}, {}^{-1})$$

be given, where $\mathcal{A} := (a_{i,j})_{1 \leq i,j \leq n}$ is an $n \times n$-matrix of constant symbols, $s_1, \ldots, s_k, c$ and $(t_M)_{\{M \subset \{1,\ldots,k\} \text{ and } |M|=5\}}$ are constant symbols, $+, \times : \mathcal{R} \times \mathcal{R} \to \mathcal{R}$ are function symbols. Further, $\cdot^{-1} : \mathcal{R}^{n \times n} \to \mathcal{R}^{n \times n}$ is another function symbol and we will often write X^{-1} for ${}^{-1}(X)$. Next, we describe a first-order theory $\mathcal{T}_{uv}$, which contains the following sentences:

1. Sentences forcing the universe $R := \mathcal{R}^{\mathcal{M}}$ of each model $\mathcal{M}$ of $\mathcal{T}_{uv}$ to be a commutative ring with respect to the functions $+^{\mathcal{M}}, \times^{\mathcal{M}}$ and with $0^{\mathcal{M}}, 1^{\mathcal{M}}$ being 0 and 1.

2. The sentence $\forall X \in \mathcal{R}^{n \times n} : (\det(X) = 1) \to (XX^{-1} = I_n)$, where I_n denotes the unit matrix in $\mathcal{R}^{n \times n}$ with entries the constant symbols $0, 1$ as appropriate.

3. $\det(\mathcal{A}) = 1$.

4. $c = a_{u,v} \sum_{M \subset \{1,\ldots,k\} \text{ and } |M|=5} t_M (\prod_{p \in M} s_p)$

5. A family of sentences $(\theta_r)_{r \in \mathbb{N}}$ as follows:

$$\begin{aligned}\theta_r :& \forall X_1, \ldots, X_r, \forall b_1^{(1)}, \ldots, b_k^{(1)}, b_1^{(2)}, \ldots, b_k^{(2)}, \ldots, b_1^{(r)}, \ldots, b_k^{(r)}, \forall e_1, \ldots, e_r \in \{0, 1, -1\} : \\ & (\det(X_1) = \cdots = \det(X_r) = 1) \to \\ & \left[(I_n + ce_{1n}) \neq (\mathcal{A}^{e_1})^{X_1 \left(I_n + e_{1,n} \sum_{p=1}^k b_p^{(1)} s_p\right) X_1^{-1}} \cdots (\mathcal{A}^{e_r})^{X_r \left(I_n + e_{1,n} \sum_{p=1}^k b_p^{(r)} s_p\right) X_r^{-1}} \right]\end{aligned}$$

Here $\mathcal{A}^1 := \mathcal{A}, \mathcal{A}^{-1} := \mathcal{A}^{-1}$ and $\mathcal{A}^0 := I_n$.

Next, let us show that the theory $\mathcal{T}_{uv}$ is inconsistent. If $\mathcal{M}$ were a model of $\mathcal{T}_{uv}$ and $R := (\mathcal{R})^{\mathcal{M}}$ would be its universe, than R is a commutative ring with 1 according to the sentences in (1) and $(a_{ij}^{\mathcal{M}})$ is an element of $\mathrm{SL}_n(R)$ according to the sentence in (2). We will abuse notation and denote $(a_{ij}^{\mathcal{M}}) = \mathcal{A}^{\mathcal{M}}$ by A.

Next, setting I as the ideal in R generated by the elements $s_1^{\mathcal{M}}, \ldots, s_k^{\mathcal{M}}$, we know according to Theorem 8.1.3, that for the subgroup H generated by the set

$$\{XAX^{-1} |\ X \in E(n, R, I)\},$$

there is an ideal J such that

$$E(n, R, I^5 J) \subset H \subset C(n, R, J).$$

However A is an element of H and thus J must contain $l(A)$ and hence $E(n, R, I^5 l(A)) \subset H$ holds. But $c^{\mathcal{M}}$ is an element of $I^5 l(A)$ according to (4). This implies that $I_n + c^{\mathcal{M}} e_{1n}$ is an element of H. But this in turn implies that there must be elements $X_1, \ldots, X_{r'} \in E(n, R, I)$ and $e_1, \ldots, e_{r'} \in \{1, -1\}$ such that

$$I_n + c^{\mathcal{M}} e_{1n} = (A^{e_1})^{X_1} \cdots (A^{e_{r'}})^{X_{r'}}.$$

But for $i = 1, \ldots, r'$ each X_i can be written as a product of m_i factors of the form $Z(I_n + y e_{1n}) Z^{-1}$ for $y \in I$ and $Z \in \mathrm{SL}_n(R)$ by definition of $E(n, R, I)$ for some m_i. Hence for $r := m_1 + \cdots + m_{r'}$, we obtain a contradiction to θ_r in (5). Thus $\mathcal{T}_{uv}$ is inconsistent.

Gödel's Compactness Theorem [37, Theorem 3.2] implies then, that a certain finite subset $\mathcal{T}_{uv}^0 \subset \mathcal{T}_{uv}$ is already inconsistent. Hence there is only a finite collection of the θ_r contained in $\mathcal{T}_{uv}^0$. So let $L_{uv} \in \mathbb{N}$ be the largest $r \in \mathbb{N}$ with $\theta_r \in \mathcal{T}_{uv}^0$. For all $r \in \mathbb{N}$, we have $\{(1) - (4), \theta_{r+1}\} \vdash \theta_r$. Hence the subset $\mathcal{T}_{uv}^1 \subset \mathcal{T}_{uv}$ that contains all sentences in (1) through (4) and the *single* sentence $\theta_{L_{uv}}$, must be inconsistent as well.

Next, let R be a commutative ring with 1, I an ideal in R generated by the elements $s_1, \ldots, s_k \in R, A = (a_{ij}) \in \mathrm{SL}_n(R)$ and c an element of the ideal $(a_{uv}) I^5$. This gives us a model $\mathcal{M}$ of (1) through (4) and hence as $\mathcal{T}_{uv}^1$ is inconsistent, this model must violate the sentence $\theta_{L_{uv}(\Phi)}$. But this implies the existence of elements $X_1, \ldots, X_{L_{uv}} \in \mathrm{SL}_n(R)$ as well as the existence of $y_1, \ldots, y_{L_{uv}} \in I$ such that

$$I_n + c e_{1n} = (A^{e_1})^{X_1 (I_n + y_1 e_{1n}) X_1^{-1}} \cdots (A^{e_{L_{uv}}})^{X_{L_{uv}} (I_n + y_{L_{uv}} e_{1n}) X_{L_{uv}}^{-1}}.$$

Hence

$$\| I_n + c e_{1n} \|_{A,I} \leq L_{uv}$$

holds for all $c \in (a_{uv})I^5$. Varying u and v as appropriate one obtains for a sufficiently large $C_n(I) \in \mathbb{N}$ that

$$\|I_n + ce_{1n}\|_{A,I} \leq C_n(I)$$

holds for all $c \in l(A)I^5$. □

From this, one can prove Theorem 8.1.1:

Proof. Let $S = \{A_1, \ldots, A_k\}$ be a normally generating set of H.

Observe that as H has finite index in $\mathrm{SL}_n(R)$, there is a normal subgroup N of $\mathrm{SL}_n(R)$ with finite index in $\mathrm{SL}_n(R)$ contained in H. But if N has finite index in $\mathrm{SL}_n(R)$, then the subgroup

$$I_1 := \{x \in R | I_n + xe_{1n} \in N\}$$

of R must have finite index in R. Furthermore, I_1 is an ideal in R as seen for example from the proof of Lemma 3.3.3. Hence $E(R, n, I_1)$ is a subgroup of N and H. But then Proposition 8.1.4 implies that the following inequality

$$\|I_n + xe_{1n}\|_{A_i,I_1} \leq C_n(I_1)$$

holds for all $x \in l(A_i)I_1^5$ and $i = 1, \ldots, k$, because I_1 is finitely generated as an ideal, as R is noetherian. But this inequality implies that

$$\|I_n + xe_{1n}\|_S \leq C_n(I_1)k$$

holds for all $x \in (l(A_1) + \cdots + l(A_k))I_1^5$.

However, observe that I_1 must be contained in $l(A_1) + \cdots + l(A_k) := J$, because clearly $\pi_J(A)$ must be scalar for each $A \in H$. Thus the ideal I_1^6 is contained in the ideal $(l(A_1) + \cdots + l(A_k))I_1^5$. Hence

$$\|I_n + xe_{1n}\|_S \leq C_n(I_1)k$$

holds for all $x \in I_1^6 =: I_2$. Next, define the subgroup $\widetilde{H}$ of H generated by

$$Q := \{A(I_n + xe_{1n})A^{-1} |\ A \in H, x \in I_2\}.$$

Obviously, Q induces a conjugation generated word norm $\|\cdot\|_Q$ on $\widetilde{H}$. Next, observe that as H contains the subgroup $E(n, R, I_1)$, one obtains from Theorem 8.1.3 that there is an ideal I_3 in R such that

$$E(n, R, I_2^5 I_3) \subset \widetilde{H} \subset C(n, R, I_3).$$

But I_3 cannot be trivial, because $\widetilde{H}$ contains non-central elements. Hence $\widetilde{H}$ contains the finite index subgroup $E(n, R, I_2^5 I_3)$ of $\mathrm{SL}_n(R)$ and hence $\widetilde{H}$ itself has finite index in

$\mathrm{SL}_n(R)$. But this implies according to a Theorem by myself, Kędra and Gal [19, Theorem], that $\widetilde{H}$ is bounded and hence there is a natural number $M(R, I_2)$ such that

$$\|\widetilde{H}\|_Q \leq M(R, I_2).$$

But each element X in Q satisfies $\|X\|_S \leq C_n(I_1)k$ and hence

$$\|\widetilde{H}\|_S \leq M(R, I_2)C_n(I_1)k \tag{8.1}$$

holds.

Further, observe that $\widetilde{H}$ is a normal subgroup of finite index of H. Thus the group $G := H/\widetilde{H}$ is finite. Let $\pi : H \to G$ be the corresponding quotient map and set

$$\Sigma(G) := \{T \subset G |\ T \text{ normally generates } G\}.$$

Then $\Sigma(G)$ is finite as G is finite and hence there is a $L(G) \in \mathbb{N}$ such that for each $T \in \Sigma(G)$, one has for the corresponding conjugation generated word norm $\|\cdot\|_T$ on G induced by T that

$$\|G\|_T \leq L(G).$$

But S normally generates H and hence $\pi(S)$ is an element of $\Sigma(G)$. Thus $\|G\|_{\pi(S)} \leq L(G)$ holds. Hence for $A \in H$ there are $X_1, \ldots, X_{L(G)} \in H$ and $B_1, \ldots, B_{L(G)} \in S \cup S^{-1} \cup \{I_n\}$ with $\pi(\prod_{i=1}^{L(G)} B_i^{X_i}) = \pi(A)$. Hence $A \left(\prod_{i=1}^{L(G)} B_i^{X_i}\right)^{-1} \in \widetilde{H}$ holds and thus (8.1) implies

$$\|A\|_S \leq \|\prod_{i=1}^{L(G)} B_i^{X_i}\|_S + \|\widetilde{H}\|_S \leq \sum_{i=1}^{L(G)} \|B_i\|_S + M(R, I_2)C_n(I_1)k.$$

However as each B_i is an element of $S \cup S^{-1} \cup \{I_n\}$, one obtains

$$\|A\|_S \leq L(G) + M(R, I_2)C_n(I_1)k$$

for all $A \in H$ and thus the theorem is proven, because the constants $L(G), M(R, I_2)$ and $C_n(I_1)$ depend on the ideals I_1 and $I_2 = I_1^6$ in R and the ideal I_1 does not depend on the set S but only on the subgroup H. □

Remark 8.1.5. The main problem in proving strong boundedness for a finite index subgroup H of arithmetic Chevalley groups is that even if the group H is normal, one cannot just study the normal subgroup generated by conjugacy classes with respect to $\mathrm{SL}_n(R)$ as we did in Chapter 3, because two elements of H might be conjugate with respect to $\mathrm{SL}_n(R)$ without being conjugate in H itself. This makes it necessary to study normal subgroups of H instead or in other words, subnormal subgroups of $\mathrm{SL}_n(R)$.

8.2 Possible generalizations and potential future research

Generalizing strong boundedness to other groups

There are four clear avenues to generalize the statements about strong boundedness of this thesis to other groups:

The first and most obvious one is to find other rings R that also satisfy the bounded generation assumption by root elements property for $G(\Phi, R)$ with Φ an irreducible root system of rank at least 2. A clear candidate for such rings are rings of S-algebraic integers in global fields of positive characteristic. As mentioned before, Nica has shown in [34] that $\mathrm{SL}_n(\mathbb{F}[T])$ is boundedly generated for $\mathbb{F}$ a finite field and it seems quite likely to me that this holds for all rings of S-algebraic integers in global fields of positive characteristic. Luckily enough, there has been quite a lot of research in the characteristic 0-case already and it should be possible to find such generalizations with similar arguments.

The second possible avenue is to consider the last remaining root system $\Phi = A_1$ and $\mathrm{SL}_2(R)$ for suitable rings R. For example, one could attempt to prove Conjecture 6.5.1. The structure of normal subgroups of $\mathrm{SL}_2(R)$ is more complicated than the one of the higher rank Chevally groups. Yet normal subgroups of $\mathrm{SL}_2(R)$ for R a Dedekind domains with infinitely many units have been described completely by Costa and Keller [13] in terms of so-called radices. It seems likely to me that the validity of the classification results in [13] can be shown under certain first order conditions as well and not only under the assumption that R is a Dedekind domain with infinitely many units. This would enable one to apply a compactness argument in a similar manner as done in Chapter 3 to obtain certain root elements. Then one could probably finish the proof in a similar manner as done in the case of $\mathrm{Sp}_4(R)$ and $G_2(R)$ using bounded generation results for $\mathrm{SL}_2(R)$ like Theorem 6.1.5. Obviously, this second avenue, if successful could likely also work for R a ring of S-algebraic integers in a global field of positive characteristic with infinitely many units. This however would require to show a version of Theorem 6.1.5 in this case.

The third avenue for generalizations is to consider other arithmetic groups entirely. I have shown such a result in Theorem 8.1.1, but the proof of Theorem 8.1.1 also highlights the problem with the proof strategy as I presented it: My strategy requires to understand the normal subgroup structure of the arithmetic group in question and I am not aware of general results of this form. Furthermore, the corresponding result about normal subgroups would have to be rephrased in first-order terms somehow as to enable the application of a compactness argument and I am doubtful that this is always possible. But using Bak's concept of form ideals and form rings, one can describe the subnormal structure of certain other matrix groups, mostly higher rank symplectic and even orthogonal groups [47] and so one can show some further generalizations of my results for these certain special cases. Ultimately though, the strategy of understanding the normal subgroup structure does not seem the most promising to me and I would like to prove strong

boundedness results from results about arithmetic lattices directly. I should also mention, that lattices can in general not be uniformly bounded as observed by Kedra, Libman and Martin [24, Theorem 5.5]

Fourth, twisted arithmetic Chevalley groups, which according to [42] are also boundedly generated by root elements, might admit strong boundedness statements as well, but I am not aware of a description of normal subgroups of these groups.

Last, there is the problem of the existence of normally generating subsets. As mentioned before, Theorem 7.2.1 and Corollary 7.2.7 imply together that for R a ring of S-algebraic integers in a number field, the minimal number $w(\mathrm{Sp}_4(R))$ of conjugacy classes needed to generate $\mathrm{Sp}_4(R)$, agrees with

$$w(\mathrm{Sp}_4(R)/(\mathrm{Sp}_4(R), \mathrm{Sp}_4(R))),$$

if the latter number is at least 1. I suspect that the same holds for more general arithmetic lattices. For example, I think that for R a ring of S-algebraic integers with infinitely many units, one still has $w(\mathrm{SL}_2(R)) = \max\{1, r(R)\}$. Note in this context that using Margulis Superrigidity [31, Theorem 16.1.11], one can show that the abelianization of many arithmetic lattices is always finite though.

Asymptotics of strong boundedness

Roughly, speaking my strategy to prove strong boundedness for arithmetic Chevalley groups consisted of first deconstructing a given finite set of conjugacy classes to obtain a sufficiently 'large' subgroup of root elements in a ball of finite diameter with respect to the corresponding conjugation generated norm and then reconstructing arbitrary elements of the group using bounded generation results. However, both of these steps seem to require linear in the rank of the root systems many factors. Together, this results in upper bounds on $\Delta_k(G(\Phi, R))/k$ quadratic in the rank of Φ as seen in Corollary 7.1.7. However, results for covering numbers [25, Theorem 1] as well as Corollary 7.1.6 and Theorem 5.2.7 for the semi-local case indicate an asymptotic linear in the rank of Φ and I believe that the true asymptotic of $\Delta_k(G(\Phi, R))/k$ should be linear in the rank of Φ even in the case of R a ring of S-algebraic integers.

To show this however, one would have to explain how to write arbitrary elements of $G(\Phi, R)$ as products of the initial normally generating set S from the start without the detour of root elements or explain how one can for $A \in G(\Phi, R)$, write root elements with arguments in $l(A)$ as products of conjugates of A with a number of factors that does not depend on the rank of Φ. The latter is reasonably easy to do if R is a field, however for R a ring of S-algebraic integers one arrives at a bit of an impasse, because of the absence of a neat decomposition like the Bruhat decomposition for fields. Furthermore, excluding the case $G_2(\mathbb{F}_2)$, for a field it is enough to find a single none-trivial root element to get

all the other ones and this is not the case for more general rings.

If R is at least a principal ideal domain, one can actually use the Bruhat decomposition [41, Chapter 8, p. 68, Corollary 1] that I use in Chapter 4 to cut down on the number of root elements needed to write a group element in a similar manner as in the field case and I will do so in a future paper. However, in contrast to the field or the semi-local case this is not enough, because one must also accumulate enough root elements such that their respective arguments are coprime.

Ultimately, the asymptotics for Δ_k provided in this thesis are difficult to improve outright for Chevalley groups, not only because the underlying ring might fail to be a principal ideal domain, which is a minor issue, but also because one needs to potentially involve all entries of a given element of a normally generating set and cannot focus on a number of entries independent of the rank of Φ as was done in the proof of Theorem 5.2.7 and more research is needed to resolve these questions.

Appendix A

Root systems and Weyl groups

This section is very similar to the appendices of Humphreys book [21] and Steinbergs book [41]. First, root systems are defined as follows:

Definition A.0.1. Let $(V, (\cdot,\cdot))$ be a finite-dimensional, euclidean vector space. Then a subset $\Phi \subset V - \{0\}$ is called a *root system*, if Φ satisfies the following assumptions:

1. The set Φ spans V as a $\mathbb{R}$-vector space.
2. If $\alpha \in \Phi$, then $\mathbb{R}\alpha \cap \Phi = \{\alpha, -\alpha\}$.
3. For any $\alpha, \beta \in \Phi$, the element $w_\alpha(\beta) := \beta - 2\frac{(\alpha,\beta)}{(\alpha,\alpha)}\alpha$ is also an element of Φ.
4. For any $\alpha, \beta \in \Phi$, the number $\langle \beta, \alpha \rangle := 2\frac{(\alpha,\beta)}{(\alpha,\alpha)}$ is an integer.

The elements of Φ are called *roots* and the dimension of V the rank of the root system.

Let Φ be a root system of rank n. A $\mathbb{Z}$-linear independent subset $\Pi = \{\alpha_1, \ldots, \alpha_n\}$ of Φ with the property

$$\Phi = \left(\Phi \cap \bigoplus_{i=1}^{n} \mathbb{N}_0\alpha_i\right) \cup \left(\Phi \cap \bigoplus_{i=1}^{n} (-\mathbb{N}_0)\alpha_i\right)$$

is called a *system of (positive) simple roots*. Fixing a system of simple roots Π in a root system Φ, the elements in $(\Phi \cap \bigoplus_{i=1}^{n} \mathbb{N}_0\alpha_i)$ are called the *positive roots of* Φ, the set usually denoted by Φ^+, and the elements of $(\Phi \cap \bigoplus_{i=1}^{n}(-\mathbb{N}_0)\alpha_i)$ are called *the negative roots of* Φ, the set usually denoted by Φ^-. Furthermore, if $\phi \in \Phi$ is equal to $\sum_{i=1}^{n} k_i\alpha_i$, then $\mathrm{wt}(\phi) := |\sum_{i=1}^{n} k_i|$ is called the *weight of the root* ϕ.

Further, for Φ a root system and $\alpha \in \Phi$, the maps $w_\alpha : V \to V, v \mapsto v - 2\frac{(\alpha,v)}{(\alpha,\alpha)}\alpha$ are isometries of $(V, (\cdot,\cdot))$. The subgroup $W(\Phi)$ of $\mathrm{Isom}(V, (\cdot,\cdot))$ is the group generated by elements of the form $\{w_\alpha | \alpha \in \Phi\}$. The Weyl group $W(\Phi)$ acts on the root system Φ. It is clear that for Π a system of simple roots and $w \in W(\Phi)$, the set $w(\Pi)$ is also a system of simple roots.

Proposition A.0.2. *Let Φ be a root system of rank n.*

1. *Systems of simple roots in Φ exist and the Weyl group $W(\Phi)$ acts simply, transitively on $\{\Pi | \Pi$ a system of simple roots in $\Phi\}$.*

2. *For $\Pi = \{\alpha_1, \ldots, \alpha_n\}$ a system of simple roots, set for $1 \leq i \neq j \leq n$*

$$m_{ij} := \begin{cases} 2, & \text{if } \sphericalangle(\alpha_i, \alpha_j) = \pi/2 \\ 3, & \text{if } \sphericalangle(\alpha_i, \alpha_j) = 2\pi/3 \\ 4, & \text{if } \sphericalangle(\alpha_i, \alpha_j) = 3\pi/4 \\ 6, & \text{if } \sphericalangle(\alpha_i, \alpha_j) = 5\pi/6 \end{cases}$$

and $m_{i,i} = 1$. Then $W(\Phi)$ is generated by $\{w_{\alpha_1}, \ldots, w_{\alpha_n}\}$ and

$$W(\Phi) \cong \langle w_{\alpha_1}, \ldots, w_{\alpha_n} | \forall 1 \leq i \leq j \leq n : (w_{\alpha_i} w_{\alpha_j})^{m_{ij}} = 1 \rangle$$

Remark A.0.3. For $\Pi = \{\alpha_1, \ldots, \alpha_n\}$ a system of simple roots in Φ the reflections $w_{\alpha_1}, \ldots, w_{\alpha_n}$ are called *fundamental reflections* in $W(\Phi)$.

There is an obvious concept of isomorphism of root systems and direct sums of root systems. A root system is called *irreducible* if it is not isomorphic to the direct sum of two non-trivial root systems. A common tool to describe root systems are Dynkin diagrams:

Definition A.0.4. Let Φ be a root system and Π a system of simple roots of Φ. Then the *Dynkin diagram* $D(\Phi)$ of (Φ, Π) is the directed multigraph defined as follows:

1. The vertices of $D(\Phi)$ are the elements of Π.

2. For $\alpha, \beta \in \Pi$ with $\alpha \neq \beta$ the edge $\{\alpha, \beta\}$ is contained in $D(\Phi)$, if $\langle \alpha, \beta \rangle \neq 0$ and the multiplicity of the edge is $|\langle \alpha, \beta \rangle \cdot \langle \beta, \alpha \rangle|$.

3. All edges are undirected except the ones connecting simple roots of unequal length. They are directed to start in longer roots and are marked by arrows.

All systems of simple roots of Φ differ by an element $W(\Phi)$ according to Proposition A.0.2 and hence the isomorphism type of the multigraph $D(\Phi)$ does not depend on the particular system of simple roots Π used to define it. Hence we will usually omit specifying the system of simple roots.

For a semi-simple, complex Lie group G there is an action of its Lie algebra $\mathfrak{g}$ on itself denoted by $\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$. For a maximal abelian subalgebra $\mathfrak{h}$ of $\mathfrak{g}$, the elements of $\mathfrak{h}$ map to diagonalizable elements of $\mathrm{End}(\mathfrak{g})$ under ad and as $\mathfrak{h}$ is abelian they are simultaneously diagonalizable. Phrased differently, $\mathfrak{g}$ decomposes as the direct sum of

simultaneous eigenspaces

$$\mathfrak{g}_\alpha := \{X \in \mathfrak{g} | \forall H \in \mathfrak{h} : \mathrm{ad}(H)(X) = \alpha(H)X\}$$

for certain linear maps $\alpha : \mathfrak{h} \to \mathbb{C}$. Set

$$\Phi := \{\alpha : \mathfrak{h} \to \mathbb{C} | \mathfrak{g}_\alpha \neq \{0\} \text{ and } \alpha \neq 0\}$$

and let V be the subspace of $\mathrm{Hom}(\mathfrak{h}, \mathbb{C})$ generated by Φ as an $\mathbb{R}$-vector space. Further, one defines the Killing-form $B : \mathfrak{g} \times \mathfrak{g} \to \mathbb{C}$ for $X, Y \in \mathfrak{g}$ as follows:

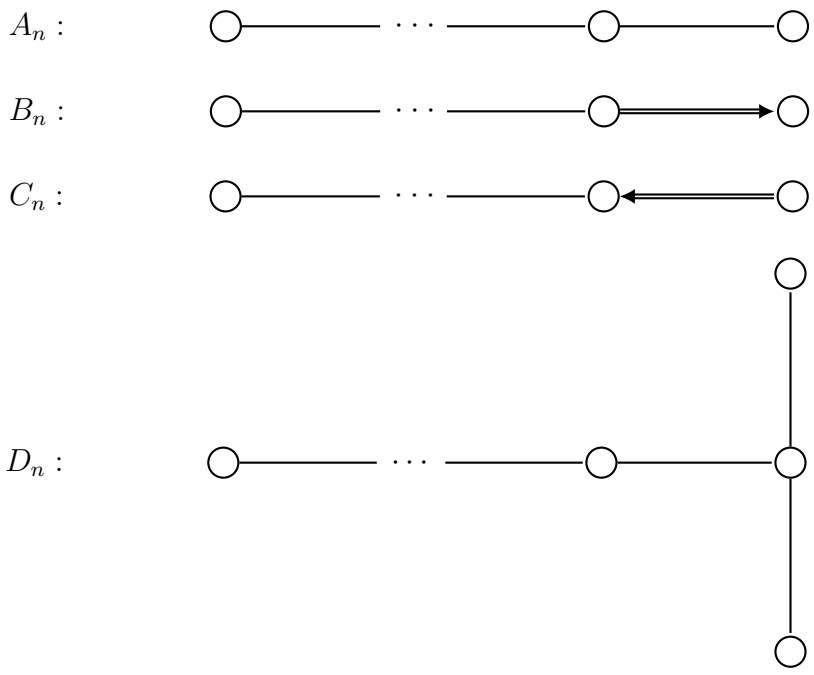
$$B(X, Y) := tr_{\mathbb{C}}^{\mathfrak{g}}(\mathrm{ad}(X)\mathrm{ad}(Y)).$$

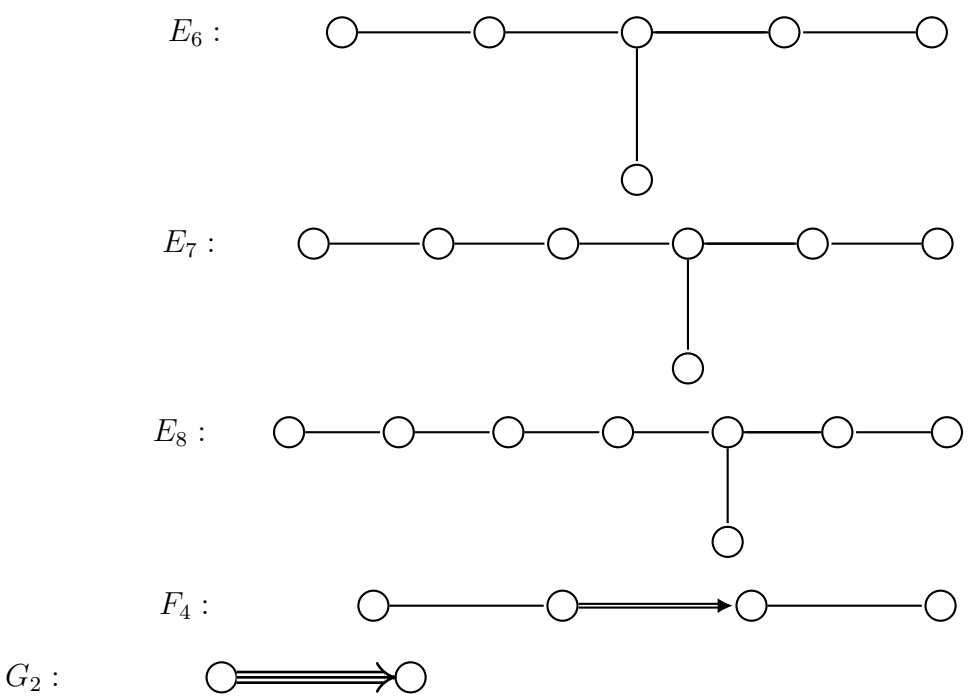
The bilinear form B defines a non-degenerate, symmetric bilinear form on $\mathfrak{h}$. Thus for each $\alpha \in \Phi$ there is a unique $H'_\alpha \in \mathfrak{h}$ such that $B(H, H'_\alpha) = \alpha(H)$ holds for all $H \in \mathfrak{h}$. Then one can define a scalar product on V by setting

$$(\alpha, \beta) := B(H'_\alpha, H'_\beta)$$

for $\alpha, \beta \in \Phi$. This yields Φ as a root system in $(V, (\cdot, \cdot))$ and if G is simple, then Φ is irreducible.

Proposition A.0.5. *An irreducible root system Φ is determined up to isomorphism by its Dynkin diagram $D(\Phi)$. Further, the Dynkin diagram $D(\Phi)$ of any irreducible root systems Φ is one of the following:*





Further each of those diagrams is realized as the Dynkin diagram of the root system Φ *of a simple, complex Lie-group.*

Remark A.0.6. The root systems A_n, B_n, C_n and D_n are commonly referred to as *classical root systems*, because they arise as root systems of familiar complex matrix groups. On the other hand, the root systems E_6, E_7, E_8, F_4 and G_2 are commonly referred to as *exceptional root systems.*

Proposition A.0.7. *Let* Φ *be a root system,* Π *a system of simple roots and* $\alpha \in \Phi$*. Then there is a* $\beta \in \Pi$ *and a* $w \in W(\Phi)$ *with* $w(\alpha) = \beta$*.*

Proposition A.0.8. *Let* Φ *be an irreducible root system,* $\alpha \in \Phi$ *be given and let* $\beta \in \Pi$ *have the same length as* α*. Then there is a* $w \in W(\Phi)$ *with* $w(\alpha) = \beta$*. Phrased differently, the equivalence relation on* Φ *induced by the action of* $W(\Phi)$ *has one equivalence class for each root length present in* Φ*.*

Proof. According to Proposition A.0.7, there is an element $w \in W(\Phi)$ with $w(\alpha) \in \Pi$. Thus we may assume that α is also an element of Π. According to Proposition A.0.5, for two elements of Π of the same length, there is a path in $D(\Phi)$ only passing through elements of Π of the same length. Phrased differently, we may assume that $\Phi = A_n$ for $n \geq 2$. So by induction on n, it suffices to consider the case $\Phi = A_2$ and that the positive roots in $\Phi = A_2$ are α, β and $\alpha + \beta$. The claim now follows as

$$w_\alpha w_\beta(\alpha) = w_\alpha(\alpha + \beta) = \beta$$

holds. □

Appendix B

The permutation group S_6 and $\mathrm{Sp}_4(\mathbb{F}_2)$

In this Appendix, we collect some statements about $\mathrm{Sp}_4(\mathbb{F}_2)$ that are used in Chapter 6 and Chapter 7. First, we note the following fact:

Proposition B.0.1. *There is a unique isomorphism $\theta : \mathrm{Sp}_4(\mathbb{F}_2) \to S_6$ with*

$$\theta(\varepsilon_\alpha(1)) = (1,2)(3,4)(5,6), \theta(\varepsilon_{-\alpha}(1)) = (1,3)(2,5)(4,6)$$
$$\theta(\varepsilon_\beta(1)) = (4,6), \theta(\varepsilon_{-\beta}) = (5,6).$$

Proof. We define $\theta(\varepsilon_\phi(1))$ as follows:

$$\theta(\varepsilon_\alpha(1)) = (1,2)(3,4)(5,6), \theta(\varepsilon_{-\alpha}(1)) = (1,3)(2,5)(4,6)$$
$$\theta(\varepsilon_\beta(1)) = (4,6), \theta(\varepsilon_{-\beta}(1)) = (5,6)$$
$$\theta(\varepsilon_{\alpha+\beta}(1)) = (1,2)(3,5)(4,6), \theta(\varepsilon_{-\alpha-\beta}(1)) = (1,3)(2,4)(5,6)$$
$$\theta(\varepsilon_{2\alpha+\beta}(1)) = (1,2), \theta(\varepsilon_{-2\alpha-\beta}(1)) = (1,3)$$

Using [41, Chapter 6, p. 43, Theorem 8] and the fact that all permutations described above have order 2, we obtain that θ extends to a homomorphism $\theta : \mathrm{Sp}_4(\mathbb{F}_2) \to S_6$ if the following conditions are satisfied for all $\phi, \psi \in C_2$:

$(\theta(\varepsilon_\phi(1)), \theta(\varepsilon_\psi(1))) = 1$, if $\{\phi + \psi\} = (\mathbb{Z}_{>0}\phi \oplus \mathbb{Z}_{>0}\psi) \cap C_2$
$(\theta(\varepsilon_\phi(1)), \theta(\varepsilon_\psi(1))) = 1$, if $\phi + \psi \notin C_2$ and $\phi + \psi \neq 0$
$(\theta(\varepsilon_\phi(1)), \theta(\varepsilon_\psi(1))) = \theta(\varepsilon_{\phi+\psi}(1))\theta(\varepsilon_\tau(1))$, if $\phi + \psi \in C_2$ and $\tau = \phi + 2\psi$ or $2\phi + \psi \in C_2$.

But note that $\theta(\varepsilon_{-\phi}(1)) = [(2,3)(5,4)]\theta(\varepsilon_\phi(1))[(2,3)(5,4)]^{-1}$ holds for all $\phi \in C_2$. Thus to see that θ extends to a homomorphism it suffices to show the conditions in the case that ϕ is positive.

Next, we go through the various possibilities for ϕ and ψ. First assume $\phi = \alpha$. For

the first case take $\psi = \beta$. But indeed

$$\begin{aligned}(\theta(\varepsilon_\alpha(1)), \theta(\varepsilon_\beta(1))) &= [(1,2)(3,4)(5,6)] \cdot (4,6) \cdot [(1,2)(3,4)(5,6)]^{-1} \cdot (4,6) = (3,5)(4,6) \\ &= (1,2)(3,5)(4,6) \cdot (1,2) = \theta(\varepsilon_{\alpha+\beta}(1))\theta(\varepsilon_{2\alpha+\beta}(1))\end{aligned}$$

holds. Further, for $\psi = \alpha + \beta$

$$\begin{aligned}(\theta(\varepsilon_\alpha(1)), \theta(\varepsilon_{\alpha+\beta}(1))) &= [(1,2)(3,4)(5,6)] \cdot (1,2)(3,5)(4,6) \cdot [(1,2)(3,4)(5,6)]^{-1} \\ &\quad \cdot (1,2)(3,5)(4,6) = (2,1)(4,6)(3,5) \cdot (1,2)(3,5)(4,6) = 1\end{aligned}$$

holds and for $\psi = 2\alpha + \beta$ one obtains

$$(\theta(\varepsilon_\alpha(1)), \theta(\varepsilon_{2\alpha+\beta}(1))) = [(1,2)(3,4)(5,6)] \cdot (1,2) \cdot [(1,2)(3,4)(5,6)]^{-1} \cdot (1,2) = 1.$$

For the case $\psi = -\beta$, we obtain

$$(\theta(\varepsilon_\alpha(1)), \theta(\varepsilon_{-\beta}(1))) = [(1,2)(3,4)(5,6)] \cdot (5,6) \cdot [(1,2)(3,4)(5,6)]^{-1} \cdot (5,6) = 1$$

and for the case $\psi = -\alpha - \beta$, we obtain

$$\begin{aligned}(\theta(\varepsilon_\alpha(1)), \theta(\varepsilon_{-\alpha-\beta}(1))) &= [(1,2)(3,4)(5,6)] \cdot (1,3)(2,4)(5,6) \cdot [(1,2)(3,4)(5,6)]^{-1} \\ &\quad \cdot (1,3)(2,4)(5,6) = (2,4)(1,3)(6,5) \cdot (1,3)(2,4)(5,6) = 1.\end{aligned}$$

Lastly, we obtain for the case $\psi = -2\alpha - \beta$ that

$$\begin{aligned}(\theta(\varepsilon_\alpha(1)), \theta(\varepsilon_{-2\alpha-\beta}(1))) &= [(1,2)(3,4)(5,6)] \cdot (1,3) \cdot [(1,2)(3,4)(5,6)]^{-1} \cdot (1,3) \\ &= (2,4)(1,3) = (1,3)(2,4)(5,6) \cdot (5,6) = \theta(\varepsilon_{-\alpha-\beta}(1))\theta(\varepsilon_{-\beta}(1)).\end{aligned}$$

This finishes the case $\phi = \alpha$. Further note that all of the previous commutators have order at most two and hence the previous calculations also settle the cases

$$(\phi, \psi) \in \{(\beta, \alpha), (\alpha + \beta, \alpha), (2\alpha + \beta, \alpha)\}.$$

Further, using the fact that $\theta(\varepsilon_{-\phi}(1)) = [(2,3)(5,4)]\theta(\varepsilon_\phi(1))[(2,3)(5,4)]^{-1}$ holds for all $\phi \in C_2$, these calculations also settle the cases

$$(\phi, \psi) \in \{(\beta, -\alpha), (\alpha + \beta, -\alpha), (2\alpha + \beta, -\alpha)\}.$$

Next, consider the case $\phi = \beta$. First, consider $\psi = \alpha + \beta$. Indeed

$$(\theta(\varepsilon_\beta(1)), \theta(\varepsilon_{\alpha+\beta}(1))) = [(1,2)(3,5)(4,6)] \cdot (4,6) \cdot [(1,2)(3,5)(4,6)]^{-1} \cdot (4,6) = 1$$

holds. Next, assume $\psi = 2\alpha + \beta$ and indeed

$$(\theta(\varepsilon_\beta(1)), \theta(\varepsilon_{2\alpha+\beta}(1))) = (1,2) \cdot (4,6) \cdot (1,2) \cdot (4,6) = 1$$

holds. Next, assume $\psi = -\alpha - \beta$. Then

$$\begin{aligned}(\theta(\varepsilon_\beta(1)), \theta(\varepsilon_{-\alpha-\beta}(1))) &= [(1,3)(2,4)(5,6)] \cdot (4,6) \cdot [(1,3)(2,4)(5,6)]^{-1} \cdot (4,6) = (2,5)(4,6) \\ &= (1,3) \cdot (1,3)(2,5)(4,6) = \theta(\varepsilon_{-2\alpha-\beta}(1))\theta(\varepsilon_{-\alpha}(1))\end{aligned}$$

holds. For $\psi = -2\alpha - \beta$, we obtain

$$(\theta(\varepsilon_\beta(1)), \theta(\varepsilon_{-2\alpha-\beta}(1))) = (1,3) \cdot (4,6) \cdot (1,3) \cdot (4,6) = 1.$$

This settles the case $\phi = \beta$ and similarly to the case $\phi = \alpha$, also settles the cases of

$$(\phi, \psi) \in \{(\alpha+\beta, \beta), (2\alpha+\beta, \beta), (\alpha+\beta, -\beta), (2\alpha+\beta, -\beta)\}.$$

Next, consider the case $\phi = \alpha + \beta$. The only remaining cases for this ϕ are $\psi = 2\alpha + \beta$ and $\psi = -2\alpha - \beta$. Observe that

$$(\theta(\varepsilon_{\alpha+\beta}(1)), \theta(\varepsilon_{2\alpha+\beta}(1))) = [(1,2)(3,5)(4,6)] \cdot (1,2) \cdot [(1,2)(3,5)(4,6)]^{-1} \cdot (1,2) = 1$$

and

$$\begin{aligned}(\theta(\varepsilon_{\alpha+\beta}(1)), \theta(\varepsilon_{-2\alpha-\beta}(1))) &= [(1,2)(3,5)(4,6)] \cdot (1,3) \cdot [(1,2)(3,5)(4,6)]^{-1} \cdot (1,3) \\ &= (2,5)(1,3) = (4,6) \cdot (1,3)(2,5)(4,6) = \theta(\varepsilon_\beta(1))\theta(\varepsilon_{-\alpha}(1)).\end{aligned}$$

This settles the case $\phi = \alpha + \beta$ and by similar considerations as in the previous cases also settles the cases

$$(\phi, \psi) \in \{(2\alpha+\beta, \alpha+\beta), (2\alpha+\beta, -\alpha-\beta)\}.$$

We are left with the case $\phi = 2\alpha + \beta$, but all possibilities for ψ have already been addressed in the previous cases. Thus indeed θ extends to a homomorphism

$$\theta : \mathrm{Sp}_4(\mathbb{F}_2) \to S_6.$$

Next, we will show that θ is surjective. To this end let H be the image of θ. Note that S_6 is generated by its transpositions (i,j) for $1 \leq i < j \leq 6$ and further

$$(6,i)(6,j) = (i,j)$$

holds. Thus S_6 is generated by the transpositions $(6,i)$ for $1 \leq i < 6$. Hence to show the surjectivity of θ it suffices to prove that all transpositions $(6,i)$ for $1 \leq i < 6$ are elements of H. To this end observe that

$$\theta(\varepsilon_\alpha(1))\theta(\varepsilon_\beta(1))\theta(\varepsilon_\alpha(-1)) = [(1,2)(3,4)(5,6)] \cdot (4,6) \cdot [(1,2)(3,4)(5,6)]^{-1} = (3,5)$$

is an element of H and hence $(3,5) \cdot \theta(\varepsilon_{-\beta}(1)) \cdot (3,5) = (3,5)(5,6)(3,5) = (3,6)$ is an element of H. Thus $\theta(\varepsilon_\beta(1)) \cdot (3,6) \cdot \theta(\varepsilon_\beta(1)) = (4,6)(3,6)(4,6) = (3,4)$ is an element of H as well. But then

$$\theta(\varepsilon_\alpha(1))\theta(\varepsilon_{-\beta}(1))(3,4) = (1,2)(3,4)(5,6)(5,6)(3,4) = (1,2)$$

is an element of H. Similarly

$$\theta(\varepsilon_{-\alpha}(1))\theta(\varepsilon_{-\beta}(1))\theta(\varepsilon_{-\alpha}(-1)) = (1,3)(2,5)(4,6)(5,6)[(1,3)(2,5)(4,6)]^{-1} = (2,4)$$

is an element of H and so $(2,4) \cdot \theta(\varepsilon_\beta(1)) \cdot (2,4) = (2,4)(4,6)(2,4) = (2,6)$ is an element of H. Thus $(1,2)(2,6)(1,2) = (1,6)$ is an element of H. So, indeed all the transpositions $(1,6), (2,6), (3,6), (4,6) = \theta(\varepsilon_\beta(1))$ and $(5,6) = \theta(\varepsilon_{-\beta}(1))$ are elements of H and so $H = S_6$ holds and hence θ is surjective.

On the other hand, S_6 has $6! = 720$ elements and $\mathrm{Sp}_4(\mathbb{F}_2)$ has

$$2^4 \cdot (2^2 - 1) \cdot (2^4 - 1) = 16 \cdot 3 \cdot 15 = 720$$

elements according to [41, Chapter 9, p. 77, Theorem 25]. Thus θ can only be surjective, if it is injective as well and this finishes the proof of the existence of an isomorphism θ as described. However, uniqueness of θ is clear, because $\{\varepsilon_\phi(1) | \phi \in C_2\}$ is a generating set of $\mathrm{Sp}_4(\mathbb{F}_2)$ and θ is defined on them. □

Remark B.0.2. The group $\mathrm{Sp}_4(\mathbb{F}_2)$ acts on the set of maximal subsets M of $\mathbb{F}_2^4 - \{0\}$ with the property that any two distinct v, w elements of M have the property $\omega(v,w) = 1$ for ω the symplectic structure fixed by $\mathrm{Sp}_4(\mathbb{F}_2)$. There are six such sets M and an analysis of the permutation of those six sets by $\mathrm{Sp}_4(\mathbb{F}_2)$ yields the isomorphism θ.

Next, we show:

Proposition B.0.3. *Let $\phi \in C_2$ be given. Then $E := \varepsilon_\phi(1)$ normally generates $\mathrm{Sp}_4(\mathbb{F}_2)$ and $\|\mathrm{Sp}_4(\mathbb{F}_2)\|_E = 5$ holds.*

Proof. First, after conjugation we may assume that ϕ is a positive simple root in C_2. Then as mentioned in [41, Chapter 86, p. 86, Example (b)], there is an automorphism j

of $\mathrm{Sp}_4(\mathbb{F}_2)$ with

$$j(\varepsilon_\alpha(1)) = \varepsilon_\beta(1).$$

Hence we may further assume that $\phi = \beta$. Then using the isomorphism θ from Proposition B.0.1 implies that we have to show that the transposition $(4,6)$ normally generates S_6 and that $\|S_6\|_{(4,6)} = 5$ holds. However, the group S_6 has only three normal subgroups $\{1\}, A_6$ and S_6 and $(4,6)$ is not an element of A_6 and hence must normally generate S_6. Obviously, any conjugate of a transposition is again a transposition and hence it suffices to show that there are elements in S_6 that cannot be written as a product of four transpositions to prove $\|S_6\|_{(4,6)} \geq 5$. To see this, observe that for $\sigma \in S_6$, the number of orbits of the induced group action of $\langle\sigma\rangle$ on $\{1,\ldots,6\}$ only depends on the conjugacy class of σ in S_6 instead of on the permutation σ itself. However, for $k \in \{1,\ldots,5\}$, a product of k transpositions in S_6 has at least $6-k$ such orbits in $\{1,\ldots,6\}$. Thus the cyclce $(1,2,3,4,5,6)$, which gives rise to just one such orbit, cannot be written as a product of at most 4 transpositions.

But [7, Lemma 2.05, Lemma 2.06, Lemma 3.01] implies that the covering number of S_6 is at most 5 and hence

$$5 \leq \|S_6\|_{(4,6)} \leq \Delta_1(S_6) \leq \mathrm{cn}(S_6) \leq 5.$$

This finishes the proof. □

Appendix C

Various proofs

Lemma 4.4.4. *The sequences*

1. $s_1 := (w_\alpha, w_\epsilon, w_\phi, w_\beta, w_\delta, w_\gamma, w_\alpha, w_\epsilon, w_\phi, w_\beta, w_\delta, w_\gamma, w_\alpha, w_\epsilon, w_\phi, w_\beta, w_\delta, w_\gamma, w_\alpha, w_\epsilon)$ *and*
2. $s_2 := (w_\delta, w_\beta, w_\phi, w_\epsilon, w_\alpha, w_\gamma, w_\delta, w_\beta, w_\phi, w_\epsilon, w_\alpha, w_\gamma, w_\delta, w_\beta, w_\phi)$

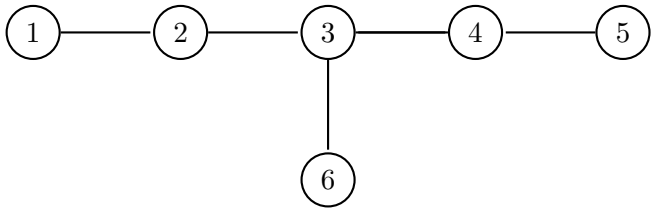
of fundamental reflections in $W(E_6)$ give minimal expressions with respect to the fundamental reflections for the corresponding Weyl group elements $w_1, w_2 \in W(E_6)$. Further, for χ the positive root of highest weight in E_6 :

$$w_1(\chi) = w_2(\chi) = \gamma \text{ and } T(\gamma) = E_6^+ - \{\alpha, \beta, \delta, \epsilon, \phi, \alpha + \beta, \delta + \epsilon\}.$$

Proof. One notes that the sequence s_1 appears as a subexpression made up of consecutive letters of the sequence given for the longest word $w_0 \in W(E_6)$ in Lemma 4.4.3. Thus s_1 must be a minimal expression for its corresponding Weyl group element w_1, because otherwise the expression for w_0 could be shortened. Similarly, it follows from the reverse expression for w_0 that s_2 is a minimal expression.

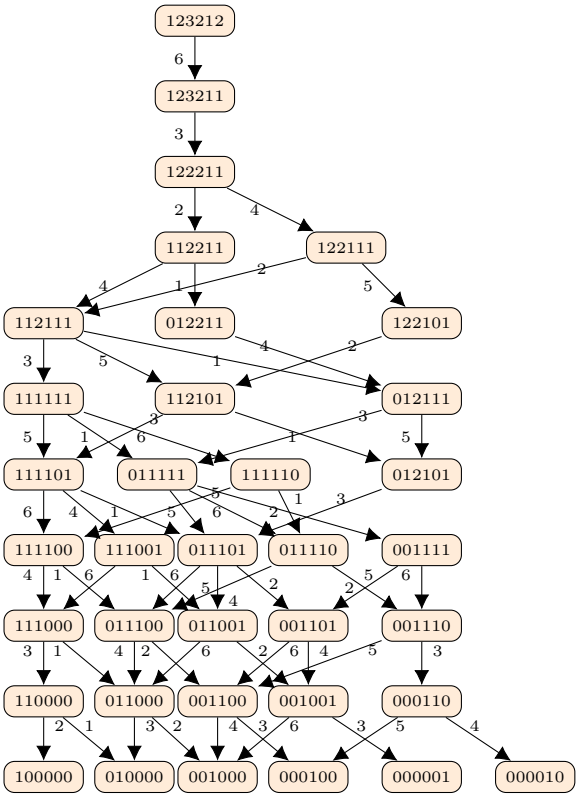
We denote the positive simple roots $\alpha, \beta, \gamma, \delta, \epsilon, \phi$ by $1, 2, 3, 4, 5$ and 6 and recall that the corresponding Dynkin diagram looks as follows

E_6 :



For convenience and later reference, we arrange the positive, roots of E_6 into a Hasse-diagram where the vertices are the positive, roots of E_6 and the label of the vertex denotes how often the corresponding simple root appears in the expression of the root/vertex in question. For example, the label $(1, 1, 1, 0, 0, 1)$ denotes the root $\alpha + \beta + \gamma + \phi$. The label on the edge of the diagram denotes what simple root is the difference between the two

roots adjacent to the edge. Furthermore, the list of positive roots of E_6 to be arranged into the Hasse-diagram is taken from [18, Appendix, Table B, p. 528].



Because the calculation of $w_1(\chi)$ is rather lengthy, we will only show $w_1(\chi) = \gamma$. To this end, observe that E_6 is simply-laced and so if $\psi \in E_6^+$ and θ a positive, simple root are given, then $w_\theta(\psi)$ is either $\psi+\theta, \psi-\theta$ or ψ, depending on whether the vertex of the Hasse-diagram corresponding to ψ is incident to an edge labeled by θ connecting to a vertex of higher or lower weight respectively or is not incident to an edge labeled by θ at all. Observe as a consequence for the positive root of highest weight $\chi = \alpha + 2\beta + 3\gamma + 2\delta + \epsilon + 2\phi$,

that:

$$
\begin{aligned}
w_1(\chi) &= w_\alpha w_\epsilon w_\phi w_\beta w_\delta w_\gamma w_\alpha w_\epsilon w_\phi w_\beta w_\delta w_\gamma w_\alpha w_\epsilon w_\phi w_\beta w_\delta w_\gamma w_\alpha w_\epsilon(\alpha + 2\beta + 3\gamma + 2\delta + \epsilon + 2\phi) \\
&= w_\alpha w_\epsilon w_\phi w_\beta w_\delta w_\gamma w_\alpha w_\epsilon w_\phi w_\beta w_\delta w_\gamma w_\alpha w_\epsilon w_\phi(\alpha + 2\beta + 3\gamma + 2\delta + \epsilon + 2\phi) \\
&= w_\alpha w_\epsilon w_\phi w_\beta w_\delta w_\gamma w_\alpha w_\epsilon w_\phi w_\beta w_\delta w_\gamma w_\alpha w_\epsilon(\alpha + 2\beta + 3\gamma + 2\delta + \epsilon + \phi) \\
&= w_\alpha w_\epsilon w_\phi w_\beta w_\delta w_\gamma w_\alpha w_\epsilon w_\phi w_\beta w_\delta w_\gamma(\alpha + 2\beta + 3\gamma + 2\delta + \epsilon + \phi) \\
&= w_\alpha w_\epsilon w_\phi w_\beta w_\delta w_\gamma w_\alpha w_\epsilon w_\phi w_\beta w_\delta(\alpha + 2\beta + 2\gamma + 2\delta + \epsilon + \phi) \\
&= w_\alpha w_\epsilon w_\phi w_\beta w_\delta w_\gamma w_\alpha w_\epsilon w_\phi w_\beta(\alpha + 2\beta + 2\gamma + \delta + \epsilon + \phi) \\
&= w_\alpha w_\epsilon w_\phi w_\beta w_\delta w_\gamma w_\alpha w_\epsilon w_\phi(\alpha + \beta + 2\gamma + \delta + \epsilon + \phi) \\
&= w_\alpha w_\epsilon w_\phi w_\beta w_\delta w_\gamma w_\alpha w_\epsilon(\alpha + \beta + 2\gamma + \delta + \epsilon + \phi) \\
&= w_\alpha w_\epsilon w_\phi w_\beta w_\delta w_\gamma w_\alpha(\alpha + \beta + 2\gamma + \delta + \phi) \\
&= w_\alpha w_\epsilon w_\phi w_\beta w_\delta w_\gamma(\beta + 2\gamma + \delta + \phi) \\
&= w_\alpha w_\epsilon w_\phi w_\beta w_\delta(\beta + \gamma + \delta + \phi) = w_\alpha w_\epsilon w_\phi w_\beta(\beta + \gamma + \phi) = w_\alpha w_\epsilon w_\phi(\gamma + \phi) = w_\alpha w_\epsilon(\gamma) = \gamma
\end{aligned}
$$

The fact that $T(\gamma) = E_6^+ - \{\alpha, \beta, \delta, \epsilon, \phi, \alpha + \beta, \delta + \epsilon\}$ can be seen by inspection of the Hasse diagram. □

Lemma 5.1.6. *Let R be a principal ideal domain. Then R has stable range at most 2.*

Proof. Let $m \geq 2$ be given and let $v_0, \ldots, v_m \in R$ be given with $(v_0, \ldots, v_m) = R$. Let S be the prime divisors of $(v_1, \ldots, v_m)$ and let T be the prime divisors of $(v_2, \ldots, v_m)$ that are not also prime divisors of v_1. Obviously T and S do not intersect. So according to the Chinese Remainder Theorem, there is an $x \in R$ such that

$$
\begin{aligned}
\forall p \in S &: x \equiv 1 \bmod p \\
\forall q \in T &: x \equiv 0 \bmod q
\end{aligned}
$$

Then consider the ideal $I := (v_1 - xv_0, v_2, \ldots, v_m)$ in R. Assume that this is ideal is not R. Then there is a prime divisor h of I. Then clearly h divides the ideal $(v_2, \ldots, v_m)$ as well. But if h would not divide v_1, then it would be an element of T and hence h divides x. But h also divides $v_1 - xv_0$, so we can conclude that h must also divide v_1, a contradiction. So h must divide v_1.

But if h divides v_1, then h is an element of S. Further, h also divides $v_1 - xv_0$ and so h also divides xv_0. But h is an element of S, so we have $x \equiv 1 \bmod h$ and so in particular

$$v_0 = 1 \cdot v_0 \equiv xv_0 \equiv 0 \bmod h$$

follows. So h also divides v_0 and so h divides $(v_0, v_1, \ldots, v_m) = R$, a contradiction. Thus $(v_1 - xv_0, v_2, \ldots, v_m) = R$ holds. □

Bibliography

[1] Eiichi Abe. Chevalley groups over local rings. *Tohoku Math. J. (2)*, 21:474–494, 1969.

[2] Eiichi Abe. Normal subgroups of Chevalley groups over commutative rings. In *Algebraic K-theory and algebraic number theory (Honolulu, HI, 1987)*, volume 83 of *Contemp. Math.*, pages 1–17. Amer. Math. Soc., Providence, RI, 1989.

[3] Eiichi Abe and Kazuo Suzuki. On normal subgroups of Chevalley groups over commutative rings. *Tohoku Math. J. (2)*, 28(2):185–198, 1976.

[4] Z. Arad, J. Stavi, and M. Herzog. Powers and products of conjugacy classes in groups. In *Products of conjugacy classes in groups*, volume 1112 of *Lecture Notes in Math.*, pages 6–51. Springer, Berlin, 1985.

[5] H. Bass. K-theory and stable algebra. *Inst. Hautes Études Sci. Publ. Math.*, (22):5–60, 1964.

[6] H. Bass, J. Milnor, and J.-P. Serre. Solution of the congruence subgroup problem for SL_n ($n \geq 3$) and Sp_{2n} ($n \geq 2$). *Inst. Hautes Études Sci. Publ. Math.*, (33):59–137, 1967.

[7] J. L. Brenner. Covering theorems for FINASIGs. VIII. Almost all conjugacy classes in $\mathcal{A}_n$ have exponent ≤ 4. *J. Austral. Math. Soc. Ser. A*, 25(2):210–214, 1978.

[8] Dmitri Burago, Sergei Ivanov, and Leonid Polterovich. Conjugation-invariant norms on groups of geometric origin. In *Groups of diffeomorphisms*, volume 52 of *Adv. Stud. Pure Math.*, pages 221–250. Math. Soc. Japan, Tokyo, 2008.

[9] M. Burger and N. Monod. Bounded cohomology of lattices in higher rank Lie groups. *J. Eur. Math. Soc. (JEMS)*, 1(2):199–235, 1999.

[10] David Carter and Gordon Keller. Bounded elementary generation of $\mathrm{SL}_n(\mathcal{O})$. *Amer. J. Math.*, 105(3):673–687, 1983.

[11] Claude Chevalley. Certains schémas de groupes semi-simples. In *Séminaire Bourbaki, Vol. 6*, pages Exp. No. 219, 219–234. Soc. Math. France, Paris, 1995.

[12] Maurice Chiodo. Finitely annihilated groups. *Bull. Aust. Math. Soc.*, 90(3):404–417, 2014.

[13] Douglas L. Costa and Gordon E. Keller. The $E(2, A)$ sections of $\mathrm{SL}(2, A)$. *Ann. of Math. (2)*, 134(1):159–188, 1991.

[14] Douglas L. Costa and Gordon E. Keller. Radix redux: normal subgroups of symplectic groups. *J. Reine Angew. Math.*, 427:51–105, 1992.

[15] Douglas L. Costa and Gordon E. Keller. On the normal subgroups of $G_2(A)$. *Trans. Amer. Math. Soc.*, 351(12):5051–5088, 1999.

[16] Filippo De Mari and Mark A. Shayman. Lie algebraic generalizations of Hessenberg matrices and the topology of Hessenberg varieties. In *Realization and modelling in system theory (Amsterdam, 1989)*, volume 3 of *Progr. Systems Control Theory*, pages 141–148. Birkhäuser Boston, Boston, MA, 1990.

[17] R. K. Dennis and L. N. Vaserstein. On a question of M. Newman on the number of commutators. *J. Algebra*, 118(1):150–161, 1988.

[18] Hans Freudenthal and H. de Vries. Linear Lie groups. pages xxii+547, 1969.

[19] Światosław R. Gal, Jarek Kędra, and Alexander Trost. Finite index subgroups in Chevalley groups are bounded: an addendum to 'on bi-invariant word metrics'. https://arxiv.org/abs/1808.06376.

[20] Roger A. Horn and Charles R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, second edition, 2013.

[21] James E. Humphreys. *Linear algebraic groups, corrected fifth printing*. Springer-Verlag, New York-Heidelberg, 1975. Graduate Texts in Mathematics, No. 21.

[22] James E. Humphreys. *Reflection groups and Coxeter groups*, volume 29 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1990.

[23] S. Karni. Covering numbers of groups of small order and sporadic groups. In *Products of conjugacy classes in groups*, volume 1112 of *Lecture Notes in Math.*, pages 52–196. Springer, Berlin, 1985.

[24] Jarek Kędra, Assaf Libman, and Ben Martin. On boundedness properties of groups. *Submitted*, https://arxiv.org/abs/1808.01815.

[25] R. Lawther and Martin W. Liebeck. On the diameter of a Cayley graph of a simple group of Lie type based on a conjugacy class. *J. Combin. Theory Ser. A*, 83(1):118–137, 1998.

[26] Arieh Lev. The covering number of the group $\mathrm{PSL}_n(F)$. *J. Algebra*, 182(1):60–84, 1996.

[27] Martin W. Liebeck and Aner Shalev. Diameters of finite simple groups: sharp bounds and applications. *Ann. of Math. (2)*, 154(2):383–406, 2001.

[28] Daniel A. Marcus. *Number fields*. Universitext. Springer, Cham, 2018. Second edition of [MR0457396], With a foreword by Barry Mazur.

[29] Aleksander V. Morgan, Andrei S. Rapinchuk, and Balasubramanian Sury. Bounded generation of SL_2 over rings of S-integers with infinitely many units. *Algebra & Number Theory*, 12(8):1949–1974, 2018.

[30] Dave Witte Morris. Bounded generation of $\mathrm{SL}(n, A)$ (after D. Carter, G. Keller, and E. Paige). *New York J. Math.*, 13:383–421, 2007.

[31] Dave Witte Morris. *Introduction to arithmetic groups*. Deductive Press, Place of publication not identified, 2015.

[32] Jürgen Neukirch. *Algebraic number theory*, volume 322 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999. Translated from the 1992 German original and with a note by Norbert Schappacher, With a foreword by G. Harder.

[33] Morris Newman. *Integral matrices*. Academic Press, New York-London, 1972. Pure and Applied Mathematics, Vol. 45.

[34] Bogdan Nica. On bounded elementary generation for SL_n over polynomial rings. *Israel J. Math.*, 225(1):403–410, 2018.

[35] Yaron Ostrover and Roy Wagner. On the extremality of Hofer's metric on the group of Hamiltonian diffeomorphisms. *Int. Math. Res. Not.*, (35):2123–2141, 2005.

[36] Leonid Polterovich. *The geometry of the group of symplectic diffeomorphisms*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2001.

[37] Wolfgang Rautenberg. *A concise introduction to mathematical logic*. Universitext. Springer, New York, third edition, 2010. With a foreword by Lev Beklemishev.

[38] Sheila Sandon. An integer-valued bi-invariant metric on the group of contactomorphisms of $\mathbb{R}^{2n} \times S^1$. *J. Topol. Anal.*, 2(3):327–339, 2010.

[39] Sheila Sandon. Bi-invariant metrics on the contactomorphism groups. *São Paulo J. Math. Sci.*, 9(2):195–228, 2015.

[40] Michael R. Stein. Stability theorems for K_1, K_2 and related functors modeled on Chevalley groups. *Japan. J. Math. (N.S.)*, 4(1):77–108, 1978.

[41] Robert Steinberg. *Lectures on Chevalley groups*, volume 66 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2016.

[42] O. I. Tavgen. Bounded generability of Chevalley groups over rings of S-integer algebraic numbers. *Izv. Akad. Nauk SSSR Ser. Mat.*, 54(1):97–122, 221–222, 1990.

[43] Leonid N. Vaserstein. On normal subgroups of Chevalley groups over commutative rings. *Tohoku Math. J. (2)*, 38(2):219–230, 1986.

[44] N. A. Vavilov. A note on the subnormal structure of general linear groups. *Math. Proc. Cambridge Philos. Soc.*, 107(2):193–196, 1990.

[45] N. A. Vavilov, A. V. Smolenskiĭ, and B. Sury. Unitriangular factorizations of Chevalley groups. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 388(Voprosy Teorii Predstavleniĭ Algebr i Grupp. 21):17–47, 309–310, 2011.

[46] Claude Viterbo. Symplectic topology as the geometry of generating functions. *Math. Ann.*, 292(4):685–710, 1992.

[47] Hong You. Subgroups of classical groups normalized by relative elementary groups. *J. Pure Appl. Algebra*, 216(5):1040–1051, 2012.