PRESENTATION OF AFFINE KAC-MOODY GROUPS OVER RINGS

DANIEL ALLCOCK

Abstract. Tits has defined Steinberg groups and Kac-Moody groups for any root system and any commutative ring R. We establish a Curtis-Tits-style presentation for the Steinberg group $\mathfrak{St}$ of any rank ≥ 3 irreducible affine root system, for any R. Namely, $\mathfrak{St}$ is the direct limit of the Steinberg groups coming from the 1- and 2-node subdiagrams of the Dynkin diagram. This leads to a completely explicit presentation. Using this we show that $\mathfrak{St}$ is finitely presented if the rank is ≥ 4 and R is finitely generated as a ring, or if the rank is 3 and R is finitely generated as a module over a subring generated by finitely many units. Similar results hold for the corresponding Kac-Moody groups when R is a Dedekind domain of arithmetic type.

1. Introduction

Suppose R is a commutative ring and A is one of the $ABCDEFG$ Dynkin diagrams, or equivalently its Cartan matrix. Steinberg defined what is now called the Steinberg group $\mathfrak{St}_A(R)$, by generators and relations [18]. It plays a central role in K-theory and some aspects of Lie theory.

Kac-Moody algebras are infinite-dimensional generalizations of the semisimple Lie algebras. When $R = \mathbb{R}$ and A is an irreducible affine Dynkin diagram, the corresponding Kac-Moody group is a central extension of the loop group of a finite-dimensional Lie group. For a general ring R and any generalized Cartan matrix A, the definition of the Kac-Moody group is due to Tits [20]. He first constructed a group functor $R \mapsto \mathfrak{St}_A(R)$ generalizing Steinberg's, also by generators and relations. Then he defined another functor $R \mapsto \tilde{\mathfrak{G}}_A(R)$ as a certain quotient of this. In this paper we will omit the tilde and refer to $\mathfrak{G}_A(R)$ as the Kac-Moody group of type A over R. (Tits actually defined $\tilde{\mathfrak{G}}_D(R)$ where D is a root datum; by $\mathfrak{G}_A(R)$ we intend

Date: September 23, 2014.
2000 *Mathematics Subject Classification.* Primary: 20G44; Secondary: 14L15, 22E67, 19C99.
Supported by NSF grant DMS-1101566.

the root datum whose generalized Cartan matrix is A and which is "simply-connected in the strong sense" [20, p. 551].)

The meaning of "Kac-Moody group" is far from standardized. In [20] Tits wrote down axioms (KMG1)–(KMG9) that one could demand of a functor from rings to groups before calling it a Kac-Moody functor. He showed [20, Thm. 1′] that any such functor admits a natural homomorphism from $\mathfrak{G}_A$, which is an isomorphism at every field. So Kac-Moody groups over fields are well-defined, and over general rings $\mathfrak{G}_A$ approximates the yet-unknown ultimate definition. This is why we refer to $\mathfrak{G}_A$ as the Kac-Moody group. But $\mathfrak{G}_A$ does not quite satisfy Tits' axioms, so ultimately some other language may be better. See section 5 for more remarks on this.

The purpose of this paper is to simplify Tits' presentations of $\mathfrak{St}_A(R)$ and $\mathfrak{G}_A(R)$ when A is an irreducible affine Dynkin diagram of rank (number of nodes) at least 3. In particular, we show that these groups are finitely presented under quite weak hypotheses on R. This is surprising because there is no obvious reason for an infinite-dimensional group over (say) $\mathbb{Z}$ to be finitely presented, and Tits' presentations are "very" infinite. His generators are indexed by all pairs (root, ring element), and his relations specify the commutators of certain pairs of these generators. Subtle implicitly-defined coefficients appear throughout his relations.

To prove our finite presentation results, we will first establish explicit presentations for $\mathfrak{St}_A(R)$ and $\mathfrak{G}_A(R)$ that are much smaller than Tits' presentations, though still infinite if R is infinite. Namely, in [2] we defined for any A a group functor we called the pre-Steinberg group $\mathfrak{PSt}_A$. The definition is the same as Tits' definition of $\mathfrak{St}_A$, as modified by Morita-Rehmann [14], except with some relations omitted. The main result of this paper is that the omitted relations are redundant. More formally, theorem 1 states that $\mathfrak{PSt}_A \to \mathfrak{St}_A$ is an isomorphism. This is useful because in [2, thm. 1.2] we gave a simple closed-form presentation of $\mathfrak{PSt}_A$.

This presentation has generators S_i and $X_i(t)$, with i varying over the nodes of the Dynkin diagram and t over R, and the relations are (70)–(95) in [2]. A full list of the relations is not needed in this paper, but for concreteness we give in table 1 the complete presentation when A is simply-laced (and not A_1). If multiple bonds are present then the relations are more complicated but similar. Or see [2, sec. 2] for the A_1, A_2, B_2 and G_2 cases, which are enough to present $\mathfrak{PSt}_A(R)$ for any A.

$$\left.\begin{aligned} X_i(t)X_i(u) &= X_i(t+u) \\ [S_i^2, X_i(t)] &= 1 \\ S_i &= X_i(1)S_iX_i(1)S_i^{-1}X_i(1) \end{aligned}\right\} \text{all } i$$

$$\left.\begin{aligned} S_iS_j &= S_jS_i \\ [S_i, X_j(t)] &= 1 \\ [X_i(t), X_j(u)] &= 1 \end{aligned}\right\} \text{all unjoined } i \neq j$$

$$\left.\begin{aligned} S_iS_jS_i &= S_jS_iS_j \\ S_i^2S_jS_i^{-2} &= S_j^{-1} \\ X_i(t)S_jS_i &= S_jS_iX_j(t) \\ S_i^2X_j(t)S_i^{-2} &= X_j(t)^{-1} \\ [X_i(t), S_iX_j(u)S_i^{-1}] &= 1 \\ [X_i(t), X_j(u)] &= S_iX_j(tu)S_i^{-1} \end{aligned}\right\} \text{all joined } i \neq j$$

TABLE 1. Defining relations for $\mathfrak{PSt}_A(R)$ when A is simply-laced. The generators are $X_i(t)$ and S_i where i varies over the nodes of the Dynkin diagram and t over R. When A is also affine, this also presents $\mathfrak{St}_A(R)$, by theorem 1.

Theorem 1 (Presentation of affine Steinberg and Kac-Moody groups)**.** *Suppose A is an irreducible affine Dynkin diagram of rank ≥ 3 and R is a commutative ring. Then the natural map $\mathfrak{PSt}_A(R) \to \mathfrak{St}_A(R)$ is an isomorphism.*

In particular, $\mathfrak{St}_A(R)$ has a presentation with generators S_i and $X_i(t)$, with i varying over the simple roots and t over R, and relators (70)–(95) from [2]*. (Or see table 1 when A is simply-laced.)*

One obtains $\mathfrak{G}_A(R)$ by adjoining the relations

$$\tilde{h}_i(u)\tilde{h}_i(v) = \tilde{h}_i(uv) \tag{1}$$

for every simple root i and all units u, v of R, where

$$\begin{aligned} \tilde{s}_i(u) &:= X_i(u)S_iX_i(1/u)S_i^{-1}X_i(u) \\ \tilde{h}_i(u) &:= \tilde{s}_i(u)\tilde{s}_i(-1). \end{aligned}$$

We remark that if A is a spherical diagram (that is, its Weyl group is finite) then $\mathfrak{PSt}_A \to \mathfrak{St}_A$ is an isomorphism, essentially by the definition of $\mathfrak{PSt}_A$. See [2] for details. So theorem 1 extends the isomorphism $\mathfrak{PSt}_A \cong \mathfrak{St}_A$ to the irreducible affine case (except for rank 2). See [3] for a further extension, to the simply-laced hyperbolic case.

The key property of our presentation of $\mathfrak{PSt}_A(R)$ is its description in terms of the Dynkin diagram rather than the full (usually infinite) root system, as in Tits' construction. Furthermore, every relation involves just one or two subscripts. In the simply-laced case one can check this by examining table 1, and in the general case one examines the relators (70)–(95) in [2]. Also, the extra relations (1) needed to define the Kac-Moody group also involve only single subscripts. Therefore we obtain the following corollary:

Corollary 2 (Curtis-Tits presentation)**.** *Suppose A is an irreducible affine Dynkin diagram of rank ≥ 3 and R is a commutative ring. Consider the groups $\mathfrak{St}_B(R)$ and the obvious maps between them, as B varies over the singletons and pairs of nodes of A. The direct limit of this family of groups equals $\mathfrak{St}_A(R)$.*

The same result holds with $\mathfrak{St}$ replaced by $\mathfrak{G}$ throughout. □

A consequence of the corollary 2 is that $\mathfrak{St}_A(R)$'s presentation can be got by gathering together the presentations for the $\mathfrak{St}_B(R)$'s. When the latter groups are finitely presented, we can therefore expect that $\mathfrak{St}_A(R)$ is too. The finite presentability of $\mathfrak{St}_B(R)$ was examined by Splitthoff [17]. Using his work and some additional arguments, we obtain the following:

Theorem 3 (Finite presentability)**.** *Suppose A is an irreducible affine Dynkin diagram of rank ≥ 3 and R is a commutative ring. Then $\mathfrak{St}_A(R)$ is finitely presented if either*

(*i*) $\operatorname{rk} A > 3$ *and R is finitely generated as a ring, or*
(*ii*) $\operatorname{rk} A \geq 3$ *and R is finitely generated as a module over a subring generated by finitely many units.*

In either case, if the unit group of R is finitely generated as an abelian group, then $\mathfrak{G}_A(R)$ is also finitely presented.

One of the main motivations for Splitthoff's work was to understand whether the Chevalley-Demazure groups over Dedekind domains of interest in number theory, are finitely presented. This was settled by Behr [5][6], capping a long series of works by many authors. The following analogue of these results follows immediately from theorem 3. How close the analogy is depends on how well $\mathfrak{G}_A$ approximates whatever plays the role of the Chevalley-Demazure group scheme in the setting of Kac-Moody algebras.

Corollary 4 (Finite presentation in arithmetic contexts)**.** *Suppose K is a global field, meaning a finite extension of $\mathbb{Q}$ or $\mathbb{F}_q(t)$. Suppose S is a nonempty finite set of places of K, including all infinite places in the number field case. Let R be the ring of S-integers in K.*

Suppose A is an irreducible affine Dynkin diagram. Then Tits' Kac-Moody group $\mathfrak{G}_A(R)$ is finitely presented if

(*i*) $\operatorname{rk} A > 3$ *when K is a function field and $|S| = 1$;*
(*ii*) $\operatorname{rk} A \geq 3$ *otherwise.* □

In a companion paper with Carbone [3] we prove analogues of all four theorems for the simply-laced hyperbolic Dynkin diagrams, such as E_{10}.

We remark that if R is a field then the $\mathfrak{G}_A$ case of corollary 2 is due to Abramenko-Mühlherr [1][15]. Namely, suppose A is any generalized Cartan matrix which is 2-spherical (its Dynkin diagram has no edges labeled ∞), and that R is a field (but not $\mathbb{F}_2$ if A has a double bond, and neither $\mathbb{F}_2$ nor $\mathbb{F}_3$ if A has a multiple bond). Then $\mathfrak{G}_A(R)$ is the direct limit of the groups $\mathfrak{G}_B(R)$. See also Caprace [8]. Abramenko-Mühlherr [1, p. 702] state that if A is irreducible affine then one can remove the restrictions $R \neq \mathbb{F}_2, \mathbb{F}_3$.

One of our goals in this work is to bring Kac-Moody groups into the world of geometric and combinatorial group theory, which mostly addresses finitely presented groups. For example: which Kac-Moody groups admit classifying spaces with finitely many cells below some chosen dimension? What other finiteness properties do they have? Do they have Kazhdan's property T? What isoperimetric inequalities do they satisfy in various dimensions? Are there (non-split) Kac-Moody groups over local fields whose uniform lattices (suitably defined) are word hyperbolic? Are some Kac-Moody groups (or classes of them) quasi-isometrically rigid? We find the last question very attractive, since the corresponding answer [9][10][13][16] for lattices in Lie groups is deep.

The author is very grateful to the Japan Society for the Promotion of Science and to Kyoto University, for their support and hospitality. He would also like to thank Lisa Carbone for useful comments on an draft version of the paper.

2. Steinberg and pre-Steinberg Groups

We work in the setting of [20] and [2], so R is a commutative ring and A is a generalized Cartan matrix. This matrix determines a complex Lie algebra $\mathfrak{g} = \mathfrak{g}_A$ called the Kac-Moody algebra, and we write Φ for the set of real roots of $\mathfrak{g}$. Tits' definition of the Steinberg group $\mathfrak{St}_A(R)$

starts with the free product $*_{\alpha\in\Phi}\,\mathfrak{U}_\alpha$, where each root group $\mathfrak{U}_\alpha$ is a copy of the additive group of R.

Tits calls a pair $\alpha, \beta \in \Phi$ *prenilpotent* if some element of the Weyl group W sends both α, β to positive roots, and some other element of W sends both to negative roots. A consequence of this condition is that every root in $\mathbb{N}\alpha + \mathbb{N}\beta$ is real, which enabled Tits to write down Chevalley-style relators for α, β. That is, for every prenilpotent pair α, β he imposes relations of the form

$$\big[\text{element of } \mathfrak{U}_\alpha, \text{element of } \mathfrak{U}_\beta\big] = \prod_{\gamma\in\theta(\alpha,\beta)-\{\alpha,\beta\}} (\text{element of } \mathfrak{U}_\gamma) \tag{2}$$

where $\theta(\alpha,\beta) := (\mathbb{N}\alpha + \mathbb{N}\beta) \cap \Phi$ and $\mathbb{N} = \{0, 1, 2, \dots\}$. In the general case, the exact relations are given in a rather implicit form in [20, §3.6]. We give them explicitly in section 4, but only in the cases we need and only as we need them. For Tits, and for purposes of this paper, this is the end of the definition of $\mathfrak{St}_A(R)$.

The pre-Steinberg group $\mathfrak{PSt}_A(R)$, defined in [2, sec. 7], imposes these relations only for the *classically nilpotent* pairs α, β. This means that α, β are linearly independent and $(\mathbb{Q}\alpha \oplus \mathbb{Q}\beta) \cap \Phi$ is finite. This is equivalent to α, β being non-antipodal and lying in some A_1^2, A_2, B_2 or G_2 root system. As the name suggests, such a pair is prenilpotent. So $\mathfrak{PSt}_A(R)$ is defined the same way as $\mathfrak{St}_A(R)$, just omitting the Chevalley relations for prenilpotent pairs that are not classically prenilpotent. In particular, $\mathfrak{St}_A(R)$ is a quotient of $\mathfrak{PSt}_A(R)$, hence the prefix "pre-".

There are some additional relations in the pre-Steinberg group, and in the Steinberg group as defined by Morita-Rehmann [14], whom we follow. These relations are natural, and required to recover Steinberg's original definition in the A_1 case. But they are irrelevant to this paper. This is because they follow from the relations in $\mathfrak{PSt}_A(R)$ that we have already described, whenever A is 2-spherical without A_1 components. So we will make only the following remarks. First, the extra relations are labeled (69) in [2] and (B′) in [14, p. 538]. Second, Tits imposed them when defining the Kac-Moody group $\mathfrak{G}_A(R)$ as a quotient of $\mathfrak{St}_A(R)$, in [20, eqn. (6)]. Third, Tits' remark (a) in [20, p. 549] explains why they follow from the relations in $\mathfrak{PSt}_A(R)$.

We have not really defined $\mathfrak{St}_A(R)$ and $\mathfrak{PSt}_A(R)$, just discussed the general form of the definitions. The reason is that there are delicate signs present that are not important in this paper. This has nothing to do with Kac-Moody theory and is already present in Steinberg's original groups. Tits' description of the relations (2) elegantly circumvents this problem, but at the cost of making all the relations implicit rather than explicit.

One reason we introduced $\mathfrak{PSt}_A$ is that we could write down a completely explicit presentation for it. The philosophy is that in many cases of interest, the natural map $\mathfrak{PSt}_A(R) \to \mathfrak{St}_A(R)$ is an isomorphism. In fact we presume that $\mathfrak{PSt}_A$ is interesting only when this isomorphism holds. When it does, we get the completely explicit presentation of $\mathfrak{St}_A$ referred to in theorem 1. Its fine details are unimportant in this paper.

3. New nomenclature for affine root systems

Our proof of theorem 1, appearing in the next section, refers to the root system as a whole, with the simple roots playing no special role. It is natural in this setting to use a nomenclature for the affine root systems that emphasizes this global perspective.

The names we use are $\widetilde{A}_n, \ldots, \widetilde{G}_2$, $\widetilde{B}_n^{\text{even}}$, $\widetilde{C}_n^{\text{even}}$, $\widetilde{F}_4^{\text{even}}$, $\widetilde{G}_2^{0 \bmod 3}$ and $\widetilde{BC}_n^{\text{odd}}$. Their virtues are: the subscript is the rank (minus 1, as usual), the corresponding finite root system is indicated by the capital letter(s), and the construction of $\widetilde{X}_n^{\cdots}$ is visible in the superscript. We now give the constructions along with the correspondence with Kac's nomenclature [12, pp. 54–55].

Suppose $\overline{\Phi}$ is a root system of type $X_n = A_{n\geq 1}$, $B_{n\geq 2}$, $C_{n\geq 2}$, $D_{n\geq 2}$, $E_{6,7,8}$, F_4 or G_2, and $\overline{\Lambda}$ is its root lattice. We define the root system Φ of type $\widetilde{X}_n$ as $\overline{\Phi} \times \mathbb{Z} \subseteq \Lambda := \overline{\Lambda} \oplus \mathbb{Z}$. It corresponds to Kac's $X_n^{(1)}$. This can be seen by taking simple roots for $\overline{\Phi}$ and adjoining $(\bar{\alpha}, 1)$ where $\bar{\alpha}$ is the lowest long root of $\overline{\Phi}$.

For $X_n = B_n$, C_n or F_4 we define $\widetilde{X}_n^{\text{even}}$ as the set of $(\bar{\alpha}, m) \in \widetilde{X}_n$ such that $m \equiv 0 \bmod 2$ if $\bar{\alpha}$ is long. These correspond to Kac's $D_{n+1}^{(2)}$, $A_{2n-1}^{(2)}$ and $E_6^{(2)}$ respectively. This can be seen as in the previous paragraph, using the lowest short root instead of the lowest long root.

We define $\widetilde{G}_2^{0 \bmod 3}$ by the same construction, with the condition $m \equiv 0 \bmod 2$ replaced by $m \equiv 0 \bmod 3$. This corresponds to Kac's $D_4^{(3)}$, by the same recipe as the previous paragraph.

For the last affine root system we recall that $BC_{n\geq 2}$ is the union of the B_n and C_n root systems. It is non-reduced and has 3 lengths of roots, called short, middling and long. Taking $\overline{\Phi} = BC_n$, we define the root system Φ of type $\widetilde{BC}_n^{\text{odd}}$ as the set of all $(\bar{\alpha}, m) \in \overline{\Phi} \times \mathbb{Z}$ such that m is odd if $\bar{\alpha}$ is long. Although $\overline{\Phi}$ is non-reduced, Φ is reduced because of this parity condition. Kac's notation is $A_{2n}^{(2)}$. To check this, begin with the set of roots in $\widetilde{BC}_n^{\text{odd}}$ having $m = 0$, which has type B_n, take simple roots for it, and adjoin $(2\bar{\alpha}, 1)$ where $\bar{\alpha}$ is the lowest short root of B_n. The Dynkin diagram of these roots is then Kac's $A_{2n}^{(2)}$.

In all cases, the root system of type $\widetilde{X}_n^{\cdots}$ is the set of pairs (root of X_n, $m \in \mathbb{Z}$), satisfying the condition that if the root is long then m has the property indicated in the superscript.

4. The isomorphism $\mathfrak{PSt}_A(R) \to \mathfrak{St}_A(R)$

This section is devoted to proving theorem 1, whose hypotheses we assume throughout. Our goal is to show that the Chevalley relations for the classically prenilpotent pairs imply those of the remaining prenilpotent pairs. We will begin by saying which pairs of real roots are prenilpotent and which are classically prenilpotent. Then we will analyze the pairs that are prenilpotent but not classically prenilpotent.

We fix the affine Dynkin diagram A, write $\Phi, \overline{\Phi}, \Lambda, \overline{\Lambda}$ as in section 3, and use an overbar to indicate projections of roots from Φ to $\overline{\Phi}$. It is easy to see that $\alpha, \beta \in \Phi$ are classically prenilpotent just if their projections $\bar{\alpha}, \bar{\beta} \in \overline{\Phi}$ are linearly independent. The following lemma describes which pairs of roots are prenilpotent but not classically prenilpotent, and what their Chevalley relations are (except for one special case discussed later).

Lemma 5. *Suppose $\alpha, \beta \in \Phi$ are distinct. Then the following are equivalent:*

(*i*) *α, β are prenilpotent but not classically prenilpotent;*
(*ii*) *$\bar{\alpha}, \bar{\beta}$ differ by a positive scalar factor;*
(*iii*) *$\bar{\alpha}, \bar{\beta}$ are equal, or one is twice the other and $\Phi = \widetilde{BC}_n^{\text{odd}}$.*

When these equivalent conditions hold, the Chevalley relations between $\mathfrak{U}_\alpha, \mathfrak{U}_\beta$ in Tits' definition of $\mathfrak{St}_A(R)$ are $[\mathfrak{U}_\alpha, \mathfrak{U}_\beta] = 1$, unless $\Phi = \widetilde{BC}_n^{\text{odd}}$, $\bar{\alpha}$ and $\bar{\beta}$ are the same short root of $\overline{\Phi} = BC_n$, and $\alpha + \beta \in \Phi$.

Proof. Recall that two roots $\alpha, \beta \in \Phi$ form a prenilpotent pair if some element of the affine Weyl group W sends both to positive roots, and some other element of W sends both to negative roots. We gave simple roots for Φ in section 3, including a choice of simple roots for the subset Φ_0 having $m = 0$. With respect to these, $\alpha = (\bar{\alpha}, m) \in \Phi$ is positive just if either $m > 0$, or $m = 0$ and $\bar{\alpha}$ is positive in Φ_0.

The "translation" part of the affine Weyl group acts on Φ by shears $(\bar{\alpha}, m) \mapsto (\bar{\alpha}, m + \phi(\bar{\alpha}))$, with ϕ a linear function $\overline{\Lambda} \to \mathbb{Z}$. The set of ϕ occurring in this way is a subgroup of finite index in $\mathrm{Hom}(\overline{\Lambda}, \mathbb{Z})$. This makes it easy to see that if $\bar{\alpha}, \bar{\beta}$ are linearly independent, then some shear in W sends α, β to positive roots, and another one sends them to negative roots. The same argument works if $\bar{\alpha}, \bar{\beta}$ differ by a positive scalar factor. So α, β are prenilpotent in these cases.

If $\bar{\alpha}, \bar{\beta}$ differ by a negative scalar factor then some positive linear combination of α, β has the form $(0, m)$. The affine Weyl group acts trivially on the second summand in $\Lambda = \overline{\Lambda} \oplus \mathbb{Z}$. It follows that if some element of W sends α, β to positive (resp. negative) roots, then m is positive (resp. negative). Since m cannot be both positive and negative, α, β cannot be a prenilpotent pair.

We have shown that $\alpha, \beta \in \Phi$ fail to be prenilpotent just if $\bar{\alpha}, \bar{\beta}$ are negative multiples of each other. We remarked above that α, β are classically prenilpotent just if $\bar{\alpha}, \bar{\beta}$ are linearly independent. This proves the equivalence of (*i*) and (*ii*). To see the equivalence of (*ii*) and (*iii*) we refer to the fact that $\overline{\Phi}$ is a reduced root system (i.e, the only positive multiple of a root that can be a root is that root itself) except in the case $\Phi = \widetilde{BC}_n^{\text{odd}}$. In this last case, the only way one root of $\overline{\Phi} = BC_n$ can be a positive multiple of a different root is that the long roots are got by doubling the short roots.

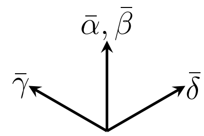
The proof of the final claim is similar. Except in the excluded case, we have $\overline{\Phi} \cap (\mathbb{N}\bar{\alpha} + \mathbb{N}\bar{\beta}) = \{\bar{\alpha}, \bar{\beta}\}$. The corresponding claim for Φ follows, so $\theta(\alpha, \beta) - \{\alpha, \beta\}$ is empty and the right hand side of (2) is the identity. That is, the Chevalley relations for α, β read $[\mathfrak{U}_\alpha, \mathfrak{U}_\beta] = 1$. (In the excluded case we remark that $\Phi \cap (\mathbb{N}\alpha + \mathbb{N}\beta) = \{\alpha, \beta, \alpha + \beta\}$. So the Chevalley relations set the commutators of elements of $\mathfrak{U}_\alpha$ with elements of $\mathfrak{U}_\beta$ equal to certain elements of $\mathfrak{U}_{\alpha+\beta}$. See case 6 below.) □

Recall that $\mathfrak{PSt}_A(R)$ is defined by the Chevalley relations of the classically prenilpotent pairs, and $\mathfrak{St}_A(R)$ by those of all prenilpotent pairs. So to prove theorem 1 it suffices to show that if α, β are prenilpotent but not classically prenilpotent, then their Chevalley relations already hold in $\mathfrak{PSt} := \mathfrak{PSt}_A(R)$.

The cases we must address are given in the lemma above, and in every case but one we must establish $[\mathfrak{U}_\alpha, \mathfrak{U}_\beta] = 1$. Each case begins by choosing two roots in Φ, of which β is a specified linear combination, and whose projections to $\overline{\Phi}$ are specified. Given the global description of Φ from section 3, this is always easy. Then we use the Chevalley relations for various classically prenilpotent pairs to deduce the Chevalley relations for α, β. When necessary we refer to [2, (79)–(92)] for the exact forms of the Chevalley relations.

The proof of theorem 1 now falls into seven cases, according to Φ and the relative position of α and β.

Case 1 of theorem 1. Assume $\bar{\alpha} = \bar{\beta}$ is a root of $\overline{\Phi} = A_{n\geq 2}$, D_n or E_n, or a long root of $\overline{\Phi} = G_2$. Choose $\bar{\gamma}, \bar{\delta} \in \overline{\Phi}$ as shown, and choose lifts $\gamma, \delta \in \Phi$ summing to β. (Choose any $\gamma \in \Phi$ lying over $\bar{\gamma}$, define $\delta = \beta - \gamma$, and use the global description of Φ to check that $\delta \in \Phi$. This is trivial except in the case $\Phi = \widetilde{G}_2^{0 \bmod 3}$, when it is easy.)



Because $\bar{\alpha} + \bar{\gamma}, \bar{\alpha} + \bar{\delta} \notin \overline{\Phi}$, it follows that $\alpha + \gamma, \alpha + \delta \notin \Phi$. So the Chevalley relations $[\mathfrak{U}_\alpha, \mathfrak{U}_\gamma] = [\mathfrak{U}_\alpha, \mathfrak{U}_\delta] = 1$ hold. The Chevalley relations for γ, δ imply $[\mathfrak{U}_\gamma, \mathfrak{U}_\delta] = \mathfrak{U}_{\gamma+\delta} = \mathfrak{U}_\beta$. (These relations are [2, (89)] in the G_2 case and [2, (81)] in the others. One can write them as $[X_\gamma(t), X_\delta] = X_{\gamma+\delta}(tu)$ in the notation of the next paragraph.) Since $\mathfrak{U}_\alpha$ commutes with $\mathfrak{U}_\gamma$ and $\mathfrak{U}_\delta$, it also commutes with $\mathfrak{U}_\beta$, as desired. □

The other cases use the same strategy: express an element of $\mathfrak{U}_\beta$ in terms of other root groups, and then evaluate its commutator with an element of $\mathfrak{U}_\alpha$. But the calculations are more delicate. We will work with explicit elements $X_\gamma(t) \in \mathfrak{U}_\gamma$ for various roots $\gamma \in \Phi$. Here t varies over R, and the definition of $X_\gamma(t)$ depends on a choice of basis vector e_γ for the corresponding root space $\mathfrak{g}_\gamma \subseteq \mathfrak{g}$. One should think of $X_\gamma(t)$ as "$\exp(te_\gamma)$". Following Tits [20, §3.3], there are two equally canonical choices for e_γ, differing by a sign. Changing one's choice corresponds to negating t. This choice of sign is irrelevant to the arguments below, except that a choice is required in order to write down the relations explicitly.

The Chevalley relations in [2] are given in a form originally due to Demazure. For example, if λ, σ are long and short simple roots for a B_2 root system, then their relations are given in [2, (85)] as

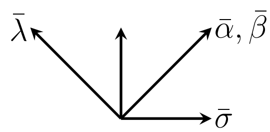
$$[X_\sigma(t), X_\lambda(u)] \cdot S_\sigma X_\lambda(-t^2u) S_\sigma^{-1} \cdot S_\lambda X_\sigma(tu) S_\lambda^{-1} = 1, \tag{3}$$

for all $t, u \in R$. The advantage of this form is technical: to write down the relation, one only needs to specify generators for $\mathfrak{g}_\sigma$ and $\mathfrak{g}_\lambda$, not the other root spaces involved. But for explicit computation one must choose generators for these other root spaces. Because S_σ and S_λ permute the root spaces by the reflections in σ and λ, the second and third terms in (3) lie in $\mathfrak{U}_{\lambda+2\sigma}$ and $\mathfrak{U}_{\lambda+\sigma}$ respectively. Therefore, after choosing suitable generators $e_{\lambda+2\sigma}$ and $e_{\lambda+\sigma}$ for $\mathfrak{g}_{\lambda+2\sigma}$ and $\mathfrak{g}_{\lambda+\sigma}$, we may rewrite (3) as

$$[X_\sigma(t), X_\lambda(u)] \cdot X_{\lambda+2\sigma}(-t^2u) \cdot X_{\lambda+\sigma}(tu) = 1. \tag{4}$$

We could just as well replace $e_{\lambda+2\sigma}$ (resp. $e_{\lambda+\sigma}$) by its negative; then we would also replace $-t^2u$ (resp. tu) by its negative.

Case 2 of theorem 1. Assume $\bar\alpha = \bar\beta$ is a long root of $\overline{\Phi} = B_{n\geq 2}$, $C_{n\geq 2}$, $BC_{n\geq 2}$ or F_4. Our first step is to choose roots $\bar\lambda, \bar\sigma \in \overline{\Phi}$ as pictured:



This is easily done using any standard description of $\overline{\Phi}$. (Note: although $\bar\lambda$ stands for "long" and $\bar\sigma$ for "short", $\bar\sigma$ is actually a middling root in the case $\overline{\Phi} = BC_n$.)

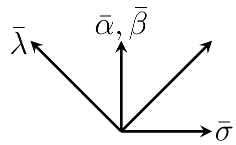
Our second step is to choose lifts $\lambda, \sigma \in \Phi$ of them with $\beta = \lambda + 2\sigma$. If $\Phi = \widetilde{B}_n$, $\widetilde{C}_n$ or $\widetilde{F}_4$ then one chooses any lift σ of $\bar\sigma$ and defines λ as $\beta - 2\sigma$. This works since every element of Λ lying over a root of $\overline{\Phi}$ is a root of Φ. If $\Phi = \widetilde{B}_n^{\text{even}}$, $\widetilde{C}_n^{\text{even}}$, $\widetilde{F}_4^{\text{even}}$ or $\widetilde{BC}_n^{\text{odd}}$ then this argument might fail since Φ is "missing" some long roots. Instead, one chooses any $\lambda \in \Phi$ lying over $\bar\lambda$ and defines σ as $(\beta - \lambda)/2$. Now, $\beta - \lambda = (\bar\beta - \bar\lambda, m)$ with the second entry being even by the meaning of the superscript $^{\text{even}}$ or $^{\text{odd}}$. Also, $\bar\beta - \bar\lambda$ is divisible by 2 in $\overline{\Lambda}$ by the figure above. It follows that $\sigma \in \Lambda$. Then, as an element of Λ lying over a short (or middling) root of $\overline{\Phi}$, σ lies in Φ.

Because σ, λ are simple roots for a B_2 root system inside Φ, their Chevalley relation (4) holds in $\mathfrak{PSt}$. It shows that any element of $\mathfrak{U}_\beta = \mathfrak{U}_{\lambda+2\sigma}$ can be written in the form

$$\big[(\text{some } x_\lambda \in \mathfrak{U}_\lambda), (\text{some } x_\sigma \in \mathfrak{U}_\sigma)\big] \cdot (\text{some } x_{\lambda+\sigma} \in \mathfrak{U}_{\lambda+\sigma}). \tag{5}$$

Referring to the picture of $\overline{\Phi}$ shows that $\alpha + \lambda \notin \Phi$. Therefore the Chevalley relations in $\mathfrak{PSt}$ include $[\mathfrak{U}_\alpha, \mathfrak{U}_\lambda] = 1$. In particular, $\mathfrak{U}_\alpha$ commutes with the first term in the commutator in (5). The same argument shows that $\mathfrak{U}_\alpha$ also commutes with the other terms. This shows that the Chevalley relations present in $\mathfrak{PSt}$ imply $[\mathfrak{U}_\alpha, \mathfrak{U}_\beta] = 1$, as desired. □

Case 3 of theorem 1. Assume $\bar\alpha = \bar\beta$ is a short root of $\overline{\Phi} = B_{n\geq 2}$, $C_{n\geq 2}$ or F_4, or a middling root of $\overline{\Phi} = BC_{n\geq 2}$. We may choose $\lambda, \sigma \in \Phi$ with sum β and the following projections to $\overline{\Phi}$ (by a simpler argument than in the previous case):



The Chevalley relations for σ, λ are (4), showing that any element of $\mathfrak{U}_\beta = \mathfrak{U}_{\sigma+\lambda}$ can be written in the form

$$(\text{some } x_{\lambda+2\sigma} \in \mathfrak{U}_{\lambda+2\sigma}) \cdot \big[(\text{some } x_\lambda \in \mathfrak{U}_\lambda), (\text{some } x_\sigma \in \mathfrak{U}_\sigma)\big]. \tag{6}$$

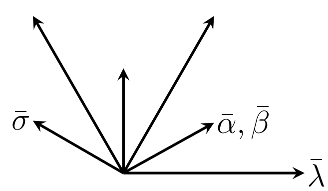
As in the previous case, we will conjugate this by an arbitrary element of $\mathfrak{U}_\alpha$. This requires the following Chevalley relations. We have $[\mathfrak{U}_\alpha, \mathfrak{U}_{\lambda+2\sigma}] = 1$ and $[\mathfrak{U}_\alpha, \mathfrak{U}_\lambda] = 1$ by the same argument as before. What is new is that the Chevalley relations for α, σ depend on whether $\alpha+\sigma$ is a root. If it is, then we get $[\mathfrak{U}_\alpha, \mathfrak{U}_\sigma] \subseteq \mathfrak{U}_{\alpha+\sigma}$, and if not then we get $[\mathfrak{U}_\alpha, \mathfrak{U}_\sigma] = 1$. In the second case we see that $\mathfrak{U}_\alpha$ commutes with (6), proving $[\mathfrak{U}_\alpha, \mathfrak{U}_\beta] = 1$ and therefore finishing the proof.

In the first case, conjugating (6) by a element of $\mathfrak{U}_\alpha$ yields

$$x_{\lambda+2\sigma} \cdot \big[x_\lambda, (\text{some } x_{\alpha+\sigma} \in \mathfrak{U}_{\alpha+\sigma}) \cdot x_\sigma\big]$$

which we can simplify by further use of Chevalley relations. Namely, neither $\lambda+\alpha+\sigma$ nor $\alpha+2\sigma$ is a root, so $\mathfrak{U}_{\alpha+\sigma}$ centralizes $\mathfrak{U}_\lambda$ and $\mathfrak{U}_\sigma$. So $x_{\alpha+\sigma}$ centralizes the other terms in the commutator, hence drops out, leaving (6). This shows that conjugation by any element of $\mathfrak{U}_\alpha$ leaves invariant every element of $\mathfrak{U}_\beta$. That is, $[\mathfrak{U}_\alpha, \mathfrak{U}_\beta] = 1$. □

Case 4 of theorem 1. Assume $\bar\alpha = \bar\beta$ is a short root of $\overline{\Phi} = G_2$. This is the hardest case by far. Begin by choosing roots $\bar\sigma, \bar\lambda \in \overline{\Phi}$ as shown, with lifts $\sigma, \lambda \in \Phi$ summing to β.



Many different root groups appear in the argument, so we choose a generator e_γ of γ's root space, for each $\gamma \in \Phi$ which is a nonnegative linear combination of α, σ, λ.

Next we write down the G_2 Chevalley relations in $\mathfrak{PSt}$ that we will need, derived from [2, (86)–(92)]. We will write them down in the $\Phi = \widetilde{G}_2$ case and then comment on the simplifications that occur if $\Phi = \widetilde{G}_2^{0 \bmod 3}$. After negating some of the e_γ, for γ involving σ and λ

but not α, we may suppose that the Chevalley relations [2, (92)] for σ, λ read

$$[X_\sigma(t), X_\lambda(u)] = X_{\sigma+\lambda}(-tu)X_{2\sigma+\lambda}(t^2u)X_{3\sigma+\lambda}(t^3u)X_{3\sigma+2\lambda}(2t^3u^2). \tag{7}$$

Then we may negate $e_{\alpha+2\sigma+\lambda}$ if necessary, to suppose the Chevalley relations [2, (90)] for $\alpha, 2\sigma+\lambda$ read

$$[X_\alpha(t), X_{2\sigma+\lambda}(u)] = X_{\alpha+2\sigma+\lambda}(3tu). \tag{8}$$

After negating some of the e_γ for γ involving α and σ but not λ, we may suppose that the Chevalley relations [2, (91)] for α and σ read

$$[X_\alpha(t), X_\sigma(u)] = X_{\alpha+\sigma}(-2tu)X_{\alpha+2\sigma}(-3tu^2)X_{2\alpha+\sigma}(-3t^2u) \tag{9}$$

We know the Chevalley relations [2, (90)] for σ and $\alpha+\sigma$ have the form

$$[X_\sigma(t), X_{\alpha+\sigma}(u)] = X_{\alpha+2\sigma}(3\varepsilon tu) \tag{10}$$

where $\varepsilon = \pm 1$. We cannot choose the sign because we've already used our freedom to negate $e_{\alpha+2\sigma}$ in order to get (9). Similarly, we know that the Chevalley relations [2, (89)] for λ and $\alpha+2\sigma$ are

$$[X_\lambda(t), X_{\alpha+2\sigma}(u)] = X_{\alpha+2\sigma+\lambda}(\varepsilon' tu) \tag{11}$$

for some $\varepsilon' = \pm 1$. (We will see at the very end that $\varepsilon = 1$ and $\varepsilon' = -1$.)

We were able to write down these relations because we could work out the roots in the positive span of any two given roots. This used the assumption $\Phi = \widetilde{G}_2$, but now suppose $\Phi = \widetilde{G}_2^{0 \bmod 3}$. It may happen that some of the vectors appearing in the previous paragraph, projecting to long roots of $\overline{\Phi} = G_2$, are not roots of Φ. One can check that if $\alpha - \beta$ is divisible by 3 in Λ then there is no change. On the other hand, if $\alpha - \beta \not\equiv 0 \bmod 3$ then $\alpha + 2\sigma + \lambda$, $\alpha + 2\sigma$ and $2\alpha + \sigma$ are not roots. Because $\alpha + 2\sigma + \lambda$ is not a root, (8) is replaced by $[\mathfrak{U}_\alpha, \mathfrak{U}_{2\sigma+\lambda}] = 1$. And because $(\mathbb{Q}\alpha \oplus \mathbb{Q}\sigma) \cap \Phi$ is now a root system of type A_2 rather than G_2, (9) is replaced by $[X_\alpha(t), X_\sigma(t)] = X_{\alpha+\sigma}(tu)$ and (10) by $[\mathfrak{U}_\sigma, \mathfrak{U}_{\alpha+\sigma}] = 1$. Finally, there is no relation (11) because there is no longer a root group $\mathfrak{U}_{\alpha+2\sigma}$. The calculations below use the relations (7)–(11). To complete the proof, one must also carry out a similar calculation using (7) together with the altered versions of (8)–(10). This calculation is so much easier that we omit it.

Since $\beta = \sigma + \lambda$, we may use (7) with $t = 1$ to express any element of $\mathfrak{U}_\beta$ as

$$X_\beta(u) = X_{2\sigma+\lambda}(u)X_{3\sigma+\lambda}(u)X_{3\sigma+2\lambda}(2u^2)[X_\lambda(u), X_\sigma(1)].$$

We use this to express the commutators generating $[\mathfrak{U}_\alpha, \mathfrak{U}_\beta]$:

$$[X_\alpha(t), X_\beta(u)] = X_\alpha(t)X_{2\sigma+\lambda}(u)X_\alpha(t)^{-1} \cdot X_\alpha(t)X_{3\sigma+\lambda}(u)X_\alpha(t)^{-1}$$

$$\begin{aligned}&\cdot X_\alpha(t)X_{3\sigma+2\lambda}(2u^2)X_\alpha(t)^{-1}\\&\cdot \big[X_\alpha(t)X_\lambda(u)X_\alpha(t)^{-1}, X_\alpha(t)X_\sigma(1)X_\alpha(t)^{-1}\big]\\&\cdot [X_\sigma(1), X_\lambda(u)]\\&\cdot X_{3\sigma+2\lambda}(-2u^2)X_{3\sigma+\lambda}(-u)X_{2\sigma+\lambda}(-u).\end{aligned}$$

By (8) we may rewrite the first term (i.e., before the first dot) as $X_{\alpha+2\sigma+\lambda}(3tu)X_{2\sigma+\lambda}(u)$. By the Chevalley relations $[\mathfrak{U}_\alpha, \mathfrak{U}_{3\sigma+\lambda}] = 1$ we may cancel the $X_\alpha(t)^{\pm 1}$ in the second term. And similarly in the third term, and in the first term of the first commutator. Then we rewrite the second term of that commutator using the Chevalley relations (9). The result is

$$\begin{aligned}[X_\alpha(t), X_\beta(u)] &= X_{\alpha+2\sigma+\lambda}(3tu)X_{2\sigma+\lambda}(u)X_{3\sigma+\lambda}(u)X_{3\sigma+2\lambda}(2u^2)\\&\cdot \Big[X_\lambda(u),\ X_{\alpha+\sigma}(-2t)X_{\alpha+2\sigma}(-3t)X_{2\alpha+\sigma}(-3t^2)X_\sigma(1)\Big]\\&\cdot \big[X_\sigma(1), X_\lambda(u)\big]\cdot X_{3\sigma+2\lambda}(-2u^2)X_{3\sigma+\lambda}(-u)X_{2\sigma+\lambda}(-u).\end{aligned} \tag{12}$$

Our next goal is the rewrite the first commutator $[\cdots,\cdots]$ on the right side. The first simplification is that all terms appearing in it centralize $\mathfrak{U}_{2\alpha+\sigma}$. So we may drop the $X_{2\alpha+\sigma}(-3t^2)$ term. Then we expand the commutator:

$$\begin{aligned}[\cdots,\cdots] &= X_\lambda(u)X_{\alpha+\sigma}(-2t)X_{\alpha+2\sigma}(-3t)X_\sigma(1)\\&\quad\cdot X_\lambda(-u)X_\sigma(-1)X_{\alpha+2\sigma}(3t)X_{\alpha+\sigma}(2t).\end{aligned}$$

We will gather the X_λ and X_σ terms at the right end by repeated use of the Chevalley relations in $\mathfrak{PSt}$. We move $X_\sigma(-1)$ across $X_{\alpha+2\sigma}(3t)$ using $[\mathfrak{U}_\sigma, \mathfrak{U}_{\alpha+2\sigma}] = 1$. Then we move it across $X_{\alpha+\sigma}(2t)$ using the special case

$$X_\sigma(-1)X_{\alpha+\sigma}(2t) = X_{\alpha+2\sigma}(-6\varepsilon t)X_{\alpha+\sigma}(2t)X_\sigma(-1)$$

of (10). The result is

$$\begin{aligned}[\cdots,\cdots] &= X_\lambda(u)X_{\alpha+\sigma}(-2t)X_{\alpha+2\sigma}(-3t)X_\sigma(1)\\&\quad\cdot X_\lambda(-u)X_{\alpha+2\sigma}(3t-6\varepsilon t)X_{\alpha+\sigma}(2t)X_\sigma(-1).\end{aligned}$$

Now we move $X_\lambda(-u)$ across $X_{\alpha+2\sigma}(3t-6\varepsilon t)$ using the special case

$$\begin{aligned}X_\lambda(-u)X_{\alpha+2\sigma}(3t-6\varepsilon t) &= X_{\alpha+2\sigma+\lambda}\big(-3\varepsilon' tu(1-2\varepsilon)\big)\\&\quad\cdot X_{\alpha+2\sigma}(3t-6\varepsilon t)X_\lambda(-u)\end{aligned}$$

of (11). Then we move it further right using $[\mathfrak{U}_\lambda, \mathfrak{U}_{\alpha+\sigma}] = 1$:

$$\begin{aligned}[\cdots,\cdots] &= X_\lambda(u)X_{\alpha+\sigma}(-2t)X_{\alpha+2\sigma}(-3t)X_\sigma(1)\\&\quad\cdot X_{\alpha+2\sigma+\lambda}\big(-3\varepsilon' tu(1-2\varepsilon)\big)X_{\alpha+2\sigma}(3t-6\varepsilon t)X_{\alpha+\sigma}(2t)\end{aligned}$$

$$\cdot X_\lambda(-u)X_\sigma(-1).$$

Now, $X_\sigma(1)$ commutes with the two terms following it, and we appeal to (10) to move it across $X_{\alpha+\sigma}(2t)$:

$$\begin{aligned}[\cdots,\cdots] = X_\lambda(u)X_{\alpha+\sigma}(-2t)X_{\alpha+2\sigma}(-3t)X_{\alpha+2\sigma+\lambda}\bigl(-3\varepsilon' tu(1-2\varepsilon)\bigr)\\ \cdot X_{\alpha+2\sigma}(3t)X_{\alpha+\sigma}(2t)X_\sigma(1)X_\lambda(-u)X_\sigma(-1).\end{aligned}$$

The second through fifth terms commute with each other, leading to much cancellation:

$$[\cdots,\cdots] = X_\lambda(u)X_{\alpha+2\sigma+\lambda}\bigl(-3\varepsilon' tu(1-2\varepsilon)\bigr)X_\sigma(1)X_\lambda(-u)X_\sigma(-1).$$

Since $X_\lambda(u)$ commutes with the term following it, we may rewrite this as

$$[\cdots,\cdots] = X_{\alpha+2\sigma+\lambda}\bigl(-3\varepsilon' tu(1-2\varepsilon)\bigr)[X_\lambda(u), X_\sigma(1)].$$

We plug this into (12) and cancel the commutators to get

$$\begin{aligned}[X_\alpha(t), X_\beta(u)] &= X_{\alpha+2\sigma+\lambda}(3tu)X_{2\sigma+\lambda}(u)X_{3\sigma+\lambda}(u)X_{3\sigma+2\lambda}(2u^2)\\ &\quad\cdot X_{\alpha+2\sigma+\lambda}\bigl(-3\varepsilon' tu(1-2\varepsilon)\bigr)X_{3\sigma+2\lambda}(-2u^2)\\ &\quad\cdot X_{3\sigma+\lambda}(-u)X_{2\sigma+\lambda}(-u).\end{aligned}$$

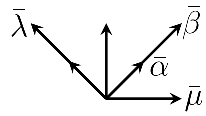
Now, $\alpha+2\sigma+\lambda$ and $3\sigma+2\lambda$ are distinct roots, both projecting to the same long root of $\overline{\Phi}$. In case 1 we established the Chevalley relations in $\mathfrak{PSt}$ for two such roots, so the $X_{\alpha+2\sigma+\lambda}$ term in the middle commutes with the $X_{3\sigma+2\lambda}$ term that precedes it. It centralizes the two terms before that by the Chevalley relations in $\mathfrak{PSt}$. So all the terms on the right cancel except the $X_{\alpha+2\sigma+\lambda}$ terms, leaving

$$\begin{aligned}[X_\alpha(t), X_\beta(u)] &= X_{\alpha+2\sigma+\lambda}\bigl(3tu(1-\varepsilon'+2\varepsilon\varepsilon')\bigr)\\ &= X_{\alpha+2\sigma+\lambda}(Ctu)\end{aligned}$$

where $C = 0$, ± 6 or 12 depending on $\varepsilon, \varepsilon' \in \{\pm 1\}$.

If $C = 0$ (i.e., $\varepsilon = 1$ and $\varepsilon' = -1$) then we have established the desired Chevalley relation $[\mathfrak{U}_\alpha, \mathfrak{U}_\beta] = 1$ and the proof is complete. Otherwise we pass to the quotient $\mathfrak{St}$ of $\mathfrak{PSt}$. Here $\mathfrak{U}_\alpha$ and $\mathfrak{U}_\beta$ commute, so we derive the relation $X_{\alpha+2\sigma+\lambda}(12t) = 1$ in $\mathfrak{St}$. Since this identity holds universally, it holds for $R = \mathbb{C}$, so the image of $\mathfrak{U}_{\alpha+2\sigma+\lambda}(\mathbb{C})$ in $\mathfrak{St}(\mathbb{C})$ is the trivial group. This is a contradiction, since $\mathfrak{St}(\mathbb{C})$ acts on the Kac-Moody algebra $\mathfrak{g}$, with $X_{\alpha+2\sigma+\lambda}(t)$ acting (nontrivially for $t \neq 0$) by $\exp \operatorname{ad}(te_{\alpha+2\sigma+\lambda})$. Since $C \neq 0$ leads to a contradiction, we must have $C = 0$ and so the Chevalley relation $[\mathfrak{U}_\alpha, \mathfrak{U}_\beta] = 1$ holds in $\mathfrak{PSt}$. $\square$

Case 5 of theorem 1. Assume $\bar\beta = 2\bar\alpha$ *in* $\overline{\Phi} = BC_{n\geq 2}$. Choose $\bar\mu, \bar\lambda \in \overline{\Phi}$ as shown, and lift them to $\mu, \lambda \in \Phi$ with $\mu + \lambda = \beta$. (Mnemonic: μ is middling and λ is long.)

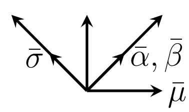


As in the case 2 (when $\bar\alpha$ and $\bar\beta$ were the same long root of $\overline{\Phi} = B_n$), we can express any element of $\mathfrak{U}_\beta$ in the form

$$\big[(\text{some } x_\lambda \in \mathfrak{U}_\lambda), (\text{some } x_\mu \in \mathfrak{U}_\mu)\big] \cdot (\text{some } x_{\mu+\lambda} \in \mathfrak{U}_{\mu+\lambda}).$$

Among the Chevalley relations in $\mathfrak{PSt}$ is the commutativity of $\mathfrak{U}_\alpha$ with $\mathfrak{U}_\lambda$, $\mathfrak{U}_\mu$ and $\mathfrak{U}_{\mu+\lambda}$. So $\mathfrak{U}_\alpha$ also centralizes $\mathfrak{U}_\beta$. □

Case 6 of theorem 1. Assume $\bar\alpha = \bar\beta$ *is a short root of* $\overline{\Phi} = BC_{n\geq 2}$ *and* $\alpha + \beta$ *is a root.* This is the exceptional case of lemma 5, and the Chevalley relation we must establish is not $[\mathfrak{U}_\alpha, \mathfrak{U}_\beta] = 1$. We will determine the correct relation during the proof. We begin by choosing $\bar\mu, \bar\sigma \in \overline{\Phi}$ as shown and lifting them to $\mu, \sigma \in \Phi$ with $\mu + \sigma = \beta$, so σ, μ generate a B_2 root system.



We choose a generator e_γ for the root space of each nonnegative linear combination $\gamma \in \Phi$ of α, σ, μ. By changing the signs of $e_{\sigma+\mu}$ and $e_{2\sigma+\mu}$ if necessary, we may suppose that the Chevalley relations [2, (85)] for σ, μ are

$$[X_\sigma(t), X_\mu(u)] = X_{\sigma+\mu}(-tu) X_{2\sigma+\mu}(t^2 u), \tag{13}$$

Since $\sigma + \mu = \beta$ we may take $t = 1$ in (13) to express any element of $\mathfrak{U}_\beta$:

$$X_\beta(u) = X_{2\sigma+\mu}(u)[X_\mu(u), X_\sigma(1)]. \tag{14}$$

Using this one can express any generator for $[\mathfrak{U}_\alpha, \mathfrak{U}_\beta]$:

$$\begin{aligned} [X_\alpha(t), X_\beta(u)] = X_\alpha(t) X_{2\sigma+\mu}(u) X_\alpha(t)^{-1} & \\ \cdot \big[X_\alpha(t) X_\mu(u) X_\alpha(t)^{-1}, X_\alpha(t) X_\sigma(1) X_\alpha(t)^{-1}\big] & \\ \cdot [X_\sigma(1), X_\mu(u)] \cdot X_{2\sigma+\mu}(-u). & \end{aligned} \tag{15}$$

By the Chevalley relations $[\mathfrak{U}_\alpha, \mathfrak{U}_{2\sigma+\mu}] = [\mathfrak{U}_\alpha, \mathfrak{U}_\mu] = 1$, the $X_\alpha(t)^{\pm 1}$'s cancel in the first term and in the first term of the first commutator.

Now we consider the Chevalley relations of α and σ. Since $\bar{\alpha}+\bar{\sigma}$ is a middling root of $\overline{\Phi}$, and Φ contains every element of Λ lying over every such root, we see that $\alpha+\sigma$ is a middling root of Φ. In particular, $(\mathbb{Q}\alpha \oplus \mathbb{Q}\sigma) \cap \Phi$ is a B_2 root system, in which α and σ are orthogonal short roots. The Chevalley relations [2, (84)] for α, σ are therefore

$$[X_\alpha(t), X_\sigma(u)] = X_{\alpha+\sigma}(-2tu), \tag{16}$$

after changing the sign of $e_{\alpha+\sigma}$ if necessary.

Next, $\mu+\sigma+\alpha = \alpha+\beta$ is a root by hypothesis. We choose $e_{\mu+\sigma+\alpha}$ so that the Chevalley relations [2, (84)] for μ, $\alpha+\sigma$ are

$$[X_\mu(t), X_{\alpha+\sigma}(u)] = X_{\mu+\alpha+\sigma}(-2tu). \tag{17}$$

Now we rewrite (15), applying the cancellations mentioned above and rewriting the second term in the first commutator using (16):

$$\begin{aligned}[X_\alpha(t), X_\beta(u)] = X_{2\sigma+\mu}(u) &\cdot \big[X_\mu(u), X_{\alpha+\sigma}(-2t)X_\sigma(1)\big] \\ &\cdot [X_\sigma(1), X_\mu(u)] \cdot X_{2\sigma+\mu}(-u).\end{aligned} \tag{18}$$

We write out the first commutator on the right side and use the Chevalley relations $[\mathfrak{U}_{\alpha+\sigma}, \mathfrak{U}_\sigma] = 1$ and (17) to obtain

$$\begin{aligned}\big[X_\mu(u), X_{\alpha+\sigma}(-2t)X_\sigma(1)\big] &= X_\mu(u)X_{\alpha+\sigma}(-2t) \cdot X_\sigma(1) \\ &\quad \cdot X_\mu(-u)X_\sigma(-1)X_{\alpha+\sigma}(2t) \\ &= X_{\mu+\alpha+\sigma}(4tu)X_{\alpha+\sigma}(-2t)X_\mu(u) \cdot X_\sigma(1) \\ &\quad \cdot X_{\mu+\alpha+\sigma}(4tu)X_{\alpha+\sigma}(2t)X_\mu(-u)X_\sigma(-1).\end{aligned}$$

Since $\overline{\mu+\alpha+\sigma} = 2\bar{\sigma}$ is a long root of $\overline{\Phi} = BC_n$, we know from case 5 that $\mathfrak{U}_\sigma$ and $\mathfrak{U}_{\mu+\alpha+\sigma}$ commute in $\mathfrak{PSt}$. And $\mathfrak{U}_{\mu+\alpha+\sigma}$ centralizes all the other terms by Chevalley relations in $\mathfrak{PSt}$. So we may gather the $X_{\mu+\alpha+\sigma}(4tu)$ terms at the beginning. Next, $[\mathfrak{U}_\sigma, \mathfrak{U}_{\alpha+\sigma}] = 1$, so we may move $X_\sigma(1)$ to the right across $X_{\alpha+\sigma}(2t)$. Then we can use (17) again to move $X_\mu(u)$ rightward across $X_{\alpha+\sigma}(2t)$. The result is

$$\begin{aligned}\big[X_\mu(u), X_{\alpha+\sigma}(-2t)X_\sigma(1)\big] &= X_{\mu+\alpha+\sigma}(8tu)X_{\alpha+\sigma}(-2t) \\ &\quad \cdot X_{\mu+\sigma+\alpha}(-4tu)X_{\alpha+\sigma}(2t)[X_\mu(u), X_\sigma(1)] \\ &= X_{\mu+\alpha+\sigma}(4tu)[X_\mu(u), X_\sigma(1)].\end{aligned}$$

Plugging this into (18) and canceling the commutators gives

$$\begin{aligned}[X_\alpha(t), X_\beta(u)] &= X_{2\sigma+\mu}(u)X_{\mu+\alpha+\sigma}(4tu)X_{2\sigma+\mu}(-u) \\ &= X_{\alpha+\beta}(4tu).\end{aligned}$$

Tits' Chevalley relation in his definition of $\mathfrak{St}$ has the same form, with the factor 4 replaced by some integer C. If $C \neq 4$ then in $\mathfrak{St}$ we deduce $X_{\alpha+\beta}\big((C-4)tu\big) = 1$ for all $t, u \in R$ and all rings R, leading to the same contradiction we found in case 4. Therefore $C = 4$ and we have established that Tits' relation already holds in $\mathfrak{PSt}$. □

Case 7 of theorem 1. Assume $\bar{\alpha} = \bar{\beta}$ is a short root of $\overline{\Phi} = BC_{n\geq 2}$ and $\alpha + \beta$ is not a root. This is similar to the previous case but much easier. We choose μ, σ and the e_γ in the same way, except that $\mu + \sigma + \alpha$ is no longer a root, so the Chevalley relation (17) is replaced by $[\mathfrak{U}_\mu, \mathfrak{U}_{\alpha+\sigma}] = 1$. We expand $X_\beta(u)$ as in (14) and obtain (18) as before. But this time the $X_{\alpha+\sigma}(-2t)$ term centralizes both $\mathfrak{U}_\mu$ and $\mathfrak{U}_\sigma$, so it vanishes from the commutator. The right side of (18) then collapses to 1 and we have proven $[\mathfrak{U}_\alpha, \mathfrak{U}_\beta] = 1$ in $\mathfrak{PSt}$. □

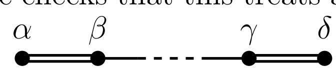
5. Finite presentations

In this section we prove theorem 3, that various Steinberg and Kac-Moody groups are finitely presented. At the end we make several remarks about possible variations on the definition of Kac-Moody groups.

Proof of theorem 3. We must show that $\mathfrak{St}_A(R)$ is finitely presented under either of the two stated hypotheses. By theorem 1 it suffices to prove this with $\mathfrak{PSt}$ in place of $\mathfrak{St}$.

(ii) We are assuming $\operatorname{rk} A \geq 3$ and that R is finitely generated as a module over a subring generated by finitely many units. Theorem 1.4(ii) of [2] shows that if R satisfies this hypothesis and A is 2-spherical, then $\mathfrak{PSt}_A(R)$ is finitely presented. This proves (ii).

(i) Now we are assuming $\operatorname{rk} A > 3$ and that R is finitely generated as a ring. Theorem 1.4(iii) of [2] gives the finite presentability of $\mathfrak{PSt}_A(R)$ if every pair of nodes of the Dynkin diagram lies in some irreducible spherical diagram of rank ≥ 3. By inspecting the list of affine Dynkin diagrams of rank > 3, one checks that this treats all cases of (i) except

$A =$ 

(with some orientations of the double edges). In this case, no irreducible spherical diagram contains α and δ.

For this case we use a variation on the proof of theorem 1.4(iii) of [2]. Consider the direct limit G of the groups $\mathfrak{St}_B(R)$ as B varies over all irreducible spherical diagrams of rank ≥ 2. If $\operatorname{rk} B \geq 3$ then $\mathfrak{St}_B(R)$ is finitely presented by theorem I of Splitthoff [17]. If $\operatorname{rk} B = 2$ then $\mathfrak{St}_B(R)$ is finitely generated by [2, thm. 11.5]. Since every irreducible rank 2 diagram lies in one of rank > 2, it follows that G is finitely presented. Now, G satisfies all the relations of $\mathfrak{St}_A(R)$ except for the commutativity of $\mathfrak{St}_{\{\alpha\}}$ with $\mathfrak{St}_{\{\delta\}}$. Because these groups may not be finitely generated, we might need infinitely many additional relations to impose commutativity in the obvious way.

So we proceed indirectly. Let Y_α be a finite subset of $\mathfrak{St}_{\{\alpha\}}$ which together with $\mathfrak{St}_{\{\beta\}}$ generates $\mathfrak{St}_{\{\alpha,\beta\}}$. This is possible since $\mathfrak{St}_{\{\alpha,\beta\}}$ is finitely generated. We define Y_δ similarly, with γ in place of β. We define H as the quotient of G by the finitely many relations $[Y_\alpha, Y_\delta] = 1$, and claim that the images in H of $\mathfrak{St}_{\{\alpha\}}$ and $\mathfrak{St}_{\{\delta\}}$ commute.

The following computation in H establishes this. First, every element of Y_δ centralizes $\mathfrak{St}_{\{\beta\}}$ by the definition of G, and every element of Y_α by definition of H. Therefore it centralizes $\mathfrak{St}_{\{\alpha,\beta\}}$, hence $\mathfrak{St}_{\{\alpha\}}$. We've shown that $\mathfrak{St}_{\{\alpha\}}$ centralizes Y_δ, and it centralizes $\mathfrak{St}_{\{\gamma\}}$ by the definition of G. Therefore it centralizes $\mathfrak{St}_{\{\gamma,\delta\}}$, hence $\mathfrak{St}_{\{\delta\}}$.

H has the same generators as $\mathfrak{PSt}_A(R)$, and its defining relations are among those defining $\mathfrak{PSt}_A(R)$. On the other hand, we have shown that the generators of H satisfy all the relations in $\mathfrak{PSt}_A(R)$. So $H \cong \mathfrak{PSt}_A(R)$. In particular, $\mathfrak{PSt}_A(R)$ is finitely presented.

It remains to prove the finite presentability of $\mathfrak{G}_A(R)$ under the extra hypothesis that the unit group of R is finitely generated as an abelian group. This follows immediately from [2, Lemma 11.2], which says that for any generalized Cartan matrix A, and any commutative ring R with finitely generated unit group, the kernel of $\mathfrak{St}_A(R) \to \mathfrak{G}_A(R)$ is finitely generated. □

Remark (Completions). We have worked with the "minimal" or "algebraic" forms of Kac-Moody groups. One can consider various completions of it, such as those surveyed in [19]. None of these completions can possibly be finitely presented, so no analogue of theorem 3 exists. But it is reasonable to hope for an analogue of corollary 2.

Remark (Chevalley-Demazure group schemes). If A is spherical then we write $\mathfrak{CD}_A$ for the associated Chevalley-Demazure group scheme, say the simply-connected version. This is the unique most natural (in a certain technical sense) algebraic group over $\mathbb{Z}$ of type A. The kernel of $\mathfrak{St}_A(R) \to \mathfrak{CD}_A(R)$ is called $K_2(A; R)$ and contains the relators (1). There is considerable interest in when $\mathfrak{CD}_A(R)$ is finitely presented, for example [5][6]. We want to emphasize that our theorem 3 does *not* resolve this question, because $\mathfrak{CD}_A(R)$ may be a proper quotient of $\mathfrak{G}_A(R)$. Indeed, $K_2(A; R)$ can be extremely complicated.

For a non-spherical Dynkin diagram A, the functor $\mathfrak{CD}_A$ is not defined. The question of whether there is a good definition, and what it would be, seems to be completely open. Only when R is a field is there known to be a unique "best" definition of a Kac-Moody group [20, theorem 1(i), p. 553]. The main problem would be to specify what extra relations to impose on $\mathfrak{G}_A(R)$. The remarks below discuss the possible forms of some additional relations.

Remark (Kac-Moody groups over integral domains). If R is an integral domain with fraction field k, then it is open whether $\mathfrak{G}_A(R) \to \mathfrak{G}_A(k)$ is injective. If $\mathfrak{G}_A$ satisfies Tits' axioms then this would follow from (KMG4), but Tits does not assert that $\mathfrak{G}_A$ satisfies his axioms. If $\mathfrak{G}_A(R) \to \mathfrak{G}_A(k)$ is not injective, then the image seems better than $\mathfrak{G}_A(R)$ itself as a candidate for the right Kac-Moody group.

Remark (Kac-Moody groups via representations). Fix a root datum D and a commutative ring R. By using Kostant's $\mathbb{Z}$-form of the universal enveloping algebra of $\mathfrak{g}$, one can construct a $\mathbb{Z}$-form $V_{\mathbb{Z}}^{\lambda}$ of any integrable highest-weight module V^{λ} of $\mathfrak{g}$. Then one defines V_R^{λ} as $V_{\mathbb{Z}}^{\lambda} \otimes R$. For each real root α, one can exponentiate $\mathfrak{g}_{\alpha,\mathbb{Z}} \otimes R \cong R$ to get an action of $\mathfrak{U}_\alpha \cong R$ on V_R^{λ}. One can define the action of the torus $(R^*)^n$ directly. Then one can take the group $\mathfrak{G}_D^{\lambda}(R)$ generated by these transformations and call it a Kac-Moody group. This approach is extremely natural and not yet fully worked out. The first such work for Kac-Moody groups over rings is Garland's landmark paper [11] treating affine groups; see also Tits' survey [19, §5], its references, and the recent articles [4][7].

Tits [20, p. 554] asserts that this construction allows one to build a Kac-Moody functor satisfying all his axioms (KMG1)–(KMG9). We imagine that he reasoned as follows. First, show that each $\mathfrak{G}_D^{\lambda}$ is a Kac-Moody functor and therefore by Tits' theorem admits a canonical functorial homomorphism from $\mathfrak{G}_A$, where A is the generalized Cartan matrix of D. (One cannot directly apply Tits' theorem, because $\mathfrak{G}_D^{\lambda}(R)$ only comes equipped with the homomorphisms $\mathrm{SL}_2(R) \to \mathfrak{G}_D^{\lambda}(R)$ required by Tits when $\mathrm{SL}_2(R)$ is generated by its subgroups $\left(\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 & 0 \\ * & 1 \end{smallmatrix}\right)$. Presumably this difficulty can be overcome.) Second, define I as the intersection of the kernels of all the homomorphisms $\mathfrak{G}_A \to \mathfrak{G}_D^{\lambda}$, and then define the desired Kac-Moody functor as $\mathfrak{G}_A/I$. (This also does not quite make sense, since $\mathfrak{G}_A$ may also lack the required homomorphisms from SL_2. As before, presumably this difficulty can be overcome.)

Remark (Loop groups). Suppose X is one of the $ABCDEFG$ diagrams, $\widetilde{X}$ is its affine extension as in section 4, and R is a commutative ring. The well-known description of affine Kac-Moody algebras and loop groups makes it natural to expect that $\mathfrak{G}_{\widetilde{X}}(R)$ is a central extension of $\mathfrak{G}_X(R[t^{\pm 1}])$ by R^*. The most general results along these lines that I know of are Garland's theorems 10.1 and B.1 in [11], although they concern slightly different groups. Instead, one might simply *define* the loop group $G_{\widetilde{X}}(R)$ as a central extension of $\mathfrak{CD}_X(R[t^{\pm1}])$ by R^*, where the 2-cocycle defining the extension would have to be made explicit. Then one could try to show that $G_{\widetilde{X}}$ satisfies Tits' axioms.

It is natural to ask whether such a group $G_{\widetilde{X}}(R)$ would be finitely presented if R is finitely generated. If R^* is finitely generated then this is equivalent to the finite presentation of the quotient $\mathfrak{CD}_X(R[t^{\pm1}])$. If $\operatorname{rk} X \geq 3$ then $\mathfrak{St}_X(R[t^{\pm1}])$ is finitely presented by Splitthoff's theorem I of [17]. Then, as Splitthoff explains in [17, §7], the finite presentability of $\mathfrak{CD}_X(R[t^{\pm1}])$ boils down to properties of $K_1(X, R[t^{\pm1}])$ and $K_2(X, R[t^{\pm1}])$.

References

[1] Abramenko, P. and Mühlherr, B., Preśentations de certaines BN-paires jumelées comme sommes amalgamées, *C.R.A.S. Série I* **325** (1997) 701–706.

[2] Allcock, Steinberg groups as amalgams, preprint *arXiv:1307.2689v2*.

[3] Allcock, D. and Carbone, L., Presentation of hyperbolic Kac-Moody groups over rings, preprint *arXiv:1409.5918*.

[4] Bao, L. and Carbone, L., Integral forms of Kac-Moody groups and Eisenstein series in low dimensional supergravity theories, *arXiv:1308.6194v1*.

[5] Behr, H., Über die endliche Definierbarkeit verallgemeinerter Einheitsgruppen II, *Invent. Math.* **4** (1967) 265–274.

[6] Behr, H., Arithmetic groups over function fields I, *J. reine angew. Math.* bf 495 (1998) 79–118.

[7] Carbone, L. and Garland, H., Infinite dimensional Chevalley groups and Kac-Moody groups over $\mathbb{Z}$, in preparation (2013).

[8] Caprace, P.-E., On 2-spherical Kac-Moody groups and their central extensions, *Forum Math.* **19** (2007) 763–781.

[9] Eskin, A., and Farb, B., Quasi-flats and rigidity in higher rank symmetric spaces, *J.A.M.S.* **10** (1997) 653–692.

[10] Farb, B. and Schwartz, R., The large-scale geometry of Hilbert modular groups, *J. Diff. Geom.* **44** (1996) 435–478.

[11] Garland, H. The arithmetic theory of loop groups, *Publ. Math. I.H.E.S.* **52** (1980) 5–136.

[12] Kac, V., *Infinite dimensional Lie algebras* (3rd edition), Cambridge University Press, Cambridge, 1990.

[13] Kleiner, B. and Leeb, B., Rigidity of quasi-isometries for symmetric spaces and Euclidean buildings, *Publ. math. I.H.E.S.* **86** (1997) 115–197.

[14] Morita, J. and Rehmann, U., A Matusumoto-type theorem for Kac-Moody groups, *Tôhoku Math. J.* **42** (1990) 537–560.

[15] Muhlherr, B., On the simple connectedness of a chamber system associated to a twin building, unpublished (1999).

[16] Schwartz, R., The quasi-isometry classification of rank one lattices, *Publ. Math. I.H.E.S.* **82** (1995) 133–168.

[17] Splitthoff, S., Finite presentability of Steinberg groups and related Chevalley groups. in *Applications of algebraic K-theory to algebraic geometry and number*

theory, Part I, II (Boulder, Colo., 1983), pp. 635–687, *Contemp. Math.* **55**, Amer. Math. Soc., Providence, RI, 1986.

[18] Steinberg, R., *Lectures on Chevalley groups*. Notes prepared by John Faulkner and Robert Wilson, Yale University, New Haven, Conn., 1968.

[19] Tits, J., Groups and group functors attached to Kac-Moody data, in *Workshop Bonn 1984*, pp. 193–223, *Lecture Notes in Math.* **1111**, Springer, Berlin, 1985.

[20] Tits, J., Uniqueness and presentation of Kac-Moody groups over fields, *J. Algebra* **105** (1987) 542–573.

Department of Mathematics, University of Texas, Austin
E-mail address: `allcock@math.utexas.edu`
URL: `http://www.math.utexas.edu/~allcock`