Journal of Algebra 322 (2009) 703–721

Contents lists available at ScienceDirect

Journal of Algebra

www.elsevier.com/locate/jalgebra

Computing Chevalley bases in small characteristics

Arjeh M. Cohen *, Dan Roozemond

Department of Mathematics and Computer Science, Technische Universiteit Eindhoven, PO Box 513, 5600 MB Eindhoven, Netherlands

ARTICLE INFO

Article history:
Received 9 January 2009
Available online 19 May 2009
Communicated by Jon Carlson

On the occasion of the distinguished birthdays of our distinguished colleagues John Cannon and Derek Holt

Keywords:
Lie algebras
Chevalley bases
Conjugacy problems
Isomorphism problems
Toral subalgebras
Algorithms
Groups of Lie type

ABSTRACT

Let L be the Lie algebra of a simple algebraic group defined over a field $\mathbb{F}$ and let H be a split maximal toral subalgebra of L. Then L has a Chevalley basis with respect to H. If $\mathrm{char}(\mathbb{F}) \neq 2, 3$, it is known how to find it. In this paper, we treat the remaining two characteristics. To this end, we present a few new methods, implemented in Magma, which vary from the computation of centralizers of one root space in another to the computation of a specific part of the Lie algebra of derivations of L.

1. Introduction

1.1. The main result

For computational problems regarding a split reductive algebraic group G defined over a field $\mathbb{F}$, it is often useful to calculate within its Lie algebra L over $\mathbb{F}$. For instance, the conjugacy question for two split maximal tori in G can often be translated to a conjugacy question for two split Cartan subalgebras of L. Here, a *Cartan subalgebra* H of L is understood to be a maximal toral subalgebra, that is, it is commutative, left multiplication by each of its elements is semisimple (i.e., has a diagonal form with respect to a suitable basis over a large enough extension field of $\mathbb{F}$), and it is maximal (with respect to inclusion) among subalgebras of L with these properties; it is called *split* (or $\mathbb{F}$-*split*)

* Corresponding author.
E-mail addresses: a.m.cohen@tue.nl (A.M. Cohen), d.a.roozemond@tue.nl (D. Roozemond).

0021-8693/$ – see front matter
doi:10.1016/j.jalgebra.2009.04.038

if left multiplication by h, denoted ad_h, has a diagonal form with respect to a suitable basis over $\mathbb{F}$ for every $h \in H$. Such a Cartan subalgebra is the Lie algebra of a split maximal torus in G.

The conjugacy question mentioned above can be answered by finding Chevalley bases with respect to each split Cartan subalgebra, so the transformation from one basis to the other is an automorphism of L, and subsequently adjusting the automorphism with the normalizer of one Cartan subalgebra so as to obtain an element of G. In this light, it is of importance to have an algorithm finding a Chevalley basis (see Section 1.3 for a precise definition). Such algorithms have been discussed for the case where the characteristic of the underlying field is distinct from 2 and from 3. However, the latter two characteristics are the most important ones for finite simple groups arising from algebraic groups, so there is a need for dealing with these special cases as well. This is taken care of by the following theorem, which is the main result of this paper.

Theorem 1. *Let L be the Lie algebra of a split simple algebraic group of rank n defined over an effective field $\mathbb{F}$. Suppose that H is an $\mathbb{F}$-split Cartan subalgebra of L. If L is given by means of a multiplication table with respect to an $\mathbb{F}$-basis of L and H is given by means of a spanning set, then there is a Las Vegas algorithm that finds a Chevalley basis of L with respect to H. If $\mathbb{F} = \mathbb{F}_q$, this algorithm needs $O^{\sim}(n^{10}\log(q)^4)$ elementary operations.*

Here $O^{\sim}(N)$ means $O(N(\log(N))^c)$ for some constant c. Recall (e.g., from [17, Introduction]) that arithmetic operations in $\mathbb{F}$ are understood to be addition, subtraction, multiplication, division, and equality testing. If $\mathbb{F}$ is the field $\mathbb{F}_q$ of size q, these all take $O^{\sim}(\log(q))$ elementary operations. Performing standard linear algebra arithmetic, that is, operations on matrices of size m, like multiplication, determinant, and kernel (solving linear equations), takes $O(m^3)$ arithmetic operations.

Better estimates than those of the theorem are conceivable, for instance because better bounds on matrix multiplication exist. However, our primary goal was to establish that the algorithm is polynomial in $n\log(q)$. Moreover, in comparison to the dimension $O(n^2)$ of L or the estimate $O(n^6)$ for arithmetic operations needed for multiplying two elements of L, or, for that matter, the size of the input $(O(n^3\log(q)))$, the high exponent of n in the timing looks more reasonable than may seem at first sight.

The proof of Theorem 1 rests on Algorithm 1, which is really an outline of an algorithm further specified in the course of the paper. The algorithm is implemented in Magma [3]. We intend to make the implementation public as a Magma package once the code has been cleaned up.

The algorithm is mostly deterministic. However, in some instances where $\mathbb{F}$ is of characteristic 2 (such as method $[\mathrm{B}_2^{\mathrm{sc}}]$ and the case where L is of type D_4; see Sections 3.3, 3.5, and 3.6) we use the Meat-axe (cf. [11,12]) for finding a particular submodule of a given module. For finite fields, the Meat-axe algorithm is analyzed in [15] and [11, Section 2]: irreducible submodules of a finite L-module of dimension m over $\mathbb{F}_q$ can be found in Las Vegas time $O^{\sim}(m^3\log(q))$ (in Section 1.6 below it is explained why this result can be applied to Lie algebras). For infinite fields, Meat-axe procedures are known; however, we know of no proof of polynomiality in the literature.

Algorithm 1 assumes that besides L and H the root datum R of the underlying group is known (see Section 1.2). However, in Section 5 we show that this root datum can be determined by running the algorithm a small number of times.

Thanks to the characterization of Lie algebras of split reductive algebraic groups described in Theorem 2 below, we can view the Lie algebras in Theorem 1 as Chevalley Lie algebras, which are defined below.

Our treatment of Lie algebras and the corresponding algebraic groups rests on the theory developed mainly by Chevalley and available in the excellent books by Borel [1], Humphreys [13], and Springer [18]. Our computational set-up is as in [7].

1.2. Root data

Split reductive algebraic groups are determined by their fields of definition and their root data. The latter is of importance to the corresponding Lie algebra and will therefore be discussed first. Throughout this paper we let $R = (X, \Phi, Y, \Phi^\vee)$ be a *root datum of rank n* as defined in [7]. This means X and Y are dual free $\mathbb{Z}$-modules of dimension n with a bilinear pairing $\langle\cdot,\cdot\rangle : X \times Y \to \mathbb{Z}$; furthermore, Φ is a finite subset of X and $\Phi^\vee$ a finite subset of Y, called the *roots* and *coroots*, respectively, and there is a one-to-one correspondence ${}^\vee : \Phi \to \Phi^\vee$ such that $\langle \alpha, \alpha^\vee \rangle = 2$ for all $\alpha \in \Phi$.

If $\alpha \in \Phi$ then $s_\alpha : x \mapsto x - \langle x, \alpha^\vee \rangle \alpha$ is a reflection on X leaving Φ invariant and $W = \langle s_\alpha \mid \alpha \in \Phi \rangle$ is a Coxeter group. Similarly, $s_\alpha^\vee : y \mapsto y - \langle \alpha, y \rangle \alpha^\vee$ is a reflection on Y leaving $\Phi^\vee$ invariant and the group generated by all these is isomorphic to W. In particular there are $\alpha_1, \dots, \alpha_l \in \Phi$, linearly independent in $X \otimes \mathbb{Q}$, such that $\Phi = \Phi^+ \dot\cup \Phi^-$, where $\Phi^+ = \Phi \cap (\mathbb{N}\alpha_1 + \cdots + \mathbb{N}\alpha_l)$ and $\Phi^- = -\Phi^+$. The roots $\alpha_1, \dots, \alpha_l$ and the coroots $\alpha_1^\vee, \dots, \alpha_l^\vee$ are called *simple*. The number l is called the *semisimple rank* of L (and of G).

The pair (W, S), where $S = \{s_{\alpha_1}, \dots, s_{\alpha_l}\}$, is a Coxeter system. The *Cartan matrix* C of R is the $l \times l$ matrix whose (i, j) entry is $\langle \alpha_i, \alpha_j^\vee \rangle$. The matrix C is related to the Coxeter type of (W, S) as follows: $s_{\alpha_i} s_{\alpha_j}$ has order m_{ij} where

$$\cos\left(\frac{\pi}{m_{ij}}\right)^2 = \frac{\langle \alpha_i, \alpha_j^\vee \rangle \langle \alpha_j, \alpha_i^\vee \rangle}{4}.$$

The *Coxeter matrix* is $(m_{ij})_{1 \leqslant i,j \leqslant l}$ and the *Coxeter diagram* is a graph-theoretic representation thereof: it is a graph with vertex set $\{1, \dots, l\}$ whose edges are the pairs $\{i, j\}$ with $m_{ij} > 2$; such an edge is labeled m_{ij}. The Cartan matrix C determines the Dynkin diagram (and vice versa). For, the *Dynkin diagram* is the Coxeter diagram with the following extra information about root lengths: $\langle \alpha_i, \alpha_j^\vee \rangle < \langle \alpha_j, \alpha_i^\vee \rangle$ if and only if the Coxeter diagram edge $\{i, j\}$ (labeled m_{ij}) is replaced by the directed edge (i, j) in the Dynkin diagram (so that the arrow head serves as a mnemonic for the inequality sign indicating that the root length of α_i is larger than the root length of α_j).

A root datum is called *irreducible* if its Coxeter diagram is connected. A root datum is called *semisimple* if its rank is equal to its semisimple rank. Each semisimple root datum can be decomposed uniquely into irreducible root data. The Dynkin diagrams of irreducible root systems are well known, and described in Cartan's notation A_n $(n \geqslant 1)$, B_n $(n \geqslant 2)$, C_n $(n \geqslant 3)$, D_n $(n \geqslant 4)$, E_n $(n \in \{6, 7, 8\})$, F_4, G_2. The nodes are usually labeled as in [2].

For computations, we fix $X = Y = \mathbb{Z}^n$ and set $\langle x, y \rangle = xy^\top$, which is an element of $\mathbb{Z}$ since x and y are row vectors. Now take A to be the integral $l \times n$ matrix containing the simple roots as row vectors; this matrix is called the *root matrix* of R. Similarly, let B be the $l \times n$ matrix containing the simple coroots; this matrix is called the *coroot matrix* of R. Then $C = AB^\top$ and $\mathbb{Z}\Phi = \mathbb{Z}A$ and $\mathbb{Z}\Phi^\vee = \mathbb{Z}B$. For $\alpha \in \Phi$ we define c^α to be the $\mathbb{Z}$-valued size l row vector satisfying $\alpha = c^\alpha A$. In the greater part of this paper, including Theorems 1 and 2, we will let G be a simple group, so $l = n$.

1.3. Chevalley Lie algebras

Given a root datum R we consider the free $\mathbb{Z}$-module

$$L_{\mathbb{Z}}(R) = Y \oplus \bigoplus_{\alpha \in \Phi} \mathbb{Z} X_\alpha,$$

where the X_α are formal basis elements. The rank of $L_{\mathbb{Z}}(R)$ is $n + |\Phi|$. We denote by $[\cdot,\cdot]$ the alternating bilinear map $L_{\mathbb{Z}}(R) \times L_{\mathbb{Z}}(R) \to L_{\mathbb{Z}}(R)$ determined by the following rules:

$$\text{for } y, z \in Y: \qquad [y, z] = 0, \tag{CB$\mathbb{Z}$1}$$

$$\text{for } y \in Y, \alpha \in \Phi: \qquad [X_\alpha, y] = \langle \alpha, y \rangle X_\alpha, \tag{CB$\mathbb{Z}$2}$$

$$\text{for } \alpha \in \Phi: \qquad [X_{-\alpha}, X_\alpha] = \alpha^\vee, \tag{CB$\mathbb{Z}$3}$$

$$\text{for } \alpha, \beta \in \Phi, \alpha \neq \pm\beta: \qquad [X_\alpha, X_\beta] = \begin{cases} N_{\alpha,\beta} X_{\alpha+\beta} & \text{if } \alpha + \beta \in \Phi, \\ 0 & \text{otherwise.} \end{cases} \tag{CB$\mathbb{Z}$4}$$

The $N_{\alpha,\beta}$ are integral structure constants chosen to be $\pm(p_{\alpha,\beta}+1)$, where $p_{\alpha,\beta}$ is the biggest number such that $\alpha - p_{\alpha,\beta}\beta$ is a root and the signs are chosen (once and for all) so as to satisfy the Jacobi identity. It is easily verified that $N_{\alpha,\beta} = -N_{-\alpha,-\beta}$ and it is a well-known result (see for example [4]) that such a product exists. $L_{\mathbb{Z}}(R)$ is called a *Chevalley Lie algebra*.

A basis of $L_{\mathbb{Z}}(R)$ that consists of a basis of Y and the formal elements X_α and satisfies (CB$\mathbb{Z}$1)–(CB$\mathbb{Z}$4) is called a *Chevalley basis* of the Lie algebra $L_{\mathbb{Z}}(R)$ with respect to the split Cartan subalgebra Y and the root datum R. If no confusion is imminent we just call this a *Chevalley basis* of $L_{\mathbb{Z}}(R)$.

For the remainder of this section, we let $L_{\mathbb{Z}}(R)$ be a Chevalley Lie algebra with root datum R, we fix $X = Y = \mathbb{Z}^n$, a basis of row vectors $e_1,\dots,e_n$ of X, and a basis of row vectors $f_1,\dots,f_n$ of Y dual to $e_1,\dots,e_n$ with respect to the pairing $\langle\cdot,\cdot\rangle$. Moreover, we let $\mathbb{F}$ be a field, we set $h_i = y_i \otimes 1$, $i = 1,\dots,n$, and $H = Y \otimes \mathbb{F}$. Now tensoring $L_{\mathbb{Z}}(R)$ with $\mathbb{F}$ yields a Lie algebra denoted $L_{\mathbb{F}}(R)$ over $\mathbb{F}$, and the integral Chevalley basis relations (CB$\mathbb{Z}$1)–(CB$\mathbb{Z}$4) can be rephrased as:

$$\text{for } i,j \in \{1,\dots,n\}: \qquad [h_i,h_j] = 0, \tag{CB1}$$

$$\text{for } i \in \{1,\dots,n\}, \alpha \in \Phi: \qquad [X_\alpha,h_i] = \langle\alpha, f_i\rangle X_\alpha, \tag{CB2}$$

$$\text{for } \alpha \in \Phi: \qquad [X_{-\alpha},X_\alpha] = \sum_{i=1}^{n} \langle e_i,\alpha^\vee\rangle h_i, \tag{CB3}$$

$$\text{for } \alpha,\beta \in \Phi, \alpha \neq \pm\beta: \qquad [X_\alpha,X_\beta] = \begin{cases} N_{\alpha,\beta}X_{\alpha+\beta} & \text{if } \alpha+\beta \in \Phi, \\ 0 & \text{otherwise.} \end{cases} \tag{CB4}$$

A Lie algebra is called *split* if it has a split Cartan subalgebra. The Cartan subalgebra H of each Chevalley Lie algebra $L_{\mathbb{F}}(R)$ is split. The image of a Chevalley basis with respect to Y and R in $L_{\mathbb{F}}(R)$ is called a *Chevalley basis of L with respect to H and R*. The interest in Chevalley Lie algebras comes from the following result.

Theorem 2. *(See Chevalley [5].) Suppose that L is the Lie algebra of a split simple algebraic group G over $\mathbb{F}$ with root datum $R = (X,\Phi,Y,\Phi^\vee)$, and that H is a split Cartan subalgebra of L. Then $L \cong L_{\mathbb{F}}(R)$ and so it has a Chevalley basis with respect to H and R. Furthermore, any two split Cartan subalgebras of L are conjugate under G. Finally, if G is simple then R is irreducible.*

In light of this theorem, for the proof of Theorem 1, it suffices to deal with Chevalley Lie algebras corresponding to an irreducible root datum.

1.4. Some difficulties

So we will deal with the construction of a Chevalley basis for a Chevalley Lie algebra L over a field $\mathbb{F}$, given only a split Cartan subalgebra H. Algorithms for finding such an H have been constructed by the first author and Murray [6] and, independently, Ryba [16]. These algorithms work for char($\mathbb{F}$) distinct from 2 and 3, and partly for char($\mathbb{F}$) = 3. The first algorithm has been implemented in the Magma computer algebra system [3]. For now, we assume that we are also given the appropriate irreducible root datum R, but in Section 5 we argue that R can be found from L and H without much effort. The output of our algorithm is an ordered basis $\{X_\alpha, h_i \mid \alpha \in \Phi, i \in \{1,\dots,n\}\}$ of $L_{\mathbb{F}}$ (based on some ordering of the elements of Φ) satisfying (CB1)–(CB4).

For fields of characteristic distinct from 2, 3, an algorithm for finding Chevalley bases given split Cartan subalgebras has been implemented in several computer algebra systems, for example Magma [3] and GAP [8]. For details, see for example [9, Section 5.11]; the algorithm CanonicalGenerators described there produces a Chevalley basis only up to scalars. The scaling, however, can be accomplished by straightforwardly solving linear equations.

If, however, we consider Lie algebras of simple algebraic groups over a field $\mathbb{F}$ of characteristic 2 or 3, the current algorithms break down in several places. Firstly, the root spaces (joint eigenspaces) of the split Cartan subalgebra H acting on L are no longer necessarily one-dimensional. This means that we will have to take extra measures in order to identify which vectors in these root spaces are root elements. This problem will be dealt with in Section 3. Secondly, we can no longer always use root chains to compute Cartan integers $\langle\alpha,\beta^\vee\rangle$, which are the most important piece of information for the root identification algorithm in the general case. We will deal with this problem in Section 4. Thirdly, when computing the Chevalley basis elements for nonsimple roots, we cannot always obtain $X_{\alpha+\beta}$ from (CB4) by $X_{\alpha+\beta} = \frac{1}{N_{\alpha,\beta}}[X_\alpha, X_\beta]$ as $N_{\alpha,\beta}$ may be a multiple of $\mathrm{char}(\mathbb{F})$. This problem, however, is easily dealt with by using a different order of the roots, so we will not discuss this any further.

1.5. Roots

Let p be zero or a prime and suppose for the remainder of this section that $\mathbb{F}$ is a field of characteristic p. We fix a root datum $R = (X, \Phi, Y, \Phi^\vee)$ and write $L = L_{\mathbb{F}}(R)$. We define roots and their multiplicities in L as follows. A *root of H on L* is the function

$$\overline{\alpha} : h \mapsto \sum_{i=1}^{n} \langle\alpha, y_i\rangle t_i, \quad \text{where } h = \sum_{i=1}^{n} y_i \otimes t_i = \sum_{i=1}^{n} t_i h_i,$$

for some $\alpha \in \Phi$; here $\langle\alpha, y_i\rangle$ is interpreted in $\mathbb{Z}$ (if $p = 0$) or $\mathbb{Z}/p\mathbb{Z}$ (if $p \neq 0$). Note that this implies that $\langle\alpha, h\rangle := \overline{\alpha}(h)$ for $h \in H$ is completely determined by the values $\langle\alpha, y_i\rangle$, $i = 1, \dots, n$. We write $\Phi(L, H)$ for the set of roots of H on L.

For $\overline{\alpha} \in \Phi(L, H)$ we define the *root space corresponding to $\overline{\alpha}$* to be

$$L_{\overline{\alpha}} = \bigcap_{i=1}^{n} \mathrm{Ker}\big(\mathrm{ad}_{h_i} - \overline{\alpha}(h_i)\big).$$

If $\overline{\alpha} \neq 0$ for all $\alpha \in \Phi$ then L is a direct sum of H and its root spaces $\{L_{\overline{\alpha}} \mid \alpha \in \Phi\}$. If on the other hand there exists an $\alpha \in \Phi$ such that $\overline{\alpha} = 0$, then L is a direct sum of $L_0 = \mathrm{C}_L(H)$ and $\{L_{\overline{\alpha}} \mid \alpha \in \Phi, \overline{\alpha} \neq 0\}$.

Given a root α, we define the *multiplicity* of α in L to be the number of $\beta \in \Phi$ such that $\overline{\alpha} = \overline{\beta}$. Observe that if $\overline{\alpha} \neq 0$ the multiplicity of $\overline{\alpha} \in \Phi(L, H)$ is equal to $\dim(L_{\overline{\alpha}})$. If $\overline{\alpha} = 0$ this multiplicity is equal to $\dim(L_0) - n$. Note that $\alpha \mapsto \overline{\alpha}$ is a surjective map $\Phi \to \Phi(L, H)$, so in what follows we abbreviate $\Phi(L, H)$ to $\overline{\Phi}$.

If each root has multiplicity 1, there is a bijection between $\overline{\Phi}$ and Φ. Our first order of business is to decide in which cases higher multiplicities occur. Observe that $\overline{\alpha} = 0$ if and only if $\overline{-\alpha} = 0$ so the multiplicity of the 0-root space is never 1. If $\mathrm{char}(\mathbb{F}) = 2$, then all nonzero multiplicities are at least 2 as $\overline{\alpha}$ and $\overline{-\alpha}$ coincide. Steinberg [19, Sections 5.1, 7.4] studied part of the classification of Chevalley Lie algebras L for which higher multiplicities occur (the simply connected case with Dynkin type A_n, D_n, $\mathrm{E}_{6,7,8}$ if $\mathrm{char}(\mathbb{F}) = 2$) in a search for all Lie algebras L with $\mathrm{Aut}(L/Z(L))$ strictly larger than G. In Section 2 of this paper we prove the following proposition, which generalizes Steinberg's result to arbitrary root data. As the multiplicity of a root of H on the Lie algebra L of a central product of split reductive linear algebraic groups is equal to the minimum over all multiplicities of its restrictions to summands of the corresponding central sum decomposition of L, the study of multiplicities of roots can easily be reduced to the case where G is simple.

Proposition 3. *Let L be the Lie algebra of a split simple algebraic group over a field $\mathbb{F}$ of characteristic p with root datum R. Then the multiplicities of the roots in $\overline{\Phi}$ are either all* 1 *or as indicated in Table* 1.

In Table 1, the Dynkin type R of L and the characteristic p of $\mathbb{F}$ are indicated by $R(p)$ in the first column. Further details regarding the table (such as the isogeny type of R appearing as a superscript on $R(p)$) are explained in the beginning of Section 2. This description uniquely determines the root datum R and hence the corresponding connected algebraic group G up to isomorphism; see [18, Chapter 9].

Table 1
Multidimensional root spaces.

$R(p)$	Mults	Soln	$R(p)$	Mults	Soln
$\mathrm{A}_2^{\mathrm{sc}}(3)$	3^2	[Der]	$\mathrm{C}_n^{\mathrm{ad}}(2)$ $(n \geqslant 3)$	$2n, 2^{n(n-1)}$	[C]
$\mathrm{G}_2(3)$	$1^6, 3^2$	[C]	$\mathrm{C}_n^{\mathrm{sc}}(2)$ $(n \geqslant 3)$	$\mathbf{2n}, 4^{\binom{n}{2}}$	[$\mathrm{B}_2^{\mathrm{sc}}$]
$\mathrm{A}_3^{\mathrm{sc},(2)}(2)$	4^3	[Der]	$\mathrm{D}_4^{(1),(n-1),(n)}(2)$	4^6	[Der]
$\mathrm{B}_2^{\mathrm{ad}}(2)$	$2^2, 4$	[C]	$\mathrm{D}_4^{\mathrm{sc}}(2)$	8^3	[Der]
$\mathrm{B}_n^{\mathrm{ad}}(2)$ $(n \geqslant 3)$	$2^n, 4^{\binom{n}{2}}$	[C]	$\mathrm{D}_n^{(1)}(2)$ $(n \geqslant 5)$	$4^{\binom{n}{2}}$	[Der]
$\mathrm{B}_2^{\mathrm{sc}}(2)$	$\mathbf{4}, 4$	[$\mathrm{B}_2^{\mathrm{sc}}$]	$\mathrm{D}_n^{\mathrm{sc}}(2)$ $(n \geqslant 5)$	$4^{\binom{n}{2}}$	[Der]
$\mathrm{B}_3^{\mathrm{sc}}(2)$	6^3	[Der]	$\mathrm{F}_4(2)$	$2^{12}, 8^3$	[C]
$\mathrm{B}_4^{\mathrm{sc}}(2)$	$2^4, 8^3$	[Der]	$\mathrm{G}_2(2)$	4^3	[Der]
$\mathrm{B}_n^{\mathrm{sc}}(2)$ $(n \geqslant 5)$	$2^n, 4^{\binom{n}{2}}$	[C]	all remaining(2)	$2^{\lvert\Phi^+\rvert}$	[A_2]

ChevalleyBasis
in: The Lie algebra L over a field $\mathbb{F}$ of a split reductive algebraic group,
a split Cartan subalgebra H of L, and
a root datum $R = (X, \Phi, Y, \Phi^\vee)$.
out: A Chevalley basis B for L with respect to H and R.
begin
let $E, \overline{\Phi} =$ FindRootSpaces(L, H),
let $\mathcal{X} =$ FindFrame$(L, H, R, \overline{\Phi}, E)$,
let $\iota =$ IdentifyRoots$(L, H, R, \overline{\Phi}, \mathcal{X})$,
let $X^0, H^0 =$ ScaleToBasis$(L, H, R, \mathcal{X}, \iota)$,
return X^0, H^0.
end

Algorithm 1. Finding a Chevalley basis.

1.6. Computing ideals of Lie algebras

Finding an ideal I of a given Lie algebra L is equivalent to finding the submodule I of the A-module L, where A is the associative subalgebra of $\mathrm{End}(L)$ generated by all ad_x for x running over a basis of L. Hence, such an ideal I can be found by application of the Meat-axe algorithm to the A-module L. We will apply the Meat-axe only to modules of bounded dimension, so that the factor n^6 resulting from the occurrence of $\dim(L)^3$ in the above-mentioned estimate for the Meat-axe running time when $\mathbb{F} = \mathbb{F}_q$ plays no role in the asymptotic timing analysis.

1.7. Algorithm 1

In the remainder of this section we give a brief overview of the inner workings of Algorithm 1. It is assumed that L is isomorphic to $L_{\mathbb{F}}(R)$. The FindRootSpaces algorithm consists of simultaneous diagonalization of L with respect to $\mathrm{ad}_{h_1}, \ldots, \mathrm{ad}_{h_n}$, where $\{h_1, \ldots, h_n\}$ is a basis of H. Its output is a basis E of H-eigenvectors of L and the set $\overline{\Phi}$ of roots of H on L. This is feasible over $\mathbb{F}$ because the elements are semisimple and H is split. As $\dim(L) = O(n^2)$, these operations need time $O^{\sim}(n^6 \log(q))$ for each basis element of H, so the total cost is $O^{\sim}(n^7 \log(q))$ elementary operations.

The algorithm called FindFrame is more involved, and solves the difficulties mentioned in Section 1.4 by various methods. The output $\mathcal{X}$ is a *Chevalley frame*, that is, a set of the form $\{\mathbb{F}X_\alpha \mid \alpha \in \Phi\}$, where X_α $(\alpha \in \Phi)$ belong to a Chevalley basis of L with respect to H and R. If all multiplicities are 1 then FindFrame is trivial, meaning that $\mathcal{X} = \{\mathbb{F}x \mid x \in E \setminus H\}$ is the required result. The remaining cases are identified by Proposition 3, and the algorithms for these cases are indicated by [A_2], [C], [Der], [$\mathrm{B}_2^{\mathrm{sc}}$] in Table 1 and explained in Section 3.

In IdentifyRoots we compute Cartan integers and use these to make the identification ι between the root system Φ of R and the Chevalley frame $\mathcal{X}$ computed previously. This identification is again made on a case-by-case basis depending on the root datum R. See Section 4 for details.

The algorithm ends with ScaleToBasis where the vectors X_α $(\alpha \in \Phi)$ belonging to members of the Chevalley frame $\mathcal{X}$ are picked in such a way that $X^0 = (X_\alpha)_{\alpha\in\Phi}$ is part of a Chevalley basis with respect to H and R, and a suitable basis $H^0 = \{h_1, \ldots, h_n\}$ of H is computed, so that they satisfy the Chevalley basis multiplication rules. This step involves the solving of several systems of linear equations, similar to the procedure explained in [6], which takes time $O^{\sim}(n^8 \log(q))$.

Finally, in Section 5, we finish the proof of Theorem 1 and discuss some further problems for which our algorithm may be of use.

2. Multidimensional root spaces

In this section we prove Proposition 3, but first we explain the notation in Table 1. As already mentioned, the first column contains the root datum R specified by means of the Dynkin type with a superscript for the isogeny type, as well as (between parentheses) the characteristic p. A root datum of type A_3 can have any of three isogeny types: adjoint, simply connected, or an intermediate one, corresponding to the subgroup of order 1, 4, and 2 of its fundamental group $\mathbb{Z}/4\mathbb{Z}$, respectively. We denote the intermediate type by $\mathrm{A}_3^{(2)}$. For computations we fix root and coroot matrices for each isomorphism class of root data, as indicated at the end of Section 1.2. For A_3, for example, the Cartan matrix is

$$C = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}.$$

As always, for the adjoint isogeny type $\mathrm{A}_3^{\mathrm{ad}}$ the root matrix A is equal to the identity matrix I and the coroot matrix B is equal to C. Similarly, for $\mathrm{A}_3^{\mathrm{sc}}$ we have $A = C$ and $B = I$. For the intermediate case $\mathrm{A}_3^{(2)}$ for instance, we take

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & 0 \\ 0 & -1 & 1 \end{pmatrix}.$$

It is straightforward to check that indeed $\det(A) = 2 = \det(B)$ and $AB^\top = C$.

A root datum of type D_n has fundamental group isomorphic to $\mathbb{Z}/4\mathbb{Z}$ if n is odd, and to $(\mathbb{Z}/2\mathbb{Z})^2$ if n is even. The unique intermediate type in the odd case is denoted by $\mathrm{D}_n^{(1)}$, and the three possible intermediate types in the even case by $\mathrm{D}_n^{(1)}$, $\mathrm{D}_n^{(n-1)}$, and $\mathrm{D}_n^{(n)}$.

The multiplicities appear in the second column under Mults. Those shown in bold correspond to the root 0. For instance, for $\mathrm{B}_2^{\mathrm{sc}}(2)$ we have $\dim(\mathrm{C}_L(H)) = 6$, so the multiplicity equals $6 - 2 = 4$.

The third column, with header Soln, indicates the method chosen by our algorithm. Further details appear later, in Section 3.

Assume the setting of Proposition 3. By Theorem 2 there is an irreducible root datum $R = (X, \Phi, Y, \Phi^\vee)$ such that $L = \mathrm{Lie}(G)$ satisfies $L \cong L_{\mathbb{F}}(R)$. Also, all split Cartan subalgebras H of L are conjugate under G, so the multiplicities of $L_{\mathbb{F}}(R)$ do not depend on the choice of H. For the proof of the proposition, there is no harm in identifying L with $L_{\mathbb{F}}(R)$ and H with the Lie algebra of a fixed split maximal torus of G.

As all multiplicities are known to be 1 if $\mathrm{char}(\mathbb{F}) = 0$, we will assume that $p := \mathrm{char}(\mathbb{F})$ is a prime. We will write $\equiv$ for equality mod p (to prevent confusion we will sometimes add: mod p). We begin with two lemmas.

Lemma 4. *Let $\alpha, \beta \in \Phi$. Then $\bar\alpha = \bar\beta$ if and only if $(c^\alpha - c^\beta)A \equiv 0$.*

Proof. For $h \in H$, by definition, $\langle \alpha, h\rangle = \langle c^\alpha A, h\rangle = c^\alpha A h^\top$. This implies that $\bar\alpha = \bar\beta$ if and only if $c^\alpha A h^\top \equiv c^\beta A h^\top$ for all $h \in H$, which is equivalent to $(c^\alpha - c^\beta)A \equiv 0$. $\square$

Lemma 5. *Let R_1, R_2 be irreducible root data of the same rank and with the same Cartan matrix C and denote their root matrices by A_1 and A_2, respectively.*

(i) *If* $\det(A_2)$ *strictly divides* $\det(A_1)$, *then the multiplicities in $L_{\mathbb{F}}(R_1)$ are greater than or equal to those in $L_{\mathbb{F}}(R_2)$.*
(ii) *If $p \nmid \det(C)$, then the multiplicities of $L_{\mathbb{F}}(R_1)$ and $L_{\mathbb{F}}(R_2)$ are the same.*

Proof. (i) Without loss of generality, we identify the ambient lattices X and Y with $\mathbb{Z}^n$ and choose the same bilinear pairing (as in Section 1.2) for each of the two root data R_1 and R_2. The condition that $\det(A_2)$ strictly divides $\det(A_1)$ then implies that the columns of A_1 belong to the lattice spanned by the columns of A_2. Hence $A_1 = A_2 M$ for a certain integral $n \times n$ matrix M. Thus $(c^\alpha - c^\beta)A_2 \equiv 0$ implies $(c^\alpha - c^\beta)A_1 \equiv (c^\alpha - c^\beta)A_2 M \equiv 0$, proving the lemma in view of Lemma 4.

(ii) As $\det(C) \not\equiv 0$, the determinants of the coroot matrices B_1 and B_2 are nonzero modulo p, and $A_1 = A_2(B_2 B_1^{-1})$ and $A_2 = A_1(B_1 B_2^{-1})$. It follows that $(c^\alpha - c^\beta)A_2 \equiv 0$ is equivalent to $(c^\alpha - c^\beta)A_1 \equiv 0$. $\square$

A typical case where part (i) of this lemma can be applied is when the adjoint and simply connected case have the same multiplicities, for then every intermediate type will have those multiplicities as well. It immediately follows from Lemma 5 that the root space dimensions are biggest in the simply connected case, and least in the adjoint case. Thus considering root data of the adjoint and simply connected isogeny types often suffices to understand the intermediate cases. Part (ii) indicates that in many cases even one isogeny type will do.

The proof of Proposition 3 follows a division of cases according to the different Dynkin types of the root datum R. For each type, we need to determine when distinct roots α, β exist in Φ such that $\overline{\alpha} = \overline{\beta}$. By Lemma 5(ii), there are deviations from the adjoint case only if p divides $\det(C)$.

As W embeds in $N_G(H)/T$, and acts equivariantly on Φ and $\overline{\Phi} = \Phi(L, H)$, the multiplicity of a root $\overline{\alpha} \in \overline{\Phi}$ only depends on the W-orbit of $\alpha \in \Phi$. By transitivity of the Weyl group on roots of the same length in Φ, it suffices to consider only $\alpha = \alpha_1$ in the cases where all roots in Φ have the same length ($\mathrm{A}_n, \mathrm{D}_n, \mathrm{E}_{6,7,8}$) and $\alpha = \alpha_1$ or α_n if there are multiple root lengths ($\mathrm{B}_n, \mathrm{C}_n, \mathrm{F}_4, \mathrm{G}_2$).

In the adjoint cases, the simple roots $\alpha_1, \ldots, \alpha_n$ are the standard basis vectors $e_1, \ldots, e_n$, since then the root matrix A and the coroot matrix B are I and $C^\top$, respectively. Similarly, in the simply connected cases, the simple roots $\alpha_1, \ldots, \alpha_n$ are the rows of the Cartan matrix C, since then $A = C$ and $B = I$. We write $c = c^\beta$ so $\beta = cA$ and either all $c_i \in \mathbb{N}$ or all $c_i \in -\mathbb{N}$.

We give the proofs of the cases where R is of type A_n, B_n, or G_2. The other cases are proved in a similar way. For V a linear subspace of L and $x \in L$, we write $\mathrm{C}_V(x)$ for the null space of ad_x on V, i.e.,

$$\mathrm{C}_V(x) := \{ v \in V \mid [x, v] = 0 \}.$$

2.1. A_n $(n \geqslant 1)$

The root datum of type A_n has Cartan matrix

$$C = \begin{pmatrix} 2 & -1 & 0 & \ldots & 0 \\ -1 & 2 & -1 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & -1 & 2 & -1 \\ 0 & \ldots & 0 & -1 & 2 \end{pmatrix},$$

and the roots are

$$\pm(\alpha_j + \cdots + \alpha_k), \quad j \in \{1, \ldots, n\},\ k \in \{j, \ldots, n\},$$

where $\{\alpha_1, \ldots, \alpha_n\}$ are the simple roots, thus giving a total of $2 \cdot \frac{1}{2} n(n+1)$ roots.

For the adjoint case, suppose $\overline{\alpha_1} = \overline{\beta}$. Observe that all $c_i \in \{0, \pm 1\}$. Since $A = I$, we must have $c_1 \equiv 1$ and $c_j \equiv 0$ ($j = 2, \ldots, n$), which implies either $p \neq 2$, $c_1 = 1$, and $c_2 = \cdots = c_n = 0$, or $p = 2$, $c_1 = \pm 1$, and $c_2 = \cdots = c_n = 0$. Since we assumed $\beta \neq \alpha_1$ we find $p = 2$ and $\beta = -\alpha_1$, giving $\frac{n^2+n}{2}$ root spaces of dimension 2.

In the simply connected case the simple roots are equal to the rows of C, so that $\overline{\alpha_1} = \overline{\beta}$ implies $2c_1 - c_2 \equiv 2$, $-c_1 + 2c_2 - c_3 \equiv -1$, $-c_{j-2} + 2c_{j-1} - c_j \equiv 0$ for $j = 4, \ldots, n$, and $-c_{n-1} + 2c_n \equiv 0$.

We distinguish three possibilities: $c_1 = 1$, $c_1 = 0$, and $c_1 = -1$. If $c_1 = 1$, then $c_2 \equiv 0$, so $c_2 = 0$. As $c_1\alpha_1 + \cdots + c_n\alpha_n$ must be a root, this implies $c_3 = \cdots = c_n = 0$, forcing $\overline{\beta} = \overline{\alpha_1}$, a contradiction.

If $c_1 = 0$, then $-c_2 \equiv 2$, so that either $p = 2$ and $c_2 = 0$, or $p = 3$ and $c_2 = 1$. In the first case, we find $c_3 \equiv 1$, giving a contradiction if $n \geqslant 5$ (because then $c_4 \equiv 0$ and $c_5 \equiv 1$), a contradiction if $n = 4$ (because then the last relation becomes $0 = -c_3 + 2c_4$, which is not satisfied). Consequently, $n = 3$ and $p = 2$; the resulting case is discussed below. In the second case, where $p = 3$ and $c_2 = 1$, we find $-1 \equiv 2 - c_3$, so that $c_3 \equiv 0$, giving a contradiction if $n \geqslant 4$ (because then $c_4 \equiv 1$), a contradiction if $n = 3$ (because then the last relation becomes $0 = -c_2 + 2c_3$, which is not satisfied). It follows that $n = 2$ and $p = 3$; this case is also discussed below.

If $c_1 = -1$, then $-c_2 \equiv 4$, so that either $p = 2$ and $c_2 = 0$, or $p = 3$ and $c_2 = -1$. In the first case, we find $c_3 = \cdots = c_n = 0$, so $\beta = -\alpha_1$. In the second case, we find that either $n = 2$ (the special case below), or $c_3 = 0$, which leads to a contradiction if $n \geqslant 4$ (because then $c_3 = 0$ but $c_4 \neq 0$), and also if $n = 3$ (because then the last equation becomes $0 = -c_2 + 2c_3$).

We next determine the multiplicities in the two cases found to occur for $\mathrm{A}_n^{\mathrm{sc}}$. For $n = 3$ and $p = 2$ we have

$$A = C = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \equiv \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \mod 2.$$

This gives $\overline{\alpha_1} = \overline{\alpha_3}$, as well as $\overline{\alpha_1 + \alpha_2} = \overline{\alpha_2 + \alpha_3}$ and $\overline{\alpha_2} = \overline{\alpha_1 + \alpha_2 + \alpha_3}$, accounting for 3 root spaces of dimension 4.

For $n = 2$ and $p = 3$ we have

$$A = C = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \equiv \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix} \mod 3,$$

which implies $\overline{\alpha_1} = \overline{\alpha_2}$ and $\overline{\alpha_1} = \overline{-(\alpha_1 + \alpha_2)}$. Similarly, $\overline{-\alpha_1} = \overline{-\alpha_2} = \overline{\alpha_1 + \alpha_2}$, giving 2 root spaces of dimension 3.

For the intermediate cases observe that by Lemma 5(i) we need only consider $(n, p) = (2, 3)$ and $(3, 2)$. But the former case has no intermediate isogeny types, and the latter case is readily checked to be as stated. This finishes the proof for the A_n case.

2.2. B_n ($n \geqslant 2$)

The root datum of type B_n has Cartan matrix

$$C = \begin{pmatrix} 2 & -1 & 0 & \ldots & 0 \\ -1 & 2 & -1 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & -1 & 2 & -2 \\ 0 & \ldots & 0 & -1 & 2 \end{pmatrix},$$

and the roots are

$$\pm(\alpha_j+\cdots+\alpha_l), \qquad j\in\{1,\dots,n\}, l\in\{j,\dots,n\},$$
$$\pm(\alpha_j+\cdots+\alpha_{l-1}+2\alpha_l+\cdots+2\alpha_n), \qquad j\in\{1,\dots,n-1\}, l\in\{j+1,\dots,n\},$$

giving a total of $2\cdot\frac{1}{2}n(n+1)+2\cdot\frac{1}{2}n(n-1)=2n^2$ roots.

In the adjoint case we have $A=I$. For the long roots, suppose $\overline{\alpha_1}=\overline{\beta}$, so $c_1\equiv 1$ and $c_2\equiv\cdots\equiv c_n\equiv 0$. If $c_1=1$, then $c_2\neq 0$ (for otherwise $\beta=\alpha_1$), which implies $p=2$ and $\beta=\alpha_1+2\alpha_2+\cdots+2\alpha_n$. If $c_1=-1$, then $p=2$, and either $c_2=0$, which gives $\beta=-\alpha_1$, or $c_2\neq 0$, which implies $\beta=-\alpha_1-2\alpha_2-\cdots-2\alpha_n$. In this case the long roots have multiplicities 4.

In the adjoint case, for the short roots, suppose $\overline{\alpha_n}=\overline{\beta}$, so $c_n\equiv 1$ and $c_1\equiv\cdots\equiv c_{n-1}\equiv 0$. This yields three possibilities for c_n: If $c_n=-2$, then $p=3$, implying c_{n-1} is either 0 or -3, neither of which give rise to roots. If $c_n=-1$, then $p=2$; now either $c_{n-1}=0$ (yielding $\beta=-\alpha_n$), or $c_{n-1}=-2$ (not giving any roots). If $c_n=1$ we must have $c_{n-1}=\cdots=c_1=0$, giving the contradiction $\beta=\alpha_n$. This shows that $p=2$ and all multiplicities are 2.

In the simply connected case we have $A=C$. We will first consider $n\geqslant 5$, and then treat $n=2,3,4$ separately. By Lemma 5(ii), we may assume $p=2$.

For the long roots, suppose $\overline{\alpha_1}=\overline{\beta}$, so $c_2\equiv 0$, $c_1+c_3\equiv 1$, and $c_{j-2}+c_j\equiv 0$ $(j=4,\dots,n)$. This forces $c_4\equiv 0$. If $c_1\equiv 0$ then $c_1=0$ and hence $c_2=0$, so $c_3=\pm 1$. Replacing β by β if needed, we may assume $c_3=1$. As $c_4\equiv 0$ and $c_5\equiv 1$, we must have $c_4=2$ and $c+5=1$, which is never satisfied by a root. If on the other hand $c_1\equiv 1$ then $c_3\equiv c_4\equiv\cdots\equiv c_n\equiv 0$, so $\beta=-\alpha_1$ or $\beta=\pm(\alpha_1+2\alpha_2+\cdots+2\alpha_n)$. This shows that, for $n\geqslant 5$, the multiplicities of $\overline{\beta}$ for β a long root are 4.

For the short roots, suppose $\overline{\alpha_n}=\overline{\beta}$, so $c_2\equiv 0$, $c_{j-2}+c_j\equiv 0$ $(j=3,\dots,n-1)$, and $c_{n-2}+c_n\equiv 1$. If $c_1\equiv 1$ then $c_3\equiv 1$, but since $c_2\equiv 0$ this contradicts that β is a root. If on the other hand $c_1\equiv 0$, then $c_2\equiv c_3\equiv\cdots\equiv c_{n-1}\equiv 0$, so $c_n\equiv 1$ and we find $\beta=-\alpha_n$. Hence, for $n\geqslant 5$, the multiplicities of $\overline{\beta}$ for β a short root are 2.

If $n=2$ then

$$C=\begin{pmatrix}2&-2\\-1&2\end{pmatrix}\equiv\begin{pmatrix}0&0\\1&0\end{pmatrix}.$$

If $\overline{\alpha_1}=\overline{\beta}$ we have $c_2\equiv 0$. Since $-2\leqslant c_2\leqslant 2$ we must have either $c_2=0$ (hence $\beta=-\alpha_1$), or $c_2=\pm 2$ (hence $c_1=\pm 1$), giving $\beta=\pm\alpha_1$ or $\beta=\pm(\alpha_1+2\alpha_2)$. If on the other hand $\overline{\alpha_2}=\overline{\beta}$ we find $c_2\equiv 1$ hence $\beta=\pm\alpha_2$ or $\beta=\pm(\alpha_1+\alpha_2)$. This shows that B_2^{sc} has 2 root spaces of dimension 4 if $p=2$.

If $n=3$ then

$$C=\begin{pmatrix}2&-1&0\\-1&2&-2\\0&-1&2\end{pmatrix}\equiv\begin{pmatrix}0&1&0\\1&0&0\\0&1&0\end{pmatrix}.$$

From a straightforward case distinction on the roots of B_3 and the fact that $\overline{\alpha_1}=\overline{\alpha_3}$ we immediately see that $\overline{\alpha_1}=\overline{\alpha_3}=\overline{\alpha_1+2\alpha_2+2\alpha_3}$, $\overline{\alpha_2}=\overline{\alpha_1+\alpha_2+\alpha_3}=\overline{\alpha_2+2\alpha_3}$, and $\overline{\alpha_1+\alpha_2}=\overline{\alpha_2+\alpha_3}=\overline{\alpha_1+\alpha_2+2\alpha_3}$. This gives the 3 required root spaces of dimension 6.

If $n=4$ then

$$C=\begin{pmatrix}2&-1&0&0\\-1&2&-1&0\\0&-1&2&-2\\0&0&-1&2\end{pmatrix}\equiv\begin{pmatrix}0&1&0&0\\1&0&1&0\\0&1&0&0\\0&0&1&0\end{pmatrix}.$$

From a straightforward case distinction on the roots of B_4 and the fact that $\overline{\alpha_1}=\overline{\alpha_3}$, we find $\overline{\alpha_1}=\overline{\alpha_3}=\overline{\alpha_3+2\alpha_4}=\overline{\alpha_1+2\alpha_2+2\alpha_3+2\alpha_4}$, as well as $\overline{\alpha_2}=\overline{\alpha_1+\alpha_2+\alpha_3}=\overline{\alpha_1+\alpha_2+\alpha_3+2\alpha_4}=\overline{\alpha_2+2\alpha_3+2\alpha_4}$ and $\overline{\alpha_1+\alpha_2}=\overline{\alpha_2+\alpha_3}=\overline{\alpha_2+2\alpha_3+2\alpha_4}=\overline{\alpha_1+\alpha_2+2\alpha_3+2\alpha_4}$. The remaining $32-24=8$ roots $(\pm(\alpha_j+\cdots+\alpha_n),\ j=1,\dots,4)$ are in 2-dimensional spaces, giving $2^4, 8^3$, as required.

2.3. G_2

The root datum of type G_2 has Cartan matrix

$$C = \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix},$$

and the roots are

$$\begin{array}{ll} \pm\alpha_1, \pm(\alpha_1+\alpha_2), \pm(2\alpha_1+\alpha_2) & \text{(6 short roots)}, \\ \pm\alpha_2, \pm(3\alpha_1+\alpha_2), \pm(3\alpha_1+2\alpha_2) & \text{(6 long roots)}, \end{array}$$

giving a total of 12 roots. As $\det(C) = 1$, we take $A = I$. All components of c are in $\{-3, \ldots, 3\}$, so all components of the differences $\alpha_1 - \beta$ and $\alpha_2 - \beta$ are in $\{-4, \ldots, 4\}$. Hence, if multidimensional root spaces occur, we must have $p \leqslant 3$.

If $p = 3$ we see $\overline{3\alpha_1+\alpha_2} = \overline{\alpha_2} = \overline{-(3\alpha_1+2\alpha_2)}$ and $\overline{-(3\alpha_1+\alpha_2)} = \overline{-\alpha_2} = \overline{3\alpha_1+2\alpha_2}$, and the remaining 6 roots all have distinct root spaces.

If $p = 2$ we find $\overline{\alpha_1+\alpha_2} = \overline{3\alpha_1+\alpha_2}$, $\overline{\alpha_1} = \overline{3\alpha_1+2\alpha_2}$ and $\overline{\alpha_2} = \overline{2\alpha_1+\alpha_2}$, giving 3 root spaces of dimension 4.

This finishes the proof of Proposition 3.

3. Finding frames

Let L be a Chevalley Lie algebra over $\mathbb{F}$ with root datum R, a fixed split Cartan subalgebra H, and given decomposition E into root spaces with respect to the set $\overline{\Phi} = \Phi(L, H)$ of roots of H on L. In this section we discuss the procedure of Algorithm 1 referred to as FindFrame. It determines the set $\mathcal{X} = \{\mathbb{F}X_\alpha \mid \alpha \in \Phi\}$, i.e., the one-dimensional root spaces with respect to Φ, to which we refer as the *Chevalley frame*. Note that we do not yet *identify* the root spaces: finding a suitable bijection between Φ and the Chevalley frame $\mathcal{X}$ is discussed in the next section. We set $p = \mathrm{char}(\mathbb{F})$.

We require that R be given, since we execute different algorithms depending on R, for example $\mathrm{B}_2^{\mathrm{ad}}$ needs [C] whereas $\mathrm{B}_2^{\mathrm{sc}}$ needs [$\mathrm{B}_2^{\mathrm{sc}}$].

For $p = 2$, we use the procedure described in Section 3.1 to find the frame once we have computed all spaces $\mathbb{F}X_\alpha + \mathbb{F}X_{-\alpha}$ for $\alpha \in \Phi$. To this algorithm we will refer as [A_2]. As an auxiliary result, this procedure stores the unordered pairs $\{\{\alpha, -\alpha\} \mid \alpha \in \Phi^+\}$, to be used in the IdentifyRoots procedure discussed in Section 4 (notably, the proof of Lemma 9).

The general method in characteristic 2 is to reduce the root spaces of dimension greater than 2 to such 2-dimensional spaces, and apply [A_2]. For this purpose, and for the two cases of characteristic 3, we distinguish three methods:

- [C]: Given two root spaces M, M' compute $C_M(M')$ to break down M. Often, but not always, $\dim(M') = 2$. An example of this method is given in Section 3.2.
- [Der]: Compute the Lie algebra $\mathrm{Der}(L)$ of derivations of L, and calculate in there. This is a useful approach if $\mathrm{Der}(L)$ is strictly larger than L, for then we can often extend H to a larger split Cartan subalgebra, so we find new semisimple elements acting on the root spaces. Examples of this method are given in Sections 3.3 and 3.4.
- [$\mathrm{B}_2^{\mathrm{sc}}$]: The case where $R(p) = \mathrm{B}_2^{\mathrm{sc}}(2)$ is slightly more involved than the other cases because $\overline{\alpha} = 0$ for some $\alpha \in \Phi$. We use the Meat-axe to split the action of the long roots on the short roots. Examples of this method are given in Sections 3.5 and 3.6.

The method chosen depends on the root datum R and the characteristic p, as indicated in the third column of Table 1.

3.1. A_2 *in characteristic 2*

First, we consider the Lie algebras L with $R(p) = \mathrm{A}_2(2)$, as this procedure is used inside various other cases. The isogeny type of the root datum is of no importance here. For clarity, we write α, β for the two simple roots of the root system of type A_2.

As indicated in Table 1, we have 3 root spaces of dimension 2. They correspond to $\langle X_\gamma, X_{-\gamma}\rangle_{\mathbb{F}}$ for $\gamma \in \{\alpha, \beta, \alpha+\beta\}$. Without loss of generality we consider $L_{\overline{\alpha}} = \langle X_\alpha, X_{-\alpha}\rangle_{\mathbb{F}}$ and $L_{\overline{\beta}} = \langle X_\beta, X_{-\beta}\rangle_{\mathbb{F}}$. Observe that the squared adjoint action $\mathrm{ad}^2_{X_\alpha}$ of X_α sends any element of $L_{\overline{\beta}}$ to zero: $[X_\alpha, [X_\alpha, X_\beta]] = [X_\alpha, N_{\alpha,\beta}X_{\alpha+\beta}] = 0$ as $2\alpha+\beta \notin \Phi$, and $[X_\alpha, X_{-\beta}] = 0$ since $\alpha - \beta \notin \Phi$. Similarly, $\mathrm{ad}^2_{X_{-\alpha}}(L_{\overline{\beta}}) = 0$.

However, the quadratic action ad^2_x of a general element $x = t_1X_\alpha + t_2X_{-\alpha}$ ($t_1, t_2 \in \mathbb{F}$, both nonzero) of $L_{\overline{\alpha}}$ does not centralize $L_{\overline{\beta}}$. Indeed:

$$\begin{aligned}\big[x, [x, X_\beta]\big] &= t_1t_2\big(\big[X_{-\alpha}, [X_\alpha, X_\beta]\big] + \big[X_\alpha, [X_{-\alpha}, X_\beta]\big]\big)\\ &= t_1t_2N_{-\alpha,\alpha+\beta}N_{\alpha,\beta}X_\beta,\end{aligned}$$

which is nonzero since $N_{-\alpha,\alpha+\beta}$ and $N_{\alpha,\beta}$ are both equal to 1 modulo 2.

Recall that we are given $L_{\overline{\alpha}}$ and $L_{\overline{\beta}}$. Fix a basis r_1, r_2 of $L_{\overline{\alpha}}$ and consider the element $x = r_1 + tr_2$, where $t \in \mathbb{F}$. It follows from the above observations that $\mathrm{ad}^2_x(L_{\overline{\beta}}) = 0$ if and only if x is a scalar multiple of X_α or $X_{-\alpha}$, so in order to find the frame elements among the $\mathbb{F}x$ for $t \in \mathbb{F}$ we have to solve

$$\begin{aligned}0 = \big[x, [x, y]\big] &= \big[r_1 + tr_2, [r_1 + tr_2, y]\big]\\ &= \big[r_1, [r_1, y]\big] + t\big(\big[r_1, [r_2, y]\big] + \big[r_2, [r_1, y]\big]\big) + t^2\big[r_2, [r_2, y]\big],\end{aligned}$$

for every $y \in L_{\overline{\beta}}$ in the unknown t. We know there is a solution as H is split. Solving this system is equivalent to solving a system of $2 \cdot 3 = 6$ quadratic equations in t (note that the $[r_i, [r_j, y]]$ are in $\langle L_{\overline{\beta}}\rangle_L$, which is at most 3-dimensional). If $\mathbb{F} = \mathbb{F}_q$, solving such a quadratic equation is equivalent to solving $\log(q)$ equations in $\log(q)$ variables over $\mathbb{F}_2$ (as $p = 2$ is fixed), requiring $O^\sim(\log(q)^3)$ arithmetic operations, or $O^\sim(\log(q)^4)$ elementary operations.

For more general Lie algebras L, the solutions for Lie subalgebras of type A_2 normalized by H will be part of a Chevalley frame. These parts can be found inside any two-dimensional root space $V \in E$ provided there is at least one other two-dimensional root space $V' \in E$ such that $\langle V, V'\rangle_L$ is of type A_2. So, if all root spaces in E are 2-dimensional and $\mathbb{F} = \mathbb{F}_q$, this method needs $O(n^2)$ root spaces V to be analyzed (at a cost of $O^\sim(n^8\log(q)^4)$ each), so that $\mathcal{X}$ will be found in $O^\sim(n^{10}\log(q)^4)$ elementary operations.

3.2. G_2 *in characteristic 3*

Secondly, we consider the Lie algebra $L = L_{\mathbb{F}}(\mathrm{G}_2)$ of the root datum of type G_2 over a field $\mathbb{F}$ of characteristic 3. By Proposition 3 there are 8 root spaces. It is readily verified that $\dim(L_{\overline{\alpha}}) = 1$ if α is a short root and $\dim(L_{\overline{\alpha}}) = 3$ if α is a long root of Φ. In particular, the short root spaces belong to $\mathcal{X}$ and it remains to split the two long root spaces.

Consider one of the two three-dimensional root spaces in E, say $V = \mathbb{F}X_{\alpha_2} + \mathbb{F}X_{3\alpha_1+\alpha_2} + \mathbb{F}X_{-3\alpha_1-2\alpha_2}$. The left multiplications on V by the short roots are easily obtained from (CB1)–(CB4); these are given in Table 2.

Although we have not yet identified the roots, we can identify the three pairs of one-dimensional root spaces $\{\mathbb{F}X_\alpha, \mathbb{F}X_{-\alpha}\}$, for $\alpha \in \Phi$ short, since $L_{\overline{-\alpha}}$ is the unique one-dimensional root space with root $-\overline{\alpha}$. From this observation and Table 2 it follows that we can obtain the triple $\mathbb{F}X_\beta$ ($\beta \in \{\alpha_2, 3\alpha_1+\alpha_2, -3\alpha_1+2\alpha_2\}$) as follows:

$$\mathbb{F}X_{\alpha_2} = \mathrm{C}_V(L_{\overline{2\alpha_1+\alpha_2}} + L_{\overline{-2\alpha_1-\alpha_2}}),$$
$$\mathbb{F}X_{3\alpha_1+\alpha_2} = \mathrm{C}_V(L_{\overline{\alpha_1+\alpha_2}} + L_{\overline{-\alpha_1-\alpha_2}}),$$
$$\mathbb{F}X_{-3\alpha_1-2\alpha_2} = \mathrm{C}_V(L_{\overline{\alpha_1}} + L_{\overline{-\alpha_1}}).$$

For the other three-dimensional space, the same approach is used. This completes the search for the Chevalley frame $\mathcal{X}$.

Table 2
Part of the G_2 multiplication table.

	X_{α_2}	$X_{3\alpha_1+\alpha_2}$	$X_{-3\alpha_1-2\alpha_2}$
X_{α_1}	$X_{\alpha_1+\alpha_2}$	0	0
$X_{-\alpha_1}$	0	$X_{2\alpha_1+\alpha_2}$	0
$X_{\alpha_1+\alpha_2}$	0	0	$X_{-2\alpha_1-\alpha_2}$
$X_{-\alpha_1-\alpha_2}$	$-X_{\alpha_1}$	0	0
$X_{2\alpha_1+\alpha_2}$	0	0	$-X_{-\alpha_1}$
$X_{-2\alpha_1-\alpha_2}$	0	$-X_{\alpha_1}$	0

3.3. D_4 *in characteristic 2*

Thirdly, we consider the Lie algebras with Dynkin diagram of type D_4 over a field $\mathbb{F}$ of characteristic 2. As mentioned in Section 2, there are three cases:

L^{ad}: the adjoint root datum (12 two-dimensional root spaces),
L^{sc}: the simply connected root datum (3 eight-dimensional root spaces),
$L^{(1)}, L^{(3)}, L^{(4)}$: the intermediate root data (6 four-dimensional root spaces).

The three intermediate root data all give rise to the same Lie algebra up to isomorphism (by triality), so we will restrict ourselves to the study of L^{ad}, L^{sc}, and $L^{(1)}$. It is straightforward to verify that L^{ad} has a 26-dimensional ideal I^{ad}, linearly spanned by X_α ($\alpha \in \Phi$), $(\alpha_1^\vee + \alpha_3^\vee + \alpha_4^\vee) \otimes 1$, and $\alpha_2^\vee \otimes 1$. This ideal can be found, for example, by use of the Meat-axe.

Similarly, L^{sc} has a 2-dimensional ideal I (spanned by $(\alpha_1^\vee + \alpha_4^\vee) \otimes 1$ and $(\alpha_3^\vee + \alpha_4^\vee) \otimes 1$). Let $I^{\mathrm{sc}} = L^{\mathrm{sc}}/I$ be the 26-dimensional Lie algebra obtained by computing in L^{sc} modulo I. Finally, $L^{(1)}$ has a 1-dimensional ideal I (spanned by $\alpha_4 \otimes 1$), and a 27-dimensional ideal I' (spanned by $\alpha_4 \otimes 1$ and X_α, $\alpha \in \Phi$). We let $I^{(\mathrm{a})} = I'/I$. Again, the 26-dimensional ideal is easily found by means of the Meat-axe.

Thus we have constructed three 26-dimensional Lie algebras: I^{ad}, I^{sc}, and $I^{(\mathrm{a})}$. By results of Chevalley (cf. [14, Part 2, Cor. 2.7]) they are isomorphic, so from now on we let I be one of these 26-dimensional Lie algebras. The Lie algebra I is simple. Its derivation algebra $\mathrm{Der}(I)$ is a Lie algebra of type F_4, and thus has 12 two-dimensional root spaces and 3 eight-dimensional root spaces.

Using a procedure similar to the one for G_2 over characteristic 3 described in Section 3.2, we can break up the eight-dimensional spaces of E into two-dimensional spaces, giving us 24 two-dimensional spaces. These two-dimensional spaces may then be broken up into one-dimensional spaces by the procedure [A_2]. The last step in the process is "pulling back" the relevant one-dimensional spaces from $\mathrm{Der}(I)$ to I. But this is straightforward, since I is an ideal of $\mathrm{Der}(I)$ by construction.

3.4. G_2 *in characteristic 2*

As noted in [19, Section 2.6], in the exceptional case $R(p) = \mathrm{G}_2(2)$, the Lie algebra L is isomorphic to the unique 14-dimensional ideal of the Chevalley Lie algebra L^{A} of adjoint type A_3 over $\mathbb{F}$.

In particular, $\mathrm{Der}(L)$ contains a copy of L^{A}. We use this fact by finding a split Cartan subalgebra H' inside $\mathrm{C}_{\mathrm{Der}(L)}(H)$ so that $H \subset H'$. For then we can calculate the Chevalley frame $\mathcal{X}^{\mathrm{A}}$ inside the Lie subalgebra $\langle L, H'\rangle_{\mathrm{Der}(L)}$ of $\mathrm{Der}(L)$ with respect to H', which is of type A_3 by the above observation.

The Chevalley frame $\mathcal{X}$ of L is now simply the part of $\mathcal{X}^{\mathrm{A}}$ that lies inside L.

3.5. $\mathrm{B}_2^{\mathrm{sc}}$ *in characteristic 2*

We consider the Chevalley Lie algebra L of type $\mathrm{B}_2^{\mathrm{sc}}$ over a field $\mathbb{F}$ of characteristic 2 with split Cartan subalgebra $H = \mathbb{F}h_1 + \mathbb{F}h_2$. This is a particularly difficult case, as the automorphism group of L is quite big: $\mathrm{Aut}(L) = G \ltimes (\mathbb{F}^+)^4$ [10, Theorem 14.1], where G is the Chevalley group of adjoint type B_2 over $\mathbb{F}$ and $\mathbb{F}^+$ refers to the additive group of $\mathbb{F}$. As a consequence, there is more choice in finding the frame than in the previous cases.

To begin, we take L_0 to be the $(0,0)$-root space of H on L, and L_1 to be the $(1,0)$-root space of H on L. It is easily verified that $L_0 = \langle H, X_{\pm\alpha_1}, X_{\pm(\alpha_1+2\alpha_2)}\rangle_{\mathbb{F}}$ (that is, the linear span of H and the long root elements) and $L_1 = \langle X_{\pm\alpha_2}, X_{\pm(\alpha_1+\alpha_2)}\rangle_{\mathbb{F}}$ (the linear span of the short root elements). We proceed in three steps.

[$\mathbf{B_2^{sc}.1}$]. The subalgebra L_0 has Dynkin type $\mathrm{A}_1 \oplus \mathrm{A}_1$. We may split it (nonuniquely) into two subalgebras of type A_1 using a direct sum decomposition procedure. This is a procedure that can be carried out with standard linear algebra arithmetic for a fixed dimension (6, in this case); see e.g., [9, Section 1.15].

[$\mathbf{B_2^{sc}.2}$]. Let A be one of these subalgebras of L_0 of type A_1. Assume for the sake of reasoning that $A = \langle X_{\pm\alpha_1}\rangle_L$, the Lie subalgebra of L generated by X_{α_1} and $X_{-\alpha_1}$. Since $[A, L_1] = L_1$ we may view L_1 as a four-dimensional A-module, and hence apply the Meat-axe [11,12] to find a proper irreducible A-submodule M of L_1. This will be a submodule of the form

$$M = \langle t_1X_{\alpha_2} + t_2X_{-\alpha_1-\alpha_2}, t_1X_{\alpha_1+\alpha_2} + t_2X_{-\alpha_2}\rangle_{\mathbb{F}}, \quad t_1, t_2 \in \mathbb{F}.$$

We take b_1, b_2 to be a basis of M, and add $\mathrm{C}_A(b_2)$ and $\mathrm{C}_A(b_1)$ to $\mathcal{X}$. These two spaces are indeed one-dimensional and coincide with the original $\mathbb{F}X_{\pm\alpha_1}$ if $b_1 \in \mathbb{F}(t_1X_{\alpha_2} + t_2X_{-\alpha_1-\alpha_2})$ and $b_2 \in \mathbb{F}(t_1X_{\alpha_1+\alpha_2} + t_2X_{-\alpha_2})$. This exhibits part of the freedom of choice induced by the factor $(\mathbb{F}^+)^4$ in $\mathrm{Aut}(L)$.

We repeat this procedure for both subalgebras of type A_1 found in the first step. The result is the part of the Chevalley frame $\mathcal{X}$ inside L_0. In fact, due to our method, we can make an identification of the long roots $\pm\alpha_1$, $\pm(\alpha_1 + 2\alpha_2)$ with the four elements of $\mathcal{X}$ found. In what follows we will work with such a choice so that we have the elements $\mathbb{F}X_{\alpha_1}$, $\mathbb{F}X_{-\alpha_1}$, $\mathbb{F}X_{\alpha_1+2\alpha_2}$, $\mathbb{F}X_{-\alpha_1-2\alpha_2}$ in $\mathcal{X}$ as well as the correspondence with the roots in Φ suggested by the subscripts.

[$\mathbf{B_2^{sc}.3}$]. We find the part of $\mathcal{X}$ inside L_1 as follows. $\mathbb{F}X_{\alpha_1+\alpha_2}$ coincides with $\mathrm{C}_{L_1}(\mathbb{F}X_{\alpha_1}, \mathbb{F}X_{\alpha_1+2\alpha_2})$. Having computed this element of $\mathcal{X}$, we finish by taking

$$\begin{aligned}
\mathbb{F}X_{\alpha_2} &= [\mathbb{F}X_{\alpha_1+\alpha_2}, \mathbb{F}X_{-\alpha_1}],\\
\mathbb{F}X_{-\alpha_1-\alpha_2} &= [\mathbb{F}X_{\alpha_2}, \mathbb{F}X_{-\alpha_1-2\alpha_2}],\\
\mathbb{F}X_{-\alpha_2} &= [\mathbb{F}X_{\alpha_1-\alpha_2}, \mathbb{F}X_{\alpha_1}].
\end{aligned}$$

This completes the search for $\mathcal{X}$ in the case $\mathrm{B}_2^{\mathrm{sc}}(2)$ and establishes that its running time is $O^\sim(\log(q))$.

3.6. $\mathrm{C}_n^{\mathrm{sc}}$ *in characteristic 2*

We consider the Chevalley Lie algebra L of type $\mathrm{C}_n^{\mathrm{sc}}$ over a field $\mathbb{F}$ of characteristic 2. Here $n \geqslant 3$, so that the multiplicity of $\bar{0}$ is strictly larger than 4. Let h_z be a basis of the 1-dimensional center of L, inside the split Cartan subalgebra H of L. This case is a generalization of the $\mathrm{B}_2^{\mathrm{sc}}$ case described in Section 3.5. We again take L_0 to be the 0-root space of H on L, so that L_0 is $3n$-dimensional and consists of H and the root spaces corresponding to the long roots. Similar to the previous case, $L_0 \cong \mathrm{A}_1 \oplus \cdots \oplus \mathrm{A}_1$ (n constituents), and again the decomposition is not unique. We describe how to find such a decomposition.

We let $\mathcal{F}$ be the set of $\binom{n}{2}$ four-dimensional root spaces (cf. Table 1). In the root system of type C_n each of these corresponds to the four roots $\pm\epsilon_i \pm \epsilon_j$ for some $i, j \in \{1, \dots, n\}$ with $i \neq j$. Our first task is to split L_0 into subalgebras of type A_1 in a way compatible with $\mathcal{F}$. To this end, we let Γ be the graph with vertex set $\mathcal{F}$, and edges $f \sim g$ whenever $f \neq g$ and $[f, g] \neq 0$.

Let Δ be a maximal coclique of Γ of size $n-1$, so that Δ consists of $n-1$ elements of $\mathcal{F}$ such that $[f,g]=0$ for all $f,g\in\Delta$. This means that, for a particular $i\in\{1,\dots,n\}$, the set $\Delta\subseteq\mathcal{F}$ corresponds to those four-spaces in $\mathcal{F}$ that arise from the roots $\pm\epsilon_i\pm\epsilon_j$, where $j\in\{1,\dots,n\}\setminus\{i\}$. Let $\overline{\Delta}=\Gamma-\Delta$, so that $\overline{\Delta}$ contains precisely the four-dimensional spaces corresponding to $\pm\epsilon_k\pm\epsilon_l$ with $k,l\neq i$.

Now compute the centralizer A in L_0 of all spaces in $\overline{\Delta}$. Then A coincides with $\langle X_{\pm\gamma},\gamma^\vee\otimes 1,h_z\rangle_{\mathbb{F}}$ for the long root $\gamma=2\epsilon_i$. Using a direct sum decomposition procedure we find the Lie subalgebra A' of A such that $A=A'\oplus\mathbb{F}h_z$, where $A'=\langle X_{\pm\gamma},\gamma^\vee\otimes 1\rangle_{\mathbb{F}}$. The subalgebra A' is one of the type A_1 constituents of L_0 we are after. Thus, by repeating this procedure for each maximal coclique of Γ of size $n-1$, we obtain a decomposition of L_0 into n subalgebras of type A_1. We will denote by $\mathcal{A}$ the set of these n subalgebras.

Now we continue as in the $\mathrm{B}_2^{\mathrm{sc}}$ case: For each element of $\mathcal{A}$ we use the procedure labeled [$\mathrm{B}_2^{\mathrm{sc}}$.2] to find suitable elements $\mathbb{F}X_{\pm\gamma}$ for $\mathcal{X}$. For each four-dimensional space $K\in\mathcal{F}$ we then use distinct $S_1,S_2\in\mathcal{A}$ satisfying $[K,S_1]\neq 0$, $[K,S_2]\neq 0$ and these $\mathbb{F}X_{\pm\gamma}$ to execute a [$\mathrm{B}_2^{\mathrm{sc}}$.3] procedure. Thus, we find the part of the frame inside K.

If $n=3$ splitting L_0 has to be done in a slightly different way, but as this is only a slight modification of the algorithm we will not go into details here. This completes the Chevalley frame finding in the case $\mathrm{C}_n^{\mathrm{sc}}(2)$. Its running time involves $O(n^2)$ executions of parts of the algorithm of Section 3.5, which is however dominated by the time $O^\sim(n^{10}\log(q)^4)$ needed for method [A_2].

We summarize the results of this section.

Proposition 6. *Given L, H, R, the set $\overline{\Phi}$ of roots of H on L, and the root spaces E, the Las Vegas procedure* FindFrame *finds a Chevalley frame. For $\mathbb{F}=\mathbb{F}_q$, it runs in time $O^\sim(n^{10}\log(q)^4)$.*

Proof. As mentioned in Section 1.7 this procedure is trivial in all cases except those mentioned in Table 1, and for each of the cases in Table 1 we have presented a solution. Recall that $|\Phi|\leqslant\dim(L)=O(n^2)$.

The timing of method [A_2] is dealt with in Section 3.1, which produces the bound stated in the proposition.

Method [C] concerns $O(n^2)$ instances of standard linear algebra arithmetic on spaces of bounded dimension, and so its running time is dominated again by time spent on the [A_2] method.

Method [Der] involves the computation of parts of the Lie algebra of derivations. Computing the full Lie algebra of derivations in instances like $D_n^{\mathrm{sc}}(2)$ would take running time $O^\sim(n^{12}\log(q))$. However, we only carry out this procedure for Lie algebras of bounded dimension (the bound being 28, which occurs for type D_4) or compute the part of $\mathrm{Der}(L)$ that leaves invariant H and the corresponding decomposition into root spaces (which reduces the running time to $O^\sim(n^8\log(q))$). Therefore, the stated bound suffices.

Finally, according to Table 1, method [$\mathrm{B}_2^{\mathrm{sc}}$] with unbounded n only occurs in the cases treated in Section 3.6, where the time analysis is already given. □

4. Root identification

In this section we clarify Step 3 of the ChevalleyBasis Algorithm 1. The routine IdentifyRoots takes as input a Chevalley Lie algebra L, a split Cartan subalgebra H of L, the root datum R and the set of roots $\overline{\Phi}=\Phi(L,H)$, and the Chevalley frame $\mathcal{X}$ found in the previous step (Section 3). It returns a bijection $\iota:\Phi\to\mathcal{X}$ so that, up to scaling, $(X_\alpha)_{\alpha\in\Phi}$ will be the root element part of a Chevalley basis.

An important tool to make this identification are the Cartan integers $\langle\alpha,\beta^\vee\rangle$. Cartan integers may be computed using root chains; see, for instance, [4].

Lemma 7. *Let $\alpha,\beta\in\Phi$. Suppose p and q are the largest nonnegative integers such that $\alpha-p\beta\in\Phi$ and $\alpha+q\beta\in\Phi$. Then $\langle\alpha,\beta^\vee\rangle=p-q$.*

We may use this lemma by computing such a chain in the set of roots $\overline{\Phi}$ corresponding to the Chevalley frame $\mathcal{X} = \{\mathbb{F}X_\alpha \mid \alpha \in \Phi\}$. However, as these roots are computed from the Lie algebra L over $\mathbb{F}$ itself, they live in the n-dimensional vector space $\mathbb{F}^n$ rather than over $\mathbb{Z}^n$.

A straightforward verification of cases for Chevalley Lie algebras arising from root systems of rank 2 shows that the chain can simply be computed in terms of the roots over $\mathbb{F}^n$, except if the characteristic is 2 or 3. So in the latter two cases, we a different method for computing $\langle \alpha, \beta^\vee \rangle$ is needed.

Lemma 8. *Suppose that $L = L_{\mathbb{F}}(R)$ is a Chevalley Lie algebra with respect to a root datum $R = (X, \Phi, Y, \Phi^\vee)$ over the field $\mathbb{F}$ of characteristic 2 or 3. Let H be the standard split Cartan subalgebra of L. Suppose furthermore that $X_\alpha, X_{-\alpha}, X_\beta, X_{-\beta}$ are four vectors spanning root spaces corresponding to $\alpha, -\alpha, \beta, -\beta \in \Phi$, respectively, and $\alpha \neq \pm\beta$.*

If Φ is simply laced, then $\langle \alpha, \beta^\vee \rangle = P - Q$, where

$$P = \begin{cases} 0 & \text{if } [X_{-\beta}, X_\alpha] = 0, \\ 1 & \text{if } [X_{-\beta}, X_\alpha] \neq 0, \end{cases} \qquad Q = \begin{cases} 0 & \text{if } [X_\beta, X_\alpha] = 0, \\ 1 & \text{if } [X_\beta, X_\alpha] \neq 0. \end{cases}$$

If Φ is doubly laced and $\operatorname{char}(\mathbb{F}) \neq 2$, *then $\langle \alpha, \beta^\vee \rangle = P - Q$, where*

$$P = \begin{cases} 0 & \text{if } [X_{-\beta}, X_\alpha] = 0, \\ 1 & \text{if } [X_{-\beta}, X_\alpha] \neq 0, [X_{-\beta}, [X_{-\beta}, X_\alpha]] = 0, \\ 2 & \text{if } [X_{-\beta}, [X_{-\beta}, X_\alpha]] \neq 0, \end{cases}$$

$$Q = \begin{cases} 0 & \text{if } [X_\beta, X_\alpha] = 0, \\ 1 & \text{if } [X_\beta, X_\alpha] \neq 0, [X_\beta, [X_\beta, X_\alpha]] = 0, \\ 2 & \text{if } [X_\beta, [X_\beta, X_\alpha]] \neq 0. \end{cases}$$

Proof. For any $\gamma, \gamma' \in \Phi$, let $p_{\gamma,\gamma'}$ and $q_{\gamma,\gamma'}$ be the biggest nonnegative integers such that $\gamma - p_{\gamma,\gamma'}\gamma' \in \Phi$ and $\gamma + q_{\gamma,\gamma'}\gamma' \in \Phi$. Recall from (CB4) and [4] that, if $\gamma + \gamma' \in \Phi$, then $[X_\gamma, X_{\gamma'}] = N_{\gamma,\gamma'} X_{\gamma+\gamma'}$, where $N_{\gamma,\gamma'} = \pm(p_{\gamma,\gamma'} + 1)$.

If Φ is simply laced, the subsystem of Φ generated by $\pm\alpha, \pm\beta$ is of type A_1A_1 or of type A_2. Then $\alpha + \beta \in \Phi$ implies $\alpha - \beta \notin \Phi$, so $N_{\alpha,\beta} = \pm 1$. Similarly, $N_{\beta,\alpha} = \pm 1$. This means that, regardless of the characteristic, we can reconstruct $p_{\alpha,\beta}$ and $q_{\alpha,\beta}$ by the procedure described in the lemma, and thus compute $\langle \alpha, \beta^\vee \rangle = p_{\alpha,\beta} - q_{\alpha,\beta}$ by Lemma 7.

If Φ is doubly laced and $\operatorname{char}(\mathbb{F}) \neq 2$, the subsystem of Φ generated by $\pm\alpha, \pm\beta$ is of type A_1A_1, A_2, or B_2. (Note that G_2 never occurs inside a bigger Lie algebra.) In the first two cases the previous argument applies, so assume $\pm\alpha, \pm\beta$ generate a subsystem of Φ of type B_2. Similarly to the previous case, if $\alpha + \beta \in \Phi$ then $\alpha - 2\beta \notin \Phi$, so that $N_{\alpha,\beta}, N_{\beta,\alpha} \in \{\pm 1, \pm 2\}$. In particular, since $\operatorname{char}(\mathbb{F}) \neq 2$, we find that both $N_{\alpha,\beta}$ and $N_{\beta,\alpha}$ are nonzero, so that we can reconstruct $p_{\alpha,\beta}$ and $q_{\alpha,\beta}$ by the procedure described in the theorem, and thus compute $\langle \alpha, \beta^\vee \rangle = p_{\alpha,\beta} - q_{\alpha,\beta}$ by Lemma 7. $\square$

Lemma 9. *Suppose that L is a Chevalley Lie algebra over $\mathbb{F}$ with respect to a root datum $R = (X, \Phi, Y, \Phi^\vee)$, H is a split Cartan subalgebra of L, and X_α and X_β are two root elements whose roots with respect to H are $\overline{\alpha}$ and $\overline{\beta}$ for certain $\overline{\alpha}, \overline{\beta} \in \overline{\Phi}$. Suppose, furthermore, that one of the following statements holds.*

(1) $\operatorname{char}(\mathbb{F}) \notin \{2, 3\}$,
(2) *Φ is simply laced,*
(3) *Φ is doubly laced and* $\operatorname{char}(\mathbb{F}) \neq 2$.

Then $\langle \alpha, \beta^\vee \rangle$ can be computed from the available data in $O^\sim(n^{10} \log(q))$ elementary operations.

Proof. Observe first of all that the case where $\alpha = \beta$ is easily caught, for example by computing $\dim(\langle \mathbb{F}X_\alpha, \mathbb{F}X_\beta \rangle_{\mathbb{F}})$. Obviously then $\langle \alpha, \beta^\vee \rangle = 2$.

Moreover, we can distinguish the case where $\alpha = -\beta$ as follows. If $\mathrm{char}(\mathbb{F}) \neq 2$ we may simply test whether $\overline{\alpha} = -\overline{\beta}$. If on the other hand $\mathrm{char}(\mathbb{F}) = 2$, we find the sets $\{\{\gamma, -\gamma\} \mid \gamma \in \Phi^+\}$ as an auxiliary result of the algorithm FindFrames described in introduction of Section 3.1. If $\alpha = -\beta$, then of course $\langle \alpha, \beta^\vee \rangle = -2$.

So assume $\alpha \neq \pm\beta$. Now if (1) holds we compute $\langle \alpha, \beta^\vee \rangle$ from the roots $\overline{\alpha}$ and $\overline{\beta}$ using Lemma 7, as mentioned earlier.

Suppose, therefore, (2) or (3) holds. We can find $\mathbb{F}X_{-\alpha}$ and $\mathbb{F}X_{-\beta}$ either simply by considering $\{\overline{\gamma} \mid \gamma \in \Phi\}$ (if $\mathrm{char}(\mathbb{F}) \neq 2$) or as an auxiliary result of FindFrames (if $\mathrm{char}(\mathbb{F}) = 2$). This leaves us in a position where we may apply Lemma 8, and thus find $\langle \alpha, \beta^\vee \rangle$.

Finally, the time needed does not exceed the time needed for standard linear algebra arithmetic for each pair of roots, that is, $O^\sim(n^4 \cdot n^6 \log(q))$. □

The last lemma enables us to compute Cartan integers in many cases. For the cases not covered by Lemma 9 we proceed as follows to construct a direct identification ι.

- $\mathrm{B}_n(2)$: The short root spaces generate an ideal, I say, of L found by the Meat-axe, and the root eigenspaces of H that do not lie in I belong to long roots. These root spaces generate a subalgebra of type D_n. This Lie algebra is simply laced, so the root identification problem can be solved there. This identifies the long root spaces. Now, for $i = 1, \dots, n$, let the short root γ_i be $\alpha_i + \alpha_{i+1} + \dots + \alpha_n$ and let $\alpha_0 = \alpha_1 + 2\alpha_2 + 2\alpha_3 + \dots + 2\alpha_n$ be the (long) highest root. Observe then that $[X_{\alpha_0}, X_{-\gamma_1}] = X_{\gamma_2}$ and $[X_{\alpha_0}, X_{-\gamma_2}] = X_{\gamma_1}$, and $X_{-\gamma_1}$ and $X_{-\gamma_2}$ are the only short root elements that do not commute with X_{α_0}. This fact, together with the set of pairs $\{\{\gamma, -\gamma\} \mid \gamma \in \Phi^+\}$ obtained in FindFrames, allows us to find $X_{\pm\gamma_1}$ and $X_{\pm\gamma_2}$. Note that we have to execute this procedure at most twice, since there are only elements of $\mathcal{X}$ that could be identified with $X_{-\gamma_1}$, and the other short root elements are fixed once $X_{-\gamma_1}$ is fixed.
  The other short root elements may now simply be found by using relations such as $[X_{\gamma_i}, X_{-\alpha_i}] = X_{\gamma_{i+1}}$.
- $\mathrm{C}_n(2)$: The short root spaces generate an ideal of L of type D_n, so we execute a similar procedure as in the previous case.
- $\mathrm{F}_4(2)$: The short roots generate an ideal of L of dimension 26 which together with the Cartan subalgebra H gives a 28-dimensional subalgebra of type D_4, allowing the same procedure as before.
- $\mathrm{G}_2(3)$: Similarly to the previous cases, we use the fact that the short roots generate an ideal of L of type A_2, which is again simply laced.
- $\mathrm{G}_2(2)$: As described in Section 3.4, the manner in which the root spaces in L^{A} correspond to those in L is completely determined. Therefore, we may use the roots identified in L^{A}, which is simply laced, to identify the roots in L.

These methods lead to the following conclusion.

Proposition 10. *Given L over $\mathbb{F}$, H, $R = (X, \Phi, Y, \Phi^\vee)$, the set $\overline{\Phi}$ of roots of H on L, and a Chevalley frame $\mathcal{X}$, the routine* IdentifyRoots *finds a bijection $\iota : \Phi \to \mathcal{X}$ such that for all $\alpha, \beta \in \Phi$, $\alpha \neq \pm\beta$,*

$$\bigl[\iota(\alpha), \iota(\beta)\bigr] = \begin{cases} \iota(\alpha+\beta) & \text{if } \alpha+\beta \in \Phi \text{ and } N_{\alpha,\beta} \not\equiv 0 \pmod p, \\ \{0\} & \text{otherwise}. \end{cases}$$

For $\mathbb{F} = \mathbb{F}_q$, the routine needs $O^\sim(n^{10} \log(q))$ elementary operations.

Proof. Lemma 9 shows that in many cases we can compute Cartan integers. To this end, we need to compute $\langle \alpha, \beta^\vee \rangle$ for all $O(n^4)$ pairs of roots, and every computation of this type involves at most 6 multiplications in L, requiring a total of $O^\sim(n^{4+6} \log(q))$ elementary operations. Once these numbers are computed, it takes $O(n^4)$ steps to select a set of simple roots and subsequently to complete the bijection between Φ and $\mathcal{X}$. These last two steps use techniques similar to those described by De

Table 3
Algorithm 1 timings.

R	$\mathbb{Q}$	17	3^3	2^6	R	$\mathbb{Q}$	17	3^3	2^6	R	$\mathbb{Q}$	17	3^3	2^6
A_1	0	0	0	0	B_6	0.9	0.6	3.2	20	D_1	0	0	0	0
A_2	0	0	0	0	B_7	2.2	1.6	10	54	D_3	0	0	0	0.3
A_3	0	0	0	0.7	B_8	5.3	3.9	27	172	D_4	0.1	0.1	0.1	3.2
A_4	0.1	0	0.1	0.1	B_9	12	8.8	68	493	D_5	0.2	0.1	0.3	22
A_5	0.1	0.1	0.1	0.2	C_1	0	0	0	0	D_6	0.6	0.4	0.9	121
A_6	0.3	0.2	0.4	0.6	C_2	0	0	0	0	D_7	1.6	1.1	2.8	545
A_7	0.6	0.5	0.9	1.5	C_3	0	0	0.1	0.1	D_8	3.8	2.8	7.7	1994
A_8	1.4	1	2	3.6	C_4	0.1	0.1	0.2	1.1	D_9	8.6	6.4	19	6396
A_9	2.8	2	4.2	7.9	C_5	0.3	0.2	0.9	10	E_6	0.9	0.6	1.6	3.3
B_1	0	0	0	0	C_6	0.9	0.6	3.2	40	E_7	4.1	3	11	27
B_2	0	0	0	0	C_7	2.2	1.6	10	177	E_8	28	21	112	398
B_3	0	0	0.1	0.4	C_8	5.2	3.9	27	693	F_4	0.2	0.2	0.7	3.3
B_4	0.1	0.1	0.2	1.9	C_9	12	8.8	69	2212	G_2	0	0	0.1	0.5
B_5	0.3	0.2	0.9	4.8										

Graaf [9, Section 5.11]: the creation of a set of simple roots Π starts with taking an arbitrary root to be the first member of Π. We then iteratively pick a suitable additional simple root β having Cartan integer $\langle\alpha,\beta^\vee\rangle \leqslant 0$ with the members α of Π. This proves that we can make the required bijection in $O^\sim(n^{10}\log(q))$ time for the cases covered by Lemma 9.

For the remainder of the proof, we can restrict ourselves to the cases not covered by Lemma 9. Here the procedure described provides ι directly, so we only need prove the last assertion of the proposition. As $G_2(2)$ is directly reduced to a case already treated, it needs no further consideration. In each of the remaining cases, we need to compute a subalgebra or an ideal of L. Although this is hard in general, the fact that we have already found the Chevalley frame $\mathcal{X}$ and the fact that the subalgebra or ideal is a sum of elements from $\mathcal{X}$ imply that the computations take $O^\sim(n^{10}\log(q))$ elementary operations. A bijection ι' from the relevant subsystem of Φ to the subset of $\mathcal{X}$ of root spaces lying in the ideal may then be identified in time $O^\sim(n^{10}\log(q))$. Finally, extending ι' to the entirety of Φ is a straightforward task, requiring only standard linear algebra arithmetic in L.

This shows that we can make the required bijection in the time stated for all cases. □

5. Conclusion

As discussed in Section 1.7 the more difficult steps of Algorithm 1 are FindFrame and IdentifyRoots. In Sections 3 (Proposition 6) and 4 (Proposition 10) we established that these steps can be dealt with in time $O^\sim(n^{10}\log(q)^4)$. This proves Theorem 1 for a given root datum.

We emphasize that this estimate is only asymptotic. Additionally, in Table 3 we present timings of Algorithm 1 for various root data and for four different fields: $\mathbb{Q}$, GF(17), GF(3^3), and GF(2^6). The times given are in seconds, for the most time-consuming root datum with the specified Lie type. Input for the algorithm were a Chevalley Lie algebra and its splitting Cartan subalgebra, to which a random basis transformation was applied, and the root datum. The timings were produced using Magma 2.15-5 on an Intel Core 2 Quad CPU running at 2.4 GHz with 8 GB of memory available, although only one core and at most 2.7 GB of memory were used.

As hinted at earlier, Algorithm 1 can easily be used to produce an algorithm that takes only L and H and produces the root datum R and a Chevalley basis. To see this, note first that, because H is given and the underlying algebraic group is assumed to be simple, we may use $\dim(H) = \mathrm{rk}(R)$, the dimension of L, and the classification of simple Lie algebras to narrow down the root system to one or two possibilities (or three, but only if $\dim(L) = 78$ and $\dim(H) = 6$). Therefore, the number of possible root systems it at most 3.

Second, given a root system, the number of possible root data is small as well. If the root system is not of type A or D, the number of possible isogeny types is at most 2. If the root system is of type D the number of possible isogeny types is at most 5, as explained in Section 2. So suppose Φ is of type A_n, and fix $p = \mathrm{char}(\mathbb{F})$. Note that the fundamental group is $\mathbb{Z}/(n+1)\mathbb{Z}$. Since two root data for A_n

lead to isomorphic Lie algebras if both have the same exponent of p in $[X : \mathbb{Z}\Phi]$, we need consider at most $\log_p(n+1)+1 = O(\log(n))$ different isogeny types. Thus, in order to identify the correct root datum, we run Algorithm 1 a sufficiently small number of times for the polynomiality bound given in the theorem to remain intact. This finishes the proof of Theorem 1.

A problem hinted at, but not solved satisfactorily, is finding a split Cartan subalgebra of a Chevalley Lie algebra L if $p = 2$. Nevertheless, verifying that a given subalgebra is indeed a split Cartan subalgebra is easy. So our results are still useful, since one is often able to obtain such a subalgebra by other means, for example as part of the original problem. Moreover, experimental implementations of randomized algorithms for finding split Cartan subalgebras look promising. We intend to publish about these algorithms in forthcoming work.

A primary goal in writing the Chevalley basis algorithm is to use it for conjugacy questions in simple algebraic groups G or finite groups $G(\mathbb{F}_q)$ of rational points over $\mathbb{F}_q$. One of the complications in this application is the fact that the group $\mathrm{Aut}(L)$ may be much larger than $G(\mathbb{F}_q)$. For this purpose, a method is needed to write an arbitrary automorphism of L as a product of an element from $G(\mathbb{F}_q)$ and a particular coset representative of $G(\mathbb{F}_q)$ in $\mathrm{Aut}(L)$. Such a method is in [7] and is also used in [6].

Once Algorithm 1 completes successfully we have a certificate for a Lie algebra to be of type R: when presented with a candidate Chevalley basis X^0, H^0, we only need to carry out the straightforward and quick task of verifying that X^0, H^0 is indeed a Chevalley basis for L with respect to H and R. In this way, our work also contributes to a recognition procedure for modular simple Lie algebras. Obviously, Algorithm 1 can be used for establishing an isomorphism between two Chevalley Lie algebras over the same field and of the same root datum.

References

[1] Armand Borel, Linear Algebraic Groups, W.A. Benjamin, Inc., New York, 1969.
[2] Nicolas Bourbaki, Éléments de mathématique, Masson, Paris, 1981, Groupes et algèbres de Lie. Chapitres 4, 5 et 6 [Lie groups and Lie algebras. Chapters 4, 5 and 6].
[3] John J. Cannon, Wieb Bosma (Eds.), Handbook of Magma Functions, Edition 2.15, School of Mathematics and Statistics, University of Sydney, 2008, http://magma.maths.usyd.edu.au/.
[4] Roger W. Carter, Simple Groups of Lie Type, Pure Appl. Math. (N. Y.), John Wiley & Sons Inc., New York, 1972.
[5] Claude Chevalley, Classification des groupes algébriques semi-simples, Collected Works, Springer-Verlag, Heidelberg, 2005.
[6] Arjeh M. Cohen, Scott H. Murray, An algorithm for Lang's theorem, J. Algebra 322 (2009) 675–702.
[7] Arjeh M. Cohen, Scott H. Murray, D.E. Taylor, Computing in groups of Lie type, Math. Comp. 73 (247) (2004) 1477–1498.
[8] The GAP Group, GAP – Groups, Algorithms, and Programming, Version 4.4.12, http://www.gap-system.org/, 2008.
[9] Willem A. de Graaf, Lie Algebras: Theory and Algorithms, North-Holland Math. Library, vol. 56, Elsevier Science, 2000.
[10] Gerrit M.D. Hogeweij, Ideals and automorphisms of almost-classical lie algebras, PhD thesis, Universiteit Utrecht, 1978.
[11] Derek F. Holt, The Meataxe as a tool in computational group theory, in: The Atlas of Finite Groups: Ten Years On, Birmingham, 1995, in: London Math. Soc. Lecture Note Ser., vol. 249, Cambridge Univ. Press, Cambridge, 1998, pp. 74–81.
[12] Derek F. Holt, Bettina Eick, Eamonn A. O'Brien, Handbook of Computational Group Theory, Discrete Math. Appl. (Boca Raton), Chapman & Hall/CRC, Boca Raton, FL, 2005.
[13] James E. Humphreys, Linear Algebraic Groups, Grad. Texts in Math., Springer-Verlag, New York, 1975.
[14] Jens Carsten Jantzen, Representations of Algebraic Groups, Math. Surveys Monogr., vol. 107, American Mathematical Society, Providence, RI, 2003.
[15] Lajos Rónyai, Computing the structure of finite algebras, J. Symbolic Comput. 9 (1990) 355–373.
[16] Alexander J.E. Ryba, Computer construction of split Cartan subalgebras, J. Algebra 309 (2) (2007) 455–483.
[17] Igor E. Shparlinski, Finite Fields: Theory and Computation, Math. Appl., vol. 477, Kluwer Academic Publishers, Dordrecht, 1999, The meeting point of number theory, computer science, coding theory and cryptography.
[18] Tonny A. Springer, Linear Algebraic Groups, second ed., Progr. Math., vol. 9, Birkhäuser, Boston, MA, 1998.
[19] Robert Steinberg, Automorphisms of classical Lie algebras, Pacific J. Math. 11 (1961) 1119–1129.