Lie algebras and Chevalley groups

Vorlesung Master oder Bachelor Veriefung

Meinolf Geck

Lehrstuhl für Algebra, Universität Stuttgart

Email: `meinolf.geck@mathematik.uni-stuttgart.de`

`https://pnp.mathematik.uni-stuttgart.de/idsr/idsr1/geckmf`

Die Theorie der Lie-Algebren und Lie-Gruppen ist ein zentrales Gebiet der modernen Mathematik, mit Bezügen zur Algebra, Analysis, Geometrie sowie zahlreichen Anwendungen etwa in der mathematischen Physik. Die Vorlesung gibt eine elementare Einführung in den algebraischen Teil dieser Theorie. Ziele sind die allgemeine Strukturtheorie der einfachen Lie-Algebren, ihre Klassifikation durch Dynkin-Diagramme sowie die Konstruktion von Chevalley-Gruppen (= algebraische Analoga von Lie-Gruppen). Die Vorlesung versucht, möglichst direkt auch einige neuere Entwicklungen mit einzubeziehen:

- Lusztig's "kanonische" Basen von einfachen Lie-Algebren,
- und die damit vereinfachte Konstruktion der Chevalley-Gruppen.

(Dies beruht auf Arbeiten, die seit 1990 erschienen sind.)

Die Vorlesung ist geeignet als Bachelor-Vertiefung oder Master-Vorlesung; sie wird auf Englisch gehalten. (Zuletzt im SoS 2020.)

Voraussetzung sind ein gutes Verständnis des Stoffes von LAAG I und II, inkl. Grundbegriffe zu Gruppen und Ringen; ansonsten werden keine besonderen Vorkenntnisse benöntigt. Basierend auf dieser Vorlesung können Bachelor-, Master- und Staatsexamensarbeiten vergeben werden.

Kommentare sehr willkommen ! (Insbesondere Druckfehler im Skript, sonstige Unklarheiten, Verbesserungsvorschläge etc.)

Stuttgart, Mai 2023

Bibliography

[1] M. ARTIN, Algebra, Prentice-Hall, 1991.

[2] C. T. BENSON AND L. C. GROVE, *Finite reflection groups*, Graduate Texts in Mathematics, vol. 99, Springer-Verlag, New York-Berlin, 1971, 1985.

[3] N. BOURBAKI, *Algèbre, Chap. 1 à 3*, Masson, Paris, 1970.

[4] N. BOURBAKI, *Groupes et algèbres de Lie, chap. 4, 5 et 6*, Hermann, Paris, 1968.

[5] N. BOURBAKI, *Groupes et algèbres de Lie, chap. 7 et 8*, Hermann, Paris, 1975.

[6] R. W. CARTER, *Simple groups of Lie type*, Wiley, New York, 1972; reprinted 1989 as Wiley Classics Library Edition.

[7] R. W. CARTER, *Lie algebras of finite and affine type*, Cambridge University Press, 2005.

[8] V. CHARI AND A. PRESSLEY, *A guide to quantum groups*, Cambridge Univ. Press, 1994.

[9] C. CHEVALLEY, Sur certains groupes simples, Tôhoku Math. J. **7** (1955), 14–66.

[10] A. J. COLEMAN, The greatest mathematical paper of all time, Math. Intelligencer **11** (1989), 29–38.

[11] K. ERDMANN AND M. WILDON, *Introduction to Lie algebras*, Springer Undergraduate Mathematics Series, Springer-Verlag London, Ltd., London, 2006.

[12] The GAP Group, *GAP – Groups, Algorithms, and Programming.* Version 4.11.0; 2020. (`http://www.gap-system.org`)

[13] S. Garibaldi, E_8, the most exceptional group, Bull. Amer. Math. Soc. (N.S.) **53** (2016), 643–671.

[14] M. Geck, On the construction of semisimple Lie algebras and Chevalley groups, Proc. Amer. Math. Soc. **145** (2017), 3233–3247.

[15] M. Geck, ChevLie: Constructing Lie algebras and Chevalley groups, J. Softw. Algebra Geom. **10** (2020), 41–49; see `https://github.com/geckmf/ChevLie.jl`.

[16] L. C. Grove, *Classical groups and geometric algebra*. Graduate Studies in Math., vol. 39, Amer. Math. Soc., Providence, RI, 2002.

[17] R. Howe, Very basic Lie theory, Amer. Math. Monthly **90** (1983), 600–623.

[18] J. E. Humphreys, *Introduction to Lie algebras and representation theory*, Graduate Texts in Mathematics, vol. 9, Springer-Verlag, New York-Berlin, 1972.

[19] N. Jacobson, Lie algebras, Dover Publications, New York, 1962.

[20] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, Julia: A fresh approach to numerical computing, SIAM Review **59** (2017), 65–98; see `https://julialang.org`.

[21] V. Kac, Infinite dimensional Lie algebras, Cambridge University Press, 1985.

[22] G. Lusztig, On quantum groups, J. Algebra **131** (1990), 466–475.

[23] G. Lusztig, *Introduction to quantum groups*, Progress in Math., vol. 110, Birkhäuser, Boston, 1993.

[24] G. Lusztig, The canonical basis of the quantum adjoint representation, J. Comb. Algebra **1** (2017), 45–57.

[25] R. V. Moody and A. Pianzola, *Lie algebras with triangular decompositions*, Canad. Math. Soc. Ser. Monographs and Advanced Texts, Wiley, New York, 1995.

[26] H. Samelson, Notes on Lie algebras, Universitext, Springer-Verlag, 1990.

[27] J-P. Serre, *Complex semisimple Lie algebras*, translated from the French by G. A. Jones, Springer–Verlag Berlin, Heidelberg, New York, 2001.

[28] R. Steinberg, *Lectures on Chevalley groups*, University Lecture Series, vol. 66, Amer. Math. Soc., 2016. (Formerly: Mimeographed notes, Department of Math., Yale University, 1967/68.)

[29] J. Tits, Sur les constantes de structure et le théorème d'existence des algèbres de Lie semi-simples, Publ. Math. I.H.E.S. **31** (1966), 21–58.

Chapter 1

Introducing Lie algebras

This chapter introduces Lie algebras and describes some fundamental constructions related to them, e.g., representations and derivations. This is illustrated with a number of examples, most notably certain matrix Lie algebras. As far as the general theory is concerned, we will arrive at the point where we can single out the important class of "semisimple" Lie algebras.

Throughout this chapter, k denotes a fixed base field. All vector spaces will be understood to be vector spaces over this field k. We use standard notions from Linear Algebra: dimension (finite or infinite), linear and bilinear maps, matrices, eigenvalues. Everything else will be formally defined but we will assume a basic familiarity with general algebraic constructions, e.g., substructures and homomorphisms.

1.1. Non-associative algebras

Let A be a vector space (over k). If we are also given a bilinear map

$$A \times A \to A, \qquad (x, y) \mapsto x \cdot y,$$

then A is called an *algebra* (over k). Familiar examples from Linear Algebra are the algebra $A = M_n(k)$ of all $n \times n$-matrices with entries in k (and the usual matrix product), or the algebra A= $k[T]$ of polynomials with coefficients in k (where T denotes an indeterminate). In these examples, the product in A is associative; in the second example, the product is also commutative. But for us here, the term “algebra” does not imply any further assumptions on the product in A (except bi-linearity). — If the product in A happens to be associative (or commutative or ...), then we say explicitly that A is an “associative algebra” (or “commutative algebra” or ...).

The usual basic algebraic constructions also apply in this general setting. We will not completely formalize all this, but assume that the reader will fill in some (easy) details if required. Some examples:

• If A is an algebra and $B \subseteq A$ is a subspace, then B is called a *subalgebra* if $x \cdot y \in B$ for all $x, y \in B$. In this case, B itself is an algebra (with product given by the restriction of $A \times A \to A$ to $B \times B$). One easily checks that, if $\{B_i\}_{i \in I}$ is a family of subalgebras (where I is any indexing set), then $\bigcap_{i \in I} B_i$ is a subalgebra.

• If A is an algebra and $B \subseteq A$ is a subspace, then B is called an *ideal* if $x \cdot y \in B$ and $y \cdot x \in B$ for all $x \in A$ and $y \in B$. In particular, B is a subalgebra in this case. Furthermore, the quotient vector space $A/B = \{x + B \mid x \in A\}$ is an algebra with product given by

$$A/B \times A/B \to A/B, \qquad (x + B, y + B) \mapsto x \cdot y + B.$$

(One checks as usual that this product is well-defined and bilinear.) Again, one easily checks that, if $\{B_i\}_{i \in I}$ is a family of ideals (where I is any indexing set), then $\bigcap_{i \in I} B_i$ is an ideal.

• If A, B are algebras, then a linear map $\varphi\colon A \to B$ is called an *algebra homomorphism* if $\varphi(x \cdot y) = \varphi(x) * \varphi(y)$ for all $x, y \in A$. (Here, “$\cdot$” is the product in A and “$*$” is the product in B.) If, furthermore, φ is bijective, then we say that φ is an *algebra isomorphism*. In this case, the inverse map $\varphi^{-1}\colon B \to A$ is also an algebra homomorphism and we write $A \cong B$ (saying that A and B are isomorphic).

• If A, B are algebras and $\varphi\colon A \to B$ is an algebra homomorphism, then the kernel $\ker(\varphi)$ is an ideal in A and the image $\varphi(A)$ is a subalgebra of B. Furthermore, we have a canonical induced homomorphism $\bar{\varphi}\colon A/\ker(\varphi) \to B$, $x + \ker(\varphi) \mapsto \varphi(x)$, which is injective and whose image equals $\varphi(A)$. Thus, we have $A/\ker(\varphi) \cong \varphi(A)$.

Some further pieces of general notation. If V is a vector space and $X \subseteq V$ is a subset, then we denote by $\langle X \rangle_k \subseteq V$ the subspace spanned by X. Now let A be an algebra. Given $X \subseteq A$, we denote by

$\langle X\rangle_{\text{alg}} \subseteq A$ the subalgebra generated by X, that is, the intersection of all subalgebras of A which contain X. One easily checks that $\langle X\rangle_{\text{alg}} = \langle \hat{X}\rangle_k$ where $\hat{X} = \bigcup_{n\geqslant 1} X_n$ and the subsets $X_n \subseteq A$ are inductively defined by $X_1 := X$ and

$$X_n := \{x\cdot y \mid x \in X_i, y \in X_{n-i} \text{ for } 1 \leqslant i \leqslant n-1\} \quad \text{for } n \geqslant 2.$$

Thus, the elements in X_n are obtained by taking the iterated product, in any order, of n elements of X. We call the elements of X_n *monomials* in X (of level n). For example, if $X = \{x,y,z\}$, then $((z\cdot(x\cdot y))\cdot z)\cdot((z\cdot y)\cdot(x\cdot x))$ is a monomial of level 8 and, in general, we have to respect the parentheses in working with such products.

Example 1.1.1. Let M be a non-empty set and $\mu\colon M\times M \to M$ be a map. Then the pair (M,μ) is called a *magma*. Now the set of all functions $f\colon M \to k$ is a vector space over k, with pointwise defined addition and scalar multiplication. Let $k[M]$ be the subspace consisting of all $f\colon M\to k$ such that $\{x \in M \mid f(x) \neq 0\}$ is finite. For $x \in M$, let $\varepsilon_x \in k[M]$ be defined by $\varepsilon_x(y) = 1$ if $x = y$ and $\varepsilon_x(y) = 0$ if $x \neq y$. Then one easily sees that $\{\varepsilon_x \mid x \in M\}$ is a basis of $k[M]$. Furthermore, we can uniquely define a bilinear map

$$k[M]\times k[M] \to k[M] \quad \text{such that} \quad (\varepsilon_x, \varepsilon_y) \mapsto \varepsilon_{\mu(x,y)}.$$

Then $A = k[M]$ is an algebra, called the *magma algebra* of M over k.

Example 1.1.2. Let $r \geqslant 1$ and $A_1,\ldots,A_r$ be algebras (all over k). Then the cartesian product $A := A_1\times\ldots\times A_r$ is a vector space with component-wise defined addition and scalar multiplication. But then A also is an algebra with product

$$A\times A \to A, \quad \big((x_1,\ldots,x_r),(y_1,\ldots,y_r)\big) \mapsto (x_1\cdot y_1,\ldots,x_r\cdot y_r),$$

where, in order to simplify the notation, we denote the product in each A_i by the same symbol "$\cdot$". For a fixed i, we have an injective algebra homomorphism $\iota_i\colon A_i \to A$ sending $x \in A_i$ to $(0,\ldots,0,x,0,\ldots,0) \in A$ (where x appears in the i-th position). If $\underline{A}_i \subseteq A$ denotes the image of ι_i, then we have a direct sum $A = \underline{A}_1 \oplus\ldots\oplus \underline{A}_r$ where each $\underline{A}_i$ is an ideal in A and, for $i \neq j$, we have $\underline{x}\cdot\underline{y} = 0$ for all $\underline{x} \in \underline{A}_i$ and $\underline{y} \in \underline{A}_j$. The algebra A is called the *direct product* of $A_1,\ldots,A_r$.

Remark 1.1.3. Let A be an algebra. For $x \in A$, we have linear maps $L_x\colon A \to A$, $y \mapsto x\cdot y$, and $R_x\colon A\to A$, $y\mapsto y\cdot x$. Then note:

$$A \text{ is associative} \quad \Leftrightarrow \quad L_x \circ R_y = R_y \circ L_x \text{ for all } x, y \in A.$$

This simple observation is a useful "trick" in proving certain identities. Here is one example. For $x \in A$, we denote $\mathrm{ad}_A(x) := L_x - R_x \in \mathrm{End}(A)$. Thus, $\mathrm{ad}_A(x)(y) = x \cdot y - y \cdot x$ for all $x, y \in A$. The following result may be regarded as a *generalised binomial formula*; it will turn out to be useful at several places in the sequel.

Lemma 1.1.4. *Let A be an associative algebra with identity element 1_A. Let $x, y \in A$, $a, b \in k$ and $n \geqslant 0$. Then*

$$\begin{aligned}&(x + (a+b)1_A)^n \cdot y \\ &\quad = \sum_{i=0}^{n} \binom{n}{i} (\mathrm{ad}_A(x) + b\,\mathrm{id}_A)^i(y) \cdot (x + a\,1_A)^{n-i}.\end{aligned}$$

(Here, $\mathrm{id}_A \colon A \to A$ denotes the identity map.)

Proof. As above, we have $\mathrm{ad}_A(x) = L_x - R_x$. Now $L_{x+(a+b)1_A}(y) = x \cdot y + (a+b)y = (L_x + (a+b)\mathrm{id}_A)(y)$ for all $y \in A$ and so

$$L_{x+(a+b)1_A} = L_x + (a+b)\mathrm{id}_A = (R_x + a\,\mathrm{id}_A) + (\mathrm{ad}_A(x) + b\,\mathrm{id}_A).$$

Since A is associative, L_x and R_x commute with each other and, hence, $\mathrm{ad}_A(x)$ commutes with both L_x and R_x. Consequently, the maps $\mathrm{ad}_A(x) + b\,\mathrm{id}_A$ and $R_{x+a1_A} = R_x + a\,\mathrm{id}_A$ commute with each other. Hence, working in $\mathrm{End}(A)$, we can apply the usual binomial formula to $L_{x+(a+b)1_A} = R_{x+a1_A} + (\mathrm{ad}_A(x) + b\,\mathrm{id}_A)$ and obtain:

$$L^n_{x+(a+b)1_A} = \sum_{i=0}^{n} \binom{n}{i} R^{n-i}_{x+a1_A} \circ (\mathrm{ad}_A(x) + b\,\mathrm{id}_A)^i.$$

Evaluating at y yields the desired formula. □

After these general considerations, we now introduce the particular (non-associative) algebras that we are interested in here.

Definition 1.1.5. Let A be an algebra (over k), with product $x \cdot y$ for $x, y \in A$. We say that A is a *Lie algebra* if this product has the following two properties:

- (Anti-symmetry) We have $x \cdot x = 0$ for all $x \in A$. Note that, using bi-linearity, this implies $x \cdot y = -y \cdot x$ for all $x, y \in A$.

- (Jacobi identity) We have $x \cdot (y \cdot z) + y \cdot (z \cdot x) + z \cdot (x \cdot y) = 0$ for all $x, y, z \in A$.

The above two rules imply the formula $x \cdot (y \cdot z) = (x \cdot y) \cdot z + y \cdot (x \cdot z)$ which has some resemblance to the rule for differentiating a product.

Usually, the product in a Lie algebra is denoted by $[x, y]$ (instead of $x \cdot y$) and called *bracket*. So the above formulae read as follows.

$$[x, x] = 0 \qquad \text{and} \qquad [x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0.$$

Usually, we will use the symbol "L" (or "$\mathfrak{g}$") to denote a Lie algebra.

Example 1.1.6. Let $L = \mathbb{R}^3$ (row vectors). Let (x, y) be the usual scalar product of $x, y \in \mathbb{R}^3$, and $x \times y$ be the "vector product" (perhaps known from a Linear Algebra course). That is, given $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$ in L, we have $x \times y = (v_1, v_2, v_3) \in L$ where

$$v_1 = x_2 y_3 - x_3 y_2, \qquad v_2 = x_3 y_1 - x_1 y_3, \qquad v_3 = x_1 y_2 - x_2 y_1.$$

One easily verifies the "Grassmann identity" $x \times (y \times z) = (x, z)\, y - (x, y)\, z$ for $x, y, z \in \mathbb{R}^3$. Setting $[x, y] := x \times y$ for $x, y \in L$, a straightforward computation shows that L is a Lie algebra over $k = \mathbb{R}$.

Example 1.1.7. Let L be a Lie algebra. If $V \subseteq L$ is any subspace, the *normalizer* of V is defined as

$$I_L(V) := \{x \in L \mid [x, v] \in V \text{ for all } v \in V\}.$$

Clearly, $I_L(V)$ is a subspace of L. We claim that $I_L(V)$ is a Lie subalgebra of L. Indeed, let $x, y \in I_L(V)$ and $v \in V$. By the Jacobi identity and anti-symmetry, we have

$$[[x, y], v] = -[v, [x, y]] = [x, \underbrace{[y, v]}_{\in V}] - [y, \underbrace{[x, v]}_{\in V}] \in V.$$

If V is a Lie subalgebra, then $V \subseteq I_L(V)$ and V is an ideal in $I_L(V)$.

Exercise 1.1.8. Let L be a Lie algebra and $X \subseteq L$ be a subset.

(a) Let $V \subseteq L$ be a subspace such that $[x, v] \in V$ for all $x \in X$ and $v \in V$. Then show that $[y, v] \in V$ for all $y \in \langle X \rangle_{\text{alg}}$ and $v \in V$. Furthermore, if $X \subseteq V$, then $\langle X \rangle_{\text{alg}} \subseteq V$.

(b) Let $I := \langle X \rangle_{\text{alg}} \subseteq L$. Assume that $[y, x] \in I$ for all $x \in X$, $y \in L$. Then show that I is an ideal of L.

(c) Let L' be a further Lie algebra and $\varphi\colon L \to L'$ be a linear map. Assume that $L = \langle X\rangle_{\text{alg}}$. Then show that φ is a Lie algebra homomorphism if $\varphi([x, y]) = [\varphi(x), \varphi(y)]$ for all $x \in X$ and $y \in L$.

Example 1.1.9. (a) Let V be a vector space. We define $[x, y] := 0$ for all $x, y \in V$. Then, clearly, V is a Lie algebra. A Lie algebra in which the bracket is identically 0 is called an *abelian Lie algebra.*

(b) Let A be an algebra that is associative. Then we define a new product on A by $[x, y] := x \cdot y - y \cdot x$ for all $x, y \in A$. Clearly, this is bilinear and we have $[x, x] = 0$; furthermore, for $x, y, z \in A$, we have

$$\begin{aligned}
[x,[y,z]]+&[y,[z,x]] + [z,[x,y]] \\
&= [x, y\cdot z - z\cdot y] + [y, z\cdot x - x\cdot z] + [z, x\cdot y - y\cdot x] \\
&= x\cdot(y\cdot z - z\cdot y) - (y\cdot z - z\cdot y)\cdot x \\
&\qquad + y\cdot(z\cdot x - x\cdot z) - (z\cdot x - x\cdot z)\cdot y \\
&\qquad\quad + z\cdot(x\cdot y - y\cdot x) - (x\cdot y - y\cdot x)\cdot z.
\end{aligned}$$

By associativity, we have $x\cdot(y\cdot z) = (x\cdot y)\cdot z$ and so on. We then leave it to the reader to check that the above sum collapses to 0. Thus, every associative algebra becomes a Lie algebra by this construction.

A particular role in the general theory is played by those algebras that do not have non-trivial ideals. This leads to:

Definition 1.1.10. Let A be an algebra such that $A \neq \{0\}$ and the product of A is not identically zero. Then A is called a *simple algebra* if $\{0\}$ and A are the only ideals of A.

We shall see first examples in the following section.

Exercise 1.1.11. This exercise (which may be skipped on a first reading) presents a very general method for constructing algebras with prescribed properties. Recall from Example 1.1.1 the definition of a magma. Given a non-empty set X, we want to define the "most general magma" containing X, following Bourbaki [**3**, Chap. I, §7, no. 1]. For this purpose, we define inductively sets X_n for $n = 1, 2, \ldots$, as follows. We set $X_1 := X$. Now let $n \geqslant 2$ and assume that X_i is already defined for $1 \leqslant i \leqslant n-1$. Then define X_n to the disjoint union of the sets $X_i \times X_{n-i}$ for $1 \leqslant i \leqslant n-1$. Finally, we define $M(X)$ to be the disjoint union of all the sets X_n, $n \geqslant 1$.

Now let $w, w' \in M(X)$. Since $M(X)$ is the disjoint union of all X_n, there are unique $p, p' \geqslant 1$ such that $w \in X_p$ and $w' \in X_{p'}$. Let $n := p+p'$. By the definition of X_n, we have $X_p \times X_{p'} \subseteq X_n$. Then define $w * w' \in X_n$ to be the pair $(w, w') \in X_p \times X_{p'} \subseteq X_n$. In this way, we obtain a product $M(X) \times M(X) \to M(X)$, $(w, w') \mapsto w * w'$. So $M(X)$ is a magma, called the *free magma* on X.

Thus, one may think of the elements of $M(X)$ as arbitrary "non-associative words" formed using X. For example, if $X = \{a, b\}$, then $(a*b)*a, (b*a)*a, a*(b*a), (a*(a*b))*b, (a*a)*(b*b)$ are pairwise distinct elements of $M(X)$; and all elements of $M(X)$ are obtained by forming such products.

(a) Show the following *universal property of the free magma*. For any magma (N, ν) and any map $\varphi\colon X \to N$, there exists a unique map $\hat{\varphi}\colon M(X) \to N$ such that $\hat{\varphi}|_X = \varphi$ and $\hat{\varphi}$ is a magma homomorphism (meaning that $\hat{\varphi}(w * w') = \nu(\hat{\varphi}(w), \hat{\varphi}(w'))$ for all $w, w' \in M(X)$).

(b) As in Example 1.1.1, let $F_k(X) := k[M(X)]$ be the magma algebra over k of the free magma $M(X)$. Note that, as an algebra, $F_k(X)$ is generated by $\{\varepsilon_x \mid x \in X\}$. We denote the product of two elements $a, b \in F_k(X)$ by $a \cdot b$. Let I be the ideal of $F_k(X)$ which is generated by all elements of the form

$$a \cdot a \qquad \text{or} \qquad a \cdot (b \cdot c) + b \cdot (c \cdot a) + c \cdot (a \cdot b),$$

for $a, b, c \in F_k(X)$. (Thus, I is the intersection of all ideals of $F_k(X)$ that contain the above elements.) Let $L(X) := F_k(X)/I$ and $\iota\colon X \to L(X)$, $x \mapsto \varepsilon_x + I$. Show that $L(X)$ is a Lie algebra over k which has the following *universal property*. For any Lie algebra L' over k and any map $\varphi\colon X \to L'$, there exists a unique Lie algebra homomorphism $\hat{\varphi}\colon L(X) \to L'$ such that $\varphi = \hat{\varphi} \circ \iota$. Deduce that ι is injective.

The Lie algebra $L(X)$ is called the *free Lie algebra* over X. By taking factor algebras of $L(X)$ by an ideal, we can construct Lie algebras in which prescribed relations hold. (See, e.g., Exercise 1.2.11.)

1.2. Matrix Lie algebras and derivations

We have just seen that every associative algebra can be turned into a Lie algebra. This leads to the following concrete examples.

Example 1.2.1. Let V be a vector space. Then $\mathrm{End}(V)$ denotes as usual the vector space of all linear maps $\varphi\colon V \to V$. In fact, $\mathrm{End}(V)$ is an associative algebra where the product is given by the composition of maps; the identity map $\mathrm{id}_V\colon V \to V$ is the identity element for this product. Applying the construction in Example 1.1.9, we obtain a bracket on $\mathrm{End}(V)$ and so $\mathrm{End}(V)$ becomes a Lie algebra, denoted $\mathfrak{gl}(V)$. Thus, $\mathfrak{gl}(V) = \mathrm{End}(V)$ as vector spaces and

$$[\varphi, \psi] = \varphi \circ \psi - \psi \circ \varphi \qquad \text{for all } \varphi, \psi \in \mathfrak{gl}(V).$$

Now assume that $\dim V < \infty$ and let $B = \{v_i \mid i \in I\}$ be a basis of V. We denote by $M_I(k)$ the algebra of all matrices with entries in k and rows and columns indexed by I, with the usual matrix product. For $\varphi \in \mathrm{End}(V)$, we denote by $M_B(\varphi)$ the matrix of φ with respect to B; thus, $M_B(\varphi) = (a_{ij})_{i,j\in I} \in M_I(k)$ where $\varphi(v_j) = \sum_{i\in I} a_{ij} v_i$ for all j. Now applying the construction in Example 1.1.9, we obtain a bracket on $M_I(k)$ and so $M_I(k)$ also becomes a Lie algebra, denoted $\mathfrak{gl}_I(k)$. Thus, $\mathfrak{gl}_I(k) = M_I(k)$ as vector spaces and

$$[X, Y] = X \cdot Y - Y \cdot X \qquad \text{for all } X, Y \in \mathfrak{gl}_I(k).$$

The map $\varphi \mapsto M_B(\varphi)$ defines an isomorphism of associative algebras $\mathrm{End}(V) \cong M_I(k)$. Consequently, this map also defines an isomorphism of Lie algebras $\mathfrak{gl}(V) \cong \mathfrak{gl}_I(k)$. (Of course, if $I = \{1, \ldots, n\}$ where $n = \dim V$, then we write as usual $M_n(k)$ and $\mathfrak{gl}_n(k)$ instead of $M_I(k)$ and $\mathfrak{gl}_I(k)$, respectively.)

Example 1.2.2. Let $\mathfrak{gl}(V)$ be as in the previous example, where $\dim V < \infty$. Then consider the map $\mathrm{Trace}\colon \mathfrak{gl}(V) \to k$ which sends each $\varphi \in \mathfrak{gl}(V)$ to the trace of φ (that is, the sum of the diagonal entries of $M_B(\varphi)$, for some basis $B = \{v_i \mid i \in I\}$ of V). Since $\mathrm{Trace}(\varphi \circ \psi) = \mathrm{Trace}(\psi \circ \varphi)$ for all $\varphi, \psi \in \mathfrak{gl}(V)$, we deduce that

$$\mathfrak{sl}(V) := \{\varphi \in \mathfrak{gl}(V) \mid \mathrm{Trace}(\varphi) = 0\}$$

is a Lie subalgebra of $\mathfrak{gl}(V)$. (Note that $\mathfrak{sl}(V)$ is not a subalgebra with respect to the matrix product!) Considering matrices as above, we have analogous definitions of $\mathfrak{sl}_I(k)$ and $\mathfrak{sl}_n(k)$ (where $I = \{1, \ldots, n\}$).

Exercise 1.2.3. Let L be a Lie algebra. If $\dim L = 1$, then L is clearly abelian. Now assume that $\dim L = 2$ and that L is not abelian. Show that L has a basis $\{x, y\}$ such that $[x, y] = y$; in particular, $\langle y \rangle_k$

is a non-trivial ideal of L and so L is not simple. Show that L is isomorphic to the following Lie subalgebra of $\mathfrak{gl}_2(k)$:

$$\left\{\begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} \;\middle|\; a, b \in k\right\}.$$

In particular, if L is a simple Lie algebra, then $\dim L \geqslant 3$.

Exercise 1.2.4. This is a reminder of a basic result from Linear Algebra. Let V be a vector space and $\varphi\colon V \to V$ be a linear map. Let $v \in V$. We say that φ is *locally nilpotent* at v if there exists some $d \geqslant 1$ (which may depend on v) such that $\varphi^d(v) = 0$. We say that φ is *nilpotent* if $\varphi^d = 0$ for some $d \geqslant 1$. Assume now that $\dim V < \infty$.

(a) Let $X \subseteq V$ be a subset such that $V = \langle X \rangle_k$. Assume that φ is locally nilpotent at every $v \in X$. Show that φ is nilpotent.

(b) Show that, if φ is nilpotent, then there is a basis B of V such that the matrix of φ with respect to B is triangular with 0 on the diagonal; in particular, we have $\varphi^{\dim V} = 0$ and the trace of φ is 0.

Example 1.2.5. Let L be a Lie algebra. In analogy to Remark 1.1.3 and Example 1.1.9(b), we define for $x \in L$ the linear map

$$\mathrm{ad}_L(x)\colon L \to L, \qquad y \mapsto [x, y].$$

Hence, we obtain a linear map $\mathrm{ad}_L\colon L \to \mathrm{End}(L)$, $x \mapsto \mathrm{ad}_L(x)$. By the Jacobi identity and anti-symmetry, we have

$$\begin{aligned}\mathrm{ad}_L([x, y])(z) &= [[x, y], z] = -[z, [x, y]] \\ &= [x, [y, z]] + [y, [z, x]] = [x, [y, z]] - [y, [x, z]] \\ &= (\mathrm{ad}_L(x) \circ \mathrm{ad}_L(y) - \mathrm{ad}_L(y) \circ \mathrm{ad}_L(x))(z)\end{aligned}$$

for all $z \in L$ and so $\mathrm{ad}_L([x, y]) = [\mathrm{ad}_L(x), \mathrm{ad}_L(y)]$. Thus, we obtain a Lie algebra homomorphism $\mathrm{ad}_L\colon L \to \mathfrak{gl}(L)$. (See also Example 1.4.3 below.) The kernel of ad_L is called the *center* of L and will be denoted by $Z(L)$; thus, $Z(L)$ is an ideal of L and

$$Z(L) = \ker(\mathrm{ad}_L) = \{x \in L \mid [x, y] = 0 \text{ for all } y \in L\}.$$

Finally, for $x, y, z \in L$, we also have the identity

$$\begin{aligned}\mathrm{ad}_L(z)([x, y]) &= [z, [x, y]] = -[x, [y, z]] - [y, [z, x]] \\ &= [x, [z, y]] + [[z, x], y] = [x, \mathrm{ad}_L(z)(y)] + [\mathrm{ad}_L(z)(x), y]\end{aligned}$$

which shows that $\mathrm{ad}_L(z)$ is a derivation in the following sense.

Definition 1.2.6. Let A be an algebra. A linear map $\delta\colon A \to A$ is called a *derivation* if $\delta(x \cdot y) = x \cdot \delta(y) + \delta(x) \cdot y$ for all $x, y \in A$. Let $\mathrm{Der}(A)$ be the set of all derivations of A. One immediately checks that $\mathrm{Der}(A)$ is a subspace of $\mathrm{End}(A)$.

Exercise 1.2.7. Let A be an algebra.

(a) Show that $\mathrm{Der}(A)$ is a Lie subalgebra of $\mathfrak{gl}(A)$.

(b) Let $\delta\colon A \to A$ be a derivation. Show that, for any $n \geqslant 0$, we have the *Leibniz rule*

$$\delta^n(x \cdot y) = \sum_{i=0}^{n} \binom{n}{i} \delta^i(x) \cdot \delta^{n-i}(y) \quad \text{for all } x, y \in A.$$

Derivations are a source for Lie algebras which do not arise from associative algebras as in Example 1.1.9; see Example 1.2.9 below. The following construction with nilpotent derivations will play a major role in Chapter 3.

Lemma 1.2.8. *Let A be an algebra where the ground field k has characteristic 0. If $d\colon A \to A$ is a derivation such that $d^n = 0$ for some $n > 0$ (that is, d is nilpotent), we obtain a map*

$$\exp(d)\colon A \to A, \qquad x \mapsto \sum_{0 \leqslant i < n} \frac{d^i(x)}{i!} = \sum_{i \geqslant 0} \frac{d^i(x)}{i!}.$$

Then $\exp(d)$ is an algebra isomorphism, with inverse $\exp(-d)$.

Proof. Since d^i is linear for all $i \geqslant 0$, it is clear that $\exp(d)\colon A \to A$ is a linear map. For $x, y \in A$, we have

$$\begin{aligned}
\exp(d)(x) \cdot \exp(d)(y) &= \Big(\sum_{i \geqslant 0} \frac{d^i}{i!}(x)\Big) \cdot \Big(\sum_{j \geqslant 0} \frac{d^j}{j!}(y)\Big) \\
&= \sum_{i,j \geqslant 0} \frac{d^i}{i!}(x) \cdot \frac{d^j}{j!}(y) = \sum_{m \geqslant 0} \Big(\sum_{\substack{i,j \geqslant 0 \\ i+j=m}} \frac{d^i}{i!}(x) \cdot \frac{d^j}{j!}(y)\Big) \\
&= \sum_{m \geqslant 0} \frac{1}{m!} \Big(\sum_{0 \leqslant i \leqslant m} \binom{m}{i} d^i(x) \cdot d^{m-i}(y)\Big) = \sum_{m \geqslant 0} \frac{d^m}{m!}(x \cdot y),
\end{aligned}$$

where the last equality holds by the Leibniz rule. Hence, the right side equals $\exp(d)(x \cdot y)$. Thus, $\exp(d)$ is an algebra homomorphism.

Now, we can also form $\exp(-d)$ and $\exp(0)$, where the definition immediately shows that $\exp(0) = \mathrm{id}_A$. So, for any $x \in A$, we obtain:

$$x = \exp(0)(x) = \exp(d+(-d))(x) = \sum_{m \geqslant 0} \frac{(d+(-d))^m(x)}{m!}.$$

Since d and $-d$ commute with each other, we can apply the binomial formula to $(d+(-d))^m$. So the right hand side evaluates to

$$\begin{aligned}\sum_{m\geqslant 0}\frac{1}{m!}\sum_{\substack{i,j\geqslant 0\\ i+j=m}}\frac{m!}{i!\,j!}(d^i\circ(-d)^j)(x) &= \sum_{i,j\geqslant 0}\frac{(d^i\circ(-d)^j)(x)}{i!\,j!}\\ = \sum_{i,j\geqslant 0}\frac{d^i}{i!}\Big(\frac{(-d)^j}{j!}(x)\Big) &= \sum_{i\geqslant 0}\frac{d^i}{i!}\Big(\sum_{j\geqslant 0}\frac{(-d)^j}{j!}(x)\Big)\\ = \sum_{i\geqslant 0}\frac{d^i}{i!}\big(\exp(-d)(x)\big) &= \exp\big(\exp(-d)(x)\big).\end{aligned}$$

Hence, we see that $\exp(d) \circ \exp(-d) = \mathrm{id}_A$; similarly, $\exp(-d) \circ \exp(d) = \mathrm{id}_A$. So $\exp(d)$ is invertible, with inverse $\exp(-d)$. $\square$

Example 1.2.9. Let $A = k[T, T^{-1}]$ be the algebra of Laurent polynomials in the indeterminate T. Let us determine $\mathrm{Der}(A)$. Since $A = \langle T, T^{-1}\rangle_{\mathrm{alg}}$, the product rule for derivations implies that every $\delta \in \mathrm{Der}(A)$ is uniquely determined by $\delta(T)$ and $\delta(T^{-1})$. Now $\delta(1) = \delta(T \cdot T^{-1}) = T\delta(T^{-1}) + \delta(T)T^{-1}$. Since $\delta(1) = \delta(1) + \delta(1)$, we have $\delta(1) = 0$ and so $\delta(T^{-1}) = -T^{-2}\delta(T)$. Hence, we conclude:

(a) Every $\delta \in \mathrm{Der}(A)$ is uniquely determined by its value $\delta(T)$.

For $m \in \mathbb{Z}$ we define a linear map $L_m \colon A \to A$ by

$$L_m(f) = -T^{m+1}D(f) \qquad \text{for all } f \in A,$$

where $D \colon A \to A$ denotes the usual formal derivate with respect to T, that is, D is linear and $D(T^n) = nD(T^{n-1})$ for all $n \in \mathbb{Z}$. Now $D \in \mathrm{Der}(A)$ (by the product rule for formal derivates) and so $L_m \in \mathrm{Der}(A)$. We have $L_m(T) = -T^{m+1}D(T) = -T^{m+1}$. Hence, if $\delta \in \mathrm{Der}(A)$ and $\delta(T) = \sum_i a_i T^i$ with $a_i \in k$, then $-\delta$ and the sum $\sum_i a_i L_{i-1}$ have the same value on T. So $-\delta$ must be equal to that sum by (a). Thus, we have shown that

(b) $$\mathrm{Der}(A) = \langle L_m \mid m \in \mathbb{Z}\rangle_k.$$

In fact, $\{L_m \mid m \in \mathbb{Z}\}$ is a basis of $\mathrm{Der}(A)$. (Just apply a linear combination of the L_m's to T and use the fact that $L_m(T) = -T^{m+1}$.) Now let $m, n \in \mathbb{Z}$. Using the bracket in $\mathfrak{gl}(A)$, we obtain that

$$[L_m, L_n](T) = (L_m \circ L_n - L_n \circ L_m)(T) = \ldots = (n-m)T^{m+n+1},$$

which is also the result of $(m-n)L_{m+n}(T)$. By Exercise 1.2.7(a), we have $[L_m, L_n] \in \mathrm{Der}(A)$. So (a) shows again that

(c) $$[L_m, L_n] = (m-n)L_{m+n} \quad \text{for all } m, n \in \mathbb{Z}.$$

Thus, $\mathrm{Der}(A)$ is an infinite-dimensional Lie subalgebra of $\mathfrak{gl}(A)$, with basis $\{L_m \mid m \in \mathbb{Z}\}$ and bracket determined as above; this Lie algebra is called a *Witt algebra* (or *centerless Virasoro algebra*; see also the notes at the end of this chapter).

Proposition 1.2.10. *Let $L = \mathrm{Der}(A)$ be the Witt algebra in Example 1.2.9. If $\mathrm{char}(k) = 0$, then L is a simple Lie algebra.*

Proof. Let $I \subseteq L$ be a non-zero ideal and $0 \neq x \in I$. Then we can write $x = c_1 L_{m_1} + \ldots + c_r L_{m_r}$ where $r \geqslant 1$, $m_1 < \ldots < m_r$ and all $c_i \in k$ are non-zero. Choose x such that r is as small as possible. We claim that $r = 1$. Assume, if possible, that $r \geqslant 2$. Since $[L_0, L_m] = -mL_m$ for all $m \in \mathbb{Z}$, we obtain that $[L_0, x] = -c_1 m_1 L_{m_1} - \ldots - c_r m_r L_{m_r} \in I$. Hence,

$$m_r x + [L_0, x] = c_1(m_r - m_1)L_{m_1} + \ldots + c_{r-1}(m_r - m_{r-1})L_{m_{r-1}}$$

is a non-zero element of I, contradiction to the minimality of r. Hence, $r = 1$ and so $L_{m_1} \in I$. Now $[L_{m-m_1}, L_{m_1}] = (m - 2m_1)L_m$ and so $L_m \in I$ for all $m \in \mathbb{Z}$, $m \neq 2m_1$. But $[L_{m_1+1}, L_{m_1-1}] = 2L_{2m_1}$ and so we also have $L_{2m_1} \in I$. Hence, we do have $I = L$, as desired. $\square$

Exercise 1.2.11. Let $L = \mathfrak{sl}_2(k)$, as in Example 1.2.2. Then $\dim L = 3$ and L has a basis $\{e, h, f\}$ where

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

(a) Check that $[e, f] = h$, $[h, e] = 2e$, $[h, f] = -2f$. Show that L is simple if $\mathrm{char}(k) \neq 2$. What happens if $\mathrm{char}(k) = 2$? Consider also the Lie algebra L' in Example 1.1.6. Is $L' \cong \mathfrak{sl}_2(\mathbb{R})$? Is L' simple? What happens if we work with $\mathbb{C}$ instead of $\mathbb{R}$?

(b) Let $\hat{L}$ be the free Lie algebra over the set $X = \{E, H, F\}$; see Exercise 1.1.11. Let $I \subseteq \hat{L}$ be the ideal generated by $[E, F] - H$, $[H, E] - 2E$, $[H, F] + 2F$ (that is, the intersection of all ideals containing those elements). By the universal property, there is a unique homomorphism of Lie algebras $\varphi \colon \hat{L} \to L$ such that $\varphi(E) = e$, $\varphi(H) = h$ and $\varphi(F) = f$. By (a), we have $I \subseteq \ker(\varphi)$. Show that the induced homomorphism $\bar{\varphi} \colon \hat{L}/I \to L$ is an isomorphism.

Exercise 1.2.12. (a) Show that $Z(\mathfrak{gl}_n(k)) = \{aI_n \mid a \in k\}$ (where I_n denotes the $n \times n$-identity matrix). What happens for $Z(\mathfrak{sl}_n(k))$?

(b) Let $X \subseteq L$ be a subset. Let $z \in L$ be such that $[x, z] = 0$ for all $x \in X$. Then show that $[y, z] = 0$ for all $y \in \langle X \rangle_{\text{alg}}$.

Exercise 1.2.13. This exercise describes a useful method for constructing new Lie algebras out of two given ones. So let S, I be Lie algebras over k and $\theta \colon S \to \text{Der}(I)$, $s \mapsto \theta_s$, be a homomorphism of Lie algebras. Consider the vector space $L = S \times I = \{(s, x) \mid s \in S, x \in I\}$ (with component-wise defined addition and scalar multiplication). For $s_1, s_2 \in S$ and $x_1, x_2 \in I$ we define

$$[(s_1, x_1), (s_2, x_2)] := \big([s_1, s_2], [x_1, x_2] + \theta_{s_1}(x_2) - \theta_{s_2}(x_1)\big).$$

Show that L is a Lie algebra such that $L = \underline{S} \oplus \underline{I}$, where

$$\underline{S} := \{(s, 0) \mid s \in S\} \subseteq L \quad \text{is a subalgebra,}$$
$$\underline{I} := \{(0, x) \mid x \in I\} \subseteq L \quad \text{is an ideal.}$$

We also write $L = S \ltimes_\theta I$ and call L the *semidirect product* of I by S (via θ). If $\theta(s) = 0$ for all $s \in S$, then $[(s_1, x_1), (s_2, x_2)] = ([s_1, s_2], [x_1, x_2])$ for all $s_1, s_2 \in S$ and $x_1, x_2 \in I$. Hence, in this case, L is the *direct product* of S and I, as in Example 1.1.2.

Exercise 1.2.14. Let A be an algebra where the ground field k has characteristic 0. Let $d \colon A \to A$ and $d' \colon A \to A$ be nilpotent derivations such that $d \circ d' = d' \circ d$. Show that $d + d'$ also is a nilpotent derivation and that $\exp(d + d') = \exp(d) \circ \exp(d')$.

Exercise 1.2.15. This exercise gives a first outlook to some constructions that will be studied in much greater depth and generality in Chapter 3. Let $L \subseteq \mathfrak{gl}(V)$ be a Lie subalgebra, where V is a finite-dimesional $\mathbb{C}$-vector space. Let $\text{Aut}(L)$ be the group of all Lie algebra automorphisms of L.

(a) Assume that $a \in L$ is nilpotent (as linear map $a\colon V \to V$). Then show that the linear map $\mathrm{ad}_L(a)\colon L \to L$ is nilpotent. (Hint: use the "trick" in Remark 1.1.3.) Is the converse also true?

(b) Let $L = \mathfrak{sl}_2(\mathbb{C})$ with basis elements e, h, f as in Exercise 1.2.11. Note that e and f are nilpotent matrices. Hence, by (a), the derivations $\mathrm{ad}_L(e)\colon L \to L$ and $\mathrm{ad}_L(f)\colon L \to L$ are nilpotent. Consequently, $t\,\mathrm{ad}_L(e)$ and $t\,\mathrm{ad}_L(f)$ are nilpotent derivations for all $t \in \mathbb{C}$. By Lemma 1.2.8, we obtain Lie algebra automorphisms

$$\exp\bigl(t\,\mathrm{ad}_L(e)\bigr)\colon L \to L \qquad \text{and} \qquad \exp\bigl(t\,\mathrm{ad}_L(f)\bigr)\colon L \to L;$$

we will denote these by $x(t)$ and $y(t)$, respectively. Determine the matrices of these automorphisms with respect to the basis $\{e, h, f\}$ of L. Check that $x(t+t') = x(t)x(t')$ and $y(t+t') = y(t)y(t')$ for all $t, t' \in \mathbb{C}$. The subgroup $G := \langle x(t), y(t') \mid t, t' \in \mathbb{C}\rangle \subseteq \mathrm{Aut}(L)$ is called the *Chevalley group* associated with the Lie algebra $L = \mathfrak{sl}_2(\mathbb{C})$. The elements of G are completely described as follows. First, compute the matrices of the following elements of G, where $u \in \mathbb{C}^\times$:

$$w(u) := x(u)y(-u^{-1})x(u) \qquad \text{and} \qquad h(u) := w(u)w(-1).$$

Check the relations $w(u)x(t)w(u)^{-1} = y(-u^{-2}t)$ and $h(u)h(u') = h(u')h(u)$ for all $t \in \mathbb{C}$ and $u, u' \in \mathbb{C}^\times$. In particular, we have

$$G = \langle x(t), w(u) \mid t \in \mathbb{C}, u \in \mathbb{C}^\times\rangle.$$

Finally, show that every element $g \in G$ can be written uniquely as either $g = x(t)h(u)$ (with $t \in \mathbb{C}$ and $u \in \mathbb{C}^\times$) or $g = x(t)w(u)x(t')$ (with $t, t' \in \mathbb{C}$ and $u \in \mathbb{C}^\times$).

1.3. Solvable and semisimple algebras

ab hier Woche 2

Let A be an algebra. If $U, V \subseteq A$ are subspaces, then we denote

$$U \cdot V := \langle u \cdot v \mid u \in U, v \in V\rangle_k \subseteq A.$$

In general, $U \cdot V$ will only be a subspace of A, even if U, V are subalgebras or ideals. On the other hand, taking $U = V = A$, then

$$A^2 := A \cdot A = \langle x \cdot y \mid x, y \in A\rangle_k$$

clearly is an ideal of A, and the induced product on A/A^2 is identically zero. So we can iterate this process: Let us set $A^{(0)} := A$ and then

$$A^{(1)} := A^2, \quad A^{(2)} := (A^{(1)})^2, \quad A^{(3)} := (A^{(2)})^2, \quad \ldots.$$

Thus, we obtain a chain of subalgebras $A = A^{(0)} \supseteq A^{(1)} \supseteq A^{(2)} \supseteq \ldots$ such that $A^{(i+1)}$ is an ideal in $A^{(i)}$ for all i and the induced product on $A^{(i)}/A^{(i+1)}$ is identically zero. An easy induction on j shows that $A^{(i+j)} = (A^{(i)})^{(j)}$ for all $i, j \geqslant 0$.

Definition 1.3.1. We say that A is a *solvable algebra* if $A^{(m)} = \{0\}$ for some $m \geqslant 0$ (and, hence, $A^{(l)} = \{0\}$ for all $l \geqslant m$.)

Note that the above definitions are only useful if A does not have an identity element which is, in particular, the case for Lie algebras by the anti-symmetry condition in Definition 1.1.5.

Example 1.3.2. (a) All Lie algebras of dimension $\leqslant 2$ are solvable; see Exercise 1.2.3.

(b) Let $n \geqslant 1$ and $\mathfrak{b}_n(k) \subseteq \mathfrak{gl}_n(k)$ be the subspace consisting of all upper triangular matrices, that is, all $(a_{ij})_{1\leqslant i,j\leqslant n} \in \mathfrak{gl}_n(k)$ such that $a_{ij} = 0$ for all $i > j$. Since the product of two upper triangular matrices is again upper triangular, it is clear that $\mathfrak{b}_n(k)$ is a Lie subalgebra of $\mathfrak{gl}_n(k)$. An easy matrix calculation shows that $\mathfrak{b}_n(k)^{(1)} = [\mathfrak{b}_n(k), \mathfrak{b}_n(k)]$ consists of upper tiangular matrices with 0 on the diagonal. More generally, $\mathfrak{b}_n(k)^{(r)}$ for $1 \leqslant r \leqslant n$ consists of upper triangular matrices (a_{ij}) such that $a_{ij} = 0$ for all $i \leqslant j < i+r$. In particular, we have $\mathfrak{b}_n(k)^{(n)} = \{0\}$ and so $\mathfrak{b}_n(k)$ is solvable.

Exercise 1.3.3. For a fixed $0 \neq \delta \in k$, we define

$$L_\delta := \left\{ \begin{pmatrix} a & b & 0 \\ 0 & 0 & 0 \\ 0 & c & a\delta \end{pmatrix} \;\middle|\; a, b, c \in k \right\} \subseteq \mathfrak{gl}_3(k).$$

Show that L_δ is a solvable Lie subalgebra of $\mathfrak{gl}_3(k)$, where $[L_\delta, L_\delta]$ is abelian. Show that, if $L_\delta \cong L_{\delta'}$, then $\delta = \delta'$ or $\delta^{-1} = \delta'$. Hence, if $|k| = \infty$, then there are infinitely many pairwise non-isomorphic solvable Lie algebras of dimension 3. (See [**11**, Chap. 3] for a further discussion of "low-dimensional" examples of solvable Lie algebras.)

[*Hint.* A useful tool to check that two Lie algebras cannot be isomorphic is as follows. Let L_1, L_2 be finite-dimensional Lie algebras over k. Let $\varphi\colon L_1 \to L_2$ be an isomorphism. Show that $\varphi \circ \mathrm{ad}_{L_1}(x) = \mathrm{ad}_{L_2}(\varphi(x)) \circ \varphi$ for $x \in L_1$. Deduce that

$\mathrm{ad}_{L_1}(x)\colon L_1 \to L_1$ and $\mathrm{ad}_{L_2}(\varphi(x))\colon L_2 \to L_2$ must have the same characteristic polynomial. Try to apply this with the element $x \in L_\delta$ where $a = 1$, $b = c = 0$.]

Exercise 1.3.4. Let L be a Lie algebra over k with $\dim L = 2n+1$, $n \geqslant 1$. Suppose that L has a basis $\{z\} \cup \{e_i, f_i \mid 1 \leqslant i \leqslant n\}$ such that $[e_i, f_i] = z$ for $1 \leqslant i \leqslant r$ and all other Lie brackets between basis vectors are 0. Then L is called a *Heisenberg Lie algebra* (see [**25**, §1.4] for further background). Check that $[L, L] = Z(L) = \langle z \rangle_k$; in particular, L is solvable. Show that, for $n = 1$,

$$L := \left\{ \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix} \;\middle|\; a, b, c \in k \right\} \subseteq \mathfrak{gl}_3(k)$$

is a Heisenberg Lie algebra; find a basis $\{z\} \cup \{e_1, f_1\}$ as above.

Lemma 1.3.5. *Let A be an algebra.*

(a) *Let B be an algebra and $\varphi\colon A \to B$ be a surjective algebra homomorphism. Then $\varphi(A^{(i)}) = B^{(i)}$ for all $i \geqslant 0$.*

(b) *Let $B \subseteq A$ be a subalgebra. Then $B^{(i)} \subseteq A^{(i)}$ for all $i \geqslant 0$.*

(c) *Let $I \subseteq A$ be an ideal. Then A is solvable if and only if I and A/I are solvable.*

Proof. (a) Induction on i. If $i = 0$, then this holds by assumption. Let $i \geqslant 0$. Then $\varphi(A^{(i+1)}) = \varphi(A^{(i)} \cdot A^{(i)}) = \langle \varphi(x) \cdot \varphi(y) \mid x, y \in A^{(i)} \rangle_k$ which equals $B^{(i)} \cdot B^{(i)}$ since $\varphi(A^{(i)}) = B^{(i)}$ by induction.

(b) Induction on i. If $i = 0$, then this is clear. Now let $i \geqslant 0$. By induction, $B^{(i)} \subseteq A^{(i)}$ and so $B^{(i+1)} = (B^{(i)})^2 \subseteq (A^{(i)})^2 = A^{(i+1)}$.

(c) If A is solvable, then I and A/I are solvable by (a), (b). Conversely, let $m, l \geqslant 0$ be such that $I^{(l)} = \{0\}$ and $(A/I)^{(m)} = \{0\}$. Let $\varphi\colon A \to A/I$ be the canonical map. Then $\varphi(A^{(m)}) = (A/I)^{(m)} = \{0\}$ by (a), hence, $A^{(m)} \subseteq \ker(\varphi) = I$. Using (b), we obtain $A^{(m+l)} = (A^{(m)})^{(l)} \subseteq I^{(l)} = \{0\}$ and so A is solvable. □

Corollary 1.3.6. *Let A be an algebra with* $\dim A < \infty$. *Then the set of all solvable ideals of A is non-empty and contains a unique maximal element (with respect to inclusion). This unique maximal solvable ideal will be denoted* $\mathrm{rad}(A)$ *and called the radical of A. We have* $\mathrm{rad}(A/\mathrm{rad}(A)) = \{0\}$.

Proof. First note that $\{0\}$ is a solvable ideal of A. Now let $I \subseteq A$ be a solvable ideal such that $\dim I$ is as large as possible. Let $J \subseteq A$ be another solvable ideal. Clearly, $B := \{x+y \mid x \in I, y \in J\} \subseteq A$ also is an ideal. We claim that B is solvable. Indeed, we have $I \subseteq B$ and so I is a solvable ideal of B; see Lemma 1.3.5(b). Let $\varphi\colon B \to B/I$ be the canonical map. By restriction, we obtain an algebra homomorphism $\varphi'\colon J \to B/I$, $x \mapsto x+I$. By the definition of B, this map is surjective. Hence, since J is solvable, then so is B/I by Lemma 1.3.5(a). But then B itself is solvable by Lemma 1.3.5(c). Hence, since $\dim I$ was maximal, we must have $B = I$ and so $J \subseteq I$. Thus, $I = \text{rad}(A)$ is the unique maximal solvable ideal of A.

Now consider $B := A/\text{rad}(A)$ and the canonical map $\varphi\colon A \to B$. Let $J \subseteq B$ be a solvable ideal. Then $\varphi^{-1}(J)$ is an ideal of A containing $\text{rad}(A)$. Now $\varphi^{-1}(J)/\text{rad}(A) \cong J$ is solvable. Hence, $\varphi^{-1}(J)$ itself is solvable by Lemma 1.3.5(c). So $\varphi^{-1}(J) = \text{rad}(A)$ and $J = \{0\}$. □

Now let L be a Lie algebra with $\dim L < \infty$.

Definition 1.3.7. We say that L is a *semisimple Lie algebra* if $\text{rad}(L) = \{0\}$. By Corollary 1.3.6, L itself or $L/\text{rad}(L)$ is semisimple.

Note that $L = \{0\}$ is considered to be semisimple. Clearly, simple Lie algebras are semisimple. For example, $L = \mathfrak{sl}_2(\mathbb{C})$ is semisimple.

Remark 1.3.8. Since the center $Z(L)$ is an abelian ideal of L, we have $Z(L) \subseteq \text{rad}(L)$. Hence, if L semisimple, then $Z(L) = \{0\}$ and so the homomorphism $\text{ad}_L\colon L \to \mathfrak{gl}(L)$ in Example 1.2.5 is injective. Thus, if L is semisimple and $n = \dim L$, then L is isomorphic to a Lie subalgebra of $\mathfrak{gl}_n(k) \cong \mathfrak{gl}(L)$.

Lemma 1.3.9. *Let $H \subseteq L$ be an ideal. Then $H^{(i)}$ is an ideal of L for all $i \geqslant 0$. In particular, if $H \neq \{0\}$ is solvable, then there exists a non-zero abelian ideal $I \subseteq L$ with $I \subseteq H$.*

Proof. To show that $H^{(i)}$ is an ideal for all i, we use induction on i. If $i = 0$, then $H^{(0)} = H$ is an ideal of L by assumption. Now let $i \geqslant 0$; we have $H^{(i+1)} = [H^{(i)}, H^{(i)}]$. So we must show that $[z,[x,y]] \in [H^{(i)}, H^{(i)}]$ and $[[x,y],z] \in [H^{(i)}, H^{(i)}]$, for all $x, y \in H^{(i)}$, $z \in L$. By anti-symmetry, it is enough to show this for $[z,[x,y]]$. By induction,

$[z, x] \in H^{(i)}$ and $[z, y] \in H^{(i)}$. Using anti-symmetry and the Jacobi identity, $[z, [x, y]] = -[x, [y, z]] - [y, [z, x]] \in [H^{(i)}, H^{(i)}]$, as required.

Now assume that $H = H^{(0)} \neq \{0\}$ is solvable. So there is some $m > 0$ such that $I := H^{(m-1)} \neq \{0\}$ and $I^2 = H^{(m)} = \{0\}$. We have just seen that I is an ideal of L, which is abelian since $I^2 = \{0\}$. $\square$

By Lemma 1.3.9, L is semisimple if and only if L has no non-zero abelian ideal: this is the original definition of semisimplicity given by Killing. This now sets the programme that we will have to pursue:

1) Obtain some idea of how solvable Lie algebras look like.

2) Study in more detail semisimple Lie algebras.

In order to attack 1) and 2), the representation theory of Lie algebras will play a crucial role. This is introduced in the following section.

1.4. Representations of Lie algebras

A fundamental tool in the theory of groups is the study of actions of groups on sets. There is an analogous notion for the action of Lie algebras on vector spaces, taking into account the Lie bracket. Throughout, let L be a Lie algebra over our given field k.

Definition 1.4.1. Let V be a vector space (also over k). Then V is called an *L-module* if there is a bilinear map

$$L \times V \to V, \qquad (x, v) \mapsto x.v$$

such that $[x, y].v = x.(y.v) - y.(x.v)$ for all $x, y \in L$ and $v \in V$. In this case, we obtain for each $x \in L$ a linear map

$$\rho_x \colon V \to V, \qquad v \mapsto x.v,$$

and one immediately checks that $\rho \colon L \to \mathfrak{gl}(V)$, $x \mapsto \rho_x$, is a Lie algebra homomorphism, that is, $\rho_{[x,y]} = [\rho_x, \rho_y] = \rho_x \circ \rho_y - \rho_y \circ \rho_x$ for all $x, y \in L$. This homomorphism ρ will also be called the corresponding *representation* of L on V. If $\dim V < \infty$ and $B = \{v_i \mid i \in I\}$ is a basis of V, then we obtain a *matrix representation*

$$\rho_B \colon L \to \mathfrak{gl}_I(k), \qquad x \mapsto M_B(\rho(x)),$$

where $M_B(\rho(x))$ denotes the matrix of $\rho(x)$ with respect to B. Thus, we have $M_B(\rho(x)) = (a_{ij})_{i,j \in I}$ where $x.v_j = \sum_{i \in I} a_{ij} v_i$ for all j.

If V is an L-module with $\dim V < \infty$, then all the known techniques from Linear Algebra can be applied to the study of the maps $\rho_x : V \to V$: these have a trace, a determinant, eigenvalues and so on.

Remark 1.4.2. Let $\rho : L \to \mathfrak{gl}(V)$ be a Lie algebra homomorphism, where V is a vector space over k; then ρ is called a *representation* of L. One immediately checks that V is an L-module via

$$L \times V \to V, \qquad (x, v) \mapsto \rho(x)(v);$$

furthermore, ρ is the homomorphism associated with this L-module structure on V as in Definition 1.4.1. Thus, speaking about "L-modules" or "representations of L" are just two equivalent ways of expressing the same mathematical fact.

Example 1.4.3. (a) If V is a vector space and L is a Lie subalgebra of $\mathfrak{gl}(V)$, then the inclusion $L \hookrightarrow \mathfrak{gl}(V)$ is a representation. So V is an L-module in a canonical way, where $\rho_x : V \to V$ is given by $v \mapsto x(v)$, that is, we have $\rho_x = x$ for all $x \in L$.

(b) The map $\mathrm{ad}_L : L \to \mathfrak{gl}(L)$ in Example 1.2.5 is a Lie algebra homomorphism, called the *adjoint representation* of L. So L itself is an L-module via this map.

Exercise 1.4.4. Let V be an L-module and $V^* = \mathrm{Hom}(V, k)$ be the dual vector space. Show that V^* is an L-module via $L \times V^* \to V^*$, $(x, \mu) \mapsto \mu_x$, where $\mu_x \in V^*$ is defined by $\mu_x(v) = -\mu(x.v)$ for $v \in V$.

Example 1.4.5. Let V be an L-module and $\rho : L \to \mathfrak{gl}(V)$ be the corresponding representation. Now V is an abelian Lie algebra with Lie bracket $[v, v'] = 0$ for all $v, v' \in V$. Hence, we have $\mathrm{Der}(V) = \mathfrak{gl}(V)$ and we can form the *semidirect product* $L \ltimes_\rho V$, see Exercise 1.2.13. We have $[(x, 0), (0, v)] = (0, x.v)$ for all $x \in L$ and $v \in V$.

Definition 1.4.6. Let V be an L-module; for $x \in L$, we denote by $\rho_x : V \to V$ the linear map defined by x. Let $U \subseteq V$ be a subspace. We say that U is an *L-submodule* (or an *L-invariant subspace*) if $\rho_x(U) \subseteq U$ for all $x \in L$. If $V \neq \{0\}$ and $\{0\}, V$ are the only L-invariant subspaces of V, then V is called an *irreducible module.*

Assume now that U is an L-invariant subspace. Then U itself is an L-module, via the restriction of $L \times V \to V$ to a bilinear map

$L \times U \to U$. Furthermore, V/U is an L-module via

$$L \times V/U \to V/U, \qquad (x, v+U) \mapsto x.v + U.$$

(One checks as usual that this is well-defined and bilinear.) Finally, assume that $n = \dim V < \infty$ and let $d = \dim U$. Let $B = \{v_1, \ldots, v_n\}$ be a basis of V such that $B' = \{v_1, \ldots, v_d\}$ is a basis of U. Since $x.v_i \in U$ for $1 \leqslant i \leqslant d$, the corresponding matrix representation has the following block triangular shape:

$$\rho_B(x) = \left(\begin{array}{c|c} \rho'(x) & * \\ \hline 0 & \rho''(x) \end{array}\right) \qquad \text{for all } x \in L,$$

where $\rho' \colon L \to \mathfrak{gl}_d(k)$ is the matrix representation corresponding to U (with respect to the basis B' of U) and $\rho'' \colon L \to \mathfrak{gl}_{n-d}(k)$ is the matrix representation corresponding to V/U (with respect to the basis $B'' = \{v_{d+1} + U, \ldots, v_n + U\}$ of V/U).

Corollary 1.4.7. *Let $V \neq \{0\}$ be an L-module with $\dim V < \infty$. There is a sequence of L-submodules $\{0\} = V_0 \subsetneqq V_1 \subsetneqq V_2 \subsetneqq \ldots \subsetneqq V_r = V$ such that V_i/V_{i-1} is irreducible for $1 \leqslant i \leqslant r$. Let $n_i = \dim(V_i/V_{i-1})$ for all i. Then there is a basis B of V such that the matrices of the representation $\rho \colon L \to \mathfrak{gl}(V)$ have the following shape*

$$\rho_B(x) = \begin{pmatrix} \rho_1(x) & * & \ldots & * \\ 0 & \rho_2(x) & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \ldots & 0 & \rho_r(x) \end{pmatrix} \qquad \textit{for all } x \in L,$$

where $\rho_i \colon L \to \mathfrak{gl}_{n_i}(k)$ is an irreducible matrix representation corresponding to the L-module V_i/V_{i-1}.

Proof. Let $U \subsetneqq V$ be an L-submodule with $\dim U$ as large as possible. If $W \subseteq V/U$ is a submodule, then one easily checks that $\{v \in V \mid v + U \in W\} \subseteq V$ is a submodule containing U, so $W = \{0\}$ or $W = V/U$. Hence, V/U is irreducible and we continue with U. $\square$

Example 1.4.8. If V is an L-module with $\dim V = 1$, then V is obviously irreducible. Let $V = \langle v \rangle_k$ where $0 \neq v \in V$. Then, for all $x \in L$, we have $x.v = \varphi(x)v$ where $\varphi(x) \in k$. It follows that $\varphi \colon L \to k$ is linear. Furthermore, $\varphi([x,y])v = [x,y].v = x.(y.v) - y.(x.v) = \varphi(y)x.v - \varphi(x)y.v = 0$ and so $\varphi([x,y]) = 0$ for all $x, y \in L$.

Exercise 1.4.9. Let k be a field of characteristic 2 and L be the Lie algebra over k with basis $\{x, y\}$ such that $[x, y] = y$ (see Exercise 1.2.3). Show that the linear map defined by

$$\rho\colon L \to \mathfrak{gl}_2(k), \quad x \mapsto \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad y \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

is a Lie algebra homomorphism and so $V = k^2$ is an L-module. Show that V is an irreducible L-module. Check that L is solvable.

There is a version for modules of the *generalised binomial formula*:

Lemma 1.4.10. *Let V be an L-module. Let $v \in V$, $x, y \in L$ and $c \in k$. Then, for all $n \geqslant 0$, we have*

$$(\rho_x - c\,\mathrm{id}_V)^n(y.v) = \sum_{i=0}^{n} \binom{n}{i} \underbrace{\mathrm{ad}_L(x)^i(y)}_{\in L} . \underbrace{\big((\rho_x - c\,\mathrm{id}_V)^{n-i}(v)\big)}_{\in V}.$$

Proof. Consider the associative algebra $A := \mathrm{End}(V)$. Then $\rho_x, \rho_y \in A$ and $y.v = \rho_y(v)$. So Lemma 1.1.4 (with $a := -c$ and $b := 0$) implies that the left hand side of the desired identity equals

$$\big((\rho_x - c\,\mathrm{id}_V)^n \circ \rho_y\big)(v) = \sum_{i=0}^{n} \binom{n}{i} \psi_i\big((\rho_x - c\,\mathrm{id}_V)^{n-1}(v)\big).$$

where $\psi_i := \mathrm{ad}_A(\rho_x)^i(\rho_y) \in A$ for $i \geqslant 0$. Now note that

$$\rho\big(\mathrm{ad}_L(x)(y)\big) = \rho([x, y]) = \rho_{[x,y]} = [\rho_x, \rho_y] = \mathrm{ad}_A(\rho_x)(\rho_y).$$

A simple induction on i shows that $\rho\big(\mathrm{ad}_L(x)^i(y)\big) = \mathrm{ad}_A(\rho_x)^i(\rho_y)$ for all $i \geqslant 0$. Thus, we have $\psi_i = \rho\big(\mathrm{ad}_L(x)^i(y)\big)$, as desired. □

Up to this point, k could be any field (of any characteristic). Stronger results will hold if k is algebraically closed.

Lemma 1.4.11 (Schur's Lemma). *Assume that k is algebraically closed. Let V be an irreducible L-module,* $\dim V < \infty$. *If $\varphi \in \mathrm{End}(V)$ is such that $\varphi \circ \rho_x = \rho_x \circ \varphi$ for all $x \in L$, then $\varphi = c\,\mathrm{id}_V$ where $c \in k$.*

Proof. We check that $\ker(\varphi)$ is an L-submodule of V. Indeed, let $v \in \ker(\varphi)$ and $x \in L$. Then $\varphi(x.v) = \varphi(\rho_x(v)) = \rho_x(\varphi(v)) = 0$ and so $x.v \in \ker(\varphi)$. Since V is irreducible, $\varphi = 0$ or $\ker(\varphi) = \{0\}$. If $\varphi = 0$, then the desired assertion holds with $c = 0$. Now assume that $\varphi \neq 0$.

Then $\ker(\varphi) = \{0\}$ and φ is bijective. Since k is algebraically closed, there is an eigenvalue $c \in k$ for φ. Setting $\psi := \varphi - c\,\mathrm{id}_V \in \mathrm{End}(V)$, we also have $\psi(x.v) = x.(\psi(v))$ for all $x \in L$ and $v \in V$. Hence, the previous argument shows that either $\psi = 0$ or ψ is bijective. But an eigenvector for c lies in $\ker(\psi)$ and so $\psi = 0$. $\square$

Proposition 1.4.12. *Assume that k is algebraically closed and L is abelian. Let $V \neq \{0\}$ be an L-module with $\dim V < \infty$. Then there exists a basis B of V such that, for any $x \in L$, the matrix of the linear map $\rho_x \colon V \to V$, $v \mapsto x.v$, with respect to B has the following shape:*

$$M_B(\rho_x) = \begin{pmatrix} \lambda_1(x) & * & \dots & * \\ 0 & \lambda_2(x) & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & \lambda_n(x) \end{pmatrix} \qquad (n = \dim V),$$

where $\lambda_i \colon L \to k$ are linear maps for $1 \leqslant i \leqslant n$. In particular, if V is irreducible, then $\dim V = 1$.

Proof. Assume first that V is irreducible. We show that $\dim V = 1$. Let $x \in L$ be fixed and $\varphi := \rho_x$. Since L is abelian, we have $0 = \rho_0 = \rho_{[x,y]} = \varphi \circ \rho_y - \rho_y \circ \varphi$ for all $y \in L$. By Schur's Lemma, $\varphi = \lambda(x)\,\mathrm{id}_V$ where $\lambda(x) \in k$. Hence, if $0 \neq v \in V$, then $x.v = \lambda(x)v$ for all $x \in L$ and so $\langle v \rangle_k \subseteq V$ is an L-submodule. Clearly, $\lambda \colon L \to k$ is linear. Since V is irreducible, $V = \langle v \rangle_k$ and so $\dim V = 1$. The general case follows from Corollary 1.4.7. $\square$

Example 1.4.13. Assume that k is algebraically closed. Let V be a vector space over k with $\dim V < \infty$. Let $\mathfrak{X} \subseteq \mathrm{End}(V)$ be a subset such that $\varphi \circ \psi = \psi \circ \varphi$ for all $\varphi, \psi \in \mathfrak{X}$. Then there exists a basis B of V such that the matrix of any $\varphi \in \mathfrak{X}$ with respect to B is upper triangular. Indeed, just note that $L := \langle \mathfrak{X} \rangle_k \subseteq \mathfrak{gl}(V)$ is an abelian Lie subalgebra and V is an L-module; then apply Proposition 1.4.12. (Of course, one could also prove this more directly.)

Exercise 1.4.14. This exercise establishes an elementary result from Linear Algebra that will be useful at several places. Let k be an infinite field and V be a k-vector space with $\dim V < \infty$. Let $V^* := \mathrm{Hom}(V, k)$ be the dual space.

(a) Show that, if $X \subseteq V$ is a finite subset such that $0 \notin X$, then there exists $\mu_0 \in V^*$ such that $\mu_0(x) \neq 0$ for all $x \in X$.

(b) Similarly, if $\Lambda \subseteq V^*$ is a finite subset such that $\underline{0} \notin \Lambda$ (where $\underline{0}\colon V \to k$ denotes the linear map with value 0 for all $v \in V$), then there exists $v_0 \in V$ such that $f(v_0) \neq 0$ for all $f \in \Lambda$.

1.5. Lie's Theorem

The content of Lie's Theorem is that Proposition 1.4.12 (which was concerned with representations of abelian Lie algebras) remains true for the more general class of solvable Lie algebras, assuming that k is not only algebraically closed but also has characteristic 0. (Exercice 1.4.9 shows that this will definitely not work in positive characteristic.) So, in order to use the full power of the techniques developed so far, we will assume that $k = \mathbb{C}$.

Let L be a Lie algebra over $k = \mathbb{C}$. If V is an L-module, then we denote as usual by $\rho_x\colon V \to V$ the linear map defined by $x \in L$. Our approach to Lie's Theorem is based on the following technical result.

Lemma 1.5.1. *Let V be an irreducible L-module (over $k = \mathbb{C}$), with $\dim V < \infty$. Let $H \subseteq L$ be an abelian ideal in L such that* $\mathrm{Trace}(\rho_x) = 0$ *for all $x \in H$. Then $\rho_x = 0$ for all $x \in H$.*

Proof. Let $x \in H$ and consider the linear map $\rho_x\colon V \to V$. Since $k = \mathbb{C}$, this map has eigenvalues. Let $c \in \mathbb{C}$ be an eigenvalue and consider the generalised eigenspace

$$V_c(\rho_x) := \{v \in V \mid (\rho_x - c\,\mathrm{id}_V)^l(v) = 0 \text{ for some } l \geqslant 1\} \neq \{0\}.$$

We claim that $V_c(\rho_x) \subseteq V$ is an L-submodule. To see this, let $v \in V_c(\rho_x)$ and $y \in L$. We must show that $y.v \in V_c(\rho_x)$. Let $l \geqslant 1$ be such that $(\rho_x - c\,\mathrm{id}_V)^l(v) = 0$. Using Lemma 1.4.10, we obtain

$$(\rho_x - c\,\mathrm{id}_V)^{l+1}(y.v) = \sum_{i=0}^{l+1} \binom{l+1}{i} \mathrm{ad}_L(x)^i(y).(\rho_x - c\,\mathrm{id}_V)^{l+1-i}(v).$$

If $i = 0, 1$, then $l + 1 - i \geqslant l$ and so $(\rho_x - c\,\mathrm{id}_V)^{l+1-i}(v) = 0$. Now let $i \geqslant 2$. Then $\mathrm{ad}_L(x)^i(y) = \mathrm{ad}_L(x)^{i-2}([x, [x, y]])$. But $[x, y] \in H$ because H is an ideal, and $[x, [x, y]] = 0$ because H is abelian. So $\mathrm{ad}_L(x)^i(y) = 0$. We conclude that $y.v \in V_c(\rho_x)$, as desired.

Now, since V is irreducible and $V_c(\rho_x) \neq \{0\}$, we conclude that $V = V_c(\rho_x)$. Let $\psi_x := \rho_x - c\,\mathrm{id}_V$. Then, for $v \in V$, there exists some $l \geqslant 1$ with $\psi_x^l(v) = 0$. So Exercise 1.2.4 shows that ψ_x is nilpotent and $\mathrm{Trace}(\psi_x) = 0$. But then $\mathrm{Trace}(\rho_x) = \mathrm{Trace}(\psi_x + c\,\mathrm{id}_V) = (\dim V)c$. So our assumption on $\mathrm{Trace}(\rho_x)$ implies that $c = 0$. Thus, we have seen that 0 is the only eigenvalue of ρ_x, for any $x \in H$.

Finally, regarding V as an H-module (by restricting the action of L on V to H), we can apply Proposition 1.4.12. This yields a basis B of V such that, for any $x \in H$, the matrix of ρ_x with respect to B is upper triangular; by the above discussion, the entries along the diagonal are all 0. Let v_1 be the first vector in B. Then $x.v_1 = \rho_x(v_1) = 0$ for all $x \in H$. Hence, the subspace

$$U := \{v \in V \mid x.v = 0 \text{ for all } x \in H\}$$

is non-zero. Now we claim that U is an L-submodule. Let $v \in V$ and $y \in L$. Then, for $x \in H$, we have $x.(y.v) = [x,y].v + y.(x.v) = [x,y].v = 0$ since $v \in U$ and $[x,y] \in H$. Since V is irreducible, we conclude that $U = V$ and so $\rho_x = 0$ for all $x \in H$. □

Proposition 1.5.2 (Semisimplicity criterion). *Let $k = \mathbb{C}$ and V be a vector space with $\dim V < \infty$. Let $L \subseteq \mathfrak{sl}(V)$ be a Lie subalgebra such that V is an irreducible L-module. Then L is semisimple.*

Proof. If $\mathrm{rad}(L) \neq \{0\}$ then, by Lemma 1.3.9, there exists a non-zero abelian ideal $H \subseteq L$ such that $H \subseteq \mathrm{rad}(L)$. Since $L \subseteq \mathfrak{sl}(V)$, Lemma 1.5.1 implies that $x = \rho_x = 0$ for all $x \in H$, contradiction. □

Example 1.5.3. Let $k = \mathbb{C}$ and V be a vector space with $\dim V < \infty$. Clearly (!), V is an irreducible $\mathfrak{gl}(V)$-module. Next note that $\mathfrak{gl}(V) = \mathfrak{sl}(V) + \mathbb{C}\,\mathrm{id}_V$. Hence, if $U \subseteq V$ is an $\mathfrak{sl}(V)$-invariant subspace, then U will also be $\mathfrak{gl}(V)$-invariant. Consequently, V is an irreducible $\mathfrak{sl}(V)$-module. Hence, Proposition 1.5.2 shows that $\mathfrak{sl}(V)$ is semisimple.

Note that, if $\mathrm{char}(k) = p > 0$ and $L = \mathfrak{sl}_p(k)$, then $Z := \{aI_p \mid a \in k\}$ is an abelian ideal in L and so L is not semisimple in this case.

Theorem 1.5.4 (Lie's Theorem). *Let $k = \mathbb{C}$. Let L be solvable and $V \neq \{0\}$ be an L-module with $\dim L < \infty$ and $\dim V < \infty$. Then the conclusions in Proposition 1.4.12 still hold, that is, there exists a* *basis B of V such that, for any $x \in L$, the matrix of the linear map $\rho_x \colon V \to V$, $v \mapsto x.v$, with respect to B has the following shape:*

$$M_B(\rho_x) = \begin{pmatrix} \lambda_1(x) & * & \dots & * \\ 0 & \lambda_2(x) & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & \lambda_n(x) \end{pmatrix} \qquad (n = \dim V),$$

where $\lambda_i \colon L \to k$ are linear maps such that $[L, L] \subseteq \ker(\lambda_i)$ for $1 \leqslant i \leqslant n$. In particular, if V is irreducible, then $\dim V = 1$.

Proof. First we show that, if V is irreducible, then $\dim V = 1$. We use induction on $\dim L$. If $\dim L = 0$, there is nothing to prove. Now assume that $\dim L > 0$. If L is abelian, then see Proposition 1.4.12. Now assume that $[L, L] \neq \{0\}$. By Lemma 1.3.9, there exists a non-zero abelian ideal $H \subseteq L$ such that $H \subseteq [L, L]$. Let $x \in H$. Since $H \subseteq [L, L]$, we can write x as a finite sum $x = \sum_i [y_i, z_i]$ where $y_i, z_i \in L$ for all i. Consequently, we also have $\rho_x = \sum_i (\rho_{y_i} \circ \rho_{z_i} - \rho_{z_i} \circ \rho_{y_i})$ and, hence, $\mathrm{Trace}(\rho_x) = 0$. By Lemma 1.5.1, $\rho_x = 0$ for all $x \in H$. Let $L_1 := L/H$. Then V also is an L_1-module via

$$L_1 \times V \to V, \qquad (y + H, v) \mapsto y.v.$$

(This is well-defined since $x.v = 0$ for $x \in H$, $v \in V$.) If $V' \subseteq V$ is an L_1-invariant subspace, then V' is also L-invariant. Hence, V is an irreducible L_1-module. By Lemma 1.3.5(c), L_1 is solvable. So, by induction, $\dim V = 1$.

The general case follows again from Corollary 1.4.7. The fact that $[L, L] \subseteq \ker(\lambda_i)$ for all i is seen as in Example 1.4.8. □

Lemma 1.5.5. *In the setting of Theorem 1.5.4, the set of linear maps $\{\lambda_1, \dots, \lambda_n\}$ does not depend on the choice of the basis B of V.*
We shall call $P(V) := \{\lambda_1, \dots, \lambda_n\}$ the set of *weights* of L on V.

Proof. Let B' be another basis of V such that, for any $\in L$, the matrix of $\rho_x \colon V \to V$ with respect to B' has a triangular shape with $\lambda'_1(x)$, ..., $\lambda'_n(x)$ along the diagonal, where $\lambda'_i \colon L \to k$ are linear maps such that $[L, L] \subseteq \ker(\lambda'_i)$ for $1 \leqslant i \leqslant n$. We must show that $\{\lambda_1, \dots, \lambda_n\} = \{\lambda'_1, \dots, \lambda'_n\}$. Assume, if possible, that there exists some j such that $\lambda'_j \neq \lambda_i$ for $1 \leqslant i \leqslant n$. Let $\Lambda := \{\lambda_i - \lambda'_j \mid 1 \leqslant i \leqslant n\}$.

Then Λ is a finite subset of $\mathrm{Hom}(L, \mathbb{C})$ such that $\underline{0} \notin \Lambda$. So, by Exercise 1.4.14(b), there exists some $x_0 \in L$ such that $\lambda'_j(x_0) \neq \lambda_i(x_0)$ for $1 \leqslant i \leqslant n$. But then $\lambda'_j(x_0)$ is an eigenvalue of $M_{B'}(\rho_{x_0})$ that is not an eigenvalue of $M_B(\rho_{x_0})$, contractiction since $M_B(\rho_{x_0})$ and $M_{B'}(\rho_{x_0})$ are similar matrices and, hence, they have the same characteristic polynomials. Thus, we have shown that $\{\lambda'_1, \ldots, \lambda'_n\} \subseteq \{\lambda_1, \ldots, \lambda\}$. The reverse inclusion is proved analogously. □

Exercise 1.5.6. Let $k = \mathbb{C}$ and L be solvable with $\dim L < \infty$. Let V be a finite-dimensional L-module and $U \subseteq V$ be a non-zero, proper L-submodule. Show that $P(V) = P(U) \cup P(V/U)$ (where the set of weights of a module is defined by Lemma 1.5.5).

Exercise 1.5.7. Assume that $k \subseteq \mathbb{C}$. Show that

$$L = \left\{ \begin{pmatrix} 0 & t & x \\ -t & 0 & y \\ 0 & 0 & 0 \end{pmatrix} \;\middle|\; t, x, y \in k \right\}$$

is a solvable Lie subalgebra of $\mathfrak{gl}_3(k)$. Regard $V = k^3$ as an L-module via the inclusion $L \hookrightarrow \mathfrak{gl}_3(k)$ (cf. Example 1.4.3). If $k = \mathbb{C}$, find a basis B of V such that the corresponding matrices of L will be upper triangular. Does this also work with $k = \mathbb{R}$?

ab hier Woche 3

Finally, we develop some very basic aspects of the representation theory of $\mathfrak{sl}_2(\mathbb{C})$. As pointed out in [**25**, §2.4], this is of the utmost importance for the general theory of semisimple Lie algebras. So, for the remainder of this section, let $L = \mathfrak{sl}_2(\mathbb{C})$, with standard basis

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

where $[e, f] = h$, $[h, e] = 2e$, $[h, f] = -2f$ (see Exercise 1.2.11). The following result is obtained by an easy application of Lie's Theorem.

Lemma 1.5.8. *Let V be an $\mathfrak{sl}_2(\mathbb{C})$-module with $\dim V < \infty$. Then there exists a non-zero vector $v^+ \in V$ such that $e.v^+ = 0$ and $h.v^+ = cv^+$ for some $c \in \mathbb{C}$.*

Proof. Let $S := \langle h, e \rangle_{\mathbb{C}} \subseteq \mathfrak{sl}_2(\mathbb{C})$. This is precisely the subalgebra of $\mathfrak{sl}_2(\mathbb{C})$ consisting of all upper triangular matrices with trace 0. Since $[h, e] = 2e$, we have $[S, S] = \langle e \rangle_{\mathbb{C}}$ and so S is solvable. By restricting the action of $\mathfrak{sl}_2(\mathbb{C})$ on V to S, we can regard V as S-module. So, by

Theorem 1.5.4, there exist a basis B of V and $\lambda_1, \ldots, \lambda_n \in S^*$ (where $n = \dim V$) such that, for any $x \in S$, the matrix of $\rho_x \colon V \to V$ is upper triangular with $\lambda_1(x), \ldots, \lambda_n(x)$ along the diagonal; furthermore, $[S, S] \subseteq \ker(\lambda_i)$ for $1 \leqslant i \leqslant n$. Let v^+ be the first vector in B. Then $\rho_x(v^+) = \lambda_1(x)v^+$ for all $x \in S$. So v^+ has the required properties, where $c := \lambda_1(h) \in \mathbb{C}$; we have $e.v^+ = 0$ since $e \in [S, S]$. □

Remark 1.5.9. Let $V \neq \{0\}$ be an $\mathfrak{sl}_2(\mathbb{C})$-module with $\dim V < \infty$. Let $v^+ \in V$ be as in Lemma 1.5.8; any such vector will be called a *primitive vector* of V. Then we define a sequence $(v_n)_{n \geqslant 0}$ in V by

$$v_0 := v^+ \qquad \text{and} \qquad v_{n+1} := \tfrac{1}{n+1} f.v_n \quad \text{for all } n \geqslant 0.$$

Let $V' := \langle v_n \mid n \geqslant 0 \rangle_{\mathbb{C}} \subseteq V$. We claim that the following relations hold for all $n \geqslant 0$ (where we also set $v_{-1} := 0$):

(a) $$h.v_n = (c - 2n)v_n \qquad \text{and} \qquad e.v_n = (c - n + 1)v_{n-1}.$$

We use induction on n. If $n = 0$, the formulae hold by definition. Now let $n \geqslant 0$. First note that $f.v_{n-1} = nv_n$. We compute:

$$\begin{aligned}(n+1)e.v_{n+1} &= e.(f.v_n) = [e, f].v_n + f.(e.v_n) = h.v_n + f.(e.v_n)\\ &= (c - 2n)v_n + (c - n + 1)f.v_{n-1} \qquad \text{(by induction)}\\ &= (c - 2n)v_n + (c - n + 1)nv_n = ((n + 1)c - n^2 - n)v_n,\end{aligned}$$

and so $e.v_{n+1} = (c - n)v_n$, as required. Next, we compute:

$$\begin{aligned}(n + 1)h.v_{n+1} &= h.(f.v_n) = [h, f].v_n + f.(h.v_n)\\ &= -2f.v_n + (c - 2n)f.v_n = (c - 2n - 2)(n + 1)v_{n+1},\end{aligned}$$

so (a) holds. Now, if $v_n \neq 0$ for all n, then $v_0, v_1, v_2, \ldots$ are eigenvectors for $\rho_h \colon V \to V$ with distinct eigenvalues (see (a)) and so $v_0, v_1, v_2, \ldots$ are linearly independent, contradiction to $\dim V < \infty$. So there is some $n_0 \geqslant 0$ such that $v_0, v_1, \ldots, v_{n_0}$ are linearly independent and $v_{n_0+1} = 0$. But then, by the definition of the v_n, we have $v_n = 0$ for all $n > n_0$ and so $V' = \langle v_0, v_1, \ldots, v_{n_0} \rangle_{\mathbb{C}}$. Furthermore, $0 = e.0 = e.v_{n_0+1} = (c - n_0)v_{n_0}$ and so $c = n_0$. Thus, we obtain:

(b) $$h.v^+ = cv^+ \qquad \text{where} \qquad c = \dim V' - 1 \in \mathbb{Z}_{\geqslant 0}.$$

So, the eigenvalue of our primitive vector v^+ has a very special form!

If $c \geqslant 1$, then the above formulae also yield an expression of $v^+ = v_0$ in terms of $v_c = v_{n_0}$; indeed, by (a), we have $[e, v_c] = v_{c-1}$, $[e, v_{c-1}] = 2v_{c-2}$, $[e, v_{c-2}] = 3v_{c-3}$ and so on. Thus, we obtain:

(c) $$\underbrace{[e, [e, [\ldots, [e,}_{c \text{ times}} v_c] \ldots]]] = (1{\cdot}2{\cdot}3{\cdot}\ldots{\cdot}c)\, v^+.$$

We now state some useful consequences of the above discussion.

Corollary 1.5.10. *In the setting of Remark 1.5.9, assume that V is irreducible. Write $\dim V = m + 1$, $m \geqslant 0$. Then ρ_h is diagonalisable with eigenvalues $\{m - 2i \mid 0 \leqslant i \leqslant m\}$ (each with multiplicity 1).*

Proof. Using the formulae in Remark 1.5.9 and an induction on n, one sees that $h.v_n \in V'$, $e.v_n \in V'$, $f.v_n \in V'$ for all $n \geqslant 0$. Thus, $V' \subseteq V$ is an $\mathfrak{sl}_2(\mathbb{C})$-submodule. Since $V' \neq \{0\}$ and V is irreducible, we conclude that $V' = V$ and $m = c$. By Remark 1.5.9(a), we have $h.v_n = (c - 2n)v_n$ for all $n \geqslant 0$. Hence, ρ_h is diagonalisable, with eigenvalues as stated above. □

Proposition 1.5.11. *Let V be any finite-dimensional $\mathfrak{sl}_2(\mathbb{C})$-module, with e, h, f as above. Then all the eigenvalues of $\rho_h \colon V \to V$ are integers and we have $\mathrm{Trace}(\rho_h) = 0$. Furthermore, if $n \in \mathbb{Z}$ is an eigenvalue of ρ_h, then so is $-n$ (with the same multiplicity as n).*

Proof. Note that the desired statements can be read off the characteristic polynomial of $\rho_h \colon V \to V$. If V is irreducible, then these hold by Corollary 1.5.10. In general, let $\{0\} = V_0 \subsetneqq V_1 \subsetneqq V_2 \subsetneqq \ldots \subsetneqq V_r = V$ be a sequence of L-submodules as in Corollary 1.4.7, such that V_i/V_{i-1} is irreducible for $1 \leqslant i \leqslant r$. It remains to note that the characteristic polynomial of $\rho_h \colon V \to V$ is the product of the characteristic polynomials of the actions of h on V_i/V_{i-1} for $1 \leqslant i \leqslant r$. □

1.6. The classical Lie algebras

Let V be a vector space over k and $\beta \colon V \times V \to k$ be a bilinear map. Then we define $\mathfrak{go}(V, \beta)$ to be the set of all $\varphi \in \mathrm{End}(V)$ such that

$$\beta(\varphi(v), w) + \beta(v, \varphi(w)) = 0 \qquad \text{for all } v, w \in V.$$

(The symbol "$\mathfrak{go}$" stands for "general orthogonal".) One checks that $\mathfrak{go}(V, \beta)$ is a Lie subalgebra of $\mathfrak{gl}(V)$ (see exercises), called a *classical*

Lie algebra. The further developments will show that these form an important class of semisimple Lie algebras (for certain β, over $k = \mathbb{C}$).

We assume throughout that β either is a *symmetric bilinear form* or an *alternating bilinear form*. This means that there is a sign $\epsilon = \pm 1$ such that $\beta(v, w) = \epsilon\beta(w, v)$ for all $v, w \in V$. (If $\epsilon = +1$, then β is symmetric; if $\epsilon = -1$, then β is alternating.) We shall also assume throughout that $\mathrm{char}(k) \neq 2$. (This avoids the consideration of some special cases that are not relevant to us here.)

For any subset $X \subseteq V$, we can define

$$X^\perp := \{v \in V \mid \beta(v, x) = 0 \text{ for all } x \in X\},$$

where it does not matter if we write "$\beta(v, x) = 0$" or "$\beta(x, v) = 0$". Note that $X^\perp$ is a subspace of V (even if X is not a subspace). We say that β is a *non-degenerate bilinear form* if $V^\perp = \{0\}$.

As in Example 1.4.3(a), the vector space V is a $\mathfrak{go}(V, \beta)$-module in a natural way. Again, this module turns out to be irreducible.

Proposition 1.6.1. *Assume that* $3 \leqslant \dim V < \infty$ *and* β *is non-degenerate. Then* V *is an irreducible* $\mathfrak{go}(V, \beta)$*-module.*

Proof. First we describe a method for producing elements in $\mathfrak{go}(V, \beta)$. For fixed $x, y \in V$ we define a linear map $\varphi_{x,y} \colon V \to V$ by $\varphi_{x,y}(v) := \beta(v, x)y - \beta(y, v)x$ for all $v \in V$. We claim that $\varphi_{x,y} \in \mathfrak{go}(V, \beta)$. Indeed, for all $v, w \in V$, we have

$$\begin{aligned} \beta(\varphi_{x,y}(v), w) &+ \beta(v, \varphi_{x,y}(w)) \\ &= \big(\beta(v, x)\beta(y, w) - \beta(y, v)\beta(x, w)\big) \\ &\qquad + \big(\beta(w, x)\beta(v, y) - \beta(y, w)\beta(v, x)\big) \\ &= -\beta(y, v)\beta(x, w) + \beta(w, x)\beta(v, y), \end{aligned}$$

which is 0 since $\beta(v, y) = \epsilon\beta(y, v)$ and $\beta(w, x) = \epsilon\beta(w, x)$.

Now let $W \subseteq V$ be a $\mathfrak{go}(V, \beta)$-submodule and assume, if possible, that $\{0\} \neq W \neq V$. Let $0 \neq w \in W$. Since β is non-degenerate, we have $\beta(y, w) \neq 0$ for some $y \in V$. If $x \in V$ is such that $\beta(x, w) = 0$, then $\varphi_{x,y}(w) = \beta(w, x)y - \beta(y, w)x = -\beta(y, w)x$. But then $\varphi_{x,y}(w) \in W$ (since W is a submodule) and so $x \in W$. Thus,

$$U_w := \{x \in V \mid \beta(x, w) = 0\} \subseteq W.$$

Since U_w is defined by a single, non-trivial linear equation, we have $\dim U_w = \dim V - 1$ and so $\dim W \geqslant \dim V - 1$. Since $W \neq V$, we have $\dim W = \dim U_w$ and $U_w = W$. This holds for all $0 \neq w \in W$ and so $W \subseteq W^\perp$. Since β is non-degenerate, we have $\dim V = \dim W + \dim W^\perp$ (by a general result in Linear Algebra); hence,

$$\dim V = \dim W + \dim W^\perp \geqslant 2\dim W \geqslant 2(\dim V - 1)$$

and so $\dim V \leqslant 2$, a contradiction. □

In the sequel, it will be convenient to work with matrix descriptions of $\mathfrak{go}(V,\beta)$; these are provided by the following exercise.

Exercise 1.6.2. Let $n = \dim V < \infty$ and $B = \{v_1, \ldots, v_n\}$ be a basis of V. We form the corresponding Gram matrix

$$Q = \big(\beta(v_i, v_j)\big)_{1 \leqslant i,j \leqslant n} \in M_n(k).$$

The following equivalences are well-known from Linear Algebra:

$$\begin{aligned} Q^{\mathrm{tr}} = Q \quad &\Leftrightarrow \quad \beta \text{ symmetric},\\ Q^{\mathrm{tr}} = -Q \quad &\Leftrightarrow \quad \beta \text{ alternating},\\ \det(Q) \neq 0 \quad &\Leftrightarrow \quad \beta \text{ non-degenerate}. \end{aligned}$$

Recall that we are assuming $\mathrm{char}(k) \neq 2$.

(a) Let $\varphi \in \mathrm{End}(V)$ and $A = (a_{ij}) \in M_n(k)$ be the matrix of φ with respect to B. Then show that $\varphi \in \mathfrak{go}(V,\beta) \Leftrightarrow A^{\mathrm{tr}}Q + QA = 0$, where A^{tr} denotes the transpose matrix. Hence, we obtain a Lie subalgebra

$$\mathfrak{go}_n(Q,k) := \{A \in M_n(k) \mid A^{\mathrm{tr}}Q + QA = 0\} \subseteq \mathfrak{gl}_n(k).$$

Deduce that $V = k^n$ is an irreducible $\mathfrak{go}_n(Q,k)$-module if $Q^{\mathrm{tr}} = \pm Q$, $\det(Q) \neq 0$ and $n \geqslant 3$.

(b) Show that if $\det(Q) \neq 0$, then $\mathfrak{go}_n(Q,k) \subseteq \mathfrak{sl}_n(k)$. (In particular, for $n = 1$, we have $\mathfrak{go}_1(Q,k) = \{0\}$ in this case.)

Proposition 1.6.3. *Let $n \geqslant 3$ and $k = \mathbb{C}$. If $Q^{\mathrm{tr}} = \pm Q$ and $\det(Q) \neq 0$, then $\mathfrak{go}_n(Q,\mathbb{C})$ is semisimple.*

Proof. This follows from Exercise 1.6.2 and the semisimplicity criterion in Proposition 1.5.2. □

Depending on what Q looks like, computations in $\mathfrak{go}_n(Q,k)$ can be more, or less complicated. Let us assume from now on that $k=\mathbb{C}$, $n=\dim V<\infty$ and Q is given by[1]

$$Q=Q_n:=\begin{pmatrix} 0 & \cdots & 0 & \delta_n \\ \vdots & \iddots & \iddots & 0 \\ 0 & \delta_2 & \iddots & \vdots \\ \delta_1 & 0 & \cdots & 0 \end{pmatrix}\in M_n(\mathbb{C}) \qquad (\delta_i=\pm 1),$$

where $\delta_i\delta_{n+1-i}=\epsilon$ for all i and, hence, $Q_n=\epsilon Q_n^{\text{tr}}$, $\det(Q_n)\neq 0$.

Exercise 1.6.4. (a) Assume that $n=2$. Determine $\mathfrak{go}_2(Q_2,\mathbb{C})$.

(b) Assume that $n=3$ and $Q_3=Q_3^{\text{tr}}$. Show that

$$\mathfrak{go}_3(Q_3,\mathbb{C})=\left\{\begin{pmatrix} a & b & 0 \\ c & 0 & -\delta b \\ 0 & -\delta c & -a \end{pmatrix} \;\middle|\; a,b,c\in\mathbb{C}\right\} \qquad (\delta:=\delta_1\delta_2)$$

is isomorphic to $\mathfrak{sl}_2(\mathbb{C})$.

(c) Assume that $n=4$ and $Q_4=Q_4^{\text{tr}}$. Show that

$$L_1:=\left\{\begin{pmatrix} a & 0 & b & 0 \\ 0 & a & 0 & -b \\ c & 0 & -a & 0 \\ 0 & -c & 0 & -a \end{pmatrix} \;\middle|\; a,b,c\in\mathbb{C}\right\}\subseteq \mathfrak{go}_4(Q_4,\mathbb{C})$$

is an ideal and $L_1\cong\mathfrak{sl}_2(\mathbb{C})$. Show that $\mathfrak{go}_4(Q_4,\mathbb{C})\cong\mathfrak{sl}_2(\mathbb{C})\times\mathfrak{sl}_2(\mathbb{C})$ (where the direct product of two algebras is defined in Example 1.1.2).

Example 1.6.5. We have the following implication:

$$A\in\mathfrak{go}_n(Q_n,\mathbb{C}) \quad\Rightarrow\quad A^{\text{tr}}\in\mathfrak{go}_n(Q_n,\mathbb{C}).$$

Indeed, if $A^{\text{tr}}Q_n+Q_nA=0$, then $Q_n^{-1}A^{\text{tr}}+AQ_n^{-1}=0$. Now note that $Q_n^{-1}=Q_n^{\text{tr}}=\epsilon Q_n$. Hence, we also have $Q_nA^{\text{tr}}+AQ_n=0$.

Finally, we determine a vector space basis of $\mathfrak{go}_n(Q_n,\mathbb{C})$. We set

$$A_{ij}:=\delta_iE_{ij}-\delta_jE_{n+1-j,n+1-i}\in M_n(\mathbb{C})$$

for $1\leqslant i,j\leqslant n$, where E_{ij} denotes the elementary matrix with 1 at position (i,j), and 0 otherwise. With this notation, we have:

[1]If $k=\mathbb{C}$ and β is non-degenerate, then one can always find a basis B of V such that Q has this form. For β alternating, this holds even over any field k; see [**16**, Theorem 2.10]. For β symmetric, this follows from the fact that, over $\mathbb{C}$, any two non-degenerate symmetric bilinear forms are equivalent; see [**16**, Theorem 4.4].

Proposition 1.6.6. *Recall that $k = \mathbb{C}$ and $Q = Q_n$ is as above.*

(a) *If $Q_n^{\text{tr}} = Q_n$, then $\{A_{ij} \mid 1 \leqslant i,j \leqslant n, i+j \leqslant n\}$ is a basis of $\mathfrak{go}_n(Q_n, \mathbb{C})$ and so $\dim \mathfrak{go}_n(Q_n, \mathbb{C}) = n(n-1)/2$.*

(b) *If $Q_n^{\text{tr}} = -Q_n$, then $\{A_{ij} \mid 1 \leqslant i,j \leqslant n, i+j \leqslant n+1\}$ is a basis of $\mathfrak{go}_n(Q_n, \mathbb{C})$ and so $\dim \mathfrak{go}_n(Q_n, \mathbb{C}) = n(n+1)/2$.*

Proof. Let $A \in M_n(\mathbb{C})$. We have $A \in \mathfrak{go}_n(Q_n, \mathbb{C})$ if and only if $A^{\text{tr}} Q_n = -Q_n A$. Since $A^{\text{tr}} Q_n = \epsilon (Q_n A)^{\text{tr}}$, this is equivalent to the condition $(Q_n A)^{\text{tr}} = -\epsilon Q_n A$. Thus, we have a bijective linear map

$$\mathfrak{go}_n(Q_n, \mathbb{C}) \to \{S \in M_n(\mathbb{C}) \mid S^{\text{tr}} = -\epsilon S\}, \quad A \mapsto Q_n A.$$

If $\epsilon = -1$, then the space on the right hand side consists precisely of all symmetric matrices in $M_n(\mathbb{C})$; hence, its dimension equals $n(n+1)/2$; similarly, if $\epsilon = 1$, then its dimension equals $n(n-1)/2$.

It remains to prove the statements about bases. All we need to do now is to find the appropriate number of linearly independent elements. First note that $Q_n E_{ij} = \delta_i E_{n+1-i,j}$. Hence, we have

$$\begin{aligned} Q_n A_{ij} &= \delta_i Q_n E_{ij} - \delta_j Q_n E_{n+1-j,n+1-i} \\ &= \delta_i^2 E_{n+1-i,j} - \delta_j \delta_{n+1-j} E_{j,n+1-i} = E_{n+1-i,j} - \epsilon E_{j,n+1-i}. \end{aligned}$$

Furthermore, $A_{ij}^{\text{tr}} Q_n = \epsilon (Q_n A_{ij})^{\text{tr}} = \epsilon (E_{n+1-i,j}^{\text{tr}} - \epsilon E_{j,n+1-i}^{\text{tr}})$ and so $A_{ij}^{\text{tr}} Q_n + Q_n A_{ij} = 0$, that is, $A_{ij} \in \mathfrak{go}_n(Q_n, \mathbb{C})$ for all $1 \leqslant i,j \leqslant n$.

Consider the set $I := \{(i,j) \mid 1 \leqslant i,j \leqslant n, i+j \leqslant n\}$; note that $|I| = n(n-1)/2$. Furthermore, if $(i,j) \in I$, then $(n+1-i)+(n+1-j) \geqslant n+2$ and so $(n+1-j, n+1-i) \notin I$. This implies that the set $\{A_{ij} \mid (i,j) \in I\} \subseteq \mathfrak{go}_n(Q_n, \mathbb{C})$ is linearly independent. Furthermore, for $1 \leqslant i \leqslant n$, we have $(i, n+1-i) \notin I$, $(n+1-i, i) \notin I$ and

$$A_i := A_{i,n+1-i} = \delta_i (1-\epsilon) E_{i,n+1-i}.$$

Hence, if $\epsilon = -1$, then $A_i \neq 0$ and $\{A_{ij} \mid (i,j) \in I\} \cup \{A_i \mid 1 \leqslant i \leqslant n\}$ is linearly independent. Thus, (a) and (b) are proved. $\square$

Remark 1.6.7. Denote by $\operatorname{diag}(x_1, \ldots, x_n) \in M_n(\mathbb{C})$ the diagonal matrix with diagonal coefficients $x_1, \ldots, x_n \in \mathbb{C}$. Let H be the subspace of $\mathfrak{go}_n(Q_n, \mathbb{C})$ consisting of all matrices in $\mathfrak{go}_n(Q_n, \mathbb{C})$ that are diagonal. Let $m \geqslant 1$ be such that $n = 2m+1$ (if n is odd) or $n = 2m$

(if n is even). Then H consists precisely of all diagonal matrices of the form

$$\begin{cases} \operatorname{diag}(x_1, \ldots, x_m, 0, -x_m, \ldots, -x_1) & \text{if } n \text{ is odd,} \\ \operatorname{diag}(x_1, \ldots, x_m, -x_m, \ldots, -x_1) & \text{if } n \text{ is even.} \end{cases}$$

(See exercises.) In particular, $\dim H = m$. With the above definition of m, the dimension formulae in Proposition 1.6.6 are re-written as follows:

$$\dim \mathfrak{go}_n(Q_n, \mathbb{C}) = \begin{cases} 2m^2 - m & \text{if } n = 2m \text{ and } Q_n^{\text{tr}} = Q_n, \\ 2m^2 + m & \text{otherwise.} \end{cases}$$

Corollary 1.6.8 (Triangular decomposition). *Let $L = \mathfrak{go}_n(Q_n, \mathbb{C})$, as above. Then every $x \in L$ has a unique expression $x = h + n^+ + n^-$ where $h \in L$ is a diagonal matrix, $n^+ \in L$ is a strictly upper triangular matrix, and $n^- \in L$ is a strictly lower triangular matrix.*

Proof. Note that A_{ij} is diagonal if $i = j$, strictly upper triangular if $i < j$, and strictly lower triangular if $i > j$. So the assertion follows from Proposition 1.6.6. □

We shall see later that the algebras $\mathfrak{sl}_n(\mathbb{C})$ and $\mathfrak{go}_n(Q_n, \mathbb{C})$ are not only semisimple but simple (with the exceptions in Exercise 1.6.4). The following result highlights the importance of these algebras.

Theorem 1.6.9 (Cartan–Killing Classification). *Let L be a semisimple Lie algebra over $\mathbb{C}$ with $\dim L < \infty$. Then L is a direct product of simple Lie algebras, each of which is isomorphic to either $\mathfrak{sl}_n(\mathbb{C})$ ($n \geqslant 2$), or $\mathfrak{go}_n(Q_n, \mathbb{C})$ ($n \geqslant 3$ and Q_n as above), or to one of five "exceptional" algebras that are denoted by G_2, F_4, E_6, E_7, E_8 and are of dimension 14, 52, 78, 133, 248, respectively.*

This classification result is proved in textbooks like those of Carter [**7**], Erdmann–Wildon [**11**] or Humphreys [**18**], to mention just a few (see also Bourbaki [**5**]). It is achieved as the culmination of an elaborate chain of arguments. Here, we shall take a shortcut around that proof. Following Moody–Pianzola [**25**], we will work in a setting where the existence of something like a "triangular decomposition" (as in Corollary 1.6.8) is systematically adopted at the outset.

Chapter 2

Semisimple Lie algebras

In this chapter we develop the theory of semisimple Lie algebras using the aproach mentioned at the end of Chapter 1. This approach provides a uniform framework for studying the various Lie algebras appearing in Theorem 1.6.9. It is completely self-contained; no prior knowledge about simple Lie algebras is required. One advantage is that it allows us to reach more directly the point where we can deal with certain more modern aspects of the theory of Lie algebras, and with the construction of Chevalley groups.

The last section contains the highlight of this chapter: the construction of Lusztig's "canonical basis" for a semisimple Lie algebra. This is a relatively recent development in the theory of Lie algebras, dating from around 1990.

Throughout this chapter, we work over the base field $k = \mathbb{C}$.

2.1. Weights and weight spaces

Let H be a finite-dimensional Lie algebra, and $\rho\colon H \to \mathfrak{gl}(V)$ be a representation of H on a finite-dimensional vector space $V \neq \{0\}$ (all over $k = \mathbb{C}$). Thus, V is an H-module as in Section 1.4. Assume that H is solvable. By Lie's Theorem 1.5.4, there exists a basis B of V such that, for any $x \in H$, the matrix of the linear map $\rho_x\colon V \to V$,

$v \mapsto x.v$, with respect to B has an upper triangular shape as follows:

$$M_B(\rho_x) = \begin{pmatrix} \lambda_1(x) & * & \dots & * \\ 0 & \lambda_2(x) & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & \lambda_n(x) \end{pmatrix} \qquad (n = \dim V),$$

where $\lambda_i \in H^* := \mathrm{Hom}(H, \mathbb{C})$ are linear maps for $1 \leqslant i \leqslant n$. By Lemma 1.5.5, the set $P(V) := \{\lambda_1, \dots, \lambda_n\} \subseteq H^*$ does not depend on the choice of the basis B and is called the set of *weights* of H on V. We will from now on make the stronger assumption that

$$\boxed{H \text{ is abelian.}}$$

A particularly favourable situation occurs when the matrices $M_B(\rho_x)$ are diagonal for all $x \in H$. This leads to the following definition.

Definition 2.1.1. In the above setting (with H abelian), we say that the H-module V is *H-diagonalisable* if, for each $x \in H$, the linear map $\rho_x \colon V \to V$ is diagonalisable, that is, there exists a basis of V such that the corresponding matrix of ρ_x is a diagonal matrix (but, a priori, the basis may depend on the element $x \in H$).

A linear map $\rho \colon H \to \mathrm{End}(V)$ is a representation of Lie algebras if and only if $\rho([x, x']) = \rho(x) \circ \rho(x') - \rho(x') \circ \rho(x)$ for all $x, x' \in H$. Since H is abelian, this just means that the maps $\{\rho(x) \mid x \in H\} \subseteq \mathrm{End}(V)$ commute with each other. Thus, the following results are really statements about commuting matrices, but it is useful to formulate them in terms of the abstract language of modules for Lie algebras in view of the later applications to "weight space decompositions".

Lemma 2.1.2. *Assume that V is H-diagonalisable. Let $U \subseteq V$ be an H-submodule. Then U is also H-diagonalisable.*

Proof. Let $x \in H$ and $\lambda_1, \dots, \lambda_r \in \mathbb{C}$ (where $r \geqslant 1$) be the distinct eigenvalues of $\rho_x \colon V \to V$. Then $V = V_1 + \ldots + V_r$ where V_i is the λ_i-eigenspace of ρ_x. Setting $U_i := U \cap V_i$ for $1 \leqslant i \leqslant r$, we claim that $U = U_1 + \ldots + U_r$. Now, let $u \in U$ and write $u = v_1 + \ldots + v_r$ where $v_i \in V_i$ for $1 \leqslant i \leqslant r$. We must show that $v_i \in U$ for all i. For this purpose, we define a sequence of vectors $(u_j)_{j \geqslant 1}$ by $u_1 := u$ and

$u_j := x.u_{j-1}$ for $j \geqslant 2$. Then a simple induction on j shows that

$$u_j = \lambda_1^{j-1} v_1 + \ldots + \lambda_r^{j-1} v_r \qquad \text{for all } j \geqslant 1.$$

Since the Vandermonde matrix $\big(\lambda_i^{j-1}\big)_{1 \leqslant i,j \leqslant r}$ is invertible, we can invert the above equations (for $j = 1, \ldots, r$) and find that each v_i is a linear combination of $u_1, \ldots, u_r$. Since U is an H-submodule of V, we have $u_j \in U$ for all j, and so $v_i \in U$ for all i, as claimed.

Now $U_i = U \cap V_i = \{u \in U \mid x.u = \lambda_i u\}$ for all i. Hence, all non-zero vectors in U_i are eigenvectors of the restricted map $\rho_x|_U : U \to U$. Consequently, $U = U_1 + \ldots + U_r$ is spanned by eigenvectors for $\rho_x|_U$ and, hence, $\rho_x|_U$ is diagonalisable. □

Proposition 2.1.3. *Assume that V is H-diagonalisable; let $n = \dim V \geqslant 1$. Then there exist $\lambda_1, \ldots, \lambda_n \in H^*$ and one basis B of V such that, for all $x \in H$, the matrix of $\rho_x : V \to V$ with respect to B is diagonal, with $\lambda_1(x), \ldots, \lambda_n(x)$ along the diagonal.*

Proof. We proceed by induction on $\dim V$. If $\dim V = 1$, the result is clear. Now assume that $\dim V > 1$. If ρ_x is a scalar multiple of the identity for all $x \in H$ then, again, the result is clear. Now assume that there exists some $y \in H$ such that ρ_y is not a scalar multiple of the identity. Since ρ_y is diagonalisable by assumption, there are at least two distinct eigenvalues. So let $\lambda_1, \ldots, \lambda_r \in \mathbb{C}$ be the distinct eigenvalues of ρ_y, where $r \geqslant 2$. Then $V = V_1 \oplus \ldots \oplus V_r$ where V_i is the λ_i-eigenspace of ρ_y. We claim that each V_i is an H-submodule of V. Indeed, let $v \in V_i$ and $x \in H$. Since H is abelian, we have $\rho_x \circ \rho_y = \rho_y \circ \rho_x$. This yields $\rho_y(x.v) = (\rho_y \circ \rho_x)(v) = (\rho_x \circ \rho_y)(v) = \rho_x(y.v) = \lambda_i(y)\rho_x(v) = \lambda_i(y)(x.v)$ and so $x.v \in V_i$. By Lemma 2.1.2, each subspace V_i is H-diagonalisable. Now $\dim V_i < \dim V$ for all i. So, by induction, there exist bases B_i of V_i such that the matrices of $\rho_x|_{V_i} : V_i \to V_i$ are diagonal for all $x \in H$. Since $V = V_1 \oplus \ldots \oplus V_r$, the set $B := B_1 \cup \ldots \cup B_r$ is a basis of V with the required property. □

Given $\lambda \in H^*$, a non-zero vector $v \in V$ is called a *weight vector* (with weight λ) if $x.v = \lambda(x)v$ for all $x \in H$. We set

$$V_\lambda := \{v \in V \mid x.v = \lambda(x)v \text{ for all } x \in H\}.$$

Clearly, V_λ is a subspace of V. If $V_\lambda \neq \{0\}$, then V_λ is called a *weight space* for H on V. In the setting of Proposition 2.1.3, write $B = \{v_1, \ldots, v_n\}$. Then $x.v_i = \lambda_i(x)v_i$ for all $x \in H$ and so $v_i \in V_{\lambda_i}$. Thus, we have $V = \sum_{1 \leqslant i \leqslant n} V_{\lambda_i}$, that is, V is a sum of weight spaces.

Proposition 2.1.4. *Assume that V is H-diagonalisable. Recall the definition of the set of weights $P(V) \subseteq H^*$ above.*

(a) *For $\lambda \in H^*$, we have $\lambda \in P(V)$ if and only if $V_\lambda \neq \{0\}$.*

(b) *We have $V = \bigoplus_{\lambda \in P(V)} V_\lambda$.*

(c) *If $U \subseteq V$ is an H-submodule, then $U = \bigoplus_{\lambda \in P(U)} U_\lambda$ where $P(U) \subseteq P(V)$ and $U_\lambda = U \cap V_\lambda$ for all $\lambda \in P(U)$.*

Proof. Let B and $\lambda_1, \ldots, \lambda_n \in H^*$ as in Proposition 2.1.3. Then $P(V) = \{\lambda_1, \ldots, \lambda_n\}$. Writing $B = \{v_1, \ldots, v_n\}$, we already observed above that $v_i \in V_{\lambda_i}$ for all i. Consequently, $V = \sum_{1 \leqslant i \leqslant n} V_{\lambda_i}$.

(a) Assume first that $\lambda \in P(V)$. By definition, this means that $\lambda = \lambda_i$ for some i. Then $x.v_i = \lambda_i(x)v_i$ for all $x \in H$ and so $v_i \in V_{\lambda_i}$. Conversely, if $V_\lambda \neq \{0\}$, then there exists some $0 \neq v \in V$ such that $x.v = \lambda(x)v$ for all $x \in H$. Then $v \in V = \sum_{1 \leqslant i \leqslant n} V_{\lambda_i}$ and so Exercise 2.1.5 below shows that $\lambda = \lambda_i$ for some i.

(b) The λ_i need not be distinct. So assume that $|P(V)| = r \geqslant 1$ and write $P(V) = \{\mu_1, \ldots, \mu_r\}$; then $V = \sum_{1 \leqslant i \leqslant r} V_{\mu_i}$. We now show that the sum is direct. If $r = 1$, there is nothing to prove. So assume now that $r \geqslant 2$ and consider the finite subset

$$\{\mu_i - \mu_j \mid 1 \leqslant i < j \leqslant r\} \subseteq H^*.$$

By Exercice 1.4.14, we can choose $x_0 \in H$ such that all elements of that subset have a non-zero value on x_0. Thus, $\mu_1(x_0), \ldots, \mu_r(x_0)$ are all distinct. Then $V = V_1 \oplus \ldots \oplus V_r$ where V_i is the $\mu_i(x_0)$-eigenspace of V. Now, we certainly have

$$V = \sum_{1 \leqslant i \leqslant r} V_{\mu_i} \subseteq \bigoplus_{1 \leqslant i \leqslant r} V_i = V;$$

note that $V_{\mu_i} = \{v \in V \mid x.v = \mu_i(x)v \text{ for all } x \in H\} \subseteq V_i$. Hence, we must have $V_{\mu_i} = V_i$ for all i.

(c) By Lemma 2.1.2, U is H-diagonalisable. So, applying (b) to U, we obtain that $U = \bigoplus_{\lambda\in P(U)} U_\lambda$. Now, we certainly have

$$U_\lambda = \{u \in U \mid x.u = \lambda(x)u \text{ for all } x \in H\} = U \cap V_\lambda$$

for any $\lambda \in H^*$. Using (a), this shows that $P(U) \subseteq P(V)$. $\square$

Exercise 2.1.5. Let H be abelian and V be an H-module. Let $r \geqslant 1$ and $\lambda, \lambda_1, \ldots, \lambda_r \in H^*$. Assume that $0 \neq v \in V_\lambda$ and $v \in \sum_{1\leqslant i\leqslant r} V_{\lambda_i}$. Then show that $\lambda = \lambda_i$ for some i. (This generalises the familiar fact that eigenvectors corresponding to pairwise distinct eigenvalues are linearly independent.)

ab hier Woche 4

Now assume that H is a subalgebra of a larger Lie algebra L with $\dim L < \infty$. Then L becomes an H-module via the restriction of $\mathrm{ad}_L\colon L \to \mathfrak{gl}(L)$ to H. So, for any $\lambda \in H^*$, we have

$$L_\lambda = \{y \in L \mid [x, y] = \lambda(x)y \text{ for all } x \in H\}.$$

In particular, $L_{\underline{0}} = C_L(H) := \{y \in L \mid [x, y] = 0 \text{ for all } x \in H\} \supseteq H$, where $\underline{0} \in H^*$ denotes the 0-map. If L is H-diagonalisable, then we can apply the above discussion and obtain a decomposition

$$L = \bigoplus_{\lambda\in P(L)} L_\lambda \quad \text{where } P(L) \text{ is the set of weights of } H \text{ on } L.$$

Proposition 2.1.6. *We have $[L_\lambda, L_\mu] \subseteq L_{\lambda+\mu}$ for all $\lambda, \mu \in H^*$; furthermore, $L_{\underline{0}}$ is a subalgebra of L. If L is H-diagonalisable, then we have the implication: $L_{\underline{0}} = \sum_{\lambda\in P(L)}[L_\lambda, L_{-\lambda}] \Rightarrow L = [L, L]$.*

Proof. Let $v \in L_\lambda$ and $w \in L_\mu$. Thus, $[x, v] = \lambda(x)v$ and $[x, w] = \mu(x)w$ for all $x \in H$. Using anti-symmetry and the Jacobi identity, we obtain that

$$\begin{aligned}[x, [v, w]] &= -[v, [w, x]] - [w, [x, v]] = [v, [x, w]] + [[x, v], w] \\ &= \mu(x)[v, w] + \lambda(x)[v, w] = (\lambda(x) + \mu(x))[v, w]\end{aligned}$$

for all $x \in H$ and so $[v, w] \in L_{\lambda+\mu}$. Furthermore, since H is abelian, $H \subseteq L_{\underline{0}} = \{y \in L \mid [x, y] = 0 \text{ for all } x \in H\}$. We have $[L_{\underline{0}}, L_{\underline{0}}] \subseteq L_{\underline{0}}$ and so $L_{\underline{0}} \subseteq L$ is a subalgebra. Now assume that L is H-diagonalisable and that $L_{\underline{0}} = \sum_{\lambda\in P(L)}[L_\lambda, L_{-\lambda}]$. Then $L_{\underline{0}} \subseteq [L, L]$. Now let $\lambda \in P(L)$, $\lambda \neq \underline{0}$. Then there exists some $h \in H$ such that $\lambda(h) \neq 0$. For any $v \in L_\lambda$ we have $[h, v] = \lambda(h)v$. So v is a non-zero multiple of $[h, v] \in [L, L]$. It follows that $L_\lambda \subseteq [L, L]$. Consequently, we have $L = \sum_{\lambda \in P(L)} L_\lambda \subseteq [L, L]$ and so $L = [L, L]$. $\square$

The following result will be useful to verify H-diagonalisability.

Lemma 2.1.7. *Let $H \subseteq L$ be abelian and $X \subseteq L$ be a subset such that $L = \langle X \rangle_{\text{alg}}$. Assume that there is a subset $\{\lambda_x \mid x \in X\} \subseteq H^*$ such that $x \in L_{\lambda_x}$ for all $x \in X$. Then L is H-diagonalisable, where every $\lambda \in P(L)$ is a $\mathbb{Z}_{\geqslant 0}$-linear combination of $\{\lambda_x \mid x \in X\}$.*

Proof. Recall from Section 1.1 that $\langle X \rangle_{\text{alg}} = \langle X_n \mid n \geqslant 1 \rangle_{\mathbb{C}}$, where X_n consists of all Lie monomials in X of level n. Let us also set

$$\Lambda_n := \{\lambda \in H^* \mid \lambda = \lambda_{x_1} + \ldots + \lambda_{x_n} \text{ for some } x_i \in X\}.$$

We show by induction on n that, for each $x \in X_n$, there exists some $\lambda \in \Lambda_n$ such that $x \in L_\lambda$. If $n = 1$, then this is clear by our assumptions on X. Now let $n \geqslant 2$ and $x \in X_n$. So $x = [y, z]$ where $y \in X_i$, $z \in X_{n-i}$ and $1 \leqslant i \leqslant n-1$. By induction, there are $\lambda \in \Lambda_i$ and $\mu \in \Lambda_{n-i}$ such that $y \in L_\lambda$ and $z \in L_\mu$. By Proposition 2.1.6, we have $x = [y, z] \in [L_\lambda, L_\mu] \subseteq L_{\lambda+\mu}$, where $\lambda + \mu \in \Lambda_{i+(n-i)} = \Lambda_n$, as desired. We conclude that L is H-diagonalisable; more precisely,

$$L = \langle X_n \mid n \geqslant 1 \rangle_{\mathbb{C}} = \sum_{n \geqslant 1} \sum_{\lambda \in \Lambda_n} L_\lambda,$$

and each $\lambda \in P(L)$ is a non-negative sum of various λ_x $(x \in X)$. $\square$

The following result will allow us to apply the exponential construction in Lemma 1.2.8 to many elements in L.

Lemma 2.1.8. *Let $H \subseteq L$ be abelian and L be H-diagonalisable. Let $\underline{0} \neq \lambda \in P(L)$ and $y \in L_\lambda$. Then $\operatorname{ad}_L(y) \colon L \to L$ is nilpotent.*

Proof. Let $\mu \in P(L)$ and $v \in L_\mu$. Then $\operatorname{ad}_L(y)(v) = [y, v] \in L_{\lambda+\mu}$ by Proposition 2.1.6. A simple induction on m shows that $\operatorname{ad}_L(y)^m(v) \in L_{m\lambda+\mu}$ for all $m \geqslant 0$. Since $\{m\lambda + \mu \mid m \geqslant 0\} \subseteq H^*$ is an infinite subset and $P(L)$ is finite, there is some $m > 0$ such that $m\lambda + \mu \notin P(L)$ and so $\operatorname{ad}_L(y)^m(v) = 0$. Hence, since $L = \langle L_\mu \mid \mu \in P(L) \rangle_{\mathbb{C}}$, we conclude that $\operatorname{ad}_L(y)$ is nilpotent (see Exercise 1.2.4(a)). $\square$

Exercise 2.1.9. In the setting of Lemma 2.1.8, let $y \in L_\lambda$ where $\underline{0} \neq \lambda \in P(L)$. Then $\mathrm{ad}_L(y)\colon L \to L$ is a nilpotent derivation and so we can form $\varphi := \exp\big(\mathrm{ad}_L(y)\big) \in \mathrm{Aut}(L)$. Show that, if $J \subseteq L$ is an ideal, then $\varphi(J) \subseteq J$.

Example 2.1.10. Let $L = \mathfrak{gl}_n(\mathbb{C})$, the Lie algebra of all $n \times n$-matrices over $\mathbb{C}$. A natural candidate for an abelian subalgebra is

$$H := \{x \in L \mid x \text{ diagonal matrix}\} \qquad (\dim H = n).$$

For $1 \leqslant i \leqslant n$, let $\varepsilon_i \in H^*$ be the map that sends a diagonal matrix to its i-th diagonal entry. Then $\{\varepsilon_1, \ldots, \varepsilon_n\}$ is a basis of H^*. If $n = 1$, then $L = H$. Assume now that $n \geqslant 2$; then $H \subsetneqq L$. For $i \neq j$ let $e_{ij} \in L$ be the matrix with entry 1 at position (i,j), and 0 everywhere else. Then a simple matrix calculation shows that

(a) $$[x, e_{ij}] = (\varepsilon_i(x) - \varepsilon_j(x))e_{ij} \qquad \text{for all } x \in H.$$

Thus, $\varepsilon_i - \varepsilon_j \in P(L)$ and $e_{ij} \in L_{\varepsilon_i - \varepsilon_j}$. Furthermore, we have

(b) $$L = \underbrace{H}_{\subseteq L_{\underline{0}}} \oplus \bigoplus_{\substack{1 \leqslant i,j \leqslant n \\ i \neq j}} \underbrace{\mathbb{C}e_{ij}}_{\subseteq L_{\varepsilon_i - \varepsilon_j}}.$$

So L is H-diagonalisable, where $P(L) = \{\underline{0}\} \cup \{\varepsilon_i - \varepsilon_j \mid i \neq j\}$. Next, note that the weights $\varepsilon_i - \varepsilon_j$ for $i \neq j$ are pairwise distinct and non-zero. Since there are $n^2 - n$ of them, Proposition 2.1.4 shows that $\dim L = \dim L_{\underline{0}} + \sum_{i \neq j} \dim L_{\varepsilon_i - \varepsilon_j} \geqslant n + (n^2 - n) = n^2 = \dim L$. Hence, all the above inequalities and inclusions must be equalities. We conclude that

(c) $$L_{\underline{0}} = H \qquad \text{and} \qquad L_{\varepsilon_i - \varepsilon_j} = \langle e_{ij} \rangle_{\mathbb{C}} \quad \text{for all } i \neq j.$$

Finally, as in Corollary 1.6.8, we have a *triangular decomposition* $L = N^+ \oplus H \oplus N^-$ where N^+ is the subalgebra consisting of all strictly upper triangular matrices in $\mathfrak{gl}_n(\mathbb{C})$ and N^- is the subalgebra consisting of all stricly lower triangular matrices in $\mathfrak{gl}_n(\mathbb{C})$. This decomposition is reflected in properties of $P(L)$ as follows. We set

$$\Phi^+ = \{\varepsilon_i - \varepsilon_j \mid 1 \leqslant i < j \leqslant n\} \quad \text{and} \quad \Phi^- := -\Phi^+.$$

Then $P(L) = \{\underline{0}\} \sqcup \Phi^+ \sqcup \Phi^-$ (disjoint union) and $N^\pm = \bigoplus_{\alpha \in \Phi^\pm} L_\alpha$. Thus, the decomposition $L = N^+ \oplus H \oplus N^-$ gives rise to a partition of $P(L) \setminus \{\underline{0}\}$ into a "positive" part Φ^+ and a "negative" part Φ^-. We also note that, for $1 \leqslant i < j \leqslant n$, we have

$$\varepsilon_i - \varepsilon_j = (\varepsilon_i - \varepsilon_{i+1}) + (\varepsilon_{i+1} - \varepsilon_{i+2}) + \ldots + (\varepsilon_{j-1} - \varepsilon_j).$$

Hence, if we set $\alpha_i := \varepsilon_i - \varepsilon_{i+1}$ for $1 \leqslant i \leqslant n-1$, then

(d) $$\Phi^{\pm} = \big\{\pm\big(\alpha_i + \alpha_{i+1} + \ldots + \alpha_{j-1}\big) \mid 1 \leqslant i < j \leqslant n\big\}.$$

Thus, setting $\Delta = \{\alpha_1, \ldots, \alpha_{n-1}\}$, every non-zero weight of H on L can be expressed uniquely as a sum of elements of Δ or of $-\Delta$. (Readers familiar with the theory of abstract root systems will recognise the concept of "simple roots" in the above properties of Δ; see, e.g., Bourbaki [**4**, Ch. VI, §1].) In any case, this picture is the prototype of what is also going on in the Lie algebras $\mathfrak{sl}_n(\mathbb{C})$ and $\mathfrak{go}_n(Q_n, \mathbb{C})$, and this is what we will formalise in Definition 2.2.1 below. For the further discussion of examples, the following remark will be useful.

Remark 2.1.11. Let $L \subseteq \mathfrak{gl}_n(\mathbb{C})$ be a subalgebra, and $H \subseteq L$ be the abelian subalgebra consisting of all diagonal matrices that are contained in L. First we claim that

(a) L is H-diagonalisable.

Indeed, by the previous example, $\mathrm{ad}_{\mathfrak{gl}_n(\mathbb{C})}(x)\colon \mathfrak{gl}_n(\mathbb{C}) \to \mathfrak{gl}_n(\mathbb{C})$ is diagonalisable for all diagonal matrices $x \in \mathfrak{gl}_n(\mathbb{C})$ and, hence, also for all $x \in H$. Thus, $\mathfrak{gl}_n(\mathbb{C})$ is H-diagonalisable. Now $[H, L] \subseteq L$ and so L is an H-submodule of $\mathfrak{gl}_n(\mathbb{C})$. So L is H-diagonalisable by Lemma 2.1.2. Furthermore, we have the following useful criterion:

(b) We have $H = C_L(H)$ if there exists some $x_0 \in H$ with distinct diagonal entries.

Indeed, let $x_0 = \mathrm{diag}(x_1, \ldots, x_n) \in H$ with distinct entries $x_i \in \mathbb{C}$ and $y = (y_{ij}) \in L$ be such that $[x_0, y] = x_0 \cdot y - y \cdot x_0 = 0$. Then $x_i y_{ij} = y_{ij} x_j$ for all i, j and so $y_{ij} = 0$ for $i \neq j$. Thus, y is a diagonal matrix. Since $y \in L$, we have $y \in H$, as required.

For example, if $L = \mathfrak{sl}_n(\mathbb{C})$, then H will consist of all diagonal matrices with trace 0. In this case, we can take

$$x_0 = \mathrm{diag}(1, 2, \ldots, n-1, -n(n-1)/2) \in H.$$

If $L = \mathfrak{go}_n(Q_n, \mathbb{C})$, then the diagonal matrices in L are described in Remark 1.6.7. In these cases, writing $n = 2m+1$ (if m if odd) or

$n = 2m$ (if n is even), we may take

$$x_0 = \begin{cases} \operatorname{diag}(1, \ldots, m, 0, -m, \ldots, -1) & \text{if } n \text{ is odd}, \\ \operatorname{diag}(1, \ldots, m, -m, \ldots, -1) & \text{if } n \text{ is even}. \end{cases}$$

Example 2.1.12. Consider the subalgebra $L_\delta \subseteq \mathfrak{gl}_3(\mathbb{C})$ in Exercise 1.3.3, where $0 \neq \delta \in \mathbb{C}$. Then the elements

$$e = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad h := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \delta \end{pmatrix}, \quad f = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

form a basis of L_δ and one checks by an explicit computation that

$$[h, e] = e, \qquad [h, f] = \delta f, \qquad [e, f] = 0.$$

Hence, we have a triangular decomposition $L_\delta = N^+ \oplus H \oplus N^-$, where $N^+ = \langle e \rangle_{\mathbb{C}}$, $N^- = \langle f \rangle_{\mathbb{C}}$ and $H := \langle h \rangle_{\mathbb{C}}$. We have $C_{L_\delta}(H) = H$ since h satisfies the condition (b) in Remark 2.1.11. The corresponding set of weights is given by $P(L_\delta) = \{\underline{0}, \alpha, \delta\alpha\}$, where $\alpha \in H^*$ is defined by $\alpha(h) = 1$. Thus, if $\delta = -1$, then we have a partition of $P(L_\delta) \setminus \{\underline{0}\}$ into a "positive" and a "negative" part (symmetrical to each other). On the other hand, if $\delta = 1$, then we only have a "positive" part but no "negative" part at all. So this example appears to differ from that of $\mathfrak{gl}_n(\mathbb{C})$ in a crucial way. We shall see that this difference has to do with the property that $[e, f] = 0$, that is, $[N^+, N^-] = \{0\}$. We also know from Exercise 1.3.3 that L_δ is solvable, while $\mathfrak{gl}_n(\mathbb{C})$ is not.

2.2. Lie algebras of Cartan–Killing type

Let L be a finite-dimensional Lie algebra over $k = \mathbb{C}$, and $H \subseteq L$ be an abelian subalgebra. Then we regard L as an H-module via the restriction of $\operatorname{ad}_L \colon L \to \mathfrak{gl}(L)$ to H. Let $P(L) \subseteq H^*$ be the corresponding set of weights. Motivated by the examples and the discussion in the previous section, we introduce the following definition.

Definition 2.2.1 (Cf. Kac [**21**, Chap. 1] and Moody–Pianzola [**25**, §2.1 and §4.1]). We say that (L, H) is of *Cartan–Killing type* if there exists a linearly independent subset $\Delta = \{\alpha_i \mid i \in I\} \subseteq H^*$ (where I is a finite index set) such that the following conditions are satisfied.

(CK1) L is H-diagonalisable, where $L_{\underline{0}} = H$.

(CK2) Each $\lambda \in P(L)$ is a $\mathbb{Z}$-linear combination of $\Delta = \{\alpha_i \mid i \in I\}$ where the coefficients are either all $\geqslant 0$ or all $\leqslant 0$.

(CK3) We have $L_{\underline{0}} = \sum_{i \in I}[L_{\alpha_i}, L_{-\alpha_i}]$.

We set $\Phi := \{\alpha \in P(L) \mid \alpha \neq \underline{0}\}$. Thus, $L = H \oplus \bigoplus_{\alpha \in \Phi} L_\alpha$, which is called the *Cartan decomposition* of L. Then H is called a *Cartan subalgebra* and Φ the set of *roots* of L with respect to H. We may also speak of (Φ, Δ) as a *based root system.*

We say that $\alpha \in \Phi$ is a *positive root* if $\alpha = \sum_{i \in I} n_i \alpha_i$ where $n_i \geqslant 0$ for all $i \in I$; similarly, $\alpha \in \Phi$ is a *negative root* if $\alpha = \sum_{i \in I} n_i \alpha_i$ where $n_i \leqslant 0$ for all $i \in I$. Let Φ^+ be the set of all positive roots and Φ^- be the set of all negative roots. Thus, $\Phi = \Phi^+ \sqcup \Phi^-$ (disjoint union).

Remark 2.2.2. We will see later that a Lie algebra L as in Definition 2.2.1 is semisimple; so all of the above notions ("Cartan subalgebra", "roots" etc.) are consistent with the common usage in the general theory of semisimple Lie algebras. Conversely, any semisimple Lie algebra is of Cartan–Killing type. This result is in fact proved along the proof of the classification result in Theorem 1.6.9.

The further theory will now be developed from the axioms in Definition 2.2.1. We begin with the following two basic results.

Lemma 2.2.3. *Assume that L is H-diagonalisable. Let $\lambda \in H^*$ be such that $[L_\lambda, L_{-\lambda}] \subseteq H$. If the restriction of λ to $[L_\lambda, L_{-\lambda}]$ is zero, then* $\text{ad}_L(x) = 0$ *for all $x \in [L_\lambda, L_{-\lambda}]$.*

Proof. Let $y \in L_\lambda$, $z \in L_{-\lambda}$, and set $x := [y, z] \in [L_\lambda, L_{-\lambda}] \subseteq H$. Consider the subspace $S := \langle x, y, z\rangle_{\mathbb{C}} \subseteq L$. Since $\lambda(x) = 0$, we have $[x, y] = \lambda(x)y = 0$, $[x, z] = -\lambda(x)z = 0$ and $[y, z] = x$. Thus, S is a subalgebra of L; furthermore, $[S, S] = \langle x\rangle_{\mathbb{C}}$ and so S is solvable. We regard L as an S-module via the restriction of $\text{ad}_L \colon L \to \mathfrak{gl}(L)$ to S. Since S is solvable, Lie's Theorem 1.5.4 shows that there is a basis B of L such that, for any $s \in S$, the matrix of $\text{ad}_L(s)$ with respect to B is upper triangular. Now $x = [y, z]$ and so $\text{ad}_L(x) = \text{ad}_L(y)\circ\text{ad}_L(z) - \text{ad}_L(z)\circ\text{ad}_L(y)$. Hence, the matrix of $\text{ad}_L(x)$ is upper triangular with 0 along the diagonal. But $\text{ad}_L(x)$ is diagonalisable and so $\text{ad}_L(x) = 0$, as desired. $\square$

Lemma 2.2.4. *Assume that L is H-diagonalisable. Let $\lambda \in H^*$ be such that $[L_\lambda, L_{-\lambda}] \subseteq H$ and the restriction of λ to $[L_\lambda, L_{-\lambda}]$ is non-zero; in particular, $\lambda \neq \underline{0}$ and $L_\lambda \neq \{0\}$. Then we have* $\dim L_{\pm\lambda} = 1$ *and $P(L) \cap \{n\lambda \mid n \in \mathbb{Z}\} = \{\underline{0}, \pm\lambda\}$.*

Proof. By assumption, there exist elements $e \in L_\lambda$ and $f \in L_{-\lambda}$ such that $h := [e, f] \in [L_\lambda, L_{-\lambda}] \subseteq H$ and $\lambda(h) \neq 0$. Note that $e \neq 0$, $f \neq 0$, $h \neq 0$. Replacing f by a scalar multiple if necessary, we may assume that $\lambda(h) = 2$. Then we have the relations

$$[e, f] = h, \quad [h, e] = \lambda(h)e = 2e, \quad [h, f] = -\lambda(x)f = -2f.$$

Thus, $S := \langle e, h, f\rangle_{\mathbb{C}}$ is a 3-dimensional subalgebra of L that is isomorphic to $\mathfrak{sl}_2(\mathbb{C})$ (see Exercise 1.2.11). Let $p := \max\{n \geqslant 1 \mid L_{n\lambda} \neq \{0\}\}$ and consider the subspace

$$M := \mathbb{C}f \oplus H \oplus L_\lambda \oplus L_{2\lambda} \oplus \ldots \oplus L_{p\lambda} \subseteq L,$$

where $\mathbb{C}f \subseteq L_{-\lambda}$, $H \subseteq L_{\underline{0}}$ and some terms $L_{n\lambda}$ may be $\{0\}$ for $2 \leqslant n < p$. By Proposition 2.1.6, we have $[L_{n\lambda}, L_{m\lambda}] \subseteq L_{(n+m)\lambda}$ for all $n, m \in \mathbb{Z}$. Furthermore, $[f, y] \in H$ for all $y \in L_\lambda$ (by assumption), $[x, f] = -\lambda(x)f \in \mathbb{C}f$ for all $x \in H$, and $[H, L_{n\lambda}] \subseteq L_{n\lambda}$ for all $n \in \mathbb{Z}$. It follows that $[S, M] \subseteq M$ and so M may be regarded as an S-module via the restriction of $\mathrm{ad}_L \colon L \to \mathfrak{gl}(L)$ to S. The set of eigenvalues of h on M is contained in $\{-2, 0, 2, 4, \ldots, 2p\}$, where -2 has multiplicity 1 as an eigenvalue and $0, 2, 2p$ have multiplicity at least 1. Now, if we had $p \geqslant 2$, then $-2p$ should also be an eigenvalue by Proposition 1.5.11, contradiction. So we have $p = 1$. But then the trace of h on M is $-2 + 2m$ where $m \geqslant 1$ is the multiplicity of 2 as an eigenvalue. By Proposition 1.5.11, that trace is 0 and so $m = 1$. Thus, we have shown that $\dim L_\lambda = 1$ and $n\lambda \notin P(L)$ for all $n \geqslant 2$.

Finally, since $[L_\lambda, L_{-\lambda}] \neq \{0\}$, we have $L_{-\lambda} \neq \{0\}$ and so we can repeat the whole argument with the roles of λ and $-\lambda$ reversed. Thus, we also have $\dim L_{-\lambda} = 1$ and $L_{-n\lambda} = \{0\}$ for all $n \geqslant 2$. □

Proposition 2.2.5. *Assume that the conditions in Definition 2.2.1 hold. Then, for each $i \in I$, we have*

$$\dim L_{\alpha_i} = \dim L_{-\alpha_i} = \dim[L_{\alpha_i}, L_{-\alpha_i}] = 1,$$

and there is a unique $h_i \in [L_{\alpha_i}, L_{-\alpha_i}]$ with $\alpha_i(h_i) = 2$. Furthermore, $\Delta = \{\alpha_i \mid i \in I\}$ is a basis of H^* and $\{h_i \mid i \in I\}$ is a basis of H.

Proof. Let I' be the set of all $i \in I$ such that the restriction of α_i to $[L_{\alpha_i}, L_{-\alpha_i}]$ is non-zero; in particular, $[L_{\alpha_i}, L_{-\alpha_i}] \neq \{0\}$ and $L_{\pm\alpha_i} \neq \{0\}$ for $i \in I'$. Now let us fix $i \in I'$. By Lemma 2.2.4, we have $\dim L_{\alpha_i} = \dim L_{-\alpha_i} = 1$. So there are elements $e_i \neq 0$ and $f_i \neq 0$ such that $L_{\alpha_i} = \langle e_i \rangle_{\mathbb{C}}$, $L_{-\alpha_i} = \langle f_i \rangle_{\mathbb{C}}$. Consequently, we have $[L_{\alpha_i}, L_{-\alpha_i}] = \langle h_i \rangle_{\mathbb{C}}$ where $0 \neq h_i := [e_i, f_i]$ and $\alpha_i(h_i) \neq 0$. So, replacing f_i by a scalar multiple if necessary, we can assume that $\alpha_i(h_i) = 2$; then h_i is uniquely determined (since $\dim[L_{\alpha_i}, L_{-\alpha_i}] = 1$). Thus, by (CK3), we have

$$H = H' + \langle h_i \mid i \in I' \rangle_{\mathbb{C}} \quad \text{where} \quad H' := \sum_{j \in I \setminus I'} [L_{\alpha_j}, L_{-\alpha_j}].$$

Now let $j \in I \setminus I'$. Then the restriction of α_j to $[L_{\alpha_j}, L_{-\alpha_j}]$ is zero and so Lemma 2.2.3 shows that $\mathrm{ad}_L(x) = 0$ for all $x \in [L_{\alpha_j}, L_{-\alpha_j}] \subseteq H$. On the other hand, if $x \in H$, then $\mathrm{ad}_L(x)$ is diagonalisable, with eigenvalues given by $\lambda(x)$ for $\lambda \in P(L)$. We conclude that, if $x \in [L_{\alpha_j}, L_{-\alpha_j}]$, then $\lambda(x) = 0$ for all $\lambda \in P(L)$. In particular, the restrictions of all α_i $(i \in I)$ to $[L_{\alpha_j}, L_{-\alpha_j}]$ are zero.

Assume, if possible, that $I' \subsetneqq I$. Then the restrictions of the linear maps α_i $(i \in I)$ to the subspace $\langle h_j \mid j \in I' \rangle_{\mathbb{C}}$ are linearly dependent. So there are scalars $c_i \in \mathbb{C}$, not all 0, such that $\sum_{i \in I} c_i \alpha_i(h_j) = 0$ for all $j \in I'$. But, we have just seen that $\alpha_i(x) = 0$ for all $x \in H'$. Hence, $\sum_{i \in I} c_i \alpha_i(x) = 0$ for all $x \in H$, contradiction to $\{\alpha_i \mid i \in I\}$ being linearly independent. So we must have $I' = I$, which shows that $H = \langle h_i \mid i \in I \rangle_{\mathbb{C}}$. On the other hand, since $\{\alpha_i \mid i \in I\}$ is linearly independent, we have $\dim H = \dim H^* \geqslant |I|$. Hence, $\{h_i \mid i \in I\}$ is a basis of H and $\{\alpha_i \mid i \in I\}$ is a basis of H^*. □

Definition 2.2.6. Assume that the conditions in Definition 2.2.1 hold. Let $h_i \in H$ $(i \in I)$ be as in Proposition 2.2.5. Then

$$A = \big(\alpha_j(h_i)\big)_{i,j \in I}$$

is called the *structure matrix* of L (with respect to Δ).

Note that, since $\{h_i \mid i \in I\}$ is a basis of H and $\{\alpha_i \mid i \in I\}$ is a basis of H^*, we certainly have $\det(A) \neq 0$.

Example 2.2.7. Let $L = \mathfrak{sl}_n(\mathbb{C})$ $(n \geqslant 2)$ and $H \subseteq L$ be the abelian subalgebra of all diagonal matrices in L; we have $\dim H = \dim H^* =$

$n-1$. By Remark 2.1.11, L is H-diagonalisable and $C_L(H) = H$. Thus, (CK1) holds. For $1 \leqslant i \leqslant n$, let $\varepsilon_i \in H^*$ be the map which sends a diagonal matrix to its i-th diagonal entry. (Note that, now, we have the linear relation $\varepsilon_1 + \ldots + \varepsilon_n = \underline{0}$.) For $i \neq j$ let $e_{ij} \in L$ be the matrix with entry 1 at position (i,j), and 0 everywhere else. Then we have again $L = H \oplus \bigoplus_{i \neq j} \mathbb{C}e_{ij}$. By the same computations as in Example 2.1.10, we see that $P(L) = \{\underline{0}\} \cup \Phi$, where

$$\Phi := \{\pm(\alpha_i + \alpha_{i+1} + \ldots + \alpha_{j-1}) \mid 1 \leqslant i < j \leqslant n\}$$

and $\alpha_i = \varepsilon_i - \varepsilon_{i+1}$ for $1 \leqslant i \leqslant n-1$. Thus, (CK2) holds, but we still need to check that $\{\alpha_1, \ldots, \alpha_{n-1}\} \subseteq H^*$ is linearly independent. For this purpose, let $U := \langle \alpha_1, \ldots, \alpha_{n-1} \rangle_{\mathbb{C}}$. If we had $U \subsetneqq H^*$, then $U^\circ := \{x \in H \mid \lambda(x) = 0 \text{ for all } \lambda \in U\} \neq \{0\}$, by standard duality properties in Linear Algebra. Let $0 \neq x \in U^\circ$. Then $\alpha_1(x) = 0$ and so the first two diagonal entries of x are equal. Next, since $\alpha_2(x) = 0$, the second and third diagonal entries are equal. Hence, we conclude that all diagonal entries are equal and so $\mathrm{Trace}(x) \neq 0$, contradiction. Hence, since $\dim H^* = n-1$, the set $\{\alpha_1, \ldots, \alpha_{n-1}\}$ is a basis of H^*. Given the above description of Φ, this now shows that $|\Phi| = n^2 - n$, and so a dimension argument as in Example 2.1.10 yields that

$$L_{\underline{0}} = H \qquad \text{and} \qquad \dim L_\alpha = 1 \quad \text{for all } \alpha \in \Phi.$$

Finally, we set $e_i := e_{i,i+1} \in L_{\alpha_i}$ and $f_i := e_{i+1,i} \in L_{-\alpha_i}$ for $1 \leqslant i \leqslant n-1$. Then $h_i := [e_i, f_i] \in H$ is the diagonal matrix with entries $1, -1$ at positions $i, i+1$ (and 0 otherwise). We see that $\{h_1, \ldots, h_{n-1}\}$ is a basis of H and, hence, that $H = \sum_{1 \leqslant i \leqslant n-1} [L_{\alpha_i}, L_{-\alpha_i}]$. Thus, (CK3) also holds and so (L, H) is of Cartan–Killing type with respect to $\Delta = \{\alpha_1, \ldots, \alpha_{n-1}\}$. We compute that

$$A = \big(\alpha_j(h_i)\big) = \begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix} \in M_{n-1}(\mathbb{Z})$$

where all non-specified entries are 0. Note that $h_i \in [L_{\alpha_i}, L_{-\alpha_i}]$ and $\alpha_i(h_i) = 2$. Hence, the above elements $\{h_1, \ldots, h_{n-1}\}$ are indeed the elements whose existence and uniqueness is proved in Proposition 2.2.5. We know that $\det(A) \neq 0$ but we leave it as an exercise to compute that $\det(A) = n$.

Assume from now on that (L, H) is of Cartan–Killing type with respect to $\Delta = \{\alpha_i \mid i \in I\}$, as in Definition 2.2.1.

Lemma 2.2.8. *Let $\alpha \in \Phi^+$ and $i \in I$. If $\alpha + m\alpha_i \in \Phi$ for some $m \in \mathbb{Z}$, then $\alpha = \alpha_i$ or $\alpha + m\alpha_i \in \Phi^+$.*

Proof. Write $\alpha = \sum_{j \in I} n_j \alpha_j$ where $n_j \in \mathbb{Z}_{\geqslant 0}$ for all j. Assume that $\alpha \neq \alpha_i$; since $\alpha \in \Phi^+$, we also have $\alpha \neq -\alpha_i$. By Proposition 2.2.5, the restriction of α_i to $[L_{\alpha_i}, L_{-\alpha_i}]$ is non-zero and so Lemma 2.2.4 implies that $\alpha \notin \mathbb{Z}\alpha_i$. Hence, we must have $n_{i_0} > 0$ for some $i_0 \neq i$. But then $n_{i_0} > 0$ is also the coefficient of α_{i_0} in $\alpha + m\alpha_i$. Since every root is either in Φ^+ or in Φ^-, we conclude that $\alpha + m\alpha_i \in \Phi^+$. □

Remark 2.2.9. Let $i \in I$ and $h_i \in [L_{\alpha_i}, L_{-\alpha_i}]$ be as in Proposition 2.2.5. Let $e_i \in L_{\alpha_i}$ and $f_i \in L_{-\alpha_i}$ be such that $h_i = [e_i, f_i]$. Since $\dim L_{\pm\alpha_i} = 1$, we have $L_{\alpha_i} = \langle e_i \rangle_{\mathbb{C}}$ and $L_{-\alpha_i} = \langle f_i \rangle_{\mathbb{C}}$. Furthermore, since $\alpha_i(h_i) = 2$, we have $[h_i, e_i] = 2e_i$ and $[h_i, f_i] = -2f_i$. Thus, $S_i := \langle e_i, h_i, f_i \rangle_{\mathbb{C}} \subseteq L$ is a 3-dimensional subalgebra isomorphic to $\mathfrak{sl}_2(\mathbb{C})$. We call $\{e_i, h_i, f_i\}$ an $\mathfrak{sl}_2$-*triple* in L. This will provide a powerful tool in the study of L. The elements $\{e_i, f_i \mid i \in I\}$ are called *Chevalley generators* of L. Note that the f_i are determined once the e_i are chosen (via the relations $h_i = [e_i, f_i]$); the e_i are only unique up to non-zero scalar multiples.

Remark 2.2.10. In the proof of Lemma 2.2.4, we used the results on representations of $\mathfrak{sl}_2(\mathbb{C})$ that we obtained in Section 1.5. We can now push this argument much further. So let us fix $i \in I$ and let $\{e_i, h_i, f_i\}$ be a corresponding $\mathfrak{sl}_2$-triple, as above. Then $\mathfrak{sl}_2(\mathbb{C}) \cong S_i := \langle e_i, h_i, f_i \rangle_{\mathbb{C}} \subseteq L$. Let us also fix $\beta \in \Phi$ such that $\beta \neq \pm\alpha_i$. Since Φ is finite, there are well-defined integers $p, q \geqslant 0$ such that

$$\beta - q\alpha_i, \quad \ldots, \quad \beta - \alpha_i, \quad \beta, \quad \beta + \alpha_i, \quad \ldots, \quad \beta + p\alpha_i$$

are all contained in Φ, but $\beta + (p+1)\alpha_i \notin \Phi$ and $\beta - (q+1)\alpha_i \notin \Phi$. (It could be that $p = 0$ or $q = 0$.) The above sequence of roots is called the α_i-*string through* β. Now consider the subspace

$$M := L_{\beta - q\alpha_i} \oplus \ldots \oplus L_{\beta - \alpha_i} \oplus L_\beta \oplus L_{\beta + \alpha_i} \oplus \ldots \oplus L_{\beta + p\alpha_i} \subseteq L.$$

We claim that M is an S_i-submodule of L. Now, we certainly have $[H, M] \subseteq M$ and so M is invariant under h_i. By Proposition 2.1.6, we have $[L_{\pm\alpha_i}, L_{\beta+n\alpha_i}] \subseteq L_{\beta+(n\pm1)\alpha_i}$ for all $n \in \mathbb{Z}$. This shows that all subspaces $L_{\beta+n\alpha_i}$ with $-q < n < p$ are invariant under e_i and f_i. Finally, by Lemma 2.2.4 (applied to $\lambda = \alpha_i$), we have $\beta \neq n\alpha_i$ for all $n \in \mathbb{Z}$. Hence, $\underline{0} \neq \beta + (p+1)\alpha_i \notin \Phi$ and so $[L_{\alpha_i}, L_{\beta+p\alpha_i}] \subseteq L_{\beta+(p+1)\alpha_i} = \{0\}$. Similarly, we have $[L_{-\alpha_i}, L_{\beta-q\alpha_i}] \subseteq L_{\beta-(q+1)\alpha_i} = \{0\}$. Thus, M is an S_i-submodule of L, as claimed. Now recall that the module action is given by $\mathrm{ad}_L \colon L \to \mathfrak{gl}(L)$. Since L is H-diagonalisable, the eigenvalues of any $x \in H$ are given by $\lambda(x)$ for $\lambda \in P(L)$ (each with multiplicity $\dim L_\lambda \geqslant 1$). So the eigenvalues of h_i on M are given by $(\beta + n\alpha_i)(h_i)$ for $-q \leqslant n \leqslant p$, each with multiplicity $\dim L_{\beta+n\alpha_i} \geqslant 1$. Explicitly, the list of eigenvalues (not counting multiplicities) is given by

$$\beta(h_i) - 2q,\ \ldots,\ \beta(h_i) - 2,\ \beta(h_i),\ \beta(h_i) + 2,\ \ldots,\ \beta(h_i) + 2p.$$

By Proposition 1.5.11, all eigenvalues of h_i are integers, and if $m \in \mathbb{Z}$ is an eigenvalue, then so is $-m$. In particular, the largest eigenvalue is the negative of the smallest eigenvalue. First of all, this implies that $\beta(h_i) + 2p = -(\beta(h_i) - 2q)$ and so

(a) $$\beta(h_i) = q - p \in \mathbb{Z}.$$

Furthermore, $-q \leqslant p - q = -\beta(h_i) \leqslant p$. Thus, we conclude that

(b) $\beta - \beta(h_i)\alpha_i \in \Phi$ belongs to the α_i-string through β.

We can go even one step further. Let $0 \neq v^+ \in L_{\beta+p\alpha_i}$ be fixed. Then $h_i.v^+ = cv^+$ where $c = \beta(h_i) + 2p = (q-p) + 2p = p + q$. Since $[e_i, v^+] \in L_{\beta+(p+1)\alpha_i} = \{0\}$, we have $e_i.v^+ = \{0\}$ and so $v^+ \in M$ is a *primitive vector*, as in Remark 1.5.9. Correspondingly, we have a subspace $E := \langle v_n \mid n \geqslant 0\rangle_{\mathbb{C}} \subseteq M$, where

$$v_0 := v^+ \quad \text{and} \quad v_{n+1} := \tfrac{1}{n+1}[f_i, v_n] \quad \text{for all } n \geqslant 0.$$

(We also set $v_{-1} := 0$.) As shown in Remark 1.5.9, we have

$$\dim E = c + 1 = p + q + 1 \quad \text{and} \quad E = \langle v_0, v_1, \ldots, v_{p+q}\rangle_{\mathbb{C}}.$$

In particular, $v_0, v_1, \ldots, v_{p+q}$ are all non-zero. We can exploit this as follows. First, $v_0 = v^+ \in L_{\beta+p\alpha_i}$. Hence, if $p \geqslant 1$, then $v_1 = [f_i, v_0] \in [L_{-\alpha_i}, L_{\beta+p\alpha_i}] \subseteq L_{\beta+(p-1)\alpha_i}$; furthermore, if $p \geqslant 2$, then

$v_2 = \frac{1}{2}[f_i, v_1] \in [L_{-\alpha_i}, L_{\beta+(p-1)\alpha_i}] \subseteq L_{\beta+(p-2)\alpha_i}$. Going on in this way, we find that $0 \neq v_p \in L_\beta$. Since $[e_i, v_p] = (c-p+1) = (q+1)v_{p-1}$ (see Remark 1.5.9), we conclude that

$$\text{(c)} \qquad \begin{aligned} [f_i,[e_i,v_p]] &= (q+1)\,[f_i,v_{p-1}] = p(q+1)v_p,\\ [e_i,[f_i,v_p]] &= (p+1)\,[e_i,v_{p+1}] = q(p+1)v_p. \end{aligned}$$

In particular, since $0 \neq v_p \in L_\alpha$, this implies that

$$\text{(c')} \qquad \begin{aligned} &\{0\} \neq [L_{\alpha_i}, L_\beta] \subseteq L_{\beta+\alpha_i} && \text{if } p>0, \text{ that is, } \beta+\alpha_i \in \Phi,\\ &\{0\} \neq [L_{-\alpha_i}, L_\beta] \subseteq L_{\beta-\alpha_i} && \text{if } q>0, \text{ that is, } \beta-\alpha_i \in \Phi. \end{aligned}$$

Remark 2.2.11. For future reference, we note that $\beta(h_i) \in \mathbb{Z}$ for all $\beta \in \Phi$ and all $i \in I$. Indeed, if $\beta \neq \pm\alpha_i$, then this holds by Remark 2.2.10(a). But if $\beta = \pm\alpha_i$, then $\beta(h_i) = \pm\alpha_i(h_i) = \pm 2$.

Corollary 2.2.12. *Consider the matrix $A = (a_{ij})_{i,j\in I}$ in Definition 2.2.6, where $a_{ij} = \alpha_j(h_i)$ for $i, j \in I$. Then the following hold.*

(a) *$a_{ij} \in \mathbb{Z}$ and $a_{ii} = 2$ for all $i, j \in I$.*

(b) *$a_{ij} \leqslant 0$ for all $i, j \in I$, $i \neq j$.*

(c) *$a_{ij} \neq 0 \Leftrightarrow a_{ji} \neq 0$ for all $i, j \in I$.*

Proof. (a) See Proposition 2.2.5 and Remark 2.2.11.

(b) Assume, if possible, that $a_{ij} > 0$. Then, by Remark 2.2.10(b), we have $\alpha_j - n\alpha_i \in \Phi$, where $n = \alpha_j(h_i) > 0$, contradiction to (CK2).

(c) This is clear for $i = j$. Now assume that $i \neq j$ and $a_{ji} \neq 0$; then $a_{ji} < 0$ by (b). By Remark 2.2.10(b), we have $\alpha_i + n\alpha_j \in \Phi$, where $n = -\alpha_i(h_j) = -a_{ji} > 0$; furthermore, $\alpha_i + n\alpha_j$ belongs to the α_j-string through α_i. Hence, since $n > 0$, we also have that $\alpha_i + \alpha_j \in \Phi$ belongs to that α_j-string. Now we reverse the roles of α_i and α_j and consider the α_i-string through α_j. Let $p, q \geqslant 0$ in Remark 2.2.10 be defined with respect to α_i and $\alpha := \alpha_j$. Since $\alpha_j + \alpha_i \in \Phi$, we have $p \geqslant 1$. By (CK2), we have $\alpha_j - \alpha_i \notin \Phi$ and so $q = 0$. Hence, Remark 2.2.10(a) shows that $a_{ij} = \alpha_j(h_i) = -p < 0$. □

Exercise 2.2.13. In the setting of Remark 2.2.10, show that $p = \max\{n \geqslant 0 \mid \beta + n\alpha_i \in \Phi\}$ and $q = \max\{n \geqslant 0 \mid \beta - n\alpha_i \in \Phi\}$. Deduce that, if $\beta \pm n\alpha_i \in \Phi$ for some $n > 0$, then $\beta \pm \alpha_i \in \Phi$.

2.3. The Weyl group

ab hier Woche 5

We keep the basic setting of the previous section, where (L, H) is of Cartan–Killing type with respect to $\Delta = \{\alpha_i \mid i \in I\} \subseteq H^*$. The formula in Remark 2.2.10(b) suggests the following definition.

Definition 2.3.1. For $i \in I$, let $h_i \in [L_{\alpha_i}, L_{-\alpha_i}]$ be as in Propositon 2.2.5. We define a linear map $s_i \colon H^* \to H^*$ by

$$s_i(\lambda) := \lambda - \lambda(h_i)\alpha_i \qquad \text{for } \lambda \in H^*.$$

Note that $s_i(\alpha_i) = \alpha_i - 2\alpha_i = -\alpha_i$ and $s_i(\lambda) = \lambda$ for all $\lambda \in H^*$ with $\lambda(h_i) = 0$. Since $H^* = \langle \alpha_i \rangle_{\mathbb{C}} \oplus \{\lambda \in H^* \mid \lambda(h_i) = 0\}$, we conclude that s_i is diagonalisable, with one eigenvalue equal to -1 and $|I| - 1$ eigenvalues equal to 1. In particular, $s_i^2 = \mathrm{id}_{H^*}$, $\det(s_i) = -1$ and $s_i \in \mathrm{GL}(H^*)$. The subgroup

$$W := \langle s_i \mid i \in I \rangle \subseteq \mathrm{GL}(H^*)$$

is called the *Weyl group* of L (with respect to Δ). Note that, since $s_i^{-1} = s_i$ for all $i \in I$, every element $w \in W$ can be written as a product $w = s_{i_1} \cdots s_{i_r}$ where $r \geqslant 0$ and $i_1, \ldots, i_r \in I$. (Such an expression for w is by no means unique; we have $w = \mathrm{id}$ if $r = 0$.)

Remark 2.3.2. By Remark 2.2.10, we have $s_i(\alpha) \in \Phi$ for all $\alpha \in \Phi$ with $\alpha \neq \pm\alpha_i$. But we also have $s_i(\alpha_i) = -\alpha_i$ and so $s_i(\Phi) = \Phi$. Consequently, we have $w(\Phi) = \Phi$ for all $w \in W$. So we have an action of the group W on the finite set Φ via

$$W \times \Phi \to \Phi, \qquad (w, \alpha) \mapsto w(\alpha).$$

Let $\mathrm{Sym}(\Phi)$ denote the symmetric group on Φ. Then we obtain a group homomorphism $\pi \colon W \to \mathrm{Sym}(\Phi)$, $w \mapsto \pi_w$, where $\pi_w(\alpha) := w(\alpha)$ for all $\alpha \in \Phi$. If $\pi_w = \mathrm{id}_\Phi$, then $w(\alpha) = \alpha$ for all $\alpha \in \Phi$. In particular, $w(\alpha_i) = \alpha_i$ for all $i \in I$. Since $\{\alpha_i \mid i \in I\}$ is a basis of H^*, it follows that $w = \mathrm{id}_{H^*}$. Thus, π is injective and W is isomorphic to a subgroup of $\mathrm{Sym}(\Phi)$; in particular, W is a finite group.

In order to prove the "Key Lemma" below, we shall use a construction that essentially relies on the fact that W is a finite group. For this purpose, let $E := \langle \alpha_i \mid i \in I \rangle_{\mathbb{R}} \subseteq H^*$. Then E is an $\mathbb{R}$-vector space, and $\{\alpha_i \mid i \in I\}$ still is a basis of E. By (CK2), we have $\Phi \subseteq E$. Since $\alpha(h_i) \in \mathbb{Z}$ for all $\alpha \in \Phi$ and $i \in I$ (see Remark 2.2.11), we also have $s_i(E) \subseteq E$ for all $i \in I$ and so $w(E) \subseteq E$ for all $w \in W$. Thus, we may regard W as a subgroup of $\mathrm{GL}(E)$ (but we will not introduce a separate notation for this). Let $\langle\ ,\ \rangle_0\colon E \times E \to \mathbb{R}$ be the standard scalar product for which $\{\alpha_i \mid i \in I\}$ is an orthonormal basis. Thus, for $v, v' \in E$ we have $\langle v, v'\rangle_0 = \sum_{i,j\in I} x_i x_j'$ where $v = \sum_{i\in I} x_i\alpha_i$ and $v' = \sum_{j\in I} x_j'\alpha_j$, with $x_i, x_j' \in \mathbb{R}$ for all $i, j \in I$. Then we define a new map $\langle\ ,\ \rangle\colon E \times E \to \mathbb{R}$ by

$$\langle v, v'\rangle := \sum_{w\in W} \langle w(v), w(v')\rangle_0 \qquad \text{for } v, v' \in E.$$

Since $E \to E$, $v \mapsto w(v)$, is linear for each $w \in W$, it is clear that $\langle\ ,\ \rangle$ is a symmetric bilinear form. For $v \in E$, we have

$$\langle v, v\rangle = \sum_{w\in W} \underbrace{\langle w(v), w(v)\rangle_0}_{\geqslant 0} \geqslant 0.$$

If $\langle v, v\rangle = 0$, then $\langle w(v), w(v)\rangle_0 = 0$ for all $w \in W$. In particular, this holds for $w = \mathrm{id}_E$ and so $\langle v, v\rangle_0 = 0$. But $\langle\ ,\ \rangle_0$ is positive-definite and so $v = 0$. Thus, $\langle\ ,\ \rangle$ is also positive-definite. Finally, taking the sum over all $w \in W$ implies the following invariance property:

$$\langle w(v), w(v')\rangle = \langle v, v'\rangle \qquad \text{for all } w \in W \text{ and } v, v' \in E.$$

Indeed, for a fixed $w \in W$, we have

$$\langle w(v), w(v')\rangle = \sum_{w'\in W} \langle w'w(v), w'w(v')\rangle_0.$$

Now, since W is a group, the map $W \to W$, $w' \mapsto w'w$, is a bijection. Hence, up to reordering terms, the sum on the right hand side is the same as the sum in the definition of $\langle v, v'\rangle$.

Remark 2.3.3. Let $i \in I$ and $\lambda \in E$; recall that $E = \langle \alpha_i \mid i \in I\rangle_{\mathbb{R}} \subseteq H^*$. Using the relation $s_i(\alpha_i) = -\alpha_i$, the defining formula for $s_i(\lambda)$, and the above invariance property, we obtain the following identities:

$$\begin{aligned} -\langle \alpha_i, \lambda\rangle &= \langle s_i(\alpha_i), \lambda\rangle = \langle s_i^2(\alpha_i), s_i(\lambda)\rangle = \langle \alpha_i, s_i(\lambda)\rangle \\ &= \langle \alpha_i, \lambda - \lambda(h_i)\alpha_i\rangle = \langle \alpha_i, \lambda\rangle - \lambda(h_i)\langle \alpha_i, \alpha_i\rangle. \end{aligned}$$

Since $\langle \alpha_i, \alpha_i\rangle \in \mathbb{R}_{>0}$, this yields the fomula

$$\lambda(h_i) = 2\frac{\langle \alpha_i, \lambda\rangle}{\langle \alpha_i, \alpha_i\rangle} \qquad \text{for all } \lambda \in E \text{ and } i \in I.$$

This formula shows that each $s_i\colon E \to E$ is a *reflection* with root α_i.

Lemma 2.3.4 (Key Lemma)**.** *Let $\alpha \in \Phi^+$ but $\alpha \notin \Delta$. Then there exists some $i \in I$ such that $\alpha(h_i) \in \mathbb{Z}_{>0}$. Furthermore, we have $s_i(\alpha) = \alpha - \alpha(h_i)\alpha_i \in \Phi^+$ and $\alpha - \alpha_i \in \Phi^+$.*

Proof. We write $\alpha = \sum_{i \in I} n_i \alpha_i$ where $n_i \in \mathbb{Z}_{\geqslant 0}$ for all i. Since $\underline{0} \neq \alpha \in E$, we can apply the above discussion and obtain

$$\sum_{i \in I} n_i \underbrace{\langle \alpha_i, \alpha \rangle}_{\in \mathbb{R}} = \langle \alpha, \alpha \rangle > 0.$$

Since $n_i \geqslant 0$ for all i, there must be some $i \in I$ such that $n_i > 0$ and $\langle \alpha_i, \alpha \rangle > 0$. Furthermore, since $\langle \alpha_i, \alpha_i \rangle > 0$, the formula in Remark 2.3.3 shows that we also have $\alpha(h_i) > 0$. By Remark 2.2.11, $\alpha(h_i) \in \mathbb{Z}$ and so $\alpha(h_i) \in \mathbb{Z}_{>0}$, as desired. Now, since $\alpha \in \Phi^+ \setminus \Delta$, we have $\alpha \neq \pm\alpha_i$. Hence, Remark 2.2.10(b) shows that $\alpha - \alpha(h_i)\alpha_i \in \Phi$ belongs to the α_i-string through α. Since $\alpha(h_i) \in \mathbb{Z}_{>0}$, we conclude that $\alpha - \alpha_i$ also belongs to that α_i-string and so $\alpha - \alpha_i \in \Phi$. It remains to show that $\alpha - \alpha_i \in \Phi^+$ and $\alpha - \alpha(h_i)\alpha_i \in \Phi^+$. But this follows from Lemma 2.2.8, since $\alpha \neq \alpha_i$. □

Remark 2.3.5. Since $\{\alpha_i \mid i \in I\}$ is a basis of H^*, we can define a linear map $\mathrm{ht}\colon H^* \to \mathbb{C}$ by $\mathrm{ht}(\alpha_i) := 1$ for $i \in I$. Let $\alpha \in \Phi$ and write $\alpha = \sum_{i \in I} n_i \alpha_i$ where $n_i \in \mathbb{Z}$ for all i. Then $\mathrm{ht}(\alpha) = \sum_{i \in I} n_i \in \mathbb{Z}$ is called the *height* of α. Since $\Phi = \Phi^+ \sqcup \Phi^-$, we have

$$\mathrm{ht}(\alpha) = 1 \Leftrightarrow \alpha \in \Delta;\ \mathrm{ht}(\alpha) \geqslant 1 \Leftrightarrow \alpha \in \Phi^+;\ \mathrm{ht}(\alpha) \leqslant -1 \Leftrightarrow \alpha \in \Phi^-.$$

The "Key Lemma" often allows us to argue by induction on the height of roots; here is a first example. Let $\alpha \in \Phi^+$ and $n = \mathrm{ht}(\alpha) \geqslant 1$. Then we can write $\alpha = \alpha_{i_1} + \ldots + \alpha_{i_n}$ where $i_j \in I$ for all j and, for each $j \in \{1, \ldots, n\}$, we also have $\alpha_{i_j} + \ldots + \alpha_{i_n} \in \Phi^+$.

We argue by induction on $n := \mathrm{ht}(\alpha) \geqslant 1$. If $n = 1$, then $\alpha = \alpha_i$ for some $i \in I$ and we are done. Now let $n \geqslant 2$. Then $\alpha \notin \Delta$ and so, by Lemma 2.3.4, we have $\beta := \alpha - \alpha_{i_1} \in \Phi^+$ for some $i_1 \in I$. Now $\mathrm{ht}(\beta) = n - 1$. By induction, there exist $i_2, \ldots, i_n \in I$ such that the required conditions hold for β. But then $\alpha = \alpha_{i_1} + \alpha_{i_2} + \ldots + \alpha_{i_n}$ and the required conditions hold for α.

Theorem 2.3.6. *Recall that (L, H) is of Cartan–Killing type with respect to $\Delta = \{\alpha_i \mid i \in I\}$. Then the following hold.*

(a) $\Phi = \{w(\alpha_i) \mid w \in W, i \in I\}$ *and* $\Phi^- = -\Phi^+$.

(b) *If* $\alpha \in \Phi$ *and* $0 \neq c \in \mathbb{C}$ *are such that* $c\alpha \in \Phi$, *then* $c = \pm 1$.

Proof. (a) Let $\Phi_0 := \{w(\alpha_i) \mid w \in W, i \in I\}$. By Remark 2.3.2, $\Phi_0 \subseteq \Phi$. Next, let $\alpha \in \Phi^+$. We show by induction on $n := \mathrm{ht}(\alpha) \geqslant 1$ that $\alpha \in \Phi_0$. If $n = 1$, then $\alpha = \alpha_i$ for some $i \in I$ and so $\alpha = \mathrm{id}(\alpha_i) \in \Phi_0$. Now let $n \geqslant 2$. By Lemma 2.3.4, there is some $j \in I$ such that $\alpha(h_j) \in \mathbb{Z}_{>0}$ and $\beta := s_j(\alpha) = \alpha - \alpha(h_j)\alpha_j \in \Phi^+$. We have $\mathrm{ht}(\beta) = n - \alpha(h_j) < n$. By induction, $\beta \in \Phi_0$ and so $\beta = w'(\alpha_i)$ for some $w' \in W$ and $i \in I$. But then $\alpha = s_j^2(\alpha) = s_j\big(s_j(\alpha)\big) = s_j(\beta) = s_j w'(\alpha_i) \in \Phi_0$, as required. Thus, we have shown that $\Phi^+ \subseteq \Phi_0$.

Next, let $\alpha \in \Phi^+$. Since $\alpha \in \Phi_0$, we can write $\alpha = w(\alpha_i)$, where $w \in W$ and $i \in I$, as above. Since $s_i(\alpha_i) = -\alpha_i$, we obtain $-\alpha = w(-\alpha_i) = ws_i(\alpha_i) \in \Phi_0 \subseteq \Phi$. Furthermore, since $\alpha \in \Phi^+$, we have $-\alpha \in \Phi^-$. Thus, we have shown that $-\Phi^+ \subseteq \Phi^- \cap \Phi_0$.

Now, there is a symmetry in Definition 2.2.1. If we set $\alpha_i' := -\alpha_i$ for all $i \in I$, then (L, H) also is of Cartan–Killing type with respect to $\Delta' := \{\alpha_i' \mid i \in I\}$. Then, clearly, $\Psi^+ := \Phi^-$ is the corresponding set of positive roots and $\Psi^- := \Phi^+$ is the set of negative roots. Now, the previous argument applied to Δ' instead of Δ shows that $-\Phi^- = -\Psi^+ \subseteq \Psi^- \cap \Phi_0 = \Phi^+ \cap \Phi_0 \subseteq \Phi^+$ and, hence, $\Phi^- \subseteq -\Phi^+ \subseteq \Phi_0$. Consequently, $\Phi = \Phi^+ \cup \Phi^- \subseteq \Phi_0$ and, hence, $\Phi = \Phi_0$.

(b) Assume that $\alpha \in \Phi$ and $c\alpha \in \Phi$, where $0 \neq c \in \mathbb{C}$. By (a) we can write $\alpha = w(\alpha_i)$ for some $w \in W$ and $i \in I$. Then $c\alpha_i = cw^{-1}(\alpha) = w^{-1}(c\alpha) \in \Phi$ and so $c\alpha_i(h_i) \in \mathbb{Z}$ by Remark 2.2.11. But $\alpha_i(h_i) = 2$ and so $2c \in \mathbb{Z}$; thus, $c\alpha_i \in \Phi$, where $c = n/2$ with $n \in \mathbb{Z}$. On the other hand, we also have $\beta := c\alpha \in \Phi$ and $c^{-1}\beta = \alpha \in \Phi$. Hence, a similar argument shows that $c^{-1}\alpha_j \in \Phi$ for some $j \in I$, where $c^{-1} = m/2$ for some $m \in \mathbb{Z}$. Thus, we have $nm = 4$. If $m = \pm 1$, then $n = \pm 4$ and so $c = \pm 2$; hence, $\pm 2\alpha_i \in \Phi$, contradiction to Lemma 2.2.4 (applied to $\lambda = \alpha_i$). Similarly, if $n = \pm 1$, then $m = \pm 4$ and so $c^{-1} = \pm 2$; hence, $\pm 2\alpha_j \in \Phi$, contradiction to Lemma 2.2.4 (applied to $\lambda = \alpha_j$). Thus, we must have $n = \pm 2$ and so $c = \pm 1$. $\square$

We would like to make it completely explicit that W and Φ are determined by the single knowledge of the structure matrix A of L.

Remark 2.3.7. Recall that $A = (a_{ij})_{i,j\in I}$, where $a_{ij} = \alpha_j(h_i) \in \mathbb{Z}$ for all $i, j \in I$. Thus, the defining equation of s_i yields that

$$s_i(\alpha_j) = \alpha_j - a_{ij}\alpha_i \quad \text{for all } i, j \in I.$$

Hence, if $\lambda \in H^*$ and $\lambda = \sum_{i\in I} \lambda_i\alpha_i \in H^*$ with $\lambda_i \in \mathbb{C}$, then we have

$$(\clubsuit) \qquad s_i(\lambda) = \sum_{j\in I} \lambda_j\big(\alpha_j - a_{ij}\alpha_i\big) = \lambda - \Big(\sum_{j\in I} a_{ij}\lambda_j\Big)\alpha_i.$$

This shows that the action of s_i on H^* is completely determined by A. For each $w \in W$, let $M_w \in \mathrm{GL}_I(\mathbb{C})$ be the matrix of w with respect to the basis $\{\alpha_i \mid i \in I\}$ of H^*. We have $w = s_{i_1} \cdots s_{i_l}$ for some $i_1, \ldots, i_l \in I$ and, hence, also $M_w = M_{s_{i_1}} \cdot \ldots \cdot M_{s_{i_l}}$. The above formulae show that each M_{s_i} is completely determined by A, and has entries in $\mathbb{Z}$. Hence, the set of matrices $\{M_w \mid w \in W\} \subseteq \mathrm{GL}_I(\mathbb{Z})$ is also determined by A. Finally, by Theorem 2.3.6(a), every $\alpha \in \Phi$ can be written as $\alpha := w(\alpha_i)$ where $w \in W$ and $i \in I$. Then $\alpha = \sum_{i\in I} n_i\alpha_i$ where $(n_i)_{i\in I} \in \mathbb{Z}^I$ is the i-th column of M_w. Thus,

$$\mathscr{C}(A) := \Big\{(n_i)_{i\in I} \in \mathbb{Z}^I \;\Big|\; \sum_{i\in I} n_i\alpha_i \in \Phi\Big\} \subseteq \mathbb{Z}^I$$

is completely determined by A. More concretely, every $\alpha \in \Phi$ is obtained by repeatedly applying the generators s_j of W to the various α_i, using formula ($\clubsuit$). If, in the process, we avoid the relation $s_i(\alpha_i) = -\alpha_i$, then we just obtain the set

$$\mathscr{C}^+(A) := \Big\{(n_i)_{i\in I} \in \mathbb{Z}_{\geqslant 0}^I \;\Big|\; \sum_{i\in I} n_i\alpha_i \in \Phi^+\Big\} \subseteq \mathbb{Z}^I.$$

(See the proof of Theorem 2.3.6.) Here are a few examples.

Example 2.3.8. Let $L = \mathfrak{sl}_3(\mathbb{C})$, where $\Delta = \{\alpha_1, \alpha_2\}$ and

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}; \qquad \text{see Example 2.2.7.}$$

The matrices of $s_1, s_2 \in W$ with respect to the basis Δ are given by:

$$s_1 : \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad s_2 : \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix};$$

see ($\clubsuit$). A direct computation shows that the product $s_1s_2 \in W$ has order 3 and so $W \cong \mathfrak{S}_3$. Applying s_1, s_2 repeatedly to α_1, α_2

(avoiding $w_i(\alpha_i) = -\alpha_i$ for $i = 1, 2$), we obtain that

$$\mathscr{C}^+(A) = \{(1,0), (0,1), (1,1)\} \qquad \text{or} \qquad \Phi^+ = \{\alpha_1, \alpha_2, \alpha_1 + \alpha_2\}$$

which is, of course, consistent with the general description of the set of roots Φ for $\mathfrak{sl}_n(\mathbb{C})$, $n \geqslant 2$, in Example 2.2.7.

Example 2.3.9. Let $L = \mathfrak{go}_4(Q_4, \mathbb{C})$ where $Q_4^{\text{tr}} = -Q_4$, as in Section 1.6. We will see later that L is of Cartan–Killing type with respect to a set $\Delta = \{\alpha_1, \alpha_2\}$ and structure matrix

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

Using (♣), the matrices of $s_1, s_2 \in W$ with respect to Δ are:

$$s_1 : \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad s_2 : \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix}.$$

Now $s_1 s_2 \in W$ has order 4 and W consists of 8 elements with matrices:

$$\pm\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \pm\begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}, \pm\begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix}, \pm\begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix}.$$

As above, we obtain that $\mathscr{C}^+(A) = \{\,(1,0),\ (0,1),\ (1,1),\ (1,2)\,\}$. Of course, this will turn out to be consistent with the general description of the set of roots Φ for $\mathfrak{go}_n(Q_n, \mathbb{C})$ (to be obtained later).

Example 2.3.10. Consider the matrix $A = \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix}$. We have not yet seen a corresponding Lie algebra but we can just formally apply the above procedure, where $\{\alpha_1, \alpha_2\}$ denotes the standard basis of $\mathbb{C}^2$. Using (♣), the matrices of $s_1, s_2 \in \mathrm{GL}_2(\mathbb{C})$ are:

$$s_1 : \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad s_2 : \begin{pmatrix} 1 & 0 \\ 3 & -1 \end{pmatrix}.$$

The product $s_1 s_2$ has order 6 and so $\langle s_1, s_2 \rangle \subseteq \mathrm{GL}_2(\mathbb{C})$ is a dihedral group of order 12. Applying s_1, s_2 repeatedly to α_1, α_2 (avoiding $s_i(\alpha_i) = -\alpha_i$ for $i = 1, 2$), we find the following set $\mathscr{C}^+(A)$:

$$\{\ (1,0), \quad (0,1), \quad (1,1), \quad (1,2), \quad (1,3), \quad (2,3)\ \}$$

(or $\{\alpha_1, \alpha_2, \alpha_1 + \alpha_2, \alpha_1 + 2\alpha_2, \alpha_1 + 3\alpha_2, 2\alpha_1 + 3\alpha_2\} \subseteq \mathbb{C}^2$).

Table 1. A Python program for computing $\mathscr{C}^+(A)$

```
>>> def refl(A,n,r,i):          # apply s_i to root r
...   nr=r[:]                   # make a copy of the root r
...   nr[i]-=sum(A[i][j]*nr[j] for j in range(n))
...   return nr
>>> def rootsystem(A):             # A=structure matrix
...   n=len(A)
...   R=[[0]*n for i in range(n)]  # initialise R with
...   for i in range(n):           # unit basis vectors
...     R[i][i]=1
...   for r in R:
...     for i in range(n):
...       if R[i]!=r:          # avoid s_i(alpha_i)=-alpha_i
...         nr=refl(A,n,r,i)   # apply s_i to r
...         if not nr in R:    # check if we get something new
...           R.append(nr)
...   R.sort(reverse=True)     # sort list nicely
...   R.sort(key=sum)
...   return R
>>> rootsystem([[2, -1], [-3, 2]])  # see Example 2.3.10
[[1, 0], [0, 1], [1, 1], [1, 2], [1, 3], [2, 3]]
```

The above examples illustrate how $\Phi = \Phi^+ \cup (-\Phi^+)$ can be computed by a purely mechanical procedure from the structure matrix A. In fact, we do not have to do this by hand, but we can simply write a computer program for this purpose. Table 1 contains such a program written in the Python language; it outputs the set $\mathscr{C}^+(A)$. (The function `refl(A,|I|,r,i)` implements the formula (♣) in Remark 2.3.7.) If we apply the program to an arbitrary matrix A, then it will either return some nonsense or run into an infinite loop.

Exercise 2.3.11. Of course, the above procedure will not work with any integer matrix A, even if the entries of A satisfy the various conditions that we have seen so far. For example, let A be

$$\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 2 & -1 & 0 \\ -2 & 2 & -1 \\ 0 & -3 & 2 \end{pmatrix}.$$

Define $s_1, s_2, s_3 \in \mathrm{GL}_3(\mathbb{C})$ using (♣); show that $|\langle s_1, s_2, s_3 \rangle| = \infty$.

Remark 2.3.12. Consider the structure matrix $A = (a_{ij})_{i,j\in I}$. The formula in Remark 2.3.3 shows that

$$(*) \qquad a_{ij} = \alpha_j(h_i) = 2\frac{\langle \alpha_i, \alpha_j\rangle}{\langle \alpha_i, \alpha_i\rangle} \qquad \text{for all } i, j \in I.$$

This has the following implication on A. Let us set $d_i := \langle \alpha_i, \alpha_i\rangle$ for $i \in I$. Since all elements $w \in W$ are represented by integer matrices with respect to the basis Δ of H^* (see Remark 2.3.7), we see from the above definition of $\langle\ ,\ \rangle$ that $d_i \in \mathbb{Z}_{>0}$. Then $(*)$ implies that

$$d_i a_{ij} = 2\langle \alpha_i, \alpha_j\rangle = 2\langle \alpha_j, \alpha_i\rangle = a_{ji} d_j \quad \text{for all } i, j \in I.$$

Hence, if we denote by $D \in M_I(\mathbb{Z})$ the diagonal matrix with diagonal entries d_i $(i \in I)$, then $D \cdot A \in M_I(\mathbb{Z})$ is a symmetric matrix. In fact, $D{\cdot}A$ is (up to the factor 2) the Gram matrix of $\langle\ ,\ \rangle$ with respect to the basis Δ of E. Since $\langle\ ,\ \rangle$ is positive-definite, a well-known result from Linear Algebra shows that $\det(D \cdot A) > 0$; since also $\det(D) > 0$, we have $\det(A) > 0$. Even more is true: Let $I' \subseteq I$ be any (non-empty) subset and consider the matrix $A_{I'} := (a_{ij})_{i,j\in I'}$; similarly, let $D_{I'} \in M_{I'}(\mathbb{Z})$ be the diagonal matrix with diagonal entries d_i $(i \in I')$. Then, by the same argument as above, $D_{I'} \cdot A_{I'}$ is the Gram matrix of the restriction of $\langle\ ,\ \rangle$ to the subspace $\langle \alpha_i \mid i \in I'\rangle_{\mathbb{R}} \subseteq E$. That restriction is still positive-definite and so $\det(A_{I'}) > 0$. Thus, all principal minors of A are positive integers.

2.4. Semisimplicity

We continue to assume that (L, H) is of Cartan–Killing type with respect to $\Delta = \{\alpha_i \mid i \in I\}$. In this section, we establish the main structural properties of L. For each $i \in I$ let $\{e_i, h_i, f_i\}$ be a corresponding $\mathfrak{sl}_2$-triple, as in Remark 2.2.9; then $\mathfrak{sl}_2(\mathbb{C}) \cong S_i = \langle e_i, h_i, f_i\rangle_{\mathbb{C}} \subseteq L$.

The first step consists of "lifting" the generators s_i of W to Lie algebra automorphisms of L. Let $i \in I$. By Lemma 2.1.8, the derivations $\operatorname{ad}_L(e_i)\colon L \to L$ and $\operatorname{ad}_L(f_i)\colon L \to L$ are nilpotent. Hence, $t\operatorname{ad}_L(e_i)$ and $t\operatorname{ad}_L(f_i)$ are nilpotent derivations for all $t \in \mathbb{C}$. So we can apply the exponential construction in Lemma 1.2.8, and set

$$x_i(t) := \exp\bigl(t\operatorname{ad}_L(e_i)\bigr) \in \operatorname{Aut}(L) \quad \text{for all } t \in \mathbb{C},$$
$$y_i(t) := \exp\bigl(t\operatorname{ad}_L(f_i)\bigr) \in \operatorname{Aut}(L) \quad \text{for all } t \in \mathbb{C}.$$

Lemma 2.4.1. *With the above notation, we set*

$$n_i(t) := x_i(t) \circ y_i(-t^{-1}) \circ x_i(t) \in \mathrm{Aut}(L) \quad \textit{for } 0 \neq t \in \mathbb{C}.$$

Then the following hold.

(a) *We have* $n_i(t)(h) = h - \alpha_i(h)h_i \in H$ *for all* $h \in H$.

(b) *We have* $\lambda\big(n_i(t)(h)\big) = s_i(\lambda)(h)$ *for all* $\lambda \in H^*$ *and* $h \in H$.

(c) *We have* $n_i(t)(L_\alpha) = L_{s_i(\alpha)}$ *for all* $\alpha \in \Phi$.

Proof. (a) Let $h \in H$. Let us first determine $x_i(t)(h)$. For this purpose, we need to work out $\mathrm{ad}_L(e_i)^m(h)$ for all $m \geqslant 1$. Now, we have $\mathrm{ad}_L(e_i)(h) = [e_i, h] = -[h, e_i] = -\alpha_i(h)e_i$ and, consequently, $\mathrm{ad}_L(e_i)^m(h) = 0$ for all $m \geqslant 2$. This already shows that

$$x_i(t)(h) = \sum_{m \geqslant 0} \frac{(t\,\mathrm{ad}_L(e_i))^m(h)}{m!} = h - \alpha_i(h)te_i.$$

Similarly, we have $\mathrm{ad}_L(f_i)(h) = [f_i, h] = -[h, f_i] = \alpha_i(h)f_i$ and, consequently, $\mathrm{ad}_L(f_i)^m(h) = 0$ for all $m \geqslant 2$. This shows that

$$y_i(t)(h) = \sum_{m \geqslant 0} \frac{(t\,\mathrm{ad}_L(e_i))^m(h)}{m!} = h + \alpha_i(h)tf_i.$$

Next, we determine $y_i(t)(e_i)$. We have $\mathrm{ad}_L(f_i)(e_i) = -[e_i, f_i] = -h_i$, $\mathrm{ad}_L^2(f_i)(e_i) = -[f_i, h_i] = -2f_i$ and, consequently, $\mathrm{ad}_L(f_i)^m(e_i) = 0$ for all $m \geqslant 3$. This yields that

$$y_i(t)(e_i) = \sum_{m \geqslant 0} \frac{(t\,\mathrm{ad}_L(f_i))^m(e_i)}{m!} = e_i - th_i - t^2 f_i.$$

Combining the above formulae, we obtain that

$$\begin{aligned}
\big(y_i(-t^{-1}) \circ x_i(t)\big)(h) &= y_i(-t^{-1})\big(h - \alpha_i(h)te_i\big) \\
&= \big(h - \alpha_i(h)t^{-1}f_i\big) - \alpha_i(h)t\big(e_i + t^{-1}h_i - t^{-2}f_i\big) \\
&= h - \alpha_i(h)h_i - \alpha_i(h)te_i.
\end{aligned}$$

Finally, $\mathrm{ad}_L(e_i)^m(e_i) = 0$ for all $m \geqslant 1$ and so $x_i(t)(e_i) = e_i$. Hence,

$$\begin{aligned}
n_i(t)(h) &= x_i(t)\big(h - \alpha_i(h)h_i - \alpha_i(h)te_i\big) \\
&= \big(h - \alpha_i(h)te_i\big) - \alpha_i(h)\big(h_i - 2te_i\big) - \alpha_i(h)te_i \\
&= h - \alpha_i(h)h_i.
\end{aligned}$$

(b) Recall that $s_i(\lambda) = \lambda - \lambda(h_i)\alpha_i$. Using (a), this yields:

$$\begin{aligned}\lambda\big(n_i(t)(h)\big) &= \lambda\big(h - \alpha_i(h)h_i\big) = \lambda(h) - \alpha_i(h)\lambda(h_i)\\ &= (\lambda - \lambda(h_i)\alpha_i)(h) = s_i(\lambda)(h)\end{aligned}$$

for all $h \in H$, as desired.

(c) Let $h \in H$ and set $h' := n_i(t)(h) \in H$. Since $\alpha_i(h_i) = 2$, we see using (a) that $n_i(t)(h_i) = -h_i$; furthermore,

$$n_i(t)(h') = n_i(t)\big(h - \alpha_i(h)h_i\big) = n_i(t)(h) + \alpha_i(h)h_i = h.$$

Now let $y \in L_\alpha$ and set $y' := n_i(t)(y) \in L$. Then

$$\begin{aligned}[h, y'] &= [n_i(t)(h'), n_i(t)(y)] = n_i(t)\big([h', y]\big)\\ &= n_i(t)\big(\alpha(h')y\big) = \alpha(h')n_i(t)(y) = \alpha(h')y',\end{aligned}$$

where the second equality holds since $n_i(t)$ is a Lie algebra automorphism. Now, by (b), we have $\alpha(h') = s_i(\alpha)(h)$ and so $y' = n_i(t)(y) \in L_{s_i(\alpha)}$. Hence, $n_i(t)(L_\alpha) \subseteq L_{s_i(\alpha)}$ and $\dim L_\alpha \leqslant \dim L_{s_i(\alpha)}$. Since $s_i^2 = \mathrm{id}_{H^*}$, we also obtain that $n_i(t)(L_{s_i(\alpha)}) \subseteq L_{s_i^2(\alpha)} = L_\alpha$ and so $\dim L_{s_i(\alpha)} \leqslant \dim L_\alpha$. Hence, we must have $n_i(t)(L_\alpha) = L_{s_i(\alpha)}$. □

Proposition 2.4.2. *We have* $\dim L_\alpha = 1$ *and* $\dim[L_\alpha, L_{-\alpha}] = 1$ *for all* $\alpha \in \Phi$. *In particular,* $\dim L = |I| + |\Phi|$.

Proof. Let $\alpha \in \Phi$. By Theorem 2.3.6(a) we can write $\alpha = w(\alpha_i)$ for some $w \in W$ and $i \in I$. Furthermore, we can write $w = s_{i_1} \cdots s_{i_r}$, where $r \geqslant 0$ and $i_1, \ldots, i_r \in I$. Let us set $\varphi := n_{i_1}(1) \circ \ldots \circ n_{i_r}(1) \in \mathrm{Aut}(L)$. Now Lemma 2.4.1(c) and a simple induction on r show that

$$L_\alpha = L_{(s_{i_1}\cdots s_{i_r})(\alpha_i)} = \big(n_{i_1}(1) \circ \ldots \circ n_{i_r}(1)\big)(L_{\alpha_i}) = \varphi(L_{\alpha_i}).$$

Since $\varphi \in \mathrm{Aut}(L)$, we conclude that $\dim L_\alpha = \dim L_{\alpha_i} = 1$, where the last equality holds by Proposition 2.2.5. Furthermore, since $-\alpha = -w(\alpha_i) = w(-\alpha_i)$, the same argument shows that $L_{-\alpha} = \varphi(L_{-\alpha_i})$. Again, since $\varphi \in \mathrm{Aut}(L)$, we also have

$$[L_\alpha, L_{-\alpha}] = [\varphi(L_{\alpha_i}), \varphi(L_{-\alpha_i})] = \varphi\big([L_{\alpha_i}, L_{-\alpha_i}]\big),$$

and this is 1-dimensional by Proposition 2.2.5. Finally, the formula for $\dim L$ follows from the direct sum decomposition $L = H \oplus \bigoplus_{\alpha\in\Phi} L_\alpha$ and the fact that $\{h_i \mid i \in I\}$ is a basis of H. □

Proposition 2.4.3. *For each $\alpha \in \Phi$, there exists a unique element $h_\alpha \in [L_\alpha, L_{-\alpha}]$ such that $\alpha(h_\alpha) = 2$. (We have $h_{\alpha_i} = h_i$ for $i \in I$.) Furthermore, $h_{-\alpha} = -h_\alpha$ and $h_{s_i(\alpha)} = n_i(1)(h_\alpha)$ for all $i \in I$.*

Proof. By Proposition 2.4.2, we have $[L_\alpha, L_{-\alpha}] = \langle h\rangle_\mathbb{C}$ for some $0 \neq h \in H$. If $\alpha(h) = 0$, then Lemma 2.2.3 would imply that $\mathrm{ad}_L(h) = 0$. In particular, all eigenvalues of $\mathrm{ad}_L(h)$ are zero and so $\alpha_i(h) = 0$ for all $i \in I$, contradiction since $\{\alpha_i \mid i \in I\}$ is a basis of H^*. Thus, $\alpha(h) \neq 0$ and so there is a unique scalar multiple of h on which α takes value 2. This defines the required element h_α.

Since $-\alpha \in \Phi$ and $[L_{-\alpha}, L_\alpha] = [L_\alpha, L_{-\alpha}]$ is 1-dimensional, we have $h_{-\alpha} = \xi h_\alpha$ for some $0 \neq \xi \in \mathbb{C}$. But then we conclude that $2 = (-\alpha)(h_{-\alpha}) = -\xi\alpha(h_\alpha) = -2\xi$ and so $\xi = -1$.

Finally, let $i \in I$ and $\beta := s_i(\alpha)$. We set $n_i := n_i(1) \in \mathrm{Aut}(L)$. As in the above proof, $L_\beta = n_i(L_\alpha)$, $L_{-\beta} = n_i(L_{-\alpha})$ and so $[L_\beta, L_{-\beta}] = n_i\big([L_\alpha, L_{-\alpha}]\big) = \langle n_i(h_\alpha)\rangle_\mathbb{C}$. Hence, $h_\beta = \xi n_i(h_\alpha)$ for some $0 \neq \xi \in \mathbb{C}$. Now, by Lemma 2.4.1(b), we have $\beta(n_i(h)) = s_i(\beta)(h)$ for all $h \in H$. Since $s_i(\beta) = s_i^2(\alpha) = \alpha$, this yields $\beta(n_i(h_\alpha)) = \alpha(h_\alpha) = 2$ and so $2 = \beta(h_\beta) = \xi\beta(n_i(h_\alpha)) = 2\xi$, that is, $\xi = 1$ and $h_{s_i(\alpha)} = n_i(h_\alpha)$. □

Exercise 2.4.4. (a) By Lemma 2.4.1, we have $n_i(t)(H) \subseteq H$ for all $i \in I$ and $0 \neq t \in \mathbb{C}$. Show that $n_i(t)^2(h) = h$ for all $h \in H$. Furthermore, show that the matrix of $n_i(t)|_H \colon H \to H$ with respect to the basis $\{h_i \mid i \in I\}$ of H has integer coefficients and determinant -1.

(b) Let $\alpha \in \Phi$ and write $\alpha = w(\alpha_i)$ where $w \in W$ and $i \in I$; further write $w = s_{i_1} \cdots s_{i_r}$ where $i_1, \ldots, i_r \in I$. Show that

$$h_\alpha = \big(n_{i_1}(1) \circ \ldots \circ n_{i_r}(1)\big)(h_i) \in \langle h_j \mid j \in I\rangle_\mathbb{Z}.$$

The following result shows that the "*Chevalley generators*" in Remark 2.2.9 are indeed generators for L as a Lie algebra.

Proposition 2.4.5. *We have $L = \langle e_i, f_i \mid i \in I\rangle_{\mathrm{alg}}$.*

Proof. Let $L_0 := \langle e_i, f_i \mid i \in I\rangle_{\mathrm{alg}} \subseteq L$. Since $h_i = [e_i, f_i] \in L_0$ for all i, we have $H \subseteq L_0$. So it remains to show that $L_{\pm\alpha} \subseteq L_0$ for all $\alpha \in \Phi^+$. We proceed by induction on $\mathrm{ht}(\alpha)$.

If $\mathrm{ht}(\alpha) = 1$, then $\alpha = \alpha_i$ for some $i \in I$. Since $L_{\alpha_i} = \langle e_i\rangle_\mathbb{C}$ and $L_{-\alpha_i} = \langle f_i\rangle_\mathbb{C}$, we have $L_{\pm\alpha_i} \subseteq L_0$ by the definition of L_0. Now let $\mathrm{ht}(\alpha) > 1$. By the Key Lemma 2.3.4, there exists some $j \in I$ such that $\beta := \alpha - \alpha_j \in \Phi^+$. We have $\mathrm{ht}(\beta) = \mathrm{ht}(\alpha) - 1$ and so, by induction, $L_{\pm\beta} \subseteq L_0$. By Remark 2.2.10(c'), since $\alpha_j + \beta = \alpha \in \Phi$, we have $\{0\} \neq [L_{\alpha_j}, L_\beta] \subseteq L_{\alpha_j+\beta} = L_\alpha$. Since $\dim L_\alpha = 1$ (see Proposition 2.4.2), we conclude that $L_\alpha = [L_{\alpha_j}, L_\beta]$, and this is contained in L_0 because L_0 is a subalgebra and $L_{\alpha_j} \subseteq L_0$, $L_\beta \subseteq L_0$. Similarly, $-\alpha = -\alpha_j - \beta$ and $L_{-\alpha} = [L_{-\alpha_j}, L_{-\beta}] \subseteq L_0$. □

Proposition 2.4.6. *Let $J \subseteq L$ be an ideal. If $J \neq \{0\}$, then there exists some $i \in I$ such that $S_i \subseteq J$. In particular, J is non-abelian and so L is semisimple.*

Proof. Assume that $J \neq \{0\}$. We have $[H, J] \subseteq J$ and so J is an H-submodule of L. Hence, by Proposition 2.1.4(b), we have

$$J = (J \cap H) \oplus \bigoplus_{\alpha\in\Phi} (J \cap L_\alpha).$$

So there are two cases: $J \cap H \neq \{0\}$ or $J \cap L_\alpha \neq \{0\}$ for some $\alpha \in \Phi$. Assume first that $J \cap H \neq \{0\}$. Let $0 \neq h \in J \cap H$. Since $\{\alpha_i \mid i \in I\}$ is a basis of H^*, we have $\alpha_i(h) \neq 0$ for some $i \in I$. Then $\alpha_i(h)e_i = [h, e_i] \in J$ and so $e_i \in J$. Hence, also $h_i = [e_i, f_i] \in J$; furthermore, $-2f_i = [h_i, f_i] \in J$ and so $S_i \subseteq J$, as desired. Now assume that $J \cap L_\alpha \neq \{0\}$ for some $\alpha \in \Phi$. By Proposition 2.4.2, we have $\dim L_\alpha = 1$ and so $L_\alpha \subseteq J$. Consequently, by Proposition 2.4.3, we also have $h_\alpha \in [L_\alpha, L_{-\alpha}] \subseteq J$ and we are back in the previous case. Thus, in any case, $S_i \subseteq J$ for some $i \in I$, as claimed. Since $S_i \cong \mathfrak{sl}_2(\mathbb{C})$ is not abelian, this shows that J is not abelian. Hence, by Lemma 1.3.9, we must have $\mathrm{rad}(L) = \{0\}$ and so L is semisimple. □

ab hier Woche 6

Definition 2.4.7. Consider the structure matrix $A = (a_{ij})_{i,j\in I}$ of L or, somewhat more generally, any matrix $A = (a_{ij})_{i,j\in I}$ such that the a_{ij} satisfy the conditions (a), (b), (c) in Corollary 2.2.12. We say that A is *decomposable* if there is a partition $I = I_1 \sqcup I_2$ (where $I_1, I_2 \subsetneq I$ and $I_1 \cap I_2 = \varnothing$) such that $a_{ij} = a_{ji} = 0$ for all $i \in I_1$ and $j \in I_2$. In this case we can label I such that A has a block diagonal shape

$$A = \left(\begin{array}{c|c} A_1 & 0 \\ \hline 0 & A_2 \end{array}\right)$$

where A_1 has rows and columns labelled by I_1, and A_2 has rows and columns labelled by I_2. If no such partition of I exists, then we say that A is *indecomposable*. Note that A can be written as a block diagonal matrix where the diagonal blocks are indecomposable.

Remark 2.4.8. Given A we define a graph with vertex set I; two vertices $i, j \in I$, $i \neq j$, are joined by an edge if $a_{ij} \neq 0$. (Recall that $a_{ij} \neq 0 \Leftrightarrow a_{ji} \neq 0$.) Then a standard argument in graph theory shows that this graph is connected if and only if A is indecomposable (see, e.g., [**4**, Ch. IV, Annexe, Cor. 1]). Hence, the indecomposability of A can be alternatively expressed as follows. For any $i, j \in I$ such that $i \neq j$, there exists a sequence of distinct indices $i = i_0, i_1, \ldots, i_r = j$ in I, where $r \geqslant 1$ and $a_{i_l i_{l+1}} \neq 0$ for $0 \leqslant l \leqslant r-1$.

Lemma 2.4.9. *Assume that the structure matrix $A = (a_{ij})_{i,j \in I}$ of L is decomposable and write $I = I_1 \sqcup I_2$ (disjoint union), as above. We define $E_1 := \langle \alpha_i \mid i \in I_1 \rangle_{\mathbb{R}} \subseteq H^*$ and $E_2 := \langle \alpha_i \mid i \in I_2 \rangle_{\mathbb{R}} \subseteq H^*$. Then $E_1 \cap E_2 = \{\underline{0}\}$ and $\Phi = \Phi_1 \sqcup \Phi_2$ (disjoint union), where*

$$\Phi_1 := \Phi \cap E_1 \neq \varnothing \qquad \text{and} \qquad \Phi_2 := \Phi \cap E_2 \neq \varnothing;$$

furthermore, $\alpha \pm \beta \notin \Phi \cup \{\underline{0}\}$ for all $\alpha \in \Phi_1$ and $\beta \in \Phi_2$.

Proof. Since I is the disjoint union of I_1 and I_2, we certainly have $E_1 \cap E_2 = \{\underline{0}\}$. Let $i \in I_1$ and $j \in I$. Then $s_i(\alpha_j) = \alpha_j - \alpha_j(h_i)\alpha_i = \alpha_j - a_{ij}\alpha_i$. Hence, if $j \in I_1$, then $s_i(\alpha_j) \in E_1$; if $j \in I_2$, then $s_i(\alpha_j) = \alpha_j$, since $a_{ij} = 0$. Consequently, we have:

(a) $\quad i \in I_1 \quad \Rightarrow \quad s_i(E_1) \subseteq E_1$ and $s_i(v) = v$ for all $v \in E_2$.

Similarly, we see that

(b) $\quad i \in I_2 \quad \Rightarrow \quad s_i(E_2) \subseteq E_2$ and $s_i(v) = v$ for all $v \in E_1$.

It follows that $w(E_1) \subseteq E_1$ and $w(E_2) \subseteq E_2$ for all $w \in W$. (Indeed, by (a) and (b), the desired property holds for all generators s_i of W and, hence, it holds for all elements of W.) Now, by Theorem 2.3.6(a), we have $\Phi = \{w(\alpha_i) \mid w \in W, i \in I\}$ and so $\Phi = \Psi_1 \cup \Psi_2$, where

$$\Psi_1 := \{w(\alpha_i) \mid w \in W, i \in I_1\} \subseteq \{w(v) \mid w \in W, v \in E_1\} \subseteq E_1,$$
$$\Psi_2 := \{w(\alpha_i) \mid w \in W, i \in I_2\} \subseteq \{w(v) \mid w \in W, v \in E_2\} \subseteq E_2.$$

Thus, $\Psi_1 \subseteq \Phi \cap E_1 = \Phi_1$ and $\Psi_2 \subseteq \Phi \cap E_2 = \Phi_2$. This yields that

$$\Phi = \Psi_1 \cup \Psi_2 \subseteq \Phi_1 \cup \Phi_2 \subseteq \Phi.$$

Furthermore, since $E_1 \cap E_2 = \{\underline{0}\}$ and $\underline{0} \not\in \Phi$, we have $\Psi_1 \cap \Psi_2 = \varnothing$ and $\Phi_1 \cap \Phi_2 = \varnothing$. So we conclude that all of the above inclusions must be equalities; hence, $\Psi_1 = \Phi_1$ and $\Psi_2 = \Phi_2$.

Finally, let $\alpha \in \Phi_1$ and $\beta \in \Phi_2$. If $\alpha \pm \beta = \underline{0}$, then $\alpha = \pm\beta \in E_1 \cap E_2 = \{\underline{0}\}$, a contradiction. Now let $\alpha \pm \beta \in \Phi$. Since $\Phi = \Phi_1 \cup \Phi_2$, we have $\alpha \pm \beta \in \Phi_1$ or $\alpha \pm \beta \in \Phi_2$. In the first case, $\alpha \pm \beta \in E_1$ and so $\pm\beta = (\alpha \pm \beta) - \alpha \in E_1 \cap E_2 = \{\underline{0}\}$, a contradiction. In the second case, $\alpha \pm \beta \in E_2$ and so $\alpha = (\alpha \pm \beta) \mp \beta \in E_1 \cap E_2 = \{\underline{0}\}$, again a contradiction. Thus, we have $\alpha \pm \beta \not\in \Phi \cup \{\underline{0}\}$. □

Exercise 2.4.10. In the above setting, consider the subgroups

$$W_1 := \langle s_i \mid i \in I_1 \rangle \subseteq W \quad \text{and} \quad W_2 := \langle s_i \mid i \in I_2 \rangle \subseteq W.$$

Use (a), (b) in the above proof to show that $W = W_1 \cdot W_2 = W_2 \cdot W_1$ and $W_1 \cap W_2 = \{\text{id}\}$. Also show that $\Phi_s = \{w(\alpha_i) \mid w \in W_s, i \in I_s\}$ for $s = 1, 2$. Thus, Φ_1 and Φ_2 are entirely determined by I_1 and I_2.

Lemma 2.4.11. *Let $I' \subseteq I$ be any subset. We set*

$$H' := \langle h_i \mid i \in I' \rangle_{\mathbb{C}}, \quad E' := \langle \alpha_i \mid i \in I' \rangle_{\mathbb{R}} \subseteq H^*, \quad \Phi' := \Phi \cap E'.$$

Then $L' := H' \oplus \bigoplus_{\alpha \in \Phi'} L_\alpha \subseteq L$ is a subalgebra of L.

Proof. Since H is abelian and $[H, L_\alpha] \subseteq L_\alpha$ for all $\alpha \in \Phi$, it is clear that $[H, L'] \subseteq H'$. Now let $\alpha, \beta \in \Phi'$. Again, we have $[L_\alpha, H] = [H, L_\alpha] \subseteq L_\alpha \subseteq L'$. Finally, we must show that $[L_\alpha, L_\beta] \subseteq L'$. If $[L_\alpha, L_\beta] = \{0\}$, then this is clear. Now assume that $[L_\alpha, L_\beta] \neq \{0\}$. By Proposition 2.1.6, we have $[L_\alpha, L_\beta] \subseteq L_{\alpha+\beta}$, and so either $\alpha + \beta = \underline{0}$ or $\alpha + \beta \in \Phi$. In the second case, since $\alpha \in E'$ and $\beta \in E'$, we also have $\alpha + \beta \in \Phi \cap E' = \Phi'$ and so $L_{\alpha+\beta} \subseteq L'$, as required. Now assume that $\alpha + \beta = \underline{0}$, that is, $\beta = -\alpha$. Then $[L_\alpha, L_\beta] = \langle h_\alpha \rangle_{\mathbb{C}}$ by Propositions 2.4.2 and 2.4.3. It remains to show that $h_\alpha \in H'$ for all $\alpha \in \Phi'$. Since $h_{-\alpha} = -h_\alpha$ (see Proposition 2.4.3), it is enough to do this for $\alpha \in \Phi' \cap \Phi^+$. We now argue by induction on $\text{ht}(\alpha)$. If $\text{ht}(\alpha) = 1$, then $\alpha = \alpha_i$ for some $i \in I$. Since also $\alpha_i \in \Phi' \subseteq E'$, we have $i \in I'$ and so $h_i = h_{\alpha_i} \in H'$. Now assume that $\text{ht}(\alpha) > 1$. By the Key Lemma 2.3.4, there exists some $i \in I$ such that $n = \alpha(h_i) > 0$

and $\alpha' := s_i(\alpha) = \alpha - n\alpha_i \in \Phi^+$. Since $\alpha \in E'$ and every root is either positive or negative, we must have $i \in I'$ and so $\alpha' \in \Phi' \cap \Phi^+$. Since $\mathrm{ht}(\alpha') = \mathrm{ht}(\alpha) - n < \mathrm{ht}(\alpha)$, we can apply induction; thus, $h_{\alpha'} \in H'$. Now $\alpha = s_i(\alpha')$ and so $h_\alpha = n_i(1)(h_{\alpha'})$; see again Proposition 2.4.3. Finally, Lemma 2.4.1(a) shows that $h_\alpha = h_{\alpha'} - \alpha_i(h_{\alpha'})h_i \in H'$. □

Exercise 2.4.12. Assume that we are in the set-up of Lemma 2.4.11. For any $\lambda \in H^*$, denote by $\lambda' \in H'^*$ the restriction of λ to H'.
Use Remark 2.3.12 to show that $\Delta' := \{\alpha_i' \mid i \in I'\} \subseteq H'^*$ is linearly independent; furthermore, $\alpha' \neq \underline{0}$ and $L_\alpha = L'_{\alpha'}$ for any $\alpha \in \Phi'$.
Deduce that (L', H') is of Cartan–Killing type with respect to Δ', and that the corresponding structure matrix is $A' = (a_{ij})_{i,j\in I'}$.

Proposition 2.4.13. *Assume that A is decomposable and write $I = I_1 \sqcup I_2$, as above. Let $\Phi = \Phi_1 \sqcup \Phi_2$ as in Lemma 2.4.9 and set*

$$L_1 := H_1 \oplus \bigoplus_{\alpha\in\Phi_1} L_\alpha \quad \textit{where} \quad H_1 := \langle h_i \mid i \in I_1\rangle_{\mathbb{C}} \neq \{0\},$$

$$L_2 := H_2 \oplus \bigoplus_{\alpha\in\Phi_2} L_\alpha \quad \textit{where} \quad H_2 := \langle h_i \mid i \in I_2\rangle_{\mathbb{C}} \neq \{0\}.$$

Then L_1, L_2 are ideals such that $L = L_1 \oplus L_2$ and $[L_1, L_2] = \{0\}$.

Proof. Since $H = H_1 \oplus H_2$ and $\Phi = \Phi_1 \cup \Phi_2$ is a disjoint union, we have $L = L_1 \oplus L_2$ (direct sum of vector spaces). By Lemma 2.4.11, L_1 and L_2 are subalgebras of L. It remains to show that $[L_1, L_2] = \{0\}$. (Note that this implies that L_1 and L_2 are ideals). We can do this term by term according to the above direct sum decompositions.

First, since H is abelian, it is clear that $[H_1, H_2] = \{0\}$. Next, let $i \in I_1$ and $\beta \in \Phi_2$. We write $\beta = \sum_{j\in I_2} n_j\alpha_j$. For $y \in L_\beta$, we have

$$[h_i, y] = \beta(h_i)y = \sum_{j\in I_2} n_j\alpha_j(h_i)y = \sum_{j\in I_2} n_j a_{ij} y = 0,$$

by the conditions on $I = I_1 \cup I_2$. Thus, $[H_1, L_\beta] = \{0\}$. A completely analogous argument also shows that $[L_\alpha, H_2] = \{0\}$ for all $\alpha \in \Phi_1$. Finally, let $\alpha \in \Phi_1$ and $\beta \in \Phi_2$. If we had $[L_\alpha, L_\beta] \neq \{0\}$, then $L_{\alpha+\beta} \neq \{0\}$ (see Proposition 2.1.6) and so $\alpha + \beta \in \Phi \cup \{\underline{0}\}$, contradiction to Lemma 2.4.9. □

Theorem 2.4.14. *Assume that $L \neq \{0\}$. Then L is simple if and only if A is indecomposable.*

Proof. By Proposition 2.4.13, L is not simple if A is decomposable. Conversely, assume now that L is not simple. Since $L \neq \{0\}$, we have $I \neq \varnothing$ and L is not abelian (see Definition 2.2.1). Let $J \subseteq L$ be an ideal such that $\{0\} \neq J \neq L$. Let $i \in I$. Then $S_i \cap J$ is an ideal of S_i. So, since $S_i \cong \mathfrak{sl}_2(\mathbb{C})$ is simple, either $S_i \subseteq J$ or $S_i \cap J = \{0\}$. Thus, we have $I = I_1 \sqcup I_2$ (disjoint union) where

$$I_1 := \{i \in I \mid S_i \subseteq J\} \qquad \text{and} \qquad I_2 := \{i \in I \mid S_i \cap J = \{0\}\}.$$

Now, since $J \neq \{0\}$, we have $I_1 \neq \varnothing$ by Proposition 2.4.6. If we had $I_2 = \varnothing$, then $I = I_1$ and so $e_i, f_i \in J$ for all $i \in I$; hence, we would have $L = \langle e_i, f_i \mid i \in I\rangle_{\text{alg}} \subseteq J$ by Proposition 2.4.5, contradiction to our assumptions. Thus, we also have $I_2 \neq \varnothing$. Now let $i \in I_1$ and $j \in I_2$. We claim that $a_{ij} = 0$. To see this, consider the relation $[h_i, e_j] = \alpha_j(h_i)e_j = a_{ij}e_j$. Since $h_i \subseteq J$, we also have $a_{ij}e_j = [h_i, e_j] \in J$. Hence, if $a_{ij} \neq 0$, then $e_j \in J$ and so $J \cap S_j \neq \{0\}$, contradiction. Thus, we must have $a_{ij} = 0$, as claimed. Since this holds for all $i \in I_1$ and $j \in I_2$, we conclude that A is decomposable. □

Remark 2.4.15. The above results lead to a simple (!) method for testing if L is a simple Lie algebra. Namely, we claim that L is simple if $I \neq \varnothing$ and there exists some $\alpha_0 \in \Phi$ with

$$\alpha_0 = \sum_{i \in I} n_i \alpha_i \quad \text{and} \quad 0 \neq n_i \in \mathbb{Z} \text{ for all } i \in I.$$

Indeed, if L were not simple, then A would be decomposable by Theorem 2.4.14 and so $I = I_1 \sqcup I_2$ (disjoint union), where $I_1 \subsetneqq I$ and $I_2 \subsetneqq I$. But then Proposition 2.4.13 would imply that every $\alpha \in \Phi$ is a linear combination of $\{\alpha_i \mid i \in I_1\}$ or of $\{\alpha_i \mid i \in I_2\}$, contradiction to the existence of α_0 as above.

Example 2.4.16. Let $L = \mathfrak{sl}_n(\mathbb{C})$, where $n \geqslant 2$. In Example 1.5.3, we have already seen that L is semisimple. Now we claim that L is simple. This is seen as follows. By Example 2.2.7, L is of Cartan–Killing type with respect to $\Delta = \{\alpha_1, \ldots, \alpha_{n-1}\}$; furthermore, the explicit description of Φ shows that $\alpha_0 = \alpha_1 + \ldots + \alpha_{n-1} \in \Phi$. Hence, L is simple by Remark 2.4.15. — In the next section, we will employ a similar argument to show that the Lie algebras $\mathfrak{go}_n(Q_n, \mathbb{C})$ are simple.

If A is indecomposable, then there is always a particular root α_0 with the property in Remark 2.4.15.

Proposition 2.4.17. *Assume that A is indecomposable (and $I \neq \varnothing$). Then there is a unique positive root $\alpha_0 \in \Phi^+$ (called "highest root") such that* $\mathrm{ht}(\alpha_0) = \max\{\mathrm{ht}(\gamma) \mid \gamma \in \Phi\}$. *We have $\alpha_0 = \sum_{i\in I} n_i\alpha_i$, where $n_i > 0$ for all $i \in I$. Furthermore, $\alpha_0 + \alpha_i \notin \Phi$ for all $i \in I$, and this property characterises α_0 (among all positive roots).*

Proof. Since $|\Phi| < \infty$, there exists at least some root $\alpha_0 \in \Phi^+$ such that $\mathrm{ht}(\alpha_0) = \max\{\mathrm{ht}(\gamma) \mid \gamma \in \Phi\}$. Then $\alpha_0 + \alpha_i \notin \Phi$ for all $i \in I$. Now let $\beta \in \Phi^+$ also be such that $\beta + \alpha_i \notin \Phi$ for all $i \in I$. We must show that $\beta = \alpha_0$. For this purpose, let $0 \neq e_\beta \in L_\beta$ and define $U \subseteq L$ to be the subspace spanned by all $v \in L$ of the form

$$(*) \qquad v = [f_{i_1}, [f_{i_2}, [\ldots, [f_{i_l}, e_\beta]\ldots]]], \quad \text{where } l \geqslant 0,\ i_1, \ldots, i_l \in I.$$

First note that, by Proposition 2.1.6, any v as above belongs to the subspace $L_{\beta-(\alpha_{i_1}+\ldots+\alpha_{i_l})} \subseteq L$. Thus, since $L_\beta = \langle e_\beta\rangle_{\mathbb{C}}$, we have

$$U = \sum_{i_1,\ldots,i_l\in I;\, l\geqslant 0} L_{\beta-(\alpha_{i_1}+\ldots+\alpha_{i_l})}.$$

In particular, this shows that $[H, U] \subseteq U$. By construction, we also have $[f_i, U] \subseteq U$ for all $i \in I$. We claim that U is an ideal in L. By Proposition 2.4.5 and Exercise 1.1.8, it remains to show that $[e_i, U] \subseteq U$ for all $i \in I$. So let $i \in I$ and v be as in $(*)$. We show by induction on l that $[e_i, v] \in U$. If $l = 0$, then $v = e_\beta$ and $[e_i, e_\beta] \in L_{\beta+\alpha_i} = \{0\}$ (since $\beta+\alpha_i \notin \Phi$); so $[e_i, e_\beta] = 0 \in U$, as required. Now let $l \geqslant 1$ and set $v' := [f_{i_2}, [f_{i_3}, [\ldots, [f_{i_l}, e_\beta]\ldots]]]$. Then $v = [f_{i_1}, v']$ and so

$$[e_i, v] = [e_i, [f_{i_1}, v']] = -[f_{i_1}, [v', e_i]] - [v', [e_i, f_{i_1}]].$$

By induction, $[v', e_i] = -[e_i, v'] \in U$ and so $[f_{i_1}, [v', e_i]] \in U$. Furthermore, if $i = i_1$, then $[e_i, f_{i_1}] = h_i$ and so $[v', [e_i, f_{i_1}]] = [v', h_i] = -[h_i, v'] \in U$. Finally, if $i \neq i_1$, then $[e_i, f_{i_1}] \in L_{\alpha_i-\alpha_{i_1}} = \{0\}$ and so $[[e_1, f_{i_1}], v'] = 0$. Thus, in all cases, we have $[e_i, v] \in U$, as desired.

But then we conclude that $U = L$, since A is indecomposable and, hence, L is simple (see Theorem 2.4.14). Now we can argue as follows. For any $\alpha \in \Phi$, we have $L_\alpha \subseteq L = U$ and so the above description of U implies that $\alpha = \beta-(\alpha_{i_1}+\ldots+\alpha_{i_l})$ for some $i_1, \ldots, i_l \in I$, $l \geqslant 0$ (see Exercise 2.1.5). Taking $\alpha = \alpha_0$ yields that $\mathrm{ht}(\alpha_0) \geqslant \mathrm{ht}(\beta) = \mathrm{ht}(\alpha_0)+l$ and so $l = 0$, that is, $\beta = \alpha_0$, as desired. Taking $\alpha = \alpha_i$ for some

$i \in I$ yields that $\alpha_0 = \beta = \alpha_i + (\alpha_{i_1} + \ldots + \alpha_{i_l})$. Hence, writing $\alpha_0 = \sum_{i \in I} n_i \alpha_i$ with $n_i \in \mathbb{Z}$, we have $n_i > 0$ for all i. $\square$

2.5. Classical Lie algebras revisited

We return to the *classical Lie algebras* introduced in Section 1.6. Our aim is to show that these algebras are simple. Let

$$Q_n = \begin{pmatrix} 0 & \cdots & 0 & \delta_n \\ \vdots & \iddots & \iddots & 0 \\ 0 & \delta_2 & \iddots & \vdots \\ \delta_1 & 0 & \cdots & 0 \end{pmatrix} \in M_n(\mathbb{C}), \qquad \begin{matrix} Q_n^{\mathrm{tr}} = \epsilon Q_n \\ (\epsilon = \pm 1) \end{matrix},$$

where $\delta_i = \pm 1$ are such that $\delta_i \delta_{n+1-i} = \epsilon$ for all i. Then

$$L := \mathfrak{go}_n(Q_n, \mathbb{C}) := \{A \in M_n(\mathbb{C}) \mid A^{\mathrm{tr}} Q_n + Q_n A = 0\} \subseteq \mathfrak{gl}_n(\mathbb{C}).$$

We assume throughout that $n \geqslant 3$. Then we have already seen in Proposition 1.6.3 that $\mathfrak{go}_n(Q_n, \mathbb{C})$ is semisimple.

Let H be the subspace of diagonal matrices in L. Let $m \geqslant 1$ be such that $n = 2m + 1$ (if n is odd) or $n = 2m$ (if n is even). By the explicit description of H in Remark 1.6.7, we have $\dim H = m$ and $H = \{h(x_1, \ldots, x_m) \mid x_i \in \mathbb{C}\}$, where

$$h(x_1, \ldots, x_m) := \begin{cases} \operatorname{diag}(x_1, \ldots, x_m, 0, -x_m, \ldots, -x_1) & \text{if } n = 2m+1, \\ \operatorname{diag}(x_1, \ldots, x_m, -x_m, \ldots, -x_1) & \text{if } n = 2m. \end{cases}$$

Furthermore, by Remark 2.1.11, we have $C_L(H) = H$ and L is H-diagonalisable. Thus, we have a weight space decomposition

$$L = H \oplus \bigoplus_{\alpha \in \Phi} L_\alpha \qquad \text{where} \qquad H = L_{\underline{0}} \text{ and } \Phi \subseteq H^* \setminus \{\underline{0}\}.$$

In order to determine Φ, we use the basis elements

$$A_{ij} := \delta_i E_{ij} - \delta_j E_{n+1-j, n+1-i} \in \mathfrak{go}_n(Q_n, \mathbb{C})$$

for all $1 \leqslant i, j \leqslant n$, where E_{ij} denotes the elementary matrix with 1 at position (i, j), and 0 otherwise. (See Proposition 1.6.6.) If $x = \operatorname{diag}(x_1, \ldots, x_n) \in H$, we write $\varepsilon_l(x) = x_l$ for $1 \leqslant l \leqslant n$; this defines a linear map $\varepsilon_l \colon H \to \mathbb{C}$. Note that $\varepsilon_l + \varepsilon_{n+1-l} = 0$ for $1 \leqslant l \leqslant n$.

Lemma 2.5.1. *We have* $[x, A_{ij}] = \big(\varepsilon_i(x) - \varepsilon_j(x)\big) A_{ij}$ *for all* $x \in H$.

Proof. If $x = \operatorname{diag}(x_1, \ldots, x_n)$, then $[x, E_{ij}] = (x_i - x_j)E_{ij}$ and so

$$\begin{aligned}[x, A_{ij}] &= \delta_i[x, E_{ij}] - \delta_j[x, E_{n+1-j,n+1-i}] \\ &= \delta_i(x_i - x_j)E_{ij} - \delta_j(x_{n+1-j} - x_{n+1-i})E_{n+1-j,n+1-i}.\end{aligned}$$

But, since $x \in H$, we have $x_{n+1-l} = -x_l$ for $1 \leqslant l \leqslant n$ and so $[x, A_{ij}] = (x_i - x_j)(\delta_i E_{ij} - \delta_j E_{n+1-j,n+1-i}) = (x_i - x_j)A_{ij}$. □

Remark 2.5.2. Later on, we shall also need to know at least some Lie brackets among the elements A_{ij}. A straightforward computation yields the following formulae. If $i + j \neq n + 1$, then

$$[A_{ij}, A_{ji}] = \delta_i\delta_j(E_{ii} - E_{jj}) + \delta_j\delta_i(E_{n+1-j,n+1-j} - E_{n+1-i,n+1-i});$$

note that this is a diagonal matrix in H. Furthermore, a particular situation occurs when $i + j = n + 1$ and $\epsilon = -1$. Then

$$A_{ij} = 2\delta_i E_{ij} \qquad \text{and} \qquad [A_{ij}, A_{ji}] = 4(E_{jj} - E_{ii}) \in H.$$

Lemma 2.5.3. *Recall that $m \geqslant 1$ is such that $n = 2m+1$ or $n = 2m$.*

(a) *In all cases, $\{\pm\varepsilon_i \pm \varepsilon_j \mid 1 \leqslant i, j \leqslant m, i \neq j\} \subseteq \Phi$. This subset contains precisely $2m(m-1)$ distinct elements.*

(b) *$\{\pm\varepsilon_i \mid 1 \leqslant i \leqslant m\} \subseteq \Phi$ if $n = 2m + 1$ is odd and $Q_n^{\mathrm{tr}} = Q_n$.*

(c) *$\{\pm 2\varepsilon_i \mid 1 \leqslant i \leqslant m\} \subseteq \Phi$ if $n = 2m$ is even and $Q_n^{\mathrm{tr}} = -Q_n$.*

Proof. (a) Let $1 \leqslant i, j \leqslant m$, $i \neq j$. Then Lemma 2.5.1 shows that $\varepsilon_i - \varepsilon_j \in \Phi$, with A_{ij} as a corresponding eigenvector. (We have $A_{ij} \neq 0$ in this case.) Now set $l := n + 1 - j$. Then $l \neq i$ and so Lemma 2.5.1 also shows that $\varepsilon_i - \varepsilon_l \in \Phi$. (Note that, again, $A_{il} \neq 0$.) But $\varepsilon_l = \varepsilon_{n+1-j} = -\varepsilon_j$ and so $\varepsilon_i + \varepsilon_j \in \Phi$. Similarly, let $k := n+1-i$; then $k \neq j$ and so $\varepsilon_k - \varepsilon_j \in \Phi$. But $\varepsilon_k = \varepsilon_{n+1-i} = -\varepsilon_i$ and so $-\varepsilon_i - \varepsilon_j \in \Phi$. Since $\{\varepsilon_1, \ldots, \varepsilon_m\} \subseteq H^*$ are linearly independent, the functions $\pm\varepsilon_i \pm \varepsilon_j \in H^*$ ($1 \leqslant i < j \leqslant m$) are all distinct. So we have precisely $2m(m-1)$ such functions.

(b) Let $1 \leqslant i \leqslant m$. Then $[x, A_{i,m+1}] = (x_i - x_{m+1})A_{i,m+1}$ for all $x \in H$. But $x_{m+1} = -x_{n+1-(m+1)} = -x_{m+1}$ and so $x_{m+1} = 0$. Hence, we have $[x, A_{i,m+1}] = x_i A_{i,m+1} = \varepsilon_i(x)A_{i,m+1}$ for all $x \in H$. So $\varepsilon_i \in \Phi$ (since $A_{i,m+1} \neq 0$). Similarly, we see that $[x, A_{m+1,i}] = -\varepsilon_i(x)A_{m+1,i}$ for all $x \in H$. Hence, $-\varepsilon_i \in \Phi$.

(c) Let $1 \leqslant i \leqslant m$ and $x \in H$. Since $x_{2m+1-i} = -x_i$, we have $[x, A_{i,2m+1-i}] = (x_i - x_{2m+1-i})A_{i,2m+1-i} = 2\varepsilon_i(x)A_{i,2m+1-i}$. Since $Q_n^{\text{tr}} = -Q_n$, we have $\delta_i = -\delta_{2m+1-i}$ and so $A_{i,2m+1-i} \neq 0$. This shows that $2\varepsilon_i \in \Phi$. Similarly, we see that $[x, A_{2m+1-i,i}] = -2\varepsilon_i(x)A_{2m+1-i,i}$ for all $x \in H$. Hence, $-2\varepsilon_i \in \Phi$. □

Proposition 2.5.4. *Let $H \subseteq L = \mathfrak{go}_n(Q_n, \mathbb{C})$ as above.*

(a) *If $Q_n^{\text{tr}} = Q_n$ and $n = 2m$ is even, then we have $|\Phi| = 2(m^2 - m)$ and $\Phi = \{\pm\varepsilon_i \pm \varepsilon_j \mid 1 \leqslant i, j \leqslant m, i \neq j\}$.*

(b) *If $Q_n^{\text{tr}} = Q_n$ and $n = 2m+1$ is odd, then we have $|\Phi| = 2m^2$ and $\Phi = \{\pm\varepsilon_i \pm \varepsilon_j, \pm\varepsilon_i \mid 1 \leqslant i, j \leqslant m, i \neq j\}$.*

(c) *If $Q_n^{\text{tr}} = -Q_n$, then $n = 2m$ is necessarily even, we have $|\Phi| = 2m^2$ and $\Phi = \{\pm\varepsilon_i \pm \varepsilon_j, \pm 2\varepsilon_i \mid 1 \leqslant i, j \leqslant m, i \neq j\}$.*

Proof. By Lemma 2.5.3, $|\Phi| \geqslant 2m^2 - 2m$ (if $n = 2m$ and $Q_n^{\text{tr}} = Q_n$) and $|\Phi| \geqslant 2m^2$ (otherwise). Since $\dim H = m$, this implies that $\dim L \geqslant \dim H + |\Phi| \geqslant 2m^2 - m$ (if $n = 2m$ and $Q_n^{\text{tr}} = Q_n$) and $\dim L \geqslant 2m^2 + m$ (otherwise). Combining this with the formulae in Remark 1.6.7, we conclude that equality holds everywhere. In particular, Φ is given by the functions described in Lemma 2.5.3. In (c), note that $Q_n^{\text{tr}} = -Q_n$ implies that n must be even. □

Remark 2.5.5. In all three cases in Proposition 2.5.4, we have $\Phi' := \{\varepsilon_i - \varepsilon_j \mid 1 \leqslant i, j \leqslant m, i \neq j\} \subseteq \Phi$, which is like the set of roots of $\mathfrak{sl}_m(\mathbb{C})$ in Example 2.2.7. We reverse the notation there[3] and set

$$\alpha_i := \varepsilon_{m+1-i} - \varepsilon_{m+2-i} \qquad \text{for } 2 \leqslant i \leqslant m.$$

Thus, $\alpha_m = \varepsilon_1 - \varepsilon_2$, $\alpha_{m-1} = \varepsilon_2 - \varepsilon_3$, $\ldots$, $\alpha_2 = \varepsilon_{m-1} - \varepsilon_m$; or $\alpha_{m+2-i} = \varepsilon_{i-1} - \varepsilon_i$. For $1 \leqslant i < j \leqslant m$, we obtain:

$$\alpha_{i+1} + \alpha_{i+2} + \ldots + \alpha_j = \varepsilon_{m+1-j} - \varepsilon_{m+1-i}$$

and so $\Phi' = \{\pm(\alpha_{i+1} + \alpha_{i+2} + \ldots + \alpha_j) \mid 1 \leqslant i < j \leqslant m\}$. Furthermore, in all three cases, we have $\Phi'' := \{\pm(\varepsilon_i + \varepsilon_j) \mid 1 \leqslant i < j \leqslant m\} \subseteq \Phi$. We will now try to obtain convenient descriptions for Φ''.

---

[3] The reason for this notational reversion is to maintain consistence with the labelling of Dynkin diagrams that we will classify in Chapter 3; see also Remark 2.5.7.

• In case (a), $\Phi = \Phi' \cup \Phi''$. If we also set $\alpha_1 := \varepsilon_{m-1} + \varepsilon_m$, then $\alpha_1, \alpha_2, \ldots, \alpha_m$ are linearly independent. For $1 \leqslant i < j \leqslant m$, we have

$$\alpha_2 + \ldots + \alpha_i = \varepsilon_{m+1-i} - \varepsilon_m, \quad \alpha_3 + \ldots + \alpha_j = \varepsilon_{m+1-j} - \varepsilon_{m-1},$$

and so $(\alpha_1 + \alpha_2 + \ldots + \alpha_i) + (\alpha_3 + \alpha_4 + \ldots + \alpha_j) = \varepsilon_{m+1-i} + \varepsilon_{m+1-j}$. (Note that $m \geqslant 2$ since $n \geqslant 3$.) Hence, these expressions describe all elements of Φ''.

• In case (b), $\Phi = \Phi' \cup \Phi'' \cup \{\pm\varepsilon_i \mid 1 \leqslant i \leqslant m\}$. If we also set $\alpha_1 := \varepsilon_m$, then $\alpha_1, \alpha_2, \ldots, \alpha_m$ are linearly independent. We have

$$\alpha_1 + (\alpha_2 + \ldots + \alpha_i) = \varepsilon_m + (\varepsilon_{m+1-i} - \varepsilon_m) = \varepsilon_{m+1-i}$$

for $1 \leqslant i \leqslant m$. Furthermore, for $1 \leqslant i < j \leqslant m$, we obtain

$$\begin{aligned} 2(\alpha_1 + &\alpha_2 + \ldots + \alpha_i) + \alpha_{i+1} + \alpha_{i+2} + \ldots + \alpha_j \\ &= 2\varepsilon_{m+1-i} + (\varepsilon_{m+1-j} - \varepsilon_{m+1-i}) = \varepsilon_{m+1-i} + \varepsilon_{m+1-j}. \end{aligned}$$

Hence, the above expressions describe all elements of Φ''.

• In case (c), $\Phi = \Phi' \cup \Phi'' \cup \{\pm 2\varepsilon_i \mid 1 \leqslant i \leqslant m\}$. If we also set $\alpha_1 := 2\varepsilon_m$, then $\alpha_1, \alpha_2, \ldots, \alpha_m$ are linearly independent. We have

$$\alpha_1 + 2(\alpha_2 + \ldots + \alpha_i) = 2\varepsilon_m + 2(\varepsilon_{m+1-i} - \varepsilon_m) = 2\varepsilon_{m+1-i}$$

for $1 \leqslant i \leqslant m$. Furthermore, for $1 \leqslant i < j \leqslant m$, we obtain

$$\begin{aligned} \alpha_1 + &2(\alpha_2 + \ldots + \alpha_i) + \alpha_{i+1} + \alpha_{i+2} + \ldots + \alpha_j \\ &= 2\varepsilon_{m+1-i} + (\varepsilon_{m+1-j} - \varepsilon_{m+1-i}) = \varepsilon_{m+1-i} + \varepsilon_{m+1-j}. \end{aligned}$$

Hence, again, the above expressions describe all elements of Φ''.

Corollary 2.5.6. *Let $L = \mathfrak{go}_n(\mathbb{C})$. Then, with notation as in Remark 2.5.5, $\Delta := \{\alpha_1, \ldots, \alpha_m\}$ is a basis of H^* and each $\alpha \in \Phi$ can be written as $\alpha = \pm \sum_{1 \leqslant i \leqslant m} n_i \alpha_i$ with $n_i \in \{0, 1, 2\}$ for all i.*

Proof. We already noted that $\{\alpha_1, \ldots, \alpha_m\}$ is linearly independent. The required expressions of α are explicitly given above. □

Remark 2.5.7. Let $x \in L = \mathfrak{go}_n(Q_n, \mathbb{C})$ and write $x = h + n^+ + n^-$ as in Corollary 1.6.8. Then one easily checks that our choice of $\alpha_1, \ldots, \alpha_m$ in Remark 2.5.5 is such that $n^\pm \in \sum_\alpha L_{\pm\alpha}$ where the sum runs over all $\alpha \in \Phi$ such that $\alpha = \sum_{1 \leqslant i \leqslant m} n_i \alpha_i$ with $n_i \geqslant 0$.

Table 2. Structure matrices A for the Lie algebras $L = \mathfrak{go}_n(Q_n, \mathbb{C})$

$$\begin{pmatrix} 2 & 0 & -1 & & & & \\ 0 & 2 & -1 & & & & \\ -1 & -1 & 2 & -1 & & & \\ & & -1 & 2 & -1 & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & -1 & 2 & -1 \\ & & & & & -1 & 2 \end{pmatrix} \qquad (Q_n^{\mathrm{tr}} = Q_n \text{ and } n = 2m),$$

$$\begin{pmatrix} 2 & -2 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 2 & -1 & & & & \\ -2 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix}.$$

$(Q_n^{\mathrm{tr}} = Q_n \text{ and } n = 2m+1)$ $\qquad$ $(Q_n^{\mathrm{tr}} = -Q_n \text{ and } n = 2m)$

Proposition 2.5.8. *Let $L = \mathfrak{go}_n(Q_n, \mathbb{C})$ and $H \subseteq L$ be as above; also write $n = 2m+1$ or $n = 2m$ with $m \geqslant 1$. Then (L, H) is of Cartan–Killing type with respect to $\Delta = \{\alpha_1, \ldots, \alpha_m\} \subseteq H^*$, as defined in Remark 2.5.5; the structure matrix A is given in Table 2. (Each of those matrices has size $m \times m$.)*

Proof. We already noted that L is H-diagonalisable and $C_L(H) = H$; hence, (CK1) in Definition 2.2.1 holds. Furthermore, (CK2) holds by Corollary 2.5.6. It remains to verify (CK3). For this purpose, we specify $e_i \in L_{\alpha_i}$ and $f_i \in L_{-\alpha_i}$ such that $\alpha_i(h_i) = 2$, where $h_i := [e_i, f_i] \in H$. For $2 \leqslant i \leqslant m$, we have $\alpha_i = \varepsilon_{m+1-i} - \varepsilon_{m+2-i}$, or $\alpha_{m+2-i} = \varepsilon_{i-1} - \varepsilon_i$. So Lemma 2.5.1 shows that

$$\begin{aligned} e_{m+2-i} &:= \delta_{i-1} A_{i-1,i} \in L_{\alpha_{m+2-i}}, \\ f_{m+2-i} &:= \delta_i A_{i,i-1} \quad\ \in L_{-\alpha_{m+2-i}}. \end{aligned}$$

Using the formulae in Remark 2.5.2, we find that

$$h_{m+2-i} := [e_{m+2-i}, f_{m+2-i}] = h(0, \ldots, 0, 1, -1, 0, \ldots, 0) \in H,$$

where the entry 1 is at the $(i-1)$-th position and -1 is at the i-th position. Hence, $\alpha_i(h_i) = 2$ for $2 \leqslant i \leqslant m$, as required.

If $Q_n^{\text{tr}} = Q_n$ and $n = 2m$, then we have $\alpha_1 = \varepsilon_{m-1} + \varepsilon_m$. As in the proof of Lemma 2.5.3(a), we see that

$$e_1 := \delta_{m-1} A_{m-1,m+1} \in L_{\alpha_1} \quad \text{and} \quad f_1 := \delta_{m+1} A_{m+1,m-1} \in L_{-\alpha_1}.$$

Using Remark 2.5.2, we find that $h_1 := [e_1, f_1] = h(0, \ldots, 0, 1, 1) \in H$ and $\alpha_1(h_1) = 2$, as required. If $Q_n^{\text{tr}} = Q_n$ and $n = 2m+1$, then we have $\alpha_1 = \varepsilon_m$. As in the proof of Lemma 2.5.3(b), we see that

$$e_1 := \delta_m A_{m,m+1} \in L_{\alpha_1} \quad \text{and} \quad f_1 := 2\delta_{m+1} A_{m+1,m} \in L_{-\alpha_1}.$$

Now $h_1 := [e_1, f_1] = h(0, \ldots, 0, 2) \in H$ and $\alpha_1(h_1) = 2$, as required. Finally, if $Q_n^{\text{tr}} = -Q_n$ and $n = 2m$, then we have $\alpha_1 = 2\varepsilon_m$. As in the proof of Lemma 2.5.3(c), we see that

$$e_1 := \tfrac{1}{2}\delta_m A_{m,m+1} \in L_{\alpha_1}, \qquad f_1 := \tfrac{1}{2}\delta_{m+1} A_{m+1,m} \in L_{-\alpha_1}.$$

Now $h_1 := [e_1, f_1] = h(0, \ldots, 0, 1) \in H$ and $\alpha_1(h_1) = 2$, as required.

In all cases, we see that $H = \langle h_1, \ldots, h_m \rangle_{\mathbb{C}}$ and so (CK3) holds. Finally, A is obtained by evaluating $\alpha_j(h_i)$ for all i, j. □

Theorem 2.5.9. *Recall that $n \geqslant 3$. If $Q_n^{\text{tr}} = Q_n$ and n is even, also assume that $n \geqslant 6$. Then $L = \mathfrak{go}_n(Q_n, \mathbb{C})$ is a simple Lie algebra.* (Note that, by Exercise 1.6.4(c), we really do have to exclude the case where $n = 4$ and $Q_4 = Q_4^{\text{tr}}$.)

Proof. By Proposition 2.5.8, (L, H) is of Cartan–Killing type with respect to $\Delta = \{\alpha_1, \ldots, \alpha_m\}$. We now use Remark 2.4.15 to show that L is simple (exactly as for $L = \mathfrak{sl}_n(\mathbb{C})$ in Example 2.4.16). Assume first that $Q_n^{\text{tr}} = Q_n$ and $n = 2m$, where $m \geqslant 3$. Then the explicit description of Φ in Remark 2.5.5 shows that

$$\alpha_1 + \alpha_2 + 2(\alpha_3 + \ldots + \alpha_{m-1}) + \alpha_m \in \Phi \quad \text{if } m \geqslant 4,$$

and $\alpha_1 + \alpha_2 + \alpha_3 \in \Phi$ if $m = 3$. Similarly, we have

$$2(\alpha_1 + \alpha_2 + \ldots + \alpha_{m-1}) + \alpha_m \in \Phi \quad \text{if } n = 2m+1 \text{ and } Q_n^{\text{tr}} = Q_n,$$
$$\alpha_1 + 2(\alpha_2 + \ldots + \alpha_{m-1} + \alpha_m) \in \Phi \quad \text{if } n = 2m \text{ and } Q_n^{\text{tr}} = -Q_n.$$

Hence, in each case, L is simple. □

Using the above descriptions of Φ, it is now not too difficult to describe the Weyl group of L. We will only do this here for the Lie algebra $\mathfrak{sl}_n(\mathbb{C})$.

Remark 2.5.10. Let us determine the Weyl group of the Lie algebra $L = \mathfrak{sl}_n(\mathbb{C})$, where $n \geqslant 2$. For this purpose, we use the inclusion $L \subseteq \hat{L} = \mathfrak{gl}_n(\mathbb{C})$. Let $\hat{H} := \{\mathrm{diag}(x_1, \ldots, x_n) \mid x_i \in \mathbb{C}\} \subseteq \hat{L}$ be the subspace of all diagonal matrices in $\hat{L}$. For $1 \leqslant i \leqslant n$, let $\hat{\varepsilon}_i \in \hat{H}^*$ be the map that sends a diagonal matrix to its i-th diagonal entry. Then $\{\hat{\varepsilon}_1, \ldots, \hat{\varepsilon}_n\}$ is a basis of $\hat{H}^*$. Another basis is given by $\{\delta, \hat{\alpha}_1, \ldots, \hat{\alpha}_{n-1}\}$ where

$$\delta := \hat{\varepsilon}_1 + \ldots + \hat{\varepsilon}_n \quad \text{and} \quad \hat{\alpha}_i := \hat{\varepsilon}_i - \hat{\varepsilon}_{i+1} \quad \text{for } 1 \leqslant i \leqslant n-1.$$

Now consider the Weyl group $W = \langle s_1, \ldots, s_{n-1} \rangle \subseteq H^*$ of L, where $H = \ker(\delta) \subseteq \hat{H}$. We define a map $\pi \colon W \to \mathrm{GL}(\hat{H}^*)$ as follows. Let $w \in W$ and write $w(\alpha_j) = \sum_i m_{ij}(w)\alpha_i$ with $m_{ij}(w) \in \mathbb{Z}$ for $1 \leqslant i, j \leqslant n-1$. Thus, $M_w = \big(m_{ij}(w)\big) \in \mathrm{GL}_{n-1}(\mathbb{C})$ is the matrix of w with respect to the basis $\Delta = \{\alpha_1, \ldots, \alpha_{n-1}\} \subseteq H^*$. Then we define $\hat{w} \in \mathrm{GL}(\hat{H}^*)$ by setting

$$\hat{w}(\delta) := \delta \quad \text{and} \quad \hat{w}(\hat{\alpha}_j) := \sum_{1 \leqslant i \leqslant n-1} m_{ij}(w)\hat{\alpha}_i \quad \text{for } 1 \leqslant j \leqslant n-1.$$

Thus, the matrix of $\hat{w}$ with respect to the basis $\{\delta, \hat{\alpha}_1, \ldots, \hat{\alpha}_{n-1}\}$ of $\hat{H}^*$ is a block diagonal matrix of the following shape:

$$\left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & M_w \end{array}\right).$$

Now $\pi \colon W \to \mathrm{GL}(\hat{H}^*)$, $w \mapsto \hat{w}$, is an injective group homomorphism, and we have $\pi(W) = \langle \hat{s}_1, \ldots, \hat{s}_{n-1} \rangle$. Since $\delta(h_i) = 0$ for all i, we see that $\hat{s}_i \colon \hat{H}^* \to \hat{H}^*$ is given by the formula

$$\hat{s}_i(\mu) = \mu - \mu(h_i)\hat{\alpha}_i \qquad \text{for all } \mu \in \hat{H}^*.$$

A straightforward computation shows that

$$\hat{s}_i(\hat{\varepsilon}_i) = \hat{\varepsilon}_{i+1}, \quad \hat{s}_i(\hat{\varepsilon}_{i+1}) = \hat{\varepsilon}_i \quad \text{and} \quad \hat{s}_i(\hat{\varepsilon}_j) = \hat{\varepsilon}_j \text{ if } j \notin \{i, i+1\}.$$

Thus, the matrix of $\hat{s}_i$ with respect to the basis $\{\hat{\varepsilon}_1, \ldots, \hat{\varepsilon}_n\}$ of $\hat{H}^*$ is the permutation matrix corresponding to the transposition in the symmetric group $\mathfrak{S}_n$ that exchanges i and $i+1$. Since $\mathfrak{S}_n$ is generated by these transpositions, we conclude that $W \cong \pi(W) \cong \mathfrak{S}_n$.

ab hier Woche 7

2.6. The structure constants $N_{\alpha,\beta}$

Returning to the general situation, let again (L, H) be of Cartan–Killing type with respect to $\Delta = \{\alpha_i \mid i \in I\}$. Let $\Phi \subseteq H^*$ be the set of roots of L and fix a collection of elements

$$\{0 \neq e_\alpha \in L_\alpha \mid \alpha \in \Phi\}.$$

Then, since $\dim L_\alpha = 1$ for all $\alpha \in \Phi$, the set

$$\{h_i \mid i \in I\} \cup \{e_\alpha \mid \alpha \in \Phi\} \quad \text{is a basis of } L.$$

If $\alpha, \beta \in \Phi$ are such that $\alpha + \beta \in \Phi$, then $[L_\alpha, L_\beta] \subseteq L_{\alpha+\beta}$ and

$$[e_\alpha, e_\beta] = N_{\alpha,\beta} e_{\alpha+\beta}, \qquad \text{where} \qquad N_{\alpha,\beta} \in \mathbb{C}.$$

The knowledge of the structure constants $N_{\alpha,\beta}$ is, of course, crucial for doing explicit computations inside L. Eventually, one would hope to find purely combinatorial formulae for $N_{\alpha,\beta}$ in terms of properties of Φ. In this section, we establish some basic properties of the $N_{\alpha,\beta}$.

It will be convenient to set $N_{\alpha,\beta} := 0$ if $\alpha + \beta \notin \Phi$.

Remark 2.6.1. Let $\alpha \in \Phi$. By Proposition 2.4.3, there is a unique $h_\alpha \in [L_\alpha, L_{-\alpha}]$ such that $\alpha(h_\alpha) = 2$. Now recall that $\Phi = -\Phi$. We claim that the elements $\{e_\alpha \mid \alpha \in \Phi\}$ can be adjusted such that

(a) $$[e_\alpha, e_{-\alpha}] = h_\alpha \qquad \text{for all } \alpha \in \Phi.$$

Indeed, we have $\Phi = \Phi^+ \cup \Phi^-$ (disjoint union), where $\Phi^- = -\Phi^+$. Let $\alpha \in \Phi^+$. Then $[e_\alpha, e_{-\alpha}] = \xi h_\alpha$ for some $0 \neq \xi \in \mathbb{C}$. Hence, replacing $e_{-\alpha}$ by a suitable scalar multiple if necessary, we can achieve that $[e_\alpha, e_{-\alpha}] = h_\alpha$. Thus, the desired relation holds for all $\alpha \in \Phi^+$. Now let $\beta \in \Phi^-$; then $\alpha = -\beta \in \Phi^+$. So $[e_\beta, e_{-\beta}] = -[e_\alpha, e_{-\alpha}] = -h_\alpha = h_\beta$, where the last equality holds by Proposition 2.4.3. So (a) holds in general. Now, writing $f_\alpha := e_{-\alpha}$ we have $[e_\alpha, f_\alpha] = h_\alpha$, $[h_\alpha, e_\alpha] = \alpha(h_\alpha)e_\alpha = 2e_\alpha$ and $[h_\alpha, f_\alpha] = -2f_\alpha$. Hence, as in Remark 2.2.10, we obtain a 3-dimensional subalgebra

(b) $$S_\alpha := \langle e_\alpha, h_\alpha, f_\alpha \rangle_{\mathbb{C}} \subseteq L \quad \text{such that} \quad S_\alpha \cong \mathfrak{sl}_2(\mathbb{C}).$$

Regarding L as an S_α-module, we obtain results completely analogous to those in Remark 2.2.10. Here is a first example. As in Section 2.3, let $E := \langle \alpha_i \mid i \in I \rangle_{\mathbb{R}} \subseteq H^*$ and $\langle \;,\; \rangle \colon E \times E \to \mathbb{R}$ be a W-invariant scalar product, where W is the Weyl group of (L, H).

Lemma 2.6.2. *Let $\alpha \in \Phi$. Then we have*

$$\lambda(h_\alpha) = 2\frac{\langle \alpha, \lambda\rangle}{\langle \alpha, \alpha\rangle} \qquad \text{for all } \lambda \in E.$$

Furthermore, if $\beta \in \Phi$ is such that $\beta \neq \pm\alpha$, then $\beta(h_\alpha) = q - p \in \mathbb{Z}$, where $p, q \geqslant 0$ are defined by the condition that

$$\beta - q\alpha, \quad \ldots, \quad \beta - \alpha, \quad \beta, \quad \beta + \alpha, \quad \ldots, \quad \beta + p\alpha$$

all belong to Φ, but $\beta + (p+1)\alpha \notin \Phi$ and $\beta - (q+1)\alpha \notin \Phi$.

In analogy to Remark 2.2.10, the above sequence of roots is called the *α-string through β*. The element h_α is also called a *co-root* of L.

Proof. We write $\alpha = w(\alpha_i)$, where $w \in W$ and $i \in I$. Applying w^{-1} to the above sequence of roots and setting $\beta' := w^{-1}(\beta)$, we see that

$$\beta' - q\alpha_i, \quad \ldots, \quad \beta' - \alpha_i, \quad \beta', \quad \beta' + \alpha_i, \quad \ldots, \quad \beta' + p\alpha_i$$

all belong to Φ. If we had $\beta' + (p+1)\alpha_i \in \Phi$, then also $\beta + (p+1)\alpha = w(\beta' + (p+1)\alpha_i) \in \Phi$, contradiction. Similarly, we have $\beta' - (q+1)\alpha_i \notin \Phi$. Hence, the above sequence is the α_i-string through β' and so $\beta'(h_i) = q - p$; see Remark 2.2.10(a). Using the W-invariance of $\langle\ ,\ \rangle$ and the formula in Remark 2.3.3, we obtain that

$$2\frac{\langle \alpha, \beta\rangle}{\langle \alpha, \alpha\rangle} = 2\frac{\langle w(\alpha_i), w(\beta')\rangle}{\langle w(\alpha_i), w(\alpha_i)\rangle} = 2\frac{\langle \alpha_i, \beta'\rangle}{\langle \alpha_i, \alpha_i\rangle} = \beta'(h_i) = q - p.$$

Furthermore, using $S_\alpha = \langle h_\alpha, e_\alpha, f_\alpha\rangle_{\mathbb{C}} \subseteq L$ as above, one sees that $\beta(h_\alpha) = q - p$, exactly as in Remark 2.2.10(a) (where $e_\alpha, h_\alpha, f_\alpha$ play the role of e_i, h_i, f_i, respectively). Hence, the formula $\lambda(h_\alpha) = 2\frac{\langle \alpha, \lambda\rangle}{\langle \alpha, \alpha\rangle}$ holds for all $\lambda \in \Phi$ such that $\lambda \neq \pm\alpha$. By the definition of h_α, it also holds for $\lambda = \pm\alpha$. Finally, since $E = \langle \Phi\rangle_{\mathbb{R}}$, it holds in general. □

Lemma 2.6.3. *Let $\alpha \in \Phi$ and write $\alpha = \sum_{i \in I} n_i \alpha_i$ with $n_i \in \mathbb{Z}$. Then $h_\alpha = \sum_{i \in I} n_i^\vee h_i$, where*

$$n_i^\vee = \frac{\langle \alpha_i, \alpha_i\rangle}{\langle \alpha, \alpha\rangle} n_i \in \mathbb{Z} \qquad \text{for all } i \in I.$$

Proof. Given the expression $\alpha = \sum_{i \in I} n_i \alpha_i$, we obtain

$$\frac{2\alpha}{\langle \alpha, \alpha\rangle} = \sum_{i \in I} n_i \frac{2}{\langle \alpha, \alpha\rangle} \frac{\langle \alpha_i, \alpha_i\rangle}{\langle \alpha_i, \alpha_i\rangle} \alpha_i = \sum_{i \in I} n_i \frac{\langle \alpha_i, \alpha_i\rangle}{\langle \alpha, \alpha\rangle} \frac{2\alpha_i}{\langle \alpha_i, \alpha_i\rangle}.$$

Now let $\lambda \in E$. Using the formula in Lemma 2.6.2, we obtain:

$$\lambda(h_\alpha) = \sum_{i\in I} n_i \frac{\langle \alpha_i, \alpha_i\rangle}{\langle \alpha, \alpha\rangle} \lambda(h_i) = \lambda\Big(\sum_{i\in I} n_i \frac{\langle \alpha_i, \alpha_i\rangle}{\langle \alpha, \alpha\rangle} h_i\Big).$$

Since this holds for all λ, we obtain the desired formula. The fact that the coefficients $n_i^\vee$ are integers follows from Exercise 2.4.4. □

Remark 2.6.4. In the following discussion, we assume throughout that (a) in Remark 2.6.1 holds, that is, we have $[e_\alpha, e_{-\alpha}] = h_\alpha$ for all $\alpha \in \Phi$. This assumption leads to the following summary about the Lie brackets in L. We have:

$$\begin{aligned}
[h_i, h_j] &= 0, && \text{for all } i, j \in I,\\
[h_i, e_\alpha] &= \alpha(h_i) e_\alpha, && \text{where } \alpha(h_i) \in \mathbb{Z},\\
[e_\alpha, e_{-\alpha}] &= h_\alpha \in \langle h_i \mid i \in I\rangle_{\mathbb{Z}} && \text{(see Lemma 2.6.3)},\\
[e_\alpha, e_\beta] &= 0 && \text{if } \alpha + \beta \notin \Phi \cup \{\underline{0}\},\\
[e_\alpha, e_\beta] &= N_{\alpha,\beta} e_{\alpha+\beta} && \text{if } \alpha + \beta \in \Phi.
\end{aligned}$$

Since $\{h_i \mid i \in I\} \cup \{e_\alpha \mid \alpha \in \Phi\}$ is a basis of L, the above formulae completely determine the multiplication in L. At this point, the only unknown quantities in those formulae are the constants $N_{\alpha,\beta}$.

Lemma 2.6.5. *If $\gamma_1, \gamma_2, \gamma_3 \in \Phi$ are such that $\gamma_1 + \gamma_2 + \gamma_3 = \underline{0}$, then*

$$N_{\gamma_1,\gamma_2} = -N_{\gamma_2,\gamma_1} \qquad \text{and} \qquad \frac{N_{\gamma_1,\gamma_2}}{\langle \gamma_3, \gamma_3\rangle} = \frac{N_{\gamma_2,\gamma_3}}{\langle \gamma_1, \gamma_1\rangle} = \frac{N_{\gamma_3,\gamma_1}}{\langle \gamma_2, \gamma_2\rangle}.$$

Proof. Since $\gamma_1 + \gamma_2 = -\gamma_3 \in \Phi$, the anti-symmetry of $[\ ,\ ]$ immediately yields $N_{\gamma_1,\gamma_2} = -N_{\gamma_2,\gamma_1}$. Now, since also $\gamma_2 + \gamma_3 = -\gamma_1 \in \Phi$, we have $[e_{\gamma_2}, e_{\gamma_3}] = N_{\gamma_2,\gamma_3} e_{\gamma_2+\gamma_3} = N_{\gamma_2,\gamma_3} e_{-\gamma_1}$ and so

$$[e_{\gamma_1}, [e_{\gamma_2}, e_{\gamma_3}]] = N_{\gamma_2,\gamma_3} [e_{\gamma_1}, e_{-\gamma_1}] = N_{\gamma_2,\gamma_3} h_{\gamma_1},$$

where we used Remark 2.6.1(a). Similarly, we obtain that

$$[e_{\gamma_2}, [e_{\gamma_3}, e_{\gamma_1}]] = N_{\gamma_3,\gamma_1} h_{\gamma_2} \quad \text{and} \quad [e_{\gamma_3}, [e_{\gamma_1}, e_{\gamma_2}]] = N_{\gamma_1,\gamma_2} h_{\gamma_3}.$$

So the Jacobi identity $[e_{\gamma_1}, [e_{\gamma_2}, e_{\gamma_3}]] + [e_{\gamma_2}, [e_{\gamma_3}, e_{\gamma_1}]] + [e_{\gamma_3}, [e_{\gamma_1}, e_{\gamma_2}]] = 0$ yields the identity $N_{\gamma_2,\gamma_3} h_{\gamma_1} + N_{\gamma_3,\gamma_1} h_{\gamma_2} + N_{\gamma_1,\gamma_2} h_{\gamma_3} = 0$. Now apply any $\beta \in \Phi$ to the above relation. Using Lemma 2.6.2, we obtain

$$\begin{aligned} 2\Big\langle \beta, \frac{N_{\gamma_2,\gamma_3}}{\langle\gamma_1,\gamma_1\rangle}\gamma_1 + \frac{N_{\gamma_3,\gamma_1}}{\langle\gamma_2,\gamma_2\rangle}\gamma_2 + \frac{N_{\gamma_1,\gamma_2}}{\langle\gamma_3,\gamma_3\rangle}\gamma_3\Big\rangle \\ = \frac{2N_{\gamma_2,\gamma_3}\langle\beta,\gamma_1\rangle}{\langle\gamma_1,\gamma_1\rangle} + \frac{2N_{\gamma_3,\gamma_1}\langle\beta,\gamma_2\rangle}{\langle\gamma_2,\gamma_2\rangle} + \frac{2N_{\gamma_1,\gamma_2}\langle\beta,\gamma_3\rangle}{\langle\gamma_3,\gamma_3\rangle} \\ = \beta\big(N_{\gamma_2,\gamma_3}h_{\gamma_1} + N_{\gamma_3,\gamma_1}h_{\gamma_2} + N_{\gamma_1,\gamma_2}h_{\gamma_3}\big) = 0. \end{aligned}$$

Since this holds for all $\beta \in \Phi$ and since $E = \langle\Phi\rangle_{\mathbb{R}}$, we deduce that

$$\frac{N_{\gamma_2,\gamma_3}}{\langle\gamma_1,\gamma_1\rangle}\gamma_1 + \frac{N_{\gamma_3,\gamma_1}}{\langle\gamma_2,\gamma_2\rangle}\gamma_2 + \frac{N_{\gamma_1,\gamma_2}}{\langle\gamma_3,\gamma_3\rangle}\gamma_3 = \underline{0}.$$

Since $\gamma_3 = -\gamma_1 - \gamma_2$, we obtain

$$\Big(\frac{N_{\gamma_2,\gamma_3}}{\langle\gamma_1,\gamma_1\rangle} - \frac{N_{\gamma_1,\gamma_2}}{\langle\gamma_3,\gamma_3\rangle}\Big)\gamma_1 + \Big(\frac{N_{\gamma_3,\gamma_1}}{\langle\gamma_2,\gamma_2\rangle} - \frac{N_{\gamma_1,\gamma_2}}{\langle\gamma_3,\gamma_3\rangle}\Big)\gamma_2 = \underline{0}.$$

Now $\{\gamma_1, \gamma_2\}$ are linearly independent. For otherwise, we would have $\gamma_2 = \pm\gamma_1$ and so $\gamma_3 = -2\gamma_1$ or $\gamma_3 = \underline{0}$, contradiction. Hence, the coefficients of γ_1, γ_2 in the above equation must be zero. □

Lemma 2.6.6. *Let $\alpha, \beta \in \Phi$ be such that $\alpha + \beta \in \Phi$. Then*

$$N_{\alpha,\beta}N_{-\alpha,-\beta} = -p(q+1)\frac{\langle\alpha+\beta,\alpha+\beta\rangle}{\langle\beta,\beta\rangle},$$

where $\beta - q\alpha, \ldots, \beta-\alpha, \beta, \beta+\alpha, \ldots, \beta+p\alpha$ is the α-string through β. In particular, this shows that $N_{\alpha,\beta} \neq 0$ (since $p \geqslant 1$ by assumption).

Proof. We have $[e_{-\alpha}, [e_\alpha, e_\beta]] = N_{\alpha,\beta}[e_{-\alpha}, e_{\alpha+\beta}] = N_{\alpha,\beta}N_{-\alpha,\alpha+\beta}e_\beta$. Applying Lemma 2.6.5 with $\gamma_1 = -\alpha$, $\gamma_2 = \alpha+\beta$, $\gamma_3 = -\beta$, we obtain

$$\frac{N_{-\alpha,\alpha+\beta}}{\langle\beta,\beta\rangle} = -\frac{N_{-\alpha,-\beta}}{\langle\alpha+\beta,\alpha+\beta\rangle}.$$

On the other hand, let $\mathfrak{sl}_2(\mathbb{C}) \cong S_\alpha = \langle e_\alpha, h_\alpha, f_\alpha\rangle \subseteq L$ as in Remark 2.6.1(b). Then, arguing as in Remark 2.2.10 (where $e_\alpha, h_\alpha, f_\alpha$ play the role of e_i, h_i, f_i, respectively), we find that

$$[e_{-\alpha}, [e_\alpha, e_\beta]] = [f_\alpha, [e_\alpha, e_\beta]] = p(q+1)e_\beta.$$

This yields the desired formula. □

There is also the following result involving four roots.

Lemma 2.6.7. *Assume that* $\beta_1, \beta_2, \gamma_1, \gamma_2 \in \Phi$ *are such that* $\beta_1+\beta_2 = \gamma_1 + \gamma_2 \in \Phi$ *and* $\beta_1 - \gamma_1 \not\in \Phi \cup \{\underline{0}\}$. *Then* $\beta_2 - \gamma_1 = \gamma_2 - \beta_1 \in \Phi$ *and*

$$N_{\beta_1,\beta_2} N_{-\gamma_1,-\gamma_2} = N_{\beta_1,\gamma_2-\beta_1} N_{-\gamma_1,\gamma_1-\beta_2} \frac{\langle \gamma_2, \gamma_2\rangle}{\langle \beta_2, \beta_2\rangle} \frac{\langle \gamma_1 - \beta_2, \gamma_1 - \beta_2\rangle}{\langle \beta_1 + \beta_2, \beta_1 + \beta_2\rangle}.$$

Proof. By the Jacobi identity we have

$$[e_{\beta_2}, [e_{\beta_1}, e_{-\gamma_1}]] + [e_{\beta_1}, [e_{-\gamma_1}, e_{\beta_2}]] + [e_{-\gamma_1}, [e_{\beta_2}, e_{\beta_1}]] = 0.$$

Now $[e_{\beta_1}, e_{-\gamma_1}] \in L_{\beta_1-\gamma_1}$ and, hence, $[e_{\beta_1}, e_{-\gamma_1}] = 0$ since $\beta_1 - \gamma_1 \not\in \Phi \cup \{\underline{0}\}$. So the first of the above summands is zero and we obtain:

$$(\dagger) \qquad [e_{-\gamma_1}, [e_{\beta_1}, e_{\beta_2}]] = -[e_{-\gamma_1}, [e_{\beta_2}, e_{\beta_1}]] = [e_{\beta_1}, [e_{-\gamma_1}, e_{\beta_2}]].$$

The left hand side of (†) evaluates to

$$\begin{aligned} [e_{-\gamma_1}, [e_{\beta_1}, e_{\beta_2}]] &= N_{\beta_1,\beta_2} [e_{-\gamma_1}, e_{\beta_1+\beta_2}] \\ &= N_{\beta_1,\beta_2} [e_{-\gamma_1}, e_{\gamma_1+\gamma_2}] = N_{\beta_1,\beta_2} N_{-\gamma_1,\gamma_1+\gamma_2} e_{\gamma_2}. \end{aligned}$$

Now $N_{\beta_1,\beta_2} \neq 0$ and $N_{-\gamma_1,\gamma_1+\gamma_2} \neq 0$ by Lemma 2.6.6. Hence, the left hand side of (†) is non-zero. So we must have $[e_{-\gamma_1}, e_{\beta_2}] \neq 0$, which means that $-\gamma_1 + \beta_2 \in \Phi$. Then, similarly, we find that

$$\begin{aligned} [e_{\beta_1}, [e_{-\gamma_1}, e_{\beta_2}]] &= N_{-\gamma_1,\beta_2} [e_{\beta_1}, e_{-\gamma_1+\beta_2}] \\ &= N_{-\gamma_1,\beta_2} [e_{\beta_1}, e_{\gamma_2-\beta_1}] = N_{-\gamma_1,\beta_2} N_{\beta_1,\gamma_2-\beta_1} e_{\gamma_2}. \end{aligned}$$

This yields $N_{\beta_1,\beta_2} N_{-\gamma_1,\gamma_1+\gamma_2} = N_{-\gamma_1,\beta_2} N_{\beta_1,\gamma_2-\beta_1}$. Finally, we have

$$\frac{N_{-\gamma_1,\beta_2}}{\langle \gamma_1 - \beta_2, \gamma_1 - \beta_2\rangle} = \frac{N_{\gamma_1-\beta_2,-\gamma_1}}{\langle \beta_2, \beta_2\rangle} = -\frac{N_{-\gamma_1,\gamma_1-\beta_2}}{\langle \beta_2, \beta_2\rangle},$$

using Lemma 2.6.5 with $(-\gamma_1) + \beta_2 + (\gamma_1 - \beta_2) = \underline{0}$. Furthermore,

$$\frac{N_{-\gamma_1,\gamma_1+\gamma_2}}{\langle \gamma_2, \gamma_2\rangle} = \frac{N_{-\gamma_2,-\gamma_1}}{\langle \gamma_1 + \gamma_2, \gamma_1 + \gamma_2\rangle} = -\frac{N_{-\gamma_1,-\gamma_2}}{\langle \gamma_1 + \gamma_2, \gamma_1 + \gamma_2\rangle},$$

using Lemma 2.6.5 with $(-\gamma_1) + (\gamma_1 + \gamma_2) + (-\gamma_2) = \underline{0}$. □

As observed by Chevalley [**9**, p. 23], the right hand side of the formula in Lemma 2.6.6 can be simplified, as follows. Let $\alpha, \beta \in \Phi$ be such that $\beta \neq \pm\alpha$. Define $p, q \geqslant 0$ as in Lemma 2.6.2. Then

$$2\frac{\langle \alpha, \beta\rangle}{\langle \alpha, \alpha\rangle} = \beta(h_\alpha) = q - p \in \mathbb{Z}.$$

To simplify the notation, let us denote $\lambda^\vee := 2\lambda/\langle\lambda,\lambda\rangle \in E$ for any $0 \neq \lambda \in E$. Thus, $\langle\alpha^\vee,\beta\rangle = q-p$. Now, by the *Cauchy–Schwartz inequality*, we have $0 \leqslant \langle\alpha,\beta\rangle^2 < \langle\alpha,\alpha\rangle\cdot\langle\beta,\beta\rangle$. This yields that

$$0 \leqslant \langle\alpha^\vee,\beta\rangle\cdot\langle\alpha,\beta^\vee\rangle = 2\frac{\langle\alpha,\beta\rangle}{\langle\alpha,\alpha\rangle}\cdot 2\frac{\langle\alpha,\beta\rangle}{\langle\beta,\beta\rangle} < 4.$$

Since $\langle\alpha^\vee,\beta\rangle$ and $\langle\alpha,\beta^\vee\rangle$ are integers, we conclude that

$$(\spadesuit_1) \qquad \langle\alpha^\vee,\beta\rangle = q-p \in \{0,\pm1,\pm2,\pm3\},$$

$$(\spadesuit_2) \qquad \langle\alpha^\vee,\beta\rangle = \pm2 \text{ or } \pm3 \quad\Rightarrow\quad \langle\alpha,\beta^\vee\rangle = \pm1.$$

Now let $\gamma := \beta - q\alpha \in \Phi$; note that also $\gamma \neq \pm\alpha$. Then one immediately sees that the α-string through γ is given by

$$\gamma, \quad \gamma+\alpha, \quad ,\ldots, \quad \gamma+(p+q)\alpha.$$

Applying ($\spadesuit_1$) to α,γ yields $\langle\alpha^\vee,\gamma\rangle = -(p+q) \in \{0,\pm1,\pm2,\pm3\}$. So

$$(\spadesuit_3) \qquad p+q = -\langle\alpha^\vee,\gamma\rangle \in \{0,1,2,3\}.$$

Now assume that $\alpha+\beta\in\Phi$, as in Lemma 2.6.6. Then we claim that

$$(\spadesuit_4) \qquad r := r(\alpha,\beta) = \frac{\langle\alpha+\beta,\alpha+\beta\rangle}{\langle\beta,\beta\rangle} = \frac{q+1}{p}.$$

This can now be proved as follows. By ($\spadesuit_3$), we have $0 \leqslant p+q \leqslant 3$. Since $\alpha+\beta\in\Phi$, we have $p \geqslant 1$. This leads to the following cases.

$\boxed{p=1,\ q=0 \text{ or } p=2,\ q=1.}$ Then $\langle\alpha^\vee,\beta\rangle = q-p = -1$, which means that $2\langle\alpha,\beta\rangle = -\langle\alpha,\alpha\rangle$. So $\langle\alpha+\beta,\alpha+\beta\rangle = \langle\alpha,\alpha\rangle + 2\langle\alpha,\beta\rangle + \langle\beta,\beta\rangle = \langle\beta,\beta\rangle$. Hence, $r=1$; we also have $(q+1)/p = 1$, as required.

$\boxed{p=1,\ q=1.}$ Then $\langle\alpha^\vee,\beta\rangle = q-p = 0$ and so $\langle\alpha^\vee,\gamma\rangle = -2$, where $\gamma := \beta-\alpha$. By ($\spadesuit_2$), we must have $\langle\alpha,\gamma^\vee\rangle = -1$ and so $2\langle\alpha,\gamma\rangle = -\langle\gamma,\gamma\rangle$. Since $\gamma = \beta-\alpha$, this yields $\langle\alpha,\alpha\rangle = \langle\beta,\beta\rangle$. Now $\langle\alpha^\vee,\beta\rangle = 0$ and so $\langle\alpha,\beta\rangle = 0$. Hence, we obtain $\langle\alpha+\beta,\alpha+\beta\rangle = \langle\alpha,\alpha\rangle + \langle\beta,\beta\rangle = 2\langle\beta,\beta\rangle$. Thus, we have $r=2$ which equals $(q+1)/p = 2$ as required.

$\boxed{p=1,\ q=2.}$ Then $\langle\alpha^\vee,\beta\rangle = q-p = 1$ and so $\langle\alpha^\vee,\gamma\rangle = -3$, where $\gamma := \beta-2\alpha$. By ($\spadesuit_2$), we must have $\langle\alpha,\gamma^\vee\rangle = -1$ and so $2\langle\alpha,\gamma\rangle = -\langle\gamma,\gamma\rangle$. Since $\gamma = \beta-2\alpha$, this yields that $2\langle\alpha,\beta\rangle = \langle\beta,\beta\rangle$. Now $\langle\alpha^\vee,\beta\rangle = 1$ also implies that $2\langle\alpha,\beta\rangle = \langle\alpha,\alpha\rangle$ and so $\langle\alpha,\alpha\rangle = \langle\beta,\beta\rangle$. Hence, we obtain $\langle\alpha+\beta,\alpha+\beta\rangle = \langle\alpha,\alpha\rangle + 2\langle\alpha,\beta\rangle + \langle\beta,\beta\rangle = 3\langle\beta,\beta\rangle$ and so $r=3$, which equals $(q+1)/p = 3$, as required.

$\boxed{p \geqslant 2,\ q = 0.}$ Then $\langle \alpha^\vee, \beta \rangle = -p \leqslant -2$ and so $\langle \alpha, \beta^\vee \rangle = -1$, by (♠$_2$). This yields $-p\langle \alpha, \alpha \rangle = 2\langle \alpha, \beta \rangle = -\langle \beta, \beta \rangle$ and so $\langle \alpha+\beta, \alpha+\beta \rangle = \langle \alpha, \alpha \rangle + 2\langle \alpha, \beta \rangle + \langle \beta, \beta \rangle = \frac{1}{p}\langle \beta, \beta \rangle$. Hence, $r = \frac{1}{p} = \frac{q+1}{p}$, as required.

Thus, the identity in (♠$_4$) holds in all cases and we obtain:

Proposition 2.6.8 (Chevalley). *Let $\alpha, \beta \in \Phi$ be such that $\alpha+\beta \in \Phi$. Using the notation in Lemma 2.6.6, we have*

$$N_{\alpha,\beta} N_{-\alpha,-\beta} = -(q+1)^2.$$

Proof. Since $\alpha + \beta \in \Phi$, we have $\beta \neq \pm\alpha$. We have seen above that then (♠$_4$) holds. It remains to use the formula in Lemma 2.6.6. □

The above formula suggests that there might be a clever choice of the elements $e_\alpha \in L_\alpha$ such that $N_{\alpha,\beta} = \pm(q+1)$ whenever $\alpha+\beta \in \Phi$. We will pursue this issue further in the following section.

Example 2.6.9. Suppose we know all $N_{\alpha_j,\beta}$, where $j \in I$ and $\beta \in \Phi^+$. We claim that then all structure constants $N_{\pm\alpha_i,\alpha}$ for $i \in I$ and $\alpha \in \Phi$ can be determined, using only manipulations with roots in Φ.

(1) First, let $i \in I$ and $\alpha \in \Phi^-$. Then Proposition 2.6.8 shows how to express $N_{-\alpha_i,\alpha}$ in terms of $N_{\alpha_i,-\alpha}$ (which is known by assumption).

(2) Next, we determine $N_{-\alpha_i,\alpha}$ for $i \in I$ and $\alpha \in \Phi^+$. If $\alpha - \alpha_i \notin \Phi$, then $N_{-\alpha_i,\alpha} = 0$. Now assume that $\alpha - \alpha_i \in \Phi$. Then $(-\alpha_i) + \alpha - (\alpha - \alpha_i) = \underline{0}$ and so Lemma 2.6.5 yields that

$$\frac{N_{-\alpha_i,\alpha}}{\langle \alpha - \alpha_i, \alpha - \alpha_i \rangle} = \frac{N_{-(\alpha-\alpha_i),-\alpha_i}}{\langle \alpha, \alpha \rangle} = -\frac{N_{-\alpha_i,-(\alpha-\alpha_i)}}{\langle \alpha, \alpha \rangle}.$$

Since $-(\alpha - \alpha_i) \in \Phi^-$, the right hand side can be handled by (1).

(3) Finally, if $i \in I$ and $\alpha \in \Phi^-$, then Proposition 2.6.8 expresses $N_{\alpha_i,\alpha}$ in terms of $N_{-\alpha_i,-\alpha}$, which is handled by (2) since $-\alpha \in \Phi^+$.

Of course, if we want to do this in a concrete example, then we need to be able to perform computations with roots in Φ: check if the sum of roots is again a root, or calculate the scalar product of a root with itself. More precisely, we do not need to know the actual values of those scalar products, but rather the values of fractions $r(\alpha, \beta) = \langle \alpha + \beta, \alpha + \beta \rangle / \langle \beta, \beta \rangle$ as above; we have seen in (♠$_4$) how such fractions are determined.

To illustrate the above results, let us consider the matrix

$$\begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix}.$$

In Example 2.3.10, we have computed corresponding "roots", although we do not know (yet) if there is a Lie algebra with the above matrix as structure matrix. We can now push this discussion a bit further. First, we explain why the above matrix plays a special role.

Example 2.6.10. Let $i, j \in I$, $i \neq j$. Since $\alpha_i - \alpha_j \notin \Phi$, we have

$$a_{ij} = 2\frac{\langle \alpha_i, \alpha_j \rangle}{\langle \alpha_i, \alpha_i \rangle} = -p$$

where $p = \max\{m \geqslant 0 \mid \alpha_j + p\alpha_i \in \Phi\}$; see Lemma 2.6.2 and Exercise 2.2.13. By ($\spadesuit_3$), we have $a_{ij} = -p \in \{0, -1, -2, -3\}$. Assume that A is indecomposable and $a_{ij} = -3$; then $a_{ji} = -1$ by ($\spadesuit_2$). We claim that then $|I| = 2$ and so

$$A = \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} \qquad \text{where} \qquad I = \{j, i\}.$$

This is seen as follows. Suppose that $|I| \geqslant 3$. Since A is indecomposable, there is some $k \in I \setminus \{i, j\}$ such that $a_{ik} \neq 0$ or $a_{jk} \neq 0$ (or both). Let $I' = \{k, j, i\}$ and consider the submatrix A' of A with rows and columns labelled by I'. Then

$$A' = \begin{pmatrix} 2 & a & b \\ a' & 2 & -1 \\ b' & -3 & 2 \end{pmatrix} \qquad \text{where } a, a', b, b' \in \mathbb{Z}_{\leqslant 0};$$

furthermore $aa' \geqslant 1$ or $bb' \geqslant 1$ (or both). We compute that $\det(A) = 2 - 2aa' - 2bb' - ab' - 3a'b \leqslant 0$, contradiction to Remark 2.3.12. So we must have $|I| = 2$ and A is given as above.

Example 2.6.11. Assume that there exists a Lie algebra L with subalgebra $H \subseteq L$ such that (L, H) is of Cartan–Kiling type with respect to $\Delta = \{\alpha_1, \alpha_2\}$ and corresponding structure matrix

$$A = \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} \qquad \text{(called of type } G_2\text{)}.$$

Then, as in Example 2.3.10, W is dihedral of order 12 and

$$\Phi^+ = \{\alpha_1,\ \alpha_2,\ \alpha_1 + \alpha_2,\ \alpha_1 + 2\alpha_2,\ \alpha_1 + 3\alpha_2,\ 2\alpha_1 + 3\alpha_2\}.$$

Table 3. Structure constants for type G_2

$N_{\alpha,\beta}$	10	01	11	12	13	23	−10	−01	−11	−12	−13	−23
10	.	1	.	.	1	.	∗	.	1	.	.	−1
01	−1	.	−2	−3	.	.	.	∗	−3	2	−1	.
11	.	2	.	−3	.	.	−1	3	∗	2	.	−1
12	.	3	3	.	.	.	.	2	−2	∗	1	−1
13	−1	.	.	.	.	.	.	1	.	−1	∗	−1
23	.	.	.	.	.	.	−1	.	1	−1	1	∗
−10	∗	.	−1	.	.	1	.	−1	.	.	−1	.
−01	.	∗	3	−2	1	.	1	.	2	3	.	.
−11	1	−3	∗	−2	.	1	.	−2	.	3	.	.
−12	.	−2	2	∗	−1	1	.	−3	−3	.	.	.
−13	.	−1	.	1	∗	1	1	.	.	.	.	.
−23	1	.	−1	1	−1	∗	.	.	.	.	.	.

(Here, e.g., −12 stands for $-(\alpha_1+2\alpha_2)\in\Phi$, and "∗" for h_α.)

We have $-\langle\alpha_1,\alpha_1\rangle = 2\langle\alpha_1,\alpha_2\rangle = -3\langle\alpha_2,\alpha_2\rangle$ and so $\langle\alpha_1,\alpha_1\rangle = 3\langle\alpha_2,\alpha_2\rangle$. From the computation in Example 2.3.10, we also see that

$$\Phi_1 := \{w(\alpha_1) \mid w\in W\} = \{\alpha_1, \alpha_1+3\alpha_2, 2\alpha_1+3\alpha_2\},$$
$$\Phi_2 := \{w(\alpha_2) \mid w\in W\} = \{\alpha_2, \alpha_1+\alpha_2, \alpha_1+2\alpha_2\}.$$

Thus, $\langle\alpha,\alpha\rangle/\langle\beta,\beta\rangle$ is known for all $\alpha,\beta\in\Phi$. Let $\{e_1,e_2,f_1,f_2\}$ be Chevalley generators for L. Let us try to determine a collection of elements $\{\mathbf{e}_\alpha \mid \alpha\in\Phi\}$ and the corresponding structure constants. Anticipating what we will do in the following section, let us set

$$\mathbf{e}_{\alpha_1}=e_1, \qquad \mathbf{e}_{\alpha_2}=-e_2, \qquad \mathbf{e}_{-\alpha_1}=f_1, \qquad \mathbf{e}_{-\alpha_2}=-f_2.$$

For $i\in I$ and $\alpha\in\Phi$, let $q_{i,\alpha} := \max\{m\geqslant 0 \mid \alpha-m\alpha_i\in\Phi\}$. In view of the formula in Proposition 2.6.8, we define successively:

$$\begin{array}{llll}
\mathbf{e}_{\alpha_1+\alpha_2} := & [e_1,\mathbf{e}_{\alpha_2}] & \in L_{\alpha_1+\alpha_2} & (q_{1,\alpha_2}=0),\\
\mathbf{e}_{\alpha_1+2\alpha_2} := & \frac{1}{2}[e_2,\mathbf{e}_{\alpha_1+\alpha_2}] & \in L_{\alpha_1+2\alpha_2} & (q_{2,\alpha_1+\alpha_2}=1),\\
\mathbf{e}_{\alpha_1+3\alpha_2} := & \frac{1}{3}[e_2,\mathbf{e}_{\alpha_1+2\alpha_2}] & \in L_{\alpha_1+3\alpha_2} & (q_{2,\alpha_1+2\alpha_2}=2),\\
\mathbf{e}_{2\alpha_1+3\alpha_2} := & [e_1,\mathbf{e}_{\alpha_1+3\alpha_2}] & \in L_{2\alpha_1+3\alpha_2} & (q_{1,\alpha_1+3\alpha_2}=0).
\end{array}$$

All these are non-zero by Lemma 2.6.6. Hence, for $\alpha\in\Phi^+$, there is a unique $\mathbf{e}_{-\alpha}\in L_{-\alpha}$ such that $[\mathbf{e}_\alpha,\mathbf{e}_{-\alpha}]=h_\alpha$. Thus, we have defined elements $\mathbf{e}_\alpha\in L_\alpha$ for all $\alpha\in\Phi$, such that Remark 2.6.1(a) holds. Let $N_{\alpha,\beta}$ be the corresponding structure constants; we leave it as an exercise for the reader to check that these are given by Table 3. (In order to compute that table, one only needs arguments like those in Example 2.6.9.) Thus, without knowing that L exists at all, we are able to compute all the structure constants $N_{\alpha,\beta}$ — and we see that they are all integers! Furthermore, using Lemma 2.6.3, we obtain

$$h_{\alpha_1+\alpha_2} = 3h_1 + h_2, \qquad h_{\alpha_1+2\alpha_2} = 3h_1 + 2h_2,$$
$$h_{\alpha_1+3\alpha_2} = h_1 + h_2, \qquad h_{2\alpha_1+3\alpha_2} = 2h_1 + h_2.$$

Thus, all the Lie brackets in L are explicitly known and the whole situation is completely rigid.

2.7. Lusztig's canonical basis

We keep the general setting of the previous section and assume now that the structure matrix A of L is indecomposable. The aim of this section is to show the remarkable fact that one can single out a "canonical" collection of elements in the various weight spaces L_α.

Remark 2.7.1. Let $i \in I$ and $\beta \in \Phi$ be such that $\beta \neq \pm\alpha_i$. As in Remark 2.2.10, let $\beta - q\alpha_i, \ldots, \beta - \alpha_i, \beta, \beta + \alpha_i, \ldots, \beta + p\alpha_i$ be the *α_i-string through β*. By the exercises, we have

$$p = p_{i,\beta} := \max\{m \geqslant 0 \mid \beta + m\alpha_i \in \Phi\},$$
$$q = q_{i,\beta} := \max\{m \geqslant 0 \mid \beta - m\alpha_i \in \Phi\}.$$

Also note that, for any $m \geqslant 0$, we have $\beta - m\alpha_i \in \Phi$ if and only if $-\beta + m\alpha_i = -(\beta - m\alpha_i) \in \Phi$. Thus, we have $q_{i,\beta} = p_{i,-\beta}$.

Theorem 2.7.2 (Lusztig [**24**, Theorem 0.6]). *Given Chevalley generators $\{e_i, f_i \mid i \in I\}$ of L, there is a collection of elements $\{0 \neq \mathbf{e}_\alpha^+ \in L_\alpha \mid \alpha \in \Phi\}$ with the following properties:*

(L1) $[f_i, \mathbf{e}_{\alpha_i}^+] = [e_i, \mathbf{e}_{-\alpha_i}^+]$ *for all $i \in I$.*

(L2) $[e_i, \mathbf{e}_\alpha^+] = (q_{i,\alpha} + 1)\mathbf{e}_{\alpha+\alpha_i}^+$ *if $i \in I$, $\alpha \in \Phi$ and $\alpha + \alpha_i \in \Phi$.*

(L3) $[f_i, \mathbf{e}_\alpha^+] = (p_{i,\alpha} + 1)\mathbf{e}_{\alpha-\alpha_i}^+$ *if $i \in I$, $\alpha \in \Phi$ and $\alpha - \alpha_i \in \Phi$.*

This collection $\{\mathbf{e}_\alpha^+ \mid \alpha \in \Phi\}$ is unique up to a global constant, that is, if $\{0 \neq \mathbf{e}'_\alpha \in L_\alpha \mid \alpha \in \Phi\}$ is another collection satisfying (L1)–(L3), *then there exists some $0 \neq \xi \in \mathbb{C}$ such that $\mathbf{e}'_\alpha = \xi\mathbf{e}_\alpha^+$ for all $\alpha \in \Phi$.*

The proof will be given later in this section, after the following remarks. First note that, even for $L = \mathfrak{sl}_2(\mathbb{C})$, we have to modify the standard elements e, h, f in order to obtain the above formulae. Indeed, setting $\mathbf{e}^+ := e$ and $\mathbf{f}^+ := -f$, we have

$$[e, \mathbf{f}^+] = -[e, f] = -h = [f, e] = [f, \mathbf{e}^+].$$

Hence, $\{\mathbf{e}^+, \mathbf{f}^+\}$ is a collection satisfying (L1); the conditions in (L2) and (L3) are empty in this case.

Remark 2.7.3. Assume that a collection $\{\mathbf{e}_\alpha^+ \mid \alpha \in \Phi\}$ as in Theorem 2.7.2 exists. Since $\mathbf{e}_{\alpha_i}^+ \in L_{\alpha_i}$ for $i \in I$, we have $\mathbf{e}_{\alpha_i}^+ = c_i e_i$, where $0 \neq c_i \in \mathbb{C}$. Similarly, we have $\mathbf{e}_{-\alpha_i}^+ \in L_{-\alpha_i}$ and so $\mathbf{e}_{-\alpha_i}^+ = d_i f_i$, where $0 \neq d_i \in \mathbb{C}$. Hence, we obtain

$$\begin{aligned} [f_i, \mathbf{e}_{\alpha_i}^+] &= c_i[f_i, e_i] = -c_i[e_i, f_i] = -c_i h_i, \\ [e_i, \mathbf{e}_{-\alpha_i}^+] &= d_i[e_i, f_i] = d_i h_i, \end{aligned}$$

and so (L1) implies that $d_i = -c_i$ for all $i \in I$. This also shows that $[\mathbf{e}_{\alpha_i}^+, \mathbf{e}_{-\alpha_i}^+] = -[e_i, f_i] = -h_i$ for $i \in I$. — Thus, Remark 2.6.1(a) does not hold for the collection $\{\mathbf{e}_\alpha^+ \mid \alpha \in \Phi\}$.

Now, the possibilities for the constants c_i are severely restricted, as follows. Let $i, j \in I$ be such that $i \neq j$ and $a_{ij} \neq 0$. Then $\beta = \alpha_i + \alpha_j \in \Phi$ (see exercises). Applying (L2) twice, we obtain:

$$\begin{aligned} [e_i, e_j] &= [e_i, c_j^{-1}\mathbf{e}_{\alpha_j}^+] = (q_{i,\alpha_j} + 1)c_j^{-1}\mathbf{e}_\beta^+ = c_j^{-1}\mathbf{e}_\beta^+, \\ [e_j, e_i] &= [e_j, c_i^{-1}\mathbf{e}_{\alpha_i}^+] = (q_{j,\alpha_i} + 1)c_i^{-1}\mathbf{e}_\beta^+ = c_i^{-1}\mathbf{e}_\beta^+. \end{aligned}$$

Note that $\pm(\alpha_i - \alpha_j) \not\in \Phi$ and so $q_{j,\alpha_i} = q_{i,\alpha_j} = 0$. Since $[e_i, e_j] = -[e_j, e_i]$, we conclude that $c_j = -c_i$. Thus

$$(*) \qquad c_j = -c_i \quad \text{whenever } i, j \in I \text{ are such that } a_{ij} < 0.$$

Since A is indecomposable, this implies that $\{c_i \mid i \in I\}$ is completely determined by c_{i_0}, for one particular choice of $i_0 \in I$. Indeed, let $i \in I$, $i \neq i_0$. By Remark 2.4.8, there is a sequence of distinct indices $i_0, i_1, \ldots, i_r = i$ ($r \geqslant 1$) such that $a_{i_l i_{l+1}} \neq 0$ for $0 \leqslant l \leqslant r-1$. Hence, using $(*)$, we find that $c_i = (-1)^r c_{i_0}$. Consequently, if $\{c_i' \mid i \in I\}$ is another collection of non-zero constants satisfying $(*)$, then $c_i' = \xi c_i$ for all $i \in I$, where $\xi = c_{i_0}' c_{i_0}^{-1} \in \mathbb{C}$ is a constant.

Remark 2.7.4. Assume that a collection $\{\mathbf{e}_\alpha^+ \mid \alpha \in \Phi\}$ as in Theorem 2.7.2 exists. Using (L1), we can define

$$h_j^+ := [e_j, \mathbf{e}_{-\alpha_j}^+] = [f_j, \mathbf{e}_{\alpha_j}^+] \in H \qquad \text{for all } j \in I.$$

Writing $\mathbf{e}_{\alpha_j}^+ = c_j e_j$ as in Remark 2.7.3, we see that $h_j^+ = -c_j h_j$. So

$$\mathbf{B} := \{h_j^+ \mid j \in I\} \cup \{\mathbf{e}_\alpha^+ \mid \alpha \in \Phi\} \quad \text{is a basis of } L.$$

We claim that the action of the Chevalley generators $\{e_i, f_i \mid i \in I\}$ on this basis is given as follows, where $j \in I$ and $\alpha \in \Phi$:

$$\boxed{\begin{array}{l} [e_i, h_j^+] = |a_{ji}|\mathbf{e}_{\alpha_i}^+, \qquad [f_i, h_j^+] = |a_{ji}|\mathbf{e}_{-\alpha_i}^+, \\ [e_i, \mathbf{e}_\alpha^+] = \left\{\begin{array}{cl} (q_{i,\alpha}+1)\mathbf{e}_{\alpha+\alpha_i}^+ & \text{if } \alpha+\alpha_i \in \Phi, \\ h_i^+ & \text{if } \alpha = -\alpha_i, \\ 0 & \text{otherwise,} \end{array}\right. \\ [f_i, \mathbf{e}_\alpha^+] = \left\{\begin{array}{cl} (p_{i,\alpha}+1)\mathbf{e}_{\alpha-\alpha_i}^+ & \text{if } \alpha-\alpha_i \in \Phi, \\ h_i^+ & \text{if } \alpha = \alpha_i, \\ 0 & \text{otherwise.} \end{array}\right. \end{array}}$$

Indeed, first let $\alpha \in \Phi$. If $\alpha + \alpha_i \not\in \Phi$, then $[e_i, \mathbf{e}_\alpha^+] = 0$; otherwise, $[e_i, \mathbf{e}_{\alpha_i}^+]$ is given by (L2). Similarly, if $\alpha - \alpha_i \not\in \Phi$, then $[f_i, \mathbf{e}_\alpha^+] = 0$; otherwise, $[f_i, \mathbf{e}_\alpha^+]$ is given by (L3). Now let $j \in I$. Then

$$[e_i, h_j^+] = -[h_j^+, e_i] = c_j[h_j, e_i] = c_j\alpha_i(h_j)e_i = c_j a_{ji} e_i.$$

If $i = j$, then $a_{ji} = 2$ and $c_j e_i = c_i e_i = \mathbf{e}_{\alpha_i}^+$; thus, $[e_i, h_i^+] = 2\mathbf{e}_{\alpha_i}^+$. Now let $i \neq j$. If $a_{ji} = 0$, then $[e_i, h_j^+] = 0$. If $a_{ji} \neq 0$, then $c_i = -c_j$ by Remark 2.7.3. So $[e_i, h_j^+] = -c_i a_{ji} e_i = -a_{ji}\mathbf{e}_{\alpha_i}^+$, where $a_{ji} < 0$. This yields the above formula for $[e_i, h_j^+]$. Finally, consider f_i. We have seen in Remark 2.7.3 that $\mathbf{e}_{-\alpha_i}^+ = -c_i f_i$. This yields that

$$[f_i, h_j^+] = -[h_j^+, f_i] = c_j[h_j, f_i] = -c_j\alpha_i(h_j)f_i = -c_j a_{ji} f_i.$$

Now we argue as before to obtain the formula for $[f_i, h_j^+]$.

Thus, all the entries of the matrices of $\text{ad}_L(e_i)$ and $\text{ad}_L(f_i)$ with respect to the basis $\{h_j^+ \mid j \in I\} \cup \{\mathbf{e}_\alpha^+ \mid \alpha \in \Phi\}$ of L are non-negative integers! This is one of the remarkable features of Lusztig's theory of "canonical bases" (see [**23**], [**24**] and further references there).

Remark 2.7.5. Assume that a collection $\{\mathbf{e}_\alpha^+ \mid \alpha \in \Phi\}$ as in Theorem 2.7.2 exists. First note that, if $0 \neq \xi \in \mathbb{C}$ is fixed and we set $\mathbf{e}'_\alpha := \xi \mathbf{e}^+_\alpha$ for all $\alpha \in \Phi$, then the new collection $\{\mathbf{e}'_\alpha \mid \alpha \in \Phi\}$ also satisfies (L1)–(L3). Conversely, we show that any two collections satisfying (L1)–(L3) are related by such a global constant ξ.

Now, as above, for $i \in I$ we have $\mathbf{e}^+_{\alpha_i} = c_i e_i$, where $0 \neq c_i \in \mathbb{C}$. Then (L2) combined with the Key Lemma 2.3.4 determines $\mathbf{e}^+_\alpha$ for all $\alpha \in \Phi^+$. Furthermore, as above, we have $\mathbf{e}^+_{-\alpha_i} = -c_i f_i$ for $i \in I$. But then (L3) also determines $\mathbf{e}^+_{-\alpha}$ for all $\alpha \in \Phi^+$. Thus, the whole collection $\{\mathbf{e}^+_\alpha \mid \alpha \in \Phi\}$ is completely determined by $\{c_i \mid i \in I\}$ and properties of Φ (e.g., the numbers $p_{i,\alpha}$, $q_{i,\alpha}$).

Now assume that $\{\mathbf{e}'_\alpha \mid \alpha \in \Phi\}$ is any other collection that satisfies (L1)–(L3). For $i \in I$, we have again $\mathbf{e}'_{\alpha_i} = c'_i e_i$, where $0 \neq c'_i \in \mathbb{C}$. Now both collections of constants $\{c_i \mid i \in I\}$ and $\{c'_i \mid i \in I\}$ satisfy $(*)$ in Remark 2.7.3. So there is some $0 \neq \xi \in \mathbb{C}$ such that $c'_i = \xi c_i$ for all $i \in I$. Hence, we have $\mathbf{e}'_{\alpha_i} = \xi \mathbf{e}^+_{\alpha_i}$ for all $i \in I$. But then the previous discussion shows that $\mathbf{e}'_\alpha = \xi \mathbf{e}^+_\alpha$ for all $\alpha \in \Phi$. This proves the uniqueness part of Theorem 2.7.2.

ab hier Woche 8

We now turn to the existence part of Theorem 2.7.2. We essentially follow Lusztig's argument in [**22**, Lemma 1.4], but there are some additional complications here, since Lusztig assumes that A is symmetric and $a_{ij} \in \{0, \pm 1\}$ for all $i \neq j$ in I. (In [**24**], the proof is based on general results on canonical bases in [**23**].)

Definition 2.7.6. We fix any total order $\sqsubseteq$ on I. (For example, let $|I| = n$ and write $I = \{i_1, \ldots, i_n\}$; then define $i_k \sqsubseteq i_l$ if $k \leqslant l$.) Let $\alpha_0 \in \Phi^+$ be the highest root in Φ; see Proposition 2.4.17. Let us fix a nonzero $\mathbf{e}_{\alpha_0} \in L_{\alpha_0}$. Then we construct a specific element $\mathbf{e}_\gamma \in L_\gamma$ for any $\gamma \in \Phi^+$ by downward induction on $\mathrm{ht}(\gamma)$, as follows. For $\gamma = \alpha_0$, we take the chosen $\mathbf{e}_{\alpha_0} \in L_{\alpha_0}$. Now let $\gamma \in \Phi^+$ be such that $\mathrm{ht}(\gamma) < \mathrm{ht}(\alpha_0)$. Since $\gamma \neq \alpha_0$, there exists some $j \in I$ such that $\gamma' := \gamma + \alpha_j \in \Phi^+$ (see Proposition 2.4.17). By Remark 2.2.10(c'), we have $\{0\} \neq [L_{-\alpha_j}, L_{\gamma'}] \subseteq L_\gamma$. So, since $\mathbf{e}_{\gamma'} \in L_{\gamma'}$ is already known by induction, we can define $0 \neq \mathbf{e}_\gamma \in L_\gamma$ by the condition that

$$[f_j, \mathbf{e}_{\gamma'}] = (p_{j,\gamma'} + 1)\mathbf{e}_\gamma.$$

Note that there may be several $j \in I$ such that $\gamma + \alpha_j \in \Phi^+$. In order to make a specific choice, we let $j = k(\gamma) := \min\{l \in I \mid \gamma + \alpha_l \in \Phi^+\}$, where the minimum is taken with respect to $\sqsubseteq$.

Once $\mathbf{e}_\gamma$ is defined for each $\gamma \in \Phi^+$, there is a unique $\mathbf{e}_{-\gamma} \in L_{-\gamma}$ such that $[\mathbf{e}_\gamma, \mathbf{e}_{-\gamma}] = h_\gamma$. Thus, we obtain a complete collection

$$\boxed{\{\mathbf{e}_\gamma \mid \gamma \in \Phi\} \quad \text{such that Remark 2.6.1(a) holds.}}$$

Let $N_{\alpha,\beta}$ be the structure constants with respect to the above collection; since Remark 2.6.1(a) holds (by construction), all the results in Section 2.6 can be used.

Remark 2.7.7. Let $i \in I$. Since $0 \neq \mathbf{e}_{\alpha_i} \in L_{\alpha_i}$, we have $\mathbf{e}_{\alpha_i} = c_i e_i$, where $0 \neq c_i \in \mathbb{C}$. Similarly, $\mathbf{e}_{-\alpha_i} = c_i' f_i$, where $0 \neq c_i' \in \mathbb{C}$. Since $h_{\alpha_i} = [\mathbf{e}_{\alpha_i}, \mathbf{e}_{-\alpha_i}] = c_i c_i' [e_i, f_i] = c_i c_i' h_i$, we conclude that $c_i' = c_i^{-1}$.

Now let $i_0 \in I$ be the smallest index with respect to $\sqsubseteq$. We start the above inductive procedure all over again with $\mathbf{e}_{\alpha_0}$ replaced by $c_{i_0}^{-1}\mathbf{e}_{\alpha_0}$. Then we obtain a new collection $\{\mathbf{e}_\gamma' \mid \gamma \in \Phi\}$, where $\mathbf{e}_\gamma' = c_{i_0}^{-1}\mathbf{e}_\gamma$ for all $\gamma \in \Phi^+$, and $\mathbf{e}_\gamma' = c_{i_0}\mathbf{e}_\gamma$ for all $\gamma \in \Phi^-$. Thus, replacing each $\mathbf{e}_\gamma$ by $\mathbf{e}_\gamma'$, we can achieve that $\mathbf{e}_{\alpha_{i_0}} = e_{i_0}$ and $\mathbf{e}_{-\alpha_{i_0}} = f_{i_0}$. (This normalisation will play a role at one point further below.)

The following result is the crucial step in the proof of Theorem 2.7.2. It shows that the collection of elements $\{\mathbf{e}_\gamma \mid \gamma \in \Phi\}$ does not depend at all on the choice of the total order $\sqsubseteq$ on I.

Lemma 2.7.8. *Let $\gamma \in \Phi^+$ and $i \in I$ be arbitrary such that $\alpha := \gamma + \alpha_i \in \Phi$. Then we also have $[f_i, \mathbf{e}_\alpha] = (p_{i,\alpha} + 1)\mathbf{e}_\gamma$.*

Proof. We proceed by downward induction on $\mathrm{ht}(\gamma)$. If $\gamma = \alpha_0$, then the condition is empty and so there is nothing to prove. Now let $\mathrm{ht}(\gamma) < \mathrm{ht}(\alpha_0)$ and $i \in I$ be such that $\alpha := \gamma + \alpha_i \in \Phi^+$. We also have $\beta := \gamma + \alpha_j \in \Phi$, where $j := k(\gamma)$. If $i = j$, then the desired formula holds by construction. Now assume that $i \neq j$. Then we have two expressions $-\alpha_i + \alpha = \gamma = -\alpha_j + \beta$. Since $\beta - \alpha = \alpha_j - \alpha_i \notin \Phi \cup \{\underline{0}\}$, we can apply Lemma 2.6.7 with $\beta_1 = -\alpha_i$, $\beta_2 = \alpha$, $\gamma_1 = -\alpha_j$, $\gamma_2 = \beta$. This yields the identity:

$$(\dagger_1) \qquad N_{-\alpha_i,\alpha} N_{\alpha_j,-\beta} = N_{-\alpha_i,\gamma'} N_{\alpha_j,-\gamma'} \frac{\langle \beta, \beta\rangle}{\langle \alpha, \alpha\rangle} \frac{\langle \gamma', \gamma'\rangle}{\langle \gamma, \gamma\rangle},$$

where $\gamma' := \alpha + \alpha_j = \beta + \alpha_i = \beta_2 - \gamma_1 = \gamma_2 - \beta_1 \in \Phi$. Now, one could try to simplify the right hand side using the formulae in the previous section. But there is a simple trick (taken from [**26**, §2.9, Lemma E]) to avoid such calculations. Namely, we can also apply Lemma 2.6.7 with $\beta_1 = \alpha_i$, $\beta_2 = -\alpha$, $\gamma_1 = \alpha_j$, $\gamma_2 = -\beta$. This yields the identity:

$$(\dagger_2) \qquad N_{\alpha_i,-\alpha} N_{-\alpha_j,\beta} = N_{\alpha_i,-\gamma'} N_{-\alpha_j,\gamma'} \frac{\langle \beta,\beta\rangle}{\langle \alpha,\alpha\rangle} \frac{\langle \gamma',\gamma'\rangle}{\langle \gamma,\gamma\rangle}.$$

Now, we have $\gamma' - \alpha_i = \beta$ and $\text{ht}(\beta) = \text{ht}(\gamma)+1$; similarly, $\gamma' - \alpha_j = \alpha$ and $\text{ht}(\alpha) = \text{ht}(\gamma) + 1$. So we can apply induction and obtain that

$$[f_i, \mathbf{e}_{\gamma'}] = (p_{i,\gamma'} + 1)\mathbf{e}_\beta \qquad \text{and} \qquad [f_j, \mathbf{e}_{\gamma'}] = (p_{j,\gamma'} + 1)\mathbf{e}_\alpha.$$

Using Remarks 2.7.1 and 2.7.7, the above formulae mean that

$$\begin{aligned} N_{-\alpha_i,\gamma'} &= c_i^{-1}(p_{i,\gamma'} + 1) = c_i^{-1}(q_{i,-\gamma'} + 1), \\ N_{-\alpha_j,\gamma'} &= c_j^{-1}(p_{j,\gamma'} + 1) = c_j^{-1}(q_{j,-\gamma'} + 1). \end{aligned}$$

But then the formula in Proposition 2.6.8 shows that $N_{\alpha_i,-\gamma'} = -c_i(q_{i,-\gamma'} + 1)$ and $N_{\alpha_j,-\gamma'} = -c_j(q_{j,-\gamma'} + 1)$. Hence, the right hand side of $(\dagger_1)$, multiplied by $c_i c_j^{-1}$, is equal to the right hand side of $(\dagger_2)$, multiplied by $c_i^{-1} c_j$. Consequently, an analogous relation holds between the left hand sides. Thus, we obtain:

$$c_i c_j^{-1} N_{-\alpha_i,\alpha} N_{\alpha_j,-\beta} = c_i^{-1} c_j N_{\alpha_i,-\alpha} N_{-\alpha_j,\beta}.$$

Since $j = k(\gamma)$, we have $[f_j, \mathbf{e}_\beta] = (p_{j,\beta} + 1)\mathbf{e}_\gamma$ and so $N_{-\alpha_j,\beta} = c_j^{-1}(p_{j,\beta} + 1) = c_j^{-1}(q_{j,-\beta} + 1)$. Hence, $N_{\alpha_j,-\beta} = -c_j(q_{j,-\beta} + 1)$ by Proposition 2.6.8. Inserting this into the above identity, we deduce that $N_{\alpha_i,-\alpha} = -c_i^2 N_{-\alpha_i,\alpha}$ and so $c_i N_{-\alpha_i,\alpha} = \pm(q_{i,-\alpha} + 1) = \pm(p_{i,\alpha} + 1)$, again by Proposition 2.6.8 and Remark 2.7.1. It remains to determine the sign. But this can be done using $(\dagger_1)$ and the formulae obtained above. Indeed, we have seen that

$$\begin{aligned} N_{\alpha_j,-\beta} &= -c_j(q_{j,-\beta} + 1), \\ N_{\alpha_j,-\gamma'} &= -c_j(q_{j,-\gamma'} + 1), \\ N_{-\alpha_i,\gamma'} &= +c_i^{-1}(q_{i,-\gamma'} + 1). \end{aligned}$$

Inserting this into $(\dagger_1)$, we obtain that

$$c_i N_{-\alpha_i,\alpha} = (q_{j,-\beta} + 1)^{-1}(q_{i,-\gamma'} + 1)(q_{j,-\gamma'} + 1) \frac{\langle \beta,\beta\rangle}{\langle \alpha,\alpha\rangle} \frac{\langle \gamma',\gamma'\rangle}{\langle \gamma,\gamma\rangle}.$$

All terms on the right hand side are positive numbers and so $c_i N_{-\alpha_i,\alpha}$ must be positive. Hence, we conclude that $N_{-\alpha_i,\alpha} = c_i^{-1}(p_{i,\alpha} + 1)$, and this yields $[f_i, \mathbf{e}_\alpha] = (p_{i,\alpha} + 1)\mathbf{e}_\gamma$, as desired. □

By the discussion in Example 2.6.9, the above result should now determine all $N_{\pm\alpha_i,\alpha}$ for $i \in I$ and $\alpha \in \Phi$. Concretely, we obtain:

Lemma 2.7.9. *Let $\alpha \in \Phi^+$ and $i \in I$ be such that $\alpha+\alpha_i \in \Phi$. Then $[e_i, \mathbf{e}_\alpha] = (q_{i,\alpha}+1)\mathbf{e}_{\alpha+\alpha_i}$.*

Proof. Set $\alpha' := \alpha+\alpha_i \in \Phi^+$ and write $[e_i, \mathbf{e}_\alpha] = c\,\mathbf{e}_{\alpha'}$, where $c \in \mathbb{C}$. By Lemma 2.7.8, we have $[f_i, \mathbf{e}_{\alpha'}] = (p_{i,\alpha'}+1)\mathbf{e}_\alpha$. Next note that

$$\begin{aligned} p_{i,\alpha} &= \max\{m \geqslant 0 \mid \alpha + m\alpha_i \in \Phi\} \\ &= \max\{m \geqslant 0 \mid \alpha' + (m-1)\alpha_i \in \Phi\}\} \\ &= \max\{m' \geqslant 0 \mid \alpha' + m'\alpha_i \in \Phi\} + 1 = p_{i,\alpha'} + 1. \end{aligned}$$

Hence, we have $[f_i, \mathbf{e}_{\alpha'}] = p_{i,\alpha}\mathbf{e}_\alpha$. Consequently, we obtain the identity $[f_i,[e_i,\mathbf{e}_\alpha]] = c[f_i,\mathbf{e}_{\alpha'}] = cp_{i,\alpha}\mathbf{e}_\alpha$. Since $\alpha \neq \pm\alpha_i$, we can apply Remark 2.2.10(c). This shows that the left hand side of the identity equals $p_{i,\alpha}(q_{i,\alpha}+1)\mathbf{e}_\alpha$. Hence, we have $c = q_{i,\alpha}+1$, as desired. □

Lemma 2.7.10. *Let $i \in I$ and $\alpha \in \Phi^-$ be negative.*

(a) *If $\alpha+\alpha_i \in \Phi$, then $[e_i,\mathbf{e}_\alpha] = -(q_{i,\alpha}+1)\mathbf{e}_{\alpha+\alpha_i}$.*
(b) *If $\alpha-\alpha_i \in \Phi$, then $[f_i,\mathbf{e}_\alpha] = -(p_{i,\alpha}+1)\mathbf{e}_{\alpha-\alpha_i}$.*

Proof. (a) Set $\beta := -\alpha \in \Phi^+$. Then $\beta-\alpha_i = -(\alpha+\alpha_i) \in \Phi$. Since $\mathrm{ht}(\beta) \geqslant 1$, we have $\mathrm{ht}(\beta-\alpha_i) \geqslant 0$ and so $\beta-\alpha_i \in \Phi^+$. By Lemma 2.7.8, we have $[f_i,\mathbf{e}_{-\alpha}] = [f_i,\mathbf{e}_\beta] = (p_{i,\beta}+1)\mathbf{e}_{-(\alpha+\alpha_i)}$ and so

$$N_{-\alpha_i,-\alpha} = c_i^{-1}(p_{i,\beta}+1) = c_i^{-1}(q_{i,\alpha}+1);$$

see Remarks 2.7.1 and 2.7.7. By Proposition 2.6.8, we obtain $N_{\alpha_i,\alpha} = -c_i(q_{i,\alpha}+1)$ and, hence, $[e_i,\mathbf{e}_\alpha] = -(q_{i,\alpha}+1)\mathbf{e}_{\alpha+\alpha_i}$.

(b) Set again $\beta := -\alpha \in \Phi^+$. Then $\beta+\alpha_i = -(\alpha-\alpha_i) \in \Phi$ and so Lemma 2.7.9 yields that $[e_i,\mathbf{e}_{-\alpha}] = [e_i,\mathbf{e}_\beta] = (q_{i,\beta}+1)\mathbf{e}_{\beta+\alpha_i}$. Thus, we have $N_{\alpha_i,\beta} = c_i(q_{i,\beta}+1)$, and Proposition 2.6.8 shows that $N_{-\alpha_i,\alpha} = N_{-\alpha_i,-\beta} = -c_i^{-1}(q_{i,\beta}+1)$; note again that $q_{i,\beta} = p_{i,\alpha}$. □

Thus, we have found explicit formulae for the structure constants $N_{\pm\alpha_i,\alpha}$, for all $i \in I$ and $\alpha \in \Phi$, summarized as follows:

$$\begin{aligned} [e_i,\mathbf{e}_\alpha] &= +(q_{i,\alpha}+1)\mathbf{e}_{\alpha+\alpha_i} && \text{if } \alpha \in \Phi^+ \text{ and } \alpha+\alpha_i \in \Phi, \\ [e_i,\mathbf{e}_\alpha] &= -(q_{i,\alpha}+1)\mathbf{e}_{\alpha+\alpha_i} && \text{if } \alpha \in \Phi^- \text{ and } \alpha+\alpha_i \in \Phi, \end{aligned}$$

$$[f_i, \mathbf{e}_\alpha] = +(p_{i,\alpha}+1)\mathbf{e}_{\alpha-\alpha_i} \qquad \text{if } \alpha \in \Phi^+ \text{ and } \alpha-\alpha_i \in \Phi,$$
$$[f_i, \mathbf{e}_\alpha] = -(p_{i,\alpha}+1)\mathbf{e}_{\alpha-\alpha_i} \qquad \text{if } \alpha \in \Phi^- \text{ and } \alpha-\alpha_i \in \Phi.$$

Hence, the signs are not yet right as compared to the desired formulae in Theorem 2.7.2. To fix this, we define for $\alpha \in \Phi$:

$$\boxed{\mathbf{e}_\alpha^+ := \begin{cases} \mathbf{e}_\alpha & \text{if } \alpha \in \Phi^+, \\ (-1)^{\mathrm{ht}(\alpha)}\mathbf{e}_\alpha & \text{if } \alpha \in \Phi^-. \end{cases}}$$

We claim that (L1), (L2), (L3) in Theorem 2.7.2 hold. First consider (L2). Let $i \in I$ and $\alpha \in \Phi$ be such that $\alpha+\alpha_i \in \Phi$. If $\alpha \in \Phi^+$, then $\mathbf{e}_\alpha^+ = \mathbf{e}_\alpha$ and the required formula holds. If $\alpha \in \Phi^-$, then $[e_i, \mathbf{e}_\alpha^+] = (-1)^{\mathrm{ht}(\alpha)}[e_i, \mathbf{e}_\alpha] = -(-1)^{\mathrm{ht}(\alpha)}(q_{i,\alpha}+1)\mathbf{e}_\alpha$; so the desired formula holds again, since $\mathbf{e}_{\alpha+\alpha_i}^+ = (-1)^{\mathrm{ht}(\alpha+\alpha_i)}\mathbf{e}_{\alpha+\alpha_i}$. The argument for (L3) is analogous. Now consider (L1). This relies on the normalisation in Remark 2.7.7. Let $i \in I$. Since $\mathrm{ht}(\alpha_i) = 1$, we have

$$\mathbf{e}_{\alpha_i}^+ = \mathbf{e}_{\alpha_i} = c_i e_i \qquad \text{and} \qquad \mathbf{e}_{-\alpha_i}^+ = -\mathbf{e}_{-\alpha_i} = -c_i^{-1} f_i.$$

Since (L2) is already known to hold, we can run the argument in Remark 2.7.3 and find that the c_i are all equal to each other, up to signs. Since $c_{i_0} = 1$ for at least one $i_0 \in I$ (see Remark 2.7.7), we conclude that $c_i = \pm 1$ for all $i \in I$. But then we obtain

$$[e_i, \mathbf{e}_{-\alpha_i}^+] = -c_i^{-1}[e_i, f_i] = -c_i^{-1} h_i,$$
$$[f_i, \mathbf{e}_{\alpha_i}^+] \;= +c_i[f_i, e_i] \;= -c_i[e_i, f_i] = -c_i h_i.$$

Since $c_i = \pm 1$, we have $c_i = c_i^{-1}$ and so the above two expressions are equal, as required. Thus, eventually, the proof of Theorem 2.7.2 is complete. — As a by-product, we also obtain:

Corollary 2.7.11. *There is a collection of elements* $\{\mathbf{e}_\alpha^+ \mid \alpha \in \Phi\}$ *satisfying* (L1)–(L3) *in Theorem 2.7.2 and such that*

$$[\mathbf{e}_\alpha^+, \mathbf{e}_{-\alpha}^+] = (-1)^{\mathrm{ht}(\alpha)} h_\alpha \qquad \textit{for all } \alpha \in \Phi.$$

Such a collection $\{\mathbf{e}_\alpha^+ \mid \alpha \in \Phi\}$ *is unique up to a global sign, that is, if* $\{\mathbf{e}'_\alpha \mid \alpha \in \Phi\}$ *is another collection satisfying* (L1)–(L3) *and the above identity, then there is some* $\xi = \pm 1$ *such that* $\mathbf{e}'_\alpha = \xi \mathbf{e}_\alpha^+$ *for all* $\alpha \in \Phi$. *We have* $\mathbf{e}_{\alpha_i}^+ = c_i e_i$ *and* $\mathbf{e}_{-\alpha_i}^+ = -c_i f_i$, *with* $c_i \in \{\pm 1\}$ *for all* $i \in I$.

Proof. Since $[\mathbf{e}_\alpha, \mathbf{e}_{-\alpha}] = h_\alpha$, the formula for $[\mathbf{e}_\alpha^+, \mathbf{e}_{-\alpha}^+]$ is clear by the definition of $\mathbf{e}_\alpha^+$ and the fact that $h_{-\alpha} = -h_\alpha$ for all $\alpha \in \Phi$. Now let $\{\mathbf{e}'_\alpha \mid \alpha \in \Phi\}$ be another collection satisyfing (L1)–(L3) and the above identity. As discussed in Remark 2.7.5, there exists $0 \neq \xi \in \mathbb{C}$ such that $\mathbf{e}'_\alpha = \xi \mathbf{e}_\alpha^+$ for all $\alpha \in \Phi$. But then $(-1)^{\mathrm{ht}(\alpha)} h_\alpha = [\mathbf{e}'_\alpha, \mathbf{e}'_{-\alpha}] = \xi^2 [\mathbf{e}_\alpha^+, \mathbf{e}_{-\alpha}^+] = \xi^2 (-1)^{\mathrm{ht}(\alpha)} h_\alpha$ and so $\xi = \pm 1$, as desired. Finally, the relations $\mathbf{e}_{\alpha_i}^+ = c_i e_i$ and $\mathbf{e}_{-\alpha_i}^+ = -c_i f_i$ (with $c_i = \pm 1$ for $i \in I$) hold for the collection constructed as above; hence, they hold for any collection satisyfing (L1)–(L3) and the above identity. □

We now establish an important consequence of Theorem 2.7.2. Let also $\tilde{L}$ be a Lie algebra of Cartan–Killing type, that is, there is a subalgebra $\tilde{H} \subseteq \tilde{L}$ and a subset $\tilde{\Delta} = \{\tilde{\alpha}_i \mid i \in \tilde{I}\}$ (for some finite index set $\tilde{I}$) such that the conditions in Definition 2.2.1 hold. Let $\tilde{A} = (\tilde{a}_{ij})_{i,j \in \tilde{I}}$ be the corresponding structure matrix.

Theorem 2.7.12 (Isomorphism Theorem). *With the above notation, assume that $I = \tilde{I}$ and $A = \tilde{A}$. Then there is a unique isomorphism of Lie algebras $\varphi \colon L \to \tilde{L}$ such that $\varphi(e_i) = \tilde{e}_i$ and $\varphi(f_i) = \tilde{f}_i$ for all $i \in I$, where $\{e_i, f_i \mid i \in I\}$ and $\{\tilde{e}_i, \tilde{f}_i \mid i \in I\}$ are Chevalley generators for L and $\tilde{L}$, respectively (as in Remark 2.2.9).*

Proof. The uniqueness of φ is clear since $L = \langle e_i, f_i \mid i \in I\rangle_{\mathrm{alg}}$; see Proposition 2.4.5. The problem is to prove the existence of φ. Let $\Phi \subseteq H^*$ be the set of roots of L and $\tilde{\Phi} \subseteq \tilde{H}^*$ be the set of roots of $\tilde{L}$. Since $A = \tilde{A}$, the discussion in Remark 2.3.7 shows that we have a canonical bijection $\Phi \xrightarrow{\sim} \tilde{\Phi}$, $\alpha \mapsto \tilde{\alpha}$, given as follows. If $\alpha = \sum_{i \in I} n_i \alpha_i \in \Phi$ (with $n_i \in \mathbb{Z}$), then $\tilde{\alpha} = \sum_{i \in I} n_i \tilde{\alpha}_i \in \tilde{\Phi}$. Then this bijection has the following property: for any $\alpha, \beta \in \Phi$, we have

$$(\heartsuit) \qquad \alpha + \beta \in \Phi \quad \Leftrightarrow \quad \tilde{\alpha} + \tilde{\beta} \in \tilde{\Phi}.$$

Now fix a total order $\sqsubseteq$ on I and let $i_0 \in I$ be the smallest index, as in Remark 2.7.7. Following the above inductive procedures, both in L and in $\tilde{L}$, first yields collections $\{\mathbf{e}_\alpha \mid \alpha \in \Phi\} \subseteq L$ and $\{\tilde{\mathbf{e}}_{\tilde{\alpha}} \mid \tilde{\alpha} \in \tilde{\Phi}\} \subseteq \tilde{L}$. Consequently, we obtain bases

$$B = \{h_i \mid i \in I\} \cup \{\mathbf{e}_\alpha^+ \mid \alpha \in \Phi\} \qquad (h_i := [e_i, f_i]),$$
$$\tilde{B} = \{\tilde{h}_i \mid i \in I\} \cup \{\tilde{\mathbf{e}}_{\tilde{\alpha}}^+ \mid \tilde{\alpha} \in \tilde{\Phi}\} \qquad (\tilde{h}_i := [\tilde{e}_i, \tilde{f}_i])$$

for L and $\tilde{L}$, respectively, such that the relations (L1)–(L3) in Theorem 2.7.2 hold. We assume that both collections are normalised as in Remark 2.7.7, that is, $\mathbf{e}^+_{\alpha_{i_0}} = \mathbf{e}_{\alpha_{i_0}} = e_{i_0}$ and $\tilde{\mathbf{e}}^+_{\tilde{\alpha}_{i_0}} = \tilde{\mathbf{e}}_{\tilde{\alpha}_{i_0}} = \tilde{e}_{i_0}$. Now define a (bijective) linear map $\varphi\colon L \to \tilde{L}$ by

$$\varphi(h_i) := \tilde{h}_i \quad (i \in I) \qquad \text{and} \qquad \varphi(\mathbf{e}^+_\alpha) := \tilde{\mathbf{e}}^+_{\tilde{\alpha}} \quad (\alpha \in \Phi).$$

We have $\mathbf{e}^+_{\alpha_i} = c_i e_i$ and $\mathbf{e}^+_{-\alpha_i} = -c_i f_i$ for all $i \in I$, where $c_i \in \{\pm 1\}$; similarly, $\tilde{\mathbf{e}}^+_{\tilde{\alpha}_i} = \tilde{c}_i e_i$ and $\tilde{\mathbf{e}}^+_{-\tilde{\alpha}_i} = -\tilde{c}_i f_i$ for all $i \in I$, where $\tilde{c}_i \in \{\pm 1\}$. Since $c_{i_0} = \tilde{c}_{i_0} = 1$, we conclude using Remark 2.7.3($*$) that $c_i = \tilde{c}_i$ for all $i \in I$. Consequently, we have

$$\varphi(e_i) = \tilde{e}_i \qquad \text{and} \qquad \varphi(f_i) = \tilde{f}_i \qquad \text{for all } i \in I.$$

Furthermore, let $i \in I$ and $\alpha \in \Phi$ be such that $\alpha + \alpha_i \in \Phi$. By ($\heartsuit$), we also have $\tilde{\alpha} + \tilde{\alpha}_i \in \tilde{\Phi}$ and

$$\begin{aligned} q_{i,\alpha} &= \max\{m \geqslant 0 \mid \alpha - m\alpha_i \in \Phi\} \\ &= \max\{m \geqslant 0 \mid \tilde{\alpha} - m\tilde{\alpha}_i \in \tilde{\Phi}\} = q_{i,\tilde{\alpha}}. \end{aligned}$$

Similarly, if $\alpha - \alpha_i \in \Phi$, then $\tilde{\alpha} - \tilde{\alpha}_i \in \tilde{\Phi}$ and $p_{i,\alpha} = p_{i,\tilde{\alpha}}$. Hence, (L2) shows that the matrix of $\mathrm{ad}_L(e_i)\colon L \to L$ with respect to the basis B is equal to the matrix of $\mathrm{ad}_{\tilde{L}}(\tilde{e}_i)\colon \tilde{L} \to \tilde{L}$ with respect to the basis $\tilde{B}$; by (L3), similar statements also hold for $\mathrm{ad}_L(f_i)$ and $\mathrm{ad}_{\tilde{L}}(\tilde{f}_i)$. Since φ is linear, this implies that

$$\begin{aligned} \varphi([e_i, y]) &= [\tilde{e}_i, \varphi(y)] = [\varphi(e_i), \varphi(y)], \\ \varphi([f_i, y]) &= [\tilde{f}_i, \varphi(y)] = [\varphi(f_i), \varphi(y)] \end{aligned}$$

for all $i \in I$, $y \in L$. Since $L = \langle e_i, f_i \mid i \in I\rangle_{\mathrm{alg}}$, it follows that $\varphi([x,y]) = [\varphi(x), \varphi(y)]$ for all $x, y \in L$ (see Exercise 1.1.8). $\square$

Proposition 2.7.13 (Cf. Chevalley [**9**, §I]). *Let $\{\mathbf{e}^+_\alpha \mid \alpha \in \Phi\}$ be a collection as in Corollary 2.7.11. Then the following hold.*

(a) *We have $\omega(\mathbf{e}^+_\alpha) = -\mathbf{e}^+_{-\alpha}$ for all $\alpha \in \Phi$. (Here, $\omega\colon L \to L$ is the Chevalley involution; see Exercise Sheet 8.)*

(b) *Let $\alpha, \beta \in \Phi$ be such that $\alpha + \beta \in \Phi$. Then $[\mathbf{e}^+_\alpha, \mathbf{e}^+_\beta] = \pm(q+1)\mathbf{e}^+_{\alpha+\beta}$, where $q \geqslant 0$ is defined as in Lemma 2.6.2.*

Proof. (a) Let $\alpha \in \Phi^+$. We show the assertion by induction on $\mathrm{ht}(\alpha)$. If $\mathrm{ht}(\alpha) = 1$, then $\alpha = \alpha_i$ for some $i \in I$. We have $\mathbf{e}^+_{\alpha_i} = c_i e_i$

and $\mathbf{e}^+_{-\alpha_i} = -c_i f_i$, where $c_i \in \{\pm 1\}$ for all $i \in I$. Hence, we obtain $\omega(\mathbf{e}^+_{\alpha_i}) = c_i\omega(e_i) = c_i f_i = -\mathbf{e}^+_{-\alpha_i}$, as required. Now let $\mathrm{ht}(\alpha) > 1$. By the Key Lemma 2.3.4, there exists some $i \in I$ such that $\beta := \alpha - \alpha_i \in \Phi^+$. We have $\mathrm{ht}(\beta) = \mathrm{ht}(\alpha) - 1$ and so $\omega(\mathbf{e}^+_\beta) = -\mathbf{e}^+_{-\beta}$, by induction. By condition (L1) in Theorem 2.7.2, we have $[e_i, \mathbf{e}^+_\beta] = (q_{i,\beta}+1)\mathbf{e}^+_\alpha$. Applying ω yields that

$$(q_{i,\beta}+1)\omega(\mathbf{e}^+_\alpha) = \omega\big([e_i, \mathbf{e}^+_\beta]\big) = [\omega(e_i), \omega(\mathbf{e}^+_\beta)] = -[f_i, \mathbf{e}^+_{-\beta}].$$

Now, we have $-\beta - \alpha_i = -\alpha \in \Phi$ and so condition (L2) in Theorem 2.7.2 yields that $[f_i, \mathbf{e}^+_{-\beta}] = (p_{i,-\beta}+1)\mathbf{e}^+_{-\alpha}$. Hence, we deduce that $\omega(\mathbf{e}^+_\alpha) = -\mathbf{e}^+_{-\alpha}$, since $p_{i,-\beta} = q_{i,\beta}$ as pointed out in Remark 2.7.1. Thus, the assertion holds for all $\alpha \in \Phi^+$. But, since $\omega^2 = \mathrm{id}_L$, we then also have $\omega(\mathbf{e}^+_{-\alpha}) = \omega\big(-\omega(\mathbf{e}^+_\alpha)\big) = -\omega^2(\mathbf{e}^+_\alpha) = -\mathbf{e}^+_\alpha$, as required.

(b) We would like to use Proposition 2.6.8, but we can not do that directly because the condition in Remark 2.6.1(a) does not hold for the collection $\{\mathbf{e}^+_\alpha \mid \alpha \in \Phi\}$. So we revert the construction of $\mathbf{e}^+_\alpha$ and define a collection $\{0 \neq e_\alpha \in L_\alpha \mid \alpha \in \Phi\}$ by

$$e_\alpha := \begin{cases} \mathbf{e}^+_\alpha & \text{if } \alpha \in \Phi^+, \\ (-1)^{\mathrm{ht}(\alpha)}\mathbf{e}^+_\alpha & \text{if } \alpha \in \Phi^-. \end{cases}$$

Then $[e_\alpha, e_{-\alpha}] = h_\alpha$ for all $\alpha \in \Phi$. By (a), we also have the formula:

$$\omega(e_\alpha) = -(-1)^{\mathrm{ht}(\alpha)}e_{-\alpha} \qquad \text{for all } \alpha \in \Phi.$$

Let $N_{\alpha,\beta}$ be the structure constants with respect to $\{e_\alpha \mid \alpha \in \Phi\}$, as in Section 2.6. Writing $[e_\alpha, e_\beta] = N_{\alpha,\beta}e_{\alpha+\beta}$, we certainly have $[\mathbf{e}^+_\alpha, \mathbf{e}^+_\beta] = \pm N_{\alpha,\beta}\mathbf{e}^+_{\alpha+\beta}$. So it suffices to show that $N_{\alpha,\beta} = \pm(q+1)$. This is seen as follows. Using the above formula for ω, we obtain $\omega\big([e_\alpha, e_\beta]\big) = N_{\alpha,\beta}\omega(e_{\alpha+\beta}) = -(-1)^{\mathrm{ht}(\alpha+\beta)}N_{\alpha,\beta}e_{-(\alpha+\beta)}$. On the other hand, we can also evaluate the left hand side as follows.

$$\begin{aligned}\omega\big([e_\alpha, e_\beta]\big) &= [\omega(e_\alpha), \omega(e_\beta)] = (-1)^{\mathrm{ht}(\alpha)+\mathrm{ht}(\beta)}[e_{-\alpha}, e_{-\beta}] \\ &= (-1)^{\mathrm{ht}(\alpha)+\mathrm{ht}(\beta)}N_{-\alpha,-\beta}e_{-(\alpha-\beta)}.\end{aligned}$$

Hence, we conclude that $N_{-\alpha,-\beta} = -N_{\alpha,\beta}$ and so Proposition 2.6.8 implies that $N^2_{\alpha,\beta} = (q+1)^2$. Thus, $N_{\alpha,\beta} = \pm(q+1)$, as claimed. $\square$

Chapter 3

Generalised Cartan matrices

In the previous chapter we have seen that a Lie algebra L of Cartan–Killing type is determined (up to isomorphism) by its structure matrix $A = (a_{ij})_{i,j\in I}$. The entries of A are integers, we have $a_{ii} = 2$ and $a_{ij} \leqslant 0$ for $i \neq j$; furthermore, $a_{ij} < 0 \Leftrightarrow a_{ji} < 0$. In Section 3.1 we show that every (indecomposable) matrix satisfying those conditions has one of three possible types: (FIN), (AFF) or (IND). There is a complete classification of all such matrices of types (FIN) and (AFF). The structure matrix A of L does turn out to be of type (FIN) and, hence, it is encoded by one of the graphs in the famous list of Dynkin diagrams of type A_n, B_n, C_n, D_n, G_2, F_4, E_6, E_7 and E_8.

Once the results in Section 3.1 are established, the central theme of this chapter is as follows. We start with an arbitrary matrix A as above, of type (FIN). Then we can construct the following objects:

1) An abstract root system Φ. In Section 2.3 we already made first steps in that direction, and presented a Python program to determine Φ from A. This will be further developed in Section 3.2.

2) A Lie algebra L of Cartan–Killing type with structure matrix A and root system Φ. This will be done by a process that reverses the construction of Lusztig's canonical basis; see Section 3.3.

3) A Chevalley group G "of type L", first over $\mathbb{C}$ and then over any field K. Here we follow Lusztig's simplified construction using the canonical basis of L; see Section 3.5.

We shall emphasise the fact that the constructions are by means of purely combinatorial procedures, which do not involve any other ingredients (or choices) and, hence, can also be implemented on a computer: the single input datum for the computer programs is the matrix A (plus the field K for the Chevalley groups). We present a specific computer algebra package with these features in Section 3.4.

3.1. Classification

Let I be a finite, non-empty index set. We consider matrices $A = (a_{ij})_{i,j\in I}$ with entries in $\mathbb{R}$ satisfying the following two conditions:

(C1) $a_{ij} \leqslant 0$ for all $i \neq j$ in I;

(C2) $a_{ij} \neq 0 \Leftrightarrow a_{ji} \neq 0$, for all $i, j \in I$.

Examples of such matrices are the structure matrices of Lie algebras of Cartan–Killing type; see Corollary 2.2.12. One of our aims will be to find the complete list of all possible such structure matrices. For this purpose, it will be convenient to first work in a more general setting, where we only assume that (C1) and (C2) hold.

In analogy to Definition 2.4.7, we say that A is *indecomposable* if there is no partition $I = I_1 \sqcup I_2$ (where $I_1, I_2 \subsetneqq I$ and $I_1 \cap I_2 = \varnothing$) such that $a_{ij} = a_{ji} = 0$ for all $i \in I_1$ and $j \in I_2$.

Some further notation. Let $u = (u_i)_{i\in I} \in \mathbb{R}^I$. We write $u \geqslant 0$ if $u_i \geqslant 0$ for all $i \in I$; we write $u > 0$ if $u_i > 0$ for all $i \in I$. Finally, $Au \in \mathbb{R}^I$ is the vector with i-th component given by $\sum_{j\in I} a_{ij}u_j$ (usual product of A with u regarded as a column vector).

Lemma 3.1.1. *Assume that A satisfies* (C1), (C2) *and is indecomposable. If $u \in \mathbb{R}^I$ is such that $u \geqslant 0$, $Au \geqslant 0$, then $u = 0$ or $u > 0$.*

Proof. Let $I_1 := \{i \in I \mid u_i = 0\}$ and $I_2 := \{i \in I \mid u_i > 0\}$. Then $I = I_1 \cup I_2$, $I_1 \cap I_2 = \varnothing$. Let $i \in I_1$ and v_i be the i-th component of Au; by assumption, $v_i \geqslant 0$. On the other hand, $v_i = \sum_{j\in I} a_{ij}u_j = \sum_{j\in I_2} a_{ij}u_j$ where all terms in the sum on the right hand side are

$\leqslant 0$ since A satisfies (C1) and $u_j > 0$ for all $j \in I_2$; furthermore, if $a_{ij} < 0$ for some $j \in I_2$, then $v_i < 0$, contradiction to $v_i \geqslant 0$. So we must have $a_{ij} = 0$ for all $i \in I_1, j \in I_2$. Since A satisfies (C2), we also have $a_{ji} = 0$ for all $i \in I_1, j \in I_2$. Since A is indecomposable, either $I_1 = I$ (and so $u = 0$) or $I_2 = I$ (and so $u > 0$). □

Theorem 3.1.2 (Vinberg)**.** *Assume that A satisfies* (C1), (C2) *and is indecomposable. Let $\mathscr{K}_A := \{u \in \mathbb{R}^I \mid Au \geqslant 0\}$. Then exactly one of the following three conditions holds.*

(FIN) $\{0\} \neq \mathscr{K}_A \subseteq \{u \in \mathbb{R}^I \mid u > 0\} \cup \{0\}$.

(AFF) $\mathscr{K}_A = \{u \in \mathbb{R}^I \mid Au = 0\} = \langle u_0 \rangle_{\mathbb{R}}$ *where* $u_0 > 0$.

(IND) $\mathscr{K}_A \cap \{u \in \mathbb{R}^I \mid u \geqslant 0\} = \{0\}$.

Accordingly, we say that A is of finite, affine or indefinite type.

Proof. First we show that the three conditions are disjoint. If (FIN) or (AFF) holds, then there exists some $u \in \mathbb{R}^I$ such that $u > 0$ and $Au \geqslant 0$. Hence, (IND) does not hold. If (AFF) holds, then there exists some $u \in \mathbb{R}^I$ such that $u > 0$ and $Au = 0 \geqslant 0$. But then also $A(-u) \geqslant 0$ and so (FIN) does not hold. Hence, the conditions are indeed disjoint. It remains to show that we are always in one of the three cases. Assume that (IND) does not hold. Then there exists some $0 \neq v \in \mathscr{K}_A$ such that $v \geqslant 0$. By Lemma 3.1.1, we have $v > 0$. We want to show that (FIN) or (AFF) holds. Assume that (FIN) does not hold. Since $\mathscr{K}_A \neq \{0\}$, this means that there exists $0 \neq u \in \mathscr{K}_A$ such that $u_h \leqslant 0$ for some $h \in I$. We have $v > 0$ and so we can consider the ratios u_i/v_i for $i \in I$. Let $j \in I$ be such that $u_j/v_j \leqslant u_i/v_i$ for all $i \in I$. If $u_j \geqslant 0$, then $u_i \geqslant 0$ for all $i \in I$ and so $u \geqslant 0$. But then Lemma 3.1.1 would imply that $u > 0$, contradiction to our choice of u. Hence, $u_j < 0$ and so $s := -u_j/v_j > 0$. Now let us look at the vector $u + sv$; its i-th component is

$$(u+sv)_i = u_i + sv_i = v_i(u_i/v_i - u_j/v_j) \begin{cases} = 0 & \text{if } i = j, \\ \geqslant 0 & \text{if } i \neq j. \end{cases}$$

Hence, we have $u + sv \geqslant 0$ and $A(u + sv) = Au + sAv \geqslant 0$. By Lemma 3.1.1, either $u + sv = 0$ or $u + sv > 0$. But $(u + sv)_j = 0$ and so we must have $u + sv = 0$, that is, $u = -sv$. But then $0 \leqslant Au = (-s)Av \leqslant 0$ (since $s > 0$ and $Av \geqslant 0$) and so $Av = Au = 0$.

Finally, consider any $0 \neq w \in \mathscr{K}_A$. Again, let $j \in J$ be such that $w_j/v_j \leqslant w_i/v_i$ for all $i \in I$, and set $t := -w_j/v_j$. As above, we see that $w + tv \geqslant 0$ and $(w + tv)_j = 0$. Furthermore, $A(w + tv) = Aw + tAv = Aw \geqslant 0$ (since $Av = 0$). So Lemma 3.1.1 implies that either $w + tv > 0$ (which is not the case) or $w + tv = 0$; hence, $w = -tv \in \langle v \rangle_{\mathbb{R}}$. So $\mathscr{K}_A \subseteq \langle v \rangle_{\mathbb{R}} \subseteq \{z \in \mathbb{R}^I \mid Az = 0\}$ and the right hand side is contained in $\mathscr{K}_A$. Hence, (AFF) holds where $u_0 = v$. $\square$

Corollary 3.1.3. *Let A be as in Theorem 3.1.2.*

(a) *A is of finite type if and only if there exists $u \in \mathbb{R}^I$ such that $u \geqslant 0$, $Au \geqslant 0$ and $Au \neq 0$. In this case,* $\det(A) \neq 0$.

(b) *A is of affine type if and only if there exists $0 \neq u \in \mathbb{R}^I$ such that $u \geqslant 0$ and $Au = 0$. In this case, A has rank $|I| - 1$.*

Proof. (a) If (FIN) holds, then Theorem 3.1.2 shows that there is some $u \in \mathbb{R}^I$ such that $u > 0$ and $Au \geqslant 0$. If we had $Au = 0$, then also $A(-u) = 0$, contradiction to $\mathscr{K}_A \subseteq \{u \in \mathbb{R}^I \mid u > 0\} \cup \{0\}$. Conversely, assume that there exists $u \in \mathbb{R}^I$ such that $u \geqslant 0$, $Au \geqslant 0$ and $Au \neq 0$; in particular, $u \neq 0$ and so (IND) does not hold. Furthermore, $Au \neq 0$ and so (AFF) does not hold. Hence, the only remaining possibility is that (FIN) holds.

Assume now that (FIN) holds. Let $v \in \mathbb{R}^I$ be such that $Av = 0$. But then $v, -v \in \mathscr{K}_A$ and so we must have $v = 0$. Hence, we have $\{v \in \mathbb{R}^I \mid Av = 0\} = \{0\}$ and so $\det(A) \neq 0$.

(b) If (AFF) holds, then Theorem 3.1.2 shows that there is some $u \in \mathbb{R}^I$ such that $u > 0$ and $Au = 0$, as required. Conversely, assume that there exists $0 \neq u \in \mathbb{R}^I$ such that $u \geqslant 0$ and $Au = 0$; in particular, $u \in \mathscr{K}_A$ and $\det(A) = 0$. But then neither (FIN) nor (IND) holds, so (AFF) must hold. The statement about the rank of A is clear by condition (AFF). $\square$

Remark 3.1.4. Let $A = (a_{ij})_{i,j \in I}$ be the structure matrix of a Lie algebra L of Cartan–Killing type, as in Chapter 2. Assume that $L \neq \{0\}$ is simple; then A is indecomposable by Theorem 2.4.14. As already remarked above, A satisfies (C1) and (C2). So we can now ask whether A is of finite, affine or indefinite type. We claim that A is of finite type. To see this, let $\alpha \in \Phi^+$ be such that $\mathrm{ht}(\alpha)$ is as large as possible. Write $\alpha = \sum_{j\in I} n_j\alpha_j$ where $n_j \in \mathbb{Z}_{\geqslant 0}$. Let $i \in I$. Using the formula in Remark 2.3.7, we obtain

$$\alpha - \Big(\sum_{j\in I} a_{ij}n_j\Big)\alpha_i = \sum_{j\in I} n_j\big(\alpha_j - a_{ij}\alpha_i\big) = s_i(\alpha) \in \Phi.$$

Now $\mathrm{ht}(s_i(\alpha)) \leqslant \mathrm{ht}(\alpha)$ and so $\sum_{j\in I} a_{ij}n_j \geqslant 0$ for all $i \in I$. Hence, we have $Au \geqslant 0$ where $u := (n_i)_{i\in I} \geqslant 0$. Furthermore, $\det(A) \neq 0$ and so $Au \neq 0$. So A is of finite type by Corollary 3.1.3(a).

Definition 3.1.5 (Kac [**21**, §1.1])**.** Assume that $A = (a_{ij})_{i,j\in I}$ satisfies (C1), (C2). We say that A is a *generalised Cartan matrix* if $a_{ij} \in \mathbb{Z}$ and $a_{ii} = 2$ for all $i, j \in I$.

Our aim is to classify the generalised Cartan matrices of finite and affine type. We begin with some preparations.

Lemma 3.1.6. *Assume that A satisfies* (C1), (C2) *and is indecomposable. Let $A_J := (a_{ij})_{i,j\in J}$ where $\varnothing \neq J \subsetneqq I$. Then, clearly, A_J also satisfies* (C1), (C2). *If A is of finite or affine type and if A_J is indecomposable, then A_J is of finite type.*

Proof. Since A is of finite or affine type, there exists $u \in \mathbb{R}^I$ such that $u > 0$ and $Au \geqslant 0$. Define $u' := (u_i)_{i\in J} \in \mathbb{R}^J$. For $i \in J$ we have

$$0 \leqslant (Au)_i = \sum_{j\in I} a_{ij}u_j = \sum_{j\in J} a_{ij}u_j + \sum_{j\in I\setminus J} \underbrace{a_{ij}u_j}_{\leqslant 0} \leqslant (A_Ju')_i.$$

Hence, $u' > 0$ and $u' \in \mathscr{K}_{A_J}$ which means that A_J is of finite or affine type (see Theorem 3.1.2). By Corollary 3.1.3, it remains to show that $A_Ju' \neq 0$. Assume, if possible, that $(A_Ju')_i = 0$ for all $i \in J$. Then the above inequality shows that $a_{ij}u_j = 0$ and, hence, $a_{ij} = 0$ for all $j \in I \setminus J$. But then A is decomposable, contradiction. $\square$

Lemma 3.1.7. *Let $A := (a_{ij})_{i,j\in I}$ be an indecomposable generalised Cartan matrix of finite or affine type. Then $0 \leqslant a_{ij}a_{ji} \leqslant 4$ for all $i, j \in I$. If $|I| \geqslant 3$, then $0 \leqslant a_{ij}a_{ji} \leqslant 3$ for all $i \neq j$ in I.*

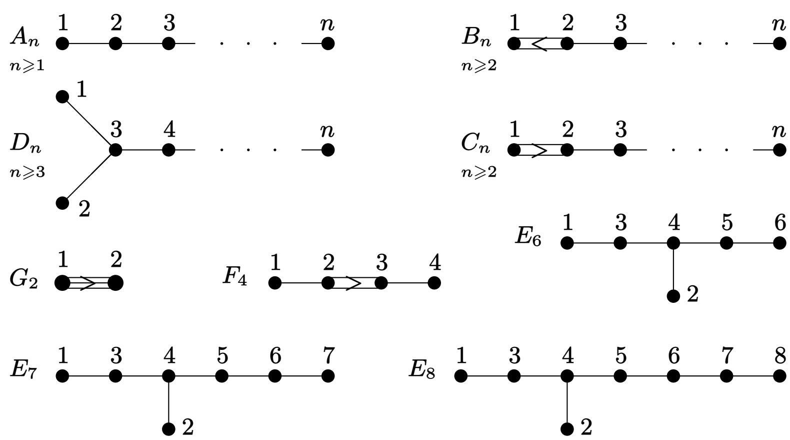
Proof. If $i = j$, then $a_{ii} = 2$ and so the assertion is clear. Now let $|I| \geqslant 2$ and $J = \{i, j\}$, where $i \neq j$ in I are such that $a_{ij} \neq 0$. Then $A_J = \begin{pmatrix} 2 & -a \\ -b & 2 \end{pmatrix}$ where $a = -a_{ij}$, $b = -a_{ji}$, $a, b > 0$. If $|I| = 2$, then $A = A_J$ is of finite or affine type; otherwise, A_J is of finite type by Lemma 3.1.6. So there exists some $u \in \mathbb{R}^J$ such that $u > 0$ and $A_J u \geqslant 0$; we can assume that u has components 1 and $c > 1$. Now

$$0 \leqslant A_J u = \begin{pmatrix} 2 & -a \\ -b & 2 \end{pmatrix} \begin{pmatrix} 1 \\ c \end{pmatrix} = \begin{pmatrix} 2 - ac \\ -b + 2c \end{pmatrix},$$

and so $b/2 \leqslant c \leqslant 2/a$. Hence, we have $ab \leqslant 4$, as desired. Finally, if $|I| \geqslant 3$, then A_J is of finite type (as already noted) and so $\det(A_J) \neq 0$ by Corollary 3.1.3(a). This implies that $ab \neq 4$, as claimed. □

ab hier Woche 9

Table 4. Dynkin diagrams of finite type

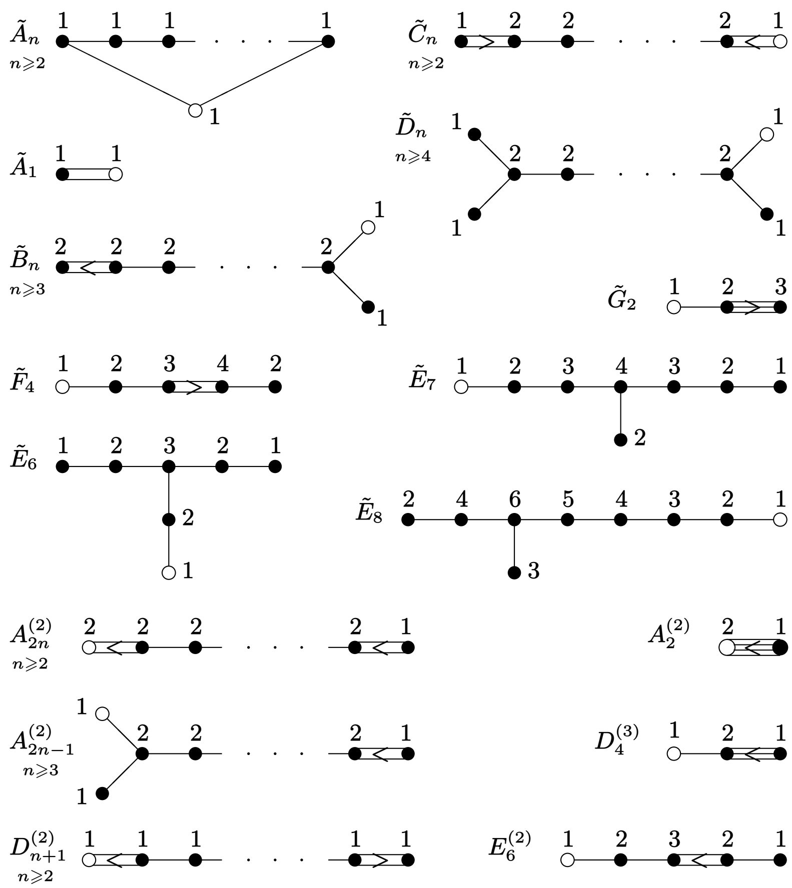


(The numbers attached to the vertices define a standard labelling of the graph.)

Definition 3.1.8. Let $A = (a_{ij})_{i,j \in I}$ be an indecomposable generalised Cartan matrix of finite or affine type. Then we encode A in a diagram, called *Dynkin diagram* and denoted by $\Gamma(A)$, as follows.

The vertices of $\Gamma(A)$ are in bijection to I. Now let $i, j \in I$, $i \neq j$. If $a_{ij} = a_{ji} = 0$, then there is no edge between the vertices labelled by i and j. Now assume that $a_{ij} \neq 0$. By Lemma 3.1.7, we have $1 \leqslant a_{ij}a_{ji} \leqslant 4$. If $a_{ij} = a_{ji} = -2$, then the vertices labelled by i, j will be joined by a double edge. Otherwise, $1 \leqslant a_{ij}a_{ji} \leqslant 4$ and we can choose the notation such that $a_{ij} = -1$; let $m := -a_{ji} \in \{1, 2, 3, 4\}$. Then the vertices labelled by i, j will be joined by m edges; if $m \geqslant 2$, then we put an additional arrow pointing towards j.

Table 5. Dynkin diagrams of affine type



(Each diagram denoted $\tilde{X}_n$ has $n+1$ vertices; $A_{2n}^{(2)}$, $A_{2n-1}^{(2)}$, $D_{n+1}^{(2)}$ have $n+1$ vertices; the numbers attached to the vertices define a vector $u = (u_i)_{i\in I}$ such that $Au = 0$.)

Note that A and $\Gamma(A)$ determine each other completely; the fact that A is indecomposable means that $\Gamma(A)$ is connected. Examples:

If $A = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$, then $\Gamma(A)$ is the graph $\tilde{A}_1$ in Table 5.

If $A = \begin{pmatrix} 2 & -4 \\ -1 & 2 \end{pmatrix}$, then $\Gamma(A)$ is the graph $\tilde{A}_2^{(2)}$ in Table 5.

If A corresponds to the Lie algebra $\mathfrak{sl}_n(\mathbb{C})$ ($n \geqslant 2$), then $\Gamma(A)$ is the graph A_{n-1} in Table 4; see Example 2.2.7. If A corresponds to a classical Lie algebra $\mathfrak{go}_n(Q_n, \mathbb{C})$, then Table 2 (p. 76) shows that

$$\Gamma(A) \text{ is the graph } \begin{cases} D_m & \text{if } Q_n^{\mathrm{tr}} = Q_n \text{ and } n = 2m \geqslant 6, \\ B_m & \text{if } Q_n^{\mathrm{tr}} = Q_n \text{ and } n = 2m+1 \geqslant 5, \\ C_m & \text{if } Q_n^{\mathrm{tr}} = -Q_n \text{ and } n = 2m \geqslant 4. \end{cases}$$

(In accordance with Exercise 1.6.4, we may identify $B_1 = C_1 = A_1$.)

Lemma 3.1.9. *The graphs in Table 4 correspond to indecomposable generalised Cartan matrices of finite type; those in Table 5 to indecomposable generalised Cartan matrices of affine type.*

Proof. Let Γ be one of the diagrams in Table 5. Let $|I| = n+1$ and write $I = \{0, 1, \ldots, n\}$ where $1, \ldots, n$ correspond to the vertices "$\bullet$" and 0 corresponds to the vertex "$\circ$". Using the conditions in Definition 3.1.8, we obtain an indecomposable generalised Cartan matrix A such that $\Gamma = \Gamma(A)$. Let $u = (u_i)_{i \in I}$ be the vector defined by the numbers attached to the vertices in Table 5. One checks in each case that $u > 0$, $Au = 0$ and so A is of affine type by Corollary 3.1.3(b). For example, the graph $D_4^{(3)}$ leads to:

$$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -3 \\ 0 & -1 & 2 \end{pmatrix}, \qquad u = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \qquad Au = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Finally, all graphs in Table 4 are obtained as $\Gamma(A_J)$ where $J = I \setminus \{0\}$. Now Lemma 3.1.6 shows, without any further calculations, that A_J is of finite type. $\square$

Lemma 3.1.10. *Let $A = (a_{ij})_{i,j \in I}$ and $A' = (a'_{ij})_{i,j \in I}$ be indecomposable generalised Cartan matrices such that $A \neq A'$ and $a_{ij} \leqslant a'_{ij}$ for all $i, j \in I$. If A is of finite or affine type, then A' is of finite type.*

Proof. Let A be of finite or affine type. There exists some $u \in \mathbb{R}^I$ such that $u > 0$ and $Au \geqslant 0$. Let $i \in I$. Then

$$\begin{aligned} (A'u)_i &= \sum_{j \in I} a'_{ij} u_j = 2u_i + \sum_{j \in I, j \neq i} a'_{ij} u_j \\ &\geqslant 2u_i + \sum_{j \in I, j \neq i} a_{ij} u_j = \sum_{j \in I} a_{ij} u_j = (Au)_i \geqslant 0. \end{aligned}$$

So $A'u \geqslant 0$ and A' is of finite or affine type, by Corollary 3.1.3. Since $A \neq A'$, there exist $i, j \in I$ such that $a_{ij} < a'_{ij}$. Then $i \neq j$ and so the above computation shows that $(A'u)_i > (Au)_i \geqslant 0$. Hence, $A'u \neq 0$ and so A' is of finite type (again, by Corollary 3.1.3). □

Lemma 3.1.11. *Let $A = (a_{ij})_{i,j \in I}$ be an indecomposable generalised Cartan matrix of finite or affine type. Assume that there is a cycle in $\Gamma(A)$, that is, there exist indices $i_1, i_2, \ldots, i_r$ in I ($r \geqslant 3$) such that*

$$(\circlearrowleft) \quad a_{i_1 i_2} a_{i_2 i_3} \cdots a_{i_{r-1} i_r} a_{i_r i_1} \neq 0 \quad \textit{and} \quad i_1, i_2, \ldots, i_r \textit{ are distinct.}$$

Then A is of affine type, $|I| = r$ and $\Gamma(A) = \tilde{A}_{r-1}$ in Table 5.

Proof. Let $J := \{i_1, \ldots, i_r\}$. By ($\circlearrowleft$) and Remark 2.4.8, the submatrix A_J is indecomposable. Define $A'_J = (a'_{ij})_{i,j \in J}$ by

$$a'_{i_1 i_2} = a'_{i_2 i_3} = \ldots = a'_{i_{r-1} i_r} = a'_{i_r i_1} = -1, \qquad a'_{jj} = 2$$

and $a'_{jj'} = 0$ for all other $j \neq j'$ in J. Then $\Gamma(A'_J)$ is the graph $\tilde{A}_{r-1}$ and so A'_J is of affine type (see Lemma 3.1.9). Furthermore, by ($\circlearrowleft$), we have $a_{ij} \leqslant a'_{ij}$ for all $i, j \in J$. So, if A is of finite type, or of affine type with $|I| > r$, then A_J is of finite type (by Lemma 3.1.6) and, hence, also A'_J (by Lemma 3.1.10), contradiction. So $|I| = r$ and $A = A_J$. If $A_J \neq A'_J$, then Lemma 3.1.10 implies that A'_J is of finite type, contradiction. □

Theorem 3.1.12. *The Dynkin diagrams of indecomposable generalised Cartan matrices of finite type are precisely those in Table 4.*

Proof. By Lemma 3.1.9, we already know that all diagrams in Table 4 satisfy this condition. Conversely, let $A = (a_{ij})_{i,j \in I}$ be an arbitrary indecomposable generalised Cartan matrix of finite type. We must show that the corresponding diagram $\Gamma(A)$ appears in Table 4. If $|I| = 1$, then $A = (2)$ and $\Gamma(A) = A_1$. Now let $|I| \geqslant 2$. By Lemma 3.1.7, there are only single, double or triple edges in $\Gamma(A)$ (and an arrow is attached to a double or triple edge). Hence, if $|I| = 2$, then $\Gamma(A)$ is one of the graphs A_2, B_2, C_2 or G_2.

Now assume that $|I| \geqslant 3$. By using Lemmas 3.1.6 and 3.1.10, one obtains further restrictions on $\Gamma(A)$ which eventually lead to the list of graphs in Table 4. We give full details for one example.

Claim: $\Gamma(A)$ does not have a triple edge. This is seen as follows. Assume, if possible, that there are $i \neq j$ in I which are connected by a triple edge. Since $|I| \geqslant 3$ and A is indecomposable, there is a further $k \in I$ connected to i or j; we choose the notation such that k is connected to i. By Lemma 3.1.11, there are no cyles in $\Gamma(A)$ and so there is no edge between j, k. Let $J := \{k, i, j\}$ and consider the submatrix A_J. We have

$$A_J = \begin{pmatrix} 2 & -a & 0 \\ -b & 2 & -c \\ 0 & -d & 2 \end{pmatrix} \qquad \text{where } a, b, c, d > 0 \text{ and } cd = 3.$$

Then A_J must also be of finite type; see Lemma 3.1.6. Let

$$A'_J = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -c \\ 0 & -d & 2 \end{pmatrix}.$$

Then A'_J is still of finite type by Lemma 3.1.10. But $\Gamma(A'_J)$ is the graph $\tilde{G}_2$ or the graph $D_4^{(3)}$, contradiction to Lemma 3.1.9.

By similar arguments one shows that, if $\Gamma(A)$ has a double edge, then there is only one double edge and no branch point (that is, a vertex connected to at least three other vertices). Hence, $\Gamma(A)$ must be one of the graphs B_n, C_n or F_4. Finally, if $\Gamma(A)$ has only single edges, then one shows that there is at most one branch point, and the remaining possibilities are A_n, D_n, E_6, E_7 and E_8. □

Remark 3.1.13. By similar arguments, one can also show that the Dynkin diagrams of indecomposable generalised Cartan matrices of affine type are precisely those in Table 5; see Kac [**21**, Chap. 4].

Exercise 3.1.14. Let A be an indecomposable generalised Cartan matrix of type (FIN). Then $\det(A) \neq 0$ and we can form A^{-1}. Use condition (FIN) to show that all entries of A^{-1} are strictly positive rational numbers. Work out some examples explicitly.

3.2. Finite root systems

Consider an arbitrary generalised Cartan matrix $A = (a_{ij})_{i,j \in I}$, where I is a non-empty finite index set. Let E be an $\mathbb{R}$-vector space with a basis $\Delta = \{\alpha_i \mid i \in I\}$. For each $i \in I$, we define a linear map

$s_i \colon E \to E$ by the formula

$$s_i(\alpha_j) := \alpha_j - a_{ij}\alpha_i \quad \text{for } j \in I \qquad \text{(cf. Remark 2.3.7).}$$

Since $a_{ii} = 2$, we have $s_i(\alpha_i) = -\alpha_i$. Furthermore, we compute $s_i^2(\alpha_j) = s_i\big(\alpha_j - a_{ij}\alpha_i\big) = s_i(\alpha_j) + a_{ij}\alpha_i = \alpha_j$ for all $j \in I$. Hence, we have $s_i^2 = \mathrm{id}_E$ and so $s_i \in \mathrm{GL}(E)$. The subgroup

$$W = W(A) := \langle s_i \mid i \in I \rangle \subseteq \mathrm{GL}(E)$$

is called the *Weyl group* associated with A. In analogy to Theorem 2.3.6(a), the corresponding *abstract root system* is defined by

$$\Phi = \Phi(A) := \{w(\alpha_i) \mid w \in W, i \in I\};$$

the roots $\{\alpha_i \mid i \in I\}$ are also called *simple roots*. Clearly, if W is finite, then so is Φ. Conversely, assume that Φ is finite. By definition, it is clear that $w(\alpha) \in \Phi$ for all $w \in W$ and $\alpha \in \Phi$. So there is an action of W on Φ. By exactly the same argument as in Remark 2.3.2, it follows that W is finite. Hence, we have:

$$|W(A)| < \infty \qquad \Leftrightarrow \qquad |\Phi(A)| < \infty.$$

In Example 2.3.10, we have computed $W(A)$ and $\Phi(A)$ for the matrix A with Dynkin diagram G_2 in Table 4; in this case, $|W(A)| = 12 < \infty$. In Exercise 2.3.11, there are two examples where $|W(A)| = \infty$. (The first of those matrices has affine type with Dynkin diagram $\tilde{A}_2$ in Table 5; the second matrix is of indefinite type.)

Remark 3.2.1. Assume that A is decomposable. So there is a partition $I = I_1 \sqcup I_2$ such that A has a block diagonal shape

$$A = \left(\begin{array}{c|c} A_1 & 0 \\ \hline 0 & A_2 \end{array}\right)$$

where A_1 has rows and columns labelled by I_1, and A_2 has rows and columns labelled by I_2. Then consider $E = E_1 \oplus E_2$, where

$$E_1 := \langle \alpha_i \mid i \in I_1 \rangle_{\mathbb{R}} \qquad \text{and} \qquad E_2 := \langle \alpha_i \mid i \in I_2 \rangle_{\mathbb{R}}.$$

By the same argument as in Lemma 2.4.9, we have $\Phi = \Phi_1 \sqcup \Phi_2$ where $\Phi_1 := \Phi \cap E_1$ and $\Phi_2 := \Phi \cap E_2$. Furthermore, as in Exercise 2.4.10, one sees that $W = W_1 \cdot W_2$ and $W_1 \cap W_2 = \{1\}$, where

$$W_1 := \langle s_i \mid i \in I_1 \rangle \subseteq W \qquad \text{and} \qquad W_2 := \langle s_i \mid i \in I_2 \rangle \subseteq W.$$

Finally, using (a) and (b) in the proof of Lemma 2.4.9, one shows that $W_1 \cong W(A_1)$ and $W_2 \cong W(A_2)$. Hence, we obtain the equivalence:

$$|W(A)| < \infty \qquad \Leftrightarrow \qquad |W(A_1)| < \infty \quad \text{and} \quad |W(A_2)| < \infty.$$

Thus, in order to characterise those A for which $W(A)$ is finite, we may assume without loss of generality that A is indecomposable.

Remark 3.2.2. Assume that $|W(A)| < \infty$. Then we can construct a $W(A)$-invariant scalar product $\langle\ ,\ \rangle\colon E \times E \to \mathbb{R}$ by the same method as in Section 2.3. (In the sequel, it will not be important how exactly $\langle\ ,\ \rangle$ is defined; it just needs to be symmetric, positive-definite and $W(A)$-invariant.) This yields the formula

$$a_{ij} = 2\frac{\langle \alpha_i, \alpha_j\rangle}{\langle \alpha_i, \alpha_i\rangle} \qquad \text{for all } i, j \in I;$$

see the argument in Remark 2.3.3. Consequently, we have

$$s_i(v) = v - 2\langle \alpha_i^\vee, v\rangle \alpha_i \qquad \text{for all } v \in E.$$

Here, we write $\alpha^\vee := 2\alpha/\langle \alpha, \alpha\rangle \in E$ for any $\alpha \in \Phi(A)$.

Lemma 3.2.3. *Assume that A is indecomposable and $|W(A)| < \infty$. Then A is of type* (FIN).

Proof. Let X be the set of all $\alpha \in \Phi$ such that α can be written as a $\mathbb{Z}$-linear combination of Δ, where all coefficients are $\geqslant 0$. Then X is non-empty; for example, $\Delta \subseteq X$. Let $\alpha_0 \in X$ be such that the sum of the coefficients is as large as possible. Write $\alpha_0 = \sum_{j\in I} n_j\alpha_j$ where $n_j \geqslant 0$ for all $j \in I$. If $m := \langle \alpha_i^\vee, \alpha_0\rangle < 0$ for some $i \in I$, then

$$s_i(\alpha_0) = \alpha_0 - \langle \alpha_i^\vee, \alpha_0\rangle\alpha_i = \underbrace{(n_i - m)}_{>n_i}\alpha_i + \sum_{\substack{j\in I\\ j\neq i}} n_j\alpha_j \in \Phi,$$

where all coefficients are still non-negative but the sum of the coefficients is strictly larger than that of α_0, contradiction. So we must have $\langle \alpha_i^\vee, \alpha_0\rangle \geqslant 0$ for all $i \in I$. But this means $\sum_{j\in I} a_{ij}n_j = \sum_{j\in I} n_j\langle \alpha_i^\vee, \alpha_j\rangle \geqslant 0$. So, if $u := (n_j)_{j\in J} \in \mathbb{R}^J$, then $u \geqslant 0$, $u \neq 0$, and $Au \geqslant 0$. Since $\det(A) \neq 0$, we also have $Au \neq 0$. So A is of type (FIN) by Corollary 3.1.3(a). □

Table 6. Positive roots for exceptional types F_4, E_6, E_7, E_8

Type F_4, $\lvert\Phi^+\rvert = 24$:	1000	0100	0010	0001	1100	0110
0011	1110	0120	0111	1120	1111	0121
1220	1121					
0122	1221	1122	1231	1222	1232	1242
1342	2342					

Type E_6, $\lvert\Phi^+\rvert = 36$:				100000	010000	001000	000100
000010	000001	101000	010100	001100	000110	000011	101100
011100	010110	001110	000111	111100	101110	011110	010111
001111	111110	101111	011210	011111	111210	111111	011211
112210	111211	011221	112211	111221	112221	112321	122321

Type E_7, $\lvert\Phi^+\rvert = 63$:

1000000	0100000	0010000	0001000	0000100	0000010	0000001
1010000	0101000	0011000	0001100	0000110	0000011	1011000
0111000	0101100	0011100	0001110	0000111	1111000	1011100
0111100	0101110	0011110	0001111	1111100	1011110	0112100
0111110	0101111	0011111	1112100	1111110	1011111	0112110
0111111	1122100	1112110	1111111	0112210	0112111	1122110
1112210	1112111	0112211	1122210	1122111	1112211	0112221
1123210	1122211	1112221	1223210	1123211	1122221	1223211
1123221	1223221	1123321	1223321	1224321	1234321	2234321

Type E_8, $\lvert\Phi^+\rvert = 120$:

10000000	01000000	00100000	00010000	00001000	00000100
00000010	00000001	10100000	01010000	00110000	00011000
00001100	00000110	00000011	10110000	01110000	01011000
00111000	00011100	00001110	00000111	11110000	10111000
01111000	01011100	00111100	00011110	00001111	11111000
10111100	01121000	01111100	01011110	00111110	00011111
11121000	11111100	10111110	01121100	01111110	01011111
00111111	11221000	11121100	11111110	10111111	01122100
01121110	01111111	11221100	11122100	11121110	11111111
01122110	01121111	11222100	11221110	11122110	11121111
01122210	01122111	11232100	11222110	11221111	11122210
11122111	01122211	12232100	11232110	11222210	11222111
11122211	01122221	12232110	11232210	11232111	11222211
11122221	12232210	12232111	11233210	11232211	11222221
12233210	12232211	11233211	11232221	12243210	12233211
12232221	11233221	12343210	12243211	12233221	11233321
22343210	12343211	12243221	12233321	22343211	12343221
12243321	22343221	12343321	12244321	22343321	12344321
22344321	12354321	22354321	13354321	23354321	22454321
23454321	23464321	23465321	23465421	23465431	23465432

For example, 2342 stands for $2\alpha_1{+}3\alpha_2{+}4\alpha_3{+}2\alpha_4$, $I = \{1,2,3,4\}$.

Proposition 3.2.4. *Assume that A is indecomposable and of type* (FIN). *Then $|W(A)| < \infty$ and $|\Phi(A)| < \infty$. Furthermore, $(\Phi(A), \Delta)$ is a based root system, that is, every $\alpha \in \Phi(A)$ can be written as a $\mathbb{Z}$-linear combination of $\Delta = \{\alpha_i \mid i \in I\}$, where the coefficients are either all $\geqslant 0$ or all $\leqslant 0$ (as in condition* (CK2) *of Definition 2.2.1). Finally, $\Phi(A)$ is reduced, that is, $\Phi(A) \cap \mathbb{R}\alpha = \{\pm\alpha\}$ for all $\alpha \in \Phi(A)$.*

Proof. We use the classification in Section 3.1 and go through the list of Dynkin diagrams in Table 4. If A has a diagram of type A_n, B_n, C_n or D_n, then $\Phi(A)$ has been explicitly described in Chapter 2; the desired properties hold by Example 2.2.7 and Corollary 2.5.6.

Now assume that A has a diagram of type G_2, F_4, E_6, E_7, or E_8. Then we take a "*computer algebra approach*", based on our Python programs in Table 1 (p. 61). We apply the program `rootsystem` to A; the program actually terminates and outputs a finite list of tuples $\mathscr{C}^+(A) \subseteq \mathbb{N}_0^I$. For example, for type G_2, we obtain:

$$\{(1,0), (0,1), (1,1), (1,2), (1,3), (2,3)\} \quad \text{(see also Example 2.3.10).}$$

For the types F_4, E_6, E_7, E_8, these vectors are explicitly listed in Table 6. Now we set $\Phi := \Phi^+ \cup (-\Phi^+)$, where

$$\Phi^+ := \Big\{\alpha := \sum_{i\in I} n_i\alpha_i \mid (n_i)_{i\in I} \in \mathscr{C}^+(A)\Big\} \subseteq E.$$

By construction, it is clear that $\Phi^+ \subseteq \Phi(A)$. Since $s_i(\alpha_i) = -\alpha_i$ for $i \in I$, it also follows that $-\Phi^+ \subseteq \Phi(A)$. Now we apply our program `refl` to all tuples in $\mathscr{C}^+(A) \cup (-\mathscr{C}^+(A))$. By inspection, we find that $\mathscr{C}^+(A) \cup (-\mathscr{C}^+(A))$ remains invariant under these operations. In other words, we have $s_i(\Phi) \subseteq \Phi$ for all $i \in I$ (recall that `refl` corresponds to applying s_i to an element of E). Since $\Delta \subseteq \Phi$, we conclude that $\Phi(A) \subseteq \Phi$ and, hence, that $\Phi(A) = \Phi$; in particular, $|\Phi(A)| < \infty$. The fact that $(\Phi(A), \Delta)$ is a based root system is clear because all tuples in $\mathscr{C}^+(A)$ have non-negative entries. The fact that $\Phi(A)$ is reduced is seen by inspection of Table 6. □

Further properties of the root system of type E_8 can be found at `https://en.wikipedia.org/wiki/E8_lattice`.

Remark 3.2.5. Of course, one can avoid the classification and the use of computer algebra methods in order to obtain the above result.

The finiteness of $W(A)$ follows from a topological argument, based on the fact that $W(A)$ is a discrete, bounded subset of $\mathrm{GL}(E)$; see, e.g., [**4**, Ch. V, §4. no. 8]. The fact that $(\Phi(A), \Delta)$ is based requires a more elaborate argument.

Let us fix an indecomposable generalised Cartan matrix $A = (a_{ij})_{i,j\in I}$ of finite type; let $W = W(A)$ and $\Phi = \Phi(A)$. We now turn to the discussion of some specific properties of W and Φ, which can be derived from the classification in Section 3.1. Let us fix a W-invariant scalar product $\langle\ ,\ \rangle\colon E \times E \to \mathbb{R}$ as in Remark 3.2.2. For $\alpha \in \Phi$, the number $\sqrt{\langle \alpha, \alpha\rangle} \in \mathbb{R}_{>0}$ will be called the *length of* α. As before, we write $\alpha^\vee := 2\alpha/\langle\alpha, \alpha\rangle \in E$ for any $\alpha \in \Phi$.

Remark 3.2.6. First we note that the arrows in the Dynkin diagrams in Table 4 indicate the relative length of the roots α_i $(i \in I)$. More precisely, let $i \neq j$ in I be joined by a possibly multiple edge; then $a_{ij} < 0$ and $a_{ji} < 0$. We choose the notation such that $a_{ij} = \langle \alpha_i^\vee, \alpha_j\rangle = -1$ and $a_{ji} = \langle\alpha_j^\vee, \alpha_i\rangle = -r$, where $r \geqslant 1$. Then

$$2\frac{\langle\alpha_i, \alpha_j\rangle}{\langle\alpha_j, \alpha_j\rangle} = a_{ji} = -r = a_{ij}r = 2\frac{\langle\alpha_i, \alpha_j\rangle}{\langle\alpha_i, \alpha_i\rangle}r$$

and so $\langle\alpha_i, \alpha_i\rangle = r\langle\alpha_j, \alpha_j\rangle$. Now set $m := \min\{\langle\alpha_i, \alpha_i\rangle \mid i \in I\}$ and $e := \max\{-a_{ij} \mid i, j \in I, i \neq j, a_{ij} \neq 0\}$. By inspection of Table 4, we conclude that we are in one of the following two cases.

(a) $e = 1$ (the *simply-laced* case). This is the case for A of type A_n, D_n, E_6, E_7, E_8. Then $\langle\alpha_i, \alpha_i\rangle = m$ for all $i \in I$.

(b) $e \in \{2, 3\}$. This is the case for A of type B_n, C_n, F_4 $(e = 2)$ or G_2 $(e = 3)$. Then $\langle\alpha_i, \alpha_i\rangle \in \{m, em\}$ for all $i \in I$.

Now consider any $\alpha \in \Phi$. By definition, we can write $\alpha = w(\alpha_i)$ where $i \in I$ and $w \in W$. So $\langle\alpha, \alpha\rangle = \langle w(\alpha_i), w(\alpha_i)\rangle = \langle\alpha_i, \alpha_i\rangle$, by the W-invariance of $\langle\ ,\ \rangle$. Hence, we conclude that

(c) $\langle\alpha, \alpha\rangle \in \{m, em\}$ for all $\alpha \in \Phi$.

Thus, in case (a), all roots in Φ have the same length; in case (b), there are precisely two root lengths in Φ and so we may speak of *short roots* and *long roots*. In case (a), we declare all roots to be long roots.

Lemma 3.2.7. *Assume that A is indecomposable. Let $e \geqslant 1$ be as in Remark 3.2.6. Then $\langle \alpha^\vee, \beta \rangle \in \{0, \pm 1, \pm e\}$ for all $\alpha, \beta \in \Phi$, $\beta \neq \pm\alpha$.*

Proof. Let $\alpha, \beta \in \Phi$. We can write $\alpha = w(\alpha_i)$ for some $w \in W$ and $i \in I$. Setting $\beta' := w^{-1}(\beta) \in \Phi$, we obtain

$$\langle \alpha^\vee, \beta \rangle = 2\frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} = 2\frac{\langle w(\alpha_i), w(\beta') \rangle}{\langle w(\alpha_i), w(\alpha_i) \rangle} = 2\frac{\langle \alpha_i, \beta' \rangle}{\langle \alpha_i, \alpha_i \rangle} = \langle \alpha_i^\vee, \beta' \rangle,$$

where we used the W-invariance property of $\langle\ ,\ \rangle$. Writing $\beta = \sum_{j \in I} n_j \alpha_j$ with $n_j \in \mathbb{Z}$, the right hand side evaluates to $\sum_{j \in I} n_j a_{ij} \in \mathbb{Z}$. Thus, $\langle \alpha^\vee, \beta \rangle \in \mathbb{Z}$ for all $\alpha, \beta \in \Phi$. Now let $\beta \neq \pm\alpha$. Assume that $|\langle \alpha^\vee, \beta \rangle| \geqslant 2$. Using the *Cauchy–Schwartz inequality* as in Section 2.6 (see ($\spadesuit_2$), p. 84), we conclude that $\langle \alpha, \beta^\vee \rangle = \pm 1$ and so

$$\langle \alpha^\vee, \beta \rangle = 2\frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} = 2\frac{\langle \alpha, \beta \rangle}{\langle \beta, \beta \rangle}\frac{\langle \beta, \beta \rangle}{\langle \alpha, \alpha \rangle} = \frac{\langle \beta, \beta \rangle}{\langle \alpha, \alpha \rangle}\langle \alpha, \beta^\vee \rangle = \pm\frac{\langle \beta, \beta \rangle}{\langle \alpha, \alpha \rangle}.$$

The left hand side is an integer and the right side equals $\pm e$ or $\pm e^{-1}$; see Remark 3.2.6(c). Hence, we must have $\langle \alpha^\vee, \beta \rangle = \pm e$. ☐

Exercise 3.2.8. Assume that A is indecomposable and Φ is simply-laced. Let $\alpha, \beta \in \Phi$ be such that $\beta \neq \pm\alpha$. By Lemma 3.2.7, we have $\langle \alpha^\vee, \beta \rangle \in \{0, \pm 1\}$. Then show the following implications:

$$\begin{aligned} \langle \alpha^\vee, \beta \rangle &= 0 \quad \Rightarrow \quad \beta - \alpha \notin \Phi \text{ and } \beta + \alpha \notin \Phi, \\ \langle \alpha^\vee, \beta \rangle &= +1 \quad \Rightarrow \quad \beta - \alpha \in \Phi,\ \beta - 2\alpha \notin \Phi \text{ and } \beta + \alpha \notin \Phi, \\ \langle \alpha^\vee, \beta \rangle &= -1 \quad \Rightarrow \quad \beta + \alpha \in \Phi,\ \beta + 2\alpha \notin \Phi \text{ and } \beta - \alpha \notin \Phi. \end{aligned}$$

Show that, if $\alpha \in \Phi$ is written as $\alpha = \sum_{i \in I} n_i \alpha_i$ with $n_i \in \mathbb{Z}$, then $\alpha^\vee = \sum_{i \in I} n_i \alpha_i^\vee$ (see also Lemma 2.6.3).

Remark 3.2.9. In Section 2.5, we have given an explicit description of the Weyl group $W(A)$ for A of type A_n. (Similar descriptions exist also for type B_n, C_n, D_n.) Now assume that A is of type G_2, F_4, E_6, E_7 or E_8. For G_2, the computation in Example 2.3.10 shows that $W(A)$ is a dihedral group of order 12. For the remaining types, we use again a "*computer algebra approach*" to determine the order $|W(A)|$. Let us write $\Phi^+ = \{\alpha_1, \ldots, \alpha_N\}$, where the roots are ordered in the same way as in Table 6. Then

$$\Phi = \Phi^+ \cup (-\Phi^+) = \{\alpha_1, \ldots, \alpha_N, \alpha_{N+1}, \ldots, \alpha_{2N}\} \subseteq E,$$

where $\alpha_{N+l} = -\alpha_l$ for $1 \leqslant l \leqslant N$. As discussed above, we can identify $W(A)$ with a subgroup of the symmetric group $\mathfrak{S}_{2N} \cong \mathrm{Sym}(\Phi)$. The permutation $\sigma_i \in \mathfrak{S}_{2n}$ corresponding to $s_i \in W(A)$ is obtained by applying s_i to a root α_l and identifying $l' \in \{1, \ldots, 2N\}$ such that $s_i(\alpha_l) = \alpha_{l'}$; then $\sigma_i(l) = l'$. Now, a computer algebra system like GAP [**12**] contains built-in algorithms to work with permutation groups; in particular, there are efficient algorithms to determine the order of such a group. In this way, we find the numbers in Table 7.

Table 7. Highest roots and $|W(A)|$ (labelling as in Table 4, p. 104)

Type	Highest root α_0	$\lvert W(A)\rvert$
A_n $(n\geqslant 1)$	$\alpha_1 + \alpha_2 + \ldots + \alpha_n$	$(n+1)!$
B_n $(n\geqslant 2)$	$2(\alpha_1 + \alpha_2 + \ldots + \alpha_{n-1}) + \alpha_n$	$2^n n!$
C_n $(n\geqslant 2)$	$\alpha_1 + 2(\alpha_2 + \ldots + \alpha_{n-1} + \alpha_n)$	$2^n n!$
D_n $(n\geqslant 3)$	$\alpha_1 + \alpha_2 + 2(\alpha_3 + \ldots + \alpha_{n-1}) + \alpha_n$	$2^{n-1} n!$
G_2	$2\alpha_1 + 3\alpha_2$	12
F_4	$2\alpha_1 + 3\alpha_2 + 4\alpha_3 + 2\alpha_4$	1152
E_6	$\alpha_1 + 2\alpha_2 + 2\alpha_3 + 3\alpha_4 + 2\alpha_5 + \alpha_6$	51840
E_7	$2\alpha_1 + 2\alpha_2 + 3\alpha_3 + 4\alpha_4 + 3\alpha_5 + 2\alpha_6 + \alpha_7$	2903040
E_8	$2\alpha_1 + 3\alpha_2 + 4\alpha_3 + 6\alpha_4 + 5\alpha_5 + 4\alpha_6 + 3\alpha_7 + 2\alpha_8$	696729600

Remark 3.2.10. As in Remark 2.3.5, we can define a linear map $\mathrm{ht}\colon E \to \mathbb{R}$ such that $\mathrm{ht}(\alpha_i) = 1$ for all $i \in I$. If $\alpha \in \Phi$ and $\alpha = \sum_{i\in I} n_i\alpha_i$ with $n_i \in \mathbb{Z}$, then $\mathrm{ht}(\alpha) = \sum_{i\in I} n_i \in \mathbb{Z}$ is called the *height* of α. Assuming that A is indecomposable, there is a unique root $\alpha_0 \in \Phi$ such that $\mathrm{ht}(\alpha_0)$ takes its maximum value; this root α_0 is called the *highest root* of Φ. One can prove this by a general argument (see, e.g., [**18**, §10.4], or Proposition 2.4.17), but here we can simply extract this from our knowledge of all root systems, using Example 2.2.7 (A_n), Remark 2.5.5 (B_n, C_n, D_n), Example 2.3.10 (G_2) and Table 6 (F_4, E_6, E_7, E_8). See Table 7 for explicit expressions of α_0 in terms of Δ.

3.3. A glimpse of Kac–Moody theory

ab hier Woche 10

Let I be a finite, non-empty index set and $A = (a_{ij})_{i,j\in I} \in M_I(\mathbb{C})$ be arbitrary with entries in $\mathbb{C}$. We would like to study Lie algebras for which A plays the role as a "structure matrix". In order to find out how this could possibly work, let us first return to the case where A is the true structure matrix of a Lie algebra L of Cartan–Killing type with respect to an abelian subalgebra $H \subseteq L$ and a subset $\Delta = \{\alpha_i \mid i \in I\}$, as in Section 2.2. Then we have

(Ch0) $\quad L = \langle e_i, h_i, f_i \mid i \in I\rangle_{\text{alg}}$

for a suitable collection of elements $\{e_i, h_i, f_i \mid i \in I\} \subseteq L$ such that the following "*Chevalley relations*" hold:

(Ch1) $\quad [e_i, f_i] = h_i$ and $[e_i, f_j] = 0$ for $i, j \in I$ such that $i \neq j$,

(Ch2) $\quad [h_i, h_j] = 0,\ [h_i, e_j] = a_{ij}e_j,\ [h_i, f_j] = -a_{ij}f_j$ for $i, j \in I$.

Indeed, let $0 \neq h_i \in H$ $(i \in I)$ as in Proposition 2.2.5. Then we have $H = \langle h_i \mid i \in I\rangle_{\mathbb{C}}$; furthermore, $h_i = [e_i, f_i]$ for suitable $e_i \in L_{\alpha_i}$ and $f_i \in L_{-\alpha_i}$. Since H is abelian, $[h_i, h_j] = 0$ for all $i, j \in I$. By Proposition 2.4.6, (Ch0) holds. By the definition of A, we have $[h_i, e_j] = \alpha_j(h_i)e_j = a_{ij}e_j$ and $[h_i, f_j] = -\alpha_j(h_i) = -a_{ji}f_j$ for all $i, j \in I$. Finally, if $i \neq j$, then $[e_i, f_j] \in [L_{\alpha_i}, L_{-\alpha_j}] \subseteq L_{\alpha_i - \alpha_j} = \{0\}$, where the last two relations hold by Proposition 2.1.6 and condition (CK2) in Definition 2.2.1. Thus, $[e_i, f_j] = 0$ for $i \neq j$. So, indeed, (Ch0), (Ch1), (Ch2) hold for L.

Now we notice that (Ch0), (Ch1), (Ch2) only refer to the collection of elements $\{e_i, h_i, f_i \mid i \in I\} \subseteq L$ and the entries of A, but not to any further structural properties of L (e.g., finite dimension or H-diagonalisability). Presenting things in this way, it seems obvious how to proceed: given any $A \in M_I(\mathbb{C})$, we try to consider a Lie algebra L for which there exist elements $\{e_i, h_i, f_i \mid i \in I\}$ such that (Ch0), (Ch1), (Ch2) hold. Two basic questions present themselves:

- Does L exist at all?
- If yes, then does L have interesting structural properties?

As Kac and Moody (independently) discovered in the 1960s, both questions have affirmative answers, and this has led to a new area of research with many interesting applications and connections, for example, to mathematical physics, especially when A is a generalised Cartan matrix of type (AFF); see the monographs [**21**], [**25**]. What we will do in this section is the following:

- exhibit the ingredients of a "triangular decomposition" in any Lie algebra L satisfying (Ch0), (Ch1), (Ch2);
- apply these ideas to prove the Existence Theorem 3.3.10.

So let us assume now that we are given any $A \in M_I(\mathbb{C})$ and a Lie algebra L, together with elements $\{e_i, h_i, f_i \mid i \in I\}$ such that the conditions (Ch0), (Ch1), (Ch2) hold. In order to avoid the discussion of trivial cases, we assume throughout that

$$e_j \neq 0 \qquad \text{or} \qquad f_j \neq 0 \qquad \text{for each } j \in I.$$

(Note that, if $e_j = f_j = 0$ for some j, then also $h_j = 0$ by (Ch1) and e_j, h_j, f_j can simply be omitted from the collection $\{e_i, h_i, f_i \mid i \in I\}$.)

Lemma 3.3.1. *In the above setting, let $H := \langle h_i \mid i \in I\rangle_{\mathbb{C}} \subseteq L$. Then H is abelian and there is a well-defined collection of linear maps*

$$\Delta := \{\alpha_j \mid j \in I\} \subseteq H^*, \quad \textit{where } \alpha_j(h_i) = a_{ij} \textit{ for all } i, j \in I.$$

The set $\Delta \subseteq H^*$ is linearly independent if and only if $\det(A) \neq 0$.

Proof. By (Ch2), H is an abelian subalgebra of L. Next we want to define $\alpha_j \in H^*$ for $j \in I$. Let $h \in H$ and write $h = \sum_{i\in I} x_i h_i$ where $x_i \in \mathbb{C}$. Then set $\alpha_j(h) := \sum_{i \in I} x_i a_{ij}$. We must show that this is well-defined. So assume that we also have $h = \sum_{i\in I} y_i h_i$ where $y_i \in \mathbb{C}$. Then $\sum_{i\in I}(x_i - y_i)h_i = 0$; using (Ch2), we obtain:

$$0 = \sum_{i\in I}(x_i - y_i)[h_i, e_j] = \Big(\sum_{i\in I}(x_i - y_i)a_{ij}\Big)e_j.$$

If $e_j \neq 0$, then this implies that $\sum_{i\in I} x_i a_{ij} = \sum_{i\in I} y_i a_{ij}$, as desired. If $f_j \neq 0$, then an analogous argument using the relation $[h_i, f_j] = -a_{ij} f_j$ yields the same conclusion. Thus, we obtain a well-defined subset $\Delta = \{\alpha_j \mid j \in I\} \subseteq H^*$ as above. Now let $x_j \in \mathbb{C}$ $(j \in I)$ be such that $\sum_{j\in I} x_j \alpha_j = \underline{0}$. Then

$$0 = \sum_{j\in I} x_j \alpha_j(h_i) = \sum_{j\in I} a_{ij} x_j \qquad \text{for all } i \in I.$$

If $\det(A) \neq 0$, then this implies $x_j = 0$ for all j and so Δ is linearly independent. Conversely, if $\det(A) = 0$, then there exist $x_j \in \mathbb{C}$ $(j \in I)$, not all equal to zero, such that $\sum_{i\in I} a_{ij} x_j = 0$ for all $i \in I$. Then we also have $\sum_{j\in I} x_j \alpha_j = \underline{0}$ and so Δ is linearly dependent. $\square$

Example 3.3.2. Let $R = \mathbb{C}[T, T^{-1}]$ be the ring of Laurent polynomials over $\mathbb{C}$ with indeterminate T. We consider the Lie algebra

$$L = \left\{ \left(\begin{array}{cc} a & b \\ c & -a \end{array} \right) \mid a, b, c \in R \right\} \quad \big(= \mathfrak{sl}_2(R)\big),$$

with the usual Lie bracket for matrices. A vector space basis of L is given by $\{T^k e_1, T^l h_1, T^m f_1 \mid k, l, m \in \mathbb{Z}\}$, where we set as usual:

$$e_1 := \left(\begin{array}{cc} 0 & 1 \\ 0 & 0 \end{array} \right), \qquad h_1 := \left(\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array} \right), \qquad f_1 := \left(\begin{array}{cc} 0 & 0 \\ 1 & 0 \end{array} \right),$$

with relations $[e_1, f_1] = h_1$, $[h_1, e_1] = 2e_1$, $[h_1, f_1] = -2f_1$. Now set

$$e_2 := Tf_1, \qquad h_2 := -h_1, \qquad f_2 := T^{-1}e_1.$$

Then you will see in the exercises that the Chevalley relations (Ch0), (Ch1), (Ch2) hold with respect to the matrix

$$A = \left(\begin{array}{cc} 2 & -2 \\ -2 & 2 \end{array} \right) \qquad \text{(affine type } \tilde{A}_1 \text{ in Table 5).}$$

Returning to the general setting, let $H \subseteq L$ be as in Lemma 3.3.1. Then $\dim H < \infty$ but we have no information at all about $\dim L$. We can still adopt a large portion of the definitions and results concerning weights and weight spaces from Section 2.1. For any $\lambda \in H^*$, we set

$$L_\lambda := \{x \in L \mid [h, x] = \lambda(h)x \text{ for all } h \in H\};$$

this is a subspace of L. If $L_\lambda \neq \{0\}$, then λ is called a *weight* and L_λ the corresponding *weight space*. Since H is abelian, we have $H \subseteq L_{\underline{0}}$, where $\underline{0} \in H^*$ is the 0-map. The same argument as in Proposition 2.1.6 shows that $[L_\lambda, L_\mu] \subseteq L_{\lambda+\mu}$ for all $\lambda, \mu \in H^*$. Let us set

$$Q_{\geqslant 0} := \big\{\lambda \in H^* \mid \lambda = \textstyle\sum_{i \in I} n_i \alpha_i \text{ where } n_i \in \mathbb{Z}_{\geqslant 0} \text{ for all } i\big\},$$
$$Q_{\leqslant 0} := \big\{\lambda \in H^* \mid \lambda = \textstyle\sum_{i \in I} n_i \alpha_i \text{ where } n_i \in \mathbb{Z}_{\leqslant 0} \text{ for all } i\big\}.$$

In the following discussion, some care is needed because Δ may be linearly dependent, and so it might happen that $Q_{\geqslant 0} \cap Q_{\leqslant 0} \neq \{\underline{0}\}$.

Lemma 3.3.3. *In the above setting, we have*

$$N^+ := \langle e_i \mid i \in I \rangle_{\text{alg}} \subseteq \textstyle\sum_{\lambda \in Q_{\geqslant 0}} L_\lambda,$$
$$N^- := \langle f_i \mid i \in I \rangle_{\text{alg}} \subseteq \textstyle\sum_{\lambda \in Q_{\leqslant 0}} L_\lambda.$$

In particular, we have $[H, N^+] \subseteq N^+$ *and* $[H, N^-] \subseteq N^-$.

Proof. Recall from Section 1.1 that $N^+ = \langle X_n \mid n \geqslant 1\rangle_{\mathbb{C}}$, where X_n consists of all Lie monomials in $\{e_i \mid i \in I\}$ of level n. By (Ch2) and the definition of α_i, we have $e_i \in L_{\alpha_i}$ for all $i \in I$. Hence, exactly as in Lemma 2.1.7, one sees that $X_n \subseteq \bigcup_\lambda L_\lambda$, where the union runs over all $\lambda \in Q_{\geqslant 0}$ that can be expressed as $\lambda = \sum_{i\in I} n_i\alpha_i$ with $\sum_{i\in I} n_i = n \geqslant 1$. This yields that

$$N^+ \subseteq \sum_{\lambda \in Q_{\geqslant 0}} L_\lambda \qquad \text{and} \qquad [H, N^+] \subseteq N^+.$$

The argument for N^- is completely analogous, starting with the fact that $f_i \in L_{-\alpha_i}$ for all $i \in I$. □

Lemma 3.3.4. *We have* $L = N^+ + H + N^-$.

Proof. The crucial property to show is that $[f_j, N^+] \subseteq N^+ + H$ for all $j \in I$. This is done as follows. As in the above proof, N^+ is spanned by Lie monomials in $\{e_i \mid i \in I\}$. So it is sufficient to show that $[f_j, x] \in N^+ + H$, where $x \in N^+$ is a Lie monomial of level, say $n \geqslant 1$. We proceed by induction on n. If $n = 1$, then $x = e_i$ for some i and so $[f_j, x] = -[e_i, f_j]$ is either zero or equal to $h_i \in H$. So the assertion holds in this case. Now let $n \geqslant 2$. Then $x = [y, z]$ where $y, z \in N^+$ are Lie monomials of level k and $n - k$, respectively; here, $1 \leqslant k \leqslant n - 1$. Using the Jacobi identity, we obtain

$$[f_j, x] = [f_j, [y, z]] = -[y, [z, f_j]] - [z, [f_j, y]] = [y, [f_j, z]] + [[f_j, y], z].$$

By induction, we can write $[f_j, z] = z' + h$, where $z' \in N^+$ and $h \in H$. This yields $[y, [f_j, z]] = [y, z'] + [y, h] = [y, z'] - [h, y] \in N^+ + H$. (We have $[y, z'] \in N^+$ by the definition of N^+, and $[h, y] \in N^+$ by Lemma 3.3.3.) Similarly, one sees that $[[f_j, y], z] \in N^+ + H$.

Thus, we have shown that $[f_j, N^+] \subseteq N^+ + H$ for all $j \in I$. By an analogous argument, one also shows that $[e_j, N^-] \subseteq N^- + H$ for all $j \in I$. Furthermore, $[e_j, H] \subseteq N^+$ and $[f_j, H] \subseteq N^-$ for all $j \in I$. Hence, setting $V := N^+ + H + N^- \subseteq L$, we conclude that

$$[e_j, V] \subseteq V \qquad \text{and} \qquad [f_j, V] \subseteq V \quad \text{for all } j \in I.$$

By Lemma 3.3.3, we also have $[h_j, V] \subseteq V$. By (Ch0), we have $L = \langle e_j, h_j, f_j \mid j \in I\rangle_{\text{alg}}$ and so Exercise 1.1.8(a) implies that $[L, V] \subseteq V$.

In particular, V is a subalgebra. Since V contains all generators of L, we must have $L = V$. $\square$

Exercise 3.3.5. In the setting of Example 3.3.2, we certainly have $H = \langle h_1, h_2 \rangle_{\mathbb{C}} = \langle h_1 \rangle_{\mathbb{C}}$. Explicitly determine the subalgebras $N^+ \subseteq L$ and $N^- \subseteq L$. Show that $L = N^+ \oplus H \oplus N^-$.

Lemma 3.3.6. *If* $\det(A) \neq 0$*, then the sum in Lemma 3.3.4 is direct; furthermore,* $H = L_{\underline{0}}$*,* $N^+ = \sum_{\lambda \in Q_{\geqslant 0}} L_\lambda$ *and* $N^- = \sum_{\lambda \in Q_{\leqslant 0}} L_\lambda$.

Proof. By Lemma 3.3.1, the assumption that $\det(A) \neq 0$ implies that $\Delta = \{\alpha_i \mid i \in I\} \subseteq H^*$ is linearly independent. This has the following consequence. In the proof of Lemma 3.3.3, we have seen that $N^+ \subseteq \sum_\lambda L_\lambda$, where the sum runs over all $\lambda \in Q_{\geqslant 0}$ that can be expressed as $\lambda = \sum_{i \in I} n_i \alpha_i$ with $\sum_{i \in I} n_i \geqslant 1$; in particular, $n_i > 0$ for at least some i, and so $\lambda \neq \underline{0}$. This shows that

$$N^+ \subseteq \sum_{\lambda \in Q_+} L_\lambda \qquad \text{where} \qquad Q_+ := \{\lambda \in Q_{\geqslant 0} \mid \lambda \neq \underline{0}\}.$$

Similary, we have $N^- \subseteq \sum_{\lambda \in Q_-} L_\lambda$, where $Q_- := \{\lambda \in Q_{\leqslant 0} \mid \lambda \neq \underline{0}\}$. Combined with Lemma 3.3.4, we obtain:

$$L = N^+ + H + N^- \subseteq \Big(\sum_{\lambda \in Q_+} L_\lambda\Big) + L_{\underline{0}} + \Big(\sum_{\mu \in Q_-} L_\mu\Big).$$

So it is sufficient to show that the sum on the right hand side is direct. Let $x \in L_{\underline{0}}$, $y \in \sum_{\lambda \in Q_+} L_\lambda$ and $z \in \sum_{\mu \in Q_-} L_\mu$ be such that $y+x+z = 0$. We must show that $x = y = z = 0$. Assume, if possible, that $x \neq 0$. Then $x \in L_{\underline{0}}$ and $x = -y - z \in L_{\lambda_1} + \ldots + L_{\lambda_r}$, where $r \geqslant 1$ and $\underline{0} \neq \lambda_i \in Q_+ \cup Q_-$ for all i. But then Exercise 2.1.5 (which also holds without any assumption on dimensions) shows that $\lambda_i = \underline{0}$ for some i, contradiction. $\square$

Even if $\det(A) = 0$, the statement of Lemma 3.3.6 remains true, but the proof requires a more subtle argument; see Kac [**21**, Theorem 1.2] or Moody–Pianzola [**25**, §4.2, Prop. 5]. The connection with Lie algebras of Cartan–Killing type is as follows.

Proposition 3.3.7. *Let* $A = (a_{ij})_{i,j \in I} \in M_I(\mathbb{C})$ *and* L *be a Lie algebra for which there exist elements* $\{e_i, h_i, f_i \mid i \in I\} \subseteq L$ *such that*

(Ch0) *and the Chevalley relations* (Ch1), (Ch2) *hold (and, for each* $j \in I$*, we have* $e_j \neq 0$ *or* $f_j \neq 0$*). Let*

$$H := \langle h_i \mid i \in I \rangle_{\mathbb{C}} \subseteq L \qquad \textit{and} \qquad \Delta := \{\alpha_j \mid j \in I\} \subseteq H^*$$

be defined as in Lemma 3.3.1. Assume that $\dim L < \infty$ *and* $\det(A) \neq 0$*. Then* (L, H) *is of Cartan–Killing type with respect to* Δ*; if* $a_{ii} = 2$ *for all* $i \in I$*, then* A *is the corresponding structure matrix.*

Proof. By Lemma 3.3.1, the set $\Delta \subseteq H^*$ is linearly independent. By Lemma 3.3.6, L is H-diagonalisable and $L_{\underline{0}} = H$; furthermore, every weight $\underline{0} \neq \lambda \in P(L)$ belongs to Q_+ or Q_-. Thus, (CK1) and (CK2) in Definition 2.2.1 hold. Finally, since $e_i \in L_{\alpha_i}$ and $f_i \in L_{-\alpha_i}$ for all $i \in I$, we have $h_i = [e_i, f_i] \in [L_{\alpha_i}, L_{-\alpha_i}]$ by (Ch1). Since $H = \langle h_i \mid i \in I \rangle_{\mathbb{C}}$, we conclude that (CK3) also holds. Now assume that $a_{ii} = 2$ for all $i \in I$. Then $\alpha_i(h_i) = 2$ and so the elements $\{h_i \mid i \in I\}$ are the elements required in Definition 2.2.6. □

We now use the above ideas to solve a question that was left open in Chapter 2. Let A be an indecomposable generalised Cartan matrix of type (FIN). We have seen that, if A is of type A_n, B_n, C_n or D_n, then A arises as the structure matrix of a Lie algebra of Cartan–Killing type (namely, from $L = \mathfrak{sl}_n(\mathbb{C})$ or $L = \mathfrak{go}_n(Q_n, \mathbb{C})$, for suitable choices of Q_n). But what about A of type G_2, F_4, E_6, E_7, or E_8 ? For example, at the end of Section 2.6, we saw that all the Lie brackets inside a Lie algebra of type G_2 are easily determined — although we did not know if such an algebra exists at all. (In principle, the same could be done for the types F_4, E_6, E_7 and E_8.) We now present a general solution of the existence problem.

Definition 3.3.8 (Cf. [**14**]). Let $A = (a_{ij})_{i,j \in I}$ be an indecomposable generalised Cartan matrix of type (FIN) (where $I \neq \varnothing$). As in Section 3.2, consider an $\mathbb{R}$-vector space E with a basis $\{\alpha_i \mid i \in I\}$, and let $\Phi = \Phi(A) \subseteq E$ be the abstract root system determined by A. (We have $|\Phi| < \infty$ by Proposition 3.2.4.) Having obtained the set Φ, let $\mathbf{M}$ be a $\mathbb{C}$-vector space with a basis

$$\mathbf{B} := \{u_i \mid i \in I\} \cup \{v_\alpha \mid \alpha \in \Phi\}; \qquad \dim \mathbf{M} = |I| + |\Phi|.$$

Taking the formulae in Lusztig's Theorem 2.7.2 as a model, we define for each $i \in I$ linear maps $\mathbf{e}_i \colon \mathbf{M} \to \mathbf{M}$ and $\mathbf{f}_i \colon \mathbf{M} \to \mathbf{M}$ as follows, where $j \in I$ and $\alpha \in \Phi$:

$$\mathbf{e}_i(u_j) := |a_{ji}| v_{\alpha_i}, \qquad \mathbf{f}_i(u_j) := |a_{ji}| v_{-\alpha_i},$$

$$\mathbf{e}_i(v_\alpha) := \begin{cases} (q_{i,\alpha}+1) v_{\alpha+\alpha_i} & \text{if } \alpha+\alpha_i \in \Phi, \\ u_i & \text{if } \alpha = -\alpha_i, \\ 0 & \text{otherwise,} \end{cases}$$

$$\mathbf{f}_i(v_\alpha) := \begin{cases} (p_{i,\alpha}+1) v_{\alpha-\alpha_i} & \text{if } \alpha-\alpha_i \in \Phi, \\ u_i & \text{if } \alpha = \alpha_i, \\ 0 & \text{otherwise.} \end{cases}$$

It is obvious that the maps $\mathbf{e}_i, \mathbf{f}_i$ are all non-zero. Now consider the Lie algebra $\mathfrak{gl}(\mathbf{M})$, with the usual Lie bracket $[\varphi, \psi] = \varphi \circ \psi - \psi \circ \varphi$ for $\varphi, \psi \in \mathfrak{gl}(\mathbf{M})$. We obtain a subalgebra by setting

$$\mathbf{L}(A) := \langle \mathbf{e}_i, \mathbf{f}_i \mid i \in I \rangle_{\text{alg}} \subseteq \mathfrak{gl}(\mathbf{M}).$$

Since $\dim \mathfrak{gl}(\mathbf{M}) < \infty$, it is clear that $\dim \mathbf{L}(A) < \infty$. Our aim is to show that $\mathbf{L}(A)$ is of Cartan–Killing type, with A as structure matrix.

Lemma 3.3.9 (Cf. [**14**, §3]). *In the setting of Definition 3.3.8, let us also define* $\mathbf{h}_i := [\mathbf{e}_i, \mathbf{f}_i] \in \mathfrak{gl}(\mathbf{M})$ *for* $i \in I$. *Then the linear maps* $\mathbf{e}_i$, $\mathbf{f}_i$, $\mathbf{h}_i \in \mathfrak{gl}(\mathbf{M})$ *satisfy the Chevalley relations* (Ch1), (Ch2)*:*

$$[\mathbf{e}_i, \mathbf{f}_j] = 0 \quad \text{for all } i, j \in I \text{ such that } i \neq j;$$

$$[\mathbf{h}_i, \mathbf{h}_j] = 0, \quad [\mathbf{h}_i, \mathbf{e}_j] = a_{ij} \mathbf{e}_j, \quad [\mathbf{h}_i, \mathbf{f}_j] = -a_{ij} \mathbf{f}_j \quad \text{for all } i, j \in I.$$

Proof. Assume first that A arises as the structure matrix of a Lie algebra L of Cartan–Killing type with respect to an abelian subalgebra $H \subseteq L$ and a subset $\Delta = \{\alpha_i \mid i \in I\} \subseteq H^*$. Thus, $A = (a_{ij})_{i,j \in I}$, where $a_{ij} = \alpha_j(h_i)$ and $h_i \in H$ is defined by Proposition 2.2.5. We already discussed at the beginning of this section that then (Ch0), (Ch1), (Ch2) hold for $\{e_i, h_i, f_i \mid i \in I\} \subseteq L$, where e_i, f_i are Chevalley generators as in Remark 2.2.9. Since $\text{ad}_L\colon L \to \mathfrak{gl}(L)$ is a homomorphism of Lie algebras, it follows that (Ch1), (Ch2) also hold for the maps $\text{ad}_L(e_i), \text{ad}_L(f_i), \text{ad}_L(h_i) \in \mathfrak{gl}(L)$. Now let $\{\mathbf{e}_\alpha^+ \mid \alpha \in \Phi\}$ be a collection of elements as in Lusztig's Theorem 2.7.2. We consider the vector space $\mathbf{M} := L$ and set

$$u_i := [e_i, \mathbf{e}_{-\alpha_i}^+] = [f_i, \mathbf{e}_{\alpha_i}^+] \quad (i \in I), \qquad v_\alpha := \mathbf{e}_\alpha^+ \quad (\alpha \in \Phi).$$

Then the above formulae defining $\mathbf{e}_i\colon \mathbf{M} \to \mathbf{M}$ and $\mathbf{f}_i\colon \mathbf{M} \to \mathbf{M}$ correspond exactly to the formulae in Remark 2.7.4; in other words, we have $\mathbf{e}_i = \mathrm{ad}_L(e_i)$ and $\mathbf{f}_i = \mathrm{ad}_L(f_i)$ for all $i \in I$. Hence, (Ch1), (Ch2) also hold for $\mathbf{e}_i, \mathbf{f}_i, \mathbf{h}_i \in \mathfrak{gl}(\mathbf{M})$.

This argument works for A of type A_n, B_n, C_n or D_n, using the fact, already mentioned, that then A arises as the structure matrix of $L = \mathfrak{sl}_n(\mathbb{C})$ or $L = \mathfrak{go}_n(\mathbb{C})$ (for suitable Q_n). It remains to consider A of type G_2, F_4, E_6, E_7 or E_8. In these cases, we use again a *computer algebra approach*: we simply write down the matrices of all the $\mathbf{e}_i$ and $\mathbf{f}_i$ with respect to the above basis $\mathbf{B}$ of $\mathbf{M}$, and explicitly verify (Ch1), (Ch2) using a computer. Note that this is a finite computation since there are only five matrices A to consider and, in each case, there are $4|I|^2 - |I|$ relations to verify; see Section 3.4 for further details and examples. — Readers who are not happy with this argument may consult [**14**, §3], where a purely theoretical argument is presented. □

Let $\mathbf{L}(A) = \langle \mathbf{e}_i, \mathbf{f}_i \mid i \in I\rangle_{\mathrm{alg}} \subseteq \mathfrak{gl}(\mathbf{M})$ be as in Definition 3.3.8 and set $\mathbf{h}_i := [\mathbf{e}_i, \mathbf{f}_i]$ for $i \in I$. By Lemma 3.3.9, the Chevalley relations (Ch1), (Ch2) hold. Let $H = \langle \mathbf{h}_i \mid i \in I\rangle_{\mathbb{C}} \subseteq \mathbf{L}(A)$; then H is an abelian subalgebra. For each $j \in I$ we define $\dot{\alpha}_j \in H^*$ as in Lemma 3.3.1, that is, $\dot{\alpha}_j(h_i) := a_{ij}$ for $i \in I$. (We write $\dot{\alpha}_j$ in order to have a notation that is separate from $\alpha_j \in \Phi = \Phi(A)$.) More generally, if $\alpha \in \Phi$, we write $\alpha = \sum_{i\in I} n_i\alpha_i$ with $n_i \in \mathbb{Z}$ and set $\dot{\alpha} := \sum_{i\in I} n_i\dot{\alpha}_i$. Thus, we obtain a subset $\dot{\Phi} := \{\dot{\alpha} \mid \alpha \in \Phi\} \subseteq H^*$.

Theorem 3.3.10 (Existence Theorem). *With the above notation, the Lie algebra $\mathbf{L}(A) \subseteq \mathfrak{gl}(\mathbf{M})$ is of Cartan–Killing type with respect to $H \subseteq \mathbf{L}(A)$ and $\dot{\Delta} = \{\dot{\alpha}_j \mid j \in I\} \subseteq H^*$, such that A is the corresponding structure matrix and $\dot{\Phi}$ is the set of roots with respect to H. In particular, $\dim \mathbf{L}(A) = |I| + |\Phi|$; furthermore, since A is indecomposable, $\mathbf{L}(A)$ is a simple Lie algebra (see Theorem 2.4.14).*

Proof. We noted in Definition 3.3.8 that $\mathbf{e}_i \neq 0$ and $\mathbf{f}_i \neq 0$ for all $i \in I$; furthermore, $\dim \mathbf{L}(A) < \infty$. Since $\mathbf{h}_i = [\mathbf{e}_i, \mathbf{f}_i] \in \mathbf{L}(A)$, it is clear that (Ch0) holds. We already noted that (Ch1), (Ch2) hold. Since A is of type (FIN), we have $\det(A) \neq 0$; furthermore, $a_{ii} = 2$ for $i \in I$. Hence, all the assumptions of Proposition 3.3.7 are satisfied and so $(\mathbf{L}(A), H)$ is of Cartan–Killing type with respect to $\dot{\Delta} = \{\dot{\alpha}_j \mid j \in I\}$ and with structure matrix A. The fact that $\dot{\Phi}$ is the set of roots with respect to H follows from Remark 2.3.7. □

3.4. Using computers: the ChevLie package

Let $A = (a_{ij})_{i,j \in I}$ be a generalised Cartan matrix with $|W(A)| < \infty$. In this section, we explain how one can systematically deal with the various constructions arising from A in an algorithmic fashion, and effectively using a computer. Various general purpose computer algebra systems contain built-in functions for dealing with root systems, Weyl groups, Lie algebras, and so on; see the online menue of GAP [**12**], for example. We introduce the basic features of the Julia [**20**] package ChevLie [**15**], which builds on the design and the conventions of the older GAP package CHEVIE. These packages are freely available and particularly well suited to the topics discussed here.

Suppose you have installed Julia on your computer and downloaded the file `chevlie1r1.jl`; then start Julia and load ChevLie into your current Julia session:

```
julia> include("chevlie1r1.jl"); using .ChevLie
```

The central function in ChevLie is the Julia constructor `LieAlg`, with holds various fields with information about a Lie algebra of a given type (a Julia symbol like `:g`) and rank (a positive integer). Let us go through an example and add further explanations as we go along (or just type `?LieAlg` for further details and examples).

```
julia> l=LieAlg(:g,2)        # Lie algebra of type G_2
#I dim = 14
LieAlg('G2')
```

In the background, the following happens. When `LieAlg` is invoked, then a few functions are applied in order to compute some basic data related to the generalised Cartan matrix A with the given type and rank, where the labelling in Table 4 is used. (If you wish to use a different labelling, then follow the instructions in the online help of `LieAlg`.) A version of the `rootsystem` program (Table 1) yields the root system Φ. This is stored in the component `roots` of `LieAlg`; the Cartan matrix A and the number $N := |\Phi^+|$ are also stored:

```
julia> l.N
6
julia> l.cartan
```

```
  2  -1
 -3   2
julia> l.roots
 [1, 0]  [0, 1]  [1, 1]  [1, 2]  [1, 3]  [2, 3]
 [-1, 0] [0, -1] [-1, -1] [-1, -2] [-1, -3] [-2, -3]
```

The roots are stored in terms of the list of tuples

$$\mathscr{C}(A) = \big\{(n_i)_{i\in I} \in \mathbb{Z}^I \bigm| \sum_{i\in I} n_i\alpha_i \in \Phi\big\} \subseteq \mathbb{Z}^I,$$

exactly as in Remark 2.3.7. This yields an explicit enumeration of the $2N$ elements of Φ as follows:

$$\underbrace{\beta_1,\ldots,\beta_{|I|}}_{\text{simple roots}},\ \underbrace{\beta_{|I|+1},\ldots,\beta_N}_{\text{further positive roots}},\underbrace{-\beta_1,\ldots,-\beta_{|I|},-\beta_{|I|+1},\ldots,-\beta_N}_{\text{negative roots}},$$

where the simple roots are those of height 1, followed by the remaining positive roots ordered by increasing height, followed by the negative roots. In particular, if A is indecomposable, then `l.roots[l.N]` is the unique highest root (see Proposition 2.4.17). Once all roots are available, the permutations induced by the generators $s_i \in W$ $(i \in I)$ of the Weyl group are computed (as explained in Remark 3.2.9) and stored. In our example:

```
julia> l.perms
 (7, 3, 2, 4, 6, 5, 1, 9, 8, 10, 12, 11)
 (5, 8, 4, 3, 1, 6, 11, 2, 10, 9, 7, 12)
```

Here, the permutation induced by any $w \in W$ is specified by the tuple of integers $(j_1,\ldots,j_{2N})$ such that $w(\beta_{j_l}) = \beta_l$ for $1 \leqslant l \leqslant 2N$. (We use that convention, and not $w(\beta_l) = \beta_{j_l}$, in order to maintain consistency with GAP and CHEVIE, where permutations act from the right; for a generator s_i, both conventions yield the same tuple, because s_i has order 2.) Working with the permutations induced by W on Φ immediately yields a test for equality of two elements (which would otherwise be difficult by working with words in the generators). Multiplication inside W is extremely efficient: if we also have an element $w' \in W$ represented by $(j_1',\ldots,j_{2N}')$, then the product $w\cdot w' \in W$ is represented by $(j'_{j_1},\ldots,j'_{j_{2N}})$. Thus, in our example, the permutation induced by the element $w = s_2 \cdot s_1 \in W$ is obtained as follows.

```
julia> p1=l.perms[1]; p2=l.perms[2];
julia> ([p1[i] for i in p2]...,)   # create a tuple
(6, 9, 4, 2, 7, 5, 12, 3, 10, 8, 1, 11)
```

We will see later how a permutation can be converted back into a word in the generators of W.

Table 8. Constructing G_2 using Julia and ChevLie

```
julia> l=LieAlg(:g,2)
julia> mats=[l.e_i[1],l.e_i[2],l.f_i[1],l.f_i[2]];
julia> [Array(m) for m in mats]
[...]
# written out as 14 x 14 - matrices
# e_1:           e_2:           f_1:           f_2:
# 01000000000000 00000000000000 00000000000000 00000000000000
# 00000000000000 00300000000000 10000000000000 00000000000000
# 00000000000000 00020000000000 00000000000000 01000000000000
# 00001000000000 00000100000000 00000000000000 00200000000000
# 00000000000000 00000012000000 00010000000000 00000000000000
# 00000023000000 00000000000000 00000000000000 00030000000000
# 00000000100000 00000000000000 00000100000000 00000000000000
# 00000000000000 00000000010000 00000000000000 00001000000000
# 00000000000000 00000000003000 00000023000000 00000000000000
# 00000000001000 00000000000000 00000000000000 00000012000000
# 00000000000000 00000000000200 00000000010000 00000000100000
# 00000000000000 00000000000010 00000000000000 00000000002000
# 00000000000001 00000000000000 00000000000000 00000000000300
# 00000000000000 00000000000000 00000000000010 00000000000000
julia> checkrels(l,l.e_i,l.f_i,l.h_i)
Relations OK
true                                   # Chevalley relations OK
```

Once Φ is available, it is then an almost trivial matter to set up the matrices of the linear maps $\mathbf{e}_i : \mathbf{M} \to \mathbf{M}$ and $\mathbf{f}_i : \mathbf{M} \to \mathbf{M}$ with respect to the basis $\mathbf{B}$ in Definition 3.3.8. These are contained in the components `l.e_i` and `l.f_i`; there is also a component `l.h_i` containing the matrices of $\mathbf{h}_i = [\mathbf{e}_i, \mathbf{f}_i]$ for $i \in I$. In our example, these matrices are printed in Table 8. Here, the following conventions are used.

- The basis $\mathbf{B}$ is always ordered as follows:

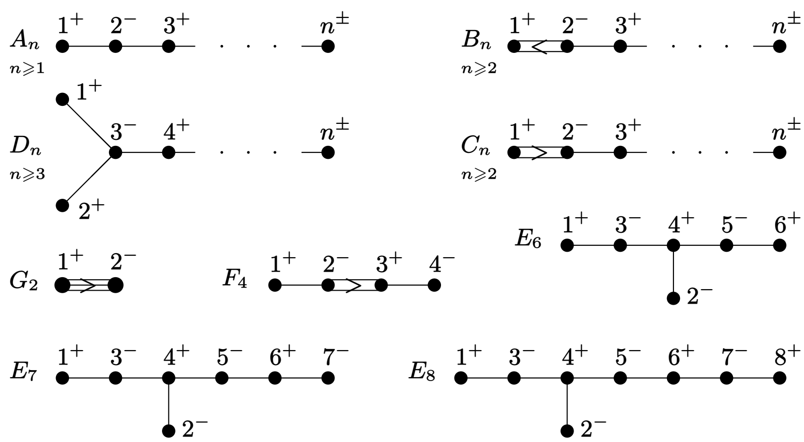
$$v_{\beta_N},\ \ldots,\ v_{\beta_1},\ u_1,\ \ldots,\ u_l,\ v_{-\beta_1},\ \ldots,\ v_{-\beta_N},$$

  where $I = \{1, \ldots, l\}$. Thus, each $\mathbf{e}_i$ is upper triangular and each $\mathbf{f}_i$ is lower triangular; each $\mathbf{h}_i$ is a diagonal matrix.

- Since the matrices representating $\mathbf{e}_i$, $\mathbf{f}_i$, $\mathbf{h}_i$ are extremely sparse, they are stored as Julia `SparseArrays`. In order to see them in full, one has to apply the Julia function `Array`.

Given the matrices of $\mathbf{e}_i$, $\mathbf{f}_i$, $\mathbf{h}_i$ for all $i \in I$, one can then check if the Chevalley relations (Ch1), (Ch2) hold; this is done by the function `checkrels`. — We rely on these programs in the proof of Lemma 3.3.9 for Lie algebras of type G_2, F_4, E_6, E_7 and E_8. (Even for type E_8, this just takes a few milliseconds.) ab hier Woche 11

Table 9. Dynkin diagrams with ϵ-function



Remark 3.4.1. Let $\{\mathbf{e}_\alpha^+ \mid \alpha \in \Phi\}$ be as in Corollary 2.7.11. We have $\mathbf{e}_{\alpha_i}^+ = c_i e_i$ and $\mathbf{e}_{-\alpha_i}^+ = -c_i f_i$ for all $i \in I$, where $c_i \in \{\pm 1\}$. Let us define $\epsilon\colon I \to \{\pm 1\}$ by $\epsilon(i) := c_i$ for $i \in I$. Then the argument in Remark 2.7.3 shows that $\epsilon(j) = -\epsilon(i)$ whenever $i, j \in I$ are such that $a_{ij} < 0$; furthermore, since A is indecomposable, there are precisely two such functions: if ϵ is one of them, then the other one is $-\epsilon$. In Table 9, we have specified a particular ϵ for each type of A. This is contained in the component `epsilon` of `LieAlg`:

```
julia> l.epsilon
  1 -1
```

Once ϵ and the elements $\mathbf{e}_{\pm\alpha_i}^+$ are fixed for $i \in I$, the whole collection of elements $\{\mathbf{e}_\alpha^+ \mid \alpha \in \Phi\}$ is uniquely determined by the conditions (L1), (L2), (L3) in Lusztig's Theorem 2.7.2 (see Remark 2.7.5). We call $\{\mathbf{e}_\alpha^+ \mid \alpha \in \Phi\}$ the *ϵ-canonical Chevalley system* of L. We shall also write $\mathbf{e}_\alpha^\epsilon = \mathbf{e}_\alpha^+$ in order to indicate the dependence on ϵ; note that, if we replace ϵ by $-\epsilon$, then $\mathbf{e}_\alpha^{-\epsilon} = -\mathbf{e}_\alpha^\epsilon$ for all $\alpha \in \Phi$.

The matrices of all $\mathbf{e}_\alpha^\epsilon$ ($\alpha \in \Phi$) are obtained using the function `canchevbasis`. For example, for type E_8, the matrices have size 248×248 but they are extremely sparse; so neither computer memory nor computing time is an issue here. (In ChevLie, they are stored as `SparseArrays`, with signed 8-bit integers as entries.) Once those matrices are available, the function `structconst` computes the corresponding structure constants $N_{\alpha,\beta}^\epsilon$ such that

$$[\mathbf{e}_\alpha^\epsilon, \mathbf{e}_\beta^\epsilon] = N_{\alpha,\beta}^\epsilon \mathbf{e}_{\alpha+\beta}^\epsilon \quad \text{for } \alpha, \beta, \alpha+\beta \in \Phi.$$

(Again, this is very efficient since one only needs to identify one nonzero entry in the matrix of $\mathbf{e}_{\alpha+\beta}^\epsilon$ and then work out only that entry in the matrix of the Lie bracket $[\mathbf{e}_\alpha^\epsilon, \mathbf{e}_\beta^\epsilon]$.) In our above example, we have:

```
julia> l.roots
 [1, 0]   [0, 1]   [1, 1]   [1, 2]   [1, 3]   [2, 3]
 [-1, 0]  [0, -1]  [-1, -1] [-1, -2] [-1, -3] [-2, -3]
julia> structconst(l,2,4)
(2, 4, -3, 5)
julia> structconst(l,1,3)
(1, 3, 0, 0)
```

Here, $(2,4,-3,5)$ means that `l.roots[2]+l.roots[4]=l.roots[5]` is a root and that $N_{\alpha,\beta}^\epsilon = -3$; the output $(1,3,0,0)$ means that `l.roots[1]+l.roots[3]` is not a root (and, hence, $N_{\alpha,\beta}^\epsilon = 0$).

Finally, we briefly discuss how one can work efficiently with the elements of the Weyl group W. Recall that $W = \langle s_i \mid i \in I\rangle$ and that $s_i^2 = \mathrm{id}$ for all $i \in I$. Thus, every element of W can be written as a product of various s_i (but inverses of the s_i are not required).

Similarly to the height of roots, the length function on W is a crucial tool for inductive arguments.

Definition 3.4.2. Let $w \in W$. We define the *length* of w, denoted $\ell(w)$, as follows. We set $\ell(\mathrm{id}) := 0$. Now let $w \in W$, $w \neq \mathrm{id}$. Then

$$\ell(w) := \min\{r \geqslant 0 \mid w = s_{i_1} \cdots s_{i_r} \text{ for some } i_1, \ldots, i_r \in I\}.$$

In particular, $\ell(s_i) = 1$ for all $i \in I$. If $r = \ell(w)$ and $i_1, \ldots, i_r \in I$ are such that $w = s_{i_1} \cdots s_{i_r}$, then we call this a *reduced expression* for w. In general, there may be several reduced expressions for w.

Remark 3.4.3. The formula in Remark 3.2.2 shows that each $s_i \in W$ ($i \in I$) is a reflection and so $\det(s_i) = -1$. Hence, we obtain

$$\det(w) = (-1)^{\ell(w)} \qquad \text{for any } w \in W.$$

Now let $w \neq \mathrm{id}$ and $w = s_{i_1} \cdots s_{i_r}$ be a reduced expression for w, where $r = \ell(w)$ and $i_1, \ldots, i_r \in I$. Since $s_i^{-1} = s_i$ for all $i \in I$, we have $w^{-1} = s_{i_r} \cdots s_{i_1}$ and so $\ell(w^{-1}) \leqslant \ell(w)$. But then also $\ell(w) = \ell((w^{-1})^{-1}) \leqslant \ell(w^{-1})$ and so $\ell(w) = \ell(w^{-1})$.

Now let $i \in I$. Then, clearly, $\ell(ws_i) \leqslant \ell(w) + 1$. Setting $w' := ws_i \in W$, we also have $w = w's_i$ and so $\ell(w) = \ell(w's_i) \leqslant \ell(w') + 1 = \ell(ws_i) + 1$. Hence, $\ell(ws_i) \geqslant \ell(w) - 1$. But, since $\det(w) = (-1)^{\ell(w)}$, we can not have $\ell(ws_i) = \ell(w)$. So we always have

$$\ell(ws_i) = \ell(w) \pm 1 \qquad \text{and} \qquad \ell(s_i w) = \ell(w) \pm 1,$$

where the second relation follows from the first by taking inverses.

Lemma 3.4.4. *Let $i \in I$ and $w \in W$. Then $\ell(ws_i) = \ell(w) + 1$ if and only if $w(\alpha_i) \in \Phi^+$. Similarly, $\ell(ws_i) = \ell(w) - 1$ if and only if $w(\alpha_i) \in \Phi^-$.*

Proof. First we show the implication: $\ell(ws_i) \geqslant \ell(w) \Rightarrow w(\alpha_i) \in \Phi^+$. This is seen as follows. Let $r := \ell(w) \geqslant 0$. If $r = 0$, then $w = \mathrm{id}$ and the assertion is clear. Now let $r \geqslant 1$ and write $w = s_{i_r} \cdots s_{i_1}$, where $i_1, \ldots, i_r \in I$. Consider the following sequence of $r + 1$ roots:

$$\alpha_i, \quad s_{i_1}(\alpha_i), \quad s_{i_2} s_{i_1}(\alpha_i), \quad \ldots, \quad s_{i_r} \cdots s_{i_1}(\alpha_i).$$

Denote them by $\beta_0, \beta_1, \ldots, \beta_r$ (from left to right). Since $\beta_0 = \alpha_i \in \Phi^+$ and $\beta_r = w(\alpha_i) \in \Phi^-$, there is some $j \in \{1, 2, \ldots, r\}$ such that

$\beta_0, \beta_1, \ldots, \beta_{j-1} \in \Phi^+$ but $\beta_j \in \Phi^-$. Now $\beta_j = s_{i_j}(\beta_{j-1})$ and so

$$\beta_j = \beta_{j-1} - m\alpha_{i_j} \in \Phi^- \quad \text{where} \quad m := \langle \alpha_{i_j}^\vee, \beta_{j-1} \rangle \in \mathbb{Z}.$$

Since $\beta_{j-1} \in \Phi^+$, this can only happen if $\beta_{j-1} = \alpha_{i_j}$; see Lemma 2.2.8. Hence, we have

$$\alpha_{i_j} = \beta_{j-1} = y(\alpha_i) \quad \text{where} \quad y := s_{i_{j-1}} s_{i_{j-2}} \cdots s_{i_1} \in W.$$

This implies that $ys_iy^{-1} = s_{i_j}$. Indeed, let $v \in E$ and write $v' := y^{-1}(v) \in E$. Using the W-invariance of $\langle\ ,\ \rangle$, we obtain

$$\langle \alpha_i^\vee, v' \rangle = 2\frac{\langle \alpha_i, v' \rangle}{\langle \alpha_i, \alpha_i \rangle} = 2\frac{\langle y(\alpha_i), y(v') \rangle}{\langle y(\alpha_i), y(\alpha_i) \rangle} = 2\frac{\langle \alpha_{i_j}, v \rangle}{\langle \alpha_{i_j}, \alpha_{i_j} \rangle} = \langle \alpha_{i_j}^\vee, v \rangle$$

and so $(ys_iy^{-1})(v) = y\big(s_i(v')\big) = y\big(v' - \langle \alpha_i^\vee, v' \rangle \alpha_i\big) = v - \langle \alpha_i^\vee, v' \rangle \alpha_{i_j} = s_{i_j}(v)$, as claimed. But this means that

$$s_{i_j} s_{i_{j-1}} \cdots s_{i_1} = s_{i_j} y = y s_i = s_{i_{j-1}} s_{i_{j-2}} \cdots s_{i_1} s_i.$$

Inserting this into the given expression for w, we obtain

$$w = (s_{i_r} \cdots s_{i_{j+1}})(s_{i_j} \cdots s_{i_1}) = (s_{i_r} \cdots s_{i_{j+1}})(s_{i_{j-1}} \cdots s_{i_1}) s_i.$$

But then $ws_i = (s_{i_r} \cdots s_{i_{j+1}})(s_{i_{j-1}} \cdots s_{i_1})$ is a product with $r-1$ factors, contradiction to the assumption that $\ell(ws_i) \geqslant \ell(w) = r$.

Thus, the above implication is proved. Conversely, let $w(\alpha_i) \in \Phi^+$ and assume, if possible, that $\ell(ws_i) \leqslant \ell(w)$. Setting $w' := ws_i$, we have $w'(\alpha_i) = w(s_i(\alpha_i)) = -w(\alpha_i) \in \Phi^-$. Hence, we must have $\ell(w's_i) < \ell(w')$. Since $w = w's_i$, this implies $\ell(w) < \ell(ws_i)$, contradiction. Hence, we must have $\ell(ws_i) > \ell(w)$.

Finally, let $w' := ws_i$. Then $\ell(w's_i) = \ell(w)$ and $w'(\alpha_i) = ws_i(\alpha) = -w(\alpha_i)$. Consequently, having established the equivalence $\ell(w's_i) = \ell(w') + 1 \Leftrightarrow w'(\alpha_i) \in \Phi^+$, we also obtain the equivalence $\ell(ws_i) = \ell(w) - 1 \Leftrightarrow w(\alpha_i) \in \Phi^+$. □

Corollary 3.4.5. *Let $w \in W$, $w \neq \mathrm{id}$. Then there exists some $i \in I$ such that $w(\alpha_i) \in \Phi^-$ and we can write $w = w's_i$, where $w' \in W$ is such that $\ell(w') = \ell(w) - 1$.*

Proof. Let $r := \ell(w) \geqslant 1$ and write $w = s_{i_1} \cdots s_{i_r}$, where $i_1, \ldots, i_r \in I$. Set $i := i_r$ and $w' := ws_i = ws_{i_r} = s_{i_1} \cdots s_{i_{r-1}} \in W$; then $w = w's_i$. We have $\ell(ws_i) = \ell(w') \leqslant r-1 < \ell(w)$ and so $\ell(ws_i) = \ell(w) - 1$. Furthermore, Lemma 3.4.4 implies that $w(\alpha_i) \in \Phi^-$. □

Remark 3.4.6. We now obtain an efficient algorithm for computing a reduced expression of an element $w \in W$, given as a permutation on the roots as above. Let $(j_1, \ldots, j_{2N})$ be the tuple representing that permutation. If $j_l = l$ for $1 \leqslant l \leqslant 2N$, then $w = \mathrm{id}$. Otherwise, by Corollary 3.4.5, there exists some $i \in I$ such that $w^{-1}(\alpha_i) \in \Phi^-$. Using the above conventions about the tuple $(j_1, \ldots, j_{2N})$, this means that there is some $i \in \{1, \ldots, |I|\}$ such that $j_i > N$. In order to make a definite choice, we take the smallest $i \in \{1, \ldots, |I|\}$ such that $j_i > N$. Then $\ell(s_i w) = \ell(w^{-1} s_i) = \ell(w) - 1$ and we can proceed with $w' := s_i w$. In ChevLie, this is implemented in the function `permword`.

```
julia> l=LieAlg(:g,2)
julia> permword(l,(6,9,4,2,7,5,12,3,10,8,1,11))
2-element Array{Int8,1}:
2 1
```

Conversion from a word (reduced or not), like $[2, 1, 2, 1]$, to a permutation is done by the function `wordperm`. Corollary 3.4.5 also shows how to produce all elements of W systematically, up to a given length. Indeed, if $W(n)$ denotes the set of all $w \in W$ such that $\ell(w) = n$, then the set of all elements of length $n + 1$ is obtained by taking the set of all products ws_i, where $w \in W(n)$ and $i \in I$ are such that $\ell(ws_i) = \ell(w) + 1$. This procedure is implemented in the function `allwords`. In our above example:

```
julia> allwords(l,3)    # elements up to length 3
#I 1 2 2 2
[] [1] [2] [1, 2] [2, 1] [1, 2, 1] [2, 1, 2]
```

(All elements are obtained by `allwords(l)`.)

3.5. Introducing Chevalley groups

Let again L be a Lie algebra and $H \subseteq L$ be an abelian subalgebra such that (L, H) is of Cartan–Killing type with respect to a subset $\Delta = \{\alpha_i \mid i \in I\} \subseteq H^*$. For each $i \in I$ let $\{e_i, h_i, f_i \mid i \in I\}$ be a corresponding $\mathfrak{sl}_2$-triple in L, as in Remark 2.2.9. Already in

Section 2.4 we introduced the automorphisms

$$x_i(t) := \exp\big(t\,\mathrm{ad}_L(e_i)\big) \in \mathrm{Aut}(L) \quad \text{for all } t \in \mathbb{C},$$
$$y_i(t) := \exp\big(t\,\mathrm{ad}_L(f_i)\big) \in \mathrm{Aut}(L) \quad \text{for all } t \in \mathbb{C}.$$

Hence, we can form the subgroup $\langle x_i(t), y_i(t) \mid i \in I, t \in \mathbb{C}\rangle \subseteq \mathrm{Aut}(L)$. In Definition 3.5.5 below we will see that one can define a similar group over *any* field instead of $\mathbb{C}$.

We will assume that A indecomposable. Then we have Lusztig's canonical basis $\mathbf{B}$ of L; see Section 2.7. We also assume that the additional conditions in Corollary 2.7.11 hold. Thus, there is a certain function $\epsilon\colon I \to \{\pm 1\}$ such that

$$\mathbf{e}^+_{\alpha_i} = \epsilon(i)e_i, \quad \mathbf{e}^+_{-\alpha_i} = -\epsilon(i)f_i, \quad h^+_j = -\epsilon(i)h_i \quad \text{for } i \in I;$$

see Remark 3.4.1. A specific choice of ϵ is defined by Table 9 (p. 127). Note that the formulae in the following theorem are independent of those choices.

Theorem 3.5.1 (Lusztig [**24**, §2]). *For $i \in I$ and $t \in \mathbb{C}$, the action of $x_i(t)$ and of $y_i(t)$ on $\mathbf{B}$ is given by the following formulae.*

$$x_i(t)(h^+_j) = h^+_j + |a_{ji}|t\mathbf{e}^+_{\alpha_i}, \qquad x_i(t)(\mathbf{e}^+_{-\alpha_i}) = \mathbf{e}^+_{-\alpha_i} + th^+_i + t^2\mathbf{e}^+_{\alpha_i},$$
$$x_i(t)(\mathbf{e}^+_{\alpha_i}) = \mathbf{e}^+_{\alpha_i}, \qquad x_i(t)(\mathbf{e}^+_{\alpha}) = \sum_{0\leqslant r\leqslant p_{i,\alpha}} \binom{q_{i,\alpha}+r}{r} t^r \mathbf{e}^+_{\alpha+r\alpha_i},$$
$$y_i(t)(h^+_j) = h^+_j + |a_{ji}|t\mathbf{e}^+_{-\alpha_i}, \qquad y_i(t)(\mathbf{e}^+_{\alpha_i}) = \mathbf{e}^+_{\alpha_i} + th^+_i + t^2\mathbf{e}^+_{-\alpha_i},$$
$$y_i(t)(\mathbf{e}^+_{-\alpha_i}) = \mathbf{e}^+_{-\alpha_i}, \qquad y_i(t)(\mathbf{e}^+_{\alpha}) = \sum_{0\leqslant r\leqslant q_{i,\alpha}} \binom{p_{i,\alpha}+r}{r} t^r \mathbf{e}^+_{\alpha-r\alpha_i},$$

where $j \in I$ and $\alpha \in \Phi$, $\alpha \neq \pm\alpha_i$. Here, $p_{i,\alpha}, q_{i,\alpha}$ are the non-negative integers defining the α_i-string through α (see Remark 2.7.1).

Proof. In the proof of Lemma 2.4.1, we already established the following formulae, where $i \in I$, $t \in \mathbb{C}$ and $h \in H$:

(a) $$x_i(t)(h) = h - \alpha_i(h)te_i,$$

(b) $$y_i(t)(h) = h + \alpha_i(h)tf_i,$$

(c) $$x_i(t)(e_i) = e_i,$$

(d) $$y_i(t)(e_i) = e_i - th_i - t^2 f_i.$$

Now, since $h_j^+ = -\epsilon(j)h_j$, we obtain using (a) that

$$x_i(t)(h_j^+) = -\epsilon(j)h_j + \epsilon(j)\alpha_i(h_j)te_i = h_j^+ + \epsilon(j)a_{ji}te_i.$$

In Remark 2.7.4, we saw that $[e_i, h_j^+] = \epsilon(j)a_{ji}e_i = |a_{ji}|\mathbf{e}_{\alpha_i}^+$. This yields the desired formula for $x_i(t)(h_j^+)$. Similarly, using (b), we obtain the desired formula for $y_i(t)(h_j^+)$. The formula for $x_i(t)(\mathbf{e}_{\alpha_i}^+)$ immediately follows from (c). Analogously to (c), we have $y_i(t)(f_i) = f_i$ and this yields the formula for $y_i(t)(\mathbf{e}_{-\alpha_i}^+)$. Next, using (d), we obtain:

$$y_i(t)(\mathbf{e}_{\alpha_i}^+) = \epsilon(i)e_i - \epsilon(i)th_i - \epsilon(i)t^2 f_i = \mathbf{e}_{\alpha_i}^+ + th_i^+ + t^2\mathbf{e}_{-\alpha_i}^+,$$

as required. Analogously to (d), we have $x_i(t)(f_i) = f_i + th_i - t^2 e_i$ and this yields the formula for $x_i(t)(\mathbf{e}_{-\alpha_i}^+)$. It remains to prove the formulae for $x_i(t)(\mathbf{e}_\alpha^+)$ and $y_i(t)(\mathbf{e}_\alpha^+)$, where $\alpha \neq \pm\alpha_i$. We only do this here in detail for $x_i(t)(\mathbf{e}_\alpha^+)$; the argument for $y_i(t)(\mathbf{e}_\alpha^+)$ is completely analogous. Now, by definition, we have

$$x_i(t)(\mathbf{e}_\alpha^+) = \mathbf{e}_\alpha^+ + \sum_{r \geqslant 1} \frac{t^r \mathrm{ad}_L(e_i)^r(\mathbf{e}_\alpha^+)}{r!}.$$

Note that $\mathrm{ad}_L(e_i)^r(\mathbf{e}_\alpha^+) \in L_{\alpha + r\alpha_i} = \{0\}$ if $r > p_{i,\alpha}$. So now assume that $1 \leqslant r \leqslant p_{i,\alpha}$. Then $\alpha + \alpha_i \in \Phi$ and $\mathrm{ad}_L(e_i)(\mathbf{e}_\alpha^+) = [e_i, \mathbf{e}_\alpha^+] = (q_{i,\alpha} + 1)\mathbf{e}_{\alpha+\alpha_i}^+$; see (L2) in Theorem 2.7.2. Furthermore,

$$\mathrm{ad}_L(e_i)^2(\mathbf{e}_\alpha^+) = [e_i, [e_i, \mathbf{e}_\alpha^+]] = (q_{i,\alpha} + 1)[e_i, \mathbf{e}_{\alpha+\alpha_i}^+].$$

If $p_{i,\alpha} \geqslant 2$, then $\alpha + 2\alpha_i \in \Phi$ and so the right hand side equals $(q_{i,\alpha} + 1)(q_{i,\alpha+\alpha_i} + 1)\mathbf{e}_{\alpha+2\alpha_i}$, again by Theorem 2.7.2. Continuing in this way, we find that

$$\mathrm{ad}_L(e_i)^r(\mathbf{e}_\alpha^+) = (q_{i,\alpha} + 1)(q_{i,\alpha+\alpha_i} + 1) \cdots (q_{i,\alpha+(r-1)\alpha_i} + 1)\mathbf{e}_{\alpha+r\alpha_i}^+$$

for $1 \leqslant r \leqslant p_{i,\alpha}$. Now note that

$$q_{i,\alpha+\alpha_i} = \max\{m \geqslant 0 \mid \alpha + \alpha_i - m\alpha_i \in \Phi\} = q_{i,\alpha} + 1.$$

Similarly, $q_{i,\alpha+r\alpha_i} = q_{i,\alpha_i} + r$ for $1 \leqslant r \leqslant p_{i,\alpha}$. Hence, we obtain that

$$\begin{aligned}(q_{i,\alpha}&+1)(q_{i,\alpha+\alpha_i} + 1) \cdots (q_{i,\alpha+(r-1)\alpha_i} + 1) \\ &= (q_{i,\alpha} + 1)(q_{i,\alpha} + 2) \cdots (q_{i,\alpha} + r) = (q_{i,\alpha} + r)!/q_{i,\alpha}!\end{aligned}$$

Inserting this into the formula for $x_i(t)(\mathbf{e}_\alpha^+)$, we obtain

$$x_i(t)(\mathbf{e}_\alpha^+) = \sum_{r \geqslant 0} \frac{t^r \mathrm{ad}_L(e_i)^r(\mathbf{e}_\alpha^+)}{r!} = \sum_{0 \leqslant r \leqslant p_{i,\alpha}} \frac{(q_{i,\alpha}+r)!}{r!\, q_{i,\alpha}!} t^r \mathbf{e}_{\alpha+r\alpha_i}^+,$$

and it remains to use the formula for binomial coefficients. □

The above result shows that the actions of $x_i(t)$ and $y_i(t)$ on L are completely determined by the structure matrix A and the (abstract) root system $\Phi = \Phi(A)$. As pointed out by Lusztig [**24**], this seems to simplify the original setting of Chevalley [**9**], where a number of signs appear in the formulae which depend on certain choices.

Example 3.5.2. Let $i \in I$ and $\alpha \in \Phi$ be such that $\alpha \neq \pm\alpha_i$. If $\alpha + \alpha_i \notin \Phi$, then the above formulae show that $x_i(t)(\mathbf{e}_\alpha^+) = \mathbf{e}_\alpha^+$. Similarly, if $\alpha - \alpha_i \notin \Phi$, then $y_i(t)(\mathbf{e}_\alpha^+) = \mathbf{e}_\alpha^+$. Now assume that $\alpha + \alpha_i \in \Phi$ and that $p_{i,\alpha} = 1$. Then

$$x_i(t)(\mathbf{e}_\alpha^+) = \mathbf{e}_\alpha^+ + \binom{q_{i,\alpha}+1}{1} t\mathbf{e}_{\alpha+\alpha_i}^+ = \mathbf{e}_\alpha^+ + (q_{i,\alpha}+1)t\mathbf{e}_{\alpha+\alpha_i}^+.$$

Similarly, if $\alpha - \alpha_i \in \Phi$ and $q_{i,\alpha} = 1$, then

$$y_i(t)(\mathbf{e}_\alpha^+) = \mathbf{e}_\alpha^+ + \binom{p_{i,\alpha}+1}{1} t\mathbf{e}_{\alpha-\alpha_i}^+ = \mathbf{e}_\alpha^+ + (p_{i,\alpha}+1)t\mathbf{e}_{\alpha-\alpha_i}^+.$$

Note that these formulae cover all cases where A is of *simply-laced* type, that is, all roots in Φ have the same length; see Exercise 3.2.8. Recall from ($\spadesuit_3$) (p. 84) that, in general, we have $p_{i,\alpha} + q_{i,\alpha} \leqslant 3$.

Remark 3.5.3. Let $N = |\Phi^+|$ and write $\Phi^+ = \{\beta_1, \ldots, \beta_N\}$ where the numbering is such that $\mathrm{ht}(\beta_1) \leqslant \mathrm{ht}(\beta_2) \leqslant \ldots \leqslant \mathrm{ht}(\beta_N)$. Let also $l = |I|$ and simply write $I = \{1, \ldots, l\}$. Then, as in Section 3.4, we order the basis $\mathbf{B}$ as follows:

$$\mathbf{e}_{\beta_N}^+,\ \ldots,\ \mathbf{e}_{\beta_1}^+,\ h_1^+,\ \ldots,\ h_l^+,\ \mathbf{e}_{-\beta_1}^+,\ \ldots,\ \mathbf{e}_{-\beta_N}^+.$$

Let $N' := 2N + l = |\mathbf{B}|$ and denote the above basis elements by $v_1, \ldots, v_{N'}$, from left to right. For $i \in I$ and $t \in \mathbb{C}$, let $X_i(t) \in M_{N'}(\mathbb{C})$ be the matrix of $x_i(t)$ with respect to the basis $\{v_1, \ldots, v_{N'}\}$; also let $Y_i(t) \in M_{N'}(\mathbb{C})$ be the matrix of $y_i(t)$ with respect to that basis. Then the formulae in Theorem 3.5.1 show that

$X_i(t)$ is an upper triangular matrix with 1 along the diagonal,

$Y_i(t)$ is a lower triangular matrix with 1 along the diagonal.

In particular, we have $\det(x_i(t)) = \det(y_i(t)) = 1$. We also notice that each entry in $X_i(t)$ or $Y_i(t)$ is of the form at^r, where the coefficient $a \in \mathbb{Z}$ and the exponent $r \in \mathbb{Z}_{\geqslant 0}$ do not depend on $t \in \mathbb{C}$. Now let

$\mathbb{Z}[T]$ be the polynomial ring over $\mathbb{Z}$ in an indeterminate T. Replacing each entry of the form at^r by aT^r, we obtain matrices

$$X_i(T) \in M_{N'}(\mathbb{Z}[T]) \qquad \text{and} \qquad Y_i(T) \in M_{N'}(\mathbb{Z}[T]).$$

Upon substituting $T \mapsto t$ for any $t \in \mathbb{C}$, we get back the original matrices $X_i(t)$ and $Y_i(t)$. The possibility of working at a polynomial level will turn out to be crucial later on.

Example 3.5.4. Let $L = \mathfrak{sl}_2(\mathbb{C})$ with standard basis $\{e, h, f\}$, such that $[e, f] = h$, $[h, e] = 2e$ and $[h, f] = -2f$. In Exercise 1.2.15, we already considered the automorphisms

$$x(t) = \exp\bigl(t\,\mathrm{ad}_L(e)\bigr) \quad \text{and} \quad y(t) = \exp\bigl(t\,\mathrm{ad}_L(e)\bigr) \qquad (t \in \mathbb{C}),$$

and worked out the corresponding matrices. Now note that $\mathbf{B} = \{e, -h, -f\}$ (see the remark just after Theorem 2.7.2). So we obtain:

$$X(t) = \begin{pmatrix} 1 & 2t & t^2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad Y(t) = \begin{pmatrix} 1 & 0 & 0 \\ t & 1 & 0 \\ t^2 & 2t & 1 \end{pmatrix}.$$

Hence, obviously, we have the following matrices over $\mathbb{Z}[T]$:

$$X(T) = \begin{pmatrix} 1 & 2T & T^2 \\ 0 & 1 & T \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad Y(T) = \begin{pmatrix} 1 & 0 & 0 \\ T & 1 & 0 \\ T^2 & 2T & 1 \end{pmatrix}.$$

We now show how the definition of G can be extended to arbitrary fields. Let K be any field. We usually attach a bar to objects defined over K. So let $\bar{L}$ be a vector space[4] over K with a basis

$$\bar{\mathbf{B}} = \{\bar{h}_j^+ \mid j \in I\} \cup \{\bar{\mathbf{e}}_\alpha^+ \mid \alpha \in \Phi\}.$$

For $i \in I$ and $\zeta \in K$ we use the formulae in Theorem 3.5.1 to define linear maps $\bar{x}_i(\zeta) \colon \bar{L} \to \bar{L}$ and $\bar{y}_i(\zeta) \colon \bar{L} \to \bar{L}$. Explicitly, we set:

$$\bar{x}_i(\zeta)(\bar{h}_j^+) := \bar{h}_j^+ + |a_{ji}|\zeta\bar{\mathbf{e}}_{\alpha_i}^+, \quad \bar{x}_i(\zeta)(\bar{\mathbf{e}}_{-\alpha_i}^+) := \bar{\mathbf{e}}_{-\alpha_i}^+ + \zeta\bar{h}_i^+ + \zeta^2\bar{\mathbf{e}}_{\alpha_i}^+,$$

$$\bar{x}_i(\zeta)(\bar{\mathbf{e}}_{\alpha_i}^+) := \bar{\mathbf{e}}_{\alpha_i}^+, \qquad \bar{x}_i(\zeta)(\bar{\mathbf{e}}_\alpha^+) := \sum_{0 \leqslant r \leqslant p_{i,\alpha}} \binom{q_{i,\alpha}+r}{r} \zeta^r \bar{\mathbf{e}}_{\alpha+r\alpha_i}^+,$$

[4]This vector space $\bar{L}$ also inherits a Lie algebra structure from L but we will not need this here.

$$\bar{y}_i(\zeta)(\bar{h}_j^+) := \bar{h}_j^+ + |a_{ji}|\zeta\bar{\mathbf{e}}_{-\alpha_i}^+, \quad \bar{y}_i(\zeta)(\bar{\mathbf{e}}_{\alpha_i}^+) := \bar{\mathbf{e}}_{\alpha_i}^+ + \zeta\bar{h}_i^+ + \zeta^2\bar{\mathbf{e}}_{-\alpha_i}^+,$$

$$\bar{y}_i(\zeta)(\bar{\mathbf{e}}_{-\alpha_i}^+) := \bar{\mathbf{e}}_{-\alpha_i}^+, \qquad \bar{y}_i(\zeta)(\bar{\mathbf{e}}_{\alpha}^+) := \sum_{0\leqslant r\leqslant q_{i,\alpha}} \binom{p_{i,\alpha}+r}{r}\zeta^r\bar{\mathbf{e}}_{\alpha-r\alpha_i}^+,$$

where $j \in I$ and $\alpha \in \Phi$, $\alpha \neq \pm\alpha_i$. (Here, the product of an integer in $\mathbb{Z}$ and an element of K is defined in the obvious way.) Let $\bar{X}_i(\zeta)$ and $\bar{Y}_i(\zeta)$ be the matrices of $\bar{x}_i(\zeta)$ and $\bar{y}_i(\zeta)$, respectively, with respect to $\bar{\mathbf{B}}$, where the elements of $\bar{\mathbf{B}}$ are arranged as in Remark 3.5.3. Then the above formulae show again that

$\bar{X}_i(\zeta)$ is upper triangular with 1 along the diagonal,

$\bar{Y}_i(\zeta)$ is lower triangular with 1 along the diagonal.

In particular, we have $\det(\bar{x}_i(\zeta)) = \det(\bar{y}_i(\zeta)) = 1$. Note that, if $K = \mathbb{C}$, then $\bar{x}_i(\zeta) = x_i(\zeta)$ and $\bar{y}_i(\zeta) = y_i(\zeta)$ for all $\zeta \in \mathbb{C}$.

Definition 3.5.5. Following Lusztig [**24**, §2], the *Chevalley group*[5] of type L over the field K is defined by

$$\bar{G}' := \langle \bar{x}_i(\zeta), \bar{y}_i(\zeta) \mid i \in I, \zeta \in K\rangle \subseteq \mathrm{GL}(\bar{L}).$$

Note again that $\bar{G}'$ is completely determined by the structure matrix A, the (abstract) root system Φ, and the field K. Also note that, if K is a finite field, then $\bar{G}'$ is a finite group.

Example 3.5.6. Let $L = \mathfrak{sl}_2(\mathbb{C})$. In Example 3.5.4, we determined the matrices of $x(t)$ and $y(t)$ for $t \in \mathbb{C}$. Now let K be any field and $\zeta \in K$. Then the matrices of $\bar{x}(\zeta)$ and $\bar{y}(\zeta)$ are given by

$$\bar{X}(\zeta) = \begin{pmatrix} 1 & 2\zeta & \zeta^2 \\ 0 & 1 & \zeta \\ 0 & 0 & 1 \end{pmatrix} \quad\text{and}\quad \bar{Y}(\zeta) = \begin{pmatrix} 1 & 0 & 0 \\ \zeta & 1 & 0 \\ \zeta^2 & 2\zeta & 1 \end{pmatrix}.$$

In the next section we will see that $\bar{G}' = \langle \bar{x}(\zeta), \bar{y}(\zeta) \mid \zeta \in K\rangle$ is isomorphic to $\mathrm{SL}_2(K)/\{\pm I_2\}$ and, hence, that $\bar{G}'$ is simple when $|K| \geqslant 4$.

Remark 3.5.7. The definition immediately shows that $\bar{x}_i(0) = \mathrm{id}_{\bar{L}}$ and $\bar{y}_i(0) = \mathrm{id}_{\bar{L}}$. Now let $0 \neq \zeta \in K$. Then

$$\bar{x}_i(\zeta)(\bar{\mathbf{e}}_{-\alpha_i}^+) = \bar{\mathbf{e}}_{-\alpha_i}^+ + \zeta\bar{h}_i^+ + \zeta^2\bar{\mathbf{e}}_{\alpha_i}^+ \neq \bar{\mathbf{e}}_{-\alpha_i}^+$$

and so $\bar{x}_i(\zeta) \neq \mathrm{id}_{\bar{L}}$. Similarly, one sees that $\bar{y}_i(\zeta) \neq \mathrm{id}_{\bar{L}}$.

[5] We denote this group by $\bar{G}'$ because it is a normal subgroup of a slightly larger group $\bar{G}$ that we will introduce in the next section; the distinction between $\bar{G}'$ and the "full" Chevalley group $\bar{G}$ already appears in Chevalley [**9**].

Of course, one would hope that the elements $\bar{x}_i(\zeta)$ and $\bar{y}_i(\zeta)$ (over K) have further properties analogous to those of $x_i(t)$ and $y_i(t)$ (over $\mathbb{C}$). In order to justify this in concrete cases, some extra argument is usually required because the definition of $\bar{x}_i(\zeta)$ or $\bar{y}_i(\zeta)$ in terms of an exponential construction is not available over K (at least not if K has positive characteristic). For this purpose, we make crucial use of the possibility of working at a "polynomial level", as already mentioned in Remark 3.5.3. Here is a simple first example.

Lemma 3.5.8. *Let $i \in I$. Then $\bar{x}_i(\zeta)^{-1} = \bar{x}_i(-\zeta)$ and $\bar{y}_i(\zeta)^{-1} = \bar{y}_i(-\zeta)$ for all $\zeta \in K$. Furthermore, $\bar{x}_i(\zeta+\zeta') = \bar{x}_i(\zeta)\bar{x}_i(\zeta')$ and $\bar{y}_i(\zeta+\zeta') = \bar{y}_i(\zeta)\bar{y}_i(\zeta')$ for all $\zeta,\zeta' \in K$.*

Proof. First we prove the assertion about $\bar{x}_i(\zeta)^{-1}$. (This would also follow from the assertion about $\bar{x}_i(\zeta+\zeta')$ and the fact that $\bar{x}_i(0) = \mathrm{id}_{\bar{L}}$, but it may be useful to run the two arguments separately, since they involve different ingredients.) Let $\mathbb{Z}[T]$ be the polynomial ring over $\mathbb{Z}$ with indeterminate T. Let $X_i(T) \in M_{N'}(\mathbb{Z}[T])$ be the matrix defined in Remark 3.5.3; upon substituting $T \mapsto t$ for any $t \in \mathbb{C}$, we obtain the matrix of the element $x_i(t) \in G$ (over $\mathbb{C}$). We claim that

$$X_i(T) \cdot X_i(-T) = I_{N'} \qquad \text{(equality in } M_{N'}(\mathbb{Z}[T])\text{)},$$

where $I_{N'}$ denotes the $N' \times N'$-times identity matrix. This is seen as follows. Let $f_{rs} \in \mathbb{Z}[T]$ be the (r,s)-entry of $X_i(T)$. Writing out the matrix product $X_i(T) \cdot X_i(-T)$, we must show that the following identities of polynomials in $\mathbb{Z}[T]$ hold for all $r,s \in \{1,\ldots,N'\}$:

$$\textstyle\sum_{r'} f_{rr'}(T) f_{r's}(-T) = \begin{cases} 1 & \text{if } r = s, \\ 0 & \text{if } r \neq s. \end{cases}$$

Since $x_i(t)x_i(-t) = \mathrm{id}_L$ (see Lemma 1.2.8), we have $X_i(t) \cdot X_i(-t) = I_{N'}$ for all $t \in \mathbb{C}$, which means that

$$\textstyle\sum_{r'} f_{rr'}(t) f_{r's}(-t) = \begin{cases} 1 & \text{if } r = s, \\ 0 & \text{if } r \neq s. \end{cases}$$

So the assertion follows from the general fact that, if $g,h \in \mathbb{Z}[T]$ are such that $g(t) = h(t)$ for infinitely many $t \in \mathbb{C}$, then $g = h$ in $\mathbb{Z}[T]$.

Now fix $\zeta \in K$. By the universal property of $\mathbb{Z}[T]$, we have a canonical ring homomorphism $\varphi_\zeta \colon \mathbb{Z}[T] \to K$ such that $\varphi_\zeta(T) = \zeta$ and $\varphi_\zeta(m) = m \cdot 1_K$ for $m \in \mathbb{Z}$. Applying φ_ζ to the entries of $X_i(T)$, we obtain the matrix $\bar{X}_i(\zeta) \in M_{N'}(K)$, by the above definition of $\bar{x}_i(\zeta)$. Similarly, applying φ_ζ to the entries of $X_i(-T)$, we obtain the matrix $\bar{X}_i(-\zeta) \in M_{N'}(K)$. Since φ_ζ is a ring homomorphism, the identity $X_i(T) \cdot X_i(-T) = I_{N'}$ over $\mathbb{Z}[T]$ implies the identity $\bar{X}_i(\zeta) \cdot \bar{X}_i(-\zeta) = \bar{I}_{N'}$ over K. Consequently, we have $\bar{x}_i(\zeta)\bar{x}_i(-\zeta) = \mathrm{id}_{\bar{L}}$, as desired. The argument for $\bar{y}_i(\zeta)$ is completely analogous.

Now consider the assertion about $\bar{x}_i(\zeta+\zeta')$. First we work over $\mathbb{C}$. For $t, t' \in \mathbb{C}$, the derivations $t\,\mathrm{ad}_L(e_i)$ and $t'\,\mathrm{ad}_L(e_i)$ of L certainly commute with each other. Hence, Exercise 1.2.14 shows that

$$\begin{aligned} x_i(t+t') &= \exp\big(t\,\mathrm{ad}_L(e_i) + t'\,\mathrm{ad}_L(e_i)\big) \\ &= \exp\big(t\,\mathrm{ad}_L(e_i)\big) \circ \exp\big(t'\mathrm{ad}_L(e_i)\big) = x_i(t)x_i(t'), \end{aligned}$$

where we omit the symbol "$\circ$" for the multiplication inside G. Now we "lift" again the above identity to a polynomial level, where we work over $\mathbb{Z}[T, T']$, the polynomial ring in two commuting indeterminates T, T' over $\mathbb{Z}$. Regarding $X_i(T)$ and $X_i(T')$ as matrices in $M_{N'}(\mathbb{Z}[T, T'])$, we claim that

$$X_i(T+T') = X_i(T) \cdot X_i(T') \qquad (\text{equality in } M_{N'}(\mathbb{Z}[T, T'])).$$

This is seen as follows. Let again $f_{rs} \in \mathbb{Z}[T]$ be the (r, s)-entry of $X_i(T)$. Writing out the above matrix product, we must show that the following identities in $\mathbb{Z}[T, T']$ hold for all $r, s \in \{1, \ldots, N'\}$:

$$f_{rs}(T+T') = \textstyle\sum_{r'} f_{rr'}(T) f_{r's}(T').$$

We have just seen that these identities do hold upon substituting $T \mapsto t$ and $T' \mapsto t'$ for any $t, t' \in \mathbb{C}$. Hence, the assertion now follows from the general fact that, if $g, h \in \mathbb{Z}[T, T']$ are any polynomials such that $g(t, t') = h(t, t')$ for all $t, t' \in \mathbb{C}$, then $g = h$ in $\mathbb{Z}[T, T']$. (Proof left as an exercise; the analogous statement is also true for polynomials in several commuting variables.) Now fix $\zeta, \zeta' \in K$. Then we have a canonical ring homomorphism $\varphi_{\zeta,\zeta'} \colon \mathbb{Z}[T, T'] \to K$ such that $\varphi_{\zeta,\zeta'}(T) = \zeta$, $\varphi_{\zeta,\zeta'}(T') = \zeta'$ and $\varphi_{\zeta,\zeta'}(m) = m \cdot 1_K$ for $m \in \mathbb{Z}$. Applying $\varphi_{\zeta,\zeta'}$ to the entries of $X_i(T)$, $X_i(T')$ and $X_i(T+T')$, we obtain the matrices $\bar{X}_i(\zeta)$, $\bar{X}_i(\zeta')$ and $\bar{X}_i(\zeta+\zeta')$. Consequently, the identity $X_i(T+T') = X_i(T) \cdot X_i(T')$ over $\mathbb{Z}[T, T']$ implies the identity $\bar{X}_i(\zeta+\zeta') = \bar{X}_i(\zeta) \cdot \bar{X}_i(\zeta')$ over K. Hence, we have $\bar{x}_i(\zeta+\zeta') = \bar{x}_i(\zeta)\bar{x}_i(\zeta')$, as desired. The argument for $\bar{y}_i(\zeta+\zeta')$ is analogous. $\square$

ab hier Woche 12

We will see similar arguments, or variations thereof, frequently in the development to follow. The following result will be very useful.

Lemma 3.5.9. *Let $x \in L$ be such that $\mathrm{ad}_L(x)\colon L \to L$ is nilpotent. Let $\theta\colon L \to L$ be any Lie algebra automorphism. Then $\mathrm{ad}_L(\theta(x))$ is nilpotent and $\exp\big(\mathrm{ad}_L(\theta(x))\big) = \theta \circ \exp(\mathrm{ad}_L(x)) \circ \theta^{-1}$.*

Proof. Let $y \in L$. Since θ is an automorphism, we have for $m \geqslant 0$:

$$\begin{aligned}\mathrm{ad}_L(\theta(x))^m(y) &= [\underbrace{\theta(x), [\theta(x), \ldots, [\theta(x)}_{m \text{ terms}}, \theta\big(\theta^{-1}(y)\big)] \ldots]]\\ &= \theta\big([\underbrace{x, [x, \ldots, [x}_{m \text{ terms}}, \theta^{-1}(y)] \ldots]]\big) = \theta\big(\mathrm{ad}_L(x)^m(\theta^{-1}(y))\big).\end{aligned}$$

Hence, since $\mathrm{ad}_L(x)^d = 0$ for some $d \geqslant 1$, we also have $\mathrm{ad}_L(\theta(x))^d = 0$, that is, $\mathrm{ad}_L(\theta(x))$ is nilpotent. The above identity also yields:

$$\begin{aligned}\big(\theta \circ \exp(\mathrm{ad}_L(x)) \circ \theta^{-1}\big)(y) &= \theta\Big(\sum_{m \geqslant 0} \frac{1}{m!} \mathrm{ad}_L(x)^m\big(\theta^{-1}(y)\big)\Big)\\ &= \sum_{m \geqslant 0} \frac{1}{m!} \theta\big(\mathrm{ad}_L(x)^m(\theta^{-1}(y))\big) = \sum_{m \geqslant 0} \frac{1}{m!} \mathrm{ad}_L(\theta(x))^m(y),\end{aligned}$$

which equals $\exp\big(\mathrm{ad}_L(\theta(x))\big)(y)$, as required. $\square$

Example 3.5.10. Consider the Chevalley involution $\omega\colon L \to L$ (see Exercise Sheet 8); we have $\omega(e_i) = f_i$, $\omega(f_i) = e_i$ and $\omega(h_i) = -h_i$ for $i \in I$. Applying Lemma 3.5.9 with $\theta = \omega$, we obtain

$$\begin{aligned}\omega \circ x_i(t) \circ \omega^{-1} &= \omega \circ \exp\big(t\,\mathrm{ad}_L(e_i)\big) \circ \omega^{-1}\\ &= \exp\big(t\,\mathrm{ad}_L(\omega(e_i))\big) = \exp\big(t\,\mathrm{ad}_L(f_i)\big) = y_i(t)\end{aligned}$$

for all $t \in \mathbb{C}$. We wish to extend this formula to any field K. For this purpose, we first consider the action of ω on $\mathbf{B}$. Since $h_j^+ = -\epsilon(j)h_j$ for $j \in I$, we have $\omega(h_j^+) = -h_j^+$. By Proposition 2.7.13, we also have $\omega(\mathbf{e}_\alpha^+) = -\mathbf{e}_{-\alpha}^+$ for $\alpha \in \Phi$. We use these formulae to define a linear map $\bar{\omega}\colon \bar{L} \to \bar{L}$; explicitly, we set:

$$\bar{\omega}(\bar{h}_j^+) := -\bar{h}_j^+ \quad (j \in I) \quad \text{and} \quad \bar{\omega}(\bar{\mathbf{e}}_\alpha^+) := -\bar{\mathbf{e}}_{-\alpha}^+ \quad (\alpha \in \Phi).$$

With this definition, we claim that

$$\bar{\omega} \circ \bar{x}_i(\zeta) \circ \bar{\omega}^{-1} = \bar{y}_i(\zeta) \qquad \text{for all } \zeta \in K.$$

To prove this, we follow the argument in Lemma 3.5.8. Let $\Omega \in M_{N'}(\mathbb{C})$ be the matrix of ω with respect to $\mathbf{B}$. The above formulae show that Ω only has entries 0 and -1; we can simply regard Ω as a matrix in $M_{N'}(\mathbb{Z}[T])$. Then the above formula over $\mathbb{C}$ implies that

$$\Omega \cdot X_i(T) = Y_i(T) \cdot \Omega \qquad \text{(equality in } M_{N'}(\mathbb{Z}[T])\text{)}.$$

Let $\bar{\Omega} \in M_{N'}(K)$ be the matrix of $\bar{\omega}$. Now fix $\zeta \in K$ and consider the canonical ring homomorphism $\varphi_\zeta \colon \mathbb{Z}[T] \to K$ with $\varphi_\zeta(T) = \zeta$. Applying φ_ζ to the entries of Ω, we obtain $\bar{\Omega}$. Hence, the above identity over $\mathbb{Z}[T]$ implies the identity $\bar{\Omega} \cdot \bar{X}_i(\zeta) = \bar{Y}_i(\zeta) \cdot \bar{\Omega}$ over K, which means that $\bar{\omega} \circ \bar{x}_i(\zeta) \circ \bar{\omega}^{-1} = \bar{y}_i(\zeta)$, as desired.

3.6. First examples and further constructions

Let us look in more detail at the example where $L = \mathfrak{sl}_n(\mathbb{C})$, $n \geqslant 2$. We use the notation in Example 2.2.7. Let $H \subseteq L$ be the abelian subalgebra of diagonal matrices. For $1 \leqslant i, j \leqslant n$ let E_{ij} be the $n \times n$-matrix with 1 as its (i,j)-entry and zeroes elsewhere. Let $e_i := E_{i,i+1}$ and $f_i := E_{i+1,i}$ for $1 \leqslant i \leqslant n-1$. Then $\{e_i, f_i \mid 1 \leqslant i \leqslant n-1\}$ are Chevalley generators of L; furthermore, $h_i = [e_i, f_i] = E_{ii} - E_{i+1,i+1}$. Also recall from Example 2.2.7 that

$$\Phi = \{\varepsilon_i - \varepsilon_j \mid 1 \leqslant i, j \leqslant n, i \neq j\}, \qquad L_{\varepsilon_i - \varepsilon_j} = \langle E_{ij} \rangle_{\mathbb{C}}.$$

We set $\mathbf{e}_\alpha^+ := (-1)^j E_{ij}$ for $\alpha = \varepsilon_i - \varepsilon_j$, $i \neq j$. By Exercise Sheet 9, the collection $\{\mathbf{e}_\alpha^+ \mid \alpha \in \Phi\}$ satisfies the conditions in Corollary 2.7.11. In particular, $\mathbf{e}_{\alpha_i}^+ = -(-1)^i e_i$ and $\mathbf{e}_{-\alpha_i}^+ = (-1)^i f_i$ for $1 \leqslant i \leqslant n-1$; furthermore, $h_i^+ = [e_i, \mathbf{e}_{-\alpha_i}^+] = (-1)^i h_i$. Now let K be any field. Following the construction in the previous section, we need to consider a vector space $\bar{L}$ over K with a basis indexed by the canonical basis $\mathbf{B}$ of L. Concretely, we may take $\bar{L} := \mathfrak{sl}_n(K)$ with basis

$$\bar{\mathbf{B}} = \{\bar{h}_j^+ \mid 1 \leqslant j \leqslant n-1\} \cup \{\bar{\mathbf{e}}_\alpha^+ \mid \alpha \in \Phi\},$$

where $\bar{h}_j \in \mathfrak{sl}_n(K)$ and $\bar{\mathbf{e}}_\alpha^+ \in \mathfrak{sl}_n(K)$ are defined exactly as above, using analogues of the matrices E_{ij} over K. For $1 \leqslant i \leqslant n-1$ and $\zeta \in K$, the actions of $\bar{x}_i(\zeta)$ and $\bar{y}_i(\zeta)$ are given as follows.

$$\bar{x}_i(\zeta)(\bar{h}_j^+) = \bar{h}_j^+ + |a_{ji}|\zeta\bar{\mathbf{e}}_{\alpha_i}^+, \quad \bar{x}_i(\zeta)(\bar{\mathbf{e}}_{-\alpha_i}^+) = \bar{\mathbf{e}}_{-\alpha_i}^+ + \zeta\bar{h}_i^+ + \zeta^2\bar{\mathbf{e}}_{\alpha_i}^+,$$
$$\bar{y}_i(\zeta)(\bar{h}_j^+) = \bar{h}_j^+ + |a_{ji}|\zeta\bar{\mathbf{e}}_{-\alpha_i}^+, \quad \bar{y}_i(\zeta)(\bar{\mathbf{e}}_{\alpha_i}^+) = \bar{\mathbf{e}}_{\alpha_i}^+ + \zeta\bar{h}_i^+ + \zeta^2\bar{\mathbf{e}}_{-\alpha_i}^+.$$

We also have $\bar{x}_i(\zeta)(\bar{\mathbf{e}}^+_{\alpha_i}) = \bar{\mathbf{e}}^+_{\alpha_i}$ and $\bar{y}_i(\zeta)(\bar{\mathbf{e}}^+_{-\alpha_i}) = \bar{\mathbf{e}}^+_{-\alpha_i}$. Now let $\alpha \in \Phi$. If $\alpha + \alpha_i \notin \Phi \cup \{\underline{0}\}$, then $\bar{x}_i(\zeta)(\bar{\mathbf{e}}^+_\alpha) = \bar{\mathbf{e}}^+_\alpha$; in particular, this applies to $\alpha = \alpha_i$. If $\alpha - \alpha_i \notin \Phi \cup \{\underline{0}\}$, then $\bar{y}_i(\zeta)(\bar{\mathbf{e}}^+_{-\alpha}) = \bar{\mathbf{e}}^+_{-\alpha}$; in particular, this applies to $\alpha = -\alpha_i$. By Example 3.5.2, we have

$$\begin{aligned} \bar{x}_i(\zeta)(\bar{\mathbf{e}}^+_\alpha) &= \bar{\mathbf{e}}^+_\alpha + t\bar{\mathbf{e}}^+_{\alpha+\alpha_i} \quad \text{if } \alpha + \alpha_i \in \Phi, \\ \bar{y}_i(\zeta)(\bar{\mathbf{e}}^+_\alpha) &= \bar{\mathbf{e}}^+_\alpha + t\bar{\mathbf{e}}^+_{\alpha-\alpha_i} \quad \text{if } \alpha - \alpha_i \in \Phi. \end{aligned}$$

(Note that $q_{i,\alpha} = 0$ in the first case, and $p_{i,\alpha} = 0$ in the second case.) Now we exploit the fact that $\bar{L} = \mathfrak{sl}_n(K)$ is not just a vector space but a Lie algebra in its own right, with the usual Lie bracket. Then the above formulae can be re-written as follows, where $\bar{b} \in \bar{\mathbf{B}}$:

$$\begin{aligned} \bar{x}_i(\zeta)(\bar{b}) &= \bar{b} + \zeta[\bar{e}_i, \bar{b}] \qquad \text{if } \bar{b} \neq \bar{\mathbf{e}}^+_{-\alpha_i}, \\ \bar{y}_i(\zeta)(\bar{b}) &= \bar{b} + \zeta[\bar{f}_i, \bar{b}] \qquad \text{if } \bar{b} \neq \bar{\mathbf{e}}^+_{\alpha_i}, \\ \bar{x}_i(\zeta)(\bar{\mathbf{e}}^+_{-\alpha_i}) &= \bar{\mathbf{e}}^+_{-\alpha_i} + \zeta[\bar{e}_i, \bar{\mathbf{e}}^+_{-\alpha_i}] + \zeta^2\bar{\mathbf{e}}^+_{\alpha_i}, \\ \bar{y}_i(\zeta)(\bar{\mathbf{e}}^+_{\alpha_i}) &= \bar{\mathbf{e}}^+_{\alpha_i} + \zeta[\bar{f}_i, \bar{\mathbf{e}}^+_{\alpha_i}] + \zeta^2\bar{\mathbf{e}}^+_{-\alpha_i}. \end{aligned}$$

(For example, arguing as in the proof of Theorem 3.5.1, we see that $[\bar{e}_i, \bar{h}^+_j] = |a_{ji}|\bar{\mathbf{e}}^+_{\alpha_i}$; if $\alpha + \alpha_i \in \Phi$, then $[\bar{e}_i, \bar{\mathbf{e}}^+_\alpha] = \bar{\mathbf{e}}^+_{\alpha+\alpha_i}$, and so on.) Now let us define the following $n \times n$-matrices over K:

$$x_i^*(\zeta) := \bar{I}_n + \zeta\bar{e}_i \quad \text{and} \quad y_i^*(\zeta) := \bar{I}_n + \zeta\bar{f}_i \quad \text{for } 1 \leqslant i \leqslant n-1$$

(where $\bar{I}_n$ is the $n \times n$-identity matrix over K). Then $x_i^*(\zeta)$ is upper triangular with 1 along the diagonal; $y_i^*(\zeta)$ is lower triangular with 1 along the diagonal. In particular, $\det(x_i^*(\zeta)) = \det(y_i^*(\zeta)) = 1$.

Lemma 3.6.1. *In the above setting, let $A \in \bar{L} = \mathfrak{sl}_n(K)$. Then*

$$\bar{x}_i(\zeta)(A) = x_i^*(\zeta) \cdot A \cdot x_i^*(\zeta)^{-1}, \qquad \bar{y}_i(\zeta)(A) = y_i^*(\zeta) \cdot A \cdot y_i^*(\zeta)^{-1}.$$

Proof. We note that $\bar{e}_i^2 = \bar{f}_i^2 = 0$. Hence, we have $x_i^*(\zeta)^{-1} = x_i^*(-\zeta)$ and $y_i^*(\zeta)^{-1} = y_i^*(-\zeta)$. This yields:

$$\begin{aligned} x_i^*(\zeta) \cdot A \cdot x_i^*(\zeta)^{-1} &= (\bar{I}_n + \zeta\bar{e}_i)A(\bar{I}_n - \zeta\bar{e}_i) = (A + \zeta\bar{e}_iA)(\bar{I}_n - \zeta\bar{e}_i) \\ &= A + \zeta\bar{e}_iA - \zeta A\bar{e}_i - \zeta^2\bar{e}_iA\bar{e}_i = A + \zeta[\bar{e}_i, A] - \zeta^2\bar{e}_iA\bar{e}_i; \end{aligned}$$

furthermore, $\bar{e}_iA\bar{e}_i = a_{i+1,i}\bar{e}_i$. Similarly, we obtain

$$y_i^*(\zeta) \cdot A \cdot y_i^*(\zeta)^{-1} = A + \zeta[\bar{f}_i, A] - \zeta^2 a_{i,i+1}\bar{f}_i.$$

We have to compare these formulae with the above ones for the actions of $\bar{x}_i(\zeta)$ and $\bar{y}_i(\zeta)$. It is sufficient to do this for matrices A that belong to the basis $\bar{\mathbf{B}}$. Hence, we must check the following implications:

$$A \neq \bar{\mathbf{e}}^+_{-\alpha_i} \;\Rightarrow\; a_{i+1,i} = 0, \qquad A = \bar{\mathbf{e}}^+_{-\alpha_i} \;\Rightarrow\; \bar{\mathbf{e}}^+_{\alpha_i} = -a_{i+1,i}\bar{e}_i,$$
$$A \neq \bar{\mathbf{e}}^+_{\alpha_i} \;\Rightarrow\; a_{i,i+1} = 0, \qquad A = \bar{\mathbf{e}}^+_{\alpha_i} \;\Rightarrow\; \bar{\mathbf{e}}^+_{-\alpha_i} = -a_{i,i+1}\bar{f}_i.$$

The first and third implications are clear by the above description of $\bar{\mathbf{B}}$. Now assume that $A = \bar{\mathbf{e}}^+_{-\alpha_i}$. Then $A = (-1)^i\bar{f}_i$ and so $a_{i+1,i} = (-1)^i$. But then $-a_{i+1,i}\bar{e}_i = -(-1)^i\bar{e}_i = \bar{\mathbf{e}}^+_{\alpha_i}$, as required. The argument for $A = \bar{\mathbf{e}}^+_{\alpha_i}$ is analogous. □

Next, we need the following result (which is independent of any theory of Lie algebras or Chevalley groups):

Proposition 3.6.2. *Let $n \geqslant 2$ and K be any field. Then*

$$\mathrm{SL}_n(K) = \langle x_i^*(\zeta), y_i^*(\zeta) \mid 1 \leqslant i \leqslant n-1, \zeta \in K\rangle.$$

Proof. We proceed by induction on n, where we start the induction with $n = 1$. Note that the assertion also holds for $\mathrm{SL}_1(K) = \{\mathrm{id}\}$. Now let $n \geqslant 2$ and assume that the assertion is already proved for $\mathrm{SL}_{n-1}(K)$. Let $G_n \subseteq \mathrm{SL}_n(K)$ be the subgroup generated by the specified generators; we must show that $G_n = \mathrm{SL}_n(K)$. We set

$$x_{ij}^*(\zeta) := \bar{I}_n + \zeta E_{ij} \qquad \text{for any } \zeta \in K \text{ and } 1 \leqslant i,j \leqslant n,\ i \neq j;$$

in particular, $x_i^*(\zeta) = x_{i,i+1}^*(\zeta)$ and $y_i^*(\zeta) = x_{i+1,i}^*(\zeta)$. First we show:

$$x_{i1}^*(\zeta) \in G_n \qquad \text{and} \qquad x_{1i}^*(\zeta) \in G_n \qquad \text{for } 2 \leqslant i \leqslant n.$$

This is seen as follows. If $n = 2$, there is nothing to show. Now let $n \geqslant 3$. Let $i,j,k \in \{1,\ldots,n\}$ be pairwise distinct; then the following commutation rule is easily checked by an explicit computation:

$$x_{jk}^*(-\zeta') \cdot x_{ij}^*(-\zeta) \cdot x_{jk}^*(\zeta') \cdot x_{ij}^*(\zeta) = x_{ik}^*(-\zeta\zeta')$$

for all $\zeta, \zeta' \in K$. Setting $\zeta' = -1$, $i = 3$, $j = 2$ and $k = 1$, we obtain:

$$x_{21}^*(1) \cdot x_{32}^*(-\zeta) \cdot x_{21}^*(-1) \cdot x_{32}^*(\zeta) = x_{31}^*(\zeta)$$

for all $\zeta \in K$. Hence, since the left hand side belongs to G_n, we also have $x_{31}^*(\zeta) \in G_n$ for all $\zeta \in K$. Next, if $n \geqslant 4$, then we set $\zeta' = -1$, $i = 4$, $j = 3$ and $k = 1$. This yields

$$x_{31}^*(1) \cdot x_{43}^*(-\zeta) \cdot x_{31}^*(-1) \cdot x_{43}^*(\zeta) = x_{41}^*(\zeta).$$

Since the left hand side is already known to belong to G_n, we also have $x_{41}^*(\zeta) \in G_n$. Continuing in this way, we find that $x_{i1}^*(\zeta) \in G_n$ for all $\zeta \in K$ and $2 \leqslant i \leqslant n$. The argument for $x_{1i}^*(\zeta)$ is analogous.

Now let $A = (a_{ij}) \in \mathrm{SL}_n(K)$ be arbitrary. It will be useful to remember that, for $i \geqslant 2$, the matrix $x_{i1}^*(\zeta) \cdot A$ is obtained by adding the first row of A, multiplied by ζ, to the i-th row of A. Similarly, the matrix $A \cdot x_{1i}^*(\zeta)$ is obtained by adding the first column of A, multiplied by ζ, to the i-th column of A. We claim that there is a finite sequence of operations of this kind that transforms A into a new matrix $B = (b_{ij})$ such that

$$B = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & B' \end{array}\right) \qquad \text{where} \qquad B' \in \mathrm{SL}_{n-1}(K).$$

Indeed, since $\det(A) \neq 0$, the first column of A is non-zero and so there exists some $i \in \{1, \ldots, n\}$ such that $a_{i1} \neq 0$. If $i > 1$, then

$$A' := x_{i1}^*\big(a_{i1}^{-1}(1 - a_{11})\big) \cdot A$$

has entry 1 at position $(1,1)$. But then we can add suitable multiplies of the first row of A' to the other rows and obtain a new matrix A'' that has entry 1 at position $(1,1)$ and entry 0 at positions $(i,1)$ for $i \geqslant 2$. Next we can add suitable multiplies of the first column of A'' to the other columns and achieve that all further entries in the first row become 0. Thus, we have transformed A into a new matrix B as required. On the other hand, if there is no $i > 1$ such that $a_{i1} \neq 0$, then $a_{11} \neq 0$ and $a_{i1} = 0$ for $i \geqslant 2$. In that case, the matrix $x_{21}^*(1) \cdot A$ has a non-zero entry at position $(2,1)$ and we are in the previous case.

Now consider B as above. By induction, we have $\mathrm{SL}_{n-1}(K) = G_{n-1}$; so the submatrix B' can be expressed as a product of the specified generators of $\mathrm{SL}_{n-1}(K)$. Under the embedding

$$\mathrm{SL}_{n-1}(K) \hookrightarrow \mathrm{SL}_n(K), \qquad C \mapsto \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & C \end{array}\right),$$

the generators of $\mathrm{SL}_{n-1}(K)$ are sent to the generators $x_i^*(\zeta) \in \mathrm{SL}_n(K)$ and $y_i^*(\zeta) \in SL_n(K)$, where $\zeta \in K$ and $2 \leqslant i \leqslant n-1$. Hence, any B as above can be expressed as a product of generators $x_i^*(\zeta)$ and $y_i^*(\zeta)$ in $\mathrm{SL}_n(K)$, for various $\zeta \in K$ and $2 \leqslant i \leqslant n-1$. Since B was obtained from A by a sequence of multiplications with matrices $x_{1i}^*(\zeta) \in G_n$

or $x_{i1}^*(\zeta) \in G_n$, we conclude that $A \in G_n$ (and we even described an algorithm for expressing A in terms of the specified generators). $\square$

Proposition 3.6.3 (Ree). *If $L = \mathfrak{sl}_n(\mathbb{C})$ and K is any field, then the Chevalley group $\bar{G}' \subseteq \mathrm{GL}(\bar{L})$ (as in Definition 3.5.5) is isomorphic to $\mathrm{SL}_n(K)/Z$, where $Z = \{\zeta \bar{I}_n \mid \zeta \in K^\times, \zeta^n = 1\}$.*

Proof. As above, let $\bar{L} = \mathfrak{sl}_n(K)$. We also set $G^* := \mathrm{SL}_n(K)$. Then G^* acts on $\bar{L}$ by conjugation. Thus, for $g \in G^*$ we define $\gamma_g \colon \bar{L} \to \bar{L}$ by $\gamma_g(A) := g \cdot A \cdot g^{-1}$; then $\gamma_g \in \mathrm{GL}(\bar{L})$. Furthermore, the map $\gamma \colon G^* \to \mathrm{GL}(\bar{L})$, $g \mapsto \gamma_g$, is a group homomorphism. By Lemma 3.6.1, we have $\gamma_g = \bar{x}_i(\zeta)$ for $g = x_i^*(\zeta)$, and $\gamma_g = \bar{y}_i(\zeta)$ for $g = y_i^*(\zeta)$. Using also Proposition 3.6.2, we conclude that the image of γ equals the Chevalley group $\bar{G}' \subseteq \mathrm{GL}(\bar{L})$. Thus, we have a surjective homomorphism $\gamma \colon G^* \to \bar{G}'$, and it remains to show that $\ker(\gamma) = Z$. So let $g \in G^*$ be such that $\gamma_g = \mathrm{id}_{\bar{L}}$. Then $g \cdot A = A \cdot g$ for all $A \in \bar{L}$; it is a standard fact from Linear Algebra that then $g = \zeta \bar{I}_n$ for some $\zeta \in K$. Since $\det(g) = 1$, we must have $\zeta^n = 1$ and so $g \in Z$. Conversely, it is clear that $Z \subseteq \ker(\gamma)$. $\square$

Remark 3.6.4. (a) It is known that $\bar{G}' \cong \mathrm{SL}_n(K)/Z$ is simple, unless $n = 2$ and K has 2 or 3 elements; see, e.g., [**16**, Theorem 1.13].

(b) The Chevalley groups associated with the classical Lie algebras $\mathfrak{go}_n(Q_n, \mathbb{C})$ can be identified with symplectic or orthogonal groups in a similar way; see Carter [**6**, Chap. 11] for further details.

(c) If K is a finite field, then $\bar{G}'$ certainly is a finite group. Even if K is small, then these groups may simply become enormous. For example, if $|K| = 2$ and L is of type E_8, then one can show that $\bar{G}'$ has $\approx 3{,}38 \times 10^{74}$ elements. Nevertheless, the groups $\bar{G}'$ have a very user-friendly internal structure, and there are highly convenient ways how to work with their elements.

Exercise 3.6.5. The purpose of this exercise is to give at least one example showing that the above procedure also works for the classical Lie algebras introduced in Section 1.6. Let $L = \mathfrak{go}_4(Q_4, \mathbb{C})$, where

$$Q_4 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \qquad Q_4^{\mathrm{tr}} = -Q_4.$$

Let $I = \{1,2\}$. We have $\Phi = \{\pm\alpha_1, \pm\alpha_2, \pm(\alpha_1+\alpha_2), \pm(\alpha_1+2\alpha_2)\}$. Chevalley generators for L are given as follows:

$$\begin{aligned} e_1 &= \tfrac{1}{2}A_{2,3}, \quad f_1 = -\tfrac{1}{2}A_{3,2}, \quad h_1 = [e_1, f_1] = \mathrm{diag}(0,1,-1,0);\\ e_2 &= -A_{1,2}, \quad f_2 = -A_{2,1}, \quad h_2 = [e_2, f_2] = \mathrm{diag}(1,-1,1,-1). \end{aligned}$$

(See the proof of Proposition 2.5.8.) We have the relations $[h_1, e_2] = -e_2$ and $[h_2, e_1] = -2e_1$; see the structure matrix in Table 2 (p. 76).

(a) Let $\epsilon\colon I \to \{\pm 1\}$ be given by $\epsilon(1) = 1$ and $\epsilon(2) = -1$, as in Table 9 (p. 127). Starting with $\mathbf{e}^+_{\alpha_i} = \epsilon(i)e_i$ and $\mathbf{e}^+_{-\alpha_i} = -\epsilon(i)f_i$ for $i \in I$, determine all the elements of the canonical basis $\mathbf{B}$, explicitly as matrices in L; observe that all those matrices have entries in $\mathbb{Z}$.

(b) Let K be any field and $\bar{L} := \mathfrak{go}_4(Q_4, K)$. The assumption in Section 1.6 that $\mathrm{char}(K) \neq 2$ is not important here; check that Proposition 1.6.6(b) also holds over K instead of $\mathbb{C}$.

(c) Define $\bar{\mathbf{B}} \subseteq \bar{L}$ by taking analogues of the matrices in (a) over K; check that $\bar{\mathbf{B}}$ is a basis of $\bar{L}$. For $i \in I$ and $\zeta \in K$, determine the matrices of $\bar{x}_i(\zeta)$ and $\bar{y}_i(\zeta)$ with respect to $\bar{\mathbf{B}}$. Check that the relations in Lemma 3.6.1 also hold here.

(d) Let $\mathrm{Sp}_4(K) := \{A \in M_4(K) \mid A^{\mathrm{tr}}Q_4A = Q_4\}$. Check that $\mathrm{Sp}_4(K)$ is a subgroup of $\mathrm{GL}_4(K)$; it is called the 4-dimensional *symplectic group*. Analogously to Proposition 3.6.3, show that $\bar{G}' \cong \mathrm{Sp}_4(K)/Z$, where $Z = \{\pm \bar{I}_4\}$. (Here, the difficult part is to show the analogue of Proposition 3.6.2; for help and further references, see [**6**, Chap. 11].)

Now let us return to the general situation, where $\bar{G}'$ is the Chevalley group (over K) associated with a Lie algebra L of Cartan–Killing type. Our next aim is to introduce the "full" Chevalley group $\bar{G}$ over K mentioned in the footnote to Definition 3.5.5. The basic idea is to add to $\bar{G}' \subseteq \mathrm{GL}(\bar{L})$ some automorphisms of $\bar{L}$ that are represented by diagonal matrices. This involves the following constructions.

Definition 3.6.6. A map $\chi\colon \Phi \to K^\times$ is called a *K-character of* Φ if $\chi(-\alpha) = \chi(\alpha)^{-1}$ and $\chi(\alpha+\beta) = \chi(\alpha)\chi(\beta)$ for all $\alpha, \beta \in \Phi$ such that $\alpha+\beta \in \Phi$. The set of all K-characters of Φ will be denoted by $\mathfrak{X}_K(\Phi)$. This set is itself an abelian group (written additively) via

$$(\chi+\chi')(\alpha) := \chi(\alpha)\chi'(\alpha) \qquad (\chi, \chi' \in \mathfrak{X}_K(\Phi), \alpha \in \Phi);$$

the neutral element is the unit character, which sends each $\alpha \in \Phi$ to $1_K \in K$. Given $\chi \in \mathfrak{X}_K(\Phi)$, we define a linear map $\bar{h}(\chi)\colon \bar{L} \to \bar{L}$ by

$$\bar{h}(\chi)(\bar{h}_j^+) := \bar{h}_j^+ \quad (j \in I) \quad \text{and} \quad \bar{h}(\chi)(\bar{\mathbf{e}}_\alpha^+) := \chi(\alpha)\bar{\mathbf{e}}_\alpha^+ \quad (\alpha \in \Phi).$$

We certainly have $\bar{h}(\chi + \chi') = \bar{h}(\chi) \circ \bar{h}(\chi')$ for all $\chi, \chi' \in \mathfrak{X}_K(\Phi)$. Furthermore, $\bar{h}(\chi)$ is invertible, where $\bar{h}(\chi)^{-1} = \bar{h}(-\chi)$. Thus, we obtain a group homomorphism

$$\mathfrak{X}_K(\Phi) \to \mathrm{GL}(\bar{L}), \qquad \chi \mapsto \bar{h}(\chi),$$

and one immediately sees that this is injective.

Remark 3.6.7. Let $\chi \in \mathfrak{X}_K(\Phi)$. We claim that, for $\alpha \in \Phi$, we have

$$\chi(\alpha) = \prod_{j \in I} \chi(\alpha_j)^{n_j} \qquad \text{where} \qquad \alpha = \sum_{j \in I} n_j \alpha_j \text{ with } n_j \in \mathbb{Z}.$$

This is seen as follows. First assume that $\alpha \in \Phi^+$. We proceed by induction on $\mathrm{ht}(\alpha)$. If $\mathrm{ht}(\alpha) = 1$, then $\alpha = \alpha_j$ for some $j \in I$ and the assertion is clear. Now let $\mathrm{ht}(\alpha) > 1$. By the Key Lemma 2.3.4, there exists some $i \in I$ such that $\alpha' := \alpha - \alpha_i \in \Phi^+$. Then the defining condition for χ implies that $\chi(\alpha) = \chi(\alpha')\chi(\alpha_i)$. Using induction, the desired formula holds for $\chi(\alpha)$. Finally, if $\alpha \in \Phi^-$, then $-\alpha \in \Phi^+$ and $\chi(\alpha) = \chi(-\alpha)^{-1}$. Hence, the desired formula holds for α as well.

The above formula shows that χ is uniquely determined by the values $\{\chi(\alpha_j) \mid j \in I\}$. Conversely, given any collection of elements $\underline{\zeta} = \{\zeta_j \mid j \in I\} \subseteq K^\times$, we can define a map $\chi_{\underline{\zeta}}\colon \Phi \to K^\times$ by

$$\chi_{\underline{\zeta}}(\alpha) := \prod_{j \in I} \zeta_j^{n_j} \qquad \text{where } \alpha = \textstyle\sum_{j \in I} n_j \alpha_j \text{ with } n_j \in \mathbb{Z}.$$

One easily sees that $\chi_{\underline{\zeta}} \in \mathfrak{X}_K(\Phi)$.

Example 3.6.8. Let $i \in I$ and $\zeta \in K^\times$. Then we obtain a K-character $\chi_{i,\zeta} \in \mathfrak{X}_K(\Phi)$ by setting

$$\chi_{i,\zeta}(\alpha) := \zeta^{\langle \alpha_i^\vee, \alpha \rangle} \qquad \text{for all } \alpha \in \Phi.$$

As in Remark 3.6.7, this K-character is associated with the collection of elements $\underline{\zeta} = \{\zeta^{a_{ij}} \mid j \in I\} \subseteq K^\times$. We shall denote $\bar{h}_i(\zeta) := \bar{h}(\chi_{i,\zeta}) \in \mathrm{GL}(\bar{L})$; thus, for $j \in I$ and $\alpha \in \Phi$, we have

$$\bar{h}_i(\zeta)(\bar{h}_j^+) = \bar{h}_j^+ \qquad \text{and} \qquad \bar{h}_i(\zeta)(\bar{\mathbf{e}}_\alpha^+) = \zeta^{\langle \alpha_i^\vee, \alpha \rangle} \bar{\mathbf{e}}_\alpha^+.$$

We will see later that $\bar{h}_i(\zeta) \in \bar{G}'$. But in general, there can exist $\chi \in \mathfrak{X}_K(\Phi)$ such that $\bar{h}(\chi) \notin \bar{G}'$; see Example 3.6.11 below. (This is one subtlety of the definition of Chevalley groups over arbitrary fields K; it disappears when K is algebraically closed.)

Proposition 3.6.9. *Let $i \in I$, $\zeta \in K$ and $\chi \in \mathfrak{X}_K(\Phi)$. Then $\bar{h}(\chi)\bar{x}_i(\zeta)\bar{h}(\chi)^{-1} = \bar{x}_i(\chi(\alpha_i)\zeta)$ and $\bar{h}(\chi)\bar{y}_i(\zeta)\bar{h}(\chi)^{-1} = \bar{y}_i(\chi(\alpha_i)^{-1}\zeta)$.*

Proof. First let $K = \mathbb{C}$ and $\bar{L} = L$; to simplify the notation, we omit the bars over the various symbols (like $\bar{L}$, $\bar{h}(\chi)$, ...). Then the defining conditions on χ imply that $h(\chi) \in \mathrm{Aut}(L)$. Indeed, let $\alpha, \beta \in \Phi$. If $\alpha + \beta \in \Phi$, then

$$\begin{aligned} h(\chi)\big([\mathbf{e}_\alpha^+, \mathbf{e}_\beta^+]\big) &= N_{\alpha,\beta}^+ h(\chi)(\mathbf{e}_{\alpha+\beta}^+) = N_{\alpha,\beta}^+ \chi(\alpha+\beta)\mathbf{e}_{\alpha+\beta}^+ \\ &= \chi(\alpha)\chi(\beta)[\mathbf{e}_\alpha^+, \mathbf{e}_\beta^+] = [h(\chi)(\mathbf{e}_\alpha^+), h(\chi)(\mathbf{e}_\beta^+)], \end{aligned}$$

as required. If $\beta = -\alpha$, then $h(\chi)\big([\mathbf{e}_\alpha^+, \mathbf{e}_{-\alpha}^+]\big) = (-1)^{\mathrm{ht}(\alpha)} h(\chi)(h_\alpha)$. Since $h(\chi(h_j^+) = h_j^+$ for $j \in I$, we have $h(\chi(h_\alpha) = h_\alpha$. On the other hand, we also have $[h(\chi)(\mathbf{e}_\alpha^+), h(\chi)(\mathbf{e}_{-\alpha}^+)] = \chi(\alpha)\chi(-\alpha)[\mathbf{e}_\alpha^+, \mathbf{e}_{-\alpha}^+] = (-1)^{\mathrm{ht}(\alpha)} h_\alpha$. Finally, if $\alpha+\beta \notin \Phi$ and $\beta \neq -\alpha$, then $h(\chi)\big([\mathbf{e}_\alpha^+, \mathbf{e}_\beta^+]\big) = 0$ and $[h(\chi)(\mathbf{e}_\alpha^+), h(\chi)(\mathbf{e}_\beta^+)] = 0$. Similarly, one sees that $h(\chi)$ respects the brackets $[h_j^+, \mathbf{e}_\alpha^+] = \alpha(h_j^+)\mathbf{e}_\alpha^+$ and $[h_j^+, h_{j'}^+] = 0$ for $j, j' \in I$ and $\alpha \in \Phi$. Now, having shown that $h(\chi) \in \mathrm{Aut}(L)$, we can apply Lemma 3.5.9; this yields that

$$\begin{aligned} h(\chi)x_i(t)h(\chi)^{-1} &= h(\chi) \circ \exp\big(t\,\mathrm{ad}_L(e_i)\big) \circ h(\chi)^{-1} \\ &= \exp\big(t\,\mathrm{ad}_L\big(h(\chi)(e_i)\big)\big) = \exp\big(\chi(\alpha_i)t\,\mathrm{ad}_L(e_i)\big) = x_i(\chi(\alpha_i)t) \end{aligned}$$

for all $t \in \mathbb{C}$. Similarly, we see that $h(\chi)y_i(t)h(\chi)^{-1} = y_i(\chi(\alpha_i)^{-1}t)$. In order to pass from $\mathbb{C}$ to K, we have to lift these identities to the appropriate polynomial level. We work over the Laurent polynomial ring $\mathscr{O} := \mathbb{Z}[T^{\pm 1}, Z_j^{\pm 1}\ (j \in I)]$ in independent indeterminates T and Z_j ($j \in I$). Let $X_i(T)$ and $Y_i(T)$ be the matrices associated wth $x_i(t)$ and $y_i(t)$ as in Remark 3.5.3; we regard them as matrices in $M_{N'}(\mathscr{O})$. Let us write $\underline{Z} = (Z_j \mid j \in I)$. Let $H(\underline{Z}) \in M_{N'}(\mathscr{O})$ be the diagonal matrix with entry 1 at the diagonal position corresponding to a basis element h_j^+ ($j \in I$), and entry $\prod_{j \in I} Z_j^{n_j}$ at the diagonal position corresponding to a basis element $\mathbf{e}_\alpha^+$ (where $\alpha = \sum_{j \in I} n_j\alpha_j \in \Phi$).

Then we claim that we have the following identities in $M_{N'}(\mathscr{O})$:

$$H(\underline{Z}) \cdot X_i(T) = X_i(Z_i T) \cdot H(\underline{Z}),$$
$$H(\underline{Z}) \cdot Y_i(T) = Y_i(Z_i^{-1} T) \cdot H(\underline{Z}).$$

To see this, let us fix a collection of elements $\underline{z} = (z_j \mid j \in I) \subseteq \mathbb{C}^\times$. As in Remark 3.6.7, we obtain a corresponding K-character $\chi_{\underline{z}} \in \mathfrak{X}_{\mathbb{C}}(\Phi)$; we have $\chi_{\underline{z}}(\alpha_j) = z_j$ for $j \in I$. Now we note that the matrix of $h(\chi_{\underline{z}}) \in \mathrm{GL}(L)$ with respect to the basis $\mathbf{B}$ is obtained from $H(\underline{Z})$ upon substituting $Z_j \mapsto z_j$ for all $j \in I$. Let us also fix $t \in \mathbb{C}$. Then, as in the previous section, the matrix of $x_i(t) \in \mathrm{GL}(K)$ is obtained from $X_i(T)$ upon substituting $T \mapsto t$; similarly, the matrix of $y_i(t) \in \mathrm{GL}(K)$ is obtained from $y_i(T)$ upon substituting $T \mapsto t$. Hence, we have the following identities in $M_{N'}(\mathbb{C})$:

$$H(\underline{z}) \cdot X_i(t) = X_i(z_i t) \cdot H(\underline{z}),$$
$$H(\underline{z}) \cdot Y_i(T) = Y_i(z_i^{-1} t) \cdot H(\underline{z}).$$

Since this holds for all $t \in \mathbb{C}$ and all collections $\underline{z} = (z_j \mid j \in I) \subseteq \mathbb{C}^\times$, we conclude that the above identities in $M_{N'}(\mathscr{O})$ do hold, as claimed.

Now we can pass from $\mathbb{C}$ to K, by the usual argument. We fix $\zeta \in K$ and a K-character $\chi \in \mathfrak{X}_K(\Phi)$. By Remark 3.6.7, there is a collection $\underline{\xi} = (\xi_j \mid j \in I) \subseteq K^\times$ such that $\chi = \chi_{\underline{\xi}}$. We have a canonical ring homomorphism $\varphi_{\zeta,\underline{\xi}} \colon \mathscr{O} \to K$ such that $T \mapsto \zeta$, $Z_j \mapsto \xi_j$ $(j \in I)$ and $m \mapsto m{\cdot}1_K$ $(m \in \mathbb{Z})$. Applying $\varphi_{\zeta,\underline{\xi}}$ to the entries of $X_i(T)$, $Y_i(T)$ and $H(\underline{Z})$, we obtain the matrices of $\bar{x}_i(\zeta)$, $\bar{y}_i(\zeta)$ and $\bar{h}(\chi)$, respectively. Then the above identities between matrices over $M_{N'}(\mathscr{O})$ imply analogous identities between matrices over K. Finally, the latter identities mean that $\bar{h}(\chi)\bar{x}_i(\zeta) = \bar{x}_i(\chi(\alpha_i)\zeta)\bar{h}(\chi)$ and $\bar{h}(\chi)\bar{y}_i(\zeta) = \bar{y}_i(\chi(\alpha_i)^{-1}\zeta)\bar{h}(\chi)$, as desired. $\square$

Definition 3.6.10 (Chevalley [**9**, p. 37])**.** We define $\bar{G} \subseteq \mathrm{GL}(\bar{L})$ to be the subgroup generated by $\bar{G}'$ (as in Definition 3.5.5) and all the elements $\bar{h}(\chi)$, where $\chi \in \mathfrak{X}_K(\Phi)$. By Proposition 3.6.9, the generators of $\bar{G}'$ are normalised by all $\bar{h}(\chi)$ $(\chi \in \mathfrak{X}_K(\Phi))$. Consequently, $\bar{G}'$ is a normal subgroup of $\bar{G}$ and

$$\bar{G} = \{g\bar{h}(\chi) \mid g \in \bar{G}', \chi \in \mathfrak{X}_K(\Phi)\}.$$

Since all $\bar{h}(\chi)$ commute with each other, $\bar{G}/\bar{G}'$ is abelian.

Example 3.6.11. Let $L = \mathfrak{sl}_2(\mathbb{C})$ with the usual basis $\{e, h, f\}$. Then $H = \langle h \rangle_{\mathbb{C}}$ and $\Phi = \{\pm\alpha\}$, where $\alpha \in H^*$ is defined by $\alpha(h) = 2$. By Remark 3.6.7, we have $\mathfrak{X}_K(\Phi) = \{\chi_\xi \mid \xi \in K^\times\}$ where $\chi_\xi(\alpha) := \xi$. Let $\mathbf{B} = \{e, -h, -f\}$ as in Example 3.5.4. Then the matrix $\bar{H}(\xi)$ of $\bar{h}(\chi_\xi) \in \mathrm{GL}(\bar{L})$ with respect to $\bar{\mathbf{B}}$ is given by

$$\bar{H}(\xi) = \left(\begin{array}{ccc} \xi & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \xi^{-1} \end{array} \right).$$

On the other hand, using the calculations in Exercise 1.2.15, one checks that $\bar{h}(\chi_\xi) \in \bar{G}'$ if and only if ξ is a square in $K^\times$. (We leave this verification as an exercise for the reader.)

Lemma 3.6.12. *Assume that K is algebraically closed. Then, for any $\chi \in \mathfrak{X}_K(\Phi)$, we have $\bar{h}(\chi) \in \langle \bar{h}_i(\zeta) \mid i \in I, \zeta \in K^\times \rangle \subseteq \mathrm{GL}(\bar{L})$.*

Proof. As in Remark 3.6.7, we have $\chi = \chi_{\underline{\zeta}}$ for a suitable collection of elements $\underline{\zeta} = (\zeta_j \mid j \in I) \subseteq K^\times$. Then $\chi(\alpha_j) = \zeta_j$ for $j \in I$. For $l \in I$ define $\chi_l \in \mathfrak{X}_K(\Phi)$ by $\chi_l(\alpha_l) := \zeta_l$ and $\chi_l(\alpha_i) := 1$ for $i \neq l$. Then $\chi = \sum_{l \in I} \chi_l$ and, hence, $\bar{h}(\chi) = \prod_{l \in I} \bar{h}(\chi_l)$. So it is sufficient to prove the assertion for $\bar{h}(\chi_l)$, where $l \in I$ is fixed.

Now, since the structure matrix $A = (a_{ij})_{i,j \in I}$ of L has a non-zero determinant, there exists numbers $r_i \in \mathbb{Q}$ such that

$$\sum_{i \in I} r_i a_{ij} = \begin{cases} 1 & \text{if } j = l, \\ 0 & \text{if } j \neq l. \end{cases}$$

Let $n \in \mathbb{Z}_{>0}$ be such that $nr_i \in \mathbb{Z}$ for all $i \in I$. Since K is algebraically closed, there exists some $\xi \in K^\times$ such that $\xi^n = \zeta_l$. Now consider

$$\chi' := \sum_{i \in I} nr_i \chi_{i,\xi} \in \mathfrak{X}_K(\Phi) \qquad \text{(with } \chi_{i,\xi} \text{ as in Example 3.6.8).}$$

Since $\chi_{i,\xi}(\alpha_j) = \xi^{\langle \alpha_i^\vee, \alpha_j \rangle} = \xi^{a_{ij}}$ for $j \in I$, we obtain

$$\chi'(\alpha_j) = \prod_{i \in I} \xi^{nr_i a_{ij}} = \xi^{n \sum_{i \in I} r_i a_{ij}} = \begin{cases} \xi^n & \text{if } j = l, \\ 1 & \text{if } j \neq l. \end{cases}$$

Hence, we have $\chi_l = \chi'$ and so $\bar{h}(\chi_l) = \bar{h}(\chi') = \prod_{i \in I} \bar{h}_i(\xi^{nr_i})$. □

The above argument also shows that the assumption on K can be dropped if $\det(A) = \pm 1$. The exact relation between $\bar{G}'$ and $\bar{G}$ is rather subtle. It turns out that $\bar{h}_i(\zeta) \in \bar{G}'$ for all $i \in I$ and $\zeta \in K^\times$. Consequently, if K is algebraically closed, then $\bar{h}(\chi) \in \bar{G}'$ for all $\chi \in \mathfrak{X}_K(\Phi)$, and so $\bar{G}' = \bar{G}$.

Suggestions for further reading

Systematic descriptions of the irreducible root systems of the various types can be found in Bourbaki [**4**, VI, §4, no. 4.4–4.13]; see also Benson–Grove [**2**, §5.3] for explicit constructions and algorithms.

See Kac [**21**, §1.9] for some notes about the historical development of the study of Kac–Moody Lie algebras. The appendix of Moody–Pianzola [**25**] contains a much more thorough discussion of Example 3.3.2. The idea of replacing $\mathbb{C}$ by a ring of Laurent polynomials can be generalised to all Lie algebras of Cartan–Killing type; see, e.g., Carter [**7**, Chap. 18] for a detailed exposition.

There are several other proofs of the important Existence Theorem 3.3.10:

- Via free Lie algebras and definitions in terms of generators and relations. See Jacobson [**19**, Chap. VII, §4], Serre [**27**, Chap. VI, Appendix] (and also [**18**, §18] for further details).
- Via explicit descriptions of structure constants. There is an elegant way to do this for A of simply-laced type; the remaining cases are obtained by a "folding" procedure. See Kac [**21**, §7.8, §7.9]. For a general approach see Tits [**29**].
- Via explicit constructions in all cases. Historically, this is the original method. For the classical types A_n, B_n, C_n, D_n, we have seen this already.

The ChevLie package presented in Section 3.4 is one example of a whole variety of software packages for Lie theory. The computer algebra systems GAP (`http://www.gap-system.org`) and Magma (`http://magma.maths.usyd.edu.au/magma/`) contain large packages for Lie theory.

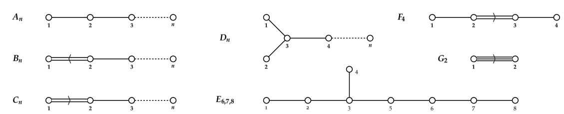
One important fact (that we did not have the time to prove here) is that $\bar{G}'$ is almost always a simple group. At the time of Chevalley's article [**9**], this gave several new classes of finite simple groups, the complete list of which is seen on the next page. The finitely many exceptions only occur when K is a finite field with 2 or 3 elements. More precisely, there are the following four cases where G' is not simple. Suppose first that $|K| = 2$. If L is of type A_1, then G' has order 6 and is isomorphic to the symmetric group $\mathfrak{S}_3$; if L is of type B_2, then G' has order 720 and is isomorphic to the symmetric group $\mathfrak{S}_6$; if L is of type G_2, then G' has order 12096 and there is a simple normal subgroup of index 2. The last exception occurs when $|K| = 3$ and L is of type A_1, in which case G' has order 12 and is isomorphic to the alternating group $\mathfrak{A}_4$. For details see Chevalley [**9**, Théorème 3 (p. 63)], Carter [**6**, §11.1] or Steinberg [**28**, Chapter 4].

In general, Carter [**6**] and Steinberg [**28**] are standard references for the further theory of Chevalley groups. In an equally famous "Séminaire" (1956–1958) directed by Chevalley, it was shown that, over an algebraically closed field k, the Chevalley groups $\bar{G}'$ are essentially the only semisimple algebraic groups, where the term "algebraic group" is meant in the context of algebraic geometry. See:

C. Chevalley, *Classification des groupes algébriques semi-simples*, Collected works. Vol. 3. Edited and with a preface by P. Cartier. With the collaboration of P. Cartier, A. Grothendieck and M. Lazard. Springer-Verlag, Berlin, 2005.

The Periodic Table Of Finite Simple Groups

Dynkin Diagrams of Simple Lie Algebras



$0, C_1, \mathbb{Z}_1$ **1** 1								
$A_1(4), A_1(5)$ A_5 60	$A_2(2)$ $A_1(7)$ 168							
$A_1(9), B_2(2)'$ A_6 360	${}^2G_2(3)'$ $A_1(8)$ 504							
A_7 2 520	$A_1(11)$ 660	$E_6(2)$ 214 841 575 522 005 575 270 400	$E_7(2)$ 7 997 476 042 075 799 759 100 487 262 680 802 918 400	$E_8(2)$ 337 804 753 143 634 806 261 388 190 614 085 595 079 991 692 242 467 651 576 160 959 909 068 800 000	$F_4(2)$ 3 311 126 603 366 400	$G_2(3)$ 4 245 696	${}^3D_4(2^3)$ 211 341 312	${}^2E_6(2^2)$ 76 532 479 683 774 853 939 200
$A_3(2)$ A_8 20 160	$A_1(13)$ 1 092	$E_6(3)$ 7 257 703 347 541 463 210 028 258 395 214 643 200	$E_7(3)$ 1 271 375 236 818 136 742 240 479 751 139 021 644 554 379 203 770 766 254 617 395 200	$E_8(3)$ [illegible]	$F_4(3)$ 5 734 420 792 816 671 844 761 600	$G_2(4)$ 251 596 800	${}^3D_4(3^3)$ 20 560 831 566 912	${}^2E_6(3^2)$ 14 636 855 916 969 695 633 965 120 680 532 377 600
A_9 181 440	$A_1(17)$ 2 448	$E_6(4)$ 85 528 710 781 342 640 103 833 619 055 142 765 466 746 880 000	$E_7(4)$ 111 131 458 114 940 385 379 597 233 477 884 941 280 664 199 527 155 056 307 251 745 263 504 588 800 000 000	$E_8(4)$ [illegible]	$F_4(4)$ 19 009 825 523 840 945 451 297 669 120 000	$G_2(5)$ 5 859 000 000	${}^3D_4(4^3)$ 67 802 350 642 790 400	${}^2E_6(4^2)$ 85 696 576 147 617 709 485 896 772 387 584 983 695 360 000 000
A_n $\frac{n!}{2}$	$PSL_{n+1}(q), L_{n+1}(q)$ $A_n(q)$ $\frac{q^{n(n+1)/2}}{(n+1,q-1)}\prod_{i=1}^{n}(q^{i+1}-1)$	$E_6(q)$ $\frac{q^{36}(q^{12}-1)(q^9-1)(q^8-1)(q^6-1)(q^5-1)(q^2-1)}{(3,q-1)}$	$E_7(q)$ $\frac{q^{63}}{(2,q-1)}\prod_{i=1, i\neq 2,8}^{9}(q^{2i}-1)$	$E_8(q)$ $q^{120}(q^{30}-1)(q^{24}-1)(q^{20}-1)(q^{18}-1)(q^{14}-1)(q^{12}-1)(q^8-1)(q^2-1)$	$F_4(q)$ $q^{24}(q^{12}-1)(q^8-1)(q^6-1)(q^2-1)$	$G_2(q)$ $q^6(q^6-1)(q^2-1)$	${}^3D_4(q^3)$ $q^{12}(q^8+q^4+1)(q^6-1)(q^2-1)$	${}^2E_6(q^2)$ $\frac{q^{36}(q^{12}-1)(q^9+1)(q^8-1)(q^6-1)(q^5+1)(q^2-1)}{(3,q+1)}$

								C_2 2
			${}^2A_3(4)$ $B_2(3)$ 25 920	$C_3(3)$ 4 585 351 680	$D_4(2)$ 174 182 400	${}^2D_4(2^2)$ 197 406 720	$G_2(2)'$ ${}^2A_2(9)$ 6 048	C_3 3
			$B_2(4)$ 979 200	$C_3(5)$ 228 501 000 000 000	$D_4(3)$ 4 952 179 814 400	${}^2D_4(3^2)$ 10 151 968 619 520	${}^2A_2(16)$ 62 400	C_5 5
${}^2B_2(2^3)$ 29 120	Tits* ${}^2F_4(2)'$ 17 971 200	${}^2G_2(3^3)$ 10 073 444 472	$B_3(2)$ 1 451 520	$C_4(3)$ 65 784 756 654 489 600	$D_5(2)$ 23 499 295 948 800	${}^2D_5(2^2)$ 25 015 379 558 400	${}^2A_2(25)$ 126 000	C_7 7
${}^2B_2(2^5)$ 32 537 600	${}^2F_4(2^3)$ 264 905 352 699 586 176 614 400	${}^2G_2(3^5)$ 49 825 657 439 340 552	$B_2(5)$ 4 680 000	$C_3(7)$ 273 457 218 604 953 600	$D_4(5)$ 8 911 539 000 000 000 000	${}^2D_4(4^2)$ 67 536 471 195 648 000	${}^2A_3(9)$ 3 265 920	C_{11} 11
${}^2B_2(2^7)$ 34 093 383 680	${}^2F_4(2^5)$ 1 318 633 155 799 591 447 702 161 609 782 722 560 000	${}^2G_2(3^7)$ 239 189 910 264 352 349 332 632	$B_2(7)$ 138 297 600	$C_3(9)$ 54 025 731 402 499 584 000	$D_5(3)$ 1 289 512 799 941 305 139 200	${}^2D_4(5^2)$ 17 880 203 250 000 000 000	${}^2A_2(64)$ 5 515 776	C_{13} 13
${}^2B_2(2^{2n+1})$ $q^2(q^2+1)(q-1)$	${}^2F_4(2^{2n+1})$ $q^{12}(q^6+1)(q^4-1)(q^3+1)(q-1)$	${}^2G_2(3^{2n+1})$ $q^3(q^3+1)(q-1)$	$O_{2n+1}(q), \Omega_{2n+1}(q)$ $B_n(q)$ $\frac{q^{n^2}}{(2,q-1)}\prod_{i=1}^{n}(q^{2i}-1)$	$PSp_{2n}(q)$ $C_n(q)$ $\frac{q^{n^2}}{(2,q-1)}\prod_{i=1}^{n}(q^{2i}-1)$	$O^+_{2n}(q)$ $D_n(q)$ $\frac{q^{n(n-1)}(q^n-1)}{(4,q^n-1)}\prod_{i=1}^{n-1}(q^{2i}-1)$	$O^-_{2n}(q)$ ${}^2D_n(q^2)$ $\frac{q^{n(n-1)}(q^n+1)}{(4,q^n+1)}\prod_{i=1}^{n-1}(q^{2i}-1)$	$PSU_{n+1}(q)$ ${}^2A_n(q^2)$ $\frac{q^{n(n+1)/2}}{(n+1,q+1)}\prod_{i=2}^{n+1}(q^i-(-1)^i)$	$\mathbb{Z}_p$ C_p p

M_{11} 7 920	M_{12} 95 040	M_{22} 443 520	M_{23} 10 200 960	M_{24} 244 823 040	$J(1), J(11)$ J_1 175 560	HJ J_2 604 800	HJM J_3 50 232 960	J_4 86 775 571 046 077 562 880	HS 44 352 000	McL 898 128 000	F_7, HHM, HTH He 4 030 387 200	Ru 145 926 144 000
Sz Suz 448 345 497 600	$O'NS, O\text{-}S$ $O'N$ 460 815 505 920	$\cdot 3$ Co_3 495 766 656 000	$\cdot 2$ Co_2 42 305 421 312 000	$\cdot 1$ Co_1 4 157 776 806 543 360 000	F_5, D HN 273 030 912 000 000	LyS Ly 51 765 179 004 000 000	F_3, E Th 90 745 943 887 872 000	$M(22)$ Fi_{22} 64 561 751 654 400	$M(23)$ Fi_{23} 4 089 470 473 293 004 800	$F_{3+}, M(24)'$ Fi'_{24} 1 255 205 709 190 661 721 292 800	F_2 B 4 154 781 481 226 426 191 177 580 544 000 000	F_1, M_1 M 808 017 424 794 512 875 886 459 904 961 710 757 005 754 368 000 000 000

- Alternating Groups
- Classical Chevalley Groups
- Chevalley Groups
- Classical Steinberg Groups
- Steinberg Groups
- Suzuki Groups
- Ree Groups and Tits Group*
- Sporadic Groups
- Cyclic Groups

Alternates† Symbol Order‡

*The Tits group ${}^2F_4(2)'$ is not a group of Lie type, but is the (index 2) commutator subgroup of ${}^2F_4(2)$. It is usually given honorary Lie type status.

The groups starting on the second row are the classical groups. The sporadic suzuki group is unrelated to the families of Suzuki groups.

†For sporadic groups and families, alternate names in the upper left are other names by which they may be known. For specific non-sporadic groups these are used to indicate isomorphisms. All such isomorphisms appear on the table except the family $B_n(2^m) \cong C_n(2^m)$.

‡Finite simple groups are determined by their order with the following exceptions: $B_n(q)$ and $C_n(q)$ for q odd, $n > 2$; $A_8 \cong A_3(2)$ and $A_2(4)$ of order 20160.