Berkeley Lectures on
Lie Groups and Quantum Groups

Richard Borcherds, Mark Haiman, Theo Johnson-Freyd, Nicolai Reshetikhin, and Vera Serganova

Last updated December 23, 2022

Contents

List of Theorems

About this book

This book on Lie Groups and Quantum Groups compiles four semesters of lectures given at the Univeristy of California, Berkeley. The lectures were given by Professors Richard Borcherds, Mark Haiman, Nicolai Reshetikhin, and Vera Serganova. Theo Johnson-Freyd, then a graduate student, transcribed the lectures and combined, edited, and reorganized them into this book. The classes, and therefore this book, would not have been possible without the students in them. We particularly thank Dustin Cartwright, Alex Fink, Dan Halpern-Leistner, Crystal Hoyt, Chul-Hee Lee, Sevak Mkrtchyan, Manuel Reyes, Noah Snyder, Matt Tucker-Simmons, and Harold Williams for their comments and supplementary lecture notes, and most importantly Anton Geraschenko for producing an earlier collection of lecture notes that inspired this volume.

The book has been divided into three parts. Part I closely follows Mark Haiman's one-semester course in Fall 2008. It covers pretty much all basics of Lie theory, up through the classification of semisimple complex Lie groups. Part II addresses some further topics, including noncompact Lie groups, the construction of E_8, and some algebraic group theory. Chapters 7 and 8 are based on lectures by Richard Borcherds in Spring 2006 and Chapter 9 follows lectures by Vera Serganova in Spring 2010. Part III is on Poisson Lie groups and quantum groups, and combines Nicolai Reshetikhin's course from Spring 2009 and Vera Serganova's course from Spring 2010.

Many chapters proceed at a rapid pace, with little discussion beyond the statements of definitions and results. Others chapter include considerably more exposition. Early results stated without proof should be understood as invitations to the reader: check this statement as a homework exercise!

The courses from which this book is derived did not follow any particular textbooks, but did suggest many books for the students to read for supplementary material. These included [Var84, Fuk86, Kac90, Bor91, FH91, Lus93, CP94, Dix96, GOV97, KS97, BK01, ES02, BG02, Kna02, Jan03, Bou05, Hum08, Lee09].

Part I

Lie Groups

Chapter 1

Motivation: Closed Linear Groups

1.1 Definition of a Lie group

1.1.1 Group objects

A *Lie group* is a group object in the category of manifolds. We will not digress too far into a discussion of categories, but we will use category theory as a language. Not every category has products, but given two objects S and T, we call the diagram

$$\begin{tikzcd} & S \\ S \times T \arrow[ur] \arrow[dr] & \\ & T \end{tikzcd}$$

the *categorical product* of S and T if for all objects X, the maps $X \to S \times T$ are in bijection with pairs of maps $X \to S$ and $X \to T$ in such a way that the following diagram commutes:

$$\begin{tikzcd} & & S \\ X \arrow[r] \arrow[urr] \arrow[drr] & S \times T \arrow[ur] \arrow[dr] & \\ & & T \end{tikzcd}$$

1.1.1.1 Definition *Let $\mathcal{C}$ be a category with finite products; denote the terminal object by* $\{\mathrm{pt}\}$*. A* group object *in $\mathcal{C}$ is an object G along with maps $\mu : G \times G \to G$, $i : G \to G$, and $e : \{\mathrm{pt}\} \to G$, such that the following diagrams commute:*

$$\begin{tikzcd} G \times G \times G \arrow[r, "1_G \times \mu"] \arrow[d, "\mu \times 1_G"] & G \times G \arrow[d, "\mu"] \\ G \times G \arrow[r, "\mu"] & G \end{tikzcd} \tag{1.1.1.2}$$

$$\begin{tikzcd} & G \times G \arrow[d, "\mu"] & \\ \{\mathrm{pt}\} \times G \arrow[ur, "e\times 1_G"] \arrow[r, "\sim"'] & G & G \times \{\mathrm{pt}\} \arrow[ul, "1_G\times e"'] \arrow[l, "\sim"] \end{tikzcd} \tag{1.1.1.3}$$

$$\begin{tikzcd} & G \times G \arrow[r, "1_G\times i"] & G \times G \arrow[dr, "\mu"] & \\ G \arrow[ur, "\Delta"] \arrow[rr] \arrow[dr, "\Delta"'] & & \{\mathrm{pt}\} \arrow[r, "e"] & G \\ & G \times G \arrow[r, "i\times 1_G"'] & G \times G \arrow[ur, "\mu"'] & \end{tikzcd} \tag{1.1.1.4}$$

In equation (1.1.1.3), the isomorphisms are the canonical ones. In equation (1.1.1.4), the map $G \to \{\mathrm{pt}\}$ *is the unique map to the terminal object, and* $\Delta : G \to G \times G$ *is the canonical diagonal map.*

If (G, μ_G, e_G, i_G) *and* (H, μ_H, e_H, i_H) *are two group objects, a map* $f : G \to H$ *is a* group object homomorphism *if the following commute:*

$$\begin{tikzcd} G \times G \arrow[r, "\mu_G"] \arrow[d, "f\times f"] & G \arrow[d, "f"] \\ H \times H \arrow[r, "\mu_H"] & H \end{tikzcd} \qquad \begin{tikzcd} & G \arrow[dd, "f"] \\ \{\mathrm{pt}\} \arrow[ur, "e_G"] \arrow[dr, "e_H"'] & \\ & H \end{tikzcd}$$

(That f *intertwines* i_G *with* i_H *is then a corollary.)*

1.1.1.5 Definition *A* (left) group action *of a group object* G *in a category* $\mathcal{C}$ *with finite products is a map* $\rho : G \times X \to X$ *such that the following diagrams commute:*

$$\begin{tikzcd} G \times G \times X \arrow[r, "1_G\times\rho"] \arrow[d, "\mu\times 1_X"] & G \times X \arrow[d, "\rho"] \\ G \times X \arrow[r, "\rho"] & X \end{tikzcd} \tag{1.1.1.6}$$

$$\begin{tikzcd} & G \times X \arrow[d, "\rho"] \\ \{\mathrm{pt}\} \times X \arrow[ur, "e\times 1_X"] \arrow[r, "\sim"'] & X \end{tikzcd} \tag{1.1.1.7}$$

(The diagram corresponding to equation (1.1.1.4) is then a corollary.) A right action *is a map* $X \times G \to X$ *with similar diagrams. We denote a left group action* $\rho : G \times X \to X$ *by* $\rho : G \curvearrowright X$*.*

Let $\rho_X : G \times X \to X$ *and* $\rho_Y : G \times Y \to Y$ *be two group actions. A map* $f : X \to Y$ *is* G-equivariant *if the following square commutes:*

$$\begin{tikzcd} G \times X \arrow[r, "\rho_X"] \arrow[d, "1_G \times f"] & X \arrow[d, "f"] \\ G \times Y \arrow[r, "\rho_Y"] & Y \end{tikzcd} \tag{1.1.1.8}$$

1.1.2 Analytic and algebraic groups

1.1.2.1 Definition *A* Lie group *is a group object in a category of manifolds. In particular, a Lie group can be infinitely differentiable (in the category* $\mathscr{C}^\infty$-MAN*) or* analytic *(in the category* $\mathscr{C}^\omega$-MAN*) when over* $\mathbb{R}$*, or* complex analytic *or* almost complex *when over* $\mathbb{C}$*. We will take "Lie group" to mean* analytic Lie group *over either* $\mathbb{C}$ *or* $\mathbb{R}$*. In fact, the different notions of real Lie group coincide, a fact that we will not directly prove, as do the different notions of complex Lie group. As always, we will use the word "smooth" for any of "infinitely differentiable", "analytic", or "holomorphic".*

A Lie action *is a group action in the category of manifolds.*

A (linear) algebraic group *over* $\mathbb{K}$ *(algebraically closed) is a group object in the category of (affine) algebraic varieties over* $\mathbb{K}$*.*

1.1.2.2 Example The *general linear group* $\mathrm{GL}(n, \mathbb{K})$ of $n \times n$ invertible matrices is a Lie group over $\mathbb{K}$ for $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. When $\mathbb{K}$ is algebraically closed, $\mathrm{GL}(n, \mathbb{K})$ is an algebraic group. It acts algebraically on $\mathbb{K}^n$ and on projective space $\mathbb{P}(\mathbb{K}^n) = \mathbb{P}^{n-1}(\mathbb{K})$. ◊

1.2 Definition of a closed linear group

We write $\mathrm{GL}(n, \mathbb{K})$ for the group of $n \times n$ invertible matrices over $\mathbb{K}$, and $\mathrm{Mat}(n, \mathbb{K})$ for the algebra of all $n \times n$ matrices. We regularly leave off the $\mathbb{K}$.

1.2.0.1 Definition *A* closed linear group *is a subgroup of* $\mathrm{GL}(n)$ *(over* $\mathbb{C}$ *or* $\mathbb{R}$*) that is closed as a topological subspace.*

1.2.1 Lie algebra of a closed linear group

1.2.1.1 Lemma / Definition *The following describe the same function* $\exp : \mathrm{Mat}(n) \to \mathrm{GL}(n)$*, called the* matrix exponential*.*

1. $\exp(a) \stackrel{\text{def}}{=} \sum_{n \geq 0} \frac{a^n}{n!}$.

2. $\exp(a) \stackrel{\text{def}}{=} e^{ta}\big|_{t=1}$, *where for fixed* $a \in \mathrm{Mat}(n)$ *we define* e^{ta} *as the solution to the initial value problem* $e^{0a} = 1$, $\frac{d}{dt}e^{ta} = ae^{t}a$.

3. $\exp(a) \stackrel{\text{def}}{=} \lim_{n\to\infty} \left(1 + \frac{a}{n}\right)^n$.

If $ab - ba = 0$*, then* $\exp(a + b) = \exp(a) + \exp(b)$*.*

The function $\exp : \mathrm{Mat}(n) \to \mathrm{GL}(n)$ *is a local isomorphism of analytic manifolds. In a neighborhood of* $1 \in \mathrm{GL}(n)$*, the function* $\log a \stackrel{\text{def}}{=} -\sum_{n>0} \frac{(1-a)^n}{n}$ *is an inverse to* $\exp$*.* □

1.2.1.2 Lemma / Definition *Let* H *be a closed linear group. The* Lie algebra of H *is the set*

$$\mathrm{Lie}(H) = \{x \in \mathrm{Mat}(n) : \exp(\mathbb{R}x) \subseteq H\}$$

1. $\mathrm{Lie}(H)$ *is a* $\mathbb{R}$*-subspace of* $\mathrm{Mat}(n)$*.*

2. $\mathrm{Lie}(H)$ *is closed under the* bracket $[,] : (a, b) \mapsto ab - ba$*.* □

1.2.1.3 Definition *A* Lie algebra *over* $\mathbb{K}$ *is a vector space* $\mathfrak{g}$ *along with an antisymmetric map* $[,] : \mathfrak{g} \otimes \mathfrak{g} \to \mathfrak{g}$ *satisfying the* Jacobi identity*:*

$$[[a, b], c] + [[b, c], a] + [[c, a], b] = 0$$

A homomorphism of Lie algebras *is a linear map preserving the bracket. A* Lie subalgebra *is a vector subspace closed under the bracket.*

1.2.1.4 Example The algebra $\mathfrak{gl}(n) = \mathrm{Mat}(n)$ of $n\times n$ matrices is a Lie algebra with $[a, b] = ab-ba$. It is $\mathrm{Lie}(\mathrm{GL}(n))$. Lemma/Definition 1.2.1.2 says that $\mathrm{Lie}(H)$ is a Lie subalgebra of $\mathrm{Mat}(n)$. ◊

1.2.2 Some analysis

1.2.2.1 Lemma *Let* $\mathrm{Mat}(n) = V \oplus W$ *as a real vector space. Then there exists an open neighborhood* $U \ni 0$ *in* $\mathrm{Mat}(n)$ *and an open neighborhood* $U' \ni 1$ *in* $\mathrm{GL}(n)$ *such that*

$$(v, w) \mapsto \exp(v) \exp(w) : V \oplus W \to \mathrm{GL}(n)$$

is a homeomorphism $U \to U'$*.* □

1.2.2.2 Lemma *Let* H *be a closed subgroup of* $\mathrm{GL}(n)$*, and* $W \subseteq \mathrm{Mat}(n)$ *be a linear subspace such that* 0 *is a limit point of the set* $\{w \in W \text{ s.t. } \exp(w) \in H\}$*. Then* $W \cap \mathrm{Lie}(H) \neq 0$*.*

Proof Fix a Euclidian norm on W. Let $w_1, w_2, \dots \to 0$ be a sequence in $\{w \in W \text{ s.t. } \exp(w) \in H\}$, with $w_i \neq 0$. Then $w_i/|w_i|$ are on the unit sphere, which is compact, so passing to a subsequence, we can assume that $w_i/|w_i| \to x$ where x is a unit vector. The norms $|w_i|$ are tending to 0, so $w_i/|w_i|$ is a large multiple of w_i. We approximate this: let $n_i = \lceil 1/|w_i| \rceil$, whence $n_i w_i \approx w_i/|w_i|$, and $n_i w_i \to x$. But $\exp w_i \in H$, so $\exp(n_i w_i) \in H$, and H is a closed subgroup, so $\exp x \in H$.

Repeating the argument with a ball of radius r to conclude that $\exp(rx)$ is in H, we conclude that $x \in \mathrm{Lie}(H)$. □

1.2.2.3 Proposition *Let H be a closed subgroup of* $\mathrm{GL}(n)$*. There exist neighborhoods* $0 \in U \subseteq \mathrm{Mat}(n)$ *and* $1 \in U' \subseteq \mathrm{GL}(n)$ *such that* $\exp : U \xrightarrow{\sim} U'$ *takes* $\mathrm{Lie}(H) \cap U \xrightarrow{\sim} H \cap U'$*.*

Proof We fix a complement $W \subseteq \mathrm{Mat}(n)$ such that $\mathrm{Mat}(n) = \mathrm{Lie}(H) \oplus W$. By Lemma 1.2.2.2, we can find a neighborhood $V \subseteq W$ of 0 such that $\{v \in V \text{ s.t. } \exp(v) \in H\} = \{0\}$. Then on $\mathrm{Lie}(H) \times V$, the map $(x, w) \mapsto \exp(x)\exp(w)$ lands in H if and only if $w = 0$. By restricting the first component to lie in an open neighborhood, we can approximate $\exp(x + w) \approx \exp(x)\exp(w)$ as well as we need to — there's a change of coordinates that completes the proof. □

1.2.2.4 Corollary *H is a submanifold of* $\mathrm{GL}(n)$ *of dimension equal to the dimension of* $\mathrm{Lie}(H)$*.* □

1.2.2.5 Corollary $\exp(\mathrm{Lie}(H))$ *generates the identity component H_0 of H.* □

1.2.2.6 Remark In any topological group, the connected component of the identity is normal. ♢

1.2.2.7 Corollary $\mathrm{Lie}(H)$ *is the tangent space* $\mathrm{T}_1 H \stackrel{\text{def}}{=} \{\gamma'(0) \text{ s.t. } \gamma : \mathbb{R} \to H, \gamma(0) = 1\} \subseteq \mathrm{Mat}(n)$*.* □

1.3 Classical Lie groups

We mention only the classical compact semisimple Lie groups and the classical complex semisimple Lie groups. There are other very interesting classical Lie groups, c.f. [Lan85].

1.3.1 Classical compact Lie groups

1.3.1.1 Lemma / Definition *The* quaternions $\mathbb{H}$ *is the unital $\mathbb{R}$-algebra generated by i, j, k with the multiplication $i^2 = j^2 = k^2 = ijk = -1$; it is a non-commutative division ring. Then $\mathbb{R} \hookrightarrow \mathbb{C} \hookrightarrow \mathbb{H}$, and $\mathbb{H}$ is a subalgebra of* $\mathrm{Mat}(4, \mathbb{R})$*. We defined the* complex conjugate *linearly by $\bar{i} = -i$, $\bar{j} = -j$, and $\bar{k} = -k$; complex conjugation is an anti-automorphism, and the fixed-point set is $\mathbb{R}$. The* Euclidean norm *of $\zeta \in \mathbb{H}$ is given by $\|\zeta\| = \bar{\zeta}\zeta$.*

The Euclidean norm *of a column vector $x \in \mathbb{R}^n, \mathbb{C}^n, \mathbb{H}^n$ is given by $\|x\|^2 = \bar{x}^T x$, where $\bar{x}$ is the component-wise complex conjugation of x.*

If $x \in \mathrm{Mat}(n, \mathbb{R}), \mathrm{Mat}(n, \mathbb{C}), \mathrm{Mat}(n, \mathbb{H})$ is a matrix, we define its Hermitian conjugate *to be the matrix $x^* = \bar{x}^T$; Hermitian conjugation is an antiautomorphism of algebras* $\mathrm{Mat}(n) \to \mathrm{Mat}(n)$*.* $\mathrm{Mat}(n, \mathbb{H}) \hookrightarrow \mathrm{Mat}(2n, \mathbb{C})$ *is a $*$-embedding.*

Let $j = \left(\begin{array}{c|c} 0 & -1 \\ \hline 1 & 0 \end{array}\right) \in \mathrm{Mat}(2, \mathrm{Mat}(n, \mathbb{C})) = \mathrm{Mat}(2n, \mathbb{C})$ *be a block matrix. We define* $\mathrm{GL}(n, \mathbb{H}) \stackrel{\text{def}}{=} \{x \in \mathrm{GL}(2n, \mathbb{C}) \text{ s.t. } jx = \bar{x}j\}$*. It is a closed linear group.* □

1.3.1.2 Lemma / Definition *The following are closed linear groups, and are compact:*

- *The* (real) special orthogonal group $\mathrm{SO}(n, \mathbb{R}) \stackrel{\text{def}}{=} \{x \in \mathrm{Mat}(n, \mathbb{R}) \text{ s.t. } x^* x = 1 \text{ and } \det x = 1\}$*.*
- *The* (real) orthogonal group $\mathrm{O}(n, \mathbb{R}) \stackrel{\text{def}}{=} \{x \in \mathrm{Mat}(n, \mathbb{R}) \text{ s.t. } x^* x = 1\}$*.*

- *The* special unitary group $\mathrm{SU}(n) \stackrel{\text{def}}{=} \{x \in \mathrm{Mat}(n, \mathbb{C}) \text{ s.t. } x^*x = 1 \text{ and } \det x = 1\}$.
- *The* unitary group $\mathrm{U}(n) \stackrel{\text{def}}{=} \{x \in \mathrm{Mat}(n, \mathbb{C}) \text{ s.t. } x^*x = 1\}$.
- *The* (compact) symplectic group $\mathrm{Sp}(n, \mathbb{R}) \stackrel{\text{def}}{=} \{x \in \mathrm{Mat}(n, \mathbb{H}) \text{ s.t. } x^*x = 1\}$. □

There is no natural quaternionic determinant.

1.3.2 Classical complex Lie groups

The following groups make sense over any field, but it's best to work over an algebraically closed field. We work over $\mathbb{C}$.

1.3.2.1 Lemma / Definition *The following are closed linear groups over $\mathbb{C}$, and are algebraic:*

- *The* (complex) special linear group $\mathrm{SL}(n, \mathbb{C}) \stackrel{\text{def}}{=} \{x \in \mathrm{GL}(n, \mathbb{C}) \text{ s.t. } \det x = 1\}$.
- *The* (complex) special orthogonal group $\mathrm{SO}(n, \mathbb{C}) \stackrel{\text{def}}{=} \{x \in \mathrm{SL}(n, \mathbb{C}) \text{ s.t. } x^T x = 1\}$.
- *The* (complex) symplectic group $\mathrm{Sp}(n, \mathbb{C}) \stackrel{\text{def}}{=} \{x \in \mathrm{GL}(2n, \mathbb{C}) \text{ s.t. } x^T jx = j\}$. □

1.3.3 The classical groups

In full, we have defined the following "classical" closed linear groups:

	Group Name	**Group Description**	**Algebra Name**	**Algebra Description**	$\dim_{\mathbb{R}}$
Compact	$\mathrm{SO}(n, \mathbb{R})$	$\{x \in \mathrm{Mat}(n, \mathbb{R}) \text{ s.t. } x^*x = 1, \det x = 1\}$	$\mathfrak{so}(n, \mathbb{R})$	$\{x \in \mathrm{Mat}(n, \mathbb{R}) \text{ s.t. } x^* + x = 0\}$	$\binom{n}{2}$
	$\mathrm{SU}(n)$	$\{x \in \mathrm{Mat}(n, \mathbb{C}) \text{ s.t. } x^*x = 1, \det x = 1\}$	$\mathfrak{su}(n)$	$\{x \in \mathrm{Mat}(n, \mathbb{C}) \text{ s.t. } x^* + x = 0, \mathrm{tr}\, x = 0\}$	$n^2 - 1$
	$\mathrm{Sp}(n, \mathbb{R})$	$\{x \in \mathrm{Mat}(n, \mathbb{H}) \text{ s.t. } x^*x = 1\}$	$\mathfrak{sp}(n, \mathbb{R})$	$\{x \in \mathrm{Mat}(n, \mathbb{H}) \text{ s.t. } x^* + x = 0\}$	$2n^2 + n$
	$\mathrm{GL}(n, \mathbb{H})$	$\{x \in \mathrm{GL}(2n, \mathbb{C}) \text{ s.t. } jx = \bar{x}j\}$	$\mathfrak{gl}(n, \mathbb{H})$	$\{x \in \mathrm{Mat}(2n, \mathbb{C}) \text{ s.t. } jx = \bar{x}j\}$	$4n^2$
Complex	$\mathrm{SL}(n, \mathbb{C})$	$\{x \in \mathrm{Mat}(n, \mathbb{C}) \text{ s.t. } \det x = 1\}$	$\mathfrak{sl}(n, \mathbb{C})$	$\{x \in \mathrm{Mat}(n, \mathbb{C}) \text{ s.t. } \mathrm{tr}\, x = 0\}$	$2(n^2 - 1)$
	$\mathrm{SO}(n, \mathbb{C})$	$\{x \in \mathrm{Mat}(n, \mathbb{C}) \text{ s.t. } x^T x = 1, \det x = 1\}$	$\mathfrak{so}(n, \mathbb{C})$	$\{x \in \mathrm{Mat}(n, \mathbb{C}) \text{ s.t. } x^T + x = 0\}$	$n(n-1)$
	$\mathrm{Sp}(n, \mathbb{C})$	$\{x \in \mathrm{Mat}(n, \mathbb{C}) \text{ s.t. } x^T jx = j\}$	$\mathfrak{sp}(n, \mathbb{C})$	$\{x \in \mathrm{Mat}(n, \mathbb{C}) \text{ s.t. } x^T j + jx = 0\}$	$2\binom{2n+1}{2}$

1.3.3.1 Proposition *Via the natural embedding* $\mathrm{Mat}(n, \mathbb{H}) \hookrightarrow \mathrm{Mat}(2n, \mathbb{C})$, *we have:*

$$\mathrm{Sp}(n) = \mathrm{GL}(n, \mathbb{H}) \cap \mathrm{U}(2n) \tag{1.3.3.2}$$
$$= \mathrm{GL}(n, \mathbb{H}) \cap \mathrm{Sp}(n, \mathbb{C}) \tag{1.3.3.3}$$
$$= \mathrm{U}(2n) \cap \mathrm{Sp}(n, \mathbb{C}) \tag{1.3.3.4}$$

1.4 Homomorphisms of closed linear groups

1.4.0.1 Definition *Let H be a closed linear group. The* adjoint action $H \curvearrowright H$ *is given by by* ${}^g h \stackrel{\text{def}}{=} ghg^{-1}$*, and this action fixes* $1 \in H$*. This induces the* adjoint action $\mathrm{Ad} : H \curvearrowright \mathrm{T}_1 H = \mathrm{Lie}(H)$*. It is given by $g \cdot y = gyg^{-1}$, where now $y \in \mathrm{Lie}(H)$.*

1.4.0.2 Lemma *Let H and G be closed linear groups and $\phi : H \to G$ a smooth homomorphism. Then $\phi(1) = 1$, so $\mathrm{d}\phi : \mathrm{T}_1 H \to \mathrm{T}_1 G$ by $X \mapsto (\phi(1 + tX))'(0)$. The diagram of actions commutes:*

$$\begin{tikzcd} H \arrow[r, phantom, "\curvearrowright"] \arrow[d, "\phi"] & \mathrm{T}_1 H \arrow[d, "\mathrm{d}\phi"] \\ G \arrow[r, phantom, "\curvearrowright"] & \mathrm{T}_1 G \end{tikzcd}$$

This is to say:

$$\mathrm{d}\phi(\mathrm{Ad}(h)Y) = \mathrm{Ad}(\phi(h))\,\mathrm{d}\phi(Y)$$

Thus $\mathrm{d}\phi\,[X, Y] = [\mathrm{d}\phi\, X, \mathrm{d}\phi\, Y]$*, so $d\phi$ is a* Lie algebra homomorphism*.*

If H is connected, the map $\mathrm{d}\phi$ determines the map ϕ. □

Exercises

1. (a) Show that the orthogonal groups $\mathrm{O}(n, \mathbb{R})$ and $\mathrm{O}(n, \mathbb{C})$ have two connected components, the identity component being the special orthogonal group SO_n, and the other consisting of orthogonal matrices of determinant -1.
   (b) Show that the center of $\mathrm{O}(n)$ is $\{\pm I_n\}$.
   (c) Show that if n is odd, then $\mathrm{SO}(n)$ has trivial center and $\mathrm{O}(n) \cong \mathrm{SO}(n) \times (\mathbb{Z}/2\mathbb{Z})$ as a Lie group.
   (d) Show that if n is even, then the center of $\mathrm{SO}(n)$ has two elements, and $\mathrm{O}(n)$ is a semidirect product $(\mathbb{Z}/2\mathbb{Z}) \ltimes \mathrm{SO}(n)$, where $\mathbb{Z}/2\mathbb{Z}$ acts on $\mathrm{SO}(n)$ by a non-trivial outer automorphism of order 2.

2. Construct a smooth group homomorphism $\Phi : \mathrm{SU}(2) \to \mathrm{SO}(3)$ which induces an isomorphism of Lie algebras and identifies $\mathrm{SO}(3)$ with the quotient of $\mathrm{SU}(2)$ by its center $\{\pm I\}$.

3. Construct an isomorphism of $\mathrm{GL}(n, \mathbb{C})$ (as a Lie group and an algebraic group) with a closed subgroup of $\mathrm{SL}(n + 1, \mathbb{C})$.

4. Show that the map $\mathbb{C}^* \times \mathrm{SL}(n, \mathbb{C}) \to \mathrm{GL}(n, \mathbb{C})$ given by $(z, g) \mapsto zg$ is a surjective homomorphism of Lie and algebraic groups, find its kernel, and describe the corresponding homomorphism of Lie algebras.

5. Find the Lie algebra of the group $U \subseteq \mathrm{GL}(n, \mathbb{C})$ of upper-triangular matrices with 1 on the diagonal. Show that for this group, the exponential map is a diffeomorphism of the Lie algebra onto the group.

6. A *real form* of a complex Lie algebra $\mathfrak{g}$ is a real Lie subalgebra $\mathfrak{g}_{\mathbb{R}}$ such that that $\mathfrak{g} = \mathfrak{g}_{\mathbb{R}} \oplus i\mathfrak{g}_{\mathbb{R}}$, or equivalently, such that the canonical map $\mathfrak{g}_{\mathbb{R}} \otimes_{\mathbb{R}} \mathbb{C} \to \mathfrak{g}$ given by scalar multiplication is an isomorphism. A real form of a (connected) complex closed linear group G is a (connected) closed real subgroup $G_{\mathbb{R}}$ such that $\mathrm{Lie}(G_{\mathbb{R}})$ is a real form of $\mathrm{Lie}(G)$.

   (a) Show that $\mathrm{U}(n)$ is a compact real form of $\mathrm{GL}(n, \mathbb{C})$ and $\mathrm{SU}(n)$ is a compact real form of $\mathrm{SL}(n, \mathbb{C})$.

   (b) Show that $\mathrm{SO}(n, \mathbb{R})$ is a compact real form of $\mathrm{SO}(n, \mathbb{C})$.

   (c) Show that $\mathrm{Sp}(n, \mathbb{R})$ is a compact real form of $\mathrm{Sp}(n, \mathbb{C})$.

Chapter 2

Mini-course in Differential Geometry

2.1 Manifolds

2.1.1 Classical definition

2.1.1.1 Definition *Let X be a (Hausdorff) topological space. A* chart *consists of the data* $U \underset{open}{\subseteq} X$ *and a homeomorphism* $\phi : U \xrightarrow{\sim} V \underset{open}{\subseteq} \mathbb{R}^n$. $\mathbb{R}^n$ *has* coordinates x_i, *and* $\xi_i \overset{\text{def}}{=} x_i \circ \phi$ *are* local coordinates *on the chart. Charts* (U, ϕ) *and* (U', ϕ') *are* compatible *if on* $U \cap U'$ *the* ξ'_i *are smooth functions of the* ξ_i *and conversely. I.e.:*

$$
\begin{tikzcd}
U \arrow[d, "\phi"] & & & U \cap U' \arrow[dl, "\bar\phi"'] \arrow[dr, "\bar\phi'"] & & & U' \arrow[d, "\phi'"] \\
V & V \supseteq & W \arrow[rr, "\bar\phi' \circ \bar\phi^{-1}", "\text{smooth with smooth inverse}"'] & & W' & \subseteq V\text{'} & V'
\end{tikzcd}
\tag{2.1.1.2}
$$

An atlas *on X is a covering by pairwise compatible charts.*

2.1.1.3 Lemma *If U and U' are compatible with all charts of $\mathcal{A}$, then they are compatible with each other.* $\square$

2.1.1.4 Corollary *Every atlas has a unique maximal extension.* $\square$

2.1.1.5 Definition *A* manifold *is a Hausdorff topological space with a maximal atlas. It can be* real, infinitely-differentiable, complex, analytic, *etc., by varying the word "smooth" in the compatibility condition equation (2.1.1.2).*

2.1.1.6 Definition *Let U be an open subset of a manifold X. A function $f : U \to \mathbb{R}$ is* smooth *if it is smooth on local coordinates in all charts. We will write $\mathscr{C}(U)$ for the space of smooth functions on U.*

Our general convention will be to use the word "smooth" as a placeholder for any of "infinitely-differentiable," "analytic," or "holomorphic," and use the symbol $\mathscr{C}(-)$ for any of these notions. The result of this is that if we prove a statement referencing "smooth" geometry, we will in fact simultaneously prove three statements, one for the infinitely-differentiable category, one for the analytic category, and one for the holomorphic category.

2.1.2 Sheafs

2.1.2.1 Definition *A* sheaf of functions $\mathscr{S}$ *on a topological space* X *assigns a ring* $\mathscr{S}(U)$ *to each open set* $U \underset{\text{open}}{\subseteq} X$ *such that:*

1. *if* $V \subseteq U$ *and* $f \in \mathscr{S}(U)$*, then* $f|_V \in \mathscr{S}(V)$*, and*
2. *if* $U = \bigcup_\alpha U_\alpha$ *and* $f : U \to \mathbb{R}$ *such that* $f|_{U_\alpha} \in \mathscr{S}(U_\alpha)$ *for each* α*, then* $f \in \mathscr{S}(U)$*.*

The stalk *of a sheaf at* $x \in X$ *is the space* $\mathcal{S}_x \stackrel{\text{def}}{=} \lim_{U \ni x} \mathscr{S}(U)$*.*

2.1.2.2 Lemma *Let* X *be a manifold, and assign to each* $U \underset{\text{open}}{\subseteq} X$ *the ring* $\mathscr{C}(U)$ *of smooth functions on* U*. Then* $\mathscr{C}$ *is a sheaf. Conversely, a topological space* X *with a sheaf of functions* $\mathscr{S}$ *is a manifold if and only if there exists a covering of* X *by open sets* U *such that* $(U, \mathscr{S}|_U)$ *is isomorphic as a space with a sheaf of functions to* $(V, \mathscr{S}^{\mathbb{R}^n}|_V)$ *for some* $V \subseteq \mathbb{R}^n$ *open.* □

2.1.3 Manifold constructions

2.1.3.1 Definition *If* X *and* Y *are smooth manifolds, then a* smooth map $f : X \to Y$ *is a continuous map such that for all* $U \subseteq Y$ *and* $g \in \mathscr{C}(U)$*, then* $g \circ f \in \mathscr{C}(f^{-1}(U))$*. Manifolds form a category* Man *with products: a* product of manifolds $X \times Y$ *is a manifold with charts* $U \times V$*.*

2.1.3.2 Definition *Let* M *be a manifold,* $p \in M$ *a point, and* $\gamma_1, \gamma_2 : \mathbb{R} \to M$ *two paths with* $\gamma_1(0) = \gamma_2(0) = p$*. We say that* γ_1 *and* γ_2 *are* tangent at p *if* $(f \circ \gamma_1)'(0) = (f \circ \gamma_2)'(0)$ *for all smooth* f *on a nbhd of* p*, i.e. for all* $f \in \mathscr{C}_p$*. Each equivalence class of tangent curves is called a* tangent vector*.*

2.1.3.3 Definition *Let* M *be a manifold and* $\mathscr{C}$ *its sheaf of smooth functions. A* point derivation *is a linear map* $\delta : \mathscr{C}_p \to \mathbb{R}$ *satisfying the* Leibniz rule*:*

$$\delta(fg) = \delta f\, g(p) + f(p)\, \delta g$$

2.1.3.4 Lemma *Any tangent vector* γ *gives a point derivation* $\delta_\gamma : f \mapsto (f \circ \gamma)'(0)$*. Conversely, every point derivation is of this form.* □

2.1.3.5 Lemma / Definition *Let* M *and* N *be manifolds, and* $f : M \to N$ *a smooth map sending* $p \mapsto q$*. The following are equivalent, and define* $(\mathrm{d}f)_p : \mathrm{T}_p M \to \mathrm{T}_q N$*, the* differential *of* f *at* p*:*

1. *If* $[\gamma] \in \mathrm{T}_p M$ *is represented by the curve* γ*, then* $(\mathrm{d}f)_p(X) \stackrel{\text{def}}{=} [f \circ \gamma]$*.*

2. *If* $X \in \mathrm{T}_p M$ *is a point-derivation on* $\mathcal{S}_{M,p}$*, then* $(\mathrm{d}f)_p(X) : \mathcal{S}_{N,q} \to \mathbb{R}$ *(or* $\mathbb{C}$*) is defined by* $\psi \mapsto X[\psi \circ f]$*.*

3. *In coordinates,* $p \in U \underset{open}{\subseteq} \mathbb{R}^m$ *and* $q \in W \underset{open}{\subseteq} \mathbb{R}^n$*, then locally* f *is given by* $f_1, \dots, f_n$ *smooth functions of* $x_1, \dots, x_m$*. The tangent spaces to* $\mathbb{R}^n$ *are in canonical bijection with* $\mathbb{R}^n$*, and a linear map* $\mathbb{R}^m \to \mathbb{R}^n$ *should be presented as a matrix:*

$$\mathrm{Jacobian}(f, x) \stackrel{\text{def}}{=} \frac{\partial f_i}{\partial x_j} \qquad \square$$

2.1.3.6 Lemma *We have the* chain rule*: if* $M \xrightarrow{f} N \xrightarrow{g} K$*, then* $\mathrm{d}(g \circ f)_p = (\mathrm{d}g)_{f(p)} \circ (\mathrm{d}f)_p$*.* $\square$

2.1.3.7 Theorem (Inverse Mapping Theorem)

1. *Given smooth* $f_1, \dots, f_n : U \to \mathbb{R}$ *where* $p \in U \underset{open}{\subseteq} \mathbb{R}^n$*, then* $f : U \to \mathbb{R}^n$ *maps some neighborhood* $V \ni p$ *bijectively to* $W \underset{open}{\subseteq} \mathbb{R}^n$ *with s/a/h inverse iff* $\det \mathrm{Jacobian}(f, x) \neq 0$*.*

2. *A smooth map* $f : M \to N$ *of manifold restricts to an isomorphism* $p \in U \to W$ *for some neighborhood* U *if and only if* $(df)_p$ *is a linear isomorphism.* $\square$

2.1.4 Submanifolds

2.1.4.1 Proposition *Let* M *be a manifold and* N *a topological subspace with the induced topology such that for each* $p \in N$*, there is a chart* $U \ni p$ *in* M *with coordinates* $\{\xi_i\}_{i=1}^m : U \to \mathbb{R}^m$ *such that* $U \cap N = \{q \in U \text{ s.t. } \xi_{n+1}(q) = \dots = \xi_m(q) = 0\}$*. Then* $U \cap N$ *is a chart on* N *with coordinates* $\xi_1, \dots, \xi_n$*, and* N *is a manifold with an atlas given by* $\{U \cap N\}$ *as* U *ranges over an atlas of* M*. The sheaf of smooth functions* $\mathscr{C}_N$ *is the sheaf of continuous functions on* N *that are locally restrictions of smooth functions on* M*. The embedding* $N \hookrightarrow M$ *is smooth, and satisfies the universal property that any smooth map* $f : Z \to M$ *such that* $f(Z) \subseteq N$ *defines a smooth map* $Z \to N$*.* $\square$

2.1.4.2 Definition *The map* $N \hookrightarrow M$ *in Proposition 2.1.4.1 is an* immersed submanifold*. A map* $Z \to M$ *is an* immersion *if it factors as* $Z \xrightarrow{\sim} N \hookrightarrow M$ *for some immersed submanifold* $N \hookrightarrow M$*.*

2.1.4.3 Proposition *If* $N \hookrightarrow M$ *is an immersed submanifold, then* N *is locally closed.* $\square$

2.1.4.4 Proposition *Any closed linear group* $H \subseteq \mathrm{GL}(n)$ *is an immersed analytic submanifold. If* $\mathrm{Lie}(H)$ *is a* $\mathbb{C}$*-subspace of* $\mathrm{Mat}(n, \mathbb{C})$*, then* H *is a holomorphic submanifold.*

Proof The following diagram defines a chart near $1 \in H$, which can be moved by left-multiplication wherever it is needed:

$$\begin{tikzcd}
\mathrm{M}(n) \supseteq \underset{}{\overset{0}{\overset{\cap}{U}}} \arrow[r, shift left, "\exp"] & \overset{1}{\overset{\cap}{V}} \subseteq \mathrm{GL}(n) \arrow[l, shift left, "\log"] \\
\mathrm{Lie}(H) \cap U \arrow[u, hook] \arrow[r, shift left] & H \cap V \arrow[u, hook] \arrow[l, shift left]
\end{tikzcd} \qquad \square$$

2.1.4.5 Lemma *Given* $\mathrm{T}_p M = V_1 \oplus V_2$*, there is an open neighborhood* $U_1 \times U_2$ *of* p *such that* $V_i = \mathrm{T}_p U_i$. $\square$

2.1.4.6 Lemma *If* $s : N \to M \times N$ *is a s/a/h section, then* s *is a (closed) immersion.* $\square$

2.1.4.7 Proposition *A smooth map* $f : N \to M$ *is an immersion on a neighborhood of* $p \in N$ *if and only if* $(df)_p$ *is injective.* $\square$

2.2 Vector Fields

2.2.1 Definition

2.2.1.1 Definition *Let* M *be a manifold. A* vector field *assigns to each* $p \in M$ *a vector* x_p*, i.e. a point derivation:*

$$x_p(fg) = f(p)\, x_p(g) + x_p(f)\, g(p)$$

We define $(xf)(p) \stackrel{\text{def}}{=} x_p(f)$*. Then* $x(fg) = f\,x(g) + x(f)\,g$*, so* x *is a* derivation*. But it might be discontinuous. A vector field* x *is* smooth *if* $x : \mathscr{C}_M \to \mathscr{C}_M$ *is a map of sheaves. Equivalently, in local coordinates the components of* x_p *must depend smoothly on* p*. By changing (the conditions on) the sheaf* $\mathscr{C}$*, we may define analytic or holomorphic vector fields.*

Henceforth, the word "vector field" will always mean "smooth (or analytic or holomorphic) vector field". Similarly, we will use the word "smooth" to mean smooth or analytic or holomorphic, depending on our category.

2.2.1.2 Lemma *The commutator* $[x, y] \stackrel{\text{def}}{=} xy - yx$ *of derivations is a derivation.*

Proof An easy calculation:

$$xy(fg) = xy(f)\,g + x(f)\,y(g) + y(f)\,x(g) + f\,xy(g) \tag{2.2.1.3}$$

Switch X and Y, and subtract:

$$[x, y](fg) = [x, y](f)\,g + f\,[x, y](g) \tag{2.2.1.4}$$

$\square$

2.2.1.5 Definition *A* Lie algebra *is a vector space* $\mathfrak{l}$ *with a bilinear map* $[,] : \mathfrak{l} \times \mathfrak{l} \to \mathfrak{l}$ *(i.e. a linear map* $[,] : \mathfrak{l} \otimes \mathfrak{l} \to \mathfrak{l}$*), satisfying*

1. *Antisymmetry:* $[x, y] + [y, x] = 0$
2. *Jacobi:* $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$

2.2.1.6 Proposition *Let* V *be a vector space. The bracket* $[x, y] \stackrel{\text{def}}{=} xy - yx$ *makes* $\mathrm{End}(V)$ *into a Lie algebra.* $\square$

2.2.1.7 Lemma / Definition *Let $\mathfrak{l}$ be a Lie algebra. The* adjoint action $\mathrm{ad} : \mathfrak{l} \to \mathrm{End}(\mathfrak{l})$ *given by* $\mathrm{ad}\, x : y \mapsto [x, y]$ *is a* derivation*:*

$$(\mathrm{ad}\, x)[y, z] = [(\mathrm{ad}\, x)y, z] + [y, (\mathrm{ad}\, x)z]$$

Moreover, $\mathrm{ad} : \mathfrak{l} \to \mathrm{End}(\mathfrak{l})$ *is a* Lie algebra homomorphism*:*

$$\mathrm{ad}([x, y]) = (\mathrm{ad}\, x)(\mathrm{ad}\, y) - (\mathrm{ad}\, y)(\mathrm{ad}\, x)$$ □

2.2.2 Integral curves

Let ∂_t be the vector field $f \mapsto \frac{\mathrm{d}}{\mathrm{d}t} f$ on $\mathbb{R}$.

2.2.2.1 Proposition *Given a smooth vector field x on M and a point $p \in M$, there exists an open interval $I \underset{open}{\subseteq} \mathbb{R}$ such that $0 \in I$ and a smooth curve $\gamma : I \to M$ satisfying:*

$$\gamma(0) = p \tag{2.2.2.2}$$

$$(d\gamma)_t(\partial_t) = x_{\gamma(t)}\, \forall t \in I \tag{2.2.2.3}$$

When M is a complex manifold and x a holomorphic vector field, we can demand that $I \underset{open}{\subseteq} \mathbb{C}$ is an open domain containing 0, and that $\gamma : I \to M$ be holomorphic.

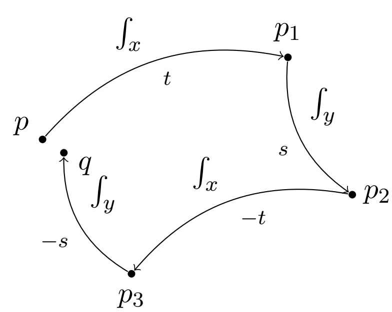
Proof In local coordinates, $\gamma : \mathbb{R} \to \mathbb{R}^n$, and we can use existence and uniqueness theorems for solutions to differential equations; then we need that a smooth (analytic, holomorphic) differential equation has a smooth (analytic, holomorphic) solution.

But there's a subtlety. What if there are two charts, and solutions on each chart, that diverge right where the charts stop overlapping? Well, since M is Hausdorff, if we have two maps $I \to M$, then the locus where they agree is closed, so if they don't agree on all of I, then we can go to the maximal point where they agree and look locally there. □

2.2.2.4 Definition *The* integral curve $\int_{x,p}(t)$ *of x at p is the maximal curve satisfying equations (2.2.2.2) and (2.2.2.3).*

2.2.2.5 Proposition *The integral curve $\int_{x,p}$ depends smoothly on $p \in M$.* □

2.2.2.6 Proposition *Let x and y be two vector fields on a manifold M. For $p \in M$ and $s, t \in \mathbb{R}$, define q by the following picture:*



Then for any smooth function f, we have $f(q) - f(p) = st[x, y]_p f + O(s, t)^3$.

Proof Let $\alpha(t) = \int_{x,p}(t)$, so that $f(\alpha(t))' = xf(\alpha(t))$. Iterating, we see that $\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^n f(\alpha(t)) = x^n f(\alpha(t))$, and by Taylor series expansion,

$$f(\alpha(t)) = \sum \frac{1}{n!}\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^n f(\alpha(0))t^n = \sum \frac{1}{n!}x^n f(p)t^n = e^{tx}f(p).$$

By varying p, we have:

$$\begin{aligned} f(q) &= \left(e^{-sy}f\right)(p_3) && (2.2.2.7)\\ &= \left(e^{-tx}e^{-sy}f\right)(p_2) && (2.2.2.8)\\ &= \left(e^{sy}e^{-tx}e^{-sy}f\right)(p_1) && (2.2.2.9)\\ &= \left(e^{tx}e^{sy}e^{-tx}e^{-sy}f\right)(p) && (2.2.2.10)\end{aligned}$$

We already know that $e^{tx}e^{sy}e^{-tx}e^{-sy} = 1 + st[x,y] + \text{higher terms}$. Therefore $f(q) - f(p) = st[x,y]_p f + O(s,t)^3$. $\square$

2.2.3 Group actions

2.2.3.1 Proposition *Let M be a manifold, G a Lie group, and $G \curvearrowright M$ a Lie group action, i.e. a smooth map $\rho : G \times M \to M$ satisfying equations (1.1.1.6) and (1.1.1.7). Let $x \in \mathrm{T}_e G$, where e is the identity element of the group G. The following descriptions of a vector field $\ell x \in \mathrm{Vect}(M)$ are equivalent:*

1. *Let $x = [\gamma]$ be the equivalence class of tangent paths, and let $\gamma : I \to G$ be a representative path. Define $(\ell x)_m = [\tilde{\gamma}]$ where $\tilde{\gamma}(t) \overset{\text{def}}{=} \rho\big(\gamma(t)^{-1}, m\big)$. On functions, ℓx acts as:*

$$(\ell x)_m f \overset{\text{def}}{=} \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} f\big(\gamma(t)^{-1}m\big)$$

2. *Arbitrarily extend x to a vector field $\tilde{x}$ on a neighborhood $U \subseteq G$ of e, and lift this to $\tilde{\tilde{x}}$ on $U \times M$ to point only in the U-direction: $\tilde{\tilde{x}}_{(u,m)} \overset{\text{def}}{=} (\tilde{x}_u, 0) \in \mathrm{T}_u U \times \mathrm{T}_m M$. Let ℓx act on functions by:*

$$(\ell x) f \overset{\text{def}}{=} -\tilde{\tilde{x}}(f \circ p)\big|_{\{e\}\times M = M}$$

3. $(\ell x)_m \overset{\text{def}}{=} -(d\rho)_{(e,m)}(x, 0)$ $\square$

2.2.3.2 Proposition *Let G be a Lie group, M and N manifolds, and $G \curvearrowright M$, $G \curvearrowright N$ Lie actions, and let $f : M \to N$ be G-equivariant. Given $x \in \mathrm{T}_e G$, define $\ell^M x$ and $\ell^N x$ vector fields on M and N as in Proposition 2.2.3.1. Then for each $m \in M$, we have:*

$$(df)_m(\ell^M x) = (\ell^N x)_{f(m)}$$ $\square$

2.2.3.3 Definition *Let $G \curvearrowright M$ be a Lie action. We define the* adjoint action *of G on* $\mathrm{Vect}(M)$ *by ${}^g y \overset{\text{def}}{=} dg(y)_{gm} = (dg)_m(y_m)$. Equivalently, $G \curvearrowright \mathscr{C}_M$ by $g : f \mapsto f \circ g^{-1}$, and given a vector field thought of as a derivation $y : \mathscr{C}_M \to \mathscr{C}_M$, we define ${}^g y \overset{\text{def}}{=} gyg^{-1}$.*

2.2.3.4 Example Let $G \curvearrowright G$ by right multiplication: $\rho(g,h) \stackrel{\text{def}}{=} hg^{-1}$. Then $G \curvearrowright \mathrm{T}_e G$ by the *adjoint action* $\mathrm{Ad}(g) = d(g - g^{-1})_e$, i.e. if $x = [\gamma]$, then $\mathrm{Ad}(g)x = [g\,\gamma(t)\,g^{-1}]$. $\lozenge$

2.2.3.5 Definition *Let $\rho : G \curvearrowright M$ be a Lie action. For each $g \in G$, we define ${}^g M$ to be the manifold M with the action ${}^g\rho : (h,m) \mapsto \rho(ghg^{-1}, m)$.*

2.2.3.6 Corollary *For each $g \in G$, the map $g : M \to {}^g M$ is G-equivariant. We have:*

$$ {}^g \ell x = \mathrm{d}g(\ell x) = \ell^{{}^g M} x = \ell(\mathrm{Ad}(g)x) \qquad \square $$

2.2.3.7 Proposition *Let $\rho : G \curvearrowright G$ by $\rho_g : h \mapsto hg^{-1}$. Then $\ell : \mathrm{T}_e G \to \mathrm{Vect}(G)$ is an isomorphism from $\mathrm{T}_e G$ to* left-invariant *vector fields, such that $(\ell x)_e = x$.*

Proof Let $\lambda : G \curvearrowright G$ be the action by left-multiplication: $\lambda_g(h) = gh$. Then for each g, λ_g is ρ-equivariant. Thus $\mathrm{d}\lambda_g(\ell x) = {}^{\lambda_g}(\ell x) = \ell x$, so ℓx is left-invariant, and $(\ell x)_e = x$ since $\rho(g,e) = g^{-1}$. Conversely, a left-invariant field is determined by its value at a point:

$$ (\ell x)_g = (\mathrm{d}\lambda_g)_e(\ell x_e) = (\mathrm{d}\lambda_g)_e(x) \qquad \square $$

2.2.4 Lie algebra of a Lie group

2.2.4.1 Lemma / Definition *Let $G \curvearrowright M$ be a Lie action. The subspace of $\mathrm{Vect}(M)$ of G-invariant derivations is a Lie subalgebra of $\mathrm{Vect}(M)$.*

Let G be a Lie group. The Lie algebra of G *is the Lie subalgebra* $\mathrm{Lie}(G)$ *of* $\mathrm{Vect}(G)$ *consisting of left-invariant vector fields, i.e. vector fields invariant under the action $\lambda : G \curvearrowright G$ given by $\lambda_g : h \mapsto gh$.*

We identify $\mathrm{Lie}(G) \stackrel{\text{def}}{=} \mathrm{T}_e G$ as in Proposition 2.2.3.7. $\square$

2.2.4.2 Lemma *Given $G \curvearrowright M$ a Lie action, $x \in \mathrm{Lie}(G)$ represented by $x = [\gamma]$, and $y \in \mathrm{Vect}(M)$, we have:*

$$ \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} {}^{\gamma(t)} y f = [\ell x, y] f $$

Proof

$$ \begin{aligned} \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} {}^{\gamma(t)} y f\,(p) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \gamma(t)\, y\, \gamma(t)^{-1} f\,(p) && (2.2.4.3)\\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \gamma(t)\, y\, f(\gamma(t)\, p) && (2.2.4.4)\\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \gamma(t)\, y\, f(\gamma(0)\, p) + \gamma(0)\, \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} y f(\gamma(t) p) && (2.2.4.5)\\ &= \ell x (y f)(p) + y\, \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} f(\gamma(t) p) && (2.2.4.6)\\ &= \ell x (y f)(p) + y(-\ell x\, f)(p) && (2.2.4.7)\\ &= [\ell x, y] f\,(p) && (2.2.4.8) \end{aligned} $$

$\square$

2.2.4.9 Corollary *Let $G \curvearrowright M$ be a Lie action. If $x, y \in \mathrm{Lie}(G)$, where $x = [\gamma]$, then*

$$\ell^M\big(\ell^{\mathrm{Ad}}(-x)\, y\big) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \ell\big(\mathrm{Ad}(\gamma(t))y\big) f = [\ell x, \ell y]\, f$$

2.2.4.10 Lemma *The Lie bracket defined on $\mathrm{Lie}(\mathrm{GL}(n)) = \mathfrak{gl}(n) = \mathrm{T}_e\,\mathrm{GL}(n) = \mathrm{M}(n)$ defined in Lemma/Definition 2.2.4.1 is the matrix bracket $[x, y] = xy - yx$.*

Proof We represent $x \in \mathfrak{gl}(n)$ by $[e^{tx}]$. The adjoint action on $\mathrm{GL}(n)$ is given by $\mathrm{Ad}_G(g)\, h = ghg^{-1}$, which is linear in h and fixes e, and so passes immediately to the action $\mathrm{Ad} : \mathrm{GL}(n) \curvearrowright \mathrm{T}_e\,\mathrm{GL}(n)$ given by $\mathrm{Ad}_{\mathfrak{g}}(g)\, y = gyg^{-1}$. Then

$$[x, y] = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} e^{tx} y e^{-tx} = xy - yx.$$

□

2.2.4.11 Corollary *If H is a closed linear group, then Lemma/Definitions 1.2.1.2 and 2.2.4.1 agree.*

Exercises

1. (a) Show that the composition of two immersions is an immersion.

   (b) Show that an immersed submanifold $N \subseteq M$ is always a closed submanifold of an open submanifold, but not necessarily an open submanifold of a closed submanifold.

2. Prove that if $f : N \to M$ is a smooth map, then $(\mathrm{d}f)_p$ is surjective if and only if there are open neighborhoods U of p and V of $f(p)$, and an isomorphism $\psi : V \times W \to U$, such that $f \circ \psi$ is the projection on V.

   In particular, deduce that the fibers of f meet a neighborhood of p in immersed closed submanifolds of that neighborhood.

3. Prove the implicit function theorem: a map (of sets) $f : M \to N$ between manifolds is smooth if and only if its graph is an immersed closed submanifold of $M \times N$.

4. Prove that the curve $y^2 = x^3$ in $\mathbb{R}^2$ is not an immersed submanifold.

5. Let M be a complex holomorphic manifold, p a point of M, X a holomorphic vector field. Show that X has a complex integral curve γ defined on an open neighborhood U of 0 in $\mathbb{C}$, and unique on U if U is connected, which satisfies the usual defining equation but in a complex instead of a real variable t.

   Show that the restriction of γ to $U \cap \mathbb{R}$ is a real integral curve of X, when M is regarded as a real analytic manifold.

6. Let $\mathrm{SL}(2,\mathbb{C})$ act on the Riemann sphere $\mathbb{P}^1(\mathbb{C})$ by fractional linear transformations $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) z = (az+b)/(cz+d)$. Determine explicitly the vector fields $f(z)\partial_z$ corresponding to the infinitesimal action of the basis elements

$$E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad F = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

of $\mathfrak{sl}(2,\mathbb{C})$, and check that you have constructed a Lie algebra homomorphism by computing the commutators of these vector fields.

7. (a) Describe the map $\mathfrak{gl}(n,\mathbb{R}) = \mathrm{Lie}(\mathrm{GL}(n,\mathbb{R})) = \mathrm{Mat}(n,\mathbb{R}) \to \mathrm{Vect}(\mathbb{R}^n)$ given by the infinitesimal action of $\mathrm{GL}(n,\mathbb{R})$.

   (b) Show that $\mathfrak{so}(n,\mathbb{R})$ is equal to the subalgebra of $\mathfrak{gl}(n,\mathbb{R})$ consisting of elements whose infinitesimal action is a vector field tangential to the unit sphere in $\mathbb{R}^n$.

8. (a) Let X be an analytic vector field on M all of whose integral curves are unbounded (i.e., they are defined on all of $\mathbb{R}$). Show that there exists an analytic action of $R = (\mathbb{R},+)$ on M such that X is the infinitesimal action of the generator ∂_t of $\mathrm{Lie}(\mathbb{R})$.

   (b) More generally, prove the corresponding result for a family of n commuting vector fields X_i and action of $\mathbb{R}^n$.

9. (a) Show that the matrix $\left(\begin{smallmatrix} -a & 0 \\ 0 & -b \end{smallmatrix}\right)$ belongs to the identity component of $\mathrm{GL}(2,\mathbb{R})$ for all positive real numbers a, b.

   (b) Prove that if $a \neq b$, the above matrix is not in the image $\exp(\mathfrak{gl}(2,\mathbb{R}))$ of the exponential map.

Chapter 3

General Theory of Lie groups

3.1 From Lie algebra to Lie group

3.1.1 The exponential map

We state the following results for Lie groups over $\mathbb{R}$. When working with complex manifolds, we can replace $\mathbb{R}$ by $\mathbb{C}$ throughout, whence the interval $I \underset{\text{open}}{\subseteq} \mathbb{R}$ is replaced by a connected open domain $I \underset{\text{open}}{\subseteq} \mathbb{C}$. As always, the word “smooth” may mean “infinitely differentiable” or “analytic” or ....

3.1.1.1 Lemma *Let G be a Lie group and $x \in \mathrm{Lie}(G)$. Then there exists a unique Lie group homomorphism $\gamma_x : \mathbb{R} \to G$ such that $(\mathrm{d}\gamma_x)_0(\partial t) = x$. It is given by $\gamma_x(t) = (\int_e \ell x)(t)$.*

Proof Let $\gamma : I \to G$ be the maximal integral curve of ℓx passing through e. Since ℓx is left-invariant, $g\gamma(t)$ is an integral curve through q. Let $g = \gamma(s)$ for $s \in I$; then $\gamma(t)$ and $\gamma(s)\gamma(t)$ are integral curves through $\gamma(s)$, to they must coincide: $\gamma(s+t) = \gamma(s)\gamma(t)$, and $\gamma(-s) = \gamma(s)^{-1}$ for $s \in I \cap (-I)$. So γ is a groupoid homomorphism, and by defining $\gamma(s+t) \overset{\text{def}}{=} \gamma(s)\gamma(t)$ for $s, t \in I$, $s + t \notin I$, we extend γ to $I + I$. Since $\mathbb{R}$ is archimedean, this allows us to extend γ to all of $\mathbb{R}$; it will continue to be an integral curve, so really I must have been $\mathbb{R}$ all along. ☐

3.1.1.2 Corollary *There is a bijection between one-parameter subgroups of G (homomorphisms $\mathbb{R} \to G$) and elements of the Lie algebra of G.* ☐

3.1.1.3 Definition *The* exponential map $\exp : \mathrm{Lie}(G) \to G$ *is given by* $\exp x \overset{\text{def}}{=} \gamma_x(1)$*, where γ_x is as in Lemma 3.1.1.1.*

3.1.1.4 Proposition *Let $x^{(b)}$ be a* smooth family *of vector fields on M parameterized by $b \in B$ a manifold, i.e. the vector field $\tilde{x}$ on $B \times M$ given by $\tilde{x}_{(b,m)} = (0, x^{(b)}_m)$ is smooth. Then $(b, p, t) \mapsto \left(\int_p x^{(b)}\right)(t)$ is a smooth map from an open neighborhood of $B \times M \times \{0\}$ in $B \times M \times \mathbb{R}$ to M. When each $x^{(b)}$ has infinite-time solutions, we can take the open neighborhood to be all of $B \times M \times \mathbb{R}$.*

Proof Note that

$$\left(\int_{(b,p)} \tilde{x}\right)(t) = \left(b, \left(\int_p x^{(b)}\right)(t)\right)$$

So $B \times M \times \mathbb{R} \to B \times M \xrightarrow{\pi} M$ by $(b,p,t) \mapsto \left(\int_{(b,p)} \tilde{x}\right)(t) \mapsto \left(\int_p x^{(b)}\right)(t)$ is a composition of smooth functions, hence is smooth. $\square$

Summarizing the above remarks, we have:

3.1.1.5 Theorem (Exponential Map)
For each Lie group G, there is a unique smooth map $\exp : \mathrm{Lie}(G) \to G$ such that for $x \in \mathrm{Lie}(G)$, the map $t \mapsto \exp(tx)$ is the integral curve of ℓx through e. The map $t \mapsto \exp(tx)$ is a Lie group homomorphism $\mathbb{R} \to G$. $\square$

3.1.1.6 Example When $G = \mathrm{GL}(n)$, the map $\exp : \mathfrak{gl}(n) \to \mathrm{GL}(n)$ is the matrix exponential. $\Diamond$

3.1.1.7 Proposition *The differential at the origin $(d\exp)_0$ is the identity map $1_{\mathrm{Lie}(G)}$.*

Proof $d(\exp tx)_0(\partial_t) = x$. $\square$

3.1.1.8 Corollary $\exp$ *is a local homeomorphism.* $\square$

3.1.1.9 Definition *The local inverse of* $\exp : \mathrm{Lie}(G) \to G$ *is called "*$\log$*".*

One can show that if G is connected, then any open neighborhood of the identity generates G. In particular:

3.1.1.10 Proposition *If G is connected, then $\exp(\mathrm{Lie}(G))$ generates G.* $\square$

3.1.1.11 Proposition *If $\phi : H \to G$ is a group homomorphism, then the following diagram commutes:*

$$\begin{tikzcd} H \arrow[r, "\phi"] & G \\ \mathrm{Lie}(H) \arrow[u, "\exp"] \arrow[r, "(d\phi)_e"] & \mathrm{Lie}(G) \arrow[u, "\exp"'] \end{tikzcd}$$

If H is connected, then $d\phi$ determines ϕ. $\square$

3.1.2 The Fundamental Theorem

Like all good algebraists, we assume the Axiom of Choice. The main goal of this chapter and the next is to prove:

3.1.2.1 Theorem (Fundamental Theorem of Lie Groups and Algebras)

1. *The functor $G \mapsto \mathrm{Lie}(G)$ gives an equivalence of categories between the category* scLieGp *of simply-connected Lie groups (over $\mathbb{R}$ or $\mathbb{C}$) and the category* LieAlg *of finite-dimensional Lie algebras (over $\mathbb{R}$ or $\mathbb{C}$).*

2. *"The" inverse functor* $\mathfrak{h} \mapsto \mathrm{Grp}(\mathfrak{h})$ *is left-adjoint to* $\mathrm{Lie} : \text{LieGp} \to \text{LieAlg}$.

Let us outline the proof that we will give. The hard part is to construct, for each Lie algebra $\mathfrak{h}$, some Lie group H with Lie algebra $\mathfrak{h}$; then, using Proposition 3.4.2.9, a simply connected Lie group $\mathrm{Grp}(\mathfrak{h})$ with Lie algebra $\mathfrak{h}$ is easy to build.

To build H, consider open neighborhoods U and V so that the horizontal maps are a homeomorphism:

$$\begin{array}{ccc} \mathrm{Lie}(G) & & G \\ \cup\mathrm{I} & & \cup\mathrm{I} \\ U & \underset{\log}{\overset{\exp}{\rightleftarrows}} & V \\ \Cup & & \Cup \\ 0 & & e \end{array}$$

Consider the restriction $\mu : G \times G \to G$ to $(V \times V) \cap \mu^{-1}(V) \to V$, and use this to define a "partial group law" $b : \text{open} \to U$, where open $\subseteq U \times U$, via

$$b(x, y) \stackrel{\text{def}}{=} \log(\exp x \exp y) \tag{3.1.2.2}$$

We will show that the Lie algebra structure of $\mathrm{Lie}(G)$ determines b.

Conversely, given $\mathfrak{h}$ a finite-dimensional Lie algebra, we will need to define b and build $\tilde{H}$ as the group freely generated by U modulo the relations $xy = b(x, y)$ if $x, y, b(x, y) \in U$. We will need to prove that $\tilde{H}$ is a Lie group, with U as an open submanifold.

Given Theorem 3.1.2.1, the following much easier corollary is immediate:

3.1.2.3 Corollary *Every Lie subalgebra* $\mathfrak{h}$ *of* $\mathrm{Lie}(G)$ *is* $\mathrm{Lie}(H)$ *for a unique connected subgroup* $H \hookrightarrow G$*, up to equivalence.*

We will instead prove Corollary 3.1.2.3 as Theorem 3.4.1.2.

The standard proof of Theorem 3.1.2.1 is to first prove Corollary 3.1.2.3 and then use Theorem 4.5.0.10, which asserts that every finite-dimensional Lie algebra is a Lie subalgebra of some $\mathfrak{gl}(n, \mathbb{C})$, which is in turn the Lie algebra of the Lie group $\mathrm{GL}(n, \mathbb{C})$. We will use Theorem 4.4.4.15, which asserts directly that every Lie algebra is the Lie algebra of some Lie group, rather than Theorem 4.5.0.10.

As explained above, the first step is to define the function b of equation (3.1.2.2). This function is the main topic of the following theorem. We will prove the second part now. We will restate this theorem as Theorem 3.3.0.3 and prove both parts at that time.

3.1.2.4 Theorem (Baker–Campbell–Hausdorff Formula (second part only))

1. *Let* $\mathcal{T}(x, y)$ *be the free tensor algebra generated by* x *and* y*, and* $\mathcal{T}(x, y)[[s, t]]$ *the (non-commutative) ring of formal power series in two commuting variables* s *and* t*. Define* $b(tx, sy) \stackrel{\text{def}}{=} \log(\exp(tx)\exp(sy)) \in \mathcal{T}(x, y)[[s, t]]$*, where* $\exp$ *and* $\log$ *are the usual formal power series. Then*

$$b(tx, sy) = tx + sy + st\frac{1}{2}[x, y] + st^2\frac{1}{12}[x, [x, y]] + s^2t\frac{1}{12}[y, [y, x]] + \dots \tag{3.1.2.5}$$

has all coefficients given by Lie bracket polynomials in x *and* y*.*

2. *Given a Lie group G, there exists a neighborhood $U' \ni 0$ in $\mathrm{Lie}(G)$ such that $U' \subseteq U \underset{\log}{\overset{\exp}{\rightleftarrows}} V \subseteq G$ and $b(x,y)$ converges on $U' \times U'$ to $\log(\exp x \exp y)$.*

We work with analytic manifolds; on $\mathscr{C}^\infty$ manifolds, we can make an analogous argument using the language of differential equations.

Proof (of part 2.) For a clearer exposition, we distinguish the maps $\exp : \mathrm{Lie}(G) \to G$ from $e^x \in \mathbb{R}[[x]]$.

We begin with a basic identity. $\exp(tx)$ is an integral curve to ℓx through e, so by left-invariance, $t \mapsto g\exp(tx)$ is the integral curve of ℓx through g. Thus, for f analytic on G,

$$\frac{\mathrm{d}}{\mathrm{d}t}\big[f(g\exp tx)\big] = \big((\ell x)f\big)(g\exp tx)$$

We iterate:

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^n \big[f(g\exp tx)\big] = \big((\ell x)^n f\big)(g\exp tx)$$

If f is analytic, then for small t the Taylor series converges:

$$\begin{aligned}
f(g\exp tx) &= \sum_{n=0}^{\infty}\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^n \big[f(g\exp tx)\big]\Big|_{t=0}\frac{t^n}{n!} && (3.1.2.6)\\
&= \sum_{n=0}^{\infty}\big((\ell x)^n f\big)(g\exp tx)\big|_{t=0}\frac{t^n}{n!} && (3.1.2.7)\\
&= \sum_{n=0}^{\infty}\big((\ell x)^n f\big)(g)\frac{t^n}{n!} && (3.1.2.8)\\
&= \sum_{n=0}^{\infty}\left(\frac{(t\,\ell x)^n}{n!}f\right)(g) && (3.1.2.9)\\
&= \big(e^{t\,\ell x}f\big)(g) && (3.1.2.10)
\end{aligned}$$

We repeat the trick:

$$f(\exp tx\exp sy) = \big(e^{s\,Ly}f\big)(\exp tx) = \big(e^{t\,\ell x}e^{s\,Ly}f\big)(e) = \big(e^{tx}e^{sy}f\big)(e)$$

The last equality is because we are evaluating the derivations at e, where $\ell x = x$.

We now let $f = \log : V \to U$, or rather a coordinate of log. Then the left-hand-side is just $\log(\exp tx \exp sy)$, and the right hand side is $\big(e^{tx}e^{sy}\log\big)(e) = \big(e^{b(tx,sy)}\log\big)(e)$, where b is the formal power series from part 1. — we have shown that the right hand side converges. But by interpreting the calculations above as formal power series, and expanding log in Taylor series, we see that the formal power series $\big(e^{b(tx,sy)}\log\big)(e)$ agrees with the formal power series $\log\big(e^{b(tx,sy)}\big) = b(tx,sy)$. This completes the proof of part 2. $\square$

3.2 Universal enveloping algebras

3.2.1 The definition

3.2.1.1 Definition *A* representation *of a Lie group is a homomorphism* $G \to \mathrm{GL}(n, \mathbb{R})$ *(or* $\mathbb{C}$*). A* representation *of a Lie algebra is a homomorphism* $\mathrm{Lie}(G) \to \mathfrak{gl}(n) = \mathrm{End}(V)$*; the space* $\mathrm{End}(V)$ *is a Lie algebra with the bracket given by* $[x, y] = xy - yx$*.*

3.2.1.2 Definition *Let* V *be a vector space. The* tensor algebra *over* V *is the free unital non-commuting algebra* $\mathcal{T}V$ *generated by a basis of* V*. Equivalently:*

$$\mathcal{T}V \stackrel{\text{def}}{=} \bigoplus_{n \geq 0} V^{\otimes n}$$

The multiplication is given by $\otimes : V^{\otimes k} \times V^{\otimes l} \to V^{\otimes(k+l)}$*.* $\mathcal{T}$ *is a functor, and is left-adjoint to* Forget : $\textsc{Alg} \to \textsc{Vect}$*.*

3.2.1.3 Lemma / Definition *Let* $\mathfrak{g}$ *be a Lie algebra. The* universal enveloping algebra *is*

$$\mathcal{U}\mathfrak{g} \stackrel{\text{def}}{=} \mathcal{T}\mathfrak{g} / \big\langle [x, y] - (xy - yx) \big\rangle$$

$\mathcal{U} : \textsc{LieAlg} \to \textsc{Alg}$ *is a functor, and is left-adjoint to* Forget : $\textsc{Alg} \to \textsc{LieAlg}$*.*

3.2.1.4 Corollary *The category of* $\mathfrak{g}$*-modules is equal to the category of* $\mathcal{U}\mathfrak{g}$*-modules.* $\square$

3.2.1.5 Example A Lie algebra $\mathfrak{g}$ is *abelian* if the bracket is identically 0. If $\mathfrak{g}$ is abelian, then $\mathcal{U}\mathfrak{g} = \mathcal{S}\mathfrak{g}$, where $\mathcal{S}V$ is the symmetric algebra generated by the vector space V (so that $\mathcal{S}$ is left-adjoint to Forget : $\textsc{ComAlg} \to \textsc{Vect}$). $\diamondsuit$

3.2.1.6 Example If $\mathfrak{f}$ is the free Lie algebra on generators $x_1, \dots, x_d$, defined in terms of a universal property, then $\mathcal{U}\mathfrak{f} = \mathcal{T}(x_1, \dots, x_d)$. $\diamondsuit$

3.2.1.7 Definition *A vector space* V *is* graded *if it comes with a direct-sum decomposition* $V = \bigoplus_{n \geq 0} V_n$*. A morphism of graded vector spaces preserves the grading. A* graded algebra *is an algebra object in the category of graded vector spaces. I.e. it is a vector space* $V = \bigoplus_{n \geq 0}$ *along with a unital associative multiplication* $V \otimes V \to V$ *such that if* $v_n \in V_n$ *and* $v_m \in V_m$*, then* $v_n v_m \in V_{n+m}$*.*

A vector space V *is* filtered *if it comes with an increasing sequence of subspaces*

$$\{0\} \subseteq V_{\leq 0} \subseteq V_{\leq 1} \subseteq \cdots \subseteq V$$

such that $V = \bigcup_{n \geq 0} V_n$*. A morphism of graded vector spaces preserves the filtration. A* filtered algebra *is an algebra object in the category of filtered vector spaces. I.e. it is a filtered vector space along with a unital associative multiplication* $V \otimes V \to V$ *such that if* $v_n \in V_{\leq n}$ *and* $v_m \in V_{\leq m}$*, then* $v_n v_m \in V_{\leq (n+m)}$*.*

Given a filtered vector space V*, we define* $\mathrm{gr}\, V \stackrel{\text{def}}{=} \bigoplus_{n \geq 0} \mathrm{gr}_n V$*, where* $\mathrm{gr}_n V \stackrel{\text{def}}{=} V_{\leq n} / V_{\leq (n-1)}$*.*

3.2.1.8 Lemma gr *is a functor. If V is a filtered algebra, then* $\operatorname{gr} V$ *is a graded algebra.* □

3.2.1.9 Example Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{K}$. Then $\mathcal{U}\mathfrak{g}$ has a natural filtration inherited from the filtration of $\mathcal{T}\mathfrak{g}$, since the ideal $\langle xy - yx = [x,y]\rangle$ preserves the filtration. Since $\mathcal{U}\mathfrak{g}$ is generated by $\mathfrak{g}$, so is $\operatorname{gr}\mathcal{U}\mathfrak{g}$; since $xy - yx = [x,y] \in \mathcal{U}_{\leq 1}$, $\operatorname{gr}\mathcal{U}\mathfrak{g}$ is commutative, and so there is a natural projection $\mathcal{S}\mathfrak{g} \twoheadrightarrow \operatorname{gr}\mathcal{U}\mathfrak{g}$. ♢

3.2.2 Poincaré–Birkhoff–Witt theorem

3.2.2.1 Theorem (Poincaré–Birkhoff–Witt)
The map $\mathcal{S}\mathfrak{g} \to \operatorname{gr}\mathcal{U}\mathfrak{g}$ is an isomorphism of algebras.

Proof Pick an ordered basis $\{x_\alpha\}$ of $\mathfrak{g}$; then the monomials $x_{\alpha_1} \dots x_{\alpha_n}$ for $\alpha_1 \leq \dots \leq \alpha_n$ are an ordered basis of $\mathcal{S}\mathfrak{g}$, where we take the "deg-lex" ordering: a monomial of lower degree is immediately smaller than a monomial of high degree, and for monomials of the same degree we alphabetize. Since $\mathcal{S}\mathfrak{g} \twoheadrightarrow \operatorname{gr}\mathcal{U}\mathfrak{g}$ is an algebra homomorphism, the set $\{x_{\alpha_1} \dots x_{\alpha_n} \text{ s.t. } \alpha_1 \leq \dots \leq \alpha_n\}$ spans $\operatorname{gr}\mathcal{U}\mathfrak{g}$. It suffices to show that they are independent in $\operatorname{gr}\mathcal{U}\mathfrak{g}$. For this it suffices to show that the set $S \overset{\text{def}}{=} \{x_{\alpha_1} \dots x_{\alpha_n} \text{ s.t. } \alpha_1 \leq \dots \leq \alpha_n\}$ is independent in $\mathcal{U}\mathfrak{g}$.

Let $I = \langle xy - yx - [x,y]\rangle$ be the ideal of $\mathcal{T}\mathfrak{g}$ such that $\mathcal{U}\mathfrak{g} = \mathcal{T}\mathfrak{g}/I$. Define $J \subseteq \mathcal{T}\mathfrak{g}$ to be the span of expressions of the form

$$\xi = x_{\alpha_1} \cdots x_{\alpha_k} \left(x_\beta x_\gamma - x_\gamma x_\beta - [x_\beta, x_\gamma]\right) x_{\nu_1} \cdots x_{\nu_l} \tag{3.2.2.2}$$

where $\alpha_1 \leq \dots \leq \alpha_k \leq \beta > \gamma$, and there are no conditions on ν_i, so that J is a right ideal. We take the deg-lex ordering in $\mathcal{T}\mathfrak{g}$. The leading monomial in equation (3.2.2.2) is $x_{\vec\alpha} x_\beta x_\gamma x_{\vec\nu}$. Thus S is an independent set in $\mathcal{T}\mathfrak{g}/J$. We need only show that $J = I$.

The ideal I is generated by expressions of the form $x_\beta x_\gamma - x_\gamma x_\beta - [x_\beta, x_\gamma]$ as a two-sided ideal. If $\beta > \gamma$ then $\left(x_\beta x_\gamma - x_\gamma x_\beta - [x_\beta, x_\gamma]\right) \in J$; by antisymmetry, if $\beta < \gamma$ we switch them and stay in J. If $\beta = \gamma$, then $\left(x_\beta x_\gamma - x_\gamma x_\beta - [x_\beta, x_\gamma]\right) = 0$. Thus J is a right ideal contained in I, and the two-sided ideal generated by J contains I. Thus the two-sided ideal generated by J is I, and it suffices to show that J is a two-sided ideal.

We multiply $x_\delta \xi$. If $k > 0$ and $\delta \leq \alpha_1$, then $x_\delta \xi \in J$. If $\delta > \alpha_1$, then $x_\delta \xi \equiv x_{\alpha_1} x_\delta x_{\alpha_2} \cdots + [x_\delta, x_{\alpha_1}] x_{\alpha_2} \dots \mod J$. And both $x_\delta x_{\alpha_2} \dots$ and $[x_\delta, x_{\alpha_1}] x_{\alpha_2} \dots$ are in J by induction on degree. Then since $\alpha_1 < \delta$, $x_{\alpha_1} x_\delta x_{\alpha_2} \cdots \in J$ by (transfinite) induction on δ.

So suffice to show that if $k = 0$, then we're still in J. I.e. if $\alpha > \beta > \gamma$, then we want to show that $x_\alpha \left(x_\beta x_\gamma - x_\gamma x_\beta - [x_\beta, x_\gamma]\right) \in J$. Well, since $\alpha > \beta$, we see that $x_\alpha x_\beta - x_\beta - [x_\alpha, x_\beta] \in J$, and same with $\beta \leftrightarrow \gamma$. So, working modulo J, we have

$$\begin{aligned}
x_\alpha\left(x_\beta x_\gamma - x_\gamma x_\beta - [x_\beta, x_\gamma]\right) &\equiv \left(x_\beta x_\alpha + [x_\alpha, x_\beta]\right) x_\gamma - \left(x_\gamma x_\alpha + [x_\alpha, x_\gamma]\right) x_\beta - x_\alpha [x_\beta, x_\gamma] \\
&\equiv x_\beta\left(x_\gamma x_\alpha + [x_\alpha, x_\gamma]\right) + [x_\alpha, x_\beta] x_\gamma - x_\gamma\left(x_\beta x_\alpha + [x_\alpha, x_\beta]\right) \\
&\quad - [x_\alpha, x_\gamma] x_\beta - x_\alpha [x_\beta, x_\gamma] \\
&\equiv x_\gamma x_\beta x_\alpha + [x_\beta, x_\gamma] x_\alpha + x_\beta [x_\alpha, x_\gamma] + [x_\alpha, x_\beta] x_\gamma - x_\gamma\left(x_\beta x_\alpha + [x_\alpha, x_\beta]\right) \\
&\quad - [x_\alpha, x_\gamma] x_\beta - x_\alpha [x_\beta, x_\gamma] \\
&= [x_\beta, x_\gamma] x_\alpha + x_\beta [x_\alpha, x_\gamma] + [x_\alpha, x_\beta] x_\gamma - x_\gamma [x_\alpha, x_\beta] - [x_\alpha, x_\gamma] x_\beta - x_\alpha [x_\beta, x_\gamma] \\
&\equiv -[x_\alpha, [x_\beta, x_\gamma]] + [x_\beta, [x_\alpha, x_\gamma]] - [x_\gamma, [x_\alpha, x_\beta]] \\
&= 0 \text{ by Jacobi.} \qquad \square
\end{aligned}$$

3.2.2.3 Corollary $\mathfrak{g} \hookrightarrow \mathcal{U}\mathfrak{g}$. *Thus every Lie algebra is isomorphic to a Lie subalgebra of some* $\mathrm{End}(V)$, *namely* $V = \mathcal{U}\mathfrak{g}$.

Proof The canonical map $\mathfrak{g} \to \mathcal{T}\mathfrak{g} \to \mathcal{U}\mathfrak{g}$ is the map $\mathfrak{g} \to \mathcal{S}\mathfrak{g}$ as the degree-one piece, which is an injection. $\square$

3.2.3 $\mathcal{U}\mathfrak{g}$ is a bialgebra

The usual definition of "associative algebra" over $\mathbb{K}$ can be encoded by saying that an *algebra* over $\mathbb{K}$ is a vector space U along with a $\mathbb{K}$-linear "multiplication" map $\mu : U \underset{\mathbb{K}}{\otimes} U \to U$ which is *associative*, i.e. the following diagram commutes:

$$\begin{tikzcd}
U \otimes U \otimes U \arrow[r, "1_U \otimes \mu"] \arrow[d, "\mu \otimes 1_U"'] & U \otimes U \arrow[d, "\mu"] \\
U \otimes U \arrow[r, "\mu"] & U
\end{tikzcd} \qquad (3.2.3.1)$$

We demand that all our algebras be *unital*, meaning that there is a linear map $e : \mathbb{K} \to U$ such that the maps $U = \mathbb{K} \otimes U \overset{e \otimes 1_U}{\to} U \otimes U \overset{\mu}{\to} U$ and $U = U \otimes K \overset{1_U \otimes e}{\to} U \otimes U \overset{\mu}{\to} U$ are the identity maps. We will call the image of $1 \in \mathbb{K}$ under e simply $1 \in U$. Reversing the direction of arrows gives:

3.2.3.2 Definition *A* coalgebra *is an algebra in the opposite category. I.e. it is a vector space* U *along with a "comultiplication" map* $\Delta : U \to U \otimes U$ *so that the following commutes:*

$$\begin{tikzcd}
U \arrow[r, "\Delta"] \arrow[d, "\Delta"'] & U \otimes U \arrow[d, "1_U \otimes \Delta"] \\
U \otimes U \arrow[r, "\Delta \otimes 1_U"] & U \otimes U \otimes U
\end{tikzcd} \qquad (3.2.3.3)$$

In elements, if $\Delta x = \sum x_{(1)} \otimes x_{(2)}$*, then we demand that* $\sum x_{(1)} \otimes \Delta(x_{(2)}) = \sum \Delta(x_{(1)}) \otimes x_{(1)}$*. We demand that our coalgebras be* counital*, meaning that there is a linear map* $\epsilon : U \to \mathbb{K}$ *such that the maps* $U \overset{\Delta}{\to} U \otimes U \overset{\epsilon \otimes 1_U}{\mathbb{K}} \otimes U = U$ *and* $U \overset{\Delta}{\to} U \otimes U \overset{1_U \otimes \epsilon}{U} \otimes \mathbb{K} = U$ *are the identity maps.*

A bialgebra *is an algebra in the category of coalgebras, or equivalently a coalgebra in the category of algebras. I.e. it is a vector space* U *with maps* $\mu : U \otimes U \to U$ *and* $\Delta : U \to U \otimes U$ *satisfying equations (3.2.3.1) and (3.2.3.3) such that* Δ *and* ϵ *are (unital) algebra homomorphisms or equivalently such that* μ *and* e *are (counital) coalgebra homomorphism. We have defined the multiplication on* $U \otimes U$ *by* $(x \otimes y)(z \otimes w) = (xz) \otimes (yw)$*, and the comultiplication by* $\Delta(x \otimes y) = \sum \sum x_{(1)} \otimes y_{(1)} \otimes x_{(2)} \otimes y_{(2)}$*, where* $\Delta x = \sum x_{(1)} \otimes x_{(2)}$ *and* $\Delta y = \sum y_{(1)} \otimes y_{(2)}$*; the unit and counit are* $e \otimes e$ *and* $\epsilon \otimes \epsilon$*.*

3.2.3.4 Definition *Let* U *be a bialgebra, and* $x \in U$*. We say that* x *is* primitive *if* $\Delta x = x \otimes 1 + 1 \otimes x$*, and that* x *is* grouplike *if* $\Delta x = x \otimes x$*. The set of primitive elements of* U *we denote by* $\mathrm{prim}\, U$*.*

3.2.3.5 Proposition $\mathcal{U}\mathfrak{g}$ *is a bialgebra with* $\mathrm{prim}\,\mathcal{U}\mathfrak{g} = \mathfrak{g}$*.*

Proof To define the comultiplication, it suffices to show that $\Delta : \mathfrak{g} \to \mathcal{U}\mathfrak{g} \otimes \mathcal{U}\mathfrak{g}$ given by $x \mapsto x \otimes 1 + 1 \otimes x$ is a Lie algebra homomorphism, whence it uniquely extends to an algebra homomorphism by the universal property. We compute:

$$[x \otimes 1 + 1 \otimes x, y \otimes 1 + 1 \otimes y]_{\mathcal{U}\otimes\mathcal{U}} = [x \otimes 1, y \otimes 1] + [1 \otimes x, 1 \otimes y] \quad (3.2.3.6)$$
$$= [x, y] \otimes 1 + 1 \otimes [x, y] \quad (3.2.3.7)$$

To show that Δ thus defined is coassiciative, it suffices to check on the generating set $\mathfrak{g}$, where we see that $\Delta^2(x) = x \otimes 1 \otimes 1 + 1 \otimes x \otimes 1 + 1 \otimes 1 \otimes x$.

By definition, $\mathfrak{g} \subseteq \mathcal{U}\mathfrak{g}$. To show equality, we use Theorem 3.2.2.1. We filter $\mathcal{U}\mathfrak{g} \otimes \mathcal{U}\mathfrak{g}$ in the obvious way, and since Δ is an algebra homomorphism, we see that $\Delta\big(\mathcal{U}\mathfrak{g}_{\leq 1}\big) \subseteq \big(\mathcal{U}\mathfrak{g} \otimes \mathcal{U}\mathfrak{g}\big)_{\leq 1}$, whence $\Delta\big(\mathcal{U}\mathfrak{g}_{\leq n}\big) \subseteq \big(\mathcal{U}\mathfrak{g} \otimes \mathcal{U}\mathfrak{g}\big)_{\leq n}$. Thus Δ induces a map $\bar{\Delta}$ on $\mathrm{gr}\,\mathcal{U}\mathfrak{g} = \mathcal{S}\mathfrak{g}$, and $\bar{\Delta}$ makes $\mathcal{S}\mathfrak{g}$ into a bialgebra.

Let $\xi \in \mathcal{U}\mathfrak{g}_{\leq n}$ be primitive, and define its image to be $\bar{\xi} \in \mathrm{gr}_n\,\mathcal{U}\mathfrak{g}$; then $\bar{\xi}$ must also be primitive. But $\mathcal{S}\mathfrak{g} \otimes \mathcal{S}\mathfrak{g} = \mathbb{K}[y_\alpha, z_\alpha]$, where $\{x_\alpha\}$ is a basis of $\mathfrak{g}$ (whence $\mathcal{S}\mathfrak{g} = \mathbb{K}[x_\alpha]$), and we set $y_\alpha = x_\alpha \otimes 1$ and $z_\alpha = 1 \otimes x_\alpha]$. We check that $\bar{\Delta}(x_\alpha) = y_\alpha + z_\alpha$, and so if $f(x) \in \mathcal{S}\mathfrak{g}$, we see that $\Delta f(x) = f(y+z)$. So $f \in \mathcal{S}\mathfrak{g}$ is primitive if and only if $f(y+z) = f(y) + f(z)$, i.e. iff f is homogenoues of degree 1. Therefore $\mathrm{prim}\,\mathrm{gr}\,\mathcal{U}\mathfrak{g} = \mathrm{gr}_1\,\mathcal{U}\mathfrak{g}$, and so if $\xi \in \mathcal{U}\mathfrak{g}$ is primitive, then $\bar{\xi} \in \mathrm{gr}_1\,\mathcal{U}\mathfrak{g}$ so $\xi = x + c$ for some $x \in \mathfrak{g}$ and some $c \in \mathbb{K}$. Since x is primitive, c must be also, and the only primitive constant is 0.□

3.2.4 Geometry of the universal enveloping algebra

3.2.4.1 Definition *Let* X *be a space and* $\mathscr{S}$ *a sheaf of functions on* X*. We define the sheaf* $\mathscr{D}$ *of* Grothendieck differential operators *inductively. Given* $U \underset{open}{\subseteq} X$*, we define* $\mathscr{D}_{\leq 0}(U) = \mathscr{S}(U)$*, and* $\mathscr{D}_{\leq n}(U) = \{x : \mathscr{S}(U) \to \mathscr{S}(U) \text{ s.t. } [x, f] \in \mathscr{D}_{\leq (n-1)}(U)\, \forall f \in \mathscr{S}(U)\}$*, where* $\mathscr{S}(U) \curvearrowright \mathscr{S}(U)$ *by left-multiplication. Then* $\mathscr{D}(U) = \bigcup_{n \geq 0} \mathscr{D}_{\leq n}(U)$ *is a filtered sheaf; we say that* $x \in \mathscr{D}_{\leq n}(U)$ *is an "*n*th-order differential operator on* U*".*

3.2.4.2 Lemma *$\mathscr{D}$ is a sheaf of filtered algebras, with the multiplication on $\mathscr{D}(U)$ inherited from* $\mathrm{End}(\mathscr{S}(U))$. *For each n, $\mathscr{D}_{\leq n}$ is a sheaf of Lie subalgebras of $\mathscr{D}$.* □

We will not prove:

3.2.4.3 Theorem (Grothendieck Differential Operators)
Let X be a manifold, $\mathscr{C}$ the sheaf of smooth functions on X, and $\mathscr{D}$ the sheaf of differential operators on $\mathscr{C}$ as in Definition 3.2.4.1. Then $\mathscr{D}(U)$ is generated as a noncommutative algebra by $\mathscr{C}(U)$ and $\mathrm{Vect}(U)$, *and* $\mathscr{D}_{\leq 1} = \mathscr{C}(U) \oplus \mathrm{Vect}(U)$. □

3.2.4.4 Proposition *Let G be a Lie group, and $\mathscr{D}(G)^G$ the subalgebra of left-invariant differential operators on G. The natural map $\mathcal{U}\mathfrak{g} \to \mathscr{D}(G)^G$ generated by the identification of $\mathfrak{g}$ with left-invariant vector fields is an isomorphism of algebras.* □

We will revisit this algebraic notion of differential operator in Section 9.5.3.

3.3 The Baker–Campbell–Hausdorff Formula

3.3.0.1 Lemma *Let U be a bialgebra with comultiplication Δ. Define $\hat{\Delta} : U[[s]] \to (U \otimes U)[[s]]$ by linearity; then $\hat{\Delta}$ is an s-adic-continuous algebra homomorphism, and so commutes with formal power series.*

Let $\psi \in U[[s]]$ with $\psi(0) = 0$. Then ψ is primitive term-by-term — $\hat{\Delta}(\psi) = \psi \otimes 1 + 1 \otimes \psi$, if and only if e^{ψ} is "group-like" in the sense that $\hat{\Delta}(e^{\psi}) = e^{\psi} \otimes e^{\psi}$, where we have defined $\otimes :$ $U[[s]] \otimes U[[s]] \to (U \otimes U)[[s]]$ by $s^n \otimes s^m \mapsto s^{n+m}$.

Proof $e^{\psi} \otimes e^{\psi} = (1 \otimes e^{\psi})(e^{\psi} \otimes 1) = e^{1\otimes\psi} e^{\psi\otimes 1} = e^{1\otimes\psi + \psi\otimes 1}$ □

3.3.0.2 Lemma *Let G be a Lie group, $\mathfrak{g} = \mathrm{Lie}(G)$, and identify $\mathcal{U}\mathfrak{g}$ with the left-invariant differential operators on G, as in Proposition 3.2.4.4. Let $\mathscr{C}(G)_e$ be the stalk of smooth functions defined in some open set around e (we write $\mathscr{C}$ for the sheaf of functions on G; when G is analytic, we really mean the sheaf of analytic functions on G). Then if $u \in \mathcal{U}\mathfrak{g}$ satisfied $uf(e) = 0$ for each $f \in \mathscr{C}(G)_e$, then $u = 0$.*

Proof For $g \in G$, we have $uf(g) = u(\lambda_{g^{-1}} f)(e) = \lambda_{g^{-1}}(uf)(e) = 0$. □

We are now prepared to give a complete proof of Theorem 3.1.2.4, which we restate for convenience:

3.3.0.3 Theorem (Baker–Campbell–Hausdorff Formula)

1. *Let $\mathfrak{f}$ be the free Lie algebra on two generators x, y; recall that $\mathcal{U}\mathfrak{f} = \mathcal{T}(x, y)$. Define the formal power series $b(tx, sy) \in \mathcal{T}(x, y)[[s, t]]$, where s and t are commuting variables, by*

$$e^{b(tx,sy)} \stackrel{\text{def}}{=} e^{tx} e^{sy}$$

Then $b(tx, sy) \in \mathfrak{f}[[s, t]]$, i.e. b is a series all of whose coefficients are Lie algebra polynomials in the generators x and y.

2. *If G is a Lie group (in the analytic category), then there are open neighborhoods $0 \in U' \underset{open}{\subseteq} U \underset{open}{\subseteq} \mathrm{Lie}(G) = \mathfrak{g}$ and $0 \in V' \underset{open}{\subseteq} V \underset{open}{\subseteq} G$ such that $U \underset{\log}{\overset{\exp}{\rightleftarrows}} V$ and $U' \rightleftarrows V'$ and such that $b(x, y)$ converges on $U' \times U'$ to $\log(\exp x \exp y)$.*

Proof 1. Let $\hat{\Delta} : \mathcal{T}(x, y)[[s, t]] = \mathcal{U}\mathfrak{f}[[s, t]] \to (\mathcal{U}\mathfrak{f} \otimes \mathcal{U}\mathfrak{f})[[s, t]]$ as in Lemma 3.3.0.1. Since $e^{tx}e^{sy}$ is grouplike —

$$\hat{\Delta}(e^{tx}e^{sy}) = \hat{\Delta}(e^{tx})\,\hat{\Delta}(e^{sy}) = \left(e^{tx} \otimes e^{tx}\right)(e^{sy} \otimes e^{sy}) = e^{tx}e^{sy} \otimes e^{tx}e^{ty}$$

— we see that $b(tx, sy)$ is primitive term-by-term.

2. Let U, V be open neighborhoods of $\mathrm{Lie}(G)$ and G respectively, and pick V' so that $\mu : G \times G \to G$ restricts to a map $V' \times V' \to V$; let $U' = \log(V')$. Define $\beta(x, y) = \log(\exp x \exp y)$; then β is an analytic function $U' \times U' \to U'$.

   Let $x, y \in \mathrm{Lie}(G)$ and $f \in \mathscr{C}(G)_e$. Then $\left(e^{tx}e^{sy}f\right)(e)$ is the Taylor series expansion of $f(\exp tx \exp sy)$, as in the proof of Theorem 3.1.2.4. Let $\tilde{\beta}$ be the formal power series that is the Taylor expansion of β; then $e^{\tilde{\beta}(tx,sy)}f(e)$ is also the Taylor series expansion of $f(\exp tx \exp sy)$. This implies that for every $f \in \mathscr{C}(G)_e$, $e^{\tilde{\beta}(tx,sy)}f(e)$ and $e^{tx}e^{sy}f(e)$ have the same coefficients. But the coefficients are left-invariant differential operators applied to f, so by Lemma 3.3.0.2 the series $e^{\tilde{\beta}(tx,sy)}$ and $e^{tx}e^{sy}$ must agree. Upon applying the formal logarithms, we see that $b(tx, sy) = \tilde{\beta}(tx, sy)$.

   But $\tilde{\beta}$ is the Taylor series of the analytic function β, so by shrinking U' (and hence V') we can assure that it converges. □

3.4 Lie subgroups

3.4.1 Relationship between Lie subgroups and Lie subalgebras

3.4.1.1 Definition *Let G be a Lie group. A* Lie subgroup *of G is a subgroup H of G with its own Lie group structure, so that the inclusion $H \hookrightarrow G$ is a local immersion. We will write "$H \leq G$" when H is a Lie subgroup of G.*

Just to emphasize the point, a subgroup $H \leq G$ does not need to be a (closed) submanifold: the manifold structures on H and G must be compatible in that $H \hookrightarrow G$ should be an immersion, but the manifold structure on H is not necessarily the restriction of a manifold structure on G. The issue is that a subgroup $H \leq G$ might be dense: for example, the injection $\mathbb{R} \hookrightarrow (\mathbb{R}/\mathbb{Z})^2$ given by $t \mapsto (t, \alpha t)$ for α irrational is a Lie subgroup but not a (closed) submanifold.

We stated the following result as Corollary 3.1.2.3:

3.4.1.2 Theorem (Identification of Lie subalgebras and Lie subgroups)
Every Lie subalgebra of $\mathrm{Lie}(G)$ is $\mathrm{Lie}(H)$ for a unique connected Lie subgroup $H \leq G$.

Proof We first prove uniqueness. If H is a Lie subgroup of G, with $\mathfrak{h} = \mathrm{Lie}(H)$ and $\mathfrak{g} = \mathrm{Lie}(G)$, then the following diagram commutes:

$$\begin{tikzcd} H \arrow[r, hook] & G \\ \mathfrak{h} \arrow[u, "\exp"] \arrow[r, hook] & \mathfrak{g} \arrow[u, "\exp"'] \end{tikzcd}$$

This shows that $\exp_G(\mathfrak{h}) \subseteq H$, and so $\exp_G(\mathfrak{h}) = \exp_H(\mathfrak{h})$, and if H is connected, this generates H. So H is uniquely determined by $\mathfrak{h}$ as a group. Its manifold structure is also uniquely determined: we pick U, V so that the vertical arrows are an isomorphism:

$$\begin{tikzcd} e \in V \subseteq G \arrow[d, shift left, "\log"] \\ 0 \in U \subseteq \mathfrak{g} \arrow[u, shift left, "\exp"] \end{tikzcd} \tag{3.4.1.3}$$

Then $\exp(U \cap \mathfrak{h}) \underset{\log}{\overset{\sim}{\to}} U \cap \mathfrak{h}$ is an immersion into $\mathfrak{g}$, and this defines a chart around $e \in H$, which we can push to any other point $h \in H$ by multiplication by h. This determines the topology and manifold structure of H.

We turn now to the question of existence. We pick U and V as in equation (3.4.1.3), and then choose $V' \underset{\text{open}}{\subseteq} V$ and $U' \overset{\text{def}}{=} \log V'$ such that:

1. $(V')^2 \subseteq V$ and $(V')^{-1} = V'$
2. $b(x, y)$ converges on $U' \times U'$ to $\log(\exp x \exp y)$
3. $hV'h^{-1} \subseteq V$ for $h \in V'$
4. $e^{\mathrm{ad}\, x} y$ converges on $U' \times U'$ to $\log\big((\exp x)(\exp y)(\exp x)^{-1}\big)$
5. $b(x, y)$ and $e^{\mathrm{ad}\, x} y$ are elements of $h \cap U$ for $x, y \in \mathfrak{h} \cap U'$

Each condition can be independently achieved on a small enough open set. In condition 4., we extend the formal power series e^t to operators, and remark that in a neighborhood of $0 \in \mathfrak{g}$, if $h = \exp x$, then $\mathrm{Ad}\, h = e^{\mathrm{ad}\, x}$. Moreover, the following square always commutes:

$$\begin{tikzcd} G \arrow[r, "g \mapsto hgh^{-1}"] & G \\ \mathfrak{g} \arrow[u, "\exp"] \arrow[r, "\mathrm{Ad}(h)"] & \mathfrak{g} \arrow[u, "\exp"'] \end{tikzcd}$$

Thus, we define $W = \exp(\mathfrak{h} \cap U')$, which is certainly an immersed submanifold of G, as $\mathfrak{h} \cap U'$ is an open subset of the immersed submanifold $\mathfrak{h} \hookrightarrow \mathfrak{g}$. We define H to be the subgroup generated by W. Then H and W satisfy the hypotheses of Proposition 3.4.1.4. $\square$

3.4.1.4 Proposition *Following the conventions of this book, we use the word "manifold" to mean "object in a particular chosen category of sheaves of functions," and we use the word "smooth" to mean "morphism in this category." So "manifold" might mean "$\mathscr{C}^\infty$ manifold," "analytic manifold," "holomorphic manifold," etc.*

Let H be a group and $U \subseteq H$ such that $e \in U$ and U has the structure of a manifold. Assume further that the maps $U \times U \to H$, ${}^{-1} : U \to H$, and (for each h in a generating set of H) $\mathrm{Ad}(h) : U \to H$ mapping $u \mapsto huh^{-1}$ have the following properties:

1. *The preimage of $U \subseteq H$ under each map is open in the domain.*

2. *The restriction of the map to this preimage is smooth.*

Then H has a unique structure as a group manifold such that U is an open submanifold.

Proof The conditions 1. and 2. are preserved under compositions, so $\mathrm{Ad}(x)$ satisfies both conditions for any $x \in H$. Let $e \in U' \underset{\text{open}}{\subseteq} U$ so that $(U')^3 \subseteq U$ and $(U')^{-1} = U'$.

For $x \in H$, view each coset xU' as a manifold via $U' \xrightarrow{x\cdot} xU'$. For any $U'' \underset{\text{open}}{\subseteq} U'$ and $x, y \in G$, consider $yU'' \cap xU'$; as a subset of xU', it is isomorphic to $x^{-1}yU'' \cap U'$. If this set is empty, then it is open. Otherwise, $x^{-1}yu_2 = u_1$ for some $u_2 \in U''$ and $u_1 \in U'$, so $y^{-1}x = u_2u_1^{-1} \in (U')^2$ and so $y^{-1}xU' \subseteq U$. In particular, the $\{y^{-1}x\} \times U' \subseteq \mu^{-1}(U) \cap (U \times U)$. By the assumptions, $U' \to y^{-1}xU$ is smooth, and so $x^{-1}yU'' \cap U'$, the preimage of U'', is open in U'. Thus the topologies and smooth structures on xU' and yU' agree on their overlap.

In this way, we can put a manifold structure on H by declaring that $S \underset{\text{open}}{\subseteq} H$ if $S \cap xU' \underset{\text{open}}{\subseteq} xU'$ for all $x \in H$ — the topology is locally the topology of $U \ni e$, and so is Hausdorff — , and that a function f on $S \underset{\text{open}}{\subseteq} H$ is smooth if its restriction to each $S \cap xU'$ is smooth.

If we were to repeat this story with right cosets rather than left cosets we would get the a similar structure: all the left cosets xU' are compatible, and all the right cosets $U'x$ are compatible. To show that a right coset is compatible with a left coset, it suffices to show that for each $x \in H$, xU' and $U'x$ have compatible smooth structures. We consider $xU' \cap U'x \subseteq xU'$, which we transport to $U' \cap x^{-1}U'x \subseteq U'$. Since we assumed that conjugation by x was a smooth map, we see that right and left cosets are compatible.

We now need only check that the group structure is by smooth maps. We see that $(xU')^{-1} = (U')^{-1}x^{-1} = U'x^{-1}$, and multiplication is given by $\mu : xU' \times U'y \to xUy$. Left- and right-multiplication maps are smooth with respect to the left- and right-coset structures, which are compatible, and we assumed that $\mu : U' \times U' \to U$ was smooth. □

3.4.2 Review of algebraic topology

3.4.2.1 Definition *A* groupoid *is a category all of whose morphisms are invertible.*

3.4.2.2 Definition *A space X is* connected *if the only subsets of X that are both open and closed are $\emptyset$ and X.*

3.4.2.3 Definition *Let X be a space and $x, y \in X$. A* path *from x to y, which we write as $x \rightsquigarrow y$, is a continuous function $[0,1] \to X$ such that $0 \mapsto x$ and $1 \mapsto y$. Given $p : x \rightsquigarrow y$ and $q : y \rightsquigarrow z$, we define the* concatenation *$p \cdot q$ by*

$$p \cdot q(t) \stackrel{\text{def}}{=} \begin{cases} p(2t), & 0 \leq t \leq \frac{1}{2} \\ q(2t-1), & \frac{1}{2} \leq t \leq 1 \end{cases}$$

We write $x \sim y$ if there is a path connecting x to y; $\sim$ is an equivalence relation, and the equivalence classes are path components *of X. If X has only one path component, then it is* path connected.

Let A be a distinguished subset of Y and $f, g : Y \to X$ two functions that agree on A. A homotopy *$f \underset{A}{\sim} g$ relative to A is a continuous map $h : Y \times [0,1] \to X$ such that $h(0,y) = f(y)$, $h(1,y) = g(y)$, and $h(t,a) = f(a) = g(a)$ for $a \in A$. If $f \sim g$ and $g \sim h$, then $f \sim h$ by concatenation. The* fundamental groupoid *$\pi_1(X)$ of X has objects the points of X and arrows $x \to y$ the homotopy classes of paths $x \rightsquigarrow y$. We write $\pi_1(X,x)$ for the set of morphisms $x \to x$ in $\pi_1(X)$. The space X is* simply connected *if $\pi_1(X,x)$ is trivial for each $x \in X$.*

3.4.2.4 Example A path connected space is connected, but a connected space is not necessarily path connected. A path is a homotopy of maps $\{\text{pt}\} \to X$, where A is empty. $\Diamond$

3.4.2.5 Definition *Let X be a space. A* covering space *of X is a space E along with a "projection" $\pi : E \to X$ such that there is a non-empty discrete space S and a covering of X by open sets such that for each U in the covering, there exists an isomorphism $\pi^{-1}(U) \xrightarrow{\sim} S \times U$ such that the following diagram commutes:*

$$\begin{tikzcd} \pi^{-1}(U) \arrow[dr, "\pi"'] \arrow[r, phantom, "\cong"] & S \times U \arrow[d, "\text{project}"] \\ & U \end{tikzcd}$$

3.4.2.6 Proposition *Let $\pi_E : E \to X$ be a covering space.*

1. *Given any path $x \rightsquigarrow y$ and a lift $e \in \pi^{-1}(x)$, there is a unique path in E starting at e that projects to $x \rightsquigarrow y$.*

2. *Given a homotopy $\underset{A}{\sim} : Y \rightrightarrows X$ and a choice of a lift of the first arrow, there is a unique lift of the homotopy, provided Y is locally compact.*

Thus E induces a functor $E : \pi_1(X) \to \textsc{Set}$, sending $x \mapsto \pi_E^{-1}(X)$. $\square$

This Proposition, as well as Proposition 3.4.2.8, are standard in any topology course. A sketch of the proof is that you cover X so as to locally trivialize E, and lift the paths in each open set, and use compactness of $[0,1]$.

3.4.2.7 Definition *A space X is* locally path connected *if each $x \in X$ has arbitrarily small path connected neighborhoods. A space X is* locally simply connected *if it has a covering by simply connected open sets.*

3.4.2.8 Proposition *Assume that X is path connected, locally path connected, and locally simply connected. Then:*

1. *X has a simply connected covering space $\tilde{\pi} : \tilde{X} \to X$.*

2. *$\tilde{X}$ satisfies the following universal property: Given $f : X \to Y$ and a covering $\pi : E \to Y$, and given a choice of $x \in X$, an element of $\tilde{x} \in \tilde{\pi}^{-1}(x)$, and an element $e \in \pi^{-1}(f(x))$, then there exists a unique $\tilde{f} : \tilde{X} \to E$ sending $\tilde{x} \mapsto e$ such that the following diagram commutes:*

$$\begin{tikzcd} \tilde{X} \arrow[r, dotted, "\tilde{f}"] \arrow[d, "\tilde{\pi}"] & E \arrow[d, "\pi"] \\ X \arrow[r, "f"] & Y \end{tikzcd}$$

3. *If X is a manifold, so is $\tilde{X}$. If f is smooth, so is $\tilde{f}$.* □

3.4.2.9 Proposition 1. *Let G be a connected Lie group, and $\tilde{G}$ its simply-connected cover. Pick a point $\tilde{e} \in \tilde{G}$ over the identity $e \in G$. Then $\tilde{G}$ in its given manifold structure is uniquely a Lie group with identity $\tilde{e}$ such that $\tilde{G} \to G$ is a homomorphism. This induces an isomorphism of Lie algebras $\mathrm{Lie}(\tilde{G}) \overset{\sim}{\to} \mathrm{Lie}(G)$.*

2. *$\tilde{G}$ satisfies the following universal property: Given any Lie algebra homomorphism $\alpha : \mathrm{Lie}(G) \to \mathrm{Lie}(H)$, there is a unique homomorphism $\phi : \tilde{G} \to H$ inducing α.*

Proof 1. If X and Y are simply-connected, then so is $X \times Y$, and so by the universal property $\tilde{G} \times \tilde{G}$ is the universal cover of $G \times G$. We lift the functions $\mu : G \times G \to G$ and $i : G \to G$ to $\tilde{G}$ by declaring that $\tilde{\mu}(\tilde{e}, \tilde{e}) = \tilde{e}$ and that $i(\tilde{e}) = \tilde{e}$; the group axioms (equations (1.1.1.2) to (1.1.1.4)) are automatic.

2. Write $\mathfrak{g} = \mathrm{Lie}(G)$ and $\mathfrak{h} = \mathrm{Lie}(H)$, and let $\alpha : \mathfrak{g} \to \mathfrak{h}$ be a Lie algebra homomorphism. Then the graph $\mathfrak{f} \subseteq \mathfrak{g} \times \mathfrak{h}$ is a Lie subalgebra. By Theorem 3.4.1.2, $\mathfrak{f}$ corresponds to a subgroup $F \leq \tilde{G} \times H$. We check that the map $F \hookrightarrow \tilde{G} \times H \to G$ induces the map $\mathfrak{f} \to \mathfrak{g}$ on Lie algebras. F is connected and simply connected, and so by the universal property, $F \cong \tilde{G}$. Thus F is the graph of a homomorphism $\phi : \tilde{G} \to H$. □

3.5 A dictionary between algebras and groups

We have completed the proof of Theorem 3.1.2.1, the equivalence between the category of finite-dimensional Lie algebras and the category of simply-connected Lie groups, subject only to Theorem 4.4.4.15. Thus a Lie algebra includes most of the information of a Lie group. We foreshadow a dictionary, most of which we will define and develop later:

Lie Algebra $\mathfrak{g}$	Lie Group G (with $\mathfrak{g} = \mathrm{Lie}(G)$)
Subalgebra $\mathfrak{h} \leq \mathfrak{g}$	Connected Lie subgroup $H \leq G$
Homomorphism $\mathfrak{h} \to \mathfrak{g}$	$\tilde{H} \to G$ provided $\tilde{H}$ simply connected
Module/representation $\mathfrak{g} \to \mathfrak{gl}(V)$	Representation $\tilde{G} \to \mathrm{GL}(V)$ ($\tilde{G}$ simply connected)
Submodule $W \leq V$ with $\mathfrak{g} : W \to W$	Invariant subspace $G : W \to W$
$V^{\mathfrak{g}} \stackrel{\text{def}}{=} \{v \in V \text{ s.t. } \mathfrak{g}v = 0\}$	$V^{\tilde{G}} = \{v \in V \text{ s.t. } Gv = v\}$
$\mathrm{ad} : \mathfrak{g} \curvearrowright \mathfrak{g}$ via $\mathrm{ad}(x)y = [x, y]$	$\mathrm{Ad} : G \curvearrowright G$ via $\mathrm{Ad}(x)y = xyx^{-1}$
An *ideal* $\mathfrak{a}$, i.e. $[\mathfrak{g}, \mathfrak{a}] \leq \mathfrak{a}$, i.e. sub-$\mathfrak{g}$-module	A is a normal Lie subgroup, provided G is connected
$\mathfrak{g}/\mathfrak{a}$ is a Lie algebra	G/A is a Lie group only if A is closed in G
Center $Z(\mathfrak{g}) = \mathfrak{g}^{\mathfrak{g}}$	$Z_0(G)$ the identity component of center; this is closed
Derived subalgebra $\mathfrak{g}' \stackrel{\text{def}}{=} [\mathfrak{g}, \mathfrak{g}]$, an ideal	Should be commutator subgroup, but that's not closed: the closure also doesn't work, although if G is compact, then the commutator subgroup is closed.
Semidirect product $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{a}$ with $\mathfrak{h} \curvearrowright \mathfrak{a}$ and $\mathfrak{a}$ an ideal	If A and H are closed, then $A \cap H$ is discrete, and $\tilde{G} = \tilde{H} \ltimes \tilde{A}$

3.5.1 Basic examples: one- and two-dimensional Lie algebras

We classify the one- and two-dimensional Lie algebras and describe their corresponding Lie groups. We begin by working over $\mathbb{R}$.

3.5.1.1 Example The only one-dimensional Lie algebra is abelian. Its connected Lie groups are the line $\mathbb{R}$ and the circle S^1. $\lozenge$

3.5.1.2 Example There is a unique abelian two-dimensional Lie algebra, given by a basis $\{x, y\}$ with relation $[x, y] = 0$. This integrates to three possible groups: $\mathbb{R}^2$, $\mathbb{R} \times (\mathbb{R}/\mathbb{Z})$, and $(\mathbb{R}/\mathbb{Z})^2$. $\lozenge$

3.5.1.3 Example There is a unique nonabelian Lie algebra up to isomorphism, which we call $\mathfrak{b}$. It has a basis $\{x, y\}$ and defining relation $[x, y] = y$:

$$\begin{tikzcd} -x \arrow[r, "\mathrm{ad}\, y"] & y \arrow[loop below, "\mathrm{ad}\, x"] \end{tikzcd}$$

We can represent $\mathfrak{b}$ as a subalgebra of $\mathfrak{gl}(2)$ by $x = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ and $y = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$. Then $\mathfrak{b}$ exponentiates under $\exp : \mathfrak{gl}(2, \mathbb{R}) \to \mathrm{GL}(2, \mathbb{R})$ to the group

$$B = \left\{ \left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right) \text{ s.t. } a \in \mathbb{R}_+, b \in \mathbb{R} \right\}$$

We check that $B = \mathbb{R}_+ \ltimes \mathbb{R}$, and B is connected and simply connected. ◊

3.5.1.4 Lemma *A discrete normal subgroup A of a connected Lie group G is in the center. In particular, any discrete normal subgroup is abelian.* □

3.5.1.5 Corollary *The group B defined above is the only connected group with Lie algebra $\mathfrak{b}$.*

Proof Any other must be a quotient of B by a discrete normal subgroup, but the center of B is trivial. □

We turn now to the classification of one- and two-dimensional Lie algebras and Lie groups over $\mathbb{C}$. Again, there is only the abelian one-dimensional algebra, and there are two two-dimensional Lie algebras: the abelian one and the nonabelian one.

3.5.1.6 Example The simply connected abelian one-(complex-)dimensional Lie group is $\mathbb{C}$ under $+$. Any quotient factors (up to isomorphism) through the cylinder $\mathbb{C} \to \mathbb{C}^\times : z \mapsto e^z$. For any $q \in \mathbb{C}^\times$ with $|q| \neq 1$, we have a discrete subgroup $q^{\mathbb{Z}}$ of $\mathbb{C}^\times$, by which we can quotient out; we get a torus $E(q) = \mathbb{C}^\times / q^{\mathbb{Z}}$. For each q, $E(q)$ is isomorphic to $(\mathbb{R}/\mathbb{Z})^2$ as a real Lie group, but the holomorphic structure depends on q. This exhausts the one-dimensional complex Lie groups. ◊

3.5.1.7 Example The groups that integrate the abelian two-dimensional complex Lie algebra are combinations of one-dimensional Lie groups: $\mathbb{C}^2, \mathbb{C} \times E, \mathbb{C}^\times \times \mathbb{C}^\times$, etc.

In the non-abelian case, the Lie algebra $\mathfrak{b}_+ \leq \mathfrak{gl}(2, \mathbb{C})$ integrates to $B_{\mathbb{C}} \leq \mathrm{GL}(2, \mathbb{C})$ given by:

$$B_{\mathbb{C}} = \left\{ \left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right) \text{ s.t. } a \in \mathbb{C}^\times, b \in \mathbb{C} \right\} = \mathbb{C}^\times \ltimes \mathbb{C}$$

This is no longer simply connected. $\mathbb{C} \curvearrowright \mathbb{C}$ by $z \cdot w = e^z w$, and the simply-connected cover of B is

$$\tilde{B}_{\mathbb{C}} = \mathbb{C} \ltimes \mathbb{C} \quad (w, z)(w', z') \overset{\text{def}}{=} (w + e^z w', z + z')$$

This is an extension:

$$0 \to \mathbb{Z} \to \tilde{B}_{\mathbb{C}} \to B_{\mathbb{C}} \to 0$$

with the generator of $\mathbb{Z}$ being $2\pi i$. Other quotients are $\tilde{B}_{\mathbb{C}} / n\mathbb{Z}$. ◊

Exercises

1. (a) Let S be a commutative $\mathbb{K}$-algebra. Show that a linear operator $d : S \to S$ is a derivation if and only if it annihilates 1 and its commutator with the operator of multiplication by every function is the operator of multiplication by another function.

(b) Grothendieck's inductive definition of differential operators on S goes as follows: the differential operators of order zero are the operators of multiplication by functions; the space $D_{\leq n}$ of operators of order at most n is then defined inductively for $n > 0$ by $D_{\leq n} = \{d \text{ s.t. } [d, f] \in D_{\leq n-1} \text{ for all } f \in S\}$. Show that the differential operators of all orders form a filtered algebra D, and that when S is the algebra of smooth functions on an open set in $\mathbb{R}^n$ [or $\mathbb{C}^n$], D is a free left S-module with basis consisting of all monomials in the coordinate derivations $\partial/\partial x^i$.

2. Calculate all terms of degree ≤ 4 in the Baker–Campbell–Hausdorff formula (equation (3.1.2.5)).

3. Let $F(d)$ be the free Lie algebra on generators $x_1, \ldots, x_d$. It has a natural $\mathbb{N}^d$ grading in which $F(d)_{(k_1,\ldots,k_d)}$ is spanned by bracket monomials containing k_i occurences of each generator X_i. Use the PBW theorem to prove the generating function identity

$$\prod_{\mathbf{k}} \frac{1}{(1 - t_1^{k_1} \ldots t_d^{k_d})^{\dim F(d)_{(k_1,\ldots,k_d)}}} = \frac{1}{1 - (t_1 + \cdots + t_d)}$$

4. Words in the symbols $x_1, \ldots, x_d$ form a monoid under concatentation, with identity the empty word. Define a *primitive word* to be a non-empty word that is not a power of a shorter word. A *primitive necklace* is an equivalence class of primitive words under rotation. Use the generating function identity in Problem 3 to prove that the dimension of $F(d)_{k_1,\ldots,k_d}$ is equal to the number of primitive necklaces in which each symbol x_i appears k_i times.

5. A *Lyndon word* is a primitive word that is the lexicographically least representative of its primitive necklace.

   (a) Prove that w is a Lyndon word if and only if w is lexicographically less than v for every factorization $w = uv$ such that u and v are non-empty.

   (b) Prove that if $w = uv$ is a Lyndon word of length > 1 and v is the longest proper right factor of w which is itself a Lyndon word, then u is also a Lyndon word. This factorization of w is called its *right standard factorization.*

   (c) To each Lyndon word w in symbols $x_1, \ldots, x_d$ associate the bracket polynomial $p_w = x_i$ if $w = x_i$ has length 1, or, inductively, $p_w = [p_u, p_v]$, where $w = uv$ is the right standard factorization, if w has length > 1. Prove that the elements p_w form a basis of $F(d)$.

6. Prove that if q is a power of a prime, then the dimension of the subspace of total degree $k_1 + \cdots + k_q = n$ in $F(q)$ is equal to the number of monic irreducible polynomials of degree n over the field with q elements.

7. This problem outlines an alternative proof of the PBW theorem (Theorem 3.2.2.1).

   (a) Let $L(d)$ denote the Lie subalgebra of $\mathcal{T}(x_1, \ldots, x_d)$ generated by $x_1, \ldots, x_d$. Without using the PBW theorem—in particular, without using $F(d) = L(d)$—show that the value given for $\dim F(d)_{(k_1,\ldots,k_d)}$ by the generating function in Problem 3 is a lower bound for $\dim L(d)_{(k_1,\ldots,k_d)}$.

(b) Show directly that the Lyndon monomials in Problem 5(b) span $F(d)$.

(c) Deduce from (a) and (b) that $F(d) = L(d)$ and that the PBW theorem holds for $F(d)$.

(d) Show that the PBW theorem for a Lie algebra $\mathfrak{g}$ implies the PBW theorem for $\mathfrak{g}/\mathfrak{a}$, where $\mathfrak{a}$ is a Lie ideal, and so deduce PBW for all finitely generate Lie algebras from (c).

(e) Show that the PBW theorem for arbitrarty Lie algebras reduces to the finitely generated case.

8. Let $b(x, y)$ be the Baker–Campbell–Hausdorff series, i.e., $e^{b(x,y)} = e^x e^y$ in noncommuting variables x, y. Let $F(x, y)$ be its linear term in y, that is, $b(x, sy) = x + sF(x, y) + O(s^2)$.

(a) Show that $F(x, y)$ is characterized by the identity

$$\sum_{k,l\geq 0} \frac{x^k\, F(x, y)\, x^l}{(k + l + 1)!} = e^x y. \tag{3.5.1.8}$$

(b) Let λ, ρ denote the operators of left and right multiplication by x, and let f be the series in two commuting variables such that $F(x, y) = f(\lambda, \rho)(y)$. Show that

$$f(\lambda, \rho) = \frac{\lambda - \rho}{1 - e^{\rho-\lambda}}$$

(c) Deduce that

$$F(x, y) = \frac{\operatorname{ad} x}{1 - e^{-\operatorname{ad} x}}(y).$$

9. Let G be a Lie group, $\mathfrak{g} = \mathrm{Lie}(G)$, $0 \in U' \subseteq U \subseteq \mathfrak{g}$ and $e \in V' \subseteq V \subseteq G$ open neighborhoods such that exp is an isomorphism of U onto V, $\exp(U') = V'$, and $V'V' \subseteq V$. Define $\beta : U' \times U' \to U$ by $\beta(x, y) = \log(\exp(x) \exp(y))$, where $\log : V \to U$ is the inverse of exp.

(a) Show that $\beta(x, (s + t)y) = \beta(\beta(x, ty), sy)$ whenever all arguments are in U'.

(b) Show that the series $(\operatorname{ad} x)/(1-e^{-\operatorname{ad} x})$, regarded as a formal power series in the coordinates of x with coefficients in the space of linear endomorphisms of $\mathfrak{g}$, converges for all x in a neighborhood of 0 in $\mathfrak{g}$.

(c) Show that on some neighborhood of 0 in $\mathfrak{g}$, $\beta(x, ty)$ is the solution of the initial value problem

$$\beta(x, 0) = x \tag{3.5.1.9}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\beta(x, ty) = F(\beta(x, ty), y), \tag{3.5.1.10}$$

where $F(x, y) = \big((\operatorname{ad} x)/(1 - e^{-\operatorname{ad} x})\big)\,(y)$.

(d) Show that the Baker–Campbell–Hausdorff series $b(x, y)$ also satisfies the identity in part (a), as an identity of formal power series, and deduce that it is the formal power series solution to the IVP in part (c), when $F(x, y)$ is regarded as a formal series.

(e) Deduce from the above an alternative proof that $b(x,y)$ is given as the sum of a series of Lie bracket polynomials in x and y, and that it converges to $\beta(x,y)$ when evaluated on a suitable neighborhood of 0 in $\mathfrak{g}$.

(f) Use part (c) to calculate explicitly the terms of $b(x,y)$ of degree 2 in y.

10. (a) Show that the Lie algebra $\mathfrak{so}(3,\mathbb{C})$ is isomorphic to $\mathfrak{sl}(2,\mathbb{C})$.

    (b) Construct a Lie group homomorphism $\mathrm{SL}(2,\mathbb{C}) \to \mathrm{SO}(3,\mathbb{C})$ which realizes the isomorphism of Lie algebras in part (a), and calculate its kernel.

11. (a) Show that the Lie algebra $\mathfrak{so}(4,\mathbb{C})$ is isomorphic to $\mathfrak{sl}(2,\mathbb{C}) \times \mathfrak{sl}(2,\mathbb{C})$.

    (b) Construct a Lie group homomorphism $\mathrm{SL}(2,\mathbb{C}) \times \mathrm{SL}(2,\mathbb{C}) \to \mathrm{SO}(4,\mathbb{C})$ which realizes the isomorphism of Lie algebras in part (a), and calculate its kernel.

12. Show that every closed subgroup H of a Lie group G is a Lie subgroup, so that the inclusion $H \hookrightarrow G$ is a closed immersion.

13. Let G be a Lie group and H a closed subgroup. Show that G/H has a unique manifold structure such that the action of G on it is smooth.

14. Show that the intersection of two Lie subgroups H_1, H_2 of a Lie group G can be given a canonical structure of Lie subgroup so that its Lie algebra is $\mathrm{Lie}(H_1) \cap \mathrm{Lie}(H_2) \subseteq \mathrm{Lie}(G)$.

15. Find the dimension of the closed linear group $\mathrm{SO}(p,q,\mathbb{R}) \subseteq \mathrm{SL}(p+q,\mathbb{R})$ consisting of elements which preserve a non-degenerate symmetric bilinear form on $\mathbb{R}^{p+q}$ of signature (p,q). When is this group connected?

16. Show that the kernel of a Lie group homomorphism $G \to H$ is a closed subgroup of G whose Lie algebra is equal to the kernel of the induced map $\mathrm{Lie}(G) \to \mathrm{Lie}(H)$.

17. Show that if H is a normal Lie subgroup of G, then $\mathrm{Lie}(H)$ is a Lie ideal in $\mathrm{Lie}(G)$.

Chapter 4

General Theory of Lie algebras

4.1 $\mathcal{U}\mathfrak{g}$ is a Hopf algebra

4.1.0.1 Definition *A* Hopf algebra *over* $\mathbb{K}$ *is a (unital, counital) bialgebra* $(U, \mu, e, \Delta, \epsilon)$ *along with a bialgebra map* $\mathcal{S} : U \to U^{\mathrm{op}}$ *called the* antipode*, where* U^{op} *is* U *as a vector space, with the opposite multiplication and the opposite comultiplication. I.e. we define* $\mu^{\mathrm{op}} : U \otimes U \to U$ *by* $\mu^{\mathrm{op}}(x \otimes y) = \mu(y \otimes x)$*, and* $\Delta^{\mathrm{op}} : U \to U \otimes U$ *by* $\Delta^{\mathrm{op}}(x) = \sum x_{(2)} \otimes x_{(1)}$*, where* $\Delta(x) = \sum x_{(1)} \otimes x_{(2)}$*. The antipode* $\mathcal{S}$ *is required to make the following pentagons commute:*

$$\begin{tikzcd}
 & U \otimes U \arrow[rr, "1_G \otimes \mathcal{S}"] & & U \otimes U \arrow[dr, "\mu"] & \\
U \arrow[ur, "\Delta"] \arrow[rr, "\epsilon"] \arrow[dr, "\Delta"'] & & \mathbb{K} \arrow[rr, "e"] & & U \\
 & U \otimes U \arrow[rr, "\mathcal{S} \otimes 1_G"'] & & U \otimes U \arrow[ur, "\mu"'] &
\end{tikzcd} \tag{4.1.0.2}$$

In fact, the antipode $\mathcal{S}$ is uniquely determined by equation (4.1.0.2) if it exists. This is analogous to the fact that the inverse to an element in an associative algebra might not exist, but if it does exist it is unique.

4.1.0.3 Definition *An algebra* (U, μ, e) *is* commutative *if* $\mu^{\mathrm{op}} = \mu$*. A coalgebra* (U, Δ, ϵ) *is* cocommutative *if* $\Delta^{\mathrm{op}} = \Delta$.

4.1.0.4 Example Let G be a finite group and $\mathscr{C}(G)$ the algebra of functions on it. Then $\mathscr{C}(G)$ is a commutative Hopf algebra with $\Delta(f)(x, y) = f(xy)$, where we have identified $\mathscr{C}(G) \otimes \mathscr{C}(G)$ with $\mathscr{C}(G \times G)$, and $\mathcal{S}(f)(x) = f(x^{-1})$.

Let G be an algebraic group, and $\mathscr{C}(G)$ the algebra of polynomial functions on it. Then Then $\mathscr{C}(G)$ is a commutative Hopf algebra with $\Delta(f)(x, y) = f(xy)$, where we have identified $\mathscr{C}(G) \otimes \mathscr{C}(G)$ with $\mathscr{C}(G \times G)$, and $\mathcal{S}(f)(x) = f(x^{-1})$.

Let G be a group and $\mathbb{K}[G]$ the group algebra of G, with mutliplication defined by $\mu(x \otimes y) = xy$ for $x, y \in G$. Then G is a cocommutative Hopf algebra with $\Delta(x) = x \otimes x$ and $\mathcal{S}(x) = x^{-1}$ for $x \in G$.

Let $\mathfrak{g}$ be a Lie algebra and $\mathcal{U}\mathfrak{g}$ its universal enveloping algebra. We have seen already (Proposition 3.2.3.5) that $\mathcal{U}\mathfrak{g}$ is naturally a bialgebra with $\Delta(x) = x \otimes 1 + 1 \otimes x$ for $x \in \mathfrak{g}$; we make $\mathcal{U}\mathfrak{g}$ into a Hopf algebra by defining $\mathcal{S}(x) = -x$ for $x \in \mathfrak{g}$. $\Diamond$

4.1.0.5 Lemma / Definition *Let U be a cocommutative Hopf algebra. Then the antipode is an involution. Moreover, the category of (algebra-) representations of U has naturally the structure of a symmetric monoidal category with duals. In particular, to each pair of representations V, W of U, there are natural ways to make $V \otimes_{\mathbb{K}} W$ and $\mathrm{Hom}_{\mathbb{K}}(V, W)$ into U-modules. Any (di)natural functorial contruction of vector spaces — for example $V \otimes W \cong W \otimes V$, $\mathrm{Hom}(U \otimes V, W) \cong \mathrm{Hom}(U, \mathrm{Hom}(V, W))$, and $W : V \mapsto \mathrm{Hom}(\mathrm{Hom}(V, W), W)$ — in fact corresponds to a homomorphism of U-modules.*

Proof Proving the last part would require we go further into category theory than we would like. We describe the U-action on $V \otimes_{\mathbb{K}} W$ and on $\mathrm{Hom}_{\mathbb{K}}(V, W)$ when V and W are U-modules. For $u \in U$, let $\Delta(u) = \sum u_{(1)} \otimes u_{(2)} = \sum u_{(2)} \otimes u_{(1)}$, and write the actions of u on $v \in V$ and on $w \in W$ as $u \cdot v \in V$ and $u \cdot w \in W$. Let $\phi \in \mathrm{Hom}_{\mathbb{K}}(V, W)$. Then we define:

$$u \cdot (v \otimes w) \stackrel{\text{def}}{=} \sum (u_{(1)} \cdot v) \otimes (u_{(2)} \cdot w) \tag{4.1.0.6}$$

$$u \cdot \phi \stackrel{\text{def}}{=} \sum u_{(1)} \circ \phi \circ \mathcal{S}(u_{(2)}) \tag{4.1.0.7}$$

Moreover, the counit map $\epsilon : U \to \mathbb{K}$ makes $\mathbb{K}$ into U-module, and it is the unit of the monoidal structure. $\square$

4.1.0.8 Remark Equation (4.1.0.6) makes the category of U-modules into a monoidal category for any bialgebra U. One can define duals via equation (4.1.0.7), but if U is not cocommutative, then $\mathcal{S}$ may not be an involution, so a choices is required as to which variation of equation (4.1.0.7) to take — in the language of monoidal categories, this choice corresponds to the possible difference between left and right duals. Moreover, when U is not cocommutative, we do not, in general, have an isomorphism $V \otimes W \cong W \otimes V$. $\Diamond$

4.1.0.9 Example When $U = \mathcal{U}\mathfrak{g}$ and $x \in \mathfrak{g}$, then x acts on $V \otimes W$ by $v \otimes w \mapsto xv \otimes w + v \otimes w$, and on $\mathrm{Hom}(V, W)$ by $\phi \mapsto x \circ \phi - \phi \circ x$. $\Diamond$

4.1.0.10 Definition *Let (U, μ, e, ϵ) be a "counital algebra" over $\mathbb{K}$, i.e. an algebra along with an algebra map $\epsilon : U \to \mathbb{K}$ (such algebras are also called* augmented*). Thus ϵ makes $\mathbb{K}$ into a U-module. Let V be a U-module. An element $v \in V$ is U*-invariant *if the linear map $\mathbb{K} \to V$ given by $1 \mapsto v$ is a U-module homomorphism. We write V^U for the vector space of U-invariant elements of V.*

4.1.0.11 Lemma *When U is a cocommutative Hopf algebra, the space $\mathrm{Hom}_{\mathbb{K}}(V, W)^U$ of U-invariant linear maps is the same as the space $\mathrm{Hom}_U(V, W)$ of U-module homomorphisms.* $\square$

4.1.0.12 Example The $\mathcal{U}\mathfrak{g}$-invariant elements of a $\mathfrak{g}$-module V form the set

$$V^{\mathfrak{g}} = \{v \in V \text{ s.t. } x \cdot v = 0 \, \forall x \in \mathfrak{g}\}.$$

We shorten the word "$\mathcal{U}\mathfrak{g}$-invariant" to "$\mathfrak{g}$-invariant." A linear map $\phi \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ is $\mathfrak{g}$-invariant if and only if $x \circ \phi = \phi \circ x$ for every $x \in \mathfrak{g}$. $\lozenge$

4.1.0.13 Definition *The* center *of a Lie algebra $\mathfrak{g}$ is the space of $\mathfrak{g}$-invariant elements of $\mathfrak{g}$ under the adjoint action:*

$$Z(\mathfrak{g}) \stackrel{\text{def}}{=} \mathfrak{g}^{\mathfrak{g}} = \{x \in \mathfrak{g} \textit{ s.t. } [\mathfrak{g}, x] = 0\}.$$

4.2 Structure theory of Lie algebras

4.2.1 Many definitions

As always, we write "$\mathfrak{g}$-module" for "$\mathcal{U}\mathfrak{g}$-module."

4.2.1.1 Definition *A $\mathfrak{g}$-module V is* simple *or* irreducible *if there is no submodule $W \subseteq V$ with $0 \neq W \neq V$. A Lie algebra is* simple *if it is simple as a $\mathfrak{g}$-module under the adjoint action. An* ideal *of $\mathfrak{g}$ is a $\mathfrak{g}$-submodule of $\mathfrak{g}$ under the adjoint action.*

An application of the Jacobi identity gives:

4.2.1.2 Lemma *If $\mathfrak{a}$ and $\mathfrak{b}$ are ideals in $\mathfrak{g}$, then so is $[\mathfrak{a}, \mathfrak{b}]$.* $\square$

4.2.1.3 Definition *The* upper central series *of a Lie algebra $\mathfrak{g}$ is the series $\mathfrak{g} \geq \mathfrak{g}_1 \geq \mathfrak{g}_2 \geq \dots$ where $\mathfrak{g}_0 \stackrel{\text{def}}{=} \mathfrak{g}$ and $\mathfrak{g}_{n+1} \stackrel{\text{def}}{=} [\mathfrak{g}, \mathfrak{g}_n]$. The Lie algebra $\mathfrak{g}$ is* nilpotent *if $\mathfrak{g}_n = 0$ for some n.*

4.2.1.4 Definition *The* derived subalgebra *of a Lie algebra $\mathfrak{g}$ is the algebra $\mathfrak{g}' \stackrel{\text{def}}{=} [\mathfrak{g}, \mathfrak{g}]$. The* derived series *of $\mathfrak{g}$ is the series $\mathfrak{g} \geq \mathfrak{g}' \geq \mathfrak{g}'' \geq \dots$ given by $\mathfrak{g}^{(0)} \stackrel{\text{def}}{=} \mathfrak{g}$ and $\mathfrak{g}^{(n+1)} \stackrel{\text{def}}{=} [\mathfrak{g}^{(n)}, \mathfrak{g}^{(n)}]$. The Lie algebra $\mathfrak{g}$ is* solvable *if $\mathfrak{g}^{(n)} = 0$ for some n. An ideal $\mathfrak{r}$ in $\mathfrak{g}$ is solvable if it is solvable as a subalgebra. By Lemma 4.2.1.2, if $\mathfrak{r}$ is an ideal of $\mathfrak{g}$, then so is $\mathfrak{r}^{(n)}$.*

4.2.1.5 Example The Lie algebra of upper-triangular matrices in $\mathfrak{gl}(n)$ is solvable. A converse to this statement is Corollary 4.2.3.5. The Lie algebra of strictly upper triangular matrices is nilpotent. $\lozenge$

4.2.1.6 Definition *A Lie algebra $\mathfrak{g}$ is* semisimple *if its only solvable ideal is 0.*

4.2.1.7 Remark If $\mathfrak{r}$ is a solvable ideal of $\mathfrak{g}$ with $\mathfrak{r}^{(n)} = 0$, then $\mathfrak{r}^{(n-1)}$ is abelian. Conversely, any abelian ideal of $\mathfrak{g}$ is solvable. Thus it is equivalent to replace the word "solvable" in Definition 4.2.1.6 with the word "abelian". $\lozenge$

4.2.1.8 Lemma *Any nilpotent Lie algebra is solvable. A non-zero nilpotent Lie algebra has non-zero center.*

Proof Clearly $\mathfrak{g}_n \supseteq \mathfrak{g}^{(n)}$. Let $\mathfrak{g} \neq 0$ be nilpotent and $m \in \mathbb{N}$ the largest number such that $\mathfrak{g}_m \neq 0$. Then $[\mathfrak{g}, \mathfrak{g}_m] = 0$, so $\mathfrak{g}_m \subseteq Z(\mathfrak{g})$. □

4.2.1.9 Proposition *A subquotient of a solvable Lie algebra is solvable. A subquotient of a nilpotent algebra is nilpotent. If $\mathfrak{a}$ is an ideal of $\mathfrak{g}$ and if $\mathfrak{a}$ and $\mathfrak{g}/\mathfrak{a}$ are both solvable, then $\mathfrak{g}$ is solvable. If $\mathfrak{a}$ is an ideal of $\mathfrak{g}$ and $\mathfrak{a}$ is nilpotent and if $\mathfrak{g} \curvearrowright \mathfrak{a}$ nilpotently, then $\mathfrak{g}$ is nilpotent. Thus a central extension of a nilpotent Lie algebra is nilpotent.*

Proof The derived and upper central series of subquotients are subquotients of the derived and upper central series. For the second statement, we start taking the derived series of $\mathfrak{g}$, eventually landing in $\mathfrak{a}$ (since $\mathfrak{g}/\mathfrak{a} \to 0$), which is solvable. The nilpotent claim is similar.

4.2.1.10 Example Let $\mathfrak{g} = \langle x, y : [x, y] = y \rangle$ be the two-dimensional nonabelian Lie algebra. Then $\mathfrak{g}^{(1)} = \langle y \rangle$ and $\mathfrak{g}^{(2)} = 0$, but $\mathfrak{g}_2 = [\mathfrak{g}, \langle y \rangle] = \langle y \rangle$ so $\mathfrak{g}$ is solvable but not nilpotent.

4.2.1.11 Definition *The* lower central series *of a Lie algebra $\mathfrak{g}$ is the series $0 \leq Z(\mathfrak{g}) \leq \mathfrak{z}_2 \leq \dots$ defined by $\mathfrak{z}_0 = 0$ and $\mathfrak{z}_{k+1} = \{x \in \mathfrak{g} \text{ s.t. } [\mathfrak{g}, x] \subseteq \mathfrak{z}_k\}$.*

4.2.1.12 Proposition *For any of the derived series, the upper central series, and the lower central series, quotients of consecutive terms are abelian.* □

4.2.1.13 Proposition *Let $\mathfrak{g}$ be a Lie algebra and $\{\mathfrak{z}_k\}$ its lower central series. Then $\mathfrak{z}_n = \mathfrak{g}$ for some n if and only if $\mathfrak{g}$ is nilpotent.* □

4.2.2 Nilpotency: Engel's theorem and corollaries

4.2.2.1 Lemma / Definition *A matrix $x \in \mathrm{End}(V)$ is* nilpotent *if $x^n = 0$ for some n. A Lie algebra $\mathfrak{g}$* acts by nilpotents *on a vector space V if for each $x \in \mathfrak{g}$, its image under $\mathfrak{g} \to \mathrm{End}(V)$ is nilpotent. If $\mathfrak{g} \curvearrowright V, W$ by nilpotents, then $\mathfrak{g} \curvearrowright V \otimes W$ and $g \curvearrowright \mathrm{Hom}(V, W)$ by nilpotents. If $v \in V$ and $\mathfrak{g} \curvearrowright V$, define the* annihilator *of v to be $\mathrm{ann}_{\mathfrak{g}}(v) = \{x \in \mathfrak{g} \text{ s.t. } xv = 0\}$. For any $v \in V$, $\mathrm{ann}_{\mathfrak{g}}(v)$ is a Lie subalgebra of $\mathfrak{g}$.* □

4.2.2.2 Theorem (Engel's Theorem)
If $\mathfrak{g}$ is a finite-dimensional Lie algebra acting on V (possibly infinite-dimensional) by nilpotent endomorphisms, and $V \neq 0$, then there exists a non-zero vector $v \in V$ such that $\mathfrak{g}v = 0$.

Proof It suffices to look at the image of $\mathfrak{g}$ in $\mathfrak{gl}(V) = \mathrm{Hom}(V, V)$. Then $\mathrm{ad} : \mathfrak{g} \curvearrowright \mathfrak{g}$ is by nilpotents.

Pick $v_0 \in V$ so that $\mathrm{ann}_{\mathfrak{g}}(v_0)$ has maximal dimension and let $\mathfrak{h} = \mathrm{ann}_{\mathfrak{g}}(v_0)$. It suffices to show that $\mathfrak{h} = \mathfrak{g}$. Suppose to the contrary that $\mathfrak{h} \subsetneq \mathfrak{g}$. By induction on dimension, the theorem holds for $\mathfrak{h}$. Consider the vector space $\mathfrak{g}/\mathfrak{h}$; then $\mathfrak{h} \curvearrowright \mathfrak{g}/\mathfrak{h}$ by nilpotents, so we can find $x \in \mathfrak{g}/\mathfrak{h}$ nonzero with $\mathfrak{h}x = 0$. Let $\hat{x}$ be a preimage of x in $\mathfrak{g}$. Then $\hat{x} \in \mathfrak{g} \smallsetminus \mathfrak{h}$ and $[\mathfrak{h}, \hat{x}] \subseteq \mathfrak{h}$. Then $\mathfrak{h}_1 \stackrel{\text{def}}{=} \langle \hat{x} \rangle + \mathfrak{h}$ is a subalgebra of $\mathfrak{g}$.

The space $U \stackrel{\text{def}}{=} \{u \in V \text{ s.t. } \mathfrak{h}u = 0\}$ is non-zero, since $v_0 \in U$. We see that U is an $\mathfrak{h}_1$-submodule of $\mathfrak{h}_1 \curvearrowright V$: $hu = 0 \in U$ for $h \in \mathfrak{h}$, and $h(\hat{x}u) = [h, \hat{x}]u + \hat{x}hu = 0u + \hat{x}0 = 0$ so $\hat{x}u \in U$. All of $\mathfrak{g}$ acts on all of V by nilpotents, so in particular $x|_U$ is nilpotent, and so there is some vector $v_1 \in U$ with $xv_1 = 0$. But then $\mathfrak{h}_1 v_1 = 0$, contradicting the maximality of $\mathfrak{h} = \mathrm{ann}(v_0)$. □

4.2.2.3 Corollary *1. If $\mathfrak{g} \curvearrowright V$ by nilpotents and V is finite dimension, then V has a basis in which $\mathfrak{g}$ is strictly upper triangular.*

2. If $\operatorname{ad} x$ is nilpotent for all $x \in \mathfrak{g}$ finite-dimensional, then $\mathfrak{g}$ is a nilpotent Lie algebra.

3. Let V be a simple $\mathfrak{g}$-module. If an ideal $\mathfrak{a} \leq \mathfrak{g}$ acts nilpotently on V then $\mathfrak{a}$ acts as 0 on V. □

4.2.2.4 Lemma / Definition *If V is a finite-dimensional $\mathfrak{g}$ module, then there exists a* Jordan-Holder series *$0 = M_0 < M_1 < M_2 < \cdots < M_n = V$ such that each M_i is a $\mathfrak{g}$-submodule and each M_{i+1}/M_i is simple.*

Proof Pick $M'_{i+1} < V/M_i$ a submodule of minimal dimension; it is automatically simple. Set $M_{i+1} < V$ the preimage of M'_{i+1}. □

4.2.2.5 Corollary *Let V be a finite-dimensional $\mathfrak{g}$-module and $0 = M_0 < M_1 < M_2 < \cdots < M_n = V$ a Jordan-Holder series for V. An ideal $\mathfrak{a} \leq \mathfrak{g}$ acts by nilpotents on V if and only if $\mathfrak{a}$ acts by 0 on each M_{i+1}/M_i. Thus there is a largest ideal of $\mathfrak{g}$ that acts by nilpotents on V.* □

4.2.2.6 Definition *The largest ideal of $\mathfrak{g}$ that acts by nilpotents on V is the* nilpotency ideal *of the action $\mathfrak{g} \curvearrowright V$.*

4.2.2.7 Proposition *Any nilpotent ideal $\mathfrak{a} \leq \mathfrak{g}$ acts nilpotently on $\mathfrak{g}$.* □

4.2.2.8 Corollary *Any finite-dimensional Lie algebra has a largest nilpotent ideal: the nilpotency ideal of* ad. □

4.2.2.9 Remark Not every ad-nilpotent element of a Lie algebra is necessarily in the nilpotency ideal of ad. ◊

4.2.2.10 Definition *Let $\mathfrak{g}$ be a Lie algebra and V a finite-dimensional $\mathfrak{g}$-module. Then V defines a* trace form *β_V: a symmetric bilinear form on $\mathfrak{g}$ given by $\beta_V(x, y) \overset{\text{def}}{=} \operatorname{tr}_V(x, y)$. The* radical *or* kernel *of β_V is the set $\ker \beta_V \overset{\text{def}}{=} \{x \in \mathfrak{g} \text{ s.t. } \beta_V(x, \mathfrak{g}) = 0\}$.*

4.2.2.11 Remark The more standard notation seems to be $\operatorname{rad} \beta$ for what we call $\ker \beta$. We prefer the term "kernel" largely to avoid the conflict of notation with Lemma/Definition 4.2.3.1. Any bilinear form β on $\mathfrak{g}$ defines two linear maps $\mathfrak{g} \to \mathfrak{g}^*$, where $\mathfrak{g}^*$ is the dual vector space to $\mathfrak{g}$, given by $x \mapsto \beta(x, -)$ and $x \mapsto \beta(-, x)$. Of course, when β is symmetric, these are the same map, and we can unambiguously call the map $\beta : \mathfrak{g} \to \mathfrak{g}^*$. Then $\ker \beta_V$ defined above is precisely the kernel of the map $\beta_V : \mathfrak{g} \to \mathfrak{g}^*$. ◊

The following proposition follows from considering Jordan-Holder series:

4.2.2.12 Proposition *If and ideal $\mathfrak{a} \leq \mathfrak{g}$ of a finite-dimensional Lie algebra acts nilpotently on a finite-dimensional vector space V, then $\mathfrak{a} \leq \ker \beta_V$.* □

4.2.3 Solvability: Lie's theorem and corollaries

4.2.3.1 Lemma / Definition *Let $\mathfrak{g}$ be a finite-dimensional Lie algebra. Then $\mathfrak{g}$ has a largest solvable ideal, the* radical $\operatorname{rad}\mathfrak{g}$.

Proof If ideals $\mathfrak{a}, \mathfrak{b} \leq \mathfrak{g}$ are solvable, then $\mathfrak{a} + \mathfrak{b}$ is solvable, since we have an exact sequence of $\mathfrak{g}$-modules

$$0 \to \mathfrak{a} \to \mathfrak{a} + \mathfrak{b} \to (\mathfrak{a} + \mathfrak{b})/\mathfrak{a} \to 0$$

which is also an extension of a solvable algebra (a quotient of $\mathfrak{b}$) by a solvable ideal. □

4.2.3.2 Theorem (Lie's Theorem)
Let $\mathfrak{g}$ be a finite-dimensional solvable Lie algebra over $\mathbb{K}$ of characteristic 0, and V a non-zero $\mathfrak{g}$-module. Assume that $\mathbb{K}$ contains eigenvalues of the actions of all $x \in \mathfrak{g}$. Then V has a one-dimensional $\mathfrak{g}$-submodule.

Proof Without loss of generality $\mathfrak{g} \neq 0$; then $\mathfrak{g}' \neq \mathfrak{g}$ by solvability. Pick any $\mathfrak{g} \geq \mathfrak{h} \geq \mathfrak{g}'$ a codimension-1 subspace. Since $\mathfrak{h} \geq \mathfrak{g}'$, $\mathfrak{h}$ is an ideal of $\mathfrak{g}$. Pick $x \in \mathfrak{g} \setminus \mathfrak{h}$, whence $\mathfrak{g} = \langle x \rangle + \mathfrak{h}$.

Being a subalgebra of $\mathfrak{g}$, $\mathfrak{h}$ is solvable, and by induction on dimension $\mathfrak{h} \curvearrowright V$ has a one-dimensional $\mathfrak{h}$-submodule $\langle w \rangle$. Thus there is some linear map $\lambda : \mathfrak{h} \to \mathbb{K}$ so that $h \cdot w = \lambda(h) w$ for each $h \in \mathfrak{h}$. Let $W = \mathbb{K}[x]w$ for $x \in \mathfrak{g} \setminus \mathfrak{h}$ as above. Then $W = \mathcal{U}(\mathfrak{g})w$, as $\mathfrak{g} = \mathfrak{h} + \langle x \rangle$ and $\mathfrak{h}w \subseteq \mathbb{K}w$.

By induction on m, each $\langle 1, x, \ldots, x^m \rangle w$ is an $\mathfrak{h}$-submodule of W:

$$h(x^m w) = x^m hw + \sum_{k+l=m-1} x^k [h, x] x^l w \tag{4.2.3.3}$$

$$= \lambda(h) x^m w + x^k h' x^l w \tag{4.2.3.4}$$

where $h' = [h, x] \in \mathfrak{h}$. Thus $h' x^l w \in \langle 1, \ldots, x^l \rangle w$ by induction, and so $x^k h' x^l w \in \langle 1, \ldots, x^{k+l} \rangle w = \langle 1, \ldots, x^{m-1} \rangle w$.

Moreover, we see that W is a generalized eigenspace with eigenvalue $\lambda(h)$ for all $h \in \mathfrak{h}$, and so $\operatorname{tr}_W h = (\dim W)\lambda(h)$, by working in a basis where h is upper triangular. But for any a, b, $\operatorname{tr}[a, b] = 0$; thus $\operatorname{tr}_W [h, x] = 0$ so $\lambda([h, x]) = 0$. Then equations (4.2.3.3) to (4.2.3.4) and induction on m show that W is an actual eigenspace.

Thus we can pick $v \in W$ an eigenvector of x, and then v generates a one-dimensional eigenspace of $x + \mathfrak{h} = \mathfrak{g}$, i.e. a one-dimensional $\mathfrak{g}$-submodule. □

4.2.3.5 Corollary *Let $\mathfrak{g}$ and V satisfy the conditions of Theorem 4.2.3.2. Then V has a basis in which $\mathfrak{g}$ is upper-diagonal.* □

4.2.3.6 Corollary *Let $\mathfrak{g}$ be a solvable finite-dimensional Lie algebra over an algebraically closed field of characteristic 0. Then every simple finite-dimensional $\mathfrak{g}$-module is one-dimensional.* □

4.2.3.7 Corollary *Let $\mathfrak{g}$ be a solvable finite-dimensional Lie algebra over a field of characteristic 0. Then $\mathfrak{g}'$ acts nilpotently on any finite-dimensional $\mathfrak{g}$-module.* □

4.2.3.8 Remark In spite of the condition on the ground field in Theorem 4.2.3.2, Corollary 4.2.3.7 is true over any field of characteristic 0. Indeed, let $\mathfrak{g}$ be a Lie algebra over $\mathbb{K}$ and $\mathbb{K} \leq \mathbb{L}$ a field extension. The upper central, lower central, and derived series are all preserved under $\mathbb{L}\otimes_{\mathbb{K}}$, so $\mathbb{L} \otimes_{\mathbb{K}} \mathfrak{g}$ is solvable if and only if $\mathfrak{g}$ is. Moreover, $\mathfrak{g} \curvearrowright V$ nilpotently if and only if $\mathbb{L} \otimes_{\mathbb{K}} \mathfrak{g} \curvearrowright \mathbb{L} \otimes_{\mathbb{K}} V$ nilpotently. Thus we may as well "extend by scalars" to an algebraically closed field. $\Diamond$

4.2.3.9 Corollary *Corollary 4.2.2.8 asserts that any Lie algebra $\mathfrak{g}$ has a largest ideal that acts nilpotently on $\mathfrak{g}$. When $\mathfrak{g}$ is solvable, then any element of $\mathfrak{g}'$ is* ad*-nilpotent. Hence the set of* ad*-nilpotent elements of $\mathfrak{g}$ is an ideal.* $\square$

4.2.4 The Killing form

4.2.4.1 Proposition *Let $\mathfrak{g}$ be a Lie algebra and V a finite-dimensional $\mathfrak{g}$-module. The trace form $\beta_V : (x, y) \mapsto \mathrm{tr}_V(xy)$ on $\mathfrak{g}$ is invariant under the $\mathfrak{g}$-action:*

$$\beta_V\big([z, x], y\big) + \beta_V\big(x, [z, y]\big) = 0 \qquad \square$$

4.2.4.2 Definition *Let $\mathfrak{g}$ be a finite-dimensional Lie algebra. The* Killing form $\beta \overset{\text{def}}{=} \beta_{(\mathfrak{g},\mathrm{ad})}$ *on $\mathfrak{g}$ is the trace form of the adjoint representation $\mathfrak{g} \curvearrowright \mathfrak{g}$.*

4.2.4.3 Proposition *Suppose V is a $\mathfrak{g}$-module and $W \subseteq V$ is a $\mathfrak{g}$-submodule. Then $\beta_V = \beta_W + \beta_{V/W}$.* $\square$

4.2.4.4 Corollary *Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{a} \leq \mathfrak{g}$ an ideal. Then $\beta_{(\mathfrak{g}/\mathfrak{a},\mathrm{ad})}|_{\mathfrak{a}\times\mathfrak{g}} = 0$, so $\beta|_{\mathfrak{a}\times\mathfrak{g}} = \beta_{\mathfrak{a}}|_{\mathfrak{a}\times\mathfrak{g}}$. In particular, the Killing form of $\mathfrak{a}$ is $\beta|_{\mathfrak{a}\times\mathfrak{a}}$.* $\square$

4.2.4.5 Proposition *Let V be a $\mathfrak{g}$-module of a Lie algebra $\mathfrak{g}$. Then* $\ker \beta_V$ *is an ideal of $\mathfrak{g}$.*

Proof The invariance of β_V implies that the map $\beta_V : \mathfrak{g} \to \mathfrak{g}^*$ given by $x \mapsto \beta_V(x, -)$ is a $\mathfrak{g}$-module homomorphism, whence $\ker \beta_V$ is a submodule a.k.a. and ideal of $\mathfrak{g}$. $\square$

The following is a corollary to Theorem 4.2.2.2, using the Jordan-form decomposition of matrices:

4.2.4.6 Proposition *Let $\mathfrak{g}$ be a Lie algebra, V a $\mathfrak{g}$-module, and $\mathfrak{a}$ an ideal of $\mathfrak{g}$ that acts nilpotently on V. Then $\mathfrak{a} \subseteq \ker \beta_V$.* $\square$

4.2.4.7 Corollary *If the Killing form β of a Lie algebra $\mathfrak{g}$ is* nondegenerate *(i.e. if $\ker \beta = 0$), then $\mathfrak{g}$ is semisimple.* $\square$

4.2.5 Jordan form

4.2.5.1 Theorem (Jordan decomposition)
Let V be a finite-dimensional vector space over an algebraically closed field $\mathbb{K}$. Then:

1. *Every $a \in \mathfrak{gl}(V)$ has a unique* Jordan decomposition *$a = s + n$, where s is diagonalizable, n is nilpotent, and they commute.*

2. $s, n \in \mathbb{K}[a]$, in the sense that they are linear combinations of powers of a.

Proof We write a in Jordan form; since strictly-upper-triangular matrices are nilpotent, existence of a Jordan decomposition of a is guaranteed. In particular, the diagonal part s clearly commutes with a, and hence with $n = a - s$. We say this again more specifically, showing that s, n constructed this way live in the polynomial subalgera $\mathbb{K}[a] \subset \mathfrak{gl}(V)$ generated by a:

Let the characteristic polynomial of a be $\prod_i (x - \lambda_i)^{n_i}$. In particular, $(x - \lambda_i)$ are relatively prime, so by the Chinese Remainder Theorem, there is a polynomial f such that $f(x) = \lambda_i \mod (x - \lambda_i)^{n_i}$. Choose a basis of V in which a is in Jordan form; since restricting to a Jordan block b of a is an algebra homomorphism $\mathbb{K}[a] \twoheadrightarrow \mathbb{K}[b]$, we can compute $f(a)$ block-by-block. Let b be a block of a with eigenvalue λ_i. Then $(b - \lambda_i)^{n_i} = 0$, so $f(b) = \lambda_i$. Thus $s = f(a)$ is diagonal in this basis, and $n = a - f(a)$ is nilpotent.

For uniqueness in part 1., let $x = n' + s'$ be any other Jordan decomposition of a. Then n' and s' commute with a and hence with any polynomial in a, and in particular n' commutes with n and s' commutes with s. But $n' + s' = a = n + s$, so $n' - n = s' - s$. Since everything commutes, $n' - n$ is nilpotent and $s' - s$ is diagonalizable, but the only nilpotent diagonal is 0. □

4.2.5.2 Remark Even though for each a, its diagonalizable and nilpotent pieces s, n live in $\mathbb{K}[a]$, the decomposition $a = s + n$ is not polynomial: as a varies, s and n need not depend polynomially on a. ♢

We now move to an entirely unmotivated piece of linear algebra:

4.2.5.3 Lemma *Let V be a finite-dimensional vector space over an algebraically closed field $\mathbb{K}$ of characteristic 0. Let $B \subseteq A \subseteq \mathfrak{gl}(V)$ be any subspaces, and define $T = \{x \in \mathfrak{gl}(V) : [x, A] \subseteq B\}$. If $t \in T$ satisfies $\mathrm{tr}_v(tu) = 0 \ \forall u \in T$, then t is nilpotent.*

We can express this as follows: Let β_V be the trace form on $\mathfrak{gl}(V) \curvearrowright V$. Then $\ker \beta_V|_{T \times T}$ consists of nilpotents.

Proof Let $t = s + n$ be the Jordan decomposition; we wish to show that $s = 0$. We fix a basis $\{e_i\}$ in which s is diagonal: $se_i = \lambda_i e_i$. Let $\{e_{ij}\}$ be the corresponding basis of matrix units for $\mathfrak{gl}(V)$. Then $(\mathrm{ad}\, s)e_{ij} = (\lambda_i - \lambda_j)e_{ij}$.

Now let $\Lambda = \mathbb{Q}\{\lambda_i\}$ be the finite-dimensional $\mathbb{Q}$-vector-subspace of $\mathbb{K}$. We consider an arbitrary $\mathbb{Q}$-linear functional $f : \Lambda \to \mathbb{Q}$; we will show that $f = 0$, and hence that $\Lambda = 0$.

By $\mathbb{Q}$-linearity, $f(\lambda_i) - f(\lambda_j) = f(\lambda_i - \lambda_j)$, and we chose a polynomial $p(x) \in \mathbb{K}[x]$ so that $p(\lambda_i - \lambda_j) = f(\lambda_i) - f(\lambda_j)$; in particular, $p(0) = 0$.

Now we define $u \in \mathfrak{gl}(V)$ by $ue_i = f(\lambda_i)e_i$, and then $(\mathrm{ad}\, u)e_{ij} = (f(\lambda_i) - f(\lambda_j))e_{ij} = p(\mathrm{ad}\, s)e_{ij}$. So $\mathrm{ad}\, u = p(\mathrm{ad}\, s)$.

Since $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is a Lie algebra homomorphism, $\mathrm{ad}\, t = \mathrm{ad}\, s + \mathrm{ad}\, n$, and $\mathrm{ad}\, s, \mathrm{ad}\, n$ commute, and $\mathrm{ad}\, s$ is diagonalizable and $\mathrm{ad}\, n$ is nilpotent. So $\mathrm{ad}\, s + \mathrm{ad}\, n$ is the Jordan decomposition of $\mathrm{ad}\, t$, and hence $\mathrm{ad}\, s = q(\mathrm{ad}\, t)$ for some polynomial $q \in \mathbb{K}[x]$. Then $\mathrm{ad}\, u = (p \circ q)(\mathrm{ad}\, t)$, and since every power of t takes A into B, we have $(\mathrm{ad}\, u)A \subseteq B$, so $u \in T$.

But by construction u is diagonal in the $\{e_i\}$ basis and t is upper-triangular, so tu is upper-triangular with diagonal $\mathrm{diag}(\lambda_i f(\lambda_i))$. Thus $0 = \mathrm{tr}(tu) = \sum \lambda_i f(\lambda_i)$. We apply f to this: $0 = \sum (f(\lambda_i))^2 \in \mathbb{Q}$, so $f(\lambda_i) = 0$ for each i. Thus $f = 0$. □

4.2.6 Cartan's criteria

4.2.6.1 Proposition *Let V be a finite-dimensional vector space over a field $\mathbb{K}$ of characteristic 0. Then a subalgebra $\mathfrak{g} \le \mathfrak{gl}(V)$ is solvable if and only if $\beta_V(\mathfrak{g}, \mathfrak{g}') = 0$, i.e. $\mathfrak{g}' \le \ker \beta_V$.*

Proof We can extend scalars and assume that $\mathbb{K}$ is algebraically closed.

The forward direction follows by Lie's theorem (Theorem 4.2.3.2): we can find a basis of V in which $\mathfrak{g}$ acts by upper-triangular matrices, and hence $\mathfrak{g}'$ acts by strictly upper-triangular matrices.

For the reverse, we'll show that $\mathfrak{g}'$ acts nilpotently, and hence is nilpotent by Engel's theorem (Theorem 4.2.2.2). We use Lemma 4.2.5.3, taking $V = V$, $A = \mathfrak{g}$, and $B = \mathfrak{g}'$. Then $T = \{t \in \mathfrak{gl}(V) \text{ s.t. } [t, \mathfrak{g}] \le \mathfrak{g}'\}$, and in particular $\mathfrak{g} \le T$, and so $\mathfrak{g}' \le T$.

So if $[x, y] = t \in \mathfrak{g}'$, then $\mathrm{tr}_V(tu) = \mathrm{tr}_V([x, y]u) = \mathrm{tr}_V(y[x, u])$ by invariance, and $y \in \mathfrak{g}$ and $[x, u] \in \mathfrak{g}'$ so $\mathrm{tr}_V(y[x, u]) = 0$. Hence t is nilpotent. $\square$

The following is a straightforward corollary:

4.2.6.2 Theorem (Cartan's First Criterion)
Let $\mathfrak{g}$ be a Lie algebra over a field of characteristic 0. Then $\mathfrak{g}$ is solvable if and only if $\mathfrak{g}' \le \ker \beta$.

Proof We have not yet proven Theorem 4.5.0.10, so we cannot assume that $\mathfrak{g} \hookrightarrow \mathfrak{gl}(V)$ for some V. Rather, we let $V = \mathfrak{g}$ and $\tilde{\mathfrak{g}} = \mathfrak{g}/Z(\mathfrak{g})$, whence $\tilde{\mathfrak{g}} \hookrightarrow \mathfrak{gl}(V)$ by the adjoint action. Then $\mathfrak{g}$ is a central extension of $\tilde{\mathfrak{g}}$, so by Proposition 4.2.1.9 $\mathfrak{g}$ is solvable if and only if $\tilde{\mathfrak{g}}$ is. By Proposition 4.2.6.1, $\tilde{\mathfrak{g}}$ is solvable if and only if $\beta_{\tilde{\mathfrak{g}}}(\tilde{\mathfrak{g}}, \tilde{\mathfrak{g}}') = 0$. But $\beta_{\mathfrak{g}}$ factors through $\beta_{\tilde{\mathfrak{g}}}$:

$$\beta_{\mathfrak{g}} = \{\mathfrak{g} \times \mathfrak{g} \overset{/Z(\mathfrak{g})}{\longrightarrow} \tilde{\mathfrak{g}} \times \tilde{\mathfrak{g}} \overset{\beta_{\tilde{\mathfrak{g}}}}{\to} \mathbb{K}\}$$

Moreover, $\mathfrak{g}' \overset{/Z(\mathfrak{g})}{\twoheadrightarrow} \tilde{\mathfrak{g}}'$, and so $\beta_{\mathfrak{g}}(\mathfrak{g}, \mathfrak{g}') = \beta_{\tilde{\mathfrak{g}}}(\tilde{\mathfrak{g}}, \tilde{\mathfrak{g}}')$. $\square$

4.2.6.3 Corollary *For any Lie algebra $\mathfrak{g}$ in characteristic zero with Killing form β, we have that $\ker \beta$ is solvable, i.e. $\ker \beta \le \mathrm{rad}\, \mathfrak{g}$.* $\square$

The reverse direction of the following is true in any characteristic (Corollary 4.2.4.7). The forward direction is an immediate corollary of Corollary 4.2.6.3.

4.2.6.4 Theorem (Cartan's Second Criterion)
Let $\mathfrak{g}$ be a Lie algebra over characteristic 0, and β its Killing form. Then $\mathfrak{g}$ is semisimple if and only if $\ker \beta = 0$. $\square$

4.2.6.5 Corollary *Let $\mathfrak{g}$ be a Lie algebra over characteristic 0. Then $\mathfrak{g}$ is semisimple if and only if any extension by scalars of $\mathfrak{g}$ is semisimple.* $\square$

4.2.6.6 Remark For any Lie algebra $\mathfrak{g}$, $\mathfrak{g}/\mathrm{rad}\, \mathfrak{g}$ is semisimple. We will see in Theorem 4.4.4.12 that in characteristic 0, $\mathrm{rad}\, \mathfrak{g}$ is a direct summand of $\mathfrak{g}$. $\Diamond$

4.3 Examples: three-dimensional Lie algebras

The classification of three-dimensional Lie algebras over $\mathbb{R}$ or $\mathbb{C}$ is long but can be done by hand [Bia]. The classification of four-dimensional Lie algebras has been completed, but beyond this it is hopeless: there are too many extensions of one algebra by another. In Chapter 5 we will classify all semisimple Lie algebras. For now we list two important Lie algebras:

4.3.0.1 Lemma / Definition *The* Heisenberg algebra *is a three-dimensional Lie algebra with a basis x, y, z, in which z is central and $[x, y] = z$. The Heisenberg algebra is nilpotent.* □

4.3.0.2 Lemma / Definition *We define $\mathfrak{sl}(2)$ to be the three-dimensional Lie algebra with a basis e, h, f and relations $[h, e] = 2e$, $[h, f] = -2f$, and $[e, f] = h$. So long as we are not working over characteristic 2, $\mathfrak{sl}(2)$ is semisimple; simplicity follows from Corollary 4.3.0.5.*

Proof Just compute the Killing form $\beta_{\mathfrak{sl}(2)}$. □

We conclude this section with two propositions and two corollaries; these will play an important role in Chapter 5.

4.3.0.3 Proposition *Let $\mathfrak{g}$ be a Lie algebra such that every (proper) ideal $\mathfrak{a}$ of $\mathfrak{g}$ and every quotient $\mathfrak{g}/\mathfrak{a}$ of $\mathfrak{g}$ is semisimple. Then $\mathfrak{g}$ is semisimple. Conversely, let $\mathfrak{g}$ be a semisimple Lie algebra over characteristic 0. Then all ideals and all quotients of $\mathfrak{g}$ are semisimple.*

Proof We prove only the converse direction. Let $\mathfrak{g}$ be semisimple, so that β is nondegenerate. Let $\alpha^\perp$ be the orthogonal subspace to $\mathfrak{a}$ with respet to β. Then $\mathfrak{a}^\perp = \ker\{x \mapsto \beta(-, x) : \mathfrak{g} \to \mathrm{Hom}(\mathfrak{a}, \mathfrak{g})\}$, so $\mathfrak{a}^\perp$ is an ideal. Then $\mathfrak{a} \cap \mathfrak{a}^\perp = \ker \beta|_{\mathfrak{a}} \leq \mathrm{rad}\, \mathfrak{a}$, and hence it's solvable and hence is 0. So $\mathfrak{a}$ is semisimple, and also $\mathfrak{a}^\perp$ is. In particular, the projection $\mathfrak{a}^\perp \overset{\sim}{\to} \mathfrak{g}/\mathfrak{a}$ is an isomorphism of Lie algebras, so $\mathfrak{g}/\mathfrak{a}$ is semisimple. □

4.3.0.4 Corollary *Every finite-dimensional semisimple Lie algebra $\mathfrak{g}$ over characteristic 0 is a direct product $\mathfrak{g} = \mathfrak{g}_1 \times \cdots \times \mathfrak{g}_m$ of simple nonabelian algebras.*

Proof Let $\mathfrak{a}$ be a minimal and hence simple idea. Then $[\mathfrak{a}, \mathfrak{a}^\perp] \subseteq \mathfrak{a} \cap \mathfrak{a}^\perp = 0$. Rinse and repeat. □

4.3.0.5 Corollary $\mathfrak{sl}(2)$ *is simple.* □

4.4 Some homological algebra

We will not need too much homological algebra; any standard textbook on the subject, e.g. [CE99, GM03, Wei94], will contain fancier versions of many of these constructions.

4.4.1 The Casimir

The following piece of linear algebra is a trivial exercise in definition-chasing:

4.4.1.1 Proposition *Let $\langle,\rangle$ be a nondegenerate not-necessarily-symmetric bilinear form on finite-dimensional V. Let (x_i) and (y_i) be dual bases, so $\langle x_i, y_j\rangle = \delta_{ij}$. Then $\theta = \sum x_i \otimes y_i \in V \otimes V$ depends only on the form $\langle,\rangle$. If $z \in \mathfrak{gl}(V)$ leaves $\langle,\rangle$ invariant, then θ is also invariant.* □

4.4.1.2 Corollary *Let β be a nondegenerate invariant (symmetric) form on a finite-dimensional Lie algebra $\mathfrak{g}$, and define $c_\beta = \sum x_i y_i$ to be the image of θ in Proposition 4.4.1.1 under the multiplication map $\mathfrak{g} \otimes \mathfrak{g} \to \mathcal{U}\mathfrak{g}$. Then c_β is a central element of $\mathcal{U}\mathfrak{g}$.* □

4.4.1.3 Lemma / Definition *Let $\mathfrak{g}$ be a finite-dimensional Lie algebra and V a $\mathfrak{g}$-module so that the trace form β_V is nondegenerate. Define the* Casimir operator *$c_V = c_{\beta_V}$ as in Corollary 4.4.1.2. Then c_V has the following properties:*

1. *c_V only depends on β_V.*
2. *$c_V \in Z(\mathcal{U}(\mathfrak{g}))$*
3. *$c_V \in \mathcal{U}(\mathfrak{g})\mathfrak{g}$, i.e. it acts as 0 on $\mathbb{K}$.*
4. *$\mathrm{tr}_V(c_V) = \sum \mathrm{tr}_V(x_i y_i) = \dim \mathfrak{g}$.*

In particular, c_V distinguishes V from the trivial representation. □

4.4.2 Review of Ext

4.4.2.1 Definition *Let $\mathcal{C}$ be an abelian category. A* complex *(with* homological indexing*) in $\mathcal{C}$ is a sequence $A_\bullet = \dots A_k \overset{d_k}{\to} A_{k-1} \to \dots$ of maps in $\mathcal{C}$ such that $d_k \circ d_{k+1} = 0$ for every k. The* homology *of $A_\bullet$ are the objects $H_k(A_\bullet) \overset{\text{def}}{=} \ker d_k / \operatorname{Im} d_{k+1}$. For each k, $\ker d_k$ is the object of* k-cycles*, and $\operatorname{Im} d_{k+1}$ is the object of* k-boundaries*.*

We can write the same complex with cohomological indexing *by writing $A^k \overset{\text{def}}{=} A_{-k}$, whence the arrows go $\dots \to A^{k-1} \overset{\delta^k}{\to} A^k \to \dots$. The* cohomology *of a complex is $H^k(A^\bullet) \overset{\text{def}}{=} H_{-k}(A_\bullet) = \ker \delta^{k+1} / \operatorname{Im} \delta^k$. The* k-cocycles *are $\ker \delta^{k+1}$ and the* k-coboundaries *are $\operatorname{Im} \delta^k$.*

A complex is exact *at k if $H_k = 0$. A* long exact sequence *is a complex, usually infinite, that is exact everywhere. A* short exact sequence *is a three-term exact complex of the form $0 \to A \to B \to C \to 0$. In particular, $A = \ker(B \to C)$ and $C = A/B$.*

4.4.2.2 Definition *Let U be an associative algebra and U-*MOD *its category of left modules. A* free module *is a module $U \curvearrowright F$ that is isomorphic to a possibly-infinite direct sum of copies of $U \curvearrowright U$. Let M be a U-module. A* free resolution *of M is a complex $F_\bullet = \dots \to F_1 \to F_0 \to 0$ that is exact everywhere except at $k = 0$, where $H_k(F_\bullet) = M$. Equivalently, the augmented complex $F_\bullet \to M \to 0$ is exact.*

4.4.2.3 Lemma *Given any module M, a free resolution $F_\bullet$ of M exists.*

Proof Let $F_{-1} \overset{\text{def}}{=} M_0 \overset{\text{def}}{=} M$ and $M_{k+1} \overset{\text{def}}{=} \ker(F_k \to F_{k-1})$. Define F_k to be the free module on a generating set of M_k. □

4.4.2.4 Lemma / Definition *Let U be an associative algebra and M, N two left U modules. Let $F_\bullet$ be a free resolution of M, and construct the complex*

$$\mathrm{Hom}_U(F_\bullet, N) = \mathrm{Hom}_U(F_0, N) \overset{\delta^1}{\to} \mathrm{Hom}_U(F_1, N) \overset{\delta^2}{\to} \dots$$

by applying the contravariant functor $\mathrm{Hom}_U(-, N)$ to the complex $F_\bullet$. Define $\mathrm{Ext}^i_U(M, N) \overset{\text{def}}{=} H^i(\mathrm{Hom}_U(F_\bullet, N))$. Then $\mathrm{Ext}^0_U(M, N) = \mathrm{Hom}(M, N)$. Moreover, $\mathrm{Ext}^i_U(M, N)$ does not depend on the choice of free resolution $F_\bullet$, and is functorial in M and N.

Proof It's clear that for each choice of a free-resolution of M, we get a functor $\mathrm{Ext}^\bullet(M, -)$.

Let $M \to M'$ be a U-morphism, and $F'_\bullet$ a free resolution of M'. By freeness we can extend the morphism $M \to M'$ to a chain morphism, unique up to chain homotopy:

$$\begin{tikzcd}
\cdots \arrow[r] & F_1 \arrow[r] \arrow[d, dotted] & F_0 \arrow[r] \arrow[d, dotted] & M \arrow[d] \\
\cdots \arrow[r] & F'_1 \arrow[r] & F'_0 \arrow[r] & M'
\end{tikzcd}$$

Chain homotopies induce isomorphisms on Hom, so $\mathrm{Ext}^\bullet(M, N)$ is functorial in M; in particular, letting $M' = M$ with a different free resolution shows that $\mathrm{Ext}^\bullet(M, N)$ is well-defined. $\square$

4.4.2.5 Lemma / Definition *The functor* $\mathrm{Hom}(-, N)$ *is* left-exact *but not* right-exact*, i.e. if $0 \to A \to B \to C \to 0$ is a short exact sequence then* $\mathrm{Hom}(A, N) \leftarrow \mathrm{Hom}(B, N) \leftarrow \mathrm{Hom}(C, N) \leftarrow 0$ *is exact, but* $0 \leftarrow \mathrm{Hom}(A, N) \leftarrow \mathrm{Hom}(B, N)$ *is not necessarily exact. Rather, we get a* long exact sequence in Ext*:*

$$\begin{tikzcd}
 & \mathrm{Ext}^0(A,N) \arrow[drr, bend right, start anchor=west, end anchor=east] & \mathrm{Ext}^0(B,N) \arrow[l] & \mathrm{Ext}^0(C,N) \arrow[l] & 0 \arrow[l] \\
 & \mathrm{Ext}^1(A,N) \arrow[drr, bend right, start anchor=west, end anchor=east] & \mathrm{Ext}^1(B,N) \arrow[l] & \mathrm{Ext}^1(C,N) \arrow[l] & \\
\cdots & \mathrm{Ext}^2(A,N) \arrow[l] & \mathrm{Ext}^2(B,N) \arrow[l] & \mathrm{Ext}^2(C,N) \arrow[l] &
\end{tikzcd}$$

When $N = A$, the image of $1_A \in \mathrm{Hom}(A, A) = \mathrm{Ext}^0(A, A)$ in $\mathrm{Ext}^1(C, A)$ is the characteristic class *of the extension $0 \to A \to B \to C \to 0$. The characteristic class determines B up to equivalence; in particular, when $1_A \mapsto 0$, then $B \cong A \oplus C$.* $\square$

4.4.3 Complete reducibility

4.4.3.1 Lemma *Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{K}$, $\mathcal{U}\mathfrak{g}$ its universal enveloping algebra, N a $\mathfrak{g}$-module, and F a free $\mathfrak{g}$-module. Then $F \otimes_{\mathbb{K}} N$ is free.*

Proof Let $F = \bigoplus \mathcal{U}\mathfrak{g}$; then $F \otimes N = \left(\bigoplus \mathcal{U}\mathfrak{g}\right) \otimes_{\mathbb{K}} N = \bigoplus (\mathcal{U}\mathfrak{g} \otimes_{\mathbb{K}} N)$, so it suffices to show that $G \overset{\text{def}}{=} \mathcal{U}\mathfrak{g} \otimes_{\mathbb{K}} N$ is free.

We understand the $\mathcal{U}\mathfrak{g}$-action on G: let $x \in \mathfrak{g}$ and $u \otimes n \in G$, then $x \cdot (u \otimes n) = (xu) \otimes n + u \otimes (x \cdot n)$ as $\Delta x = x \otimes 1 + 1 \otimes x$. Here xu is the product in $\mathcal{U}\mathfrak{g}$ and $x \cdot n$ is the action $\mathfrak{g} \curvearrowright N$.

We can put a filtration on G by $G_{\leq n} = \mathcal{U}\mathfrak{g}_{\leq n} \otimes_{\mathbb{K}} N$. This makes G into a *filtered module*:

$$\mathcal{U}(\mathfrak{g})_{\leq k} G_{\leq l} \subseteq G_{\leq k+l}$$

Thus $\operatorname{gr} G$ is a $\operatorname{gr} \mathcal{U}\mathfrak{g}$-module, but $\operatorname{gr} \mathcal{U}\mathfrak{g} = \mathcal{S}\mathfrak{g}$, and $\mathcal{S}\mathfrak{g}$ acts through the first term, so $\mathcal{S}\mathfrak{g} \otimes N$ is a free $\mathcal{S}\mathfrak{g}$-module, by picking any basis of N.

let $\{n_\beta\}$ be a basis of N and $\{x_\alpha\}$ a basis of $\mathfrak{g}$. Then $\{x_{\vec{\alpha}} n_\beta\}$ is a basis of $\operatorname{gr} G = \mathcal{S}\mathfrak{g} \otimes N$, hence also a basis of $\mathcal{U}(\mathfrak{g}) \otimes N$. Thus $\mathcal{U}(\mathfrak{g}) \otimes N$ is free. We have used Theorem 3.2.2.1 implicitly multiple times. ☐

4.4.3.2 Corollary *If M and N are finite-dimensional $\mathfrak{g}$-modules, then:*

$$\operatorname{Ext}^i(M, N) \cong \operatorname{Ext}^i(\operatorname{Hom}(N, M), \mathbb{K}) \cong \operatorname{Ext}^i(\mathbb{K}, \operatorname{Hom}(M, N))$$

Proof Let $F_\bullet \to \mathbb{K}$ be a free resolution of $\mathfrak{g}$-modules. By Lemma 4.4.3.1, $F_\bullet \otimes M \to \mathbb{K} \otimes M \cong M$ and $F_\bullet \otimes M \otimes N^* \to M \otimes N^* = \operatorname{Hom}(M, N)$ are free resolutions. A $\mathcal{U}(\mathfrak{g})$-module homomorphism is exactly a $\mathfrak{g}$-invariant linear map:

$$\begin{aligned}
\operatorname{Hom}_{\mathcal{U}(\mathfrak{g})}(F_\bullet \otimes M, N) &= \operatorname{Hom}_{\mathbb{K}}(F_\bullet \otimes M, N)^{\mathfrak{g}} \\
&= \operatorname{Hom}_{\mathbb{K}}(F_\bullet \otimes M \otimes N^*, \mathbb{K})^{\mathfrak{g}} = \operatorname{Ext}^\bullet(M \otimes N^*, \mathbb{K}) \\
&= \operatorname{Hom}_{\mathbb{K}}(F_\bullet, M^* \otimes N)^{\mathfrak{g}} = \operatorname{Ext}^\bullet(\mathbb{K}, \operatorname{Hom}(M, N))
\end{aligned}$$ ☐

4.4.3.3 Lemma *If M, N are finite-dimensional $\mathfrak{g}$-modules and $c \in Z(\mathcal{U}\mathfrak{g})$ such that the characteristic polynomials f and g of c on M and N are relatively prime, then $\operatorname{Ext}^i(M, N) = 0$ for all i.*

Proof By functoriality, c acts on $\operatorname{Ext}^i(M, N)$. By centrality, the action of c on $\operatorname{Ext}^i(M, N)$ must satisfy both the characteristic polynomials: $f(c), g(c)$ annihilate $\operatorname{Ext}^i(M, N)$. If f and g are relatively prime, then $1 = af + bg$ for some polynomials a, b; thus 1 annihilates $\operatorname{Ext}^i(M, N)$, which must therefore be 0. ☐

4.4.3.4 Theorem (Schur's Lemma)
Let U be an algebra and N a simple non-zero U-module, and let $\alpha : N \to N$ a U-homomorphism; then $\alpha = 0$ or α is an isomorphism.

Proof The image of α is a submodule of N, hence either 0 or N. If $\operatorname{Im} \alpha = 0$, then we're done. If $\operatorname{Im} \alpha = N$, then $\ker \alpha \neq 0$, so $\ker \alpha = N$ by simplicity, and α is an isom. ☐

4.4.3.5 Corollary *Let M, N be finite-dimensional simple U-modules such that $c \in Z(U)$ annihilates M but not N; then $\operatorname{Ext}^i(M, N) = 0$ for every i.*

Proof By Theorem 4.4.3.4, c acts invertibly on N, so all its eigenvalues (over the algebraic closure) are non-zero. But the eigenvalues of c on M are all 0, so the characteristic polynomials are relatively prime. ☐

4.4.3.6 Theorem (Ext^1 vanishes over a semisimple Lie algebra)
Let $\mathfrak{g}$ be a semisimple Lie algebra over a field $\mathbb{K}$ of characteristic 0, and let M and N be finite-dimensional $\mathfrak{g}$-modules. Then $\mathrm{Ext}^1(M,N) = 0$.

Proof Using Corollary 4.4.3.2 we may assume that $M = \mathbb{K}$. Assume that N is not a trivial module. Then $\mathfrak{g} = \mathfrak{g}_1 \times \cdots \times \mathfrak{g}_k$ by Corollary 4.3.0.4 for $\mathfrak{g}_i$ simple, and some $\mathfrak{g}_i$ acts non-trivially on N. Then β_N does not vanish on $\mathfrak{g}_i$ by Theorem 4.2.6.4, and so $\ker_{\mathfrak{g}_i} \beta_N = 0$ by simplicity. Thus we can find a Casimir $c \in Z(\mathcal{U}\mathfrak{g}_i) \subseteq Z(\mathcal{U}\mathfrak{g})$. In particular, $\mathrm{tr}_N(c) = \dim \mathfrak{g}_i \neq 0$, but c annihilates $\mathbb{K}$, and so by Corollary 4.4.3.5 $\mathrm{Ext}^1(\mathbb{K},N) = 0$.

If N is a trivial module, then we use the fact that $\mathrm{Ext}^1(\mathbb{K},N)$ classifies extensions $0 \to N \to L \to \mathbb{K} \to 0$, which we will classify directly. (See Example 4.4.4.6 for a direct verification that Ext^1 classifies extensions in the case of $\mathfrak{g}$-modules.) Writing L in block form (as a vector space, $L = N \oplus \mathbb{K}$), we see that $\mathfrak{g}$ acts on L like $\begin{bmatrix} 0 & * \\ 0 & 0 \end{bmatrix}$. Then $\mathfrak{g}$ acts by nilpotent matrices, but $\mathfrak{g}$ is semisimple, so $\mathfrak{g}$ annihilates L. Thus the only extension is the trivial one, and $\mathrm{Ext}^1(\mathbb{K},N) = 0$. □

We list two corollaries, which are important enough to call theorems. We recall the following definition:

4.4.3.7 Definition *An object in an abelian category is* simple *if it has no non-zero proper subobjects. An object is* completely reducible *if it is a direct sum of simple objects.*

4.4.3.8 Theorem (Weyl's Complete Reducibility Theorem)
Every finite-dimensional representation of a semisimple Lie algebra over characteristic zero is completely reducible. □

4.4.3.9 Theorem (Whitehead's Theorem)
If $\mathfrak{g}$ is a semisimple Lie algebra over characteristic zero, and M and N are finite-dimensional non-isomorphic simple $\mathfrak{g}$-modules, then $\mathrm{Ext}^i(M,N)$ vanishes for all i. □

4.4.4 Computing $\mathrm{Ext}^i(\mathbb{K}, M)$

4.4.4.1 Proposition *Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{K}$, and $\mathbb{K}$ the trivial representation. Then $\mathbb{K}$ has a free $\mathcal{U}\mathfrak{g}$ resolution given by:*

$$\cdots \to \mathcal{U}(\mathfrak{g}) \otimes_{\mathbb{K}} \textstyle\bigwedge^2 \mathfrak{g} \overset{d_2}{\to} \mathcal{U}(\mathfrak{g}) \otimes_{\mathbb{K}} \mathfrak{g} \overset{d_1}{\to} \mathcal{U}(\mathfrak{g}) \overset{\epsilon}{\to} \mathbb{K} \to 0 \tag{4.4.4.2}$$

The maps $d_k : \mathcal{U}(\mathfrak{g}) \otimes \bigwedge^k \mathfrak{g} \to \mathcal{U}(\mathfrak{g}) \otimes \bigwedge^{k-1} \mathfrak{g}$ for $k \leq 1$ are given by:

$$\begin{aligned} d_k(x_1 \wedge \cdots \wedge x_k) = & \sum_i (-1)^{i-1} x_i \otimes (x_1 \wedge \cdots \hat{x}_i \cdots \wedge x_k) \\ & - \sum_{i<j} (-1)^{i-j+1} 1 \otimes ([x_i,x_j] \wedge x_1 \wedge \cdots \hat{x}_i \cdots \hat{x}_j \cdots \wedge x_k) \end{aligned} \tag{4.4.4.3}$$

Proof That d_k is well-defined requires only checking that it is antisymmetric. That $d_{k-1} \circ d_k = 0$ is more or less obvious: terms cancel either by being sufficiently separated to appear twice with opposite signs (like in the free resolution of the symmetric polynomial ring), or by syzygy, or by Jacobi.

For exactness, we quote a general principle: Let $F_\bullet(t)$ be a t-varying complex of vector spaces, and choose a basis for each one. Assume that the vector spaces do not change with t, but that the matrix coefficients of the differentials d_k depend algebraically on t. Then the dimension of H^i can jump for special values of t, but does not fall at special values of t. In particular, exactness is a Zariski open condition.

Thus consider the complex with the vector spaces given by equation (4.4.4.2), but with the differential given by

$$\begin{aligned} d_k(x_1 \wedge \cdots \wedge x_k) = & \sum_i (-1)^{i-1} x_i \otimes (x_1 \wedge \cdots \hat{x_i} \cdots \wedge x_k) \\ & - t \sum_{i<j} (-1)^{i-j+1} 1 \otimes ([x_i, x_j] \wedge x_1 \wedge \cdots \hat{x_i} \cdots \hat{x_j} \cdots \wedge x_k) \end{aligned} \tag{4.4.4.4}$$

This corresponds to the Lie algebra $\mathfrak{g}_t = (\mathfrak{g}, [x,y]_t \stackrel{\text{def}}{=} t[x,y])$. When $t \neq 0$, $\mathfrak{g}_t \cong \mathfrak{g}$, by $x \mapsto tx$, but when $t = 0$, $\mathfrak{g}_0$ is abelian, and the complex consists of polynomial rings and is obviously exact.

Thus the t-varying complex is exact at $t = 0$ and hence in an open neighborhood of 0. If $\mathbb{K}$ is not finite, then an open neighborhood of 0 contains non-zero terms, and so the complex is exact for some $t \neq 0$ and hence for all t. If $\mathbb{K}$ is finite, we replace it by its algebraic closure. $\square$

4.4.4.5 Corollary $\mathrm{Ext}^\bullet(\mathbb{K}, M)$ *is the cohomology of the* Chevalley complex *with coefficients in M:*

$$0 \to M \xrightarrow{\delta^1} \mathrm{Hom}_{\mathbb{K}}(\mathfrak{g}, M) \xrightarrow{\delta^2} \mathrm{Hom}(\textstyle\bigwedge^2 \mathfrak{g}, M) \to \dots$$

If $g \in \mathrm{Hom}_{\mathbb{K}}(\bigwedge^{k-1}\mathfrak{g}, M)$, then the differential $\delta^k g$ is given by:

$$\begin{aligned} \delta^k g(x_1 \wedge \cdots \wedge x_k) = & \sum_i (-1)^{i-1} x_i g(x_1 \wedge \cdots \hat{x_i} \cdots \wedge x_k) \\ & - \sum_{i<j} (-1)^{i-j+1} g([x_i, x_j] \wedge x_1 \wedge \cdots \hat{x_i} \cdots \hat{x_j} \cdots \wedge x_k) \end{aligned}$$ $\square$

4.4.4.6 Example Let M and N be finite-dimensional $\mathfrak{g}$-modules. Then $\mathrm{Ext}^i(M,N)$ is the cohomology of $\cdots \xrightarrow{\delta^k} \mathrm{Hom}_{\mathbb{K}}(\bigwedge^l \mathfrak{g}, M^* \otimes N) \xrightarrow{\delta^{k+1}} \cdots$. We compute $\mathrm{Ext}^1(M,N)$.

If $\phi \in M^* \otimes N$ and $x \in \mathfrak{g}$, then the action of x on ϕ is given by $x \cdot \phi = x_N \circ \phi - \phi \circ x_M =$ "$[x, \phi]$". A 1-cocycle is a map $f : \mathfrak{g} \to M^* \otimes N$ such that $0 = \delta^1 f(x \wedge y) = f([x,y]) - \big((x \cdot f)(y) - (y \cdot f)(x)\big) =$ "$[x, f(y)] - [y, f(x)]$".

Let $0 \to N \to V \to M \to 0$ be a $\mathbb{K}$-vector space, and choose a splitting $\sigma : M \to V$ as vector spaces. Then $\mathfrak{g}$ acts on $M \oplus N$ by $x \mapsto \begin{bmatrix} x_N & f(x) \\ 0 & x_M \end{bmatrix}$, and the cocycles f exactly classify the possible ways to put something in the upper right corner.

The ways to change the splitting $\sigma \mapsto \sigma' = \sigma + h$ correspond to $\mathbb{K}$-linear maps $h : M \to N$. This changes $f(x)$ by $x_N \circ h - h \circ x_M = \delta^1(h)$.

We have seen that the 1-cocycles classify the splitting, and changing the 1-cocycle by a 1-coboundary changes the splitting but not the extension. So $\mathrm{Ext}^1(M,N)$ classifies extensions up to isomorphism. $\lozenge$

4.4.4.7 Example Consider abelian extensions of Lie algebras $0 \to \mathfrak{m} \to \hat{\mathfrak{g}} \to \mathfrak{g} \to 0$ where $\mathfrak{m}$ is an abelian ideal of $\hat{\mathfrak{g}}$. Since $\mathfrak{m}$ is abelian, the action $\hat{\mathfrak{g}} \curvearrowright \mathfrak{m}$ factors through $\mathfrak{g} = \hat{\mathfrak{g}}/\mathfrak{m}$. Conversely, we can classify abelian extensions $0 \to \mathfrak{m} \to \hat{\mathfrak{g}} \to \mathfrak{g} \to 0$ given $\mathfrak{g}$ and a $\mathfrak{g}$-module $\mathfrak{m}$.

We pick a $\mathbb{K}$-linear splitting $\sigma : \mathfrak{g} \to \hat{\mathfrak{g}}$; then $\hat{\mathfrak{g}} = \{\sigma(x) + m\}$ as x ranges over $\mathfrak{g}$ and m over $\mathfrak{m}$, and the bracket is

$$[\sigma(x) + m, \sigma(y) + n] = \sigma([x, y]) + [\sigma(x), n] - [\sigma(y), m] + g(x, y)$$

where g is the error term measuring how far off σ is from being a splitting of $\mathfrak{g}$-modules. There is no $[m, n]$ term, because $\mathfrak{m}$ is assumed to be an abelian ideal of $\hat{\mathfrak{g}}$.

Then g is antisymmetric. The Jacobi identity on $\hat{\mathfrak{g}}$ is equivalent to g satisfying:

$$\begin{aligned} 0 &= x\, g(y \wedge z) - y\, g(x \wedge z) + z\, g(x \wedge y) - g([x, y] \wedge z) + g([x, z] \wedge y) - g([y, z] \wedge x) && (4.4.4.8) \\ &= x\, g(y \wedge z) - g([x, y] \wedge z) + \text{cycle permutations} && (4.4.4.9) \end{aligned}$$

I.e. g is a 2-cocycle in $\mathrm{Hom}_{\mathbb{K}}(\mathfrak{g}, \mathfrak{m})$. In particular, the 2-cocycles classify extensions of $\mathfrak{g}$ by $\mathfrak{m}$ along with a splitting. If we change the splitting by $f : \mathfrak{g} \to \mathfrak{m}$, then $\mathfrak{g}$ changes by $(x \cdot f)(y) - (y \cdot f)(x) - f([x, y]) = \delta^2(f)$. We have proved: $\Diamond$

4.4.4.10 Proposition $\mathrm{Ext}^2_{\mathcal{U}\mathfrak{g}}(\mathbb{K}, \mathfrak{m})$ *classifies abelian extensions* $0 \to \mathfrak{m} \to \hat{\mathfrak{g}} \to \mathfrak{g} \to 0$ *up to isomorphism. The element* $0 \in \mathrm{Ext}^2$ *corresponds to the semidirect product* $\hat{\mathfrak{g}} = \mathfrak{g} \ltimes \mathfrak{m}$. $\square$

4.4.4.11 Corollary *Abelian extensions of semisimple Lie groups are semidirect products.* $\square$

4.4.4.12 Theorem (Levi's Theorem)
Let $\mathfrak{g}$ *be a finite-dimensional Lie algebra over characteristic* 0*, and let* $\mathfrak{r} = \mathrm{rad}(\mathfrak{g})$*. Then* $\mathfrak{g}$ *has a* Levi decomposition*: semisimple* Levi subalgebra $\mathfrak{s} \subseteq \mathfrak{g}$ *such that* $\mathfrak{g} = \mathfrak{s} \ltimes \mathfrak{r}$*.*

Proof Without loss of generality, $\mathfrak{r} \neq 0$, as otherwise $\mathfrak{g}$ is already semisimple.

Assume first that $\mathfrak{r}$ is not a minimal non-zero ideal. In particular, let $\mathfrak{m} \neq 0$ be an ideal of $\mathfrak{g}$ with $\mathfrak{m} \subsetneq \mathfrak{r}$. Then $\mathfrak{r}/\mathfrak{m} = \mathrm{rad}(\mathfrak{g}/\mathfrak{m}) \neq 0$, and by induction on dimension $\mathfrak{g}/\mathfrak{m}$ has a Levi subalgebra. Let $\tilde{\mathfrak{s}}$ be the preimage of this subalgebra in $\mathfrak{g}/\mathfrak{m}$. Then $\tilde{\mathfrak{s}} \cap \mathfrak{r} = \mathfrak{m}$ and $\tilde{\mathfrak{s}}/\mathfrak{m} \xrightarrow{\sim} (\mathfrak{g}/\mathfrak{m})/(\mathfrak{r}/\mathfrak{m}) = \mathfrak{g}/\mathfrak{r}$. Hence $\mathfrak{m} = \mathrm{rad}(\tilde{\mathfrak{s}})$. Again by induction on dimension, $\tilde{\mathfrak{s}}$ has a Levi subalgebra $\mathfrak{s}$; then $\tilde{\mathfrak{s}} = \mathfrak{s} \oplus \mathfrak{m}$ and $\mathfrak{s} \cap \mathfrak{r} = 0$, so $\mathfrak{s} \xrightarrow{\sim} \mathfrak{g}/\mathfrak{r}$. Thus $\mathfrak{s}$ is a Levi subalgebra of $\mathfrak{g}$.

We turn now to the case when $\mathfrak{r}$ is minimal. Being a radical, $\mathfrak{r}$ is solvable, so $\mathfrak{r}' \neq \mathfrak{r}$, and by minimality $\mathfrak{r}' = 0$. So $\mathfrak{r}$ is abelian. In particular, the action $\mathfrak{g} \curvearrowright \mathfrak{r}$ factors through $\mathfrak{g}/\mathfrak{r}$, and so $0 \to \mathfrak{r} \to \mathfrak{g} \to \mathfrak{g}/\mathfrak{r} \to 0$ is an abelian extension of $\mathfrak{g}/\mathfrak{r}$, and thus must be semidirect by Corollary 4.4.4.11. $\square$

4.4.4.13 Remark We always have $Z(\mathfrak{g}) \leq \mathfrak{r}$, and when $\mathfrak{r}$ is minimal, $Z(\mathfrak{g})$ is either 0 or $\mathfrak{r}$. When $Z(\mathfrak{g}) = \mathfrak{r}$, then $0 \to \mathfrak{r} \to \mathfrak{g} \to \mathfrak{g}/\mathfrak{r} \to 0$ is in fact an extension of $\mathfrak{g}/\mathfrak{r}$-modules, and so is a direct product by Example 4.4.4.6. $\Diamond$

We will not prove the following, but instead suggest it as Exercise 11.

4.4.4.14 Theorem (Malcev–Harish-Chandra Theorem)
All Levi subalgebras of a given Lie algebra are conjugate under the action of the group $\exp \mathrm{ad}\, \mathfrak{n} \subseteq \mathrm{GL}(V)$*, where* $\mathfrak{n}$ *is the largest nilpotent ideal of* $\mathfrak{g}$*. (In particular,* $\mathrm{ad} : \mathfrak{n} \curvearrowright \mathfrak{g}$ *is nilpotent, so the power series for* $\exp$ *terminates.)* ☐

We are now ready to complete the proof of Theorem 3.1.2.1, with a theorem of Cartan:

4.4.4.15 Theorem (Lie's Third Theorem)
Let $\mathfrak{g}$ *be a Lie algebra over* $\mathbb{R}$*. Then* $\mathfrak{g} = \mathrm{Lie}(G)$ *for some analytic Lie group* G*.*

Proof Find a Levi decomposition $\mathfrak{g} = \mathfrak{s} \ltimes \mathfrak{r}$. If $\mathfrak{s} = \mathrm{Lie}(S)$ and $\mathfrak{r} = \mathrm{Lie}(R)$ where S and R are connected and simply connected, then the action $\mathfrak{s} \curvearrowright \mathfrak{r}$ lifts to an action $S \curvearrowright R$. Thus we can construct $G = S \ltimes R$, and it is a direct computation that $\mathfrak{g} = \mathrm{Lie}(G)$ in this case.

So it suffices to find groups S and R with the desired Lie algebras. We need not even assure that the groups we find are simply-connected; we can always take universal covers. In any case, $\mathfrak{s}$ is semisimple, so the action $\mathfrak{s} \to \mathfrak{gl}(\mathfrak{s})$ is faithful, and thus we can find $S \subseteq \mathrm{GL}(\mathfrak{s})$ with $\mathrm{Lie}(S) = \mathfrak{s}$.

On the other hand, $\mathfrak{r}$ is solvable: the chain $\mathfrak{r} \geq \mathfrak{r}' \geq \mathfrak{r}'' \geq \dots$ eventually gets to 0. We can interpolate between $\mathfrak{r}$ and $\mathfrak{r}'$ by one-co-dimensional vector spaces, which are all necessarily ideals of some $\mathfrak{r}^{(k)}$, and the quotients are one-dimensional and hence abelian. Thus any solvable Lie algebra is an extension by one-dimensional algebras, and this extension also lifts to the level of groups. So $\mathfrak{r} = \mathrm{Lie}(R)$ for some Lie group R. ☐

4.5 From Zassenhaus to Ado

Ado's Theorem (Theorem 4.5.0.10) normally is not proven in a course in Lie Theory. For example, [Kna02] relegates Ado's Theorem to an appendix (B.3). In fact, we will see that Ado's Theorem is a direct consequence of Theorem 4.4.4.12, although we will need to develop some preliminary facts.

4.5.0.1 Lemma / Definition *A* Lie derivation *of a Lie algebra* $\mathfrak{a}$ *is a linear map* $f : \mathfrak{a} \to \mathfrak{a}$ *such that* $f([x, y]) = [f(x), y] + [x, f(y)]$*. Equivalently, a derivation is a one-cocycle in the Chevalley complex with coefficients in* $\mathfrak{a}$*.*

A derivation *of an associative algebra* A *is a linear map* $f : A \to A$ *so that* $f(xy) = f(x)\, y + x\, f(x)$*.*

The product (composition) of (Lie) derivations is not necessarily a (Lie) derivation, but the commutator of derivations is a derivation. We write $\mathrm{Der}\, \mathfrak{a}$ *for the Lie algebra of Lie derivations of* $\mathfrak{a}$*, and* $\mathrm{Der}\, A$ *for the algebra of associative derivations of* A*. Henceforth, we drop the adjective "Lie", talking about simply derivations of a Lie algebra.*

We say that $\mathfrak{h} \curvearrowright \mathfrak{a}$ by derivations *if the map* $\mathfrak{h} \to \mathfrak{gl}(\mathfrak{a})$ *in fact lands in* $\mathrm{Der}\, \mathfrak{a}$*.* ☐

In very general language, let A and B be vector spaces, and $a : A^{\otimes n} \to A$ and $b : B^{\otimes n} \to B$ n-linear maps. Then a *homomorphism* from (A, a) to (B, b) is a linear map $\phi : A \to B$ so that $\phi \circ a = b \circ \phi^{\otimes n}$, and a *derivation* from (A, a) to (B, b) is a linear map $\phi : A \to B$ such that $\phi \circ a = b \circ (\sum_{i=1}^{n} \phi_i)$, where $\phi_i \stackrel{\text{def}}{=} 1 \otimes \cdots \otimes \phi \otimes \cdots \otimes 1$, with the ϕ in the ith spot. The space $\mathrm{Hom}(A, B)$ of homomorphisms is not generally a vector space, but the space $\mathrm{Der}(A, B)$ of derivations is. If $(A, a) = (B, b)$, then $\mathrm{Hom}(A, A)$ is closed under composition and hence a monoid, whereas

$\mathrm{Der}(A, A)$ is closed under the commutator and hence a Lie algebra. The notions of "derivation" and "homomorphism" agree for $n = 1$, whence the map ϕ must intertwine a with b. The difference between derivations and homomorphisms is the difference between grouplike and primitive elements of a bialgebra.

4.5.0.2 Proposition *Let $\mathfrak{a}$ be a Lie algebra.*

1. *Every derivation of $\mathfrak{a}$ extends uniquely to a derivation of $\mathcal{U}(\mathfrak{a})$.*
2. $\mathrm{Der}\,\mathfrak{a} \to \mathrm{Der}\,\mathcal{U}(\mathfrak{a})$ *is a Lie algebra homomorphism.*
3. *If $\mathfrak{h} \curvearrowright \mathfrak{a}$ by derivations, then $\mathfrak{h}(\mathcal{U}(\mathfrak{a})) \subseteq \mathcal{U}(\mathfrak{a}) \cdot \mathfrak{h}(\mathfrak{a}) \cdot \mathcal{U}(\mathfrak{a})$ the two-sided ideal of $\mathcal{U}(\mathfrak{a})$ generated by the image of the $\mathfrak{h}$ action in $\mathfrak{a}$.*
4. *If $N \leq \mathcal{U}(\mathfrak{a})$ is an $\mathfrak{h}$-stable two-sided ideal, so is N^n.*

Proof 1. Let $d \in \mathrm{Der}\,\mathfrak{a}$, and define $\hat{\mathfrak{a}} \stackrel{\text{def}}{=} \mathbb{K}d \oplus \mathfrak{a}$; then $\mathcal{U}(\mathfrak{a}) \subseteq \mathcal{U}(\hat{\mathfrak{a}})$. The commutative $[d, -]$ preserves $\mathcal{U}(\mathfrak{a})$ and is the required derivation. Uniqueness is immediate: once you've said how something acts on the generators, you've defined it on the whole algebra.

2. This is an automatic consequence of the uniqueness: the commutator of two derivations is a derivation, so if it's unique, it must be the correct derivation.

3. Let $a_1, \dots, a_k \in \mathfrak{a}$ and $h \in \mathfrak{h}$. Then $h(a_1 \cdots a_k) = \sum_{i=1}^k a_1 \cdots h(a_i) \cdots a_k \in \mathcal{U}(\mathfrak{a})\, \mathfrak{h}(\mathfrak{a})\, \mathcal{U}(\mathfrak{a})$.

4. N^n is spanned by monomials $a_1 \cdots a_n$ where all $a_i \in N$. Assuming that $h(a_i) \in N$ for each $h \in \mathfrak{h}$, we see that $h(a_1 \cdots a_k) = \sum_{i=1}^k a_1 \cdots h(a_i) \cdots a_k \in N^n$. □

4.5.0.3 Lemma / Definition *Let $\mathfrak{h}$ and $\mathfrak{a}$ be Lie algebras and let $\mathfrak{h} \curvearrowright \mathfrak{a}$ by derivations. The* semidirect product $\mathfrak{h} \ltimes \mathfrak{a}$ *is the vector space $\mathfrak{h} \oplus \mathfrak{a}$ with the bracket given by $[h_1 + a_1, h_2 + a_2] = [h_1, h_2]_{\mathfrak{h}} + [a_1, a_2]_{\mathfrak{a}} + h_1 \cdot a_2 - h_2 \cdot a_1$, where be $h \cdot a$ we mean the action of h on a. Then $\mathfrak{h} \ltimes \mathfrak{a}$ is a Lie algebra, and $0 \to \mathfrak{a} \to \mathfrak{h} \ltimes \mathfrak{a} \to \mathfrak{h} \to 0$ is a split short exact sequence in* LieAlg. □

4.5.0.4 Proposition *Let $\mathfrak{h} \curvearrowright \mathfrak{a}$ by derivations, and let $\mathfrak{g} = \mathfrak{h} \ltimes \mathfrak{a}$ be the semidirect product. Then the actions $\mathfrak{h} \curvearrowright \mathcal{U}\mathfrak{a}$ by derivations and $\mathfrak{a} \curvearrowright \mathcal{U}\mathfrak{a}$ by left-multiplication together make a $\mathfrak{g}$-action on $\mathcal{U}\mathfrak{a}$.*

Proof We need only check the commutator of $\mathfrak{h}$ with $\mathfrak{a}$. Let $u \in \mathcal{U}(\mathfrak{a})$, $h \in \mathfrak{h}$, and $a \in \mathfrak{a}$. Then $(h \circ a)u = h(au) = h(a)\,u + a\,h(u) = [h, a]u + a\,h(u)$. Thus $[h, a] \in \mathfrak{g}$ acts as the commutator of operators h and a on $\mathcal{U}(\mathfrak{a})$. □

4.5.0.5 Definition *An algebra U is* left-noetherian *if left ideals of U satisfy the* ascending chain condition*. I.e. if any chain of left ideals $I_1 \leq I_2 \leq \dots$ of U stabilizes.*

We refer the reader to any standard algebra textbook for a discussion of noetherian rings. For noncommutative ring theory see [Lam01, MR01, GW04].

4.5.0.6 Proposition *Let U be a filtered algebra. If $\operatorname{gr} U$ is left-noetherian, then so is U.*

In particular, $\mathcal{U}(\mathfrak{a})$ is left-noetherian, since $\operatorname{gr}\mathcal{U}(\mathfrak{a})$ is a polynomial ring on $\dim\mathfrak{a}$ generators.

Proof Let $I \leq U$ be a left ideal. We define $I_{\leq n} = I \cap U_{\leq n}$, and hence $I = \bigcup I_{\leq n}$. We define $\operatorname{gr} I = \bigoplus I_{\leq n}/I_{\leq n-1}$, and this is a left ideal in $\operatorname{gr} U$. If $I \leq J$, then $\operatorname{gr} I \leq \operatorname{gr} J$, using the fact that U injects into $\operatorname{gr} U$ as vector spaces.

So if we have an ascending chain $I_1 \leq I_2 \leq \dots$, then the corresponding chain $\operatorname{gr} I_1 \leq \operatorname{gr} I_2 \leq \dots$ eventually terminate by assumption: $\operatorname{gr} I_n = \operatorname{gr} I_{n_0}$ for $n \geq n_0$. But if $\operatorname{gr} I = \operatorname{gr} J$, then by induction on n, $I_{\leq n} = J_{\leq n}$, and so $I = J$. Hence the original sequence terminates. $\square$

4.5.0.7 Lemma *Let $\mathfrak{j} = \mathfrak{h} + \mathfrak{n}$ be a finite-dimensional Lie algebra, where $\mathfrak{h}$ is a subalgebra of $\mathfrak{j}$ and $\mathfrak{n}$ an ideal. Assume that $\mathfrak{g} \curvearrowright W$ is a finite-dimensional representation such that $\mathfrak{h}, \mathfrak{n} \curvearrowright W$ nilpotently. Then $\mathfrak{g} \curvearrowright W$ nilpotently.*

Proof If $W = 0$ there is nothing to prove. Otherwise, by Theorem 4.2.2.2 there is some $w \in W^{\mathfrak{n}}$, where $W^{\mathfrak{n}}$ is the subspace of W annihilated by $\mathfrak{n}$. Let $h \in \mathfrak{h}$ and $x \in \mathfrak{n}$. Then:

$$xhw = \underbrace{[x,h]}_{\in\mathfrak{n}} w + h \underbrace{xw}_{=0} = 0$$

Thus $hw \in V^{\mathfrak{n}}$, and so $w \in V^{\mathfrak{g}}$. By modding out and iterating, we see that $\mathfrak{g} \curvearrowright V$ nilpotently. $\square$

4.5.0.8 Theorem (Zassenhaus's Extension Lemma)
Let $\mathfrak{h}$ and $\mathfrak{a}$ be finite-dimensional Lie algebras so that $\mathfrak{h} \curvearrowright \mathfrak{a}$ by derivations, and let $\mathfrak{g} = \mathfrak{h} \ltimes \mathfrak{a}$. Moreover, let V be a finite-dimensional $\mathfrak{a}$-module, and let $\mathfrak{n}$ be the nilpotency ideal of $\mathfrak{a} \curvearrowright V$. If $[\mathfrak{h}, \mathfrak{a}] \leq \mathfrak{n}$, then there exists a finite-dimensional $\mathfrak{g}$-module W and a surjective $\mathfrak{a}$-module map $W \twoheadrightarrow V$, and so that the nilpotency ideal $\mathfrak{m}$ of $\mathfrak{g} \curvearrowright W$ contains $\mathfrak{n}$. If $\mathfrak{h} \curvearrowright \mathfrak{a}$ by nilpotents, then we can arrange for $\mathfrak{m} \subseteq \mathfrak{h}$ as well.

Proof Consider a Jordan-Holder series $0 = M_0 \subseteq M_1 \subseteq \dots \subseteq M_n = V$. Then $\mathfrak{n} = \bigcap \ker(M_i/M_{i-1})$ by Corollary 4.2.2.5. We define $N = \bigcap \ker\big(\mathcal{U}\mathfrak{a} \to \operatorname{End}(M_i/M_{i-1})\big)$, an ideal of $\mathcal{U}\mathfrak{a}$. Then $N \supseteq \mathfrak{n} \supseteq [\mathfrak{h}, \mathfrak{a}]$, and so N is an $\mathfrak{h}$-stable ideal of $\mathcal{U}\mathfrak{a}$ by the third part of Proposition 4.5.0.2, and N^k is $\mathfrak{h}$-stable by the fourth part.

Since $\mathcal{U}\mathfrak{a}$ is left-noetherian (Proposition 4.5.0.6), N^k is finitely generated for each k, and hence N^k/N^{k+1} is a finitely generated $\mathcal{U}\mathfrak{a}$ module. But the action $\mathcal{U}\mathfrak{a} \curvearrowright (N^k/N^{k+1})$ factors through $\mathcal{U}\mathfrak{a}/N$, so in fact N^k/N^{k+1} is a finitely generated $(\mathcal{U}\mathfrak{a}/N)$-module. But $\mathcal{U}\mathfrak{a}/N \cong \bigoplus \operatorname{Im}\big(\mathcal{U}\mathfrak{a} \to \operatorname{End}(M_i/M_{i-1})\big) \subseteq \bigoplus \operatorname{End}(M_i/M_{i-1})$, which is finite-dimensional. So $\mathcal{U}\mathfrak{a}/N$ is finite-dimensional, N^k/N^{k+1} a finitely-generated $(\mathcal{U}\mathfrak{a}/N)$-module, and hence N^k/N^{k+1} is finite-dimensional.

By construction, $N(M_k) \subseteq M_{k-1}$, so N^n annihilates V, where n is the length of the Jordan-Holder series $0 = M_0 \subseteq M_1 \subseteq \dots \subseteq \mathrm{M}(n) = V$. Let $\dim V = d$, and define

$$W \stackrel{\text{def}}{=} \bigoplus_{i=1}^{d} \mathcal{U}\mathfrak{a}/N^n$$

Then W is finite-dimensional since $\mathcal{U}\mathfrak{a}/N^n \cong \mathcal{U}\mathfrak{a}/N \oplus N/N^2 \oplus \cdots \oplus N^{n-1}/N^n$ as a vector space, and each summand is finite-dimensional. To construct the map $W \twoheadrightarrow V$, we pick a basis $\{v_i\}_{i=1}^d$ of V, and send $(0,\ldots,1,\ldots,0) \mapsto v_i$, where 1 is the image of $1 \in \mathcal{U}\mathfrak{a}$ in $\mathcal{U}\mathfrak{a}/N^n$, and it is in the ith spot. By construction $\mathcal{U}\mathfrak{a}_{\leq 0}$ acts as scalars, and so N does not contain $\mathcal{U}\mathfrak{a}_{\leq 0}$; thus the map is well-defined. Moreover, $\mathfrak{g} \curvearrowright \mathcal{U}\mathfrak{a}$ by Proposition 4.5.0.2, and N is $\mathfrak{h}$-stable and hence $\mathfrak{g}$-stable. Thus $\mathfrak{g} \curvearrowright W$ naturally, and the action $\mathcal{U}\mathfrak{a} \curvearrowright V$ factors through N^n, and so $W \twoheadrightarrow V$ is a map of $\mathfrak{g}$-modules.

By construction, N and hence $\mathfrak{n}$ acts nilpotently on W. But $\mathfrak{n}$ is an ideal of $\mathfrak{g}$: a general element of $\mathfrak{g}$ is of the form $h+a$ for $h \in \mathfrak{h}$ and $a \in \mathfrak{a}$, and $[h+a,\mathfrak{n}] = [h,\mathfrak{n}]+[a,\mathfrak{n}] \subseteq [h,\mathfrak{a}]+[a,\mathfrak{n}] \subseteq \mathfrak{n}+\mathfrak{n} = \mathfrak{n}$. So $\mathfrak{m} \supseteq \mathfrak{n}$, as $\mathfrak{m}$ is the largest nilpotency ideal of $\mathfrak{g} \curvearrowright W$.

We finish by considering the case when $\mathfrak{h} \curvearrowright \mathfrak{a}$ nilpotently. Then $\mathfrak{h} \curvearrowright W$ nilpotently, and since $[\mathfrak{h},\mathfrak{a}] \subseteq \mathfrak{n}$, $\mathfrak{h}+\mathfrak{n}$ is an ideal of $\mathfrak{g}$. By Lemma 4.5.0.7, $\mathfrak{h}+\mathfrak{n}$ acts nilpotently on W, and so is a subideal of $\mathfrak{m}$. $\square$

4.5.0.9 Corollary *Let $\mathfrak{r}$ be a solvable Lie algebra over characteristic 0, and let $\mathfrak{n}$ be its largest nilpotent ideal. Then every derivation of $\mathfrak{r}$ has image in $\mathfrak{n}$. In particular, if $\mathfrak{r}$ is an ideal of some larger Lie algebra $\mathfrak{g}$, then $[\mathfrak{g},\mathfrak{r}] \subseteq \mathfrak{n}$.*

Proof Let d be a derivation of $\mathfrak{r}$ and $\mathfrak{h} = \mathbb{K}d \oplus \mathfrak{r}$. Then $\mathfrak{h}$ is solvable by Proposition 4.2.1.9, and $\mathfrak{h}' \curvearrowright \mathfrak{h}$ nilpotently by Corollary 4.2.3.7. But $d(\mathfrak{r}) \subseteq \mathfrak{h}'$ and $\mathfrak{r}$ is an ideal of $\mathfrak{h}$, and so $d(\mathfrak{r})$ acts nilpotently on $\mathfrak{r}$, and is thus a subideal of $\mathfrak{n}$.

The second statement follows from the fact that $[g,-]$ is a derivation; this follows ultimately from the Jacobi identity. $\square$

4.5.0.10 Theorem (Ado's Theorem)
Let $\mathfrak{g}$ be a finite-dimensional Lie algebra over characteristic 0. Then $\mathfrak{g}$ has a faithful representation $\mathfrak{g} \hookrightarrow \mathfrak{gl}(V)$, and this representation can be chosen so that the largest nilpotent ideal $\mathfrak{n} \leq \mathfrak{g}$ acts nilpotently on V.

Ado originally proved a weaker version of Theorem 4.5.0.10 over $\mathbb{R}$. Harish-Chandra [HC49] gave essentially the version we present. A year earlier, Iwasawa [Iwa48] removed the dependence on characteristic, but without the nilpotency refinement. In fact, the theorem as stated holds over an arbitrary field in arbitrary characteristic, although our proof requires characteristic zero — the general theorem is due to Hochschild [Hoc66].

Proof We induct on $\dim \mathfrak{g}$. The $\mathfrak{g} = 0$ case is trivial, and we break the induction step into cases:

Case I: $\mathfrak{g} = \mathfrak{n}$ is nilpotent. Then $\mathfrak{g} \neq \mathfrak{g}'$, and so we choose a subspace $\mathfrak{a} \supseteq \mathfrak{g}'$ of codimensional 1 in $\mathfrak{g}$. This is automatically an ideal, and we pick $x \notin \mathfrak{a}$, and $\mathfrak{h} = \langle x \rangle$. Any one-dimensional subspace is a subalgebra, and $\mathfrak{g} = \mathfrak{h} \ltimes \mathfrak{a}$. By induction, $\mathfrak{a}$ has a faithful module V' and acts nilpotently.

The hypotheses of Theorem 4.5.0.8 are satisfied, and we get an $\mathfrak{a}$-module homomorphism $W \twoheadrightarrow V'$ with $\mathfrak{g} \curvearrowright W$ nilpotently. As yet, this might not be a faithful representation of $\mathfrak{g}$: certainly $\mathfrak{a}$ acts faithfully on W because it does so on V', but x might kill W. We pick a faithful nilpotent $\mathfrak{g}/\mathfrak{a} = \mathbb{K}$-module W_1, e.g. $x \mapsto \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \in \mathfrak{gl}(2)$. Then $V = W \oplus W_1$ is a faithful nilpotent $\mathfrak{g}$ representation.

Case II: $\mathfrak{g}$ is solvable but not nilpotent. Then $\mathfrak{g}' \leq \mathfrak{n} \lneq \mathfrak{g}$. We pick an ideal $\mathfrak{a}$ of codimension 1 in $\mathfrak{g}$ such that $\mathfrak{n} \subseteq \mathfrak{a}$, and x and $\mathfrak{h} = \mathbb{K}x$ as before, so that $\mathfrak{g} = \mathfrak{h} \ltimes \mathfrak{a}$. Then $\mathfrak{n}(\mathfrak{a}) \supseteq \mathfrak{n}$ — if a matrix acts nilpotently on $\mathfrak{g}$, then certainly it does so on $\mathfrak{a}$, and by construction $\mathfrak{n} \subseteq \mathfrak{a}$ — and we have a faithful module $\mathfrak{a} \curvearrowright V'$ by induction, with $\mathfrak{n}(\mathfrak{a}) \curvearrowright V'$ nilpotently. Then $[\mathfrak{h}, \mathfrak{a}] \curvearrowright V'$ nilpotently, since $[\mathfrak{h}, \mathfrak{a}] \subseteq \mathfrak{n}(\mathfrak{a})$ by Corollary 4.5.0.9, so we use Theorem 4.5.0.8 to get $\mathfrak{g} \curvearrowright W$ and an $\mathfrak{a}$-module map $W \twoheadrightarrow V'$, such that $\mathfrak{n} \curvearrowright W$ nilpotently. We form $V = W \oplus W_1$ as before so that $\mathfrak{g} \curvearrowright V$ is faithful. Since $\mathfrak{n}$ is contained in $\mathfrak{a}$ and $\mathfrak{a}$ acts as 0 on W_1, $\mathfrak{n}$ acts nilpotently on W.

Case III: general. By Theorem 4.4.4.12, there is a splitting $\mathfrak{g} = \mathfrak{s} \ltimes \mathfrak{r}$ with $\mathfrak{s}$ semisimple and $\mathfrak{r}$ solvable. By Case II, we have a faithful $\mathfrak{r}$-representation V' with $\mathfrak{n}(\mathfrak{r}) \curvearrowright V'$ nilpotently. By Corollary 4.5.0.9 the conditions of Theorem 4.5.0.8 apply, so we have $\mathfrak{g} \curvearrowright W$ and an $\mathfrak{r}$-module map $W \twoheadrightarrow V'$, and since $\mathfrak{n} \leq \mathfrak{r}$ we have $\mathfrak{n} \leq \mathfrak{n}(\mathfrak{r})$ so $\mathfrak{n} \curvearrowright W$ nilpotently. We want to get a faithful representation, and we need to make sure it is faithful on $\mathfrak{s}$. But $\mathfrak{s} = \mathfrak{g}/\mathfrak{r}$ is semisimple, so has no center, so $\mathrm{ad} : \mathfrak{s} \curvearrowright \mathfrak{s}$ is faithful. So we let $W_1 = \mathfrak{s} = \mathfrak{g}/\mathfrak{r}$ as $\mathfrak{g}$-modules, and $\mathfrak{g} \curvearrowright V = W \oplus W_1$ is faithful with $\mathfrak{n}$ acting as 0 on W_1 and nilpotently on W. $\square$

Exercises

1. Classify the 3-dimensional Lie algebras $\mathfrak{g}$ over an algebraically closed field $\mathbb{K}$ of characteristic zero by showing that if $\mathfrak{g}$ is not a direct product of smaller Lie algebras, then either
    - $\mathfrak{g} \cong \mathfrak{sl}(2, \mathbb{K})$,
    - $\mathfrak{g}$ is isomorphic to the nilpotent Heisenberg Lie algebra $\mathfrak{h}$ with basis X, Y, Z such that Z is central and $[X, Y] = Z$, or
    - $\mathfrak{g}$ is isomorphic to a solvable algebra $\mathfrak{s}$ which is the semidirect product of the abelian algebra $\mathbb{K}^2$ by an invertible derivation. In particular $\mathfrak{s}$ has basis X, Y, Z such that $[Y, Z] = 0$, and $\mathrm{ad}\, X$ acts on $\mathbb{K}Y + \mathbb{K}Z$ by an invertible matrix, which, up to change of basis in $\mathbb{K}Y + \mathbb{K}Z$ and rescaling X, can be taken to be either $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} \lambda & 0 \\ 0 & 1 \end{smallmatrix}\right)$ for some nonzero $\lambda \in \mathbb{K}$.

2. (a) Show that the Heisenberg Lie algebra $\mathfrak{h}$ in Problem 1 has the property that Z acts nilpotently in every finite-dimensional module, and as zero in every simple finite-dimensional module.

    (b) Construct a simple infinite-dimensional $\mathfrak{h}$-module in which Z acts as a non-zero scalar. [Hint: take X and Y to be the operators $\frac{\mathrm{d}}{\mathrm{d}t}$ and t on $\mathbb{K}[t]$.]

3. Construct a simple 2-dimensional module for the Heisenberg algebra $\mathfrak{h}$ over any field $\mathbb{K}$ of characteristic 2. In particular, if $\mathbb{K} = \bar{\mathbb{K}}$, this gives a counterexample to Lie's theorem in non-zero characteristic.

4. Let $\mathfrak{g}$ be a finite-dimensional Lie algebra over $\mathbb{K}$.

   (a) Show that the intersection $\mathfrak{n}$ of the kernels of all finite-dimensional simple $\mathfrak{g}$-modules can be characterized as the largest ideal of $\mathfrak{g}$ which acts nilpotently in every finite-dimensional $\mathfrak{g}$-module. It is called the *nilradical* of $\mathfrak{g}$.

   (b) Show that the nilradical of $\mathfrak{g}$ is contained in $\mathfrak{g}' \cap \operatorname{rad}(\mathfrak{g})$.

   (c) Let $\mathfrak{h} \subseteq \mathfrak{g}$ be a subalgebra and V a $\mathfrak{g}$-module. Given a linear functional $\lambda : \mathfrak{h} \to \mathbb{K}$, define the associated weight space to be $V_\lambda = \{v \in V : Hv = \lambda(H)v \text{ for all } H \in \mathfrak{h}\}$. Assuming $\operatorname{char}(\mathbb{K}) = 0$, adapt the proof of Lie's theorem to show that if $\mathfrak{h}$ is an ideal and V is finite-dimensional, then V_λ is a $\mathfrak{g}$-submodule of V.

   (d) Show that if $\operatorname{char}(\mathbb{K}) = 0$ then the nilradical of $\mathfrak{g}$ is equal to $\mathfrak{g}' \cap \operatorname{rad}(\mathfrak{g})$. [Hint: assume without loss of generality that $\mathbb{K} = \bar{\mathbb{K}}$ and obtain from Lie's theorem that any finite-dimensional simple $\mathfrak{g}$-module V has a non-zero weight space for some weight λ on $\mathfrak{g}' \cap \operatorname{rad}(\mathfrak{g})$. Then use (c) to deduce that $\lambda = 0$ if V is simple.]

5. Let $\mathfrak{g}$ be a finite-dimensional Lie algebra over $\mathbb{K}$, $\operatorname{char}(\mathbb{K}) = 0$. Prove that the largest nilpotent ideal of $\mathfrak{g}$ is equal to the set of elements of $\mathfrak{r} = \operatorname{rad}\mathfrak{g}$ which act nilpotently in the adjoint action on $\mathfrak{g}$, or equivalently on $\mathfrak{r}$. In particular, it is equal to the largest nilpotent ideal of $\mathfrak{r}$.

6. Prove that the Lie algebra $\mathfrak{sl}(2, \mathbb{K})$ of 2×2 matrices with trace zero is simple, over a field $\mathbb{K}$ of any characteristic $\neq 2$. In characteristic 2, show that it is nilpotent.

7. In this exercise, we'll deduce from the standard functorial properties of Ext groups and their associated long exact sequences that $\operatorname{Ext}^1(N, M)$ bijectively classifies extensions $0 \to M \to V \to N \to 0$ up to isomorphism, for modules over any associative ring with unity.

   (a) Let F be a free module with a surjective homomorphism onto N, so we have an exact sequence $0 \to K \to F \to N \to 0$. Use the long exact sequence to produce an isomorphism of $\operatorname{Ext}^1(N, M)$ with the cokernel of $\operatorname{Hom}(F, M) \to \operatorname{Hom}(K, M)$.

   (b) Given $\phi \in \operatorname{Hom}(K, M)$, construct V as the quotient of $F \oplus M$ by the graph of $-\phi$ (note that this graph is a submodule of $K \oplus M \subseteq F \oplus M$).

   (c) Use the functoriality of Ext and the long exact sequences to show that the characteristic class in $\operatorname{Ext}^1(N, M)$ of the extension constructed in (b) is the element represented by the chosen ϕ, and conversely, that if ϕ represents the characteristic class of a given extension, then the extension constructed in (b) is isomorphic to the given one.

8. Calculate $\operatorname{Ext}^i(\mathbb{K}, \mathbb{K})$ for all i for the trivial representation $\mathbb{K}$ of $\mathfrak{sl}(2, \mathbb{K})$, where $\operatorname{char}(\mathbb{K}) = 0$. Conclude that the theorem that $\operatorname{Ext}^i(M, N) = 0$ for $i = 1, 2$ and all finite-dimensional representations M, N of a semi-simple Lie algebra $\mathfrak{g}$ does not extend to $i > 2$.

9. Let $\mathfrak{g}$ be a finite-dimensional Lie algebra. Show that $\operatorname{Ext}^1(\mathbb{K}, \mathbb{K})$ can be canonically identified with the dual space of $\mathfrak{g}/\mathfrak{g}'$, and therefore also with the set of 1-dimensional $\mathfrak{g}$-modules, up to isomorphism.

10. Let $\mathfrak{g}$ be a finite-dimensional Lie algebra. Show that $\operatorname{Ext}^1(\mathbb{K}, \mathfrak{g})$ can be canonically identified with the quotient $\operatorname{Der}(\mathfrak{g})/\operatorname{Inn}(\mathfrak{g})$, where $\operatorname{Der}(\mathfrak{g})$ is the space of derivations of $\mathfrak{g}$, and $\operatorname{Inn}(\mathfrak{g})$ is the subspace of inner derivations, that is, those of the form $d(x) = [y, x]$ for some $y \in \mathfrak{g}$. Show that this also classifies Lie algebra extensions $\hat{\mathfrak{g}}$ containing $\mathfrak{g}$ as an ideal of codimension 1.

11. Let $\mathfrak{g}$ be a finite-dimensional Lie algebra over $\mathbb{K}$, $\operatorname{char}(\mathbb{K}) = 0$. The Malcev-Harish-Chandra theorem says that all Levi subalgebras $\mathfrak{s} \subseteq \mathfrak{g}$ are conjugate under the action of the group $\exp \operatorname{ad} \mathfrak{n}$, where $\mathfrak{n}$ is the largest nilpotent ideal of $\mathfrak{g}$ (note that $\mathfrak{n}$ acts nilpotently on $\mathfrak{g}$, so the power series expression for $\exp \operatorname{ad} X$ reduces to a finite sum when $X \in \mathfrak{n}$).

    (a) Show that the reduction we used to prove Levi's theorem by induction in the case where the radical $\mathfrak{r} = \operatorname{rad} \mathfrak{g}$ is not a minimal ideal also works for the Malcev-Harish-Chandra theorem. More precisely, show that if $\mathfrak{r}$ is nilpotent, the reduction can be done using any nonzero ideal $\mathfrak{m}$ properly contained in $\mathfrak{r}$. If $\mathfrak{r}$ is not nilpotent, use Problem 4 to show that $[\mathfrak{g}, \mathfrak{r}] = \mathfrak{r}$, then make the reduction by taking $\mathfrak{m}$ to contain $[\mathfrak{g}, \mathfrak{r}]$.

    (b) In general, given a semidirect product $\mathfrak{g} = \mathfrak{h} \ltimes \mathfrak{m}$, where $\mathfrak{m}$ is an abelian ideal, show that $\operatorname{Ext}^1_{\mathcal{U}(\mathfrak{h})}(\mathbb{K}, \mathfrak{m})$ classifies subalgebras complementary to $\mathfrak{m}$, up to conjugacy by the action of $\exp \operatorname{ad} \mathfrak{m}$. Then use the vanishing of $\operatorname{Ext}^1(M, N)$ for finite-dimensional modules over a semi-simple Lie algebra to complete the proof of the Malcev-Harish-Chandra theorem.

Chapter 5

Classification of Semisimple Lie Algebras

Henceforth every Lie algebra, except when otherwise marked, is finite-dimensional over a field of characteristic 0.

5.1 Classical Lie algebras over $\mathbb{C}$

5.1.1 Reductive Lie algebras

5.1.1.1 Lemma / Definition *A Lie algebra* $\mathfrak{g}$ *is* reductive *if* $(\mathfrak{g}, \mathrm{ad})$ *is completely reducible.*

A Lie algebra is reductive if and only if it is of the form $\mathfrak{g} = \mathfrak{s} \times \mathfrak{a}$ *where* $\mathfrak{s}$ *is semisimple and* $\mathfrak{a}$ *is abelian. Moreover,* $\mathfrak{a} = Z(\mathfrak{g})$ *and* $\mathfrak{s} = \mathfrak{g}'$.

Proof Let $\mathfrak{g}$ be a reductive Lie algebra; then $\mathfrak{g} = \bigoplus \mathfrak{a}_i$ as $\mathfrak{g}$-modules, where each $\mathfrak{a}_i$ is an ideal of $\mathfrak{g}$ and $[\mathfrak{a}_i, \mathfrak{a}_j] \subseteq \mathfrak{a}_i \cap \mathfrak{a}_j = 0$ for $i \neq j$. Thus $\mathfrak{g} = \prod \mathfrak{a}_i$ as Lie algebras, and each $\mathfrak{a}_i$ is either simple non-abelian or one-dimensional. $\square$

5.1.1.2 Proposition *Any $*$-closed subalgebra of* $\mathfrak{gl}(n, \mathbb{C})$ *is reductive.*

Proof We define a symmetric real-valued bilinear form $(,)$ on $\mathfrak{gl}(n, \mathbb{C})$ by $(x, y) = \mathrm{real}\big(\mathrm{tr}(xy^*)\big)$. Then $(x, x) = \sum |x_{ij}|^2$, so $(,)$ is positive-definite. Moreover:

$$([z, x], y) = -(x, [z^*, y])$$

so $[z^*, -]$ is adjoint to $[-, z]$.

Let $\mathfrak{g} \subseteq \mathfrak{gl}(n, \mathbb{C})$ be any subalgebra and $\mathfrak{a} \leq \mathfrak{g}$ an ideal. Then $\mathfrak{a}^\perp \subseteq \mathfrak{g}^*$ by invariance, where $\mathfrak{g}^*$ is the Lie algebra of Hermitian conjugates of elements of $\mathfrak{g}$. If $\mathfrak{g}$ is $*$-closed, then $\mathfrak{g}^* = \mathfrak{g}$ and $\mathfrak{a}^\perp$ is an ideal of $\mathfrak{g}$. By positive-definiteness, $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^\perp$, and we rinse and repeat to write $\mathfrak{g}$ is a sum of irreducibles. $\square$

5.1.1.3 Example The classical Lie algebras $\mathfrak{sl}(n, \mathbb{C}) \stackrel{\text{def}}{=} \{x \in \mathfrak{gl}(n) \text{ s.t. } \mathrm{tr}\, x = 0\}$, $\mathfrak{so}(n, \mathbb{C}) \stackrel{\text{def}}{=} \{x \in \mathfrak{gl}(n) \text{ s.t. } x + x^T = 0\}$, and $\mathfrak{sp}(n, \mathbb{C}) \stackrel{\text{def}}{=} \{x \in \mathfrak{gl}(2n) \text{ s.t. } jx + x^T j = 0\}$ are reductive. Indeed, since they have no center except in very low dimensions, they are all semisimple. We will see later that they are all simple, except in a few low dimensions.

Since a real Lie algebra $\mathfrak{g}$ is semisimple if $\mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C}$ is, the real Lie algebras $\mathfrak{sl}(n, \mathbb{R})$, $\mathfrak{so}(n, \mathbb{R})$, and $\mathfrak{sp}(n, \mathbb{R})$ also are semisimple. $\Diamond$

5.1.2 Guiding examples: $\mathfrak{sl}(n)$ and $\mathfrak{sp}(n)$ over $\mathbb{C}$

Let $\mathfrak{g} = \mathfrak{sl}(n)$ or $\mathfrak{sp}(n)$. We extract an abelian subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$. For $\mathfrak{sl}_n$ we use the diagonal traceless matrices:

$$\mathfrak{h} \stackrel{\text{def}}{=} \left\{ \begin{bmatrix} z_1 & & \\ & \ddots & \\ & & z_n \end{bmatrix} \text{ s.t. } \sum z_i = 0 \right\}$$

For $\mathfrak{sp}(n) \stackrel{\text{def}}{=} \{x \in \mathfrak{gl}(2n) \text{ s.t. } jx + x^T j = 0\}$, it will be helpful to redefine j. We can use any j which defines a non-degenerate antisymmetric bilinear form, and we take:

$$j = \left[\begin{array}{ccc|ccc} & & & & & 1 \\ & 0 & & & \iddots & \\ & & & 1 & & \\ \hline & & -1 & & & \\ & \iddots & & & 0 & \\ -1 & & & & & \end{array}\right]$$

Let a^R be the matrix a reflected across the antidiagonal. Then we can define $\mathfrak{sp}(n)$ in block diagonal form:

$$\mathfrak{sp}(n) = \left\{ \left[\begin{array}{c|c} a & b \\ \hline c & d \end{array}\right] \in \text{Mat}(2n) = \text{Mat}(2, \text{Mat}(n)) \text{ s.t. } d = -a^R, b = b^R, c = c^R \right\} \tag{5.1.2.1}$$

In this basis, we take as our abelian subalgebra

$$\mathfrak{h} \stackrel{\text{def}}{=} \left\{ \left[\begin{array}{ccc|ccc} z_1 & & & & & \\ & \ddots & & & 0 & \\ & & z_n & & & \\ \hline & & & -z_n & & \\ & 0 & & & \ddots & \\ & & & & & -z_1 \end{array}\right] \right\}$$

5.1.2.2 Proposition *Let $\mathfrak{g} = \mathfrak{sl}(n)$ or $\mathfrak{sp}(n)$. For $\mathfrak{h} \leq \mathfrak{g}$ defined above, the adjoint action $\text{ad} : \mathfrak{h} \curvearrowright \mathfrak{g}$ is diagonal.*

Proof We make explicit the basis of $\mathfrak{g}$. For $\mathfrak{g} = \mathfrak{sl}(n)$, the natural basis is $\{e_{ij}\}_{i \neq j} \cup \{e_{ii} - e_{i+1,i+1}\}_{i=1}^{n-1}$, where e_{ij} is the matrix with a 1 in the (ij) spot and 0s elsewhere. In particular,

$\{e_{ii} - e_{i+1,i+1}\}_{i=1}^{n-1}$ is a basis of $\mathfrak{h}$. Let $h \in \mathfrak{h}$ be

$$h = \begin{bmatrix} z_1 & & \\ & \ddots & \\ & & z_n \end{bmatrix}$$

Then $[h, e_{ij}] = (z_i - z_j)e_{ij}$, and $[h, h'] = 0$ when $h' \in \mathfrak{h}$.

For $\mathfrak{g} = \mathfrak{sp}(n)$, the natural basis suggested by equation (5.1.2.1) is

$$\left\{a_{ij} \overset{\text{def}}{=} \left[\begin{array}{c:c} e_{ij} & 0 \\ \hdashline 0 & -e_{n+1-j,n+1-i} \end{array}\right]\right\} \cup \left\{b_{ij} \overset{\text{def}}{=} \left[\begin{array}{c:c} 0 & e_{ij} + e_{n+1-j,n+1-i} \\ \hdashline 0 & 0 \end{array}\right] \text{ s.t. } i + j \leq n+1\right\}$$
$$\cup \left\{c_{ij} \overset{\text{def}}{=} \left[\begin{array}{c:c} 0 & 0 \\ \hdashline e_{ij} + e_{n+1-j,n+1-i} & 0 \end{array}\right] \text{ s.t. } i + j \leq n+1\right\} \tag{5.1.2.3}$$

Of course, when $i = j$, then $\left\{\left[\begin{array}{c:c} e_{ii} & 0 \\ \hdashline 0 & -e_{n+1-i,n+1-i} \end{array}\right]\right\}$ is a basis of $\mathfrak{h}$. Let $h \in \mathfrak{h}$ be given by

$$h = \left[\begin{array}{ccc:ccc} z_1 & & & & & \\ & \ddots & & & 0 & \\ & & z_n & & & \\ \hdashline & & & -z_n & & \\ & 0 & & & \ddots & \\ & & & & & -z_1 \end{array}\right]$$

Then $[h, a_{ij}] = (z_i - z_j)a_{ij}$, $[h, b_{ij}] = (z_i + z_j)b_{ij}$, and $[h, c_{ij}] = (-z_i - z_j)c_{ij}$. □

5.1.2.4 Definition *Let $\mathfrak{h}$ be a maximal abelian subalgebra of a finite-dimensional Lie algebra $\mathfrak{g}$ so that* $\operatorname{ad} : \mathfrak{h} \curvearrowright \mathfrak{g}$ *is diagonalizable, so diagonal in an eigenbasis. Write $\mathfrak{h}^*$ for the vector space dual to $\mathfrak{h}$. Each eigenbasis element of $\mathfrak{g}$ defines an eigenvalue to each $h \in \mathfrak{h}$, and this assignment is linear in $\mathfrak{h}$; thus, the eigenbasis of $\mathfrak{g}$ picks out a vector $\alpha \in \mathfrak{h}^*$. The set of such vectors are the* roots *of the pair $(\mathfrak{g}, \mathfrak{h})$.*

We will refine this definition in Definition 5.4.1.1, and we will prove that the set of roots of a semisimple Lie algebra $\mathfrak{g}$ is determined up to isomorphism by $\mathfrak{g}$ (in particular, it does not depend on the subalgebra $\mathfrak{h}$).

5.1.2.5 Example When $\mathfrak{g} = \mathfrak{sl}(n)$ and $\mathfrak{h}$ is as above, the roots are $\{0\} \cup \{z_i - z_j\}_{i \neq j}$, where $\{z_i\}_{i=1}^n$ are the natural linear functionals $\mathfrak{h} \to \mathbb{C}$. When $\mathfrak{g} = \mathfrak{sp}(n)$ and $\mathfrak{h}$ is as above, the roots are $\{0\} \cup \{\pm 2z_i\} \cup \{\pm z_i \pm z_j\}_{i \neq j}$. ♢

5.1.2.6 Lemma / Definition *Let $\mathfrak{g}$ and $\mathfrak{h} \leq \mathfrak{g}$ as in Definition 5.1.2.4. Then the roots break $\mathfrak{g}$ into eigenspaces:*

$$\mathfrak{g} = \bigoplus_{\alpha \text{ a root}} \mathfrak{g}_\alpha = \mathfrak{h} \oplus \bigoplus_{\alpha \neq 0 \text{ a root}} \mathfrak{g}_\alpha$$

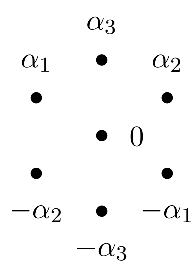
In particular, since $\mathfrak{h}$ *is a maximal abelian subalgebra, the* 0*-eigenspace of* $\mathfrak{h} \curvearrowright \mathfrak{g}$ *is precisely* $\mathfrak{g}_0 = \mathfrak{h}$*. Then the spaces* $\mathfrak{g}_\alpha$ *are the* root spaces *of the pair* $(\mathfrak{g}, \mathfrak{h})$*. By the Jacobi identity,* $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}$*.*□

5.1.2.7 Lemma *When* $\mathfrak{g} = \mathfrak{sl}(n)$ *or* $\mathfrak{sp}(n)$ *and* $\mathfrak{h}$ *is as above, then for* $\alpha \neq 0$ *the root space* $\mathfrak{g}_\alpha \subseteq \mathfrak{g}$ *is one-dimensional. Let* $\mathfrak{h}_\alpha \stackrel{\text{def}}{=} [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$*. Then* $\mathfrak{h}_\alpha = \mathfrak{h}_{-\alpha}$ *is one-dimensional, and* $\mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha} \oplus \mathfrak{h}_\alpha$ *is a subalgebra of* $\mathfrak{g}$ *isomorphic to* $\mathfrak{sl}(2)$*.*

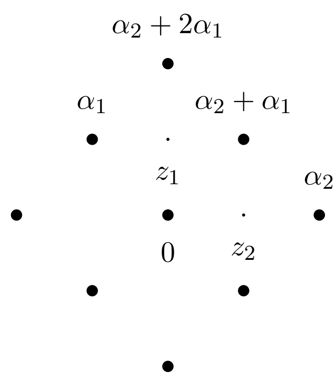
Proof For each root α, pick a basis element $g_{\pm\alpha} \in \mathfrak{g}_{\pm\alpha}$ (in particular, we can use the eigenbasis of $\mathfrak{h} \curvearrowright \mathfrak{g}$ given above), and define $h_\alpha \stackrel{\text{def}}{=} [g_\alpha, g_{-\alpha}]$. Define $\alpha(h_\alpha) = a$ so that $[h_\alpha, g_{\pm\alpha}] = \pm a g_\alpha$; one can check directly that $a \neq 0$. For the isomorphism, we use the fact that $\mathbb{C}$ is algebraically closed.□

5.1.2.8 Definition *Let* $\mathfrak{g}$ *and* $\mathfrak{h} \leq \mathfrak{g}$ *as in Definition 5.1.2.4. The* rank *of* $\mathfrak{g}$ *is the dimension of* $\mathfrak{h}$*, or equivalently the dimension of the dual space* $\mathfrak{h}^*$*.*

5.1.2.9 Example The Lie algebras $\mathfrak{sl}(3)$ and $\mathfrak{sp}(2)$ are rank-two. For $\mathfrak{g} = \mathfrak{sl}(3)$, the dual space $\mathfrak{h}^*$ to $\mathfrak{h}$ spanned by the vectors $z_1 - z_2$ and $z_2 - z_3$ naturally embeds in a three-dimensional vector space spanned by $\{z_1, z_2, z_3\}$, and we choose an inner product on this space in which $\{z_i\}$ is an orthonormal basis. Let $\alpha_1 = z_1 - z_2$, $\alpha_2 = z_2 - z_3$, and $\alpha_3 = z_1 - z_3$. Then the roots $\{0, \pm\alpha_i\}$ form a perfect hexagon:



For $\mathfrak{g} = \mathfrak{sp}(2)$, we have $\mathfrak{h}^*$ spanned by $\{z_1, z_2\}$, and we choose an inner product in which this is an orthonormal basis. Let $\alpha_1 = z_1 - z_2$ and $\alpha_2 = 2z_2$. The roots form a diamond:

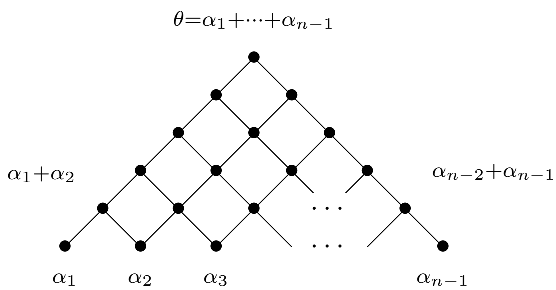


◊

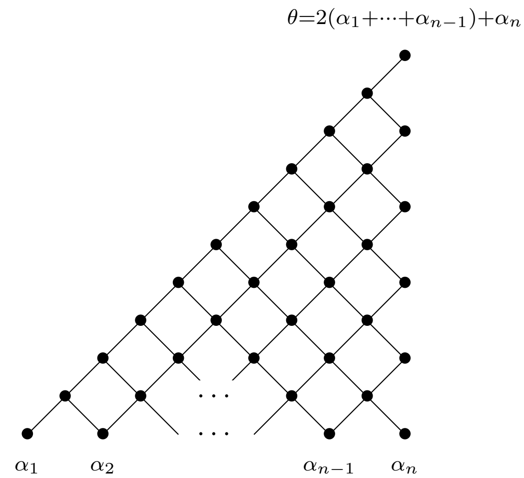
5.1.2.10 Lemma *Let* $\mathfrak{g}, \mathfrak{h} \leq \mathfrak{g}$ *be as in Definition 5.1.2.4. Let* $v \in \mathfrak{h}$ *be chosen so that* $\alpha(v) \neq 0$ *for every root* α*. Then* v *divides the roots into* positive roots *and* negative roots *according to the sign of* $\alpha(v)$*. A* simple root *is any positive root that is not expressible as a sum of positive roots.*□

5.1.2.11 Example Let $\mathfrak{g} = \mathfrak{sl}(n)$ or $\mathfrak{sp}(n)$, and choose $v \in \mathfrak{h}$ so that $z_1(v) > z_2(v) > \cdots > z_n(v) > 0$. The positive roots of $\mathfrak{sl}(n)$ are $\{z_i - z_j\}_{i<j}$, and the positive roots of $\mathfrak{sp}(n)$ are $\{z_i - z_j\}_{i<j} \cup \{z_i + z_i\} \cup \{2z_i\}$. The simple roots of $\mathfrak{sl}(n)$ are $\{\alpha_i = z_i - z_{i+1}\}_{i=1}^{n-1}$, and the simple roots of $\mathfrak{sp}(n)$ are $\{\alpha_i = z_i - z_{i+1}\}_{i=1}^{n-1} \cup \{2z_n\}$. In each case, the simple roots are a basis of $\mathfrak{h}^*$. Moreover, the roots are in the $\mathbb{Z}$-span of the simple roots, i.e. the lattice generated by the simple roots, and the positive roots are in the intersection of this lattice with the positive cone, so that the positive roots are in the $\mathbb{N}$-span of the simple roots.

We partially order the positive roots by saying that $\alpha < \beta$ if $\beta - \alpha$ is a positive root. Under this partial order there is a unique maximal positive root θ, the *highest root*; for $\mathfrak{sl}(n)$ we have $\theta = z_1 - z_n = \alpha_1 + \cdots + \alpha_{n-1}$, and for $\mathfrak{sp}(n)$ we have $\theta = 2z_1 = 2(\alpha_1 + \cdots + \alpha_{n-1}) + \alpha_n$. We draw these partial orders:



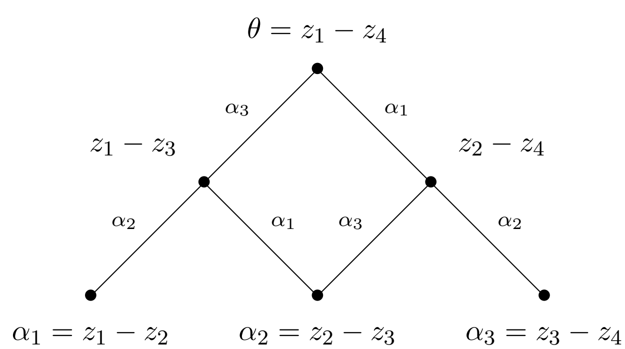
$\mathfrak{sl}(n)$



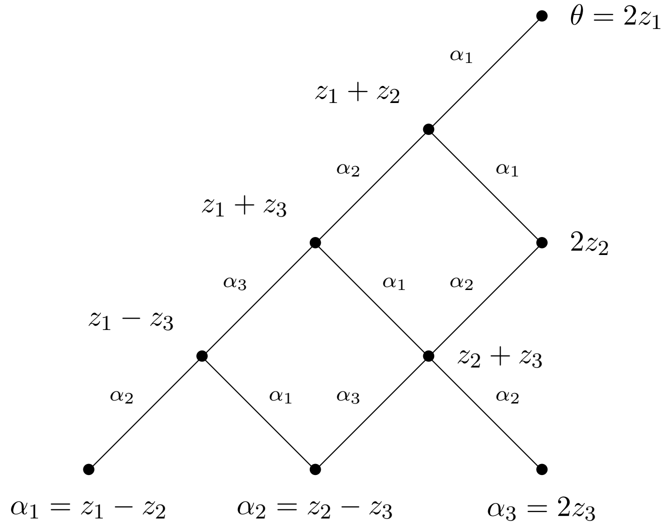
$\mathfrak{sp}(n)$

$\lozenge$

To make this very clear, we draw the rank-three pictures fully labeled (edges by the difference between consecutive nodes):



$\mathfrak{sl}(4)$



$\mathfrak{sp}(3)$

Using these pictures of $\mathfrak{sl}(n)$ and $\mathfrak{sp}(n)$, we can directly prove the following:

5.1.2.12 Proposition *The Lie algebras $\mathfrak{sl}(n,\mathbb{C})$ and $\mathfrak{sp}(n,\mathbb{C})$ are simple.*

Proof Let $\mathfrak{g} = \mathfrak{sl}(n)$ or $\mathfrak{sp}(n)$, and $\mathfrak{h}$, the systems of positive and simple roots, and θ the highest root as above. Recall that for each root $\alpha \neq 0$, the root space $\mathfrak{g}_\alpha$ is one-dimensional, and we pick an eigenbasis $\{g_\alpha\}_{\alpha\neq 0} \cup \{h_i\}_{i=1}^{\text{rank}}$ for the action $\mathfrak{h} \curvearrowright \mathfrak{g}$.

Let $x \in \mathfrak{g}$. It is a standard exercise from linear algebra that $\mathfrak{h}x$ is the span of the eigenvectros g_α, $\alpha \neq 0$, for which the coefficient of x in the eigenbasis is non-zero. In particular, if $x \in \mathfrak{g} \smallsetminus \mathfrak{h}$, then $[\mathfrak{h}, x]$ includes some g_α. By switching the roles of positive and negative roots if necessary, we can assure that α is positive; thus $[\mathfrak{h}, x] \supseteq \mathfrak{g}_\alpha$ for some positive α.

One can check directly that if $\alpha, \beta, \alpha+\beta$ are all nonzero roots, then $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \mathfrak{g}_{\alpha+\beta}$. In particular, for any positive root α, $\theta - \alpha$ is a positive root, and so $[\mathfrak{g}, \mathfrak{g}_\alpha] \supseteq [\mathfrak{g}_{\theta-\alpha}, \mathfrak{g}_\alpha] = \mathfrak{g}_\theta$. In particular, $g_\theta \in [\mathfrak{g}, x]$.

But $[\mathfrak{g}_{\theta-\alpha}, \mathfrak{g}_\theta] = \mathfrak{g}_\alpha$, and so $[\mathfrak{g}, g_\theta]$ generates all g_α for α a positive root. We saw already (Lemma 5.1.2.7) that $[\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] = \mathfrak{h}_\alpha$ is non-zero, and that $[\mathfrak{g}_{\pm\alpha}, \mathfrak{h}_\alpha] = \mathfrak{g}_{\pm\alpha}$. Thus $[\mathfrak{g}, \mathfrak{g}_\alpha] \supseteq \mathfrak{g}_{-\alpha}$, and in particular g_θ generates every g_α for $\alpha \neq 0$, and every h_α. Then g_θ generates all of $\mathfrak{g}$.

Thus x generates all of $\mathfrak{g}$ for any $x \in \mathfrak{g} \smallsetminus \mathfrak{h}$. If $x \in \mathfrak{h}$, then $\alpha(x) \neq 0$ for some α, and then $[\mathfrak{g}_\alpha, x] = \mathfrak{g}_\alpha$, and we repeat the proof with some nonzero element of $\mathfrak{g}_\alpha$. Hence $\mathfrak{g}$ is simple. $\square$

When $\mathfrak{g} = \mathfrak{sl}(n)$, let ϵ_i refer to the matrix e_{ii}, and when $\mathfrak{g} = \mathfrak{sp}(n)$, let ϵ_i refer to the matrix $\left[\begin{array}{c|c} e_{ii} & 0 \\ \hline 0 & -e_{n+1-i,n+1-i} \end{array}\right]$. We construct a linear isomorphism $\mathfrak{h}^* \xrightarrow{\sim} \mathfrak{h}$ by assigning an element $\alpha_i^\vee$ of $\mathfrak{h}$ to each simple root α_i as follows: to $\alpha_i = z_i - z_{i+1}$ for $1 \leq i \leq n-1$ we assign $\alpha_i^\vee = \epsilon_i - \epsilon_{i+1}$, and to $\alpha_n = 2z_n$ a root of $\mathfrak{sp}(n)$ we assign $\alpha_n^\vee = \epsilon_n$. In particular, $\alpha_i(h_i) = 2$ for each simple root. We define the *Cartan matrix* a by $a_{ij} \stackrel{\text{def}}{=} \alpha_i(h_j)$.

5.1.2.13 Example For $\mathfrak{sl}(n)$, we have the following $(n-1)\times(n-1)$ matrix:

$$a = \begin{bmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & & \vdots \\ 0 & -1 & \ddots & \ddots & 0 \\ \vdots & & \ddots & 2 & -1 \\ 0 & \dots & 0 & -1 & 2 \end{bmatrix}$$

For $\mathfrak{sp}(n)$, we have the following $n \times n$ matrix:

$$a = \left[\begin{array}{ccccc:c} 2 & -1 & 0 & \dots & 0 & 0 \\ -1 & 2 & -1 & & \vdots & \vdots \\ 0 & -1 & \ddots & \ddots & 0 & \vdots \\ \vdots & & \ddots & 2 & -1 & 0 \\ 0 & \dots & 0 & -1 & 2 & -1 \\ \hdashline 0 & \dots & \dots & 0 & -2 & 2 \end{array}\right]$$

$\Diamond$

To each of the above matrices we associate a *Dynkin diagram*. This is a graph with a node for each simple root, and edges assigned by: i and j are not connected if $a_{ij} = 0$; they are singly connected if a_{ij} is a block $\left[\begin{smallmatrix} 2 & -1 \\ -1 & 2 \end{smallmatrix}\right]$; we put a double arrow from j to i when the (i,j)-block is $\left[\begin{smallmatrix} 2 & -1 \\ -2 & 2 \end{smallmatrix}\right]$. So for $\mathfrak{sl}(n)$ we get the graph $\bullet\!-\!\!-\!\bullet\!-\!\!-\cdots-\!\!-\!\bullet$, and for $\mathfrak{sp}(n)$ we get $\bullet\!-\!\!-\!\bullet\!-\!\!-\cdots-\!\!-\!\bullet\Leftarrow\bullet$.

5.1.2.14 Lemma / Definition *The identification $\mathfrak{h}^* \xrightarrow{\sim} \mathfrak{h}$ lets us construct reflections of $\mathfrak{h}^*$ by $s_i : \alpha \mapsto \alpha - \langle \alpha, \alpha_i^\vee \rangle \alpha$, where $\langle , \rangle$ is the pairing $\mathfrak{h}^* \otimes \mathfrak{h} \to \mathbb{C}$ that we had earlier written as $\langle \alpha, \beta \rangle = \alpha(\beta)$. These reflections generated the* Weyl group W.

For each of $\mathfrak{sl}(n)$ and $\mathfrak{sp}(n)$, let $R \subseteq \mathfrak{h}^*$ be the set of roots and W the Weyl group. Then $W \curvearrowright R \smallsetminus \{0\}$. In particular, for $\mathfrak{sl}(n)$, we have $W = S_n$ the symmetric group on n letters, where the reflection $(i, i+1)$ acts as s_i; $W \curvearrowright R \smallsetminus \{0\}$ is transitive. For $\mathfrak{sp}(n)$, we have $W = S_n \ltimes (\mathbb{Z}/2)^n$, the hyperoctahedral group, *generated by the reflections $s_i = (i, i+1) \in S_n$ and s_n the sign change, and the action $W \curvearrowright R \smallsetminus \{0\}$ has two orbits.* $\square$

We will spend the rest of this chapter showing that the pictures of $\mathfrak{sl}(n)$ and $\mathfrak{sp}(n)$ in this section is typical of simple Lie algebras over $\mathbb{C}$.

5.2 Representation theory of $\mathfrak{sl}(2)$

Our hero for this section is the Lie algebra $\mathfrak{sl}(2, \mathbb{C}) \overset{\text{def}}{=} \langle e, h, f : [e, f] = h, [h, e] = 2e, [h, f] = -2f \rangle = \{x \in \mathrm{Mat}(2, \mathbb{C}) \text{ s.t. } \operatorname{tr} x = 0\}$.

5.2.0.1 Example As a subalgebra of $\mathrm{Mat}(2, \mathbb{C})$, $\mathfrak{sl}(2)$ has a tautological representation on $\mathbb{C}^2$, given by $E \mapsto \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$, $F \mapsto \left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right]$, and $H \mapsto \left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right]$. Let v_0 and v_1 be the basis vectors of $\mathbb{C}^2$. Then the representation $\mathfrak{sl}(2) \curvearrowright \mathbb{C}^2$ has the following picture:

$$\begin{tikzcd} v_0 \bullet \arrow[loop right, "H"] \arrow[d, bend left, "F"] \\ v_1 \bullet \arrow[loop right, "H"] \arrow[u, bend left, "E"] \end{tikzcd}$$

This is the infinitesimal verion of the action $\mathrm{SL}(2) \curvearrowright \mathbb{C}^2$ given by

$$\begin{aligned} (\exp(-te)) \begin{bmatrix} x \\ y \end{bmatrix} &= \begin{bmatrix} 1 & -t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} x - ty \\ y \end{bmatrix} & (5.2.0.2) \\ \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \exp(-te) &= \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} -y \\ 0 \end{bmatrix} \right\} & (5.2.0.3) \\ \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \exp(-tf) &= \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} 0 \\ -x \end{bmatrix} \right\} & (5.2.0.4) \end{aligned}$$

$\Diamond$

5.2.0.5 Example Since $\mathrm{SL}(2) \curvearrowright \mathbb{C}^2$, it acts also on the space of functions on $\mathbb{C}^2$; by the previous calculations, we see that the action is:

$$e = -y\partial_x, \quad f = -x\partial_y, \quad h = -x\partial_x + y\partial_y$$

These operations are homogenous — they preserve the total degree of any polynomial — and so the symmetric tensor product $\mathcal{S}^n(\mathbb{C}^2) = \{$homogeneous polynomials of degree n in x and $y\}$ is a submodule of $\mathrm{SL}(2) \curvearrowright \{$functions$\}$. Let $v_i \stackrel{\text{def}}{=} \binom{n}{i} x^i y^{n-i}$ be a basis vector in $\mathcal{S}^n(\mathbb{C}x \oplus \mathbb{C}y)$. Then the action $\mathrm{SL}(2) \curvearrowright \mathcal{S}^2(\mathbb{C}^2)$ has the following picture:

$$\begin{tikzcd} y^n = v_0 \bullet \arrow[loop right, "h=n"] \arrow[d, bend left, "f=1"] \\ nxy^{n-1} = v_1 \bullet \arrow[loop right, "h=n-2"] \arrow[u, bend left, "n=e"] \arrow[d, bend left, "f=2"] \\ \binom{n}{2}x^2y^{n-2} = v_2 \bullet \arrow[loop right, "h=n-4"] \arrow[u, bend left, "n-1=e"] \arrow[d, bend left] \\ \vdots \arrow[loop right] \arrow[u, bend left] \arrow[d, bend left] \\ \binom{n}{n-1}x^{n-1}y = v_{n-1} \bullet \arrow[loop right, "h=2-n"] \arrow[u, bend left] \arrow[d, bend left, "f=n"] \\ x^n = v_n \bullet \arrow[loop right, "h=-n"] \arrow[u, bend left, "1=e"] \end{tikzcd} \tag{5.2.0.6}$$

Let us call this module V_n. Then V_n is irreducible, because applying e enough times to any non-zero element results in a multiple of v_0, and v_0 generates the module. $\Diamond$

5.2.0.7 Proposition *Let V be any $(n+1)$-dimensional irreducible module over $\mathfrak{sl}(2)$. Then $V \cong V_n$.*

Proof Suppose that $v \in V$ is an eigenvector of h, so that $hv = \lambda v$. Then $hev = [h,e]v + ehv = 2ev + \lambda ev$, so ev is an h-eigenvector with eigenvalue $\lambda + 2$. Similarly, fv is an h-eigenvector with eigenvalue $\lambda - 2$. So the space spanned by h-eigenvectors of V is a submodule of V; by the irreducibility of V, and using the fact that h has at least one eigenvector, this submodule must be the whole of V, and so h acts diagonally.

By finite-dimensionality, there is an eigenvector v_0 of h with the highest eigenvalue, and so $ev_0 = 0$. By Theorem 3.2.2.1, $\{f^k e^l h^m\}$ spans $\mathcal{U}\mathfrak{sl}(2)$, and so $\{v_i \overset{\text{def}}{=} f^i v_0/i!\}$ is a basis of V (by irreducibility, V is generated by v_0). In particular, $v_n = f^n v_0/n!$, the $(n+1)$st member of the basis, has $fv_n = 0$, since V is $(n+1)$-dimensional.

We compute the action of e by induction, using the fact that $hv_k = (\lambda_0 - 2k)v_k$:

$$\begin{aligned}
ev_0 &= 0 && (5.2.0.8)\\
ev_1 &= efv_0 = [e,f]v_0 + fev_0 = hv_0 = \lambda_0 v_0 && (5.2.0.9)\\
ev_2 &= efv_1/2 = [e,f]v_1/2 + fev_1/2 = hv_1/2 + f\lambda_0 v_0/2\\
&= (\lambda_0 - 2)v_1/2 + \lambda_0 v_1/2 = (\lambda_0 - 1)v_1 && (5.2.0.10)\\
&\dots\\
ev_k &= efv_{k-1}/k = hv_{k-1}/k + fev_{k-1}/k = (\lambda_0 - 2k + 2)v_{k-1}/k + (\lambda_0 - k + 2)fv_{k-2}/k\\
&= \big((\lambda_0 - 2k + 2)/k + (k-1)(\lambda_0 - k + 2)/k\big)v_{k-1} = (\lambda_0 - k + 1)v_{k-1} && (5.2.0.11)
\end{aligned}$$

But $fv_n = 0$, and so:

$$\begin{aligned}
0 = efv_n &= [e,f]v_n + fev_n = hv_n + (\lambda_0 - n + 1)fv_{n-1}\\
&= (\lambda_0 - 2n)v_n + (\lambda_0 - n + 1)nv_n = \big((n+1)\lambda_0 - (n+1)n\big)v_n && (5.2.0.12)
\end{aligned}$$

Thus $\lambda_0 = n$ and V is isomorphic to V_n defined in equation (5.2.0.6). $\square$

5.3 Cartan subalgebras

5.3.1 Definition and existence

5.3.1.1 Lemma *Let $\mathfrak{h}$ be a nilpotent Lie algebra over a field $\mathbb{K}$, and $\mathfrak{h} \curvearrowright V$ a finite-dimensional representation. For $h \in \mathfrak{h}$ and $\lambda \in \mathbb{C}$, define $V_{\lambda,h} = \{v \in V \text{ s.t. } \exists n \text{ s.t. } (h-\lambda)^n v = 0\}$. Then $V_{\lambda,h}$ is an $\mathfrak{h}$-submodule of V.*

Proof Let $\mathrm{ad} : \mathfrak{h} \curvearrowright \mathfrak{h}$ be the adjoint action; since $\mathfrak{h}$ is nilpotent, $\mathrm{ad}\, h \in \mathrm{End}(\mathfrak{h})$ is a nilpotent endomorphism. Define $\mathfrak{h}_{(m)} \overset{\text{def}}{=} \ker\big((\mathrm{ad}\, h)^m\big)$; then $\mathfrak{h}_{(m)} = \mathfrak{h}$ for m large enough. We will show that $\mathfrak{h}_{(m)} V_{\lambda,h} \subseteq V_{\lambda,h}$ by induction on m; when $m = 0$, $\mathfrak{h}_{(0)} = 0$ and the statement is trivial.

Let $y \in \mathfrak{h}_{(m)}$, whence $[h, y] \in \mathfrak{h}_{(m-1)}$, and let $v \in V_{\lambda,h}$. Then $(h-\lambda)^n v = 0$ for n large enough, and so

$$(h-\lambda)^n yv = y(h-\lambda)^n v + [(h-\lambda)^n, y]v \tag{5.3.1.2}$$
$$= 0 + [(h-\lambda)^n, y]v \tag{5.3.1.3}$$
$$= \sum_{k+1=n-1} (h-\lambda)^k [h, y](h-\lambda)^l v \tag{5.3.1.4}$$

since $[\lambda, y] = 0$. By increasing n, we can assure that for each term in the sum at least one of the following happens: l is large enough that $(h-\lambda)^l v = 0$, or k is large enough that $(h-\lambda)^k V_{\lambda,h} = 0$. The large-$l$ terms vanish immediately; the large-k terms vanish upon realizing that $(h-\lambda)V_{h,\lambda} \subseteq V_{h,\lambda}$ by definition and $[h, y]V_{h,\lambda} \subseteq \mathfrak{h}_{(m-1)}V_{h,\lambda} \subseteq V_{h,\lambda}$ by induction on m. □

5.3.1.5 Corollary *Let $\mathfrak{h}$ be a nilpotent Lie algebra over $\mathbb{K}$, $\mathfrak{h} \curvearrowright V$ a finite-dimensional representation, and $\lambda : \mathfrak{h} \to \mathbb{K}$ a linear map. Then $V_\lambda \overset{\text{def}}{=} \bigcap_{h\in\mathfrak{h}} V_{\lambda(h),h}$ is an $\mathfrak{h}$-submodule of V.* □

5.3.1.6 Proposition *Let $\mathfrak{h}$ be a finite-dimensional nilpotent Lie algebra over an algebraically closed field $\mathbb{K}$ of characteristic 0, and V a finite-dimensional $\mathfrak{h}$-module. For each $\lambda \in \mathfrak{h}^*$, define V_λ as in Corollary 5.3.1.5. Then $V = \bigoplus_{\lambda\in\mathfrak{h}^*} V_\lambda$.*

Proof Let $h_1, \dots, h_k \in \mathfrak{h}$, and let $H_k \subseteq \mathfrak{h}$ be the linear span of the h_i. Let $W \overset{\text{def}}{=} \bigcap_{i=1}^k V_{\lambda(h_i),h_i}$. It follows from Theorem 4.2.3.2 that $W = \bigcap_{h\in H} V_{\lambda(h),h}$, since we can choose a basis of V in which $\mathfrak{h} \curvearrowright V$ by upper-triangular matrices.

We have seen already that W is a submodule of V. Let $h_{k+1} \notin H_k$; then we can decompose W into generalized eigenspaces of h_{k+1}. We proceed by induction on k until we have a basis of $\mathfrak{h}$. □

5.3.1.7 Definition *For $\lambda \in \mathfrak{h}^*$, the space V_λ in Corollary 5.3.1.5 is a* weight space *of V, and $V = \bigoplus_{\lambda\in\mathfrak{h}^*} V_\lambda$ the* weight space decomposition*.*

5.3.1.8 Lemma *Let $\mathfrak{h}$ be a finite-dimensional nilpotent Lie algebra over an algebraically closed field of characteristic 0, and let V and W be two finite-dimensional $\mathfrak{h}$ modules. Then the weight spaces of $V \otimes W$ are given by $(V \otimes W)_\lambda = \bigoplus_{\alpha+\beta=\lambda} V_\alpha \otimes W_\beta$.*

Proof $h(v \otimes w) = hv \otimes w + v \otimes hw$. □

5.3.1.9 Corollary *Let $\mathfrak{g}$ be a finite-dimensional Lie algebra over an algebraically closed field of characteristic 0, and $\mathfrak{h} \subseteq \mathfrak{g}$ a nilpotent subalgebra. Then the weight spaces of* $\mathrm{ad} : \mathfrak{h} \curvearrowright \mathfrak{g}$ *satisfy $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}$.* □

5.3.1.10 Proposition *Let $\mathfrak{g}$ be a finite-dimensional Lie algebra over an algebraically closed field of characteristic 0, and let $\mathfrak{h} \subseteq \mathfrak{g}$ be a nilpotent subalgebra. The following are equivalent:*

1. *$\mathfrak{h} = N(\mathfrak{h}) \overset{\text{def}}{=} \{x \in \mathfrak{g} \text{ s.t. } [x, \mathfrak{h}] \subseteq \mathfrak{h}\}$, the* normalizer *of $\mathfrak{h}$ in $\mathfrak{g}$.*

2. *$\mathfrak{h} = \mathfrak{g}_0$ is the 0-weight space of* $\mathrm{ad} : \mathfrak{h} \curvearrowright \mathfrak{g}$.

Proof Define $N^{(i)} \stackrel{\text{def}}{=} \{x \in \mathfrak{g} \text{ s.t. } (\operatorname{ad}\mathfrak{h})^i x \subseteq \mathfrak{h}\}$. Then $N^{(0)} = \mathfrak{h}$ and $N^{(1)} = N(\mathfrak{h})$, and $N^{(i)} \subseteq N^{(i+1)}$. By finite-dimensionality, the sequence $N^{(0)} \subseteq N^{(1)} \subseteq \dots$ must eventually stabilize. By definition $\bigcup N^{(i)} = \mathfrak{g}_0$, so 2. implies 1. But $N^{(i+1)} = N(N^{(i)})$, and so 1. implies 2. $\square$

5.3.1.11 Definition *A subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ satisfying the equivalent conditions of Proposition 5.3.1.10 is a* Cartan subalgebra *of $\mathfrak{g}$.*

5.3.1.12 Theorem (Existence of a Cartan Subalgebra)
Every finite-dimensional Lie algebra over an algebraically closed field of characteristic 0 has a Cartan subalgebra.

Before we prove this theorem, we will need some definitions and lemmas.

5.3.1.13 Definition *Let $\mathbb{K}$ be a field; we say that $X \subseteq \mathbb{K}^n$ is* Zariski closed *if $X = \{x \in \mathbb{K}^n \text{ s.t. } p_i(x) = 0\,\forall i\}$ for some possibly infinite set $\{p_i\}$ of polynomials in $\mathbb{K}[x_1, \dots, x_n]$. A subset $X \subseteq \mathbb{K}^n$ is* Zariski open *if $\mathbb{K}^n \smallsetminus X$ is Zariski closed.*

5.3.1.14 Lemma *If $\mathbb{K}$ is infinite and $U, V \subseteq \mathbb{K}^n$ are two non-empty Zariski open subsets, then $U \cap V$ is non-empty.*

Proof Let $\bar{U} \stackrel{\text{def}}{=} \mathbb{K}^n \smallsetminus U$ and similarly for $\bar{V}$. Let $u \in U$ and $v \in V$. If $u = v$ we're done, and otherwise consider the line $L \subseteq \mathbb{K}^n$ passing through u and v, parameterized $\mathbb{K} \stackrel{\sim}{\to} L$ by $t \mapsto tu + (1-t)v$. Then $L \cap \bar{U}$ and $L \cap \bar{V}$ are finite, as their preimages under $\mathbb{K} \to L$ are loci of polynomials. Since $\mathbb{K}$ is infinite, L contains infinitely many points in $U \cap V$. $\square$

5.3.1.15 Lemma / Definition *Let $\mathfrak{g}$ be a finite-dimensional Lie algebra over an algebraically closed field of characteristic 0. An element $x \in \mathfrak{g}$ is* regular *if $\mathfrak{g}_{0,x}$ has minimal dimension. If x is regular, then $\mathfrak{g}_{0,x}$ is a nilpotent subalgebra of $\mathfrak{g}$.*

Proof We will write $\mathfrak{h}$ for $\mathfrak{g}_{0,x}$. That $\mathfrak{h}$ is a subalgebra follows from Corollary 5.3.1.9. Suppose that $\mathfrak{h}$ is not nilpotent, and let $U \stackrel{\text{def}}{=} \{h \in \mathfrak{h} \text{ s.t. } \operatorname{ad} h|_{\mathfrak{h}} \text{ if not nilpotent}\} \neq 0$. Then $U = \{h \in \mathfrak{h} \text{ s.t. } (\operatorname{ad} h|_{\mathfrak{h}})^d \neq 0\}$ is a Zariski-open subset of $\mathfrak{h}$. Moreover, $V \stackrel{\text{def}}{=} \{h \in \mathfrak{h} \text{ s.t. } h \text{ acts invertibly on } \mathfrak{g}/\mathfrak{h}\}$ is also a non-empty Zariski-open subset of $\mathfrak{h}$, where V is the quotient of $\mathfrak{h}$-modules; it is non-empty because $x \in V$. By Lemma 5.3.1.14 (recall that any algebraically closed field is infinite), there exists $y \in U \cap V$. Then $\operatorname{ad} y$ preserves $\mathfrak{g}_{\alpha,x}$ for every α, as $y \in \mathfrak{h} = \mathfrak{g}_{0,x}$, and y acts invertibly on every $\mathfrak{g}_{\alpha,x}$ for $\alpha \neq 0$. Then$\mathfrak{g}_{0,y} \subseteq \mathfrak{g}_{0,x} = \mathfrak{h}$, but $y \in U$ and so $\mathfrak{g}_{0,y} \neq \mathfrak{h}$. This contradicts the minimality of $\mathfrak{h}$. $\square$

Proof (of Theorem 5.3.1.12) We let $\mathfrak{g}$, $x \in \mathfrak{g}$, and $\mathfrak{h} = \mathfrak{g}_{0,x}$ be as in Lemma/Definition 5.3.1.15. Then $\mathfrak{h} \subseteq \mathfrak{g}_{0,\mathfrak{h}}$ because $\mathfrak{h}$ is nilpotent, and $\mathfrak{g}_{0,\mathfrak{h}} \subseteq \mathfrak{g}_{0,x} = \mathfrak{h}$ because $x \in \mathfrak{h}$. Thus $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$. $\square$

We mention one more fact about the Zariski topology:

5.3.1.16 Lemma *Let U be a Zariski open set over $\mathbb{C}$. Then U is path connected.*

Proof Let $u, v \in U$ and construct the line L as in the proof of Lemma 5.3.1.14. Then $L \cap U$ is isomorphic to $\mathbb{C} \smallsetminus \{\text{finite}\}$, and therefore is path connected. $\square$

5.3.1.17 Proposition *Let $\mathfrak{g}$ be a finite-dimensional Lie algebra over $\mathbb{C}$. Then all Cartan subalgebras of $\mathfrak{g}$ are conjugate by automophisms of $\mathfrak{g}$.*

Proof Consider $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$. Then $\mathrm{ad}\,\mathfrak{g} \subseteq \mathfrak{gl}(\mathfrak{g})$ is a Lie subalgebra, and so corresponds to a connected Lie subgroup $\mathrm{Int}\,\mathfrak{g} \subseteq \mathrm{GL}(\mathfrak{g})$ generated by $\exp(\mathrm{ad}\,\mathfrak{g})$. Since $\mathfrak{g} \curvearrowright \mathfrak{g}$ be derivations, $\exp(\mathrm{ad}\,\mathfrak{g}) \curvearrowright \mathfrak{g}$ by automorphisms, and so $\mathrm{Int}\,\mathfrak{g} \subseteq \mathrm{Aut}\,\mathfrak{g}$.

Let $\mathfrak{h} \subseteq \mathfrak{g}$ be a Cartan subalgebra, and $\mathfrak{g} = \bigoplus \mathfrak{g}_{\alpha,\mathfrak{h}}$ the corresponding weight-space decomposition. Since $\mathfrak{g}$ is finite-dimensional, the set

$$R_{\mathfrak{h}} \overset{\text{def}}{=} \{h \in \mathfrak{h} \text{ s.t. } \alpha(h) \neq 0 \text{ if } \alpha \neq 0 \text{ and } \mathfrak{g}_{\alpha,\mathfrak{h}} \neq 0\} = \{h \in \mathfrak{h} \text{ s.t. } \mathfrak{g}_{0,h} = \mathfrak{h}\}$$

is non-empty and open, since we can take α to range over a finite set (by finite-dimensionality).

Let $\sigma : \mathrm{Int}\,\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ be the canonical action, and consider the restriction to $\sigma : \mathrm{Int}\,\mathfrak{g} \times R_{\mathfrak{h}} \to \mathfrak{g}$. Pick $y \in R_{\mathfrak{h}}$ and let $e \in \mathrm{Int}\,\mathfrak{g}$ be the identity element. We compute the image of the infinitesimal action $d\sigma\left(\mathrm{T}_{(e,y)}(\mathrm{Int}\,\mathfrak{g} \times R_{\mathfrak{h}})\right) \subseteq \mathrm{T}_y\mathfrak{g} \cong \mathfrak{g}$. By construction, varying the first component yields an action by conjugation: $x \mapsto [x, y]$. Thus the image of $\mathrm{T}_e\,\mathrm{Int}\,\mathfrak{g} \times \{0 \in \mathrm{T}_y R_{\mathfrak{h}}\}$ is $(\mathrm{ad}\,y)(\mathfrak{g})$. Since y acts invertibly, $(\mathrm{ad}\,y)(\mathfrak{g}) \supseteq \bigoplus_{\alpha\neq 0} \mathfrak{g}_{\alpha,\mathfrak{h}}$. By varying the second coordinate (recall that $R_{\mathfrak{h}}$ is open), we see that $d\sigma\left(\mathrm{T}_{(e,y)}(\mathrm{Int}\,\mathfrak{g} \times R_{\mathfrak{h}})\right) \supseteq \mathfrak{h} = \mathfrak{g}_{0,\mathfrak{h}}$ also. Thus $d\sigma\left(\mathrm{T}_{(e,y)}(\mathrm{Int}\,\mathfrak{g} \times R_{\mathfrak{h}})\right) = \mathfrak{g} = \mathrm{T}_y\mathfrak{g}$, and so the image $(\mathrm{Int}\,\mathfrak{g})(R_{\mathfrak{h}})$ contains a neighborhood of y and therefore is open.

For each $y \in \mathfrak{g}$, consider the generalized nullspace $\mathfrak{g}_{0,y}$; the dimension of $\mathfrak{g}_{0,y}$ depends on the characteristic polynomial of y, and the coefficients of the characteristic polynomial depend polynomially on the matrix entries of $\mathrm{ad}\,y$. In particular, $\dim \mathfrak{g}_{0,y} \geq r$ if and only if the last r coefficients of the characteristic polynomial of $\mathrm{ad}\,y$ are 0, and so $\{y \in \mathfrak{g} \text{ s.t. } \dim \mathfrak{g}_{0,y} \geq r\}$ is Zariski closed. Therefore $y \mapsto \dim \mathfrak{g}_{0,y}$ is upper semi-continuous in the Zariski topology. In particular, let r be the minimum value of $\dim \mathfrak{g}_{0,y}$, which exists since $\dim \mathfrak{g}_{0,y}$ takes values in integers. Then $\mathrm{Reg} \overset{\text{def}}{=} \{y \in \mathfrak{g} \text{ s.t. } \dim \mathfrak{g}_{0,y} = r\}$, the set of regular elements, is Zariski open and therefore dense. In particular, Reg intersects $(\mathrm{Int}\,\mathfrak{g})(R_{\mathfrak{h}})$.

But if $y \in (\mathrm{Int}\,\mathfrak{g})(R_{\mathfrak{h}})$ then $\dim \mathfrak{g}_{0,y} = \dim \mathfrak{h}$. Therefore $\dim \mathfrak{h}$ is the minimal value of $\dim \mathfrak{g}_{0,y}$ and in particular $R_{\mathfrak{h}} \subseteq \mathrm{Reg}$. Conversely, $\mathrm{Reg} = \bigcup_{\mathfrak{h}' \text{ a Cartan}} R_{\mathfrak{h}'} = \bigcup_{\mathfrak{h}' \text{ a Cartan}} (\mathrm{Int}\,\mathfrak{g}) R_{\mathfrak{h}'}$.

However, $\mathrm{Int}\,\mathfrak{g}$ is a connected group, $R_{\mathfrak{h}}$ is connected being $\mathbb{C}^n$ minus some hyperplanes, and Reg is connected on account of being Zariski open. But the orbits of $(\mathrm{Int}\,\mathfrak{g})R_{\mathfrak{h}}$ are disjoint, and their union is all of Reg, so Reg must consist of a single orbit.

To review: $\mathfrak{h}$ is Cartan and so contains regular elements of $\mathfrak{g}$, and any other regular element of $\mathfrak{g}$ is in the image under $\mathrm{Int}\,\mathfrak{g}$ of a regular element of $\mathfrak{h}$. Thus every Cartan subalgebra is in $(\mathrm{Int}\,\mathfrak{g})\mathfrak{h}$. $\square$

5.3.2 More on the Jordan decomposition and Schur's lemma

Recall Theorem 4.2.5.1 that every $x \in \mathrm{End}(V)$, where V is a finite-dimensional vector spce over an algebraically closed field, has a unique decomposition $x = x_s + x_n$ where x_s is diagonalizable and x_n is nilpotent. We will strengthen this result in the case when $x \in \mathfrak{g} \to \mathrm{End}(V)$ and $\mathfrak{g}$ is semisimple.

5.3.2.1 Lemma *Let $\mathfrak{g}$ be a finite-dimensional Lie algebra over an algebraically closed field, and* $\operatorname{Der}\mathfrak{g} \subseteq \operatorname{End}\mathfrak{g}$ *the algebra of derivations of $\mathfrak{g}$. If $x \in \operatorname{Der}\mathfrak{g}$, then $x_s, x_n \in \operatorname{Der}\mathfrak{g}$.*

Proof For $x \in \operatorname{Der}\mathfrak{g}$, construct the weight-space decomposition $\mathfrak{g} = \bigoplus_\lambda \mathfrak{g}_{\lambda,x}$ of generalized eigenspaces of x. Since x is a derivation, the weight spaces add: $[\mathfrak{g}_{\mu,x}, \mathfrak{g}_{\nu,x}] \subseteq \mathfrak{g}_{\mu+\nu,x}$. Let $y \in \operatorname{End}\mathfrak{g}$ act as λ on $\mathfrak{g}_{\lambda,x}$; then y is a derivation by the additive property. But y is diagonalizable and commutes with x, and $x - y$ is nilpotent because all its eigenvalues are 0 so $y = x_s$. $\square$

We have an immediate corollary:

5.3.2.2 Lemma / Definition *If $\mathfrak{g}$ is a semisimple finite-dimensional Lie algebra over an algebraically closed field $\mathbb{K}$, then every $x \in \mathfrak{g}$ has a unique* Jordan decomposition $x = x_s + x_n$ *such that $[x_s, x_n] = 0$, $\operatorname{ad} x_s$ is diagonalizable, and $\operatorname{ad} x_n$ is nilpotent.*

Proof If $\mathfrak{g}$ is semisimple then $\operatorname{ad} : \mathfrak{g} \to \operatorname{Der}\mathfrak{g}$ is injective as $Z(\mathfrak{g}) = 0$ and surjective because $\operatorname{Der}\mathfrak{g}/\operatorname{ad}\mathfrak{g} = \operatorname{Ext}^1(\mathfrak{g}, \mathbb{K}) = 0$. $\square$

5.3.2.3 Theorem (Schur's Lemma over an algebraically closed field)
Let U be an algebra over $\mathbb{K}$ an algebraically closed field, and let V be a finite-dimensional (over $\mathbb{K}$) irreducible U-module. Then $\operatorname{End}_U(V) = \mathbb{K}$.

Proof Let $\phi \in \operatorname{End}_U(V)$ and $\lambda \in \mathbb{K}$ an eigenvalue of ϕ. Then $\phi - \lambda$ is singular and hence 0 by Theorem 4.4.3.4. $\square$

5.3.2.4 Proposition *Let $\mathfrak{g}$ be a finite-dimensional semisimple Lie algebra over an algebraically closed field of characteristic 0, and let $\sigma : \mathfrak{g} \curvearrowright V$ be a finite-dimensional $\mathfrak{g}$ module. For $x \in \mathfrak{g}$, write x_s and x_n as in Lemma/Definition 5.3.2.2, and write $\sigma(x)_s$ and $\sigma(x)_n$ for the diagonalizable and nilpotent parts of $\sigma(x) \in \mathfrak{gl}(\mathfrak{g})$ as given by Theorem 4.2.5.1. Then $\sigma(x)_s = \sigma(x_s)$ and $\sigma(x)_n = \sigma(x_n)$.*

Proof We reduce to the case when V is an irreducible $\mathfrak{g}$-module using Theorem 4.4.3.8, and we write $\mathfrak{g} = \prod \mathfrak{g}_i$ a product of simples using Corollary 4.3.0.4. Then $\mathfrak{g}_i \curvearrowright V$ as 0 for every i except one, for which the action $\mathfrak{g}_i \curvearrowright V$ is faithful. We replace $\mathfrak{g}$ by that $\mathfrak{g}_i$, whence $\sigma : \mathfrak{g} \hookrightarrow \mathfrak{gl}(V)$ with $\mathfrak{g}$ simple.

It suffices to show that $\sigma(x)_s \in \sigma(\mathfrak{g})$, since then $\sigma(x_s) = \sigma(s)$ for some $s \in \mathfrak{g}$, $\sigma(x)_n = \sigma(x) - \sigma(s) = \sigma(x - s)$, and s and $n = x - s$ commute, sum to x, and act diagonalizably and nilpotently since the adjoint action $\operatorname{ad} : \mathfrak{g} \curvearrowright \mathfrak{g}$ is a submodule of $\mathfrak{g} \curvearrowright \mathfrak{gl}(V)$, so $s = x_s$ and $n = x_n$.

By semisimplicity, $\mathfrak{g} = \mathfrak{g}' \subseteq \mathfrak{sl}(V)$. By Theorem 5.3.2.3, the centralizer of $\mathfrak{g}$ in $\mathfrak{gl}(V)$ consists of scalars. In characteristic 0, the only scalar in $\mathfrak{sl}(V$ is 0, so the centralizer of $\mathfrak{g}$ in $\mathfrak{sl}(V)$ is 0. Define the normalizer $N(\mathfrak{g}) = \{x \in \mathfrak{sl}(V) \text{ s.t. } [x, \mathfrak{g}] \subseteq \mathfrak{g}\}$; then $N(\mathfrak{g})$ is a Lie subalgebra of $\mathfrak{sl}(V)$ containing $\mathfrak{g}$, and $N(\mathfrak{g})$ acts faithfully on $\mathfrak{g}$ since the centralizer of $\mathfrak{g}$ in $\mathfrak{sl}(V)$ is 0, and this action is by derivations. But all derivations are inner, as in the proof of Lemma/Definition 5.3.2.2, and so $N(\mathfrak{g}) \curvearrowright \mathfrak{g}$ factors through $\mathfrak{g} \curvearrowright \mathfrak{g}$, and hence $N(\mathfrak{g}) = \mathfrak{g}$.

So it suffices to show that $\sigma(x)_s \in N(\mathfrak{g})$ for $x \in \mathfrak{g}$. Since $\sigma(x)_n$ is nilpotent, it's traceless, and hence in $\mathfrak{sl}(V)$; then $\sigma(x)_s \in \mathfrak{sl}(V)$ as well. We construct a generalized eigenspace decomposition of V with respect to $\sigma(x) : V = \bigoplus V_{\lambda,x}$. Then $\sigma(x)_s$ acts on $V_{\lambda,x}$ by the scalar λ. We also construct a generalized eigenspace decomposition $\mathfrak{g} = \bigoplus \mathfrak{g}_{\alpha,x}$ with respect to the adjoint action $\mathrm{ad} : \mathfrak{g} \curvearrowright \mathfrak{g}$. Since $\mathfrak{g} \subseteq \mathfrak{gl}(V)$, we have $\mathfrak{g}_{\alpha,x} = \mathfrak{g} \cap \mathrm{End}_{\mathbb{K}}(V)_\alpha = \bigoplus \mathrm{Hom}_{\mathbb{K}}(V_{\lambda,x}, V_{\lambda+\alpha,x})$, by tracking the eigenvalues of the right and left actions of $\mathfrak{g}$ on V.

Moreover, $\mathrm{ad}(\sigma(x)_s) = \mathrm{ad}(\sigma(x_s))$ because both act by α on $\mathrm{Hom}_{\mathbb{K}}(V_{\lambda,x}, V_{\lambda+\alpha,x})$ and hence on $\mathfrak{g}_\alpha$. Thus $\sigma(x)_s$ fixes $\mathfrak{g}$ since $\sigma(x_s)$ does. Therefore $\sigma(x)_s \in N(\mathfrak{g})$. □

5.3.3 Precise description of Cartan subalgebras

5.3.3.1 Lemma *Let $\mathfrak{g}$ be a semisimple Lie algebra over characteristic 0, $\mathfrak{h} \subseteq \mathfrak{g}$ a nilpotent subalgebra, and $\mathfrak{g} = \bigoplus \mathfrak{g}_\alpha$ the root space decomposition with respect to $\mathfrak{h}$. Then the Killing form β pairs $\mathfrak{g}_\alpha$ with $\mathfrak{g}_{-\alpha}$ nondegenerately, and $\beta(\mathfrak{g}_\alpha, \mathfrak{g}_{\alpha'}) = 0$ if $\alpha + \alpha' \neq 0$.*

Proof Let $x \in \mathfrak{g}_\alpha$ and $y \in \mathfrak{g}_{\alpha'}$. For any $h \in \mathfrak{h}$, $(\mathrm{ad}\, h - \alpha(h))^n x = 0$ for some n. So

$$0 = \beta\big((\mathrm{ad}\, h - \alpha(h))^n\, x, y\big) = \beta\big(x, (-\,\mathrm{ad}\, h - \alpha(h))^n\, y\big)$$

but $(-\,\mathrm{ad}\, h - \alpha(h))^n$ is invertible on $\mathfrak{g}_{\alpha'}$ unless $\alpha' = -\alpha$. Nondegeneracy follows from nondegeneracy of β on all of $\mathfrak{g}$. □

5.3.3.2 Corollary *If $\mathfrak{g}$ is a finite-dimensional semisimple Lie algebra over characteristic 0, and let $\mathfrak{h} \subseteq \mathfrak{g}$ a nilpotent subalgebra, then the largest nilpotency ideal in $\mathfrak{g}_0$ of the action $\mathrm{ad} : \mathfrak{g}_0 \curvearrowright \mathfrak{g}$ is the 0 ideal.*

Proof The Killing form β pairs $\mathfrak{g}_0$ with itself nondegenerately. As β is the trace form of $\mathrm{ad} : \mathfrak{g}_0 \curvearrowright \mathfrak{g}$, and $\mathrm{ad}(\mathfrak{g})$-nilpotent ideal of $\mathfrak{g}_0$ must be in $\ker \beta = 0$. □

5.3.3.3 Proposition *Let $\mathfrak{g}$ be a finite-dimensional semisimple Lie algebra over an algebraically closed field $\mathbb{K}$ of characteristic 0, and let $\mathfrak{h} \subseteq \mathfrak{g}$ be a Cartan subalgebra. Then $\mathfrak{h}$ is abelian and $\mathrm{ad} : \mathfrak{h} \curvearrowright \mathfrak{g}$ is diagonalizable.*

Proof By definition, $\mathfrak{h}$ is nilpotent and hence solvable, and by Theorem 4.2.3.2 we can find a basis of $\mathfrak{g}$ in which $\mathfrak{h} \curvearrowright \mathfrak{g}$ by upper triangular matrices. Thus $\mathfrak{h}' = [\mathfrak{h}, \mathfrak{h}]$ acts by strictly upper triangular matrices and hence nilpotently on $\mathfrak{g}$. But $\mathfrak{h} = \mathfrak{g}_0$, and so $\mathfrak{h}' = 0$ by Corollary 5.3.3.2. This proves that $\mathfrak{h}$ is abelian.

Let $x \in \mathfrak{h}$. Then $\mathrm{ad}\, x_s = (\mathrm{ad}\, x)_s$ acts as $\alpha(x)$ on $\mathfrak{g}_\alpha$, and in particular x_s centralizes $\mathfrak{h}$. So $x_s \in \mathfrak{g}_0 = \mathfrak{h}$ and so $x_n = x - x_s \in \mathfrak{h}$. But if $n \in \mathfrak{h}$ acts nilpotently on $\mathfrak{g}$, then $\mathbb{K}n$ is an ideal of $\mathfrak{h}$, since $\mathfrak{h}$ is abelian, and acts nilpotently on $\mathfrak{g}$, so $\mathbb{K}n = 0$ by Corollary 5.3.3.2. Thus $x_n = 0$ and $x = x_s$. In particular, x acts diagonalizably on $\mathfrak{g}$. To show that $\mathfrak{h}$ acts diagonalizably, we use finite-dimensionality and the classical fact that if n diagonalizable matrices commute, then they can by simultaneously diagonalized. □

5.3.3.4 Corollary *Let $\mathfrak{g}$ be a finite-dimensional semisimple Lie algebra over an algebraically closed field of characteristic 0. Then a subalgebra $\mathfrak{h} \subseteq \mathfrak{g}$ is Cartan if and only if $\mathfrak{h}$ is a maximal diagonalizable abelian subalgebra.*

Proof We first show the maximality of a Cartan subalgebra. Let $\mathfrak{h}$ be a Cartan subalgebra and $\mathfrak{h}_1 \supseteq \mathfrak{h}$ abelian. Then $\mathfrak{h}_1 \subseteq \mathfrak{g}_0 = \mathfrak{h}$ because it normalizes $\mathfrak{h}$.

Conversely, let $\mathfrak{h}$ be a maximal diagonalizable abelian subalgebra of $\mathfrak{g}$, and write $\mathfrak{g} = \bigoplus \mathfrak{g}_\alpha$ the weight space decomposition of $\mathfrak{h} \curvearrowright \mathfrak{g}$. We want to show that $\mathfrak{h} = \mathfrak{g}_0$, the centralizer of $\mathfrak{h}$. Pick $x \in \mathfrak{g}_0$; then $x_s, x_n \in \mathfrak{g}_0$, and so $x_s \in \mathfrak{h}$ by maximality. In particular, $\mathfrak{g}_0$ is spanned by $\mathfrak{h}$ and ad-nilpotent elements. Thus $\mathfrak{g}_0$ is nilpotent by Theorem 4.2.2.2 and therefore solvable, so $\mathfrak{g}_0'$ acts nilpotently on $\mathfrak{g}$. But $\mathfrak{g}_0'$ is an ideal of $\mathfrak{g}_0$ that acts nilpotently, so $\mathfrak{g}_0' = 0$, so $\mathfrak{g}_0$ is abelian. Then any one-dimensional subspace of $\mathfrak{g}_0$ is an ideal of $\mathfrak{g}_0$, and a subspace spanned by a nilpotent acts nilpotently, so $\mathfrak{g}_0$ doesn't have any nilpotents. Therefore $\mathfrak{g}_0 = \mathfrak{h}$. □

5.4 Root systems

5.4.1 Motivation and a quick computation

In any semisimple Lie algebra over $\mathbb{C}$ we can choose a Cartan subalgebra, to which we assign combinatorial data. Since all Cartan subalgebras are conjugate, this data, called a *root system*, will not depend on our choice. Conversely, this data will uniquely describe the Lie algebra, based on the representation theory of $\mathfrak{sl}(2)$.

5.4.1.1 Definition *Let $\mathfrak{g}$ be a semisimple Lie algebra over $\mathbb{C}$, and $\mathfrak{h}$ a Cartan subalgebra. The* root space decomposition *of $\mathfrak{g}$ is the weight decomposition $\mathfrak{g} = \bigoplus_\alpha \mathfrak{g}_\alpha$ of* ad $: \mathfrak{h} \curvearrowright \mathfrak{g}$*; each $\mathfrak{g}_\alpha$ is a* root space*, and the set of weights $\alpha \in \mathfrak{h}^*$ that appear in the root space decomposition comprise the* roots *of $\mathfrak{g}$. By Proposition 5.3.1.17 the structure of the set of roots depends up to isomorphism only on $\mathfrak{g}$.*

5.4.1.2 Lemma / Definition *Let $\mathfrak{g}$ be a semisimple Lie algebra over $\mathbb{C}$ with Killing form β, $\mathfrak{h}$ a Cartan subalgebra, and $x_\alpha \in \mathfrak{g}_\alpha$ for $\alpha \neq 0$. To x_α we can associate $y_\alpha \in \mathfrak{g}_{-\alpha}$ with $\beta(x_\alpha, y_\alpha) = -1$ and to the root α we associate a* coroot *h_α with $\beta(h_\alpha, -) = \alpha$. Then $\{x_\alpha, y_\alpha, h_\alpha\}$ span a subalgebra $\mathfrak{sl}(2)_\alpha$ of $\mathfrak{g}$ isomorphic to $\mathfrak{sl}(2)$.*

Proof That h_α and y_α are well-defined follows from the nondegeneracy of β. For any $h \in \mathfrak{h}$, $x \in \mathfrak{g}_\alpha$, and $y \in \mathfrak{g}_{-\alpha}$, we have

$$\beta(h, [x, y]) = \beta([x, h], y) \tag{5.4.1.3}$$
$$= -\alpha(h)\, \beta(x, y) \tag{5.4.1.4}$$

Thus $[x, y] = -\beta(x, y) h_\alpha$. Moreover, since $x_\alpha \in \mathfrak{g}_\alpha$, $[h_\alpha, x_\alpha] = \alpha(h_\alpha) x_\alpha$, and since $y_\alpha \in \mathfrak{g}_{-\alpha}$, $[h_\alpha, y_\alpha] = -\alpha(h_\alpha) y_\alpha$.

Thus $x_\alpha, y_\alpha, h_\alpha$ span a three-dimensional Lie subalgebra of $\mathfrak{g}$ isomorphic to either $\mathfrak{sl}(2)$ or the Heisenberg algebra. But in every finite-dimensional representation the Heisenberg algebra acts nilpotently, whereas $\mathrm{ad}(h_\alpha) \in \mathrm{End}(\mathfrak{g})$ is diagonalizable. Therefore this subalgebra is isomorphic to $\mathfrak{sl}(2)$, and $\alpha(h_\alpha) \neq 0$. □

5.4.1.5 Corollary *Let α be a root of $\mathfrak{g}$. Then $\pm\alpha$ are the only non-zero roots of $\mathfrak{g}$ in $\mathbb{C}\alpha$, and $\dim \mathfrak{g}_\alpha = 1$. In particular, $\mathfrak{sl}(2)_\alpha = \mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha} \oplus \mathbb{C} h_\alpha$.*

Proof We consider $\mathfrak{j}_\alpha \stackrel{\text{def}}{=} \bigoplus_{\alpha' \in \mathbb{C}\alpha \smallsetminus \{0\}} \mathfrak{g}_{\alpha'} \oplus \mathbb{C}h_\alpha$; it is a subalgebra of $\mathfrak{g}$ and an $\mathfrak{sl}(2)_\alpha$-submodule, since $[\mathfrak{g}_\alpha, \mathfrak{g}_{\alpha'}] \subseteq \mathfrak{g}_{\alpha+\alpha'}$, and $h_{c\alpha} = ch_\alpha$ for $c \in \mathbb{C}$. Let $\alpha' \in \mathbb{C}\alpha \smallsetminus \{0\}$ be a root; as a weight of the $\mathfrak{sl}(2)_\alpha$ representation, we see that $\alpha' \in \mathbb{Z}\alpha/2$. If any half-integer multiple of α actually appears, then $\alpha/2$ appears, and by switching α to $\alpha/2$ if necessary we can assure that $\mathfrak{j}_\alpha$ contains only representations V_{2m}. But each V_{2m} has contributes a basis vector in weight 0, and the only part of $\mathfrak{j}_\alpha$ in weight 0 is $\mathbb{C}h_\alpha$. Therefore $\mathfrak{j}_\alpha$ is irreducible as an $\mathfrak{sl}(2)_\alpha$ module, contains $\mathfrak{sl}(2)_\alpha$, and so equals $\mathfrak{sl}(2)_\alpha$. □

5.4.1.6 Corollary *The roots α span $\mathfrak{h}^*$, and the coroots h_α span $\mathfrak{h}$.*

Proof We let α range over the non-zero roots. Then

$$\bigcap_{\alpha \neq 0} \ker \alpha = Z(\mathfrak{g}) = 0 \tag{5.4.1.7}$$

$$\sum_{\alpha \neq 0} \mathbb{C}h_\alpha = \mathfrak{g}' \cap \mathfrak{h} = \mathfrak{h} \tag{5.4.1.8}$$

That $\beta^{-1} : \alpha \mapsto h_\alpha$ is a linear isomorphism $\mathfrak{h}^* \to \mathfrak{h}$ completes the proof. □

5.4.1.9 Proposition *Let $\mathfrak{g}$ be a finite-dimensional semisimple Lie algebra over $\mathbb{C}$, and $\mathfrak{h} \subseteq \mathfrak{g}$ a Cartan subalgebra. Let $R \subseteq \mathfrak{h}^*$ be the set of nonzero roots and $R^\vee \subseteq \mathfrak{h}$ the set of nonzero coroots. Then $\alpha \mapsto \alpha^\vee \stackrel{\text{def}}{=} \frac{2h_\alpha}{\alpha(h_\alpha)}$ defines a bijection $\vee : R \to R^\vee$, and the triple $(R, R^\vee, \vee)$ comprise a* root system *in $\mathfrak{h}$.*

We will define the words "root system" in the next section to generalize the data already computed.

5.4.2 The definition

5.4.2.1 Definition *A* root system *is a complex vector space $\mathfrak{h}$, a finite subset $R \subseteq \mathfrak{h}^*$, a subset $R^\vee \subseteq \mathfrak{h}$, a bijection $\vee : R \to R^\vee$, subject to*

RS1 $\langle \alpha, \beta^\vee \rangle \in \mathbb{Z}$

RS2 *$R = -R$ and $R^\vee = -R^\vee$, with $(-\alpha)^\vee = -(\alpha^\vee)$*

RS3 $\langle \alpha, \alpha^\vee \rangle = 2$

RS4 *If $\alpha, \beta \in R$ are not proportional, then $(\beta + \mathbb{C}\alpha) \cap R$ consists of a "string":*

$$\left\langle (\beta + \mathbb{C}\alpha) \cap R, \alpha^\vee \right\rangle = \{m, m-2, \dots, -m+2, -m\}$$

Nondeg *R spans $\mathfrak{h}^*$ and $R^\vee$ spans $\mathfrak{h}$*

Reduced *$\mathbb{C}\alpha \cap R = \{\pm\alpha\}$ for $\alpha \in R$.*

Two root systems are isomorphic *if there is a linear isomorphism of the underlying vector spaces, inducing an isomorphism on dual spaces, that carries each root system to the other. The* rank *of a root system is the dimension of $\mathfrak{h}$.*

5.4.2.2 Definition *Given a root system $(R, R^\vee)$ on a vector space $\mathfrak{h}$, the* Weyl group $W \subseteq \mathrm{GL}(\mathfrak{h}^*)$ *is the group generated by the reflections $s_\alpha : \lambda \mapsto \lambda - \langle \lambda, \alpha^\vee \rangle \alpha$ as α ranges over R.*

5.4.2.3 Proposition

1. *It follows from* **RS3** *that $s_\alpha^2 = e \in W$ for each root α.*
2. *It follows from* **RS4** *that $WR = R$. Thus W is finite. Moreover, W preserves $\mathfrak{h}^*_\mathbb{R}$, the $\mathbb{R}$-span of R.*
3. *The W-average of any positive-definite inner product on $\mathfrak{h}^*_\mathbb{R}$ is a W-invariant positive-definite inner product. Let $(,)$ be a W-invariant positive-definite inner product. Then s_α is orthogonal with respect to $(,)$, and so $s_\alpha : \lambda \mapsto \lambda - \frac{2(\lambda,\alpha)}{(\alpha,\alpha)}\alpha$. This inner product establishes an isomorphism $\mathfrak{h} \xrightarrow{\sim} \mathfrak{h}^*$, under which $\alpha^\vee \mapsto 2\alpha/(\alpha,\alpha)$.*
4. *Therefore* **Reduced** *holds with R replaced by $R^\vee$ if it holds at all.*
5. *Let W act on $\mathfrak{h}$ dual to its action on $\mathfrak{h}^*$. Then $w(\alpha^\vee) = (w\alpha)^\vee$ for $w \in W$ and $\alpha \in R$. Thus $s_{w\alpha} = w s_\alpha w^{-1}$.*
6. *If $V \subseteq \mathfrak{h}^*$ is spanned by any subset of R, then $R \cap V$ and its image under $\vee$ form another nondegenerate root system.*
7. *Two root systems with the same Weyl group and lattices are related by an isomorphism.* $\square$

5.4.2.4 Definition *Let R be a root system in $\mathfrak{h}^*$. Define the* weight lattice *to be $P \overset{\text{def}}{=} \{\lambda \in \mathfrak{h}^* \text{ s.t. } \langle \lambda, \alpha^\vee \rangle \in \mathbb{Z} \forall \alpha^\vee \in R^\vee\}$ and the* root lattice *Q to be the $\mathbb{Z}$-span of R. Then* **RS1** *implies that $R \subseteq Q \subseteq P \subseteq \mathfrak{h}^*$; by* **Nondeg**, *both P and Q are of full rank and so the index $P : Q$ is finite. We define the* coweight lattice *to be $P^\vee$ and the* coroot lattice *to be $Q^\vee$.*

5.4.3 Classification of rank-two root systems

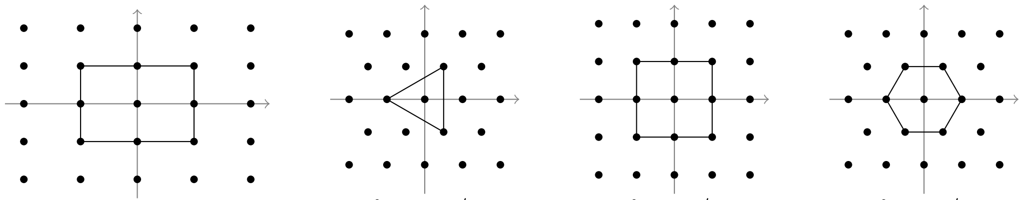
By **Reduced**, there is a unique rank-one root system up to isomorphism, the root system of $\mathfrak{sl}(2)$.

Let R be a rank-two root system; then its Weyl group W is a finite subgroup $W \subseteq \mathrm{GL}(2,\mathbb{R})$ generated by reflections. The only finite subgroups of $\mathrm{GL}(2,\mathbb{R})$ are the cyclic and dihedral groups; only the dihedral groups are generated by reflections, and so $W \cong D_{2m}$ for some m. Moreover, W preserves the root lattice Q.

5.4.3.1 Lemma *The only dihedral groups that preserve a lattice are D_4, D_6, D_8, and D_{12}.*

Proof Let r_θ be a rotation by θ. Its eigenvalues are $e^{\pm i\theta}$, and so $\mathrm{tr}(r_\theta) = 2\cos\theta$. If r_θ preserves a lattice, its trace must be an integer, and so $2\cos\theta \in \{1, 0, -1, -2\}$, as $2\cos\theta = 2$ corresponds to the identity rotation, and $|\cos\theta| \leq 1$. Therefore $\theta \in \{\pi, 2\pi/3, \pi/2, \pi/3\}$, i.e. $\theta = 2\pi/m$ for $m \in \{2, 3, 4, 6\}$, and the only valid dihedral groups are D_{2m} for these values of m. $\square$

5.4.3.2 Corollary *There are four rank-2 root systems, corresponding to the rectangular lattice, the square lattice, and the hexagonal lattice twice:*



$\theta = \pi$ $\quad$ $\theta = 2\pi/3$ $\quad$ $\theta = \pi/2$ $\quad$ $\theta = \pi/3$

□

For each dihedral group, we can pick two reflections α_1, α_2 with a maximally obtuse angle; these generate W and the lattice. On the next page we list the four rank-two root systems with comments on their corresponding Lie groups.

5.4.3.3 Lemma / Definition *The axioms of a finite root system are symmetric under the interchange $R \leftrightarrow R^\vee$. This interchange assigns a* dual *to each root system.*

Proof Only **RS4** is not obviously symmetric. We did not use **RS4** to classify the two-dimensional root systems; we needed only a corollary:

RS4' $W(R) = R$,

which is obviously symmetric. But **RS4** describes only the two-dimensional subspaces of a root system, and every rank-two root system with **RS4'** replacing **RS4** in fact satisfies **RS4**. This suffices to show that **RS4'** implies **RS4** for finite root systems. □

We remark that the statement is false for infinite root systems, and we presented the definition we did to accommodate the infinite case. We will not discuss infinite root systems further.

5.4.4 Positive roots

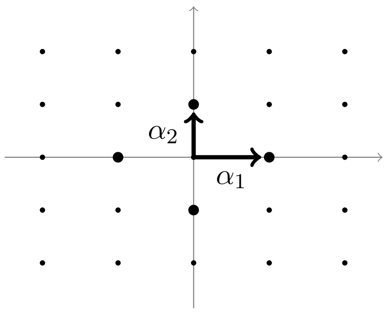
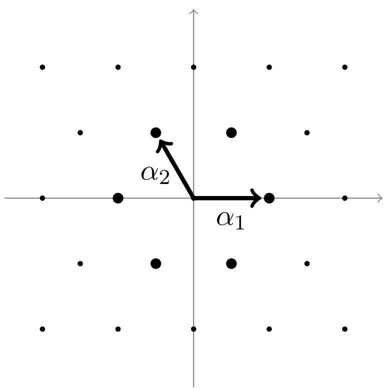
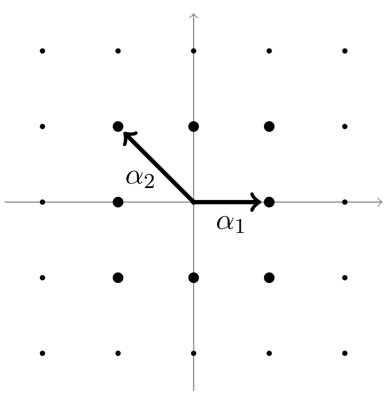
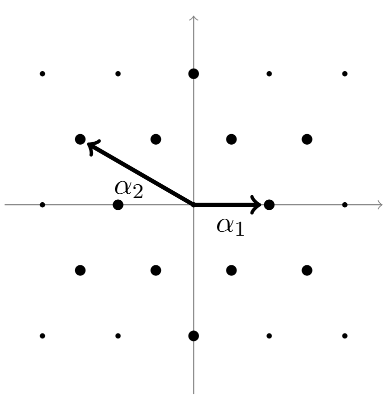
5.4.4.1 Definition *A* positive root system *consists of a (finite) root system $R \subseteq \mathfrak{h}_{\mathbb{R}}^*$ and a vector $v \in \mathfrak{h}_{\mathbb{R}}$ so that $\alpha(v) \neq 0$ for every root $\alpha \in R$. A root $\alpha \in R$ is* positive *if $\alpha(v) > 0$, and* negative *otherwise. Let R_+ be the set of positive roots and R_- the set of negative ones; then $R = R_+ \sqcup R_-$, and by* **RS2**, *$R_+ = -R_-$.*

The $\mathbb{R}_{\geq 0}$-span of R_+ is a cone in $\mathfrak{h}_{\mathbb{R}}^*$, and we let Δ be the set of extremal rays in this cone. Since the root system is finite, extremal rays are generated by roots, and we use **Reduced** *to identify extremal rays with positive roots. Then $\Delta \subseteq R$ is the set of* simple roots.

5.4.4.2 Lemma *If α and β are two simple roots, then $\alpha - \beta$ is not a root. Moreover, $(\alpha, \beta) \leq 0$ for $\alpha \neq \beta$.*

Proof If $\alpha - \beta$ is a positive root, then $\alpha = \beta + (\alpha - \beta)$ is not simple; if $\alpha - \beta$ is negative then β is not simple.

For the second statement, assume that α and β are any two roots with $(\alpha, \beta) > 0$. If $\alpha \neq \beta$, then they cannot be proportional, and we assume without loss of generality that $(\alpha, \alpha) \leq (\beta, \beta)$.

m	Picture	Name	notes
2		$A_1 \times A_1 = D_2$	corresponding to the Lie algbera $\mathfrak{sl}(2) \times \mathfrak{sl}(2) = \mathfrak{so}(4)$
3		A_2	corresponding to $\mathfrak{sl}(3)$ acting on the traceless diagonals
4		$B_2 = C_2$	$\mathfrak{so}(5) = \mathfrak{sp}(4)$. (When we get higher up, the Bs and Cs will separate, and we will have a new sequence of Ds.)
6		G_2	a new simple algebra of dimension 14 = number of roots plus dimension of root space. We will see later that its smallest representation has dimension 7. There are many descriptions of this representation and the corresponding Lie algebra; the seven-dimensional representation comes from the Octonians, a non-associative, non-commutative "field", and G_2 is the automorphism group of the pure-imaginary part of the Octonians.

The rank-2 root systems

Then $s_\beta(\alpha) = \alpha - \frac{2(\alpha,\beta)}{(\beta,\beta)}\beta = \alpha - \beta$, because $2(\alpha,\beta)/(\beta,\beta) = \langle \alpha, \beta^\vee \rangle$ is a positive integer strictly less than 2. Thus $\alpha - \beta$ is a root if $(\alpha,\beta) > 0$. □

5.4.4.3 Lemma *Let $\mathbb{R}^n$ have a positive definite inner product $(,)$, and suppose that $v_1, \dots, v_n \in \mathbb{R}^n$ satisfy $(v_i, v_j) \le 0$ if $i \neq j$, and such that there exists v_0 with $(v_0, v_i) > 0$ for every i. Then $\{v_1, \dots, v_n\}$ is an independent set.*

Proof Suppose that $0 = c_1v_1 + \dots + c_nv_n$. Renumbering as necessary, we assume that $c_1, \dots, c_k \ge 0$, and $c_{k+1}, \dots, c_n \le 0$. Let $v = c_1v_1 + \dots + c_kv_k = |c_{k+1}|v_{k+1} + \dots + |c_n|v_n$. Then $0 \le (v,v) = (\sum_{i=1}^k c_iv_i, \sum_{j=k+1}^n -c_kv_k) = \sum_{i,j} |c_ic_j|(v_i,v_j) \le 0$, which can happen only if $v = 0$. But then $0 = (v, v_0) = \sum_{i=1}^k c_i(v_i, v_0) > 0$ unless all c_i are 0 for $i \le k$. Similarly we must have $c_j = 0$ for $j \ge k+1$, and so $\{v_i\}$ is independent. □

5.4.4.4 Corollary *In any positive root system, the set Δ of simple roots is a basis of $\mathfrak{h}^*$.*

Proof By Lemma 5.4.4.2, Δ satisfies the conditions of Lemma 5.4.4.3 and so is independent. But Δ generates R_+ and hence R, and therefore spans $\mathfrak{h}^*$. □

5.4.4.5 Lemma *Let $\Delta = \{\alpha_1, \dots, \alpha_n\}$ be a set of vectors in $\mathbb{R}^m$ with inner product $(,)$, and assume that α_i are all on one side of a hyperplane: there exists v such that $(\alpha_i, v) > 0\ \forall i$. Let W be the group generated by reflections S_{α_i}. Let R_+ be any subset of $\mathbb{R}_{\ge 0}\Delta \smallsetminus \{0\}$ such that $s_i(R_+ \smallsetminus \{\alpha_i\}) \subseteq R_+$ for each i, and such that the set of heights $\{(\alpha, v)\}_{\alpha \in R_+} \subseteq \mathbb{R}_{\ge 0}$ is well-ordered. Then $R_+ \subseteq W(\Delta)$.*

Proof Let $\beta \in R_+$. We proceed by induction on its height.

There exists i such that $(\alpha_i, \beta) > 0$, because if $(\beta, \alpha_i) \le 0\ \forall i$, then $(\beta,\beta) = 0$ since β is a positive combination of the α_is. Thus $s_i(\beta) = \beta - (\text{positive})\alpha_i$; in particular, $\bigl(v, s_i(\beta)\bigr) < (v, \beta)$.

If $\beta \neq \alpha_i$, then $s_i(\beta) \in R_+$ by hypothesis, so by induction $s_i(\beta) \in W(\Delta)$, and hence $\beta = s_i(s_i(\beta)) \in W(\Delta)$. If $\beta = \alpha_i$, it's already in $W(\Delta)$. □

5.4.4.6 Corollary *Let R be a finite root system, R_+ a choice of positive roots, and Δ the corresponding set of simple roots. Then $R = W(\Delta)$, and the set $\{s_{\alpha_i}\}_{\alpha_i \in \Delta}$ generates W.* □

5.4.4.7 Corollary *Let R be a finite root system, R_+ a choice of positive roots, and Δ the corresponding set of simple roots. Then $R \subseteq \mathbb{Z}\Delta$ and $R_+ \subseteq \mathbb{Z}_{\ge 0}\Delta$.* □

5.4.4.8 Proposition *Let R be a finite root system, and R_+ and R'_+ two choices of positive roots. Then R_+ and R'_+ are W-conjugate.*

Proof Let Δ be the set of simple roots corresponding to R_+. If $\Delta \subseteq R'_+$, then $R_+ \subseteq R'_+$. Then $R_- \subseteq R'_-$ by negating, and $R_+ \supseteq R'_+$ by taking complements, so $R_+ = R'_+$.

Suppose $\alpha_i \in \Delta$ but $\alpha_i \notin R'_+$, and consider the new system of positive roots $s_i(R'_+)$, where $s_i = s_{\alpha_i}$ is the reflection corresponding to α_i. Then $s_i(R'_+) \cap R_+ \supseteq s_i(R'_+ \cap R_+)$, because a system of roots that does not contain α_i does not lose anything under s_i. But $\alpha_i \in R'_-$, so $-\alpha_i \in R'_+$, and so $\alpha_i \in s_i(R'_+)$ and hence in $s_i(R'_+) \cap R_+$. Therefore $\bigl|s_i(R'_+) \cap R_+\bigr| > \bigl|R'_+ \cap R_+\bigr|$.

If $s_i(R'_+) \neq R_+$, then we can find $\alpha_j \in \Delta \smallsetminus s_i(R'_+)$. We repeat the argument, at each step making the set $\bigl|w(R'_+) \cap R_+\bigr|$ strictly bigger, where $w = \cdots s_js_i \in W$. Since R_+ is a finite set, eventually we cannot get any bigger; this can only happen when $\Delta \subseteq w(R'_+)$, and so $R_+ = w(R'_+)$. □

5.5 Cartan matrices and Dynkin diagrams

5.5.1 Definitions

5.5.1.1 Definition *A finite-type* Cartan matrix *of rank n is an $n \times n$ matrix a_{ij} satisfying the following:*

- $a_{ii} = 2$ *and* $a_{ij} \in \mathbb{Z}_{\leq 0}$ *for* $i \neq j$.
- a *is* symmetrizable*: there exists an invertible diagonal matrix d with da symmetric.*
- a *is* positive*: all principle minors of a are positive.*

An isomorphism *between Cartan matrices a_{ij} and b_{ij} is a permutation $\sigma \in S_n$ such that $a_{ij} = b_{\sigma i, \sigma j}$.*

5.5.1.2 Lemma / Definition *Let R be a finite root system, R_+ a system of positive roots, and $\Delta = \{\alpha_1, \dots, \alpha_n\}$ the corresponding simple roots. The* Cartan matrix *of R is the matrix $a_{ij} \overset{\text{def}}{=} \langle \alpha_j, \alpha_i^\vee \rangle = 2(\alpha_i, \alpha_j)/(\alpha_i, \alpha_i)$.*

The Cartan matrix of a root system is a Cartan matrix. It depends (up to isomorphism) only on the root system. Conversely, a root system is determined up to isomorphism by its Cartan matrix.

Proof That the Cartan matrix depends only on the root system follows from Proposition 5.4.4.8. That the Cartan matrix determines the root system follows from Corollary 5.4.4.6.

Given a choice of root system and simple roots, let $d_i \overset{\text{def}}{=} (\alpha_i, \alpha_i)/2$, and let $d_{ij} \overset{\text{def}}{=} d_i \delta_{ij}$ be the diagonal matrix with the d_is on the diagonal. Then d is invertible because $d_i > 0$, and $da = (\alpha_i, \alpha_j)$ is obviously symmetric. Let $I \subseteq \{1, \dots, n\}$; then the $I \times I$ principle minor of da is just $\prod_{i \in I} d_i$ times the corresponding principle minor of a. Since $d_i > 0$ for each i and da is the matrix of a positive-definite symmetric bilinear form, we see that a is positive. $\square$

5.5.2 Classification of finite-type Cartan matrices

We classify (finite-type) Cartan matrices by encoding their information in graph-theoretic form ("Dynkin diagrams") and then classifying (indecomposable) Dynkin diagrams.

5.5.2.1 Definition *Let a be an integer matrix so that every principle 2×2 sub-matrix has the form $\left[\begin{smallmatrix} 2 & -k \\ -l & 2 \end{smallmatrix}\right]$ with $k, l \in \mathbb{Z}_{\geq 0}$ and either both k and l are 0 or one of them is 1. Let us call such a matrix* generalized Cartan*.*

5.5.2.2 Lemma *A Cartan matrix is generalized Cartan. A generalized Cartan matrix is not Cartan if any entry is -4 or less.*

Proof Consider a 2×2 sub-matrix $\left[\begin{smallmatrix} 2 & -k \\ -l & 2 \end{smallmatrix}\right]$. Then if one of k and l is non-zero, the other must also be non-zero by symmetrizability. Moreover, $kl < 4$ by positivity, and so one of k and l must be 1. $\square$

5.5.2.3 Definition *Let a be a rank-n generalized Cartan matrix. Its* diagram *is a graph on n vertices with (labeled, directed) edges determined as follows:*

Let $1 \leq i, j \leq n$, and consider the $\{i, j\} \times \{i, j\}$ submatrix of a. By definition, either k and l are both 0, or one of them is 1 and the other is a positive integer. We do not draw an edge between vertices i and j if $k = l = 0$. We connect i and j with a single undirected edge if $k = l = 1$. For $k = 2, 3$, we draw an arrow with k edges from vertex i to vertex j if the $\{i, j\}$ block is $\left[\begin{smallmatrix} 2 & -1 \\ -k & 2 \end{smallmatrix}\right]$.

5.5.2.4 Definition *A diagram is* Dynkin *if its corresponding generalized Cartan matrix is in fact Cartan.*

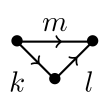
5.5.2.5 Lemma / Definition *The diagram of a generalized Cartan matrix a is disconnected if and only if a is block diagonal, and connected components of the diagram correspond to the blocks of a. A block diagonal matrix a is Cartan if and only if each block is. A connected diagram is* indecomposable*. We write "$\times$" for the disjoint union of Dynkian diagrams.* □

5.5.2.6 Example There is a unique indecomposable rank-1 diagram, and it is Dynkin: $A_1 = \bullet$.

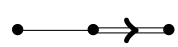
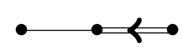
The indecomposable rank-2 Dynkin diagrams are:

$A_2 =$ $B_2 = C_2 =$  $G_2 =$  ♢

5.5.2.7 Lemma / Definition *A* subdiagram *of a diagram is a subset of the vertices, with edges induced from the parent diagram. Subdiagrams of a Dynkin diagram correspond to principle submatrices of the corresponding Cartan matrix. Any subdiagram of a Dynkin diagram is Dynkin.* □

By symmetrizability, if we have a triangle , then the multiplicities must be related: $m = kl$. So k or l is 1, and you can check that the three possibilities all have determinant ≤ 0. Moreover, a triple edge cannot attach to an edge, and two double edges cannot attach, again by positivity. As such, we will never need to discuss the triple-edge again.

5.5.2.8 Example There are three indecomposable rank-3 Dynkin diagrams:

$A_3 =$ $B_3 =$  $C_3 =$  ♢

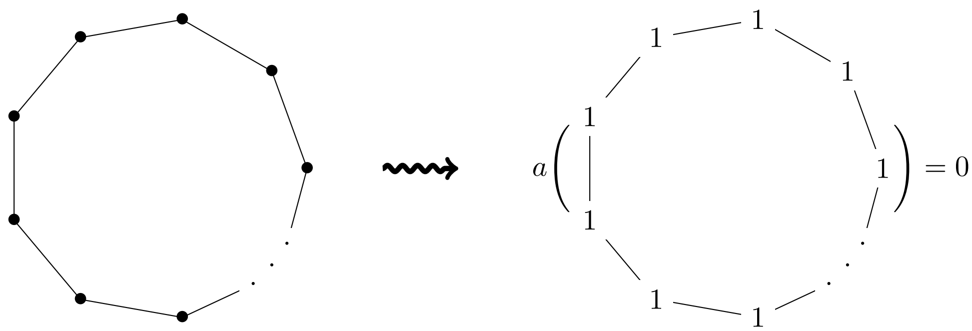
5.5.2.9 Definition *Let a be a generalized Cartan rank-n matrix. We can specify a vector in $\mathbb{R}^n$ by assigning a "weight" to each vertex of the corresponding diagram. The* neighbors *of a vertex are counted with multiplicity: an arrow leaving a vertex contributes only one neighbor to that vertex, but an arrow arriving contributes as many neighbors as the arrow has edges. Naturally, each vertex of a weighted diagram has some number of "weighted neighbors": each neighbor is counted with multiplicity and multiplied by its weight, and these numbers are summed.*

5.5.2.10 Lemma *Let a be a generalized Cartan matrix, and think of a vector $\vec{x}$ as a weighting of the corresponding diagram. With the weighted-neighbor conventions in Definition 5.5.2.9, the* *multiplication $a\vec{x}$ can be achieved by subtracting the number of weighted neighbors of each vertex from twice the weight of that vertex.*

Thus, a generalized Cartan matrix is singular if its corresponding diagram has a weighting such that each vertex has twice as many (weighted) neighbors as its own weight. □

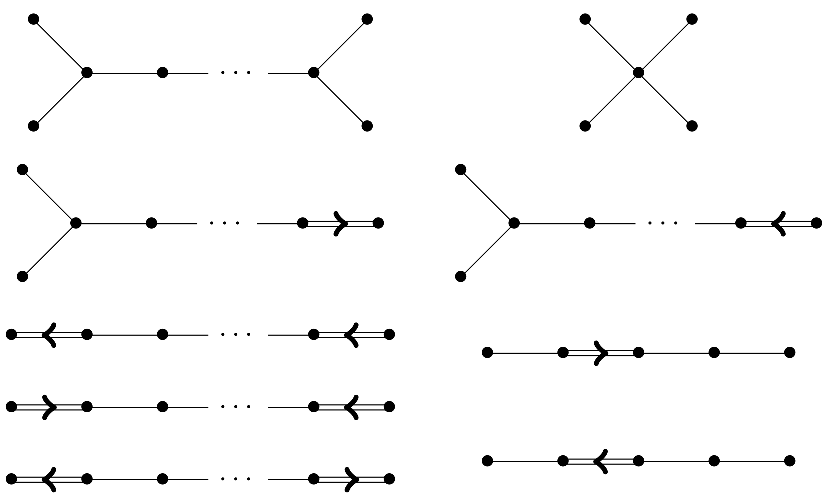
5.5.2.11 Corollary *A ring of single edges, and hence any diagram with a ring as a subdiagram, is not Dynkin.*

Proof We assign weight 1 to each vertex; this shows that the determinant of the ring is 0:



□

5.5.2.12 Corollary *The following diagrams correspond to singular matrices and hence are not Dynkin:*



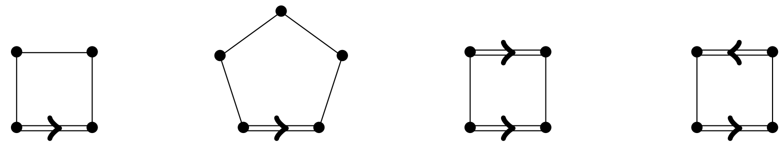
Proof For example, we can show the last two as singular with the following weightings:

$$1 — 2 \Rightarrow 3 — 2 — 1$$

$$2 — 4 \Leftarrow 3 — 2 — 1$$

□

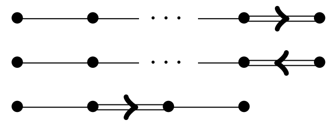
5.5.2.13 Lemma *The following diagrams are not Dynkin:*



det $= -4$ det $= -4$ det $= -8$ *not symmetrizable* □

5.5.2.14 Corollary *The indecomposable Dynkin diagrams with double edges are the following:*

$B_n =$, $C_n =$, $F_4 =$



Proof Any indecomposable Dynkin diagram with a double edge is a chain. The double edge must come at the end of the chain, unless the diagram has rank 4. □

5.5.2.15 Lemma *Consider a Y-shaped indecomposable diagram. Let the lengths of the three arms, including the middle vertex, be k, l, m. Then the diagram is Dynkin if and only if $\frac{1}{k} + \frac{1}{l} + \frac{1}{m} > 1$.*

Proof One can show directly that the determinant of such a diagram is $klm(\frac{1}{k} + \frac{1}{l} + \frac{1}{m} - 1)$. We present null-vectors for the three "Egyptian fraction" decompositions of 1 — triples k, l, m such that $\frac{1}{k} + \frac{1}{l} + \frac{1}{m} = 1$:

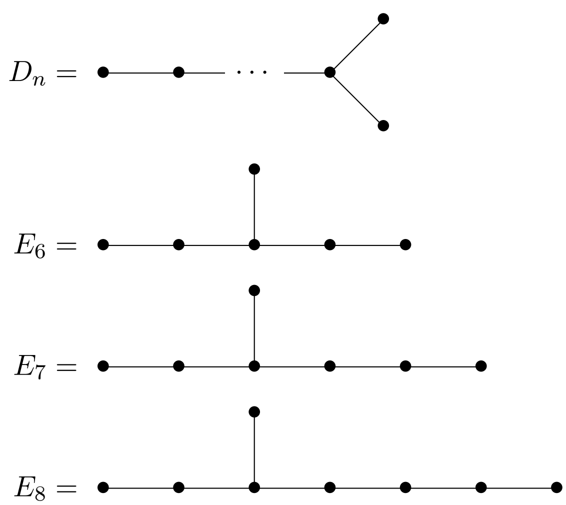
$$\begin{array}{ccccccccc} & & & & 2 & — & 1 & & \\ & & & & | & & & & \\ 1 & — & 2 & — & 3 & — & 2 & — & 1 \end{array}$$

$$\begin{array}{ccccccccc} & & 3 & — & 2 & — & 1 & & \\ & & | & & & & & & \\ 2 & — & 4 & — & 3 & — & 2 & — & 1 \end{array}$$

$$\begin{array}{ccccccccccccccc} & & & & 3 & & & & & & & & & & \\ & & & & | & & & & & & & & & & \\ 2 & — & 4 & — & 6 & — & 5 & — & 4 & — & 3 & — & 2 & — & 1 \end{array}$$

□

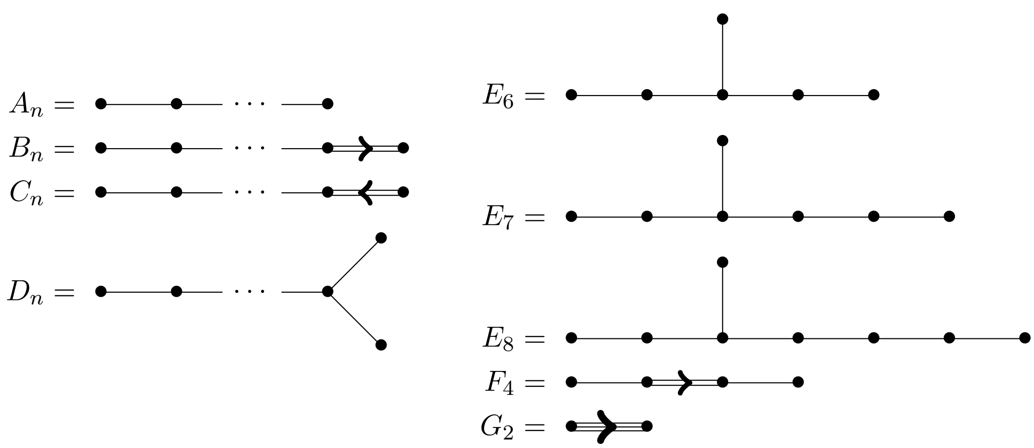
5.5.2.16 Corollary *The indecomposable Dynkin diagrams made entirely of single edges are:*



□

All together, we have proven:

5.5.2.17 Theorem (Classification of indecomposable Dynkin diagrams) *A diagram is Dynkin if and only if it is a disjoint union of indecomposable Dynkin diagrams. The indecomposable Dynkin diagrams comprise four infinite families and five "sporadic" cases:*



□

5.5.2.18 Example We mention the small-rank coindidences. We can continue the E series for smaller n: $E_5 = D_5$, $E_4 = A_4$, and E_3 is sometimes defined as the disjoint union $A_1 \times A_2$ (E_1, E_2 are never defined). The B, C, and D series make sense for $n \geq 2$, whence $B_2 = C_2$ and $D_2 = A_1 \times A_1$ and $D_3 = A_3$. Some diagrams have nontrivial symmetries: for $n \geq 1$, the symmetry group of A_n has order 2, and similarly for D_n for $n \neq 4$. The diagram D_4 has an unexpected symmetry: its symmetry group is S_3, with order 6. The symmetry group of E_6 is order-2. ◊

5.6 From Cartan matrix to Lie algebra

In Theorem 5.5.2.17, we classified indecomposable finite-type Cartan matrices, and therefore all finite-type Cartan matrices. We can present generators and relations showing that each indecomposable Cartan matrix is the Cartan matrix of some simple Lie algebra — indeed, the infinite families A_n, B_n, C_n, and D_n correspond respectively to the classical Lie algebras $\mathfrak{sl}(n, \mathbb{C})$, $\mathfrak{so}(2n+1, \mathbb{C})$, $\mathfrak{sp}(n, \mathbb{C})$, and $\mathfrak{so}(2n, \mathbb{C})$ — and it is straightforward to show that a disjoint union of Cartan matrices corresponds to a direct product of Lie algebras.

In this section, we explain how to construct a semisimple Lie algebra for any finite-type Cartan matrix, and we show that a semisimple Lie algebra is determined by its Cartan matrix. This will complete the proof of the classification of semisimple Lie algebras. Most, but not all, of the construction applies to generalized Cartan matrices; the corresponding Lie algebras are *Kac–Moody*, which are infinite-dimensional versions of semisimple Lie algebras. We will not discuss Kac–Moody algebras here.

5.6.0.1 Lemma / Definition *Let Δ be a rank-n Dynkin diagram with vertices labeled a basis $\{\alpha_1, \dots, \alpha_n\}$ of a vector space $\mathfrak{h}^*$, and let a_{ij} be the corresponding Cartan matrix. Since a_{ij} is nondegenerate, it defines a map $\vee : \mathfrak{h}^* \to \mathfrak{h}$ by $a_{ij} = \langle \alpha_j, \alpha_i^\vee \rangle$. We define $\tilde{\mathfrak{g}} = \tilde{\mathfrak{g}}_\Delta$ to be the Lie algebra generated by $\{e_i, f_i, h_i\}_{i=1}^n$ subject to the relations*

$$[h_i, e_j] = a_{ij} e_j \tag{5.6.0.2}$$

$$[h_i, f_j] = -a_{ij} f_j \tag{5.6.0.3}$$

$$[e_i, f_j] = \delta_{ij} h_i \tag{5.6.0.4}$$

$$[h_i, h_j] = 0 \tag{5.6.0.5}$$

For each i, we write $\mathfrak{sl}(2)_i$ for the subalgebra spanned by $\{e_i, f_i, h_i\}$; clearly $\mathfrak{sl}(2)_i \cong \mathfrak{sl}(2)$.

Let $Q = \mathbb{Z}\Delta$ be the root lattice of Δ. Then the free Lie algebra generated by $\{e_i, f_i, h_i\}_{i=1}^n$ has a natural Q-grading, by $\deg e_i = \alpha_i$, $\deg f_i = -\alpha_i$, and $\deg h_i = 0$; under this grading, the relations are homogeneous, so the grading passes to the quotient $\tilde{\mathfrak{g}}_\Delta$.

Let $\tilde{\mathfrak{h}} \subseteq \tilde{\mathfrak{g}}$ be the subalgebra generated by $\{h_i\}_{i=1}^n$; then it is abelian and spanned by $\{h_i\}_{i=1}^n$. The adjoint action $\mathrm{ad} : \mathfrak{h} \curvearrowright \tilde{\mathfrak{g}}$ is diagonalized by the grading: h_i acts on anything of degree $q \in Q$ by $\langle q, \alpha_i^\vee \rangle$.

Let $\tilde{\mathfrak{n}}_+$ be the subalgebra of $\tilde{\mathfrak{g}}$ generated by $\{e_i\}_{i=1}^n$ and let $\tilde{\mathfrak{n}}_-$ be the subalgebra of $\tilde{\mathfrak{g}}$ generated by $\{f_i\}_{i=1}^n$; the algebras $\tilde{\mathfrak{n}}_\pm$ are called the "upper-" and "lower-triangular" subalgebras. □

5.6.0.6 Proposition *Let $\Delta, \tilde{\mathfrak{g}}, \tilde{\mathfrak{h}}, \tilde{\mathfrak{n}}_\pm$ be as in Lemma/Definition 5.6.0.1. Then $\tilde{\mathfrak{g}} = \tilde{\mathfrak{n}}_- \oplus \tilde{\mathfrak{h}} \oplus \tilde{\mathfrak{n}}_+$ as vector spaces; this is the "triangular decomposition" of $\tilde{\mathfrak{g}}$.*

Proof That $\tilde{\mathfrak{n}}_-, \tilde{\mathfrak{h}}, \tilde{\mathfrak{n}}_+$ intersect trivially follows form the grading, so it suffices to show that $\tilde{\mathfrak{g}} = \tilde{\mathfrak{n}}_- + \tilde{\mathfrak{h}} + \tilde{\mathfrak{n}}_+$. By inspecting the relations, we see that $(\mathrm{ad}\, f_i)\tilde{\mathfrak{n}}_- \subseteq \tilde{\mathfrak{n}}_-$, $(\mathrm{ad}\, f_i)\tilde{\mathfrak{h}} \subseteq \langle f_i \rangle \subseteq \tilde{\mathfrak{n}}_-$, and $(\mathrm{ad}\, f_i)\tilde{\mathfrak{n}}_+ \subseteq \tilde{\mathfrak{h}} + \tilde{\mathfrak{n}}_+$. Therefore $\mathrm{ad}\, f_i$ preserves $\tilde{\mathfrak{n}}_- + \tilde{\mathfrak{h}} + \tilde{\mathfrak{n}}_+$, $\tilde{\mathfrak{h}}$ does so obviously, and $\mathrm{ad}\, e_i$ does so by the obvious symmetry $f_i \leftrightarrow e_i$. Therefore $\tilde{\mathfrak{n}}_- + \tilde{\mathfrak{h}} + \tilde{\mathfrak{n}}_+$ is an ideal of $\tilde{\mathfrak{g}}$ and therefore a subalgebra, but it contains all the generators of $\tilde{\mathfrak{g}}$. □

5.6.0.7 Proposition *Let $\Delta, \tilde{\mathfrak{g}}$ be as in Lemma/Definition 5.6.0.1, and let $\lambda \in \mathfrak{h}^*$. Write $\mathbb{C}\langle f_1, \dots, f_n\rangle$ for the free algebra generated by noncommuting symbols $f_1, \dots, f_n$ and $M_\lambda \overset{\text{def}}{=} \mathbb{C}\langle f_1, \dots, f_n\rangle v_\lambda$ for its free module generated by the symbol v_λ. Then there exists an action of $\tilde{\mathfrak{g}}$ on M_λ such that:*

$$f_i\left(\prod f_{j_k} v_\lambda\right) = \left(f_i \prod f_{j_k}\right) v_\lambda \tag{5.6.0.8}$$

$$h_i\left(\prod f_{j_k} v_\lambda\right) = \left(\lambda(h_i) - \sum_k a_{i,j_k}\right)\left(\prod f_{j_k} v_\lambda\right) \tag{5.6.0.9}$$

$$e_i\left(\prod f_{j_k} v_\lambda\right) = \sum_{k \text{ s.t. } j_k = i} f_{j_1} \cdots f_{j_{k-1}} h_i f_{j_{k+1}} \cdots f_{j_l} v_\lambda \tag{5.6.0.10}$$

Proof We have only to check that the action satisfies the relations equations (5.6.0.2) to (5.6.0.5). The Q-grading verifies equations (5.6.0.2), (5.6.0.3), and (5.6.0.5); we need only to check equation (5.6.0.4). When $i \neq j$, the action by e_i ignores any action by f_j, and so we need only check that $[e_i, f_i]$ acts by h_i. Write $\underline{f}$ for some monomial $f_{j_1} \cdots f_{j_n}$. Then $e_i f_i(\underline{f} v_\lambda) = e_i(f_i \underline{f} v_\lambda) = h_i \underline{f} v_\lambda + f_i e_i(\underline{f} v_\lambda)$, clear by the construction. □

5.6.0.11 Definition *The $\tilde{\mathfrak{g}}$-module M_λ defined in Proposition 5.6.0.7 is the* Verma module *of $\tilde{\mathfrak{g}}$ with weight λ.*

5.6.0.12 Corollary *The map $\mathfrak{h} \to \tilde{\mathfrak{h}}$ is an isomorphism, so $\mathfrak{h} \hookrightarrow \tilde{\mathfrak{h}}$. The upper- and lower-triangular algebras $\tilde{\mathfrak{n}}_-$ and $\tilde{\mathfrak{n}}_+$ are free on $\{f_i\}$ and $\{e_i\}$ respectively.* □

5.6.0.13 Proposition *Assume that Δ is an indecomposable system of simple roots, in the sense that the Dynkin diagram of the Cartan matrix of Δ is connected. Construct $\tilde{\mathfrak{g}}$ as in Lemma/Definition 5.6.0.1. Then any proper ideal of $\tilde{\mathfrak{g}}$ is graded, contained in $\tilde{\mathfrak{n}}_- + \tilde{\mathfrak{n}}_+$, and does not contain any e_i or f_i.*

Proof The grading on $\tilde{\mathfrak{g}}$ is determined by the adjoint action of $\mathfrak{h} = \tilde{\mathfrak{h}}$. Let $\mathfrak{a}$ be an ideal of $\tilde{\mathfrak{g}}$ and $a \in \mathfrak{a}$. Let $a = \sum a_q g_q$ where g_q are homogeneous of degree $q \in Q$. Then $[h_i, a] = \sum \langle q, \alpha_i^\vee \rangle a_q g_q$, and so $[\mathfrak{h}, a]$ has the same dimension as the number of non-zero coefficients a_q; in particular, $g_q \in [\mathfrak{h}, a]$. Thus $\mathfrak{a}$ is graded.

Suppose that $\mathfrak{a}$ has a degree-0 part, i.e. suppose that there is some $h \in \mathfrak{h} \cap \mathfrak{a}$. Since the Cartan matrix a is nonsingular, there exists $\alpha_i \in \Delta$ with $\alpha_i(h) \neq 0$. Then $[f_i, h] = \alpha_i(h) f_i \neq 0$, and so $f_i \in \mathfrak{a}$.

Now let $\mathfrak{a}$ be any ideal with $f_i \in \mathfrak{a}$ for some i. Then $h_i = [e_i, f_i] \in \mathfrak{a}$ and $e_i = -\frac{1}{2}[e_i, h_i] \in \mathfrak{a}$. But let α_j be any neighbor of α_i in the Dynkin diagram. Then $a_{ij} \neq 0$, and so $[f_j, h_i] = a_{ij} f_j \neq 0$; then $f_j \in \mathfrak{a}$. Therefore, if the Dynkin diagram is connected, then any ideal of $\tilde{\mathfrak{g}}$ that contains some f_i (or some e_i by symmetry) contains every generator of $\tilde{\mathfrak{g}}$. □

5.6.0.14 Corollary *Under the conditions of Proposition 5.6.0.13, $\tilde{\mathfrak{g}}$ has a unique maximal proper ideal.*

Proof Let $\mathfrak{a}$ and $\mathfrak{b}$ be any two proper ideals of $\tilde{\mathfrak{g}}$. Then the ideal $\mathfrak{a} + \mathfrak{b}$ does not contain $\mathfrak{h}$ or any e_i or f_i, and so is a proper ideal. □

5.6.0.15 Definition *Let Δ be a system of simple roots with connected Dynkin diagram, and let $\tilde{\mathfrak{g}} = \tilde{\mathfrak{g}}_\Delta$ be defined as in Lemma/Definition 5.6.0.1. We define $\mathfrak{g} = \mathfrak{g}_\Delta$ as the quotient of $\tilde{\mathfrak{g}}$ by its unique maximal proper ideal. Then $\langle h_i, e_i, f_i\rangle \hookrightarrow \mathfrak{g}$, where by $\langle h_i, e_i, f_i\rangle$ we mean the linear span of the generators of $\tilde{\mathfrak{g}}$. Since we quotiented by a maximal ideal, $\mathfrak{g}$ is simple.*

5.6.0.16 Theorem (Serre Relations)
Let $\mathfrak{g}$ be as in Definition 5.6.0.15, and e_i, f_i the images of the corresponding generators of $\tilde{\mathfrak{g}}$. Then:

$$(\operatorname{ad} e_j)^{1-a_{ji}} e_i = 0 \tag{5.6.0.17}$$

$$(\operatorname{ad} f_j)^{1-a_{ji}} f_i = 0 \tag{5.6.0.18}$$

Proof We will check equation (5.6.0.18); equation (5.6.0.17) is exactly analogous. Let s be the left-hand-side of equation (5.6.0.18), interpreted as an element of $\tilde{\mathfrak{g}}$. We will show that the ideal generated by s is proper.

When $i = j$, $s = 0$, and when $i \neq j$, $a_{ji} \leq 0$, and so the degree of s is $-\alpha_i - (\geq 1)\alpha_j$. In particular, bracketing with f_k and h_k only moves the degree further from 0. Therefore, the claim follows from the following equation:

$$[e_k, s]_{\tilde{\mathfrak{g}}} = 0 \text{ for any } k \tag{5.6.0.19}$$

When $k \neq i, j$, $[e_k, f_i] = [e_k, f_j] = 0$. So it suffices to check equation (5.6.0.19) when $k = i, j$. Let $m = -a_{ji}$. When $k = j$, we compute:

$$\begin{aligned}
(\operatorname{ad} e_j)(\operatorname{ad} f_j)^{1+m} f_i &= \bigl[\operatorname{ad} e_j, (\operatorname{ad} f_j)^{1+m}\bigr] f_i + (\operatorname{ad} f_j)^{1+m} (\operatorname{ad} e_j) f_i && (5.6.0.20)\\
&= \bigl[\operatorname{ad} e_j, (\operatorname{ad} f_j)^{1+m}\bigr] f_i + 0 && (5.6.0.21)\\
&= \sum_{l=0}^{m} (\operatorname{ad} f_j)^{m-l} \bigl(\operatorname{ad}[e_j, f_j]\bigr) (\operatorname{ad} f_j)^l f_i && (5.6.0.22)\\
&= \sum_{l=0}^{m} (\operatorname{ad} f_j)^{m-l} (\operatorname{ad} h_j) (\operatorname{ad} f_j)^l f_i && (5.6.0.23)\\
&= \sum_{l=0}^{m} \bigl(l(-\langle \alpha_j, \alpha_j^\vee\rangle) - \langle \alpha_i, \alpha_j^\vee\rangle\bigr) (\operatorname{ad} f_j)^m f_i && (5.6.0.24)\\
&= \left(\sum_{l=0}^{m} (-2l + m)\right) (\operatorname{ad} f_j)^m f_i && (5.6.0.25)\\
&= \left(-2\frac{m(m+1)}{2} + (m+1)m\right) (\operatorname{ad} f_j)^m f_i = 0 && (5.6.0.26)
\end{aligned}$$

where equation (5.6.0.21) follows by $[e_i, f_j] = 0$, equation (5.6.0.22) by the fact that ad is a Lie algebra homomorphism, and the rest is equations (5.6.0.3) and (5.6.0.4), that $m = -a_{ji}$, and arithmetic.

When $k = i$, e_i and f_j commute, and we have:

$$\begin{aligned}
(\operatorname{ad} e_i)(\operatorname{ad} f_j)^{1+m} f_i &= \bigl[\operatorname{ad} e_j, (\operatorname{ad} f_j)^{1+m}\bigr] f_i + (\operatorname{ad} f_j)^{1+m} (\operatorname{ad} e_i) f_i && (5.6.0.27)\\
&= 0 + (\operatorname{ad} f_j)^{1+m} (\operatorname{ad} e_i) f_i && (5.6.0.28)\\
&= (\operatorname{ad} f_j)^{1+m} h_i = 0 && (5.6.0.29)
\end{aligned}$$

provided that $m \geq 1$. When $m = 0$, we use the symmetrizability of the Cartan matrix: if $a_{ji} = 0$ then $a_{ij} = 0$. Therefore

$$(\operatorname{ad} e_i)(\operatorname{ad} f_j)^{1-a_{ji}} f_i = (\operatorname{ad} e_i)[f_j, f_i] = -(\operatorname{ad} e_i)(\operatorname{ad} f_i)^{1-a_{ij}} f_j$$

which vanishes by the first computation. □

We have defined for each indecomposable Dynkin diagram Δ a simple Lie algebra $\mathfrak{g}_\Delta$. If $\Delta = \Delta_1 \times \Delta_2$ is a disjoint union of Dynkin diagrams, we define $\mathfrak{g}_\Delta \stackrel{\text{def}}{=} \mathfrak{g}_{\Delta_1} \times \mathfrak{g}_{\Delta_2}$.

5.6.0.30 Definition *Let V be a (possibly-infinite-dimensional) $\mathfrak{g}$-module. An element $v \in V$ is* integrable *if for each i, the $\mathfrak{sl}(2)_i$-submodule of V generated by v is finite-dimensional. We write $I(V)$ for the set of integrable elements of V.*

5.6.0.31 Lemma *Let V be a $\mathfrak{g}$-module. Then $I(V)$ is a $\mathfrak{g}$-submodule.*

Proof Let $N \subseteq V$ be an $(n+1)$-dimensional irreducible representation of $\mathfrak{sl}(2)_i$; then it is isomorphic to V_n defined in Example 5.2.0.5. It suffices to show that $e_j N$ is contained within some finite-dimensional $\mathfrak{sl}(2)_i$ submodule of V for $i \neq j$; the rest follows by switching $e \leftrightarrow f$ and permuting the indices, using the fact that $\{e_j, f_j\}$ generate $\mathfrak{g}$.

Then N is spanned by $\{{f_i}^k v_0\}_{k=0}^n$ where $v_0 \in N$ is the vector annihilated by e_i; in particular, ${f_i}^{n+1} v_0 = 0$. Since e_j and f_i commute, $e_j N$ is spanned by $\{{f_i}^k e_j v_0\}_{k=0}^n$. It suffices to compute the $\mathfrak{sl}(2)_i$ module generated by $e_j v_0$, or at least to show that it is finite-dimensional. The action of h_i on $e_j v_0$ is $h_i e_j v_0 = ([h_i, e_j] + e_j h_i) v_0 = (a_{ij} + n) e_j v_0$. For $k \neq n+1$, $f_i^k e_j v_0 = e_j f_i^k v_0 = 0$. Moreover, by Theorem 5.6.0.16, $e_i^k e_j v_0 = [e_i^k, e_j] v_0 + e_j e_i^k v_0 = (\operatorname{ad} e_i)^k (e_j) v_0 + 0$, which vanishes for large enough k. Then the result follows by Theorem 3.2.2.1 and the fact that $[e_i, f_i] = h_i$. □

5.6.0.32 Corollary *Let Δ be a Dynkin diagram and define $\mathfrak{g}$ as above. Then $\mathfrak{g}$ is* ad*-integrable.*

Proof Since $\{e_k, f_k\}$ generate $\mathfrak{g}$, it suffices to show that e_k and f_k are ad-integrable for each k. But the $\mathfrak{sl}(2)_i$-module generated by f_k has f_k as its highest-weight vector, since $[e_i, f_k] = 0$, and is finite-dimensional, since $(\operatorname{ad} f_i)^n f_k = 0$ for large enough n by Theorem 5.6.0.16. □

5.6.0.33 Corollary *The non-zero weights R of $\operatorname{ad} : \mathfrak{g} \curvearrowright \mathfrak{g}$ form a root system.*

Proof Axioms **RS1**, **RS2**, **RS3**, **RS4**, and **Nondeg** of Definition 5.4.2.1 follow from the ad-integrability. Axiom **Reduced** and that R is finite follow from Lemma 5.4.4.5. □

5.6.0.34 Theorem (Classification of finite-dimensional simple Lie algebras)
The list given in Theorem 5.5.2.17 classifies the finite-dimensional simple Lie algebras over $\mathbb{C}$.

Proof A Lie algebra with an indecomposable root system is simple, because any such system has a highest root, and linear combination of roots generates the highest root, and the highest root generates the entire algebra. So it suffices to show that two simple Lie algebras with isomorphic root systems are isomorphic.

Let Δ be an indecomposable root system, and define $\tilde{\mathfrak{g}}$ and $\mathfrak{g}$ as above. Let $\mathfrak{g}_1$ be a Lie algebra with root system Δ. Then the relations defining $\tilde{\mathfrak{g}}$ hold in $\mathfrak{g}_1$, and so there is a surjection $\tilde{\mathfrak{g}} \twoheadrightarrow \mathfrak{g}_1$; if $\mathfrak{g}_1$ is simple, then the kernel of this surjection is a maximal ideal of $\tilde{\mathfrak{g}}$. But $\tilde{\mathfrak{g}}$ has a unique maximal ideal, and $\mathfrak{g}$ is the quotient by this ideal; thus $\mathfrak{g}_1 \cong \mathfrak{g}$. □

5.6.0.35 Example The families $ABCD$ correspond to the classical Lie algebras: $A_n \leftrightarrow \mathfrak{sl}(n+1)$, $B_n \leftrightarrow \mathfrak{so}(2n+1)$, $C_n \leftrightarrow \mathfrak{sp}(n)$, and $D_n \leftrightarrow \mathfrak{so}(2n)$. We recall that we have defined $\mathfrak{sp}(n)$ as the Lie algebra that fixes the nondegenerate antisymmetric $2n \times 2n$ bilinear form: $\mathfrak{sp}(n) \subseteq \mathfrak{gl}(2n)$. The EFG Lie algebras are new.

The coincidences in Example 5.5.2.18 correspond to coincidences of classical Lie algebras: $\mathfrak{so}(6) \cong \mathfrak{sl}(4)$, $\mathfrak{so}(5) \cong \mathfrak{sp}(2)$, and $\mathfrak{so}(4) \cong \mathfrak{sl}(2) \times \mathfrak{sl}(2)$. The identities $\mathfrak{so}(3) \cong \mathfrak{sp}(1) \cong \mathfrak{sl}(2)$ suggest that we define $B_1 = C_1 = A_1 = \bullet$, but $\mathfrak{sl}(2)$ is not congruent to $\mathfrak{so}(2)$, so we do not assign meaning to D_1. $\lozenge$

Exercises

1. (a) Show that $\mathrm{SL}(2,\mathbb{R})$ is topologically the product of a circle and two copies of $\mathbb{R}$, hence it is not simply connected.

   (b) Let S be the simply connected cover of $\mathrm{SL}(2,\mathbb{R})$. Show that its finite-dimensional complex representations, i.e., real Lie group homomorphisms $S \to \mathrm{GL}(n,\mathbb{C})$, are determined by corresponding complex representations of the Lie algebra $\mathrm{Lie}(S)^{\mathbb{C}} = \mathfrak{sl}(2,\mathbb{C})$, and hence factor through $\mathrm{SL}(2,\mathbb{R})$. Thus S is a simply connected real Lie group with no faithful finite-dimensional representation.

2. (a) Let U be the group of 3×3 upper-unitriangular complex matrices. Let $\Gamma \subseteq U$ be the cyclic subgroup of matrices
$$\begin{bmatrix} 1 & 0 & m \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$
   where $m \in \mathbb{Z}$. Show that $G = U/\Gamma$ is a (non-simply-connected) complex Lie group that has no faithful finite-dimensional representation.

   (b) Adapt the solution to Set 4, Problem 2(b) to construct a faithful, irreducible infinite-dimensional linear representation V of G.

3. Following the outline below, prove that if $\mathfrak{h} \subseteq \mathfrak{gl}(n,\mathbb{C})$ is a real Lie subalgebra with the property that every $X \in \mathfrak{h}$ is diagonalizable and has purely imaginary eigenvalues, then the corresponding connected Lie subgroup $H \subseteq GL(n,\mathbb{C})$ has compact closure (this completes the solution to Set 1, Problem 7).

   (a) Show that $\operatorname{ad} X$ is diagonalizable with imaginary eigenvalues for every $X \in \mathfrak{h}$.

   (b) Show that the Killing form of $\mathfrak{h}$ is negative semidefinite and its radical is the center of $\mathfrak{h}$. Deduce that $\mathfrak{h}$ is reductive and the Killing form of its semi-simple part is negative definite. Hence the Lie subgroup corresponding to the semi-simple part is compact.

   (c) Show that the Lie subgroup corresponding to the center of $\mathfrak{h}$ is a dense subgroup of a compact torus. Deduce that the closure of H is compact.

   (d) Show that H is compact — that is, closed — if and only if it further holds that the center of $\mathfrak{h}$ is spanned by matrices whose eigenvalues are rational multiples of i.

4. Let $V_n = \mathcal{S}^n(\mathbb{C}^2)$ be the $(n+1)$-dimensional irreducible representation of $\mathfrak{sl}(2,\mathbb{C})$.

   (a) Show that for $m \leq n$, $V_m \otimes V_n \cong V_{n-m} \oplus V_{n-m+2} \oplus \cdots \oplus V_{n+m}$, and deduce that the decomposition into irreducibles is unique.

   (b) Show that in any decomposition of $V_1^{\otimes n}$ into irreducibles, the multiplicity of V_n is equal to 1, the multiplicity of V_{n-2k} is equal to $\binom{n}{k} - \binom{n}{k-1}$ for $k = 1, \ldots, \lfloor n/2 \rfloor$, and all other irreducibles V_m have multiplicity zero.

5. Let a be a symmetric generalized Cartan matrix, i.e. a is symmetric with diagonal entries 2 and off-diagonal entries 0 or -1. Let Γ be a subgroup of the automorphism group of the Dynkin diagram D of a, such that every edge of D has its endpoints in distinct Γ orbits. Define the *folding* D' of D to be the diagram with a node for every Γ orbit I of nodes in D, with edge weight k from I to J if each node of I is adjacent in D to k nodes of J. Denote by a' the generalized Cartan matrix with diagram D'.

   (a) Show that a' is symmetrizable and that every symmetrizable generalized Cartan matrix (not assumed to be of finite type) can be obtained by folding from a symmetric one.

   (b) Show that every folding of a finite type symmetric Cartan matrix is of finite type.

   (c) Verify that every non-symmetric finite type Cartan matrix is obtained by folding from a unique symmetric finite type Cartan matrix.

6. An indecomposable symmetrizable generalized Cartan matrix a is said to be of *affine type* if $\det(a) = 0$ and all the proper principal minors of a are positive.

   (a) Classify the affine Cartan matrices.

   (b) Show that every non-symmetric affine Cartan matrix is a folding, as in the previous problem, of a symmetric one.

   (c) Let $\mathfrak{h}$ be a vector space, $\alpha_i \in \mathfrak{h}^*$ and $\alpha_i^\vee \in \mathfrak{h}$ vectors such that a is the matrix $\langle \alpha_j, \alpha_i^\vee \rangle$. Assume that this realization is non-degenerate in the sense that the vectors α_i are linearly independent. Define the *affine Weyl group* W to be generated by the reflections s_{α_i}, as usual. Show that W is isomorphic to the semidirect product $W_0 \ltimes Q$ where Q and W_0 are the root lattice and Weyl group of a unique finite root system, and that every such $W_0 \ltimes Q$ occurs as an affine Weyl group.

   (d) Show that the affine and finite root systems related as in (c) have the property that the affine Dynkin diagram is obtained by adding a node to the finite one, in a unique way if the finite Cartan matrix is symmetric.

7. Work out the root systems of the orthogonal Lie algebras $\mathfrak{so}(m,\mathbb{C})$ explicitly, thereby verifying that they correspond to the Dynkin diagrams B_n if $m = 2n+1$, or D_n if $m = 2n$. Deduce the isomorphisms $\mathfrak{so}(4,\mathbb{C}) \cong \mathfrak{sl}(2,\mathbb{C}) \times \mathfrak{sl}(2,\mathbb{C})$, $\mathfrak{so}(5,\mathbb{C}) \cong \mathfrak{sp}(4,\mathbb{C})$, and $\mathfrak{so}(6,\mathbb{C}) \cong \mathfrak{sl}(4,\mathbb{C})$.

8. Show that the Weyl group of type B_n or C_n (they are the same because these two root systems are dual to each other) is the group $S_n \ltimes (\mathbb{Z}/2\mathbb{Z})^n$ of signed permutations, and that the Weyl group of type D_n is its subgroup of index two consisting of signed permutations

with an even number of sign changes, i.e., the semidirect factor $(\mathbb{Z}/2\mathbb{Z})^n$ is replaced by the kernel of S_n-invariant summation homomorphism $(\mathbb{Z}/2\mathbb{Z})^n \to \mathbb{Z}/2\mathbb{Z}$

9. Let $(\mathfrak{h}, R, R^\vee)$ be a finite root system, $\Delta = \{\alpha_1, \ldots, \alpha_n\}$ the set of simple roots with respect to a choice of positive roots R_+, $s_i = s_{\alpha_i}$ the corresponding generators of the Weyl group W. Given $w \in W$, let $l(w)$ denote the minimum length of an expression for w as a product of the generators s_i.
   (a) If $w = s_{i_1} \ldots s_{i_r}$ and $w(\alpha_j) \in R_-$, show that for some k we have $\alpha_{i_k} = s_{i_{k+1}} \ldots s_{i_r}(\alpha_j)$, and hence $s_{i_k} s_{i_{k+1}} \ldots s_{i_r} = s_{i_{k+1}} \ldots s_{i_r} s_j$. Deduce that $l(ws_j) = l(w) - 1$ if $w(\alpha_j) \in R_-$.
   (b) Using the fact that the conclusion of (a) also holds for $v = ws_j$, deduce that $l(ws_j) = l(w) + 1$ if $w(\alpha_j) \notin R_-$.
   (c) Conclude that $l(w) = |w(R_+) \cup R_-|$ for all $w \in W$. Characterize $l(w)$ in more explicit terms in the case of the Weyl groups of type A and B/C.
   (d) Assuming that $\mathfrak{h}$ is over $\mathbb{R}$, show that the dominant cone $X = \{\lambda \in \mathfrak{h} : \langle \lambda, \alpha_i^\vee \rangle \geq 0 \text{ for all } i\}$ is a fundamental domain for W, i.e., every vector in $\mathfrak{h}$ has a unique element of X in its W orbit.
   (e) Deduce that $|W|$ is equal to the number of connected regions into which $\mathfrak{h}$ is separated by the removal of all the root hyperplanes $\langle \lambda, \alpha^\vee \rangle$, $\alpha^\vee \in R^\vee$.

10. Let $h_1, \ldots, h_r$ be linear forms in variables $x_1, \ldots, x_n$ with integer coefficients. Let $\mathbb{F}_q$ denote the finite field with $q = p^e$ elements. Prove that except in a finite number of "bad" characteristics p, the number of vectors $v \in \mathbb{F}_q^n$ such that $h_i(v) \neq 0$ for all i is given for all q by a polynomial $\chi(q)$ in q with integer coefficients, and that $(-1)^n \chi(-1)$ is equal to the number of connected regions into which $\mathbb{R}^n$ is separated by the removal of all the hyperplanes $h_i = 0$.

    Pick your favorite finite root system and verify that in the case where the h_i are the root hyperplanes, the polynomial $\chi(q)$ factors as $(q - e_1) \ldots (q - e_n)$ for some positive integers e_i called the *exponents* of the root system. In particular, verify that the sum of the exponents is the number of positive roots, and that (by Problem 9(e)) the order of the Weyl group is $\prod_i (1 + e_i)$

11. The *height* of a positive root α is the sum of the coefficients c_i in its expansion $\alpha = \sum_i c_i \alpha_i$ on the basis of simple roots.

    Pick your favorite root system and verify that for each $k \geq 1$, the number of roots of height k is equal to the number of the exponents e_i in Problem 10 for which $e_i \geq k$.

12. Pick your favorite root system and verify that if h denotes the height of the highest root plus one, then the number of roots is equal to h times the rank. This number h is called the *Coxeter number*. Verify that, moreover, the multiset of exponents (see Problem 10) is invariant with respect to the symmetry $e_i \mapsto h - e_i$.

13. A *Coxeter element* in the Weyl group W is the product of all the simple reflections, once each, in any order. Prove that a Coxeter element is unique up to conjugacy. Pick your favorite root system and verify that the order of a Coxeter element is equal to the Coxeter number (see Problem 12).

14. The *fundamental weights* λ_i are defined to be the basis of the weight lattice P dual to the basis of simple coroots in $Q^\vee$, i.e., $\langle \lambda_i, \alpha_j^\vee \rangle = \delta_{ij}$.

    (a) Prove that the stabilizer in W of λ_i is the Weyl group of the root system whose Dynkin diagram is obtained by deleting node i of the original Dynkin diagram.

    (b) Show that each of the root systems E_6, E_7, and E_8 has the property that its highest root is a fundamental weight, and identify the corresponding simple root. Deduce that the order of the Weyl group $W(E_k)$ in each case is equal to the number of roots times the order of the Weyl group $W(G)$, where G is the root system formed from E_k by deleting the identified root. Use this to calculate the orders of the Weyl groups $W(E_k)$.

15. Let $e_1, \ldots, e_8$ be the usual orthonormal basis of coordinate vectors in Euclidean space $\mathbb{R}^8$. The root system of type E_8 can be realized in $\mathbb{R}^8$ with simple roots $\alpha_i = e_i - e_{i+1}$ for $i = 1, \ldots, 7$ and
$$\alpha_8 = \left(-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right).$$
Show that the root lattice Q is equal to the weight lattice P, and that in this realization, Q consists of all vectors $\beta \in \mathbb{Z}^8$ such that $\sum_i \beta_i$ is even and all vectors $\beta \in \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right) + \mathbb{Z}^8$ such that $\sum_i \beta_i$ is odd. Show that the root system consists of all vectors of squared length 2 in Q, namely, the vectors $\pm e_i \pm e_j$ for $i < j$, and all vectors with coordinates $\pm\frac{1}{2}$ and an odd number of coordinates with each sign.

16. Show that the root system of type F_4 has 24 long roots and 24 short roots, and that the roots of each length form a root system of type D_4. Show that the highest root and the highest short root are the fundamental weights at the end nodes of the diagram. Then use Problem 14(a) to calculate the order of the Weyl group $W(F_4)$. Show that $W(F_4)$ acts on the set of short (resp. long roots) as the semidirect product $S_3 \ltimes W(D_4)$, where the symmetric group S_3 on three letters acts on $W(D_4)$ as the automorphism group of its Dynkin diagram.

17. Pick your favorite root system and verify that the generating function $W(t) = \sum_{w \in W} t^{l(w)}$ is equal to $\prod_i (1 + t + \cdots + t^{e_i})$, where e_i are the exponents as in Problem 10.

18. Let S be the subring of W-invariant elements in the ring of polynomial functions on $\mathfrak{h}$. Pick your favorite root system and verify that S is a polynomial ring generated by homogeneous generators of degrees $e_i + 1$, where e_i are the exponents as in Problem 10.

Chapter 6

Representation Theory of Semisimple Lie Groups

6.1 Irreducible Lie-algebra representations

Any representation of a Lie group induces a representation of its Lie algebra, so we start our story there. We recall Theorem 4.4.3.8: any finite-dimensional representation of a semisimple Lie algebra is the direct sum of simple representations. In Section 5.2 we computed the finite-dimensional simple representations of $\mathfrak{sl}(2)$; we now generalize that theory to arbitrary finite-dimensional semisimple Lie algebras.

6.1.0.1 Lemma / Definition *Let $\mathfrak{g}$ be a semisimple Lie algebra with Cartan subalgebra $\mathfrak{h}$ and root system R, and choose a system of positive roots R_+. Let $\mathfrak{n}_\pm = \bigoplus_{\alpha \in R_\pm} \mathfrak{g}_\alpha$ be the upper- and lower-triangular subalgebras; then $\mathfrak{g} = \mathfrak{n}_- \oplus \mathfrak{h} \oplus \mathfrak{n}_+$ as a a vector space. We define the* Borel subalgebra *$\mathfrak{b} = \mathfrak{h} \oplus \mathfrak{n}_+$, and $\mathfrak{n}_+$ is an ideal of $\mathfrak{b}$ with $\mathfrak{h} = \mathfrak{b}/\mathfrak{n}_+$.*

Pick $\lambda \in \mathfrak{h}^*$; then $\mathfrak{b}$ has a one-dimensional module $\mathbb{C}v_\lambda$, where $hv_\lambda = \lambda(h)\, v_\lambda$ for $h \in \mathfrak{h}$ and $\mathfrak{n}_+ v_\lambda = 0$.

As a subalgebra, $\mathfrak{b}$ acts on $\mathfrak{g}$ from the right, and so we define the Verma module *of $\mathfrak{g}$ with weight λ by:*

$$M_\lambda \overset{\text{def}}{=} \mathcal{U}\mathfrak{g} \otimes_{\mathcal{U}\mathfrak{b}} \mathbb{C}v_\lambda$$

As a vector space, $M_\lambda \cong \mathcal{U}\mathfrak{n}_- \otimes_{\mathbb{C}} \mathbb{C}v_\lambda$. It is generated as a $\mathfrak{g}$-module by v_λ with the relations $hv_\lambda = \lambda(h)v_\lambda$, $\mathfrak{n}_+ v_\lambda = 0$, and no relations on the action of $\mathfrak{n}_-$ except those from $\mathfrak{g}$.

Proof The explicit description of M_λ follows from Theorem 3.2.2.1: $\mathcal{U}\mathfrak{g} = \mathcal{U}\mathfrak{n}_- \otimes \mathcal{U}\mathfrak{h} \otimes \mathcal{U}\mathfrak{n}_+$ as vector spaces. $\square$

6.1.0.2 Corollary *Any module with highest weight λ is a quotient of M_λ.* $\square$

6.1.0.3 Lemma *Let $\Delta = \{\alpha_1, \dots, \alpha_n\}$ be the simple roots of $\mathfrak{g}$, and let $Q \overset{\text{def}}{=} \mathbb{Z}\Delta$ be the root lattice and $Q_+ \overset{\text{def}}{=} \mathbb{Z}_{\geq 0}\Delta$. Then the* weight grading *given by the action of $\mathfrak{h}$ on the Verma module M_λ is:*

$$M_\lambda = \bigoplus_{\beta \in Q_+} (M_\lambda)_{\lambda - \beta}$$

Moreover, let $N \subseteq M_\lambda$ be a proper submodule. Then $N \subseteq \bigoplus_{\beta \in Q_+ \smallsetminus \{0\}} (M_\lambda)_{\lambda-\beta}$.

Proof The description of the weight grading follows directly from the description of M_λ given in Lemma/Definition 6.1.0.1. Any submodule is graded by the action of $\mathfrak{h}$. Since $(M_\lambda)_\lambda = \mathbb{C} v_\lambda$ is one-dimensional and generates M_λ, a proper submodule cannot intersect $(M_\lambda)_\lambda$. □

6.1.0.4 Corollary *For any $\lambda \in \mathfrak{h}^*$, the Verma module M_λ has a unique maximal proper submodule. The quotient $M_\lambda \twoheadrightarrow L_\lambda$ is an irreducible $\mathfrak{g}$-module. Conversely, any irreducible $\mathfrak{g}$-module with highest weight λ is isomorphic to L_λ, since it must be a quotient of M_λ by a maximal ideal.* □

6.1.0.5 Definition *Let $\mathfrak{g}$ be a semisimple Lie algebra and $\Delta = \{\alpha_1, \dots, \alpha_n\}$. We recall the* root lattice $Q \overset{\text{def}}{=} \mathbb{Z}\Delta$ *and the* weight lattice $P \overset{\text{def}}{=} \{\lambda \in \mathfrak{h}^* \text{ s.t. } \langle \lambda, Q^\vee \rangle \subseteq \mathbb{Z}\}$. *A* dominant integral weight *is an element of $P_+ \overset{\text{def}}{=} \{\lambda \in P \text{ s.t. } \langle \lambda, \alpha_i^\vee \rangle \geq 0 \, \forall i\}$.*

Recall Definition 5.6.0.30: an element v in a possibly-infinite-dimensional $\mathfrak{g}$-module V is *integrable* if for each i, the $\mathfrak{sl}(2)_i$-submodule of V generated by v is finite-dimensional.

6.1.0.6 Proposition *If $\lambda \in P_+$, then L_λ consists of integrable elements.*

Proof Since L_λ is irreducible, its submodule of integrable elements is either 0 or the whole module. So it suffices to show that if $\lambda \in P_+$, then v_λ is integrable. Pick a simple root α_i. By construction, $e_i v_\lambda = 0$ and $h_i v_\lambda = \langle \lambda, \alpha_i^\vee \rangle v_\lambda$. Since $\lambda \in P_+$, $\langle \lambda, \alpha_i^\vee \rangle = m \geq 0$ is an integer. Consider the $\mathfrak{sl}(2)_i$-submodule of M_λ generated by v_λ; if m is a nonnegative integer, from the representation theory of $\mathfrak{sl}(2)$ we know that $e_i f_i^{m+1} v_\lambda = 0$. But if $j \neq i$, then $e_j f_i^{m+1} v_\lambda = f_i^{m+1} e_j v_\lambda = 0$. Recalling the grading, we see then that $f_i^{m+1} v_\lambda$ generates a submodule of M_λ, and so $f_i^{m+1} v_\lambda \mapsto 0$ in L_λ. Hence the $\mathfrak{sl}(2)_i$-submodule of L_λ generated by v_λ is finite, and so v_λ is integrable. □

6.1.0.7 Definition *Let $\mathfrak{g}$ be a semisimple Lie algebra. We define the category $\hat{\mathcal{O}}$ to be a full subcategory of the category $\mathfrak{g}$-MOD of (possibly-infinite-dimensional) $\mathfrak{g}$ modules. The objects $X \in \hat{\mathcal{O}}$ are required to satisfy the following conditions:*

- *The action $\mathfrak{h} \curvearrowright X$ is diagonalizable.*
- *For each $\lambda \in \mathfrak{h}^*$, the weight space X_λ is finite-dimensional.*
- *There exists a finite set $S \subseteq \mathfrak{h}^*$ such that the weights of X lie in $S + (-Q_+)$.*

The more commonly used "category $\mathcal{O}$" imposes an extra technical condition.

6.1.0.8 Lemma *The category $\hat{\mathcal{O}}$ is closed under submodules, quotients, extensions, and tensor products.*

Proof The $\mathfrak{h}$-action grades subquotients of any graded module, and acts diagonally. An extension of graded modules is graded, with graded components extensions of the corresponding graded components. Since $\mathfrak{g}$ is semisimple, any extension of finite-dimensional modules is a direct sum, and so the $\mathfrak{h}$-action is diagonal on any extension of objects in $\hat{\mathcal{O}}$. Finally, tensor products are handled by Lemma 5.3.1.8. □

6.1.0.9 Definition *Write the additive group $\mathfrak{h}^*$ multiplicatively: $\lambda \mapsto x^\lambda$. The* group algebra $\mathbb{Z}[\mathfrak{h}^*]$ *is the algebra of "polynomials" $\sum c_i x^{\lambda_i}$, with the obvious addition and multiplication. I.e. $\mathbb{Z}[\mathfrak{h}^*]$ is the free abelian group $\bigoplus_{\lambda \in \mathfrak{h}^*} \mathbb{Z}x^\lambda$, with multiplication given on a basis by $x^\lambda x^\mu = x^{\lambda+\mu}$.*

Let $\mathbb{Z}[-Q_+]$ be the subalgebra of $\mathbb{Z}[\mathfrak{h}^*]$ generated by $\{x^\lambda \text{ s.t. } -\lambda \in Q_+\}$. This has a natural topology given by setting $\|x^{-\alpha_i}\| = c^{-\alpha_i}$ for α_i a simple root and c some real constant with $c > 1$. We let $\mathbb{Z}[\![-Q_+]\!]$ be the completion of $\mathbb{Z}[-Q_+]$ with respect to this topology. Equivalently, $\mathbb{Z}[\![-Q_+]\!]$ is the algebra of formal power series *in the variables $x^{-\alpha_1}, \dots, x^{-\alpha_n}$ with integer coefficients.*

Then $\mathbb{Z}[-Q_+]$ is a subalgebra of both $\mathbb{Z}[\mathfrak{h}^*]$ and $\mathbb{Z}[\![-Q_+]\!]$. We will write $\mathbb{Z}[h^*, -Q_+]\!]$ for the algebra $\mathbb{Z}[\mathfrak{h}^*] \otimes_{\mathbb{Z}[-Q_+]} \mathbb{Z}[\![-Q_+]\!]$.

The algebra $\mathbb{Z}[h^*, -Q_+]\!]$ is a formal gadget, consisting of formal fractional Laurant series. We use it as a space of generating functions.

6.1.0.10 Definition *Given $X \in \hat{\mathcal{O}}$, its* character *is* $\operatorname{ch}(X) \in \mathbb{Z}[h^*, -Q_+]\!]$ *by:*

$$\operatorname{ch}(X) \stackrel{\text{def}}{=} \sum_{\lambda \text{ a weight of } X} \dim(X_\lambda) x^\lambda$$

We remark that every coefficient of $\operatorname{ch}(X)$ is a nonnegative integer. It is also clear that ch is additive for extensions.

6.1.0.11 Example Let M_λ be the Verma module with weight λ, and let R_+ be the set of positive roots of $\mathfrak{g}$. Then

$$\operatorname{ch}(M_\lambda) = \frac{x^\lambda}{\prod_{\alpha \in R_+}(1 - x^{-\alpha})} \stackrel{\text{def}}{=} x^\lambda \prod_{\alpha \in R_+} \sum_{l=0}^{\infty} x^{-l\alpha}$$

This follows from Theorem 3.2.2.1, the explicit description of $M_\lambda \cong \mathcal{U}\mathfrak{n}_- \otimes \mathbb{C}v_\lambda$, and some elementary combinatorics. $\Diamond$

6.1.0.12 Proposition *Let $\mathfrak{g}$ be simple Lie algebra, P_+ the set of dominant integral weights, and W the Weyl group. Let $\lambda \in P_+$, and L_λ the irreducible quotient of M_λ given in Corollary 6.1.0.4. Then:*

1. $\operatorname{ch}(L_\lambda)$ *is W-invariant.*
2. *If μ is a weight of L_λ, then $\mu \in W(\nu)$ for some $\nu \in P_+ \cap (\lambda - Q_+)$.*
3. *L_λ is finite-dimensional.*

Conversely, every finite-dimensional irreducible $\mathfrak{g}$-module is L_λ for a unique $\lambda \in P_+$.

Proof 1. We use Proposition 6.1.0.6: L_λ consists of integrable elements. Let α_i be a root of $\mathfrak{g}$; then L_λ splits as an $\mathfrak{sl}(2)_i$ module: $L_\lambda = \bigoplus V_a$, where each V_a is an irreducible $\mathfrak{sl}(2)_i$ submodule. In particular, $V_a = \mathbb{C}v_{a,m} \oplus \mathbb{C}v_{a,m-2} \cdots \oplus \mathbb{C}v_{a,-m}$ for some m depending on a, where h_i acts on $\mathbb{C}v_{a,l}$ by l. But $\operatorname{ch}(L_\lambda) = \sum_a \operatorname{ch}(V_a) = \sum_a \sum_{j=-m,-m+2,\dots,m} \operatorname{ch}(\mathbb{C}v_{a,m})$. Let $\operatorname{ch}\big(\mathbb{C}(v_a)_l\big) = x^{\mu_{a,l}}$; then $\langle \mu_{a,l}, \alpha_i^\vee \rangle = l$ by definition, and $v_{a,l-2} \in f_i \mathbb{C}v_{a,l}$, and so $s_i \mu_{a,l} = \mu_{a,-l}$. This shows that $\operatorname{ch}(V_a)$ is fixed under the action of s_i, and so $\operatorname{ch}(L_\lambda)$ is also s_i-invariant. But the reflections s_i generate W, and so $\operatorname{ch}(L_\lambda)$ is W-invariant.

2. We partially order P: $\nu \leq \mu$ if $\mu - \nu \in Q_+$. In particular, the weights of L_λ are all less than or equal to λ.

   Let $\lambda \in P$. Then $s_i(\lambda) = \lambda - \langle \lambda, \alpha_i^\vee \rangle \alpha_i$, and so $W(\lambda) \subseteq \lambda + Q$. If $\lambda \in P_+$ then $\langle \lambda, \alpha_i^\vee \rangle \geq 0$ for every i and so $s_i \leq \lambda$; if $\lambda \in P \smallsetminus P_+$ then there is some i with $\langle \lambda, \alpha_i^\vee \rangle < 0$, i.e. some i with $s_i(\lambda) > \lambda$. But W is finite, so for any $\lambda \in P$, $W(\lambda)$ has a maximal element, which must be in P_+. This proves that $P = W(P_+)$.

   Thus, if μ is a weight of L_λ, then $\mu \in W(\nu)$ for some $\nu \in P_+$. But by 1., ν is a weight of L_λ, and so $\nu \leq \lambda$. This proves statements 2.

   Moreover, the W-invariance of $\mathrm{ch}(L_\lambda)$ shows that if $\lambda \in P_+$, then $W(\lambda) \subseteq \lambda - Q_+$, and moreover that P_+ is a fundamental domain of W.

3. The Weyl group W is finite. Consider the two cones $\mathbb{R}_{\geq 0} P_+$ and $-\mathbb{R}_{\geq 0} Q_+$. Since the inner product (the symmetrization of the Cartan matrix) is positive definite and by construction the inner product of anything in $\mathbb{R}_{\geq 0} P_+$ with anything in $-\mathbb{R}_{\geq 0} Q_+$ is negative, the two cones intersect only at 0. Thus there is a hyperplane separating the cones: i.e. there exists a linear functional $\eta : \mathfrak{h}^*_\mathbb{R} \to \mathbb{R}$ such that its value is positive on P_+ but negative on $-Q_+$. Then $\lambda - Q_+$ is below the $\eta = \eta(\lambda)$ hyperplane. But $-Q_+$ is generated by $-\alpha_i$, each of which has a negative value under η, and so $\lambda - Q_+$ contains only finitely many points μ with $\eta(\mu) \geq 0$. Thus $P_+ \cap (\lambda - Q_+)$ is finite, and hence so is its image under W.

For the converse statement, let L be a finite-dimensional irreducible $\mathfrak{g}$-module, and let $v \in L$ be any vector. Then consider $\mathfrak{n}_+ v$, the image of v under repeated application of various e_is. By finite-dimensionality, $\mathfrak{n}_+ v$ must contain a vector $l \in \mathfrak{n}_+ v$ so that $e_i l = 0$ for every i. By the $\mathfrak{sl}(2)$ representation theory, l must be homogeneous, and indeed a top-weight vector of L, and by the irreducibility l generates L. Let the weight of l be λ; then the map $v_\lambda \to l$ generates a map $M_\lambda \twoheadrightarrow L$. But M_λ has a unique maximal submodule, and since L is irreducible, this maximal submodule must be the kernel of the map $M_\lambda \twoheadrightarrow L$. Thus $L \cong L_\lambda$. □

6.1.1 Weyl Character Formula

In this section we compute the characters of the irreducible representations of a semisimple Lie algebra.

6.1.1.1 Lemma / Definition *Let $\mathfrak{g}$ be a semisimple Lie algebra, $\mathfrak{h}$ its Cartan subalgebra, and $\Delta = \{\alpha_1, \dots, \alpha_n\}$ its simple root system. For each $i = 1, \dots, n$, we define a* fundamental weight *$\Lambda_i \in \mathfrak{h}^*$ by $\langle \Lambda_i, \alpha_j^\vee \rangle = \delta_{ij}$. Then $P_+ = \mathbb{Z}_{\geq 0}\{\Lambda_1, \dots, \Lambda_n\}$.*

The following are equivalent, and define the Weyl vector *ρ:*

1. $\rho = \sum_{i=1}^n \Lambda_i$. *I.e.* $\langle \rho, \alpha_j^\vee \rangle = 1$ *for every* j.

2. $\rho = \frac{1}{2} \sum_{\alpha \in R_+} \alpha$.

Proof Let $\rho_2 = \frac{1}{2} \sum_{\alpha \in R_+} \alpha$. Since $s_i(R_+ \smallsetminus \{\alpha_i\}) = R_+$ but $s_i(\alpha_i) = -\alpha_i$, we see that $s_i(\tilde\rho) = \tilde\rho - \alpha_i$, and so $\langle \tilde\rho, \alpha_i^\vee \rangle = 1$ for every i. The rest is elementary linear algebra. □

6.1.1.2 Theorem (Weyl Character Formula)
Let $\mathrm{sign} : W \to \{\pm 1\}$ *be given by* $\mathrm{sign}(w) = \det_{\mathfrak{h}} w$*; i.e.* sign *is the group homomorphism generated by* $s_i \mapsto -1$ *for each* i*. Let* $\lambda \in P_+$*. Then:*

$$\mathrm{ch}(L_\lambda) = \frac{\sum_{w \in W} \mathrm{sign}(w)\, x^{w(\lambda+\rho)-\rho}}{\prod_{\alpha \in R_+} (1 - x^{-\alpha})} = \frac{\sum_{w \in W} \mathrm{sign}(w)\, x^{w(\lambda+\rho)}}{\prod_{\alpha \in R_+} (x^{\alpha/2} - x^{-\alpha/2})} \tag{6.1.1.3}$$

The equality of fractions follows simply from the description $\rho = \frac{1}{2} \sum_{\alpha \in R_+} \alpha$.

6.1.1.4 Remark The sum in equation (6.1.1.3) is finite. Indeed, the numerator and denominator on the right-hand-side fraction are obiously antisymmetric in W, and so the whole expression is W-invariant. The numerator on the left-hand-side fraction is a polynomial, and each $(1 - x^{-\alpha})$ is invertible as a power series: $(1 - x^{-\alpha})^{-1} = \sum_{n=0}^{\infty} x^{-n\alpha}$. So the fraction is a W-invariant power series, and hence a polynomial. $\diamondsuit$

To prove Theorem 6.1.1.2 we will need a number of lemmas. In Example 6.1.0.11 we computed the character of the Verma module M_λ. Then Theorem 6.1.1.2 asserts:

$$\mathrm{ch}(L_\lambda) = \sum_{w \in W} \mathrm{sign}(w)\ \mathrm{ch}\big(M_{w(\lambda+\rho)-\rho}\big) \tag{6.1.1.5}$$

As such, we will begin by understanding M_λ better. We recall Lemma/Definition 4.4.1.3: Let $(,)$ be the Killing form on $\mathfrak{g}$, and $\{x_i\}$ any basis of $\mathfrak{g}$ with dual basis $\{y_j\}$, i.e. $(x_i, y_j) = \delta_{ij}$ for every i, j; then $c = \sum x_i y_i \in \mathcal{U}\mathfrak{g}$ is central, and does not depend on the choice of basis.

6.1.1.6 Lemma *Let* $\lambda \in \mathfrak{h}^*$ *and* M_λ *the Verma module with weight* λ*. Let* $c \in \mathcal{U}\mathfrak{g}$ *be the Casimir, corresponding to the Killing form on* $\mathfrak{g}$*. Then* c *acts on* M_λ *by multiplication by* $(\lambda, \lambda + 2\rho)$*.*

Proof Let $\mathfrak{g}$ have rank n. Write R for the set of roots of $\mathfrak{g}$, R_+ for the positive roots, and Δ for the simple roots, as we have previously.

Recall Lemma 5.3.3.1. We construct a basis of $\mathfrak{g}$ as follows: we pick an orthonormal basis $\{u_i\}_{i=1}^n$ of $\mathfrak{h}$. For each α a non-zero root of $\mathfrak{g}$, the space $\mathfrak{g}_\alpha$ is one-dimensional; let x_α be a basis vector in $\mathfrak{g}_\alpha$. Then the dual basis to $\{u_i\}_{i=1}^n \cup \{x_\alpha\}_{\alpha \in R \smallsetminus \{0\}}$ is $\{u_i\}_{i=1}^n \cup \{y_\alpha\}_{\alpha \in R \smallsetminus \{0\}}$, where $y_\alpha = \frac{x_{-\alpha}}{(x_\alpha, x_{-\alpha})}$. Then:

$$c = \sum_{i=1}^n u_i^2 + \sum_{\alpha \in R \smallsetminus \{0\}} x_\alpha y_\alpha = \sum_{i=1}^n u_i^2 + \sum_{\alpha \in R \smallsetminus \{0\}} \frac{x_\alpha x_{-\alpha}}{(x_\alpha, x_{-\alpha})} = \sum_{i=1}^n u_i^2 + \sum_{\alpha \in R_+} \frac{x_\alpha x_{-\alpha} + x_{-\alpha} x_\alpha}{(x_\alpha, x_{-\alpha})}$$

Since M_λ is generated by its highest weight vector v_λ, and c is central, to understand the action of c on M_λ it suffices to compute $c v_\lambda$. We use the fact that for $\alpha \in R_+$, $x_\alpha v_\lambda = 0$; then

$$x_\alpha x_{-\alpha} v_\lambda = h_\alpha v_\lambda + x_{-\alpha} x_\alpha v_\lambda = h_\alpha v_\lambda = \lambda(h_\alpha)\, v_\lambda$$

where for each $\alpha \in R_+$ we have defines $h_\alpha \in \mathfrak{h}$ by $h_\alpha = [x_\alpha, x_{-\alpha}]$. Moreover, $(,)$ is $\mathfrak{g}$-invariant, and $[h_\alpha, x_\alpha] = \alpha(h_\alpha)x_\alpha$, where $\alpha(h_\alpha) \neq 0$. So:

$$(x_\alpha, x_{-\alpha}) = \frac{1}{\alpha(h_\alpha)}\big([h_\alpha, x_\alpha], x_{-\alpha}\big) = \frac{1}{\alpha(h_\alpha)}\big(h_\alpha, [x_\alpha, x_{-\alpha}]\big) = \frac{(h_\alpha, h_\alpha)}{\alpha(h_\alpha)}$$

We also have that $u_i v_\lambda = \lambda(u_i)v_\lambda$, and since $\{u_i\}$ is an orthonormal basis, $(\lambda,\lambda) = \sum_{i=1}^n \big(\lambda(u_i)\big)^2$. Thus:

$$cv_\lambda = \sum_{i=1}^n \big(\lambda(u_i)\big)^2 v_\lambda + \sum_{\alpha\in R_+} \frac{\lambda(h_\alpha)}{\frac{(h_\alpha,h_\alpha)}{\alpha(h_\alpha)}} v_\lambda = \left((\lambda,\lambda) + \sum_{\alpha\in R_+} \frac{\lambda(h_\alpha)\,\alpha(h_\alpha)}{(h_\alpha,h_\alpha)} \right) v_\lambda$$

We recall that h_α is proportional to $\alpha^\vee$, that $(\alpha,\alpha) = 4/(\alpha^\vee,\alpha^\vee)$, and that $\lambda(\alpha^\vee) = (\lambda,\alpha)/(\alpha,\alpha)$. Then $\frac{\lambda(h_\alpha)\,\alpha(h_\alpha)}{(h_\alpha,h_\alpha)} = (\lambda,\alpha)$, and so:

$$(\lambda,\lambda) + \sum_{\alpha\in R_+} \frac{\lambda(h_\alpha)\,\alpha(h_\alpha)}{(h_\alpha,h_\alpha)} = (\lambda,\lambda) + \sum_{\alpha\in R_+} (\lambda,\alpha) = (\lambda,\lambda+2\rho)$$

Thus c acts on M_λ by multiplication by $(\lambda,\lambda+2\rho)$. □

6.1.1.7 Lemma / Definition *Let X be a $\mathfrak{g}$-module. A weight vector $v \in X$ is* singular *if $\mathfrak{n}_+ v = 0$. In particular, any highest-weight vector is singular, and conversely any singular vector is the highest weight vector in the submodule it generates.* □

6.1.1.8 Corollary *Let $\lambda \in P$, and M_λ the Verma module with weight λ. Then M_λ contains finitely many singular vectors, in the sense that their span is finite-dimensional.*

Proof Let C^λ be the set $C^\lambda \overset{\text{def}}{=} \{\mu \in P \text{ s.t. } (\mu+\rho,\mu+\rho) = (\lambda+\rho,\lambda+\rho)\}$. Then C^λ is a sphere in P centered at $-\rho$, and in particular it is a finite set. On the other hand, since $(\mu+\rho,\mu+\rho) = (\mu,\mu+2\rho) + (\rho,\rho)$, we see that:

$$C^\lambda = \{\mu \in P \text{ s.t. } c \text{ acts on } M_\mu \text{ by } (\lambda,\lambda+2\rho)\}$$

Recall that any module with highest weight μ is a quotient of M_μ. Let $v \in M_\lambda$ be a non-zero singular vector with weight μ. Then on the one hand $cv = (\lambda,\lambda+2\rho)v$, since $v \in M_\lambda$, and on the other hand $cv = (\mu,\mu+2\rho)$, since v is in a quotient of M_μ. In particular, $\mu \in C^\lambda$. But the weight spaces $(M_\lambda)_\mu$ of M_λ are finite-dimensional, and so the dimension of the space of singular vectors is at most $\sum_{\mu\in C^\lambda} \dim\big((M_\lambda)_\mu\big) < \infty$. □

6.1.1.9 Corollary *Let $\lambda \in P$. Then there are nonnegative integers $b_{\lambda,\mu}$ such that*

$$\operatorname{ch} M_\lambda = \sum b_{\lambda,\mu} \operatorname{ch} L_\mu \tag{6.1.1.10}$$

and $b_{\lambda,\mu} = 0$ unless $\mu \le \lambda$ and $\mu \in C^\lambda$. Moreover, $b_{\lambda,\lambda} = 0$.

Proof We construct a filtration on M_λ. Since M_λ has only finitely many non-zero singular vectors, we choose w_1 a singular vector of minimal weight μ_1, and let $F_1 M_\lambda$ be the submodule of M_λ generated by w_1. Then $F_1 M_\lambda$ is irreducible with highest weight μ_1. We proceed by induction, letting w_i be a singular vector of minimal weight in $M_\lambda / F_{i-1} M_\lambda$, and $F_i M_\lambda$ the primage of the subrepresentation generated by w_i. This filters M_λ:

$$0 = F_0 M_\lambda \subseteq F_1 M_\lambda \subseteq \dots \tag{6.1.1.11}$$

Moreover, since M_λ has only finitely many weight vectors all together, the filtration must terminate:

$$0 = F_0 M_\lambda \subseteq F_1 M_\lambda \subseteq \dots \subseteq F_k M_\lambda = M_\lambda$$

By construction, the quotients are all irreducible: $F_i M_\lambda / F_{i-1} M_\lambda = L_{\mu_i}$ for some $\mu_i \in C^\lambda$, $\mu_i \leq \lambda$.

We recall that ch is additive for extensions. Therefore

$$\operatorname{ch} M_\lambda = \sum_{i=1}^{k} \operatorname{ch}(F_i M_\lambda / F_{i-1} M_\lambda) = \sum_{i=1}^{k} \operatorname{ch} L_{\mu_i}$$

Then $b_{\lambda,\mu}$ is the multiplicity of μ appearing as the weight of a singular vector of M_λ, and we have equation (6.1.1.10). The conditions stated about $b_{\lambda,\mu}$ are immediate: we saw that μ can only appear as a weight of M_λ if $\mu \in C^\lambda$ and $\mu \leq \lambda$; moreover, L_λ appears as a subquotient of M_λ exactly once, so $b_{\lambda,\lambda} = 1$. □

6.1.1.12 Definition *The coefficients $b_{\lambda,\mu}$ in equation (6.1.1.10) are the* Kazhdan–Luztig multiplicities.

6.1.1.13 Lemma *If $\lambda \in P_+$, $\mu \leq \lambda$, $\mu \in C^\lambda$, and $\mu + \rho \geq 0$, then $\mu = \lambda$.*

Proof We have that $(\mu + \rho, \mu + \rho) = (\lambda + \rho, \lambda + \rho)$ and that $\lambda - \mu = \sum_{i=1}^n k_i \alpha_i$, where all k_i are nonnegative. Then

$$\begin{aligned} 0 &= (\lambda + \rho, \lambda + \rho) - (\mu + \rho, \mu + \rho) \\ &= \big((\lambda + \rho) - (\mu + \rho), (\lambda + \rho) + (\mu + \rho)\big) \\ &= (\lambda - \mu, \lambda + \mu + 2\rho) \\ &= \sum_{i=1}^{n} k_i (\alpha, \lambda + \mu + 2\rho) \end{aligned}$$

But $\lambda, \mu + \rho \geq 0$, and $(\alpha_i, \rho) > 0$, so $(\alpha, \lambda + \mu + 2\rho) > 0$, and so all $k_i = 0$ since they are nonnegative. □

Proof (of Theorem 6.1.1.2) We have shown (Corollary 6.1.1.9) that $\operatorname{ch} M_\lambda = \sum b_{\lambda,\mu} \operatorname{ch} L_\mu$, where $b_{\lambda,\mu}$ is a lower-triangular matrix on $C^\lambda = C^\mu$ with 1s on the diagonal. Thus it has a lower-triangular inverse with 1s on the diagonal:

$$\operatorname{ch} L_\lambda = \sum_{\mu \leq \lambda, \mu \in C^\lambda} c_{\lambda,\mu} \operatorname{ch} M_\mu$$

But by Proposition 6.1.0.12 statement 1., $\operatorname{ch} L_\lambda$ is W-invariant, provided that $\lambda \in P_+$, thus so is $\sum c_{\lambda,\mu} \operatorname{ch} M_\mu$. We recall Example 6.1.0.11:

$$\operatorname{ch} M_\mu = \frac{x^\mu}{\prod_{\alpha \in R_+}(1 - x^{-\alpha})} = \frac{x^{\mu+\rho}}{\prod_{\alpha \in R_+}(x^{\alpha/2} - x^{-\alpha/2})}$$

Therefore

$$\operatorname{ch} L_\lambda = \frac{\sum_{\mu \leq \lambda, \mu \in C^\lambda} c_{\lambda,\mu} x^{\mu+\rho}}{\prod_{\alpha \in R_+}(x^{\alpha/2} - x^{-\alpha/2})}$$

But the denominator if W-antisymmetric, and so the numerator must be as well:

$$\sum_{\mu \leq \lambda, \mu \in C^\lambda} c_{\lambda,\mu} x^{w(\mu+\rho)} = \sum_{\mu \leq \lambda, \mu \in C^\lambda} \operatorname{sign}(w) c_{\lambda,\mu} x^{\mu+\rho} \text{ for every } w \in W$$

This is equivalent to the condition that $c_{\lambda,\mu} = \operatorname{sign}(w) c_{\lambda, w(\mu+\rho)-\rho}$. By the proof of Proposition 6.1.0.12 statement 2., we know that P_+ is a fundamental domain of W; since $c_{\lambda,\lambda} = 1$, if $\mu + \rho \in W(\lambda + \rho)$, then $c_{\lambda,\mu} = \operatorname{sign}(w)$, and so:

$$\sum_{\mu \leq \lambda, \mu \in C^\lambda} c_{\lambda,\mu} x^{\mu+\rho} = \sum_{w \in W} \left(x^{w(\lambda+\rho)} + \sum_{\substack{\mu < \lambda, \mu \in C^\lambda \\ \mu+\rho \in P^+}} c_{\lambda,\mu} x^{w(\mu+\rho)} \right)$$

But the rightmost sum is empty by Lemma 6.1.1.13. □

6.1.1.14 Remark Specializing to the trivial representation L_0, Theorem 6.1.1.2 says that

$$1 = \frac{\sum_{w \in W} \operatorname{sign}(w) x^{w(\rho)}}{\prod_{\alpha \in R_+}(x^{\alpha/2} - x^{-\alpha/2})} \tag{6.1.1.15}$$

So we can rewrite equation (6.1.1.3) as

$$\chi^\lambda = \frac{\sum_{w \in W} \operatorname{sign}(w) x^{w(\lambda+\rho)}}{\sum_{w \in W} \operatorname{sign}(w) x^{w(\rho)}}$$

◊

The following is an important corollary:

6.1.1.16 Theorem (Weyl Dimension Formula)
Let $\lambda \in P_+$. Then $\dim L_\lambda = \prod_{\alpha \in R_+} (\alpha, \lambda + \rho)/(\alpha, \rho)$.

Proof The formula $\operatorname{ch}(L_\lambda) = \frac{\sum_{w \in W} \operatorname{sign}(w)\, x^{w(\lambda+\rho)}}{\prod_{\alpha \in R_+}(x^{\alpha/2} - x^{-\alpha/2})}$ is a polynomial in x. In particular, it defines a real-valued function on $\mathbb{R}_{>0} \times \mathfrak{h}$ given by $x^\alpha \mapsto a^{\alpha(h)}$ — when $a = 1$ or $h = 0$, the formula as written is the indeterminate form $\frac{0}{0}$, but the function clearly returns $\sum_\mu \dim\big((L_\lambda)_\mu\big) = \dim L_\lambda$. We will calculate this value of the function by taking a limit, using l'Hôpital's rule.

In particular, letting $x^\alpha \mapsto e^{t(\alpha,\lambda+\rho)}$ in equation (6.1.1.15) gives

$$\prod_{\alpha\in R_+}\left(e^{t(\alpha/2,\lambda+\rho)} - e^{-t(\alpha/2,\lambda+\rho)}\right) = \sum_{w\in W}\operatorname{sign}(w)e^{t(w(\rho),\lambda+\rho)} = \sum_{w\in W}\operatorname{sign}(w)e^{t(\rho,w(\lambda+\rho))}$$

where the second equality comes from $w \mapsto w^{-1}$ and $(w^{-1}x,y) = (x,wy)$. On the other hand, we let $x^\alpha \mapsto e^{t(\alpha,\rho)}$ in equation (6.1.1.3). Then

$$\operatorname{ch} L_\lambda|_{x=e^{t\rho}} = \frac{\sum_{w\in W}\operatorname{sign}(w)e^{t(w(\lambda+\rho),\rho)}}{\prod_{\alpha\in R_+}\left(e^{t(\alpha/2,\rho)} - e^{-t(\alpha/2,\rho)}\right)} =$$
$$= \frac{\prod_{\alpha\in R_+}\left(e^{t(\alpha/2,\lambda+\rho)} - e^{-t(\alpha/2,\lambda+\rho)}\right)}{\prod_{\alpha\in R_+}\left(e^{t(\alpha/2,\rho)} - e^{-t(\alpha/2,\rho)}\right)} = \prod_{\alpha\in R_+}\frac{\left(e^{t(\alpha/2,\lambda+\rho)} - e^{-t(\alpha/2,\lambda+\rho)}\right)}{\left(e^{t(\alpha/2,\rho)} - e^{-t(\alpha/2,\rho)}\right)}$$

Therefore:

$$\dim L_\lambda = \lim_{t\to 0}\prod_{\alpha\in R_+}\frac{\left(e^{t(\alpha/2,\lambda+\rho)} - e^{-t(\alpha/2,\lambda+\rho)}\right)}{\left(e^{t(\alpha/2,\rho)} - e^{-t(\alpha/2,\rho)}\right)} = \prod_{\alpha\in R_+}\lim_{t\to 0}\frac{\left(e^{t(\alpha/2,\lambda+\rho)} - e^{-t(\alpha/2,\lambda+\rho)}\right)}{\left(e^{t(\alpha/2,\rho)} - e^{-t(\alpha/2,\rho)}\right)} \overset{\text{l'H}}{=}$$
$$\overset{\text{l'H}}{=} \prod_{\alpha\in R_+}\lim_{t\to 0}\frac{\left((\alpha/2,\lambda+\rho)e^{t(\alpha/2,\lambda+\rho)} + (\alpha/2,\lambda+\rho)e^{-t(\alpha/2,\lambda+\rho)}\right)}{\left((\alpha/2,\rho)e^{t(\alpha/2,\rho)} + (\alpha/2,\rho)e^{-t(\alpha/2,\rho)}\right)} = \prod_{\alpha\in R_+}\frac{(\alpha,\lambda+\rho)}{(\alpha,\rho)} \qquad \square$$

6.1.1.17 Example Let us compute the dimensions of the irreducible representations of $\mathfrak{g} = \mathfrak{sl}(n+1)$. We work with the standard the simple roots be $\Delta = \{\alpha_1,\dots,\alpha_n\}$, whence $R_+ = \{\alpha_i + \alpha_{i+1} + \dots + \alpha_j\}_{1\le i<j\le n}$. Let us write λ and ρ in terms of the fundamental weights Λ_i, defined by $(\Lambda_i,\alpha_j) = \delta_{ij}$: $\rho = \sum_{i=1}^n \Lambda_i$ and $\lambda+\rho = \sum_{i=1}^n a_i\Lambda_i$. Then:

$$\dim L_\lambda = \prod_{\alpha\in R_+}\frac{(\lambda+\rho,\alpha)}{(\rho,\alpha)} = \prod_{1\le i\le j\le n}\frac{a_i + a_{i+1} + \dots + a_{j-1} + a_j}{j-i+1} = \frac{1}{n!!}\prod_{1\le i\le j\le n}\sum_{k=i}^{j} a_k$$

where we have defined $n!! \overset{\text{def}}{=} n!\,(n-1)!\cdots 3!\,2!\,1!$. For example, the irrep of $\mathfrak{sl}(3)$ with weight $\lambda+\rho = 3\Lambda_1 + 2\Lambda_2$ has dimension $\frac{1}{2!!}2\cdot 3\cdot(2+3) = 15$. $\diamondsuit$

6.1.2 Some applications of the Weyl Character Formula

Let λ be a positive weight. Recall equation (6.1.1.5):

$$\operatorname{ch} L(\lambda) = \sum_{w\in W}\operatorname{sign}(w)\operatorname{ch} M\big(w(\lambda+\rho)-\rho\big)$$

6.1.2.1 Definition *The* Kostant partition function *measures the number of ways to write* λ *as a sum of positive roots:*

$$\mathcal{P}(\lambda) = \#\big\{\{m_\alpha\}\in\mathbb{Z}_{\ge 0} \text{ s.t. } \lambda = \textstyle\sum_{\alpha\in\Delta^+} m_\alpha\alpha\big\}$$

The continuous Kostant partition function *is the following piecewise polynomial:*

$$\mathcal{P}_{cont}(\gamma) = \mathrm{Vol}\big\{m_\alpha \in \mathbb{R}_{\geq 0} \ s.t.\ \gamma = \sum m_\alpha \alpha\big\}$$

Then the following is clear:

6.1.2.2 Lemma $\dim M(\lambda)_\mu = \mathcal{P}(\lambda - \mu)$, *so that* $\mathrm{ch}\, M(\lambda) = \sum_\mu \mathcal{P}(\lambda-\mu)x^\mu$. □

6.1.2.3 Example The dimensions of the weight spaces for the Verma module for $\mathfrak{sl}(3)$ are:

$$\begin{array}{ccccccccccc}
1 & & 1 & & 1 & & 1 & & & & \\
& 2 & & 2 & & 2 & & 1 & & & \\
3 & & 3 & & 3 & & 2 & & 1 & & \\
& 4 & & 4 & & 3 & & 2 & & 1 &
\end{array}$$

◊

Then we can calculate the dimension at weight μ in $L(\lambda)$ as the alternating sum of partition functions:

$$\mathcal{M}(\lambda,\mu) = \dim L(\lambda)_\mu = \sum_{w\in W} \mathrm{sign}(w)\, \mathcal{P}\big(w(\lambda+\rho) - \mu - \rho\big)$$

In particular, on the boundary all multiplicities are 1. Unfortunately, the formula is not very useful in applications, because the order of the Weyl group is very big. See [FH91] for more details.

6.1.2.4 Remark Any alternating sum of dimensions ought to come from a complex, and equation (6.1.1.5) is no exception. In fact, in the category $\mathcal{O}$ there is a resolution of each irreducible such that each term is a direct sum of Verma modules, the *BGG resolution*:

$$0 \to M\big(w_0(\lambda+\rho)-\rho\big) \to \cdots \to \bigoplus_{w\in W \text{ s.t. } \ell(w)=k} M\big(w(\lambda+\rho)-\rho\big) \to \cdots \to$$
$$\to \bigoplus M\big(s_i(\lambda+\rho)-\rho\big) \to M(\lambda) \to L(\lambda) \to 0$$

Here w_0 is the longest element of W; it maps positive roots to negative roots. By definition, $\mathrm{sign}(w) = (-1)^{\ell(w)}$, so equation (6.1.1.5) has some nice algebra behind it. ◊

We will conclude this section with some applications to the tensor product in the category of finite-dimensional $\mathfrak{g}$-modules. Since $\mathfrak{g}$ is semisimple, there must be "fusion" coefficients $\Gamma^\mu_{\lambda,\mu}$ such that:

$$L(\lambda)\otimes L(\mu) = \bigoplus_\nu L(\nu)^{\oplus \Gamma^\nu_{\lambda,\mu}} \tag{6.1.2.5}$$

The question is to find a formula for the $\Gamma^\nu_{\lambda,\mu}$s.

Recall from the theory of finite groups that the characters are orthogonal. Similarly, we have an orthogonality condition for formal characters. Let R denote the ring $R = \bigoplus_{\lambda\in P}\mathbb{Z}[x^\lambda]$. It carries a W-action, and we denote the W-invariant subring by R^W.

6.1.2.6 Proposition *The characters* $\operatorname{ch} L(\lambda)$ *form a basis in* R^W*, which is orthonormal with respect to the pairing* $(,)$ *given by:*

$$(\phi, \psi) = \textit{constant coefficient of } \frac{\mathcal{D}\bar{\mathcal{D}}}{|w|}\phi\bar{\psi}$$

where $\overline{x^\lambda} = x^{-\lambda}$*, so that* $\overline{\operatorname{ch} L(\lambda)} = \operatorname{ch} L(\lambda)^*$*, and* $\mathcal{D} \stackrel{\text{def}}{=} \sum \operatorname{sign}(w)\, x^{w(\rho)}$*.*

6.1.2.7 Remark We describe the geometric meaning of Proposition 6.1.2.6. We will discuss in Section 7.2 the compact forms of semisimple groups, but for now let's restrict to $G = \mathrm{GL}(n, \mathbb{C})$ and $K = \mathrm{U}(n)$. Let $T = K \cap H$ be the maximal torus. Then we know from linear algebra that every element of K is conjugate to something in T. Let's identify x^λ with $h \mapsto \exp(\lambda(h))$. If we scale things correctly, this depends only on $\exp(h) \in H$, and agrees with $\operatorname{tr}_{L(\lambda)} \exp(h)$. Thus the characters give class functions, and:

$$(\phi, \psi) = \int_K \phi\, \bar{\psi}\, dg = \frac{1}{|W|} \int_T \phi\, \bar{\psi}\ \mathrm{Vol}_t\ dt$$

Here Vol_t is the volume of the conjugacy class of $t \in T$ in K. The $1/|W|$ counts the redundancy of how we diagonalize unitary matrices. So the idea is that $\mathcal{D}\bar{\mathcal{D}}(h) = \mathrm{Vol}_{\exp h}$. ◊

Proof (of Proposition 6.1.2.6) An improvement of the discussion in the previous remark explains why the $\operatorname{ch} L(\lambda)$ are orthonormal. We will check that they are a basis. There is another obvious basis of R^W: each W-orbit intersects P^+ once, so for each $\lambda \in P^+$, set $E_\lambda = c_\lambda \sum_{w\in W} x^{w(\lambda)}$, where c_λ is some coefficient so that $(E_\lambda, E_\lambda) = 1$. Let $d_{\mu,\lambda}$ be coefficients so that $\operatorname{ch}(L(\lambda)) = \sum_\mu d_{\mu,\lambda} E_\mu$. Then it's clear that d is a lower-triangular matrix with 1s on the diagonal, and hence invertible. Therefore $\operatorname{ch} L(\lambda)$ is a basis. □

We can now calculate the fusion coefficients Γ in equation (6.1.2.5). We have:

$$\begin{aligned}
\Gamma^\nu_{\lambda,\mu} &= \big(\operatorname{ch} L(\lambda)\ \operatorname{ch} L(\mu), \operatorname{ch} L(\nu)\big) = \\
&= \text{constant coef of } \frac{1}{|W|}\frac{1}{\mathcal{D}} \sum_{w,u,v\in W} \operatorname{sign}(w)\, x^{w(\lambda+\rho)}\ \operatorname{sign}(u)\, x^{u(\mu+\rho)}\ \operatorname{sign}(v)\, x^{-v(\nu+\rho)} = \\
&= \text{const coef of} \frac{1}{\mathcal{D}} \sum_{w,\sigma\in W} \operatorname{sign}(w)\, x^{w(\lambda+\rho)}\ \operatorname{sign}(\sigma)\, x^{\sigma(\mu+\rho)-(\nu+\rho)} = \\
&= \sum_{\sigma\in W} \operatorname{sign}(\sigma)\, \mathcal{M}\big(\lambda, \nu+\rho-\sigma(\mu+\rho)\big) = \\
&= \sum_{\sigma,w\in W} \operatorname{sign}(\sigma w) \mathcal{P}\big(w(\lambda+\rho)+\sigma(\mu+\rho)-\nu-2\rho\big) \qquad (6.1.2.8)
\end{aligned}$$

where we substituted $\sigma = uv^{-1}$, and used various facts. This is the *Steinberg formula*. Unfortunately, this is still not very effective for actual calculations. Much better is the Littlewood-Richardson rule, but that only works for $\mathfrak{gl}(n)$.

6.2 Algebraic Lie groups

We have classified the representations of any semisimple Lie algebra, and therefore the representations of its simply connected Lie group. But a Lie algebra corresponds to many (connected) Lie groups, quotients of the simply connected group by (necessarily central) discrete subgroups, and a representation of the Lie algebra is a representation of one of these groups only if the corresponding discrete normal subgroup acts trivially in the representation. We will see that the simply connected Lie group of any semisimple Lie algebra is algebraic, and that its algebraic quotients are determined by the finite-dimensional representation theory of the Lie algebra.

6.2.1 Guiding example: $\mathrm{SL}(n)$ and $\mathrm{PSL}(n)$

Our primary example, as always, is the Lie algebra $\mathfrak{sl}(2,\mathbb{C})$, consisting of traceless 2×2 complex matrices. It is the Lie algebra of $\mathrm{SL}(2,\mathbb{C})$, the group of 2×2 complex matrices with determinant 1.

6.2.1.1 Lemma / Definition *The group* $\mathrm{SL}(2,\mathbb{C})$ *has a non-trivial center:* $Z(\mathrm{SL}(2,\mathbb{C})) = \{\pm 1\}$. *We define the* projective special linear group *to be* $\mathrm{PSL}(2,\mathbb{C}) \overset{\text{def}}{=} \mathrm{SL}(2,\mathbb{C})/\{\pm 1\}$. *Equivalently,* $\mathrm{PSL}(2,\mathbb{C}) = \mathrm{PGL}(2,\mathbb{C}) \overset{\text{def}}{=} \mathrm{GL}(2,\mathbb{C})/\{\textit{scalars}\}$, *the* projective general linear group. □

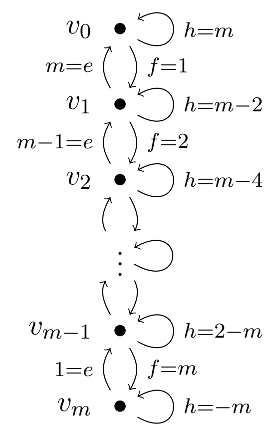
6.2.1.2 Proposition *The group* $\mathrm{SL}(2,\mathbb{C})$ *is connected and simply connected. The kernel of the map* $\mathrm{ad} : \mathrm{SL}(2,\mathbb{C}) \to \mathrm{GL}(\mathfrak{sl}(2,\mathbb{C}))$ *is precisely the center, and so* $\mathrm{PSL}(2,\mathbb{C})$ *is the connected component of the group of automoraphisms of* $\mathfrak{sl}(2,\mathbb{C})$. *The groups* $\mathrm{SL}(2,\mathbb{C})$ *and* $\mathrm{PSL}(2,\mathbb{C})$ *are the only connected Lie groups with Lie algebra* $\mathfrak{sl}(2,\mathbb{C})$.

Proof The only nontrivial statement is that $\mathrm{SL}(2,\mathbb{C})$ is simply connected. Consider the subgroup $U \overset{\text{def}}{=} \left\{\begin{bmatrix}1 & * \\ 0 & 1\end{bmatrix} \in \mathrm{SL}(2,\mathbb{C})\right\}$. Then U is the stabilizer of the vector $\begin{bmatrix}1\\0\end{bmatrix} \in \mathbb{C}^2 \smallsetminus \{0\}$, and $\mathrm{SL}(2,\mathbb{C})$ acts transitively on $\mathbb{C}^2 \smallsetminus \{0\}$. Thus the space of left cosets $\mathrm{SL}(2,\mathbb{C})/U$ is isomorphic to the space $\mathbb{C}^2 \smallsetminus \{0\} \cong \mathbb{R}^4 \smallsetminus \{0\}$ as a real manifold. But $U \cong \mathbb{C}$, so $\mathrm{SL}(2,\mathbb{C})$ is connected and simply connected. □

6.2.1.3 Lemma *The groups* $\mathrm{SL}(2,\mathbb{C})$ *and* $\mathrm{PSL}(2,\mathbb{C})$ *are algebraic.*

Proof The determinant of a matrix is a polynomial in the coefficients, so $\{x \in \mathrm{M}(2,\mathbb{C}) \text{ s.t. } \det x = 1\}$ is an algebraic group. Any automorphism of $\mathfrak{sl}(2,\mathbb{C})$ preserves the Killing form, a nondegenerate symmetric pairing on the three-dimensional vector space $\mathfrak{sl}(2,\mathbb{C})$. Thus $\mathrm{PSL}(2,\mathbb{C})$ is a subgroup of $\mathrm{O}(3,\mathbb{C})$. It is connected, and so a subgroup of $\mathrm{SO}(3,\mathbb{C})$, and three-dimensional, and so is all of $\mathrm{SO}(3,\mathbb{C})$. Moreover, $\mathrm{SO}(3,\mathbb{C})$ is algebraic: it consists of matrices $x \in \mathrm{M}(3,\mathbb{C})$ that preserve the nondegenerate form (a system of quadratic equations in the coefficients) and have unit determinant (a cubic equation in the coefficients). □

Recall that any irreducible representation of $\mathfrak{sl}(2,\mathbb{C})$ looks like a chain: e moves up the chain, f down, and h acts diagonally with eigenvalues changing by 2 from m at the top to $-m$ at the bottom:



The exponential map $\exp : \mathfrak{sl}(2,\mathbb{C}) \to \mathrm{SL}(2,\mathbb{C})$ acts on the Cartan by $th = \begin{bmatrix} t & \\ & -t \end{bmatrix} \mapsto \begin{bmatrix} e^t & \\ & e^{-t} \end{bmatrix}$. Let $T = \exp(\mathfrak{h})$; then the kernel of $\exp : \mathfrak{h} \to T$ is $2\pi i \mathbb{Z} h$. On the other hand, when $t = \pi i$, $\exp(th) = -1$, which maps to 1 under $\mathrm{SL}(2,\mathbb{C}) \twoheadrightarrow \mathrm{PSL}(2,\mathbb{C})$; therefore the kernel of the exponential map $\mathfrak{h} \to \mathrm{PSL}(2,\mathbb{C})$ is just $\pi i \mathbb{Z} h$.

In particular, the $(m+1)$-dimensional representation V_m of $\mathfrak{sl}(2,\mathbb{C})$ is a representation of $\mathrm{PSL}(2,\mathbb{C})$ if and only if m is even, because $-1 \in \mathrm{SL}(2,\mathbb{C})$ acts on V_m as $(-1)^m$. We remark that $\ker\{\exp : \mathfrak{h} \to \mathrm{SL}(2,\mathbb{C})\}$ is precisely $2\pi i Q^\vee$, where $Q^\vee$ is the coroot lattice of $\mathfrak{sl}(2)$, and $\ker\{\exp : \mathfrak{h} \to \mathrm{PSL}(2,\mathbb{C})\}$ is precisely the coweight lattice $2\pi i P^\vee$

6.2.1.4 Remark This will be the model for any semisimple Lie algebra $\mathfrak{g}$ with Cartan subalgebra $\mathfrak{h}$. We will understand the exponential map from $\mathfrak{h}$ to the simply connected Lie group G corresponding to $\mathfrak{g}$, and we will also understand the map to $G/Z(G)$, the simplest quotient. Every group with Lie algebra $\mathfrak{g}$ is a quotient of G, and hence lies between G and $G/Z(G)$. The kernels of the maps $\mathfrak{h} \to G$ and $\mathfrak{h} \to G/Z(G)$ will be precisely $2\pi i Q^\vee$ and $2\pi i P^\vee$, respectively, and every other group will correspond to a lattice between these two. $\lozenge$

Let us consider one further example: $\mathrm{SL}(n,\mathbb{C})$. It is simply-connected, and its center is $Z(\mathrm{SL}(n,\mathbb{C})) = \{n\text{th roots of unity}\}$. We define the *projective special linear group* to be $\mathrm{PSL}(n,\mathbb{C}) \overset{\text{def}}{=} \mathrm{SL}(n,\mathbb{C})/Z(\mathrm{SL}(n,\mathbb{C}))$; the groups with Lie algebra $\mathfrak{sl}(n,\mathbb{C})$ live between these two, and so correspond to subgroups of $Z(\mathrm{SL}(n,\mathbb{C})) \cong \mathbb{Z}/n$, the cyclic group with n elements.

We now consider the Cartan $\mathfrak{h} \subseteq \mathfrak{sl}(n,\mathbb{C})$, thought of as the space of traceless diagonal matrices: $\mathfrak{h} = \{\langle z_1,\dots,z_n\rangle \in \mathbb{C}^n \text{ s.t. } \sum z_i = 0\}$. In particular, $\mathfrak{sl}(n,\mathbb{C})$ is of A-type, and so we can identify roots and coroots: $\alpha_i = \alpha_i^\vee = \langle 0,\dots,0,1,-1,0,\dots,0\rangle$, where the non-zero terms are in the $(i,i+1)$th spots. Then the coroot lattice $Q^\vee$ is the span of $\alpha_i^\vee$: if $\sum z_i = 0$, then we can write $\langle z_1,\dots,z_n\rangle \in \mathbb{Z}^n$ as $z_1\alpha_1 + (z_1+z_2)\alpha_2 + \dots + (z_1+\dots+z_{n-1})\alpha_{n-1}$, since $z_n = -(z_1+\dots+z_{n-1})$. The coweight lattice $P^\vee$, on the other hand, is the lattice of vectors $\langle z_1,\dots,z_n\rangle$ with $\sum z_i = 0$ and with $z_i - z_{i+1}$ an integer for each $i \in \{1,\dots,n-1\}$. In particular, $\sum z_i = z_1 + \big(z_1 + (z_2 - z_1)\big) + \dots + \big(z_1 + (z_2-z_1) + \dots + (z_n - z_{n-1})\big) = nz_1 + \text{integer}$. Therefore

$z_1 \in \mathbb{Z}\frac{1}{n}$, and $z_i \in z_1 + \mathbb{Z}$. So $P^\vee = Q^\vee \sqcup (\langle \frac{1}{n}, \dots, \frac{1}{n} \rangle + Q^\vee) \sqcup \dots \sqcup (\langle \frac{n-1}{n}, \dots, \frac{n-1}{n} \rangle + Q^\vee)$. In this way, $P^\vee/Q^\vee$ is precisely $\mathbb{Z}/n$, in agreement with the center of $\mathrm{SL}(n, \mathbb{C})$.

6.2.2 Definition and general properties of algebraic groups

We have mentioned already (Definition 1.1.2.1) the notion of an "algebraic group", and we have occasionally used some algebraic geometry (notably in the proof of Theorem 5.3.1.12), but we have not developed that story. We do so now.

6.2.2.1 Definition *A subset $X \subseteq \mathbb{C}^n$ is an* affine variety *if it is the vanishing set of a set $P \subseteq \mathbb{C}[x_1, \dots, x_n]$ of polynomials:*

$$X = V(P) \overset{\text{def}}{=} \{x \in \mathbb{C}^n \text{ s.t. } p(x) = 0 \,\forall p \in P\}$$

Equivalently, X is Zariski closed *(see Definition 5.3.1.13). To any affine variety X we associate an ideal $I(X) \overset{\text{def}}{=} \{p \in \mathbb{C}[x_1, \dots, x_n] \text{ s.t. } p|_X = 0\}$. The* coordinate ring *of, or the* ring of polynomial functions *on, X is the ring $\mathscr{O}(X) \overset{\text{def}}{=} \mathbb{C}[x]/I(X)$.*

6.2.2.2 Lemma *If X is an affine variety, then $I(X)$ is a radical ideal. If $I \subseteq J$, then $V(I) \supseteq V(J)$, and conversely if $X \subseteq Y$ then $I(X) \supseteq I(Y)$. It is clear from the definition that if X is an affine varienty, then $V(I(X)) = X$; more generally, we can define $I(X)$ for any subset $X \supseteq \mathbb{C}^n$, whence $V(I(X))$ is the Zariski closure of X.* □

6.2.2.3 Definition *A* morphism of affine varieties *is a function $f : X \to Y$ such that the coordinates on Y are polynomials in the coordinates of X. Equivalently, any function $f : X \to Y$ gives a homomorphism of algebras $f^\# : \mathrm{Fun}(Y) \to \mathrm{Fun}(X)$, where $\mathrm{Fun}(X)$ is the space of all $\mathbb{C}$-valued functions on X. A function $f : X \to Y$ is a morphism of affine varieties if $f^\#$ restricts to a map $f^\# : \mathscr{O}(Y) \to \mathscr{O}(X)$.*

6.2.2.4 Lemma / Definition *Any point $a \in \mathbb{C}^n$ gives an* evaluation map *$\mathrm{ev}_a : p \mapsto p(a) : \mathbb{C}[x_1, \dots, x_n] \to \mathbb{C}$. If X is an affine variety, then $a \in X$ if and only if $I(X) \subseteq \ker \mathrm{ev}_a$ if and only if $\mathrm{ev}_a : \mathscr{O}(X) \to \mathbb{C}$ is a morphism of affine varieties.* □

6.2.2.5 Corollary *The algebra $\mathscr{O}(X)$ determines the set of evaluation maps $\mathscr{O}(X) \to \mathbb{C}$, and if $\mathscr{O}(X)$ is presented as a quotient of $\mathbb{C}[x_1, \dots, x_n]$, then it determines $X \subseteq \mathbb{C}^n$. A morphism f of affine varieties is determined by the algebra homomorphism $f^\#$ of coordinate rings, and conversely any such algebra homomorphism determines a morphism of affine varieties. Thus the category of affine varieties precisely the opposite category to the category of finitely generated commutative reduced algebras over $\mathbb{C}$.* □

Recall that an algebra is *reduced* if $x^2 = 0$ implies $x = 0$. The condition that $\mathscr{O}(X)$ be reduced is needed: $\mathbb{C}[x]/(x^2)$ is not the coordinate ring of any algebraic variety.

6.2.2.6 Lemma / Definition *The category of affine varieties contains all finite products. The* product *of affine varieties $X \subseteq \mathbb{C}^m$ and $Y \subseteq \mathbb{C}^l$ is $X \times Y \subseteq \mathbb{C}^{m+l}$ with $\mathscr{O}(X \times Y) \cong \mathscr{O}(X) \otimes_{\mathbb{C}} \mathscr{O}(Y)$.*

Proof The maps $\mathscr{O}(X), \mathscr{O}(Y) \to \mathscr{O}(X \times Y)$ are given by the projections $X \times Y \to X, Y$. The map $\mathscr{O}(X) \otimes \mathscr{O}(Y) \to \mathscr{O}(X \times Y)$ is an isomorphism because all three algebras are finitely generated and the evaluation maps separate functions. $\square$

We recall Definition 1.1.2.1:

6.2.2.7 Definition *An* affine algebraic group *is a group object in the category of affine varieties. We will henceforth drop the adjective "affine" from the term "algebraic group", as we will never consider non-affine algebraic groups.*

Equivalently, an algebraic group is a reduced finitely generated commutative algebra $\mathscr{O}(G)$ along with algebra maps

comultiplication $\Delta : \mathscr{O}(G) \to \mathscr{O}(G) \otimes_{\mathbb{C}} \mathscr{O}(G)$ *dual to the multiplication $G \times G \to G$*

antipode $\mathcal{S} : \mathscr{O}(G) \to \mathscr{O}(G)$ *dual to the inverse map $G \to G$*

counit $\epsilon = \mathrm{ev}_e : \mathscr{O}(G) \to \mathbb{C}$

The group axioms equations (1.1.1.2) to (1.1.1.4) are equivalent to the axioms of a commutative Hopf algebra (Definition 4.1.0.1).

6.2.2.8 Lemma / Definition *Let A be a Hopf algebra. An algebra ideal $B \subseteq A$ is a* Hopf ideal *if $\Delta(B) \subseteq B \otimes A + A \otimes B \subseteq A \otimes A$. An ideal $B \subseteq A$ is Hopf if and only if the Hopf algebra structure on A makes the quotient B/A into a Hopf algebra.* $\square$

6.2.2.9 Definition *A commutative but not necessarily reduced Hopf algebra is an affine* group scheme*.*

6.2.2.10 Definition *An affine variety X over $\mathbb{C}$ is* smooth *if X is a manifold.*

6.2.2.11 Proposition *An algebraic group over $\mathbb{C}$ is smooth.*

6.2.2.12 Corollary *Since smooth schemes are reduced, affine group schemes are automatically algebraic groups.* $\square$

Proof (of Proposition 6.2.2.11) Let $E = \ker \epsilon$. Since $e \cdot e = e$, we see that the following diagram commutes:

$$\begin{tikzcd} \mathscr{O}(G) \arrow[r, "\Delta"] \arrow[d, "\epsilon"] & \mathscr{O}(G) \otimes \mathscr{O}(G) \arrow[d, "\epsilon\otimes\epsilon"] \\ \mathbb{C} \arrow[r, leftrightarrow, "\sim"] & \mathbb{C} \otimes \mathbb{C} \end{tikzcd}$$

In particular,

$$\Delta E \subseteq E \otimes \mathscr{O}(G) + \mathscr{O}(G) \otimes E, \tag{6.2.2.13}$$

and so E is a Hopf ideal, and $\mathscr{O}(G)/E$ is a Hopf algebra. Moreover, equation (6.2.2.13) implies that $\Delta(E^n) \subseteq \sum_{k+l=n} E^k \otimes E^l$, and so Δ and $\mathcal{S}$ induce maps $\tilde{\Delta}$ and $\tilde{\mathcal{S}}$ on $R = \mathrm{gr}_E \mathscr{O}(G) \stackrel{\text{def}}{=}$

$\bigoplus_{k\in\mathbb{N}} E^k/E^{k+1}$. In particular, R is a *graded Hopf algebra*, and is generated as an algebra by $R_1 = E/E^2$. Moreover, if $x \in R_1$, then x is primitive: $\Delta x = x\otimes 1 + 1\otimes x$.

Since $R_1 = E/E^2$ is finitely dimensional, R is finitely generated; let $R = \mathbb{C}[y_1,\dots,y_n]/J$ where $n = \dim G$ and J is a Hopf ideal of the Hopf algebra $\mathbb{C}[y_1,\dots,y_n]$ with the generators y_i all primitive. We can take the y_is to be a basis of R_1, and so $J_1 = 0$. We use the fact that $\mathbb{C}[y_1,\dots,y_n]\otimes\mathbb{C}[y_1,\dots,y_n] = \mathbb{C}[y_1,\dots,y_n,z_1,\dots,z_n]$, and that the antipode Δ is given by $\Delta : f(y) \mapsto f(y+z)$. Then a minimal-degree homogeneous element of J must be primitive, so $f(y+z) = f(y)+f(z)$, which in characteristic zero forces f to be homogeneous of degree 1. A similar calculation with the antipode forces the minimal-degree homogeneous elements $f \in J$ to satisfy $\mathcal{S}f = -f$.

In particular, $\mathrm{gr}_E\,\mathscr{O}(G)$ is a polynomial ring. We leave out the fact from algebraic geometry that this is equivalent to G being smooth at e. But we have shown that the Hopf algebra maps are smooth, whence G is smooth at every point. ☐

6.2.2.14 Corollary *An algebraic group over $\mathbb{C}$ is a Lie group.* ☐

Recall that if G is a Lie group with $\mathscr{C}(G)$ the algebra of smooth functions on G, and if $\mathfrak{g} = \mathrm{Lie}(G)$, then $\mathcal{U}\mathfrak{g}$ acts on $\mathscr{C}(G)$ by left-invariant differential operators, and indeed is isomorphic to the algebra of left-invariant differential operators.

6.2.2.15 Definition *Let G be a group. A subalgebra $S \subseteq \mathrm{Fun}(G)$ is* left-invariant *if for any $s \in S$ and any $g \in G$, the function $h \mapsto s(g^{-1}h)$ is an element of S. Equivalently, we define the action $G \curvearrowright \mathrm{Fun}(G)$ by $gs = s\circ g^{-1}$; then a subalgebra is left-invariant if it is fixed by this action.*

6.2.2.16 Lemma *Let $S \subseteq \mathrm{Fun}(G)$ be a left-invariant subalgebra, and let $s \in S$ be a function such that $\Delta s = \{(x,y)\mapsto s(xy)\} \subseteq \mathrm{Fun}(G\times G)$ is in fact an element of $S\otimes S \subseteq \mathrm{Fun}(G)\otimes\mathrm{Fun}(G) \hookrightarrow \mathrm{Fun}(G\times G)$. Then let $\Delta s = \sum s_1\otimes s_2$, where we suppress the indices of the sum. The action $G \curvearrowright S$ is given by*

$$g : s \mapsto \sum s_1(g^{-1})s_2$$ ☐

6.2.2.17 Corollary *Let u be a left-invariant differential operator and $s \in S$ as in Lemma 6.2.2.16, where $S \subseteq \mathscr{C}(G)$ is a left-invariant algebra of smooth functions. Then $us \in S$.*

Proof The left-invariance of u implies that $u(gs) = g\,u(s)$. Since $s(g^{-1}) \in \mathbb{C}$, we have:

$$u(gs)(h) = u\left(\sum s_1(g^{-1})s_2\right)(h) = \sum s_1(g^{-1})\,u(s_2)(h)$$

Let $h = e$. Then $\sum s_1(g^{-1})\,u(s_2)(e) = u(gs)(e) = g(us)(e) = (us)(g^{-1})$. In particular:

$$(us)(g) = \sum s_1(g)\,u(s_2)(e) \tag{6.2.2.18}$$

But $(us_2)(e)$ are numbers. Thus $us \in S$. ☐

6.2.2.19 Corollary *Let G be an algebraic group, with Lie algebra $\mathfrak{g} = \mathrm{Lie}(G)$. Then $\mathcal{U}\mathfrak{g}$ acts on $\mathscr{O}(G)$ by left-invariant differential operators.*

Since a differential operator is determined by its action on polynomials, we have a natural embedding $\mathcal{U}\mathfrak{g} \hookrightarrow \mathscr{O}(G)^*$ of vector spaces. ☐

6.2.2.20 Lemma *Let $u, v \in \mathcal{U}\mathfrak{g}$, where G is an algebraic group. Then $uv\, s = \sum u(s_1)\, v(s_2)(e)$.*

Proof This follows from equation (6.2.2.18). $\square$

6.2.2.21 Corollary *For each differential operator $u \in \mathcal{U}\mathfrak{g}$, let $\lambda_u \in \mathscr{O}(G)^*$ be the map $\lambda_u : s \mapsto u(s)(e)$. Then $\lambda_{uv}(s) = \sum \lambda_u(s_1)\, \lambda_v(s_2)$.* $\square$

6.2.2.22 Lemma *Let A be any (counital) coalgebra, for example a Hopf algebra. Then A^* is naturally an algebra: the map $A^* \otimes A^* \to A^*$ is given by $\langle \mu\nu, a\rangle \overset{\text{def}}{=} \langle \mu \otimes \nu, \Delta a\rangle$, and $\epsilon : A \to \mathbb{C}$ is the unit $\epsilon \in A^*$.* $\square$

6.2.2.23 Remark The dual to an algebra is not necessarily a coalgebra; if A is an algebra, then it defines a map $\Delta : A^* \to (A \otimes A)^*$, but if A is infinite-dimensional, then $(A \otimes A)^*$ properly contains $A^* \otimes A^*$. $\Diamond$

6.2.2.24 Remark Following the historical precedent, we take the pairing $(A^* \otimes A^*) \otimes (A \otimes A)$ to be $\langle \mu \otimes \nu, a \otimes b\rangle = \langle \mu, a\rangle\langle \nu, b\rangle$. This is in some sense the wrong pairing — it corresponds to writing $(A \otimes B)^* = A^* \otimes B^*$ for finite-dimensional vector spaces A, B, whereas $B^* \otimes A^*$ would be more natural — and is "wrong" in exactly the same way that the "-1" in the definition of the left action of G on $\mathrm{Fun}(G)$ is wrong. $\Diamond$

6.2.2.25 Proposition *The embedding $\mathcal{U}\mathfrak{g} \hookrightarrow \mathscr{O}(G)^*$ is given by the map $u \mapsto \lambda_u$ in Corollary 6.2.2.21, and is an algebra homomorphism.* $\square$

6.2.2.26 Definition *Let G be any group; then we define the* group algebra $\mathbb{C}[G]$ *of G to be the free vector space on the set G, with the multiplication given on the basis by the multiplication in G. The unit $e \in G$ becomes the unit $1 \cdot e \in \mathbb{C}[G]$.*

6.2.2.27 Lemma *If G is an algebraic group, then $\mathbb{C}[G] \hookrightarrow \mathscr{O}(G)^*$ is an algebra homomorphism given on the basis $g \mapsto \mathrm{ev}_g$.* $\square$

6.2.3 Constructing G from $\mathfrak{g}$

A Lie algebra $\mathfrak{g}$ does not determine the group G with $\mathfrak{g} = \mathrm{Lie}(G)$. We will see that the correct extra data consists of prescribed representation theory. Throughout the discussion, we gloss the details, merely waving at the proofs of various statements.

6.2.3.1 Lemma / Definition *Let G be an algebraic group. A finite-dimensional module $G \curvearrowright V$ is* algebraic *if the map $G \to \mathrm{GL}(V)$ is a morphism of affine varieties.*

Any finite-dimensional algebraic (left) action $G \curvearrowright V$ of an algebraic group G gives rise to a (left) coaction *$V^* \to \mathscr{O}(G) \otimes V^*$:*

$$\begin{tikzcd}
V^* \arrow[r, "\text{coact}"] \arrow[d, "\text{coact}"] & \mathcal{O}(G) \otimes V^* \arrow[d, "\text{comult} \otimes \text{id}"] \\
\mathcal{O}(G) \otimes V^* \arrow[r, "\text{id} \otimes \text{coact}"] & \mathcal{O}(G) \otimes \mathcal{O}(G) \otimes V^*
\end{tikzcd}$$

This in turn gives rise to a (left) action $\mathscr{O}(G)^* \curvearrowright V$, which specializes to the actions $G \curvearrowright V$ and $\mathcal{U}\mathfrak{g} \curvearrowright V$ under $G \hookrightarrow \mathbb{C}[G] \hookrightarrow \mathscr{O}(G)^*$ and $\mathcal{U}\mathfrak{g} \hookrightarrow \mathscr{O}(G)^*$. □

We will take the following definition; see also Section 12.3.1 or [BK01, ES02].

6.2.3.2 Definition *A* rigid category *is an abelian category $\mathcal{M}$ with a (unital) monoidal product and duals. We will write the monoidal product as $\otimes$.*

A rigid subcategory *of $\mathcal{M}$ is a full subcategory that is a tensor category with the induced abelian and tensor structures. I.e. it is a full subcategory containing the zero object and the monoidal unit, and closed under extensions, tensor products, and duals.*

6.2.3.3 Definition *A rigid category $\mathcal{M}$ is* finitely generated *if for some finite set of objects $V_1, \dots, V_n \in \mathcal{M}$, any object is a subquotient of some tensor product of V_is (possibly with multiplicities). Of course, by letting $V_0 = V_1 \oplus \dots \oplus V_n$, we see that any finitely generated rigid category is in fact generated by a single object.*

6.2.3.4 Example For any Lie algebra $\mathfrak{g}$, the category $\mathfrak{g}$-MOD of finite-dimensional representations of $\mathfrak{g}$ is a tensor category; indeed, if U is any Hopf algebra, then U-MOD is a tensor category. ◊

6.2.3.5 Definition *Let $\mathfrak{g}$ be a finite-dimensional Lie algebra over $\mathbb{C}$, and let $\mathcal{M}$ be a rigid subcategory of $\mathfrak{g}$-MOD. By definition, for each $V \in \mathcal{M}$, we have a linear map $\mathcal{U}\mathfrak{g} \to \operatorname{End} V$. Thus for each linear map $\phi : \operatorname{End} V \to \mathbb{C}$ we can construct a map $\{\mathcal{U}\mathfrak{g} \to \operatorname{End} V \overset{\phi}{\to} \mathbb{C}\} \in \mathcal{U}\mathfrak{g}^*$; we let $A_{\mathcal{M}} \subseteq \mathcal{U}\mathfrak{g}^*$ be the set of all such maps. Then $A_{\mathcal{M}}$ is the set of* matrix coefficients *of $\mathcal{M}$. Indeed, for each V, the maps $\mathcal{U}\mathfrak{g} \to \operatorname{End} V \to \mathbb{C}$ are the matrix coefficients of the action $\mathfrak{g} \curvearrowright V$. In particular, for each $V \in \mathcal{M}$, the space* $(\operatorname{End} V)^*$ *is naturally a subspace of $A_{\mathcal{M}}$, and $A_{\mathcal{M}}$ is the union of such subspaces.*

6.2.3.6 Lemma *If $\mathcal{M}$ is a rigid subcategory of $\mathfrak{g}$-MOD, then $A_{\mathcal{M}}$ is a subalgebra of the commutative algebra $\mathcal{U}\mathfrak{g}^*$. Moreover, $A_{\mathcal{M}}$ is a Hopf algebra, with comultiplication dual to the multiplication in $\mathcal{U}\mathfrak{g} \otimes \mathcal{U}\mathfrak{g} \to \mathcal{U}\mathfrak{g}$.*

Proof The algebra structure on $A = A_{\mathcal{M}}$ is straightforward: the multiplication and addition stem from the rigidity of $\mathcal{M}$, the unit is $\epsilon : \mathcal{U}\mathfrak{g} \to \mathbb{C} \overset{\sim}{\to} \mathcal{C}$, and the subtraction is not obvious but is straightforward; it relies on the fact that $\mathcal{M}$ is abelian, and so contains all subquotients.

We will explain where the Hopf structure on A comes from — since $\mathcal{U}\mathfrak{g}$ is infinite-dimensional, $\mathcal{U}\mathfrak{g}^*$ does not have a comultiplication in general. But $\mathcal{M}$ consists of finite-dimensional representations; if $V \in \mathcal{M}$, then we send $\{\mathcal{U}\mathfrak{g} \to \operatorname{End} V \overset{\phi}{\to} \mathbb{C}\} \in A$ to $\{(\operatorname{End} V \otimes \operatorname{End} V) \overset{\text{multiply}}{\longrightarrow} \operatorname{End} V \overset{\phi}{\to} \mathbb{C}\} \in (\operatorname{End} V)^* \otimes (\operatorname{End} V)^* \subseteq A \otimes A$.

That this is dual to the multiplication in $\mathcal{U}\mathfrak{g}$ comes from the fact that $\mathcal{U}\mathfrak{g} \to \operatorname{End} V$ is an algebra homomorphism. □

6.2.3.7 Corollary *The map $\mathcal{U}\mathfrak{g} \to A^*$ dual to $A \hookrightarrow \mathcal{U}\mathfrak{g}^*$ is an algebra homomorphism.* □

6.2.3.8 Proposition *Let $\mathcal{M}$ be a finitely-generated rigid subcategory of $\mathfrak{g}$-MOD. Then $A_{\mathcal{M}} = \mathscr{O}(G)$ for some algebraic group G.*

Proof If $\mathcal{M}$ is finitely generated, then there is some finite-dimensional representation $V_0 \in \mathcal{M}$ so that $(\operatorname{End} V_0)^*$ generates $A_{\mathcal{M}}$. Then $A_{\mathcal{M}}$ is a finitely generated commutative Hopf algebra, and so $\mathscr{O}(G)$ for some algebraic group G. $\square$

6.2.3.9 Lemma *Let $\mathfrak{g}$ be a finite-dimensional Lie algebra, $\mathcal{M}$ a finitely-generated rigid subcategory of $\mathfrak{g}$-MOD , and G the algebraic group corresponding to the algebra $A_{\mathcal{M}}$ of matrix coefficients of $\mathcal{M}$. We will henceforth write $\mathscr{O}(G)$ for $A_{\mathcal{M}}$. Then G acts naturally on each $V \in \mathcal{M}$.*

Proof Let $\{v^1, \dots, v^n\}$ be a basis of V and $\{\xi_1, \dots, \xi_n\}$ the dual basis of V^*. For each i, we define $\lambda_i : V \to \mathscr{O}(G)$ by $v \mapsto \{u \mapsto \langle \xi_i, uv\rangle\}$ where $v \in V$ and $u \in \operatorname{End} V$. Then we define $\sigma : V \to V \otimes \mathscr{O}(G)$ a right coaction of $\mathscr{O}(G)$ on V by $v \mapsto \sum_{i=1}^n v^i \otimes \lambda_i(v)$. It is a coaction because $uv = \sum_{i=1}^n v^i \lambda_i(v)(u)$ by construction. In particular, it induces an action $G \curvearrowright V$. $\square$

6.2.3.10 Proposition *Let $\mathcal{M}$ be a finitely-generated rigid subcategory of $\mathfrak{g}$-MOD that contains a faithful representation of $\mathfrak{g}$. Then the map $\mathcal{U}\mathfrak{g} \to A_{\mathcal{M}}^*$ is an injection.*

Proof Let $\sigma : G \curvearrowright V$ as in the proof of Lemma 6.2.3.9. Then the induced representation $\operatorname{Lie}(G) \curvearrowright V$ is by contracting σ with point derivations. But $\mathfrak{g} \curvearrowright V$ and the map $\mathcal{U}\mathfrak{g} \to \mathscr{O}(G)^*$ maps $x \in \mathfrak{g}$ to a point derivation since $x \in \mathfrak{g}$ is primitive. Thus the following diagram commutes for each $V \in \mathcal{M}$:

$$\begin{tikzcd} \mathfrak{g} \arrow[r] \arrow[d] & \mathscr{O}(G)^* \arrow[d] \\ \operatorname{Lie}(G) \arrow[r] & \mathfrak{gl}(V) \end{tikzcd}$$

The map $\mathfrak{g} \to \operatorname{Lie}(G)$ does not depend on V. Thus, if $\mathcal{M}$ contains a faithful $\mathfrak{g}$-module, then $\mathcal{U}\mathfrak{g} \hookrightarrow \mathcal{U}\operatorname{Lie}(G) \hookrightarrow \mathscr{O}(G)^*$. $\square$

6.2.3.11 Example Let $\mathfrak{g} = \mathbb{C}$ be one-dimensional, and let $\mathcal{M}$ be generated by one-dimensional representations V_α and V_β, where the generator $x \in \mathfrak{g}$ acts on V_α by multiplication by α, and on V_β by β. Then $\mathcal{M}$ is generated by $V_\alpha \oplus V_\beta$, and x acts as the diagonal matrix $\left(\begin{smallmatrix}\alpha & \\ & \beta\end{smallmatrix}\right)$. Let $\alpha, \beta \neq 0$, and let $\alpha \notin \mathbb{Q}\beta$. Then $\operatorname{Lie}(G)$ will contain all diagonal matrices, since $\alpha/\beta \notin \mathbb{Q}$, but $\mathfrak{g} \hookrightarrow \operatorname{Lie}(G)$ as a one-dimensional subalgebra. The group G is the complex torus, and the subgroup corresponding to $\mathfrak{g} \subseteq \operatorname{Lie}(G)$ is the irrational line. $\lozenge$

6.2.3.12 Proposition *Let V_0 be the generator of $\mathcal{M}$ satisfying the conditions of Proposition 6.2.3.8, and let W be a neighborhood of $0 \in \mathfrak{g}$. Then the image of $\exp(W)$ is Zariski dense in G.*

Proof Assume that $\mathcal{M}$ contains a faithful representation of $\mathfrak{g}$; otherwise, mod out $\mathfrak{g}$ by the kernel of the map $\mathfrak{g} \to \operatorname{Lie}(G)$. Thus, we may consider $\mathfrak{g} \subseteq \operatorname{Lie}(G)$, and let $H \subseteq G$ be a Lie subgroup with $\mathfrak{g} = \operatorname{Lie}(H)$. Let $f \in \mathscr{O}(G)$ and $u \in \mathcal{U}\mathfrak{g}$; then the pairing $\mathcal{U}\mathfrak{g} \otimes \mathscr{O}(G) \to \mathbb{C}$ sends $u \otimes f \mapsto u(f|_H)(e)$. In particular, the pairing depends only on a neighborhood of $e \in H$, and hence only on a neighborhood $W \ni 0$ in $\mathfrak{g}$. But the pairing is nondegenerate; if the Zariski closure of $\exp W$ in G were not all of G, then we could find $f, g \in \mathscr{O}(G)$ that agree on $\exp W$ but that have different behaviors under the pairing. $\square$

6.2.3.13 Definition *Let $\mathcal{M}$ be a finitely-generated rigid subcategory of $\mathfrak{g}$-MOD, and let G be the corresponding algebraic group as in Proposition 6.2.3.8. Then $\mathfrak{g}$ is* algebraically integrable *with respect to $\mathcal{M}$ if the map $\mathfrak{g} \to \mathrm{Lie}(G)$ is an isomorphism. In particular, $\mathcal{M}$ must contain a faithful representation of $\mathfrak{g}$.*

6.2.3.14 Example Let $\mathfrak{g}$ be a finite-dimensional abelian Lie algebra over $\mathbb{C}$, and let $X \subseteq \mathfrak{g}^*$ be a lattice of full rank, so that $X \otimes_{\mathbb{Z}} \mathbb{C} = \mathfrak{g}^*$. Let $\{\xi_1, \dots, \xi_n\}$ be a $\mathbb{Z}$-basis of X and hence a $\mathbb{C}$-basis of $\mathfrak{g}^*$, and let $\mathcal{M} = \{\bigoplus \mathbb{C}_\lambda \text{ s.t. } \lambda \in X\}$, where $\mathfrak{g} \curvearrowright \mathbb{C}_\lambda$ by $z \mapsto \lambda(z)\times$. Then $V_0 = \bigoplus \mathbb{C}_{\xi_i}$ is a faithful representation of $\mathfrak{g}$ in $\mathcal{M}$ and generates $\mathcal{M}$.

Then $G \subseteq \mathrm{GL}(V_0)$ is the Zariski closure if $\exp \mathfrak{g}$, and for $z \in \mathfrak{g}$, $\exp(z_1, \dots, z_n)$ is the diagonal matrix whose (i,i)th entry is $e^{\xi_i(z)}$. Thus G is a torus $T \cong (\mathbb{C}^\times)^n$, with $\mathcal{O}(T) = \mathbb{C}[t_1^{\pm 1}, \dots, t_n^{\pm 1}]$. In particular, $\mathfrak{g}$ is algebraically integrable with respect to $\mathcal{M}$, since X is a lattice. $\Diamond$

6.2.3.15 Proposition *Let $\mathfrak{g}$ be a finite-dimensional Lie algebra and $\mathcal{M}$ a finitely generated rigid subcategory of $\mathfrak{g}$-MOD containing a faithful representation. Suppose that $\mathfrak{g} = \bigoplus_{i=1}^r \mathfrak{g}_i$ as a vector space, where each $\mathfrak{g}_i$ is a Lie subalgebra of $\mathfrak{g}$; then $\mathcal{M}$ embeds in $\mathfrak{g}_i$-MOD for each i. If each $\mathfrak{g}_i$ is algebraically integrable with respect to (the image of) $\mathcal{M}$, then so is $\mathfrak{g}$.*

Proof Let G, G_i be the algebraic groups corresponding to $\mathfrak{g} \curvearrowright \mathcal{M}$ and to $\mathfrak{g}_i \curvearrowright \mathcal{M}$. Then for each i we have a map $G_i \to G$. Let $H \subseteq G$ be the subgroup of G corresponding to $\mathfrak{g} \subseteq \mathrm{Lie}(G)$. Consider the map $m : G_1 \times \cdots \times G_r \to G$ be the function that multiplies in the given order; it is not a group homomorphism, but it is a morphism of affine varieties. Since each $G_i \to G$ factors through H, and since H is a subgroup of G, the map m factors through H. Indeed, the differential of m at the identity is the sum map $\bigoplus \mathfrak{g}_i \to \mathfrak{g}$.

Thus we have an algebraic map m, with Zariski dense image. But it is a general fact that any such map (a *dominant morphism*) is dimension non-increasing. Therefore $\dim G \leq \dim(G_1 \times \cdots \times G_r) = \dim \mathfrak{g}$, and so $\mathfrak{g} = \mathrm{Lie}(G)$. $\square$

6.2.3.16 Theorem (Semisimple Lie algebras are algebraically integrable)
Let $\mathfrak{g}$ be a semisimple finite-dimensional Lie algebra over $\mathbb{C}$, and let $\mathfrak{h} \subseteq \mathfrak{g}$ be its Cartan subalgebra and Q and P the root and weight lattices. Let X be any lattice between these: $Q \subseteq X \subseteq P$. Let $\mathcal{M}$ be the category of finite-dimensional $\mathfrak{g}$-modules with highest weights in X. Then $\mathcal{M}$ is finitely generated rigid and contains a faithful representation of $\mathfrak{g}$, and $\mathfrak{g}$ is algebraically integrable with respect to $\mathcal{M}$.

Proof Let $V \in \mathcal{M}$; then its highest weights are all in X, and so all its weights are in X since $X \supseteq Q$. Moreover, the decomposition of V into irreducible $\mathfrak{g}$-modules writes $V = \bigoplus L_\lambda$, where each $\lambda \in P_+ \cap X$. This shows that $\mathcal{M}$ is rigid. It contains a faithful representation because the representation of $\mathfrak{g}$ corresponding to the highest root is the adjoint representation, and the highest root is an element of Q and hence of X. Moreover, $\mathcal{M} = \{\bigoplus V_\lambda \text{ s.t. } \lambda \in P_+ \cap X\}$ is finitely generated: $P_+ \cap X$ is $\mathbb{Z}_{\geq 0}$-generated by finitely many weights.

We recall the triangular decomposition (c.f. Proposition 5.6.0.6) of $\mathfrak{g}$: $\mathfrak{g} = \mathfrak{n}_- \oplus \mathfrak{h} \oplus \mathfrak{n}_+$. Then $\mathfrak{h}$ is abelian and acts on modules in $\mathcal{M}$ diagonally; in particular, $\mathfrak{h}$ is algebraically integrable by Example 6.2.3.14. On the other hand, on any $\mathfrak{g}$-module, $\mathfrak{n}_+$ and $\mathfrak{n}_-$ act by strict upper- and strict lower-triangular matrices, and the matrix exponential restricted to strict upper- (lower-) triangular matrices is a polynomial. In particular, by finding a faithful generator of $\mathcal{M}$ (for example, the sum of the generators plus the adjoint representation), we see that $\mathfrak{n}_\pm$ are algebraically integrable. The conclusion follows by Proposition 6.2.3.15. $\square$

6.2.3.17 Theorem (Classification of Semisimple Lie Groups over $\mathbb{C}$)
Let $\mathfrak{g}$ be a finite-dimensional semisimple Lie algebra over $\mathbb{C}$. Any connected Lie group G with $\mathrm{Lie}(G)$ *is semisimple; in particular, the algebraic groups constructed in Theorem 6.2.3.16 comprise all integrals of $\mathfrak{g}$.*

Proof Let $\tilde{G}$ be the connected and simply connected Lie group with $\mathrm{Lie}(G) = \mathfrak{g}$; then any integral of $\mathfrak{g}$ is a quotient of $\tilde{G}$ be a discrete and hence central subgroup of $\tilde{G}$, and the integrals are classified by the kernels of these quotients and hence by the subgroups of the center $Z(\tilde{G})$. Let G_X be the algebraic group corresponding to X. Since $Z(G_P) = P/Q$, it suffices to show that G_P is connected and simply connected.

We show first that G_X is connected. It is an affine variety; G_X is connected if and only if $\mathscr{O}(G_X)$ is an integral domain. Since G_X is the Zariski closure of $\exp W$ for a neighborhood W of $0 \in \mathfrak{g}$, and $\exp W$ is connected, so is G_X.

Let $U_\pm$ be the image of $\exp(\mathfrak{n}_\pm)$ in G_X, and let $T = \exp(\mathfrak{h})$. But $\exp : \mathfrak{n}_\pm \to U_\pm$ is the matrix exponential on strict triangular matrices, and hence polynomial with polynomial inverse; thus $U_\pm$ are simply connected.

We quote a fact from algebraic geometry: the image of an algebraic map contains a set Zariski open in its Zariski closure. In particular, since the image of $U_- \times T \times U_+$ is Zariski dense, it contains a Zariski open set, and so the complement of the image must live inside some closed subvariety of G_X with complex codimension at least 1, and hence real codimension at least 2, since locally this subvariety is the vanishing set of some polynomials in $\mathbb{C}^n$. So in any one-complex-dimensional slice transverse to this subvariety, the subvariety consists of just some points. Therefore any path in G_X can be moved off this subvariety and hence into the image of $U_- \times T \times U_+$.

It suffices to consider paths in G_X from e to e, and by choosing for each such path a nearby path in U_-TU_+, we get a map $\pi_1(U_-TU_+) \twoheadrightarrow \pi_1(G_X)$. On the other hand, by the LU decomposition (see any standard Linear Algebra textbook), the map $U_- \times T \times U_+ \to U_-TU_+$ is an isomorphism. Since $U_\pm$ are isomorphic as affine varieties to $\mathfrak{n}_\pm$, we have:

$$\pi_1(U_-TU_+) = \pi_1(U_- \times T \times U_+) = \pi_1(T)$$

And $\pi_1(T) = X^*$, the co-lattice to X, i.e. the points in $\mathfrak{g}$ on which all of X takes integral values.

Thus, it suffices to show that the map $\pi_1(T) \twoheadrightarrow \pi_1(G_P)$ collapses loops in T when $X = P$. But then $\pi_1(T) = P^* = Q^\vee$ is generated by the simple coroots $\alpha_i^\vee$. For each generator $\alpha_i^\vee = h_i$, we take $\mathfrak{sl}(2)_i \subseteq \mathfrak{g}$ and exponentiate to a map $\mathrm{SL}(2,\mathbb{C}) \to G$. Then the loops in $\exp(\mathbb{R}h_i)$, which generate $\pi_1(T)$, go to loops in $\mathrm{SL}(2,\mathbb{C})$ before going to G. But $\mathrm{SL}(2,\mathbb{C})$ is simply connected. This shows that the map $\pi_1(T) \twoheadrightarrow \pi_1(G_P)$ collapses all such loops, and G_P is simply connected. $\square$

Exercises

1. Show that the simple complex Lie algebra $\mathfrak{g}$ with root system G_2 has a 7-dimensional matrix representation with the generators shown below.

$$e_1 = \begin{bmatrix} 0&0&0&0&0&0&0\\ 0&0&1&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&1&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0 \end{bmatrix} \qquad f_1 = \begin{bmatrix} 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&1&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&1&0&0\\ 0&0&0&0&0&0&0 \end{bmatrix}$$
$$e_2 = \begin{bmatrix} 0&1&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&2&0&0&0\\ 0&0&0&0&1&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&1\\ 0&0&0&0&0&0&0 \end{bmatrix} \qquad f_2 = \begin{bmatrix} 0&0&0&0&0&0&0\\ 1&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&1&0&0&0&0\\ 0&0&0&2&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&1&0 \end{bmatrix} \tag{6.2.3.18}$$

2. (a) Show that there is a unique Lie group G over $\mathbb{C}$ with Lie algebra of type G_2.
   (b) Find explicit equations of G realized as the algebraic subgroup of $\mathrm{GL}(7,\mathbb{C})$ whose Lie algebra is the image of the matrix representation in Problem 1.

3. Show that the simply connected complex Lie group with Lie algebra $\mathfrak{so}(2n,\mathbb{C})$ is a double cover $\mathrm{Spin}(2n,\mathbb{C})$ of $\mathrm{SO}(2n,\mathbb{C})$, whose center Z has order four. Show that if n is odd, then Z is cyclic, and there are three connected Lie groups with this Lie algebra: $\mathrm{Spin}(2n,\mathbb{C})$, $\mathrm{SO}(2n,\mathbb{C})$ and $\mathrm{SO}(2n,\mathbb{C})/\{\pm I\}$. If n is even, then $Z \cong (\mathbb{Z}/2\mathbb{Z})^2$, and there are two more Lie groups with the same Lie algebra.

4. If G is an affine algebraic group, and $\mathfrak{g}$ its Lie algebra, show that the canonical algebra homormorphism $\mathcal{U}\mathfrak{g} \to \mathscr{O}(G)^*$ identifies $\mathcal{U}\mathfrak{g}$ with the set of linear functionals on $\mathscr{O}(G)$ whose kernel contains a power of the maximal ideal $\mathfrak{m} = \ker(\mathrm{ev}_e)$.

5. Show that there is a unique Lie group over $\mathbb{C}$ with Lie algebra of type E_8. Find the dimension of its smallest matrix representation.

6. Construct a finite dimensional Lie algebra over $\mathbb{C}$ which is not the Lie algebra of any algebraic group over $\mathbb{C}$. [Hint: the adjoint representation of an algebraic group on its Lie algebra is algebraic.]

Part II

Further Topics

Chapter 7

Real Lie Groups

7.1 (Over/Re)view of Lie groups

7.1.1 Lie groups in general

In general, a Lie group G can be broken up into a number of pieces.

The connected component of the identity, $G_{\text{conn}} \subseteq G$, is a normal subgroup, and G/G_{conn} is a discrete group.

$$1 \to G_{\text{conn}} \to G \to G_{\text{discrete}} \to 1$$

The maximal connected normal solvable subgroup of G_{conn} is called G_{sol}. Recall that a group is *solvable* if there is a chain of subgroups $G_{\text{sol}} \supseteq \cdots \supseteq 1$, where consecutive quotients are abelian. The Lie algebra of a solvable group is solvable, so Lie's theorem (Theorem 4.2.3.2) tells us that G_{sol} is isomorphic to (an extension by a discrete subgroup of) a subgroup of the group of upper triangular matrices.

Every normal solvable subgroup of $G_{\text{conn}}/G_{\text{sol}}$ is discrete, and therefore in the center (which is itself discrete). We call the pre-image of the center G_*. Then G/G_* is a product of simple groups (groups with no normal subgroups).

Define $G_{\text{nil}} \stackrel{\text{def}}{=} [G_{\text{sol}}, G_{\text{sol}}]$ to be the commutator subgroup. Since G_{sol} is solvable, G_{nil} is *nilpotent*: there is a chain of subgroups $G_{\text{nil}} \supseteq G_1 \supseteq \cdots \supseteq G_k = 1$ such that G_i/G_{i+1} is in the center of G_{nil}/G_{i+1}. In fact, G_{nil} must be isomorphic to a subgroup of the group of upper triangular matrices with ones on the diagonal. Such a group is called *unipotent*.

$$G_{\text{sol}} \subseteq \left\{ \begin{pmatrix} * & & & * \\ & \ddots & & \\ & & * & \\ 0 & & & * \end{pmatrix} \right\} \qquad G_{\text{nil}} \subseteq \left\{ \begin{pmatrix} 1 & & & * \\ & \ddots & & \\ & & 1 & \\ 0 & & & 1 \end{pmatrix} \right\}$$

All together, we have the following picture:

$$
\begin{array}{r}
\begin{array}{cl} G & \\ | & \Big\}\ \text{discrete; classification hopeless} \end{array}\\
\text{connected}\left\{\begin{array}{cl} G_{\mathrm{conn}} & \\ | & \Big\}\ \text{product of connected simples; classified}\\ G_* & \\ | & \Big\}\ \text{abelian discrete; classification trivial}\\ G_{\mathrm{sol}} & \\ | & \Big\}\ \text{abelian; classification trivial}\\ G_{\mathrm{nil}} & \\ | & \Big\}\ \text{nilpotent; classification a mess}\\ 1 & \end{array}\right.
\end{array}
\tag{7.1.1.1}
$$

The classification of connected simple Lie groups is quite long. There are many infinite series and a lot of exceptional cases. Some infinite series are $\mathrm{PSU}(n)$, $\mathrm{PSL}(n,\mathbb{R})$, and $\mathrm{PSL}(n,\mathbb{C})$. The "P" means "mod out by the center".

There are many connected simple Lie groups, but the classification is made easier by the following observation: there is a unique connected simple group for each simple Lie algebra. We've already classified complex semisimple Lie algebras, and it turns out that there a finite number of real Lie algebras which complexify to any given complex semisimple Lie algebra — such a real Lie algebra is a *real form* of the corresponding complex algebra. One warning is that tensoring with $\mathbb{C}$ preserves semisimplicity, but not simplicity. For example, $\mathfrak{sl}_2(\mathbb{C})$ is simple as a real Lie algebra, but its complexification is $\mathfrak{sl}_2(\mathbb{C}) \oplus \mathfrak{sl}_2(\mathbb{C})$, which is not simple.

7.1.1.2 Example Let G be the group of all "shape-preserving" transformations of $\mathbb{R}^4$: translations, reflections, rotations, scaling, etc. It is sometimes called $\mathbb{R}^4 \cdot \mathrm{GO}(4,\mathbb{R})$. The $\mathbb{R}^4$ stands for translations, the G means that you can multiply by scalars, and the O means that you can reflect and rotate. The $\mathbb{R}^4$ is a normal subgroup. In this case, the picture in equation (7.1.1.1) is:

$$
\begin{array}{r}
\begin{array}{cl} G = \mathbb{R}^4 \cdot \mathrm{GO}(4,\mathbb{R}) & \\ | & \Big\}\ G/G_{\mathrm{conn}} = \mathbb{Z}/2\mathbb{Z} \end{array}\\
\begin{array}{c} G_{\mathrm{conn}}/G_{\mathrm{sol}} \\ = \mathrm{SO}_4(\mathbb{R}) \end{array}\left\{\begin{array}{cl} G_{\mathrm{conn}} = \mathbb{R}^4 \cdot GO^+(4,\mathbb{R}) & \\ | & \Big\}\ G_{\mathrm{conn}}/G_* = \mathrm{PSO}(4,\mathbb{R}) \cong \mathrm{SO}(3,\mathbb{R}) \times \mathrm{SO}(3,\mathbb{R})\\ G_* = \mathbb{R}^4 \cdot \mathbb{R}^\times & \\ | & \Big\}\ G_*/G_{\mathrm{sol}} = \mathbb{Z}/2\mathbb{Z}\\ G_{\mathrm{sol}} = \mathbb{R}^4 \cdot \mathbb{R}^+ & \end{array}\right.\\
\begin{array}{cl} | & \Big\}\ G_{\mathrm{sol}}/G_{\mathrm{nil}} = \mathbb{R}^+\\ G_{\mathrm{nil}} = \mathbb{R}^4 & \\ | & \Big\} \\ 1 & \end{array}
\end{array}
$$

Here $\mathrm{GO}^+(4,\mathbb{R})$ is the connected component of the identity of $\mathrm{GO}(4,\mathbb{R})$ (those transformations that preserve orientation), $\mathbb{R}^\times$ is scaling by something other than zero, and $\mathbb{R}^+$ is scaling by something positive. Note that $\mathrm{SO}(3,\mathbb{R}) = \mathrm{PSO}(3,\mathbb{R})$ is simple.

$\mathrm{SO}(4,\mathbb{R})$ is "almost" the product $\mathrm{SO}(3,\mathbb{R})\times\mathrm{SO}(3,\mathbb{R})$. To see this, consider the associative (but not commutative) algebra of quaternions, $\mathbb{H}$. Since $q\bar{q} = a^2+b^2+c^2+d^2 > 0$ whenever $q \neq 0$, any non-zero quaternion has an inverse (namely, $\bar{q}/q\bar{q}$). Thus, $\mathbb{H}$ is a division algebra. Think of $\mathbb{H}$ as $\mathbb{R}^4$ and let S^3 be the unit sphere, consisting of the quaternions such that $\|q\| = q\bar{q} = 1$. It is easy to check that $\|pq\| = \|p\|\cdot\|q\|$, from which we get that left (right) multiplication by an element of S^3 is a norm-preserving transformation of $\mathbb{R}^4$. So we have a map $S^3\times S^3 \to \mathrm{O}(4,\mathbb{R})$. Since $S^3\times S^3$ is connected, the image must lie in $\mathrm{SO}(4,\mathbb{R})$. It is not hard to check that $\mathrm{SO}(4,\mathbb{R})$ is the image. The kernel is $\{(1,1),(-1,-1)\}$. So we have $S^3\times S^3/\{(1,1),(-1,-1)\} \cong \mathrm{SO}(4,\mathbb{R})$.

Conjugating a purely imaginary quaternion by some $q \in S^3$ yields a purely imaginary quaternion of the same norm as the original, so we have a homomorphism $S^3 \to \mathrm{O}(3,\mathbb{R})$. Again, it is easy to check that the image is $\mathrm{SO}(3,\mathbb{R})$ and that the kernel is ± 1, so $S^3/\{\pm 1\} \simeq \mathrm{SO}(3,\mathbb{R})$.

So the universal cover of $\mathrm{SO}(4,\mathbb{R})$ (a double cover) is the cartesian square of the universal cover of $\mathrm{SO}(3,\mathbb{R})$ (also a double cover). (One can also see the statement about universal covers by considering the corresponding Lie algebras.) Orthogonal groups in dimension 4 have a strong tendency to split up like this. Orthogonal groups in general tend to have these double covers, as we shall see in Section 7.3. These double covers are important if you want to study fermions. $\Diamond$

7.1.2 Lie groups and Lie algebras

Let $\mathfrak{g}$ be a Lie algebra. We set $\mathfrak{g}_{\mathrm{sol}} = \operatorname{rad}\mathfrak{g}$ to be the maximal solvable ideal (normal subalgebra), and $\mathfrak{g}_{\mathrm{nil}} = [\mathfrak{g}_{\mathrm{sol}}, \mathfrak{g}_{\mathrm{sol}}]$. Then we get the chain similar to the one in equation (7.1.1.1):

$$\begin{array}{cl} \mathfrak{g} & \\ | & \Big\}\ \text{product of simples; classified} \\ \mathfrak{g}_{\mathrm{sol}} & \\ | & \Big\}\ \text{abelian; classification trivial} \\ \mathfrak{g}_{\mathrm{nil}} & \\ | & \Big\}\ \text{nilpotent; classification a mess} \\ 0 & \end{array}$$

We have an equivalence of categories between simply connected Lie groups and Lie algebras. The correspondence cannot detect:

- Non-trivial components of G. For example, SO_n and O_n have the same Lie algebra.
- Discrete normal (therefore central, Lemma 3.5.1.4) subgroups of G. If $Z \subseteq G$ is any discrete normal subgroup, then G and G/Z have the same Lie algebra. For example, $\mathrm{SU}(2)$ has the same Lie algebra as $\mathrm{PSU}(2) \cong \mathrm{SO}(3,\mathbb{R})$.

If $\tilde{G}$ is a connected and simply connected Lie group with Lie algebra $\mathfrak{g}$, then any other connected group G with Lie algebra $\mathfrak{g}$ must be isomorphic to $\tilde{G}/Z$, where Z is some discrete subgroup of the center. Thus, if you know all the discrete subgroups of the center of $\tilde{G}$, you can read off all the connected Lie groups with the given Lie algebra.

7.1.2.1 Example Let's find all the connected groups with the algebra $\mathfrak{so}(4,\mathbb{R})$. First let's find a simply connected group with this Lie algebra. You might guess $\mathrm{SO}(4,\mathbb{R})$, but that isn't simply connected. The simply connected one is $S^3 \times S^3$ as we saw in Example 7.1.1.2 (it is a product of two simply connected groups, so it is simply connected). The center of S^3 is generated by -1, so the center of $S^3 \times S^3$ is $(\mathbb{Z}/2\mathbb{Z})^2$, the *Klein four-group*. There are three subgroups of order 2:

$$
\begin{tikzcd}
 & (\mathbb{Z}/2\mathbb{Z})^2 \arrow[dl, dash] \arrow[d, dash] \arrow[dr, dash] & \\
(-1,1) \arrow[dr, dash] & (-1,-1) \arrow[d, dash] & (1,-1) \arrow[dl, dash] \\
 & 1 &
\end{tikzcd}
\qquad
\begin{tikzcd}
 & \mathrm{PSO}(4,\mathbb{R}) \arrow[dl, dash] \arrow[d, dash] \arrow[dr, dash] & \\
\mathrm{SO}(3,\mathbb{R}) \times S^3 \arrow[dr, dash] & \mathrm{SO}(4,\mathbb{R}) \arrow[d, dash] & S^3 \times \mathrm{SO}(3) \arrow[dl, dash] \\
 & S^3 \times S^3 &
\end{tikzcd}
$$

Therefore, there are five groups with Lie algebra $\mathfrak{so}(4,\mathbb{R})$. Note that we are counting these groups "categorically", or "with symmetries". The automorphisms of $\mathfrak{so}(4,\mathbb{R})$ induce automorphisms on $\mathrm{PSO}(4,\mathbb{R})$, $\mathrm{SO}(4,\mathbb{R})$, and $S^3 \times S^3$. The *inner* automorphisms of $\mathfrak{so}(4,\mathbb{R})$ induce automorphisms of $\mathrm{SO}(3,\mathbb{R}) \times S^3$ and $S^3 \times \mathrm{SO}(3,\mathbb{R})$, but the isomorphism relating these two corresponds to the *outer* automorphism of $\mathfrak{so}(4,\mathbb{R})$. $\Diamond$

7.1.3 Lie groups and finite groups

The classification of finite simple groups resembles the classification of connected simple Lie groups. For example, $\mathrm{PSL}(n,\mathbb{R})$ is a simple Lie group, and $\mathrm{PSL}(n,\mathbb{F}_q)$ is a finite simple group except when $n = q = 2$ or $n = 2, q = 3$. Simple finite groups form about 18 series similar to Lie groups, and 26 or 27 exceptions, called sporadic groups, which don't seem to have analogues among Lie groups, although collectively one might compare "the sporadic simple groups" to "the exceptional Lie groups".

Moreover, finite groups and Lie groups are both built up from simple and abelian groups. However, the way that finite groups are built is much more complicated than the way Lie groups are built. Finite groups can contain simple subgroups in very complicated ways; not just as direct factors.

7.1.3.1 Example Within the theory of finite groups there are *wreath products*. Let G and H be finite simple groups with an action of H on a set of n points. Then H acts on G^n by permuting the factors. We can form the semi-direct product $G^n \rtimes H$, sometimes denoted $G \wr H$. There is no analogue for finite dimensional Lie groups, although there is an analogue for infinite dimensional Lie groups, which is why the theory becomes hard in infinite dimensions. $\Diamond$

7.1.3.2 Remark One important difference between (connected) Lie groups and finite groups is that the commutator subgroup of a solvable finite group need not be a nilpotent group. For example, the symmetric group S_4 has commutator subgroup A_4, which is not nilpotent. Also, nilpotent finite groups are almost never subgroups of upper triangular matrices (with ones on the diagonal). $\Diamond$

7.1.4 Lie groups and real algebraic groups

By "algebraic group", we mean an affine algebraic variety which is also a group, such as $\mathrm{GL}(n, \mathbb{R})$. Any algebraic group is a Lie group. Probably all the Lie groups you've come across have been algebraic groups. Since they are so similar, we'll list some differences. We will see in Section 7.2 that although in general the theories of Lie and algebraic groups are quite different, algebraic groups behave very similarly to *compact* Lie groups.

7.1.4.1 Remark Unipotent and semisimple abelian algebraic groups are totally different, but for Lie groups they are nearly the same. For example $\mathbb{R} \simeq \{\left(\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix}\right)\}$ is unipotent and $\mathbb{R}^\times \simeq \left\{\left(\begin{smallmatrix} a & 0 \\ 0 & a^{-1} \end{smallmatrix}\right)\right\}$ is semisimple. As Lie groups, they are closely related (nearly the same), but the Lie group homomorphism $\exp : \mathbb{R} \to \mathbb{R}^\times$ is not algebraic (polynomial), so they look quite different as algebraic groups. $\diamondsuit$

7.1.4.2 Remark Abelian varieties are different from affine algebraic groups. For example, consider the (projective) elliptic curve $y^2 = x^3 + x$ with its usual group operation and the group of matrices of the form $\left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right)$ with $a^2 + b^2 = 1$. Both are isomorphic to S^1 as Lie groups, but they are completely different as algebraic groups; one is projective and the other is affine. $\diamondsuit$

Some Lie groups do not correspond to any algebraic group. We describe two such groups in Examples 7.1.4.3 and 7.1.4.7.

7.1.4.3 Example The *Heisenberg group* is the subgroup of symmetries of $\mathrm{L}^2(\mathbb{R}, \mathbb{C})$ generated by translations ($f(t) \mapsto f(t + x)$), multiplication by $e^{2\pi i t y}$ ($f(t) \mapsto e^{2\pi i t y} f(t)$), and multiplication by $e^{2\pi i z}$ ($f(t) \mapsto e^{2\pi i z} f(t)$), for $x, y, z \in \mathbb{R}$. The general element is of the form $f(t) \mapsto e^{2\pi i (yt+z)} f(t+x)$. This can also be modeled as

$$\left\{\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}\right\} \Big/ \left\{\begin{pmatrix} 1 & 0 & n \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \middle| n \in \mathbb{Z}\right\}$$

It has the property that in any finite dimensional representation, the center (elements with $x = y = 0$) acts trivially, so it cannot be isomorphic to any algebraic group: by Proposition 9.1.0.3, algebraic groups always have finite-dimensional representations. $\diamondsuit$

For the second example, we quote without proof:

7.1.4.4 Theorem (Iwasawa decomposition)
If G is a (connected) semisimple Lie group, then there are closed subgroups K, A, and N, with K compact, A abelian, and N unipotent, such that the multiplication map $K \times A \times N \to G$ is a surjective diffeomorphism. Moreover, A and N are simply connected. $\square$

See also Proposition 8.3.2.1, where we give the proof for $G = \mathrm{GL}(n, \mathbb{R})$, and Theorem 11.1.1.3, where we give a proof for arbitrary G using Poisson geometry.

7.1.4.5 Example When $G = \mathrm{SL}(n, \mathbb{R})$, Theorem 7.1.4.4 says that any basis can be obtained uniquely by taking an orthonormal basis ($K = \mathrm{SO}(n)$), scaling by positive reals ($A = (\mathbb{R}_{>0})^n$ is the group of diagonal matrices with positive real entries), and shearing (N is the group of upper triangular matrices with ones on the diagonal). This is exactly the result of the Gram–Schmidt process. $\diamondsuit$

7.1.4.6 Corollary *As manifolds, $G = K \times A \times N$. In particular, $\pi_1(G) = \pi_1(K)$.* $\square$

7.1.4.7 Example Let's now try to find all connected groups with Lie algebra $\mathfrak{sl}(2,\mathbb{R}) = \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) | a+d=0\}$. There are two obvious ones: $\mathrm{SL}(2,\mathbb{R})$ and $\mathrm{PSL}(2,\mathbb{R})$. There aren't any other ones that can be represented as groups of finite dimensional matrices. However, $\pi_1(\mathrm{SL}(2,\mathbb{R})) = \pi_1(\mathrm{SO}(2,\mathbb{R})) = \pi_1(S^1) = \mathbb{Z}$, and so the universal cover $\widetilde{\mathrm{SL}(2,\mathbb{R})}$ has center $\mathbb{Z}$. Any finite dimensional representation of $\widetilde{\mathrm{SL}(2,\mathbb{R})}$ factors through $\mathrm{SL}(2,\mathbb{R})$, so none of the covers of $\mathrm{SL}(2,\mathbb{R})$ can be written as a group of finite dimensional matrices. Representing such groups is a pain.

The most important case is the *metaplectic group* $\mathrm{Mp}(2,\mathbb{R})$, which is the connected double cover of $\mathrm{SL}(2,\mathbb{R})$. It turns up in the theory of modular forms of half-integral weight and has an infinite-dimensional representation called the metaplectic representation. $\diamondsuit$

7.1.5 Important Lie groups

We now list some important Lie groups. See also Sections 1.3 and 4.3.

7.1.5.1 Example (Dimension 1) The only one-dimensional connected Lie groups are $\mathbb{R}$ and $S^1 = \mathbb{R}/\mathbb{Z}$. $\diamondsuit$

7.1.5.2 Example (Dimension 2) The abelian two-dimensional Lie groups are quotients of $\mathbb{R}^2$ by some discrete subgroup; there are three cases: $\mathbb{R}^2$, $\mathbb{R}^2/\mathbb{Z} = \mathbb{R} \times S^1$, and $\mathbb{R}^2/\mathbb{Z}^2 = S^1 \times S^1$.

There is also a non-abelian group, the group of all matrices of the form $\left(\begin{smallmatrix} a & b \\ 0 & a^{-1} \end{smallmatrix}\right)$, where $a > 0$. The Lie algebra is the subalgebra of 2×2 matrices of the form $\left(\begin{smallmatrix} h & x \\ 0 & -h \end{smallmatrix}\right)$, which is generated by two elements H and X, with $[H, X] = 2X$. $\diamondsuit$

7.1.5.3 Example (Dimension 3) There are some boring abelian and solvable groups, such as $\mathbb{R}^2 \ltimes \mathbb{R}^1$, or the direct sum of $\mathbb{R}^1$ with one of the two dimensional groups. As the dimension increases, the number of boring solvable groups gets huge, and nobody can do anything about them, so we ignore them from here on.

You also get the group $\mathrm{SL}(2,\mathbb{R})$, which is the most important Lie group of all. We saw in Example 7.1.4.7 that $\mathrm{SL}(2,\mathbb{R})$ has fundamental group $\mathbb{Z}$. The double cover $\mathrm{Mp}(2,\mathbb{R})$ is important. The quotient $\mathrm{PSL}(2,\mathbb{R})$ is simple, and acts on the open upper half plane by linear fractional transformations.

Closely related to $\mathrm{SL}(2,\mathbb{R})$ is the compact group $\mathrm{SU}(2)$. We know that $\mathrm{SU}(2) \simeq S^3$, and it covers $\mathrm{SO}(3,\mathbb{R})$, with kernel ± 1. After we learn about Spin groups, we will see that $\mathrm{SU}(2) \cong \mathrm{Spin}(3,\mathbb{R})$. The Lie algebra $\mathfrak{su}(2)$ is generated by three elements X, Y, and Z with relations $[X, Y] = 2Z$, $[Y, Z] = 2X$, and $[Z, X] = 2Y$. An explicit representation is given by $X = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, $Y = \left(\begin{smallmatrix} 0 & i \\ i & 0 \end{smallmatrix}\right)$, and $Z = \left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)$. Another specific presentation is given by $\mathfrak{su}(2) \simeq \mathbb{R}^3$ with the standard cross-product as the Lie bracket. The Lie algebras $\mathfrak{sl}(2,\mathbb{R})$ and $\mathfrak{su}(2)$ are non-isomorphic, but when you complexify, they both become isomorphic to $\mathfrak{sl}(2,\mathbb{C})$.

There is another interesting three-dimensional algebra. The *Heisenberg algebra* is the Lie algebra of the Heisenberg group. It is generated by X, Y, Z, with $[X, Y] = Z$ and Z central. You can think of this as strictly upper-triangular three-by-three matrices; c.f. Example 7.1.4.3. $\Diamond$

Nothing interesting happens in dimensions four and five. There are just lots of extensions of previous groups. We mention just a few highlights in higher dimensions.

7.1.5.4 Example (Dimension 6) We get the group $\mathrm{SL}(2, \mathbb{C})$. Later, we will see that it is also called $\mathrm{Spin}(1, 3; \mathbb{R})$. $\Diamond$

7.1.5.5 Example (Dimension 8) We have $\mathrm{SU}(3)$ and $\mathrm{SL}(3, \mathbb{R})$. This is the first time we get a non-trivial root system. $\Diamond$

7.1.5.6 Example (Dimension 14) The first exceptional group G_2 shows up. $\Diamond$

7.1.5.7 Example (Dimension 248) The last exceptional group E_8 shows up. We will discuss E_8 in detail in Section 8.1. $\Diamond$

7.1.5.8 Example (Dimension ∞) These lectures are mostly about finite-dimensional algebras, but let's mention some infinite dimensional Lie groups and Lie algebras.

1. Automorphisms of a Hilbert space form a Lie group.

2. Diffeomorphisms of a manifold form a Lie group. There is some physics stuff related to this.

3. *Gauge groups* are (continuous, smooth, analytic, or whatever) maps from a manifold M to a group G.

4. The *Virasoro algebra* is generated by L_n for $n \in \mathbb{Z}$ and c, with relations

$$[L_n, L_m] = (n - m)L_{n+m} + \delta_{n+m,0}\frac{n^3 - n}{12}c,$$

   where c is central (called the *central charge*). If you set $c = 0$, you get (complexified) vector fields on S^1, where we think of L_n as $ie^{in\theta}\frac{\partial}{\partial\theta}$. Thus, the Virasoro algebra is a central extension

$$0 \to \mathbb{C} \cdot c \to \text{Virasoro} \to \mathrm{Vect}(S^1) \to 0.$$

5. *Affine Kac–Moody algebras* are more or less central extensions of certain gauge groups over the circle. $\Diamond$

7.2 Compact Lie groups

7.2.1 Basic properties

In Chapter 5 we classified semisimple Lie algebras over an algebraically closed field characteristic 0. Now we will discuss the connection to compact groups. Representations of Lie groups are always taken to be smooth.

7.2.1.1 Example $\mathrm{SU}(n) = \{x \in \mathrm{GL}(n, \mathbb{C}) | x^* x = \mathrm{id} \text{ and } \det x = 1\}$ is a compact connected Lie group over $\mathbb{R}$. It is the group of linear transformations of $\mathbb{C}^n$ preserving a positive-definite hermitian form. $\mathrm{SU}(2)$ is topologically a 3-sphere. $\lozenge$

7.2.1.2 Lemma *Let G be compact. Then there exists the G-invariant* volume form *(a nowhere-vanishing top degree form) ω satisfying:*

1. *The volume of G is one: $\int_G \omega = 1$, and*
2. *ω is left invariant: $\int_G f\omega = \int_G L_h^* f\, \omega$ for all $h \in G$. Recall that $L_h^* f$ is defined by $(L_h^* f)(g) = f(hg)$.*

Proof To construct ω pick $\omega_e \in \bigwedge^{\text{top}} (\mathrm{T}_e G)^*$ and define $\omega_g = L_{g^{-1}}^* \omega_e$. $\square$

In fact, ω is also right-invariant if G is connected. If G is not connected, the right translations of ω can disagree with ω only by a sign, and in particular define the same measure $|\omega|$. See the exercises.

7.2.1.3 Proposition *If G is a compact group and V is a real representation of G, then there exists a positive definite G-invariant inner product on V. That is, $(gv, gw) = (v, w)$.*

Proof Pick any positive definite inner product $\langle , \rangle$ on V, e.g. by picking a basis and declaring it to be orthonormal. Define $(,)$ by:

$$(v, w) = \int_G \langle gv, gw \rangle \omega.$$

It is positive definite and invariant. $\square$

7.2.1.4 Corollary *Any finite dimensional representation of a compact group G is* completely reducible *— it splits into a direct sum of irreducibles.*

Proof The orthogonal complement to a subrepresentation is a subrepresentation. $\square$

In particular, the representation $\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g})$ is completely reducible. Thus, we can decompose $\mathfrak{g}$ into a direct sum of irreducible ideals: each is either simple or one-dimensional. We dump all the one-dimensional ideals into the center $\mathfrak{a}$ of $\mathfrak{g}$, and write $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_k \oplus \mathfrak{a}$. Thus, the Lie algebra of a compact group is the direct sum of its center and a semisimple Lie algebra. Recall Lemma/Definition 5.1.1.1: such a Lie algebra is called *reductive*.

7.2.1.5 Proposition *If G is simply connected and compact, then $\mathfrak{a}$ is trivial.*

Proof It suffices to consider the case that G is connected. Recall that $\mathrm{Grp} : \textsc{LieAlg} \to \textsc{scLieGp}$ is an equivalence of categories. In particular, $G = G_{\mathrm{ss}} \times A$, where $\mathrm{Lie}(A) = \mathfrak{a}$ and A is simply connected. But the only simply connected abelian Lie groups are $\mathbb{R}^n$, and so $\mathfrak{a}$ cannot be nontrivial. □

7.2.1.6 Proposition *If the Lie group G of $\mathfrak{g}$ is compact, then the Killing form β on $\mathfrak{g}$ is negative semi-definite. If the Killing form on $\mathfrak{g}$ is negative definite, then there is some compact group G with Lie algebra $\mathfrak{g}$.*

In fact, by the proof of Proposition 7.2.2.12, in the latter case every Lie group G with $\mathrm{Lie}(G) = \mathfrak{g}$ is compact.

Proof By Proposition 7.2.1.3, $\mathfrak{g}$ has an ad-invariant positive definite product, so the map $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ has image in $\mathfrak{so}(\mathfrak{g})$. Then $\mathrm{ad}(x)^T = -\mathrm{ad}(x)$, and so all eigenvalues of x are imaginary and $\mathrm{tr}_{\mathfrak{g}}(\mathrm{ad}\, x)^2 \leq 0$.

Conversely, if β is negative definite, then it is non-degenerate, so $\mathfrak{g}$ is semisimple by Theorem 4.2.6.4. Moreover,

$$-\beta(\mathrm{ad}(x)y, z) = \beta(y, \mathrm{ad}(x)z)$$

and so $\mathrm{ad}(x) = -\mathrm{ad}(x)^T$ with respect to this inner product. That is, the image of ad lies in $\mathfrak{so}(\mathfrak{g})$. It follows that the image under Ad of the simply connected group $\tilde{G}$ with Lie algebra $\mathfrak{g}$ lies in $\mathrm{SO}(\mathfrak{g})$. Thus, the image is a closed subgroup of a compact group, so it is compact. Since Ad has a discrete kernel, the image has the same Lie algebra. □

This motivates the following:

7.2.1.7 Definition *A real Lie algebra is* compact *if its Killing form is negative definite.*

How to classify compact Lie algebras? We know the classification of semisimple Lie algebras over $\mathbb{C}$, so we can always *complexify*: $\mathfrak{g} \rightsquigarrow \mathfrak{g}_{\mathbb{C}} \overset{\text{def}}{=} \mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C}$, which is again semisimple. However, this process might not be injective. For example, $\mathfrak{su}(2) = \left\{ \left(\begin{smallmatrix} a & b \\ -b^* & a \end{smallmatrix}\right) \text{ s.t. } a \in i\mathbb{R}, b \in \mathbb{C} \right\}$ and $\mathfrak{sl}(2,\mathbb{R})$ both complexify to $\mathfrak{sl}(2,\mathbb{C})$.

If $\mathfrak{g}_{\mathbb{C}} = \mathfrak{g}_{\mathbb{R}} \otimes_{\mathbb{R}} \mathbb{C}$, then $\mathfrak{g}_{\mathbb{R}}$ is a *real form* of $\mathfrak{g}_{\mathbb{C}}$. The following is due to Cartan:

7.2.1.8 Theorem (Cartan's classification of compact Lie algebras)
Every semisimple Lie algebra over $\mathbb{C}$ has exactly one (up to isomorphism) compact real form.

For example, the classical Lie groups $\mathrm{SL}(n,\mathbb{C})$, $\mathrm{SO}(n,\mathbb{C})$, and $\mathrm{Sp}(2n,\mathbb{C})$ have as their compact real forms $\mathrm{SU}(n)$, $\mathrm{SO}(n,\mathbb{R})$, and $\mathrm{Sp}(2n)$ from Lemma/Definition 1.3.1.2.

Proof The idea of the proof is as follows. Recall that if $\mathfrak{g}_{\mathbb{C}} = \mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C} = \mathfrak{g} \oplus i\mathfrak{g}$, then there is a "complex conjugation" $\sigma : \mathfrak{g}_{\mathbb{C}} \to \mathfrak{g}_{\mathbb{C}}$ that is an automorphism of real Lie algebras but a $\mathbb{C}$-antilinear involution, with $\mathfrak{g} = \mathfrak{g}_{\mathbb{C}}^{\sigma}$ the fixed points. So classifying real forms amounts to classifying all antilinear involutions. Moreover, suppose that we have two $\mathbb{C}$-antilinear involutions $\sigma_1, \sigma_2 : \mathfrak{g} \to \mathfrak{g}$. If $\sigma_1 = \phi\sigma_2\phi^{-1}$ for some $\phi \in \mathrm{Aut}\,\mathfrak{g}$, then $\mathfrak{g}^{\sigma_1} \overset{\phi}{\leftrightarrows} \mathfrak{g}^{\sigma_2}$. Conversely, any isomorphism $\mathfrak{g}_1 \overset{\sim}{\to} \mathfrak{g}_2$ of real Lie algebras lifts to an isomorphism $\phi : \mathfrak{g}_1 \otimes_{\mathbb{R}} \mathbb{C} \overset{\sim}{\to} \mathfrak{g}_2 \otimes_{\mathbb{R}} \mathbb{C}$, and if $\mathfrak{g}_a = \mathfrak{g}^{\sigma_a}$ for antilinear involutions σ_1, σ_2, then ϕ conjugates σ_1 to σ_2.

Existence Let $\mathfrak{g}$ be a semisimple Lie algebra over $\mathbb{C}$, and let $x_1, \dots, x_n$, $h_1, \dots, h_n$, $y_1, \dots, y_n$ be the Chevalley generators. We define the *Cartan involution* $\sigma : \mathfrak{g} \to \mathfrak{g}$ to be the $\mathbb{C}$-antilinear automorphism extending $\sigma(x_i) = -y_i$, $\sigma(y_i) = -x_i$, and $\sigma(h_i) = -h_i$. Let $\mathfrak{k} = \mathfrak{g}^\sigma = \{x \in \mathfrak{g} \text{ s.t. } \sigma(x) = x\}$. Then $\mathfrak{k}$ is an $\mathbb{R}$-Lie subalgebra of $\mathfrak{g}$ and $\mathfrak{g} = \mathfrak{k} \oplus i\mathfrak{k}$.

We claim that the Killing form on $\mathfrak{k}$ is negative definite. Suppose that $h = \sum_{j=1}^n a_j h_j \in \mathfrak{h} \cap \mathfrak{k}$. Then all a_i are pure-imaginary, and so the eigenvalues of h are imaginary and $\beta(h,h) < 0$. On the other hand, $\mathfrak{k} \cap (\mathfrak{n}^- \oplus \mathfrak{n}^+) = \{\sum_{j=1}^n (a_j x_j - \bar{a}_j y_j)\}$ for $a \in \mathbb{C}$, and the Weyl group action shows that β is negative on all of the root space.

Uniqueness Let $\mathfrak{g}$ be semisimple over $\mathbb{C}$ with Killing form β, and let $\theta : \mathfrak{g} \to \mathfrak{g}$ be a $\mathbb{C}$-antilinear involution for which $\mathfrak{g}^\theta$ is compact. We can skew the Killing form to $\beta_\theta(v,w) = \beta(\theta(v), w)$; this is the unique $\mathfrak{g}^\theta$-invariant Hermitian form on $\mathfrak{g}$.

Thus β_θ determines a *polar decomposition* $\mathrm{GL}(\mathfrak{g}) = \mathrm{U}_\theta(\mathfrak{g}) \times \mathrm{H}^+_\theta(\mathfrak{g})$, where $\mathrm{U}_\theta(\mathfrak{g}) = \{\phi \in \mathrm{GL}(\mathfrak{g}) \text{ s.t. } \phi\theta = \theta\phi\}$ are the unitary matrices with respect to β_θ and $\mathrm{H}^+_\theta(\mathfrak{g})$ are the symmetric positive-definite matrices — by *symmetric* we mean that for $\phi \in \mathrm{H}^+_\theta(\mathfrak{g})$ we have $\phi\theta = \theta\phi^{-1}$. If $\phi \in \mathrm{H}^+_\theta(\mathfrak{g})$, then it is of the form $\phi = \exp(\alpha)$ for some Hermitian matrix $\alpha \in \mathfrak{gl}(\mathfrak{g})$. Define $(\mathrm{Aut}\,\mathfrak{g})^+_\theta \overset{\text{def}}{=} \mathrm{Aut}\,\mathfrak{g} \cap \mathrm{H}^+_\theta(\mathfrak{g})$.

Choose an orthonormal basis $\{e_1, \dots, e_N\}$ for $\mathfrak{g}$ and define the *structure constants* via $[e_i, e_j] = \sum c^k_{ij} e_i e_j$. Let $\alpha = \sum \alpha_i e_i \in \mathfrak{gl}(\mathfrak{g})$ be Hermitian. Then $\exp(\alpha) \in \mathrm{Aut}\,\mathfrak{g}$ if and only if $\alpha_i + \alpha_j = \alpha_k$ whenever $c^k_{ij} \neq 0$. In particular, if $\exp(\alpha) \in \mathrm{Aut}\,\mathfrak{g}$, then so is $\exp(t\alpha)$ for any $t \in \mathbb{R}$. For $\phi \in (\mathrm{Aut}\,\mathfrak{g})^+_\theta$, by ϕ^t, $t \in \mathbb{R}$, we will mean $\exp(t\alpha)$, where $\phi = \exp(\alpha)$.

We now suppose that $\sigma : \mathfrak{g} \to \mathfrak{g}$ is any other $\mathbb{C}$-antilinear involutive Lie automorphism. Let $\omega = \sigma\theta \in \mathrm{Aut}\,\mathfrak{g}$. Then

$$\beta_\theta(\omega x, y) = \beta(\theta\omega x, y) = \beta(\omega^{-1}\theta x, y) = \beta(\theta x, \omega y) = \beta_\theta(x, \omega y)$$

as $\sigma^2 = \theta^2 = 1$. So ω is symmetric, so $\rho = \omega^2 \in (\mathrm{Aut}\,\mathfrak{g})^+_\theta$. By diagonalizing ω, it is clear that ρ^t and ω commute for all $t \in \mathbb{R}$. Moreover, $\rho^t\theta = \theta\rho^{-t}$ for any $t \in \mathbb{R}$, as $\rho^t \in (\mathrm{Aut}\,\mathfrak{g})^+_\theta$. Then:

$$\begin{aligned}(\rho^{1/4}\theta\varphi^{-1/4})\sigma = \rho^{1/2}\theta\sigma = \rho^{1/2}\omega^{-1} = \rho^{-1/2}\rho\omega^{-1} = \\ = \rho^{-1/2}\omega^2\omega^{-1} = \rho^{-1/2}\omega = \omega\rho^{-1/2} = \sigma\theta\rho^{-1/2} = \sigma(\rho^{1/4}\theta\rho^{1/4})\end{aligned}$$

In particular, θ is conjugate to some antilinear involution that commutes with σ.

Moreover, if $\mathfrak{g}^\theta$ is compact, then so is $\mathfrak{g}^{\theta'}$ for any conjugate of $\theta' = \phi\theta\phi^{-1}$. Now suppose that $\mathfrak{g}^\sigma$ is also compact. We will prove that if $\mathfrak{g}^\sigma$, $\mathfrak{g}^\theta$ are both compact and σ, θ commute, then $\sigma = \theta$.

Indeed, we decompose into eigenspaces $\mathfrak{g} = \mathfrak{g}^\sigma \oplus i\mathfrak{g}^\sigma$. Then since θ, σ commute, θ preserves the decomposition and we can write $\mathfrak{g}^\sigma = (\mathfrak{g}^\sigma)^\theta \oplus (\mathfrak{g}^\sigma)'$, where the latter is $(\mathfrak{g}^\sigma)' = \{x \text{ s.t. } \theta x = -x\}$. But by skew-linearity, we have $\mathfrak{g}^\theta = (\mathfrak{g}^\sigma)^\theta + i(\mathfrak{g}^\sigma)'$.

However, if $\mathfrak{g}^\sigma$ is compact, then $\mathrm{ad}\,x$ has pure-imaginary eigenvalues for all $x \in \mathfrak{g}^\sigma$, and in particular for $x \in (\mathfrak{g}^\sigma)'$. On the other hand, since $\mathfrak{g}^\theta$ is compact, $\mathrm{ad}\,x$ has pure-imaginary eigenvalues for $x \in i(\mathfrak{g}^\sigma)'$. Thus $\mathrm{ad}\,x = 0$ for $x \in (\mathfrak{g}^\sigma)'$, and by semisimplicity $(\mathfrak{g}^\sigma)' = 0$. Therefore $\sigma = \theta$. □

7.2.2 Unitary representations

Unitary representations are very important, and for the last 50 years people have wanted to classify unitary representations of specific groups. The whole subject was started by Hermann Weyl, and is motivated by quantum mechanics. In fact, the unitary representation theory of real Lie groups is an ongoing project.

7.2.2.1 Definition *A* Hilbert space V *is a vector space over* $\mathbb{C}$ *with a positive-definite Hermitian form* $(,)$*, which induces a* norm $\|v\| = \sqrt{(v,v)}$*, and* V *is required to be complete with respect to* $\|\cdot\|$*. The* operator norm *of* $x \in \mathrm{End}(V)$ *is* $|x| \overset{\text{def}}{=} \sup\{\frac{\|xv\|}{\|v\|} \text{ s.t. } v \in V \smallsetminus \{0\}\}$*. The* bounded operators *on* V *are* $\mathrm{B}(V) \overset{\text{def}}{=} \{x \in \mathrm{End}(V) \text{ s.t. } |x| < \infty\}$*. The* unitary operators *are* $\mathrm{U}(V) \overset{\text{def}}{=} \{x \in \mathrm{End}(V) \text{ s.t. } \|xv\| = \|v\| \ \forall v \in V\}$*.*

7.2.2.2 Remark $\mathrm{B}(V)$ is an associative unital algebra over $\mathbb{C}$ whereas $\mathrm{U}(V)$ is a group. ♢

7.2.2.3 Definition *Let* G *be a Lie group. A* unitary representation *of* G *is a homomorphism* $G \to \mathrm{U}(V)$ *such that* (gx, y) *is continuous in each variable.* V *is* (topologically) irreducible *if any closed invariant subspace is either* 0 *or* V*. Given a unitary representation* $G \to \mathrm{U}(V)$*, we define* $\mathrm{B}_G(V) \overset{\text{def}}{=} \{x \in \mathrm{B}(V) \text{ s.t. } xg = gx \ \forall g \in G\}$*.*

In fact, the continuity condition in the definition of "unitary representation" is a little subtle, because $\mathrm{U}(V)$ has multiple topologies, but we don't want to go into this.

7.2.2.4 Theorem (Schur's lemma for unitary representations)
If V *is an irreducible unitary representation of* G*, then* $\mathrm{B}_G(V) = \mathbb{C}$*.*

Proof Pick $x \in \mathrm{B}_G(V)$, and think about $a = x + x^*$ and $b = (x - x^*)/i$. These are Hermitian and commute with G. Then by some functional analysis:

$$a = \int_{\mathrm{Spec}\, a} x \, \mathrm{d}P(x)$$

The point is that if $E \subseteq \mathrm{Spec}\, a$ is a Borel subset, then $P(E)$ is a projector and commutes with a and also with G, and now the standard kernel-and-image argument works: $\ker P(E)$ is an invariant closed subspace, so $P(E) = \lambda \mathrm{id}$, and therefore a is scalar. A similar argument works for b, so $x = (a + ib)/2 \in \mathbb{C}$. □

Let K be a compact Lie group.

7.2.2.5 Example $\mathrm{L}^2(K)$ is an example of a unitary representation, where the action is $g\phi(x) = \phi(g^{-1}x)$. ♢

7.2.2.6 Example Any finite-dimensional representation of K is unitary, by averaging to get the invariant form. ♢

In fact, for a compact group K, any continuous representation on a Hilbert space can be made into a unitary representation. But these don't give more examples:

7.2.2.7 Proposition *Any irreducible unitary representation of K is finite-dimensional.*

Proof Pick $v \in V$ with $\|v\| = 1$. Define a projection $T : V \to V$ by $T(x) = (x, v)v$. Take the average $\bar{T} = \int g\, T\, g^{-1}\, \mathrm{d}g$. Then T is self-adjoint and compact, so $\bar{T}$ is as well. Moreover, $(Tx, x) \geq 0$, so $(\bar{T}x, x) \geq 0$. But $\bar{T}$ is compact and self-adjoint, and so has an eigenvalue. Then $\ker(\bar{T} - \lambda \mathrm{id})$ is an invariant subspace. So $\bar{T} = \lambda \mathrm{id}$, but it is also compact, so this is only possible if $\dim V < \infty$. ☐

This also proves:

7.2.2.8 Proposition *Any unitary representation of K has an irreducible subrepresentation.* ☐

By taking orthogonal complements we have:

7.2.2.9 Proposition *Every unitary representation of K is the closure of the direct sum of its irreducible subrepresentations.* ☐

7.2.2.10 Remark Proposition 7.2.2.9 does not hold for K noncompact. ◊

7.2.2.11 Theorem (Ado's theorem for compact groups)
Every compact group has a faithful finite-dimensional representation.

A statement similar to Theorem 7.2.2.11 holds also for algebraic groups: algebraic and compact groups are very similar. Compare also Theorems 7.2.2.13 and 9.1.1.4.

Proof The representation $K \curvearrowright \mathrm{L}^2(K)$ is faithful. As t ranges over some indexing set, let $V_t \subseteq \mathrm{L}^2(K)$ comprise all the irreducible subrepresentations of $\mathrm{L}^2(K)$, and let $\pi_t : K \to \mathrm{U}(V_t)$ be the corresponding homomorphisms. Then $\bigcap \ker \pi_t$ is trivial. But in a compact group, any set of closed subgroups will eventually stop: we have $\ker \pi_1 \supseteq (\ker \pi_1 \cap \ker \pi_2) \supseteq \dots$ eventually stops at $\ker \pi_1 \cap \dots \cap \ker \pi_s = \{1\}$. So then $V = V_1 \oplus \dots \oplus V_s$ is a faithful finite-dimensional representation of K. ☐

7.2.2.12 Proposition *Let $\mathfrak{k}$ be a semisimple compact Lie algebra and $\mathfrak{g} = \mathfrak{k}_{\mathbb{C}} = \mathfrak{k} \otimes_{\mathbb{R}} \mathbb{C}$. Then $\mathfrak{g}$ and $\mathfrak{k}$ have the same finite-dimensional complex representations, and by the Cartan classification we know its finite-dimensional complex irreducible representations. In particular, we have fundamental weights $\omega_1, \dots, \omega_n$, and the corresponding representations $V_{\omega_1}, \dots, V_{\omega_n}$ tensor-generate the full finite-dimensional representation theory. Let $V = V_{\omega_1} \oplus \dots \oplus V_{\omega_n}$, and construct an algebraic group $G \subseteq \mathrm{GL}(V)$ with $\mathrm{Lie}(G) = \mathfrak{g}$, and let $K \subseteq G$ correspond to $\mathfrak{k} \subseteq \mathfrak{g}$. Then K is simply connected.*

Proof Let $\tilde{K} \to K$ be the simply-connected cover. So $K = \tilde{K}/\Gamma$. If Γ is finite, set $K' = \tilde{K}$, and otherwise pick $\Gamma' \subsetneq \Gamma$ of finite index — it is an abelian discrete group — and $K' = K/\Gamma'$. Then we have a finite cover $K' \to K$. So K' has a faithful representation, as it is compact, but all the faithful representations are already there, so $K' = K$. ☐

We know that the center $Z(K) = \ker \mathrm{Ad}$. But also $Z(K) = P/Q$, the quotient of the weight lattice by the root lattice: inside K we have the maximal torus T, whose group of characters is P; in the adjoint form we have $\mathrm{Ad}\, T \subseteq \mathrm{Ad}\, K$, and its characters are Q; but then the center is the quotient of one by the other.

7.2.2.13 Theorem (Peter–Weyl theorem for compact groups)
If K is a compact group, then:

$$\mathrm{L}^2(K) = \overline{\bigoplus_{L(\lambda)\in \mathrm{Irr}(K)} L(\lambda)\otimes L(\lambda)^*}$$

The bar denotes closure.

Proof (Sketch) For semisimples, we use Theorem 9.1.1.4, and for arbitrary compacts we use that any compact is a quotient of a torus times a semisimple by a discrete group. The only thing to prove is that $\bigoplus_{\lambda\in P^+} L(\lambda)\otimes L(\lambda)^*$ is dense in $\mathrm{L}^2(K)$. And this follows from the fact that polynomial functions are dense in L^2. □

7.3 Orthogonal groups and related topics

With Lie algebras of small dimensions, and especially with the orthogonal groups, there are accidental isomorphisms. Almost all of these can be explained with Clifford algebras and Spin groups. The motivational examples that we'd like to explain are:

$\mathrm{SO}(2,\mathbb{R}) = S^1$ can double cover itself.

$\mathrm{SO}(3,\mathbb{R})$ has a simply connected double cover S^3.

$\mathrm{SO}(4,\mathbb{R})$ has a simply connected double cover $S^3\times S^3$.

$\mathrm{SO}(5,\mathbb{C})$: Look at $\mathrm{Sp}(4,\mathbb{C})$, which acts on $\mathbb{C}^4$ and on $\bigwedge^2(\mathbb{C}^4)$, which is 6 dimensional, and decomposes as $5\oplus 1$. $\bigwedge^2(\mathbb{C}^4)$ has a symmetric bilinear form given by $\bigwedge^2(\mathbb{C}^4)\otimes\bigwedge^2(\mathbb{C}^4)\to\bigwedge^4(\mathbb{C}^4)\simeq\mathbb{C}$, and $\mathrm{Sp}(4,\mathbb{C})$ preserves this form. You get that $\mathrm{Sp}(4,\mathbb{C})$ acts on $\mathbb{C}^5$, preserving a symmetric bilinear form, so it maps to $\mathrm{SO}(5,\mathbb{C})$. You can check that the kernel is ± 1. So $\mathrm{Sp}(4,\mathbb{C})$ is a double cover of $\mathrm{SO}(5,\mathbb{C})$.

$\mathrm{SO}(6,\mathbb{C})$: $\mathrm{SL}(4,\mathbb{C})$ acts on $\mathbb{C}^4$, and we still have our 6 dimensional $\bigwedge^2(\mathbb{C}^4)$, with a symmetric bilinear form. So you get a homomorphism $\mathrm{SL}(4,\mathbb{C})\to\mathrm{SO}(6,\mathbb{C})$, which you can check is surjective, with kernel ± 1.

So we have double covers S^1, S^3, $S^3\times S^3$, $\mathrm{Sp}(4,\mathbb{C})$, $\mathrm{SL}(4,\mathbb{C})$ of the orthogonal groups in dimensions 2,3,4,5, and 6, respectively. All of these look completely unrelated. In fact, we will put them all into a coherent framework, and rename them $\mathrm{Spin}(2,\mathbb{R})$, $\mathrm{Spin}(3,\mathbb{R})$, $\mathrm{Spin}(4,\mathbb{R})$, $\mathrm{Spin}(5,\mathbb{C})$, and $\mathrm{Spin}(6,\mathbb{C})$, respectively.

7.3.1 Clifford algebras

7.3.1.1 Example We have not yet defined Clifford algebras, but here are some motivational examples of Clifford algebras over $\mathbb{R}$.

- $\mathbb{C}$ is generated by $\mathbb{R}$, together with i, with $i^2=-1$

- $\mathbb{H}$ is generated by $\mathbb{R}$, together with i, j, each squaring to -1, with $ij + ji = 0$.
- Dirac wanted a square root for the operator $\nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} - \frac{\partial^2}{\partial t^2}$ (the wave operator in 4 dimensions). He supposed that the square root is of the form $A = \gamma_1 \frac{\partial}{\partial x} + \gamma_2 \frac{\partial}{\partial y} + \gamma_3 \frac{\partial}{\partial z} + \gamma_4 \frac{\partial}{\partial t}$ and compared coefficients in the equation $A^2 = \nabla$. Doing this yields $\gamma_1^2 = \gamma_2^2 = \gamma_3^2 = 1$, $\gamma_4^2 = -1$, and $\gamma_i\gamma_j + \gamma_j\gamma_i = 0$ for $i \neq j$.

  Dirac solved this by taking the γ_i to be 4×4 complex matrices. Then A operates on vector-valued functions on space-time. $\Diamond$

Generalizing the examples, we might define a *Clifford algebra* over $\mathbb{R}$ to be an associative algebra generated by some elements $\{\gamma_1, \dots, \gamma_n\}$ with relations prescribing $\gamma_i^2 \in \mathbb{R}$ and $\gamma_i\gamma_j + \gamma_j\gamma_i = 0$ for $i \neq j$. A better definition is:

7.3.1.2 Definition *Suppose that V is a vector space over a field $\mathbb{K}$ with some* quadratic form*, i.e. a homogeneous degree-two polynomial, $N : V \to \mathbb{K}$ function $V \to \mathbb{K}$. The* Clifford algebra $\mathrm{Cliff}(V, N)$ *is the $\mathbb{K}$-algebra generated by V with relations $v^2 = N(v)$. We define* $\mathrm{Cliff}(m, n; \mathbb{R})$ *to be the Clifford algebra over $\mathbb{K} = \mathbb{R}$ for $V = \mathbb{R}^{m+n}$ with $N(x_1, \dots, x_{m+n}) = x_1^2 + \dots + x_m^2 - x_{m+1}^2 - \dots - x_{m+n}^2$.*

7.3.1.3 Remark We know that $N(\lambda v) = \lambda^2 N(v)$ and that the expression $(a, b) \overset{\text{def}}{=} N(a + b) - N(a) - N(b)$ is bilinear. If the characteristic of $\mathbb{K}$ is not 2, we have $N(a) = \frac{(a,a)}{2}$. Thus, you can work with symmetric bilinear forms instead of quadratic forms so long as the characteristic of $\mathbb{K}$ is not 2. We'll use quadratic forms so that everything works in characteristic 2. $\Diamond$

7.3.1.4 Remark A few authors (mainly in index theory) use the relations $v^2 = -N(v)$. Some people add a factor of 2, which usually doesn't matter, but is wrong in characteristic 2. $\Diamond$

7.3.1.5 Example Take $V = \mathbb{R}^2$ with basis i, j, and with $N(xi + yj) = -x^2 - y^2$. Then the relations are $(xi + yj)^2 = -x^2 - y^2$ are exactly the relations for the quaternions: $i^2 = j^2 = -1$ and $(i + j)^2 = i^2 + ij + ji + j^2 = -2$, so $ij + ji = 0$. $\Diamond$

7.3.1.6 Remark If the characteristic of $\mathbb{K}$ is not 2, a "completing the square" argument shows that any quadratic form is isomorphic to $c_1x_1^2 + \dots + c_nx_n^2$, and if one can be obtained from another other by permuting the c_i and multiplying each c_i by a non-zero square, the two forms are isomorphic.

It follows that every quadratic form on a vector space over $\mathbb{C}$ is isomorphic to $x_1^2 + \dots + x_n^2$, and that every quadratic form on a vector space over $\mathbb{R}$ is isomorphic to $x_1^2 + \dots + x_m^2 - x_{m+1}^2 - \dots - x_{m+n}^2$ (m pluses and n minuses) for some m and n. One can check that these forms over $\mathbb{R}$ are non-isomorphic.

We will always assume that N is non-degenerate (i.e. that the associated bilinear form is non-degenerate), but one could study Clifford algebras arising from degenerate forms.

The reader should be warned, though, that the above criterion is not sufficient for classifying quadratic forms. For example, over the field $\mathbb{F}_3$, the forms $x^2 + y^2$ and $-x^2 - y^2$ are isomorphic via the isomorphism $\left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right) : \mathbb{F}_3^2 \to \mathbb{F}_3^2$, but -1 is not a square in $\mathbb{F}_3$. Also, completing the square doesn't work in characteristic 2. $\Diamond$

7.3.1.7 Remark The tensor algebra $\mathcal{T}V$ has a natural $\mathbb{Z}$-grading, and to form the Clifford algebra $\mathrm{Cliff}(V, N)$, we quotient by the ideal in $\mathcal{T}V$ generated by the even elements $v^2 - N(v)$. Thus, the algebra $\mathrm{Cliff}(V) = \mathrm{Cliff}(V)^0 \oplus \mathrm{Cliff}(V)^1$ is $\mathbb{Z}/2\mathbb{Z}$-graded. (The subspace $\mathrm{Cliff}(V)^0$ consists of the *even* elements, and $\mathrm{Cliff}(V)^1$ the *odd* ones.) A $\mathbb{Z}/2\mathbb{Z}$-graded algebra is called a *superalgebra*. A superalgebra is *supercommutative* if even elements commute with everything and the odd ones *anticommute*, i.e. for x, y homogeneous with respect to the $\mathbb{Z}/2\mathbb{Z}$ grading, we should have $xy = (-1)^{\deg x \cdot \deg y} yx$. The Clifford algebra $\mathrm{Cliff}(V, N)$ is supercommutative only if $N = 0$. ◊

7.3.1.8 Example

$\mathrm{Cliff}(0, 0; \mathbb{R}) = \mathbb{R}$.

$\mathrm{Cliff}(1, 0; \mathbb{R}) = \mathbb{R}\langle\varepsilon\rangle/(\varepsilon^2 - 1) = \mathbb{R}(1 + \varepsilon) \oplus \mathbb{R}(1 - \varepsilon) = \mathbb{R} \oplus \mathbb{R}$, with ϵ odd. Note that in the given basis, this is a direct sum of *algebras* over $\mathbb{R}$. Note also that this basis is not homogeneous with respect to the $\mathbb{Z}/2\mathbb{Z}$ grading.

$\mathrm{Cliff}(0, 1; \mathbb{R}) = \mathbb{R}\langle i\rangle/(i^2 + 1) = \mathbb{C}$, with i odd.

$\mathrm{Cliff}(2, 0; \mathbb{R}) = \mathbb{R}\langle\alpha, \beta\rangle/(\alpha^2 - 1, \beta^2 - 1, \alpha\beta + \beta\alpha)$. We get a homomorphism $\mathrm{Cliff}(2, 0; \mathbb{R}) \to \mathrm{Mat}(2, \mathbb{R})$, given by $\alpha \mapsto \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ and $\beta \mapsto \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. The homomorphism is onto because the two given matrices generate $\mathrm{Mat}(2, \mathbb{R})$ as an algebra. The dimension of $\mathrm{Mat}(2, \mathbb{R})$ is 4, and the dimension of $\mathrm{Cliff}(2, 0; \mathbb{R})$ is at most 4 because it is spanned by 1, α, β, and $\alpha\beta$. So we have that $\mathrm{Cliff}(2, 0, \mathbb{R}) \simeq \mathrm{Mat}(2, \mathbb{R})$.

$\mathrm{Cliff}(1, 1; \mathbb{R}) = \mathbb{R}\langle\alpha, \beta\rangle/(\alpha^2 - 1, \beta^2 + 1, \alpha\beta + \beta\alpha)$. Again, we get an isomorphism with $\mathrm{Mat}(2, \mathbb{R})$, given by $\alpha \mapsto \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ and $\beta \mapsto \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$. Another isomorphism is $\alpha \mapsto \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, $\beta \mapsto \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$. This latter isomorphism is compatible with the standard $\mathbb{Z}/2\mathbb{Z}$-grading on $\mathrm{Mat}(2, \mathbb{R})$ in which matrices of the form $\left(\begin{smallmatrix} * & 0 \\ 0 & * \end{smallmatrix}\right)$ are even and matrices of the form $\left(\begin{smallmatrix} 0 & * \\ * & 0 \end{smallmatrix}\right)$ are odd.

Thus, we've computed the Clifford algebras:

$m\backslash n$	0	1	2
0	$\mathbb{R}$	$\mathbb{C}$	$\mathbb{H}$
1	$\mathbb{R} \oplus \mathbb{R}$	$\mathrm{Mat}(2, \mathbb{R})$	
2	$\mathrm{Mat}(2, \mathbb{R})$		

◊

7.3.1.9 Remark If $\{v_1, \ldots, v_n\}$ is a basis for V, then $\{v_{i_1} \cdots v_{i_k} | i_1 < \cdots < i_k,\ k \le n\}$ spans $\mathrm{Cliff}(V)$, so the dimension of $\mathrm{Cliff}(V)$ is less than or equal to $2^{\dim V}$. In fact, it is always equal; one show this by following the proof of Theorem 3.2.2.1. ◊

7.3.1.10 Remark What is $\mathrm{Cliff}(U \oplus V)$ in terms of $\mathrm{Cliff}\, U$ and $\mathrm{Cliff}\, V$? One might reasonably guess $\mathrm{Cliff}(U \oplus V) \cong \mathrm{Cliff}\, U \otimes \mathrm{Cliff}\, V$. For the usual definition of tensor product, this is false: $\mathrm{Mat}(2, \mathbb{R}) = \mathrm{Cliff}(1, 1; \mathbb{R}) \neq \mathrm{Cliff}(1, 0; \mathbb{R}) \otimes \mathrm{Cliff}(0, 1; \mathbb{R}) = \mathbb{C} \oplus \mathbb{C}$. However, for the *superalgebra tensor product*, this is correct. The superalgebra tensor product is the regular tensor product of vector spaces, with the product given by $(a \otimes b)(c \otimes d) = (-1)^{\deg b \cdot \deg c} ac \otimes bd$ for homogeneous elements a, b, c, and d. ◊

Ignoring the previous remark and specializing to $\mathbb{K} = \mathbb{R}$, let's try to compute $\mathrm{Cliff}(U \oplus V)$ when $\dim U = m$ is even. Let $\alpha_1, \ldots, \alpha_m$ be an orthogonal basis for U and let $\beta_1, \ldots, \beta_n$ be an orthogonal basis for V. Then set $\gamma_i = \alpha_1\alpha_2\cdots\alpha_m\beta_i$. What are the relations between the α_i and the γ_j? We have:

$$\alpha_i\gamma_j = \alpha_i\alpha_1\alpha_2\cdots\alpha_m\beta_j = \alpha_1\alpha_2\cdots\alpha_m\beta_i\alpha_i = \gamma_j\alpha_i$$

We used that $\dim U$ is even and that α_i anti-commutes with everything except itself. Then:

$$\begin{aligned}
\gamma_i\gamma_j &= \gamma_i\alpha_1\cdots\alpha_m\beta_j = \alpha_1\cdots\alpha_m\gamma_i\beta_j \\
&= \alpha_1\cdots\alpha_m\alpha_1\cdots\alpha_m \underbrace{\beta_i\beta_j}_{-\beta_j\beta_i} = -\gamma_j\gamma_i, \quad i \neq j \\
\gamma_i^2 &= \alpha_1\cdots\alpha_m\alpha_1\cdots\alpha_m\beta_i\beta_i = (-1)^{\frac{m(m-1)}{2}}\alpha_1^2\cdots\alpha_m^2\beta_i^2 \\
&= (-1)^{m/2}\alpha_1^2\cdots\alpha_m^2\beta_i^2 && (m \text{ even})
\end{aligned}$$

So the γ_i's commute with the α_i and satisfy the relations of some Clifford algebra. Thus, we've shown that, for the ordinary (non-super) tensor product, $\mathrm{Cliff}(U \oplus V) \cong \mathrm{Cliff}(U) \otimes \mathrm{Cliff}(W)$, where W is V with the quadratic form multiplied by $(-1)^{\frac{1}{2}\dim U}\alpha_1^2\cdots\alpha_m^2 = (-1)^{\frac{1}{2}\dim U}\cdot\mathrm{discriminant}(U)$.

Taking $\dim U = 2$, we find that

$$\begin{aligned}
\mathrm{Cliff}(m+2, n; \mathbb{R}) &\cong \mathrm{Mat}(2, \mathbb{R}) \otimes \mathrm{Cliff}(n, m; \mathbb{R}) \\
\mathrm{Cliff}(m+1, n+1; \mathbb{R}) &\cong \mathrm{Mat}(2, \mathbb{R}) \otimes \mathrm{Cliff}(m, n; \mathbb{R}) \\
\mathrm{Cliff}(m, n+2; \mathbb{R}) &\cong \mathbb{H} \otimes \mathrm{Cliff}(n, m, \mathbb{R})
\end{aligned}$$

where the indices switch whenever the discriminant is positive. Using these formulas, we can reduce any Clifford algebra to tensor products of things like $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, and $\mathrm{Mat}(2, \mathbb{R})$.

Recall the rules for taking tensor products of matrix algebras (all tensor products are over $\mathbb{R}$, and are not super):

- $\mathbb{R} \otimes X \cong X$.
- $\mathbb{C} \otimes \mathbb{H} \cong \mathrm{Mat}(2, \mathbb{C})$. This follows from the isomorphism $\mathbb{C} \otimes \mathrm{Cliff}(m, n, \mathbb{R}) \cong \mathrm{Cliff}(m+n, \mathbb{C})$.
- $\mathbb{C} \otimes \mathbb{C} \cong \mathbb{C} \oplus \mathbb{C}$.
- $\mathbb{H} \otimes \mathbb{H} \cong \mathrm{Mat}(4, \mathbb{R})$. Consider the action on $\mathbb{H} \cong \mathbb{R}^4$ given by $(x \otimes y) \triangleright z = xzy^{-1}$.
- $\mathrm{Mat}\big(m, \mathrm{Mat}_n(X)\big) \cong \mathrm{Mat}(mn, X)$.
- $\mathrm{Mat}(m, X) \otimes \mathrm{Mat}(n, Y) \cong \mathrm{Mat}(mn, X \otimes Y)$.

Thus we can compute all Clifford algebras over $\mathbb{R}$. We will write down the interesting ones. Filling in the middle of the table is easy because you can move diagonally by tensoring with $\mathrm{Mat}(2, \mathbb{R})$.

	0	1	2	3	4	5	6	7	8
0	$\mathbb{R}$	$\mathbb{C}$	$\mathbb{H}$	$\mathbb{H} \oplus \mathbb{H}$	$\mathbb{M}_2(\mathbb{H})$	$\mathbb{M}_4(\mathbb{C})$	$\mathbb{M}_8(\mathbb{R})$	$\mathbb{M}_8(\mathbb{R}) \oplus \mathbb{M}_8(\mathbb{R})$	$\mathbb{M}_{16}(\mathbb{R})$
1	$\mathbb{R} \oplus \mathbb{R}$						$\mathbb{M}_8(\mathbb{C})$		
2	$\mathbb{M}_2(\mathbb{R})$							etc.	
3	$\mathbb{M}_2(\mathbb{C})$								
4	$\mathbb{M}_2(\mathbb{H})$								
5	$\mathbb{M}_2(\mathbb{H}) \oplus \mathbb{M}_2(\mathbb{H})$								
6	$\mathbb{M}_4(\mathbb{H})$								
7	$\mathbb{M}_8(\mathbb{C})$	$\mathbb{M}_8(\mathbb{H})$							
8	$\mathbb{M}_{16}(\mathbb{R})$		$\mathbb{M}_{16}(\mathbb{H})$.						
9	$\mathbb{M}_{16}(\mathbb{R}) \oplus \mathbb{M}_{16}(\mathbb{R})$			etc.					

To fit it on the page, we have abbreviated $\mathrm{Mat}(n, \mathbb{K}) = \mathbb{M}_n(\mathbb{K})$.

It is easy to see that $\mathrm{Cliff}(8+m, n) \cong \mathrm{Cliff}(m, n+8) \cong \mathrm{Cliff}(m, n) \otimes \mathrm{Mat}(16, \mathbb{R})$, which gives the table a kind of mod 8 periodicity. There is a more precise way to state this: $\mathrm{Cliff}(m, n, \mathbb{R})$ and $\mathrm{Cliff}(m', n', \mathbb{R})$ are *super Morita equivalent* if and only if $m - n \equiv m' - n' \mod 8$.

7.3.1.11 Remark This mod 8 periodicity turns up in several other places:

1. Real Clifford algebras $\mathrm{Cliff}(m, n; \mathbb{R})$ and $\mathrm{Cliff}(m', n'; \mathbb{R})$ are super Morita equivalent if and only if $m - n \equiv m' - n' \mod 8$.

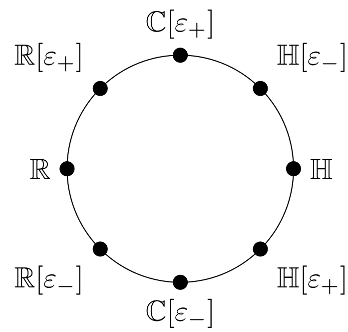
2. *Bott periodicity* says that stable homotopy groups of orthogonal groups are periodic mod 8.

3. Real K-theory is periodic with a period of 8.

4. Even unimodular lattices (such as the E_8 lattice) exist in $\mathbb{R}^{m,n}$ if and only if $m - n \equiv 0 \mod 8$.

5. The *super Brauer group* of $\mathbb{R}$ is $\mathbb{Z}/8\mathbb{Z}$. The super Brauer group is defined as follows. Take all real superalgebras over $\mathbb{R}$ up to super Morita equivalence; this forms a monoid under the super tensor product, and the super Brauer group is the group of invertible elements in this monoid. It turns out that every element of the super Brauer group is represented by a super

division algebra over $\mathbb{R}$ (a $\mathbb{Z}/2\mathbb{Z}$-graded algebra in which every non-zero homogeneous element is invertible), c.f. [Tri05]. Writing $\varepsilon_\pm$ for an odd generator satisfying $\varepsilon_\pm^2 = \pm 1$, and letting $i \in \mathbb{C}$ be odd[1] but $i, j, k \in \mathbb{H}$ even, this group is:



Note that the purely-even $\mathbb{C}$ is not invertible in the monoid of real superalgebras, because $\mathbb{C} \otimes_\mathbb{R} \mathbb{C} = \mathbb{C} \oplus \mathbb{C}$ and there is no way to tensor this to something with only one simple representation. $\Diamond$

Recall that $\mathrm{Cliff}(V) = \mathrm{Cliff}^0(V) \oplus \mathrm{Cliff}^1(V)$, where $\mathrm{Cliff}^1(V)$ is the odd part and $\mathrm{Cliff}^0(V)$ is the even part. We will need to know the structure of $\mathrm{Cliff}^0(m, n; \mathbb{R})$, which is, fortunately, easy to compute in in terms of smaller Clifford algebras. Working over an arbitrary field $\mathbb{K}$, let $\dim U = 1$ with γ a basis vector and let $\gamma_1, \dots, \gamma_n$ an orthogonal basis for V. Then $\mathrm{Cliff}^0(U \oplus V)$ is generated by $\gamma\gamma_1, \dots, \gamma\gamma_n$. We compute the relations:

$$\gamma\gamma_i \cdot \gamma\gamma_j = \begin{cases} -\gamma\gamma_j \cdot \gamma\gamma_i, & i \neq j \\ (-\gamma^2)\gamma_i^2, & i = j \end{cases}$$

So $\mathrm{Cliff}^0(U \oplus V)$ is itself the (evenization of the) Clifford algebra $\mathrm{Cliff}(W)$, where W is V with the quadratic form multiplied by $-\gamma^2 = -\operatorname{discriminant}(U)$. Over $\mathbb{R}$, this tells us that:

$$\mathrm{Cliff}^0(m+1, n; \mathbb{R}) \cong \mathrm{Cliff}(n, m; \mathbb{R})$$
$$\mathrm{Cliff}^0(m, n+1; \mathbb{R}) \cong \mathrm{Cliff}(m, n; \mathbb{R})$$

Mind the indices.

7.3.1.12 Remark For complex Clifford algebras, the situation is similar, but easier. One finds that $\mathrm{Cliff}(2m, \mathbb{C}) \cong \mathrm{Mat}(2^m, \mathbb{C})$ and $\mathrm{Cliff}(2m+1, \mathbb{C}) \cong \mathrm{Mat}(2^m, \mathbb{C}) \oplus \mathrm{Mat}(2^m, \mathbb{C})$, with $\mathrm{Cliff}^0(n, \mathbb{C}) \cong \mathrm{Cliff}(n-1, \mathbb{C})$. You could figure these out by tensoring the real algebras with $\mathbb{C}$ if you wanted. We see a mod 2 periodicity now. Bott periodicity for the unitary group is mod 2. $\Diamond$

[1] One could make i even since $\mathbb{R}[i, \varepsilon_\pm] = \mathbb{R}[\mp\varepsilon_\pm i, \varepsilon_\pm]$, and $\mathbb{R}[\mp\varepsilon_\pm i] \cong \mathbb{C}$ is entirely even. What really matters is that $i\epsilon_\pm = -\epsilon_\pm i$. If i is odd, then this is just the statement that i and $\epsilon_\pm$ supercommute. If i is even, then we are insisting that the odd element $\epsilon_\pm$ not commute with i, but rather that conjugating by it acts by complex conjugation.

7.3.2 Clifford groups, Spin groups, and Pin groups

In this section, we define Clifford groups, denoted $\mathrm{CLG}(V,N)$, and find an exact sequence

$$1 \to \mathbb{K}^\times \xrightarrow{\text{central}} \mathrm{CLG}(V,N) \to \mathrm{O}(V,N) \to 1.$$

7.3.2.1 Remark A standard notation for the Clifford group is to write $\Gamma_V K$ for what we call $\mathrm{CLG}(V,N)$, highlighting the dependence on the field and suppressing the dependence on the norm. Sometimes you see $C_V\mathbb{K}$ for our $\mathrm{Cliff}(V,N)$. We prefer this rarer notation, as we overuse the letter Γ elsewhere in this text. $\Diamond$

7.3.2.2 Definition *Let V be a $\mathbb{K}$-vector space and $N : V \to \mathbb{K}$ a quadratic form on it. Let α : $\mathrm{Cliff}(V,N) \to \mathrm{Cliff}(V,N)$ be the automorphism induced by $-1 : V \to V$, i.e. it is the automorphism that acts as $+1$ on $\mathrm{Cliff}^0(V,N)$ and as -1 on $\mathrm{Cliff}^1(V,N)$. The* Clifford group *is:*

$$\mathrm{CLG}(V,N) \stackrel{\text{def}}{=} \{x \in \mathrm{Cliff}(V,N) \text{ invertible s.t. } xV\alpha(x)^{-1} \subseteq V\}$$

In particular, $\mathrm{CLG}(V,N)$ acts on V.

7.3.2.3 Remark Many books leave out the α, which is a mistake, though not a serious one. They use xVx^{-1} instead of $xV\alpha(x)^{-1}$. Our definition is better for the following reasons:

1. It is the correct superalgebra sign. The superalgebra convention says that whenever you exchange two elements of odd degree, you pick up a minus sign, and V is odd.
2. Putting α in makes the theory much cleaner in odd dimensions. For example, we will see that the described action gives a map $\mathrm{CLG}(V) \to \mathrm{O}(V)$ which is onto if we use α, but not if we do not. (You get $\mathrm{SO}(V)$ without the α, which isn't too bad, but is still annoying.) $\Diamond$

7.3.2.4 Lemma *The elements of $\mathrm{CLG}(V)$ that act trivially on V are precisely $\mathbb{K}^\times \subseteq \mathrm{CLG}(V) \subseteq \mathrm{Cliff}(V)$.*

We will give the proof when $\operatorname{char}\mathbb{K} \neq 2$. Lemma 7.3.2.4 and the rest of the results are also true in characteristic 2, but you have to work harder: you can't go around choosing orthogonal bases because they may not exist.

Proof Suppose that $a = a_0 + a_1 \in \mathrm{CLG}(V)$ acts trivially on V, with a_0 even and a_1 odd. Then $(a_0 + a_1)v = v\alpha(a_0 + a_1) = v(a_0 - a_1)$. Matching up even and odd parts, we get $a_0 v = va_0$ and $a_1 v = -va_1$. Choose an orthogonal basis $\gamma_1, \dots, \gamma_n$ for V. We may write

$$a_0 = x + \gamma_1 y$$

where $x \in \mathrm{CLG}^0(V)$ and $y \in \mathrm{CLG}^1(V)$ and neither x nor y contain a factor of γ_1, so $\gamma_1 x = x\gamma_1$ and $\gamma_1 y = -y\gamma_1$. Applying the relation $a_0 v = va_0$ with $v = \gamma_1$, we see that $y = 0$, so a_0 contains no monomials with a factor γ_1.

Repeat this procedure with v equal to the other basis elements to show that $a_0 \in \mathbb{L}^\times$ (since it cannot have any γ's in it). Similarly, write $a_1 = y + \gamma_1 x$, with x and y not containing a factor of

γ_1. Then the relation $a_1\gamma_1 = -\gamma_1 a_1$ implies that $x = 0$. Repeating with the other basis vectors, we conclude that $a_1 = 0$, as y is odd but cannot have any factors.

So $a_0 + a_1 = a_0 \in \mathbb{K} \cap \mathrm{CLG}(V) = \mathbb{K}^\times$. □

7.3.2.5 Corollary *All elements of* $\mathrm{CLG}(V, N)$ *are homogeneous.*

Proof Suppose that $x \in \mathrm{Cliff}(V, N)$ is invertible and $xV\alpha(x)^{-1} \subseteq V$. Then for each $v \in V$, $xv\alpha(x)^{-1}$ is odd, so $-xv\alpha(x)^{-1} = \alpha(xv\alpha(x)^{-1}) = \alpha(x)vx^{-1}$. This implies that $x^{-1}\alpha(x)$ commutes with v; since V generates $\mathrm{Cliff}(V, N)$, we learn that $x^{-1}\alpha(x)$ is central in $\mathrm{Cliff}(V, N)$. By Lemma 7.3.2.4, $x^{-1}\alpha(x) \in \mathbb{K}^\times$, i.e. x is an eigenvector for α. But the eigenvectors of α are precisely the homogeneous elements. □

We will denote by $(-)^T$ the anti-automorphism of $\mathrm{Cliff}(V)$ induced by the identity on V ("anti" means that $(ab)^T = b^T a^T$). Do not confuse $a \mapsto \alpha(a)$ (automorphism), $a \mapsto a^T$ (anti-automorphism), and $a \mapsto \alpha(a^T)$ (anti-automorphism).

7.3.2.6 Definition *The* spinor norm *of* $a \in \mathrm{Cliff}(V)$ *is* $N(a) \stackrel{\text{def}}{=} a\,a^T \in \mathrm{Cliff}(V)$. *The* twisted spinor norm *is* $N^\alpha(a) \stackrel{\text{def}}{=} a\,\alpha(a)^T$.

7.3.2.7 Remark On $V \subseteq \mathrm{Cliff}(V, N)$, the spinor norm N coincides with the quadratic form N. Many authors seem not to have noticed this, and use different letters. Sometimes they use a sign convention which makes them different. ◊

7.3.2.8 Proposition

1. *The restriction of N to* $\mathrm{CLG}(V)$ *is a homomorphism whose image lies in* $\mathbb{K}^\times$. *(N is a mess on the rest of* $\mathrm{Cliff}(V)$.*)*
2. *The action of* $\mathrm{CLG}(V)$ *on V is orthogonal (with respect to N). That is, the map* $\mathrm{CLG}(V) \to \mathrm{GL}(V)$ *factors through* $\mathrm{O}(V, N)$.

Proof First we show that if $a \in \mathrm{CLG}(V)$, then $N^\alpha(a)$ acts trivially on V:

$$\begin{aligned}
N^\alpha(a)\, v\, \alpha\big(N^\alpha(a)\big)^{-1} &= a\alpha(a)^T v \Big(\alpha(a)\, \underbrace{\alpha\big(\alpha(a)^T\big)}_{=a^T}\Big)^{-1} \\
&= a\,\underbrace{\alpha(a)^T v (a^{-1})^T}_{=(a^{-1}v^T\alpha(a))^T}\alpha(a)^{-1} \\
&= aa^{-1}v\alpha(a)\alpha(a)^{-1} && (T|_V = \mathrm{id}_V \text{ and } a^{-1}v\alpha(a) \in V) \\
&= v
\end{aligned}$$

So by Lemma 7.3.2.4, $N^\alpha(a) \in \mathbb{K}^\times$. This implies that N^α is a homomorphism on $\mathrm{CLG}(V)$ because:

$$\begin{aligned}
N^\alpha(a)N^\alpha(b) &= a\alpha(a)^T N^\alpha(b) \\
&= aN^\alpha(b)\alpha(a)^T && (N^\alpha(b) \text{ is central}) \\
&= ab\alpha(b)^T\alpha(a)^T \\
&= (ab)\alpha(ab)^T = N^\alpha(ab)
\end{aligned}$$

After all this work with N^α, what we're really interested is N. On the even elements of $\mathrm{CLG}(V)$, N agrees with N^α, and on the odd elements, $N = -N^\alpha$. Since $\mathrm{CLG}(V)$ consists of homogeneous elements by Corollary 7.3.2.5, N is also a homomorphism from $\mathrm{CLG}(V)$ to $\mathbb{K}^\times$. This proves the first statement.

Finally, since N is a homomorphism on $\mathrm{CLG}(V)$, the action on V preserves the quadratic form $N|_V$. Thus, we have a homomorphism $\mathrm{CLG}(V) \to \mathrm{O}(V)$. □

Now we analyze the homomorphism $\mathrm{CLG}(V) \to \mathrm{O}(V)$. Lemma 7.3.2.4 says exactly that the kernel is $\mathbb{K}^\times$. Next we will show that the image is all of $\mathrm{O}(V)$. Suppose that we have $r \in V$ with $N(r) \neq 0$. Then:

$$\begin{aligned} rv\alpha(r)^{-1} &= -rv\frac{r}{N(r)} = v - \frac{vr^2 + rvr}{N(r)} \\ &= v - \frac{(v,r)}{N(r)}r && (7.3.2.9) \\ &= \begin{cases} -r & \text{if } v = r \\ v & \text{if } (v,r) = 0 \end{cases} && (7.3.2.10) \end{aligned}$$

Thus $r \in \mathrm{CLG}(V)$, and it acts on V by a reflection through the hyperplane $r^\perp$. One might deduce that the homomorphism $\mathrm{CLG}(V) \to \mathrm{O}(V)$ is onto because $\mathrm{O}(V)$ is generated by reflections. However, this would be incorrect: $\mathrm{O}(V)$ is *not* always generated by reflections!

7.3.2.11 Example Let $\mathbb{K} = \mathbb{F}_2$, $H = \mathbb{K}^2$ with the quadratic form $x^2 + y^2 + xy$, and $V = H \oplus H$. Then $\mathrm{O}(V, \mathbb{K})$ is not generated by reflections. See Exercise 5. ◊

7.3.2.12 Remark It turns out that this is the *only* counterexample. For any other vector space and/or any other non-degenerate quadratic form, $\mathrm{O}(V, \mathbb{K})$ is generated by reflections. The map $\mathrm{CLG}(V) \to \mathrm{O}(V)$ is surjective even when $V, \mathbb{K}$ are as in Example 7.3.2.11. Also, in every case except Example 7.3.2.11, $\mathrm{CLG}(V)$ is generated as a group by non-zero elements of V (i.e. every element of $\mathrm{CLG}(V)$ is a monomial). ◊

7.3.2.13 Remark Equation (7.3.2.9) is the definition of the reflection of v through r. It is only possible to reflect through vectors of non-zero norm. Reflections in characteristic 2 are strange; strange enough that people don't call them reflections, they call them *transvections*. ◊

We have proven that we have the following diagram:

$$\begin{tikzcd} & 1 \arrow[r] & \mathbb{K}^\times \arrow[r] \arrow[d, equal] & \mathrm{CLG}(V) \arrow[r] \arrow[d, "N"] & \mathrm{O}(V) \arrow[r] \arrow[d, dotted, "N"] & 1 \\ 1 \arrow[r] & \{\pm 1\} \arrow[r] & \mathbb{K}^\times \arrow[r, "x \mapsto x^2"] & \mathbb{K}^\times \arrow[r] & \mathbb{K}^\times/\{\text{squares}\} \arrow[r] & 1 \end{tikzcd} \tag{7.3.2.14}$$

The rows are exact, $\mathbb{K}^\times$ is in the center of $\mathrm{CLG}(V)$ (this is obvious, since $\mathbb{K}^\times$ is in the center of $\mathrm{Cliff}(V)$), and $N : \mathrm{O}(V) \to \mathbb{K}^\times/\{\text{squares}\}$ is the unique homomorphism sending reflection through $r^\perp$ to $N(r)$ modulo squares in $\mathbb{K}^\times$.

7.3.2.15 Definition *Given a $\mathbb{K}$-vector space V with a quadratic form $N : V \to \mathbb{K}$, the corresponding* pin *and* spin *groups are* $\mathrm{Pin}(V, N) \overset{\mathrm{def}}{=} \{x \in \mathrm{CLG}(V, N) \text{ s.t. } N(x) = 1\}$ *and its even part* $\mathrm{Spin}(V, N) \overset{\mathrm{def}}{=} \mathrm{Pin}^0(V, N)$.

7.3.2.16 Remark A word of warning: it's fairly common for people working over $\mathbb{R}$ to replace the condition $N(x) = 1$ with $N(x) = \pm 1$; c.f. Example 7.3.3.5. This is reasonable only because -1 is not a square in $\mathbb{R}$ — in particular, it is not a good choice if you want to work with algebraic groups. ♢

On $\mathbb{K}^\times$, the spinor norm is given by $x \mapsto x^2$, so the elements of spinor norm 1 are just ± 1. By restricting the top row of (7.3.2.14) to elements of norm 1 and even elements of norm 1, respectively, we get two exact sequences:

$$1 \longrightarrow \{\pm 1\} \longrightarrow \mathrm{Pin}(V) \longrightarrow \mathrm{O}(V) \xrightarrow{N} \mathbb{K}^\times/\{\text{squares}\}$$
$$1 \longrightarrow \{\pm 1\} \longrightarrow \mathrm{Spin}(V) \longrightarrow \mathrm{SO}(V) \xrightarrow{N} \mathbb{K}^\times/\{\text{squares}\}$$

To see exactness of the top sequence, note that the kernel of $\mathrm{Pin}(V) \to \mathrm{O}(V)$ is $\mathbb{K}^\times \cap \mathrm{Pin}(V) = \{\pm 1\}$, and that the image of $\mathrm{Pin}(V)$ in $\mathrm{O}(V)$ is exactly the elements of norm 1. The bottom sequence is similar, except that the image of $\mathrm{Spin}(V)$ is not all of $\mathrm{O}(V)$, it is only $\mathrm{SO}(V)$; by Remark 7.3.2.12, every element of $\mathrm{CLG}(V)$ is a product of elements of V, so every element of $\mathrm{Spin}(V)$ is a product of an even number of elements of V. Thus, its image is a product of an even number of reflections, so it is in $\mathrm{SO}(V)$.

7.3.2.17 Example Take V to be a positive-definite vector space over $\mathbb{R}$. Then N maps to $+1$ in $\mathbb{R}^\times/\{\text{squares}\} = \{\pm 1\}$ (because N is positive definite). So the spinor norm on $\mathrm{O}(V, \mathbb{R})$ is trivial.

So if $V = \mathbb{R}^n$ is equipped with a positive-definite metric, we get double covers:

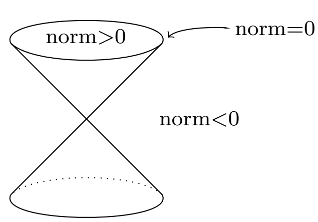
$$1 \longrightarrow \{\pm 1\} \longrightarrow \mathrm{Pin}(n, \mathbb{R}) \longrightarrow \mathrm{O}(n, \mathbb{R}) \longrightarrow 1$$
$$1 \longrightarrow \{\pm 1\} \longrightarrow \mathrm{Spin}(n, \mathbb{R}) \longrightarrow \mathrm{SO}(n, \mathbb{R}) \longrightarrow 1$$

This will account for the weird double covers we saw at the start of Section 7.3. ♢

7.3.2.18 Example What if the metric on V is negative-definite? Then every reflection maps to $-1 \in \mathbb{R}/\{\text{squares}\}$, so the spinor norm N is the same as the determinant map $\mathrm{O}(V) \to \pm 1$. In particular, $\mathrm{Pin}(0, n; \mathbb{R})$ is a double cover of $\mathrm{SO}(0, n; \mathbb{R}) = \mathrm{SO}(n)$, rather than of $\mathrm{O}(n)$. ♢

So in order to find interesting examples of the spinor norm, you have to look at cases that are neither positive definite nor negative definite.

7.3.2.19 Example Consider the *Lorentz space* $\mathbb{R}^{1,3}$, i.e. $\mathbb{R}^4$ with a metric with signature $\{+---\}$.



Reflection through a *spacelike* vector — a vector with norm < 0, also called a "parity reversal" P — has spinor norm -1 and determinant -1, and reflection through a *timelike* vector — norm > 0, "time reversal" T — has spinor norm $+1$ and determinant -1. So $\mathrm{O}(1,3;\mathbb{R})$ has four components (it is not hard to check that these are all the components), usually called 1, P, T, and PT. $\Diamond$

7.3.2.20 Remark We mention a few things for those who know Galois cohomology. We have an exact sequence of algebraic groups:

$$1 \to \mathrm{GL}(1) \to \mathrm{CLG}(V) \to \mathrm{O}(V) \to 1$$

"Algebraic group" means you don't put in a field. You do not necessarily get an exact sequence at any given field.

In general, if $1 \to A \to B \to C \to 1$ is exact, then $1 \to A(\mathbb{K}) \to B(\mathbb{K}) \to C(\mathbb{K})$ is exact, but $B(\mathbb{K}) \to C(\mathbb{K})$ need not be onto. What you actually get is a long exact sequence:

$$1 \to \mathrm{H}^0\big(\mathrm{Gal}(\bar{\mathbb{K}}/\mathbb{K}), A\big) \to \mathrm{H}^0\big(\mathrm{Gal}(\bar{\mathbb{K}}/\mathbb{K}), B\big) \to \mathrm{H}^0\big(\mathrm{Gal}(\bar{\mathbb{K}}/\mathbb{K}), C\big) \to H^1\big(\mathrm{Gal}(\bar{\mathbb{K}}/\mathbb{K}), A\big) \to \cdots$$

It turns out that $\mathrm{H}^1\big(\mathrm{Gal}(\bar{\mathbb{K}}/\mathbb{K}), \mathrm{GL}(1)\big) = 1$. However, $\mathrm{H}^1\big(\mathrm{Gal}(\bar{\mathbb{K}}/\mathbb{K}), \{\pm 1\}\big) = \mathbb{K}^\times/\{\text{squares}\}$. So from $1 \to \mathrm{GL}(1) \to \mathrm{CLG}(V) \to \mathrm{O}(V) \to 1$ we get:

$$1 \to \mathbb{K}^\times \to \mathrm{CLG}(V,\mathbb{K}) \to \mathrm{O}(V,\mathbb{K}) \to 1 = \mathrm{H}^1\big(\mathrm{Gal}(\bar{\mathbb{K}}/\mathbb{K})\,\mathrm{GL}(1)\big)$$

However, taking $1 \to \{\pm 1\} \to \mathrm{Spin}(V) \to \mathrm{SO}(V) \to 1$ we get:

$$1 \to \{\pm 1\} \to \mathrm{Spin}(V,\mathbb{K}) \to \mathrm{SO}(V,\mathbb{K}) \xrightarrow{N} \mathbb{K}^\times/\{\text{squares}\} = \mathrm{H}^1\big(\mathrm{Gal}(\bar{\mathbb{K}}/\mathbb{K}), \{\pm 1\}\big)$$

So we see that the non-surjectivity of N is some kind of higher Galois cohomology.

It is important to remember that $\mathrm{Spin}(V) \to \mathrm{SO}(V)$ is an onto map of *algebraic* groups, but $\mathrm{Spin}(V,\mathbb{K}) \to \mathrm{SO}(V,\mathbb{K})$ need *not* be an onto map of *groups*. $\Diamond$

7.3.2.21 Example Since 3 is odd, $\mathrm{O}(3,\mathbb{R}) \cong \mathrm{SO}(3,\mathbb{R}) \times \{\pm 1\}$. So we do have an exact sequence:

$$1 \to \{\pm 1\} \to \mathrm{Spin}(3,\mathbb{R}) \to \mathrm{SO}(3,\mathbb{R}) \to 1$$

Notice that $\mathrm{Spin}(3,\mathbb{R}) \subseteq \mathrm{Cliff}^0(3,\mathbb{R}) \cong \mathbb{H}$, so $\mathrm{Spin}(3,\mathbb{R}) \subseteq \mathbb{H}^\times$, and in fact we saw that it is the sphere S^3. $\Diamond$

7.3.3 Examples of Spin and Pin groups and their representations

Notice that $\mathrm{Pin}(V,\mathbb{K}) \subseteq \mathrm{Cliff}(V,\mathbb{K})^\times$, so any module over $\mathrm{Cliff}(V,\mathbb{K})$ gives a representation of $\mathrm{Pin}(V,\mathbb{K})$. We already figured out that the $\mathrm{Cliff}(V,\mathbb{K})$s are direct sums of matrix algebras over $\mathbb{R}, \mathbb{C}$, and $\mathbb{H}$ when $\mathbb{K} = \mathbb{R}, \mathbb{C}$.

What are the representations (modules) of complex Clifford algebras? Recall that $\mathrm{Cliff}(2n,\mathbb{C}) \cong \mathrm{Mat}(2^n,\mathbb{C})$, which has a representations of dimension 2^n, which is called the *spin representation* of $\mathrm{Pin}(2n,\mathbb{C})$; and $\mathrm{Cliff}(2n+1,\mathbb{C}) \cong \mathrm{Mat}(2^n,\mathbb{C}) \oplus \mathrm{Mat}(2^n,\mathbb{C})$, which has 2 representations, called the *spin representations* of $\mathrm{Pin}(2n+1,\mathbb{C})$.

What happens if we restrict these to $\mathrm{Spin}(V,\mathbb{C}) \subseteq \mathrm{Pin}_V(\mathbb{C})$? To do that, we have to recall that $\mathrm{Cliff}^0(2n,\mathbb{C}) \cong \mathrm{Mat}(2^{n-1},\mathbb{C}) \times \mathrm{Mat}(2^{n-1},\mathbb{C})$ and $\mathrm{Cliff}^0(2n+1,\mathbb{C}) \cong \mathrm{Mat}(2^n,\mathbb{C})$. So in even dimensions $\mathrm{Pin}(2n,\mathbb{C})$ has one spin representation of dimension 2^n splitting into two *half spin representations* of dimension 2^{n-1} and in odd dimensions, $\mathrm{Pin}(2n+1,\mathbb{C})$ has two spin representations of dimension 2^n which become the same upon restriction to $\mathrm{Spin}(V,\mathbb{C})$.

Now we'll give a second description of spin representations. We'll just do the even dimensional case (odd is similar). Say $\dim V = 2n$, and say we're over $\mathbb{C}$. Choose an orthonormal basis $\gamma_1,\dots,\gamma_{2n}$ for V, so that $\gamma_i^2 = 1$ and $\gamma_i\gamma_j = -\gamma_j\gamma_i$ in $\mathrm{Cliff}(V)$. Now look at the group G generated by $\gamma_1,\dots,\gamma_{2n}$, which is finite, with order 2^{1+2n} (you can write all its elements explicitly). Then the representations of $\mathrm{Cliff}(V,\mathbb{C})$ correspond to representations of G in which $-1 \in G$ acts as $-1 \in \mathbb{C}$ (as opposed to acting as 1 — it must square to 1). So another way to look at representations of the Clifford algebra is by looking at representations of G.

Let's look at the structure of G. First, the center is $\{\pm 1\}$: this uses the fact that we are in even dimensions, lest the product of all the generators also be central. Using this, we count the conjugacy classes. There are two conjugacy classes of size 1 ($\{1\}$ and $\{-1\}$) and $2^{2n}-1$ conjugacy classes of size 2 ($\{\pm\gamma_{i_1}\cdots\gamma_{i_k}\}$ for nonempty subsets $\{i_1,\dots,i_k\} \subseteq \{1,\dots,2n\}$). So G has a total of $2^{2n}+1$ conjugacy classes, and hence $2^{2n}+1$ irreducible representations. By inspection, G/center is abelian, isomorphic to $(\mathbb{Z}/2\mathbb{Z})^{2n}$, and this gives us 2^{2n} one-dimensional representations. So there is only one more representation left to find! We can figure out its dimension by recalling that the sum of the squares of the dimensions of irreducible representations gives us the order of G, which is 2^{2n+1}. So $2^{2n} \times 1^1 + 1 \times d^2 = 2^{2n+1}$, where d is the dimension of the mystery representation. Thus $d = 2^n$. So G, and therefore $\mathrm{Cliff}(2n,\mathbb{C})$, has an irreducible representation of dimension 2^n (as we found earlier in another way).

7.3.3.1 Example Consider $\mathrm{O}(2,1;\mathbb{R})$. As before, $\mathrm{O}(2,1;\mathbb{R}) \cong \mathrm{SO}(2,1;\mathbb{R}) \times \{\pm 1\}$, and $\mathrm{SO}(2,1,\mathbb{R})$ is not connected: it has two components, separated by the spinor norm N.

$$1 \to \{\pm 1\} \to \mathrm{Spin}(2,1;\mathbb{R}) \to \mathrm{SO}(2,1;\mathbb{R}) \xrightarrow{N} \{\pm 1\}$$

Since $\mathrm{Spin}(2,1;\mathbb{R}) \subseteq \mathrm{Cliff}^0(2,1;\mathbb{R}) \cong \mathrm{Mat}(2,\mathbb{R})$, $\mathrm{Spin}(2,1;\mathbb{R})$ has one two-dimensional spin representation. So there is a map $\mathrm{Spin}(2,1;\mathbb{R}) \to \mathrm{SL}(2,\mathbb{R})$. By counting dimensions, the map is a surjection, and we mentioned already that no nontrivial connected cover of $\mathrm{SL}(2,\mathbb{R})$ has a faithful representation. So $\mathrm{Spin}(2,1;\mathbb{R}) \cong \mathrm{SL}(2,\mathbb{R})$. $\Diamond$

Now let's look at some 4 dimensional orthogonal groups

7.3.3.2 Example Look at $\mathrm{SO}(4,\mathbb{R})$, which is compact. It has a complex spin representation of dimension $2^{4/2} = 4$, which splits into two half spin representations of dimension 2. We have the sequence

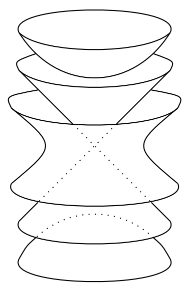
$$1 \to \pm 1 \to \mathrm{Spin}(4,\mathbb{R}) \to \mathrm{SO}(4,\mathbb{R}) \to 1 \qquad (N = 1)$$

$\mathrm{Spin}(4,\mathbb{R})$ is also compact, so the image in any complex representation is contained in some unitary group. So we get two maps $\mathrm{Spin}(4,\mathbb{R}) \to \mathrm{SU}(2)$, which is to say a map $\mathrm{Spin}(4,\mathbb{R}) \to \mathrm{SU}(2) \times \mathrm{SU}(2)$, and both sides have dimension 6 and centers of order 4. Thus, we find that $\mathrm{Spin}(4,\mathbb{R}) \cong \mathrm{SU}(2) \times \mathrm{SU}(2) \cong S^3 \times S^3$, which give you the two half spin representations. $\Diamond$

7.3.3.3 Example What about $\mathrm{SO}(3,1;\mathbb{R})$? Notice that $\mathrm{O}(3,1;\mathbb{R})$ has four components distinguished by the maps $\det, N : \mathrm{O}(3,1;\mathbb{R}) \to \pm 1$. So we get:

$$1 \to \pm 1 \to \mathrm{Spin}(3,1;\mathbb{R}) \to \mathrm{SO}(3,1;\mathbb{R}) \xrightarrow{N} \pm 1 \to 1$$

We expect two half spin representations, which give us two homomorphisms $\mathrm{Spin}(3,1;\mathbb{R}) \to \mathrm{SL}(2,\mathbb{C})$. This time, each of these homomorphisms is an isomorphism. The $\mathrm{SL}(2,\mathbb{C})$s are double covers of simple groups. Here, we don't get the splitting into a product as in the positive definite case. This isomorphism is heavily used in quantum field theory because $\mathrm{Spin}(3,1,\mathbb{R})$ is a double cover of the connected component of the Lorentz group (and $\mathrm{SL}(2,\mathbb{C})$ is easy to work with). Note also that the center of $\mathrm{Spin}(3,1;\mathbb{R})$ has order 2, not 4, as for $\mathrm{Spin}(4,0,\mathbb{R})$. Also note that the group $\mathrm{PSL}(2,\mathbb{C})$ acts on the compactified $\mathbb{C} \cup \{\infty\}$ by $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)(\tau) = \frac{a\tau+b}{c\tau+d}$. Subgroups of this group are called *Kleinian groups*. On the other hand, the group $\mathrm{SO}(3,1;\mathbb{R})^+$ (identity component) acts on $\mathbb{H}^3$ (three dimensional hyperbolic space). To see this, look at the following picture:



Each sheet of norm -1 is a hyperboloid isomorphic to $\mathbb{H}^3$ under the induced metric. In fact, we'll define hyperbolic space that way. If you're a topologist, you're very interested in hyperbolic 3-manifolds, which are $\mathbb{H}^3/(\text{discrete subgroup of } \mathrm{SO}(3,1;\mathbb{R}))$. If you use the fact that $\mathrm{SO}(3,1;\mathbb{R}) \cong \mathrm{PSL}(2,\mathbb{R})$, then you see that these discrete subgroups are in fact Kleinian groups. $\Diamond$

There are lots of exceptional isomorphisms in small dimension, all of which are very interesting, and almost all of them can be explained by spin groups.

7.3.3.4 Example $\mathrm{O}(2,2;\mathbb{R})$ has four components (given by $\det, N$). $\mathrm{Cliff}^0(2,2;\mathbb{R}) \cong \mathrm{Mat}(2,\mathbb{R}) \times \mathrm{Mat}(2,\mathbb{R})$, which induces an isomorphism $\mathrm{Spin}(2,2;\mathbb{R}) \to \mathrm{SL}(2,\mathbb{R}) \times \mathrm{SL}(2,\mathbb{R})$, which gives you the two half spin representations. Both sides have dimension 6 with centers of order 4. So this time we get two non-compact groups. Let's look at the fundamental group of $\mathrm{SL}(2,\mathbb{R})$, which is $\mathbb{Z}$, so the fundamental group of $\mathrm{Spin}(2,2;\mathbb{R})$ is $\mathbb{Z} \oplus \mathbb{Z}$. Recall, $\mathrm{Spin}(4,0;\mathbb{R})$ and $\mathrm{Spin}(3,1;\mathbb{R})$ were both simply connected. $\mathrm{Spin}(2,2;\mathbb{R})$ shows that spin groups need not be simply connected.

So we can take covers of $\mathrm{Spin}(2,2;\mathbb{R})$. What do these covers (e.g. the universal cover) look like? This is hard to describe because for finite dimensional complex representations, you get finite dimensional representations of the Lie algebra $\mathfrak{g} = \mathfrak{spin}(2,2;\mathbb{R}) = \mathfrak{sl}(2,\mathbb{R}) \oplus \mathfrak{sl}(2,\mathbb{R})$, which correspond to the finite dimensional representations of $\mathfrak{g} \otimes \mathbb{C}$, which correspond to the finite dimensional representations of $\mathfrak{spin}(4,0;\mathbb{R}) = $ Lie algebra of $\mathrm{Spin}(4,0;\mathbb{R})$, which correspond to the finite dimensional representations of $\mathrm{Spin}(4,0;\mathbb{R})$, since this group is simply connected. This means that any finite dimensional representation of a cover of $\mathrm{Spin}(2,2;\mathbb{R})$ actually factors through $\mathrm{Spin}(2,2;\mathbb{R})$.

So there is no way you can talk about these things with finite matrices, and infinite dimensional representations are hard.

To summarize, the *algebraic group* $\mathrm{Spin}(2,2)$ is simply connected (as an algebraic group, i.e. as a functor from rings to groups), which means that it has no algebraic central extensions. However, the *Lie group* $\mathrm{Spin}(2,2;\mathbb{R})$ is not simply connected; it has fundamental group $\mathbb{Z}\oplus\mathbb{Z}$. This problem does not happen for compact Lie groups (where every finite cover is algebraic). We saw this phenomenon already with $\mathrm{SL}(2)$. ♢

7.3.3.5 Example We've done $\mathrm{O}(4,0)$, $\mathrm{O}(3,1)$ and $\mathrm{O}(2,2)$, from which we can obviously get $\mathrm{O}(1,3)$ and $\mathrm{O}(0,4)$. Note that $\mathrm{O}(4,0;\mathbb{R}) \cong \mathrm{O}(0,4;\mathbb{R})$, $\mathrm{SO}(4,0;\mathbb{R}) \cong \mathrm{SO}(0,4;\mathbb{R})$, and $\mathrm{Spin}(4,0;\mathbb{R}) \cong \mathrm{Spin}(0,4;\mathbb{R})$. However, $\mathrm{Pin}(4,0;\mathbb{R}) \not\cong \mathrm{Pin}(0,4;\mathbb{R})$, for the simple reason, mentioned in Example 7.3.2.18, that $\mathrm{Pin}(4,0;\mathbb{R})$ double covers $\mathrm{O}(4;\mathbb{R})$ but $\mathrm{Pin}(0,4;\mathbb{R})$ only double covers $\mathrm{SO}(4;\mathbb{R})$.

There is an alternate definition of real Pin groups, frequently used, in which $\mathrm{Pin}(m,n;\mathbb{R})$ always double covers $\mathrm{O}(m,n;\mathbb{R})$ — replace $N(x)=1$ in Definition 7.3.2.15 with $N(x)=\pm 1$, and use the fact that -1 is not a square in $\mathbb{R}$. For this alternate definition, $\mathrm{Pin}(n,0)$ and $\mathrm{Pin}(0,n)$ are never isomorphic as double covers of $\mathrm{O}(n)$.

$$\begin{array}{ccc} \mathrm{Pin}(n,0;\mathbb{R}) & & \mathrm{Pin}(0,n;\mathbb{R}) \\ \downarrow & & \downarrow \\ \mathrm{O}(n,0;\mathbb{R}) & = & \mathrm{O}(0,n;\mathbb{R}) \end{array}$$

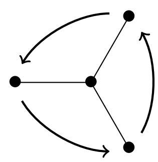
Take a reflection (of order 2) in $\mathrm{O}(n,0;\mathbb{R}) = \mathrm{O}(0,n;\mathbb{R})$, and lift it to the Pin groups. What is the order of the lift? The reflection vector v, with $v^2=\pm 1$, lifts to the element $v \in \mathrm{CLG}(V,\mathbb{R}) \subseteq \mathrm{Cliff}^1(V,\mathbb{R})$. Notice that $v^2 = 1$ in the case $V=\mathbb{R}^{n,0}$ and $v^2=-1$ in the case of $V=\mathbb{R}^{0,n}$, so in $\mathrm{Pin}(n,0;\mathbb{R})$, the reflection lifts to something of order 2, but in $\mathrm{Pin}(0,n;\mathbb{R})$, you get an element of order 4. So these two groups are different.

Two groups are *isoclinic* if they are confusingly similar. A similar phenomenon is common for groups of the form $2\cdot G\cdot 2$, which means it has a center of order 2, then some group G, and the abelianization has order 2. Watch out.

Remarkably, the groups $\mathrm{Pin}(4m,0)$ and $\mathrm{Pin}(0,4m)$ (with the modified definition of Pin groups) *are* isomorphic as abstract groups. Indeed, there is an isomorphism $\mathrm{Cliff}(4m,0;\mathbb{R}) \cong \mathrm{Cliff}(0,4m;\mathbb{R})$ given by sending $v \in V = \mathbb{R}^n$ to $v\theta$, where $\theta \in \mathrm{Cliff}^0(0,n;\mathbb{R})$ is the product of all the generators, i.e. the volume form on V, and this isomorphism relates the two Pin groups. The dependence on n mod 4 is as follows: we want θ to be an even element so as to have an isomorphism of superalgebras; we want $\theta^2=+1$, but in general $\theta^2 = (-1)^{\binom{n}{2}}$.

More generally, [BDMGK01] show that there is an abstract isomorphism $\mathrm{Pin}(m,n) \cong \mathrm{Pin}(m',n')$ iff $m+n = m'+n'$ and $m-n \cong m'-n' \mod 8$. ♢

7.3.3.6 Example There is a special property of the eight-dimensional orthogonal groups called *triality*. Recall that $\mathrm{O}(8,\mathbb{C})$ has Dynkin diagram D_4, which has a symmetry of order three:



But $\mathrm{O}(8,\mathbb{C})$ and $\mathrm{SO}(8,\mathbb{C})$ do not have corresponding symmetries of order three. The thing that does have the "extra" symmetry of order three is the spin group $\mathrm{Spin}(8,\mathbb{R})$!

You can see it as follows. Look at the half spin representations of $\mathrm{Spin}(8,\mathbb{R})$. Since this is a spin group in even dimension, there are two. $\mathrm{Cliff}(8,0;\mathbb{R}) \cong \mathrm{Mat}(2^{8/2-1},\mathbb{R}) \times \mathrm{Mat}(2^{8/2-1},\mathbb{R}) \cong \mathrm{Mat}(8,\mathbb{R}) \times \mathrm{Mat}(8,\mathbb{R})$. So $\mathrm{Spin}(8,\mathbb{R})$ has two eight-dimensional real half spin representations. But the spin group is compact, so it preserves some quadratic form, so you get two homomorphisms $\mathrm{Spin}(8,\mathbb{R}) \to \mathrm{SO}(8,\mathbb{R})$. So $\mathrm{Spin}(8,\mathbb{R})$ has three eight-dimensional representations: the half spins, and the one from the map to $\mathrm{SO}(8,\mathbb{R})$. These maps $\mathrm{Spin}(8,\mathbb{R}) \to \mathrm{SO}(8,\mathbb{R})$ lift to the triality automorphisms $\mathrm{Spin}(8,\mathbb{R}) \to \mathrm{Spin}(8,\mathbb{R})$.

The center of $\mathrm{Spin}(8,\mathbb{R})$ is the *Klein four-group* $(\mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/2\mathbb{Z})$ because the center of the Clifford group is $\{\pm 1, \pm\gamma_1 \cdots \gamma_8\}$. There are three non-trivial elements of the center, and quotienting by any of these gives you something isomorphic to $\mathrm{SO}(8,\mathbb{R})$. This is special to eight dimensions: in odd dimensions the center of $\mathrm{Spin}(n,\mathbb{R})$ is just $\mathbb{Z}/2\mathbb{Z}$; in dimensions $4k+2$ it is $\mathbb{Z}/4\mathbb{Z}$; and in dimensions $4k \neq 8$ the center is $(\mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/2\mathbb{Z})$ but the quotients are not all isomorphic. $\Diamond$

7.3.3.7 Remark Is $\mathrm{O}(V,\mathbb{K})$ a simple group? No, for the following reasons:

1. There is a determinant map $\mathrm{O}(V,\mathbb{K}) \to \pm 1$, which is usually onto, so it can't be simple.

2. There is a spinor norm map $\mathrm{O}(V,\mathbb{K}) \to \mathbb{K}^\times/\{\text{squares}\}$. Again this is often nontrivial.

3. $-1 \in$ center of $\mathrm{O}(V,\mathbb{K})$, and so the center is a nontrivial normal subgroup.

4. $\mathrm{SO}(V,\mathbb{K})$ tends to split if $\dim V = 4$, tends to be abelian if $\dim V = 2$, and tends to be trivial if $\dim V = 1$.

It turns out that the orthogonal groups are usually simple apart from these four reasons why they're not. Let's take the kernel of the determinant, to get to SO, then look at $\mathrm{Spin}(V,\mathbb{K})$, then quotient by the center (which tends to have order 1, 2, or 4), and assume that $\dim V \geq 5$. Then this is usually simple. If $\mathbb{K}$ is a finite field, then this gives many of the finite simple groups. $\Diamond$

7.3.3.8 Remark $\mathrm{SO}(V,\mathbb{K})$ is not (best defined as) the subgroup of $\mathrm{O}(V,\mathbb{K})$ of elements of determinant 1 in general. Rather it is the image of $\mathrm{CLG}^0(V,\mathbb{K}) \subseteq \mathrm{Cliff}(V,\mathbb{K}) \to \mathrm{O}(V,\mathbb{K})$, which is the correct definition. Let's look at why this is right and the definition you know is wrong. There is a homomorphism, called the *Dickson invariant*, $\mathrm{CLG}(V,K) \to \mathbb{Z}/2\mathbb{Z}$, which takes $\mathrm{CLG}^0(V,\mathbb{K})$ to 0 and $\mathrm{CLG}^1(V,\mathbb{K})$ to 1. It is easy to check that $\det(v) = (-1)^{\mathrm{dickson}(v)}$. So if the characteristic of $\mathbb{K}$ is not 2, $\det = 1$ is equivalent to $\mathrm{dickson} = 0$, but in characteristic 2, determinant is the wrong invariant (because the determinant is always 1). $\Diamond$

7.3.3.9 Example Let us conclude by mentioning some special properties of $\mathrm{O}(1, n; \mathbb{R})$ and $\mathrm{O}(2, n; \mathbb{R})$. First, $\mathrm{O}(1, n; \mathbb{R})$ acts on hyperbolic space $\mathbb{H}^n$, which is a component of norm -1 in $\mathbb{R}^{n,1}$.

Second, $\mathrm{O}(2, n; \mathbb{R})$ acts on the *Hermitian symmetric space* (where *Hermitian* means that it has a complex structure, and *symmetric* means "really nice"). There are three ways to construct this space:

1. It is the set of positive definite two-dimensional subspaces of $\mathbb{R}^{2,n}$.
2. It is the norm-zero vectors $\omega \in \mathbb{PC}^{2,n}$ with $(\omega, \bar{\omega}) = 0$.
3. It is the vectors $x + iy \in \mathbb{R}^{1,n-1}$ with $y \in C$, where the cone C is the interior of the norm-zero cone. $\Diamond$

7.4 $\mathrm{SL}(2, \mathbb{R})$

7.4.1 Finite dimensional representations

The finite-dimensional (complex) representations of the following are essentially the same: $\mathrm{SL}(2, \mathbb{R})$, $\mathfrak{sl}(2, \mathbb{R})$, $\mathrm{SL}(2, \mathbb{C})$ (as a complex Lie group), $\mathfrak{sl}(2, \mathbb{C})$ (as a complex Lie algebra), $\mathrm{SU}(2, \mathbb{R})$, and $\mathfrak{su}(2, \mathbb{R})$. This is because finite dimensional representations of a simply connected Lie group are in bijection with representations of the Lie algebra, and because complex representations of a real Lie algebra $\mathfrak{g}$ correspond to complex representations of its complexification $\mathfrak{g} \otimes \mathbb{C}$ considered as a complex Lie algebra.

7.4.1.1 Remark Representations of a complex Lie algebra $\mathfrak{g} \otimes \mathbb{C}$ are not the same as the representations of the real Lie algebra $\mathfrak{g} \otimes \mathbb{C} = \mathfrak{g} \oplus \mathfrak{g}$. The representations of $\mathfrak{g} \oplus \mathfrak{g}$ correspond roughly to (reps of $\mathfrak{g}$) $\otimes$ (reps of $\mathfrak{g}$). $\Diamond$

7.4.1.2 Remark If $\mathrm{SL}(2, \mathbb{R})$ were simply connected, it would follow from Theorem 3.1.2.1 that the finite-dimensional $\mathbb{C}$- or $\mathbb{R}$-representation theory of $\mathrm{SL}(2, \mathbb{R})$ matched the finite-dimensional representation of theory of $\mathfrak{sl}(2, \mathbb{R})$. Strictly speaking, $\mathrm{SL}(2, \mathbb{R})$ is not simply connected, but as we saw in Example 7.1.4.7, the finite-dimensional representation theory cannot see that $\mathrm{SL}(2, \mathbb{R})$ is not simply connected. $\Diamond$

The finite-dimensional representation theory of $\mathfrak{sl}(2, \mathbb{R})$ is completely described in the following theorem:

7.4.1.3 Theorem (Finite-dimensional representation theory of $\mathfrak{sl}(2, \mathbb{R})$)
For each positive integer n, $\mathfrak{sl}(2, \mathbb{R})$ has one irreducible complex representation of dimension n. All finite-dimensional complex representations of $\mathfrak{sl}(2, \mathbb{R})$ are completely reducible.

Proof (Sketch) One good proof of Theorem 7.4.1.3 is to prove the corresponding statements for $\mathrm{SU}(2)$. Another good proof is essentially the one we gave in Proposition 5.2.0.7 for $\mathfrak{sl}(2, \mathbb{C})$. Complete reducibility follows from the existence of a Casimir: in the basis $H = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, $E = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, and $F = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$ for $\mathfrak{sl}(2, \mathbb{R})$, one choice of Casimir is $2EF + 2FE + H^2 \in \mathcal{U}\mathfrak{sl}(2, \mathbb{R})$. $\square$

Recall that a representation of a group G is *irreducible* if it has no proper subrepresentations, and *completely reducible* if it splits as a direct sum of irreducible G-representations. Complete reducibility makes a representation theory much easier. Some theories with complete reducibility include:

1. Complex representations of a finite group.
2. Unitary representations of any group G (you can take orthogonal complements: if $U \subseteq V$ then $V = U \oplus U^{\perp}$).
3. Hence, representations of any compact group (by averaging, every representation is isomorphic to a unitary one).
4. Finite-dimensional representations of a semisimple Lie group.

See Section 7.2.2 for the full story about unitary representations of compact groups. See Chapter 6 for the full story about complex semisimple Lie groups.

Some theories without complete reducibility include:

1. Representations of a finite group G over fields of characteristic dividing $|G|$.
2. Infinite-dimensional representations of non-compact Lie groups (even if they are semisimple).

In particular, the real Lie group SL$(2,\mathbb{R})$ is not compact. Hence its full representation theory is much more complicated than that of SU(2).

7.4.2 Background about infinite dimensional representations

What is an infinite-dimensional representation of a Lie group G? The most naive guess is that a G-representation should be a Banach space with a (continuous) G action. But from a physical point of view, the actions of $\mathbb{R}$ on L^2 functions, L^1, functions, etc., are all the same, whereas they are complete different as Banach spaces. The second guess is to restrict from Banach spaces to Hilbert spaces, which has the disadvantage that the finite-dimensional representations of SL$(2,\mathbb{R})$ are not Hilbert-space representations, so we would have to throw away some interesting representations.

The solution was found by Harish-Chandra. The point is that if G is a Lie group with Lie algebra $\mathfrak{g} = \mathrm{Lie}(G)$, then $\mathfrak{g}$ acts on any finite-dimensional G-representation, but not, usually, on the infinite-dimensional ones — for example, the $\mathbb{R}$ action on $L^2(\mathbb{R})$ by left translation is infinitesimally generated by $\frac{\mathrm{d}}{\mathrm{d}x}$ acting on $L^2(\mathbb{R})$, but $\frac{\mathrm{d}}{\mathrm{d}x}$ of an L^2 function is not in general L^2. So to get a good category of representations, we add the $\mathfrak{g}$-action back in:

7.4.2.1 Definition *Let G be a Lie group, $\mathfrak{g} = \mathrm{Lie}(G)$, and K a maximal compact subgroup in G. A $(\mathfrak{g}, K)$-module is a vector space V along with actions by $\mathfrak{g}$ and K such that:*

1. *They give the same representations of $\mathfrak{k} = \mathrm{Lie}(K)$.*
2. $\mathrm{Ad}_k(u)\, v = k(u(k^{-1}v))$ *for $k \in K$, $u \in \mathfrak{g}$, and $v \in V$.*

We write $(\mathfrak{g}, K)$-MOD for the category of $(\mathfrak{g}, K)$-modules. A $(\mathfrak{g}, K)$-module is admissible *if as a K-representation, each K-irrep appears as a direct summand only a finite number of times.*

7.4.2.2 Proposition *Let V be a Hilbert space with a G-action. A K-finite vector $v \in V$ is an element of some finite-dimensional sub-K-representation of V, and we set V_ω to be the collection of all K-finite vectors. Then V_ω is a $(\mathfrak{g}, K)$-module, although in general it does not carry an action by G. If V was an irreducible unitary G-module, then V_ω is admissible.* □

Our goal in the next section is to classify the unitary irreducibly representations of $G = \mathrm{SL}(2, \mathbb{R})$. We will do this in several steps:

1. Classify all irreducible admissible $(\mathfrak{g}, K)$ modules. This was solved (for arbitrary simple G) by Langlands, Harish-Chandra, et. al.

2. Figure out which ones have Hermitian inner products. This is easy.

3. Figure out which ones are positive definite. This is very hard, and we'll only do it for $\mathrm{SL}(2, \mathbb{R})$.

7.4.3 The unitary representations of $\mathrm{SL}(2, \mathbb{R})$

Let $G = \mathrm{SL}(2, \mathbb{R})$ with Lie algebra $\mathfrak{g} = \mathfrak{sl}(2, \mathbb{R})$. A maximal compact subgroup is the rotation group $K = \left\{ \left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix} \right) \right\}$, generated by $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix} \right)$. All representations will be complex, and hence the $\mathfrak{g}$ action extends to a $\mathfrak{g} \otimes \mathbb{C} = \mathfrak{sl}(2, \mathbb{C})$ action. We take the basis $H = -i \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix} \right)$, $E = \frac{1}{2} \left(\begin{smallmatrix} 1 & i \\ i & -1 \end{smallmatrix} \right)$, and $F = \frac{1}{2} \left(\begin{smallmatrix} 1 & -i \\ -i & -1 \end{smallmatrix} \right)$. These satisfy $[H, E] = 2E$, $[H, F] = -2F$, and $[E, F] = H$, and iH generates K. We begin by studying the irreducible $(\mathfrak{g}, K)$ modules.

7.4.3.1 Remark The group $\mathrm{SL}(2, \mathbb{R})$ has two different classes of Cartan subgroups — the rotations $\left(\begin{smallmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{smallmatrix} \right)$ and the scalings $\left(\begin{smallmatrix} a & 0 \\ 0 & a^{-1} \end{smallmatrix} \right)$ — and the rotation Cartan is the maximal compact subgroup. Non-compact abelian groups need not have eigenvectors in infinite-dimensional spaces, whereas compact ones do, and our strategy, as always, is to study a representation by studying its eigenvalues on the Cartan subalgebra. ♢

Given an irreducible $V \in (\mathfrak{g}, K)$-MOD, we can write it as a direct sum of eigenspaces of $H = -i \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix} \right)$, as iH generates the compact group $K = S^1$. Moreover, *all eigenvalues of H are integers.* If V were finite-dimensional, the highest eigenvalue would give us complete control. Instead, we look at the Casimir $\Omega = 2EF + 2FE + H^2 + 1$ — we have added 1 to the usual Casimir so that some numerology works out in integers later. Since Ω commutes with G and V is irreducible, Ω acts as a scalar on V; we set $\lambda \overset{\text{def}}{=} \sqrt{\Omega|_V}$ to be the square root of this scalar.

Set V_n to be the subspace of V on which H has eigenvalue $n \in \mathbb{Z}$. A standard calculation shows that $HEv = (n+2)Ev$ and $HFv = (n-2)Fv$, so that E, F move $V_n \to V_{n\pm 2}$. Since $\Omega = 4FE + H^2 + 2H + 1$ (using $[E, F] = H$), and since $\Omega v = \lambda^2 v$, we see that $FEv = \frac{1}{4}\big(\lambda^2 - (n+1)^2\big)v$, and in particular we have shown that any H-eigenvector in an irreducible $(\mathfrak{g}, K)$-module is also an FE-eigenvector.

Moreover, if V is irreducible, then there cannot be more than one dimension at each weight n, and V_ω is spanned by the weight spaces. Notice that the weight spaces are linearly independent, as they support different eigenvalues. The picture of any irreducible $(\mathfrak{g}, K)$-module is a chain:

$$\begin{tikzcd}
\cdots \arrow[r, bend left, "E"] & v_{n-4} \arrow[l, bend left, "F"] \arrow[r, bend left, "E"] & v_{n-2} \arrow[l, bend left, "F"] \arrow[r, bend left, "E"] & v_n \arrow[l, bend left, "F"] \arrow[r, bend left, "E"] & v_{n+2} \arrow[l, bend left, "F"] \arrow[r, bend left, "E"] & v_{n+4} \arrow[l, bend left, "F"] \arrow[r, bend left, "E"] & \cdots \arrow[l, bend left, "F"]
\end{tikzcd}$$

Each v_n is an eigenvector of H with weight n, and a basis for the corresponding weight space. The map E moves us up the chain $n \mapsto n+2$, and F moves us down, and we should pick a normalization for bases so that $FEv_n = \frac{1}{4}\big(\lambda^2 - (n+1)^2\big)v_n$ and $EFv_n = \frac{1}{4}\big(\lambda^2 - (n-1)^2\big)v_n$: for example, supposing neither acts as 0, we could make E take basis vectors to basis vectors and F multiply by the correct eigenvalue.

There are four possible shapes for such a chain. It might be infinite in both directions (neither a highest weight nor a lowest weight), infinite to only the left (a highest weight but no lowest weight), infinite to only the right (a lowest weight but no highest weight), or finite (both a highest and a lowest weight). We'll see that all these show up. We also see that an irreducible representation is completely determined once we know λ and some n for which $V_n \neq 0$. The remaining question is to construct representations with all possible values of $\lambda \in \mathbb{C}$ and $n \in \mathbb{Z}$.

7.4.3.2 Example If n is even, it is easy to check that the following is a $(\mathfrak{g}, K)$-representation, although it might not be irreducible:

$$\begin{tikzcd}
\cdots \arrow[r, bend left, "\frac{\lambda-5}{2}"] & v_{-4} \arrow[l, bend left, "\frac{\lambda+5}{2}"] \arrow[r, bend left, "\frac{\lambda-3}{2}"] & v_{-2} \arrow[l, bend left, "\frac{\lambda+3}{2}"] \arrow[r, bend left, "\frac{\lambda-1}{2}"] & v_0 \arrow[l, bend left, "\frac{\lambda+1}{2}"] \arrow[r, bend left, "\frac{\lambda+1}{2}"] & v_2 \arrow[l, bend left, "\frac{\lambda-1}{2}"] \arrow[r, bend left, "\frac{\lambda+3}{2}"] & v_4 \arrow[l, bend left, "\frac{\lambda-3}{2}"] \arrow[r, bend left, "\frac{\lambda+5}{2}"] & \cdots \arrow[l, bend left, "\frac{\lambda-5}{2}"]
\end{tikzcd}$$

I.e. our representation is spanned by basis vectors v_n for $n \in 2\mathbb{Z}$, with $Hv_n = nv_n$, $Ev_n = \frac{\lambda+n+1}{2}v_{n+2}$, and $Fv_n = \frac{\lambda-n+1}{2}v_{n-2}$. $\Diamond$

How can a chain fail to be infinite? Alternately, how can an infinite chain fail to be irreducible? These can only happen when some Ev_n or some Fv_n vanishes — otherwise, from any vector you can generate the whole space. E, F can act as zero only when:

n is even and λ an odd integer.
n is odd and λ an even integer.

To illustrate what happens, we give two examples:

7.4.3.3 Example Take n even and $\lambda = 3$. Then the chain from Example 7.4.3.2 looks like:

$$\begin{tikzcd}
V_- & & & & & & & V_+ \\
\cdots \arrow[r, bend left, "-2"] & v_{-6} \arrow[l, bend left, "5"] \arrow[r, bend left, "-1"] & v_{-4} \arrow[l, bend left, "4"] \arrow[r, bend left, "0"] & v_{-2} \arrow[l, bend left, "3"] \arrow[r, bend left, "1"] & v_0 \arrow[l, bend left, "2"] \arrow[r, bend left, "2"] & v_2 \arrow[l, bend left, "1"] \arrow[r, bend left, "3"] & v_4 \arrow[l, bend left, "0"] \arrow[r, bend left, "4"] & v_6 \arrow[l, bend left, "-1"] \arrow[r, bend left, "5"] & \cdots \arrow[l, bend left, "-2"]
\end{tikzcd}$$

The two rays $V_\pm$ are irreducible subrepresentations, and $V/(V_+ \oplus V_-)$ is a three-dimensional irreducible representation. $\Diamond$

7.4.3.4 Example Take n even and $\lambda = -3$. Then our picture is:

$$\begin{tikzcd}
\cdots \arrow[r, bend left, "-5"] & v_{-6} \arrow[l, bend left, "2"] \arrow[r, bend left, "-4"] & v_{-4} \arrow[l, bend left, "1"] \arrow[r, bend left, "-3"] & v_{-2} \arrow[l, bend left, "0"] \arrow[r, bend left, "-2"] & v_0 \arrow[l, bend left, "-1"] \arrow[r, bend left, "-1"] & v_2 \arrow[l, bend left, "-2"] \arrow[r, bend left, "0"] & v_4 \arrow[l, bend left, "-3"] \arrow[r, bend left, "1"] & v_6 \arrow[l, bend left, "-4"] \arrow[r, bend left, "2"] & \cdots \arrow[l, bend left, "-5"]
\end{tikzcd}$$

So V has a three-dimensional irreducible subrepresentation, and the quotient is a direct sum of two rays. $\Diamond$

All together, we have:

7.4.3.5 Proposition *The irreducible $(\mathfrak{g}, K)$-representations consist of:*

1. *For each $\lambda \in \mathbb{C} \smallsetminus \mathbb{Z}$, modulo $\lambda \equiv -\lambda$, we have two both-ways-infinite irreducible representations: one for even weights and one for odd weights. For $\lambda \in 2\mathbb{Z}$ there is a both-ways-infinite irreducible representation with even weights, and for $\lambda \in 2\mathbb{Z}+1$ there is a both-ways-infinite irreducible representation with odd weights.*

2. *For each $\lambda \in \mathbb{Z}_{\geq 0}$, there are two half-infinite* discrete series representations*: one with highest weight $-\lambda-1$ and one with lowest weight $\lambda+1$.*

3. *For each $\lambda \in \mathbb{Z}_{\leq -1}$, we have a $(-\lambda)$-dimensional irreducible representation, with weights in $\{\lambda+1, \lambda+3, \ldots, -\lambda-1\}$.* $\square$

7.4.3.6 Remark One can index the two discrete series for $\lambda \neq 0$ by calling one the "positive-λ" series and the other the "negative-λ" series, thereby using numbers $\lambda \in \mathbb{Z} \smallsetminus \{0\}$. Then the two series for $\lambda = 0$ are called "limits of discrete series". $\Diamond$

Which of these can be made into *unitary* representations? Recall that we have been working with the basis $H = -i\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, $E = \frac{1}{2}\left(\begin{smallmatrix} 1 & i \\ i & -1 \end{smallmatrix}\right)$, and $F = \frac{1}{2}\left(\begin{smallmatrix} 1 & -i \\ -i & -1 \end{smallmatrix}\right)$ for $\mathfrak{sl}(2,\mathbb{C})$. Recall that in a unitary representation of G, we must have $x^* = -x$ for $x \in \mathfrak{g}$; hence if a $(\mathfrak{g}, K)$-module restricts to a unitary representation of $\mathfrak{sl}(2,\mathbb{R})$, then it must satisfy $H^* = H$, $E^* = -F$, and $F^* = -E$.

Therefore if we have a Hermitian inner product $(,)$ on an irreducible representation, it must satisfy:

$$\begin{aligned}(v_{n+2}, v_{n+2}) &= \left(\frac{2}{\lambda+n+1}Ev_n, \frac{2}{\lambda+n+1}Ev_n\right)\\ &= \frac{4}{(\bar{\lambda}+n+1)(\lambda+n+1)}(Ev_n, Ev_n)\\ &= \frac{4}{(\bar{\lambda}+n+1)(\lambda+n+1)}(v_n, -FEv_n)\\ &= \frac{4}{(\bar{\lambda}+n+1)(\lambda+n+1)}\left(v_n, -\frac{(\lambda+n+1)(\lambda-n-1)}{4}v_n\right)\\ &= -\frac{\lambda-n-1}{\bar{\lambda}+n+1}(v_n, v_n)\end{aligned}$$

Thus, if we are to have a unitary representation, we must have $-(\lambda-n-1)(\bar{\lambda}+n+1)^{-1} \in \mathbb{R}_{>0}$, or equivalently $(n+1)^2 - \lambda^2 \in \mathbb{R}_{>0}$, for all weights n other than the top weight (if it exists). Conversely, if $(n+1)^2 - \lambda^2 \in \mathbb{R}_{>0}$ for all non-top weights n, then the corresponding irreducible $(\mathfrak{g}, K)$-module can be made unitary. Inspecting the list in Proposition 7.4.3.5, we find:

7.4.3.7 Proposition *The irreducible unitary $(\mathfrak{g}, K)$-modules consist of:*

1. *Both-ways-infinite irreducible chains with $\lambda^2 \le 0$. These are called the* principal series representations. *(When $\lambda = 0$ and n is odd, the both-ways infinite chain splits as a direct sum of two limits of discrete series representations; both are unitary.)*

2. *Both-ways-infinite chains with j even and $0 < \lambda < 1$. These are called* complementary series representations. *They are annoying, and you spend a lot of time trying to show that they don't occur.*

3. *The* discrete series representations*: half-infinite chains with $\lambda \in \mathbb{Z}_{\ge 0}$ (λ and n must have opposite parity). (When $\lambda = 0$, the half-infinite chains are called* limits of discrete series representations.*)*

4. *The one-dimensional representation.*

In particular, finite-dimensional representations that are not the trivial representation are not unitary. □

7.4.3.8 Remark The nice stuff that happened for SL(2, ℝ) breaks down for more complicated Lie groups. ◊

7.4.3.9 Remark Representations of finite covers of SL(2, ℝ) are similar, except that the weights n need not be integral. For example, for the *metaplectic group* Mp(2, ℝ), the double cover of SL(2, ℝ), the weights (eigenvalues of H) must be half-integers. ◊

Exercises

1. Show that if G is an abelian compact connected Lie group, then it is a product of circles, so it is $\mathbb{T}^n$.

2. If G is compact and connected, show that its left-invariant volume form ω is also right invariant. Even if G is not compact but not connected, show that the measure $|\omega|$ obtained from a left invariant form $\omega \in \bigwedge^{\text{top}} \mathrm{T}G$ agrees with the measure obtained from a right invariant form.

   Show that the left- and right-invariant volume forms on G do not agree for G a non-abelian connected Lie group of dimension 2.

3. If you haven't already, prove that the Lie algebra of a solvable group is solvable.

4. Find the structure of $\mathrm{Cliff}(m, n; \mathbb{R})$, the Clifford algebra over $\mathbb{R}^{n+m}$ with the form $x_1^2 + \cdots + x_m^2 - x_{m+1}^2 - \cdots - x_{m+n}^2$.

5. Let $\mathbb{K} = \mathbb{F}_2$, $H = \mathbb{K}^2$ with the quadratic form $x^2 + y^2 + xy$, and $V = H \oplus H$. Prove the assertion in Example 7.3.2.11 that $\mathrm{O}(V, \mathbb{K})$ is not generated by reflections.

6. Prove that $\mathrm{Spin}(3, 3; \mathbb{R}) \cong \mathrm{SL}(4, \mathbb{R})$.

7. Show that the three descriptions in Example 7.3.3.9 of the Hermitian symmetric space are the same.

8. Check the assertion in Example 7.4.3.2, and find a similar representation for n odd.

9. Prove that when n is odd and $\lambda = 0$, then the both-ways-infinite chain V from the previous exercise (corresponding to the one in Example 7.4.3.2) splits as a direct sum of a "negative ray" and a "positive ray".

10. Classify the irreducible unitary representations of $\mathrm{Mp}(2, \mathbb{R})$.

Chapter 8

From Dynkin diagram to Lie group, revisited

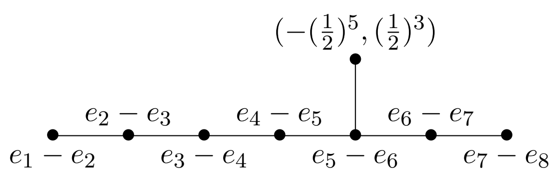
In Section 5.6 we described a construction that begins with a Dynkin diagram (i.e. Cartan matrix) and constructed a Lie algebra — Lie algebra in hand, one can use Chapters 3 and 4 to construct a real Lie group, and we gave a different construction of the complex algebraic group for a given Dynkin diagram in Section 6.2.3. Our construction of the Lie algebra required the somewhat unenlightening Serre relations. We will try now to explain the construction in more detail, with E_8 as our primary example. (In the E_8 case specifically, one can construct the E_8 Lie algebra as a sum of the D_8 Lie algebra and a half-spin representation.) We will go on to describe how to find real forms of a given complex semisimple Lie algebra, and conclude by describing all simple real Lie groups.

8.1 E_8

8.1.1 The E_8 lattice

We introduce the following notation for vectors: we denote repetitions by exponents, so that $(1^8) = (1,1,1,1,1,1,1,1)$ and $(\pm(\frac{1}{2})^2, 0^6) = (\pm\frac{1}{2}, \pm\frac{1}{2}, 0,0,0,0,0,0)$.

Recall the Dynkin diagram for E_8:



Each vertex is a simple root r with $(r,r) = 2$; $(r,s) = 0$ when r and s are not joined, and $(r,s) = -1$ when r and s are joined. We choose an orthonormal basis $e_1, \ldots, e_8$, in which the simple roots are as given.

8.1.1.1 Example We want to figure out what the *root lattice* L of E_8 is (this is the lattice generated by the roots). If you take $\{e_i - e_{i+1}\} \cup (-1^5, 1^3)$ (all the A_7 vectors plus twice the strange vector), they generate the D_8 lattice $= \{(x_1, \ldots, x_8)$ s.t. $x_i \in \mathbb{Z}$ and $\sum x_i$ is even$\}$. So the E_8 lattice consists of two cosets of this lattice, where the other coset is $\{(x_1, \ldots, x_8)$ s.t. $x_i \in \mathbb{Z} + \frac{1}{2}$ and $\sum x_i$ is odd$\}$.

Alternative version: If you reflect this lattice through the hyperplane $e_1^\perp$, then you get the same thing except that $\sum x_i$ is always even. We will freely use both characterizations, depending on which is more convenient for the calculation at hand. $\diamondsuit$

8.1.1.2 Example We should also work out the *weight lattice*, which consists of the vectors s such that $(r,r)/2$ divides (r,s) for all roots r. Notice that the weight lattice of E_8 is contained in the weight lattice of D_8, which is the union of four cosets of D_8: D_8, $D_8 + (1, 0^7)$, $D_8 + ((\frac{1}{2})^8)$ and $D_8 + (-\frac{1}{2}, (\frac{1}{2})^7)$. Which of these have integral inner product with the vector $(-(\frac{1}{2})^5, (\frac{1}{2})^3)$? They are the first and the last, so the weight lattice of E_8 is $D_8 \cup D_8 + (-\frac{1}{2}, (\frac{1}{2})^7)$, which is equal to the root lattice of E_8. $\diamondsuit$

8.1.1.3 Definition *The* dual *of a lattice $L \in \mathbb{R}^n$ is the lattice consisting of vectors having integral inner product with all lattice vectors. A lattice is* unimodular *if it is equal to its dual. It is* even *if the inner product of any vector with itself is always even.*

So Examples 8.1.1.1 and 8.1.1.2 show that E_8 is unimodular (as are G_2 and F_4 but not general Lie algebra lattices) and even.

8.1.1.4 Remark Even unimodular lattices in $\mathbb{R}^n$ only exist if $8|n$ (this 8 is the same 8 that shows up in the periodicity of Clifford groups). The E_8 lattice is the only example in dimension equal to 8 (up to isomorphism, of course). There are two in dimension 16 (one of which is $L \oplus L$, the other is $D_{16}\cup$ some coset). There are 24 in dimension 24, which are the *Niemeier lattices*. In 32 dimensions, there are more than a billion! $\diamondsuit$

8.1.2 The E_8 Weyl group

The *Weyl group* $\mathfrak{W}(E_8)$ of E_8 is generated by the reflections through $s^\perp$ where $s \in L$ and $(s,s) = 2$ (these are called *roots*).

8.1.2.1 Example First, let's find all the roots: $(x_1, \ldots, x_8)$ such that $\sum x_i^2 = 2$ with all $x_i \in \mathbb{Z}$ or all in $\mathbb{Z} + \frac{1}{2}$ and $\sum x_i$ even. If $x_i \in \mathbb{Z}$, obviously the only solutions are permutations of $(\pm 1, \pm 1, 0^6)$, of which there are $\binom{8}{2} \times 2^2 = 112$ choices. In the $\mathbb{Z} + \frac{1}{2}$ case, you can choose the first 7 places to be $\pm\frac{1}{2}$, and the last coordinate is forced, so there are 2^7 choices. Thus, you get 240 roots. $\diamondsuit$

8.1.2.2 Example Let's find the orbits of the roots under the action of the Weyl group. We don't yet know what the Weyl group looks like, but we can find a large subgroup that is easy to work with. Let's use the $\mathfrak{W}(D_8)$ (the Weyl group of D_8), which consists of the following: we can apply all permutations of the coordinates, or we can change the sign of an even number of coordinates: e.g., reflection in $(1, -1, 0^6)$ swaps the first two coordinates, and reflection in $(1, -1, 0^6)$ followed by reflection in $(1, 1, 0^6)$ changes the sign of the first two coordinates.

Notice that under $\mathfrak{W}(D_8)$, the roots form two orbits: the set which is all permutations of $(\pm 1^2, 0^6)$, and the set $(\pm(\frac{1}{2})^8)$. Do these become the same orbit under the Weyl group of E_8? Yes; to show this, we just need one element of $\mathfrak{W}(E_8)$ taking some element of the first orbit to the second orbit. Take reflection in $((\frac{1}{2})^8)^\perp$ and apply it to $(1^2, 0^6)$: you get $((\frac{1}{2})^2, -(\frac{1}{2})^6)$, which is in the second orbit. So there is just one orbit of roots under the Weyl group. $\Diamond$

What do orbits of $\mathfrak{W}(E_8)$ on other vectors look like? We're interested in this because we might want to do representation theory. The character of a representation is a map from weights to integers, and it is $\mathfrak{W}(E_8)$-invariant.

8.1.2.3 Example Let's look at vectors of norm 4. So $\sum x_i^2 = 4$, $\sum x_i$ even, and all $x_i \in \mathbb{Z}$ or all $x_i \in \mathbb{Z} + \frac{1}{2}$. There are 8×2 possibilities which are permutations of $(\pm 2, 0^7)$. There are $\binom{8}{4} \times 2^4$ permutations of $(\pm 1^4, 0^4)$, and there are 8×2^7 permutations of $(\pm\frac{3}{2}, \pm(\frac{1}{2})^7)$. So there are a total of 240×9 of these vectors. There are 3 orbits under $\mathfrak{W}(D_8)$, and as before, they are all one orbit under the action of $\mathfrak{W}(E_8)$. To see this, just reflect $(2, 0^7)$ and $(1^3, -1, 0^4)$ through $((\frac{1}{2})^8)$. $\Diamond$

In Exercise 1 you will prove that there are 240×28 vectors of norm 6, and that they all form one orbit. For norm 8 there are two orbits, because you have vectors that are twice a norm 2 vector, and vectors that aren't. As the norm gets bigger, you'll get a large number of orbits.

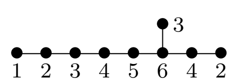
8.1.2.4 Remark If you've seen a course on modular forms, you'll know that the number of vectors of norm $2n$ is given by $240 \times \sum_{d|n} d^3$. If you call these c_n, then $\sum c_n q^n$ is a modular form of level 1 (E_8 is even and unimodular) and weight 4 ($= \dim E_8/2$). $\Diamond$

What is the order of the Weyl group of E_8? We'll do this by 4 different methods, which illustrate the different techniques for this kind of thing:

8.1.2.5 Example (Order of $\mathfrak{W}(E_8)$, method 1) This is a good one as a mnemonic. The order of E_8 is given by:

$$|\mathfrak{W}(E_8)| = 8! \times \prod \begin{pmatrix} \text{numbers on the} \\ \text{affine } E_8 \text{ diagram} \end{pmatrix} \times \frac{\text{Weight lattice of } E_8}{\text{Root lattice of } E_8}$$

$$= 8! \times \left(\quad \right) \times 1$$

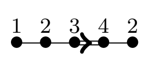


$$= 2^{14} \times 3^5 \times 5^2 \times 7$$

By "numbers on the affine diagram" we mean: take the corresponding affine diagram, and write down the coefficients on it of the highest root. Notice that the "affine E_8" is the diagram E_9.

We can do the same thing for any other Lie algebra, for example:

$$|\mathfrak{W}(F_4)| = 4! \times \left(\quad \right) \times 1 = 2^7 \times 3^2 \qquad \Diamond$$



8.1.2.6 Example (Order of $\mathfrak{W}(E_8)$, method 2) The order of a reflection group is equal to the products of degrees of the fundamental invariants. For E_8, the fundamental invariants are of degrees $2, 8, 12, 14, 18, 20, 24, 30$. Incidentally, other than the 2, these are all one more than primes. $\Diamond$

8.1.2.7 Example (Order of $\mathfrak{W}(E_8)$, method 3) This one is actually an honest method (without quoting weird facts). The only fact we will use is the following: suppose G acts transitively on a set X with $H =$ the group fixing some point; then $|G| = |H| \cdot |X|$.

This is a general purpose method for working out the orders of groups. First, we need a set acted on by the Weyl group of E_8. Let's take the root vectors (vectors of norm 2). This set has 240 elements, and the Weyl group of E_8 acts transitively on it. So $|\mathfrak{W}(E_8)| = 240 \times |$subgroup fixing $(1, -1, 0^6)|$. But what is the order of this subgroup (call it G_1)? Let's find a set acted on by this group. It acts on the set of norm 2 vectors, but the action is not transitive. What are the orbits? G_1 fixes $r = (1, -1, 0^6)$. For other roots s, G_1 obviously fixes (r, s). So how many roots are there with a given inner product with r?

(s,r)	number	choices
2	1	r
1	56	$(1, 0, \pm 1^6), (0, -1, \pm 1^6), (\frac{1}{2}, -\frac{1}{2}, (\frac{1}{2})^6)$
0	126	some list
-1	56	some list
-2	1	$-r$

So there are at least five orbits under G_1. In fact, each of these sets is a single orbit under G_1. We can see this by finding a large subgroup of G_1. Take $\mathfrak{W}(D_6)$, which is all permutations of the last six coordinates and all even sign changes of the last six coordinates. It is generated by reflections associated to the roots orthogonal to e_1 and e_2 (those that start with two 0s). The three cases with inner product 1 are each orbits under $\mathfrak{W}(D_6)$. So to see that there is a single orbit under G_1, we just need some reflections that mess up these orbits. If you take a vector $(\frac{1}{2}, \frac{1}{2}, \pm(\frac{1}{2})^6)$ and reflect norm-2 vectors through it, you will get exactly 5 orbits. So G_1 acts transitively on the sets of roots with a prescribed inner product with r.

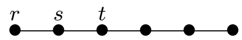
We'll use the orbit of vectors s with $(r, s) = -1$. Let G_2 be the vectors fixing s and r: $\overset{r}{\bullet}\!\!-\!\!\!-\!\!\overset{s}{\bullet}$. We have that $|G_1| = |G_2| \cdot 56$.

We will press on, although it get's tedious. Our plan is to chose vectors acted on by G_i and fixed by G_{i+1} which give us the Dynkin diagram of E_8. So the next step is to try to find vectors t that give us the picture $\overset{r}{\bullet}\!\!-\!\!\!-\!\!\overset{s}{\bullet}\!\!-\!\!\!-\!\!\overset{t}{\bullet}$, which is to say they have inner product -1 with s and 0 with r. The possibilities for t are $(-1, -1, 0, 0^5)$ (one of these), $(0, 0, 1, \pm 1, 0^4)$ and permutations of its last five coordinates (10 of these), and $(-\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, \pm(\frac{1}{2})^5)$ (there are 16 of these), so we get 27 total. Then we could check that they form one orbit, which is boring.

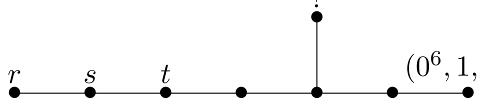
Next find vectors which go next to t in our picture $\overset{r}{\bullet}\!\!-\!\!\!-\!\!\overset{s}{\bullet}\!\!-\!\!\!-\!\!\overset{t}{\bullet}\!\!-\!\!\!-\!\!\overset{?}{\bullet}$, i.e. vectors whose inner product is -1 with t and zero with r, s. The possibilities are permutations of the last four coordinates of $(0, 0, 0, 1, \pm 1, 0^3)$ (8 of these) and $(-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, \pm(\frac{1}{2})^4)$ (8 of these), so there are 16 total. Again check transitivity.

Find a fifth vector: the possibilities are $(0^4, 1, \pm 1, 0^2)$ and permutations of the last three coordinates (6 of these), and $(-(\frac{1}{2})^4, \frac{1}{2}, \pm(\frac{1}{2})^3)$ (4 of these) for a total of 10.

For the sixth vector, we can have $(0^5, 1, \pm 1, 0)$ or $(0^5, 1, 0, \pm 1)$ (4 possibilites) or $(-(\frac{1}{2})^5, \frac{1}{2}, \pm(\frac{1}{2})^2)$ (2 possibilities), so we get 6 total.

The next case — finding the seventh vector — is tricky. The possibilities are $(0^6, 1, \pm 1)$ (2 of these) and $((-\frac{1}{2})^6, \frac{1}{2}, \frac{1}{2})$ (just 1). The proof of transitivity fails at this point. By now the group we're using ($\mathfrak{W}(D_6)$ and one more reflection) doesn't even act transitively on the pair (you can't get between them by changing an even number of signs). What elements of $\mathfrak{W}(E_8)$ fix all these first six points  ?

We want to find roots perpendicular to all of these vectors, and the only possibility is $((\frac{1}{2})^8)$. How does reflection in this root act on the three possible seventh vectors? $(0^6, 1^2) \mapsto ((-\frac{1}{2})^6, (\frac{1}{2})^2)$ and $(0^6, 1, -1)$ maps to itself. Is this last vector in the same orbit? In fact they are in different orbits. To see this, look for vectors that complete the E_8 diagram:



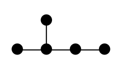
In the $(0^6, 1, 1)$ case, you can take the vector $((-\frac{1}{2})^5, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2})$. But in the other case, you can show that there are no possibilities. So these really are different orbits.

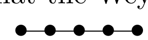
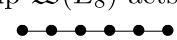
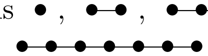
Use the orbit with two elements, and you get

$$|\mathfrak{W}(E_8)| = 240 \times \underbrace{56 \times \overbrace{27 \times 16 \times 10 \times 6 \times 2 \times 1}^{\text{order of } \mathfrak{W}(E_6)}}_{\text{order of } \mathfrak{W}(E_7)}$$

because the group fixing all 8 vectors must be trivial. You also get that

$$|\mathfrak{W}(\text{“}E_5\text{”})| = 16 \times \underbrace{10 \times \overbrace{6 \times 2 \times 1}^{|\mathfrak{W}(A_2 \times A_1)|}}_{|\mathfrak{W}(A_4)|}$$

where "E_5" is the algebra with diagram , also known as D_5. Similarly, $E_4 = A_4$ and $E_3 = A_2 \times A_1$.

We got some other information. We found that the Weyl group $\mathfrak{W}(E_8)$ acts transitively on all the configurations , , , , , and  (A_1 through A_6), but not on $A_7 =$ . ♢

8.1.2.8 Example (Order of $\mathfrak{W}(E_8)$, method 4) Let L denote the E_8 lattice. Look at $L/2L$, which has 256 elements, as a set acted on by $\mathfrak{W}(E_8)$. There is an orbit of size one (represented by 0). There is an orbit of size $240/2 = 120$, consisting of the roots (a root is congruent mod $2L$ to its negative). Left over are 135 elements. Let's look at norm-4 vectors. Each norm-4 vector r satisfies $r \equiv -r \mod 2$, and there are $240 \cdot 9$ of them, which is a lot, so norm-4 vectors must be congruent to a bunch of stuff. Let's look at $r = (2, 0^7)$. Notice that it is congruent to vectors of the form $(0^a, \pm 2, 0^b)$, of which there are sixteen. It is easy to check that these are the only norm-4 vectors congruent to r mod 2. So we can partition the norm-4 vectors into $240 \cdot 9/16 = 135$ subsets of 16 elements. So $L/2L$ has $1 + 120 + 135$ elements, where 1 is the zero, each element among the 120 is represented by two elements of norm 2, and each of the 135 is represented by sixteen elements of norm 4. A set of sixteen elements of norm 4 which are all congruent is called a *frame*. It consists of elements $\pm v_1, \dots, \pm v_8$, where $v_i^2 = 4$ and $(v_i, v_j) = 1$ for $i \neq j$, so up to sign it is an orthogonal basis.

We know that $\mathfrak{W}(E_8)$ acts transitively on frames, and so:

$$|\mathfrak{W}(E_8)| = (\#\text{ of frames}) \times |\text{subgroup fixing a frame}|$$

So we need to know what are the automorphisms of a frame. A frame consists of eight subsets of the form $(r, -r)$, and isometries of a frame form the group $(\mathbb{Z}/2\mathbb{Z})^8 \cdot S_8$ (warning: sometimes $A \cdot B$ specifically means an extension that does not split; we are using it to mean any extension), but these may not all be in the Weyl group. In the Weyl group, we found a $(\mathbb{Z}/2\mathbb{Z})^7 \cdot S_8$, where the first part is the group of sign changes of an *even* number of coordinates. So the subgroup fixing a frame must be in between $(\mathbb{Z}/2\mathbb{Z})^7 \cdot S_8$ and $(\mathbb{Z}/2\mathbb{Z})^8 \cdot S_8$, and since these groups differ by a factor of 2, it must be one of them. Observe that changing an odd number of signs doesn't preserve the E_8 lattice, so the subgroup fixing a frame must be $(\mathbb{Z}/2\mathbb{Z})^7 \cdot S_8$, which has order $2^7 \cdot 8!$. So the order of the Weyl group is

$$135 \cdot 2^7 \cdot 8! = |2^7 \cdot S_8| \times \frac{\#\text{ norm-4 elements}}{2 \times \dim L}$$

◊

8.1.2.9 Remark Similarly, if Λ denotes the Leech lattice, you actually get the order of Conway's group is:

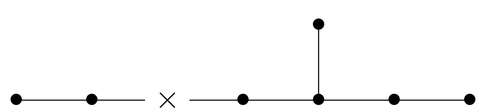
$$\left|2^{12} \cdot M_{24}\right| \cdot \frac{\#\text{ norm-8 elements}}{2 \times \dim \Lambda}$$

Here M_{24} is the Mathieu group (one of the sporadic simple groups). The Leech lattice seems very much to be trying to be the root lattice of the monster group, or something like that. There are a lot of analogies, but nobody can make sense of it. ◊

8.1.2.10 Remark $\mathfrak{W}(E_8)$ acts on $(\mathbb{Z}/2\mathbb{Z})^8$, which is a vector space over $\mathbb{F}_2$, with quadratic form $N(a) = \frac{(a,a)}{2} \mod 2$. Thus we get a map $\mathfrak{W}(E_8) \to \mathrm{O}^+(8, \mathbb{F}_2)$ with kernel $\{\pm 1\}$, and it is surjective. Here $\mathrm{O}^+(8, \mathbb{F}_2)$ denotes one of the 8-dimensional orthogonal groups over $\mathbb{F}_2$. So $\mathfrak{W}(E_8)$ is very close to being an orthogonal group of a characteristic-2 vector space. ◊

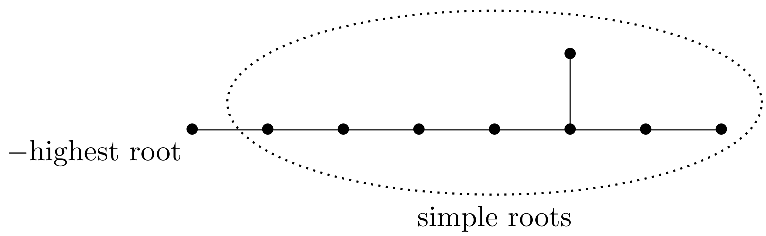
What is inside the root lattice/Lie algebra/Lie group E_8? One obvious way to find things inside is to cover nodes of the E_8 diagram.

8.1.2.11 Example



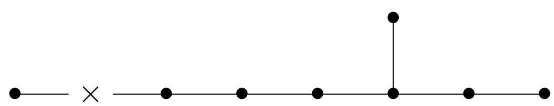
If we remove the shown node, you see that E_8 contains $A_2 \times D_5$. ◊

We can do better by showing that we can embed the affine $\tilde{E}_8$ root system into the E_8 lattice.



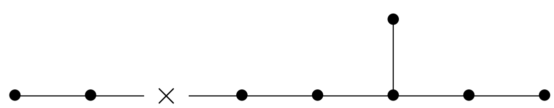
Now you can remove nodes here and get some bigger sub-diagrams.

8.1.2.12 Example Work with $\tilde{E}_8$ as above, and cover:



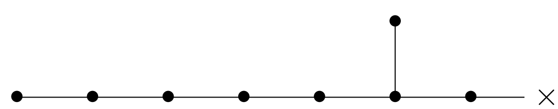
We get that an $A_1 \times E_7$ in E_8. The E_7 consisted of 126 roots orthogonal to a given root. This gives an easy construction of the E_7 root system as all the elements of the E_8 lattice perpendicular to $(1, -1, 0 \ldots)$. $\diamondsuit$

8.1.2.13 Example Alternately, we can cover:



Then we get an $A_2 \times E_6$, where the E_6 are all the vectors with the first 3 coordinates equal. So we get the E_6 lattice for free too. $\diamondsuit$

8.1.2.14 Example



We see that there is a D_8 in E_8, which is all vectors of the E_8 lattice with integer coordinates. We sort of constructed the E_8 lattice this way in the first place. $\diamondsuit$

We can ask questions like: What is the E_8 Lie algebra as a representation of D_8? To answer this, we look at the weights of the E_8 algebra, considered as a module over D_8: the 112 roots of the form $(0^a, \pm 1, 0^b, \pm 1, 0^c)$, the 128 roots of the form $(\pm\frac{1}{2}, \pm\frac{1}{2}, \ldots)$, and the vector 0 with multiplicity 8. These give you the Lie algebra of D_8. Recall that D_8 is the Lie algebra of $\mathrm{SO}(16)$. The double cover has a half-spin representation of dimension $2^{16/2-1} = 128$. So E_8 decomposes as a representation of D_8 as the adjoint representation (of dimension 120) plus a half-spin representation of dimension 128. This is often used to construct the Lie algebra E_8. We'll do a better construction in Section 8.2.1.

8.1.2.15 Example We've found that the Lie algebra of D_8, which is the Lie algebra of SO(16), is contained in the Lie algebra of E_8. Which *group* is contained in the compact form of the E_8? The simply-connected group with Lie algebra $\mathfrak{so}(16,\mathbb{R})$ is $\mathrm{Spin}(16,\mathbb{R})$, and so the full list of groups corresponds to the list of subgroups of the center $(\mathbb{Z}/2\mathbb{Z})^2$ (c.f. Example 7.1.2.1):

$$\begin{tikzcd}[arrows=dash] & 1 \arrow[dl] \arrow[d] \arrow[dr] & \\ (1,-1) \arrow[dr] & (-1,-1) \arrow[d] & (-1,1) \arrow[dl] \\ & (\mathbb{Z}/2\mathbb{Z})^2 & \end{tikzcd} \qquad \begin{tikzcd}[arrows=dash] & \mathrm{Spin}(16,\mathbb{R}) \arrow[dl] \arrow[d] \arrow[dr] & \\ \mathrm{Spin}(16,\mathbb{R})/(\mathbb{Z}/2\mathbb{Z}) \arrow[dr] & \mathrm{SO}(16,\mathbb{R}) \arrow[d] & \mathrm{Spin}(16,\mathbb{R})/(\mathbb{Z}/2\mathbb{Z}) \arrow[dl] \\ & \mathrm{PSO}(16,\mathbb{R}) & \end{tikzcd}$$

We have a homomorphism $\mathrm{Spin}(16,\mathbb{R}) \to$ compact form of E_8. The kernel consists of those elements that act trivially on the Lie algebra of E_8, which is equal to the Lie algebra of D_8 plus the half-spin representation. On the Lie algebra of D_8, everything in the center acts trivially, and on the half-spin representation, one of the order-two elements is trivial and the other is not. So the image of the homomorphism is a subgroup of E_8 isomorphic to $\mathrm{Spin}(16,\mathbb{R})/(\mathbb{Z}/2\mathbb{Z})$. $\Diamond$

8.2 Constructions

8.2.1 From lattice to Lie algebra

In this section, we will try to find a natural map from root lattices to Lie algebras. Our construction will apply, at the minimum, to the root lattices corresponding to simply-laced Dynkin diagrams. The idea is simple: take as a basis the formal symbols e^α for each root α, add in $L \otimes \mathbb{K}$ where L is the root lattice, and define the Lie bracket by setting $[e^\alpha, e^\beta] = e^{\alpha+\beta}$. Except that this has a sign problem, because $[e^\alpha, e^\beta] \neq -[e^\beta, e^\alpha]$.

Is there some good way to resolve the sign problem? Not really. Suppose we had a nice functor from root lattices to Lie algebras. Then we would get that the automorphism group of the lattice has to be contained in the automorphism group of the Lie algebra (which is contained in the Lie group), and the automorphism group of the lattice contains the Weyl group of the lattice. But the Weyl group is not usually a subgroup of the Lie group.

8.2.1.1 Example We can see this going wrong even in the case of $\mathfrak{sl}(2,\mathbb{R})$. Remember that the Weyl group is $\mathcal{N}(T)/T$ where $T = \left(\begin{smallmatrix} a & 0 \\ 0 & a^{-1}\end{smallmatrix}\right)$ and its normalizer is $\mathcal{N}(T) = T \cup \left(\begin{smallmatrix} 0 & b \\ -b^{-1} & 0\end{smallmatrix}\right)$. This second part consists of stuff having order four, so you cannot possibly write this as a semi-direct product of T and the Weyl group. $\Diamond$

So the Weyl group $\mathfrak{W}$ is not usually a subgroup of the normalizer of the torus $\mathcal{N}(T)$. The best we can do is to find a group of the form $2^n \cdot \mathfrak{W} \subseteq \mathcal{N}(T)$ where n is the rank (the dimension of the torus). For example, let's do it for $\mathrm{SL}(n+1,\mathbb{R})$ Then $T = \mathrm{diag}(a_1, \dots, a_n)$ with $a_1 \cdots a_n = 1$. Then we take the normalizer of the torus is $\mathcal{N}(T) = T \cdot \{\text{permutation matrices with entries } = \pm 1 \text{ and } \det = 1\}$, and the second factor is a $2^n \cdot S_n$, and it does not split. The problem we had earlier with signs can be traced back to the fact that this group doesn't split.

In fact, we *can* construct the Lie algebra from something acted on by $2^n \cdot W$, although not from something acted on by W. Let's take a *central extension* of the lattice by a group of order 2. Notation is a pain because the lattice L is written additively and the extension will be nonabelian; instead, we will write the lattice multiplicatively, by assigning $\alpha \mapsto e^\alpha$, and we will emphasize this change by writing the abelian group L as e^L. Then we write $\hat{e}^L$ for the central extension, and insist that the kernel $\hat{e}^L \to e^L$ be $\{\pm 1\}$ (which is central in $\hat{e}^L$, of course):

$$1 \to \{\pm 1\} \to \hat{e}^L \to e^L \to 1$$

We will take as our extension $\hat{e}^L$ the one satisfying $\hat{e}^\alpha \hat{e}^\beta = (-1)^{(\alpha,\beta)} \hat{e}^\beta \hat{e}^\alpha$ for each α, β, where $\pm \hat{e}^\alpha$ are the two elements of $\hat{e}^L$ mapping to e^α.

What do the automorphisms of $\hat{e}^L$ look like?

$$1 \to \underbrace{(L/2L)}_{(\mathbb{Z}/2)^{\mathrm{rank}(L)}} \to \mathrm{Aut}(\hat{e}^L) \to \mathrm{Aut}(e^L)$$

For each $\alpha \in L/2L$, we get an (inner) automorphism $\hat{e}^\beta \to (-1)^{(\alpha,\beta)} \hat{e}^\beta$, and hence the map $(L/2L) \to \mathrm{Aut}(\hat{e}^L)$. For our extension this map makes the above sequence exact, and the extension is usually non-split.

Now we define a Lie algebra on $(L \otimes \mathbb{K}) \oplus \bigoplus_{\alpha^2=2} \mathbb{K}\hat{e}^\alpha$, modulo the convention that $-1 \in \hat{e}^L$ acts as -1 in the vector space: you take $\bigoplus_{\ell \in \hat{e}^L} \mathbb{K}\ell$ and quotient by identifying $\ell \in \mathbb{K}\ell$ with $-(-\ell) \in \mathbb{K}(-\ell)$. Then declare the following "obvious" rules:

- $[\alpha, \beta] \stackrel{\text{def}}{=} 0$ for $\alpha, \beta \in L$, so that the Cartan subalgebra is abelian;
- $[\alpha, \hat{e}^\beta] \stackrel{\text{def}}{=} (\alpha, \beta)\hat{e}^\beta$, so that $\hat{e}^\beta$ is in the β root space;
- $[\hat{e}^\alpha, \hat{e}^\beta] \stackrel{\text{def}}{=} \hat{e}^\alpha \hat{e}^\beta$ if $(\alpha, \beta) < 0$ and $\alpha \neq -\beta$ — by this, we mean product in the group $\hat{e}^L$, and if that leaves $\bigoplus_{\alpha^2=2} \mathbb{K}\hat{e}^\alpha$, then the bracket is 0;
- $[\hat{e}^\alpha, (\hat{e}^\alpha)^{-1}] \stackrel{\text{def}}{=} \alpha$.

Note that $[\hat{e}^\alpha, \hat{e}^\beta] = 0$ if $(\alpha, \beta) \geq 0$, since $(\alpha + \beta)^2 > 2$. We also have $[\hat{e}^\alpha, \hat{e}^\beta] = 0$ if $(\alpha, \beta) \leq -2$, since then $(\alpha+\beta)^2 = \alpha^2 + \beta^2 + 2(\alpha, \beta) \geq 2 + 2 - 2(2) = 0$ and so again $\alpha + \beta$ is not a root.

8.2.1.2 Proposition *If $(,) : L \times L \to \mathbb{Z}$ is positive-definite, then the bracket defined above defines a Lie algebra (i.e. it is skew-symmetric and satisfies the Jacobi identity).*

The proof is easy but tiresome, because there are lots of cases. We'll do (most of) them, to show that it's not as tiresome as you might think.

Proof Antisymmetry is almost immediate. The only condition that must be checked is $[\hat{e}^\alpha, \hat{e}^\beta] = \hat{e}^\alpha \hat{e}^\beta$, which is non-zero only if $(\alpha, \beta) = -1$.

For the Jacobi identity — $[[a, b], c] + [[b, c], a] + [[c, a], b] = 0$ — we check many cases.

1. All of a, b, c are in L. The Jacobi identity is trivial because all brackets are zero.

2. Two of a, b, c are L, so say $\{a, b, c\} = \{\alpha, \beta, e^\gamma\}$. Then:

$$[[\alpha,\beta],\hat{e}^\gamma] + [[\beta,\hat{e}^\gamma],\alpha] + [[\hat{e}^\gamma,\alpha],\beta] = 0 + (\beta,\gamma)[\hat{e}^\gamma,\alpha] - (\alpha,\gamma)[\hat{e}^\gamma,\beta] = \\ = -(\beta,\gamma)(\alpha,\gamma)e^\gamma + (\alpha,\gamma)(\beta,\gamma)\hat{e}^\gamma = 0$$

3. One of a, b, c in L, so $\{a, b, c\} = \{\alpha, \hat{e}^\beta, \hat{e}^\gamma\}$. Since $\hat{e}^\beta$ has weight β and $\hat{e}^\gamma$ has weight γ, and $\hat{e}^\beta\hat{e}^\gamma$ has weight $\beta+\gamma$.

$$\begin{aligned}
&[[\alpha,\hat{e}^\beta],\hat{e}^\gamma] = (\alpha,\beta)[\hat{e}^\beta,\hat{e}^\gamma] \\
&[[\hat{e}^\beta,\hat{e}^\gamma],\alpha] = -[\alpha,[\hat{e}^\beta,\hat{e}^\gamma]] = -(\alpha,\beta+\gamma)[\hat{e}^\beta,\hat{e}^\gamma] \\
&[[\hat{e}^\gamma,\alpha],\hat{e}^\beta] = -[[\alpha,\hat{e}^\gamma],\hat{e}^\beta] = (\alpha,\gamma)[\hat{e}^\beta,\hat{e}^\gamma]
\end{aligned}$$

   The sum is zero.

4. The really tiresome case is when none of a, b, c are in L. Let the terms now be $\hat{e}^\alpha, \hat{e}^\beta, \hat{e}^\gamma$. By positive-definiteness, the dot products (α,β), (α,γ), and (β,γ) lie in $\{-2,\dots,2\}$, and $(\alpha,\beta) = \pm 2$ iff $\alpha = \pm\beta$.

   (a) If two of these are values are zero, then all the $[[*,*],*]$ are zero.

   (b) Suppose that $\alpha = -\beta$. By (a), γ cannot be orthogonal to them. In one case, $(\alpha,\gamma) = 1$ and $(\gamma,\beta) = -1$. Adjust signs so that $\hat{e}^\alpha\hat{e}^\beta = 1$ and then calculate:

$$[[\hat{e}^\gamma,\hat{e}^\beta],\hat{e}^\alpha] - [[\hat{e}^\alpha,\hat{e}^\beta],\hat{e}^\gamma] + [[\hat{e}^\alpha,\hat{e}^\gamma],\hat{e}^\beta] = \hat{e}^\alpha\hat{e}^\beta\hat{e}^\gamma - (\alpha,\gamma)\hat{e}^\gamma + 0 = \hat{e}^\gamma - \hat{e}^\gamma = 0$$

   (c) The case when $\alpha = -\beta = \gamma$ is easy because $[\hat{e}^\alpha,\hat{e}^\gamma] = 0$ and $[[\hat{e}^\alpha,\hat{e}^\beta],\hat{e}^\gamma] = -[[\hat{e}^\gamma,\hat{e}^\beta],\hat{e}^\alpha]$.

   (d) So we have reduced to the case when each dot product is $\{-1, 0, 1\}$, and at most one of them is 0. If some $(\alpha,\beta) = 1$, then neither $(\alpha+\gamma,\beta)$ nor $(\alpha,\beta+\gamma)$ is -1, and so all brackets are 0.

   (e) Suppose that $(\alpha,\beta) = (\beta,\gamma) = (\gamma,\alpha) = -1$, in which case $\alpha+\beta+\gamma = 0$. Then $[[\hat{e}^\alpha,\hat{e}^\beta],e^\gamma] = [\hat{e}^\alpha\hat{e}^\beta,\hat{e}^\gamma] = \alpha+\beta$. By symmetry, the other two terms are $\beta+\gamma$ and $\gamma+\alpha$;the sum of all three terms is $2(\alpha+\beta+\gamma) = 0$.

   (f) Suppose that $(\alpha,\beta) = (\beta,\gamma) = -1$, $(\alpha,\gamma) = 0$, in which case $[\hat{e}^\alpha,\hat{e}^\gamma] = 0$. We check that $[[\hat{e}^\alpha,\hat{e}^\beta],\hat{e}^\alpha] = [\hat{e}^\alpha\hat{e}^\beta,\hat{e}^\gamma] = \hat{e}^\alpha\hat{e}^\beta\hat{e}^\gamma$ (since $(\alpha+\beta,\gamma) = -1$). Similarly, we have $[[\hat{e}^\beta,\hat{e}^\gamma],\hat{e}^\alpha] = [\hat{e}^\beta\hat{e}^\gamma,\hat{e}^\alpha] = \hat{e}^\beta\hat{e}^\gamma\hat{e}^\alpha$. We notice that $\hat{e}^\alpha\hat{e}^\beta = -\hat{e}^\beta\hat{e}^\alpha$ and $\hat{e}^\gamma\hat{e}^\alpha = \hat{e}^\alpha\hat{e}^\gamma$ so $\hat{e}^\alpha\hat{e}^\beta\hat{e}^\gamma = -\hat{e}^\beta\hat{e}^\gamma\hat{e}^\alpha$; again, the sum of all three terms in the Jacobi identity is 0.

This concludes the verification of the Jacobi identity, so we have a Lie algebra. □

8.2.1.3 Remark Is there a proof avoiding case-by-case check? Good news: yes! Bad news: it's actually more work. We really have functors as follows:

$$\begin{tikzcd}
\text{Dynkin diagrams} \arrow[r] \arrow[dr] & \text{Double cover } \hat{L} \arrow[rr, "\substack{\text{elementary,}\\ \text{but tedious}}", "\substack{\text{only for positive-}\\ \text{definite lattices}}"'] \arrow[d] \arrow[dr] & & \text{Lie algebras} \\
& \text{Root lattice } L & \text{Vertex algebras} \arrow[ur] &
\end{tikzcd}$$

these work for any *even* lattice

The double cover $\hat{L}$ is not a lattice; it is generated as a group by symbols $\hat{e}^{\alpha_i}$ for simple roots α_i, with relations $\hat{e}^{\alpha_i}\hat{e}^{\alpha_j} = (-1)^{(\alpha_i,\alpha_j)}\hat{e}^{\alpha_j}\hat{e}^{\alpha_i}$ and that -1 is central of order 2.

Unfortunately, you have to spend several weeks learning vertex algebras. In fact, the construction we did was the vertex algebra approach, with all the vertex algebras removed. Vertex algebras provide a more general construction which gives a much larger class of infinite dimensional Lie algebras. $\Diamond$

Now we should study the double cover $\hat{L}$, and in particular prove its existence. Given a Dynkin diagram, we defined $\hat{L}$ as generated by the elements $\hat{e}^{\alpha_i}$ for α_i simple roots with the given relations. It is easy to check that we get a surjective homomorphism $\hat{L} \to L$ with kernel generated by z with $z^2 = 1$. What's a little harder to show is that $z \neq 1$ (i.e., show that $\hat{L} \neq L$). The easiest way to do it is to use cohomology of groups, but since we have such an explicit case, we'll do it bare hands.

Our challenge then is: Given Z, H groups with Z abelian, construct extensions $1 \to Z \to G \to H \to 1$ where Z lands in the center of G. As a set, G consists of pairs (z, h), and we consider the product $(z_1, h_1)(z_2, h_2) \stackrel{\text{def}}{=} (z_1\, z_2\, c(h_1, h_2), h_1\, h_2)$ for some $c : H \times H \to Z$ (which will end up being a *cocycle in group cohomology*). There is an obvious homomorphism $(z, h) \mapsto h$, and we normalize c so that $c(1, h) = c(h, 1) = 1$, whence $z \mapsto (z, 1)$ is a homomorphism from Z to the center of G. In particular, $(1, 1)$ is the identity. We'll leave it as an exercise to figure out what the inverses are.

But when is the multiplication we've defined on $G = Z \times H$ even associative? Let's just write everything out:

$$\begin{aligned}
\big((z_1, h_1)(z_2, h_2)\big)(z_3, h_3) &= (z_1\, z_2\, z_3\, c(h_1, h_2)\, c(h_1\, h_2, h_3), h_1\, h_2\, h_3) \\
(z_1, h_1)\big((z_2, h_2)(z_3, h_3)\big) &= (z_1\, z_2\, z_3\, c(h_1, h_2\, h_3)\, c(h_2, h_3), h_1\, h_2\, h_3)
\end{aligned}$$

So we win only if c satisfies the *cocycle identity*:

$$c(h_1, h_2)\, c(h_1\, h_2, h_3) = c(h_1, h_2\, h_3)\, c(h_2, h_3).$$

This identity is immediate when c is *bimultiplicative*: $c(h_1, h_2h_3) = c(h_1, h_2)c(h_1, h_3)$ and $c(h_1h_2, h_3) = c(h_1, h_3)c(h_2, h_3)$. Not all cocycles come from such maps, but this is the case we care about.

To construct the double cover, let $Z = \{\pm 1\}$ and $H = L$ (free abelian). If we write H additively, we want c to be a bilinear map $L \times L \to \{\pm 1\}$. It is really easy to construct bilinear maps on free abelian groups. Just take any basis $\alpha_1, \ldots, \alpha_n$ of L, choose $c(\alpha_1, \alpha_j)$ arbitrarily for each i, j and extend c via bilinearity to $L \times L$. In our case, we want to find a double cover $\hat{L}$ satisfying $\hat{e}^\alpha \hat{e}^\beta = (-1)^{(\alpha,\beta)} \hat{e}^\beta \hat{e}^\alpha$ where $\hat{e}^\alpha$ is a lift of e^α. This just means that $c(\alpha, \beta) = (-1)^{(\alpha,\beta)} c(\beta, \alpha)$. To satisfy this, just choose $c(\alpha_i, \alpha_j)$ on the basis $\{\alpha_i\}$ so that $c(\alpha_i, \alpha_j) = (-1)^{(\alpha_i,\alpha_j)} c(\alpha_j, \alpha_i)$. This is trivial to do as $(-1)^{(\alpha_i,\alpha_i)} = 1$, since the lattice is even. There is no canonical way to choose this 2-cocycle (otherwise, the central extension would split as a product), but all the different double covers are isomorphic because we can specify $\hat{L}$ by generators and relations. Thus, we have constructed $\hat{L}$ (or rather, verified that the kernel of $\hat{L} \to L$ has order 2, not 1).

8.2.1.4 Remark Let's now look at lifts of automorphisms of L to $\hat{L}$. There are two special cases:

1. Multiplication by -1 is an automorphism of L, and we want to lift it to $\hat{L}$ explicitly. As a first attempt, try sending $\hat{e}^\alpha$ to $\hat{e}^{-\alpha} := (\hat{e}^\alpha)^{-1}$. This doesn't work because $a \mapsto a^{-1}$ is not an automorphism on non-abelian groups.

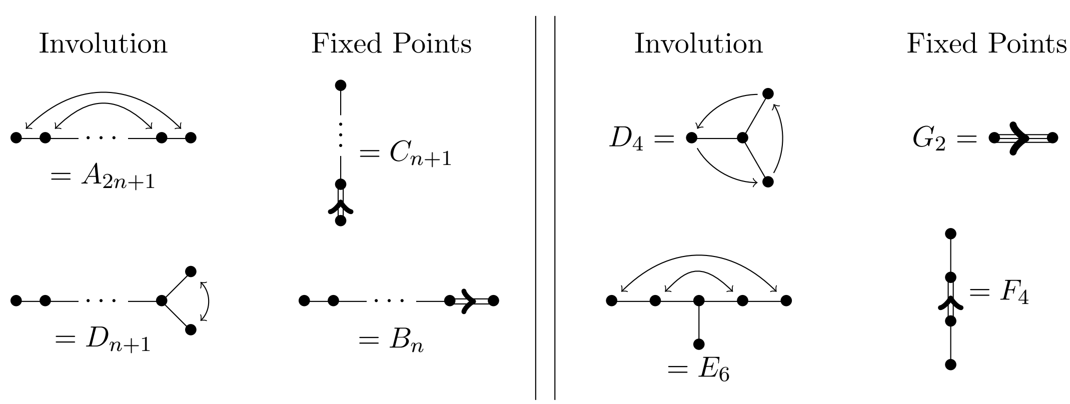
   Instead, we define $\omega : \hat{e}^\alpha \mapsto (-1)^{\alpha^2/2} (\hat{e}^\alpha)^{-1}$, which is an automorphism of $\hat{L}$:

$$\begin{aligned} \omega(\hat{e}^\alpha)\omega(\hat{e}^\beta) &= (-1)^{(\alpha^2+\beta^2)/2} (\hat{e}^\alpha)^{-1} (\hat{e}^\beta)^{-1} \\ \omega(\hat{e}^\alpha \hat{e}^\beta) &= (-1)^{(\alpha+\beta)^2/2} (\hat{e}^\beta)^{-1} (\hat{e}^\alpha)^{-1} \end{aligned}$$

2. If $r^2 = 2$, then reflection through $r^\perp$, $\alpha \mapsto \alpha - (\alpha, r)r$, is an automorphism of L. This lifts to $\hat{e}^\alpha \mapsto \hat{e}^\alpha (\hat{e}^r)^{-(\alpha,r)} \times (-1)^{\binom{(\alpha,r)}{2}}$. This is a homomorphism, but usually of order 4, not 2!

So although automorphisms of L lift to automorphisms of $\hat{L}$, the lift might have larger order. ♢

The construction given above works for the root lattices of A_n, D_n, E_6, E_7, and E_8; these lattices are all even, positive definite, and generated by vectors of norm 2 (in fact, any such lattices is a sum of A_n–E_8). What about B_n, C_n, F_4 and G_2? The reason the construction doesn't work for these cases is because there are roots of different lengths. These all occur as fixed points of diagram automorphisms of A_n, D_n and E_6. In fact, we presented a *functor* from simply-laced Dynkin diagrams to Lie algebras, so an automorphism of the diagram gives an automorphism of the algebra.



A_{2n} doesn't give you a new algebra, but rather some superalgebra that we will not describe.

8.2.2 From lattice to Lie group

First, let's work over $\mathbb{R}$. We start with a simply-laced Dynkin diagram, and as in the previous section construct $L \oplus \mathbb{R}\hat{e}^L$. Then we can form its Lie group by looking at those automorphisms generated by the elements $\exp(\lambda \operatorname{Ad}(\hat{e}^\alpha))$, where λ is some real number, $\hat{e}^\alpha$ is one of the basis elements of the Lie algebra corresponding to the root α, and $\operatorname{Ad}(\hat{e}^\alpha)(a) = [\hat{e}^\alpha, a]$. In other words:

$$\exp(\lambda \operatorname{Ad}(\hat{e}^\alpha))(a) = 1 + \lambda[\hat{e}^\alpha, a] + \frac{\lambda^2}{2}[\hat{e}^\alpha, [\hat{e}^\alpha, a]]$$

To see that $\operatorname{Ad}(\hat{e}^\alpha)^3 = 0$, note that if β is a root, then $\beta + 3\alpha$ is not a root (or 0).

8.2.2.1 Remark In general, the group generated by these automorphisms is not the whole automorphism group of the Lie algebra. There might be extra diagram automorphisms, for example. $\lozenge$

8.2.2.2 Remark In fact, the construction of a Lie algebra above works over any commutative ring, e.g. over $\mathbb{Z}$ — one way to say this is that it defines a "*group scheme* over $\mathbb{Z}$". The only place we used division is in $\exp(\lambda \operatorname{Ad}(\hat{e}^\alpha))$, where we divided by 2 in the quadratic term. The only time this term is non-zero is when we apply $\exp(\lambda \operatorname{Ad}(\hat{e}^\alpha))$ to $\hat{e}^{-\alpha}$, in which case we find that $[\hat{e}^\alpha, [\hat{e}^\alpha, \hat{e}^{-\alpha}]] = [\hat{e}^\alpha, \alpha] = -(\alpha, \alpha)\hat{e}^\alpha$, and the fact that $(\alpha, \alpha) = 2$ cancels the division by 2. So we can in fact construct the E_8 group, for example, over *any* commutative ring. In particular, we have groups of type E_8 over *finite fields*, which are actually finite simple groups. These are called *Chevalley groups*; it takes work to show that they are simple, c.f. [Car72]. $\lozenge$

8.3 Every possible simple Lie group

8.3.1 Real forms

So far we've constructed a Lie algebra and a Lie group of type E_8. (Our construction works starting with any simply-laced diagram, and over any ring, as we observed above.) But for a given field, there are in fact usually several different groups of type E_8. In particular, there is only one complex Lie algebra of type E_8, which corresponds to several different real Lie algebras of type E_8. We discussed real forms a little bit in Section 7.2.1, and review that discussion below.

A special case of Theorem 7.2.1.8 guarantees that E_8 has a unique compact form. For comparison, the form we constructed in Proposition 8.2.1.2 is the *split form* of E_8. That a given Dynkin diagram supports multiple Lie algebras is not special to E_8:

8.3.1.1 Example Recall the algebra $\mathfrak{sl}(2, \mathbb{R}) = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ with a, b, c, d real $a + d = 0$; this does not integrate to a compact group. On the other hand, $\mathfrak{su}(2, \mathbb{R}) = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ with $d = -a$ imaginary and $b = -\bar{c}$, is compact. These have the same Lie algebra over $\mathbb{C}$. $\lozenge$

Suppose that $\mathfrak{g}$ is a Lie algebra with complexification $\mathfrak{g} \otimes \mathbb{C}$. How can we find another Lie algebra $\mathfrak{h}$ with the same complexification? On $\mathfrak{g} \otimes \mathbb{C}$ there is an anti-linear involution $\omega_{\mathfrak{g}} : g \otimes z \mapsto g \otimes \bar{z}$. Similarly, $\mathfrak{g} \otimes \mathbb{C} = \mathfrak{h} \otimes \mathbb{C}$ has an anti-linear involution $\omega_{\mathfrak{h}}$. Notice that $\omega_{\mathfrak{g}}\omega_{\mathfrak{h}}$ is a linear involution of $\mathfrak{g} \otimes \mathbb{C}$. Conversely, if we know this (linear) involution, we can reconstruct $\mathfrak{h}$ from it. Indeed, given an involution ω of $\mathfrak{g} \otimes \mathbb{C}$, we can get $\mathfrak{h}$ as the fixed points of the map $a \mapsto \omega_{\mathfrak{g}}\, \omega(a)$"$=$"$\overline{\omega(a)}$.

Equivalently, break $\mathfrak{g}$ into the ± 1 eigenspaces of ω, so that $\mathfrak{g} = \mathfrak{g}^+ \oplus \mathfrak{g}^-$. Then $\mathfrak{h} = \mathfrak{g}^+ \oplus i\mathfrak{g}^-$. Notice that $\omega_\mathfrak{g}$ is a (real) Lie algebra automorphism of $\mathfrak{g} \otimes \mathbb{C}$; that $\mathfrak{h}$ is also a Lie algebra is equivalent to ω being a Lie algebra map (rather than just a linear involution).

Thus, to find other real forms, we have to study the involutions of the complexification of $\mathfrak{g}$. The exact relation between involutions and is kind of subtle, but this is a good way to go. We used a similar argument to construct the compact form of each simple Lie algebra in Theorem 7.2.1.8.

8.3.1.2 Example Let $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{R})$. It has an involution $\omega(m) = -m^T$. By definition, $\mathfrak{su}(2, \mathbb{R})$ is the set of fixed points of the involution which is ω times complex conjugation on $\mathfrak{sl}(2, \mathbb{C})$. ♢

So to construct real forms of E_8, we want some involutions of the Lie algebra E_8 which we constructed. What involutions do we know about? There are two obvious ways to construct involutions:

1. Lift -1 on the lattice L to $\omega : \hat{e}^\alpha \mapsto (-1)^{\alpha^2/2}(\hat{e}^\alpha)^{-1}$, which induces an involution on the Lie algebra.

2. Take $\beta \in L/2L$, and look at the involution $\hat{e}^\alpha \mapsto (-1)^{(\alpha,\beta)}\hat{e}^\alpha$.

It will turn out that 2. gives nothing new on its own: we'll get the Lie algebra we started with. On the other hand, 1. only gives us one real form, which will turn out to be the compact form we already knew about. To get all real forms, we'll multiply these two kinds of involutions together.

Recall that $L/2L$ has 3 orbits under the action of the Weyl group, of size 1, 120, and 135. These will correspond to the three real forms of E_8. How do we distinguish different real forms? The answer was found by Cartan: look at the signature of an invariant quadratic form on the Lie algebra!

8.3.1.3 Definition *A bilinear form $(,)$ on a Lie algebra is called* invariant *if $([a, b], c)+(b, [a, c]) = 0$ for all a, b, c. Such a form is called "invariant" because it corresponds to the form being invariant under the corresponding group action.*

8.3.1.4 Lemma *We construct an invariant bilinear form on the split form of E_8 (the one constructed in Proposition 8.2.1.2) as follows:*

- $(\alpha, \beta)_{\textit{in the Lie algebra}} = (\alpha, \beta)_{\textit{in the lattice}}$
- $(\hat{e}^\alpha, (\hat{e}^\alpha)^{-1}) = 1$
- $(a, b) = 0$ *if a and b are in root spaces α and β with $\alpha + \beta \neq 0$.*

This form is unique up to multiplication by a constant since E_8 is simple. □

Since invariant forms are unique up to scaling, the absolute values of their signatures are invariants of the corresponding Lie algebras. For the split form of E_8, what is the signature of the invariant bilinear form (the split form is the one we just constructed)? On the Cartan subalgebra L, $(\ ,\ )$ is positive definite, so we get $+8$ contribution to the signature. On $\{e^\alpha, (e^\alpha)^{-1}\}$, the form is $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, which contributes $0 \cdot 120$ to the signature. Thus, the signature is $+8$. So if we find any real form with a different signature, we'll have found a new Lie algebra.

8.3.1.5 Example Let's first try involutions $e^\alpha \mapsto (-1)^{(\alpha,\beta)}e^\alpha$. But this doesn't change the signature. The lattice L is still positive definite, and you still have $\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)$ or $\left(\begin{smallmatrix}0&-1\\-1&0\end{smallmatrix}\right)$ on the other parts. In fact, these Lie algebras actually turn out to be isomorphic to what we started with, though we haven't shown this. $\diamondsuit$

8.3.1.6 Example Now we'll try $\omega : e^\alpha \mapsto (-1)^{\alpha^2/2}(e^\alpha)^{-1}$, $\alpha \mapsto -\alpha$. What is the signature of the form? Let's write down the $+$ and $-$ eigenspaces of ω. The $+$ eigenspace will be spanned by $e^\alpha - e^{-\alpha}$, and these vectors have norm -2 and are orthogonal. The $-$ eigenspace will be spanned by $e^\alpha + e^{-\alpha}$ and L, which have norm 2 and are orthogonal, and L is positive definite. What is the Lie algebra corresponding to the involution ω? It will be spanned by $e^\alpha - e^{-\alpha}$ where $\alpha^2 = 2$, so these basis vectors have norm -2, and by $i(e^\alpha + e^{-\alpha})$, which also have norm -2, and iL, which is negative definite. So the bilinear form is negative definite, with signature -248. In particular, $|-248| \neq |8|$, and so ω gives a real form of E_8 that is not the split real form! In particular, since the bilinear form is negative definite, we have found the *compact real form* of E_8. $\diamondsuit$

Finally, let's look at involutions of the form $\hat{e}^\alpha \mapsto (-1)^{(\alpha,\beta)}\omega(\hat{e}^\alpha)$. Notice that ω commutes with $\hat{e}^\alpha \mapsto (-1)^{(\alpha,\beta)}\hat{e}^\alpha$. The βs in (α,β) correspond to $L/2L$ modulo the action of the Weyl group $\mathfrak{W}(E_8)$. Remember this has three orbits, with one norm-0 vector, 120 norm-2 vectors, and 135 norm-4 vectors. The norm-0 vector gives us the compact form. Let's look at the other cases and see what we get.

First, suppose V has a negative definite symmetric inner product $(,)$, and suppose σ is an involution of $V = V_+ \oplus V_-$ (eigenspaces of σ). What is the signature of the invariant inner product on $V_+ \oplus iV_-$? On V_+, it is negative definite, and on iV_- it is positive definite. Thus, the signature is $\dim V_- - \dim V_+ = -\operatorname{tr}(\sigma)$. So, letting V be the compact form of E_8, we want to work out the traces of the involutions $\hat{e}^\alpha \mapsto (-1)^{(\alpha,\beta)}\omega(\hat{e}^\alpha)$.

8.3.1.7 Example Given some $\beta \in L/2L$, what is $\operatorname{tr}(\hat{e}^\alpha \mapsto (-1)^{(\alpha,\beta)}\hat{e}^\alpha)$? If $\beta = 0$, the trace is obviously 248, as the involution is the identity map.

If $\beta^2 = 2$, we need to figure how many roots have a given inner product with β. We counted these in Example 8.1.2.7:

(α,β)	# of roots α with given inner product	eigenvalue
2	1	1
1	56	-1
0	126	1
-1	56	-1
-2	1	1

Thus, the trace is $1 - 56 + 126 - 56 + 1 + 8 = 24$ (the 8 is from the Cartan subalgebra). So the signature of the corresponding form on the Lie algebra is -24. We've found a third Lie algebra. $\diamondsuit$

8.3.1.8 Example If we also look at the case when $\beta^2 = 4$, what happens? How many α with $\alpha^2 = 2$ and with given (α,β) are there? In this case, we have:

(α, β)	# of roots α with given inner product	eigenvalue
2	14	1
1	64	-1
0	84	1
-1	64	-1
-2	14	1

The trace will be $14 - 64 + 84 - 64 + 14 + 8 = -8$. This is just the split form again. ◊

In summary, we've found three forms of E_8, corresponding to the three classes in $L/2L$, with signatures 8, -24, and -248. In fact, these are the only real forms of E_8, but we won't prove this. In general, if $\mathfrak{g}$ is the compact form of a simple Lie algebra, then Cartan showed that the other forms correspond exactly to the conjugacy classes of involutions in the automorphism group of $\mathfrak{g}$. Be warned, though, that this doesn't work if you don't start with the compact form.

8.3.2 Working with simple Lie groups

As an example of how to work with simple Lie groups, we will look at the general question: Given a simple Lie group, what is its homotopy type?

8.3.2.1 Proposition *Let G be a simple real Lie group. Then G has a unique conjugacy class of maximal compact subgroups K, and G is homotopy equivalent to K.*

Proposition 8.3.2.1 essentially follows from Theorem 7.1.4.4. We will give the proof for $G = \mathrm{GL}(n, \mathbb{R})$, in spite of the fact that $\mathrm{GL}(n, \mathbb{R})$ is not simple.

Proof (Proof for GL($n, \mathbb{R}$)) $\mathrm{GL}(n, \mathbb{R})$ an obvious compact subgroup: $\mathrm{O}(n, \mathbb{R})$. Suppose K is *any* compact subgroup of $\mathrm{GL}(n, \mathbb{R})$. Choose any positive definite form $(,)$ on $\mathbb{R}^n$. This will probably not be invariant under K, but since K is compact, we can average it over K: define a new form $(a, b)_{\text{new}} = \int_K (ka, kb)\, \mathrm{d}k$. This gives an invariant positive definite bilinear form (since the integral of something positive definite is positive definite, since the space of positive-definite forms is convex). Thus, any compact subgroup preserves some positive definite form. But any subgroup fixing some positive definite bilinear form is conjugate to some subgroup of $\mathrm{O}(n, \mathbb{R})$, since we can diagonalize the form. So K is contained in a conjugate of $\mathrm{O}(n, \mathbb{R})$.

Next we want to show that $G = \mathrm{GL}(n, \mathbb{R})$ is homotopy equivalent to $\mathrm{O}(n, \mathbb{R}) = K$. We will show that $\mathrm{GL}(n)$ splits into a *Iwasawa decomposition*, as we asserted it did in Theorem 7.1.4.4: we claim that $G = KAN$, where $K = \mathrm{O}(n)$, $A = (\mathbb{R}_{>0})^n$ consists of all diagonal matrices with positive coefficients, and $N = \mathrm{N}(n)$ consists of matrices which are upper-triangular with 1s on the diagonal. For arbitrary G, you can always assure that K is compact, N is unipotent, and A is abelian and acts semisimply on all G-representations.

The proof of this you saw before was called the *Gram–Schmidt process* for orthonormalizing a basis. Suppose $\{v_1, \ldots, v_n\}$ is any basis for $\mathbb{R}^n$.

1. Make it orthogonal by subtracting some stuff. You'll get a new basis with $w_1 = v_1$, $w_2 = v_2 - \frac{(v_2,v_1)}{(v_1,v_1)} v_1$, $w_3 = v_3 - *v_2 - *v_1$, ..., satisfying $(w_i, w_j) = 0$ if $i \neq j$.

2. Normalize by multiplying each basis vector so that it has norm 1. Now we have an orthonormal basis.

This is just another way to say that $\mathrm{GL}(n) = \mathrm{O}(n) \cdot (\mathbb{R}_{>0})^n \cdot \mathrm{N}(n)$. We made the basis orthogonal by multiplying it by something in $N = \mathrm{N}(n)$, and we normalized it by multiplying it by something in $A = (\mathbb{R}_{>0})^n$. Then we end up with an orthonormal basis, i.e. an element of $K = \mathrm{O}(n)$. Tada! This decomposition is just a topological one, not a decomposition as groups. Uniqueness is easy to check: the pairwise intersections of K, A, N are trivial.

Now we can get at the homotopy type of $\mathrm{GL}(n)$. The groups $N \simeq \mathbb{R}^{n(n-1)/2}$ and $A \cong (\mathbb{R}_{>0})^n$ are contractible, and so $\mathrm{GL}(n, \mathbb{R})$ has the same homotopy type as $K = \mathrm{O}(n, \mathbb{R})$, its maximal compact subgroup. □

8.3.2.2 Example If you wanted to know $\pi_1(\mathrm{GL}(3, \mathbb{R}))$, you could calculate $\pi_1(\mathrm{O}(3, \mathbb{R})) \cong \mathbb{Z}/2\mathbb{Z}$, so $\mathrm{GL}(3, \mathbb{R})$ has a double cover. Nobody has shown you this double cover because it is not algebraic.$\diamondsuit$

8.3.2.3 Example Let's go back to various forms of E_8 and figure out (guess) the fundamental groups. We need to know the maximal compact subgroups.

1. One of them is easy: the compact form is its own maximal compact subgroup. What is the fundamental group? We quote the fact that for compact simple groups, $\pi_1 \cong \frac{\text{weight lattice}}{\text{root lattice}}$, which is 1. So this form is simply connected.

2. We now consider the $\beta^2 = 2$ case, with signature -24. Recall that there were 1, 56, 126, 56, and 1 roots α with $(\alpha, \beta) = 2$, 1, 0, -1, and -2 respectively, and there are another 8 dimensions for the Cartan subalgebra. On the $1, 126, 1, 8$ parts, the form is negative definite. The sum of these root spaces gives a Lie algebra of type $E_7 \times A_1$ with a negative-definite bilinear form (the 126 gives you the roots of an E_7, and the 1s are the roots of an A_1; the 8 is $7+1$). So it a reasonable guess that the maximal compact subgroup has something to do with $E_7 \times A_1$.

   E_7 and A_1 are not simply connected: the compact form of E_7 has $\pi_1 = \mathbb{Z}/2$ and the compact form of A_1 also has $\pi_1 = \mathbb{Z}/2$. So the universal cover of $E_7 A_1$ has center $(\mathbb{Z}/2)^2$. Which part of this acts trivially on E_8? We look at the E_8 Lie algebra as a representation of $E_7 \times A_1$, and read off how it splits from the picture above: $E_8 \cong E_7 \oplus A_1 \oplus 56 \otimes 2$, where 56 and 2 are irreducible, and the centers of E_7 and A_1 both act as -1 on them. So the maximal compact subgroup of this form of E_8 is the simply connected compact form of $E_7 \times A_1$ modulo a $\mathbb{Z}/2$ generated by $(-1, -1)$.

   But $\pi_1(E_8)$ is the same as π_1 of the compact subgroup, which is $(\mathbb{Z}/2)^2/(-1, -1) \cong \mathbb{Z}/2$. So this simple group has a nontrivial double cover, which is non-algebraic.

3. For the split form of E_8 with signature 8, the maximal compact subgroup is $\mathrm{Spin}_{16}(\mathbb{R})/(\mathbb{Z}/2)$, and $\pi_1(E_8)$ is $\mathbb{Z}/2$. Again there is a non-algebraic double cover.

You can also compute other homotopy invariants with this method. $\diamondsuit$

8.3.2.4 Example Let's look at the 56-dimensional representation of E_7 in more detail. Recall that in E_8 we had the picture:

(α,β)	# of α's
2	1
1	56
0	126
-1	56
-2	1

The Lie algebra E_7 fixes these five spaces of dimensions $1, 56, 126+8, 56, 1$. From this we can get some representations of E_7. The $126+8$ splits as $1 \oplus 133$. But we also get a 56-dimensional representation of E_7. Let's show that this is actually an irreducible representation. Recall that when we calculated $\mathfrak{W}(E_8)$ in Example 8.1.2.7, we showed that $\mathfrak{W}(E_7)$ acts transitively on this set of 56 roots of E_8, which we identify as weights of E_7.

A representation is called *minuscule* if the Weyl group acts transitively on the weights. Minuscule representations are particularly easy to work with. They are necessarily irreducible (provided there is some weight with multiplicity one), since the weights of any summand would form a union of Weyl orbits. And to calculate the character of a minuscule representation, we just translate the 1 at the highest weight around, so every weight has multiplicity 1.

So the 56-dimensional representation of E_7 must actually be the irreducible representation with whatever highest weight corresponds to one of the vectors. $\diamondsuit$

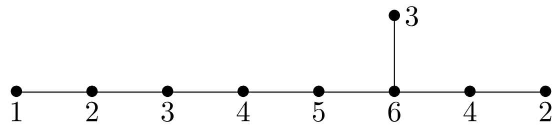
8.3.3 Finishing the story

We will construct all simple Lie groups as follows. Let $\mathfrak{g}$ be the compact form, pick an involution σ splitting $\mathfrak{g} = \mathfrak{g}^+ \oplus \mathfrak{g}^-$, and form $\mathfrak{g}^+ \oplus i\mathfrak{g}^-$. To construct the split form $\mathfrak{g}$ for A_n, D_n, E_6, E_7, we repeat the procedure we used for E_8 in Section 8.1; to construct the compact form, we use the involutions of $\omega : \hat{e}^\alpha \mapsto (-1)^{\alpha^2/2}\hat{e}^{-\alpha}$. To construct $\mathfrak{g}$ for B_n, C_n, F_4, G_2, we look at fixed points of diagram automorphisms. Thus, to list all simple Lie groups, we must understand the automorphisms of compact Lie algebras. For this, we use without proof the following theorem due to Kac [Kac69], exposited in [Hel01, Ch X, §5] (see also [Kac90]):

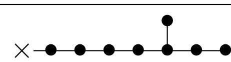
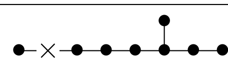
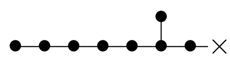
8.3.3.1 Theorem (Kac's classification of compact simple Lie algebra automorphisms) *The inner automorphisms of order N of a compact simple Lie algebra are computed as follows. Write down the corresponding* affine *Dynkin diagram with its numbering m_i. Choose integers n_i with $\sum n_i m_i = N$. Then the automorphism $\hat{e}^{\alpha_j} \mapsto \hat{e}^{2\pi i n_j/N}\hat{e}^{\alpha_j}$ has order dividing N. Up to conjugation, all inner automorphisms with order dividing N are obtained in this way, and two automorphism obtained in this way are conjugate if and only if they are conjugate by a diagram automorphism.*

The outer automorphisms of a compact simple Lie algebra are constructed as follows: pick an automorphisms of order $r|N$ of the corresponding Dynkin diagram, and use it to form a "twisted affine Dynkin diagram" for the corresponding folded diagram. Then play a similar number game: choose integers n_i for the numbered twisted affine Dynkin diagram satisfying $\sum n_i m_i = N/r$. $\square$

8.3.3.2 Example Using Theorem 8.3.3.1, let's list all the real forms of E_8. We've already found three, and it took us a long time, whereas now we can do it fast. The affine E_8 diagram is:

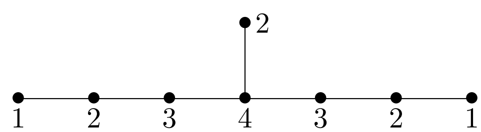


So we need to solve $\sum n_i m_i = 2$ where $n_i \geq 0$; there are only a few possibilities:

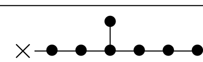
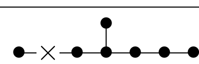
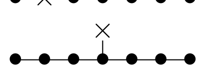
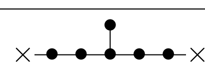
$\sum n_i m_i = 2$	# of ways	how to do it	maximal compact subgroup K
2×1	one way		E_8 (compact form)
1×2	two ways		$A_1 \times E_7$
			D_8 (split form)
$1 \times 1 + 1 \times 1$	no ways		

The points not crossed off form the Dynkin diagram of the maximal compact subgroup. So by just looking at the diagram, we can see what all the real forms are! $\Diamond$

8.3.3.3 Example Let's do E_7. Write down the affine diagram:



We get four possibilities:

$\sum n_i m_i = 2$	# of ways	how to do it	maximal compact subgroup K
2×1	one way		E_7 (compact form)
1×2	two ways		$A_1 \times D_6$
			A_7 (split form)
$1 \times 1 + 1 \times 1$	one way		$E_6 \oplus \mathbb{R}$

Some remarks:

1. When we count the number of ways, we are counting up to automorphisms of the diagram.

2. In the split real form, the maximal compact subgroup has dimension equal to half the number of roots. The roots of A_7 look like $\varepsilon_i - \varepsilon_j$ for $i, j \leq 8$ and $i \neq j$, so the dimension is $8 \cdot 7 + 7 = 56 = \frac{112}{2}$.

3. The Lie algebra of the maximal compact subgroup of the last real form on our table is is $E_6 \oplus \mathbb{R}$, because the fixed subalgebra contains the whole Cartan subalgebra whereas the visible E_6 diagram only accounts for 6 of the 7 dimensions.

You can use Remark 3 to construct the minuscule representations of E_6, by asking: How does the algebra E_7 decompose as a representation of the algebra $E_6 \oplus \mathbb{R}$? We decompose E_7 according to the eigenvalues of $\mathbb{R}$. The $E_6 \oplus \mathbb{R}$ is precisely the zero eigenvalue of $\mathbb{R}$, since $\mathbb{R}$ is central in $E_6 \oplus \mathbb{R}$, and the rest of E_7 is 54-dimensional. The easy way to see the decomposition is to look at the roots. Recall that in Example 8.1.2.7 we computed the Weyl group by looking for vectors filling in the dotted line in $\bullet\!-\!\bullet\cdots\bullet$ or $\bullet\cdots\bullet\!-\!\bullet$. For each diagram there were 27 possibilities, and they form the weights of a 27-dimensional representation of E_6. The orthogonal complement of the two nodes is an E_6 root system whose Weyl group acts transitively on these 27 vectors, since we showed that these form a single orbit. The entire E_7 root system consists of the vectors of the E_6 root system plus these 27 vectors plus the other 27 vectors. This splits up the E_7 explicitly. The two 27s form individual orbits, so the corresponding representations are irreducible. Thus, as a representation of E_6, we have split $E_7 \cong E_6 \oplus \mathbb{R} \oplus 27 \oplus 27$, and the 27s are minuscule. $\diamondsuit$

8.3.3.4 Definition *A* symmetric space *is a (simply connected) Riemannian manifold M such that for each point $p \in M$, there is an automorphism fixing p and acting as -1 on the tangent space. It is* Hermitian *if it has a complex structure.*

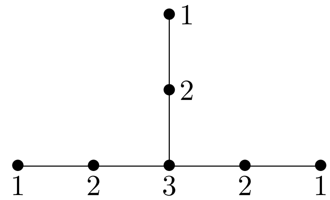
The condition to be a symmetric space looks weird, but it turns out that all kinds of nice objects you know about are symmetric spaces. Typical examples you may have seen include the spheres S^n, the hyperbolic spaces $\mathbb{H}^n$, and the Euclidean spaces $\mathbb{R}^n$.

Roughly speaking, symmetric spaces generalize this list, and have the nice properties that $S^n, \mathbb{H}^n, \mathbb{R}^n$ have. Cartan classified all symmetric spaces: depending on the details of the simply-connectedness hypotheses, the list consists of non-compact simple Lie groups modulo their maximal compact subgroups. Historically, Cartan classified simple Lie groups, and then later classified symmetric spaces, and was surprised to find the same result.

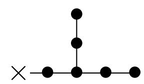
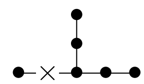
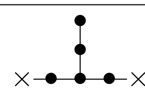
8.3.3.5 Example Let G denote the fourth real form of E_7 in Example 8.3.3.3, and K its maximal compact subgroup, with $\mathrm{Lie}(K) = \mathbb{R} \oplus E_6$. We will explain how this G/K is a symmetric space, although we'll be rather sketchy. First of all, to make it a symmetric space, we have to find a nice invariant Riemannian metric on it. It is sufficient to find a positive definite bilinear form on the tangent space at p which is invariant under K, as we can then translate it around. We can do this since K is compact (so you have the averaging trick).

Now why is G/K Hermitian? We'll show that there is an *almost-complex structure*: each tangent space is a complex vector space. The factor of $\mathbb{R}$ in $\mathrm{Lie}(K)$ corresponds to a K-invariant S^1 inside K, and the stabilizer of each point is isomorphic to K. So the tangent space at each point has an action of S^1, and by identifying this S^1 with the circle of complex numbers of unit norm we make each tangent space into a $\mathbb{C}$-vector space. This is the almost-complex structure on G/K, and it turns out to be integral, so that it comes from an actual complex structure. Notice that it is a little unexpected that G/K has a complex structure: in the case of $G = E_7$ and $K = E_6 \oplus \mathbb{R}$, each of G, K is odd-dimensional, and so there is no hope of putting separate complex structures on each and taking a quotient. $\diamondsuit$

8.3.3.6 Example Let's look at E_6, with affine Dynkin diagram:



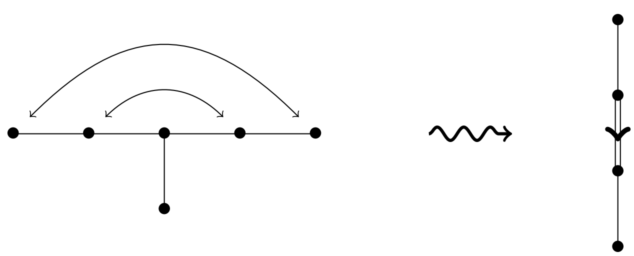
We get three possible inner involutions:

$\sum n_i m_i = 2$	# of ways	how to do it	maximal compact subgroup K
2×1	one way		E_6 (compact form)
1×2	one way		$A_1 A_5$
$1 \times 1 + 1 \times 1$	one way		$D_5 \oplus \mathbb{R}$

In the last one, the maximal compact subalgebra is $D_5 \oplus \mathbb{R}$. Just as in Example 8.3.3.5, the corresponding symmetric space G/K is Hermitian. Let's compute its dimension (over $\mathbb{C}$). The dimension will be the dimension of E_6 minus the dimension of $D_5 \oplus \mathbb{R}$, all divided by 2 (because we want complex dimension), which is $(78 - 46)/2 = 16$.

So between Example 8.3.3.5 and here we have found two non-compact simply-connected Hermitian symmetric spaces of dimensions 16 and 27. These are the only "exceptional" cases; all the others fall into infinite families!

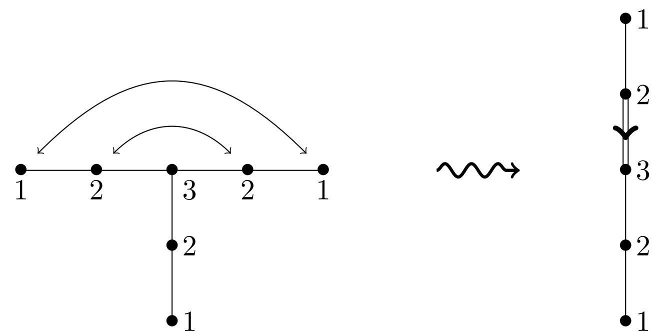
There is also an *outer* automorphisms of E_6 coming from the diagram automorphism:

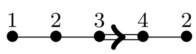


The fixed point subalgebra has Dynkin diagram obtained by folding the E_6 on itself: the F_4 Dynkin diagram. Thus the fixed points of E_6 under the diagram automorphism form an F_4 Lie algebra, and we get a real form of E_6 with maximal compact subgroup F_4. This is probably the easiest way to construct F_4, by the way. Moreover, we can decompose E_6 as a representation of F_4: $\dim E_6 = 78$ and $\dim F_4 = 52$, so $E_6 = F_4 \oplus 26$, where 26 turns out to be irreducible (the smallest non-trivial representation of F_4 — the only one anybody actually works with). The roots of F_4 look like $(0^2, \pm 1, \pm 1)$ (24 of these, counting permutations), $((\pm\frac{1}{2})^4)$ (16 of these), and $(0^3, \pm 1)$ (8 of these); the last two types are in the same orbit of the Weyl group.

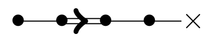
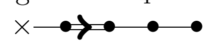
The 26-dimensional representation has the following character: it has all norm-1 roots with multiplicity one and the 0 root with multiplicity two. In particular, it is not minuscule.

There is one other real form of E_6. To get at it, we have to talk about Kac's description of *outer* automorphisms in Theorem 8.3.3.1. We use the involution of E_6 above to form the twisted affine Dynkian diagram for F_4:



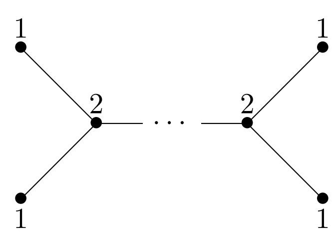
Note that this is the *twisted affine Dynkin diagram* for F_4. The *affine* Dynkin diagram for F_4 is 

. The arrow goes the other direction.

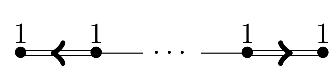
So now we need to find integers n_i with $\sum n_i m_i = 2/2 = 1$, since we are looking for involutions ($N = 2$). There are two ways to do this for E_6. Using  just gives us F_4 back, which is the involution we found more naively in the previous paragraph. Using  gives a real form of E_6 with maximal compact subalgebra C_4. This last form turns out to be the split real form of E_6. $\Diamond$

In general, the *twisted affine Dynkin diagram* of a non-simply-laced Dynkin diagram is what you get by reversing all the arrows, forming the affine Dynkin diagram, and then re-reversing all of the arrows. Reversing all of the arrows is called "Langlands duality," and the point is that Langlands duality does not commute with affinization.

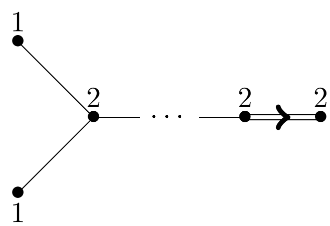
8.3.3.7 Example Consider a diagram of type D. Its affinization is:



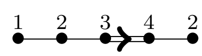
If you fold by the order-2 automorphism of the finite diagram, you get a diagram of type B. The twisted affine diagram of type B is:

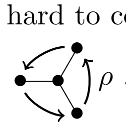


This gives the correct classification of compact forms of D_n: using the identity automorphism of the Dynkin diagram produces the compact forms with maximal compact of type $D_p \times D_{n-p}$, for $0 \le p \le \lfloor \frac{n}{2} \rfloor$ (and one form with maximal compact of type $A_{n-1} \times \mathbb{R}$); using the nonidentity automorphism produces the compact forms with maximal compact of type $B_p \times B_{n-1-p}$ with $0 \le p \le \lfloor \frac{1}{2}(n-1) \rfloor$. You would get the wrong answer if you naively folded the affine diagram to produce:



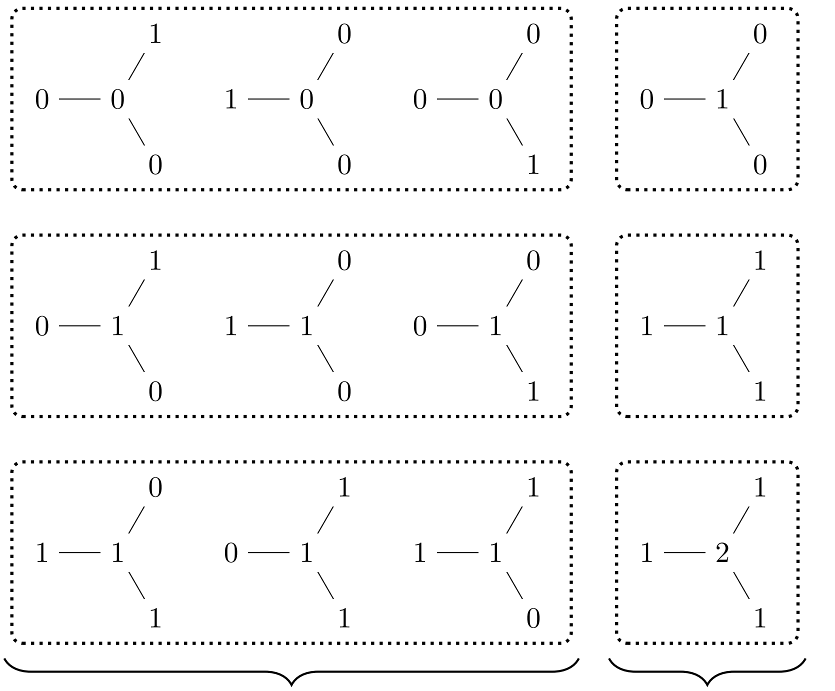
This is in fact the *untwisted* affine diagram of type B. $\Diamond$

8.3.3.8 Example The affine Dynkin of F_4 is  . We can cross out one node of weight 1, giving the compact form, or a node of weight 2 (in two ways), giving maximal compacts $A_1 \times C_3$ (which turns out to be the split form) and B_4. This gives us three real forms. $\Diamond$

8.3.3.9 Example We will conclude by listing the real forms of G_2. This is one of the only root systems we can actually draw — four-dimensional chalkboards are hard to come by. To construct G_2, we start with D_4 and look at its fixed points under *triality*:  . We will completely explicitly find the fixed points.

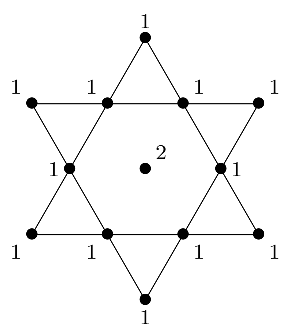
First we look at the Cartan subalgebra. The automorphism ρ fixes a two-dimensional space, and has one-dimensional eigenspaces corresponding to $\zeta, \bar{\zeta}$, where ζ is a primitive cube root of unity. The two-dimensional fixed space will be the Cartan subalgebra of G_2.

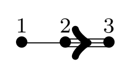
We will now list all positive roots of D_4 as linear combinations of simple roots (rather than fundamental weights), grouped into orbits under ρ:

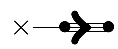
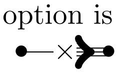


projections of norm 2/3 projections of norm 2

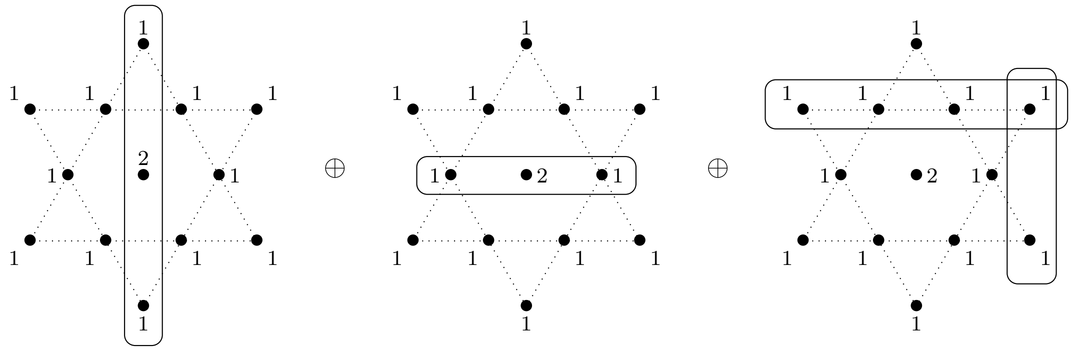
The picture for negative roots is almost the same. In the quotient, we have a root system with six roots of norm 2 and six roots of norm 2/3. Thus, the root system is G_2:



We will now work out the real forms. The affine Dynkin diagram is 

. We can delete the node with weight one, giving the compact form:  . The only other option is to delete the node with weight two, giving a real form with compact subalgebra $A_1 \times A_1$:  . So this second one must be the split form, because there is nothing else the split form can be.

We will say a bit more about the split form of G_2. What does the split G_2 Lie algebra look like as a representation of its maximal compact $A_1 \times A_1$? The answer is small enough that we can just draw a picture:



In the left and middle, we have drawn the two orthogonal A_1s; each uses one of the two elements in the Cartan at the origin. On the right, we have drawn the tensor product of an irreducible four-dimensional A_1 representation (the horizontal row) and an irreducible two-dimensional A_1 representation (the two vertical column); the total representation consists of the eight roots in the two rows of of length four. So as a representation of the compact $A_1^{(\text{horizontal})} \times A_1^{(\text{vertical})}$, we have decomposed into irreducibles $G_2 = 3 \otimes 1 + 1 \otimes 3 + 4 \otimes 2$.

All together, we can determine exactly what the maximal compact subgroup is. It is a quotient of the simply-connected compact group $\mathrm{SU}(2) \times \mathrm{SU}(2)$, with Lie algebra $A_1 \times A_1$. Just as for E_8, we need to identify which elements of the center $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ act trivially on G_2. Since we've decomposed G_2, we can compute this easily. A non-trivial element of the center of $\mathrm{SU}(2)$ acts as 1 on odd-dimensional representations and as -1 on even-dimensional representations. So the element $(-1,-1) \in \mathrm{SU}(2) \times \mathrm{SU}(2)$ acts trivially on $3 \otimes 1 + 1 \otimes 3 + 4 \times 2$. Thus the maximal compact subgroup of the non-compact simple G_2 is $(\mathrm{SU}(2) \times \mathrm{SU}(2))/(-1,-1) \cong \mathrm{SO}_4(\mathbb{R})$. $\lozenge$

8.3.3.10 Remark We have constructed $3 + 4 + 5 + 3 + 2 = 17$ (from E_8, E_7, E_6, F_4, G_2) real forms of exceptional simple Lie groups.

There are another five exceptional real Lie groups. Take *complex* groups $E_8(\mathbb{C})$, $E_7(\mathbb{C})$, $E_6(\mathbb{C})$, $F_4(\mathbb{C})$, and $G_2(\mathbb{C})$, and consider them as *real* Lie groups. As real Lie groups they are still simple, of dimensions 248×2, 133×2, 78×2, 52×2, and 14×2. $\lozenge$

Exercises

1. Show that the number of norm-6 vectors in the E_8 lattice is 240×28, and they form one orbit under the $\mathfrak{W}(E_8)$ action.

2. (a) Show that $\mathrm{SU}(2) \times E_7(\text{compact})/(-1,-1)$ is a subgroup of $E_8(\text{compact})$.

   (b) Show that $\mathrm{SU}(9)/(\mathbb{Z}/3\mathbb{Z})$ is also a subgroup of $E_8(\text{compact})$.

   C.f. Example 8.1.2.15.

3. Let L be an even lattice and $\hat{L}$ as in the discussion following Remark 8.2.1.3. Prove that any automorphism of L preserving $(,)$ lifts (noncanonically!) to an automorphism of $\hat{L}$.

4. Check that the asserted homomorphisms in Remark 8.2.1.4 are.
5. Check that the bilinear form defined in Lemma 8.3.1.4 is in fact invariant.

Chapter 9

Algebraic Groups

9.1 General facts about algebraic groups

Compact groups and reductive algebraic groups over $\mathbb{C}$ are the same thing, as we will see in the semisimple case. But this does not cover the characteristic-p case, or even nilpotent groups. Certain things are easier to do in the framework of algebraic groups, and certain things are easier in the Lie framework.

We pick $\mathbb{K}$ algebraically closed and characteristic 0. An (affine) *algebraic group* is an algebraic variety G with group structure $m : G \times G \to G$, $i : G \to G$ that are all morphisms of algebraic varieties. Then it's clear that the shift maps (left- and right-multiplication) are algebraic.

Some facts:

9.1.0.1 Proposition *If $f : G \to H$ is a homomorphism of algebraic groups, then its image is Zariski-closed.*

Henceforth, "closed" means Zariski-closed.

Proof We will use the following fact from algebraic geometry. For any algebraic map of varieties $f : X \to Y$, the image $f(X)$ contains an open dense set inside $\overline{f(X)}$.

So, let $U \subseteq f(G)$ open with $\overline{U} = \overline{f(G)}$. Then for $y \in \overline{f(G)}$, we have $yU \cap U$ non-empty, as it is again Zariski-dense open. But then $y \in U \cdot U$, and so $\overline{f(G)} = U \cdot U = f(G)$. □

Let $m : G\times G \to G$ be the multiplication, and pull it back to $\Delta : \mathbb{K}[G] \to \mathbb{K}[G\times G] = \mathbb{K}[G]\otimes\mathbb{K}[G]$ via $\Delta f = \sum_{i=1}^{s} f_i \otimes f^i$ where $f(gx) = \sum_{i=1}^{s} f_i(g)f^i(x)$. But then the image of the action of the group on $\mathbb{K}[G]$ lies in the span of finitely many functions: $g \cdot f(x) \in \mathrm{span}\{f^i(x)\}$. Therefore:

9.1.0.2 Proposition *Any finite-dimensional subspace $W \subseteq \mathbb{K}[G]$ (considered as a G-module with respect to left translation) is contained in some G-invariant finite-dimensional subspace.* □

9.1.0.3 Proposition *If G is an algebraic group, then it has a finite-dimensional faithful representation.*

Proof Pick regular functions that separate points — you can always do this with finitely many of them — and consider the finite-dimensional invariant space containing them. □

9.1.0.4 Example Every semisimple simply-connected Lie group over $\mathbb{C}$ is algebraic, as we proved in Section 6.2. However, this fails over $\mathbb{R}$. For example, it's easy to see directly that $\pi_1 \mathrm{SL}(2,\mathbb{R}) = \mathbb{Z}$, and so the simply-connected cover of $\mathrm{SL}(2,\mathbb{R})$ has infinite discrete center, which in particular cannot be Zariski-closed. So the simply-connected cover of $\mathrm{SL}(2,\mathbb{R})$ is not an algebraic group over $\mathbb{R}$. Indeed, every finite-dimensional representation of the simply-connected cover of $\mathrm{SL}(2,\mathbb{R})$ gives a finite-dimensional representation of $\mathfrak{sl}(2,\mathbb{R})$, hence of $\mathfrak{sl}(2,\mathbb{C})$, hence of $\mathrm{SL}(2,\mathbb{C})$, hence of $\mathrm{SL}(2,\mathbb{R})$ itself. So the simply-connected cover of $\mathrm{SL}(2,\mathbb{R})$ lacks a faithful finite-dimensional representation, and so cannot be algebraic by Proposition 9.1.0.3. $\Diamond$

9.1.0.5 Proposition *If $H \subseteq G$ is a (Zariski-)closed subgroup, and both are algebraic, then H is cut out by an ideal I_H in $\mathbb{K}[G]$. Then I_H is clearly an H-invariant subspace. So we can ask about the normalizer in G of I_H. In fact, $H = \{g \in G \text{ s.t. } f(gx) \in I_H \; \forall f \in I_H\}$.* $\square$

The following fact is true for semisimple Lie algebras (c.f. Lemma/Definition 5.3.2.2), but not Lie algebras in general. Recall Theorem 4.2.5.1: if you pick any $g \in \mathrm{GL}(V)$, then you can write $g = x_s + x_n$, where x_s is semisimple and x_n is nilpotent, and $[x_s, x_n] = 0$. In fact, these conditions uniquely pick out x_s, x_n, and it turns out that there are polynomials p, q depending on g so that $x_s = p(g)$ and $x_n = q(g)$. Moreover, if $g \in \mathrm{GL}(V)$, then x_s is also invertible, although x_n never is. So we write $g = x_s(1 + x_s^{-1} x_n) = g_s g_n$, and:

9.1.0.6 Theorem (Group Jordan-Chevalley decomposition)
Each $g \in \mathrm{GL}(V)$ factors uniquely as $g = g_s g_n$ where g_s is semisimple and $g_s(g_n - 1)$ is nilpotent. $\square$

9.1.0.7 Proposition *Pick an algebraic embedding $G \hookrightarrow \mathrm{GL}(V)$, and pick $g \in G$. Then it follows from Theorem 9.1.0.6 that $g_s, g_n \in G$. Indeed, you write $G = \{g \in \mathrm{GL}(V) \text{ s.t. } f(gx) = I_G \; \forall f \in I_G\}$. But then any polynomial of g leaves I_G invariant. The elements $g_s, g_n \in G$ do not depend on the embedding $G \hookrightarrow \mathrm{GL}(V)$.* $\square$

This also all works in Lie algebras, where you think in terms of the adjoint action by derivations: a Lie algebra of an algebraic group is closed under Jordan-Chevalley decompositions. So if you can present a Lie algebra that's not closed under the JC decomposition, then it is not algebraic.

9.1.0.8 Example Let $G = \left\{ \begin{pmatrix} 1 & y & z \\ 0 & e^x & xe^x \\ 0 & 0 & e^x \end{pmatrix} \text{ s.t. } x, y, z \in \mathbb{C} \right\}$. The bottom corner is $\exp\begin{pmatrix} x & x \\ 0 & x \end{pmatrix}$, so this is a closed linear group. But its Lie algebra consists of matrices of the form $\begin{pmatrix} 0 & y & z \\ 0 & x & x \\ 0 & 0 & x \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x \end{pmatrix} + \begin{pmatrix} 0 & y & z \\ 0 & 0 & x \\ 0 & 0 & 0 \end{pmatrix}$ and so is not closed under the JC composition. So G is not an algebraic group. $\Diamond$

9.1.0.9 Proposition *Let G be semisimple and connected $G \subseteq \mathrm{GL}(V)$. Then G is algebraic.*

Proof Look at the G-action in $\mathrm{End}(V)$. Then $\mathrm{End}(V) = \mathfrak{g} \oplus \mathfrak{m}$, and $[\mathfrak{g}, \mathfrak{m}] \subseteq \mathfrak{m}$, because any representation of a semisimple group is completely reducible. So take the connected component $\mathcal{N}_{\mathrm{GL}(V)}(\mathfrak{g})_0$ of the normalizer:

$$\mathcal{N}_{\mathrm{GL}(V)}(\mathfrak{g}) = \{g \in \mathrm{GL}(V) \text{ s.t. } gxg^{-1} \in \mathfrak{g} \; \forall x \in \mathfrak{g}\}$$

Then $\mathcal{N}_{\mathrm{GL}(V)}(\mathfrak{g})$ acts on $\mathfrak{g}$ by automorphisms, and $G \subseteq \mathcal{N}_{\mathrm{GL}(V)}(\mathfrak{g})_0$ as the inner automorphisms of $\mathfrak{g}$. But it follows from Theorem 4.4.3.9 that the connected component of the $\mathrm{Aut}\,\mathfrak{g}$ consists of inner automorphisms. So there is a projection $\mathcal{N}_{\mathrm{GL}(V)}(\mathfrak{g})_0 \twoheadrightarrow G/\text{center}$, and the kernel is the intersection of the centralizer $\mathcal{Z}_{\mathrm{GL}(V)}(\mathfrak{g}) = \{g \in \mathrm{GL}(V) \text{ s.t. } gxg^{-1} = x\ \forall x \in \mathfrak{g}\}$ with $\mathcal{N}_{\mathrm{GL}(V)}(\mathfrak{g})_0$. And since G is connected, $g \in \mathrm{GL}(V)$ centralizes $\mathfrak{g}$ only if it commutes with G, and we have presented the connected component of the normalizer as a product:

$$\mathcal{N}_{\mathrm{GL}(V)}(\mathfrak{g})_0 = \big(G/\text{center}\big) \times \big(\mathcal{Z}_{\mathrm{GL}(V)}(\mathfrak{g}) \cap \mathcal{N}_{\mathrm{GL}(V)}(\mathfrak{g})_0\big)$$

Moreover, $\mathcal{Z}_{\mathrm{GL}(V)}(\mathfrak{g})_0 \hookrightarrow \mathcal{Z}_{\mathrm{GL}(V)}(\mathfrak{g}) \cap \mathcal{N}_{\mathrm{GL}(V)}(\mathfrak{g})_0$ as a normal subgroup with quotient the center of G, and again this is a product. All together, we can construct a surjection $\mathcal{N}_{\mathrm{GL}(V)}(\mathfrak{g})_0 \twoheadrightarrow \mathcal{Z}_{\mathrm{GL}(V)}(\mathfrak{g})_0$ with kernel G. But both $\mathcal{N}_{\mathrm{GL}(V)}(\mathfrak{g})_0$ and $\mathcal{Z}_{\mathrm{GL}(V)}(\mathfrak{g})$ are algebraic subvarieties of $\mathrm{GL}(V)$, and hence so is the kernel. □

9.1.0.10 Proposition *If V is any finite-dimensional representation of algebraic G, then we can construct a canonical the map $V \otimes V^* \to \mathbb{K}[G]$ corresponding to the coaction $V^* \to V^* \otimes \mathbb{K}[G]$. It is injective, and gives a homomorphism of left G-modules $V \hookrightarrow \mathbb{K}[G] \otimes V$, where G acts on $\mathbb{K}[G] \otimes V$ by multiplication in $\mathbb{K}[G]$ and trivially in V: $\mathbb{K}[G] \otimes V \cong \mathbb{K}[G]^{\oplus \dim V}$.* □

9.1.0.11 Remark In general, if your group is not reductive, then you can get finite-dimensional representations of arbitrary length (in the sense of Jordan-Holder series). ◊

9.1.1 Peter–Weyl theorem

9.1.1.1 Definition *Let G be an algebraic group. The* regular representation *of G is the algebra of functions $\mathbb{C}[G]$. It is a $G \times G$ module: if $f \in \mathbb{C}[G]$, then we set $(g_1, g_2)f|_x = f(g_1^{-1} x g_2)$. Each G action centralizes the other.*

9.1.1.2 Definition *Suppose we have groups G, H and modules $G \curvearrowright M$ and $H \curvearrowright N$. The* exterior tensor product *$M \boxtimes N$ is the vector space $M \otimes N$ with the obvious $G \times H$ action.*

The following fact is well-known:

9.1.1.3 Theorem (Peter–Weyl theorem for finite groups)
Let G be a finite group. Then $\mathbb{C}[G] = \bigoplus_{\text{irreps of } G} V \boxtimes V^*$ as $G \times G$ modules. □

Certainly any finite group is algebraic, and in fact the same statement holds for affine algebraic groups. We will prove:

9.1.1.4 Theorem (Peter–Weyl theorem for algebraic groups)
Let G be a connected simply-connected semisimple Lie group over $\mathbb{C}$. As $G \times G$ modules, we have:

$$\mathbb{C}[G] = \bigoplus_{\lambda \in P^+} L(\lambda) \boxtimes L(\lambda)^*$$

9.1.1.5 Remark The statement and proof hold for reductive groups, but we haven't defined those. ◊

Before giving the proof, we recall the following:

9.1.1.6 Lemma / Definition *Let G be a semisimple Lie group over $\mathbb{C}$ with Lie algebra $\mathfrak{g}$, and let $\mathfrak{h}$ be a choice of Cartan subalgebra. The* maximal torus *is $H = (\mathcal{N}_G(\mathfrak{h}))_0$, the connected component of the normalizer in G of $\mathfrak{h}$. Then* $\mathrm{Lie}(H) = \mathfrak{h}$*, and H is a* torus*: $H \cong \mathbb{C}^\times \times \cdots \times \mathbb{C}^\times$. The* exponential map $\exp : \mathfrak{h} \to H$ *is a homomorphism of abelian groups. We set $\Gamma = \ker \exp$. It is a* lattice*, or (discrete) free finite-rank abelian group.*

We set $\hat{H}$ to be the set of all 1-dimensional (irreducible) representations of H. The set of irreps of $\mathfrak{h}$ is just $\mathfrak{h}^*$. The map $\exp^* : \hat{H} \hookrightarrow \mathfrak{h}^*$ *corresponds to the subset:*

$$\hat{H} = \{\lambda \in \mathfrak{h}^* \text{ s.t. } \langle \lambda, \Gamma \rangle \subseteq \mathbb{Z}\}$$

Moreover, $\hat{H}$ is an abelian group, because we can tensor representations. When G is simply-connected, $\hat{H} = P$.

Any torus satisfies a Peter–Weyl theorem: $\mathbb{C}[H] = \bigoplus_{\Phi \in \hat{H}} \mathbb{C}\Phi$. Suppose that G is simply-connected, and let C_λ be the one-dimensional representation of H corresponding to $\lambda \in P$. Then $C_{-\lambda} = C_\lambda^*$ and $\mathbb{C}[H] = \bigoplus_{\lambda \in P^+} C_\lambda \boxtimes C_{-\lambda}$ as $H \times H$-modules. □

Proof (of Theorem 9.1.1.4) Since G has a faithful algebraic representation $G \hookrightarrow \mathrm{End}(V)$, $\mathbb{C}[G]$ is a quotient of $\mathbb{C}[\mathrm{End}(V)]$ by some invariant ideal, and $\mathbb{C}[\mathrm{End}(V)] = \mathcal{S}(V \otimes V^*)$ is a direct sum of finite-dimensional $G \times G$ representations. So $\mathbb{C}[G]$ decomposes as a direct sum of finite-dimensional irreducible representations. If G_1, G_2 are semisimple Lie groups with weight lattices P_1, P_2, then the weight lattice for $G_1 \times G_2$ is just $P_1 \times P_2$, and the irreducible representations are $L(\lambda \times \mu) = L(\lambda) \boxtimes L(\mu)$. So we are interested in $\mathrm{Hom}_{G\times G}\big(L(\lambda) \boxtimes L(\mu), \mathbb{C}[G]\big)$ for $\lambda, \mu \in P^+$.

When $L(\mu) = L(\lambda)^*$, there is a distinguished map $j_\lambda : L(\lambda) \boxtimes L(\lambda)^* \hookrightarrow \mathbb{C}[G]$, the *matrix coefficient*, defined by:

$$j_\lambda(v \otimes \varphi)\,(g) = \langle g^{-1}v, \phi \rangle$$

Then the sum of all matrix coefficients gives an injection $\bigoplus_{\lambda \in P^+} L(\lambda) \boxtimes L(\lambda)^* \hookrightarrow \mathbb{C}[G]$. To show that there are no other direct summands requires a bit more preparation.

Let $\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}^+$, and let $N^\pm$ be the group with Lie algebra $\mathfrak{n}^\pm$. Let H denote the maximal torus of Lemma/Definition 9.1.1.6. Then multiplication gives an algebraic map $N^+ \times H \times N^- \to G$ with Zariski-dense image, the *big Bruhat cell* N^+HN^-.

Pick $\lambda, \mu \in P^+$ and $\psi \in \mathrm{Hom}_{G\times G}\big(L(\lambda) \boxtimes L(\mu), \mathbb{C}[G]\big)$. Let v be a highest vector in $L(\lambda)$ and let w be a lowest vector in $L(\mu)$. Then $N^+v = v$ and $N^-w = w$. Suppose that $\psi(v \otimes w) = f \in \mathbb{C}[G]$. Then for any $n_\pm \in N^\pm$ and any $g \in G$ we have $f(n_+ g n_-) = f(g)$. But f is determined by its values on any Zariski-dense set, e.g. N^+HN^-, and so f is determined by its values on H.

Moreover, we know what happens to the vectors when we multiply them by elements of the torus. Let $h \in \mathfrak{h}$. Then $(\exp h)v = e^{\lambda(h)}v$ and $(\exp h)w = e^{\mu'(h)}w$, where μ' is the lowest weight of $L(\mu)$. So $f\big((\exp h_1)h(\exp h_2)\big) = e^{\lambda(-h_1)}e^{\mu'(h_2)}f(h)$ for $h \in H$ and $h_1, h_2 \in \mathfrak{h}$. On the other hand, H is commutative, so we must have $\mu' = -\lambda$ and $f|_H \in C_\lambda \boxtimes C_{-\lambda}$, which is one-dimensional. But the irrep with lowest weight $-\lambda$ is $L(\lambda)^*$.

Therefore:

$$\dim \mathrm{Hom}_G\big(L(\lambda) \boxtimes L(\mu), \mathbb{C}[G]\big) = \begin{cases} 1, & L(\mu) = L(\lambda)^*, \\ 0, & \text{otherwise.} \end{cases}$$ □

As a corollary, we return to our earlier discussion of $\mathcal{U}\mathfrak{g}$ and the nilpotent cone. We will prove:

9.1.1.7 Proposition *As in Remark 9.4.3.8, choose the space Y in Step 2 of the proof of Theorem 9.4.3.7 to be $\mathfrak{g}$-invariant. Then Y decomposes as:*

$$Y = \bigoplus_{\lambda \in P^+} L(\lambda)^{\oplus m_\lambda}$$

where the multiplicities are $m_\lambda = \dim L(\lambda)_0 = \dim L(\lambda)^{\mathfrak{h}}$.

For example, $m_\lambda \neq 0$ implies that $\lambda \in Q$. This is not surprising: only the root lattice appears as weights of $\mathcal{U}\mathfrak{g}$.

9.1.1.8 Example When $\mathfrak{g} = \mathfrak{sl}(2)$, each representation of even weight appears with multiplicity 1. ◊

To prove Proposition 9.1.1.7, we first give a series of lemmas, many of which are of independent interest. The idea of the proof is as follows: $G \cdot x$ is not closed, but we can deform it to $G \cdot h$ which is, where h is from the principal $\mathfrak{sl}(2)$, and hence a semisimple element. Then we will prove that $Y \cong \mathbb{C}[G \cdot h]$. This kind of approach doesn't always work; it's rather specific to this situation.

9.1.1.9 Lemma *If $z \in \mathfrak{g}_{\text{ss}}$ (the semisimple elements), then the adjoint orbit $G \cdot z$ is closed.*

Proof Suppose $z' \in \overline{G \cdot z}$. If $p(t)$ is the minimum polynomial for $\operatorname{ad}(z)$, then it also annihilates $\operatorname{ad}(z')$. So the minimum polynomial of z' can only have smaller degree. The characteristic polynomials are the same: $\det(\operatorname{ad}(z) - t) = \det(\operatorname{ad}(z') - t)$, because the characteristic polynomial is invariant, so constant on orbits, and one is in the closure of the orbit of the other. Therefore all multiplicities of eigenvalues are the same, and in particular the multiplicities of the zero eigenvalue are the same. Then $\dim \ker \operatorname{ad}(z') = \dim \ker \operatorname{ad}(z)$. So $\dim G \cdot z = \dim G \cdot z'$, and hence $z' \in G \cdot z$. □

9.1.1.10 Lemma *Let $h \in \mathfrak{g}$ correspond to the principal $\mathfrak{sl}(2)$, and let $r : \mathbb{C}[\mathfrak{g}] \to \mathbb{C}[G \cdot h]$ be the restriction map. Then $r : Y \to \mathbb{C}[G \cdot h]$ is an isomorphism.*

Proof Surjectivity follows from the fact that on an orbit $r(f_i)$ are constants. Injectivity is the interesting part. Remember that Y is graded; so pick $q_1, \dots, q_m \in Y$ homogeneous and linearly independent. We want to show that their images $r(q_1), \dots, r(q_m)$ are also linearly independent.

So assume the opposite. We use the notation ϕ from Step 1 of the proof of Theorem 9.4.3.7; then $\phi_h : G \to \mathfrak{g}^*$ birationally. By the assumption, $\dim \phi_h^*(q_1, \dots, q_m) < m$. We can multiply h by any constant: because each q_i is homogeneous, we have for any $t \in \mathbb{C}^\times$ that $\dim \phi_{th}^*(q_1, \dots, q_m) < m$. On the other hand, you can check that $th + x \in G \cdot th$. So $\dim \phi_{th+x}^*(q_1, \dots, q_m) < m$. But this rank is a semicontinuous function, so we can take $t = 0$: $\dim \phi_x^*(q_1, \dots, q_m) < m$. But then $q_1, \dots, q_m$ are linearly dependent on $G \cdot x$, and therefore on $\mathcal{N}$. But this is a contradiction. Thus $Y \to \mathbb{C}[\mathcal{N}]$ is an isomorphism. □

9.1.1.11 Remark This was a good trick. You see what happened: you have generic orbits, which are closed, because they are maximal dimension. And then you have nongeneric orbits, but they are still in the families. ◊

9.1.1.12 Lemma $\mathrm{Stab}_G(h) = H$.

Proof First of all, $\mathfrak{h}$ is the centralizer of h in $\mathfrak{g}$. So $N_G(\mathfrak{h}) \subseteq \mathrm{Stab}_G(h)$. But $N_G(\mathfrak{h})/H \cong W$ and $\mathrm{Stab}_W h = \{e\}$ by regularity. □

Proof (of Proposition 9.1.1.7) The map $\xi^* : \mathbb{C}[G \cdot h] \hookrightarrow \mathbb{C}[G]$ dual to $\xi : G \twoheadrightarrow G \cdot h$ is a homomorphism of (left) G-modules. By Lemma 9.1.1.12 we identify $G/H \cong G \cdot h$, and $\mathrm{Im}\,\xi^* = \{f \in \mathbb{C}[G] \text{ s.t. } f(xg) = f(x)\ \forall g \in H\}$. Then:

$$\mathbb{C}[G]^{H_{\text{right}}} \cong \left(\bigoplus L(\lambda) \boxtimes L(\lambda)^*\right)^{H_{\text{right}}} \cong \bigoplus L(\lambda) \boxtimes L(\lambda)^H,$$

as $(L(\lambda)^*)^H \cong L(\lambda)^H$. □

9.2 Homogeneous spaces and the Bruhat decomposition

9.2.1 Homogeneous spaces

Suppose we have an algebraic group G and a Zariski-closed subgroup $H \subseteq G$. Then $X = G/H$ makes sense as a topological space, and even, upon realizing G, H as Lie groups, X makes sense as a manifold. How about as an algebraic variety? The problem in the algebraic case is that even if G, H are affine, then X is not affine generally, and so we cannot just write down X as the spectrum of something. Instead, we will use a trick due to Chevellay.

For the remainder of this section, we will not write the word "Zariski": "closed" means "Zariski-closed." By "algebraic group" we mean "affine algebraic group."

9.2.1.1 Proposition *Let G be an algebraic group and $H \subseteq G$ a closed subgroup. Then there exists a representation V of G and a line $\ell \subseteq V$ such that $H = \mathrm{Stab}_G\,\ell$.*

Proof Recall Proposition 9.1.0.5: if $I_H \subseteq \mathbb{K}[G]$ is the ideal of functions vanishing on H, then $H = \mathrm{Stab}_G(I_H)$ is the stabilizer of the ideal. Since $\mathbb{K}[G]$ is Noetherian, any ideal is finitely generated; we pick up generators $f_1, \dots, f_n$ of I_H, and by Proposition 9.1.0.2 there exists a finite-dimensional G-invariant subspace $\tilde{V} \subseteq \mathbb{K}[G]$ containing $f_1, \dots, f_n$.

We set $W = I_H \cap \tilde{V}$. It is an easy exercise to show that $H = \mathrm{Stab}_G(W)$: it contains all the generators. If $\dim W = d$, then to get a line we can take powers. Set $V = \tilde{V}^{\wedge d}$, and $\ell = W^{\wedge d}$. It's immediate that $H \subseteq \mathrm{Stab}_G(\ell)$, and the reverse inclusion is almost as immediate. □

As a corollary, we have:

9.2.1.2 Lemma / Definition *The quotient G/H can be defined as the algebraic space $X \cong G \cdot [\ell]$, where $[\ell]$ is the point in $\mathbb{P}(V)$ corresponding to the line ℓ in V in Proposition 9.2.1.1. It is* locally closed *— the image of an algebraic map — in $\mathbb{P}(V)$, and hence a* quasiprojective variety *— something that can be embedded in a projective space.* □

Now we will discuss the special case that H is normal. Then X is a group, and the point is that it's affine algebraic:

9.2.1.3 Proposition *If $H \subseteq G$ is closed and normal, then there exists a representation $\pi : G \to \mathrm{GL}(V)$ such that $H = \ker \pi$.*

In particular, $G/H \hookrightarrow \mathrm{GL}(V)$ will be closed, as any locally closed subgroup in a group is closed, and so affine.

Proof We start with V' and $\ell' \subseteq V'$ a line such that $H = \mathrm{Stab}_G(\ell')$, as in Proposition 9.2.1.1. This choice defines a character $\chi : H \to \mathbb{K}$ by $hv = \chi(h)v$, where $v \in \ell$. Recall that the set $\hat{H}$ of characters of a normal subgroup H of G carries a G-action by $(g \cdot \chi)(h) = \chi(g^{-1}hg)$.

For each $\eta \in \hat{H}$ we set $V'_\eta \overset{\text{def}}{=} \{v \in V' \text{ s.t. } hv = \eta(h)v \, \forall h \in H\}$, and we set $W \overset{\text{def}}{=} \bigoplus_{\eta \in \hat{H}} V'_\eta$. Then W is G-invariant, as G permutes the V'_ηs. Similarly, we construct V as a sum of matrix algebras:

$$V \overset{\text{def}}{=} \bigoplus \mathrm{End}_{\mathbb{K}}(V'_\eta)$$

We let G act via conjugation to construct the representation $\pi : G \to \mathrm{GL}(V)$, and it's an easy exercise to calculate the kernel $\ker \pi = H$. $\square$

This explains why we can quotient by any normal subgroup and get something affine. We will not prove that the constructions above do not depend on the choice of representation — that $X = G/H$ as an algebraic space does not depend on the choice of V — nor that $G/H \to \mathrm{GL}(V)$ is actually a morphism of algebraic groups, but both statements are true.

9.2.1.4 Proposition *Suppose that G is an abelian connected affine algebraic group. Then its only projective homogeneous space is a point.*

Proof Since G/H is a group, it's affine, but it is also projective, and the only connected affine projective space is a point. $\square$

9.2.1.5 Example (Warning) The (real) torus T^2 as a complex analytic space is a homogeneous space for $\mathbb{C}$, and it is (isomorphic to) a projective space, namely an elliptic curve. But the $\mathbb{C}$ and $\mathbb{C}^\times$ actions are not algebraic: $T^2 = \mathbb{C}^\times/\Gamma$, but Γ, being an infinite discrete group, is not Zariski-closed. $\lozenge$

So there are ways that the algebraic category and the Lie category are quite different.

9.2.2 Solvable groups

9.2.2.1 Proposition *If G is algebraic, then $G' = [G, G]$ is algebraic.*

In Lie groups, it can happen that G' is not Lie: you need G to be simply-connected in order to give G' a manifold structure.

Proof Let $Y_g = gGg^{-1}G$. It is *constructible* — a disjoint union of locally closed sets. Then:

$$\overline{G'} = \overline{\bigcup_{\text{finite subsets of } G} \prod_{g \in \text{subset}} Y_g} = \overline{Y_{g_1} \cdots Y_{g_n}}$$

The point is that, by the Noetherian condition, the union is finite: you order the elements, construct a chain (you might as well assume that the subsets are growing), and chains are only finitely long. Then you can take the top one. $\square$

9.2.2.2 Definition *A group G is* solvable *if the chain:*

$$G \supseteq G' \supseteq G'' \supseteq \dots$$

stops at $G^{(n)} = \{1\}$. (C.f. Definition 4.2.1.4.)

9.2.2.3 Example The fundamental example is the group $\mathrm{N}(n)$ of upper triangular matrices in $\mathrm{GL}(n)$. $\lozenge$

9.2.2.4 Proposition *Let G be a connected solvable group acting on a projective variety X. Then G has a fixed point on X.*

Proof If G is abelian, then you pick a closed orbit, which always exists, and it must be a fixed point by Proposition 9.2.1.4.

Now we do induction on dimension. Let Y be the set of points fixed by G' — it is an exercise to show that if G connected then so is G', and in particular Y is nonempty. Then since G' is normal, Y is G-invariant. So actually the G-action factors through G/G', but this is abelian, and we are done. $\square$

This has many nice consequences:

9.2.2.5 Theorem (Lie–Kolchin)
If G is connected solvable and V an n-dimensional representation of G, then there is a full flag fixed by G. In other words, $G \hookrightarrow \mathrm{N}(V)$. More precisely, we have $0 = V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_n = V$ invariant under G.

Proof Take the variety of all flags. This is a projective variety, and so we use Proposition 9.2.2.4. $\square$

This is the group version of Lie's theorem (Theorem 4.2.3.2). It's amazing how some things become so easy in the algebraic category. The hard part is algebraic geometry.

9.2.2.6 Definition *Let G be an affine algebraic group. A* Borel subgroup *is a maximal solvable connected closed subgroup.*

9.2.2.7 Proposition *If B is a Borel subgroup in G, then G/B is projective. Any two Borel subgroups in G are conjugate.*

Proof Pick a faithful representation $G \hookrightarrow \mathrm{GL}(V)$. By Theorem 9.2.2.5, there is a B-invariant flag $F \in \mathrm{Fl}(V)$. The G-stabilizer of this flag is necessarily in $\mathrm{N}(V) = \mathrm{Stab}_{\mathrm{GL}(V)}(F)$, and in particular it is solvable. Hence its connected component must be B, by maximality: $B = \mathrm{Stab}_G(F)_0$. Then $G \cdot F$ is closed, because if it is not, then we have an orbit of smaller dimension, which then has stabilizer a larger group, but we picked B maximal. So $G \cdot F = G/\mathrm{Stab}_G(F)$ is a closed subset of

$\mathrm{Fl}(V)$, and hence projective, and it is the quotient of G/B by $\pi_0\bigl(\mathrm{Stab}_G(F)\bigr)$, a finite group. This proves projectivity.

Let B' be another Borel group. Then it acts on G/B via the G action, and by Proposition 9.2.2.4 there is a B'-fixed point $x \in G/B$. Picking $g \in G$ so that $x = gB \in G/B$, we see that g conjugates $\mathrm{Stab}_G\, x$ to B. By maximality, $B' = \mathrm{Stab}_G\, x$. □

9.2.2.8 Lemma / Definition *Let P be a closed subgroup in G. Then G/P is projective iff P contains some Borel. Such a subgroup is called* parabolic.

Proof In one direction it should be clear: we have a map $G/B \to G/P$ if $B \subseteq P$, and the image of a projective variety is projective.

In the other direction, suppose that G/P is projective. Let B be some Borel in G. Then by Proposition 9.2.2.4, B has a fixed point $x \in G/P$. Then $\mathrm{Stab}_G\, x = gPg^{-1} \supseteq B$, so $P \supseteq g^{-1}Bg$. □

9.2.2.9 Lemma / Definition *We define the* nilradical *of a group G as* $\mathrm{Nil}(G) \overset{\text{def}}{=} \bigl(\bigcap_{\pi\in\mathrm{Irr}(G)} \ker\pi\bigr)_0$. *It is normal and* unipotent *— all elements are unipotent — and maximal with respect to these properties. Let V be a faithful representation of G, and pick a Jordan-Holder series $V \supsetneq V_1 \supsetneq \cdots \supsetneq V_k$ so that each V_i/V_{i+1} is irreducible. Then with respect to this flag,* $\mathrm{Nil}(G)$ *consists of upper-triangulars with* 1*s on the diagonal. The* nilradical *of the Lie algebra is* $\mathrm{nil}\,\mathfrak{g} = \mathrm{Lie}(\mathrm{Nil}\, G)$. *(C.f. Chapter 4, Exercise 4.)* □

9.2.2.10 Lemma / Definition *A group G is* reductive *if it is a quotient of $G_{\mathrm{ss}} \times T/$finite group, where G_{ss} is semisimple and T is a torus. An algebraic group G is reductive iff* $\mathrm{Nil}(G) = \{1\}$. □

9.2.2.11 Lemma / Definition *Let $p : G \to G/\mathrm{Nil}(G)$ be the projection. The* radical *of G is* $\mathrm{Rad}(G) \overset{\text{def}}{=} p^{-1}\bigl(\mathcal{Z}(G/\mathrm{Nil}\,G)_0\bigr)_0$. *It is a maximal normal connected solvable subgroup in G, and hence algebraic. We have* $\mathrm{Lie}(\mathrm{Rad}\, G) = \mathrm{rad}\,\mathfrak{g}$, *and $G/\mathrm{Rad}\, G$ is semisimple.* □

9.2.2.12 Remark This is almost the Levi decomposition, which finds G built from a solvable and a semisimple. But in the Levi decomposition, "built" means a semidirect product, whereas here we only have an extension. Indeed, it's not true that you can write $\mathrm{GL}(n)$ as a semidirect product: you need to take a quotient. ◊

9.2.2.13 Lemma *Every parabolic subgroup of G contains* $\mathrm{Rad}\, G$.

Proof $\mathrm{Rad}\, G$ has a fixed point on G/P, but since $\mathrm{Rad}\, G$ is normal it acts trivially on all of G/P. □

So for the purpose of understanding projective homogeneous spaces, you can forget about $\mathrm{Rad}\, G$ completely, since it doesn't contribute to the action of G on any G/P. In particular, replacing G by $G/\mathrm{Rad}\, G$ gives:

9.2.2.14 Corollary *Any projective homogeneous space for any group is isomorphic to G/P where G is semisimple is P is some parabolic subgroup.* □

Henceforth, we assume that G is connected. If we allow disconnected groups, then classifying homogeneous spaces is as hard as classifying finite groups. In the next few sections we will classify connected semisimple groups and their parabolic subgroups, and thereby classify all connected projective homogeneous spaces.

9.2.3 Parabolic Lie algebras

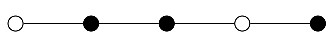
We let G be connected and semisimple.

Pick a parabolic subgroup P of G, and let $\mathfrak{p} = \mathrm{Lie}(P)$. Then $\mathfrak{p} \subseteq \mathfrak{b}$ for some Borel, and they are all conjugate, so let's pick one: $\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}^+$ with $\mathfrak{b} = \mathfrak{h} \oplus \mathfrak{n}^+$, the *standard Borel.* We also recall the generators $h_1, \ldots, h_n, x_1, \ldots, x_n, y_1, \ldots, y_n$.

9.2.3.1 Lemma *There exists a subset $S \subseteq \{1, \ldots, n\}$ such that $\mathfrak{p}$ is generated by $\{h_1, \ldots, h_n, x_1, \ldots, x_n\}$ and y_j for $j \in S$.*

Proof This follows from a very simple fact. If $\alpha, \beta \in \Delta^+$, and $[y_\alpha, y_\beta] \in \mathfrak{p}$, then take $x_\alpha \in \mathfrak{p}$, and since y_α, x_α form an $\mathfrak{sl}(2)$-triple, then $[x_\alpha, [y_\alpha, y_\beta]] = cy_\beta$ for $c \neq 0$, so $y_\alpha, y_\beta \in \mathfrak{p}$. Then do induction. □

9.2.3.2 Example To represent a parabolic subalgebra, you draw the Dynkin diagram, and shade a few of the nodes: the unshaded nodes are S. For example, for $\mathfrak{sl}(6)$:



$$\leftrightarrow \begin{pmatrix} 0 & \boxed{*} & * & * & * & * \\ 0 & 0 & \boxed{0} & 0 & * & * \\ 0 & 0 & 0 & \boxed{0} & * & * \\ 0 & 0 & 0 & 0 & \boxed{*} & * \\ 0 & 0 & 0 & 0 & 0 & \boxed{0} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

◊

9.2.3.3 Lemma *1. Let $V = L(\lambda)$. Then $\mathbb{P}(V)$ has only one closed orbit: the orbit of the line of the highest vector, which we will call ℓ_λ.*

2. If V is not irreducible, then any closed orbit in $\mathbb{P}(V)$ is contained in $\mathbb{P}(W)$ for W some irreducible invariant subspace of V.

Proof Any closed orbit is projective, and has a fixed point, i.e. a fixed line in V. But if $V = L(\lambda)$ then it has a unique fixed line, and in general a fixed line is a highest weight space. The only thing that is not clear is if there are multiple isomorphic direct summands. But this part is a simple exercise. □

9.2.3.4 Corollary *P is the stabilizer of some $\ell_\lambda \subseteq L(\lambda)$.* □

9.2.3.5 Example If $P = B$ is the standard Borel, then λ is the sum of all the fundamental weights. For example, for $\mathfrak{sl}(n+1)$, this gives an embedding of the flag variety into $\mathbb{C}^{n+1} \otimes \bigwedge^2 \mathbb{C}^{n+1} \otimes \cdots \otimes \bigwedge^n \mathbb{C}^{n+1}$. ◊

9.2.3.6 Theorem (Classification of parabolic subgroups)
Suppose that G is connected and semisimple. Pick a system of simple roots Γ.

1. *Conjugacy classes of parabolic subgroups are in bijection with subsets of Γ. For $S \subseteq \Gamma$, we denote the corresponding parabolic subgroup containing the standard Borel by P_S.*

2. *If we pick* $\lambda = \sum_{i\in S} m_i\omega_i$*, with* $m_i > 0$*, then* $P_S = \mathrm{Stab}_G(\ell_\lambda)$*. Notice: it does not depend on the coefficients, just that they are non-zero.*

3. *If* P *is parabolic, then it is connected and* $\mathcal{N}_G(P) = P$*.*

9.2.3.7 Example The biggest parabolic is the whole group, whence S is empty; the smallest is the Borel, whence $S = \Gamma$. $\diamondsuit$

Proof By Corollary 9.2.3.4, the conjugacy class of any parabolic contains $P = \mathrm{Stab}_G(\ell_\lambda)$ for some irreducible representation $L(\lambda)$. If $\lambda = \mu + \nu$, then $\mathrm{Stab}_G(\ell_\lambda) = \mathrm{Stab}_G(\ell_\mu) \cap \mathrm{Stab}_G(\ell_\nu)$. This is because $L(\mu) \otimes L(\nu)$ contains a unique canonical component isomorphic to $L(\lambda)$, and $\ell_\lambda = \ell_\mu \otimes \ell_\nu$. So something stabilizes ℓ_λ iff it stabilizes each of ℓ_μ, ℓ_ν. Therefore, if $\lambda = \sum_{i\in\Gamma} m_i\omega_i$, then $\mathrm{Stab}_G(\ell_\lambda) = \bigcap_{i \text{ s.t. } m_i\neq 0} \mathrm{Stab}_G(\ell_{\omega_i})$. So it depends only on the *support* of λ: the $i \in \Gamma$ so that $m_i \neq 0$. This proves that 1. implies 2.

Statement 1. follows from Lemma 9.2.3.1 if we can show that every parabolic subgroup is connected. Let P be a parabolic subgroup and P_0 the connected component of the identity. It is also parabolic. Using Corollary 9.2.3.4, let λ be a highest weight with $P = \mathrm{Stab}\,\ell_\lambda$, and λ_0 the highest weight with $P_0 = \mathrm{Stab}\,\ell_{\lambda_0}$. But $\mathrm{Lie}\,P = \mathrm{Lie}\,P_0$, and $\mathrm{Lie}\,\mathrm{Stab}\,\ell_\lambda$ determines λ. This proves statement 1. and the first part of statement 3.

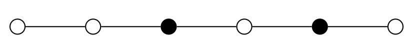
To finish 3., suppose that $P = \mathrm{Stab}_G(\ell_\lambda)$, and take $g \in \mathcal{N}_G(P)$. Then $g(\ell_\lambda)$ is fixed by P, but P has only one fixed point, because P contains B, and B has only one fixed point, so $g(\ell_\lambda) = \ell_\lambda$, so $g \in P$. $\square$

9.2.4 Flag manifolds for classical groups

In this section we discuss some of the classical flag manifolds. We will denote by G a classical semisimple Lie group, B its standard Borel subgroup, and P any parabolic that includes the standard Borel. Sometimes any G/P is called a *flag manifold*, and sometimes only G/B is the flag manifold and G/P are "partial flag manifolds".

For $S \subseteq \Gamma$, the corresponding parabolic P_S has a Levi decomposition: $P_S = G_S \rtimes \mathrm{Nil}(P_S)$, where G_S is reductive. The semisimple part $(G_S)'$ of G_S can be read from the diagram for P_S, simply by deleting the marked nodes from the diagram.

9.2.4.1 Example Let $G = \mathrm{SL}(n+1) = A_n$. Pick $k_1 < \cdots < k_s < n+1$; the flag manifolds are all of the form $\mathrm{Fl}(k_1, \ldots, k_s, n+1)$. For example, consider $\mathrm{SL}(7) = A_6$ with the third and fifth nodes marked:

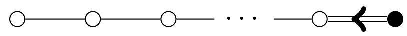


Then $(G_S)' = \mathrm{SL}(3) \times \mathrm{SL}(2) \times \mathrm{SL}(2)$, and we have the flag variety $\mathrm{Fl}(3, 5, 7)$. $\diamondsuit$

9.2.4.2 Example Now let's move to the types B_n, C_n, which are $\mathrm{SO}(2n+1)$ and $\mathrm{Sp}(2n)$. Let's work over $\mathbb{C}$. Then we have representations on $\mathbb{C}^{2n+1}$ with symmetric form $(,)$ or $\mathbb{C}^{2n}$ with antisymmetric form $\langle,\rangle$. The components of a flag are isotropic subspaces, and so never get to dimension past half the total:

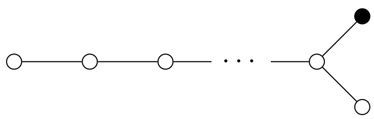
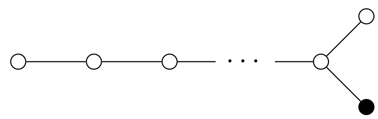
$$\mathrm{OFl}(m_1, \ldots, m_s) = \{V_1 \subsetneq \cdots \subsetneq V_s \text{ s.t. } (V_i, V_i) = 0\}$$
$$\mathrm{SpFl}(m_1, \ldots, m_s) = \{V_1 \subsetneq \cdots \subsetneq V_s \text{ s.t. } \langle V_i, V_i\rangle = 0\}$$

For example, take C_n with the last node marked:



This is the Lagrangian Grassmanian, i.e. the set of Lagrangian subspaces in $\mathbb{C}^{2n}$. $\Diamond$

9.2.4.3 Example Finally, D_n. Then $G = \mathrm{SO}(2n)$ acting on $\mathbb{C}^{2n}$, $(,)$. Then you have the same as before, but $\mathrm{OGr}(n, 2n)$, the Grassmanian of n-dimensional isotropic subspaces in $\mathbb{C}^{2n}$ has two connected components. The two components correspond to the last vertices of the Dynkin diagram:


,


How do you see that there are two components? For $n = 2$, it's clear: there are two isotropic lines $x = iy$ and $x = -iy$. In general, there is a "fermionic Fock space" construction that turns n-dimensional isotropic subspaces in $\mathbb{C}^{2n}$ into half spin representations of $\mathrm{Spin}(2n)$, of which there are two. $\Diamond$

9.2.5 Bruhat decomposition

In this section we prove the Bruhat decomposition theorem. We first fix a bit of notation. We let G denote a connected and semisimple complex Lie group. (In this section we will work over $\mathbb{C}$, but just about everything holds over an arbitrary field $\mathbb{K}$.) Its chosen maximal torus is T, the standard Borel is B, and the positive and negative parts are $N^{\pm}$. (Of course, $B = T \ltimes N^+$.) The Weyl group is $W = \mathrm{Stab}_G(T)/T$. In general, there does not exist a group embedding $W \hookrightarrow G$, but we can always find a set-theoretic map $W \hookrightarrow G$ so that the corresponding inner automorphisms act on T as they ought. For $w \in W$, the coset wB does not depend on the choice of embedding $W \hookrightarrow G$ (since the different choices differ by $t \in T \subset B$), and so we will fix an embedding once and for all and leave it out of the notation.

9.2.5.1 Theorem (Bruhat decomposition)
Every connected semisimple complex Lie group decomposes into Bruhat cells *as* $G = \bigsqcup_{w \in W} BwB = \bigsqcup_{w \in W} N^- wB$. *Thus the* flag manifold G/B *decomposes into* Shubert cells*: we set* $U_w = N^- wB$*, thought of as the* N^-*-orbit of* $wB \in G/B$*, and* $G/B = \bigsqcup_{w \in W} U_w$ *is a disjoint union of* $|W|$ *many* N^- *orbits. Each orbit is very simple as a topological space:* $U_w \cong \mathbb{C}^{\ell(w_0) - \ell(w)}$*, where* w_0 *is the longest element of* W*, and* $\ell(w)$ *is the length of* $w \in W$*.*

In fact, there are four Bruhat decomposition theorems, where we can replace one or both Bs in the first equality with the "negative" Borel B^-: the longest element of the Weyl group, swapping the positive roots for the negative ones, switches $B \leftrightarrow B^-$. We prefer the version $G = \bigsqcup B^- wB$, as it allows us to talk about highest vectors rather than lowest vectors. The second equality is obvious, using $B = T \ltimes N^+$ and $B^- = N^- \rtimes T$, and that $Tw = wT$. We first give an example and then explain the proof for $G = \mathrm{GL}(n)$, and then give the proof for general G.

9.2.5.2 Example For $G = \mathrm{SL}(2, \mathbb{C})$, the flag variety $G/B = \mathbb{P}^1$ is the Riemann sphere. There are two cells: the north pole U_{w_0}, which is fixed by N^-, and the *big Bruhat cell* U_e. $\Diamond$

In general, the Bruhat decomposition writes G and G/B as cell complexes. As the real dimensions are all even, we know the homology: $\dim H_{2i}(G/B, \mathbb{Z}) = \{w \in W \text{ s.t. } \ell(w) = i\}$.

Proof (of Theorem 9.2.5.1 for $G = \mathrm{GL}(n)$) When $G = \mathrm{GL}(n)$, the Bruhat decomposition is essentially an "LPU" decomposition, and follows from Gaussian elimination. The group B consists of upper-triangular matrices, N^- of lower-triangulars with 1s on the diagonal, the Weyl group is $W = S_n$, and we can inject $W \hookrightarrow G$ as permutation matrices. For each $x \in G$, we want to find $a \in N^-$ and $b \in B$ so that $axb = w \in W$, and we want to know how many ways we can do this.

When you think in terms of matrices, this is a very easy procedure. Multiplying on the left by a lower-triangular matrix is some operation on the rows of the matrix: in particular, you can pick any row and subtract from it any row above. Multiplying on the right by an upper-triangular is a column operation, again with the restriction that to any column you can add or subtract only columns to its left.

So we simply perform Gaussian elimination. Look at your matrix x, and look down the first column until you come to the first non-zero element. By multiplying on the left, you can make 0s all below it, and by multiplying on the right you can make zeros to the right of it and make that entry into a 1. So your matrix now looks like:

$$x = \begin{pmatrix} 0 & * & * & * \\ 0 & * & * & * \\ \neq 0 & * & * & * \\ * & * & * & * \end{pmatrix} \quad \rightsquigarrow \quad \begin{pmatrix} 0 & * & * & * \\ 0 & * & * & * \\ 1 & 0 & 0 & 0 \\ 0 & * & * & * \end{pmatrix}$$

Move over a column and repeat. At the end, you have your permutation matrix: this gives you one of your double cosets.

Why don't the double cosets intersect? The answer is that the procedure does not change which minors are non-zero. Pick up the first non-zero minor from the first column, and then find the first non-zero minor in the first two columns that contains the one you already picked up, etc. Doing this determines the shape of the permutation matrix w.

Finally, let's prove the claim made earlier that $U_w \cong \mathbb{C}^{\ell(w_0)-\ell(w)}$. Pick a permutation matrix w, and ask: "what conjugations by lower-triangulars don't break it?" Then $U_w \cong N^-/\operatorname{Stab}_{N^-} w$. Remember these are nilpotent groups, so $\exp : \operatorname{Lie}(N^-/\operatorname{Stab}_{N^-} w) \to N^-/\operatorname{Stab}_{N^-} w$ is an isomorphism, and so it suffices to count dimensions. But:

$$\begin{aligned} \dim(N^-/\operatorname{Stab}_{N^-} w) &= \#\{i > w(1)\} + \#\{i > w(2), i \neq i(1)\} + \dots = \\ &= \#\{(i < j) \text{ s.t. } w(i) < w(j)\} = \frac{n(n-1)}{2} - \ell(w) = \ell(w_0) - \ell(w) \end{aligned} \qquad \square$$

In fact, the general proof is even easier to write down, although less enlightening:

Proof (of Theorem 9.2.5.1 for general G) The outline of the proof is as follows. We use the fact that Theorem 9.2.5.1 holds for SL(2), which follows from the analysis above for $G = \mathrm{GL}(n)$, to conclude that $N^-WB = G$. We then show that the double cosets are disjoint. Finally, we will study the shape of the Shubert cells.

1. Recall that the Lie algebra $\mathfrak{g}$ is generated by special $\mathfrak{sl}(2)$-subalgebras $\mathfrak{g}_i = \langle x_i, h_i, y_i \rangle$. These lift to algebraic subgroups $G_i \subseteq G$, which may be simply connected SL(2)s or adjoint-form SO(3)s. G is generated by the G_is. So to prove that $G = N^- W B$, it suffices to show that $N^- W B$ is closed under multiplication by G_i.

   Since Theorem 9.2.5.1 holds for SL(2), we know that $G_i = N_i^- B_i \sqcup N_i^- s_i B_i$ — the Weyl group W_i for SL(2) consists of only the two elements $\{e, s_i\}$, where $s_i \in W_i \hookrightarrow W$ is the ith simple reflection. We consider the parabolic subgroup $P_i \overset{\text{def}}{=} G_i B = G_i \rtimes \mathrm{Nil}(P_i)$, where $\mathrm{Nil}(P_i) \subseteq B$. Then $P_i = N_i^- B \sqcup N_i^- s_i B$. We denote the torus for G_i by K_i — it is $K_i = T \cap G_i$. Then any element $g \in G_i$ is either:

   (a) $g = \exp(ay_i)\, K_i\, \exp(bx_i)$, or

   (b) $g = \exp(ay_i)\, s_i\, K_i\, \exp(bx_i)$

   where x_i, y_i are the generators of the ith $\mathfrak{sl}(2)$ and $a, b \in \mathbb{C}$.

   We wish to show that $N^- W B\, g = N^- W B$. We will work out case (a), and case (b) is similar and an exercise. We first observe that $K_i \exp(\beta x_i) \in B$, and so it suffices to show that $N^- w B \exp(a y_i) \subseteq N^- W B$. But $B \exp(a y_i) \subseteq P_i$, and so:

   $$N^- w B \exp(a y_i) \subseteq N^- w N_i^- B \;\sqcup\; N^- w N_i^- s_i B$$

   The third factor, in N_i^-, we denote by $\exp(a' y_i)$ for some $a' \in \mathbb{C}$: $N^- w B \exp(a y_i) = N^- w \exp(a' y_i) B$ or $N^- w \exp(a' y_i) s_i B$.

   Let α_i denote the ith simple root. Suppose that $w(\alpha_i) \in \Delta^+$; then $w \exp(a' y_i) w^{-1} = \exp(a' w(y_i)) \in N^-$ and $\exp(a' y_i) s_i = s_i \exp(a'' x_i)$, and so in either case $N^- w B \exp(a y_i) \subseteq N^- W B$. Alternately, suppose that $w(\alpha_i) \in \Delta^-$. Then we must have $w = \sigma s_i \tau$ for some $\sigma, \tau \in W$ with $\sigma(\alpha_i) \in \Delta^+$ and $\tau(\alpha_i) = \alpha_i$. But then $w \exp(a' y_i) = \sigma s_i \exp(a' y_i) \tau$, which is either $\sigma \exp(a'' y_i) s_i \exp(b' x_i) \tau$ or $\sigma \exp(a'' y_i) \exp(b' x_i) \tau$ for some $a'', b' \in \mathbb{C}$. Again in either case, upon multiplying by N^- on the left and B on the right we land in $N^- W B$.

   Along with case (b), which is similar and left as an exercise, we have shown that the multiplication map $N^- \times W \times B \to G$ is onto: every element of G is in some double coset.

2. We next want to show that the double cosets are all disjoint. Of course, being cosets, they are either disjoint or equal; what we want to show is that if $N^- w B = N^- \sigma B$ for $w, \sigma \in W$, then $\sigma = w$. To do this, we study N^--orbits on G/B. Fix a regular dominant weight λ (a weight in the positive Weyl chamber off any wall). By Theorem 9.2.3.6, we have an embedding of G-sets $G/B \hookrightarrow \mathbb{P}(L(\lambda))$, given by $eB \mapsto \ell_\lambda$, the line of highest weight vectors. Since λ is regular, the W-orbit of λ has $|W|$ many elements.

   Then $U_w = N^- \ell_{w(\lambda)}$, where $\ell_{w(\lambda)}$ is the line of elements in $L(\lambda)$ of weight $w(\lambda)$. Since $N^- = \exp(\mathfrak{n}^-)$ can only move down, the weights of any element of U_w are all at most $w(\lambda)$. So if $U_w = U_\sigma$, then we must have $w(\lambda) \leq \sigma(\lambda)$ and $w(\lambda) \geq \sigma(\lambda)$, and hence $\sigma = w$.

3. Finally, we will show that $U_w \cong \mathbb{C}^{\ell(w_0) - \ell(w)}$.

As above, we think of U_w as some N^--orbit in $\mathbb{P}L(\lambda)$. Consider the extremal weight $\mu = w(\lambda)$ and pick a weight vector $v_\mu \in \ell_\mu$. The map $\overline{\exp} : \mathfrak{n}^- \to G/B = \mathbb{P}(G \cdot \ell_\lambda)$ given by $\overline{\exp}(x) = \exp(x) v_\mu$ is algebraic, as $\mathfrak{n}^-$ is nilpotent, and satisfies $\overline{\exp}^{-1}(v_\mu) = \operatorname{Stab}_{\mathfrak{n}^-}(v_\mu)$.

This subalgebra has a root decomposition:

$$\operatorname{Stab}_{\mathfrak{n}^-}(v_\mu) = \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha$$

for $\Phi = \{\alpha \in \Delta^- \text{ s.t. } \mathfrak{g}_\alpha v_\mu = 0\}$, which we would like to describe in more detail. Recall that $\mu = w(\lambda)$; we see that $x_\alpha v_\mu = 0$ iff $x_{w(\alpha)} v_\lambda = 0$ iff $w(\alpha) \in \Delta^+$. I.e. $\Phi = \{\alpha \in \Delta^- \text{ s.t. } w(\alpha) \in \Delta^+\}$, and the cardinality of this set is $\ell(w)$. Thus $\dim(\mathfrak{n}^- / \operatorname{Stab}_{\mathfrak{n}^-} w_\mu) = \ell(w_0) - \ell(w)$, which is what was to be shown. $\square$

9.2.5.3 Remark The Shubert cells are ordered by closure. The *Bruhat order* on W is defined as the opposite order: $w \leq \sigma$ iff $U_\sigma \subseteq \overline{U_w}$. For example, for $\mathfrak{sl}(3)$ with simple reflections s_1, s_2, the order is:

$$\begin{tikzcd} & e \arrow[dl, dash] \arrow[dr, dash] & \\ s_2 \arrow[d, dash] \arrow[drr, dash] & & s_1 \arrow[d, dash] \arrow[dll, dash] \\ s_2 s_1 \arrow[dr, dash] & & s_1 s_2 \arrow[dl, dash] \\ & w_0 & \end{tikzcd}$$

This leads to some nice combinatorics that we will not go into. The closures $\overline{U_\sigma}$ are *Shubert varieties*, and are singular. $\lozenge$

9.2.5.4 Remark One can repeat everything we've done for arbitrary parabolic subalgebras. In particular, Shubert-like cell decompositions are known for all projective homogeneous spaces G/P in addition to the full flag variety G/B. $\lozenge$

9.3 Frobenius Reciprocity

9.3.1 Geometric induction

We are working in the algebraic category, but we could instead work in some analytic category (e.g. complex holomorphic functions), and everything works.

9.3.1.1 Definition *Let G be an algebraic group acting on X, and suppose that $L \to X$ is a vector bundle. It is a G-*vector bundle *if there is a G-action on the bundle that extends the action on X. I.e. for each $g \in G$, there should be a map of bundles $\{L \to X\} \xrightarrow{g} \{L \to X\}$ that's linear on fibers and restricts to the map $X \xrightarrow{g} X$.*

We will study the case $X = G/H$ where H is a closed subgroup. Then there is a standard procedure for how you can construct G-vector bundles:

9.3.1.2 Lemma / Definition *Let $\pi : H \to \mathrm{GL}(V)$ be a representation of H. Define $G \times_H V = G \times V/\sim$, where the equivalence relation is that $(gh, h^{-1}v) \sim (g, v)$ for each $h \in H$, $v \in V$, and $g \in G$. This gives a bundle $G \times_H V \to G/H$ by forgetting the second part, and the fiber is clearly identified with V. It is a G-vector bundle. This construction gives a functor:*

$$\mathcal{L} : \{\textit{representations of } H\} \to \{G\textit{-vector bundles on } G/H\}$$

The inverse functor takes the fiber over $x = eH \in G/H$; it is an H-module since $H = \mathrm{Stab}_G\, x$. This is an equivalence of categories. □

9.3.1.3 Lemma / Definition *Let V be a representation of H and $\mathcal{L}(V) = G\times_H V$ the corresponding bundle on G/H. The space of global sections $\Gamma(G/H, \mathcal{L}(V))$ is a (possibly infinite-dimensional) representation of G. It has a description as a space of functions:*

$$\Gamma_{G/H}(V) = \Gamma(G/H, \mathcal{L}(V)) = \{\phi : G \to V \textit{ s.t. } \phi(gh) = h^{-1}\phi(g)\, \forall h \in H, g \in G\}$$

This gives a functor $\mathrm{Ind} = \Gamma_{G/H} : H\text{-}\textsc{mod} \to G\text{-}\textsc{mod}$. *It is an embedding of categories, and is exact on the left.* □

9.3.1.4 Remark Since we are working in the algebraic setting, by Γ we mean the algebraic sections. If you want unitary representations of real-analytic groups, you can do a similar construction with L^2 sections, and the results are similar. In the smooth category, you can write down a similar construction, but the end result is very different. ♢

9.3.1.5 Lemma *Given a chain $G \supseteq K \supseteq H$, there is a canonical isomorphism of functors:*

$$\Gamma_{G/K} \circ \Gamma_{K/H} = \Gamma_{G/H}$$ □

We will study this induction functor. As opposed to the finite case, we do not have complete reducibility. For example, the Cartan is solvable. So it's important to have the correct statement.

9.3.1.6 Proposition *Let M be a G-module, and V an H-module. There is a canonical isomorphism:*

$$\mathrm{Hom}_G\big(M, \Gamma_{G/H}(V)\big) \cong \mathrm{Hom}_H(M, V)$$

So the induction functor is *right-adjoint* to the restriction functor.

Proof You write out the definitions.

$$\mathrm{Hom}_G\big(M, \Gamma_{G/H}(V)\big) = \big\{\phi : G \to \mathrm{Hom}_{\mathbb{C}}(M, V) \text{ s.t. } \phi(g^{-1}xh) = h^{-1}\,\phi(x)\,g\ \forall x, g \in G,\ h \in H\big\}$$

So pick $\phi \in \mathrm{Hom}_G(M, \Gamma_{G/H}(V))$, and send it to $\phi(e) \in \mathrm{Hom}_H(M, V)$. We claim this is a canonical homomorphism, because we can go back: if we have $\alpha \in \mathrm{Hom}_H(M, V)$, we can move it to $\phi_\alpha : x \mapsto \alpha x^{-1}$. (We leave for you to check if this should be x or x^{-1}. The point is that the value at any point is determined by the value at e.) □

9.3.1.7 Corollary *If V was an injective H-module, then $\Gamma_{G/H}(V)$ is an injective G-module.* □

9.3.1.8 Remark Let G be reductive (e.g. semisimple). Then $\mathbb{C}[G] = \bigoplus L(\lambda) \boxtimes L(\lambda)^*$ (Theorem 9.1.1.4). Recall that we have actions both on the left and on the right. Then:

$$\Gamma_{G/H}\big(\mathbb{C}[G]\big) = \bigoplus L(\lambda) \boxtimes \big(L(\lambda)^* \otimes \mathbb{C}[G]\big)^H$$

Thus the multiplicities are:

$$= \bigoplus L(\lambda) \boxtimes \mathrm{Hom}_H\big(L(\lambda), \mathbb{C}[G]\big) \qquad \lozenge$$

9.3.2 Induction for the universal enveloping algebra

9.3.2.1 Definition *Let $\mathfrak{g} \supseteq \mathfrak{h}$ be Lie algebras and V an $\mathfrak{h}$-module. We define:*

$$\mathrm{Ind}_{\mathfrak{h}}^{\mathfrak{g}} V = \mathcal{U}(\mathfrak{g}) \otimes_{\mathcal{U}(\mathfrak{h})} V$$

9.3.2.2 Remark In general, this is a very large $\mathfrak{g}$-module. Indeed, it is so large that the $\mathfrak{g}$ action does not integrate to a G action, even when G is simply-connected. When you move away form finite groups, you have the group algebra, and the function algebra, but also the universal enveloping algebra. $\lozenge$

9.3.2.3 Proposition *The induction functor* $\mathrm{Ind}_{\mathfrak{h}}^{\mathfrak{g}}$ *is* left-adjoint *to restriction:*

$$\mathrm{Hom}_{\mathfrak{g}}\big(\mathrm{Ind}_{\mathfrak{h}}^{\mathfrak{g}}(V), M\big) = \mathrm{Hom}_{\mathfrak{h}}(V, M) \qquad \square$$

9.3.2.4 Remark We can replace $\mathcal{U}\mathfrak{h} \subseteq \mathcal{U}\mathfrak{g}$ by any inclusion of associative rings. $\lozenge$

So the induction functor for $\mathcal{U}\mathfrak{h} \subseteq \mathcal{U}\mathfrak{g}$ is on the opposite side as the induction functor for $H \subseteq G$. We would like to have the ordering in the same direction as in Proposition 9.3.1.6. We can try:

9.3.2.5 Lemma / Definition *Let $\mathfrak{h} \subseteq \mathfrak{g}$ be Lie algebras. The* coinduction *functor is:*

$$\mathrm{Coind}(V) \overset{\mathrm{def}}{=} \mathrm{Hom}_{\mathcal{U}\mathfrak{h}}(\mathcal{U}\mathfrak{g}, V)$$

It is right-adjoint to restriction:

$$\mathrm{Hom}_{\mathfrak{g}}(M, \mathrm{Coind}\, V) = \mathrm{Hom}_{\mathfrak{h}}(M, V) \qquad \square$$

9.3.2.6 Remark This is humongous. As soon as you take the dual space to a countable-dimensional space, you get something uncountable. So we need to cut it down. $\lozenge$

9.3.2.7 Definition *Let M be any $\mathfrak{g}$-module. We define $Z(M) = \{m \in M \text{ s.t. } \dim(\mathcal{U}\mathfrak{g} \cdot m) < \infty\}$.*

This is closely related to the "Zuckerman functor".

9.3.2.8 Remark When G is reductive, the composition $Z \circ \mathrm{Coind}$ gets pretty close to the group induction. $\lozenge$

9.3.3 The derived functor of induction

Being an adjoint, $\Gamma_{G/H}$ is exact on the left, but it is not exact. So we will study the derived functor. For this, we need enough injective modules.

9.3.3.1 Proposition *If H is an algebraic group, then $\mathbb{C}[H]$ is an injective H-module.*

Proof In fact, $\mathrm{Hom}_H(V, \mathbb{C}[H]) \cong V^*$. How do you see this? Think about it for a moment in a different way. You have the functions in V^*. And you think about the LHS as the algebraic maps $\phi : H \to V^*$ such that $\phi(h^{-1}x) = h\phi(x)$. But such functions are completely determined by their values at e. So this is very similar to what we did previously: if I know $\phi(e)$ I know it everywhere. So we constructed a map LHS→RHS.

And now we need the inverse map, which is also very clear: it is just the Frobenius reciprocity induced from the trivial subgroup. If we have $\xi \in V^*$, we construct $\phi(g) = g\xi$.

Now, $V \mapsto V^*$ is clearly an exact functor, so then $\mathbb{C}[H]$ is injective. □

Finally, we want to show that any V can be mapped to an injective representation. Recall:

9.3.3.2 Lemma / Definition *A* representation *V of an algebraic group G is actually a* corepresentation *of $\mathbb{C}[G]$, i.e. a map $\rho : V \to \mathbb{C}[G] \otimes V$ such that the two natural maps $V \to \mathbb{C}[G] \otimes \mathbb{C}[G] \otimes V$ are the same:*

$$\begin{tikzcd} V \arrow[r, "\rho"] \arrow[d, "\Delta \otimes \mathrm{id}"'] & \mathbb{C}[G] \otimes V \arrow[d, "\mathrm{id} \otimes \rho"] \\ \mathbb{C}[G] \otimes \mathbb{C}[G] \otimes V & \mathbb{C}[G] \otimes \mathbb{C}[G] \otimes V \end{tikzcd}$$

Then ρ is an injection of G-modules, where G acts on $\mathbb{C}[G] \otimes V$ from the left on $\mathbb{C}[G]$ and trivially on V. □

9.3.3.3 Corollary *Every H-representation V has an injective resolution*

$$0 \to I_0 \to I_1 \to I_2 \to \dots,$$

meaning that $\mathrm{H}^0(I_\bullet) = V$ and $\mathrm{H}^{>0}(I_\bullet) = 0$.

Proof Set $I_0 = V \otimes \mathbb{C}[H]$ and $V \hookrightarrow I_0$ via ρ. Then rinse and repeat for $I_0/V \hookrightarrow I_1 = (I_0/V) \otimes \mathbb{C}[H]$, etc. □

9.3.3.4 Lemma / Definition *We set $\Gamma^i_{G/H}(V) \stackrel{\text{def}}{=} R^i\Gamma_{G/H}(V) \stackrel{\text{def}}{=} \mathrm{H}^i(\Gamma_{G/H}(I_\bullet))$, where $I_\bullet$ is any injective resolution of V — i.e. $\Gamma^i_{G/H}$ is the right-derived functor for $\Gamma_{G/H}$. It satisfies $\Gamma^i_{G/H}(V) = \mathrm{H}^i(G/H, \mathcal{L}(V))$.* □

9.3.3.5 Remark If H is reductive, then $\Gamma^i_{G/H}(V) = 0$ for $i > 0$, by complete reducibility. ♢

9.4 Center of universal enveloping algebra

9.4.1 Harish-Chandra's homomorphism

9.4.1.1 Theorem (Schur's lemma for countable-dimensional algebras) *Let R be a countable-dimensional associative algebra over $\mathbb{C}$, and M a simple R-module. Then $\mathrm{End}_R(M) = \mathbb{C}$.*

The corresponding statement is well known when $\dim R < \infty$.

Proof Any non-zero endomorphism is an isomorphism, because kernel and image are invariant subspaces. So let's pick up some endomorphism $x \neq 0$, and there are two cases: either x is algebraic over $\mathbb{C}$, or it's transcendental.

1. x is transcendental. Then $\mathbb{C}(x) \subseteq \mathrm{End}_R(M)$. But $\dim \mathbb{C}(x) = 2^{|\mathbb{N}|}$, because we can take $1/(x-a)$ for all $a \in \mathbb{C}$. On the other hand, $\mathrm{End}_R(M)$ is countable dimensional: if we pick $m \in M$, $m \neq 0$, and $\phi \in \mathrm{End}_R(M)$, then ϕ is determined by $\phi(m)$, because m generates M; similarly, $\dim M$ is countable because R is countable-dimensional, and $M = Rm$. So this was impossible.

2. x is algebraic over $\mathbb{C}$. Then $p(X) = (x-\lambda_1)\dots(x-\lambda_n) = 0$, so $x = \lambda_i$ for some i. □

9.4.1.2 Remark Note that when x is transcendental, the action of $\mathbb{C}(x)$ on itself gives a counterexample to Theorem 9.4.1.1 in dimension $2^{|\mathbb{N}|}$. ◊

Let $\mathfrak{g}$ be a finite-dimensional Lie algebra. We define $\mathcal{Z}(\mathfrak{g})$ to be the center of $\mathcal{U}\mathfrak{g}$. Note that as always, taking the center is not functorial. Let M be a simple representation of $\mathfrak{g}$. Then $\mathcal{Z}\mathfrak{g}$ acts on M as scalars by Theorem 9.4.1.1. Set $\operatorname{Spec}\mathcal{Z}(\mathfrak{g}) \overset{\text{def}}{=} \mathrm{Hom}_{\text{algebras}}(\mathcal{Z}(\mathfrak{g}), \mathbb{C})$; we have defined a map $\mathrm{Irr}\,\mathfrak{g} \to \operatorname{Spec} Z(\mathfrak{g})$ by sending $M \in \mathrm{Irr}\,\mathfrak{g}$ to the algebra homomorphism $\phi : \mathcal{Z}(\mathfrak{g}) \to \mathbb{C}$ for which $z|_M = \phi(z)\mathrm{id}$.

Recall Theorem 3.2.2.1: the canonical map $\mathcal{S}\mathfrak{g} \to \mathrm{gr}\,\mathcal{U}\mathfrak{g}$ is an isomorphism. Moreover, by repeating all the constructions of the tensor, symmetric, and universal enveloping algebras in the category of $\mathfrak{g}$-modules, we see that this canonical map is a morphism of $\mathfrak{g}$-modules. In characteristic 0, we can in fact construct a $\mathfrak{g}$-module isomorphism $\mathcal{S}\mathfrak{g} \xrightarrow{\sim} \mathcal{U}\mathfrak{g}$ by symmetrizing:

$$\begin{tikzcd} \mathcal{S}\mathfrak{g} \arrow[r] \arrow[dr, "\sim"'] & \mathcal{T}\mathfrak{g} \arrow[d] \\ & \mathcal{U}\mathfrak{g} \end{tikzcd}$$

All the arrows are morphisms of $\mathfrak{g}$-modules, and by Theorem 3.2.2.1 the diagonal is an isomorphism.

But $\mathcal{Z}(\mathfrak{g}) = (\mathcal{U}\mathfrak{g})^{\mathfrak{g}}$, where for a $\mathfrak{g}$-module M we write $M^{\mathfrak{g}} = \{m \in M \text{ s.t. } xm = 0\ \forall x \in \mathfrak{g}\}$ as the fixed points. We have thus exhibited a vector-space isomorphism $\mathcal{Z}(\mathfrak{g}) \cong (\mathcal{S}\mathfrak{g})^{\mathfrak{g}}$, and so to study $\mathcal{Z}(\mathfrak{g})$ we will begin by studying $(\mathcal{S}\mathfrak{g})^{\mathfrak{g}}$. Pick any Lie group G with $\mathrm{Lie}(G) = \mathfrak{g}$. Then it acts on $\mathfrak{g}$ via the adjoint action, and hence on $\mathcal{S}\mathfrak{g}$ and $\mathcal{U}\mathfrak{g}$, and we prefer to write the fixed points as fixed points of this group action.

Henceforth we suppose that $\mathfrak{g}$ is semisimple. The Killing form identifies $\mathfrak{g} \cong \mathfrak{g}^*$ as $\mathfrak{g}$-modules, so $(\mathcal{S}(\mathfrak{g}^*))^G \cong (\mathcal{S}\mathfrak{g})^G$. We remark that $(\mathcal{S}(\mathfrak{g}^*))^G$ is precisely the space of G-invariant polynomials on $\mathfrak{g}$. We choose a triangular decomposition $\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}^+$. Let $H \subseteq G$ be the subgroup with $\mathrm{Lie}(H) = \mathfrak{h}$. Let $r : \mathcal{S}(\mathfrak{g}^*) \to \mathcal{S}(\mathfrak{h}^*)$ denote the restriction map. We suppose moreover that we are working over an algebraically closed field $\mathbb{K}$ of characteristic 0.

9.4.1.3 Lemma *1. $\mathrm{Im}\, r \subseteq \mathcal{S}(\mathfrak{h}^*)^W$, where W is the corresponding Weyl group.*

2. r is injective.

Proof 1. We denote by $\mathcal{N}_G(\mathfrak{h})$ the normalizer of $\mathfrak{h}$ under $\mathrm{Ad} : G \curvearrowright \mathfrak{g}$. As is well-known, $W \cong \mathcal{N}(\mathfrak{h})/H$; H is commutative and so acts on $\mathfrak{h} = \mathrm{Lie}(H)$ trivially, and the action $W \curvearrowright \mathfrak{h}$ is precisely the action of $\mathrm{Ad} : \mathcal{N}(\mathfrak{h}) \curvearrowright \mathfrak{h}$. If a polynomial on $\mathfrak{g}$ is G-invariant, then in particular its restriction to $\mathfrak{h}$ is $\mathcal{N}(\mathfrak{h})$-invariant.

2. Any semisimple element is conjugate under the adjoint action to some element of $\mathfrak{h}$. Denote by $\mathfrak{g}_{\mathrm{ss}}$ the set of semisimple elements; it is dense in $\mathfrak{g}$. Let f be a G-invariant polynomial on $\mathfrak{g}$. If $f(\mathfrak{h}) = 0$, then $f(\mathfrak{g}_{\mathrm{ss}}) = 0$, so $f(\mathfrak{g}) = 0$. So $\ker r = 0$. □

9.4.1.4 Proposition *The map $r : \mathcal{S}(\mathfrak{g}^*)^G \to \mathcal{S}(\mathfrak{h}^*)^W$ is an isomorphism.*

Proof After Lemma 9.4.1.3, it suffices to show that r is surjective. We pick fundamental weights $\omega_1, \dots, \omega_n$ for $\mathfrak{g}$, and think of them as coordinate functions on $\mathfrak{h}$. Then $\mathcal{S}(\mathfrak{h}^*)$ has a basis $\omega_1^{a_1} \cdots \omega_n^{a_n}$, and $\mathcal{S}(\mathfrak{h}^*)^W$ is spanned by $\left\{\sum_{w\in W} w\big(\omega_1^{a_1} \cdots \omega_n^{a_n}\big)\right\}$. We claim that $\mathcal{S}(\mathfrak{h}^*)^W$ is generated as an algebra by $\left\{\sum_{w\in W} w(\omega_i^m) \text{ s.t. } m \in \mathbb{Z}_{\geq 0}\right\}$.

To prove this claim, we use the following *polarization formula*. Let $x_1, \dots, x_n$ be coordinate functions on $\mathbb{K}^n$, and $\mathbb{K}[x_1, \dots, x_n]$ the corresponding ring of polynomials. Then $\Gamma = (\mathbb{Z}/2)^{n-1}$ acts on $\mathbb{K}[x_1, \dots, x_n]$ by $p_i : x_j \mapsto (-1)^{\delta_{ij}} x_j$, where $p_1, \dots, p_{n-1}$ are the generators of Γ. There is a homomorphism $\mathrm{sign} : \Gamma \to (\mathbb{Z}/2) = \{\pm 1\}$ given by $p_i \mapsto -1$. Then:

$$\sum_{\gamma\in\Gamma} \mathrm{sign}(\gamma)\, \gamma(x_1 + \cdots + x_n)^n = 2^{n-1}\, n!\, x_1 \cdots x_n \tag{9.4.1.5}$$

Indeed, the left-hand-side is homogeneous of degree n, but the only monomials that can appear must be of odd degree in each variable $x_1, \dots, x_{n-1}$, by anti-symmetry under the Γ action. To prove the claim, we apply equation (9.4.1.5) to $x_i = \omega_i^{a_i}$, and thus obtain $\omega_1^{a_1} \dots \omega_n^{a_n}$.

Then to prove surjectivity, it suffices to show that for each $m \in \mathbb{Z}_{\geq 0}$ and ω_i a fundamental weight, $\sum_{w\in W} w(\omega_i^m) \in \mathrm{Im}\, r$. But if V is a finite-dimensional representation of $\mathfrak{g}$, $\mathrm{tr}_V(g^m) \in \mathcal{S}(\mathfrak{g}^*)^G$, since tr is ad-invariant. Let $V = L(\lambda)$ with $\lambda \in P^+$, then:

$$\mathrm{tr}_{L(\lambda)}(h^m) = \sum_{w\in W} w\big(\lambda(h)^m\big) + \sum_{\substack{\mu\in P^+\\ \mu<\lambda}} d_{\mu,\lambda}\, w\big(\mu(h)^m\big)$$

for some constants $d_{\mu,\lambda}$. Thus we can complete the proof by induction on P^+ to conclude that each $\sum_{w\in W} w\big(\lambda(h)^m\big)$ is in $\mathrm{Im}\, r$. □

9.4.1.6 Remark We proved $\mathcal{S}(\mathfrak{g}^*)^G \cong \mathcal{S}(\mathfrak{h}^*)^W$, but we are actually interested in $\mathcal{Z}(\mathfrak{g}) = \mathcal{U}\mathfrak{g}^G$. We proved that $\mathcal{Z}(\mathfrak{g}) \cong \mathcal{S}(\mathfrak{g}^*)^G$ as graded vector spaces, but not as rings. In fact, the arguments above show that $\operatorname{gr}\mathcal{Z}(\mathfrak{g}) \cong \mathcal{S}(\mathfrak{h}^*)$ as rings, but there are many commutative filtered rings that are not isomorphic to their associated graded rings. $\Diamond$

9.4.1.7 Definition *Let $\mathfrak{g}$ be a semisimple Lie algebra over $\mathbb{K}$ an algebraically closed field of characteristic 0. Pick a triangular decomposition $\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}^+$, and use PBW to write $\mathcal{U}\mathfrak{g} = \mathcal{U}\mathfrak{n} \otimes \mathcal{U}\mathfrak{h} \otimes \mathcal{U}\mathfrak{n}^+$, as vector spaces and in fact as $\mathfrak{h}$-modules. Then $\mathcal{U}\mathfrak{g}^{\mathfrak{h}} = \mathcal{U}\mathfrak{h} \oplus \big(\mathfrak{n}^-(\mathcal{U}\mathfrak{n}^-) \otimes \mathcal{U}\mathfrak{h} \otimes (\mathcal{U}\mathfrak{n}^+)\mathfrak{n}^-\big)$. The* Harish-Chandra homomorphism *$\theta : \mathcal{Z}(\mathfrak{g}) \to \mathcal{S}\mathfrak{h}$ is:*

$$\mathcal{Z}(\mathfrak{g}) = \mathcal{U}\mathfrak{g}^G \hookrightarrow \mathcal{U}\mathfrak{g}^{\mathfrak{h}} = \mathcal{U}\mathfrak{h} \oplus \big(\mathfrak{n}^-(\mathcal{U}\mathfrak{n}^-) \otimes \mathcal{U}\mathfrak{h} \otimes (\mathcal{U}\mathfrak{n}^+)\mathfrak{n}^-\big) \twoheadrightarrow \mathcal{U}\mathfrak{h} = \mathcal{S}\mathfrak{h}$$

The surjection simply forgets the second direct summand, and the final equality uses the commutativity of $\mathfrak{h}$.

So far θ is simply a homomorphism of filtered vector spaces.

9.4.1.8 Example Let $\mathfrak{g} = \mathfrak{sl}(2)$, given by x, h, y, with the Killing form $(h, h) = 8$, $(x, y) = 4$. Then $\Omega = \frac{1}{8}h^2 + \frac{1}{2}(xy + yx) = \frac{h^2}{8} + \frac{h}{4} + \frac{yx}{2}$. But when we apply the Harish-Chandra projection, we have: $\theta(\Omega) = \frac{h^2}{8} + \frac{h}{4} = \frac{1}{8}\big((h+1)^2 - 1\big)$. $\Diamond$

Recall that $\mathcal{S}\mathfrak{h} = \mathbb{K}[\mathfrak{h}^*]$, the algebra of polynomial functions on $\mathfrak{h}^*$, and that $\max\operatorname{Spec}\mathbb{C}[\mathfrak{h}^*] = \mathfrak{h}^*$. Let $\theta^* : \mathfrak{h}^* \to \operatorname{Spec}(\mathcal{Z}(\mathfrak{g}))$ be the dual map on Spec to the Harish-Chandra homomorphism θ. I.e. $\theta^*(\lambda) = \lambda \circ \theta \in \operatorname{Hom}(\mathcal{Z}(\mathfrak{g}), \mathbb{C})$. In particular, $\theta^*(\lambda)$ is some sort of character on $\mathfrak{g}$, and so we change notation slightly writing $\chi_\lambda \stackrel{\text{def}}{=} \theta^*(\lambda)$.

9.4.1.9 Lemma *If $z \in \mathcal{Z}(\mathfrak{g})$, then $z|_{M(\lambda)} = \chi_\lambda(z)\,\mathrm{id}$, where $M(\lambda)$ is the Verma module with highest weight λ.*

Proof We have $z = \theta(z) + yx$, where $y \in \mathcal{U}\mathfrak{g}$ and $x \in \mathfrak{n}^+$, but if v is the highest vector of $M(\lambda)$, then $\mathfrak{n}^+ v = 0$, so $zv = \theta(z)\,v$. Let $\theta(z) = p(\lambda) \in \mathbb{C}[\mathfrak{h}^*]$. The claim follows from the fact that $hv = \lambda(h)\,v$ for all $h \in \mathbb{H}$. $\square$

9.4.1.10 Corollary *θ is a homomorphism of rings.* $\square$

We will now describe the image of θ.

9.4.1.11 Definition *The* shifted action *of W on $\mathfrak{h}^*$ is $\lambda^w \stackrel{\text{def}}{=} w(\lambda + \rho) - \rho$.*

9.4.1.12 Lemma *Define $S_{\alpha_i} \stackrel{\text{def}}{=} \{\lambda \in \mathfrak{h}^* \text{ s.t. } \lambda(h_i) \in \mathbb{Z}_{\geq 0}\}$. If $\lambda \in S_{\alpha_i}$ and $\mu = \lambda^{s_i}$, then:*

$$\operatorname{Hom}_{\mathfrak{g}}(M(\mu), M(\lambda)) \neq 0.$$

Proof We have $\lambda(h_i) = k_i$ and $(\lambda + \rho)(h_i) = k_i + 1$. Let $v' = y_i^{k_i+1} v$, where v is the highest weight vector of $M(\lambda)$. Then $\mathfrak{n}^+ v' = 0$, and $hv' = \mu(h)\,v'$. It follows that $0 \neq \operatorname{Hom}_{\mathfrak{b}}(C_\mu, M(\lambda)) \cong \operatorname{Hom}_{\mathcal{U}\mathfrak{g}}(\mathcal{U}\mathfrak{g} \otimes_{\mathcal{U}\mathfrak{b}} C_\mu, M(\lambda))$, but $\mathcal{U}\mathfrak{g} \otimes_{\mathcal{U}\mathfrak{b}} C_\mu = M(\mu)$. $\square$

9.4.1.13 Proposition *We have $\theta(\mathcal{Z}(\mathfrak{g})) \subseteq \mathbb{C}[\mathfrak{h}^*]^{W_{\mathrm{sh}}}$, by which we mean the fixed points of the shifted action.*

Proof Pick up a simple root α_i, and let $s_i \in W$ be its reflection. Suppose that $\lambda \in S_{\alpha_i}$ as defined in Lemma 9.4.1.12. Then $\chi_\lambda = \chi_{\lambda^{s_i}}$, as by Lemma 9.4.1.9 each acts centrally on the corresponding Verma module and there is a nontrivial homomorphism. So if $f \in \operatorname{Im}\theta$, then $f(\lambda) = f(\lambda^{s_i})$ for all $\lambda \in S_{\alpha_i}$. But the set S_{α_i} is the union of countably many hyperplanes, so is Zariski dense in $\mathfrak{h}^*$; therefore $f(\lambda) = f(\lambda^{s_i})$ for any $\lambda \in \mathfrak{h}^*$. This checks it on the simple reflections, and so W-invariance follows. $\square$

9.4.1.14 Theorem (The Harish-Chandra isomorphism)
The Harish-Chandra map $\theta : \mathcal{Z}(\mathfrak{g}) \to \mathbb{K}[\mathfrak{h}^*]^{W_{\mathrm{sh}}}$ is an isomorphism.

Proof The proof goes by going from filtered rings to graded rings. Recall that if A is a filtered $\mathbb{K}$-algebra — $A = \bigcup_{i=0}^\infty A_i$ with $A_iA_j \subseteq A_{i+j}$ and each A_i is a vector space — then we define the *associated graded ring* $\operatorname{gr} A = \bigoplus_i (A_i/A_{i-1})$. Moreover, suppose that we have two filtered rings A, B and a homomorphism $\theta : A \to B$ that preserves filtrations. Then we can define the map $\operatorname{gr}\theta : \operatorname{gr} A \to \operatorname{gr} B$. Recall the following two fundamental facts:

1. $\operatorname{gr}\theta$ is a homomorphism; i.e. gr is a functor from filtered algebra to graded algebras.

2. If $\operatorname{gr}\theta$ is an isomorphism, then so was θ, at least when all the graded components are finite-dimensional.

Neither is hard to check. But by Proposition 9.4.1.4, $r : (\mathcal{S}\mathfrak{g}^*)^G \to (\mathcal{S}\mathfrak{h}^*)^W$ is an isomorphism, and with the Killing form we have $(\mathcal{S}\mathfrak{g}^*)^G \cong (\mathcal{S}\mathfrak{g})^G$ and $(\mathcal{S}\mathfrak{h}^*)^W \cong (\mathcal{S}\mathfrak{h})^W$. But $\operatorname{gr}(\mathcal{U}\mathfrak{g})^G = (\mathcal{S}\mathfrak{g})^G$ and $\operatorname{gr}(\mathcal{S}\mathfrak{h})^{W_{\mathrm{sh}}} = (\mathcal{S}\mathfrak{h})^W$, and $\operatorname{gr}\theta = r$.

$$\begin{tikzcd}
(\mathcal{U}\mathfrak{g})^G \arrow[r, dashed, "\mathrm{gr}"] \arrow[d, "\theta"'] & (\mathcal{S}\mathfrak{g})^G \arrow[r, leftrightarrow, "\sim", "(,)"'] & (\mathcal{S}\mathfrak{g}^*)^G \arrow[d, "r"', "\sim"] \\
(\mathcal{S}\mathfrak{h})^{W_{\mathrm{sh}}} \arrow[r, dashed, "\mathrm{gr}"] & (\mathcal{S}\mathfrak{h})^W \arrow[r, leftrightarrow, "\sim", "(,)"'] & (\mathcal{S}\mathfrak{h}^*)^W
\end{tikzcd}$$

$\square$

9.4.1.15 Remark The converse to statement 2. above is not true, in the following sense: You can have a filtered homomorphism $A \to B$ that is an isomorphism of algebras but not an isomorphism of filtered algebras, and it generally will not induce an isomorphism $\operatorname{gr} A \to \operatorname{gr} B$. $\Diamond$

9.4.1.16 Remark A more general statement, the *Duflo theorem*, asserts that there is an isomorphism $(\mathcal{S}\mathfrak{g})^G \cong \mathcal{Z}(\mathfrak{g})$ of rings. See [BNLT03] for a proof using only that $\mathfrak{g}$ has an invariant form $(,)$. $\Diamond$

9.4.2 Exponents of a semisimple Lie algebra

We can now start to study $\mathcal{Z}(\mathfrak{g})$ when $\mathfrak{g}$ is semisimple in earnest. We recall the following fact from the theory of geometric invariants; c.f. [Ser09, Spr77]:

9.4.2.1 Proposition *Suppose that the action of a finite group W on some vector space $\mathfrak{h}$ is generated by reflections. Then $\mathcal{S}(\mathfrak{h})^W$ is isomorphic to a polynomial ring $\mathbb{K}[f_1, \dots, f_n]$, where each f_i is homogeneous within the grading on $\mathcal{S}\mathfrak{h}$, and $n = \dim \mathfrak{h}$.* $\square$

9.4.2.2 Corollary *$\mathcal{Z}(\mathfrak{g})$ is isomorphic to a polynomial ring of n variables, where $n = \operatorname{rank} \mathfrak{g}$.* $\square$

9.4.2.3 Lemma / Definition *Let $\mathfrak{g}$ be a semisimple Lie algebra. The degrees $m_1, \dots, m_n$ of homogeneous generators $f_1, \dots, f_n$ of $\mathcal{Z}(\mathfrak{g}) \cong (\mathcal{S}\mathfrak{g})^{W_{\text{sh}}}$ are the* exponents *of $\mathfrak{g}$. They satisfy:*

$$m_1 \cdot \dots \cdot m_n = |W| \tag{9.4.2.4}$$

$$m_1 + \dots + m_n = \frac{1}{2}(\dim \mathfrak{g} + n) \tag{9.4.2.5}$$

We will always order the exponents by increasing degree: $m_1 \leq m_2 \leq \dots \leq m_n$.

The isomorphism $\mathcal{Z}(\mathfrak{g}) \xrightarrow{\sim} (\mathcal{S}\mathfrak{h})^{W_{\text{sh}}}$ depended on choosing a triangular decomposition for $\mathfrak{g}$ — indeed, even defining $\mathfrak{h}$ and W_{sh} require such a choice. But any two triangular decompositions are conjugate, so the set of exponents is well-defined.

Proof In $\mathcal{S}(\mathfrak{h})^W$ we have:

$$R(t) = \sum_{k=0}^{\infty} \dim \mathcal{S}^k(\mathfrak{h})^W t^k = \prod_{i=1}^{n} \frac{1}{1 - t^{m_1}} \tag{9.4.2.6}$$

If V is a linear representation of W, then:

$$\dim V^W = \frac{1}{|W|} \sum_{w \in W} \operatorname{tr}_V w \tag{9.4.2.7}$$

It more or less follows that:

$$R(t) = \frac{1}{|W|} \sum_{w \in W} \frac{1}{\det(1 - wt)} \tag{9.4.2.8}$$

And comparing equations (9.4.2.6) and (9.4.2.8) gives:

$$\prod_{i=1}^{n} \frac{1}{1 - t^{m_1}} = \frac{1}{m_1 \dots m_n (1 - t^n)} + \frac{\sum (m_i - 1)}{2 m_1 \dots m_n (1 - t)^{n-1}} + \dots$$

whereas

$$\frac{1}{|W|} \sum_{w \in W} \frac{1}{\det(1 - wt)} = \frac{1}{|W|(1 - t)^n} + \frac{1}{|W|} \sum_{\text{reflections}} \frac{1}{2(1 - t)^{n-1}}$$

Thus, $\sum (m_i - 1) =$ number of reflections. $\square$

9.4.2.9 Example In G_2 we have $m_1 m_2 = 12$ and $m_1 + m_2 = \frac{1}{2}(14+2) = 8$. But we always have $m_1 = 2$, because we always have a Casimir element. So $m_2 = 6$. $\diamondsuit$

9.4.2.10 Example In $\mathfrak{sl}(n) = A_{n-1}$ we have $m_1 \cdots m_{n-1} = n!$ and $m_1 + \cdots + m_{n-1} = \frac{n^2-1+n-1}{2} = \frac{n^2+n}{2} - 1$. You can solve this. One solution is $m_1 = 2$, $m_2 = 3$, $\ldots$, $m_{n-1} = n$. One possible set of generators of $\mathcal{S}(\mathfrak{g}^*)^G$ is the traces $\operatorname{tr} x^2, \operatorname{tr} x^3, \ldots, \operatorname{tr} x^n$.

Of course the set of generators is not unique. A second option is to take the characteristic polynomial $\det(X - t\mathrm{id})$ and take the coefficients. $\diamondsuit$

9.4.2.11 Example The groups B_n and C_n must have the same exponents, because the Weyl groups are the same. We have $B_n = \mathrm{O}(2n+1)$, and $\dim \mathfrak{g} = \frac{(2n+1)2n}{2} = (2n+1)n$. What is the order of the Weyl group? Well, $W = S_n \rtimes \mathbb{Z}_2^n$. So:

$$m_1 \cdots m_n = 2^n \, n! \tag{9.4.2.12}$$

$$m_1 + \cdots + m_n = (n+1)n \tag{9.4.2.13}$$

So the obvious solution is $2, 4, \ldots, 2n$.

So what are they? We can still take the generators from $\mathfrak{sl}_n$. But some of those vanish: the traces of odd powers of skew-symmetric matrices are 0. So we have $\operatorname{tr} x^2, \ldots, \operatorname{tr} x^{2m}$. $\diamondsuit$

9.4.2.14 Example Finally, let's look at $D_n = \mathfrak{o}(2n) = \{\text{skew-symmetric matrices of size } 2n\}$. Then $\dim \mathfrak{g} = n(2n-1)$, and $m_1 + \cdots + m_n = \frac{1}{2}(\dim \mathfrak{g} + n) = n^2$. Also, $m_1 \cdots m_n = |W| = 2^{n-1} n!$. So when you start looking for the most reasonable solution, it is $2, 4, \ldots, 2(n-1)$ and one more: n, somewhere in the middle.

So, think of x as a skew-symmetric matrix. Then $f_k(x) = \operatorname{tr}(x^{2k})$ are invariant as before. But if we take $2k = 2n$, then we might as well take the determinant, but the n is actually the *Pfaffian* of x, which is a polynomial $\operatorname{Pf}(x) = \sqrt{\det x}$.

Why is $\operatorname{Pf}(x)$ a polynomial? Think of x as a matrix of some skew-symmetric form on $\mathbb{C}^{2n}$. Then by linear algebra, if x is nondegenerate, then in some basis it has a canonical form $\left(\begin{smallmatrix} 0 & 1_n \\ -1_n & 0 \end{smallmatrix}\right)$. So in general, we have $x = y^T \left(\begin{smallmatrix} 0 & 1_n \\ -1_n & 0 \end{smallmatrix}\right) y$, and so $\det x = (\det y)^2$. $\diamondsuit$

We now explain how to calculate the exponents of any semisimple Lie algbera.

9.4.2.15 Lemma / Definition *Let $\mathfrak{g}$ be a semisimple Lie algebra with its standard generators $x_1, \ldots, x_n$, $h_1, \ldots, h_n$, $y_1, \ldots, y_n$. Let $x = x_1 + \cdots + x_n$. There exists a unique $h \in \mathfrak{h}$ such that the $\alpha_i(h) = 2$ for all simple roots $\alpha_1, \ldots, \alpha_n$. Write $h = c_1 h_1 + \cdots + c_n h_n$, and let $y = c_1 y_1 + \cdots + c_n y_n$. Then $\{x, h, y\}$ is an $\mathfrak{sl}_2$ triple, the* principal $\mathfrak{sl}(2)$ *in $\mathfrak{g}$.*

In the adjoint action of the principal $\mathfrak{sl}(2)$ on $\mathfrak{g}$, $\alpha(h)$ is even, and so every irreducible $\mathfrak{sl}(2)$-representation appearing in $\mathfrak{g}$ has a one-dimensional 0-weight space. Since h is regular, $\mathfrak{g}^h = \mathfrak{h}$. Therefore the number of $\mathfrak{sl}(2)$-irreducible components is exactly the rank of $\mathfrak{g}$. $\square$

9.4.2.16 Example In $\mathfrak{sl}_3$, we have $x = \begin{pmatrix} 0 & 1 & \\ & 0 & 1 \\ & & 0 \end{pmatrix}$, $h = \begin{pmatrix} 2 & & \\ & 0 & \\ & & -2 \end{pmatrix}$, and $y = \begin{pmatrix} 0 & & \\ 2 & 0 & \\ 0 & 2 & 0 \end{pmatrix}$. So we see that $\mathfrak{sl}_3 = \mathfrak{sl}_2 \oplus V_4$. $\diamondsuit$

Decompose $\mathfrak{g}$ into irreducible representations of its principal $\mathfrak{sl}(2)$: $\mathfrak{g} = V_{p_1} \oplus \cdots \oplus V_{p_n}$ with $p_1 \leq \cdots \leq p_n$, where by definition $\dim V_p = p+1$. Then $\sum_{i=1}^n (p_i+1) = \dim \mathfrak{g}$, and so $\sum_{i=1}^n \bigl(\frac{p_i}{2}+1\bigr) = \frac{1}{2}(\dim \mathfrak{g} + n)$. So the numbers $\frac{p_i}{2} + 1$ satisfy the same relation as the exponents. In fact:

9.4.2.17 Proposition $m_i = \dfrac{p_i}{2} + 1$.

Therefore to compute the exponents of a semisimple Lie algebra $\mathfrak{g}$ you need only to decompose $\mathrm{ad} : \mathfrak{sl}(2) \curvearrowright \mathfrak{g}$ into irreducibles. This is a little bit of work, but you know all the weights, so you know how to do it.

Proof Let $v_1, \dots, v_n$ be the lowest weight vectors in the components $V_{p_1}, \dots, V_{p_n}$. These are precisely the vectors that are killed by $y \in \mathfrak{sl}(2)$, and so $\mathfrak{g}^y = \mathbb{K}v_1 \oplus \cdots \oplus \mathbb{K}v_n$. We consider the slightly bigger space $M = \mathbb{K}x \oplus \mathfrak{g}^y$. This is a linear subspace of $\mathfrak{g}$ with dimension $n+1$. Moreover, M comes with specified coordinates $t_0, \dots, t_n : M \to \mathbb{K}$, so that any vector in M is of the form $t_0 x + t_1 v_1 + \cdots + t_n v_n$. So let $\phi : M \hookrightarrow \mathfrak{g}$ be the injection, and we are going to study the map $\phi^* : \mathbb{K}[\mathfrak{g}]^G \hookrightarrow \mathbb{K}[\mathfrak{g}] \to \mathbb{K}[M] = \mathbb{K}[t_0, t_1, \dots, t_n]$, where by $\mathbb{K}[V]$, of course, we mean the algebra $\mathcal{S}V^*$ of polynomial functions on V.

We claim that ϕ^* is injective. Consider the map $\gamma : G \times M \to \mathfrak{g}$ given by first embedding and then acting: $g \times (t_0, \dots, t_n) \mapsto (\mathrm{Ad}\, g)(t_0 x + t_1 v_1 + \cdots + t_n v_n)$. We first compute the image of $\mathrm{d}\gamma|_{e,1,0,\dots,0} : \mathfrak{g} \oplus M \to \mathfrak{g}$, where e is the identity element in G:

$$\mathrm{Im}\, \mathrm{d}\gamma|_{e,1,0,\dots,0} = \mathfrak{g}^y \oplus \mathrm{ad}_{\mathfrak{g}}(x) = \mathfrak{g}^y \oplus [x, \mathfrak{g}] = \mathfrak{g}.$$

Therefore γ is locally a surjective map. If we are working over $\mathbb{K} = \mathbb{C}$, then it follows that $\mathrm{Im}\,\gamma$ contains a topologically open subset of $\mathfrak{g}$, and hence is Zariski-dense in $\mathfrak{g}$. If we are working over some other field $\mathbb{K}$, we can use the fact that if an algebraic map between affine spaces has surjective derivative, then the image is Zariski dense; see [FH91]. Either way, $\mathrm{Im}\,\gamma = G \cdot M$ is Zariski dense in $\mathfrak{g}$, and so if $f \in \mathcal{S}(\mathfrak{g}^*)^G$, then $f|_M = 0$ implies $f|_{G \cdot M} = 0$ so $f = 0$. This proves that $\phi^* : \mathbb{K}[\mathfrak{g}]^G \to \mathbb{K}[M]$ is injective.

Actually, the above argument proves a bit more. Consider the map $\gamma' : G \times \mathfrak{g}^y \to \mathfrak{g}$ given by $\gamma'(g, v) = \gamma(g, x + v)$. Then $\mathrm{d}\gamma'|_{e,0} : \mathfrak{g} \oplus \mathfrak{g}^y \to \mathfrak{g}$ is still a surjection, and so the composition $\mathbb{K}[\mathfrak{g}]^G \xrightarrow{\phi^*} \mathbb{K}[M] \xrightarrow{\mathrm{ev}\, t_0 = 1} \mathbb{K}[\mathfrak{g}^y] = \mathbb{K}[t_1, \dots, t_n]$ is an injection.

Let $f_1, \dots, f_n$ generate $\mathbb{K}[\mathfrak{g}]^G = \mathbb{K}[\mathfrak{h}]^W$; they are algebraically independent by Proposition 9.4.2.1. By injectivity, $\phi^*(f_1)(1, t_1, \dots, t_n), \dots, \phi^*(f_n)(1, t_1, \dots, t_n) \in \mathbb{K}[t_1, \dots, t_n]$ are also algebraically independent. It follows that we can set up a bijection between $\{t_1, \dots, t_n\}$ and $\{f_1, \dots, f_n\}$ so that each t_i appears with non-zero exponent in a non-zero monomial in the corresponding $\phi^*(f_j)$.

On the other hand, ad_h acts as some diagonal matrix on M, because x and all v_is are eigenvectors. As a vector field, the action is:

$$\mathrm{ad}_h = 2t_0 \frac{\partial}{\partial t_0} - \sum_{i=1}^n p_i t_i \frac{\partial}{\partial t_i} \tag{9.4.2.18}$$

where $[h, v_i] = -p_i v_i$ because v_i is the lowest vector of V_{p_i}. So:

$$\mathrm{ad}_h\bigl(\phi^*(f)\bigr) = 0 \quad \forall f \in \mathcal{S}(\mathfrak{g}^*)^G \tag{9.4.2.19}$$

So for each t_i, take the corresponding $\phi^*(f_j)$, and suppose that the monomial in which t_i appears is $c\, t_0^{d_0} \cdots t_n^{d_n}$. By equation (9.4.2.18), all monomials are eigenvectors for ad_h, and by equation (9.4.2.19), $2d_0 = \sum_{k=1}^n p_k d_k$. However, ϕ^* is degree non-increasing: $\deg f_j \geq \sum_{k=0}^n d_k \geq d_0 + 1 \geq \frac{p_i}{2} + 1$.

By assumption, we ordered our generators so that $\deg f_1 \leq \cdots \leq \deg f_n$, and $p_1 \leq \cdots \leq p_n$. Thus for each i we have $\deg f_i \geq \frac{p_i}{2} + 1$. However, $\sum \deg f_i = \frac{1}{2}(\dim \mathfrak{g} + \mathfrak{n}) = \sum \bigl(\frac{p_i}{2} + 1\bigr)$, and so we must have equality $\deg f_i = p_i$. $\square$

9.4.2.20 Corollary *We can choose $f_1, \dots, f_n$ so that $\phi^*(f_i) = t_0^{p_i/2} t_i + \mathit{poly}(t_0, t_1, \dots, t_{i-1})$.* $\square$

9.4.2.21 Corollary *The differentials $\mathrm{d}f_1, \dots, \mathrm{d}f_n$ are linearly independent at $x = (1, 0, \dots, 0)$.* $\square$

9.4.3 The nilpotent cone

9.4.3.1 Lemma / Definition *The* nilpotent cone *in a Lie algebra $\mathfrak{g}$ is*

$$\mathscr{N} \stackrel{\text{def}}{=} \{z \in \mathfrak{g} \textit{ s.t. } \mathrm{ad}(z) \textit{ is nilpotent}\}.$$

It is closed under scalar multiplication (hence the name "cone"). If $\mathfrak{g}$ is semisimple, then $z \in \mathscr{N}$ if and only if z acts nilpotently on every finite-dimensional representation. Then the traces of all powers are 0: for any $\pi : \mathfrak{g} \to \mathfrak{gl}(V)$ we have $\mathrm{tr}_V\, \pi(z)^m = 0$ for all m. $\square$

9.4.3.2 Definition *The* centralizer *of $z \in \mathfrak{g}$ is $\mathcal{C}_{\mathfrak{g}}(z) = \ker \mathrm{ad}(z) = \mathrm{Stab}_{\mathrm{ad}}(z)$. Recall from Lemma/Definition 5.3.1.15 that an element $z \in \mathfrak{g}$ is* regular *if its adjoint orbit has maximal dimension, or equivalently if its centralizer has minimal dimension. When $\mathfrak{g}$ is semisimple, this dimension is not less than the rank n. We define $\mathscr{N}_{\mathrm{reg}} = \{z \in \mathscr{N} \textit{ s.t. } \dim \mathcal{C}_{\mathfrak{g}}(z) = n\}$.*

9.4.3.3 Lemma *Let $\mathfrak{g}$ be semisimple. If $z \in \mathscr{N}$, then $G \cdot z$, the orbit under the adjoint action, intersects $\mathfrak{n}^+$ nontrivially.*

For $\mathfrak{g} = \mathfrak{sl}_n$, this is more or less trivial, by Jordan form: for any nilpotent z, there is a flag so that z moves along the flag. We essentially reproduce that argument in the general setting:

Proof We claim that there is $u \in \mathfrak{g}$ such that $[u, z] = z$. For this, we need to show that $z \in \mathrm{Im}\, \mathrm{ad}(z)$. But we have the Killing form, so $\mathrm{Im}\, \mathrm{ad}(z) = \bigl(\ker \mathrm{ad}(z)\bigr)^\perp$. Let $\mathfrak{g}_k = \mathrm{Im}\, \mathrm{ad}(z)^k$. Pick $a \in \ker \mathrm{ad}(z)$; then $\mathrm{ad}(a)$ commutes with $\mathrm{ad}(z)$, so $[a, \mathfrak{g}_k] \subseteq \mathfrak{g}_k$. So we have:

$$\mathrm{ad}(z) = \left(\begin{array}{c|c|c} 0 & * & * \\ \hline 0 & 0 & * \\ \hline 0 & 0 & 0 \end{array}\right) \qquad \mathrm{ad}(a) = \left(\begin{array}{c|c|c} * & * & * \\ \hline 0 & * & * \\ \hline 0 & 0 & * \end{array}\right)$$

But then $\mathrm{tr}\bigl(\mathrm{ad}(a)\, \mathrm{ad}(z)\bigr) = 0$, which is to say $a \perp z$, and so $z \in \bigl(\ker \mathrm{ad}(z)\bigr)^\perp = \mathrm{Im}\, \mathrm{ad}(z)$.

So pick $u \in \mathfrak{g}$ with $[u, z] = z$. Recall Lemma/Definition 5.3.2.2: we can write $u = u_s + u_n$ with u_s semisimple, u_n nilpotent, and $u_s, u_n \in \mathbb{C}[u]$. Then in particular $\mathrm{ad}(u_n)$ is some polynomial in $\mathrm{ad}(u)$, and so acts on z by some scalar, but the only nilpotent scalar is 0. So we can pick $u = u_s$ to be semisimple.

But then $u \in \mathfrak{h}'$ for some Cartan subalgebra $\mathfrak{h}'$. So there is $g \in G$ with $\mathrm{Ad}(g)u \in \mathfrak{h}$. But $\mathrm{Ad}(g)$ is an automorphism of $\mathfrak{g}$, and so $[\mathrm{Ad}(g)u, \mathrm{Ad}(g)z] = \mathrm{Ad}(g)z$, so we can pick g with $\mathrm{Ad}(g)z \in \mathfrak{n}^+$, by picking a triangular decomposition so that u is in the positive part of $\mathfrak{h}'$. □

9.4.3.4 Corollary *All G-invariant functions are constant on $\mathscr{N}$.*

Proof Pick $z \in \mathscr{N}$ and $u \in \mathfrak{g}$ such that $[u, z] = z$. Then $\mathrm{Ad}(tu)z = e^t z$, and in particular $0 \in \overline{\{\mathrm{Ad}(tu)z\}_{t\in\mathbb{K}}} \subseteq \overline{G \cdot z}$. So a G-invariant function takes the same value at z as at 0. □

9.4.3.5 Proposition *Let $\mathfrak{g}$ be a semisimple Lie algebra of rank n, $\mathscr{N}$ its nilpotent cone, and I the ideal in $\mathbb{C}[\mathfrak{g}]$ generated by the functions $f_1, \dots, f_n$ of Lemma/Definition 9.4.2.3. Then:*

1. *$\mathscr{N}$ is irreducible, with vanishing ideal $I(\mathscr{N}) = I$.*
2. *$\mathscr{N}_{\mathrm{reg}} = G \cdot x$, where $\{x, h, y\}$ is the principal $\mathfrak{sl}_2$ from Lemma/Definition 9.4.2.15.*
3. *$\dim \mathscr{N} = \dim \mathfrak{g} - n$.*

Proof Lemma 9.4.3.3 gives that $\mathscr{N} = G \cdot \mathfrak{n}^+$. Let $B = \mathcal{N}_G(\mathfrak{n}^+)$ be the normalizer of $\mathfrak{n}^+$ in G; its Lie algebra is $\mathrm{Lie}(B) = \mathfrak{b}^+ = \mathfrak{h} \oplus \mathfrak{n}^+$. Then $\dim(B \cdot x) = \dim B - n = \dim \mathfrak{n}^+$. Therefore, $\overline{B \cdot x} = \mathfrak{n}^+$, and so $\overline{G \cdot x} = \mathscr{N}$. But G is a connected group, and so $\overline{G \cdot x}$ is irreducible.

By Corollary 9.4.3.4, $f_1, \dots, f_n \in I(\mathscr{N})$, and so we only need to show that $\sqrt{I} = I$. But this follows from Corollary 9.4.2.21: if $\sqrt{I} \neq I$, then some $\mathrm{d}f_i$ would have to depend on the others at some point. □

9.4.3.6 Example When $\mathfrak{g} = \mathfrak{sl}_n$, then $\mathscr{N}$ consists of the nilpotent $n \times n$ matrices. The $\mathrm{SL}(n)$-orbit of a matrix is determined by its Jordan form. So the $\mathrm{SL}(n)$-orbits are in one-to-one correspondence with unordered partitions of n. ◊

9.4.3.7 Theorem ($\mathcal{U}\mathfrak{g}$ is free over its center)
If $\mathfrak{g}$ is semisimple, then $\mathcal{U}\mathfrak{g}$ is free as a module over its center $\mathcal{Z}(g)$.

Proof We will prove this in three steps. Our strategy will be to show in Steps 1 and 2 that $\mathcal{S}(\mathfrak{g}^*)$ is free as an $\mathcal{S}(\mathfrak{g}^*)^G$-module. Then in Step 3 we will conclude the result via the filtered-graded yoga.

Step 1 Pick $q_1, \dots, q_m \in \mathcal{S}(\mathfrak{g}^*)$ linearly independent on $G \cdot x$. We claim that there exists Zariski-open $U \subseteq \mathfrak{g}$, $U \ni x$ such that $q_1, \dots, q_m$ are linearly independent on $G \cdot z$ for any $z \in U$.

This fact uses just a little algebraic geometry. In fact, for every $z \in \mathfrak{g}$, define the obvious map $\phi_z : G \to G \cdot z$, $g \mapsto \mathrm{Ad}_g(z)$. Then there is a dual map. The linear independence is equivalent to the statement that $\mathrm{rank}\, \phi_x^*(q_1, \dots, q_m) = m$. But $\mathrm{rank}\, \phi_z^*(q_1, \dots, q_m) = m$ is a Zariski-open condition in z, as it is the statement that certain minors are non-zero.

Step 2 Let I be as in Proposition 9.4.3.5. It is a graded vector subspace of $\mathcal{S}(\mathfrak{g}^*)$, and we pick a splitting $\mathcal{S}(\mathfrak{g}^*) = I \oplus Y$ of graded vector spaces. We will prove that the multiplication map $\mu : \mathcal{S}(\mathfrak{g}^*)^G \otimes Y \to \mathcal{S}(\mathfrak{g}^*)$ is an isomoprhism.

We first prove that μ is surjective. Let $p : \mathcal{S}(\mathfrak{g}^*) \to Y$ be the projection induced by the splitting. Then for $q \in \mathcal{S}^k(\mathfrak{g}^*)$, we have $q \cdot p(q) = f_1 q_1 + \cdots + f_n q_n$ with $\deg q_i < k$. So we can proceed by induction.

To prove injectivity, we argue as follows: Any element of $\mathcal{S}(\mathfrak{g}^*)^G \otimes Y$ is of the form $\sum_{i=1}^m s_i \otimes q_i$, with $s_i \in \mathcal{S}(\mathfrak{g}^*)^G$ and $q_i \in Y$, and we can chose it so that $q_1, \dots, q_m$ are linearly independent. Then suppose that $\mu\big(\sum_{i=1}^m s_i \otimes q_i\big) = \sum s_i q_i = 0$. Since $q_1, \dots, q_m$ are linearly independent on $G \cdot x$, they are linearly independent on $G \cdot z$ for $z \in U$, by Step 1. On the other hand, the s_i are constant on any orbit, because they are invariant. So then $s_i|_{G\cdot z} = 0$ because the q_i are linearly independent. But then $s_i = 0$, because U is Zariski-open and hence dense.

Step 3 Let $\sigma : \mathcal{S}(\mathfrak{g}^*) \xrightarrow{\sim} \mathcal{S}\mathfrak{g} \hookrightarrow \mathcal{T}\mathfrak{g} \twoheadrightarrow \mathcal{U}\mathfrak{g}$, where the first map is the Killing form, the second is by symmetrization, and the last is defining. This is a homomorphism of $\mathfrak{g}$-modules. Consider the multiplication map $\tilde{\mu} : \mathcal{Z}(\mathfrak{g}) \otimes \sigma(Y) \to \mathcal{U}(\mathfrak{g})$; then $\mu = \operatorname{gr} \tilde{\mu}$. Since μ is an isomorphism, so is $\tilde{\mu}$. □

9.4.3.8 Remark When we choose the orthogonal complement Y on Step 2 above, we can make it $\mathfrak{g}$-invariant by induction on degree. If we do this, then $Y \cong \mathbb{C}[\mathcal{N}]$ as G-modules. ♢

Our motivation for studying the geometry of the nilpotent cone is the following:

9.4.3.9 Lemma / Definition *Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{C}$, $M \in \operatorname{Irr}\mathfrak{g}$, and $z \in \mathcal{Z}(\mathfrak{g})$. Then z acts on M as a scalar: M picks out an algebra homomorphism $|_M : \mathcal{Z}(\mathfrak{g}) \to \mathbb{C}$. Given $\chi \in \operatorname{Spec}\mathcal{Z}(\mathfrak{g})$, the category of irreducible representations M with $|_M = \chi$ is the* block $(\operatorname{Irr}\mathfrak{g})^\chi$*. If $M \in (\operatorname{Irr}\mathfrak{g})^\chi$, then $(\ker\chi)(M) = 0$ and so $(\mathcal{U}\mathfrak{g})(\ker\chi)(M) = 0$. We define $\mathcal{U}_\chi(\mathfrak{g}) \stackrel{\text{def}}{=} \mathcal{U}\mathfrak{g}/(\mathcal{U}\mathfrak{g}\ker\chi)$.* □

9.4.3.10 Remark The Hilbert-Poincaré series of $\mathcal{U}_\chi(\mathfrak{g})$ is independent of the choice of $\chi : \mathcal{Z}(\mathfrak{g}) \to \mathbb{C}$. In fact, $\operatorname{gr}\mathcal{U}_\chi(\mathfrak{g}) \cong \mathcal{S}(\mathfrak{g}^*)/\langle f_1, \dots, f_n\rangle = \mathbb{C}[\mathcal{N}]$, and $\mathcal{U}_\chi(\mathfrak{g})$ and $\mathbb{C}[\mathcal{N}]$ are isomorphic as G-modules. ♢

9.4.3.11 Remark As an algebra, $\mathcal{U}_\chi(\mathfrak{g})$ depends on χ and is non-commutative. Each $\chi \in \operatorname{Spec}\mathcal{Z}(\mathfrak{g})$ gives a *quantization* of $\mathbb{C}[\mathcal{N}]$ in the sense of Definition 13.0.1.6 and Remark 13.0.1.7 ♢

9.5 Borel-Weil-Bott theorem and corollaries

9.5.1 The main theorem

We are interested in the case $H = B$, but it is more straightforward to use $B^- = w_0(B)$. Then $G/B^- \cong G/B$. As in Lemma/Definition 9.3.1.2, G-line bundles on G/B^- are in bijection with one-dimensional representations of B^-. But $B^- = T \ltimes N^-$, where T is the maximal torus in G, and on any one-dimensional representation, the nilpotent part acts trivially. So one-dimensional representations of B^- are in bijection with one-dimensional representations — *characters* — of T. The characters of T are precisely the weight lattice $\mathbf{P}$. (We switched notation from the italic P before, because we want to use "P" to mean a parabolic subgroup.)

The category of line bundles on a space X is a group $\operatorname{Pic}(X)$ under tensor product. We claim without proof:

9.5.1.1 Proposition $\mathrm{Pic}(G/B^-) \cong \mathrm{Pic}_G(G/B^-) \cong \mathbf{P}$ *as groups.* $\square$

By definition, $\mathrm{Pic}_G(G/B^-)$ is the group of G-equivariant line bundles.

9.5.1.2 Definition *If $\lambda \in \mathbf{P}$, we denote by C_λ the one-dimensional representation of B^- with character λ, and we set $\mathcal{O}(\lambda) \overset{\text{def}}{=} G \times_{B^-} C_\lambda$.*

This gives the map $\mathbf{P} \to \mathrm{Pic}_G(G/B^-)$ (which of course maps to $\mathrm{Pic}(G/B^-)$ by forgetting the G-structure) asserted to be an isomorphism in Proposition 9.5.1.1.

9.5.1.3 Example Let $G = \mathrm{SL}(2)$. Then $G/B^- = \mathbb{P}^1$, and $\mathcal{O}(-1)$ is the *tautological line bundle*. When you tensor it, or take its dual, you get the other line bundles. $\lozenge$

9.5.1.4 Theorem (Borel–Weil–Bott)
Assume G is simply connected (there is a version without too). Let $\mu \in \mathbf{P}$. If $\mu + \rho$ is not regular, then $\mathrm{H}^i(G/B^-, \mathcal{O}(\mu)) = 0$ for all i. If $\mu + \rho$ is regular, then there is a unique $w \in W$ with $\mu + \rho = w(\lambda + \rho)$ for $\lambda \in \mathbf{P}^+$, and in this case:

$$\mathrm{H}^i(G/B^-, \mathcal{O}(\mu)) = \begin{cases} 0 & i \neq \ell(w) \\ L(\lambda) & i = \ell(w) \end{cases}$$

So we see in the geometry the same shifted action as in the Weyl character formula.

Proof 1. If μ is not dominant, then $\Gamma(G/B^-, \mathcal{O}(\mu)) = 0$. If μ is dominant, then $\Gamma(G/B^-, \mathcal{O}(\mu)) = L(\mu)$. Then:

$$\mathrm{Hom}_G(L(\nu), \Gamma_{G/B^-}(C_\mu)) = \begin{cases} 0 & \nu \neq \mu \\ \mathbb{C} & \nu = \mu \end{cases}$$

2. Let $G = \mathrm{SL}(2)$, and pick $n \in \mathbb{Z}$. We set $I(n) \overset{\text{def}}{=} \Gamma_{B^-/T}(C_n)$, where the maximal torus T is just the circle S^1, and C_n is the one-dimensional module on which $S^1 = T$ acts n-fold (i.e. the nth tensor power of the defining module $T \hookrightarrow \mathbb{C}$). By Corollary 9.3.1.7, $I(n)$ is an injective B^--module, since C_n is an injective T-module. Explicitly, $\mathfrak{b}^- = \langle h, y\rangle$ and $I(n) = \langle t^{\frac{n}{2}+k} \text{ s.t. } k \in \mathbb{Z}_{\geq 0}\rangle$, and h acts by $2t\frac{\partial}{\partial t}$ and y by $\frac{\partial}{\partial t}$:

$$\begin{tikzcd} \vdots \arrow[d, bend left, "y"] \\ \bullet\, n+4 \arrow[loop left, "h"] \arrow[d, bend left, "y"] \\ \bullet\, n+2 \arrow[loop left, "h"] \arrow[d, bend left, "y"] \\ \bullet\, n \arrow[loop left, "h"] \end{tikzcd}$$

Conversely, C_n is the homology of:

$$0 \to I(n) \to I(n+2) \to 0$$

The middle map is the obvious one that kills the lowest spot and leaves everything else intact.

Then $\mathrm{Hom}_{B^-}(L(m), I(n))$ is easy to write down. It is zero unless $m \geq n$, and then it is just the horizontal maps on weights. So, for positive n, we have:

$$\Gamma_{G/B^-}(I(n)) = \bigoplus_{k=0}^{\infty} L(n+2k)$$

and for negative n it is:

$$\Gamma_{G/B^-}(I(n)) = \bigoplus_{k=0}^{\infty} L(-n+2k)$$

For $n \geq 0$, the homology that we must calculate is for:

$$0 \to \bigoplus_{k=0}^{\infty} L(n+2k) \xrightarrow{\partial} \bigoplus_{k=0}^{\infty} L(n+2+2k) \to 0$$

The boundary map ∂ is bijective except at $L(n)$, and so:

$$\mathrm{H}^i(G/B^-, \mathcal{O}(n)) = \begin{cases} L(n), & i=0 \\ 0, & i=1 \end{cases}$$

When $n < 0$, we do not have sections, and so the map ∂ must be injective:

$$0 \to \bigoplus_{k=0}^{\infty} L(-n+2k) \xrightarrow{\partial} \bigoplus_{k=0}^{\infty} L(-n+2k+2) \to 0$$

Then the result is that, for $n < -1$:

$$\mathrm{H}^i(G/B^-, \mathcal{O}(n)) = \begin{cases} 0, & i=0 \\ L(-n-2), & i=1 \end{cases}$$

The picture is symmetric around -1.

And finally:

$$\mathrm{H}^i(G/B^-, \mathcal{O}(-1)) = 0 \text{ for } i = 0, 1$$

This gives the proof in the case of $\mathrm{SL}(2)$.

3. For the general case, we need some properties of $\Gamma^i_{G/H}$.

   (a) There is a long exact sequence. Suppose you have an exact sequences $0 \to A \to B \to C \to 0$ of H-modules. Then there is a long exact sequence:

   $$0 \to \Gamma^0_{G/H}(A) \to \Gamma^0_{G/H}(B) \to \Gamma^0_{G/H}(C) \to \Gamma^1_{G/H}(A) \to \Gamma^1_{G/H}(B) \to \Gamma^1_{G/H}(C) \to \dots$$

   This is general: as soon as you have a right-derived functor, you have a long exact sequence. The only thing to check is a little homology.

(b) Recall Lemma 9.3.1.5: $\Gamma_{G/H}(\Gamma_{H/K}(V)) = \Gamma_{G/K}(V)$ if $G \supseteq H \supseteq K$. It follows that if $\Gamma^i_{H/K}(V) = 0$ for all i, then $\Gamma^i_{G/K}(V) = 0$ for all i. Indeed, we start with an injective resolution, do the induction, and if we already have an exact sequence of injectives and apply the functor, we get an exact sequence, because Γ is exact on injective modules.

(c) Let M be a finite-dimensional G-module and V some H-module. We claim that $\Gamma^i_{G/H}(V \otimes M) \cong \Gamma^i_{G/H}(V) \otimes M$. Why? First of all, the functor $\otimes M$, if M is finite-dimensional representation, moves injective modules to injective modules — if I is injective, then $I \otimes M$ is injective — because $\otimes M$ has an adjoint functor $\otimes M^*$. So it suffices to prove the statement for Γ^0, which is just Γ. But, recalling Lemma/Definition 9.3.1.3, we construct:

$$M \otimes \Gamma_{G/H}(V) \to \Gamma_{G/H}(M \otimes V)$$

via $(m \otimes \phi)(g) \stackrel{\text{def}}{=} g^{-1} m \otimes \phi(g)$. To construct the inverse, we use the following trick:

$$\Gamma_{G/H}(M \otimes V) \otimes M^* \to \Gamma_{G/H}(M \otimes M^* \otimes V) \stackrel{\text{tr}}{\to} \Gamma_{G/H}(V)$$

Then pulling the M^* over to the right, we get the inverse map.

4. We are now ready to finish the proof of the theorem. We will prove the following lemma due to Bott, although the proof we give is due to Demazure:

9.5.1.5 Lemma *Let $\mu \in \mathbf{P}$, and α_i a simple root, and assume that $\mu(h_i) \geq 0$. Writing $\nu = r_i(\mu + \rho) - \rho$, we have $\mathrm{H}^i(G/B^-, \mathcal{O}(\mu)) = \mathrm{H}^{i+1}(G/B^-, \mathcal{O}(\nu))$.*

Before proving this, let's explain why Lemma 9.5.1.5 implies the theorem. Choose $w_0 = r_{i_1} \circ \cdots \circ r_{i_\ell}$. Then in the middle we count to μ: $r_{i_p} \circ \cdots \circ r_{i_\ell}(\lambda + \rho) = \mu + \rho$. Then $\mathcal{O}(w_0(\lambda + \rho))$ has cohomology only in the highest possible degree. And we can go back, using the facts above, tracking where the cohomology goes. That the highest cohomology cannot be bigger than the dimension of the flag is obvious form the geometric picture.

5. **Proof (of Lemma 9.5.1.5)** Let $\mathfrak{g}_i$ denote the $\mathfrak{sl}(2)$ corresponding to the root α_i, and let $\mathfrak{p}_i = \mathfrak{b}^- + \mathfrak{g}_i$ and $P_i \subseteq G$ the corresponding parabolic subgroup. Geometrically, we have a map of homogeneous spaces $G/B^- \to G/P_i$ with fiber P_i/B^-. But $P_i/\operatorname{Nil}(P_i) = G_i \cdot T$, where G_i is the $\mathrm{SL}(2)$ corresponding to $\mathfrak{g}_i$, and we write $\cdot$ because the product isn't direct: the groups intersect. So $P_i/B^- = \mathbb{P}^1$. So the irreducible representations of P_i are the same as of $G_i \cdot T$, namely the representations $V(\eta)$ with $\eta \in \mathbf{P}$ and $\eta(h_i) \geq 0$.

Incidentally, $\mathcal{O}(-\rho)$ does not have cohomology — this follows from the $\mathrm{SL}(2)$ case — and so $\mathrm{H}^j(P_i/B^-, \mathcal{O}(-\rho)) = 0$. The trick is to take $C_{-\rho}$ a B^- module; then $V = C_{-\rho} \otimes V(\mu + \rho)$ is acyclic everywhere:

$$\mathrm{H}^j(P_i/B^-, C_{-\rho} \otimes V(\mu + \rho)) = 0$$

Hence the same is true for G in place of P_i.

The module $V(\mu + \rho)$ has a three-set filtration with C_μ on the top, C_ν on the bottom, and $V' = C_{-\rho} \otimes V(\mu + \rho - \alpha_i)$ in the middle. So we have, for some X, two exact sequences:

$$0 \to C_\nu \to V \to X \to 0$$
$$0 \to V' \to X \to C_\mu \to 0$$

Since V, V' are acyclic, they drop out in the long exact sequences. All together, we have:

$$\mathrm{H}^i(G/B^-, \mathscr{L}(X)) \cong \mathrm{H}^i(G/B^-, \mathcal{O}(\mu))$$
$$\mathrm{H}^i(G/B^-, \mathscr{L}(X)) \cong \mathrm{H}^{i+1}(G/B^-, \mathcal{O}(\nu)) \qquad \square$$

9.5.2 Differential operators and more on the nilpotent cone

Let's think philosophically about what we did in the previous section. We gave a certain geometric construction of finite-dimensional representations: the Borel-Weil-Bott theorem allows you to realize a finite-dimensional representation of G, a semisimple group, as the sections of some line bundle. We want to push this farther. We would like to do something with infinite-dimensional representations. Thus, we are led to the following question: is it possible to get some geometric realization of the representations of the Lie algebra $\mathfrak{g}$?

The answer is yes. If a group G acts on a set X, then it acts on the space of sections of any bundle over X. If we start with $X = G/B^-$ and the line bundle $\mathcal{O}(\lambda)$, then $\Gamma(\mathcal{O}(\lambda))$ is a representation of G, and hence also of $\mathfrak{g} = \mathrm{Lie}(G)$. But we can go further: pick an open set $U \subseteq G/B^-$. Then $\Gamma(U, \mathcal{O}(\lambda))$ is not a G-module, because the G action would take your from inside U to outside it. But it is a $\mathfrak{g}$-module.

9.5.2.1 Example Let $G = \mathrm{SL}(2)$. Then $G/B^- = \mathbb{P}^1$ is the projective line. As our open subset $U \subseteq \mathbb{P}^1$, we'll take the open Schubert cell — this is the sphere without the north pole, so isomorphic to $\mathbb{C}$, c.f. Example 9.2.5.2 — and we consider $\Gamma(U, \mathcal{O}(n))$. Since any line bundle over $U \cong \mathbb{C}$ is trivializable, there is an isomorphism $\Gamma(U, \mathcal{O}(n)) \cong \mathbb{C}[z]$. One such choice of trivialization corresponds to the action of $\mathfrak{g} = \langle y, h, x \rangle$ on $\mathbb{C}[z]$ by:

$$y \mapsto \frac{\partial}{\partial z} \qquad h \mapsto 2z\frac{\partial}{\partial z} - n \qquad x \mapsto -z^2\frac{\partial}{\partial z} + nz$$

It should be noted that the action is not by vector fields, although it is an action by first-order differential operators. $\Diamond$

9.5.2.2 Example Recall the *Weyl algebra* generated by $z, \frac{\partial}{\partial z}$. It acts on $\mathbb{C}[z]$ in the usual way. Let's freely adjoint to $\mathbb{C}[z]$ a symbol δ_0, which we think of as $\delta(z-0)$, subject to the constraint $z\delta_0 = 0$, and try to extend the representation of the Weyl algebra. Of course, the best way to do this is to freely define $\delta_0', \delta_0'', \dots$ as the derivatives of δ_0, and the algebraic relationships between these and the previously defined terms will follow from the product rule.

Let $\mathfrak{sl}(2)$ act on $\mathbb{C}[z]$ as in Example 9.5.2.1. We can extend this to our module of "generalized functions" $\mathbb{C}[z][\delta_0]/(z\delta_0 = 0)$, and an easy calculation shows that $x\delta_0 = 0$, $h\delta_0 = (-n-2)\delta_0$, and y acts freely. So what we get is the Verma module $M(-n-2)$. $\Diamond$

9.5.2.3 Definition *Let X be a non-singular algebraic variety, $\mathcal{L}$ a line bundle over X, and $U \subseteq X$ and affine open set. The* differential operators on U with coefficients in $\mathcal{L}$ *is the filtered infinite-dimensional algebra $\mathscr{D}(U, \mathcal{L}) \subseteq \mathrm{End}(\Gamma(U, \mathcal{L}))$ defined inductively via:*

$$\mathscr{D}_{\leq 0}(U, \mathcal{L}) = \mathscr{O}(U)$$
$$\mathscr{D}_{\leq i}(U, \mathcal{L}) = \left\{\delta \in \mathrm{End}(\Gamma(U, \mathcal{L})) \text{ s.t. } [\delta, \phi] \in \mathscr{D}_{\leq i-1}(U, \mathcal{L})\ \forall \phi \in \mathscr{O}(U)\right\}$$

Here $\mathscr{O}(U)$ is the algebra of regular functions on U, acting linearly on fibers. Compare with Definition 3.2.4.1.

For example, upon trivializing $\Gamma(U,\mathcal{L}) \cong \mathscr{O}(U)$, the condition for whether $\delta \in \mathscr{D}_{\leq 1}(U,\mathcal{L})$ is that $[\delta, -]$ be a derivation, so $\delta - \delta(1)$ is a vector field. One can make a similar definition replacing $\mathscr{O}(U)$ with any commutative algebra and $\Gamma(U,\mathcal{L})$ with any module.

9.5.2.4 Example If $\mathcal{U} = \mathbb{C}^n$ and $\mathcal{L}$ is trivial, then $\mathscr{D}(U)$ is nothing else but the Weyl algebra. If our coordinates on $\mathbb{C}^n$ are $x_1, \dots, x_n$, then

$$\mathscr{D}(U) = \mathcal{T}(x_1, \dots, x_n, \partial_1, \dots \partial_n) / \big\langle [x_i, x_j] = 0, [\partial_i, \partial_j] = 0, [\partial_i, x_j] = \delta_{ij} \big\rangle$$

The filtration is given by $\deg \partial_i = 1$, $\deg x_i = 0$. ♢

9.5.2.5 Remark When U is not affine, $\Gamma(U,\mathcal{L})$ might have very few sections, and so Definition 9.5.2.3 will not in general give a sheaf as written. Rather, Definition 9.5.2.3 defines a sheaf of algebras $\mathscr{D}(-,\mathcal{L})$. ♢

By construction, $\operatorname{gr} \mathscr{D}(U,\mathcal{L})$ is a commutative algebra. We claim that if A is any filtered algebra so that $\operatorname{gr} A$ is commutative, then $\operatorname{gr} A$ is naturally Poisson with the bracket of total degree -1. In brief: let $x \in \operatorname{gr} A_m$ and $y \in \operatorname{gr} A_n$ be represented by $\tilde{x} \in A_{\leq m}$ and $\tilde{y} \in A_{\leq n}$. Then $[\tilde{x}, \tilde{y}] \in A_{m+n-1}$, as $\operatorname{gr} A$ is commutative, and $[\tilde{x}, \tilde{y}]$ represents $\{x, y\} \in \operatorname{gr} A_{m+n-1}$. The fact is that, if $\operatorname{gr} A$ is commutative, then $\{x, y\}$ does not depend on the choice of representatives $\tilde{x}, \tilde{y}$. In the case of $A = \mathscr{D}(U,\mathcal{L})$, more can be said (c.f. Theorem 3.2.4.3):

9.5.2.6 Proposition *If X is non-singular, then $\operatorname{gr} \mathscr{D}(X,\mathcal{L}) = \Gamma(X, \mathcal{S}^\bullet(\mathrm{T}X))$. Here $\mathrm{T}X$ is the tangent bundle, and $\mathcal{S}^\bullet(\mathrm{T}X)$ is the sheaf of symmetric polyvector fields. Geometrically, $\Gamma(X, \mathcal{S}^\bullet(\mathrm{T}X)) = \mathscr{O}(\mathrm{T}^*X)$, T^*X is a Poisson manifold, and the isomorphism is actually of Poisson algebras.*

Proof (sketch) It suffices to consider the affine case. Let $R = \mathscr{O}(X)$ with X affine. Then $\mathscr{D}_{\leq 1}(X,\mathcal{L})$ consists of the linear endomorphism $\delta : \Gamma(\mathcal{L}) \to \Gamma(\mathcal{L})$ such that $[\delta, R] \subseteq R$. Moreover, $[\delta, -] : R \to R$ is a derivation. This gives a canonical map $\operatorname{gr}_1 \mathscr{D}(X,\mathcal{L}) = \mathscr{D}_{\leq 1}(X)/R \to \operatorname{Der} R$. By working sufficiently locally that $\mathcal{L}$ trivializes, check that this map is locally, hence globally, an isomorphism. Note that $\operatorname{Der} R = \Gamma(X, \mathrm{T}X)$ since every derivation factors through the de Rham differential $\mathrm{d} : R \to \Gamma(\mathrm{T}^*X)$ and to get back to R you contract with your vector field.

The proof repeats the above analysis with 1 replaced by n. If $\delta \in \mathscr{D}_{\leq n}(X,\mathcal{L})$, then its image in $\mathscr{D}_{\leq n}(X,\mathcal{L})/\mathscr{D}_{\leq n-1}(X)$ acts as a derivation $R \to \Gamma(X, \mathcal{S}^{n-1}(\mathrm{T}X))$, and hence factors through the de Rham differential and then you contract with a polyvector field. What must be shown is that this is exactly all there are.

To see that the isomorphism is one of Poisson algebras, one can use the fact that $\Gamma(X, \mathcal{S}^\bullet(\mathrm{T}X))$ is generated by its degree-zero and -one pieces $\mathscr{O}(X)$ and $\operatorname{Der}(X)$, and check that these pieces have the desired Poisson brackets. □

In our situation, we let $X = G/B^-$ be the flag manifold and we let $\mathcal{L} = \mathcal{O}(\lambda)$. The group action determines a homomorphism $\mathfrak{g} \to \mathcal{D}_{\leq 1}(X, \mathcal{O}(\lambda))$; Example 9.5.2.1 gives the $\mathfrak{sl}(2)$ case. By universality, this extends to an algebra homomorphism $\theta_\lambda : \mathcal{U}\mathfrak{g} \to \mathscr{D}(X, \mathcal{O}(\lambda))$. Recall from Definition 9.4.1.7 the *Harish-Chandra homomorphism* $\theta : \mathcal{Z}(\mathfrak{g}) \to \mathcal{S}(\mathfrak{h}) = \mathrm{Pol}(\mathfrak{h}^*)$ and its dual map $\theta^* : \mathfrak{h}^* \to \mathrm{Hom}(\mathcal{Z}(\mathfrak{g}), \mathbb{C})$. We denote the central character $\theta^*(\lambda) : \mathcal{Z}(\mathfrak{g}) \to \mathbb{C}$ by χ_λ. It turns out that the algebraic discussion of central characters matches the representations of the geometric algebra $\mathscr{D}(X, \mathcal{O}(\lambda))$:

9.5.2.7 Proposition *Let λ be a weight. Then $\theta_\lambda(\ker \chi_\lambda) = 0$ and so θ_λ factors through $\mathcal{U}_\lambda\mathfrak{g} \overset{\text{def}}{=} \mathcal{U}\mathfrak{g}/(\mathcal{U}\mathfrak{g} \ker \chi_\lambda)$. Moreover, $\theta_\lambda : \mathcal{U}_\lambda\mathfrak{g} \to \mathscr{D}(G/B^-, \mathcal{O}(\lambda))$ is an isomorphism.*

9.5.2.8 Remark In fact, λ need not be an integral weight. Although we have not defined it, $\mathscr{D}(G/B^-, \mathcal{O}(\lambda))$ makes sense geometrically for arbitrary $\lambda \in \mathfrak{h}^*$, even though the line bundle $\mathcal{O}(\lambda)$ is not globally defined unless λ is integral. We will discuss this further in Section 9.5.3. $\Diamond$

We will prove Proposition 9.5.2.7 in a series of lemmas.

9.5.2.9 Lemma *If $z \in \mathcal{Z}(\mathfrak{g})$, then $\chi_\lambda(z) = \theta_\lambda(z)$.*

Proof Pick $x \in X = G/B^-$ such that $\mathrm{Stab}_G(x) = B^-$, and let $U \ni x$ be an open set. Let $I_x \subseteq \mathcal{O}(\lambda)|_U$ be the $\mathscr{O}(U)$-submodule of sections vanishing at x; then $B^- \cdot I_x = I_x$. For $\xi \in \mathfrak{b}^-$ and $\varphi \in \mathcal{O}(\lambda)|_U$, by definition $\xi \cdot \varphi = \theta_\lambda(\xi)[\varphi]$. But since I_x is fixed by B^-, this action descends to the quotient $\mathcal{O}(\lambda)|_U / I_x$, which is a line (or 0 if U is too big). In particular, the value of $(\xi \cdot \varphi)(x)$ depends only on $\varphi(x)$.

Recall Definition 9.4.1.7: for $z \in \mathcal{Z}(\mathfrak{g})$, we have $z = \theta(z) \mod \mathfrak{n}^-\mathcal{U}(\mathfrak{g})$ for $\theta(z) \in \mathcal{S}\mathfrak{h}$. But $\mathfrak{n}^-\mathcal{O}(\lambda) = I_x$, and so $(z \cdot \varphi)(x) = (\theta(z) \cdot \varphi)(x) = \chi_\lambda(z)\varphi(x)$. This proves that $\chi_\lambda(z) = \theta_\lambda(z)$ with respect to their actions at the point x, and so now we can do it at any point, because z is in the center. Indeed, since $z \in \mathcal{Z}(\mathfrak{g})$, z commutes with the action of G, which is transitive, so we can start with x and move it to any other point: $z \cdot g^*(\phi) = g^*(z \cdot \phi)$. $\square$

9.5.2.10 Corollary *θ_λ factors through $\mathcal{U}\mathfrak{g} \to \mathcal{U}_\lambda\mathfrak{g} = \mathcal{U}\mathfrak{g}/\mathcal{U}\mathfrak{g} \ker \chi_\lambda$.* $\square$

To prove the last part of Proposition 9.5.2.7 — that $\theta_\lambda : \mathcal{U}_\lambda\mathfrak{g} \to \mathscr{D}(G/B^-, \mathcal{O}(\lambda))$ is an isomorphism —, we will look at the associated graded algebras on both sides. Each side is naturally filtered, and θ respects the filtrations, so we will then use the standard fact that we discussed in the proof of Theorem 9.4.1.14 that lets us go back.

9.5.2.11 Lemma *Let $\mathscr{N}$ be the cone of nilpotent elements in $\mathfrak{g}$. Then $\mathrm{gr}(\mathcal{U}_\lambda\mathfrak{g}) = \mathbb{C}[\mathscr{N}]$, the ring of regular functions on $\mathscr{N}$.*

Proof Recall the following facts about this ring of functions: it is a polynomial algebra, and the center acts freely. We identify $\mathfrak{g} \cong \mathfrak{g}^*$ via the Killing form. Then remember what we did in the proof of Theorem 9.4.3.7: we took $I(\mathscr{N})$ the ideal of the cone, and then $\mathcal{S}\mathfrak{g} = I(\mathscr{N}) \oplus Y$, where Y was a homogeneous compliment. Then we proved that $\mathcal{S}\mathfrak{g}^G \otimes Y \to \mathcal{S}\mathfrak{g}$ is an isomorphism. Moreover, we have the natural maps $\gamma : \mathcal{S}\mathfrak{g} \to \mathcal{T}\mathfrak{g} \to \mathcal{U}\mathfrak{g}$, and so we have:

$$\begin{tikzcd} \mathcal{S}\mathfrak{g}^G \otimes Y \arrow[r] \arrow[d, "\sim"] & \mathcal{S}\mathfrak{g} \arrow[d, "\sim"] \\ \mathcal{Z}(\mathfrak{g}) \otimes \gamma(Y) \arrow[r] & \mathcal{U}\mathfrak{g} \end{tikzcd}$$

$\square$

In Proposition 9.5.2.6 we observed that for any X and any line bundle $\mathcal{L}$, $\mathrm{gr}(\mathscr{D}(X,\mathcal{L})) = \mathscr{O}(\mathrm{T}^*X)$. Since G acts transitively by conjugation on Borel subalgebras in $\mathfrak{g}$ and the stabilizer of a given Borel is itself, we can identify G/B^- as the set of Borel subalgebras in $\mathfrak{g}$. We denote the Borel corresponding to $x \in X = G/B^-$ by $\mathfrak{b}_x$. Therefore $\mathrm{T}^*X = \{(x,\xi) \text{ s.t. } x \in X, \xi \in (\mathfrak{g}/\mathfrak{b}_x)^*\}$, because we identify $\mathrm{T}_xX \cong \mathfrak{g}/\mathfrak{b}_x$. By the Killing form, $(\mathfrak{g}/\mathfrak{b}_x)^* \cong \mathfrak{n}_x$, where $\mathfrak{n}_x = [\mathfrak{b}_x, \mathfrak{b}_x]$, because $(\mathfrak{b}_x)^\perp$ with respect to the Killing form is just $\mathfrak{n}_x$. All together, we think of the elements of T^*X as pairs a Borel subalgebra $\mathfrak{b}_x$ and $\xi \in \mathfrak{n}_x = [\mathfrak{b}_x, \mathfrak{b}_x]$. In particular, ξ is nilpotent. We define the map $p : \mathrm{T}^*X \to \mathscr{N}$ that forgets the first factor, projecting $(\mathfrak{b}_x, \xi) \mapsto \xi \in \mathscr{N}$. Since we have exhibitted T^*X as a subbundle of $X \times \mathfrak{g}$, this is a "projection onto the fiber".

9.5.2.12 Lemma *1. p is surjective.*

2. If $\xi \in \mathscr{N}$ is regular, then $p^{-1}(\xi)$ is a single point.

3. $p^{-1}(\xi)$ is a connected projective variety.

4. $p^* : \mathbb{C}[\mathscr{N}] \to \mathcal{O}(\mathrm{T}^*X)$ is an isomorphism.

So in algebrogeometric language, p is *proper*, and is an isomorphism on open parts.

Proof 1. is by inspection. 2. Any regular ξ can be embedded in a principle $\mathrm{SL}(2) = \{\eta, h, \xi\}$. Then $\mathrm{Stab}_{\mathfrak{g}}(\xi) \subseteq \mathfrak{b}$ for some Borel $\mathfrak{b}$, and we claim it is unique. Indeed, you pick up the regular piece, look at the centralizer, and see that the centralizer of the pair $(\mathfrak{b}, \xi)$ is the same as the centralizer of ξ, and therefore there is only one $\mathfrak{b}$. The best way to see this is to pick up one particular $\mathfrak{b}$, and then construct the centralizer and see that there are only positive weights. We explained 3. and 4. already. $\square$

Proof (of Proposition 9.5.2.7) We claim that $\mathrm{gr}\,\theta_\lambda = p^*$. We have $\mathbb{C}[\mathscr{N}] \subseteq \mathcal{S}\mathfrak{g} \to \mathcal{S}(\mathrm{T}X)$. For each $\xi \in \mathfrak{g}$ we want to construct a vector field on X. To define the vector at x, we look at the image of ξ in $\mathfrak{g}/\mathfrak{b}_x = \mathrm{T}_xX$. But $p^* : \mathcal{S}\mathfrak{g} \to \mathcal{S}(\mathrm{T}X)$ is the associated graded for $\theta_\lambda : \mathcal{U}\mathfrak{g} \to \mathscr{D}(X,\mathcal{L})$. $\square$

9.5.3 Twisted differential operators and Beilinson-Bernstein

In the previous section we proved that $\mathscr{D}(X, \mathcal{O}(\lambda)) \cong \mathcal{U}_\lambda\mathfrak{g} = \mathcal{U}\mathfrak{g}/\big(\ker(\chi_\lambda)\mathcal{U}\mathfrak{g}\big)$ when $\lambda \in \mathbf{P}$ is an integral weight. But the right-hand side makes sense for arbitrary weights λ; we will now discuss the generalization of the left-hand side.

9.5.3.1 Definition *A system of* twisted differential operators *(a* TDO*) on a space X is a sheaf of filtered algebras locally isomorphic to the sheaf of differential operators with trivial coefficients (Definitions 3.2.4.1 and 9.5.2.3). We denote the space of TDOs on X by $\mathfrak{TDO}(X)$.*

We denote the sheaf of differential operators with trivial coefficients by $\mathscr{D}$. To study TDOs it is necessary to understand the automorphisms of $\mathscr{D}(U)$, as these will give possible transition maps. Since $\mathscr{D}(U)$ is generated as a filtered algebra by $\mathscr{D}_{\leq 1}(U) = \Gamma(\mathrm{T}U) \oplus \mathscr{O}(U)$ and since the automorphisms must preserve the filtration, said automorphisms must fix the $\mathscr{O}(U)$ part and for $v \in \Gamma(\mathrm{T}U)$ be of the form $v \mapsto v + \langle \alpha, v\rangle$ for some $\alpha \in \Omega^1_{\mathrm{d}}(U) = \Gamma(\mathrm{T}^*U)$. Moreover, to get the commutation relations we must have $\mathrm{d}\alpha = 0$. We have proven:

9.5.3.2 Lemma *There is a bijection $\mathfrak{TDO}(X) \leftrightarrow \mathrm{H}^1(X, \Omega^\bullet_{\mathrm{d}}(X))$.* □

9.5.3.3 Example Let $X = G/B$. We have a short exact sequence of sheaves:

$$0 \to \mathbb{C} \to \mathscr{O}_X \to (\Omega^1_{\mathrm{d}})_X \to 0$$

Taking the long exact sequence in cohomology, we know that $\mathscr{O}_X$ has only nonzero cohomology in dimension 1. Therefore:

$$\mathrm{H}^1(X, \Omega^1_d)_X \cong \mathrm{H}^2(X, \mathbb{C})$$

The right-hand side gives Schubert cells. And if an element of the left-hand side gives a sheaf of twisted differential operators on a line bundle, then the corresponding class on the right-hand side is the *Chern class*. ◊

9.5.3.4 Definition *A* Lie algebroid *on a space X is a sheaf $\mathfrak{a}$ of $\mathscr{O}_X$ modules that is simultaneously a sheaf of Lie algebras (same $\mathbb{C}$ action, but the bracket need not be $\mathscr{O}_X$-linear). The $\mathscr{O}_X$ and Lie algebra structures are required to satisfy a "Leibniz rule" compatibility condition, namely the existence a an* anchor map *$\alpha : \mathfrak{a} \to \Gamma(-, \mathrm{T}X)$ of sheaves of $\mathscr{O}_X$-modules such that, for local sections $\varphi \in \mathscr{O}_X$ and $\xi, \eta \in \mathfrak{a}$, $[\xi, \varphi\eta] = \varphi[\xi, \eta] + L_{\alpha(\xi)}(\varphi)\eta$. It follows that α is a Lie algebra homomorphism.*

9.5.3.5 Remark When doing differential geometry, one often adds the requirement that the $\mathscr{O}_X$-module structure on $\mathfrak{a}$ makes $\mathfrak{a}$ into the sheaf of sections of some vector bundle $A \to X$, i.e. that $\mathfrak{a}$ be locally free as an $\mathscr{O}_X$-module. ◊

9.5.3.6 Example Suppose that a Lie algebra $\mathfrak{g}$ acts on a space X. Then the *action algebroid* $\tilde{\mathfrak{g}}$ is, as a sheaf of left $\mathscr{O}_X$-modules, the sheaf of $\mathfrak{g}$-valued functions on X: $\tilde{\mathfrak{g}} = \mathscr{O}_X \otimes \mathfrak{g}$. The anchor map is determined by the action, and the bracket on sections of $\tilde{\mathfrak{g}}$ is:

$$[f \otimes \xi, g \otimes \eta] = fg \otimes [\xi, \eta] + f\, L_\xi(g) \otimes \eta - g\, L_\eta(f) \otimes \xi \qquad ◊$$

9.5.3.7 Definition *Let $\mathfrak{a}$ be a Lie algebroid on X. The* universal enveloping algebroid $\mathcal{U}\mathfrak{a}$ *is the sheaf of algebras on X generated by $\mathscr{O}_X$ and $\mathfrak{a}$ and subject to the relations that: the multiplication $\mathscr{O}_X \otimes \mathscr{O}_X \to \mathcal{U}\mathfrak{a}$ is the multiplication in $\mathscr{O}_X$; the multiplication $\mathscr{O}_X \otimes \mathfrak{a} \to \mathcal{U}\mathfrak{a}$ is the action by $\mathscr{O}_X$ on $\mathfrak{a}$; the commutator in $\mathcal{U}\mathfrak{a}$ between sections of $\mathfrak{a}$ is given by the Lie bracket; the commutator in $\mathcal{U}\mathfrak{a}$ between a section of $\mathfrak{a}$ and a section of $\mathscr{O}_X$ is given by the anchor map.*

9.5.3.8 Example The tangent bundle $\mathrm{T}X$ defines a Lie algebroid $\Gamma(-, \mathrm{T}X)$: the bracket is the bracket of vector fields and the anchor map is the identity. Its universal enveloping algebra is the sheaf $\mathscr{D}_X$ of differential operators on X from Definitions 3.2.4.1 and 9.5.2.3.

More generally, when $\mathcal{L}$ is a line bundle on X, then $\mathscr{D}_{\leq 1}(-,\mathcal{L})$ is a Lie algebroid containing $\mathscr{O}_X$. Its universal enveloping algebroid is a bit too big to be $\mathscr{D}(-,\mathcal{L})$, but upon identifying the $\mathscr{O}_X$ in $\mathscr{D}_{\leq 1}$ with the $\mathscr{O}_X$ in $\mathcal{U}\mathscr{D}_{\leq 1}$, we arrive at the sheaf of differential operators with coefficients in $\mathcal{L}$. ♢

We now restrict our attention to $X = G/B$ for G a semisimple connected simply-connected algebraic group and B a Borel. It carries a G-action, and so we can form the action algebroid $\tilde{\mathfrak{g}}$ on X. Writing $\alpha : \tilde{\mathfrak{g}} \to \Gamma(\mathrm{T}X)$ for the anchor map, we have:

$$\tilde{\mathfrak{b}}|_U \overset{\text{def}}{=} \ker(\alpha)|_U = \{\varphi : U \to \mathfrak{g} \text{ s.t. } \varphi(x) \in \mathfrak{b}_x \ \forall x \in U\}$$
$$[\tilde{\mathfrak{b}}, \tilde{\mathfrak{b}}]|_U = \{\varphi : U \to \mathfrak{g} \text{ s.t. } \varphi(x) \in \mathfrak{n}_x \ \forall x \in U\}$$

Moreover, there are canonical isomorphisms $\mathfrak{b}_x/\mathfrak{n}_x \cong \mathfrak{b}_y/\mathfrak{n}_y$ for all $x, y \in X$, since the choice is up to conjugation by B. But $\mathfrak{b}_x/\mathfrak{n}_x \cong \mathfrak{h}$.

We set $\tilde{\mathcal{U}} = \mathscr{O}_X \otimes \mathcal{U}\mathfrak{g}$, the sheaf of $\mathcal{U}\mathfrak{g}$-valued functions on X. Each $\lambda \in \mathfrak{h}^*$ gives a map $\lambda : \mathfrak{b}_x \to \mathbb{C}$ for each $x \in X$, since $\mathfrak{b}_x/\mathfrak{n}_x = \mathfrak{h}$. Let's denote by $\mathcal{I}_\lambda$ the ideal in $\tilde{\mathcal{U}}$ generated by $(\varphi - \lambda(\varphi))$ for $\varphi \in \tilde{\mathfrak{b}}$. Then we define $\mathscr{D}_X^\lambda \overset{\text{def}}{=} \tilde{\mathcal{U}}/\mathcal{I}_\lambda$.

9.5.3.9 Lemma *1. $\mathscr{D}_X^\lambda$ is a TDO.*

2. $\mathscr{D}^\lambda(X) \overset{\text{def}}{=} \Gamma(X, \mathscr{D}_X^\lambda) = \mathcal{U}_\lambda\mathfrak{g}$.

Proof For the second statement, go to associated graded, same as before. For the first statement, calculate: for $U \subseteq X$, $\mathscr{D}_0^\lambda(U) = \mathscr{O}(U)$, and:

$$\frac{\mathscr{D}_1^\lambda(U)}{\mathscr{D}_0^\lambda(U)} = \frac{\mathscr{O}(U) \otimes \mathfrak{g}}{\ker \alpha} = \Gamma(U, \mathrm{T}X)$$

and so we are locally isomorphic to differential operators. □

9.5.3.10 Definition *Let X be an algebraic variety with sheaf of functions $\mathscr{O}_X$. An $\mathscr{O}_X$-module $\mathcal{M}$ is* quasicoherent *if for a small enough cover, for $V \subseteq U$, we have $\mathcal{M}(V) = \mathcal{M}(U) \times_{\mathscr{O}(U)} \mathscr{O}(V)$. For X affine, quasicoherent sheaves are the same as $\mathbb{C}[X]$-modules, e.g. sections of vector bundles. However, when X is projective, sheaves have too few global sections, so quasicoherence is a better notion than "module over global sections".*

We have in front of us two interesting categories of modules. On the one hand, we have the category $\mathcal{U}_\lambda\mathfrak{g}$-MOD of modules over the algebra $\mathscr{D}^\lambda(X) \cong \mathcal{U}_\lambda\mathfrak{g}$. On the other hand, we have the category $\mathscr{D}_X^\lambda$-MOD of sheaves of $\mathscr{D}_X^\lambda$-modules that are quasicoherent as $\mathscr{O}_X$-modules. In fact, these categories are sometimes the same:

9.5.3.11 Theorem (Beilinson–Bernstein)
Assume that λ is dominant and $\lambda + \rho$ is regular, but not necessarily integral. The global sections *functor Γ sends $\mathcal{F} \in \mathscr{D}_X^\lambda$-MOD to $\Gamma(X, \mathcal{F}) \in \mathcal{U}_\lambda\mathfrak{g}$-MOD. If $F \in \mathcal{U}_\lambda\mathfrak{g}$-MOD, we define its* localization *via $(\mathrm{L}\,F)(U) = \mathscr{D}^\lambda(U) \otimes_{\mathscr{D}^\lambda(X)} F$. These functors define an equivalence of categories $\mathrm{L} : \mathcal{U}_\lambda\mathfrak{g}\text{-MOD} \rightleftarrows \mathscr{D}_X^\lambda\text{-MOD} : \Gamma$.*

9.5.3.12 Remark When λ is integral, Theorem 9.5.1.4 shows that dominance of λ and regularity of $\lambda+\rho$ are necessary. $\Diamond$

As usual, we prove Theorem 9.5.3.11 via a series of lemmas.

Let $\mathcal{E},\mathcal{F}$ be $\tilde{\mathcal{U}}$-modules. Then $\mathcal{E}\otimes_{\mathcal{O}_X}\mathcal{F}$ is again a $\tilde{\mathcal{U}}$-module. For $\mu\in\mathbf{P}$, we set $\mathcal{F}(\mu)\stackrel{\text{def}}{=}\mathcal{O}(\mu)\otimes_{\mathcal{O}}\mathcal{F}$. Since we have an infinitesimal action on the fiber, the weights add. In particular, if $\mathcal{F}$ is a $\mathcal{D}_X^\lambda$-module (i.e. if $\mathcal{I}_\lambda\subseteq\tilde{\mathcal{U}}$ acts as 0 on $\mathcal{F}$), then $\mathcal{F}(\mu)$ is a $\mathcal{D}_X^{\lambda+\mu}$-module.

Suppose now that μ is dominant and integral, and consider $V=L(\mu)$. Its induced bundle $G\times_B V$ has sections $\mathcal{V}=\mathcal{O}_X\otimes V$, and thus is a $\tilde{\mathcal{U}}$-module. Being a finite-dimensional B-representation, V has a $\mathfrak{b}^-$-invariant filtration $0\subseteq V_1\subseteq V_2\subseteq\cdots\subseteq V_s=V$, such that the quotients are all one-dimensional $\mathfrak{b}^-$-representations — indeed, each quotient V_i/V_{i-1} is of the form C_{γ_i} for some weight γ_i of V (we played a similar trick in the proof of Theorem 9.5.1.4). In particular, $\gamma_1=\nu$ is necessarily the lowest weight of V, and $\gamma_s=\mu$ is the highest weight. Moving to sheaves, we have a similar story:

$$0\subseteq\mathcal{V}_1\subseteq\mathcal{V}_2\subseteq\cdots\subseteq\mathcal{V}_s=\mathcal{V}$$

This time $\mathcal{V}_i/\mathcal{V}_{i-1}=\mathcal{O}(\gamma_i)$.

9.5.3.13 Lemma *Pick $\mathcal{F}\in\mathscr{D}_X^\lambda$-MOD. The maps $C_\nu=V_1\to V$ and $V\to V_s/V_{s-1}=C_\mu$ determine maps $\mathcal{O}(\nu)\to\mathcal{V}$ and $\mathcal{V}\to\mathcal{O}(\mu)$, and hence maps $i:\mathcal{F}\to\mathcal{F}(\nu)(-\nu)\to\mathcal{F}(-\nu)\otimes\mathcal{V}$ and $j:\mathcal{F}\otimes\mathcal{V}\to\mathcal{F}(\mu)$. Then i has a right inverse and j has a left inverse.*

Proof The idea of the proof is as follows. $\mathcal{F}(\gamma_i)$ carries an action of $\mathcal{D}_X^{\lambda+\gamma_i}$. We will prove:

1. If $\gamma_i\neq\nu$, then $\mathcal{U}_{\lambda-\nu+\gamma_i}\mathfrak{g}\neq\mathcal{U}_\lambda\mathfrak{g}$.
2. If $\gamma_i\neq\mu$, then $\mathcal{U}_{\lambda+\gamma_i}\mathfrak{g}\neq\mathcal{U}_{\lambda+\nu}\mathfrak{g}$.

Fact 1. proves that i has a right inverse, and 2. that j has a left inverse.

The claims follow from Theorem 9.4.1.14: if two weights define the same central character, then they are on the same orbit of the shifted Weyl group. For 1., we argue as follows. We must show that $\lambda+\gamma_i-\nu\neq w(\lambda+\rho)-\rho$ for any $w\in W$. Assume the opposite. Then $w(\lambda+\rho)-(\lambda+\rho)=\gamma_i-\nu$. But ν is the lowest weight of $V=L(\mu)$, so $\gamma_i-\nu>0$. On the other hand, $w(\lambda+\rho)-(\lambda+\rho)\leq 0$ as λ is dominant, a contradiction.

For 2., we give a similar argument, this time using regularity of $\lambda+\rho$ rather than dominance of λ. Assume that $w(\lambda+\gamma_i+\rho)=\lambda+\mu+\rho$; then $\lambda+\rho-w(\lambda+\rho)=w(\gamma_i)-\mu$. But $w(\gamma_i)-\mu\leq 0$ as μ is the highest weight, and $\lambda+\rho-w(\lambda+\rho)>0$ by regularity. $\square$

9.5.3.14 Lemma *For $\mathcal{F}\in\mathscr{D}_X^\lambda$-MOD, we have $\mathrm{H}^i(X,\mathcal{F})=0$ for $i\neq 0$, and $\mathrm{H}^0(X,\mathcal{F})\neq 0$ for $\mathcal{F}\neq 0$.*

Proof We will give the proof when $\mathcal{F}$ is coherent, i.e. finite-generated. For general $\mathcal{F}$ one would then need to use the usual trick, which we will skip: a quasicoherent sheaf is an inductive limit of coherent sheaves, and you must check various maps.

We will use but not prove *Serre's theorem*: if $X\hookrightarrow\mathbb{P}^n$ and $\mathcal{F}$ is a non-zero coherent sheaf on X, then $\mathcal{F}\otimes\mathcal{O}(m)$ does not have nonzero cohomology for large enough m. In our case, we have $\pi : X \hookrightarrow \mathbb{P}(L(\kappa))$ for some κ, and $\mathcal{O}_X(\kappa) = \pi^*\mathcal{O}_{\mathbb{P}}(1)$. In particular, for $i > 0$ and $-\nu$ sufficiently large (a large integer multiple of κ), $\mathrm{H}^i(X, \mathcal{F}(-\nu)) = 0$. Then $\mathrm{H}^i(X, \mathcal{F}(-\nu) \otimes_{\mathcal{O}_X} \mathcal{V}) = \mathrm{H}^i(X, \mathcal{F}(-\nu)) \otimes V = 0$. By Lemma 9.5.3.13, after tensoring with $\mathcal{O}(-\nu)$, $\mathcal{F}$ is a direct summand of $\mathcal{F}(-\nu) \otimes_{\mathcal{O}_X} \mathcal{V}$. Thus $\mathrm{H}^i(X, \mathcal{F}) = 0$ for $i \neq 0$.

The $i = 0$ case is similar. For sufficiently large μ and non-zero $\mathcal{F}$, $\mathrm{H}^0(X, \mathcal{F}(\mu)) \neq 0$ by Serre's theorem. If we did have $\mathrm{H}^0(X, \mathcal{F}) = 0$, then we would have $\mathrm{H}^0(X, \mathcal{F} \otimes_{\mathcal{O}_X} \mathcal{V}) = 0$ by the argument in the previous paragraph, and hence $\mathrm{H}^0(X, \mathcal{F}(\mu)) = 0$ as by Lemma 9.5.3.13 $\mathcal{F}(\mu)$ is a direct summand of $\mathcal{F} \otimes_{\mathcal{O}_X} \mathcal{V}$. But this is a contradiction. □

Proof (of Theorem 9.5.3.11) The functors L and Γ form an adjoint pair:

$$\mathrm{Hom}_{\mathscr{D}_X^\lambda}(\mathrm{L}\, F, \mathcal{F}) = \mathrm{Hom}_{\mathscr{D}^\lambda(X)}(F, \Gamma\mathcal{F})$$

Lemma 9.5.3.9 shows that the canonical map $\mathcal{U}_\lambda\mathfrak{g} \to \Gamma\,\mathrm{L}\,\mathcal{U}_\lambda\mathfrak{g}$ is an isomorphism, and hence $A \to \Gamma\,\mathrm{L}\,A$ is an isomorphism whenever F is free. Take $F \in \mathcal{U}_\lambda\mathfrak{g}$-MOD, and construct a free resolution of it:

$$\cdots \to A_2 \to A_1 \to A_0 \to 0 \tag{9.5.3.15}$$

and apply L:

$$\cdots \to \mathrm{L}\,A_2 \to \mathrm{L}\,A_1 \to \mathrm{L}\,A_0 \to 0 \tag{9.5.3.16}$$

and apply Γ:

$$\cdots \to \Gamma\,\mathrm{L}\,A_1 \to \Gamma\,\mathrm{L}\,A_1 \to \Gamma\,\mathrm{L}\,A_0 \to 0 \tag{9.5.3.17}$$

As equations (9.5.3.15) and (9.5.3.17) are isomorphic, the only cohomology is in the last spot. On the other hand, Lemma 9.5.3.13 shows that $\Gamma : \mathscr{D}_X^\lambda$-MOD $\to \mathcal{U}_\lambda\mathfrak{g}$-MOD is exact and faithful (being an exact functor that does not take non-zero objects to zero). Therefore the only cohomology of equation (9.5.3.16) is in the last spot. Since L is also exact, it follows that $\Gamma\,\mathrm{L}\,F \cong F$ for any $F \in \mathcal{U}_\lambda\mathfrak{g}$-MOD.

In particular, the canonical map $\mathrm{L}\,\Gamma\mathcal{F} \to \mathcal{F}$ gives rise to an isomorphism $\Gamma\,\mathrm{L}\,\Gamma\mathcal{F} \to \Gamma F$. We claim that $\mathrm{L}\,\Gamma\mathcal{F} \to \mathcal{F}$ is an isomorphism. Indeed, look at the corresponding exact sequence:

$$0 \to \mathcal{X} \to \mathrm{L}\,\Gamma\mathcal{F} \to \mathcal{F} \to \mathcal{Y} \to 0$$

Applying Γ gives

$$0 \to \Gamma\mathcal{X} \to \Gamma\,\mathrm{L}\,\Gamma\mathcal{F} \xrightarrow{\sim} \Gamma\mathcal{F} \to \Gamma\mathcal{Y} \to 0$$

Since Γ is exact, $\Gamma\mathcal{X} = 0 = \Gamma\mathcal{Y}$, and since Γ is faithful, this implies that $\mathrm{L}\,\Gamma\mathcal{F} \to \mathcal{F}$ is an iso. □

We will give two applications of Theorem 9.5.3.11. It has many others, mostly via applying the theory of D-modules to questions in representation theory.

9.5.3.18 Corollary *Let $\mu \in \mathbf{P}^+$. Then the* translation functors

$$\mathcal{D}_X^\lambda\text{-MOD} \xrightarrow{\otimes\mathcal{O}(\mu)} \mathcal{D}_X^{\lambda+\mu}\text{-MOD}$$

is an equivalence of categories. When $\lambda + \rho$ is regular dominant, the corresponding equivalence

$$\mathcal{U}_\lambda\mathfrak{g}\text{-MOD} \to \mathcal{U}_{\lambda+\mu}\mathfrak{g}\text{-MOD}$$

is the translation principle*.* □

9.5.3.19 Example If λ is itself integral dominant, then there is an equivalence $\Phi : \mathcal{U}_\lambda\mathfrak{g}\text{-MOD} \to \mathcal{U}_0\mathfrak{g}\text{-MOD}$. We can construct Φ^{-1} by hand, via $M \mapsto \big(M \otimes L(\lambda)\big)/(\ker \chi_\lambda)$. Reading off its adjoint, we have $\Phi : M \mapsto \big(M \otimes L(\lambda)^*\big)/(\ker \chi_0)$. In particular, finite-dimensionality is preserved. ♢

Our final application explores the resolution of Bernstein, Gelfand, and Gelfand, which we mentioned in Remark 6.1.2.4.

9.5.3.20 Example The following is a resolution of $L(0)$:

$$0 \to M(-2\rho) \to \cdots \to \bigoplus_{\ell(w)=k} M(w(\rho) - \rho) \to \cdots \to M(0) \to 0$$

Incidentally, $w(\rho) - \rho = \sum_{\alpha \in \Delta^- \cap w(\Delta^+} \alpha$.

Take $G/B \supseteq U_0 = N^- \cdot x$, where $\mathrm{Stab}_G\, x = B$ and $N^- \cong \mathfrak{n}^-$, and consider the de Rham complex of U_0:

$$0 \to \Omega^0(\mathcal{U}_0) \xrightarrow{d} \Omega^1(\mathcal{U}_0) \xrightarrow{d} \dots \xrightarrow{d} \Omega^\ell(U_0) \to 0$$

In fact, $\Omega^k(\mathcal{U}_0) \cong \mathcal{S}^\bullet(\mathfrak{n}^-)^* \otimes \bigwedge^k(\mathfrak{n}^-)^*$ is an isomorphism of $\mathfrak{g}$-modules, and the $\mathfrak{g}$-action commutes with the differential. So the cohomology groups are all $\mathfrak{g}$-modules. Taking the restricted dual, we set $M_k \overset{\text{def}}{=} \Omega^k(\mathcal{U}_0)^* = \mathcal{U}\mathfrak{g} \otimes_{\mathcal{U}\mathfrak{b}} \bigwedge^k(\mathfrak{n}^-)$, thereby building a complex

$$0 \to M_\ell \to \cdots \to M_2 \to M_1 \to M_0 \to 0$$

Upon quotienting, we have $M_k/(\ker \chi_0) = \bigoplus_{\ell(w)=l} M(w(\rho) - \rho)$, so that the BGG resolution is a dual to the de Rham complex. ♢

9.5.4 Kostant theorem

Let $\mathfrak{g}$ be a semisimple Lie algebra over $\mathbb{C}$ with chosen triangular decomposition $\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}^+$, and let M be a finite-dimensional $\mathfrak{g}$-module. In this section we will describe $\mathrm{H}^i(\mathfrak{n}^+, M)$. This is the cohomology of the chain complex $\bigwedge^i(\mathfrak{n}^+)^* \otimes M$, which carries an $\mathfrak{h}$ action via, on the first piece, the adjoint action, and on the second piece the action by $\mathfrak{h} \hookrightarrow \mathfrak{g}$, and this action commutes with the differential. So $\mathrm{H}^i(\mathfrak{n}^+, M)$ is an $\mathfrak{h}$-module.

For $\lambda \in \mathfrak{h}^*$, we denote by C_λ the one-dimensional $\mathfrak{h}$-module with weight λ.

9.5.4.1 Theorem (Kostant)
$\mathrm{H}^i(\mathfrak{n}^+, L(\lambda)) = \bigoplus_{\ell(w)=i} C_{w(\lambda+\rho)-\rho}$

We will give two proofs. The first proof is based on the BGG resolution (Remark 6.1.2.4 and Example 9.5.3.20), which is a pretty strong result in itself. The second proof uses the Borel-Weil-Bott theorem (Theorem 9.5.1.4).

Proof (1) Recall that the Killing form identifies $(\mathfrak{n}^+)^* \cong \mathfrak{n}^-$. From this, it follows that the homology $\mathrm{H}_i(\mathfrak{n}^-, L(\lambda))$ is the same as the cohomology $\mathrm{H}^i(\mathfrak{n}^+, L(\lambda))$. By Example 9.5.3.20,

$$0 \to M_\ell \to M_{\ell-1} \to \cdots \to M_0 \to 0$$

is a resolution of $L(\lambda)$, where $M_i = \bigoplus_{\ell(w)=i} M(w(\lambda+\rho)-\rho)$. As an $(\mathfrak{h}\oplus\mathfrak{n}^-)$-module, $M(\mu) \cong \mathcal{U}\mathfrak{n}^- \otimes C_\mu$, and in particular it is free over $\mathfrak{n}^-$ (the $\mathfrak{n}^-$ action on C_μ is trivial). Therefore:

$$\mathrm{H}_i(\mathfrak{n}^-, M(\mu)) = \begin{cases} 0 & i > 0 \\ C_\mu & i = 0 \end{cases}$$

Moreover, $\mathrm{H}_i(\mathfrak{n}^-, L(\lambda)) = \mathrm{H}_0(\mathfrak{n}^-, M_i)$, and the result follows. □

Proof (2) We define the category $(\mathfrak{b}, H)$-MOD of *Harish-Chandra modules* to be the full subcategory of the category of $\mathfrak{b}$-modules whose objects are locally nilpotent over $\mathfrak{n}^+$ and semisimple over $\mathfrak{h}$ with weights in $\mathbf{P}$ the weight lattice of $\mathfrak{g}$. Then:

$$\mathrm{H}^i(\mathfrak{n}^+, L(\lambda))_\mu = \mathrm{Ext}^i_{(\mathfrak{b},H)}(C_\mu, L(\lambda))$$

The left-hand side is the weight-μ subspace, and the right-hand side is computed in this category $(\mathfrak{b}, H)$-MOD. The point is that if you take a projective resolution, and take its semisimple part, it's still projective.

Moreover, we claim that there is an equivalence of categories B-MOD $\overset{\sim}{\leftrightarrow}$ $(\mathfrak{b}, H)$-MOD; by B-MOD we mean the category of *algebraic* representations of the affine algebraic group B, i.e. the category of corepresentations of the algebra of polynomial functions (Lemma/Definition 9.3.3.2). In the forward direction, every B-module is in $(\mathfrak{b}, H)$-MOD, and the other direction is exponentiation (since $\mathfrak{n}^+$ is nilpotent, $\exp : \mathfrak{n}^+ \to N$ is algebraic). Thus:

$$\mathrm{H}^i(\mathfrak{n}^+, L(\lambda))_\mu = \mathrm{Ext}^i_B(C_\mu, L(\lambda)) = \mathrm{Ext}^i_B(L(\lambda)^*, C_{-\mu}) \tag{9.5.4.2}$$

We pick an injective resolution of $C_{-\mu}$:

$$0 \to I_0 \to I_1 \to \cdots \to I_\ell \to 0 \tag{9.5.4.3}$$

To compute the right-hand side of equation (9.5.4.2), we apply the functor $\mathrm{Hom}_B(L(\lambda)^*, -)$ to each term in equation (9.5.4.3). Then, using Frobenius reciprocity and dualizing:

$$\begin{aligned}\mathrm{Ext}^i(\mathfrak{n}^+, L(\lambda))_\mu &= \mathrm{Hom}_B(L(\lambda)^*, I_i) = \mathrm{Hom}_G(L(\lambda)^*, \mathrm{Ind}_B^G(M_i)) = \\ &= \mathrm{Hom}_G\big(L(\lambda)^*, H^i(G/B, \mathcal{O}(-\mu))\big) = \mathrm{Hom}_G\big(L(\lambda), H^i(G/B^-, \mathcal{O}(\mu))\big)\end{aligned}$$

The theorem follows from Theorem 9.5.1.4. □

9.5.4.4 Remark The second proof can be run in reverse to give Theorem 9.5.1.4 as a corollary of Theorem 9.5.4.1: originally Kostant proved his theorem using spectral sequences. On the other hand, the existence of a BGG resolution is stronger, because cohomology doesn't know everything.

One can also use Theorem 9.5.4.1 to prove the Weyl character formula (Theorem 6.1.1.2). It is not the quickest way to prove it, but it is not very difficult. ◊

Exercises

Let $\mathfrak{g}$ be a semisimple Lie algebra over $\mathbb{C}$, G its connected simply-connected algebraic group, and B a Borel.

1. Show that the centralizer of any semisimple element in $\mathfrak{g}$ is reductive.

2. Prove the *Jacobson-Morozov Theorem*: if $x \in \mathfrak{g}$ is a nilpotent element, then there exist $h, y \in \mathfrak{g}$ such that $\{x, h, y\}$ form an $\mathfrak{sl}(2)$-triple, i.e. $[h, x] = 2x$, $[h, y] = -2y$, and $[x, y] = h$.

3. Show that any two $\mathfrak{sl}(2)$-triples containing a given x are conjugate by the action of the adjoint group.

4. Show that the nilpotent cone in $\mathfrak{g}$ has finitely many orbits with respect to the adjoint action.

5. Let $\mathfrak{p}$ be a parabolic subalgebra of $\mathfrak{g}$ associated with a subset S of simple roots. Assume that the nilpotent radical of $\mathfrak{p}$ is abelian.

   (a) Show that $S = \{\alpha_i\}$ is a single element set and hence $\mathfrak{p}$ is a maximal proper subalgebra of $\mathfrak{g}$.

   (b) Let $\mathfrak{g}$ be simple. Show that if θ is the highest weight of the adjoint representation and $\theta = \sum m_i \alpha_i$, then $m_i = 1$.

6. Let X denote the G-orbit of the B-invariant line in an irreducible representation $L(\lambda)$. Let Ω denote the Casimir element in $\mathcal{U}(\mathfrak{g})$. Show that any $x \in X$ satisfies the quadratic equations:

   $$\Omega(x \otimes x) = (2\lambda, 2\lambda + 2\rho)(x \otimes x)$$

   This is an analogue of the *Plucker relations*. (Hint: $x \otimes x \in L(2\lambda) \subseteq L(\lambda) \otimes L(\lambda)$.)

7. Let $B \subseteq P \subseteq G$. Show that if $\mathrm{H}^i(P/B, P \times_B V)$ is not zero for one i only, then:

   $$\mathrm{H}^{i+j}(G/B, G \times_B V) \cong \mathrm{H}^j(G/P, G \times_P \mathrm{H}^i(P \times_B V))$$

8. Use the previous exercise and the Borel-Weil-Bott theorem to calculate $\mathrm{H}^i(G/P, G \times_P V)$ for an irreducible P-module V.

9. For an arbitrary Lie algebra $\mathfrak{g}$ show that the first cohomology group $\mathrm{H}^1(\mathfrak{g}; \mathfrak{g})$ with coefficients in the adjoint module is isomorphic to the algebra $\mathrm{Der}(\mathfrak{g})/\operatorname{ad}(\mathfrak{g})$, where $\mathrm{Der}(\mathfrak{g})$ denotes the algebra of derivations of $\mathfrak{g}$.

10. A *Heisenberg algebra* is a 3-dimensional Lie algebra with one-dimensional center that coincides with the commutator of the algebra. Check that a Heisenberg algebra is isomorphic to the subalgebra of strictly upper triangular matrices in $\mathfrak{sl}(3)$ and calculate its cohomology with trivial coefficients. (Hint: you can use the Kostant theorem.)

Part III

Poisson and Quantum Groups

Chapter 10

Lie bialgebras

Welcome to Part III. In it we will study Quantum Groups, which are neither quantum nor groups. They are non-commutative non-cocommutative Hopf algebras, deformations of $\mathcal{U}\mathfrak{g}$ and $\mathscr{C}(G)$. Unless we say otherwise, G will always be an affine algebraic group over $\mathbb{C}$, or a real form thereof. Recall from Chapter 7 that a complex group can have many interesting real forms — for example, $\mathrm{SL}(n,\mathbb{C})$ has both $\mathrm{SL}(n,\mathbb{R})$ and $\mathrm{SU}(n)$. We will be interested in both "complex" and "real" quantum groups. Let us now outline the next few chapters.

Let $\mathfrak{g}$ be a Lie algebra. If there is a complex algebraic group G with $\mathfrak{g} = \mathrm{Lie}(G)$, then to $\mathfrak{g}$ we can associate two Hopf algebras $\mathcal{U}\mathfrak{g}$ and $\mathscr{C}(G)$, and these are dual: Proposition 3.2.4.4 provides a map $\mathcal{U}\mathfrak{g} \otimes \mathscr{C}(G) \to \mathbb{C}$ that plays well with the Hopf structures. Let $\mathfrak{g}^*$ be the dual vector space, and let's fix a Lie algebra structure on $\mathfrak{g}^*$. So we have a pair of algebras, and subject to certain compatibility restrictions, this pair $(\mathfrak{g}, \mathfrak{g}^*)$ will be called a *Lie bialgebra*.

To the bialgebra $(\mathfrak{g}, \mathfrak{g}^*)$, then $\mathfrak{g}$ gives us a connected simply-connected Lie group G, and $\mathfrak{g}^*$ gives us another (connected simply-connected) group G^*. So we get a *dual pair of Lie groups* G and G^*, and out of this we can construct, assuming everything is algebraic, a pair of Hopf algebras $\mathcal{U}\mathfrak{g}$ and $\mathscr{C}(G)$, and also the pair $\mathcal{U}\mathfrak{g}^*$ and $\mathscr{C}(G^*)$. Each is a pair of dual Hopf algebras, and the pairs are dual to each other in a different sense. Then we will have corresponding quantum groups $\mathcal{U}_q(\mathfrak{g})$ and $\mathscr{C}_q(G^*)$, deformations in the category of Hopf algebras of $\mathcal{U}(\mathfrak{g})$ and $\mathscr{C}(G)$, and also $\mathcal{U}_q(\mathfrak{g}^*)$ and $\mathscr{C}_q(G)$. In fact, the Hopf algebras $\mathcal{U}_q(\mathfrak{g})$ and $\mathscr{C}_q(G^*)$ are more or less the same — they are the same algebraically, but the topology is different. So the slogan is that "after quantization, there is no difference between Universal Enveloping Algebra and Algebra of Functions."

The general notion of quantization first appeared in physics, and then filtered to mathematics and eventually representation theory. The idea is that given a *symplectic manifold* (M, ω), and maybe using extra data, you can construct a family of associative algebras $A_\hbar$, but the center of $A_\hbar$ is usually trivial $(\mathbb{C} \cdot 1)$. But to a *Poisson manifold* (P, p), the family, which exists at least formally, can be very interesting. And there is a notion of *symplectic leaves*. We will study the case when the Poisson manifold is equipped with a compatible group structure, and call it a *Poisson Lie group*. Why "Poisson Lie" and not "Lie Poisson" you'll have to ask Drinfeld. Probably because "Lie group" sounds like one word. Then there is a general philosophy which is hard to formulate precisely, that to symplectic leaves we should associate irreducible representations.

10.1 Basic definitions

In Lie theory, we normally introduce the more natural notion of *Lie group*, then define *Lie algebra* by noticing that the tangent space at the identity has some natural structure. But we will go in the opposite direction, so to avoid having to know what a Poisson manifold is:

10.1.1 Lie bialgebras

For now, we consider only finite-dimensional Lie algebras over $\mathbb{C}$.

10.1.1.1 Definition *A pair* $(\mathfrak{g}, \delta : \mathfrak{g} \to \bigwedge^2 \mathfrak{g})$ *is a* Lie bialgebra *if* $\mathfrak{g}$ *is a Lie algebra and* δ *satisfies:*

1. δ *is a* Lie cobracket*. We can understand this condition in two ways: either that* $\delta^* : \bigwedge^2 \mathfrak{g}^* \to \mathfrak{g}^*$ *is a Lie bracket, and also by the* co-Jacobi identity*:*

$$\mathrm{Alt}(\delta \otimes \mathrm{id}) \circ \delta = 0$$

2. *a compatibility condition:*

$$\delta([x, y]) = [x, \delta(y)] + [\delta(x), y] \tag{10.1.1.2}$$

This is a cocycle *property of* δ*. We have extended the bracket from* $\mathfrak{g}$ *to wedge powers of* $\mathfrak{g}$ *by declaring on pure tensors that* $[x, y \wedge z] \stackrel{\text{def}}{=} [x, y] \wedge z + y \wedge [x, z]$.

10.1.1.3 Example Let $\mathfrak{b}_+ = \mathbb{C}H \oplus \mathbb{C}X$ with $[H, X] = 2X$ denote the upper Borel in $\mathfrak{sl}(2)$. Then you can check that $\delta(H) = 0$, $\delta(X) = H \wedge X$ makes $\mathfrak{b}_+$ into a Lie bialgebra. ♢

10.1.1.4 Example (Standard structure on $\mathfrak{sl}(2)$) The Lie algebra $\mathfrak{sl}(2, \mathbb{C})$ is spanned by X, Y, H with relations $[X, Y] = H$, $[H, X] = 2X$, and $[H, Y] = -2Y$. Its *standard structure as a Lie bialgebra* is $\delta(H) = 0$, $\delta(X) = H \wedge X$, and $\delta(Y) = H \wedge Y$. ♢

Recall Corollary 4.4.4.5:

10.1.1.5 Definition *The* Chevalley complex *or* BRST complex *for a Lie algebra is the complex* $\mathrm{C}^\bullet(\mathfrak{g}, M) = \mathrm{Hom}_{\mathbb{C}}(\bigwedge^\bullet \mathfrak{g}, M)$*, where* M *is a* $\mathfrak{g}$*-module and* $\mathrm{Hom}_{\mathbb{C}}$ *means all linear maps. Then the differential* $\mathrm{d} : \mathrm{C}^n(\mathfrak{g}, M) \to \mathrm{C}^{n+1}(\mathfrak{g}, M)$ *is given by*

$$\begin{aligned} \mathrm{d}f(x_1, \dots, x_{n+1}) = &\sum_{i<j} (-1)^{i+j-1} f([x_i, x_j], x_1, \dots, \hat{x_i}, \dots, \hat{x_j}, \dots, x_{n+1}) \\ &+ \sum_i (-1)^i x_i \cdot f(x_1, \dots, \hat{x_i}, \dots, x_{n+1}). \end{aligned}$$

Here $\cdot$ *is the action of* $\mathfrak{g}$ *on* M*.*

10.1.1.6 Example When $M = \mathbb{C}$, then $\mathrm{C}^\bullet \mathfrak{g} \stackrel{\text{def}}{=} \mathrm{C}^\bullet(\mathfrak{g}, \mathbb{C}) \cong \bigwedge^\bullet \mathfrak{g}^*$ as a graded vector space. ♢

10.1.1.7 Remark Consider the bigraded vector space $\left(\bigwedge^\bullet(\mathfrak{g}\oplus\mathfrak{g}^*)\right)^* = \bigwedge^\bullet(\mathfrak{g}\oplus\mathfrak{g}^*) \cong \bigwedge^\bullet\mathfrak{g}\otimes\bigwedge^\bullet\mathfrak{g}^*$. The nth row can be equipped with differentials making it equal to the Chevalley complex with $M = \bigwedge^n\mathfrak{g}$. But now let's say we had a bialgebra structure. Then we also have vertical maps from the cohomology of $\mathfrak{g}^*$.

$$\begin{tikzcd}
\mathbb{C} \arrow[r, "d_{\mathfrak{g}}"] \arrow[d, squiggly, "d_{\mathfrak{g}^*}"] & \mathfrak{g}^* \arrow[r, "d_{\mathfrak{g}}"] \arrow[d, squiggly, "d_{\mathfrak{g}^*}"] & \bigwedge^2\mathfrak{g}^* \arrow[r, "d_{\mathfrak{g}}"] \arrow[d, squiggly, "d_{\mathfrak{g}^*}"] & \bigwedge^3\mathfrak{g}^* \arrow[r, "d_{\mathfrak{g}}"] \arrow[d, squiggly, "d_{\mathfrak{g}^*}"] & \cdots & M=\mathbb{C} \\
\mathfrak{g} \arrow[r, "d_{\mathfrak{g}}"] \arrow[d, squiggly, "d_{\mathfrak{g}^*}"] & \mathfrak{g}\otimes\mathfrak{g}^* \arrow[r, "d_{\mathfrak{g}}"] \arrow[d, squiggly, "d_{\mathfrak{g}^*}"] & \mathfrak{g}\otimes\bigwedge^2\mathfrak{g}^* \arrow[r, "d_{\mathfrak{g}}"] \arrow[d, squiggly, "d_{\mathfrak{g}^*}"] & \mathfrak{g}\otimes\bigwedge^3\mathfrak{g}^* \arrow[r, "d_{\mathfrak{g}}"] \arrow[d, squiggly, "d_{\mathfrak{g}^*}"] & \cdots & M=\mathfrak{g} \\
\bigwedge^2\mathfrak{g} \arrow[r, "d_{\mathfrak{g}}"] \arrow[d, squiggly, "d_{\mathfrak{g}^*}"] & \bigwedge^2\mathfrak{g}\otimes\mathfrak{g}^* \arrow[r, "d_{\mathfrak{g}}"] \arrow[d, squiggly, "d_{\mathfrak{g}^*}"] & \bigwedge^2\mathfrak{g}\otimes\bigwedge^2\mathfrak{g}^* \arrow[r, "d_{\mathfrak{g}}"] \arrow[d, squiggly, "d_{\mathfrak{g}^*}"] & \bigwedge^2\mathfrak{g}\otimes\bigwedge^3\mathfrak{g}^* \arrow[r, "d_{\mathfrak{g}}"] \arrow[d, squiggly, "d_{\mathfrak{g}^*}"] & \cdots & M=\bigwedge^2\mathfrak{g} \\
\bigwedge^3\mathfrak{g} \arrow[r, "d_{\mathfrak{g}}"] \arrow[d, squiggly, "d_{\mathfrak{g}^*}"] & \bigwedge^3\mathfrak{g}\otimes\mathfrak{g}^* \arrow[r, "d_{\mathfrak{g}}"] \arrow[d, squiggly, "d_{\mathfrak{g}^*}"] & \bigwedge^3\mathfrak{g}\otimes\bigwedge^2\mathfrak{g}^* \arrow[r, "d_{\mathfrak{g}}"] \arrow[d, squiggly, "d_{\mathfrak{g}^*}"] & \bigwedge^3\mathfrak{g}\otimes\bigwedge^3\mathfrak{g}^* \arrow[r, "d_{\mathfrak{g}}"] \arrow[d, squiggly, "d_{\mathfrak{g}^*}"] & \cdots & M=\bigwedge^3\mathfrak{g} \\
\vdots & \vdots & \vdots & \vdots & &
\end{tikzcd}$$

The bracket $[,]:\bigwedge^2\mathfrak{g}\to\mathfrak{g}$ gives the differential $d_{\mathfrak{g}}$, and Jacobi is equivalent to $d^2=0$. If we also have a cobracket with coJacobi, then we get the vertical maps $d_{\mathfrak{g}^*}$. Do we have a bicomplex, i.e. do the differentials commute? (They never actually commute; do they anticommute?) The answer is that $d_{\mathfrak{g}}d_{\mathfrak{g}^*}+d_{\mathfrak{g}^*}d_g=0$ exactly if δ dual to the cobracket gives a Lie bialgebra structure for $\mathfrak{g}$. $\Diamond$

10.1.2 Poisson algebras

10.1.2.1 Definition *A* Poisson algebra *is a mixture of a Lie algebra and a commutative algebra: it is a pair* $(A,\{,\})$ *such that*

- *A is a commutative (unital, associative) algebra.*
- *$\{,\}:A^{\otimes 2}\to A$ is a Lie algebra bracket.*
- *$\{,\}$ is a biderivation: $\{a,bc\}=\{a,b\}c+b\{a,c\}$.*

The main reason Poisson algebras are so important is that they describe infinitesimal jets of deformations of commutative algebra. Indeed, consider a family of associative multiplications $\star_\hbar : A^{\otimes 2}\to A$ such that

- $a\star_0 b = ab$ is commutatiive.
- the multiplication is given by an analytic function in $\hbar$: $a\star_\hbar b = ab+\hbar m_1(a,b)+\hbar^2 m_2(a,b)+O(\hbar^3)$ as $\hbar\to 0$. (To say this precisely we must introduce some topology on A.)

10.1.2.2 Proposition *$\{a,b\}\overset{\text{def}}{=} m_1(a,b)-m_1(b,a)$ is a Poisson structure on a commutative algebra A.*

10.1.2.3 Remark Since this is a deformation, it is in a sense "quantum", and it is in this sense that Quantum Groups are quantum. $\Diamond$

10.1.2.4 Example Our primary examples of Poisson algebras come from *Poisson manifolds*. A *Poisson manifold* is a pair (M, p) where $p \in \Gamma(\bigwedge^2 \mathrm{T}M)$ and M is a manifold ($\mathscr{C}^\infty$, affine algebraic variety, etc.; and whatever M is, we take the appropriate type of section p). Then p is a *bivector field*, and in local coordinates $p(x) = \sum_{ij} p^{ij} \frac{\partial}{\partial x^i} \wedge \frac{\partial}{\partial x^j}$. Any time we choose local coordinates $\{x^i\}$ on M, then we let $\mathrm{d}x^i$ be the corresponding basis of $\mathrm{T}^*_x M$ and $\frac{\partial}{\partial x^i}$ the basis in $\mathrm{T}_x M$.

Then we define $\{f, g\} \stackrel{\text{def}}{=} \langle p, \mathrm{d}f \wedge \mathrm{d}g \rangle = \sum_{ij} p^{ij}(x) \frac{\partial f}{\partial x^i} \wedge \frac{\partial g}{\partial x^j}$, and we demand that $\{,\}$ is a Poisson bracket on $\mathscr{C}(M)$ (the space of $\mathscr{C}^\infty$ or polynomial or whatever functions). That $\{,\}$ is a biderivation is trivial, since everything is a first-order operator. That it satisfies the Jacobi identity requires p to satisfy a non-trivial condition, which in local coordinates can be written as:

$$0 = \sum_\ell \left(\frac{\partial p^{ij}(x)}{\partial x^\ell} p^{\ell k}(x) + \frac{\partial p^{jk}(x)}{\partial x^\ell} p^{\ell i}(x) + \frac{\partial p^{ki}(x)}{\partial x^\ell} p^{\ell j}(x) \right), \quad \forall i, j, k.$$

$\Diamond$

10.1.2.5 Example Let $\mathfrak{g}$ be a finite-dimensional Lie algebra and $\mathfrak{g}^*$ its dual vector space. Since these are vector spaces, we have a canonical isomorphism $\mathrm{T}^*_x \mathfrak{g}^* \cong (\mathfrak{g}^*)^* = \mathfrak{g}$ for each $x \in \mathfrak{g}^*$. Thus, for any $f \in \mathscr{C}(\mathfrak{g}^*)$ and $x \in \mathfrak{g}^*$, then $\mathrm{d}f(x) \in \mathfrak{g}$. So if $f, g \in \mathscr{C}(\mathfrak{g}^*)$, we can define $[f(x), g(x)]$, and we define the Poisson bracket by:

$$\{f, g\}(x) \stackrel{\text{def}}{=} \langle x, [\mathrm{d}f(x), \mathrm{d}g(x)] \rangle .$$

This is the *Lie-Kirollov-Kostant bracket*. It was discovered by Lie, and used to study the representation theory of Lie algebras.

It's probably more instructive to see this is local coordinates. Let $\{e_i\}$ be a basis in $\mathfrak{g}$, and $\{x_i\}$ the corresponding coordinate functions on $\mathfrak{g}^*$. Then

$$\{x_i, x_j\} = \sum_k f_{ij}^k x_k.$$

where f_{ij}^k are the *structure constants* of $\mathfrak{g}$, defined by the equation $[e_i, e_j] = \sum_k f_{ij}^k e_k$. $\Diamond$

10.1.2.6 Definition *Let (M_1, p_1) and (M_2, p_2) be Poisson manifolds. Their* product *is $(M_1 \times M_2, p_{12})$ where p_{12} is the sum of p_1 and p_2. More explicitly, $\mathrm{T}_{(x,y)}(M_1 \times M_2) = \mathrm{T}_x M_1 \oplus \mathrm{T}_y M_2$, so $\bigwedge^2 \mathrm{T}_{(x,y)}(M_1 \times M_2) = \bigwedge^2 \mathrm{T}_x M_1 \oplus (\mathrm{T}_x M_1 \otimes \mathrm{T}_y M_2) \oplus \bigwedge^2 \mathrm{T}_y M_2$, and we define $p_{12}(x, y) \stackrel{\text{def}}{=} p_1(x) \oplus 0 \oplus p_2(y)$.*

10.1.2.7 Lemma / Definition *There is a natural embedding $\mathscr{C}(M_1) \otimes \mathscr{C}(M_2) \hookrightarrow \mathscr{C}(M_1 \times M_2)$, which is an isomorphism if M_1 and M_2 are affine algebraic. For this embedding and the product of Poisson manifolds define above, we have:*

$$\{f_1 \otimes f_2, g_1 \otimes g_2\} = \{f_1, g_1\} \otimes f_2 g_2 + f_1 g_1 \otimes \{f_2, g_2\} \tag{10.1.2.8}$$

In general, if A_1 and A_2 are Poisson algebras, their tensor product *is the commutative algebra $A_1 \otimes_k A_2$ equipped with the bracket defined by 10.1.2.8.*

10.1.2.9 Definition *If A_1 and A_2 are two Poisson algebra, then $\phi : A_1 \to A_2$ is a* morphism of Poisson algebras *if $\phi(ab) = \phi(a)\phi(b)$ and $\phi(\{a,b\}) = \{\phi(a),\phi(b)\}$. We define a map $\phi : (P_1, p_1) \to (P_2, p_2)$ to be a* morphism of Poisson manifolds *if ψ is a manifold map and the pullback ψ^* is a morphism of Poisson algebras.*

10.1.3 Definition of Poisson Lie group

10.1.3.1 Definition *A* Poisson Lie group *is a Lie group G along with a Poisson structure p on the underlying manifold of G, such that the multiplication map $G \times G \to G$ is a Poisson map.*

If G is a group, then the tangent bundle $\mathrm{T}G$ is trivial: $\mathrm{T}G \cong \mathfrak{g} \times G$, and we will always choose the trivialization by left translations. I.e. $\ell_g : G \to G$ is $x \mapsto gx$. Then $\mathrm{d}\ell_g : \mathrm{T}G \overset{\sim}{\to} \mathrm{T}G$. It takes $\mathrm{T}_h G \to \mathrm{T}_{gh} G$, and in particular $\mathrm{d}\ell_{h^{-1}} : \mathrm{T}_h \overset{\sim}{\to} \mathrm{T}_e G = \mathfrak{g}$. So $\mathrm{d}\ell : \mathrm{T}G \overset{\sim}{\to} \mathfrak{g} \times G$ by $(\xi, h) \mapsto (\mathrm{d}\ell_{h^{-1}}(\xi), h)$.

If G is a Poisson Lie group, its Poisson structure is a section $p \in \Gamma(\bigwedge^2 \mathrm{T}G)$, but by trivialization this is equivalent to a map $p : G \to \bigwedge^2 \mathfrak{g}$.

10.1.3.2 Theorem (The Lie algebra of a Poisson Lie group is a Lie bialgebra) *Let (G,p) be a Poisson Lie group. Writing $e \in G$ for the identity, we have $p(e) = 0$. Trivialize $\mathrm{T}G = \mathfrak{g} \times G$ by left translations, so that $p : G \to \bigwedge^2 \mathfrak{g}$. Then $\delta = \mathrm{d}p(e) : \mathfrak{g} = \mathrm{T}_e G \to \mathrm{T}_0(\bigwedge^2 \mathfrak{g}) = \bigwedge^2 \mathfrak{g}$ is the cobracket for a Lie bialgebra structure on $\mathfrak{g}$.*

Proof In our trivialization, the condition that the multiplication map be Poisson is equivalent to the request that:

$$p(xy) = p(x) + (\mathrm{Ad}_x \otimes \mathrm{Ad}_x) p(y), \quad \forall x, y \in G. \tag{10.1.3.3}$$

In particular, $p(e^2) = p(e) + p(e)$, and so $p(e) = 0$. In order for $\delta = \mathrm{d}p$ to be the cobracket of a Lie bialgebra, we must check two conditions:

1. **The Jacobi identity:** Something will satisfy Jacobi, since p did; we need to check that the dual map $[,]_{\mathfrak{g}^*} \overset{\text{def}}{=} \delta^* : \mathfrak{g}^* \wedge \mathfrak{g}^* \to \mathfrak{g}^*$ is this something. Fix $f_1, f_2 \in \mathscr{C}(G)$, and their differentials $\mathrm{d}f_1(e), \mathrm{d}f_2(e) \in \mathrm{T}_e^* G = \mathfrak{g}^*$. Let's denote $\mathrm{d}f_i(e)$ by ξ_i. Then, letting $X \in \mathfrak{g}$, we have

$$\begin{aligned} [\xi_1, \xi_2]_{\mathfrak{g}^*}(X) &= \langle \mathrm{d}p(e)(X), \mathrm{d}f_1(e) \wedge \mathrm{d}f_2(e) \rangle \\ &= \left. \frac{\mathrm{d}}{\mathrm{d}t} \right|_{t=0} \{f_1, f_2\}(\exp(tX)). \end{aligned} \tag{10.1.3.4}$$

   Here $\{,\}$ is our Poisson bracket on G, and $\exp : \mathfrak{g} \to G$ is the exponential map.

   On the other hand, $\{f_1, f_2\}(e) = 0$. The Jacobi for $\{,\}$ says that $\{f_1, \{f_2, f_3\}\}(\exp(tX)) + \text{cyclic} = 0$. So take $\frac{\mathrm{d}^2}{\mathrm{d}t^2} \left[\{f_1, \{f_2, f_3\}\}(\exp(tX)) + \text{cyclic}\right]\Big|_{t=0}$ and conclude the Jacobi for $[,]_{\mathfrak{g}^*}$.

2. **The cocycle property:** The one-line proof reads "p is a 1-cocycle for G by equation (10.1.3.3), and so automatically induces a 1-cocycle for $\mathrm{Lie}\, G$." We proceed to prove this.

   We apply equation (10.1.3.3) twice, on a commutator:

$$p(y^{-1}zy) = p(y^{-1}) + \left(\mathrm{Ad}_{y^{-1}} \otimes \mathrm{Ad}_{y^{-1}}\right) p(z) + \left(\mathrm{Ad}_{y^{-1}z} \otimes \mathrm{Ad}_{y^{-1}z}\right) p(y) \tag{10.1.3.5}$$

Setting $z = \exp(tX)$ and differentiating in t at $t = 0$:

$$\delta\left(\mathrm{Ad}_{y^{-1}}(X)\right) = \left(\mathrm{Ad}_{y^{-1}} \otimes \mathrm{Ad}_{y^{-1}}\right)[X, p(y)] + \left(\mathrm{Ad}_{y^{-1}} \otimes \mathrm{Ad}_{y^{-1}}\right)\delta(X)$$

We've used that $\delta = \mathrm{d}p(e)$. Now we take $y = e^{tY}$ and differentiate:

$$\delta([X,Y]) = [X, \delta(Y)] - [Y, \delta(X)]$$

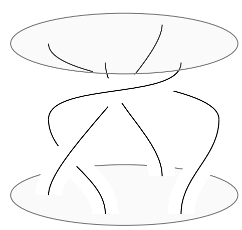
Here we used that $[X, Y \wedge Z] = [X,Y] \wedge Z + Y \wedge [X,Z]$. □

10.1.3.6 Theorem (Lie III for bialgebras)
Let $(\mathfrak{g}, \delta)$ be a Lie bialgebra. Then the connected simply-connected Lie group G with $\mathrm{Lie}(G) = \mathfrak{g}$ has a unique Poisson structure p making it into a Poisson Lie group for which $\delta = \mathrm{d}p(e)$.

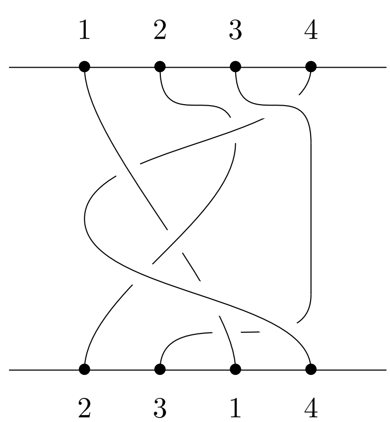
10.2 Braids and the classical Yang–Baxter equation

10.2.1 Braid groups

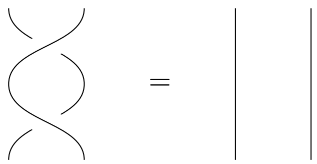
Let $X_n \stackrel{\text{def}}{=} \left\{(x_1, \dots, x_n) : x_i \neq x_j, x_i \in \mathbb{R}^2\right\}$. Then S_n acts on X_n, and we define $\tilde{X}_n \stackrel{\text{def}}{=} X_n/S_n$. The *braid group* $B_n \stackrel{\text{def}}{=} \pi_1(\tilde{X}_n)$ is the fundamental group of this space. We can draw pictures of elements in B_n as follows:

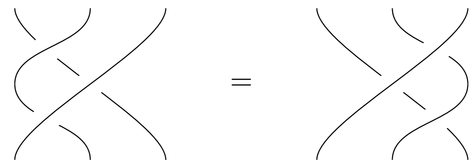


A more standard drawing is obtained by picking the points to lie on the line, and project all paths to the plane, recording overcrossings and undercrossings.



(This is not the same braid group element as the three-dimensional picture.) Since we are working with π_1, when we draw paths we mean to take them up to isotopy. The *Reidemeister moves* are the following and their mirror images:





We do not include Reidemeister 1 because we are working only with braids.

10.2.1.1 Theorem (Braid group presentation)
The braid group has the following presentation:

$$B_n \cong \langle s_i, i = 1, \dots, n-1 \text{ s.t. } s_i s_j = s_j s_i \text{ if } |i-j| > 1, \text{ and } s_i s_{i+1} s_i = s_{i+1} s_i s_{i+1} \rangle$$

The generator s_i corresponds to the braid that is trivial on all strands for a single crossing between strands i and $i+1$. □

10.2.1.2 Corollary *Let V be a vector space. Given a matrix $S \in \mathrm{Aut}(V \otimes V)$, consider assigning to each generator $s_1, \dots, s_{n-1}$ of B_n the map*

$$s_i = 1 \otimes \cdots \otimes S \otimes \cdots \otimes 1 : V^{\otimes n} \to V^{\otimes n},$$

where S acts in the i and $i+1$ spots. This assignment extends to a representation of B_n on $V^{\otimes n}$ if and only if S satisfies the Yang–Baxter equation*:*

$$(S \otimes 1)(1 \otimes S)(S \otimes 1) = (1 \otimes S)(S \otimes 1)(1 \otimes S). \tag{10.2.1.3}$$

The whole reason for developing quantum groups, Poisson Lie groups, etc., was to study these equations. Except that this didn't evolve in Topology, but rather in Statistical Mechanics.

Equation (10.2.1.3) is a hugely over-determined system: there are $(\dim V)^6$ equations for $(\dim V)^4$ unknowns. There is a trivial solution, namely $S = P : x \otimes y \mapsto y \otimes x$. This solution is blind to whether the crossing is "over" or "under". To look for interesting solutions, we try to construct a family of solutions $S(\hbar) = P \circ (1 + \hbar r + O(\hbar^2))$.

10.2.1.4 Proposition *Such $S = P \circ (1 + \hbar r + O(\hbar^2))$ satisfies the Yang–Baxter equation only if r satisfies the* classical Yang–Baxter equation*:*

$$\mathrm{CYB}(r) \stackrel{\text{def}}{=} [r_{12}, r_{13}] + [r_{12}, r_{23}] + [r_{13}, r_{23}] = 0. \tag{10.2.1.5}$$

Proof Expand equation (10.2.1.3) to order $\hbar^2$; the order-$\hbar$ stuff cancels. □

This is an equation that involves only commutators. We should consider it as an equation in $\mathfrak{gl}(V)^{\otimes 3}$ for $r \in \mathfrak{gl}(V)^{\otimes 2}$.

Recall Theorem 4.5.0.10 that every finite-dimensional Lie algebra is a subalgebra of $\mathfrak{gl}(V)$ for some finite-dimensional V. So finding solutions to equation (10.2.1.5) is the same as classifying all solutions in an arbitrary finite-dimensional Lie algebra. We will see that Lie bialgebras are a source of solutions to equation (10.2.1.5). This is why we are interested in Lie bialgebras if we are interested in knot theory.

The general questions are:

1. How to construct solutions to equation (10.2.1.5)? I.e. how to construct Lie bialgebras?
2. "Quantization": How to construct S for a given r? The answer is in the construction of a special class of quantum groups.

This second question is the historical motivation for our subject. Similarly, the main motivation for Lie was to study the solutions of differential equations. This history is almost completely forgotten.

10.2.2 Quasitriangular Lie bialgebras

10.2.2.1 Definition *Let $\mathfrak{g}$ be a Lie algebra. A* classical R-matrix *is an element $r \in \mathfrak{g} \otimes \mathfrak{g}$ satisfying the* classical Yang–Baxter equation*:*

$$\mathrm{CYB}(r) \overset{\mathrm{def}}{=} [r_{12}, r_{13}] + [r_{12}, r_{23}] + [r_{13}, r_{23}] = 0 \tag{10.2.2.2}$$

This equation lives in $\mathcal{U}(\mathfrak{g})^{\otimes 3}$. We set $r_{12} \overset{\mathrm{def}}{=} r \otimes 1$, where we have embedded $\mathfrak{g} \otimes \mathfrak{g} \hookrightarrow \mathcal{U}\mathfrak{g} \otimes \mathcal{U}\mathfrak{g}$, and $r_{23} \overset{\mathrm{def}}{=} 1 \otimes r$. We leave it as an exercise to guess r_{13}.

10.2.2.3 Remark In terms of some basis e_i of $\mathfrak{g}$, we have $r = \sum_{ij} r^{ij} e_i \otimes e_j$. So r has $(\dim \mathfrak{g})^2$ variables, and equation (10.2.2.2) is $(\dim \mathfrak{g})^3$ equations, so it's entirely nonobvious why there would be any solutions to this equation. But, indeed, the "Drinfeld double construction" says there are some. ◊

10.2.2.4 Example Let $\mathfrak{g} = \mathfrak{sl}(2)$ with standard basis H, X, Y with $[H, X] = 2X$, $[H, Y] = -2Y$, and $[X, Y] = H$. Then $r = \frac{1}{4} H \otimes H + X \otimes Y$ is a classical R-matrix. ◊

Let $\mathfrak{g}$ be a Lie algebra. Suppose that $r \in \mathfrak{g} \otimes \mathfrak{g}$ satisfies equation (10.2.2.2). We consider $r_\pm : \mathfrak{g}^* \to \mathfrak{g}$ given by

$$r_+(l) \overset{\mathrm{def}}{=} (l \otimes \mathrm{id}) r \tag{10.2.2.5}$$

$$r_-(l) \overset{\mathrm{def}}{=} -(\mathrm{id} \otimes l) r \tag{10.2.2.6}$$

The minus sign is for later convenience. We set $\mathfrak{g}_\pm \overset{\mathrm{def}}{=} \mathrm{Im}(r_\pm) \subseteq \mathfrak{g}$.

10.2.2.7 Lemma *With the notation as above, $\mathfrak{g}_\pm$ and $\tilde{\mathfrak{g}} \overset{\mathrm{def}}{=} \mathfrak{g}_+ + \mathfrak{g}_-$ are Lie subalgebras of $\mathfrak{g}$.*

10.2.2.8 Example Let $\mathfrak{g} = \mathfrak{sl}(2, \mathbb{C}) = \mathbb{C}^3$ with the basis $\{H, X, Y\}$, and $\mathfrak{g}^* = \mathbb{C}^3$ with the dual bases $\{H^\vee, X^\vee, Y^\vee\}$. We choose $r = \frac{1}{4} H \otimes H + X \otimes Y$. Then $r_+(l) = l(H)\frac{H}{4} + l(X) Y$. Letting l vary over all of $\mathfrak{g}^*$ we see that $\mathrm{Im}(r_+) = \mathbb{C}H \oplus \mathbb{C}Y$ is the lower Borel in $\mathfrak{sl}(2)$. Similarly, $\mathfrak{g}_-$ is the upper Borel. ◊

Proof (of Lemma 10.2.2.7) By standard linear algebra $r \in \mathfrak{g}_- \otimes \mathfrak{g}_+$. Then look at equation (10.2.2.2):

$$\underset{[\mathfrak{g}_-,\mathfrak{g}_-]\otimes\mathfrak{g}_+\otimes\mathfrak{g}_+}{\underset{\Cap}{[r_{12},r_{13}]}} + \underset{\mathfrak{g}_-\otimes[\mathfrak{g}_+,\mathfrak{g}_-]\otimes\mathfrak{g}_+}{\underset{\Cap}{[r_{12},r_{23}]}} + \underset{\mathfrak{g}_-\otimes\mathfrak{g}_-\otimes[\mathfrak{g}_+,\mathfrak{g}_+]}{\underset{\Cap}{[r_{13},r_{23}]}} = 0$$

So this is only possible if $[\mathfrak{g}_-, \mathfrak{g}_-] \subseteq \mathfrak{g}_-$, $[\mathfrak{g}_+, \mathfrak{g}_-] \subseteq \mathfrak{g}_+ + \mathfrak{g}_-$, and $[\mathfrak{g}_+, \mathfrak{g}_+] \in \mathfrak{g}_+$. □

10.2.2.9 Proposition *Set $t \overset{\text{def}}{=} r + \sigma(r)$, where σ is the permutation $x \otimes y \mapsto y \otimes x$. Then $t \in \mathrm{Sym}^2(\tilde{\mathfrak{g}})$ because it is symmetrized and in $(\mathfrak{g}_- \otimes \mathfrak{g}_+) + (\mathfrak{g}_+ \otimes \mathfrak{g}_-) \subseteq (\mathfrak{g}_- + \mathfrak{g}_+) \otimes (\mathfrak{g}_- + \mathfrak{g}_+)$. Moreover, t is $\tilde{\mathfrak{g}}$-invariant.*

Proof We act by $(\sigma \otimes \mathrm{id})$ on equation (10.2.1.5), which just switches the indices 1 and 2, and add. So the last term cancels:

$$\begin{array}{rl}
\sigma\otimes\mathrm{id}: & [r_{21},r_{13}] + [r_{21},r_{23}] + [r_{23},r_{13}] = 0 \\
+ & [r_{12},r_{13}] + [r_{12},r_{23}] + [r_{13},r_{23}] = 0 \\
\hline
 & [r_{12}+r_{21}, r_{13}+r_{23}] = 0
\end{array}$$

But $r_{12} + r_{21} = t_{12}$, and so $[t \otimes 1, \sum_i (r^i \otimes 1 \otimes r_i + 1 \otimes r^i \otimes r_i)] = 0$, where $r = \sum_i r^i \otimes r_i$. This is equivalent to saying that for all l, $[t, \sum_i (r^i \otimes 1 + 1 \otimes r^i) l(r_i)] = 0$. And similarly for $\mathfrak{g}_+$. □

10.2.2.10 Lemma / Definition *Assume with have a classical R-matrix r. Set $\delta_r(x) \overset{\text{def}}{=} [r, x\otimes 1 + 1\otimes x]$, where we have extended the bracket to tensors by the Leibniz rule: $[A\otimes B, C\otimes 1] \overset{\text{def}}{=} [A, C]\otimes B$. Then $\delta_r(x) \in \bigwedge^2 \mathfrak{g} \subseteq \mathfrak{g} \otimes \mathfrak{g}$.*

Moreover, (g, δ_r) is a Lie bialgebra. Lie bialgebras that arise from classical R-matrices are called quasitriangular, *because there is a triangle in the* braid relation *equation (10.2.1.3), and equation (10.2.2.2) arises as a "semiclassical limit" of equation (10.2.1.3).*

Proof We have to prove two facts.

0. $\sigma \circ \delta_r(x) = [\sigma(r), x \otimes 1 + 1 \otimes x] = [t - r, x \otimes 1 + 1 \otimes x] = 0 - \delta(x)$, so δ_r lands in the exterior square.
1. cocycle: $\delta_r[x, y] = [r, [x, y] \otimes 1 + 1 \otimes [x, y]] = [x, \delta_r y] + [\delta_r x, y]$ by Jacobi for $\tilde{\mathfrak{g}}$. Recall, $[x, y \wedge z] \overset{\text{def}}{=} [x, y] \wedge z + y \wedge [x, z]$.
2. co-Jacobi: $\mathrm{Alt}\big((\delta_r \otimes \mathrm{id}) \circ \delta_r\big) = 0$. This is equivalent to equation (10.2.2.2). □

10.2.2.11 Proposition *Let G be a Lie group with $\mathfrak{g} = \mathrm{Lie}(G)$, and suppose that $\mathfrak{g}$ comes equipped with a classical R-matrix r. Then $p_r : x \mapsto r - (\mathrm{Ad}_x \otimes \mathrm{Ad}_x)(r)$ is a Poisson Lie structure on G, and the tangent Lie bialgebra is $(\mathfrak{g}, \delta_r)$. (As always, we have trivialized $\mathrm{T}G = \mathfrak{g} \times G$ by right translation.)* □

We will not prove Proposition 10.2.2.11, but we give a hint: think about p_r as a 1-cocycle for G with coefficients in $\bigwedge^2 \mathfrak{g}$ (c.f. Section 4.4); see that it is a 1-coboundary.

10.2.3 Factorizable Lie bialgebras

10.2.3.1 Definition *A* factorizable Lie bialgebra *is a quasitriangular Lie bialgebra* $(\mathfrak{g}, r)$ *for which* $t = r + \sigma r$ *is nondegenerate in the following sense: the symmetric bilinear form on* $\mathfrak{g}^*$ *defined by* $\langle l, m\rangle_t \stackrel{\text{def}}{=} \langle l \otimes m, t\rangle$ *is nondegenerate.*

If $\mathfrak{g}$ *is factorizable, then for every* $x \in \mathfrak{g}$ *there is a unique* $l \in \mathfrak{g}^*$ *such that* $x = x_+ - x_-$ *and* $x_\pm = r_\pm(l)$*, where* $r_\pm$ *are as in equations (10.2.2.5) and (10.2.2.6); it is given by* $l = t(x)$*.*

10.2.3.2 Example Let $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$ with standard basis H, X, Y (H is the Cartan, X, Y are the root elements): $[H, X] = 2X$, $[H, Y] = -2Y$, $[X, Y] = H$. Then $r = \frac{1}{4}H \otimes H + X \otimes Y$ is a solution to equation (10.2.2.2), and $t = r + \sigma(r) = \frac{1}{2}H \otimes H + X \otimes Y + Y \otimes X \in \mathrm{Sym}^2(\mathfrak{sl}_2)^{\mathfrak{sl}_2}$. This gives the *Casimir* element $c \stackrel{\text{def}}{=} \frac{H^2}{2} + XY + YX \in \mathcal{U}\mathfrak{sl}_2$. I.e. $c \in \mathcal{Z}(\mathcal{U}\mathfrak{sl}_2)$, which is in fact freely generated by c: $\mathcal{Z}(\mathcal{U}\mathfrak{sl}_2) = \mathbb{C}[c]$.

Then $t = \frac{1}{2}\Delta(c) - c \otimes 1 - 1 \otimes c$, where $\Delta : \mathcal{U}\mathfrak{sl}_2 \to \mathcal{U}\mathfrak{sl}_2^{\otimes 2}$ is the coassociative algebra homomorphism such that $\Delta x = x \otimes 1 + 1 \otimes x$ for $x \in \mathfrak{sl}_2 \subseteq \mathcal{U}\mathfrak{sl}_2$. Since Δ is a homomorphism and $c \in \mathcal{Z}(\mathcal{U}\mathfrak{sl}_2)$, we have $[t, \Delta x] = 0$ for each $x \in \mathfrak{sl}_2$. The element t is called the *mixed Casimir*, and r then is not so strange, being roughly half of the mixed Casimir.

The basis $\{H, X, Y\}$ for $\mathfrak{sl}_2$ determines a dual basis $\{H^\vee, X^\vee, Y^\vee\}$ for $\mathfrak{sl}_2^*$, where we define $K^\vee$ (for $K = H, X, Y$) to be the linear functional that is 1 on K and 0 on the other two basis elements. Then $r_+(l) = \frac{H}{4}l(H) + X\,l(Y)$, where $l \in \mathfrak{sl}_2^*$, so $\mathrm{Im}(r_+) = \mathbb{C}H \oplus \mathbb{C}X = \mathfrak{b}_+ \subseteq \mathfrak{sl}_2$ and $\mathrm{Im}(r_-) = \mathfrak{b}_-$. The kernels are $\ker(\mathfrak{r}_+) = \mathbb{C}X^\vee$ and $\ker(r_-) = \mathbb{C}Y^\vee$, so $\ker(r_+)^\perp$, which is the collection of all elements of $\mathfrak{sl}_2$ on which $X^\vee$ vanishes, is just $\mathfrak{b}_-$.

Note that $t = \frac{1}{2}H \otimes H + X \otimes Y + Y \otimes X$ is the inverse pairing to the Killing form on $\mathfrak{sl}_2$, and so is nondegenerate; in particular, $(\mathfrak{sl}_2, r)$ is a factorizable Lie bialgebra. An element $x = \alpha H + \beta X + \gamma Y \in \mathfrak{sl}_2$ factors as $x_+ = \frac{\alpha}{2}H + \beta X$ and $x_- = -\frac{\alpha}{2}H - \gamma Y$. This is precisely the *linear Gaussian factorization*: $\left(\begin{smallmatrix} \alpha & \beta \\ \gamma & -\alpha \end{smallmatrix}\right) = \left(\begin{smallmatrix} \alpha/2 & \beta \\ 0 & -\alpha/2 \end{smallmatrix}\right) - \left(\begin{smallmatrix} -\alpha/2 & 0 \\ -\gamma & \alpha/2 \end{smallmatrix}\right)$.

We can also work out the cobracket δ_r:

$$\begin{aligned}
\delta_r(H) &= [r, H \otimes 1 + 1 \otimes H] = 0 \\
\delta_r(X) &= \frac{1}{4}[H, X] \otimes H + \frac{1}{4}H \otimes [H, X] + X \otimes [Y, X] \\
&= \frac{1}{2}X \otimes H + \frac{1}{2}H \otimes X - X \otimes H \\
&= \frac{1}{2}H \wedge X \\
\delta_r(Y) &= \frac{1}{2}H \wedge Y
\end{aligned}$$

Equivalently, we can work out the dual Lie algebra structure on $\mathfrak{sl}_2^*$. For example:

$$[H^\vee, X^\vee](a) = H^\vee \wedge X^\vee(\delta_r(a)) \tag{10.2.3.3}$$

But $\delta_r(a)$ had better be in $\mathbb{C}H \wedge X$, otherwise equation (10.2.3.3) is 0. So equation (10.2.3.3) is non-zero only if $a = cX$.

$$
\begin{aligned}
[H^\vee, X^\vee](X) &= (H^\vee \wedge X^\vee)\left(\frac{1}{2} H \wedge X\right) \\
&= \frac{1}{2}\langle H^\vee \wedge X^\vee, H \wedge X\rangle \\
&= \frac{1}{2}\langle H^\vee \otimes X^\vee - X^\vee \otimes X^\vee, H \otimes X - X \otimes H\rangle \\
&= \frac{1}{2}(2) \\
&= 1
\end{aligned}
$$

So $[H^\vee, X^\vee] = X^\vee$ and $[H^\vee, Y^\vee] = Y^\vee$, and $[X^\vee, Y^\vee] = 0$. In particular, $\mathfrak{sl}_2^*$ is very different from $\mathfrak{sl}_2$: it is solvable rather than semisimple. This is the *standard Lie bialgebra structure* on $\mathfrak{sl}_2$. There is a classification of Lie bialgebra structures, and for $\mathfrak{sl}_2$ there is only one factorizable one.

We will see counterparts of this for all simple Lie algebras. ♢

10.2.3.4 Definition *If* $\mathfrak{g}_2$ *is a Lie bialgebra, a Lie subalgebra* $\mathfrak{g}_1 \subseteq \mathfrak{g}_2$ *is a* Lie sub-bialgebra *if* $\delta(\mathfrak{g}_1) \subseteq \mathfrak{g}_1 \wedge \mathfrak{g}_1$.

10.2.3.5 Example The upper and lower Borels $b_\pm \subseteq \mathfrak{sl}_2$ are Lie sub-bialgebras, where $\mathfrak{sl}_2$ is given its standard structure. These are not quasitriangular. ♢

10.3 SL(2, ℂ) and Hopf Poisson algebras

10.3.1 The Poisson bracket on SL(2, ℂ)

We will use Proposition 10.2.2.11, which says that upon trivializing $\mathrm{T}G = \mathfrak{g} \times G$ by right translations, the Poisson structure on G determined by classical R-matrix $r \in \mathfrak{g} \otimes \mathfrak{g}$ is $p_r(x) = -\operatorname{Ad}_x \otimes \operatorname{Ad}_x(r) + r \in \Gamma(\bigwedge^2 \mathrm{T}G) \cong \mathscr{C}(G) \otimes \bigwedge^2 \mathfrak{g}$.

We give $\mathrm{SL}(2, \mathbb{C})$ its standard coordinates:

$$
\mathrm{SL}(2, \mathbb{C}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ s.t. } ad - bc = 1 \right\}
$$

Then $\mathscr{C}(\mathrm{SL}(2)) = \mathbb{C}[a, b, c, d]/(ad - bc - 1)$ is a commutative Hopf algebra, with comultiplication encoding the matrix multiplication:

$$
\begin{pmatrix} \Delta(a) & \Delta(b) \\ \Delta(c) & \Delta(d) \end{pmatrix} = \begin{pmatrix} a \otimes a + b \otimes c & a \otimes b + b \otimes d \\ c \otimes a + d \otimes b & c \otimes b + d \otimes d \end{pmatrix}
$$

The Poisson bracket on $\mathscr{C}(\mathrm{SL}(2))$ is $\{f_1, f_2\}(g) = \langle p(g), \mathrm{d}f_1(g) \wedge \mathrm{d}f_2(g)\rangle$, where

$$p(g) = \sum_{\alpha,\beta} p^{\alpha\beta}(g) e_\alpha \otimes e_\beta \tag{10.3.1.1}$$

$$\langle e_\alpha, \mathrm{d}f(g)\rangle = \frac{\mathrm{d}}{\mathrm{d}t} f(e^{te_\alpha} g)\big|_{t=0} \tag{10.3.1.2}$$

$$= \sum_{ij} \frac{\mathrm{d}}{\mathrm{d}t}(e^{te_\alpha} g)_{ij}\big|_{t=0} \frac{\partial f}{\partial g_{ij}}(g) \tag{10.3.1.3}$$

$$= \sum_{ij} (e_\alpha g)_{ij} \frac{\partial f}{\partial g_{ij}} \tag{10.3.1.4}$$

We have identified the coordinates as $\left(\begin{smallmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{smallmatrix}\right) = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. To define "$e_\alpha g$" we use the fact that SL(2) is a matrix group. Then the exponential map really is the matrix exponential. We emphasize that the formulas depend on the use of right-trivialization; if we had used left trivialization, then the formula would have included $f(ge^{te_\alpha})$ instead of $f(e^{te_\alpha} g)$.

All together, we have:

$$\{f_1, f_2\}(g) = \langle p(q), \mathrm{d}f_1(g) \wedge \mathrm{d}f_2(g)\rangle = 2 \sum_{\alpha,\beta,i,j,k,l} p^{\alpha\beta}(g)(e_\alpha g)_{ij}(e_\beta g)_{kl} \frac{\partial f_1}{\partial g_{ij}} \frac{\partial f_2}{\partial g_{kl}}.$$

The 2 is because of the wedge bracket: we must subtract $ij \leftrightarrow kl$, but everything is skew symmetric.

On the other hand,

$$p^{\alpha\beta}(g) e_\alpha g \otimes e_\beta g = p(g)(g \otimes g) \tag{10.3.1.5}$$

$$= \left(-(g \otimes g)(r)(g^{-1} \otimes g^{-1}) + r^V\right)(g \otimes g) \tag{10.3.1.6}$$

$$= -(g \otimes g)r + r(g \otimes g) \tag{10.3.1.7}$$

and so, up to that unfortunate factor of 2, we have

$$\{f_1, f_2\} = 2 \sum_{ijkl} [r, g \otimes g]_{ij,kl} \frac{\partial f_1}{\partial g_{ij}} \frac{\partial f_2}{\partial g_{kl}}.$$

In particular,

$$\{g_{ij}, g_{kl}\} = 2[r, g \otimes g]_{ij,kl}. \tag{10.3.1.8}$$

Here and above, we let g denote the matrix $g = \left(\begin{smallmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{smallmatrix}\right)$ of coordinate functions. Then $g \otimes g$ is the 4×4 matrix

$$g \otimes g = \begin{pmatrix} g_{11} \otimes g_{11} & g_{11} \otimes g_{12} & g_{12} \otimes g_{11} & g_{12} \otimes g_{12} \\ g_{11} \otimes g_{21} & g_{11} \otimes g_{22} & g_{12} \otimes g_{21} & g_{12} \otimes g_{22} \\ g_{21} \otimes g_{11} & g_{21} \otimes g_{12} & g_{22} \otimes g_{11} & g_{22} \otimes g_{12} \\ g_{21} \otimes g_{21} & g_{21} \otimes g_{22} & g_{22} \otimes g_{21} & g_{22} \otimes g_{22} \end{pmatrix}$$

which we think of as a "$2 \times 2 \times 2 \times 2$" matrix so as to write "$(g \otimes g)_{ij,kl}$". Note that $r \in \mathfrak{sl}(2) \otimes \mathfrak{sl}(2)$ is also a $2 \times 2 \times 2 \times 2$ matrix, via the standard action of $\mathfrak{sl}(2)$ on $\mathbb{C}^2$.

Finally, we summarize equation (10.3.1.8) into one matrix equation, by introducing the notation $\{\cdot \otimes\!\!\!\!\raisebox{-0.5ex}{,}\,\cdot\}$ for the combination of Poisson bracket and exterior tensor product:

$$\{g \otimes\!\!\!\!\raisebox{-0.5ex}{,}\, g\} \stackrel{\text{def}}{=} 2[r, g \otimes g]$$

We will henceforth get rid of this 2 by rescaling the Poisson bracket. This commutator should be understood as follows. On the one hand, g is a matrix of distinguished coordinate functions on $\mathrm{SL}(2)$. On the other hand, it is a variable ranging over elements of $\mathrm{SL}(2)$, and as such acts on every $\mathrm{SL}(2)$ module V by some (variable) matrix. The bracket is defined in $\mathrm{End}(V)^{\otimes 2}$ for any V. In particular, it is defined when $V = \mathbb{C}^2$ the standard representation, and this is the representation corresponding to the standard matrix coordinates we used above.

We can say this in yet another way. Introduce the notation $g_1 \stackrel{\text{def}}{=} g \otimes 1 \otimes \cdots \otimes 1$, $g_2 = 1 \otimes g \otimes 1 \otimes \ldots$, and $r_{12} = r \otimes 1 \otimes \ldots$ etc. Thinking of g has a variable ranging over $\mathrm{SL}(2)$, then each of g_1, etc. is something that can act on $V^{\otimes n}$ for V any $\mathrm{SL}(2)$-module. Then equation (10.3.1.8) could also be summarized as $\{g_1, g_2\} = [r_{12}, g_1 g_2]$. In this notation it becomes easy to verify the Jacobi identity:

$$\begin{aligned}
\{g_1, \{g_2, g_3\}\} &= \{g_1, [r_{23}, g_2 g_3]\} \\
&= [r_{23}, \{g_1, g_2 g_3\}] \\
\{g_1, g_2 g_3\} &= \{g_1, g_2\} g_3 + g_2 \{g_1, g_3\} \\
&= [r_{12}, g_1 g_2] g_3 + g_2 [r_{13}, g_1 g_3]
\end{aligned}$$

Working with $V = \mathbb{C}^2$ the defining $\mathfrak{sl}(2)$-module, we have $r = \frac{1}{4} H \otimes H + X \otimes Y$ considered as a matrix in $\mathrm{End}(\mathbb{C}^2)^{\otimes 2}$, where H, X, Y is the standard basis of $\mathfrak{sl}(2)$. Let's choose a basis e_1, e_2 of $\mathbb{C}^2$, and so $e_{ij} = e_i \otimes e_j$ is a basis of $\mathbb{C}^4 = \mathbb{C}^2 \otimes \mathbb{C}^2$. Returning to the usual letters, $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, and so:

$$g \otimes g = \left(\begin{array}{cc:cc} aa & ab & ba & bb \\ ac & ad & bc & bd \\ \hdashline ca & cb & da & db \\ cc & cd & dc & dd \end{array}\right), \quad r = \left(\begin{array}{cc:cc} 1/4 & 0 & 0 & 0 \\ 0 & -1/4 & 1 & 0 \\ \hdashline 0 & 0 & -1/4 & 0 \\ 0 & 0 & 0 & 1/4 \end{array}\right)$$

Then:

$$\{g \otimes\!\!\!\!\raisebox{-0.5ex}{,}\, g\} = \begin{pmatrix} \{a,a\} & \{a,b\} & \{b,a\} & \{b,b\} \\ \{a,c\} & \{a,d\} & \{b,c\} & \{b,d\} \\ \{c,a\} & \{c,b\} & \{d,a\} & \{d,b\} \\ \{c,c\} & \{c,d\} & \{d,c\} & \{d,d\} \end{pmatrix} = [r, g \otimes g] = \begin{pmatrix} 0 & \frac{1}{2}ab & -\frac{1}{2}ab & 0 \\ \frac{1}{2}ac & bc & 0 & -\frac{1}{2}bd \\ -\frac{1}{2}ac & 0 & -bc & \frac{1}{2}bd \\ 0 & -\frac{1}{2}cd & \frac{1}{2}cd & 0 \end{pmatrix}$$

Put another way:

$$\{a,b\} = \frac{1}{2}ab, \quad \{a,c\} = -\frac{1}{2}ac, \quad \{a,d\} = bc, \quad \{b,c\} = 0, \quad \{b,d\} = -\frac{1}{2}bd, \quad \{c,d\} = \frac{1}{2}cd.$$

You would never guess these formulas.

10.3.2 Hopf Poisson algebras

The formulas in the previous section make $\mathscr{C}(\mathrm{SL}(2))$ into an example of:

10.3.2.1 Definition *A* Hopf Poisson algebra *is a commutative algebra A equipped with a comultiplication $\Delta : A \to A^{\otimes 2}$ making it into a Hopf algebra, and a bracket $\{,\} : A^{\otimes 2} \to A$ making it into a Poisson algebra, satisfying the compatibility requirement that*

$$\Delta(\{f_1, f_2\}) = \{\Delta f_1, \Delta f_2\}.$$

This equation is in $A^{\otimes 2}$, and the right-hand side requires the tensor product of Poisson algebras*, defined by: $\{s \otimes t, u \otimes v\} \stackrel{\text{def}}{=} \{s, u\} \otimes tv + su \otimes \{t, v\}$.*

More generally:

10.3.2.2 Proposition *If G is an algebraic Poisson Lie group, then $\mathscr{C}(G)$ is Hopf Poisson.*

10.3.2.3 Remark Suppose we have an algebraic group G and an algebraic subgroup $H \subseteq G$. Then we get two Hopf algebras $\mathscr{C}(G)$ and $\mathscr{C}(H)$. They are related as follows. Let I_H be the vanishing ideal of H; since H is a subgroup, it is a Hopf ideal. Then $\mathscr{C}(H) = \mathscr{C}(G)/I_H$. $\Diamond$

10.3.2.4 Definition *Let G be a Poisson Lie group. A subgroup $H \subseteq G$ is a* Poisson Lie subgroup *if it is both a Lie subgroup and a Poisson submanifold. If G, H are algebraic, then $\mathscr{C}(H) = \mathscr{C}(G)/I_H$, and that H is a Poisson submanifold is equivalent to the condition that $\{I_H, \mathscr{C}(G)\} \subseteq I_H$. This is to say that I_H is a* Poisson ideal*. Since it is both a Hopf ideal and a Poisson ideal, it is a* Hopf Poisson ideal.

10.3.2.5 Example In $\mathscr{C}(\mathrm{SL}(2))$, what are some natural ideals? What can you vanish without going into contradictions with the Poisson bracket? Can you vanish a? No, because $\{a, d\} = bc$, and if we vanished a, we'd have $0 = bc \neq 0$. Can we vanish c? Yes: there's no problem with $\{a, d\} = 0$. Indeed, if $c = 0$, since $ad - bc = 1$ we would have $d = a^{-1}$, and so the bracket should vanish. Thus $(c \equiv 0)$ defines a Hopf Poisson ideal in $\mathscr{C}(\mathrm{SL}(2))$. And sure enough it is the vanishing ideal of a Poisson subgroup, namely the upper Borel $B_+ = \{\left(\begin{smallmatrix} a & b \\ 0 & a^{-1} \end{smallmatrix}\right)\}$, with Poisson bracket generated by $\{a, b\} = \frac{1}{2}ab$. Another Hopf Poisson ideal is generated by b, and cuts out the lower Borel. Finally, the Cartan subgroup H has vanishing ideal $\langle b, c \rangle$ and trivial Poisson structure. $\Diamond$

10.3.3 $\mathrm{SL}(2, \mathbb{C})^*$, a dual Lie group

In Example 10.2.3.2 we described the Lie algebra structure on $\mathfrak{sl}(2)^*$ dual to the standard bialgebra structure on $\mathfrak{sl}(2)$. Writing the basis as $H^\vee, X^\vee, Y^\vee$, the brackets are $[H^\vee, X^\vee] = X^\vee$, $[H^\vee, Y^\vee] = Y^\vee$, and $[X^\vee, Y^\vee] = 0$. We can realize this Lie algebra as pairs of matrices:

$$\mathfrak{sl}(2)^* = \left\{ \left(\begin{pmatrix} a & b \\ 0 & -a \end{pmatrix}, \begin{pmatrix} -a & 0 \\ c & a \end{pmatrix} \right) \right\}$$

where $H^\vee = \frac{1}{2}(H, -H) = \left(\left(\begin{smallmatrix} 1/2 & 0 \\ 0 & -1/2 \end{smallmatrix}\right), \left(\begin{smallmatrix} -1/2 & 0 \\ 0 & 1/2 \end{smallmatrix}\right) \right)$ and $X^\vee = (X, 0) = (\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right))$ and $Y^\vee = (0, Y) = (\left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right))$. So $\mathfrak{sl}(2)^*$ is naturally a subalgebra of $\mathfrak{b}_+ \oplus \mathfrak{b}_-$, and the simply connected group $\mathrm{SL}(2)^*$ that exponentiates $\mathfrak{sl}(2)^*$ is a subgroup of $B_+ \times B_-$, namely the subgroup of the form:

$$\mathrm{SL}(2)^* = \left\{ \left(\begin{pmatrix} a & b \\ 0 & a^{-1} \end{pmatrix}, \begin{pmatrix} a^{-1} & 0 \\ c & a \end{pmatrix} \right) \right\}$$

Given $b \in \mathrm{SL}(2)^*$, let $b^\pm$ be its projection in $\mathrm{SL}(2)^* \hookrightarrow B_+ \times B_- \to B_\pm$. Letting $B_\pm$ act on $\mathbb{C}^2$ in the usual way, we claim that the Poisson structure on $\mathrm{SL}(2)^*$ is:

$$\{b^+ \underset{,}{\otimes} b^+\} \overset{\text{def}}{=} [r, b^+ \otimes b^+] \tag{10.3.3.1}$$

$$\{b^+ \underset{,}{\otimes} b^-\} \overset{\text{def}}{=} [r, b^+ \otimes b^-] \tag{10.3.3.2}$$

$$\{b^- \underset{,}{\otimes} b^-\} \overset{\text{def}}{=} [r, b^- \otimes b^-] \tag{10.3.3.3}$$

As before, each side of the equations should be interpreted in $\mathrm{End}(V)^{\otimes 2}$ where V is any $\mathrm{SL}(2)$-module. The matrices $b^\pm$ act via the standard embeddings $B_\pm \hookrightarrow \mathrm{SL}(2)$, and $r = \frac{1}{4} H \otimes H + X \otimes Y \in \mathfrak{sl}(2)^{\otimes 2}$. On the left hand side, we define $\{A \underset{,}{\otimes} B\}$ to be the matrix $\{A_{ij}, B_{kl}\}$ where i,j,k,l range from 1 to $\dim V$.

In the coordinates above, where $b = (b^+, b^-) = \left(\left(\begin{smallmatrix} a & b \\ 0 & -a \end{smallmatrix}\right), \left(\begin{smallmatrix} -a & 0 \\ c & a \end{smallmatrix}\right) \right)$, the Poisson brackets are:

$$\{a, b\} = \frac{1}{2} ab \tag{10.3.3.4}$$

$$\{a, c\} = -\frac{1}{2} ac \tag{10.3.3.5}$$

$$\{b, c\} = a^2 - a^{-2} \tag{10.3.3.6}$$

The Hopf comultiplication is:

$$\Delta a = a \otimes a \tag{10.3.3.7}$$

$$\Delta c = a \otimes c + c \otimes a^{-1} \tag{10.3.3.8}$$

$$\Delta b = a \otimes b + b \otimes a^{-1} \tag{10.3.3.9}$$

Note that $\mathscr{C}(\mathrm{SL}(2)^*) = \mathbb{C}[a, a^{-1}, b, c]$. The coproduct of (and brackets with) a^{-1} is determined by $aa^{-1} = 1$ (and the Leibniz rule).

Let us suppose that equations (10.3.3.4) to (10.3.3.6) do define a Poisson structure on $\mathscr{C}(\mathrm{SL}(2)^*)$ compatible with its Hopf structure. We will check that this Poisson structure corresponds to the Lie cobracket on $\mathfrak{sl}(2)^*$. To do so, we study a neighborhood of the identity. Note that the identity is when $a = 1$, $b = 0$, and $c = 0$. An infinitesimally-near point is of the form $a = e^{H/4}$, $b = X$, and $c = Y$ for some "infinitesimal" coordinates H, X, Y. More precisely, we realize a formal neighborhood of the identity as $\mathbb{C}[\![H, X, Y]\!]$.

By the Leibniz rule, $\{e^{H/4}, X\} = \frac{1}{4} e^{H/4} \{H, X\}$. On the other hand, equations (10.3.3.4) to (10.3.3.6) say that $\{e^{H/4}, X\} = \{a, b\} = \frac{1}{2} ab = \frac{1}{2} e^{H/4} X$. So $\{H, X\} = 2X$. A similar calculation shows that $\{H, Y\} = -2Y$. Finally, $\{X, Y\} = a^2 + a^{-2} = e^{H/2} - e^{-H/2} = H + \frac{1}{24} H^3 + \dots$. Differentiating as $H, X, Y \to 0$ (which is the same as differentiating as $(a, b, c) \to (1, 0, 0)$) gives the Lie bracket for $\mathfrak{sl}(2)$.

10.4 The double construction of Drinfeld

In this section we explain the origin of the formula $r = \frac{1}{4} H \otimes H + X \otimes Y$, and hence the origin of the standard Lie bialgebra structure on $\mathfrak{sl}(2)$ and more generally on all Kac–Moody algebras.

10.4.1 Classical doubles

10.4.1.1 Definition *The Lie algebra $\mathfrak{g}_1$ acts* by derivations *on the Lie algebra $\mathfrak{g}_2$ if the underlying vector space of $\mathfrak{g}_2$ is a $\mathfrak{g}_1$-module, and also $x \cdot [l, m] = [x \cdot l, m] + [l, x \cdot m]$ for all $x \in \mathfrak{g}_1$ and $l, m \in \mathfrak{g}_2$, where $\cdot$ is the action and $[,]$ is the bracket in $\mathfrak{g}_2$.*

If $\mathfrak{g}_1$ acts on $\mathfrak{g}_2$ by derivations, the semidirect product $\mathfrak{g}_1 \ltimes \mathfrak{g}_2$ *is the vector space $\mathfrak{g}_1 \oplus \mathfrak{g}_2$ with the bracket $[(x,l),(y,m)] = ([x,y], [l,m] + x \cdot m - y \cdot l)$. It is a Lie algebra.*

10.4.1.2 Remark The semidirect product of Lie algebras is the infinitesimal version of the semidirect product of groups. One way to remember whether to write $\ltimes$ or $\rtimes$ is that the open end points towards the thing being acted on: it's a pair of hands, twisting things around. $\Diamond$

10.4.1.3 Example Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{g}^*$ its dual vector space with $[,]_{\mathfrak{g}^*} = 0$ trivial. Then $\mathfrak{g} \curvearrowright \mathfrak{g}^*$ via the ad^*-action. The double construction presented below builds the semidirect product $\mathfrak{g} \ltimes \mathfrak{g}^*$ in the case when $\mathfrak{g}$ is given its Lie bialgebra structure with trivial cobracket.

As an example, consider $\mathfrak{g} = \mathfrak{so}(3)$. Then $\mathfrak{g}^* \cong \mathbb{R}^3$, and the coadjoint action is the action by rotations. Then $\mathfrak{g} \ltimes \mathfrak{g}^* \cong \mathfrak{so}(3) \ltimes \mathbb{R}^3$ is the Lie algebra of affine transformations of $\mathbb{R}^3$. (For other values of 3, the Lie algebras $\mathfrak{so}(n) \ltimes \mathfrak{so}(n)^*$ and $\mathfrak{so}(n) \ltimes \mathbb{R}^n$ are different.) $\Diamond$

10.4.1.4 Lemma / Definition *Let $(\mathfrak{g}, \delta)$ be a Lie bialgebra and $(\mathfrak{g}^*, \delta_*)$ its dual Lie bialgebra. There are coadjoint actions $\mathrm{ad}^*_{\mathfrak{g}} : \mathfrak{g} \curvearrowright \mathfrak{g}^*$ and $\mathrm{ad}^*_{\mathfrak{g}^*} : \mathfrak{g}^* \curvearrowright \mathfrak{g}$. Neither is by derivations (unless one of δ, δ_* is trivial), but there is a version of the semidirect product construction that is more symmetrical.*

Indeed, there is a unique Lie algebra structure on the vector space $\mathcal{D}(\mathfrak{g}) \overset{\text{def}}{=} \mathfrak{g} \oplus \mathfrak{g}^*$ such that:

- *$\mathfrak{g}$ and $\mathfrak{g}^*$ are Lie subalgebras of $\mathcal{D}(\mathfrak{g})$, and*
- *the canonical symmetric bilinear form on $((x,l),(y,m)) = \langle x, m \rangle + \langle y, l \rangle$ on $\mathcal{D}(\mathfrak{g})$ is $\mathrm{ad}_{\mathcal{D}(\mathfrak{g})}$-invariant.*

The Lie algebra $\mathcal{D}(\mathfrak{g})$ is the Drinfeld double *of $\mathfrak{g}$. It has a unique Lie bialgebra structure for which $(\mathfrak{g}, \delta)$ and $(\mathfrak{g}^*, -\delta_*)$ are sub-bialgebras.*

10.4.1.5 Remark The double $\mathcal{D}(\mathfrak{g})$ is an example of a *bicrossed product.* If $\mathfrak{g}_1$ and $\mathfrak{g}_2$ act on each other, then subject to compatibility conditions of the actions one can form a Lie algebra called $\mathfrak{g}_1 \bowtie \mathfrak{g}_2$. Thus one often writes $\mathcal{D}(\mathfrak{g}) = \mathfrak{g} \bowtie \mathfrak{g}^*$. $\Diamond$

Proof (of Lemma/Definition 10.4.1.4) We will prove uniqueness of the bracket, and leave the remainder to the reader.

Let $\{e_i\}$ be a basis in $\mathfrak{g}$ and $\{e^i\}$ the dual basis of $\mathfrak{g}^*$. Define the *structure constants* by $[e_i, e_j]_{\mathfrak{g}} = \sum_k C^k_{ij} e_k$ and $\delta e_i = \sum_{jk} f^{jk}_i e_i \wedge e_k$. Then $[e^i, e^j]_{\mathfrak{g}^*} = \sum_k f^{ij}_k e^k$ and $\delta_* e^i = \sum_{jk} c^i_{jk} e^j \wedge e^k$.

The basis of $\mathcal{D}(\mathfrak{g})$ is $\{e_i, e^i\}$, and the canonical pairing is $\langle e_i, e_j\rangle = 0 = \langle e^i, e^j\rangle$ and $\langle e_i, e^j\rangle = \delta_i^j$. So that $\mathfrak{g}$ and $\mathfrak{g}^*$ are Lie subalgebras of $\mathcal{D}(\mathfrak{g})$, the brackets of basis elements must be $[e_i, e_j]_{\mathcal{D}} = \sum_k C_{ij}^k e_k$ and $[e^i, e^j]_{\mathcal{D}} = \sum_k f_k^{ij} e^k$. We have only to define the mixed brackets $[e^i, e_j] \in \mathfrak{g} \oplus \mathfrak{g}^*$, which we do by demanding that the pairing $\langle,\rangle$ be invariant. Since $\mathfrak{g}$ and $\mathfrak{g}^*$ are isotropic for $\langle,\rangle$, by computing $\langle [e^i, e_j], e^k\rangle$ we will pick up the $\mathfrak{g}$ component of $[e^i, e_j]$. Then invariance demands:

$$0 = \langle [e^i, e_j], e^k\rangle + \langle e_j, [e^i, e^k]\rangle = \langle [e^i, e_j], e^k\rangle + \langle e_j, \sum_\ell f_\ell^{ik} e^\ell\rangle = \langle [e^i, e_j], e^k\rangle + f_j^{ik}$$

Thus $[e^i, e_j] = -\sum_k f_j^{ik} e_k +$ something in $\mathfrak{g}^*$, and repeating the trick gives:

$$[e^i, e_j] = -\sum_k f_j^{ik} e_k + \sum_k C_{jk}^i e^k$$

□

10.4.1.6 Remark Here is a different way to think about the double, which makes the Jacobi identity manifest but hides the invariance of the scalar product. Recall from Remark 10.1.1.7 that to a Lie bialgebra $(\mathfrak{g}, \delta)$ we can assign a Chevalley bicomplex on $\bigwedge^\bullet \mathfrak{g} \otimes \bigwedge^\bullet \mathfrak{g}^*$ with differentials that encode the bracket and cobracket (and the coadjoint actions). The Jacobi identity is equivalent to each of the rows and each of the columns being chain complexes; the bialgebra compatibility condition is the same as the demand that the squares commute, so that the total complex is in fact a chain complex. But the underlying graded vector space of the total complex is precisely $\bigwedge^\bullet(\mathfrak{g} \oplus \mathfrak{g}^*)$, and the differential is equivalent to a Lie algebra structure on $(\mathfrak{g} \oplus \mathfrak{g}^*)^* = \mathcal{D}(\mathfrak{g})$. This is precisely the Lie algebra structure defined in Lemma/Definition 10.4.1.4. ♢

10.4.1.7 Remark We asserted in Lemma/Definition 10.4.1.4 that $\mathcal{D}(\mathfrak{g})$ has a Lie bialgebra structure such that $(\mathfrak{g}, \delta)$ and $(\mathfrak{g}^*, -\delta_*)$ are sub-bialgebras. It is clear that this coalgebra structure is uniquely determined by these requirements. Then the dual Lie bialgebra $\mathcal{D}(\mathfrak{g})^*$ has a complicated cobracket, but as a Lie algebra $\mathcal{D}(\mathfrak{g})^* = \mathfrak{g} \oplus \mathfrak{g}^*$ is the direct sum of Lie algebras. In particular, the pairing on $\mathcal{D}(\mathfrak{g})^*$ is not invariant. ♢

10.4.1.8 Proposition *This Lie bialgebra structure on $\mathcal{D}(\mathfrak{g})$ is quasitriangular with R-matrix $r = \sum_i e^i \otimes e_i \in \mathfrak{g}^* \otimes \mathfrak{g} \hookrightarrow \mathcal{D}(\mathfrak{g}) \otimes \mathcal{D}(\mathfrak{g})$. (This r does not depend on the basis: it is simply the identity map $\mathrm{id} : \mathfrak{g} \to \mathfrak{g}$ thought of as an element of $\mathfrak{g} \otimes \mathfrak{g}^*$.)*

Proof We compute:

$$\begin{aligned}
\delta_r(e_i) &= [r, e_i \otimes 1 + 1 \otimes e_i] && (10.4.1.9)\\
&= \sum_j [e^j, e_i] \otimes e_j + \sum_j e^j \otimes [e_j, e_i] && (10.4.1.10)\\
&= \sum_{j,k} \left(-f_i^{jk} e_k + C_{ik}^j e^k\right) \otimes e_j + \sum_{j,k} e^j \otimes C_{ji}^k e_k && (10.4.1.11)\\
&= \sum_k f_i^{jk} e_j \otimes e_k && (10.4.1.12)
\end{aligned}$$

In equation (10.4.1.12) we have reindexed and used the skew-symmetry to cancel two terms and change the sign of the third. A similar calculation shows $\delta_r(e^i) = -\sum_{jk} C^i_{jk} e^j \wedge e^k$, and so δ_r is the Lie cobracket for which $(\mathfrak{g}, \delta)$ and $(\mathfrak{g}^*, -\delta_*)$ are subcoalgebras.

We still need to check the classical Yang–Baxter equation. The first term $[r_{12}, r_{13}]$ is

$$[r_{12}, r_{13}] = [e^i \otimes e_i \otimes 1, e^j \otimes 1 \otimes e_j] = [e^i, e^j] \otimes e_i \otimes e_j.$$

The second term is $e^i \otimes e^j \otimes [e_i, e_j]$, and the last is $e^i \otimes [e_i, e^j] \otimes e_j$. But

$$\begin{aligned}[e^i, e^j] \otimes e_i \otimes e_j + e^i \otimes e^j \otimes [e_i, e_j] + e^i \otimes [e_i, e^j] \otimes e_j = \\ = f^{ij}_k e^k \otimes e_i \otimes e_j + C^k_{ij} e^i \otimes e^j \otimes e_j + e^i \otimes (-C^j_{ik} e^k + f^{jk}_i e_k) \otimes e_j = 0 \quad (10.4.1.13)\end{aligned}$$

□

10.4.1.14 Proposition *$\mathcal{D}(\mathfrak{g})$ is factorizable. Indeed, $r + \sigma(r) = \sum_i (e^i \otimes e_i + e_i \otimes e^i)$ defines the canonical invariant scalar product on $\mathcal{D}(\mathfrak{g}) = \mathfrak{g} \oplus \mathfrak{g}^*$, and the corresponding factorization writes each element of $\mathcal{D}(\mathfrak{g})$ as the sum of its $\mathfrak{g}$ and $\mathfrak{g}^*$ parts.* □

Let's investigate the following question. What happens when you take the double of the double? More generally, suppose $(\mathfrak{g}, \delta_r)$ is a factorizable Lie bialgebra, with R-matrix r and $t = r + \sigma(r) \in \mathrm{Sym}^2(\mathfrak{g})^{\mathfrak{g}}$ a nondegenerate invariant bilinear form on $\mathfrak{g}^*$. What is its double?

We defined $r_\pm : \mathfrak{g}^* \to \mathfrak{g}$ by $r_+ : \xi \mapsto (\mathrm{id} \otimes \xi) r$ and $r_- : \xi \mapsto -(\xi \otimes \mathrm{id}) r$. Since t is nondegenerate, it determines $t^\# : \mathfrak{g}^* \xrightarrow{\sim} \mathfrak{g}$, $\xi \mapsto (\mathrm{id} \otimes \xi) t = r_+(\xi) - r_-(\xi)$. Recall that this gave us the *factorization property*: every $x \in \mathfrak{g}$ decomposes uniquely as $x = x_+ - x_-$ such that for some $\xi \in \mathfrak{g}^*$, $x_\pm = r_\pm(\xi)$.

10.4.1.15 Theorem (Double of a factorizable Lie bialgebra)
The map $\mathcal{D}(\mathfrak{g}) = \mathfrak{g} \bowtie \mathfrak{g}^* \to \mathfrak{g} \oplus \mathfrak{g}$ given by $(x, \xi) \mapsto (x + r_+(\xi), x - r_-(\xi))$ is a Lie algebra isomorphism.

Proof It is clearly a linear isomorphism, since $r_+ - r_- = t^\#$ is an isomorphism. But it is also a Lie algebra homomorphism, since the diagonal map $\mathfrak{g} \to \mathfrak{g} \oplus \mathfrak{g}$ is are $\pm r_\pm : \mathfrak{g}^* \to \mathfrak{g}$ (and so also $(r_+, -r_-) : \mathfrak{g}^* \to \mathfrak{g} \oplus \mathfrak{g}$). □

10.4.2 Kac–Moody algebras and their standard Lie bialgebra structure

We begin by explaining the origin of standard Lie bialgebra on $\mathfrak{sl}(2)$. We then introduce the notion of "Kac–Moody algebra," and show that the $\mathfrak{sl}(2)$ story generalizes. Very briefly: Kac–Moody Lie algebras are a generalization of semisimple Lie algebras, and each is approximately the double of its upper Borel.

10.4.2.1 Example Let $\mathfrak{b}_+ \subseteq \mathfrak{sl}(2)$ denote the upper Borel subalgebra. Its basis is $\{H, X\}$ with bracket $[H, X] = 2X$. Recall that $\mathfrak{b}_+$ is in fact a sub-bialgebra: the cobracket is $\delta H = 0$ and $\delta X = \frac{1}{2} H \wedge X$. We will describe $\mathcal{D}(\mathfrak{b}_+)$.

We choose for the dual $\mathfrak{b}_+^*$ the dual basis to be $\{H^\vee, X^\vee\}$ with brackets $[H^\vee, X^\vee] = X^\vee$ and cobrackets $\delta H^\vee = 0$ and $\delta X^\vee = H^\vee \wedge X^\vee$. The double $\mathcal{D}(\mathfrak{b}_+) = \mathfrak{b}_+ \oplus \mathfrak{b}_+^\vee$ has basis $\{H, X, H^\vee, X^\vee\}$. We can work out the brackets on $\mathcal{D}(\mathfrak{b}_+)$ from the proof of Lemma/Definition 10.4.1.4:

$$[X^\vee, H] = 2X^\vee, \qquad [X^\vee, X] = -2H^\vee + H, \qquad [H^\vee, H] = 0, \qquad [H^\vee, X] = -X$$

The cobracket on $\mathcal{D}(\mathfrak{b}_+)$ is determined by the requirement that $\mathfrak{b}_+$ and $\mathfrak{b}_+^*$ be subcoalgebras. By Proposition 10.4.1.8, $\mathcal{D}(\mathfrak{b}_+)$ is quasitriangular, with R-matrix

$$r = H^\vee \otimes H + X^\vee \otimes X.$$

We change bases slightly: $H' = \frac{1}{2}H - H^\vee$, $H'' = \frac{1}{2}H + H'$, $X' = X$, and $Y' = -\frac{1}{2}X^\vee$. Then H'' is in the center of the Lie algebra $\mathcal{D}(\mathfrak{b}_+)$, and $\delta H'' = 0$, so $\mathbb{C}H''$ is a Lie bialgebra ideal. On the other hand, H', X', and Y' satisfy the $\mathfrak{sl}(2)$ relations. Therefore $\mathcal{D}(b_+) = \mathbb{C}H'' \oplus \mathfrak{sl}_2$ as a Lie algebra. Since $\mathbb{C}H''$ is a Lie bialgebra ideal, the quotient $\mathcal{D}(\mathfrak{b}_+)/\mathbb{C}H'' \cong \mathfrak{sl}(2)$ is a Lie bialgebra. The coalgebra part is the standard Lie bialgebra structure introduced in Example 10.1.1.4, up to some signs and factors of 2. It is a general fact that the quotient of a quasitriangular Lie bialgebra is quasitriangular, and doubles are always quasitriangular. Thus the standard Lie bialgebra structure on $\mathfrak{sl}(2)$ is quasitriangular, explaining the R-matrix in Example 10.2.2.4:

$$r = \frac{1}{2}H'' \otimes H'' + \frac{1}{2}\left(H' \otimes H'' - H'' \otimes H'\right) - \frac{1}{2}\left(\frac{1}{4}H' \otimes H' - Y' \otimes X'\right).$$

Note that the cobracket on $\mathfrak{b}_+$ is completely natural: the upper and lower Borels $\mathfrak{b}_\pm$ are perfectly paired by the Killing form, and the cobracket on $\mathfrak{b}_+$ is (up to some signs and factors of 2) dual to the bracket on $\mathfrak{b}_-$. ◊

This story completely generalizes to all semisimple Lie algberas; c.f. Exercise 10. We now generalize further, by axiomatizing the Cartan-Dynkin structure of semisimple Lie algebras.

10.4.2.2 Definition *Let $\mathfrak{h}$ be a finite-dimensional vector space, $\mathfrak{h}^*$ its dual, and choose a collection $h_1, \dots, h_n \in \mathfrak{h}$ of "co-roots" and $\alpha_1, \dots, \alpha_n \in \mathfrak{h}$ of "roots". The* generalized Cartan matrix *is $a_{ij} \overset{\text{def}}{=} \langle \alpha_i, h_j \rangle$. We demand the following conditions:*

- *$a_{ii} = 2$,*
- *$a_{ij} \in \mathbb{Z}_{\leq 0}$ for $i \neq j$*
- *If $a_{ij} \neq 0$, then $a_{ji} \neq 0$.*
- *There exists $d_1, \dots, d_n \in \mathbb{Z}_{>0}$ diagonalizing the matrix, i.e. $d_i a_{ij} = a_{ji} d_i$ (no sum).*
- *$\dim \mathfrak{h} = n + \dim(\ker a)$.*

Given this data, the Kac–Moody algebra *$\mathfrak{g}(a)$ is the Lie algebra generated by $\mathfrak{h}$ and generators e_i and f_i for each $i = 1, \dots, n$, with two defining relations:*

$$[h, h'] = 0, \quad [h, e_i] = \langle \alpha_i, h \rangle e_i, \quad [h, f_i] = -\langle \alpha_i, h \rangle f_i, \quad [e_i, f_j] = \delta_{ij} h_i \tag{10.4.2.3}$$

$$\left(\mathrm{ad}_{e_i}\right)^{1-a_{ij}}(e_j) = 0 \text{ and } \left(\mathrm{ad}_{f_i}\right)^{1-a_{ij}}(f_j) = 0 \tag{10.4.2.4}$$

Equation (10.4.2.4) is the Serre relation. *The Lie algebra $\mathfrak{g}(a)$ is $\mathbb{Z}$-graded with $\deg(\mathfrak{h}) = 0$, $\deg(e_i) = 1$, and $\deg(f_i) = -1$. The abelian subalgebra $\mathfrak{h}$ is the* Cartan subalgebra.

10.4.2.5 Example If a is positive definite, then $n = \dim \mathfrak{h}$ and $\mathfrak{g}(a)$ is a semisimple finite-dimensional Lie algebra. ◊

10.4.2.6 Example (Galber and Kac) Suppose that $\dim \ker a = 1$, whence $\dim \mathfrak{h} = 1 + n$, and a is positive semidefinite. It is a theorem of Galber and Kac that the matrix a has a block form as follows: a nondegenerate $(n-1) \times (n-1)$ block, and a column of zeros. The nondegenerate part defines a semisimple Lie algebra $\mathfrak{g}$, and the full Lie algebra is

$$\mathfrak{g}(a) \cong \mathfrak{g}[t, t^{-1}] \oplus \mathbb{C}K \oplus \mathbb{C}t\frac{\mathrm{d}}{\mathrm{d}t}. \tag{10.4.2.7}$$

The subalgebra $\mathfrak{g}[t, t^{-1}]$ is the *loop algebra* of $\mathfrak{g}$, because it is the Lie algebra of "algebraic loops" $S^1 \to \mathfrak{g}$. The basis vector K is central in $\mathfrak{g}(a)$, whereas $t\frac{\mathrm{d}}{\mathrm{d}t}$ acts on the loop algebra by differentiation, and hence $\mathfrak{g}[t, t^{-1}] \oplus \mathbb{C}t\frac{\mathrm{d}}{\mathrm{d}t}$ is a nontrivial semidirect product. Write $\mathfrak{h}_{\mathfrak{g}}$ for the Cartan subalgebra of the semisimple $\mathfrak{g}$. The Cartan subalgebra of $\mathfrak{g}(a)$ is $\mathfrak{h} = \mathfrak{h}_{\mathfrak{g}} \oplus \mathbb{C}K \oplus \mathbb{C}t\frac{d}{dt}$.

The translation between equation (10.4.2.7) and the generators in Definition 10.4.2.2 is that $e_1, \dots, e_{n-1}$ and $f_1, \dots, f_{n-1}$ in $\mathfrak{g}(a)$ correspond to the same generators in the finite-dimensional $\mathfrak{g}$; and writing e_0 and f_0 for the generators corresponding to the column of zeros in a, we have $e_0 \mapsto tf_\theta$ and $f_0 \mapsto t^{-1}e_\theta$, where θ is the longest root in $\mathfrak{g}$.

Such Lie algebras are called *affine.* ◊

10.4.2.8 Remark The affine Kac–Moody algebras are the most studied infinite-dimensional Lie algebras. On the one hand, they have simple presentations, and the representation theory of simple Lie algebras transfers directly. On the other hand, they have simple geometrical interpretation as central extensions of loop algebra, along with derivatives. They are intimately connected to physics: including but not limited to two-dimensional gauge theories and conformal field theories.

In the 1980s, people asked about whether there were interesting examples when $\dim \ker a \geq 2$. There don't seem to be.

There is the following picture of Kac–Moody algebras. You should think of the simple Lie algebras as spheres, because their Weyl groups are finite. The affine Lie algebras are cylinders. The Kac–Moody algebras with larger kernel are from this perspective hyperbolic, and grow exponentially. There are difficult open questions to understand even the Weyl groups of such algebras; for example, $\mathrm{SL}(2, \mathbb{Z})$ shows up. You would think that since it's a difficult problem, so there should be geniuses working on it, but no: geniuses look for difficult problems with easy solutions. We will not dwell in these lectures on even the affine Kac–Moody algebras, let alone the bigger ones, except to mention that each does have a quasitriangular Lie bialgebra structure. ◊

10.4.2.9 Proposition *If $\mathfrak{g}(a)$ is a Kac–Moody algebra, then*

$$\delta h = 0, \quad \delta e_i = \frac{d_i}{2} h_i \wedge e_i, \quad \delta f_i = \frac{d_i}{2} h_i \wedge f_i$$

is a Lie bialgebra structure, and moreover it is factorizable. □

We will explain the quasitriangularity. For simplicity, we restrict to affine Kac–Moody algebras. The idea: we learned the double construction of Lie bialgebras, and Kac–Moody algebras have Borel subalgebras, and if we apply the double construction to the Borel subalgebra, we will get back the Kac–Moody algebra. We halve and then double.

Let $a = (a_{ij})$ be a generalized Cartan matrix, and suppose that it is $n \times n$, on the index set $i, j \in I = \{0, \dots, n-1\}$. The affine case means that $\operatorname{rank}(a) = n-1$. To be even more restrictive, let's also assume that a is already symmetric, so that we are in the ADE cases.

The Borel subalgebra is $\mathfrak{b}_+(a) = \mathfrak{h} \oplus \mathfrak{n}_+(a)$. It has the following description: $\mathfrak{h}$ is generated by h_i for $i \in I$ and also a new symbol d corresponding to the kernel of a; and $\mathfrak{n}_+$ is generated by e_i for $i \in I$; the relations are $[h_i, h_j] = [h_i, d] = 0$, and $[h_i, e_j] = a_{ij} e_j$, and $[d, e_j] = \delta_{0,j} e_j$, and also $(\mathrm{ad}_{e_i})^{1-a_{ij}}(e_j) = 0$.

The Lie cobracket on $\mathfrak{b}_+(a)$ is, in turn, given by $\delta h_i = \delta d = 0$ and $\delta e_i = \frac{1}{2} h_i \wedge e_i$. The double $\mathcal{D}(\mathfrak{b}_+(a))$ is generated by $h_i, d, e_i, h_i^*, d^*, e_i^*$.

10.4.2.10 Proposition *As a Lie algebra, the double breaks up as a direct sum* $\mathcal{D}(\mathfrak{b}_+) \cong \mathfrak{g}(a) \oplus \mathfrak{h}$. *The* $\mathfrak{g}(a)$ *part is generated by the* e_i, $f_i = e_i^*$, *and by a Cartan part generated by* $H_i = \frac{1}{2}\left(\sum_{j=0}^{n-1} h_j^* a_{ij} + h_i + d^*\right)$ *and* $D = \frac{1}{2}(d + h_0^*)$*; these satisfy the defining relations of* $\mathfrak{g}(a)$*. The second copy* $\mathfrak{h}$ *of the Cartan is generated by* $\tilde{H}_i = -\sum_{j=0}^{n-1} h_j^* a_{ij} + h_i - d^*$ *and by* $\tilde{D} = d - h_0^*$*. Moreover, the quotient map* $\mathcal{D}(\mathfrak{b}_+) \to \mathfrak{g}(A)$ *lets us push the quasitriangular element from the double to the Kac–Moody algebra.* □

10.5 The Belavin–Drinfeld Classification

The *Belavin–Drinfeld classification* describes the factorizable Lie bialgebra structures on Kac–Moody Lie algebras $\mathfrak{g}$. We will restrict our attention to the case when $\mathfrak{g}$ is a simple finite-dimensional Lie algebra over $\mathbb{C}$.

10.5.0.1 Proposition *Any Lie bialgebra structure* δ *on* $\mathfrak{g}$ *is quasitriangular.*

Proof First we state two facts:

1. $\mathrm{H}^1(\mathfrak{g}, V) = 0$ for all V.
2. $\left(\bigwedge^3 \mathfrak{g}\right)^{\mathfrak{g}}$ is one-dimensional, generated by $[\Omega_{12}, \Omega_{23}]$, where Ω is the Casimir.

Recall that δ is a 1-cocycle for $\mathfrak{g}$ with coefficients in $\bigwedge^2 \mathfrak{g}^*$. Then by the first statement δ is in fact a 1-coboundary. We recall what this means: the Chevalley complex (Definition 10.1.1.5) is

$$\operatorname{Hom}_{\mathbb{C}}(\mathbb{C}, \textstyle\bigwedge^2 \mathfrak{g}) \xrightarrow{\mathrm{d}} \operatorname{Hom}_{\mathbb{C}}(\mathfrak{g}, \textstyle\bigwedge^2 \mathfrak{g}) \to \dots,$$

and so we see that there is some $\tilde{r} \in \bigwedge^2 \mathfrak{g}$ so that $\delta = \mathrm{d}\tilde{r} : x \mapsto [x \otimes 1 + 1 \otimes x, \tilde{r}]$. Then recall moreover (Lemma/Definition 10.2.2.10) that δ satisfies co-Jacobi if and only if $\mathrm{CYB}(\tilde{r}) \in (\bigwedge^3 \mathfrak{g})^{\mathfrak{g}}$, where $\mathrm{CYB}(-)$ is the classical Yang–Baxter function. The Lie bialgebra $(\mathfrak{g}, \delta)$ is quasitriangular if in fact $\mathrm{CYB}(r) = 0$, where $\mathrm{d}(\text{antisymmetrization of } r) = \delta$; we will construct such an r from $\tilde{r}$.

By the second statement, $\mathrm{CYB}(\tilde{r}) = c[\Omega_{12}, \Omega_{23}]$. Consider $r \overset{\text{def}}{=} \tilde{r} + \sqrt{c}\,\Omega$. Since Ω is central, $\mathrm{d}\Omega = 0$, and so r and $\tilde{r}$ define the same δ. □

10.5.0.2 Remark From the proof, we see that $\tilde{r}$ is antisymmetric, and so $r + r_{21} = 2\sqrt{c}\Omega$. Thus if $c \neq 0$ then $(\mathfrak{g}, \delta)$ is factorizable, as for simple $\mathfrak{g}$ the Casimir Ω is nondegenerate. It follows that to classify factorizable Lie bialgebra structures on $\mathfrak{g}$, it suffices to classify R-matrices with $r + r_{21} = \Omega$; this gets all such structures up to rescaling. ♢

10.5.0.3 Definition *Let Γ be the set of simple roots of $\mathfrak{g}$. Choose $\Gamma_1, \Gamma_2 \subseteq \Gamma$ and $\tau : \Gamma_1 \to \Gamma_2$. This data $(\Gamma_1, \Gamma_2, \tau)$ is a* Belavin–Drinfeld triple *if:*

1. *τ is an orthogonal bijection.*
2. *$\forall \alpha \in \Gamma_1$, there exists n such that $\tau^n(\alpha) \in \Gamma_2 \smallsetminus \Gamma_1$.*

10.5.0.4 Example When $\mathfrak{g} = \mathfrak{sl}_{n+1}$ there are n roots. We let Γ_1 be the leftmost $n-1$ roots in the Dynkin diagram, and Γ_2 the rightmost $n-1$ roots, and let τ be the shift map once to the right. Eventually each $\alpha \in \Gamma_1$ leaves Γ_1, and the map τ preserves angles. Recall that the angles between the simple roots are described by the number of edges connecting them in the Dynkin diagram. ♢

10.5.0.5 Remark We can easily extend τ to the lattices $\mathbb{Z}\Gamma_1$ and $\mathbb{Z}\Gamma_2$. We get a partial order on the set Δ_+ of positive roots, given by $\alpha \leq \beta$ if $\tau^n\alpha = \beta$ for some n. ♢

We denote by Ω_0 the "$\mathfrak{h}$-part" of Ω.

10.5.0.6 Theorem (Belavin–Drinfeld classification)
Suppose that $(\Gamma_1, \Gamma_2, \tau)$ is a Belavin–Drinfeld triple, and that $r_0 \in \mathfrak{h} \otimes \mathfrak{h}$ satisfies:

1. $r_0 + r_0^{21} = \Omega_0$
2. $(\tau\alpha \otimes \mathrm{id})r_0 + (\mathrm{id} \otimes \alpha)r_0 = 0$ *for all $\alpha \in \Gamma_1$.*

Then

$$r \overset{\text{def}}{=} r_0 + \sum_{\alpha \in \Delta_+} f_\alpha \otimes e_\alpha + \sum_{\substack{\alpha, \beta \in \Delta_+ \\ \alpha < \beta}} f_\alpha \wedge e_\beta \tag{10.5.0.7}$$

is an r-matrix with $r + r_{21} = \Omega$.

Conversely, all R-matrices with $r + r_{21} = \Omega$ are of this form for some choice of $\mathfrak{h}, \Gamma, (\Gamma_1, \Gamma_2, \tau)$.

We will sketch how to go from a Belavin–Drinfeld triple to an R-matrix, and provide the tools for the converse as well. Our strategy will be to study the factorizable R-matrix r by studying the induced map $f \overset{\text{def}}{=} r_- \circ j : \mathfrak{g} \to \mathfrak{g}$. Here $j : \mathfrak{g} \to \mathfrak{g}^*$ is the isomorphism induced by $r + r_{21}$, and r_- is the map $\mathfrak{g}^* \to \mathfrak{g}$ induced by r, as in equation (10.2.2.6). Then $\mathrm{CYB}(r) = 0$ iff for all $x, y \in \mathfrak{g}$ we have:

$$(f - \mathrm{id})[f(x), f(y)] = f[(f - \mathrm{id})(x), (f - \mathrm{id})(y)] \tag{10.5.0.8}$$

Suppose, for a moment, that both f and $f - \mathrm{id}$ were invertible — this can never happen. Then we would have $x = (f - \mathrm{id})^{-1}\hat{x}$ and $y = (f - \mathrm{id})^{-1}\hat{y}$ for some $\hat{x}$ and $\hat{y}$. Dropping the hats, equation (10.5.0.8) would say:

$$f(f - \mathrm{id})^{-1}[x, y] = [f(f - \mathrm{id})^{-1}x, f(f - \mathrm{id})^{-1}y]$$

Of course, this equation is nonsense, as $(f - \mathrm{id})^{-1}$ does not exist. What we can do without difficulty is define $\theta =$ "$f/(f-\mathrm{id})$" $: \operatorname{Im}(f - \mathrm{id})/\ker f \to \operatorname{Im} f/\ker(f - \mathrm{id})$. Note that if $x \in \ker f$, then $(f - \mathrm{id})(-x) = x$, and so $x \in \operatorname{Im}(f-1)$, so the quotients make sense. Certainly we have a map "$(f-\mathrm{id})^{-1}$" $: \operatorname{Im}(f - \mathrm{id}) \to \mathfrak{g}/\ker(f - \mathrm{id})$. Then $\theta = f(f - \mathrm{id})^{-1}$ would vanish on $\ker f$, and return something in $\operatorname{Im} f$.

We are particularly interested in the case when $r + r_{21} = \Omega$ on the nose, and this happens iff $f + f^* = \mathrm{id}$, where f^* is the adjoint to f with respect to the Killing form on $\mathfrak{g}$. Then $\ker f = \operatorname{Im}(f-1)^\perp$ and $\ker(f-1) = (\operatorname{Im} f)^\perp$, because $f(x) = 0$ iff $(f(x), y) = 0\,\forall y$ iff $(x, f^*(y)) = 0\,\forall y$ iff $(x, (1-f)(y)) = 0\,\forall y$ iff $x \perp \operatorname{Im}(f-1)$. Continuing to play with the formulas, one discovers:

10.5.0.9 Lemma / Definition *If $f + f^* = 1$, then equation (10.5.0.8) holds iff $\mathfrak{c}_1 \overset{\text{def}}{=} \operatorname{Im}(f-1)$ and $\mathfrak{c}_2 \overset{\text{def}}{=} \operatorname{Im}(f)$ are subalgebras and θ is an isomorphism. The pair of subalgebras $\mathfrak{c}_1$ and $\mathfrak{c}_2$ is the* Cayley transform *of f.*

So we started out being interested in classical R-matrices that symmetrize to the Casimir, and now we're interested in subalgebras.

We now connect this to Belavin–Drinfeld triples. Such a triple included two isomorphic subdiagrams Γ_1, Γ_2 of the Dynkian diagram for $\mathfrak{g}$. Let $\mathfrak{g}_i$ be the subalgebra spanned by those $\{h_\alpha, e_\alpha, f_\alpha\}$ for $\alpha \in \mathbb{Z}\Gamma_i$. Then τ induces an isomorphism $\mathfrak{g}_1 \to \mathfrak{g}_2$ by sending $h_\alpha \mapsto h_{\tau\alpha}$ and so on. Our goal will be to construct $(\mathfrak{c}_1, \mathfrak{c}_2, \theta)$ from $(\Gamma_1, \Gamma_2, \tau)$, by asking that $\mathfrak{c}_i \supseteq \mathfrak{g}_i$ and that $\theta|_{\mathfrak{g}_1} = \tau$. To fix notation, we split each $\mathfrak{g}_i = \mathfrak{n}_i^- \oplus \mathfrak{h} \oplus \mathfrak{n}_i^+$ as usual, and we also introduce the subalgebras $\mathfrak{n}^+_{\smallsetminus 1} \overset{\text{def}}{=} \langle \mathbb{C}e_\alpha \text{ s.t. } \alpha \notin \mathbb{Z}\Gamma_1\rangle$ and $\mathfrak{n}^-_{\smallsetminus 2} \overset{\text{def}}{=} \langle \mathbb{C}f_\alpha \text{ s.t. } \alpha \notin \mathbb{Z}\Gamma_2\rangle$.

In order to have an R-matrix, the data $(\mathfrak{c}_1, \mathfrak{c}_2, \theta)$ had to satisfy some strong conditions, including that $\mathfrak{c}_i \supseteq \mathfrak{c}_i^\perp$. Then a reasonable guess is $\mathfrak{c}_i = \mathfrak{g}_i \oplus \mathfrak{n}^+_{\smallsetminus i} \oplus V_i$, where $V_i \subseteq \mathfrak{h}_i^\perp$ satisfies $V_i^\perp \subseteq (\mathfrak{h}_i^\perp \cap V_i)$. Then $\mathfrak{c}_i^\perp = \mathfrak{n}^+_{\smallsetminus i} \oplus V_i^\perp \subseteq \mathfrak{c}_i$, and $\mathfrak{c}_i/\mathfrak{c}_i^\perp = \mathfrak{g}_i \oplus V_i/(V_i^\perp \cap \mathfrak{h}_i^\perp)$.

We asked for f such that $\theta|_{\mathfrak{g}_1} = \tau$. Since θ respects the decomposition $\mathfrak{g} = \mathfrak{n}_- \oplus \mathfrak{h} \oplus \mathfrak{n}_+$, we should hope that f does as well, and so set $f = f_+ + f_0 + f_-$, where $f_+ : \mathfrak{n}_+ \to \mathfrak{n}_+$, etc. For this to work, we had better have:

$$\begin{aligned} \operatorname{Im}(f - \mathrm{id}) &= \mathfrak{c}_1 = \mathfrak{g}_1 \oplus \mathfrak{n}^+_{\smallsetminus 1} \oplus V_1, & \operatorname{Im} f &= \mathfrak{c}_2, \\ \ker f &= \mathfrak{c}_1^\perp = \mathfrak{n}^+_{\smallsetminus 1} \oplus V_1/(V_1^\perp \cap \mathfrak{g}_1^\perp), & \ker(f - \mathrm{id}) &= \mathfrak{c}_2^\perp. \end{aligned}$$

Then $\ker(f - \mathrm{id}) \subseteq \mathfrak{n}_- \oplus \mathfrak{h}$, and so $(f_+ - \mathrm{id}_+) : \mathfrak{n}_+ \to \mathfrak{n}_+$ is invertible. Then $\psi \overset{\text{def}}{=} f_+/(f_+ - \mathrm{id}_+)$ must be:

$$\psi(x) = \begin{cases} 0 & x \in \mathfrak{n}^+_{\smallsetminus 1} \\ \tau(x) & x \in \mathfrak{n}_1^+ \end{cases}$$

Moreover, we want $\psi - \mathrm{id}_+ = (f_+ - \mathrm{id}_+)^{-1}$; then $\psi_+ - \mathrm{id}_+$ is invertible if and only if $(f_+ - \mathrm{id}_+)$ is invertible. This gives the second condition in the definition of Belavin–Drinfeld triple:

10.5.0.10 Lemma *$\psi - \mathrm{id}_+$ is invertible iff $\forall \alpha \in \Gamma_1$, there exists n such that $\tau^n(\alpha) \in \Gamma_2 \smallsetminus \Gamma_1$.*

Proof If the latter condition holds, then ψ is nilpotent and $(\psi - \mathrm{id}_+)^{-1} = -\sum_{n \geq 0} \psi^n$.

Conversely, suppose that for some $\alpha \in \Gamma_1$, $\tau^n(\alpha) \in \Gamma_1$ for all n. Since τ is a bijection and Γ_1 is finite, then eventually $\alpha = \tau^n(\alpha)$ for some n. So ψ has 1 as an eigenvalue, and $\psi - \mathrm{id}_+$ is not invertible. □

From all of this, we know f_+; we can figure out f_- from the demand that $f + f^* = 1$. Working it all out, we get:

$$f = f_0 - \sum_{n\geq 1} \psi^n + \mathrm{id}_- + \sum_{n\geq 1} (\psi^*)^n$$

Here ψ^* is the map that undoes ψ on $\mathfrak{n}_-$.

We have already explained how to build an R-matrix from f, and so we do get equation (10.5.0.7). This completes the sketch of one direction. We invite the reader to work out the converse. Details are available in, for example, [ES02].

Exercises

1. Check that the construction in Example 10.1.1.3 gives a Lie bialgebra.

2. (a) In Definition 10.1.1.5, show that $\mathrm{d}^2 = 0$, so that the construction does define a cochain complex.

   (b) Make sense of the following very simple description of the Chevalley complex in terms of Grassman algebra $\bigwedge^\bullet \mathfrak{g}$, being careful with upper and lower indices: Let $\{c_i\}$ be a basis of $\mathfrak{g}$; then $\bigwedge^\bullet \mathfrak{g}$ is the associative algebra generated by the c^i subject to $c^i c^j + c^j c^i = 0$. For simplicity, let $M = \mathbb{C}$. Let f_{ij}^k be the structure constants. Then $\mathrm{d} = \sum_{ijk} f_{ij}^k c^i c^j \frac{\partial}{\partial c^k}$.

   (c) Check that the construction in Remark 10.1.1.7 gives a bicomplex if and only if $(\mathfrak{g}, \mathfrak{g}^*)$ is a Lie bialgebra.

3. Prove Proposition 10.1.2.2. All you have to check is that the Jacobi identity on the commutator induces the Jacobi on $\{,\}$, and then you have to check the Leibniz rule.

4. (a) Show that the Poisson structure on a Poisson Lie group vanishes at the identity.

   (b) Let G be a Poisson Lie group. Show that the map $g \mapsto g^{-1}$ is an *anti-Poisson map*, i.e. it is a Poisson map from G to $\bar{G}$, where for any Poisson manifold M we write $\bar{M}$ for the same manifold with the negated Poisson structure.

5. Prove equation (10.1.3.4). Try to do it invariantly, but if you cannot, do it in local coordinates. On the one hand, local coordinates are very messy, and on the other hand, by making your hands dirty, you can really see what you're doing.

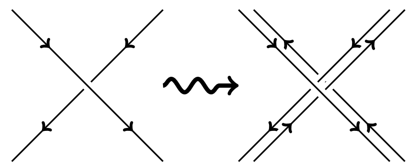
6. Formulate the notion of Lie bialgebra ideal. You must decide on the correct condition on the cobracket.

7. Show that equations (10.3.3.4) to (10.3.3.6) define a Hopf Poisson structure on $\mathscr{C}(\mathrm{SL}(2)^*)$.

8. Check directly that the Lie bracket defined in the proof of Lemma/Definition 10.4.1.4 satisfies the Jacobi identity. Also give a basis-free description of this bracket.

9. (a) Let $\mathfrak{g}$ be a factorizable Lie algebra, and $r = \sum_i e_i \otimes e^i \in \mathfrak{g} \otimes \mathfrak{g}^* \hookrightarrow \mathcal{D}(\mathfrak{g})^{\otimes 2}$ the r-matrix for the double of $\mathfrak{g}$. Let $f : \mathcal{D}(\mathfrak{g}) \to \mathfrak{g} \oplus \mathfrak{g}$ denote the map from Theorem 10.4.1.15. Compute $(f \otimes f)(r)$ in terms of f.

(b) Show that the following "doubling" operation on tangle diagrams respects the Reidemeister moves:



(c) Relate parts (a) and (b).

10. Prove:

(a) Let $\mathfrak{g}$ be a simple Lie algebra. Its upper Borel $\mathfrak{b}_+$ can be defined in terms of generators and relations: for each $i \in \Gamma$ the Dynkin diagram, we have two generators H_i and X_i, and the relations are

$$[H_i, H_j] = 0,\ [H_i, X_i] = a_{ij} X_j,\ \text{and}\ (\mathrm{ad}_{X_i})^{1-a_{ij}}(X_j) = 0 \text{ if } i \neq j$$

The dimension of $\mathfrak{b}_+$ is $r + |\Delta_+|$, where r is the rank of $\mathfrak{g}$ and Δ_+ is the set of positive roots.

(b) There is a Lie bialgebra structure on $\mathfrak{b}_+$ determined by $\delta H_i = 0$, $\delta X_i = \frac{d_i}{2} H_i \wedge X_i$, where $d_i = (\alpha_i, \alpha_i)/2$ is the length of the simple root α_i.

(c) Check that, as a Lie algebra, $\mathcal{D}(\mathfrak{b}_+) \cong \mathfrak{g} \oplus \mathfrak{h}$, where $\mathfrak{h}$ is a central copy of the Cartan.

Chapter 11

Symplectic geometry of Poisson Lie groups

In Section 9.2.5 we proved that every semisimple Lie group G has a *Bruhat decomposition* into cells indexed by the Weyl group; we will review this decomposition, and the related *Shubert decomposition* of the flag manifold, in Section 11.2. The goal of this chapter is to understand these decompositions in terms of the Poisson geometry of G.

11.1 Real forms of Lie bialgebras

Let us update the discussion from Section 8.3.1 on real forms of Lie algebras to cover the bialgebra case.

A *real Lie bialgebra* is a real vector space $\mathfrak{g}_\mathbb{R}$, real $[,] : \mathfrak{g}_\mathbb{R} \otimes \mathfrak{g}_\mathbb{R} \to \mathfrak{g}_\mathbb{R}$, and real $\delta : \mathfrak{g}_\mathbb{R} \to \mathfrak{g}_\mathbb{R} \wedge \mathfrak{g}_\mathbb{R}$, satisfying the same relations as in Definition 10.1.1.1; i.e. it is a Lie bialgebra over $\mathbb{R}$. The *complexification* of $\mathfrak{g}_\mathbb{R}$ is the complex Lie bialgebra $\mathfrak{g}_\mathbb{C} = \mathfrak{g}_\mathbb{R} \otimes_\mathbb{R} \mathbb{C}$. Then $\mathfrak{g}_\mathbb{R} \subseteq \mathfrak{g}_\mathbb{C}$ is invariant with respect to complex conjugation.

If $\mathfrak{g}_\mathbb{C}$ is a complex Lie bialgebra, a *real form* of $\mathfrak{g}_\mathbb{C}$ is any real Lie bialgebra $\mathfrak{g}_\mathbb{R}$ along with an isomorphism $\mathfrak{g}_\mathbb{C} = \mathfrak{g}_\mathbb{R} \otimes \mathbb{C}$. Let $\sigma : \mathfrak{g}_\mathbb{C} \to \mathfrak{g}_\mathbb{C}$ be a real-linear complex-antilinear ($\sigma(\lambda a) = \bar{\lambda}\sigma(a)$) Lie bialgebra automorphism ($\sigma[a,b] = [\sigma a, \sigma b]$ and $(\sigma \otimes \sigma)(\delta a) = \delta(\sigma a)$), and suppose that it is in fact an involution ($\sigma^2 = \mathrm{id}$). Denote the set of fixed points of σ by $\mathfrak{g}^\sigma$; it is a real form of $\mathfrak{g}_\mathbb{C}$. Conversely, letting σ be the usual complex conjugation on $\mathfrak{g}_\mathbb{R} \otimes \mathbb{C}$, we see that every real form arises in this way.

11.1.0.1 Example We describe the real forms of $\mathfrak{sl}(2,\mathbb{C})$ with its standard bialgebra structure. The *Killing form* $\langle x, y\rangle = \mathrm{tr}(\mathrm{Ad}_x \mathrm{Ad}_y)$ determines a quadratic form on $\mathfrak{sl}(2,\mathbb{C})$; in terms of the basis we have $x = aH + bX + cY \in \mathfrak{sl}(2,\mathbb{C})$, and then $\langle x, x\rangle = 2a^2 + bc$. By looking at the signature of this quadratic form, we can tell apart various real forms.

1. The *compact real form* $\mathfrak{su}(2)$ corresponds to $\sigma = (-1) \circ (\text{Hermetian conjugation})$, i.e. $\sigma(H) = -H$, $\sigma(X) = -Y$, $\sigma(Y) = -X$. Hermetian conjugation is an antiautomorphism of both the Lie algebra and coalgebra structures, and the minus sign makes it into an automorphism.

The σ-invariant elements are spanned (over $\mathbb{R}$) by iH, $X-Y$, and $i(X+Y)$. The Killing form is negative definite on this real subspace. The cobracket must be slightly amended: a priori, it is $\delta(iH)=0$, $\delta(X-Y)=H\wedge(X-Y)$, and $\delta(i(X+Y))=H\wedge i(X+Y)$, which does not land in $(\mathfrak{g}_{\mathbb{R}})^{\wedge 2}$, but for $i\delta$ everything works.

2. The second real form corresponds to complex conjugation in the usual matrix representation. I.e. σ acts trivially on H,X,Y, and is extended $\mathbb{C}$-antilinearly from that. The fixed subalgebra is $\mathfrak{sl}(2,\mathbb{R})$ with its standard bialgebra structure from Example 10.1.1.4. The Killing form has one negative and two positive eigenvalues.

3. There is a third real form of $\mathfrak{sl}(2,\mathbb{C})$ as a Lie bialgebra, called $\mathfrak{su}(1,1)$. The involution σ is $X\mapsto Y$, $Y\mapsto X$, and $H\mapsto -H$. Then the fixed points are iH, $X+Y$, and $i(X-Y)$, and the Killing form has one negative eigenvalue and two positive eigenvalues — indeed, this real form is isomorphic to $\mathfrak{sl}(2,\mathbb{R})$ as a Lie algebra. As with $\mathfrak{su}(2)$, the cobracket on $\mathfrak{su}(1,1)$ is not defined over $\mathbb{R}$ without multiplying $\delta\mapsto i\delta$. Note that as a Lie bialgebra, $\mathfrak{sl}(2,\mathbb{R})$ and $\mathfrak{su}(1,1)$ are not isomorphic. You can see this already that their complexifications have cobrackets differing by a factor of i. Another way to see this: $\ker\delta=\mathbb{R}H$ for $\mathfrak{sl}(2,\mathbb{R})$ and $\mathbb{R}(iH)$ for $\mathfrak{su}(1,1)$, but $\operatorname{ad}H$ acting on $\mathfrak{sl}(2,\mathbb{R})$ has eigenvalues $0,2,-2$, whereas $\operatorname{ad}(iH)$ acting on $\mathfrak{su}(1,1)$ has eigenvalues $0,\pm 2i$, which is to say it does not have as many real eigenspaces. $\lozenge$

11.1.0.2 Example Our main hero is $\mathfrak{sl}(2)$, but we will say a few words about higher-rank Lie algebras. Recall (Section 5.6) that any simple Lie algebra $\mathfrak{g}=\mathfrak{g}_{\mathbb{C}}$ has a rank r and generators $\{H_i,X_i,Y_i\}$ subject to relations that depend on the Dynkin diagram. The *standard Lie bialgebra structure* is

$$\delta H_i=0,\quad \delta X_i=\frac{d_i}{2}H_i\wedge X_i,\quad \delta Y_u=\frac{d_i}{2}H_i\wedge Y_i \tag{11.1.0.3}$$

where d_i is the length of the root i, i.e. they are the entries on the diagonal matrix that symmetrizes the Dynkin diagram.

The *compact real form* of $\mathfrak{g}$ is determined by the involution $\sigma(H_i)=-H_i$, $\sigma(X_i)=-Y_i$, and $\sigma(Y_i)=-X_i$. As a vector space,

$$\mathfrak{g}_{\mathbb{R}}^{\text{compact}}=\bigoplus_{j=1}^{r}\mathbb{R}iH_j\oplus\bigoplus_{\alpha\in\Delta_+}\big(\mathbb{R}i(X_\alpha+X_{-\alpha})\oplus\mathbb{R}(X_\alpha-X_{-\alpha})\big).$$

To make $\mathfrak{g}_{\mathbb{R}}^{\text{compact}}$ into a real Lie bialgebra requires rescaling $\delta\mapsto i\delta$.

When $\mathfrak{g}=\mathfrak{sl}(n,\mathbb{C})$, the compact real form is $\mathfrak{su}(n)$. The positive roots Δ_+ consist of differences $\{\epsilon_i-\epsilon_j\}_{i<j}$, where ϵ_i reads off the ith diagonal matrix entry. Then $X_{\epsilon_i-\epsilon_j}=e_{ij}$, $X_{-\epsilon_i+\epsilon_j}=e_{ji}$, and $H_i=e_{ii}-e_{i+1,i+1}$ for $i=1,\dots,n-1$; here e_{ij} is the matrix that's all zeros except for a 1 in the (i,j)th entry. $\lozenge$

11.1.0.4 Example We begin with $\mathfrak{su}(2)$. It has a basis $iH,i(X+Y),(X-Y)$, and the dual Lie algebra $\mathfrak{su}(2)^*$ has a dual basis h,e,f with bracket $[h,e]=e$, $[h,f]=f$, and $[e,f]=0$. We can realize this group in terms of 2×2 complex matrices as $h=\frac{1}{2}\left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right)$, $e=\left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right)$, and

$f = \left(\begin{smallmatrix} 0 & i \\ 0 & 0 \end{smallmatrix}\right)$. These span precisely the traceless complex upper-triangular matrices with real diagonal. Exponentiating, we have:

$$\mathrm{SU}(2)^* = \left\{ \begin{pmatrix} a & c \\ 0 & a^{-1} \end{pmatrix} \text{ s.t. } a > 0, c \in \mathbb{C} \right\}$$

The $\mathrm{SU}(n)$ case is similar, and we leave it as Exercise 5. The punchline is:

$$\mathrm{SU}(n)^* = \left\{ \begin{pmatrix} a_1 & & * \\ & \ddots & \\ 0 & & a_n \end{pmatrix} \text{ s.t. } a_i > 0, a_1 \dots a_n = 1, * \in \mathbb{C} \right\}$$

Note that this is not a compact group, even though $\mathrm{SU}(n)$ is — indeed, $\mathrm{SU}(n)^*$ is solvable! $\Diamond$

11.1.1 Iwasawa decomposition

We now generalize Examples 11.1.0.2 and 11.1.0.4.

Let $\mathfrak{g}$ be a semisimple complex Lie algebra and G the corresponding connected simply connected complex Lie group, with its standard Poisson structure, and let $\mathfrak{k}, K$ be their compact forms. By Proposition 10.4.2.9, $\mathfrak{g}$ is factorizable, and so by Theorem 10.4.1.15 $\mathcal{D}(\mathfrak{g}) \cong \mathfrak{g} \oplus \mathfrak{g}$. Integrate $\mathfrak{g}^*$ and $\mathcal{D}(\mathfrak{g})$ to connected simply connected groups G^* and $\mathcal{D}(G)$. Then $\mathcal{D}(G) \cong G \times G$. Moreover, $G^* = B_- \times_H B_+$, since Proposition 10.4.2.10 gives $\mathfrak{g}^* = \mathfrak{b}_- \times_{\mathfrak{h}} \mathfrak{b}_+$.

It follows that $\mathcal{D}(K)$ is a real form of $G \times G$, and K^* is a real form of $B_- \times_H B_+$. Which real forms? The compact form K of G corresponds to the involution of $\mathfrak{g}$ that switches $\mathfrak{b}_+$ with $\mathfrak{b}_-$. Track this through the double construction: you find the involution that switches the two copies of G in $\mathcal{D}(G) = G \times G$. Thus:

11.1.1.1 Proposition *$\mathcal{D}(K) \cong G$ as real Lie groups.* $\square$

Let $N_\pm \subseteq B_\pm$ denote the strictly upper- or lower-triangular matrices, so that $B_\pm = H \ltimes N_\pm$. The real form of $N_- \times N_+ \subseteq B_- \times_H B_+ = G^*$ in question is nothing but $N = N_+$ as a real Lie group, since the involution switches $N_- \leftrightarrow N_+$. The real form of $H \subseteq G^*$ is the group A consisting of "positive diagonal matrices," i.e. the connected component of the split real form of H, and so:

11.1.1.2 Proposition *$K^* \cong AN$ as real Lie groups.* $\square$

We have Lie group embeddings $G, G^* \hookrightarrow \mathcal{D}(G)$, and the multiplication map $G \times G^* \to \mathcal{D}(G)$ is a surjection, with kernel $G \cap G^*$. When G is semisimple, the intersection is seen to be trivial. Passing to compact forms, we find an isomorphism $K \times K^* \xrightarrow{\sim} \mathcal{D}(K) = G$, and so:

11.1.1.3 Theorem (Iwasawa decomposition)
The multiplication map $K \times A \times N \to G$ is an isomorphism of manifolds. $\square$

11.2 Recollections on Bruhat and Shubert cells

11.2.1 Bruhat decomposition

We now recall the Bruhat decomposition of Section 9.2.5.

We first recall the *Weyl group* in the sense of Definition 5.4.2.2. Suppose we have a root system $\Delta = \{\alpha \in \mathbb{R}^n\}$. We will be dealing only with semisimple finite-dimensional Lie algebras. We pick $\Gamma \subseteq \Delta$ the simple roots, so that $\Delta = \Delta_+ \cup \Delta_-$ and $\Gamma \subseteq \Delta_+$. It will be convenient to enumerate the simple roots: $\Gamma = \{\alpha_1, \dots, \alpha_r\}$ where r is the rank of Δ (i.e. the rank of $\mathfrak{g}$). To each root $\alpha \in \Delta_+$, we associate a reflection $s_\alpha : x \mapsto x - 2(\alpha, x)/(\alpha, \alpha)$, which is reflection with respect to the hyperplane $\alpha^\perp$. We define the *Weyl group* to be the group generated by these reflections, and it is a property of root systems that this is a finite group. We let $s_i = s_{\alpha_i}$. The following is well-known:

11.2.1.1 Theorem (Presentation of the Weyl group)
$W \cong \langle s_i,\ i = 1, \dots, r$ s.t. $s_i^2 = 1,\ (s_i s_j)^{m_{ij}} = 1\rangle$, where m_{ij} is related to the Cartan matrix: if $a_{ij} \cdot a_{ji} = (0, 1, 2, 3)$ then $m_{ij} = (2, 3, 4, 6)$. □

11.2.1.2 Example For $\mathrm{SL}(n)$, $\Delta = A_{n-1}$, and so $(s_i s_j)^2 = 1$ if $i \neq j \pm 1$, and $(s_i s_{i+1})^3 = 1$. Thus $W_{A_{n-1}} = S_n$. If $\Delta_+ = \{\epsilon_i - \epsilon_j\}_{i<j}$, then $\Gamma = \{\epsilon_i - \epsilon_{i+1}\}_{i=1}^{n-1}$, and W acts by permutations on ϵ_i. ◊

If Δ is the root system of $\mathfrak{g}$, we let $\mathfrak{h} \subseteq \mathfrak{g}$ be the Cartan subalgebra, and so $\mathfrak{h} = \bigoplus_{i=1}^r \mathbb{R}\alpha_i^*$ (roots are in $\mathfrak{h}^*$). Then $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{n}_+ \oplus \mathfrak{n}_-$, where $\mathfrak{n}_\pm$ are nilpotent and correspond to $\Delta_\pm$. By definition, W acts on $\mathfrak{h}$. Does it act naturally on $\mathfrak{g}$? Not quite. Strictly speaking, the answer is No. But anytime the answer is No, you can ask "if not this, then what?" One version of this question is the following. G acts on $\mathfrak{g}$ by the adjoint action; e.g. $\mathrm{SL}(n)$ acts on $\mathfrak{sl}(n)$ by conjugation of matrices. Is there a natural embedding $W \hookrightarrow G$? No: for example, you can embed $W = S_n$ into $GL(n)$ as permutation matrices, but some of these have negative determinant; you can add some signs to set-theoretically embed $W \hookrightarrow \mathrm{SL}(n)$, but the signs are not canonical and the embedding is not a group homomorphism.

What you can always do is to let $H \subseteq G$ denote the Cartan subgroup, and construct its normalizer $N(H) = \{g \in G \text{ s.t. } gHg^{-1} \subseteq H\} \subseteq G$. Then $N(H)/H \cong W$. So one way to try to embed $W \hookrightarrow G$ would be to split the surjection $N(H) \to W$, and the ambiguity in this is precisely H. We will denote a choice of set-theoretic section $W \hookrightarrow N(H)$ by $w \mapsto \dot{w}$. Then in particular $\dot{w}$ acts on $\mathfrak{g}$.

11.2.1.3 Example We will explain the isomorphism $N(H)/H \cong W$ for $G = \mathrm{SL}(n)$. Then H consists of the diagonal matrices. Suppose that both d and $gdg^{-1} = d'$ are diagonal. So $gd = d'g$, and if $g_{ij} \neq 0$, then $d_i = d'_j$. So $d'_i = d_{\sigma(i)}$, where σ is a permutation on the indices $1, \dots, n$. When $g_{ij} = 0$, there are no conditions, other than of course $\det g = 1$. Thus $N(H)$ consists precisely of the monomial matrices. This establishes the extension $N(H) = H.W$, since we can factor any monomial matrix as a diagonal one times a permutation:

$$\begin{pmatrix} & & a \\ b & & \\ & c & \end{pmatrix} = \begin{pmatrix} a & & \\ & b & \\ & & c \end{pmatrix} \begin{pmatrix} & & 1 \\ 1 & & \\ & 1 & \end{pmatrix} \in \text{diagonal} \rtimes S_n = H \rtimes W \subseteq \mathrm{GL}(n)$$

Note that if we really want to land in $\mathrm{SL}(n)$, then at least half of the permutations must have signs added in an ad hoc fashion, since without signs the determinant is negative. ◊

Let G be a complex semisimple Lie group, and $B_\pm \subseteq G$ the upper and lower Borel subgroups; we will often write B for B_+. We restate Theorem 9.2.5.1:

11.2.1.4 Theorem (Bruhat decomposition)
$G = \bigsqcup_{w\in W} B_- w B_- = \bigsqcup_{w\in W} B_- w B_+ = \bigsqcup_{w\in W} B_+ w B_+$

By definition, $BwB = B\dot{w}B = \{b\dot{w}b' \text{ s.t. } b, b' \in B\}$, where $\dot{w}$ is any representative of w in $N(H)$. It does not depend on the choice of $\dot{w} \in N(H)$, since if $\dot{w}$ and $\ddot{w}$ are two representatives, then $\dot{w} = \ddot{w}h$ for $h \in H$. Let $U_\pm \subset B_\pm$ denote the unipotent triangular matrices. Then, for example, any $b \in B_-$ factors as $b = nh$ where $n \in U_-$ and $h \in H$. Sine $h\dot{w} = \dot{w}h'$ for some other $h' \in H$, we could also write the Bruhat decomposition as $G = \bigsqcup_{w\in W} U_- w B_+ = \dots$.

11.2.1.5 Definition *The* double Bruhat cells *are* $G_{u,v} = BuB \cap B_- v B_-$.

Proof (of Theorem 11.2.1.4 for $G = \mathrm{SL}(n)$) Fix $g \in \mathrm{SL}(n)$. We want to put it into a unique cell BwB, where $B = B_+$ are the upper triangular matrices. Choose $b \in B$ such that bg^{-1} has maximal number of zeros in the left side of each row. Indeed, any matrix can be multiplied by a triangular from the left to get into the form $\begin{pmatrix} 0 & 0 & * \\ 0 & * & * \\ * & * & * \end{pmatrix}$. This is backwards row-eschelon reduction: if at any time two rows have the same number of 0s, then we can multiply by an upper triangular to create another 0. So each row will have different number of zeros, we can find $\sigma \in S_n$ such that $\dot{\sigma} b g^{-1} = b'$ is upper triangular. Which is to say we have found (σ, b, b') such that $g = (b')^{-1}\dot{\sigma} b \in B\sigma B$.

We now must show that these cells do not intersect. Assume that $b\dot{\sigma}b' = \tilde{b}\dot{\tau}\tilde{b}'$ for some bs$\in B$, $\sigma, \tau \in S_n$. Letting $\beta = \tilde{b}^{-1}b$ and $\beta' = \tilde{b}'(b')^{-1}$, we have $\beta\dot{\sigma} = \dot{\tau}\beta'$. But $\dot{\sigma}$ and $\dot{\tau}$ are monomial matrices, and so the only possibility is that $\sigma = \tau$ and $\beta, \beta' \in H$. □

11.2.2 Shubert cells

We continue to write G for a complex semisimple Lie group, and $B = B_+ \subseteq G$ for a choice of upper Borel subgroup. Consider the quotient G/B. For $G = \mathrm{SL}(n, \mathbb{C})$, this is naturally isomorphic to the *flag variety*, which by definition is the collection of chains of subspaces $0 \subseteq V_1 \subseteq \cdots \subseteq V_{n-1} \subseteq \mathbb{C}^n$, where $\dim V_i = i$. Indeed, $\mathrm{SL}(n)$ includes all changes of bases, and the Borel changes the basis in each flag. Thus we will call G/B the *generalized flag variety*.

The Bruhat decomposition $G = \sqcup_{w\in W} U_- w B$ provides a decomposition of G/B into *Shubert cells*:

$$G/B = \bigsqcup_{w\in W} (U_- wB)/B \stackrel{\text{def}}{=} \bigsqcup_{w\in W} U_w$$

By definition, the cell U_w is the U_--orbit of $wB \in G/B$. Unpacking the definition, we find

$$U_w \cong \{u \in U_+ \text{ s.t. } \dot{w}u\dot{u}^{-1} \in U_-\}. \tag{11.2.2.1}$$

11.2.2.2 Theorem (Shubert cells admit almost coordinates)
There are almost coordinates *on U_w, meaning coordinates on a Zariski open subset, given as follows. Choose for each w a reduced word $\tilde{w} = s_{i_1} \dots s_{i_{\ell(w)}}$, i.e. a factorization of w into simple reflections which is minimal in length. (This minimal length $\ell(w)$ is the* length *of w.) Define $\phi_{\tilde{w}} : \mathbb{C}^{\ell(w)} \to U_w$ by*

$$(t_1, \dots, t_l) \mapsto \exp(t_1 e_{i_1}) \dots \exp(t_l e_{i_l}).$$

In particular, $\dim_{\mathbb{C}} U_w = \ell(w)$. □

The proof of Theorem 11.2.2.2 is not difficult but involves involved computations.

11.2.2.3 Example Consider the case $G = \mathrm{SL}(3)$ and $w = w_0 = (13) \in S_3$ the longest element. (The simple reflections are $s_1 = (12)$ and $s_2 = (23)$.) We choose $\dot{w}_0 = \left(\begin{smallmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{smallmatrix}\right)$. Recalling equation (11.2.2.1), we see that $U_{w_0} \cong U$.

There are two reduced words for w_0: $w_0' = s_1 s_2 s_1$ and $w_0'' = s_2 s_1 s_2$. They give the following almost coordinates:

$$\phi'(t_1,t_2,t_3) = \begin{pmatrix} 1 & t_1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & t_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & t_3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & t_1+t_3 & t_1 t_2 \\ 0 & 1 & t_2 \\ 0 & 0 & 1 \end{pmatrix} \tag{11.2.2.4}$$

$$\phi''(t_1,t_2,t_3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & t_1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & t_2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & t_3 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & t_2 & t_2 t_3 \\ 0 & 1 & t_1+t_3 \\ 0 & 0 & 1 \end{pmatrix} \tag{11.2.2.5}$$

Indeed, ϕ' is a coordinate system on the Zariski open subset of U consisting of the matrices $\left(\begin{smallmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{smallmatrix}\right)$ with $c \neq 0$, and ϕ'' is a coordinate system on the subset with $a \neq 0$. ♢

Do these different coordinate systems fit together to form a chart? This is a very deep subject, and quickly gets you into the topic of cluster algebras. There is a parallel story of almost-coordinates on G, using the Bruhat decomposition. The almost coordinate systems lead to actual coordinate systems on the *nonnegative split form*. Rather than giving a precise definition, we give an example: the *nonnegative* elements $\mathrm{SL}(n,\mathbb{R})_{\geq 0} \subset \mathrm{SL}(n,\mathbb{R})$ are those in which all minors are nonnegative.

Nonnegative matrices have been studied since the nineteenth century, and interesting results were discovered in the 50s and then forgotten. The subject has come back again in representation theory with results of Lusztig, Fomin and Zelevinsky, and others.

Let us consider nonnegativity from a more theoretical point of view. What is a "minor"? The first *fundamental representation* of $\mathrm{SL}(n)$ is $\mathbb{C}^n$, and the others are $\bigwedge^k \mathbb{C}^n$, where g acts diagonally: $g(x_1 \wedge \cdots \wedge x_i) = gx_1 \wedge \cdots \wedge gx_i$. Choose a basis $\{e_i\}_{i=1}^n$ in $\mathbb{C}^n$. Then $\{e_{i_1} \wedge \cdots \wedge e_{i_k}\}_{i_1 < \cdots < i_k}$ is a basis of $\bigwedge^k \mathbb{C}^n$. Consider the matrix coefficients in this basis of the action of $g \in \mathrm{SL}(n)$:

$$g(e_{i_1} \wedge \cdots \wedge e_{i_k}) = ge_{i_1} \wedge \cdots \wedge ge_{i_k} = \sum_{j_1 < \cdots < j_k} g^{j_1 \dots j_k}_{i_1 \dots i_k} e_{j_1} \wedge e_{j_k}.$$

Then $g^{j_1 \dots j_k}_{i_1 \dots i_k}$ nothing but *minor* of g corresponding to the $k \times k$ submatrix with colums $i_1, \dots, i_k$ and rows $j_1, \dots, j_k$. In summary, $g \in \mathrm{SL}(n)$ is *nonnegative* when it has nonnegative matrix elements in the monomial basis for the full fundamental representation $\bigwedge^\bullet \mathbb{C}^n$.

If we want to move beyond $G = \mathrm{SL}(n)$, we should not talk about the standard basis e_i of $\mathbb{C}^n$. Rather, we can talk about a weight basis, i.e. a basis diagonalizing the Cartan action. In a fundamental representation, the weight spaces are all (zero- or) one-dimensional, so the weight basis is unique up to rescaling. We will discuss was first discovered in the world of quantum groups, and applies to all G; c.f. Section 13.3. The $\mathrm{SL}(n)$ version is:

11.2.2.6 Theorem (Lusztig's canonical basis for $\mathrm{SL}(n)$)
In each finite-dimensional representation of $\mathrm{SL}(n)$*, there is a unique basis, called the* canonical basis *or the* crystal base*, in which all matrix elements of nonnegative $g \in \mathrm{SL}(n)$ are nonnegative.*

The nonnegative matrices do not form a subgroup of $\mathrm{SL}(n)$ — the inverse of a nonnegative matrix is not nonnegative — but they do form a subsemigroup. In terms of Hopf algebras, one finds a subbialgebra over $\mathbb{R}_{\geq 0}$. We will quantize $\mathrm{SL}(n)$, deforming the Poisson algebra to an associative algebra. The deformation of a real Lie group will not be a real associative algebra but rather a complex $*$-algebra. The natural question is: what exactly quantizes the positive part? The answer will be some sort of "non-compact quantum group with a positivity condition," axiomatized not in terms of Hopf algebras but in terms of well-structured bialgebras. There is some work on this, but the question is largely still open.

11.3 The dressing action

We continue to let G denote a complex semisimple Lie group, and $B \subseteq G$ its upper Borel. Let $K \subseteq G$ denote the compact real form of G and $T \subseteq K$ its maximal torus. There is an isomorphism of real manifolds $G/B \cong K/T$. (This follows from the Iwasawa decomposition Theorem 11.1.1.3.) We will write C_w for the cell U_w regarded as a real manifold. By Theorem 11.2.2.2, U_w is a complex manifold of complex dimension $\ell(w)$, and so C_w is a real manifold of real dimension $2\ell(w)$. Any time a real dimension is even, you should expect a natural symplectic structure. Indeed, we will see that K/T has a natural Poisson structure, and the C_w are the symplectic leaves.

11.3.1 Symplectic leaves

Since Poisson algebras are a Lie-algebraic enhancement of commutative algebra, the study of Poisson manifolds is a geometric analog of the study of Lie or associative algebras. One can ask: what is the analogue of modules? One reasonable answer is given by the topic of this section.

11.3.1.1 Definition *A* symplectic manifold *is a pair (M, ω) where M is a manifold (in your favorite category) and ω is a nondegenerate closed 2-form. Recall that a 2-form ω on M is* nondegenerate *if the corresponding map $\mathrm{T}M \to \mathrm{T}^*M$ that sends a vector $v \in \mathrm{T}_x M$ to the dual vector $\langle \omega(x), v \otimes (-) \rangle$ is an isomorphism of vector bundles. Equivalently, we can choose local coordinates and write $\omega(x) = \sum_{ij} \omega_{ij}(x) \mathrm{d}x^i \wedge \mathrm{d}x^j$; then ω is nondegenerate iff $\det(\omega_{ij}(x)) \neq 0$ $\forall x \in M$. The 2-form ω is* closed *if* $\mathrm{d}\omega = 0$.

11.3.1.2 Remark Symplectic manifolds were invented in the 19th century as a tool for studying classical mechanics. The name "symplectic" was introduced by Weyl in 1939, and comes from the

Greek word “symplektikos” meaning “braided together”; the corresponding Latin word “coplexus” gives the word “complex”. $\diamondsuit$

11.3.1.3 Example Let $M_{2n} = (\mathbb{R}^n) \oplus (\mathbb{R}^n)^*$, with coordinates p_i, q^i. Then $\omega = \sum_{i=1}^n \mathrm{d}p_i \wedge \mathrm{d}q^i$ is a symplectic structure on M_{2n}. $\diamondsuit$

11.3.1.4 Example Let N_n be a smooth n-dimensional manifold, and $M = \mathrm{T}^*N$, which is $2n$-dimensional. Choose a chart $U \in N$ so that $\mathrm{T}^*U \cong (\mathbb{R}^n)^* \times U$ and now $U \in \mathbb{R}^n$. Then in local coordinates $\omega = \sum_i \mathrm{d}p_i \wedge \mathrm{d}q^i$. $\diamondsuit$

11.3.1.5 Proposition *Any symplectic manifold is a Poisson manifold with* $\{f, g\} \stackrel{\text{def}}{=} \langle \omega^{-1}, \mathrm{d}f \wedge \mathrm{d}g\rangle$. *The theorem also works in the opposite direction: if a Poisson manifold is suitably nondegenerate, then the inverse of the bivector gives a symplectic structure.* $\square$

11.3.1.6 Remark By the *inverse of a form* ω^{-1} we mean the following: $\omega \in \Gamma^2(\bigwedge^2 \mathrm{T}^*M)$ is nondegenerate, so it gives an isomorphism $\omega : \mathrm{T}M \to \mathrm{T}^*M$. In local coordinates, if $v = \sum v^i \frac{\partial}{\partial x^i} \in TM$ and $\omega = \sum \omega_{ij} \mathrm{d}x^i \wedge \mathrm{d}x^j$, then $\omega(v) \stackrel{\text{def}}{=} \sum \omega_{ij} \mathrm{d}x^i v^j$. Then nondegeneracy implies that there is a bivector $\omega^{-1} \in \Gamma(\bigwedge^2 \mathrm{T}M)$ giving the opposite map $\omega^{-1} : \mathrm{T}^*M \to \mathrm{T}M$, which in coordinates is given by $\omega^{-1}(x) = \sum (\omega^{-1})^{ij}(x) \frac{\partial}{\partial x^i} \wedge \frac{\partial}{\partial x^j}$, where $(\omega^{-1})^{ij}$ is the inverse matrix to ω_{ij}. $\diamondsuit$

11.3.1.7 Definition *Let (M, p) be a Poisson manifold. A* Hamiltonian vector field *is a vector field v for which there exists a* Hamiltonian function $H \in \mathscr{C}(M)$ *such that $v = p\langle \mathrm{d}H\rangle$. Recall that $p \in \Gamma(\bigwedge^2 \mathrm{T}M)$, so $p(x) : \mathrm{T}^*_x M \to \mathrm{T}_x M$. In local coordinates $p(x)\langle \mathrm{d}H(x)\rangle = \sum_{ij} p^{ij}(x) \frac{\partial H}{\partial x^i} \frac{\partial}{\partial x^j}$. The flow lines for v are* Hamiltonian flow lines. *Classical* Hamiltonian mechanics *is the study of dynamical systems defined by Hamiltonian vector fields.*

11.3.1.8 Definition *Let (M, p) be a Poisson manifold. The* symplectic leaf *through $x \in M$ is the space of points on the manifold that you can reach by piecewise Hamiltonian flow starting at x.*

11.3.1.9 Theorem (Existence of symplectic leaves)
Each symplectic leaf in a Poisson manifold is an immersed submanifold. The Poisson bivector field is tangent to the leaf, and the restriction is nondegenerate and defines a symplectic structure on the leaf. $\square$

11.3.1.10 Remark The generic local situation is the following. A *Casimir function* for a Poisson manifold (M, p) is a function $f \in \mathscr{C}(M)$ that Poisson-commutes with everything: $\{f, g\} = 0$ for all $g \in \mathscr{C}(M)$. The common level sets for all Casimirs are, roughly, the symplectic leaves. The global situation is a little worse: symplectic leaves can wrap around the manifold irrationally (like the irrational line in a torus, although being odd dimensional this is not a symplectic leaf), and also level sets may be disconnected.

Even locally the Casimirs might not quite cut out the symplectic leaves if the rank of the Poisson tensor drops. For example, consider the two-dimensional nonabelian Lie algebra with basis x, y and bracket $[x, y] = x$, and consider x, y as coordinate functions on a Poisson manifold via Example 10.1.2.5. Then each connected component of $\{(x, y) \text{ s.t. } x \neq 0\}$ is a 2-dimensional symplectic leaf, and each point with $x = 0$ is itself a 0-dimensional symplectic leaf. Thus the only Casimirs are constants, because they are constants on each leaf. On the other hand, the 0-dimensional leaves "should" be cut out by the functions $f(y)\delta(x)$, where f ranges over smooth functions in one variable, and $\delta(x)$ is the non-existent Dirac delta function.

The analogy with representation theory is that, to construct an irreducible representation, we have to first fix all the central elements to be complex numbers. $\diamondsuit$

11.3.1.11 Example Consider $G = \mathrm{SU}(2)$, the group of two-by-two unitary matrices with det = 1. This is the compact real form of $\mathrm{SL}(2, \mathbb{C})$. Its Lie algebra $\mathfrak{g} = \mathrm{Lie}(G)$ is a real three-dimensional algebra, so $\mathfrak{g}^*$ is a Poisson manifold. What are the symplectic leaves? They are the co-adjoint orbits of G acting on $\mathfrak{g}^*$.

More precisely, G acts on $\mathfrak{g}$ by the adjoint action. If G is a matrix group (so $G \subseteq \mathrm{GL}(V)$), then $\mathfrak{g} = \mathrm{Lie}(G)$ is a matrix Lie algebra (so $\mathfrak{g} \subseteq \mathfrak{gl}(V)$), and so the adjoint action is $\mathrm{Ad}_g(x) \overset{\text{def}}{=} gxg^{-1}$. We define the *co-adjoint action* by

$$\mathrm{Ad}_g^*(l)(x) \overset{\text{def}}{=} l(\mathrm{Ad}_{g^{-1}}(x)), \quad l \in \mathfrak{g}^*, \ x \in \mathfrak{g}, \ g \in G.$$

A generic element of $\mathfrak{su}(2)$, i.e. a generic traceless Hermition 2×2 matrix, is of the form:

$$X = \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & -\alpha \end{pmatrix} \text{ where } \alpha \in \mathbb{R} \text{ and } \beta \in \mathbb{C}.$$

Because the Killing form is invariant and nondegenerate, it identifies $\mathfrak{su}(2)$ and its dual as SU(2)-modules. The Ad orbit through such a matrix is:

$$\mathrm{Ad}_G \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & -\alpha \end{pmatrix} = \left\{ u \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & -\alpha \end{pmatrix} u^{-1} : u \in \mathrm{SU}(2) \right\}.$$

The moduli space of conjugacy classes is the space of sets of eigenvalues, modulo permutation of eigenvalues. So:

$$\mathrm{Ad}_G \cong \left\{ \begin{pmatrix} \lambda & 0 \\ 0 & -\lambda \end{pmatrix} \text{ s.t. } \lambda \in i\mathbb{R} \right\} / \left\{ \begin{pmatrix} \lambda & 0 \\ 0 & -\lambda \end{pmatrix} \sim \begin{pmatrix} -\lambda & 0 \\ 0 & \lambda \end{pmatrix} \right\} \cong i\mathbb{R}_{\geq 0}. \tag{11.3.1.12}$$

The function $X \mapsto \mathrm{tr}(X^2) = 2\lambda^2$ is conjugation-invariant, and its level sets are precisely the Ad-orbits. Recall the *Pauli matrices*, which are a basis of $i\mathfrak{su}(2)$:

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

Setting $X = x_1 \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right) + x_2 \left(\begin{smallmatrix} 0 & -i \\ i & 0 \end{smallmatrix}\right) + x_3 \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, we have $\mathrm{tr}(X^2) = 2(x_1^2 + x_2^2 + x_3^2)$, and so the Ad^*-orbits are each two-dimensional spheres, except for the special orbit at the origin. $\diamondsuit$

We will eventually prove that the symplectic leaves of $\mathfrak{g}^*$ are precisely the Ad^*-orbits of G acting on $\mathfrak{g}^*$.

11.3.1.13 Example The orbits of the adjoint action of $\mathrm{SU}(n)$ on $\mathfrak{su}(n)$ are the sets of possible eigenvalues, which is to say $\mathbb{R}^{n-1}$ embedded in $\mathbb{R}^n$ as the hyperplane $\{\sum_i \lambda_i = 0\}$, modulo the permutation action by the symmetric group S_n. The orbits are picked out as the common level surfaces of the Casimirs $c_i = \mathrm{tr}(X^i)$, $i = 2, \ldots, n$. (Note that $c_1 = \mathrm{tr} = 0$ identically on $\mathfrak{su}(n)$.) $\Diamond$

11.3.1.14 Example We considered above some real Poisson manifolds; we consider now the complex case. Let $G = \mathrm{SL}(2, \mathbb{C})$ and $\mathfrak{g} = \mathfrak{sl}(2, \mathbb{C})$, which is non-canonically isomorphic to $\mathbb{C}^3$. The Killing form gives a canonical isomorphism $\mathfrak{g}^* \cong \mathfrak{g}$. The theorem that symplectic leaves are Ad^*-orbits still holds, and we identify these with adjoint orbits. Let's understand the orbit:

$$G \cdot \begin{pmatrix} a & b \\ c & -a \end{pmatrix} = \left\{ g \begin{pmatrix} a & b \\ c & -a \end{pmatrix} g^{-1} : g \in \mathrm{SL}(2, \mathbb{C}) \right\}$$

Such orbits come in various forms:

1. If the matrix is diagonalizable and non-zero, then the orbits is classified by the eigenvalues $\pm\lambda$. So the set of orbits through diagonalizable matrices is $(\mathbb{C} \smallsetminus \{0\})/\mathbb{Z}_2$. Each orbit is 2-dimensional over $\mathbb{C}$.

2. If the matrix is not diagonalizable, it must have a repeated eigenvalue, which must be 0 as the matrix is traceless. Up to conjugation, there is only one nondiagonalizable 2×2 matrix, namely $x = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$. What is the dimension of its G-orbit? To compute the dimension of an orbit you subtract the dimension of the stabilizer from the dimension of the group. The stabilizer of x is the group unipotent matrices

$$\exp(ax) = \left\{ \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \right\}$$

This group is one-complex-dimensional, and so the orbit is two-complex-dimensional.

3. There is the 0-dimensional orbit through 0. $\Diamond$

11.3.2 Constructing the dressing action

Let P be a Poisson manifold. The Poisson tensor π determines a pairing on $\Omega^1(P)$: if ω, ω' are 1-forms, then $\pi(\omega \wedge \omega') \in \mathscr{C}(P)$. Given a triple of 1-forms $\omega, \omega', \omega''$, we can define

$$\pi(\omega, (\omega' \wedge \omega'')) \stackrel{\text{def}}{=} \pi(\omega \wedge \omega') \wedge \omega'' - \pi(\omega \wedge \omega'') \wedge \omega'$$

It is not hard to check that the result depends only on the 2-form $\omega' \wedge \omega''$, and extends by linearity to $\Omega^2(P)$.

11.3.2.1 Theorem (Koszul's bracket on Ω^1)
1. If π is a Poisson tensor on P, then

$$[\omega, \omega']_{\Omega^1(P)} \stackrel{\text{def}}{=} \mathrm{d}\big(\pi(\omega, \omega')\big) + \pi(\omega, \mathrm{d}\omega') - \pi(\mathrm{d}\omega, \omega') \qquad (11.3.2.2)$$

is a Lie bracket on $\Omega^1(P)$.

2. *The map* $\pi^\# : \Omega^1(P) \to \mathrm{Vect}(P)$ *defined by* $\langle \eta, \pi^\# \omega \rangle = \pi(\omega, \eta)$ *is a Lie algebra homomorphism.*

Proof We rewrite equation (11.3.2.2) as

$$[\omega, \omega']_{\Omega^1(P)} = \mathrm{d}\big(\pi(\omega, \omega')\big) + i_{\pi^\# \omega}(\mathrm{d}\omega') - i_{\pi^\# \omega'}(\mathrm{d}\omega).$$

Note that if ω, ω' are both closed, then $[\omega, \omega']$ is exact. On exact 1-forms, we have $\pi^\# \mathrm{d}f = v_f$ is the Hamiltonian vector field for f. We leave as an exercise to check that

$$[f\omega, \omega'] = f[\omega, \omega'] + \pi(\mathrm{d}f, \omega')\omega. \tag{11.3.2.3}$$

We may now show Claim 1. Antisymmetry is clear. Since $[\mathrm{d}f, \mathrm{d}g] = \mathrm{d}\{f, g\}$, the Jacobi identity holds for exact 1-forms. Finally, the claim for arbitrary 1-forms follows from the exact case together with equation (11.3.2.3).

To prove Claim 2, we observe that it is automatic on exact forms — $\pi^\# \mathrm{d}f = v_f$ and $v_{\{f,g\}} = [v_f, v_g]$ — and then apply equation (11.3.2.3) to get the general case. See Exercise 9. □

We will use Theorem 11.3.2.1 to give a nonlinear version of the coadjoint action $G \curvearrowright \mathfrak{g}^*$. More precisely, given a dual pair of Poisson Lie groups G, G^*, we will build a local action $G^* \curvearrowright G$. A *local action* of G^* on G is a Lie algebra homomorphism $\mathfrak{g}^* \to \mathrm{Vect}(G)$. The following result was first proved in [Wei88]:

11.3.2.4 Theorem (Weinstein's description of the dressing action)
Let G *be a Poisson Lie group.*

1. *The space of left-invariant one-forms on* G *is a Lie subalgebra in* $\Omega^1(G)$ *with the bracket as in equation (11.3.2.2).*

2. *We trivialize the cotangent bundle* $\mathrm{T}^*G \cong G \times \mathfrak{g}^*$ *by right-translations. This gives a natural embedding* $\mathfrak{g}^* \hookrightarrow \Omega^1(G)$*, with image exactly the left-invariant forms, which is a homomorphism of Lie algebras. With the Poisson tensor* p*, we take* $p : \Omega^1(G) \to \mathrm{Vect}(G)$*. Then* $\mathfrak{g}^* \hookrightarrow \Omega^1(G) \xrightarrow{p} \mathrm{Vect}(G)$ *is a Lie algebra homomorphism.*

Proof Trivialize $\mathrm{T}^*G \cong G \times \mathfrak{g}^*$ by right-translations. Then $\omega, \omega' \in \Omega^1(G)$ are left-invariant exactly when, in the trivialization, they are constant maps $G \to \mathfrak{g}^*$. In the trivialization, the Poisson bivector field becomes a matrix of functions $\pi^{ij} : G \to \bigwedge^2 \mathfrak{g}$, and satisfying the cocycle identity:

$$\pi(gh) = \mathrm{Ad}_g^{\otimes 2}\, \pi(h) + \pi(g). \tag{11.3.2.5}$$

We have seen already (Theorem 11.3.2.1) that $[,]_{\Omega^1(G)}$ is a valid Lie bracket on $\Omega^1(G)$. Let $w, w' \in \mathfrak{g}^*$ and ω, ω' the corresponding "constant" 1-forms. We wish to calculate $[\omega, \omega']_{\Omega^1(G)}$. Since ω, ω' are constant, the only nonzero entry in the right-hand side of equation (11.3.2.2) is $\mathrm{d}\pi(\omega, \omega')$. Note that $\mathrm{d}\pi|_e$ is precisely the bracket on $\mathfrak{g}^*$, and so

$$\mathrm{d}\pi(\omega, \omega')(e) = [w, w']_{\mathfrak{g}^*}.$$

It thus suffices to show that $\mathrm{d}\pi(\omega,\omega')$ is constant when ω,ω' are. Consider differentiating equation (11.3.2.5) in h near $h=e$. We find:

$$\partial_i \pi^{jk}(g) = \mathrm{Ad}_g^{\otimes 2}\, \partial_i \pi^{jk}(e). \tag{11.3.2.6}$$

The left-hand side is the value at g of the tensor $\mathrm{d}\pi$, so equation (11.3.2.6) says simply that, in the trivialization, $\mathrm{d}\pi$ is constant. Thus so is $\mathrm{d}\pi(\omega,\omega')$. □

The map $\mathfrak{g}^* \to \mathrm{Vect}(G)$ is the (local) *dressing action* of G^* on G. Why is it called "dressing"? What does it dress? The name came from the theory of solitons, which are waves that propagate without loosing shape. In the 1970s, a theory of solitons was developed in terms of infinite-dimensional completely integrable systems. Some examples include the KdV equation and the nonlinear Schrodinger equation. Recall that in *Hamiltonian mechanics* you have a symplectic manifold (M,ω) and a *Hamiltonian* $H \in \mathscr{C}(M)$. You invert ω to get the Poisson structure, and define the *Hamiltonian flow* to be the vector field $v_H = \omega^{-1}(\mathrm{d}H)$. The system is *integrable* when you have a Lagrangian fibration $L \hookrightarrow M \to B$ such that H is constant on the fibers. This reduces the dimension of the action: v_H is tangent to the fibers. Integrability tends to arise when the system has a lot of symmetry. Solitons are created by subgroups: the action of a Poisson Lie group on the phase space of an integrable system "dresses" soliton solutions into multiple solutions.

Returning to Poisson Lie groups, recall from Definition 11.3.1.8 that to each $H \in \mathscr{C}(G)$ we define a Hamiltonian vector field $v_H = \pi(\mathrm{d}H)$, and for each $g \in G$ we get a symplectic manifold by taking every point you can reach by piece-wise Hamiltonian flow lines. Said another way, a submanifold $S \subset G$ is a symplectic leaf exactly when $\mathrm{T}_g S = \mathrm{Im}(\pi : \mathrm{T}_g^* G \to \mathrm{T}_g G)$ for every $g \in S$.

11.3.2.7 Corollary (Semenov-Tian-Shansky, Weinstein, Lu) *The symplectic leaves in (the image of G^* inside) a Poisson Lie group G are orbits of the dressing action of G^* on G.* □

The extension to the complement $G \smallsetminus G^*$ is due to Alekseev and Manin, and can be found in [EL07, Proposition 2.12].

11.3.3 Orbits of the dressing action

Our goal now is to give an algebraic description of these orbits. Let P be a Poisson manifold and $G \times P \to P$ an action of the Lie group G on P. If G is a Poisson Lie group, the action is called a *Poisson Lie action* if $G \times P \to P$ is a Poisson map. For example, a G action preserving the Poisson structure is a Poisson Lie action when G is equipped with the trivial Poisson structure. Another example is that if G is a Poisson Lie group and $H \subseteq G$ is a Poisson subgroup, then the left and right actions $H \times G \to G$ and $G \times H \to G$ are Poisson-Lie.

11.3.3.1 Proposition *Suppose $G \curvearrowright P$ is a Poisson Lie action. Then the algebra $\mathscr{C}(P)^G$ of G-invariant functions on P is a Poisson subalgebra of P.*

Proof Given a Poisson Lie action $\alpha : G \times P \to P$, consider the pullback of $f \in \mathscr{C}(P)$ along α, defined by $\alpha^* f(g,p) = f(\alpha(g,p))$. Since α is a Poisson map, $\alpha^*\{f,g\}_P = \{\alpha^* f, \alpha^* g\}_{G\times P}$. The function f is G-invariant exactly when $\alpha^* f = 1 \otimes f \in \mathscr{C}(G) \otimes \mathscr{C}(P) \subseteq \mathscr{C}(G \times P)$. (In the smooth category it is a dense subalgebra; in the algebraic category $\mathscr{C}(G) \otimes \mathscr{C}(P) = \mathscr{C}(G \times P)$.) So if

$f, g \in \mathscr{C}(P)^G$, $\alpha^*\{f,g\} = \{\alpha^* f, \alpha^* g\} = \{1 \otimes f, 1 \otimes g\} = \{1,1\} \otimes fg + 1 \otimes \{f,g\} = 1 \otimes \{f,g\}$, and so $\{f,g\} \in \mathscr{C}(P)^G$. $\square$

The algebra $\mathscr{C}(P)^G$ is nothing but the algebra of (global) functions on the quotient P/G. Suppose that quotient is a manifold. Then Proposition 11.3.3.1 is equivalent to the assertion that P/G is Poisson whenever $G \curvearrowright P$ is Poisson Lie.

11.3.3.2 Corollary *If $H \subseteq G$ is a Poisson Lie subgroup, then both $H\backslash G$ and G/H are Poisson manifolds with Poisson maps $G \to G/H$ and $G \to H\backslash G$.* $\square$

Suppose now that G is a connected simply-connected Poisson Lie group, with Lie bialgebra $(\mathfrak{g}, \delta)$. Then we can build the dual connected simply-connected Poisson Lie group, with Lie bialgebra $(\mathfrak{g}^*, \delta_*)$. We can also build the double $\mathfrak{g} \bowtie \mathfrak{g}^*$, which we can exponentiate to a connected simply-connected Lie group that we will call $\mathcal{D}(G)$. The embeddings $\mathfrak{g}, (\mathfrak{g}^*)^{\mathrm{op}} \hookrightarrow \mathfrak{g} \bowtie \mathfrak{g}^*$ extend to Poisson Lie group embeddings $i : G \hookrightarrow \mathcal{D}(G)$ and $j : (G^*)^{\mathrm{op}} \hookrightarrow \mathcal{D}(G)$. (The $(-)^{\mathrm{op}}$ denotes that the sign of the coalgebra / Poisson structure is reversed, so that $(\mathfrak{g}^*, \delta_*)^{\mathrm{op}} = (\mathfrak{g}^*, -\delta_*)$ and $(G^*, \pi_*)^{\mathrm{op}} = (G^*, -\pi_*)$.) Because $i(\mathfrak{g}) \cap j(\mathfrak{g}^*) = \{0\}$, we see that $\Sigma \overset{\text{def}}{=} i(G) \cap j(G^*)$ is a discrete subgroup of $\mathcal{D}(G)$.

11.3.3.3 Corollary *The maps $\mathcal{D}(G) \to \mathcal{D}(G)/i(G)$ and $\mathcal{D}(G) \to \mathcal{D}(G)/j(G^{*\mathrm{op}})$ are Poisson and commute with the left $\mathcal{D}(G)$-actions.* $\square$

Consider the following sequence of Poisson maps:

$$G \overset{i}{\hookrightarrow} \mathcal{D}(G) \to \mathcal{D}(G)/j(G^{*\mathrm{op}}).$$

$G^{*\mathrm{op}}$ acts on $\mathcal{D}(G)$ by left multiplication (as a subgroup of $\mathcal{D}(G)$), and so also acts $\mathcal{D}(G)/G^{*\mathrm{op}}$. We saw in Theorem 11.3.2.4 that $G^{*\mathrm{op}}$ acts on G via the Dressing action. These actions are arranged so that the above sequence is G^*-equivariant. Of course, the composition $G \to \mathcal{D}(G)/j(G^{*\mathrm{op}})$ is a local isomorphism, identifying $\mathcal{D}(G)/j(G^{*\mathrm{op}}) = G/\Sigma$, where $\Sigma \overset{\text{def}}{=} i(G) \cap j(G^*)$. In many cases Σ is trivial.

11.3.3.4 Corollary *Symplectic leaves of G are connected components of the $G^{*\mathrm{op}}$ orbits on $\mathcal{D}(G)/G^{*\mathrm{op}}$, i.e. double cosets $G^{*\mathrm{op}} x G^{*\mathrm{op}}$ for $x \in \mathcal{D}(G)$.* $\square$

11.4 Examples

11.4.1 Symplectic leaves of K

Consider the case $G = K$ the compact real form of $G_{\mathbb{C}}$, with the standard Poisson structure. Then $\mathcal{D}(G) = G_{\mathbb{C}}$ by Theorem 11.1.1.3, and $\mathcal{D}(K)/K^{*\mathrm{op}} \cong K$, and so we have a global action of $K^{*\mathrm{op}}$ on K. Symplectic leaves of K are preimages of the double cosets $K^{*\mathrm{op}}\backslash G/K^{*\mathrm{op}}$. Note that $K^* = AN$ is almost the Borel $B \subset G$ — they differ only by a copy of the compact Cartan $T \subseteq K$.

Recall Theorem 11.2.1.4: $G = \bigsqcup_{w \in W} BwB$. Let's choose $x = b\dot{w}b' \in BwB$ and work out its double coset for $K^* = AN$, which is to say $ANb\dot{w}b'AN$. Can we absorb b into AN? No, but we can write $b = tan$ where $t \in T = K \cap H$. Similarly $b' = t'a'n'$, but we can move the t' past $\dot{w}$, and absorb it into t. The end result is that the double coset $ANxAN$ is parameterized by the pair (w, t), assembling into the normalizer $N(T)$ of T:

11.4.1.1 Proposition *The space of double cosets $K^{*\mathrm{op}}\backslash\mathcal{D}(K)/K^{*\mathrm{op}}$ is naturally isomorphic to $N(T)$.* □

Choose $w \in W$, and let $N_w(T) = T\dot{w}$ denote its preimage in $N(T)$. Identify $\mathcal{D}(K)/K^{*\mathrm{op}} \cong K$, and let K_w denote the preimage of $N_w(T)$ along the map $K \to K^{*\mathrm{op}}\backslash K \cong K^{*\mathrm{op}}\backslash\mathcal{D}(K)/K^{*\mathrm{op}} \cong N(T)$.

11.4.1.2 Corollary *$K \cong \bigsqcup_{w\in W} K_w$, where each K_w is a homogeneous Poisson submanifold fibered over the torus $T \subset K$. The fibers are symplectic leaves isomorphic to Shubert cells of K/T.* □

11.4.1.3 Example When $K = \mathrm{SU}(2)$, the Weyl group is $S_2 = \{1, w_0\}$, and so $\mathrm{SU}(2)/T$ decomposes as $C_1 \sqcup C_{w_0}$ of dimensions 0 and 2. As a manifold, $\mathrm{SU}(2) \cong S^3$ and $\mathrm{SU}(2)/T \cong S^2 \cong \mathbb{CP}^1$. (This is the *Hopf fibration.*) The cell C_1 is the north pole $\{\infty\} \subset \mathbb{CP}^1$, and the cell C_{w_0} is $\mathbb{R}^2 \subseteq \mathbb{CP}^1$ with its translation-invariant symplectic form. ♢

11.4.2 Symplectic leaves of G^*

We have two embeddings:

$$G \overset{i}{\hookrightarrow} G \times G = \mathcal{D}(G) \overset{j}{\hookleftarrow} G^{*\mathrm{op}}$$

where i is the diagonal embedding, and $G^{*\mathrm{op}} = \{(b_+, b_-) \in B_+ \times B_- \text{ s.t. } [b_+]_0 = [b_-]_0^{-1}\}$, where $[b_\pm]_0$ is the Cartan part of the element. Then $j : (b_+, b_-) \mapsto (b_+, b_-) \in G \times G$, since $B_\pm \subseteq G$. We will drop the "op," since we will only care about the Poisson structure on G^* up to rescaling. The symplectic leaves of G^* are exactly the connected components of the preimages of the left-G-orbits in $(G \times G)/G$.

Note that $(G \times G)/G = \{(g_1, g_2)G \text{ s.t. } g_i \in G\} \xrightarrow{\sim} G$ via the map sending $(g_1, g_2)G \mapsto g_1 g_2^{-1}$. Under this isomorphism, the left action of G on $(G \times G)/G$ corresponds to the conjugation action. Thus:

11.4.2.1 Corollary *Symplectic leaves in G^* are connected components of preimages of conjugation orbits in G under the map of manifolds $G^* \overset{j}{\hookrightarrow} G \times G \to G$, where the second map sends $(g_1, g_2) \mapsto g_1 g_2^{-1}$.* □

11.4.2.2 Remark Recall that $G^* = B_+ \times_H B_-$. The composition $G^* \to G$ sends $(b_+, b_-) \mapsto b_+ b_-^{-1}$. On the Cartan $H \subset G^*$, this is the map $(h, h^{-1}) \mapsto h^2$. The kernel of this map has order 2^{rank}. ♢

11.4.3 Symplectic leaves of G

The symplectic leaves of G are the connected components of the preimages of the G^*-orbits in $\mathcal{D}(G)/G^*$ under the map $G \overset{i}{\hookrightarrow} G \times G = \mathcal{D}(G) \to \mathcal{D}(G)/G^*$. Since $G^* = B_+ \times_H B_- \subset B_+ \times B_- \subset G \times G$, the symplectic leaves of G are very closely related to the Bruhat decomposition $G = \sqcup BwB$.

There were in fact four Bruhat decompositions "$\bigsqcup BwB$", depending on whether we use the upper or lower Borel. To understand $G^*\backslash G\times G/G^*$, it is convenient to use the "$++$" decomposition for the first G and the "$--$" one for the second: $G\times G = \bigsqcup_{u,v}(B_+uB_+)\times(B_-vB_-)$. What we really want to understand is the subset $G^*\backslash \operatorname{diag}(G)/G^* \subseteq G^*\backslash G\times G/G^*$. Explicitly, for each $g\in G$, we consider $(g,g)\in G\times G$, which lives in the some particular $(B_+uB_+)\times(B_-vB_-)$ for $u,v\in W$.

11.4.3.1 Definition *For each pair $u,v\in W$, the corresponding* double Bruhat cell *is*

$$G^{uv} \overset{\text{def}}{=} B_+uB_+\cap B_-vB_-.$$

We can decompose $B_\pm = N_\pm H$, and the formula $G^* = B_+\times_H B_-$ unpacks to

$$G^* = \{(N_+h, N_-h^{-1}) \text{ s.t. } h\in H\}.$$

Suppose that $g\in G^{uv}$, so that $g = b_+\dot u b'_+ = b_-\dot v b'_-$. Note that if $b'_+ = h'n'_+$, then $\dot u h' n'_+ = h''\dot u n'_+$ for some h'', so we might as well write $g = b_+\dot u n'_+ = b_-\dot v n'_-$. Then the orbit $G^*(g,g)G^*$ becomes

$$G^*(g,g)G^* = \big\{(N_+hb_+\dot u n_+N_+h', N_-h^{-1}b_-\dot v n_-N_-(h')^{-1}) \text{ s.t. } h,h'\in H\big\}.$$

Write $b_\pm = n_\pm h_\pm$, commute all hs towards the left — $hN_\pm = N_\pm h$ and $\dot u h \overset{\text{def}}{=} h_u\dot u$ — and absorb the ns in the Ns. Then

$$G^*(g,g)G^* = \big\{(hh_+h'_uN_+\dot uN_+, h^{-1}h_-(h'_v)^{-1}N_-\dot vN_-)\big) \text{ s.t. } h,h'\in H\big\}.$$

What is $N_+\dot uN_+$? Recall the discussion from Section 11.2.1. When $u=1$, it is as small as N_+, but when $u=w_0$ is the longest word, $N_+\dot uN_+$ is quite large. In general, we can try to pull as much of the left copy of N_+ across $\dot u$ as possible. We'll get stuck when we hit

$$N_+^u \overset{\text{def}}{=} \{n\in N_+ \text{ s.t. } \dot u^{-1}n\dot u\in N_-\},$$

and so $N_+\dot uN_+ = N_+^u\dot uN_+$.

So far, we have that for $g = n_+h_+\dot un'_+ = n_-h_-\dot vn'_-\in G^{u,v}$, that

$$G^*(g,g)G^* = \big\{hh'_u\,N_+^u(h_+\dot u)N_+, (hh'_v)^{-1}\,N_-^v(h_-\dot v)N_- \text{ s.t. } h,h'\in H\big\} \tag{11.4.3.2}$$

We moved the $h_\pm$s back next to $\dot u,\dot v$, because the choice of $\dot u\in N(H)$ effects the coordinate $h_\pm$.

There's still some redundancy in equation (11.4.3.2). Both h,h' range over H, so we might as well change coordinates to:

$$h_1 = hh'_u,\quad h_2 = hh'_v.$$

How are these related? We have a map $H\times H\to H\times H, (h,h')\mapsto(hh'_u,hh'_v)$. Take logarithms: we have the map $\mathfrak h\times\mathfrak h\to\mathfrak h\times\mathfrak h, (x,x')\mapsto(x+x'_u, x+x'_v)$. The image of this map consists of all (y,z) where $z-y$ is in the image of $x'\mapsto x'_v - x'_u$, which is isomorphic (via u^{-1}) to the image of $x'\mapsto x'_{vu^{-1}} - x'$. I.e. the image of $H\times H\to H\times H, (h,h')\mapsto(hh'_u,hh'_v)$ is isomorphic to $H\times\exp(\operatorname{Im}(vu^{-1}-\mathrm{id}))\subseteq H\times H$.

There is no more redundancy. Counting dimensions, we find, for $g \in G^{u,v}$:

$$\dim G^*(g,g)G^* = \dim(N_+^u) + \dim(N_-^v) + \dim \operatorname{Im}(vu^{-1} - \mathrm{id})$$
$$= \dim(N_+^u) + \dim(N_-^v) + \operatorname{rank} - \dim \mathfrak{h}^{vu^{-1}}, \quad (11.4.3.3)$$

where of course $\operatorname{rank} = \dim(\mathfrak{h}) = \dim(H)$ and $\mathfrak{h}^{vu^{-1}}$ is the subspace of $\mathfrak{h}$ fixed by $vu^{-1} \in W$.

Theorem 11.2.2.2 implies that $\dim N_+^u = \ell(u)$ and $\dim N_-^v = \ell(v)$, the lengths of the words u and v. These can be any numbers. On the other hand, $G^*(g,g)G^*$ is a symplectic leaf in G, and so $\dim G^*(g,g)G^*$ is even. This is a miracle:

11.4.3.4 Corollary $\dim \operatorname{Im}(vu^{-1} - \mathrm{id}) = \ell(u) + \ell(v) \mod 2$. □

11.4.3.5 Theorem (Symplectic geometry of double Bruhat cells)

1. *The double Bruhat cell $G^{u,v}$ is fibered over $H^{vu^{-1}} = \exp(\ker(vu^{-1} - \mathrm{id}))$.*

2. *Each fiber is isomorphic to $N_+^u \times N_-^v \times \exp(\operatorname{Im}(vu^{-1} - \mathrm{id}))$.*

3. *$G^{u,v}$ is a* homogeneous Poisson variety*: a fiber bundle whose fibers are symplectic leaves.* □

Exercises

1. Show that the symplectic structure described in Example 11.3.1.4 does not depend on the choice of coordinates. Find a natural 1-form θ on $M = \mathrm{T}^*N$ such that $\omega = \mathrm{d}\theta$.

2. Prove Proposition 11.3.1.5. Hint: the closure $\mathrm{d}\omega = 0$ is equivalent to the Jacobi for $\{,\}$.

3. Show that the symplectic leaves of $\mathfrak{su}(2)^*$ are (the images under the Killing form of) the spheres constructed in Example 11.3.1.11.

4. (a) Describe the adjoint orbits in $\mathfrak{sl}(3,\mathbb{C})$, and the dimension of each orbit.
   (b) Describe the adjoint orbits in $\mathfrak{gl}(n,\mathbb{C})$.
   (c) Describe the adjoint orbits for the group of matrices of the form:
$$\begin{pmatrix} a & a_1 & b_1 \\ 0 & b & c_1 \\ 0 & 0 & c \end{pmatrix}$$

5. Prove, as claimed in Example 11.1.0.4, that $\mathrm{SU}(n)^*$ is the group of upper-triangular complex matrices with determinant 1 and purely positive-real diagonal. Hint: the Lie algebra $\mathfrak{su}(n)^*$ can be read from the standard structure on $\mathfrak{sl}(n)$, and is roughly two copies of the upper Borel; now you just have to represent $\mathfrak{su}(n)^*$ in matrices and exponentiate.

6. Find the action of

$$\dot{s}_i = \left(\begin{array}{ccc|cc|ccc} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ \hline & & & 0 & 1 & & & \\ & & & -1 & 0 & & & \\ \hline & & & & & 1 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{array}\right)$$

on $\mathfrak{sl}_n$. On generators: $T_i \stackrel{\text{def}}{=} \mathrm{Ad}_{\dot{s}_i}$, and find $T_i(H_j = e_{jj} - e_{j+1,j+1})$, $T_j(e_{i,i+1} = e_i)$, and $T_j(e_{i+1,i} = f_i)$.

7. Construct the Bruhat decomposition (Theorem 11.2.1.4) by hand for $G = \mathrm{SL}(2)$.

8. Describe the closures $\overline{BwB}$.

9. Check equation (11.3.2.3), which asserts a version of Leibniz's rule for the bracket from equation (11.3.2.2). Use this to finish the proof of Theorem 11.3.2.1: show that the Jacobi identity and that $\pi^\#$ is a homomorphism hold for arbitrary 1-forms if they hold for exact 1-forms.

10. Given a compact semisimple Lie group K, with its standard Poisson Lie structure, provide an explicit description of the symplectic leaves of $K^* = AN$. Hint: in addition to the Iwasawa decomposition $G = KAN$, there is also the *polar decomposition* $G = K\mathscr{H}$, where $\mathscr{H}$ is the space of Hermitian matrices.

11. Let σ denote the involution of G whose fixed points are the compact real form K, and let $g \mapsto g^*$ denote the involution whose fixed points are the split real form. Show that $(N_+^u)^* = N_-^u$ and that $\sigma(N_+^u) = N_+^{u^{-1}}$. Conclude that the only symplectic leaves of G which have σ-fixed points are those in $G^{u,u}$. Identify $(G^{u,u})^\sigma = K_u$, and identify the fibration $(G^{u,u})^\sigma \to H^\sigma$ from Theorem 11.4.3.5 with the projection $K_u \to T$. Finally, identify the symplectic leaves in $(G^{u,u})^\sigma$ with the Shubert cell C_u.

Chapter 12

Quantum SL(2)

We turn now to the study of quantum groups. In this chapter, we focus on the case of SL(2), and in the next chapter we address the higher-rank case. Most of the story will be on the quantum universal enveloping algebra and its representation theory, but historically the first thing that was defined was a quantization of the *group* SL(2), meaning a quantization of its algebra of functions, and so we will begin there.

12.1 Quantum groups $\mathscr{C}_q(\mathrm{GL}(2))$ and $\mathscr{C}_q(\mathrm{SL}(2))$

12.1.1 Quantum matrices

Following Manin's approach, we will define quantizations of GL(2) and SL(2) by thinking of them as "symmetry groups" of a "quantum plane," just as the classical group GL(2) and SL(2) act on the classical plane. Pick $q \in \mathbb{C}^\times$. The *quantum plane* is the "noncommutative space" whose algebra of functions is $\mathscr{C}_q(\mathbb{C}^2) = \mathbb{C}_q[x,y] = \mathbb{C}\langle x,y\rangle/(yx - qxy)$. We sloganize the relation $yx = qxy$: the generators "q-mute" rather than commute.

We must think about all spaces in terms of their algebras of functions. For instance, the classical matrix multiplication $\mathrm{Mat}(2) \times \mathbb{C}^2 \to \mathbb{C}^2$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix} \tag{12.1.1.1}$$

in terms of algebras of functions is really a coaction $\mathscr{C}(\mathbb{C}^2) \to \mathscr{C}(\mathrm{Mat}(2)) \otimes \mathscr{C}(\mathbb{C}^2)$. Indeed, equation (12.1.1.1) tells you the coaction on generators: it is

$$x \mapsto ax + by = a \otimes x + b \otimes y, \quad y \mapsto cx + dy = c \otimes x + d \otimes y, \tag{12.1.1.2}$$

where x, y are the generators of $\mathscr{C}(\mathbb{C}^2) = \mathbb{C}[x,y]$ and a, b, c, d are the generators of $\mathscr{C}(\mathrm{Mat}(2,\mathbb{C})) = \mathbb{C}[a,b,c,d]$.

"Quantum 2×2 matrices" should act on the quantum plane, or rather their algebra of functions should coact. We will therefore try to build an algebra corresponding to the "space" of quantum matrices, which would be a noncommutative algebra $\mathscr{C}_q(\mathrm{Mat}(2))$ generated by the "coordinate functions" a, b, c, d. Let's keep equation (12.1.1.2) verbatim — matrix multiplication is already noncommutative. What relations must a, b, c, d satisfy in order to make it an algebra homomorphism? Well, we must have $y'x' = qx'y'$, where $x' = a \otimes x + b \otimes y$ and $y' = c \otimes x + d \otimes y$. Expanding in the monomials $x^k y^l$ — they are still a basis for $\mathscr{C}_q(\mathbb{C}^2)$ — and matching coefficients gives:

$$ac = qca \tag{12.1.1.3}$$

$$bd = qdb \tag{12.1.1.4}$$

$$qad + bc = q(qcb + da) \tag{12.1.1.5}$$

We also declare that quantum matrices should act from the right on row vectors, which is to say we declare that the map $\mathscr{C}_q(\mathbb{C}^2) \to \mathscr{C}_q(\mathrm{Mat}(2)) \otimes \mathscr{C}_q(\mathbb{C}^2)$ given on generators by

$$\begin{pmatrix} x & y \end{pmatrix} \mapsto \begin{pmatrix} x'' & y'' \end{pmatrix} = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a \otimes x + b \otimes y & c \otimes x + d \otimes y \end{pmatrix}$$

should be an algebra homomorphism. This leads to the following relations:

$$ab = qba \tag{12.1.1.6}$$

$$cd = qdc \tag{12.1.1.7}$$

$$qad + cb = q(qbc + da) \tag{12.1.1.8}$$

Equations (12.1.1.5) and (12.1.1.8) can be rewritten as:

$$bc = cb \tag{12.1.1.9}$$

$$ad - da = (q - q^{-1})bc \tag{12.1.1.10}$$

12.1.1.11 Definition *The noncommutative space of* 2×2 quantum matrices *is the associative algebra* $\mathscr{C}_q(\mathrm{Mat}(2)) = \mathbb{C}\langle a, b, c, d\rangle / ($*equations (12.1.1.3), (12.1.1.4), (12.1.1.6), (12.1.1.7), (12.1.1.9), and (12.1.1.10)*$)$.

Recall that a *bialgebra* is a unital associative algebra A together with a comultiplication $\Delta : A \to A \otimes A$ which is an algebra homomorphism and which is counital and coassociative. Bialgebras are "quantum monoids" because if X is a monoid, then $\mathscr{C}(X)$ is a commutative bialgebra. We traditionally use *Sweedler's notation* for the comultiplication: if $a \in A$, we will write $\Delta(a) = \sum a_1 \otimes a_2 \in A \otimes A$ instead of, say, $\sum_i a_{1,i} \otimes a_{2,i}$, and so for example if $m : A \otimes A \to A$ is the multiplication, then the composition $m \circ \Delta$ takes a to $\sum a_1 a_2$. We will write $\epsilon : A \to \mathbb{C}$ for the counit. Its defining equation is $a = \sum a_1 \, \epsilon(a_2) = \sum \epsilon(a_1) \, a_2$ for all $a \in A$.

If a bialgebra is a "quantum monoid," then a "quantum group" is a *Hopf algebra*: a bialgebra A which admits an *antipode* quantizing the map $x \mapsto x^{-1}$, i.e. a linear map $\mathcal{S} : A \to A$ (which will be automatically an algebra- and coalgebra-antiautomorphism) such that $\sum \mathcal{S}(a_1) a_2 = \sum a_1 \mathcal{S}(a_2) = \epsilon(a) \, 1_A$ for all $a \in A$, where 1_A denotes the unit.

12.1.1.12 Lemma $\mathscr{C}_q(\mathrm{Mat}(2))$ *is a bialgebra, where the coalgebra structure is given on generators by the usual matrix multiplication:*

$$\Delta : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} a \otimes a + b \otimes c & a \otimes c + b \otimes d \\ c \otimes a + d \otimes b & c \otimes b + d \otimes d \end{pmatrix} \tag{12.1.1.13}$$

The counit is $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \mapsto \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. $\square$

12.1.2 The quantum determinant

The group GL(2) is the open subspace of Mat(2) consisting of matrices with nonzero determinant, and SL(2) is the closed subspace with determinant 1. To quantize these groups, we need a "quantum determinant."

12.1.2.1 Lemma / Definition *The* quantum determinant $\det_q = ad - qbc = da - q^{-1}bc \in \mathscr{C}_q(\mathrm{Mat}(2))$ *is central and* grouplike *in the sense that* $\Delta(\det_q) = \det_q \otimes \det_q$. □

It follows that the ideal generated by $\det_q - 1$ is a bialgebra ideal, meaning that $\mathscr{C}_q(\mathrm{Mat}(2))/\langle \det_q - 1\rangle$ is a bialgebra.

12.1.2.2 Lemma / Definition Quantum SL(2) *is the noncommutative space whose algebra of functions is* $\mathscr{C}_q(\mathrm{SL}(2)) \stackrel{\text{def}}{=} \mathscr{C}_q(\mathrm{Mat}(2))/\langle \det_q - 1\rangle$. *It is a Hopf algebra with antipode given on generators by*

$$\mathcal{S} : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} d & -qb \\ -q^{-1}c & a \end{pmatrix}. \qquad \square$$

12.1.2.3 Remark The antipode is unique if it exists, just like x^{-1} is unique if it exists. So the nontrivial statement in Lemma/Definition 12.1.2.2 is that $\mathcal{S}$ exists. Unlike in the classical case, this S is not an involution. Rather, $\mathcal{S}^2$ is matrix-conjugation by $\left(\begin{smallmatrix} q & 0 \\ 0 & q^{-1} \end{smallmatrix}\right)$. ◊

12.1.2.4 Definition Quantum GL(2) *is the noncommutative space whose algebra of functions is* $\mathscr{C}_q(\mathrm{GL}(2)) \stackrel{\text{def}}{=} \mathscr{C}_q(\mathrm{Mat}(2))[\det_q^{-1}]$.

Let us discuss the ring-theoretic properties enjoyed by $\mathscr{C}_q(\mathrm{GL}(2))$. It is a localization of $\mathscr{C}_q(\mathrm{Mat}(2))$. *Localization* can be done in noncommutative land, but it is difficult and technical. However, when localizing at a central element, everything is easy, and $\mathscr{C}_q(\mathrm{Mat}(2))[\det_q^{-1}] \stackrel{\text{def}}{=} \mathscr{C}_q(\mathrm{Mat}(2))[t]/\langle t \det_q - 1\rangle$. Our goal will be to show that $\mathscr{C}_q(\mathrm{GL}(2))$ is a *quantized coordinate ring*, i.e. a notherian affine domain. By "noetherian" we mean both left- and right-notherian. By "domain" we mean it has no zero divisors. By "affine" we mean finitely generated. The third is the easiest: $\mathscr{C}_q(\mathrm{GL}(2))$ is manifestly finitely generated.

12.1.2.5 Definition *Let R be a ring and $\sigma : R \to R$ a ring endomorphism. A (left)* σ-derivation $\delta : R \to R$ *is an abelian group endomorphism satisfying*

$$\delta(xy) = \delta(x)\,y + \sigma(x)\,\delta(y).$$

If you take $\sigma = \mathrm{id}$, you get the usual notion of derivation.

Given such R, σ, δ, we define the skew polynomial ring, *also called the* Ore extension, *to be*

$$R[x;\sigma,\delta] \stackrel{\text{def}}{=} R\langle x\rangle/\langle xr = \sigma(r)x + \delta(r), r \in R\rangle. \tag{12.1.2.6}$$

Equation (12.1.2.6) allows all xs to be moved to the right of all rs. It follows in particular that $R[x;\sigma,\delta] \cong \bigoplus_{n=0}^{\infty} R \cdot x^n$ as a left R-module.

12.1.2.7 Proposition *$R[x;\sigma,\delta]$ is a domain if and only if R is a domain and σ is one-to-one.*

Proof Any nonzero element of $R[x;\sigma,\delta]$ has a leading order term of the form rx^n. On leading order terms, the multiplication is

$$rx^n\, sx^m = r\sigma^n(s)\, x^{n+m} + \text{lower order}.$$

So $rx^n\, sx^m \neq 0$ provided $r\sigma^n(s) \neq 0$, which happens provided R has no zero divisors and σ has no kernel. The "only if" direction is an easy exercise. □

12.1.2.8 Theorem (Noncommutative Hilbert Basis Theorem)
Suppose R is notherian and σ is an automorphism. Then $R[x;\sigma,\delta]$ is notherian. □

The same theorem holds with "notherian" replaced by just left- or right-notherian. The theorem quickly fails if σ is not an automorphism. The usual commutative Hilbert basis theorem is the special case when R is commutative and $(\sigma,\delta) = (\mathrm{id}, 0)$.

12.1.2.9 Corollary *$\mathscr{C}_q(\mathrm{Mat}(2))$ is notherian.*

Proof Recognize $\mathscr{C}_q(\mathrm{Mat}(2)) \cong \mathbb{C}[a;\sigma_1,\delta_1][b;\sigma_2,\delta_2][c;\sigma_3,\delta_3][d;\sigma_4,\delta_4]$ where

$$\begin{aligned}
&\sigma_1 = \mathrm{id}, && \delta_1 = 0,\\
&\sigma_2 : a \mapsto q^{-1}a, && \delta_2 = 0,\\
&\sigma_3 : (a,b) \mapsto (q^{-1}a, b), && \delta_3 = 0,\\
&\sigma_4 : (a,b,c) \mapsto (a, q^{-1}b, q^{-1}c), && \delta_4 : (a,b,c) \mapsto ((q-q^{-1})bc, 0, 0).
\end{aligned}$$ □

12.1.2.10 Corollary *$\mathscr{C}_q(\mathrm{GL}(2))$ and $\mathscr{C}_q(\mathrm{SL}(2))$ are notherian domains.*

Proof Quotients of notherian rings are notherian. $\mathscr{C}_q(\mathrm{SL}(2))$ is a quotient of $\mathscr{C}_q(\mathrm{Mat}(2))$ and $\mathscr{C}_q(\mathrm{GL}(2))$ is a quotient of $\mathscr{C}_q(\mathrm{Mat}(2))[t;\mathrm{id},0]$. □

12.2 $\mathcal{U}_q\mathfrak{sl}(2)$

Having quantized SL(2) in the previous section, our goal now is to construct its Lie algebra, or rather its "quantized universal enveloping algebra" $\mathcal{U}_q\mathfrak{sl}(2)$. How will we know we have succeeded? We will need a version of Proposition 3.2.4.4 identifying $\mathcal{U}_q\mathfrak{sl}(2)$ with differential operators on $\mathrm{SL}_q(2)$; our version is Theorem 12.2.4.3. We first state the answer:

12.2.0.1 Definition *Choose $q \in \mathbb{C}^\times$ such that $q \neq \pm 1$. The algebra $\mathcal{U}_q\mathfrak{sl}(2)$ is generated by elements $K^{\pm 1}, E, F$ with defining relations*

$$KEK^{-1} = q^2 E, \tag{12.2.0.2}$$

$$KFK^{-1} = q^{-2} F, \tag{12.2.0.3}$$

$$[E,F] = \frac{K-K^{-1}}{q-q^{-1}}. \tag{12.2.0.4}$$

$\mathcal{U}_q\mathfrak{sl}(2)$ is an example of a *quantized universal enveloping algebra*. Morally speaking, its "limit" as $q \mapsto 1$ is $\mathcal{U}\mathfrak{sl}(2)$. Of course, you cannot simply specialize $q = 1$ in equation (12.2.0.4). The idea is to write $K = q^H$. Then the first two relations simply exponentiate the relations $[H, E] = 2E$ and $[H, F] = -2F$. Formally applying L'Hospital's rule to the third relation gives $[E, F] = H + O(q-1)$, where $O(q-1)$ is something vanishing in the limit as $q \mapsto 1$.

12.2.0.5 Proposition *$\mathcal{U}_q\mathfrak{sl}(2)$ is an affine noetherian domain. The set $\{E^aK^bF^c\}$, with $a, c \in \mathbb{N}$ and $b \in \mathbb{Z}$, is a PBW basis.*

Proof Recognize $\mathcal{U}_q\mathfrak{sl}(2) \cong \mathbb{C}[K^{\pm 1}][E; \sigma_1, \delta_1][F; \sigma_2, \delta_2]$ where

$$\sigma_1 : K \mapsto q^{-2}K \qquad\qquad \delta_1 = 0$$

$$\sigma_2 : (K, E) \mapsto \left(q^2K, E\right) \qquad\qquad \delta_2 : (K, E) \mapsto \left(0, -\frac{K - K^{-1}}{q - q^{-1}}\right).$$

Apply Proposition 12.1.2.7 and Theorem 12.1.2.8. ☐

12.2.1 When q is not a root of unity

When q is not a root of unity, the representation theory of $\mathcal{U}_q\mathfrak{sl}(2)$ is almost identical to that of $\mathcal{U}\mathfrak{sl}(2)$. We will use the following notation:

12.2.1.1 Definition *The* quantum integers*, also called* q-numbers*, are* $[n]_q = 1 + \cdots + q^{n-1} = \frac{q^n - 1}{q - 1}$. *The* q-factorial *is* $[n]_q! = [1]_q \cdots [n]_q$ *and the* q-binomial coefficients *are* $\binom{n}{k}_q = \frac{[n]_q!}{[k]_q!\,[n-k]_q!}$.

Then, just as in the classical case, the first important representations we encounter are the *Verma modules* $M(\lambda)$, $\lambda \in \mathbb{C}^\times$, defined by saying that $M(\lambda)$ is generated by a *highest weight vector* v_0 on which K acts by λ. Explicitly, $M(\lambda) = \mathrm{span}\{v_i \text{ s.t. } i \geq 0\}$ and the actions are

$$Fv_n = v_{n+1}, \tag{12.2.1.2}$$

$$Kv_n = q^{-2n}\lambda v_0, \tag{12.2.1.3}$$

$$Ev_n = [n]_q \, \frac{q^{-n+1}\lambda - q^{n-1}\lambda^{-1}}{q - q^{-1}} \, v_{n-1}. \tag{12.2.1.4}$$

These relations follow from declaring that $v_n \overset{\text{def}}{=} F^n v_0$ and that $Kv_0 = \lambda v_0$ and $Ev_0 = 0$.

We are mostly interested in the finite-dimensional representations of $\mathcal{U}_q\mathfrak{sl}(2)$, and so we let V be finite-dimensional for the remainder of this section. Note first that if V is a finite-dimensional $\mathcal{U}_q\mathfrak{sl}(2)$-representation, then in particular it is a finite-dimensional $\mathbb{C}[K^{\pm 1}]$ representation, and so K acts diagonalizably on V. Suppose $v \in V$ is a K-eigenvector with eigenvalue λ — let's call it a *weight vector*, since we are thinking of K as "q^H." Then $KEv = \lambda q^2 Ev$ and $KFv = \lambda q^{-2}Fv$, and so we start to see the picture from equation (5.2.0.6).

To complete the picture, we need to know that q is not a root of unity. Then the action of E on the weight vectors has no cycles (as $q^{2m}\lambda \neq \lambda$ for $m \neq 0$), and so, if V is to be finite-dimensional, there must be some weight vector v_0 with $Ev_0 = 0$. If V is irreducible, then it is a quotient of some $M(\lambda)$.

Finally, if V is to be finite-dimensional, then it must also have a lowest-weight vector $v_n = F^n v_0$, meaning that $v_n \neq 0$ but $v_{n+1} = F v_n = 0$. Now inspect equation (12.2.1.4). We must have $0 = E v_{n+1} = (\text{coefficient}) v_n$, and $v_n \neq 0$. Since q is not a root of unity, $[n+1]_q \neq 0$. So we must have $q^{-(n+1)+1}\lambda - q^{(n+1)-1}\lambda^{-1} = 0$, which happens exactly when $\lambda = \pm q^n$.

Rescaling $w_n = v_n/[n]_q!$, we have proved:

12.2.1.5 Theorem (Rep($\mathcal{U}_q\mathfrak{sl}(2)$) when q is not a root of unity)
All finite-dimensional representations of $\mathcal{U}_q\mathfrak{sl}(2)$ are completely reducible. The finite-dimensional irreps are indexed by pairs (n, ϵ) where $n \in \mathbb{N}$ and $\epsilon \in \{\pm 1\}$. The representation $V_{n,\epsilon}$ is $(n+1)$-dimensional, with basis weight basis $\{w_0, \dots, w_n\}$ and actions

$$E = \begin{pmatrix} 0 & [n]_q & & & \\ & 0 & [n-1]_q & & \\ & & \ddots & \ddots & \\ & & & 0 & [1]_q \\ & & & & 0 \end{pmatrix} \qquad F = \epsilon \begin{pmatrix} 0 & & & & \\ [1]_q & 0 & & & \\ & [2]_q & 0 & & \\ & & \ddots & \ddots & \\ & & & [n]_q & 0 \end{pmatrix} \qquad K = \epsilon \begin{pmatrix} q^n & & & & \\ & q^{n-2} & & & \\ & & q^{n-4} & & \\ & & & \ddots & \\ & & & & q \end{pmatrix}$$

12.2.1.6 Remark Recall the quantum plane $\mathbb{C}_q[x,y] = \mathbb{C}\langle x, y\rangle/(yx = qxy)$. Just like $\mathcal{U}\mathfrak{sl}(2)$ acts on $\mathbb{C}[x,y]$ by differential operators ($E = x\partial_y$, etc.), $\mathcal{U}_q\mathfrak{sl}(2)$ acts on $\mathbb{C}_q[x,y]$ by "quantum differential operators." Indeed, define $\delta = \delta_x$ on $\mathbb{C}[x]$ by

$$\delta f(x) \stackrel{\text{def}}{=} \frac{f(qx) - f(q^{-1}x)}{qx - q^{-1}x}.$$

Then $\mathcal{U}_q\mathfrak{sl}(2)$ acts on $\mathbb{C}_q[x,y]$ by

$$E = x\delta_y, \quad F = y\delta_x, \quad K : (x,y) \mapsto (qx, q^{-1}y).$$

These operators preserve homogeneous degrees, and $\mathbb{C}_q[x,y] = \bigoplus_n V_{n,+}$. ◊

12.2.1.7 Remark There is an isomorphism $\mathcal{U}_q\mathfrak{sl}(2) \cong \mathcal{U}_{-q}\mathfrak{sl}(2)$ sending $(K, E, F) \mapsto (-K, E, F)$. This isomorphism sends $V_{n,\epsilon}$ to $V_{n,\epsilon'}$ where $\epsilon' = (-1)^n\epsilon$. ◊

12.2.1.8 Theorem (Harish-Chandra isomorphism for $\mathcal{U}_q\mathfrak{sl}(2)$)
The quadratic Casimir

$$C = EF + \frac{q^{-1}K + qK^{-1}}{(q-q^{-1})^2} = FE + \frac{qK + q^{-1}K^{-1}}{(q-q^{-1})^2}$$

is central in $\mathcal{U}_q\mathfrak{sl}(2)$. The full center is

$$\mathcal{Z}(\mathcal{U}_q\mathfrak{sl}(2)) = \mathbb{C}[C].$$

Proof The first sentence is a straightforward if tedious calculation. By considering weights under the K-action, it is clear that the centralizer of K is $\mathbb{C}[K^{\pm 1}, EF] = \mathbb{C}[K^{\pm 1}, FE] = \mathbb{C}[K^{\pm 1}, C]$. Any $u \in \mathbb{C}[K^{\pm 1}, EF]$ can be written uniquely in the form

$$u = \sum_{i=0}^{N} F^i P_i(K) E^i$$

where the P_i are Laurent polynomials. Define the *Harish-Chandra homomorphism* $\mathbb{C}[K^{\pm 1}, EF] \to \mathbb{C}[x^{\pm 1}]$ by

$$\theta_u \stackrel{\text{def}}{=} P_0.$$

Suppose that $u \in \mathcal{Z}(\mathcal{U}_q\mathfrak{sl}(2))$. Then it acts by a scalar on every $M(\lambda)$, and by $P_0(\lambda) = \theta_u(\lambda)$ on the highest-weight vector.

We will use the following fact: the direct sum of all finite-dimensional $\mathcal{U}_q\mathfrak{sl}(2)$-representations is faithful. In particular, if $u \in \mathcal{Z}(\mathcal{U}_q\mathfrak{sl}(2))$, then there is some $M(\lambda)$ such that $\theta_u \neq 0$, and so $\theta : \mathcal{Z}(\mathcal{U}_q\mathfrak{sl}(2)) \to \mathbb{C}[x^{\pm 1}]$ is an injection.

Take $\lambda = q^n\epsilon$ and consider the exact sequence

$$0 \to M(\lambda q^{-2(n+1)}) \to M(\lambda) \to V_{n,\epsilon} \to 0.$$

It follows that $\theta_u(q^{-1}\lambda) = \theta_u(q^{-1}\lambda^{-1})$. Since q is not a root of unity, this identity holds for infinitely many λ, and hence for all λ. We conclude that for every $u \in \mathcal{Z}(\mathcal{U}_q(\mathfrak{sl}(2)))$, $\theta_u(q^{-1}K)$ is invariant under $K \leftrightarrow K^{-1}$. Thus there is some polynomial $\vartheta_u \in \mathbb{C}[x]$ such that $\theta_u(q^{-1}K) = \vartheta_u(K + K^{-1})$, or equivalently $\theta_u(K) = \vartheta_u(qK + q^{-1}K^{-1})$. In particular, $\vartheta : \mathcal{Z}(\mathcal{U}_q\mathfrak{sl}(2)) \hookrightarrow \mathbb{C}[x]$.

But $\vartheta_C = x/(q - q^{-1})^2$, and so $\vartheta : \mathcal{Z}(\mathcal{U}_q\mathfrak{sl}(2)) \to \mathbb{C}[x]$ is an isomorphism. □

12.2.2 When q is a root of unity

Let us suppose that q is a primitive lth root of unity. We actually care about the degree not of q but of q^2, so we set

$$d = \begin{cases} l, & l \text{ odd}, \\ l/2, & l \text{ even}. \end{cases}$$

12.2.2.1 Lemma *If $q^{2d} = 1$, then K^d, E^d, F^d lie in the center of $\mathcal{U}_q\mathfrak{sl}(2)$.*

Proof K^d is central, since $K^dEK^{-d} = q^{2d}E = E$ and $K^dFK^{-d} = q^{-2d}F = F$, and E^d, F^d commute with K for the same reason. The most interesting relation is that F^d commutes with E. We have

$$[E, F^d] = [d]_q \frac{q^{1-d}K - q^{d-1}K^{-1}}{q - q^{-1}} F^{d-1}.$$

When $q^{2d} = 1$, $[d]_q = \frac{q^d - q^{-d}}{q - q^{-1}} = 0$. □

12.2.2.2 Corollary *If V is an irreducible representation of $\mathcal{U}_q\mathfrak{sl}(2)$, then K^d, E^d, F^d act on V by scalars.* □

12.2.2.3 Theorem (Rep($\mathcal{U}_q\mathfrak{sl}(2)$) when q is a root of unity)
Suppose $q^{2d} = 1$ and V is an irreducible representation of $\mathcal{U}_q\mathfrak{sl}(2)$. Then $\dim V \leq d$, with equality whenever any of the following equations hold:

$$F^d|_V \neq 0, \quad E^d|_V \neq 0, \quad K^d|_V \neq \pm 1.$$

Recall that for any algebra A, Schur's lemma provides a map $\gamma : \mathrm{Irr}(A) \to \mathrm{Specm}(\mathcal{Z}(A))$, where $\mathrm{Specm}(A) = \hom(A, \mathbb{C})$ *denotes the set of maximal ideals. Suppose that $\mu \in \mathrm{Specm}(\mathcal{Z}(\mathcal{U}_q\mathfrak{sl}(2)))$ sends $F^d \mapsto a \neq 0$. Then $\gamma^{-1}(\mu)$ contains at most one representation.*

Proof We start as in the previous section. The action of K is diagonalizable, so pick some $v_0 \in V$ such that $Kv_0 = \lambda v_0$ and set $v_i = F^i v_0$. Then $Kv_i = \lambda q^{-2i} v_i$.

Assume that $F^d|_V = a$. Then $v_d = av_0$, and so the v_is lie in a circle if $a \neq 0$ and a chain if $a = 0$. We claim that the v_is span a $\mathcal{U}_q\mathfrak{sl}(2)$-subrepresentation. They are clearly closed under K and F. Recall from Theorem 12.2.1.8 the Casimir $C = EF + \frac{q^{-1}K + qK^{-1}}{(q-q^{-1})^2}$. It is central and so acts as a scalar on V. But $EF^i v_0 = EF\,F^{i-1}v_0 = (\text{scalar})F^{i-1}v_0$, confirming the claim. It follows in particular that $\dim V \leq d$, and that $\dim V = d$ if $a = F^d|_V \neq 0$. A similar calculation works when $E^d|_V \neq 0$.

Suppose finally that $E^d|_V = F^d|_V = 0$. Then we can choose a highest weight vector v_0 and analyze as in the non-root case: $Ev_0 = 0$ and $EF^m v_0 = [m]_q \frac{q^{1-m}\lambda - q^{m-1}\lambda}{q-q^{-1}} F^{m-1} v_0$. Suppose that $\dim V < d$. Then $[m+1]_q \frac{q^{-m}\lambda - q^m\lambda}{q-q^{-1}} = 0$ for some $m < d-1$, for which $[m]_q \neq 0$, forcing $\lambda = \pm q^m$ hence $K^d = \pm 1$.

For the last statement, let $V \in \gamma^{-1}(\mu)$, so that the $\mathcal{U}_q\mathfrak{sl}(2)$-action factors through $\mathcal{U}_q\mathfrak{sl}(2)/(\ker\mu)$. Since $\mu(F^d) \neq 0$, using the PBW basis from Proposition 12.2.0.5 we see that $\{E^iK^j\}$ span $\mathcal{U}_q\mathfrak{sl}(2)/(\ker\mu)$, and so $\dim \mathcal{U}_q\mathfrak{sl}(2)/(\ker\mu) \leq d^2$. But V is an irrep, and so the image of $\mathcal{U}_q\mathfrak{sl}(2)$ in $\mathrm{End}_{\mathbb{C}}(V)$ is the full matrix algebra, and so $\mathcal{U}_q\mathfrak{sl}(2)/(\ker\mu) \cong \mathrm{End}_{\mathbb{C}}(V) \cong \mathrm{Mat}(d)$. $\square$

12.2.2.4 Definition *The* small quantum group *is the quotient* $\mathcal{U}_q\mathfrak{sl}(2)/\langle K^d = 1, E^d = F^d = 0\rangle$. *It is not a semisimple algebra.*

12.2.3 Hopf structure

We continue to study $\mathcal{U}_q\mathfrak{sl}(2)$. So far we have only discussed it as an associative algebra; we now describe its Hopf structure. Recall that we have generators K, E, F. We define the comultiplication by:

$$\Delta K = K \otimes K, \tag{12.2.3.1}$$
$$\Delta E = 1 \otimes E + E \otimes K, \tag{12.2.3.2}$$
$$\Delta F = K^{-1} \otimes F + F \otimes 1. \tag{12.2.3.3}$$

Then equation (12.2.3.1) tells you that K is a grouplike element. We extend it to a comultiplication $\mathcal{U}_q\mathfrak{sl}(2) \to \mathcal{U}_q\mathfrak{sl}(2)^{\otimes 2}$ by declaring it to be an algebra homomorphism; we must check that it respects equations (12.2.0.2–12.2.0.4), of which only the third requires thought:

$$[\Delta E, \Delta F] = F \otimes E - F \otimes E + EK^{-1} \otimes KF - K^{-1}E \otimes FK + \\ + K^{-1} \otimes \frac{K - K^{-1}}{q - q^{-1}} + \frac{K - K^{-1}}{q - q^{-1}} \otimes K = \Delta \frac{K - K^{-1}}{q - q^{-1}}. \tag{12.2.3.4}$$

Coassociativity can be checked on generators. The counit and antipode are uniquely determined; they are:

$$\epsilon : (K, E, F) \mapsto (1, 0, 0), \tag{12.2.3.5}$$
$$\mathcal{S} : (K, E, F) \mapsto (K^{-1}, -EK^{-1}, -KF). \tag{12.2.3.6}$$

12.2.3.7 Remark $\mathcal{S}$ is not an involution. Rather, $\mathcal{S}^2$ is conjugation by K. $\Diamond$

12.2.3.8 Remark The comultiplication Δ used above is not unique. Choose any invertible element $\phi \in \mathbb{C}[K^{\pm 1}]^{\otimes 2}$. Then $\Delta' : u \mapsto \phi\Delta(u)\phi^{-1}$ is another valid comultiplication. The element ϕ is a type of *gauge transformation*, and the comultiplication Δ is unique up to gauge equivalence. $\Diamond$

We will study the following construction in more detail in Section 12.3.1. Let V, W be finite-dimensional $\mathcal{U}_q\mathfrak{sl}(2)$-modules. Then the vector space tensor product $V \otimes W$ is automatically a $\mathcal{U}_q\mathfrak{sl}(2)^{\otimes 2}$-module, and we can make it into a $\mathcal{U}_q\mathfrak{sl}(2)$-module by restricting along $\Delta : \mathcal{U}_q\mathfrak{sl}(2) \to \mathcal{U}_q\mathfrak{sl}(2)^{\otimes 2}$. This determines a monoidal functor $\otimes : \mathcal{U}_q\mathfrak{sl}(2)\text{-MOD} \times \mathcal{U}_q\mathfrak{sl}(2)\text{-MOD} \to \mathcal{U}_q\mathfrak{sl}(2)\text{-MOD}$. Gauge-equivalent comultiplications determine isomorphic monoidal functors. By diagonalizing K, one can easily show:

12.2.3.9 Proposition *$V_{\epsilon,n} \otimes V_{\epsilon',n'} \cong \bigoplus_m V_{\epsilon\epsilon',m}$ where the sum ranges over those m such that $n + n' + m$ is even and $\{n, n', m\}$ satisfy the triangle inequality $|n - n'| \leq m \leq n + n'$.* $\Box$

12.2.3.10 Corollary *Suppose q is not a root of unity. The monoidal abelian subcategories of $\mathcal{U}_q\mathfrak{sl}(2)$-MOD are the $\oplus$-closures of:*

1. *All $V_{\pm,n}$.*
2. *$V_{+,n}$ with n arbitrary.*
3. *$V_{\epsilon,n}$ with $\epsilon = (-1)^n$ and n arbitrary.*
4. *$V_{\pm,n}$ with n even.*
5. *$V_{+,n}$ with n even.*
6. *Just $V_{\pm,0}$.*
7. *Just $V_{+,0}$.* $\Box$

We will use the subcategory consisting of modules $V_{+,n}$ often enough that we will give it a name: $\mathcal{U}_q\mathfrak{sl}(2)\text{-MOD}^+$.

12.2.4 $\mathcal{U}_q\mathfrak{sl}(2)$ and $\mathscr{C}_q(\mathrm{SL}(2))$ are dual

12.2.4.1 Definition *Let A and B be bialgebras. A* bialgebra pairing *between them is a map $\langle , \rangle : A \otimes B \to \mathbb{C}$ such that $\langle \Delta a, b_1 \otimes b_2 = \langle a, b_1 b_2 \rangle$ and $\langle a_1 \times a_2, \Delta b \rangle = \langle a_1 a_2, b \rangle$ and $\epsilon(a) = \langle a, 1_B \rangle$ and $\epsilon(b) = \langle 1_A, b \rangle$. The bialgebras A and B are* in duality *if the pairing is nondegenerate.*

Equivalently, if A is a bialgebra, then A^*, the space of linear maps $A \to \mathbb{C}$, is an algebra with $(\alpha\beta)(a) \overset{\text{def}}{=} \sum \alpha(a_1)\beta(a_2)$. A bialgebra pairing is a pairing inducing algebra homomorphisms $A \to B^*$ and $B \to A^*$. It is a duality when these homomorphisms are injections.

Recall that the bialgebra $\mathscr{C}_q(\mathrm{Mat}(2))$ of quantum matrices was defined by the relations

$$ba = qab \quad ca = qac \quad bc = cb \quad db = qbd \quad dc = qcd \quad ad - da = (q^{-1} - q)bc.$$

Consider the two-dimensional representation $V_{1,+}$ of $\mathcal{U}_q\mathfrak{sl}(2)$ given by

$$E \mapsto \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad F \mapsto \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad K \mapsto \begin{pmatrix} q & 0 \\ 0 & q^{-1} \end{pmatrix}.$$

Suppose that $u \in \mathcal{U}_q\mathfrak{sl}(2)$ acts on $V_{1,+}$ by the matrix $\left(\begin{smallmatrix} a(u) & b(u) \\ c(u) & d(u) \end{smallmatrix}\right)$. The idea is to construct a pairing between $\mathscr{C}_q(\mathrm{Mat}(2))$ and $\mathcal{U}_q\mathfrak{sl}(2)$ such that

$$\left\langle \begin{pmatrix} a & b \\ c & d \end{pmatrix}, u \right\rangle = \begin{pmatrix} A(u) & B(u) \\ C(u) & D(u) \end{pmatrix}. \tag{12.2.4.2}$$

12.2.4.3 Theorem ($\mathcal{U}_q\mathfrak{sl}(2)$ and $\mathscr{C}_q(\mathrm{SL}(2))$ are dual)
The pairing defined on generators in equation (12.2.4.2) determines a duality between $\mathscr{C}_q(\mathrm{SL}(2)) = \mathscr{C}_q(\mathrm{Mat}(2))/(\det_q - 1)$ and $\mathcal{U}_q\mathfrak{sl}(2)$.

Proof Given $u \in \mathcal{U}_q\mathfrak{sl}(2)$, write $\Delta u = \sum u_1 \otimes u_2$. We first must check that for $x, y \in \{a, b, c, d\}$, we have

$$(xy)(u) = \sum x(u_1)y(u_2). \tag{12.2.4.4}$$

We can check this on the basis $\{E^iF^jK^l\}$. The trick is that in our representation, $x(E^2) = x(F^2) = 0$ for all $x \in \{a, b, c, d\}$. So we need only to check equation (12.2.4.4) on $\{E^iF^jK^l\}$ for $i, j < 2$. Recall that $\Delta E = (E \otimes K + 1 \otimes E)$ and $\Delta F = (K^{-1} \otimes F + F \otimes 1)$ and $\Delta K^l = K^l \otimes K^l$. Then, for some constants α, β, the values of $\{a, b, c, d\}$ on $\{E^iF^jK^l\}$ are:

	K^l	FK^l	F^2K^l	EK^l	E^2K^l	EFK^l	E^2FK^l	EF^2K^l	$E^2F^2K^l$
BA							β		
AB							$q^{-1}\beta$		
CA		q^{2l}						αq^{2l-1}	
AC		q^{2l-1}						αq^{2l-2}	
DA	1					q			
AD	1					q^{-1}			
BC						1			
CB						1			
DB					q^{-2l}				
BD					q^{-2l-1}				
DC									
CD									

The empty entries vanish.

This establishes that we have a homomorphism $\mathscr{C}_q(\mathrm{Mat}(2)) \to (\mathcal{U}_q\mathfrak{sl}(2))^*$. The fact that we had a representation gave a homomorphism $\mathcal{U}_q\mathfrak{sl}(2) \to \mathscr{C}_q(\mathrm{Mat}(2))^*$. So we have a pairing $\mathscr{C}_q(\mathrm{Mat}(2)) \otimes \mathcal{U}_q\mathfrak{sl}(2) \to \mathbb{C}$ of bialgebras.

However, this pairing has kernel. Indeed, for all $u \in \mathcal{U}_q\mathfrak{sl}(2)$,

$$\langle \det_q, u\rangle = \epsilon(u),$$

where $\epsilon : \mathcal{U}_q\mathfrak{sl}(2) \to \mathbb{C}$ is the counit. To check this, it suffices to check it on generators and use that $\det_q$ is grouplike.

Thus our pairing factors through $\mathscr{C}_q(\mathrm{SL}(2)) \otimes \mathcal{U}_q\mathfrak{sl}(2) \to \mathbb{C}$. We must show that there is no further kernel. There cannot be kernel on the $\mathscr{C}_q(\mathrm{SL}(2))$ side, since it is merely the algebra of matrix coefficients of the action of $\mathcal{U}_q\mathfrak{sl}(2)$ on $V_{1,+}$. Suppose that there is kernel on the $\mathcal{U}_q\mathfrak{sl}(2)$ side. Then it must be a Hopf ideal, and must vanish on $V_{1,+}$ (since that was the module we used to define the pairing), and hence (being a Hopf ideal) on all powers of $V_{1,+}$. But Proposition 13.2.1.5, in the case $\mathfrak{g} = \mathfrak{sl}(2)$, implies that the action of $\mathcal{U}_q\mathfrak{sl}(2)$ on $\bigoplus_n V_{1,+}^{\otimes n}$ is faithful. □

12.3 The Jones polynomial

12.3.1 Hopf algebras and monoidal categories

12.3.1.1 Definition *A* monoidal category *is a category $\mathcal{C}$ equipped with a functor $\otimes : \mathcal{C} \times \mathcal{C} \to \mathcal{C}$ such that, for objects V, W, X, the two products $(V \otimes W) \otimes X \underset{a_{V,W,X}}{\overset{\sim}{\to}} V \otimes (W \otimes X)$ are naturally isomorphism (natural in all three variables V, W, X) — the natural isomorphism a, called the* associator*, is part of the data of the monoidal category. There should also be a unit object $\mathbf{1} \in \mathcal{C}$ with natural isomorphisms $V \otimes \mathbf{1} \underset{r_V}{\overset{\sim}{\to}} V \underset{l_V}{\overset{\sim}{\leftarrow}} \mathbf{1} \otimes V$ called* unitors*. These natural isomorphisms must satisfy their own associativity conditions, given by the commutativity of the following diagrams:*

$$\begin{tikzcd}
& (V \otimes W) \otimes (X \otimes Y) \arrow[dr, "a_{V,W,X\otimes Y}"] & \\
((V \otimes W) \otimes X) \otimes Y \arrow[ur, "a_{V\otimes W,X,Y}"] \arrow[d, "a_{V,W,X}\otimes \mathrm{id}_Y"'] & & V \otimes (W \otimes (X \otimes Y)) \\
(V \otimes (W \otimes X)) \otimes Y \arrow[rr, "a_{V,W\otimes X,Y}"] & & V \otimes ((W \otimes X) \otimes Y) \arrow[u, "\mathrm{id}_V \otimes a_{W,X,Y}"']
\end{tikzcd}$$

$$\begin{tikzcd}
& (V\otimes\mathbf{1})\otimes W \arrow[rr, "a_{V,\mathbf{1},W}"] \arrow[dr, "l_V\otimes\mathrm{id}_W"'] & & V\otimes(\mathbf{1}\otimes W) \arrow[dl, "\mathrm{id}_V\otimes r_W"] & \\
(\mathbf{1}\otimes V)\otimes W \arrow[rr, "r_V\otimes\mathrm{id}_W"] \arrow[dr, "a_{\mathbf{1},V,W}"] & & V\otimes W & & V\otimes(W\otimes\mathbf{1}) \arrow[ll, "\mathrm{id}_V\otimes l_W"'] \\
& \mathbf{1}\otimes(V\otimes W) \arrow[ur, "r_{V\otimes W}"] & & (V\otimes W)\otimes\mathbf{1} \arrow[ul, "l_{V\otimes W}"'] \arrow[ur, "a_{V,W,\mathbf{1}}"'] &
\end{tikzcd}$$

$$\mathbf{1}\otimes\mathbf{1} \overset{r_{\mathbf{1}}}{\underset{l_{\mathbf{1}}}{\rightrightarrows}} \mathbf{1}$$

The first relation is called the *pentagon equation* and the second are called *triangle equations* for obvious reasons.

12.3.1.2 Remark A monoidal category is *strict* if a, r, l are all identities. Every monoidal category is equivalent to a strict one, but almost no categories "in nature" are strict. For example, Vect itself isn't strict. Indeed, if V, W, X are vector spaces, then $(V\otimes W)\otimes X$ and $V\otimes(W\otimes X)$ are not *equal*, but they are canonically isomorphic. The problem goes back all the way to set theory. Even if V, W, X are sets, $(V\times W)\times X$ and $V\times(W\times X)$ are not equal: the first consists of ordered pairs (a, x) where a is an ordered pair (v, w), and the second consists of ordered pairs (v, b) where b is an ordered pair (w, x). ♢

12.3.1.3 Definition *A monoidal category* $(\mathcal{C}, \otimes, a, \dots)$ *is* rigid *if for every object* $V \in \mathcal{C}$*, there exists* dual objects V^* *and* ${}^*V$ *and maps* $i_V : \mathbf{1} \to V\otimes V^*$, $e_V : V^*\otimes V \to \mathbf{1}$, $i_{{}^*V} : \mathbf{1} \to {}^*V\otimes V$, $e_{{}^*V} : V\otimes {}^*V \to \mathbf{1}$, *such that the following compositions are identities:*

$$V = 1\otimes V \overset{i_V\otimes\mathrm{id}}{\longrightarrow} V\otimes V^*\otimes V \overset{\mathrm{id}\otimes e_V}{\longrightarrow} V\otimes 1 = V, \tag{12.3.1.4}$$

$$V^* = V^*\otimes 1 \overset{\mathrm{id}\otimes i_V}{\longrightarrow} V^*\otimes V\otimes V^* \overset{e_V\otimes\mathrm{id}}{\longrightarrow} 1\otimes V^* = V^*, \tag{12.3.1.5}$$

$$V = V\otimes 1 \overset{\mathrm{id}\otimes i_{{}^*V}}{\longrightarrow} V\otimes {}^*V\otimes V \overset{e_{{}^*V}\otimes\mathrm{id}}{\longrightarrow} 1\otimes V = V, \tag{12.3.1.6}$$

$${}^*V = 1\otimes {}^*V \overset{i_{{}^*V}\otimes\mathrm{id}}{\longrightarrow} {}^*V\otimes V\otimes {}^*V \overset{\mathrm{id}\otimes e_{{}^*V}}{\longrightarrow} {}^*V\otimes 1 = {}^*V. \tag{12.3.1.7}$$

In the equations in Definition 12.3.1.3, we left out all unitors and associators. They are easy to add in, and only clutter the formulas.

12.3.1.8 Lemma *Fix a monoidal category* $(\mathcal{C}, \otimes, \dots)$ *and an object* $V \in \mathcal{C}$*. Then the data* (V^*, i_V, e_V) *is unique-up-to-unique-isomorphism if it exists. Similarly for* $({}^*V, i_{{}^*V}, e_{{}^*V})$*. In particular, there are canonical isomorphisms* ${}^*(V^*) \cong V \cong ({}^*V)^*$. □

12.3.1.9 Proposition *Let* H *be a bialgebra over* $\mathbb{C}$*. Define a monoidal functor on the category* H-mod *of finite-dimensional left* H*-modules by declaring that if* $(V, \pi_V), (W, \pi_W) \in H$-mod *(where*

$\pi_V : H \to \mathrm{End}_{\mathbb{C}}(V)$ is the action), then H acts on the vector-space tensor product $V \otimes W$ via the comultiplication $H \xrightarrow{\Delta} H \otimes H \to \mathrm{End}_{\mathbb{C}}(V) \otimes \mathrm{End}_{\mathbb{C}}(W) \to \mathrm{End}_{\mathbb{C}}(V \otimes W)$; coassociativity of Δ means that the ordinary associator for vector spaces provides an associator for H-MOD. The unit object $\mathbf{1}$ is $\mathbb{C}$, made into an H-module via the counit $\epsilon : H \to \mathbb{C} = \mathrm{End}_{\mathbb{C}}(\mathbb{C})$.

Suppose furthermore that H is a Hopf algebra. Then H-MOD is rigid. Indeed, write V^* for the linear dual to V. It has a right H-action, i.e. a left action by H^{op}, that we will call π_V^*. Then $(V, \pi_V)^* \overset{\mathrm{def}}{=} (V^*, \pi_V^* \circ \mathcal{S})$ and ${}^*(V, \pi_V) \overset{\mathrm{def}}{=} (V^*, \pi_V^* \circ \mathcal{S}^{-1})$ do the trick. □

See, often when Hopf algebras are introduced, they're seen as generalizations of "functions on a group":

- Δ comes from $\cdot$,
- ϵ comes from e,
- $\mathcal{S}$ comes from $g \mapsto g^{-1}$,
- and the algebra structure understands the geometry of the group.

But then there's a switcheroo: we think about H-modules, and use the "co" structure to give H-MOD extra structure, via

- $\Delta \rightsquigarrow \otimes$
- $\epsilon \rightsquigarrow \mathbf{1}$
- $\mathcal{S} \rightsquigarrow *$
- and the algebra structure makes H-mod into an abelian category.

The reason we can make such a switcheroo is that the notion of "Hopf algebra" is symmetric under taking linear duals. More precisely, we can quantize the functions on a group, or the universal enveloping algebra, and both are Hopf algebras. So the switcheroo is really a version of Definition 12.2.4.1.

For vector spaces V, W, the tensor products $V \otimes W$ and $W \otimes V$ are naturally isomorphism. What about in a general monoidal category? For categories of representations of (non-quantum) groups and Lie algebras, the answer is yes: the vector-space "swap" map does the job. But there may be others.

12.3.1.10 Example The category SVECT of *super vector spaces* has objects $\mathbb{Z}/2$-graded vector spaces, and braiding is $v \otimes w \mapsto (-1)^{|v| \cdot |w|} w \otimes v$, where $|v|$ is the degree of v (either 0 or 1). Note that SVECT = REP($\mathbb{Z}/2$), but this is not the usual vector-space "swap" of representations. ◊

Given a Hopf algebra H, how could we construct a natural transformation $\sigma_{V,W} : V \otimes W \to W \otimes V$?

$$\begin{tikzcd}
\mathcal{C} \times \mathcal{C} \arrow[rr, bend left, "\otimes"] \arrow[rr, bend right, "\otimes^{\mathrm{op}}"'] \arrow[rr, phantom, "\wr\Uparrow\sigma"] & & \mathcal{C}
\end{tikzcd}$$

12.3.1.11 Proposition *Let A and B be algebras, and $f, g : A \to B$ homomorphisms. Construct functors $f^*, g^* : B\text{-MOD} \to A\text{-MOD}$ by precomposition. (So, for example, for a B-module (M, π_M), we set $f^*(M, \pi_M) \stackrel{\text{def}}{=} (M, \pi_M \circ f)$.) Suppose $b \in B$ satisfies $b\, f(a) = g(a)\, b$ for every $a \in A$. Then $\pi(b) : M \mapsto \pi_M(b)$ is a natural transformation $f^* \Rightarrow g^*$. If b is invertible, so is $\pi(b)$.*

Essentially all natural transformations $f^* \Rightarrow g^*$ are of this type. More precisely, if by "A-MOD" and "B-MOD" we mean the categories of all possibly-infinite-dimensional modules, then every natural transformation $f^* \Rightarrow g^*$ is of the form b^* for some $b \in B$ as above. If we mean the categories of finite-dimensional modules, then there may be a few more such natural transformations, coming from bs that live not in B but it its profinite-dimensional completion. □

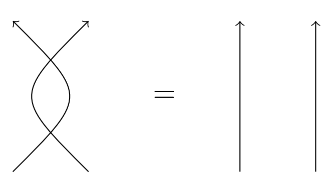
In the notation of Proposition 12.3.1.11, the functors $(V, W) \mapsto V \otimes W$ and $(V, W) \mapsto W \otimes V$ on H-MOD are, respectively, Δ^* and $(\Delta^{\mathrm{op}})^* = (\mathrm{swap} \circ \Delta)^*$, where $\Delta : H \to H \otimes H$ is the comultiplication. So to build an isomorphism $\sigma_{V,W} : V \otimes W \xrightarrow{\sim} W \otimes V$ of H-modules, we could choose $R \in H \otimes H$ invertible, and set $\sigma(v \otimes w) = \mathrm{swap}\, R(v \otimes w)$. We would need:

$$\Delta^{\mathrm{op}}(x) = R\, \Delta(x)\, R^{-1}, \quad \forall x \in H \tag{12.3.1.12}$$

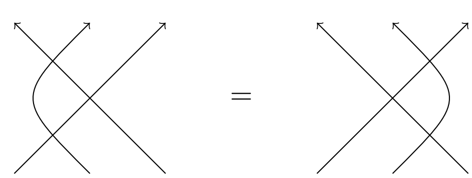
where $\Delta^{\mathrm{op}} = \mathrm{swap}\ \Delta$.

What properties do we demand of $\sigma_{V,W}$? The old fashioned answer is to satisfy the relations on the symmetric group. We will explain this in pictures. The yoga is to draw objects as labeled strands. For example, $\uparrow V$ is an object, $\downarrow V$ is its dual, and $\curvearrowright$ is the pairing $V \otimes V^* \to \mathbb{C}$. We will read diagrams so that composition goes from bottom to top. That way, if you read a diagram from top to bottom, you get the operations in left-to-right order.

Then if $g \in S_n$, we could demand that there is a well-defined $\sigma_g : V^{\otimes n} \to V^{\otimes n}$ built from σ. Specifically: choose any word for g as a product of transpositions; use $\sigma_{V,V}$ for every transposition; demand that the answer doesn't depend on choice of word. This is equivalent to the requirement that



(12.3.1.13)



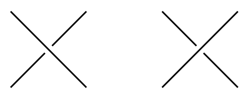
(12.3.1.14)

Actually, there's something more needed. We have two maps $V \otimes W \otimes X \to X \otimes V \otimes W$ which we should insist are equal:

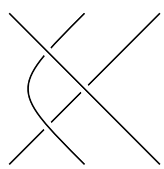
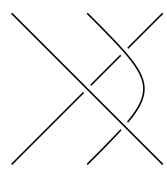
$$\sigma_{V\otimes W,X} = (\sigma_{V,X} \otimes \mathrm{id}_W) \circ (\mathrm{id}_V \otimes \sigma_{W,X}) \tag{12.3.1.15}$$

12.3.1.16 Definition *A monoidal category* $(\mathscr{C}, \otimes, \dots)$ *is* symmetric *if it is equipped with a natural isomorphism* $\sigma_{V,W} : V \otimes W \xrightarrow{\sim} W \otimes V$ *satisfying equations (12.3.1.13) and (12.3.1.15). (Equation (12.3.1.14) follows from these relations.)*

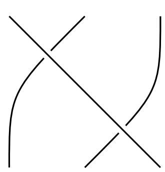
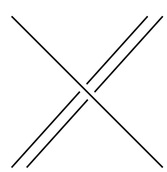
Monoidal categories are a categorical version of associative algebras, and symmetric monoidal categories are like commutative algebras. But there's another option, which is slightly less commutative, and definitely quantum: rather than demanding that the symmetric group act on tensor products, we could demand only that the braid group acts. The braid group has overcrossings and undercrossings.



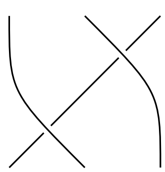
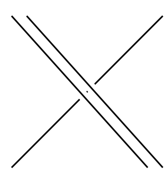
We will use the famous result of Artin's that the braid group has a presentation in terms of overcrossings, with the only relation being (that  is invertible, with inverse , and) the *braid relation*, also called Reidemeister Three:

 =  (12.3.1.17)

12.3.1.18 Definition *A monoidal category* $(\mathscr{C}, \otimes, \dots)$ *is* braided *if it is equipped with a natural isomorphism* $\beta_{V,W} =$  $: V \otimes W \xrightarrow{\sim} W \otimes V$ *satisfying the following two versions of equation (12.3.1.15) (which are no longer equivalent), called the* hexagon equations *because they look like hexagons if you restore the associators:*

 =  (12.3.1.19)

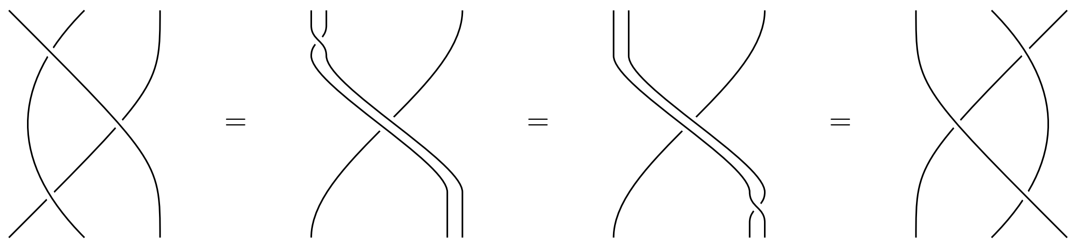
$$(\beta_{V,X} \otimes \mathrm{id}_W) \circ (\mathrm{id}_V \otimes \beta_{W,X}) = \beta_{V\otimes W,X}$$

 =  (12.3.1.20)

$$(\mathrm{id}_W \otimes \beta_{V,X}) \circ (\beta_{V,W} \otimes \mathrm{id}_X) = \beta_{V,W\otimes X}$$

12.3.1.21 Lemma *[Equation (12.3.1.17)](#) holds in any braided monoidal category.*

Proof



The first and third equalities are the hexagon equation equation (12.3.1.20) and the second one is naturality of the braiding. ☐

Recall that our braiding was $\sigma_{V,W} = \operatorname{swap} R(v \otimes w)$, corresponding to $R = \sum R_{(1)} \otimes R_{(2)} \in H \otimes H$. Write $R_{13} \stackrel{\text{def}}{=} \sum R_{(1)} \otimes 1 \otimes R_{(2)} \in H^{\otimes 3}$, and similarly R_{12} and R_{23}. After moving some "swaps" around, equations (12.3.1.19) and (12.3.1.20) become:

$$(\Delta \otimes \mathrm{id})(R) = R_{13} R_{23} \tag{12.3.1.22}$$

$$(\mathrm{id} \otimes \Delta)(R) = R_{13} R_{12} \tag{12.3.1.23}$$

The left-hand sides are what you get by applying the functions $\Delta \otimes \mathrm{id}$ and $\mathrm{id} \otimes \Delta : H^{\otimes 2} \to H^{\otimes 3}$ to $R \in H^{\otimes 2}$, and the right-hand sides are multiplication in $H^{\otimes 3}$.

12.3.1.24 Lemma / Definition *A* quasitriangular Hopf algebra *is a Hopf algebra H equipped with an* R-matrix *$R \in H^{\otimes 2}$ which is invertible and satisfies equations (12.3.1.12), (12.3.1.22), and (12.3.1.23). It is* triangular *if additionally $R^{-1} = \operatorname{swap}(R)$. If H is quasitriangular, then the category H-MOD of finite-dimensional H-modules is braided rigid monoidal. If H is triangular, then H-MOD is symmetric.* ☐

The names "triangular" and "quasitriangular" are because equations (12.3.1.14) and (12.3.1.17) have triangles in them. Those equations, when written in terms of the R-matrix, are called the *Yang–Baxter equation*:

$$R_{12} R_{13} R_{23} = R_{23} R_{13} R_{12} \tag{12.3.1.25}$$

Look at an n-dimensional representation of H. Then R evaluates to an $n^2 \times n^2$ matrix, giving n^4 unknowns, but equation (12.3.1.25) is an equality of $n^3 \times n^3$ matrices, so it is n^6 equations. Equation (12.3.1.25) first turned up in physics. The original motivation for quantum groups was to find solutions to it.

12.3.1.26 Example Let $V = \mathbb{C}^2$ with ordered basis $\{e_1, e_2\}$ and choose $q \in \mathbb{C}^\times$. Identify $V \otimes V = \mathbb{C}^4$ with ordered basis $\{e_{11}, e_{12}, e_{21}, e_{22}\}$, where $e_{ij} = e_i \otimes e_j$. The following R-matrix solves equation (12.3.1.25):

$$R = \begin{pmatrix} q & & & \\ & 1 & & \\ & q - q^{-1} & 1 & \\ & & & q \end{pmatrix}$$

We will see that this solution arises from quantum $\mathfrak{sl}(2)$. Anticipating the answer, it will turn out that $H = \mathcal{U}_q\mathfrak{sl}(2)$ is almost quasitriangular: it will have a R-matrix solving equations (12.3.1.12), (12.3.1.22), and (12.3.1.23), except R won't live in $H^{\otimes 2}$ but rather in a certain completion of it. In terms of Proposition 12.3.1.11, the category H-MOD of finite-dimensional H-modules will be braided, but the category of all possibly-infinite-dimensional modules does not have a braiding. $\Diamond$

12.3.1.27 Remark The main physical applications of the Yang–Baxter equation in fact want R to depend on a "spectral parameter." Solutions to the equation (12.3.1.25) come from quantizations of compact groups, whereas solutions with a spectral parameter come from quantizations of loop groups. $\Diamond$

12.3.2 The Temperley–Lieb algebra

We continue to fix $q \in \mathbb{C}^\times$. Recall the n-strand braid group $B_n = \langle s_1, \ldots, s_{n-1} \text{ s.t. } s_i s_{i+1} s_i = s_{i+1} s_i s_{i+1}\rangle$. Here s_i is the crossing  in the ith and $(i+1)$th spots, and the relation is equation (12.3.1.17).

12.3.2.1 Definition *The* Hecke–Iwahori algebra *is the following quotient of the group algebra of the braid group:*

$$\mathcal{H}_n(q) \stackrel{\text{def}}{=} \mathbb{C}[B_n]/\langle (s_i - q)(s_i + q^{-1})\rangle.$$

For example, when $q = 1$, the added relation is then $s_i^2 = 1$, and so $\mathcal{H}_n(1)$ is the group algebra of the symmetric group. The representation theory of $\mathcal{H}_n(q)$ is rich and well-studied.

12.3.2.2 Proposition *When q is not a root of unity, $\mathcal{H}_n(q)$ is semisimple and isomorphic as a vector space to $\mathbb{C}[S_n]$. Its irreducible representations are enumerated by the same combinatorial data as are those of S_n, namely partitions of n aka Young diagrams with n boxes. Let W_λ denote the irrep corresponding to λ. Then $\mathcal{H}_n(q) \cong \bigoplus_\lambda \mathrm{Mat}(W_\lambda)$.* $\square$

12.3.2.3 Remark When q^2 is a primitive dth root of unity, the representation theory of $\mathcal{H}_n(q)$ is very close to the charateristic-d representation theory of the symmetric group S_n. In particular, when $n < d$, $\mathcal{H}_n(q)$ is still semisimple. Full details are worked out in [Wen85]. $\Diamond$

Consider the R-matrix R from Example 12.3.1.26, and

$$S = \mathrm{swap} \circ R = \left(\begin{array}{c:cc:c} q & & & \\ \hdashline & 0 & 1 & \\ & 1 & q - q^{-1} & \\ \hdashline & & & q \end{array}\right), \tag{12.3.2.4}$$

where we've emphasized its block form. The middle block has eigenvalues q and $-q^{-1}$, and the outer blocks are simply q. So S solves $(S - q)(S + q^{-1})$. This implies:

12.3.2.5 Lemma *Set $V = \mathbb{C}^2$. The action of B_n on $V^{\otimes n}$ via $s_i \mapsto \mathrm{id}^{\otimes i-1} \otimes S \otimes \mathrm{id}^{n-i-1}$ factors through $\mathcal{H}_n(q)$.* $\square$

Note that $\dim \mathcal{H}_n(q) = n!$, which grows much faster than $\dim V^{\otimes n} = 2^n$. So the action has some interesting image and kernel.

12.3.2.6 Definition *The nth* Temperley–Lieb *algebra $\mathcal{TL}_n(q)$ is the image of $\mathcal{H}_n(q)$ in* $\mathrm{End}(V^{\otimes n})$.

We will give a presentation of $\mathcal{TL}_n(q)$. In order to do so, we explain the origin of equation (12.3.2.4) in terms of quantum SL(2). The idea should be clear: understand $V = \mathbb{C}^2 = V_{+,1}$ as the two-dimensional representation of $\mathcal{U}_q\mathfrak{sl}(2)$, and find S as a morphism of $\mathcal{U}_q\mathfrak{sl}(2)$-modules. By Proposition 12.2.3.9, $V^{\otimes 2} \cong V_{+,0} \oplus V_{+,2}$, where $V_{+,0} = \mathbb{C}$ is the monoidal unit and $V_{+,2}$ is three-dimensional. So $\mathrm{End}_{\mathcal{U}_q\mathfrak{sl}(2)}(V^{\otimes 2}) \cong \mathbb{C}P \oplus \mathbb{C}(1-P)$, where $P^2 = P$ is the projection onto $V_{+,0} = \mathbb{C}$. Choose a weight basis e_1, e_2 of V, so that

$$K = \begin{pmatrix} q & 0 \\ 0 & q^{-1} \end{pmatrix}, \quad E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad F = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

12.3.2.7 Lemma $\hom_{\mathcal{U}_q\mathfrak{sl}(2)}(V^{\otimes 2}, \mathbb{C}) \cong \mathbb{C}$ *is spanned by the pairing $\langle , \rangle : V^{\otimes 2} \to \mathbb{C}$ defined by:*

$$\langle e_1, e_1 \rangle = \langle e_2, e_2 \rangle = 0, \quad \langle e_2, e_1 \rangle = 1, \quad \langle e_1, e_2 \rangle = -q.$$

Proof For example, $\Delta(E) = 1 \otimes E + E \otimes K$, and sure enough

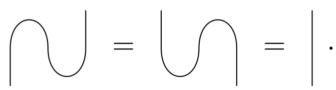
$$\langle e_1, Ee_1 \rangle + \langle Ee_1, Ke_1 \rangle = \langle e_1, e_2 \rangle + \langle e_2, qe_1 \rangle = -q + q = 0.$$

Invariance under $\Delta(K)$, $\Delta(F)$ are similar. $\square$

Let use write this pairing as $\cap : V \otimes V \to \mathbb{C}$. The inverse copairing is

$$\cup = e_1 \otimes e_2 - q^{-1} e_2 \otimes e_1 \in V \otimes V.$$

They are inverse in the sense that



12.3.2.8 Lemma *The projection $P \in \mathrm{End}(V^{\otimes 2})$ onto the one-dimensional direct summand is*

$$P = -\frac{1}{q+q^{-1}} \substack{\cup \\ \cap}.$$

Proof

$$P^2 = \frac{1}{(q+q^{-1})^2} \substack{\cup \\ \bigcirc \\ \cap} = \frac{-q-q^{-1}}{(q+q^{-1})^2} \substack{\cup \\ \cap} = P$$

since $\bigcirc = \langle e_1, e_2 \rangle - q^{-1} \langle e_2, e_1 \rangle = -q - q^{-1}$. $\square$

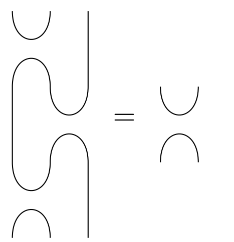
12.3.2.9 Lemma *The matrix S from equation (12.3.2.4) is*

$$S = q(1-P) - q^{-1}P = q \;\Big| \quad \Big| + \genfrac{}{}{0pt}{}{\cup}{\cap} \in \mathrm{End}_{\mathcal{U}_q\mathfrak{sl}(2)}(V^{\otimes 2}).$$

In particular, it solves the quantum Yang–Baxter equation. □

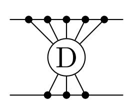
12.3.2.10 Proposition *The map $\mathcal{H}_n(q) \to \mathrm{Mat}(V^{\otimes n})$, whose image is $\mathcal{TL}_n(q)$, is surjective onto $\mathrm{End}_{\mathcal{U}_q\mathfrak{sl}(2)}(V^{\otimes n})$. Set $\tau = -q - q^{-1}$. $\mathcal{TL}_n(q)$ is generated by elements $e_i = \big|\,\big|\cdots\big|\,\genfrac{}{}{0pt}{}{\cup}{\cap}\,\big|\cdots\big|$, $i = 1, \dots, n-1$, with relations $e_i^2 = \tau e_i$, $e_i e_{i\pm 1} e_i = e_i$, and $e_i e_j = e_j e_i$ if $|i-j| \geq 2$.* □

The proof of the relation $e_i e_{i\pm 1} e_i = e_i$ is the following picture:



(12.3.2.11)

There is a better way to say this. Consider the monoidal category where objects are parameterized by nonnegative integers, which we think of as sets of points on a line $\bullet\bullet\bullet\bullet\bullet$ up to isotopy. The morphisms are diagrams that you can use to connect such points:



For example, the diagram $\cap$ makes sense an element of $\hom(\bullet\bullet, -)$, and equation (12.3.2.11) is an equation in $\hom(\bullet\bullet\bullet, \bullet\bullet\bullet)$. Composition is by vertical stacking of diagrams. We make it into a monoidal category by horizontal stacking of diagrams. We do not allow the lines in the diagram to cross, so this is the category of *planar tangles*, also called *noncrossing diagrams*.

Given any category $\mathcal{C}$, you can build a new category $\mathbb{C}[\mathcal{C}]$, called its $\mathbb{C}$-*linearization*, with the same objects and $\hom_{\mathbb{C}[\mathcal{C}]}(A, B) = \mathbb{C}[\hom(A, B)]$, by which we mean the vector space with basis $\hom(A, B)$. Let's do this to the category of planar tangles. The hom spaces are all infinite-dimensional, because planar tangles can have nested closed loops. To make something finite-dimensional, let's declare that $\bigcirc = \tau = -q - q^{-1}$.

12.3.2.12 Definition *The* Temperley–Lieb category $\mathcal{TL}(q)$ *is the linearization of the category of planar tangles modulo the relation* $\bigcirc = \tau = -q - q^{-1}$.

12.3.2.13 Remark Clearly $\mathcal{TL}(q)$ depends only on $\tau = q + q^{-1}$, and not on q itself. Also clearly it makes sense integrally: you can define a version over $\mathbb{Z}[\tau]$ rather than over $\mathbb{C}(q)$. ◊

Then Proposition 12.3.2.10 says:

12.3.2.14 Theorem (Schur–Weyl duality for $\mathcal{U}_q\mathfrak{sl}(2)$)
Suppose q is not a root of unity. The monoidal subcategory of $\mathcal{U}_q\mathfrak{sl}(2)$-MOD consisting of representations $V_{+,n}$ is equivalent to the completion of $\mathcal{TL}$ under taking direct sums and direct summands. □

We will discuss Schur–Weyl duality in more generality in Section 13.2.4. Recall that the irreps of $\mathcal{H}_n(q)$ are indexed by partitions of n, or equivalently Young diagrams with n boxes. The "2" in $\mathfrak{sl}(2)$ ends up translating to the fact that $\mathcal{TL}_n(q)$ is the quotient of $\mathcal{H}_n(q)$ acting faithfully on the sum of irreps corresponding to Young diagrams with at most two rows.

12.3.2.15 Remark Temperley and Lieb did not find their algebra by thinking about quantum groups. Rather, they found it in [TL71] as a unifying feature of three seemingly unrelated problems: the Ising model of ferromagnetism, the Potts model of ice crystalization, and chromatic polynomials of graphs. ♢

12.3.3 Ribbon tangles

The category of *ribbon tangles* combines knots and braids. We will give two definitions:

12.3.3.1 Definition *A* geometric ribbon *is a continuous piecewise-smooth embedding $\rho : [0,1]^{\times 2} \hookrightarrow [0,1] \times (0,1)^{\times 2}$ sending $(0,1) \times [0,1] \hookrightarrow (0,1)^{\times 3}$ and $\{0,1\} \times [0,1] \hookrightarrow \{0,1\} \times (0,1)^{\times 2}$. We think of the first $[0,1]$ in the domain of the ribbon as the "long" direction and the second $[0,1]$ as the "short" direction. We think of the first $[0,1]$ in the codomain of the ribbon as the "vertical" direction, the second as "horizontal", and the third as "transverse to the blackboard." The map $[0,1] \times (0,1)^2 \to [0,1] \times (0,1)$ is "the projection to the blackboard."*

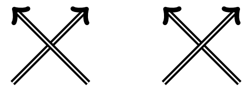
A ribbon $\rho : [0,1]^{\times 2} \hookrightarrow [0,1] \times (0,1)^2$ is blackboard framed *if after projecting to the blackboard, the result $[0,1]^{\times 2} \to [0,1] \times (0,1)$ is an oriented local isomorphism. A ribbon tangle is* regular *if every ribbon in it is blackboard framed, and further the projection to the blackboard is never worse than two-to-one.*

The ends *of the ribbon are the two intervals $\rho(\{0\} \times [0,1]) \subset \{0,1\} \times (0,1)^2$ and $\rho(\{1\} \times [0,1]) \subset \{0,1\} \times (0,1)^2$. The* bottom *and* top *of the ribbon are the ends living, respectively, in $\{0\} \times (0,1)^2$ and $\{1\} \times (0,1)^2$. An end $\rho(\{i\} \times [0,1])$ is* positive *or* negative *according to whether it lives in $\{i\} \times (0,1)^2$ or $\{1-i\} \times (0,1)^2$.*

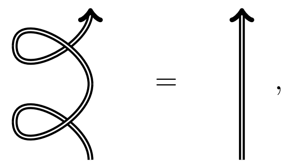
A ribbon tangle is a disjoint union of ribbons (all living in the same $[0,1] \times (0,1)^{\times 2}$). The ends *of a ribbon tangle are the disjoint union of the ends of the ribbon, labeled by whether they are positive or negative. These ends are naturally sorted into the top and bottom.*

The category TangGeo *of* geometric ribbon tangles *has as its objects all possible tops and bottoms of tangles, i.e. all possible disjoint unions of oriented intervals in $(0,1)^2$, each interval being labeled by whether it is a top or a bottom. The morphisms in* TangGeo *are isotopy classes of geometric tangles, where the isotopy is the identity on $\{0,1\} \times (0,1)^2$. Composition is by stacking: it is associative because we take isotopy classes.*

12.3.3.2 Definition *A* tangle diagram *is a diagram (graph) drawn in $[0,1] \times (0,1)$ made out of* *oriented edges* $\Uparrow$ *and crossings*

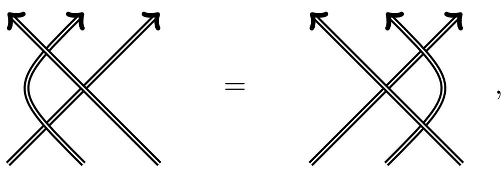


and ending on $\{0,1\} \times (0,1)$*, modulo planar isotopy and the following* framed Reidemeister moves*:*



(12.3.3.3)

(12.3.3.4)



seven more possible orientations. (12.3.3.5)

Equation (12.3.3.3) is a framed version of the first Reidemeister move. The ordinary first Reidemeister move does not hold for ribbons.

The category TangDiag *of ribbon tangle diagrams has as its objects finite sequences of signs and as its morphisms ribbon tangle diagrams. The signs indicate whether the edge points up or down at the the endpoints. For example, the second version of equation (12.3.3.4) is an equation in* $\hom(+-,+-)$*. Composition is by vertical stacking and the monoidal structure is by horizontal stacking.*

Recall that categories $\mathcal{A}$ and $\mathcal{B}$ are *equivalent* if there are functors $F : \mathcal{A} \to \mathcal{B}$ and $G : \mathcal{B} \to \mathcal{A}$ such that $F \circ G \cong \mathrm{id}_{\mathcal{B}}$ and $G \circ F \cong \mathrm{id}_{\mathcal{A}}$.

12.3.3.6 Theorem (Reidemeister theorem for ribbon tangles)
There is a canonical equivalence of categories TangGeo $\simeq$ TangDiag*.*

We can, therefore, unambiguously write simply Tang for either equivalent category TangGeo $\simeq$ TangDiag.

Proof Take any tangle diagram. Construct a geometric tangle in the obvious way: the ribbons live in $[0,1] \times (0,1) \times \{1/2\}$ except at the crossings, where they bump up or down as needed. The Reidemeister moves clearly give equivalent diagrams. This provides the functor TangDiag $\to$ TangGeo. To go the other way, use the fact that every geometric ribbon is isotopic to a blackboard framed one, and every tangle is isotopic to a regular one. Then use the Reidemeister theorem. $\square$

12.3.3.7 Definition *Let* $\mathcal{C}$ *be a monoidal category. A* $\mathcal{C}$*-valued* tangle invariant *is a monoidal functor* Tang $\to \mathcal{C}$*.*

12.3.3.8 Theorem (Jones polynomial)
There is a $\mathcal{TL}$-valued tangle invariant given by sending a sequence of $\pm$s to its length and sending the crossings to:

Composing with the functors $\mathcal{TL} \to \mathcal{U}_q\mathfrak{sl}(2)$-MOD $\to$ VECT sending $n \mapsto (\mathbb{C}^2)^{\otimes n}$ gives a VECT*-valued tangle invariant.* □

This invariant is called the *Jones polynomial* because, by using the version of $\mathcal{TL}$ defined over $\mathbb{Z}[\tau] = \mathbb{Z}[-q - q^{-1}]$, it takes every *link* — every morphism from $\emptyset$ to $\emptyset$ — to an element of $\mathbb{Z}[q^{\pm 1}]$.

12.3.3.9 Remark There are various conventions for the Jones polynomial. Our q is often called $-q^2$ in the quantum topology literature. Often one normalizes the Jones polynomial, multiplying it by some power of q and perhaps dividing by the Jones polynomial of the unknot. ♢

12.3.4 Ribbon Hopf algebras

In Section 12.3.1 we discussed braided monoidal categories and quasitriangular Hopf algebras. Having introduced framed tangles in Section 12.3.3, we can ask: what Hopf structure do they correspond to?

Let us make the question more precise. Suppose you have some categorical notion, say braided monoidal category. What does it mean to say that $\mathcal{B}$ is the "free braided monoidal category on one object"? It means that $\mathcal{B}$ is a braided monoidal category with a distinguished object $X \in \mathcal{B}$, and if you have any braided monoidal category $\mathcal{C}$ with an object $Y \in \mathcal{C}$, then there is a unique-up-to-unique-isomorphism functor $\mathcal{B} \to \mathcal{C}$ of braided monoidal categories taking $X \mapsto Y$. More generally, the free braided monoidal category on n objects $X_1, \dots, X_n$ is the category so that braided monoidal functors out of it are "the same" (meaning up to unique isomorphism) as n-tuples of objects in the target category. It is a general phenomenon that if you know what are all the "free" braided monoidal categories, then you know what it means to be a braided monoidal category.

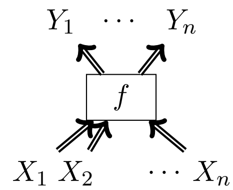
So let us declare that TANG is the "free ribbon category on one object," and ask, OK, what is a ribbon category?

12.3.4.1 Definition *Given a set S, let* TANG(S) *denote the category of S-colored ribbons. It is constructed either geometrically or diagrammatically just like* TANG, *with the change that each ribbon is colored by some element of S (and composition must be consistent with the colors).*

More generally, given a monoidal category $\mathcal{C}$, there is a category TANG$(\mathcal{C})$ *constructed as follows. Its objects are the objects of* TANG$(\mathrm{Ob}(\mathcal{C}))$*, but it has more morphisms. The idea is that morphisms are no longer just ribbon tangles, but now ribbon tangled graphs. There are two versions:*

Diagrammatic *Morphisms are diagrams in the plane, which is to say a certain type of graph, modulo combinatorial moves. In addition to the crossings, for every pair of tuples $(X_1, \dots, X_m)$*

and $(Y_1, \dots, Y_n)$ *of objects in* $\mathcal{C}$ *and for every morphism* $f \in \hom(X_1 \otimes \cdots \otimes X_m, Y_1 \otimes \dots Y_n)$*, we allow a vertex labeled by* f*, with incident edges labeled by* $X_1, \dots, X_m, Y_n, \dots, Y_1$*:*



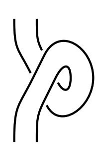
In addition to the Reidemeister moves, we add a move allowing crossings to move past vertices, above or below.

Geometric *Morphisms are geometric ribbon tangles, with an additional allowed ingredient. A* coupon *is an embedded* $\gamma : [0,1]^2 \hookrightarrow (0,1)^3$*; like a ribbon except it is not allowed to touch the boundary of the box. Ribbons and coupons must be disjoint except that ribbons may now end both at the top and bottom of the box but also on coupons, and only in an orientation-consistent way: the only allowed intersection of a ribbon* ρ *and a coupon* γ *is that, for* i *either* 0 *or* 1*, we allow* $\rho(\{i\} \times [0,1]) \subset \gamma(\{1-i\} \times [0,1])$ *preserving orientation. Ribbons should be colored by objects in* $\mathcal{C}$ *and coupons are colored by morphisms just as in the diagrammatic version; the colorings must be consistent in the sense that the ribbons ending at a coupon* $f \in \hom(X_1 \otimes \cdots \otimes X_m, Y_1 \otimes \dots Y_n)$ *must be precisely* $X_1, \dots, X_m, Y_n, \dots, Y_1$*.*

The notion of ribbon category *is whatever it must be so that* $\textsc{Tang}(\mathcal{C})$ *is the free ribbon category generated by* $\mathcal{C}$*. By this we mean: if* $\mathcal{C}$ *is a ribbon category, then there should be a tautological monoidal functor* evaluate : $\textsc{Tang}(\mathcal{C}) \to \mathcal{C}$ *that is the identity on objects and takes vertices/coupons to the corresponding morphisms.*

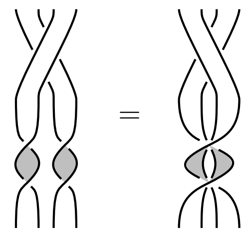
Definition 12.3.4.1 is not a very good definition, because these categories of tangles might be very complicated. How complicated are they? Clearly a ribbon category is rigid, since you can turn a ribbon around, and braided, since ribbons can braid. What else is there? You can twist a ribbon around its main axis by 360°.

See, in a rigid braided monoidal category, a "Reidemeister 1" diagram like



doesn't evaluate to a map $V \to V$. Rather, it gives an isomorphism $\phi_V : V \to {}^{**}V$. If you had crossed the other way, you'd get a different isomorphism $V \to {}^{**}V$. The 360° twist is supposed to be a map $\theta_V : V \to V$ so that $\phi_V \circ \theta_V^{-1}$ is "the" isomorphism $V \cong {}^{**}V$, because for actual ribbons the above two diagrams are isotopic.

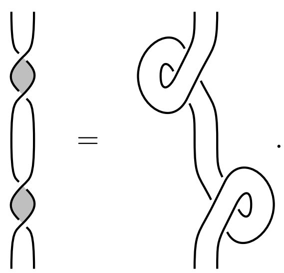
360° twist is related to the braiding in another way:

 (12.3.4.2)

12.3.4.3 Definition *A braided monoidal category* $(\mathcal{C}, \otimes, \beta, \dots)$ *is* balanced *if it is equipped with a natural automorphism* $\theta_V : V \xrightarrow{\sim} V$ *(i.e. an automorphism of the identity functor) called the "full twist" which solves equation (12.3.4.2), or, equivalently,*

$$\beta_{W,V}^{-1} \circ (\theta_V \otimes \theta_W) = \beta_{V,W} \circ \theta_{V \otimes W}.$$

A ribbon category *is a balanced braided monoidal category which is rigid and additionally*

 (12.3.4.4)

In Section 12.3.1 we translated braided monoidality into quasitriangularity by asking "what structure on a Hopf algebra H would make H-MOD braided?" We can ask the same question for ribbon. The answer is:

12.3.4.5 Lemma / Definition *A* ribbon Hopf algebra *is a quasitriangular Hopf algebra* (H, R)*, where* $H = (H, \cdot, 1, \Delta, \epsilon)$ *is a Hopf algebra and* $R \in H^{\otimes 2}$ *is the R-matrix, together with a central invertible element* $\tau \in H$ *such that*

$$\epsilon(\tau) = 1, \tag{12.3.4.6}$$

$$\mathcal{S}(\tau) = \tau, \tag{12.3.4.7}$$

$$\Delta(\tau) = (\tau \otimes \tau) \cdot (\mathrm{swap}(R) \cdot R)^{-1}. \tag{12.3.4.8}$$

If H is ribbon, then H-MOD is ribbon, where we define $\theta_V = \pi_V(\tau)$ for a representation (V, π_V).

12.3.4.9 Proposition *Suppose* (H, R) *is a quasitriangular Hopf algebra. Let* $R = \sum R_1 \otimes R_2$ *and define* $u = \sum \mathcal{S}(R_2) R_1 = m^{\mathrm{op}}\big((\mathrm{id} \otimes \mathcal{S})(R)\big) \in H$*. Assume it is invertible. Then*

$$\epsilon(u) = 1,$$
$$\mathcal{S}^2(a) = uau^{-1} \quad \forall a \in H,$$
$$\mathcal{S}(u)u = u\mathcal{S}(u) \in Z(A),$$
$$\Delta u = (u \otimes u)\big(\mathrm{swap}(R)R\big)^{-1}.$$

Suppose furthermore that we can choose an invertible $b \in H$ such that

$$\Delta(b) = b \otimes b,$$
$$\mathcal{S}^2(a) = bab^{-1} \quad \forall a \in H.$$

Then $\epsilon(b) = 1$ and $\mathcal{S}(b) = b^{-1}$ and

$$\tau \stackrel{\text{def}}{=} b^{-1}u$$

makes H into a ribbon Hopf algebra. □

Theorem 12.3.3.8 says that, up to issues of completion, $\mathcal{U}_q\mathfrak{sl}(2)$ is ribbon.

Exercises

1. (a) Define quantum exterior powers of the quantum plane in terms of q-binomial coefficients. Explain the origin of the quantum determinant $\det_q$.
   (b) Generalize $\mathscr{C}_q(\mathrm{Mat}(2))$ to $\mathscr{C}_q(\mathrm{Mat}(n))$, defined as the quantum matrices acting on

$$\mathbb{C}\langle x_1, \dots, x_n\rangle/(x_j x_i = q x_i x_j, i < j).$$

2. Prove the assertions in Lemma/Definition 12.1.2.1 by considering the category of comodules for the bialgebra $\mathscr{C}_q(\mathrm{Mat}(2))$. Formulate the proof in such a way that it generalizes well to $\mathscr{C}_q(\mathrm{Mat}(n))$ for any n.

3. Check that $\mathscr{C}_q(\mathrm{GL}(2))$ is a Hopf algebra.

4. Find a quasitriangular Hopf algebra whose braided monoidal category of modules is the category SVect of super vector spaces.

5. If $\mathcal{C}$ is a $\mathbb{C}$-linear abelian category, define its *K-group* $K(\mathcal{C})$ to be the abelian group generated by a symbol $[V]$ for each object $V \in \mathscr{C}$, modulo the relation that $[V] + [W] = [X]$ every time there is a short exact sequence $0 \to V \to X \to W \to 0$.
   (a) Suppose $\mathcal{C}$ is monoidal such that $\otimes$ is exact in each variable. Show that this happens, for example, whenever $\mathcal{C}$ is rigid. Show $K(\mathcal{C})$ is an associative ring.
   (b) Suppose furthermore that $\mathcal{C}$ is braided. Show that $K(\mathcal{C})$ is a commutative ring.
   (c) Suppose in fact that $\mathcal{C}$ is symmetric. Show in this case that $K(\mathcal{C})$ is a *λ-ring*, meaning that it has a version of divided powers. The following can be made into a definition of λ-ring: the free commutative ring on n variables is the polynomial ring $\mathbb{Z}[x_1, \dots, x_n]$; the free λ-ring on n variables is the ring of symmetric polynomials $x_1 + \dots + x_n$, $x_1x_2 + x_1x_3 + \dots + x_{n-1}x_n$, …, $x_1 \cdots x_n$.

6. Prove that Proposition 12.3.2.10 gives a complete set of relations for $\mathcal{TL}_n$. Hint: Find an explicit basis and count dimensions. You should find the Catalan numbers.

7. Show that equation (12.3.3.3) does not follow from equations (12.3.3.4) and (12.3.3.5).

8. (a) Check that the invariant presented in Theorem 12.3.3.8 is in fact invariant under the framed Reidemeister moves.
   (b) How does the value of the Jones polynomial change under a Reidemeister-1 move?
   (c) Calculate the value of the Jones polynomial on the two trefoil knots. See that the trefoil knots are not oriented-isotopic.

9. Explain why equation (12.3.4.7) is equivalent to equation (12.3.4.4) and why equation (12.3.4.8) is equivalent to equation (12.3.4.2).

Chapter 13

Higher-rank quantum groups

And so we come to the final chapter of this book. Let us begin with a brief summary of the last three chapters:

- We introduced Lie bialgebras $(\mathfrak{g}, \mathfrak{g}^*)$. They may or may not be *triangular*, *quasitriangular*, or *factorizable*. Just as we can exponentiate any Lie algebra to a Lie group, we can exponentiate any Lie bialgebra $(\mathfrak{g}, \mathfrak{g}^*)$ to a *dual pair* of Poisson Lie groups (G, p) and (G^*, p_*). When G was simple, we gave it a *standard Poisson structure* by realizing it, up to a copy of the Cartan, as the *Drinfeld double* of the upper Borel $\mathfrak{b}_+$.

- We used this standard Poisson structure to understand the geometry and combinatorics of G. In particular, we recognized its symplectic leaves as double Bruhat cells.

- We then investigated a particular quantum group — quantum SL(2) — without really defining what "quantization" means. We saw that it came in two dual forms — $\mathcal{U}_q\mathfrak{sl}(2)$ and $\mathscr{C}(\mathrm{SL}_q(2))$ — which were a *dual pair of Hopf algebras*. We discovered that, at least when q was generic, the representation theory of quantum SL(2) was very similar to the representation theory of SL(2), the major difference being that instead of being symmetric monoidal, it was braided.

Our goal in this chapter is to explain how the story of $\mathcal{U}_q\mathfrak{sl}(2)$ generalizes to other groups, and "quantizes" the Poisson Lie theory of the first two chapters. We will use the quantum theory to build essentially-canonical bases for each finite-dimensional representation.

If G is a group, then $\mathscr{C}(G)$ is a *Hopf algebra*: it is an associative unital algebra $(A, m, 1_A)$ (commutative in the case of $\mathscr{C}(G)$, but not part of the definition) along with a *comultiplication* $\Delta : A \to A^{\otimes 2}$, which is a homomorphism of unital algebras. Δ should be *coassociative* — $(\Delta \otimes \mathrm{id}) \circ \Delta = (\mathrm{id} \otimes \Delta) \circ \Delta$ — and there should be a *counit* — a linear functional $\epsilon : A \to \mathbb{C}$ that is a homomorphism of algebras satisfying $(\mathrm{id} \otimes \epsilon) \circ \Delta = \mathrm{id} = (\epsilon \otimes \mathrm{id}) \circ \Delta$. For a $A = \mathscr{C}(G)$, where G is finite or algebraic, we have $(\Delta f)(x, y) = f(xy)$, and $\epsilon(f) = f(1)$. (When G is algebraic and $\mathscr{C}$ means algebraic functions, we have an isomorphism $\mathscr{C}(G \times G) \cong \mathscr{C}(G) \otimes \mathscr{C}(G)$. If $\mathscr{C}$ means smooth functions, then the left-hand side is a completion of the right-hand side. In the non-algebraic case, Δ lands in the left-hand side, and may not factor through the right-hand side.)

So far we have defined a *bialgebra*. A *Hopf algebra* is a bialgebra along with an *antipode* $\mathcal{S} : A \to A$ that is a bialgebra antiautomorphism — $\mathcal{S}(ab) = \mathcal{S}(b)\mathcal{S}(a)$ and $(\mathcal{S} \otimes \mathcal{S}) \circ \Delta = \Delta^{\mathrm{op}} \circ \mathcal{S}$, where $\Delta^{\mathrm{op}} = \mathrm{swap} \circ \Delta$, where swap is the canonical map $X \otimes Y \to Y \otimes X$ — such that

$$m \circ (\mathcal{S} \otimes \mathrm{id}_A) \circ \Delta = m \circ (\mathrm{id}_A \otimes \mathcal{S}) \circ \Delta = 1_A \circ \epsilon.$$

In the case of a group, $\mathcal{S}(f)(x) = f(x^{-1})$, and the above equation asserts that $f(xx^{-1}) = f(x^{-1}x) = f(1)$.

If G now is an algebraic Poisson Lie group, then $\mathscr{C}(G)$ is a Poisson Hopf algebra, meaning that it is a Poisson algebra and the comultiplication is a Poisson homomorphism. The antipode $\mathcal{S}$ is an anti-Poisson map.

Given an algebraic Lie group G with Lie algebra $\mathfrak{g}$, there is another natural Hopf algebra associated to it: the *universal enveloping algebra* $\mathcal{U}\mathfrak{g}$. It is already "quantum": it quantizes $\mathrm{Sym}\,\mathfrak{g}$ in the sense of Remark 13.0.1.7. We defined a *dual pair of Hopf algebras* to be a pair of Hopf algebras (A, B) with a nondegenerate pairing $\langle,\rangle : A \otimes B \to \mathbb{C}$ that gets along with the Hopf algebra structure:

$$\langle ab, l\rangle = \langle a \otimes b, \Delta_B(l)\rangle, \qquad \langle \Delta_A(a), l \otimes m\rangle = \langle a, lm\rangle,$$
$$\langle 1_A, l\rangle = \epsilon_B(l), \qquad \langle a, 1_B\rangle = \epsilon_A(a),$$
$$\langle \mathcal{S}_A(a), l\rangle = \langle a, \mathcal{S}_B(l)\rangle.$$

13.0.0.1 Example If A is a finite-dimensional Hopf algebra, then its linear dual A^* is also a Hopf algebra, and A, A^* are a dual pair. $\Diamond$

The motivating example of a dual pair of Hopf algebras is $\mathcal{U}\mathfrak{g}$ and $\mathscr{C}(G)$, where the pairing $\langle,\rangle : \mathcal{U}\mathfrak{g} \otimes \mathscr{C}(G) \to \mathbb{C}$ is given by recognizing $\mathcal{U}\mathfrak{g}$ as the differential operators on G supported at the identity. Now if $(\mathfrak{g}, \mathfrak{g}^*)$ is a Lie bialgebra, then we get four Hopf algebras, dual to each other in various ways:

$$\begin{tikzcd}
& \mathfrak{g} \arrow[dl] \arrow[d] \arrow[r, leftrightarrow, bend left, "\text{Lie bialgebra duality}"] & \mathfrak{g}^* \arrow[d] \arrow[dr] & \\
\mathcal{U}\mathfrak{g} \arrow[r, leftrightarrow, bend right, "\text{Hopf duality}"'] \arrow[rrr, leftrightarrow, bend right=60, "\text{Lie bialgebra duality}"'] & \mathscr{C}(G) \arrow[r, leftrightarrow, bend right, "\text{Poisson duality}"'] & \mathscr{C}(G^*) \arrow[r, leftrightarrow, bend right, "\text{Hopf duality}"'] & \mathcal{U}\mathfrak{g}^*
\end{tikzcd}$$

Let's look at a special case: both the bracket and cobracket on $\mathfrak{g}$ are trivial. Then G and G^* are simply the vector spaces $\mathfrak{g}, \mathfrak{g}^*$ thought of as Lie groups, and $\mathscr{C}(G) = \mathrm{Sym}\,\mathfrak{g}^* = \mathcal{U}\mathfrak{g}^*$ and $\mathscr{C}(G^*) = \mathrm{Sym}\,\mathfrak{g} = \mathcal{U}\mathfrak{g}$. In particular, "Hopf duality" and "Poisson duality" become the same thing.

Thus we will begin a program of *quantization*, in which $\mathcal{U}\mathfrak{g}$ and $\mathscr{C}(G)$ are deformed to noncommutative noncocommutative Hopf algebras $\mathcal{U}_\hbar\mathfrak{g}$ and $\mathscr{C}_\hbar(G)$. We will unify the above picture by discovering that, algebraically but not topologically, $\mathcal{U}_\hbar\mathfrak{g} \cong \mathscr{C}_\hbar(G^*)$. I.e. we will have a single Hopf algebra with multiple "classical limits."

13.0.1 Deformation quantization

Recall that if P is a Poisson manifold, then $\mathscr{C}(P)$ is a *Poisson algebra*: it is a commutative algebra along with a Lie bracket $\{,\}$ which satisfies the *Leibniz rule* $\{a, bc\} = \{a, b\}c + b\{a, c\}$. A "quantization" of P is, roughly speaking, a deformation of $\mathscr{C}(P)$ to an associative algebra, where the deformation is "in the $\{,\}$-direction."

13.0.1.1 Definition *Let $(A, m, 1)$ be an associative unital algebra over $\mathbb{C}$. (The same definition works for other types of algebraic structures as well.) A* formal deformation *of A is a $\mathbb{C}[\![\hbar]\!]$-linear associative multiplication $\tilde{m}$, with unit $\tilde{1}$, on $A[\![\hbar]\!]$ of the form*

$$\tilde{m}(a, b) = m(a, b) + \sum_{i=1}^{\infty} \hbar^i m^{(i)}(a, b), \quad \tilde{1} = 1 + \sum_{i=1} \tilde{1}^{(i)}$$

where each $m^{(i)} : A \otimes A \to A$ is extended to $A[\![\hbar]\!]$ by linearity and each $\tilde{1}^{(i)} \in A$.

Two formal deformations $(\tilde{m}, \tilde{1}), (\hat{m}, \hat{1})$ are equivalent *if they are intertwined by an isomorphism $\phi : A[\![\hbar]\!] \to A[\![\hbar]\!]$ of the form*

$$\phi = \mathrm{id} + \sum_{i=1}^{\infty} \hbar^i \phi^{(i)}.$$

Note that any map of the form $\mathrm{id} + O(\hbar)$ is automatically invertible.

There is a better definition of equivalence class of formal deformation:

13.0.1.2 Lemma *A formal deformation of an associative unital algebra A, up to equivalence, is naturally the same as a sheaf $A_\hbar$ of associative unital algebras on the* formal disk $\mathbb{D} \overset{\text{def}}{=} \mathrm{Spec}(\mathbb{C}[\![\hbar]\!])$ *which is flat as a sheaf of vector spaces (all fibers are isomorphic) and such that the fiber $A_0 = A_\hbar|_{\hbar=0}$ is identified with A.* □

13.0.1.3 Remark What we really want are non-formal deformations, meaning flat families parameterized not by $\hbar \in \mathbb{D}$ but by some algebraic variety X. This is a general philosophy of life. If you want to study some structures, look for stable structures that come in families. When you try to classify stable structures, you hope that there is a discrete collection of them. For example, for simple Lie algebras, it is clear that simplicity is an open condition, but in fact for any connected component in the space of simple Lie algebras, all points in that component are isomorphic.

Here's another reason for looking for deformations, at least for someone indoctrinated in physics. The world, it turns out, is not commutative. Rather, the commutativity of classical mechanics arises only in a limit of quantum mechanics as the quantum parameter $\hbar$ goes to 0. The problem with quantization is that it goes in the wrong direction: you know the theory at $\hbar = 0$, and you want to work out the theory for all $\hbar$. ◊

13.0.1.4 Remark Our strategy will be to present A by generators and relations, and then define $A_\hbar$ by the same generators with deformed relations. The problem, then, is to show that the deformed relations do not drop the size of $A_\hbar$. To show that $A_\hbar$ is a flat family amounts to proving a PBW-type theorem giving a vector-space isomorphism $A_\hbar \cong A$. ◊

13.0.1.5 Lemma *Suppose that A is commutative, and choose a formal deformation of A as an associative algebra. Then $\{a,b\} \overset{\text{def}}{=} m^{(1)}(a,b) - m^{(1)}(b,a)$ is a Poisson bracket on A.*

Proof $\{a,b\} = \lim_{\hbar\to 0} \hbar^{-1}[a,b]$, where $[a,b]$ is the commutator in $A_\hbar$. The Jacobi and Leibniz identities for $\{,\}$ follow from those of $[,]$. (The Leibniz identity in a noncommutative algebra is the fact that $[ab,c] = a[b,c] + [a,c]b$.) □

13.0.1.6 Definition *Let A be a Poisson algebra, with Poisson bracket $\{,\}$. A* formal deformation quantization *of A is a formal deformation as an associative algebra such that $\{a,b\} = m^{(1)}(a,b) - m^{(1)}(b,a)$; the commutative algebra A is the* classical limit *of its deformation quantization, and as a Poisson algebra it is the* semiclassical limit*. Two formal deformation quantizations are equivalent if they are equivalent as formal deformations. A formal deformation quantization is a* star product *if additionally each $m^{(i)}$ is a differential operator of two variables. Two star products are equivalent if the intertwining isomorphism ϕ is such that all $\phi^{(i)}$ are differential operators.*

13.0.1.7 Remark Another type of deformation occurs whenever you have a filtration. Suppose that B is a filtered associative algebra, meaning that $B = \bigcup_{i=0}^\infty B_{\leq i}$ and $B_{\leq i}B_{\leq j} \subset B_{\leq i+j}$. Define its *associated graded* to be $\operatorname{gr} B \overset{\text{def}}{=} \bigoplus_{i=0}^\infty B_{\leq i}/B_{\leq i-1}$. (By convention, $B_{\leq -1} = 0$.) Then $\operatorname{gr} B$ is another associative algebra, and B should be thought of as a deformation of $\operatorname{gr} B$.

For example, suppose that $\operatorname{gr} B$ is commutative. Then $[B_{\leq i}, B_{\leq j}] \subset B_{\leq i+j-1}$, and we can define a map $\{,\} : (B_{\leq i}/B_{\leq i-1}) \otimes (B_{\leq j}/B_{\leq j-1}) \to (B_{\leq i+j-1}/B_{\leq i+j-2})$ that makes $A = \operatorname{gr} B$ into a graded Poisson algebra. The filtered algebra B is a type of quantization of A, and the graded algebra A is a type of classical limit of B.

In fact, this filtered/graded version of deformation is essentially equivalent to formal deformation quantization in the sense of Definition 13.0.1.1. The *Rees algebra* of B is

$$\operatorname{Rees}(B) \overset{\text{def}}{=} \bigcup_{i=0}^{\infty} B_{\leq i}\hbar^i[\![\hbar]\!] \subset B[\![\hbar]\!].$$

It is an associative algebra over $\mathbb{C}[\![\hbar]\!]$, since $(B_{\leq i}\hbar^i)(B_{\leq j}\hbar^j) \subset B_{\leq i+j}\hbar^{i+j}$. Moreover,

$$\operatorname{Rees}(B)|_{\hbar=0} = \operatorname{Rees}(B)/\hbar \operatorname{Rees}(B) = \operatorname{gr}(B)$$

since the $\hbar B_{\leq i}\hbar^i$ part of the denominator cancels the copy of $B_{\leq i}\hbar^{i+1}$ inside $B_{\leq i+1}\hbar^{i+1}$. Moreover, if we are working over a field, then we can choose a vector space isomorphism $B \cong \operatorname{gr}(B)$, giving a vector space isomorphism $\operatorname{Rees}(B) \cong \operatorname{gr}(B)[\![\hbar]\!]$.

Look at $B[\![\hbar]\!]$. It has a $\mathbb{C}^\times$-action that rescales $\hbar \mapsto \lambda\hbar$, and $\operatorname{Rees}(B)$ is a $\mathbb{C}^\times$-submodule. Together with the $\mathbb{C}[\![\hbar]\!]$-action, $\operatorname{Rees}(B)$ is a $\mathbb{C}^\times$*-equivariant sheaf on* $\mathbb{D}$. The $\mathbb{C}^\times$-action on $B[\![\hbar]\!]$ extends to an action by the multiplicative monoid $\mathbb{C} = \mathbb{C}^\times \cup \{0\}$, but $\operatorname{Rees}(B)$ is not a $\mathbb{C}$-submodule.

Suppose V is any $\mathbb{C}^\times$-equivariant sheaf on $\mathbb{D}$. Away from 0, the $\mathbb{C}^\times$-action identifies all the fibers of V, and so we can talk about sections of V that are constant relative to the $\mathbb{C}^\times$-action; we'll call the space of such sections $\Gamma^\flat(\mathbb{D} \smallsetminus \{0\}; V)$. This space has a filtration defined as follows: a section $v(\hbar) \in \Gamma^\flat(\mathbb{D} \smallsetminus \{0\}; V)$ is of filtration $\leq i$ if $\hbar^i v(\hbar)$ extends over $\hbar = 0$, i.e. if v has a pole of order at most i at the origin. In the case of $\mathrm{Rees}(B)$, the fibers away from the origin are all copies of B, so $\Gamma^\flat(\mathbb{D} \smallsetminus \{0\}; \mathrm{Rees}(B)) \cong B$, and we recover the filtration on B. In general, we have an adjunction

$$\{\text{filtered objects}\} \leftrightarrows \{\mathbb{C}^\times\text{-equivariant sheaves on } \mathbb{D}\}$$

where the $\to$ arrow is the Rees construction and the $\leftarrow$ arrow is this filtration. It is not an equivalence of categories because $\Gamma^\flat(\mathbb{D} \smallsetminus \{0\}; V)$ might not be the union of its finite-filtration pieces, and because a section might extend over the origin in more than one way. But it is an equivalence except for these two reasons why it isn't. ♢

13.0.1.8 Theorem (Kontsevich quantization)
If P is a finite-dimensional Poisson manifold, then $\mathscr{C}(P)$ admits a star product. □

It's worth emphasizing that Theorem 13.0.1.8 uses certain "smoothness" properties of $\mathscr{C}(P)$, and does not hold for sufficiently non-smooth Poisson algebras. It also can fail for infinite-dimensional Poisson manifolds. Moreover, there is no functor of quantization.

13.0.1.9 Definition *Let $(\mathfrak{g}, \mathfrak{g}^*)$ be a Lie bialgebra. A* formal deformation quantization *of $\mathscr{C}(G)$ is formal deformation in of Hopf algebras in the sense of Definition 13.0.1.1 such that the comultiplication on $\mathscr{C}_\hbar(G) \cong \mathscr{C}(G)[\![\hbar]\!]$ agrees with that on $\mathscr{C}(G)$ to order $\hbar$ and the multiplication deforms in the direction of the Poisson structure on $\mathscr{C}(G)$ in the sense of Definition 13.0.1.6. A* formal deformation quantization *of $\mathcal{U}\mathfrak{g}$ is a formal deformation of Hopf algebras such that the multiplication is undeformed to first order in $\hbar$ and the comultiplciation satisfies*

$$\Delta_\hbar(x) = x \otimes 1 + 1 \otimes x + \hbar\delta'(x) + O(\hbar^2), \quad \delta' - \mathrm{swap} \circ \delta' = \delta$$

for $x \in \mathfrak{g}$. Formal deformation quantizations of $\mathcal{U}\mathfrak{g}$ are also called quantized universal enveloping algebras*.*

13.0.1.10 Theorem (Etingof–Kazhdan)
There is a functor $\mathcal{U}_\hbar$ from Lie bialgebras to Hopf algebras such that $\mathcal{U}_\hbar\mathfrak{g}$ is a formal deformation quantization of $\mathcal{U}\mathfrak{g}$ and such that $\mathcal{U}_\hbar\mathfrak{g}$ and $\mathcal{U}_\hbar\mathfrak{g}^*$ are a dual pair of Hopf algebras. □

The proof of Theorem 13.0.1.10 uses deep results. Rather than using Theorem 13.0.1.10, we will construct these deformation quantizations explicitly. The explicit construction has many advantages, including the possibility of working algebraically with $q = e^\hbar$ as we did in Chapter 12.

Universal enveloping algebras are a canonical example of Remark 13.0.1.7: $\mathcal{U}\mathfrak{g}$ is filtered with associated graded $\mathrm{Sym}\,\mathfrak{g}$; the induced Poisson structure on $\mathrm{Sym}\,\mathfrak{g} = \mathscr{C}(\mathfrak{g}^*)$ is the Lie–Kirilov–Kostant one. Let us temporarily save $\hbar$, and use t as the parameter in the Rees construction. Then $\mathrm{Rees}(\mathcal{U}\mathfrak{g}) = \mathcal{U}\mathfrak{g}_t$, where $\mathfrak{g}_t = \mathfrak{g}$ as a vector space with Lie bracket rescaled to $[,]_t = t[,]$.

Now suppose that $\mathfrak{g}$ is a Lie bialgebra and that we have constructed a quantization $\mathcal{U}_\hbar\mathfrak{g}$. Assuming we can do so universally, we will in fact be able to build a two-parameter family $\mathcal{U}_\hbar\mathfrak{g}_t = \mathcal{U}_{\hbar,t}\mathfrak{g}$. The duality in Theorem 13.0.1.10 is actually between $\mathcal{U}_{\hbar,t}\mathfrak{g}$ and $\mathcal{U}_{t,\hbar}\mathfrak{g}^*$. The $\hbar \to 0$ limit of $\mathcal{U}_{\hbar,t}\mathfrak{g}$ is $\mathcal{U}\mathfrak{g}_t = \mathrm{Rees}(\mathcal{U}\mathfrak{g})$, whereas the $t \to 0$ limit is essentially $\mathrm{Rees}(\mathscr{C}(G^*))$.

13.1 Constructing quantum groups

We constructed the standard Poisson structures on semisimple Lie groups by realizing them as almost the doubles of the Borels. We will construct their quantizations in the same way. This is fairly close to the historical story. Hopf algebras were invented in the late 60s and early 70s, but there were no examples, other than Sweedler's example (isomorphic to what we will call $\mathcal{U}_q\mathfrak{b}_+$, where $\mathfrak{b}_+ \subset \mathfrak{sl}(2)$ is the upper Borel and $q = e^{\hbar/2}$), until the 80s and 90s when Drinfeld and Jimbo constructed $\mathcal{U}_q\mathfrak{g}$ using more or less the procedure below.

13.1.1 The quantum Drinfeld double

13.1.1.1 Lemma / Definition *Let $(H, H^*, \langle\rangle)$ be a dual pair of Hopf algebras. Its* Drinfeld double *$\mathcal{D}(H, H^*)$ is, as a vector space, $H \otimes H^*$. As a coalgebra it is $H \otimes H^\circ$, where H° denotes H^* with the opposite comultiplication: $\Delta_{H^\circ} = \mathrm{swap} \circ \Delta_{H^*}$. The multiplication is*

$$(a \otimes l) \cdot (b \otimes m) = \sum_{b,l} ab_{(2)} \otimes l_{(2)} m \, \langle b_{(1)}, \mathcal{S}^{-1}(l_{(1)}) \rangle \langle b_{(3)}, l_{(3)} \rangle$$

where $\mathcal{S}$ is the antipode for H^* (hence $\mathcal{S}^{-1}$ is the antipode for H°) and $b_{(i)}$ and $l_{(i)}$ are defined by $\Delta_H^{(3)}(b) = (\Delta \otimes \mathrm{id})(\Delta(b)) = (\mathrm{id} \otimes \Delta)(\Delta(b)) = \sum b_{(1)} \otimes b_{(2)} \otimes b_{(3)}$ and $\Delta_{H^*}^{(3)}(l) = \sum l_{(1)} \otimes l_{(2)} \otimes l_{(3)}$. These structures make $\mathcal{D}(H, H^*)$ into a Hopf algebra.

We will write $\mathcal{D}(H)$ for $\mathcal{D}(H, H^*)$ when the choice of dual H^* is understood. We will also write it as $H \bowtie H^\circ$, since it is a generalization of semidirect products.

Recall that a Hopf algebra H is *quasitriangular* if it is equipped with an invertible element $R \in H^{\otimes 2}$ such that

$$(\Delta \otimes \mathrm{id})R = R_{12}R_{23}$$
$$(\mathrm{id} \otimes \Delta)R = R_{13}R_{23}$$
$$\mathrm{swap}(\Delta(a)) = R\,\Delta(a)\,R^{-1}$$

13.1.1.2 Proposition *Suppose H is finite-dimensional. Choose a basis $\{e_i\}$ for H and dual basis $\{e^i\}$ for H^*, and define*

$$R = \sum_i (e_i \otimes 1) \otimes (1 \otimes e^i) \in \mathcal{D}(H)^{\otimes 2}.$$

This R makes $\mathcal{D}(H)$ into a quasitriangular Hopf algebra.

Furthermore, the Hopf algebra structure on $\mathcal{D}(H)$ is the uniquely determined by the requirements that the inclusions $H, H^\circ \hookrightarrow \mathcal{D}(H)$ are Hopf maps and that R is a quasitriangular structure. □

When H is infinite-dimensional, the formula for R does not converge in the algebraic tensor product $\mathcal{D}(H) \otimes \mathcal{D}(H)$. It may converge in an appropriately completed tensor product, and often defines a braiding on the category of finite-dimensional $\mathcal{D}(H)$-modules.

13.1.2 Quantizing $\mathfrak{sl}(2)$

Our expectation, based on the Poisson Lie case, is that we will be able to quantize a semisimple Lie algebra $\mathfrak{g}$ by quantizing its Borel subalgebra $\mathfrak{b}_+$ and then taking a double. To realize this expectation we need to guess a quantization of $\mathfrak{b}_+$. We start with the case of $\mathfrak{g} = \mathfrak{sl}(2)$, so that $\mathfrak{b}_+ = \{H, E\}$ with defining relation $[H, E] = 2E$. We wish to write down a quantized universal enveloping algebra $\mathcal{U}_\hbar\mathfrak{b}_+$, which is supposed to be a Hopf algebra over $\mathbb{C}[\![\hbar]\!]$. Let us make the following guess:

$$\mathcal{U}_\hbar\mathfrak{b}_+ = \mathcal{U}\mathfrak{b}_+[\![\hbar]\!] \text{ as algebras} \tag{13.1.2.1}$$

By this we mean that the algebra structure is undeformed.

Recall that the Lie cobracket on $\mathfrak{b}_+$ was $H \mapsto 0$, $E \mapsto E \wedge H$. Our requirement for $\mathcal{U}_\hbar\mathfrak{b}_+$ is:

$$\Delta_\hbar(H) - \text{swap} \circ \Delta_\hbar(H) = O(\hbar^2) \tag{13.1.2.2}$$

$$\Delta_\hbar(E) - \text{swap} \circ \Delta_\hbar(E) = \hbar H \wedge E + O(\hbar^2) \tag{13.1.2.3}$$

Equation (13.1.2.2) suggests the following guess:

$$\Delta_\hbar(H) = \Delta(H) = H \otimes 1 + 1 \otimes H \tag{13.1.2.4}$$

There are various ways to satisfy equation (13.1.2.3). Consider, for example, a comultiplication like

$$\Delta_\hbar : E \mapsto E \otimes f(\hbar H) + g(\hbar H) \otimes E$$

where $f(x) = 1 + f^{(1)}x + O(x^2)$ and $g(x) = 1 + g^{(2)}x + O(x^2)$. Every map of this form is an algebra homomorphism because $\Delta_\hbar(H) = H \otimes 1 + 1 \otimes H$ and H commutes with $f(\hbar H), g(\hbar H)$. Equation (13.1.2.3) is satisfied if

$$f^{(1)} - g^{(1)} = 1/2.$$

Coassociativity is satisfied if

$$\begin{aligned} E \otimes f(\hbar H) \otimes f(\hbar H) + g(\hbar H) \otimes E \otimes f(\hbar H) + \Delta g(\hbar H) \otimes E \\ = E \otimes \Delta f(\hbar H) + g(\hbar H) \otimes E \otimes f(\hbar H) + g(\hbar H) \otimes g(\hbar H) \otimes E \end{aligned}$$

which happens exactly when $f(\hbar H)$ and $g(\hbar H)$ are grouplike. If H is primitive, what functions of it are grouplike? Exponentials. So we could, for example, take $g = 1$ and $f(x) = \exp(x/2)$. This gives:

13.1.2.5 Definition *$\mathcal{U}_\hbar\mathfrak{b}_+$ is the Hopf algebra over $\mathbb{C}[\![\hbar]\!]$ generated as an algebra by symbols H, E with defining relations*

$$[H, E] = 2E,$$
$$\Delta_\hbar(H) = H \otimes 1 + 1 \otimes H,$$
$$\Delta_\hbar(E) = E \otimes e^{\hbar H/2} + 1 \otimes E.$$

13.1.2.6 Example $\mathcal{U}_\hbar\mathfrak{b}_+$ has a basis $\{H^nE^m\}_{n,m\geq 0}$, and in this basis

$$\begin{aligned}\Delta_\hbar(H^nE^m) &= (\Delta_\hbar H)^n(\Delta_\hbar E)^m \\ &= (\Delta^{(0)}H)^n(\Delta^{(0)}E + h\Delta^{(1)}E + \dots)^m \\ &= (\Delta^{(0)}H)^n\Big(\Delta^{(0)}E^m + \hbar\big((\Delta^{(0)}E)^{m-1}\Delta^{(1)}E + (\Delta^{(0)}E)^{m-2}(\Delta^{(1)}E)(\Delta^{(0)}E) + \dots\big) + O(h^2)\Big)\end{aligned}$$

where $\Delta^{(0)}$ is the undeformed comultiplication for which both E and H are primitive and $\Delta^{(1)}E = E\otimes H/2$ is the linear term in the expansion of $E\otimes e^{hH/2}$. ◊

If we are going to construct the double of $\mathcal{U}_\hbar\mathfrak{b}_+$, we will need its dual $(\mathcal{U}_\hbar\mathfrak{b}_+)^\circ$. The full algebraic dual is too big. An appropriate dual is constructed as follows. First, consider the undeformed Hopf algebra $\mathcal{U}\mathfrak{b}_+$. As a coalgebra it is simply $\mathcal{U}\mathfrak{b}_+ \cong \mathrm{Sym}(\mathfrak{b}_+)$, and so a good dual is $(\mathcal{U}\mathfrak{b}_+)^* \overset{\text{def}}{=} \mathrm{Sym}(\mathfrak{b}_+^*) \cong \mathrm{Sym}(\mathfrak{b}_-)$, so we should look for a dual about that large. Up to issues of completion, $\mathrm{Sym}(\mathfrak{b}_+^*)$ is the algebra of functions on the simply-connected group with Lie algebra $\mathfrak{b}_+$, namely $\mathbb{C}\ltimes\mathbb{C}$. Call the coordinate functions $H^\vee, E^\vee$, corresponding to the matrix

$$(H^\vee, E^\vee) \overset{\text{def}}{=} \begin{pmatrix} e^{H^\vee} & 0 \\ 0 & e^{-H^\vee}\end{pmatrix}\begin{pmatrix} 1 & E^\vee \\ 0 & 1\end{pmatrix} = \begin{pmatrix} e^{H^\vee} & e^{H^\vee}E^\vee \\ 0 & e^{-H^\vee}\end{pmatrix}.$$

Then the multiplication is

$$(H_1^\vee, E_1^\vee)(H_2^\vee, E_2^\vee) = (H_1^\vee + H_2^\vee, e^{-2H_2^\vee}E_1^\vee + E_2^\vee). \tag{13.1.2.7}$$

So $(\mathcal{U}_\hbar\mathfrak{b}_+)^\circ$ should be a deformation of this. Equation (13.1.2.7) tells us the comultiplication to leading order. In fact, since $\mathcal{U}_\hbar\mathfrak{b}_+ = \mathcal{U}\mathfrak{b}_+[\![\hbar]\!]$ as algebras, it is reasonable to try to take equation (13.1.2.7) to all orders.

$$\Delta H^\vee = H^\vee\otimes 1 + 1\otimes H^\vee, \tag{13.1.2.8}$$

$$\Delta E^\vee = E^\vee\otimes 1 + e^{-2H^\vee}\otimes E^\vee. \tag{13.1.2.9}$$

Actually, this is the opposite comultiplication — equation (13.1.2.7) is for $(\mathcal{U}\mathfrak{b}_+)^*$, and equations (13.1.2.8) and (13.1.2.9) are for $(\mathcal{U}_\hbar\mathfrak{b}_+)^\circ$. Finally, we need to make $(\mathcal{U}_\hbar\mathfrak{b}_+)^\circ$ into an algebra. When $\hbar = 0$, it should be commutative, and we can easily read off the first-order deformation:

$$[H^\vee, E^\vee] = -\frac{\hbar}{2}E^\vee + O(\hbar^2). \tag{13.1.2.10}$$

But now we see something remarkable: up to rescaling $H^\vee \rightsquigarrow -\frac{4}{\hbar}H^\vee$ and taking opposite coalgebras, equations (13.1.2.8) to (13.1.2.10) are precisely the defining relations for $\mathcal{U}_\hbar\mathfrak{b}_+$ if we take equation (13.1.2.10) to be exact (dropping the $O(\hbar^2)$ term).

13.1.2.11 Proposition *Define $(\mathcal{U}_\hbar\mathfrak{b}_+)^\circ$ by equations (13.1.2.8) to (13.1.2.10), with the $O(\hbar^2)$-term dropped. Then it is a Hopf algebra, and there is a unique Hopf pairing $\langle,\rangle : \mathcal{U}_\hbar\mathfrak{b}_+\otimes(\mathcal{U}_\hbar\mathfrak{b}_+)^\circ \to \mathbb{C}[\![\hbar]\!]$ such that $\langle H, H^\vee\rangle = \langle E, E^\vee\rangle = 1$ and $\langle H, E^\vee\rangle = \langle E, H^\vee\rangle = 0$.*

This unique pairing satisfies

$$\langle H^nE^m, (H^\vee)^{n'}(E^\vee)^{m'}\rangle = \delta_{n,n'}\,\delta_{m,m'}\,n!\,[m]!$$

where $[m]! \overset{\text{def}}{=} [m][m-1]\dots[1]$ *and* $[m] \overset{\text{def}}{=} \frac{\sinh(\hbar m/2)}{\sinh(\hbar/2)}$ *is the* m*th* quantum integer. □

13.1.2.12 Remark The second term in equation (13.1.2.9) is a problem: what is $e^{-2H^\vee}$? A priori it is a formal power series $1 - 2H^\vee + \frac{1}{2}(-2H^\vee)^2 + \dots$. What we can do is to equip $(\mathcal{U}_\hbar \mathfrak{b}_+)^\circ$ with a filtration that treats $H^\vee$ as "very small," and complete with respect to the corresponding topology.

Define $\mathcal{U}_\hbar \mathfrak{b}_-$ to be the Hopf algebra generated by H, F with $[H, F] = -2F$ and $\Delta(H) = H \otimes 1 + 1 \otimes H$ and $\Delta(F) = F \otimes 1 + e^{-\hbar H/2} \otimes F$. Then the map $(\mathcal{U}_\hbar \mathfrak{b}_+)^\circ \to \mathcal{U}_\hbar \mathfrak{b}_-$ sending $E^\vee \mapsto F$ and $H^\vee \mapsto \hbar H/4$ is an isomorphism after inverting $\hbar$. But we constructed $(\mathcal{U}_\hbar \mathfrak{b}_+)^\circ$ by thinking about $\mathscr{C}(B_+)$, where here B_+ means really its simply-connected cover. So this is an example of the philosophy mentioned at the end of the introduction: $\mathscr{C}_\hbar(B_+) \simeq \mathcal{U}_\hbar \mathfrak{b}_+^*$. ◊

In Example 10.4.2.1 we constructed the standard Lie bialgebra structure on $\mathfrak{sl}(2)$ by recognizing $\mathfrak{sl}(2) \cong \mathcal{D}(\mathfrak{b}_+)/I$, where $\mathcal{D}(\mathfrak{b}_+)$ was the classical double and $I \subset \mathcal{D}(\mathfrak{b}_+)$ was the center. Our strategy to construct $\mathcal{U}_\hbar \mathfrak{sl}(2)$ will be to describe it as

$$\mathcal{U}_\hbar \mathfrak{sl}(2) = \mathcal{D}(\mathcal{U}_\hbar \mathfrak{b}_+)/\tilde{I} = (\mathcal{U}_\hbar \mathfrak{b}_+ \bowtie \mathcal{U}_\hbar \mathfrak{b}_-)/\tilde{I}.$$

The isomorphism $(\mathcal{U}_\hbar \mathfrak{b}_+)^\circ \cong \mathcal{U}_\hbar \mathfrak{b}_-$ is as in Remark 13.1.2.12, and only holds after inverting $\hbar$.

Applying Lemma/Definition 13.1.1.1 and Proposition 13.1.1.2 gives:

13.1.2.13 Proposition *$\mathcal{D}(\mathcal{U}_\hbar \mathfrak{b}_+)$ is generated by $H, H^\vee, E, E^\vee$ with relations*

$$[H, H^\vee] = 0, \qquad [H, E] = 2E, \qquad [H, E^\vee] = -2E^\vee,$$
$$[H^\vee, E^\vee] = -\frac{\hbar}{2}E^\vee, \qquad [H^\vee, E] = \frac{\hbar}{2}E, \qquad [E, E^\vee] = e^{\hbar H/2} - e^{-2H^\vee}.$$

It is a Hopf algebra with comultiplication

$$\Delta(H) = H \otimes 1 + 1 \otimes H \qquad \Delta(E) = E \otimes e^{\hbar H/2} + 1 \otimes E$$
$$\Delta(H^\vee) = H^\vee \otimes 1 + 1 \otimes H^\vee \qquad \Delta(E^\vee) = E^\vee \otimes 1 + e^{-2H^\vee} \otimes E^\vee.$$

After completing appropriately so that the formal power series converges, $\mathcal{D}(\mathcal{U}_\hbar \mathfrak{b}_+)$ is quasitriangular with R-matrix

$$R = \sum_{n,m \geq 0} \frac{H^n E^m \otimes (H^\vee)^n (E^\vee)^m}{n! [m]!} = e^{H \otimes H^\vee} \sum_{m \geq 0} \frac{E^m \otimes (E^\vee)^m}{[m]!}. \qquad \square$$

The $[n]$!s are *q-factorials* at $q = e^{\hbar/2}$ defined by $[n] \stackrel{\text{def}}{=} \frac{\sinh(n\hbar/2)}{\sinh(\hbar/2)}$ and $[n]! \stackrel{\text{def}}{=} \prod_{i=0}^n [i]$. The formal power series $\sum_{n=0}^\infty \frac{x^n}{[n]!}$ was first introduced by Euler.

13.1.2.14 Proposition *The element $\frac{\hbar}{4}H - H^\vee \in \mathcal{D}(\mathcal{U}_\hbar \mathfrak{b}_+)$ is primitive and central, and so generates a Hopf ideal. The quotient $\mathcal{D}(\mathcal{U}_\hbar \mathfrak{b}_+)/\langle \frac{\hbar}{4}H - H^\vee \rangle$ is generated by elements $H, E, F = E^\vee / \sinh(\hbar/2)$ with relations*

$$[H, E] = 2E, \qquad [H, F] = -2F, \qquad [E, F] = \frac{\sinh(\frac{\hbar}{2}H)}{\sinh(\frac{\hbar}{2})},$$

and comultiplication is

$$\Delta(H) = H \otimes 1 + 1 \otimes H$$
$$\Delta(E) = E \otimes e^{\hbar H/2} + 1 \otimes E$$
$$\Delta(F) = E \otimes 1 + e^{-\hbar H/2} \otimes F$$

Quotients of quasitriangular Hopf algebras are quasitriangular. Hence $\mathcal{D}(\mathcal{U}_\hbar\mathfrak{b}_+)/\langle \frac{\hbar}{4}H - H^\vee\rangle$ *is quasitriangular with R-matrix*

$$R = e^{\frac{\hbar}{4}H\otimes H}\sum_{m\geq 0}\frac{\sinh(\frac{\hbar}{2})^m}{[m]!}E^m \otimes F^m$$

$\mathcal{D}(\mathcal{U}_\hbar\mathfrak{b}_+)/\langle \frac{\hbar}{4}H - H^\vee\rangle$ *is clearly a deformation of* $\mathcal{U}\mathfrak{sl}(2)$*, and so we henceforth call it* $\mathcal{U}_\hbar\mathfrak{sl}(2)$*.*

13.1.2.15 Remark The R-matrix does not live in in the algebraic tensor product $\mathcal{U}_\hbar\mathfrak{sl}(2)^{\otimes 2}$. Rather, it lives in the completion consisting of formal power series $\sum a_n\hbar^n$ where $a_n \in \mathcal{U}_\hbar\mathfrak{sl}(2)^{\otimes 2}$.$\Diamond$

Recall the Hopf algebra $\mathcal{U}_q\mathfrak{sl}(2)$ from Section 12.2, or rather the same Hopf algebra with opposite comultiplication. It was defined over $\mathbb{C}(q)$ with generators $E, F, K^{\pm 1}$ and relations

$$KEK^{-1} = q^2E, \quad KFK^{-1} = q^{-2}F, \quad [E,F] = \frac{K-K^{-1}}{q-q^{-1}}$$
$$\Delta(K) = K\otimes K, \quad \Delta(E) = E\otimes 1 + K\otimes E, \quad \Delta(F) = F\otimes K^{-1} + 1\otimes F$$

The map $q \mapsto e^{\hbar/2}$ and $K \mapsto q^H = e^{\hbar H/2}$ defines a homomorphism $\mathcal{U}_q\mathfrak{sl}(2) \to \mathcal{U}_\hbar\mathfrak{sl}(2)[\hbar^{-1}]$.

13.1.2.16 Proposition *There exists an algebra isomorphism* $\phi : \mathcal{U}_\hbar\mathfrak{sl}(2) \xrightarrow{\sim} \mathcal{U}\mathfrak{sl}(2)[\![\hbar]\!]$ *such that* $\phi(H) = H$*.* $\square$

Proving this is exercise 7.

Let us agree that a *finite-dimemsional* representation of $\mathcal{U}_\hbar\mathfrak{sl}(2)$ is a representation on $\mathbb{C}[\![\hbar]\!]$.

13.1.2.17 Corollary *The finite-dimensional representations of* $\mathcal{U}_\hbar\mathfrak{sl}(2)$ *are simply those of* $\mathcal{U}\mathfrak{sl}(2)$*, tensored with* $\mathbb{C}[\![\hbar]\!]$*. Specifically, for each* highest weight $l \in \mathbb{N}$*, there is an* $(l+1)$*-dimensional irrep with basis* $\{v_0^{(l)}, \dots, v_l^{(l)}\}$ *with actions*

$$Hv_m^{(l)} = (l-2m)v_m^{(l)}, \quad Ev_m^{(l)} = v_{m-1}^{(l)}, \quad Fv_m^{(l)} = f(m,l)v_{m+1}^{(l)},$$

for some function $f(m,l)$*.* $\square$

13.1.2.18 Remark In Section 6.2.3 we constructed the connected simply connected algebraic group G from each semisimple Lie algebra $\mathfrak{g}$ by considering the category of finite-dimensional representations and then applying Tannakian reconstruction. Specifically, we defined the ring $\mathscr{O}(G)$ of polynomial functions to consist of all "matrix coefficients" of finite-dimensional $\mathfrak{g}$-representations. Applying the same procedure to $\mathcal{U}_\hbar\mathfrak{sl}(2)$-MOD reconstructs the quantum group $\mathrm{SL}_q(2)$ from Section 12.1 with $q = e^{\hbar/2}$. $\Diamond$

13.1.3 Quantizing the Serre relations

Let $\mathfrak{g}$ be a simple Lie algebra and fix a Borel subalgebra $\mathfrak{b}_+ \subseteq \mathfrak{g}$. We will quantize $\mathfrak{g}$ by following the same steps as for $\mathfrak{sl}(2)$.

Let $\Gamma \subseteq \Delta_+ \subseteq \Delta$ denote the simple roots, the positive roots, and the roots, respectively. For convenience, enumerate the roots: $\Gamma = \{\alpha_1, \dots, \alpha_r\}$. Let a_{ij} denote the Cartan matrix and let $d_i \overset{\text{def}}{=} (\alpha_i, \alpha_i)/2$, so that $b_{ij} \overset{\text{def}}{=} d_i a_{ij}$ is the symmetrized Cartan matrix. Then $\mathcal{U}\mathfrak{b}_+$ is generated by elements H_i, E_i for $i = 1, \dots, r$ with the following relations:

$$[H_i, H_j] = 0 \tag{13.1.3.1}$$
$$[H_i, E_j] = a_{ij} E_j \tag{13.1.3.2}$$
$$\underbrace{[E_i, [E_i, \dots, [E_i}_{1-a_{ij} \text{ times}}, E_j] \dots]]] = 0 \tag{13.1.3.3}$$

Equation (13.1.3.3) is the *Serre relation*. It was first discovered by Chevalley, and Serre proved that it sufficed to define $\mathfrak{b}_+$. Before it was discovered, the only way to present $\mathfrak{b}_+$ was to work with all of Δ_+. We will write $\tilde{\mathfrak{b}}_+$ for the Lie algebra generated by $H_1, \dots, E_r$ with modulo equations (13.1.3.1) and (13.1.3.2), but without the Serre relation.

We'd like to quantize $\mathcal{U}\mathfrak{b}_+$ to a Hopf algebra $\mathcal{U}_\hbar \mathfrak{b}_+$. Our request is that for each i, the subalgebra generated by $\{H_i, E_i\}$ should be a copy of the Hopf algebra from Definition 13.1.2.5. More precisely, we make that request for the short roots, and for the long roots to modify the value of "$\hbar$":

$$\Delta_\hbar(H_i) = H_i \otimes 1 + 1 \otimes H_i \tag{13.1.3.4}$$
$$\Delta_\hbar(E_i) = E_i \otimes e^{\hbar d_i H_i/2} + 1 \otimes E_i \tag{13.1.3.5}$$

Equations (13.1.3.4) and (13.1.3.5) are consistent with equations (13.1.3.1) and (13.1.3.2) presenting $\tilde{\mathfrak{b}}_+$, so we will leave those undeformed:

13.1.3.6 Definition *$\mathcal{U}_\hbar \tilde{\mathfrak{b}}_+ = \mathcal{U}\tilde{\mathfrak{b}}_+[\![\hbar]\!]$ as an algebra: it has generators $\{E_1, \dots, E_r, H_1, \dots, H_r\}$ and relations (13.1.3.1) and (13.1.3.2). It is a Hopf algebra with comultiplication (13.1.3.4) and (13.1.3.5).*

13.1.3.7 Remark The following notation will be very convenient. We set $q \overset{\text{def}}{=} e^{\hbar/2}$ and $q_i \overset{\text{def}}{=} q^{d_i} = e^{d_i \hbar/2}$. These are elements of $\mathbb{C}[\![\hbar]\!]$, but we will soon switch to treating q as a nonperturbative variable. We also set $K_i \overset{\text{def}}{=} q_i^{H_i} = q^{d_i H_i} = e^{d_i H_i \hbar/2} \in \mathcal{U}_\hbar \tilde{\mathfrak{b}}_+$. In these variables, equations (13.1.3.1), (13.1.3.2), (13.1.3.4), and (13.1.3.5) are equivalent to

$$K_i K_j = K_j K_i,$$
$$K_i E_j K_i^{-1} = q_i^{a_{ij}} E_j = q^{b_{ij}} E_j,$$
$$\Delta(K_i) = K_i \otimes K_i,$$
$$\Delta(E_i) = E_i \otimes K_i + 1 \otimes E_i. \qquad \Diamond$$

The Serre relations, however, are not consistent with the comultiplication. Consider equation (13.1.3.3) for adjacent short roots, meaning $d_i = d_j = 1$ and $a_{ij} = a_{ji} = -1$. Is it preserved by $\Delta_\hbar$?

$$\begin{aligned}
0 &\overset{?}{=} [\Delta_\hbar(E_i), [\Delta_\hbar(E_i), \Delta_\hbar(E_j)]] \\
&= \big[E_i \otimes e^{\hbar H_i/2} + 1 \otimes E_i, \big[E_i \otimes e^{\hbar H_i/2} + 1 \otimes E_i, E_j \otimes e^{\hbar H_j/2} + 1 \otimes E_j\big]\big] \\
&= \big[E_i \otimes e^{\hbar H_i/2} + 1 \otimes E_i, \\
&\qquad [E_i, E_i] \otimes e^{\hbar H_i/2} e^{\hbar H_j/2} + E_i \otimes [e^{\hbar H_i/2}, E_j] + E_j \otimes [E_i, e^{\hbar H_j/2}] + 1 \otimes [E_i, E_j]\big] \quad = \dots
\end{aligned}$$

It doesn't seem likely. We run into a problem already with $[e^{\hbar H_i/2}, E_j]$, which isn't anything nice.

To guess a quantization of equation (13.1.3.3), we expand it out in $\mathcal{U}\tilde{\mathfrak{b}}_+$:

$$\underbrace{[E_i, [E_i, \dots, [E_i}_{n \text{ times}}, E_j] \dots]]] = \sum_{s=0}^{n} (-1)^s \binom{n}{s} E_i^{n-s} E_j E_i^s$$

This has a chance of being quantized, because we can quantize the binomial coefficient to a q-*binomial coefficient*:

$$\begin{bmatrix} n \\ s \end{bmatrix}_q \overset{\text{def}}{=} \frac{[n]_q!}{[n-s]_q! [s]_q!}, \quad [n]_q! \overset{\text{def}}{=} \prod_{m=0}^{n} [m]_q, \quad [m]_q \overset{\text{def}}{=} \frac{q^m - q^{-m}}{q - q^{-1}}.$$

For each $i \neq j$, we can then consider the following *Serre element* in $\mathcal{U}_\hbar \tilde{\mathfrak{b}}_+$:

$$\mathrm{Serre}_{ij}^+ \overset{\text{def}}{=} \sum_{s=0}^{1-a_{ij}} (-1)^s \begin{bmatrix} 1 - a_{ij} \\ s \end{bmatrix}_{q_i} E_i^{1-a_{ij}-s} E_j E_i^s$$

13.1.3.8 Remark Every Hopf algebra has an *adjoint action* on itself given by $\mathrm{ad}(x)y \overset{\text{def}}{=} \sum x_{(1)}\, y\, \mathcal{S}(x_{(2)})$, where following Sweedler's notation we have $\Delta(x) = \sum x_{(1)} \otimes x_{(2)}$. In terms of the adjoint action,

$$\mathrm{Serre}_{ij}^+ \propto \mathrm{ad}(E_i)(E_j K_j^{-1}),$$

where by $\propto$ we mean times some powers of q and Ks. $\Diamond$

13.1.3.9 Lemma Serre_{ij}^+ *is* quasiprimitive *in the sense that* $\Delta(\mathrm{Serre}_{ij}^+) = \mathrm{Serre}_{ij}^+ \otimes A + B \otimes \mathrm{Serre}_{ij}^+$ *where* A, B *are grouplike (hence invertible).* $\square$

The proof of Lemma 13.1.3.9 given in [Jan03] proceeds via a tedious and unenlightening computation involving q-binomial identities. There should be a slick proof using Remark 13.1.3.8.

13.1.3.10 Corollary *The ideal* $\langle \mathrm{Serre}^+ \rangle$ *generated by all the* Serre_{ij}^+*s is a Hopf ideal.* $\square$

13.1.3.11 Definition $\mathcal{U}_\hbar \mathfrak{b}_+ \overset{\text{def}}{=} \mathcal{U}_\hbar \tilde{\mathfrak{b}}_+ / \langle \mathrm{Serre}^+ \rangle$.

Now to define $\mathcal{U}_\hbar\mathfrak{g}$ we know what to do: we work out the dual $(\mathcal{U}_\hbar\mathfrak{b})^\circ$, which is essentially $\mathcal{U}_\hbar\mathfrak{b}_-$ up to some factors of $\hbar$; then we take the double $\mathcal{U}_\hbar\mathfrak{b} \bowtie (\mathcal{U}_\hbar\mathfrak{b})^\circ$; then we quotient by a central diagonal copy of the Cartan, and rescale $F_i = E_i^\vee / \sinh(\hbar d_i/2)$. The result is:

13.1.3.12 Definition *$\mathcal{U}_\hbar\tilde{\mathfrak{g}}$ is generated as an algebra over $\mathbb{C}[\![\hbar]\!]$ by elements $\{H_i, E_i, F_i\}_{i=1,\dots,r}$. We give two equivalent versions of the relations, continuing the notation from Remark 13.1.3.7:*

$$\begin{aligned}
[H_i, H_j] &= 0, & K_iK_j &= K_jK_i, && (13.1.3.13)\\
[H_i, E_j] &= a_{ij}E_j, & K_iE_j &= q^{b_{ij}}E_jK_i, && (13.1.3.14)\\
[H_i, F_j] &= -a_{ij}F_j, & K_iF_j &= q^{-b_{ij}}F_jK_i, && (13.1.3.15)\\
[E_i, F_j] &= \delta_{ij}\,\frac{\sinh(\hbar d_i H_i/2)}{\sinh(\hbar d_i/2)}, & [E_i, F_j] &= \delta_{ij}\,\frac{K_i - K_i^{-1}}{q_i - q_i^{-1}}. && (13.1.3.16)
\end{aligned}$$

It is a Hopf algebra with comultiplication

$$\Delta(H_i) = H_i \otimes 1 + 1 \otimes H_i, \qquad (13.1.3.17)$$
$$\Delta(K_i) = K_i \otimes K_i, \qquad (13.1.3.18)$$
$$\Delta(E_i) = E_i \otimes e^{\hbar d_i H_i/2} + 1 \otimes E_i = E_i \otimes K_i + 1 \otimes E_i, \qquad (13.1.3.19)$$
$$\Delta(F_i) = F_i \otimes 1 + e^{-\hbar d_i H_i/2} \otimes F_i = F_i \otimes 1 + K_i^{-1} \otimes F_i. \qquad (13.1.3.20)$$

Equations (13.1.3.17) and (13.1.3.18) are equivalent.

The Serre elements *are*

$$\mathrm{Serre}^+_{ij} \overset{\text{def}}{=} \sum_{s=0}^{1-a_{ij}} (-1)^s \begin{bmatrix} 1-a_{ij} \\ s \end{bmatrix}_{q_i} E_i^{1-a_{ij}-s}E_jE_i^s, \quad \mathrm{Serre}^-_{ij} \overset{\text{def}}{=} \sum_{s=0}^{1-a_{ij}} (-1)^s \begin{bmatrix} 1-a_{ij} \\ s \end{bmatrix}_{q_i} F_i^{1-a_{ij}-s}F_jF_i^s.$$

They generate a Hopf ideal $\langle\mathrm{Serre}\rangle$, and $\mathcal{U}_\hbar\mathfrak{g} \overset{\text{def}}{=} \mathcal{U}_\hbar\tilde{\mathfrak{g}}/\langle\mathrm{Serre}\rangle$.

The Drinfeld–Jimbo quantum group *$\mathcal{U}_q\mathfrak{g}$ is the Hopf algebra over $\mathbb{C}(q)$ generated by E_i, F_i, K_i^{-1} with the above relations.*

13.1.3.21 Lemma $[F_k, \mathrm{Serre}^+_{ij}] = [E_k, \mathrm{Serre}^-_{ij}] = 0$ *in $\mathcal{U}_\hbar\tilde{\mathfrak{g}}$.* □

13.1.3.22 Theorem (Quantum triangular decomposition)
Let $\mathcal{U}_q\mathfrak{n}_-$, $\mathcal{U}_q\mathfrak{h}$, and $\mathcal{U}_q\mathfrak{n}_+$ be the algebras generated just by the F_is, K_is, and E_is, respectively, subject to the pertinent relations from Definition 13.1.3.12. $\mathcal{U}_q\mathfrak{g}$ enjoys a triangular decomposition*: the multiplication map $\mathcal{U}_q\mathfrak{n}_- \otimes \mathcal{U}_q\mathfrak{h} \otimes \mathcal{U}_q\mathfrak{n}_+ \to \mathcal{U}_q\mathfrak{g}$ is an isomorphism. The same statement also holds for the perturbative quantum group $\mathcal{U}_\hbar\mathfrak{g}$.*

Proof Define $\mathcal{U}_q\tilde{\mathfrak{n}}_\pm$ to be the algebras generated by the Es and Fs without the Serre relations. They are clearly free, and $\mathcal{U}_q\tilde{\mathfrak{g}}$ clearly enjoys a triangular decomposition. Lemma 13.1.3.21 implies that

$$\langle\mathrm{Serre}\rangle = (\mathcal{U}_q\tilde{\mathfrak{n}}_-)(\mathcal{U}_q\mathfrak{h})\langle\mathrm{Serre}^+\rangle + \langle\mathrm{Serre}^-\rangle(\mathcal{U}_q\mathfrak{h})(\mathcal{U}_q\tilde{\mathfrak{n}}_+). \qquad (13.1.3.23)$$

In particular, $\langle\text{Serre}\rangle \cap \mathcal{U}_q\mathfrak{h} = 0$. Consider the commuting square

$$\begin{tikzcd} \mathcal{U}_q\tilde{\mathfrak{n}}_- \otimes \mathcal{U}_q\mathfrak{h} \otimes \mathcal{U}_q\tilde{\mathfrak{n}}_+ \arrow[r, "\sim"] \arrow[d, two heads] & \mathcal{U}_q\tilde{\mathfrak{g}} \arrow[d, two heads] \\ \mathcal{U}_q\mathfrak{n}_- \otimes \mathcal{U}_q\mathfrak{h} \otimes \mathcal{U}_q\mathfrak{n}_+ \arrow[r] & \mathcal{U}_q\mathfrak{g} \end{tikzcd}$$

The vertical arrows are the obvious quotients. But equation (13.1.3.23) implies that they have the same kernel, so the bottom arrow is an isomorphism. □

13.1.4 The quantum Weyl group

Fix a semisimple Lie algebra $\mathfrak{g}$. In fact, fix root data for it: simple, positive, and all roots $\Gamma = \{1, \dots, r\} \subset \Delta_+ \subset \Delta$; Cartan matrix a_{ij}; etc.

13.1.4.1 Definition *The* Artin braid group $\mathcal{B}(\mathfrak{g})$ *is the group generated by* $T_1, \dots, T_r$ *with relation*

$$\underbrace{T_iT_jT_iT_j\dots}_{m_{ij}\ \textit{times}} = \underbrace{T_jT_iT_jT_i\dots}_{m_{ij}\ \textit{times}}$$

where for $a_{ij} \cdot a_{ji} = 0, 1, 2, 3$ *we set* $m = 2, 3, 4, 6$ *respectively. The* Weyl group*, of course, is* $\mathcal{W}(\mathfrak{g}) = \mathcal{B}(\mathfrak{g})/\langle T_i^2 = 1\rangle$.

We've emphasized a couple times and we'll emphasize again: the Weyl group does not necessarily act on $\mathfrak{g}$. What acts is all of G, and so any subgroup of G, including the normalizer of the maximal torus; said normalizer is a possibly non-split extension (torus).$\mathcal{W}(\mathfrak{g})$. So $\mathcal{W}(\mathfrak{g})$ acts except that the action is off by some torus elements. In the root basis for $\mathfrak{g}$, i.e. the basis diagonalizing the torus action, $\mathcal{W}(\mathfrak{g})$ acts except that the action is off by some basis-dependent scalars.

So we can correct the fact that Weyl group $\mathcal{W}(\mathfrak{g})$ doesn't quite act by including a torus part. But we can also correct it in another way: after choosing root data, the braid group $\mathcal{B}(\mathfrak{g})$ acts! In fact, it even acts on the quantum group:

13.1.4.2 Theorem (Lusztig's braid group action)
The following formulas define an action of $\mathcal{B}(\mathfrak{g})$ *on* $\mathcal{U}_q\mathfrak{g}$ *as an algebra:*

$$T_i(K_j) = K_jK_i^{-a_{ij}}, \quad T_i(E_i) = -F_iK_i^{-1}, \quad T_i(F_i) = -K_iE_i,$$
$$T_i(E_j) = \sum_{s=0}^{-a_{ij}} (-1)^{-a_{ij}-s} q_i^{-s} E_i^{[-a_{ij}-s]} E_j E_i^{[s]}, \quad T_i(F_j) = \sum_{s=0}^{-a_{ij}} (-1)^{-a_{ij}-s} q_i^{s} F_i^{[-a_{ij}-s]} F_j F_i^{[s]}.$$

As above, $E_i^{[n]} \overset{\text{def}}{=} E_i^n/[n]_{q_i}!$ *is a quantum divided power and* $q_i \overset{\text{def}}{=} q^{d_i}$. □

The action does not preserve the coalgebra structure, a fact we will take advantage of in Theorem 13.2.3.12.

13.1.4.3 Remark Let us check that T_i preserves equation (13.1.3.16), with $i = i$. When $j = i$ it is clear, and so we need to check that, for $j \neq i$,

$$[T_i(E_i), T_i(F_j)] \stackrel{?}{=} 0. \tag{13.1.4.4}$$

The left-hand side of equation (13.1.4.4) is, of course,

$$\left[-F_i K_i^{-1}, \sum_{s=0}^{-a_{ij}} (-1)^{-a_{ij}-s} q_i^s F_i^{[-a_{ij}-s]} F_j F_i^{[s]} \right]$$
$$= -\sum_{s=0}^{-a_{ij}} (-1)^{-a_{ij}-s} q_i^s \left(F_i K_i^{-1} F_i^{[-a_{ij}-s]} F_j F_i^{[s]} - F_i^{[-a_{ij}-s]} F_j F_i^{[s]} F_i K_i^{-1} \right).$$

Multiply by $-(-1)^{a_{ij}}[-a_{ij}]!$ and move the K_i^{-1}s to the right. You get:

$$= \sum_{s=0}^{-a_{ij}} (-1)^s \begin{bmatrix} -a_{ij} \\ s \end{bmatrix} q_i^s \left(q_i^{(-a_{ij})(2)+a_{ij}} F_i^{1-a_{ij}-s} F_i F_i^s - F_i^{-a_{ij}-s} F_j F_i^{s+1} \right) K_i.$$

Kill the K_i, multiply by q_i, and reindex the second summand to give:

$$= \sum_{s=0}^{1-a_{ij}} (-1)^s \left(\begin{bmatrix} -a_{ij} \\ s \end{bmatrix} q_i^{1+s-a_{ij}} + \begin{bmatrix} -a_{ij} \\ s-1 \end{bmatrix} q_i^s \right) F_i^{1-a_{ij}-s} F_i F_i^s.$$

But the usual binomial identity $\binom{n}{s} + \binom{n}{s-1} = \binom{n+1}{s}$ quantizes to $\begin{bmatrix} n \\ s \end{bmatrix} q^{1+n+s} + \begin{bmatrix} n \\ s-1 \end{bmatrix} q^s = \begin{bmatrix} n+1 \\ s \end{bmatrix} q^{n+1}$, and so the we find

$$= q_i^{1-a_{ij}} \sum_{s=0}^{1-a_{ij}} (-1)^s \begin{bmatrix} 1-a_{ij} \\ s \end{bmatrix} F_i^{1-a_{ij}-s} F_i F_i^s,$$

which vanishes by the Serre relation. ◊

One application of Theorem 13.1.4.2 is to give a PBW basis for $\mathcal{U}_\hbar \mathfrak{g}$. For the non-quantum group, the PBW theorem says that $\mathcal{U}\mathfrak{g}$ has a basis consisting of monomials in H_i, E_α, F_α, where $i \in \Gamma$ is a simple root and $\alpha \in \Delta_+$ is a positive root. (The terms in a given monomial are of course required to come in some prescribed order: if $E_\alpha E_\beta^2$ is in the basis, then $E_\beta^2 E_\alpha$ is not.) If we try to give the same basis for $\mathcal{U}_\hbar \mathfrak{g}$, we run into a problem: what is the quantum version of E_α for a non-simple root α?

13.1.4.5 Theorem (Parameterization of positive roots)
A reduced word *for an element $w \in \mathcal{W}(\mathfrak{g})$ is a factorization into simple roots $w = s_{i_1} \cdots s_{i_\ell}$ with ℓ as small as possible; this ℓ is called the* length *of w. The* longest word *is the unique element $w_0 \in \mathcal{W}(\mathfrak{g})$ maximizing the length; its length is $N = |\Delta_+|$. Fix a reduced word $s_{i_1} \cdots s_{i_N}$ for w_0. Then $\{\alpha_{i_1}, s_{i_1}(\alpha_{i_2}), s_{i_1} s_{i_2}(\alpha_{i_3}), \dots, s_{i_1} \cdots s_{i_{N-1}}(\alpha_{i_N})\}$ is the set Δ_+ of positive roots.* □

13.1.4.6 Example In $\mathfrak{sl}(n)$, $s_i(\alpha_{i+1}) = [E_{i,i+1}, E_{i+1,i+2}] = E_{i,i+2}$, where $E_{i,j}$ of course denotes the elementary matrix with a 1 in the (i,j) spot and 0s elsewhere. ◊

Having fixed the reduced word $s_{i_1}\cdots s_{i_N}$ for w_0, we can therefore define the elements $\{E_\alpha\}_{\alpha\in\Delta_+}$ to be the set $\{\alpha_{i_1}, T_{i_1}(\alpha_{i_2}), T_{i_1}T_{i_2}(\alpha_{i_3}), \dots, T_{i_1}\cdots T_{i_{N-1}}(\alpha_{i_N})\}$ where the T_i act as in Theorem 13.1.4.2. Then:

13.1.4.7 Theorem (PBW theorem for $\mathcal{U}_\hbar\mathfrak{g}$)
$\mathcal{U}_\hbar\mathfrak{g}$ has a basis consisting of monomials in the elements $\{F_\alpha, H_i, E_\alpha\}$, where the F_α, E_α are defined relative to a fixed reduced word $s_{i_1}\cdots s_{i_N}$ for w_0. □

This definition for E_α is rather artificial, but there is no better definition. There are many different reduced words for w_0, and "E_α" depends on the choice of reduced word. All reduced words give the same notion of E_i for i a simple root (this is a non-obvious theorem), but they give different sets $\{E_\alpha\}$ after that. The only good news is that different reduced words are related by conjugation, and the different sets $\{E_\alpha\}$ are also conjugate. So up to conjugation the PBW basis is canonical.

13.1.4.8 Remark There is a thriving industry right now in *categorification*. It shows up in many places including "dimensional reduction" of quantum field theories.

Suppose $\mathcal{C}$ is an abelian category. The *K-group* of $\mathcal{C}$ is the abelian group $K(\mathcal{C})$ generated by symbols $[X]$ for each object $X\in\mathcal{C}$ modulo the relation that $[X]+[Y]=[Z]$ any time there is a short exact sequence $0\to X\to Z\to Y\to 0$. If $\mathcal{C}$ is rigid monoidal, then $K(\mathcal{C})$ is naturally a ring, and $\mathcal{C}$ is called a *categorification* of $K(\mathcal{C})$. For example, VECT is a categorification of $\mathbb{Z}$.

If you have a $\mathbb{C}$-algebra A, the question of categorification is to find a monoidal category $\mathcal{C}$ such that $K(\mathcal{C})\otimes_{\mathbb{Z}}\mathbb{C}\cong A$. What about for the quantum group? The category of $\mathbb{Z}$-graded vector spaces categorifies $\mathbb{Z}[q^{\pm1}]$, and so the question is to find $\mathcal{C}$ such that $K(\mathcal{C})\otimes_{\mathbb{Z}[q^{\pm1}]}\mathbb{C}(q)\cong\mathcal{U}_q\mathfrak{g}$.

In particular, $K(\mathcal{C})$ will an *integral form* of $\mathcal{U}_q\mathfrak{g}$. There are many integral forms to choose from. Consider first $\mathcal{U}_q\mathfrak{sl}(2)$. We could generate it over $\mathbb{Z}[q^{\pm1}]$ by just using E, F, $K^{\pm1}$, and $\frac{K-K^{-1}}{q-q^{-1}}$. Then in particular we have elements E^n, but not, say, $\frac{E^2}{2}$. Another natural option is to use *divided powers* $E^{(n)}=\frac{E^n}{n!}$, except that, given Exercise 9a, what we really want are the *quantum divided powers* $E^{[n]}\overset{\text{def}}{=}\frac{E^n}{[n]_q!}$ rather than divided powers. In more detail, set:

$$[m]_q \overset{\text{def}}{=} \frac{q^m-q^{-m}}{q-q^{-1}} \in \mathbb{Z}[q^{\pm1}]$$

$$[m]_q! \overset{\text{def}}{=} [m]_q\cdots[1]_q \in \mathbb{Z}[q^{\pm1}]$$

$$\begin{bmatrix} n \\ m \end{bmatrix}_q \overset{\text{def}}{=} \frac{[n]_q!}{[m]_q![n-m]_q!} \in \mathbb{Z}[q^{\pm1}]$$

$$[K;m]_q \overset{\text{def}}{=} \frac{Kq^m-K^{-1}q^{-m}}{q-q^{-1}}$$

$$\begin{bmatrix} K;c \\ r \end{bmatrix}_q \overset{\text{def}}{=} \prod_{s=1}^{r} \frac{[K;c+1-s]_q}{[s]_q}$$

Then the $\mathbb{Z}[q^{\pm1}]$-subalgebra of $\mathcal{U}_q\mathfrak{sl}(2)$ generated by $E^{[n]}$, $F^{[n]}$, $K^{\pm1}$, and $\left[\begin{smallmatrix} K;c \\ r \end{smallmatrix}\right]_q$ is a Hopf algebra over $\mathbb{Z}[q^{\pm1}]$, called the *divided power form* of $\mathcal{U}_q\mathfrak{sl}(2)$. There are other integral forms of $\mathcal{U}_q\mathfrak{sl}(2)$: for example, you could use divided powers of E but not of F.

To define integral forms of $\mathcal{U}_q\mathfrak{g}$ are similar: you make a series of choices about when to use just powers and when to use divided powers. In order to define them requires Theorems 13.1.4.2 and 13.1.4.5, since we need to define E_α for a positive root α in order to talk about its divided powers. Let $d_\alpha = (\alpha,\alpha)/2$ denote half the length of the root α, and $q_\alpha = q^{d_\alpha}$ and $q_i = q_{\alpha_i}$. The *divided power form* of $\mathcal{U}_q\mathfrak{g}$ is the $\mathbb{Z}[q^{\pm 1}]$-subalgebra of $\mathcal{U}_q\mathfrak{g}$ generated by $E_\alpha^{[n]} \stackrel{\text{def}}{=} \frac{E_\alpha^n}{[n]_{q_\alpha}!}$, $F_\alpha^{[n]} \stackrel{\text{def}}{=} \frac{F_\alpha^n}{[n]_{q_\alpha}!}$, $K_i^{\pm 1}$, and $\left[\begin{smallmatrix} K_i;c \\ r \end{smallmatrix}\right]_{q_i}$.

There's another integral form worth mentioning. In the divided power form, we handled the denominator of equation (13.1.3.16) by introducing the elements $\left[\begin{smallmatrix} K_i;c \\ r \end{smallmatrix}\right]_{q^{d_i}}$. Another option is to define $\bar{E}_\alpha = (q_\alpha - q_\alpha^{-1})E_\alpha$ and $\bar{F}_\alpha = (q_\alpha - q_\alpha^{-1})E_\alpha$. Then $[\bar{E}_\alpha, \bar{F}_\alpha] = (q_\alpha - q_\alpha^{-1})(K_\alpha - K_\alpha^{-1})$ without denominators, and so together with $K_\alpha^{\pm 1}$ they provide an integral form. This form and the divided power form are the two extremes, and there are many other integral forms that lie between them.

By the way, the study of integral forms is important even in the classical limit. The usual integral form of $\mathcal{U}\mathfrak{sl}(2)$ generated by E, F, H has a bad mod-p reduction — it's bad in the sense that the mod-p reduction has a very large center, containing E^p, F^p, and H^p. But there is a divided power form of $\mathcal{U}\mathfrak{sl}(2)$ generated by $E^{(n)} \stackrel{\text{def}}{=} \frac{E^n}{n!}$, $F^{(n)} \stackrel{\text{def}}{=} \frac{F^n}{n!}$, and $\binom{H}{n} \stackrel{\text{def}}{=} \frac{H(H-1)\cdots(H-n+1)}{n(n-1)\cdots 1}$ whose reduction mod p has a much more reasonable center. The quantum analog is that different integral forms of $\mathcal{U}_q\mathfrak{g}$ have very different behavior when q is a root of unity. $\Diamond$

13.2 Representations of quantum groups

13.2.1 Highest weight theory for $\mathcal{U}_q\mathfrak{g}$

We will now study the category of finite-dimensional $\mathcal{U}_q\mathfrak{g}$-modules when q is not a root of unity. Recall from Section 12.2.1 that in the case of $\mathcal{U}_q\mathfrak{sl}(2)$, the representation theory was almost the same as in the classical case, with an extra sign. We will see that in the general case we have a similar situation: $\textsc{Rep}(\mathcal{U}_q\mathfrak{g})$ looks like $\textsc{Rep}(\mathcal{U}\mathfrak{g})$ with some extra combinatorial data.

Let $\Gamma = \{\alpha_1, \dots, \alpha_r\} \subseteq \Delta_+ \subseteq \Delta$ denote the simple, positive, and all roots. From the classical theory, we define the *root lattice* Q to be the $\mathbb{Z}$-span of Γ, and the *weight lattice* to be $P \stackrel{\text{def}}{=} \{\mu \in \sum_{\alpha\in\Pi} m_\alpha \alpha \text{ s.t. } \frac{2(\mu,\alpha)}{(\alpha,\alpha)} \in \mathbb{Z}\}$. Unlike the classical theory, we will also choose $\epsilon \in \hom_{\text{ab qps}}(Q, \{\pm 1\}) = \hom_{\text{sets}}(\Gamma, \{\pm 1\})$. As above, we set $d_\alpha = (\alpha,\alpha)/2$ and $q_\alpha = q^{d_\alpha}$, and write $d_i = d_{\alpha_i}$ and $q_i = q_{\alpha_i}$. We also write $b_{ij} = d_i a_{ij}$ for the symmetrized Cartan matrix.

Now, in parallel with the classical story, we use our understanding of $\textsc{Rep}(\mathcal{U}_q\mathfrak{sl}(2))$ to decompose $\mathcal{U}_q\mathfrak{g}$-modules. Indeed, let M be a finite-dimensional $\mathcal{U}_q\mathfrak{g}$-module. Each $\alpha \in \Gamma$ determines an embedding $\mathcal{U}_{q_\alpha}\mathfrak{sl}(2) \to \mathcal{U}_q\mathfrak{g}$. In particular, K_α will act diagonally with eigenvalues in $\pm q^{\mathbb{Z}d_\alpha}$. Since all the different K_αs commute, we can simultaneously diagonalize all of them. Then M has a *weight decomposition*

$$M = \bigoplus M_{\lambda,\epsilon},$$
$$M_{\lambda,\epsilon} = \{m \in M \text{ s.t. } K_i m = \epsilon(i) q^{(\lambda,\alpha_i)} m \ \forall i\}.$$

Clearly $\lambda \in P$.

Applying the weight decomposition to $\mathcal{U}_q\mathfrak{g}$ itself gives a *root decomposition*

$$\mathcal{U}_q\mathfrak{g} = \bigoplus_{\mu\in Q} (\mathcal{U}_q\mathfrak{g})_\mu.$$

For $\mathcal{U}_q\mathfrak{g}$, all the signs $\epsilon(i)$ are positive, since the K_is act on the E_js and F_js with eigenvalues $+q^{\pm b_{ij}}$. If M is a $\mathcal{U}_q\mathfrak{g}$-module, we clearly have

$$(\mathcal{U}_q\mathfrak{g})_\mu M_{\lambda,\epsilon} \subseteq M_{\lambda+\mu,\epsilon}. \tag{13.2.1.1}$$

Moreover, since the K_is are grouplike, weights are additive under tensor product:

$$(M\otimes N)_{\lambda+\lambda',\epsilon\epsilon'} \supseteq M_{\lambda,\epsilon}\otimes N_{\lambda',\epsilon'}, \quad (M^*)_{\lambda,\epsilon} = (M_{-\lambda,\epsilon})^*. \tag{13.2.1.2}$$

People usually try to get right of the ϵs as soon as they can. By equation (13.2.1.1), for fixed ϵ the vector space $M_\epsilon = \bigoplus_\lambda M_{\lambda,\epsilon}$ is a $\mathcal{U}_q\mathfrak{g}$-submodule of M. Thus $\mathrm{REP}(\mathcal{U}_q\mathfrak{g})$ decomposes as a direct sum of 2^r categories indexed by the possible ϵs (where $r = |\Gamma|$ is the rank of $\mathfrak{g}$). We will write $\mathrm{REP}_\epsilon(\mathcal{U}_q\mathfrak{g}) \subset \mathrm{REP}(\mathcal{U}_q\mathfrak{g})$ for the subcategory of Ms for which $M = M_\epsilon$. Define the *ϵ-twisted trivial module* $\mathbb{C}_\epsilon$ of $\mathcal{U}_q\mathfrak{g}$ to be the one-dimensional module in which K_i acts by $\epsilon(i)$ and E_i, F_i act by 0. Then tensoring with $\mathbb{C}_\epsilon$ gives an equivalence of categories $\mathrm{REP}_+(\mathcal{U}_q\mathfrak{g}) \leftrightarrow \mathrm{REP}_\epsilon(\mathcal{U}_q\mathfrak{g})$, where REP_+ means REP_ϵ for $\epsilon \equiv +1$. By equation (13.2.1.2), $\mathrm{REP}_+(\mathcal{U}_q\mathfrak{g})$ is a monoidal subcategory of $\mathrm{REP}(\mathcal{U}_q\mathfrak{g})$.

Now suppose that $M \in \mathrm{REP}_+(\mathcal{U}_q\mathfrak{g})$. Then for some λ there exists nonzero $v \in M_\lambda$ such that $E_i v = 0$ for all i. Indeed, if $v \in M_\lambda$, then $E_i v \in M_{\lambda+\alpha_i}$, and the set of inhabited weights is finite. Such v is, of course, a *heighest weight vector*.

Now we build the *Verma module* $M(\lambda)$ for $\lambda \in P$ by declaring:

$$M(\lambda) \stackrel{\text{def}}{=} \mathcal{U}_q\mathfrak{g} \otimes_{\mathcal{U}_q\mathfrak{b}_+} \mathbb{C}_\lambda$$

where of course $\mathbb{C}_\lambda$ means the 1-dimensional $\mathcal{U}_q\mathfrak{b}_+$-module on which the E_is act by 0 and the K_is act by q^λ. Just as in the classical case, $M(\lambda)$ has a unique maximal submodule, and so $M(\lambda)$ has a unique simple quotient called $L(\lambda)$. Of course, depending on λ, $L(\lambda)$ might be infinite-dimensional. Nevertheless, so far we have proved:

13.2.1.3 Lemma *If $M \in \mathrm{REP}_+(\mathcal{U}_q\mathfrak{g})$ is an irreducible finite-dimensional $\mathcal{U}_q\mathfrak{g}$-module, then $M \cong L(\lambda)$ for some λ.* □

To continue with the classical story, we need to prove:

13.2.1.4 Proposition *$L(\lambda)$ is finite-dimensional iff $\lambda \in P^+ \stackrel{\text{def}}{=} \{\lambda \in P$ s.t. $\frac{2(\lambda,\alpha)}{(\alpha,\alpha)} \in \mathbb{Z}_{\geq 0}\, \forall \alpha \in \Gamma\}$.*

Proof The "only if" direction is immediate from the $\mathfrak{sl}(2)$-theory: restrict $L(\lambda)$ to the ith $\mathcal{U}_{q_i}\mathfrak{sl}(2) \hookrightarrow \mathcal{U}_q\mathfrak{g}$; see that $q^{(\lambda,\alpha_i)}$ must equal $q^{d_i n_i}$ for some $n_i \in \mathbb{Z}_{\geq 0}$.

The argument for the "if" direction goes as follows. Suppose $\lambda \in P^+$. It suffices to construct some finite-dimensional quotient of $M(\lambda)$, as $L(\lambda)$ will be a quotient of it. (In fact, we will construct $L(\lambda)$, but we don't need that fact. It is in any case a "fragile" fact, since our construction will not give $L(\lambda)$ in most non-semisimple generalizations of the story, e.g. super and finite-characteristic versions.) For each i, let $m_i = d_i^{-1}(\lambda, \alpha_i) \in \mathbb{Z}_{\geq 0}$. We will construct a map

$$\varphi_i : M\big(\lambda - (m_i + 1)\alpha_i\big) \to M(\lambda)$$

by sending the highest-weight vector on the left-hand side to $F_i^{m_i+1}v$, where v is the highest weight vector on the right-hand side. It suffices to show that this is in fact a highest-weight vector, as it clearly has the correct weight. But for $j \neq i$, $E_j F_i^{m_i+1} v = F_i^{m_1+i} E_j v = 0$ trivially, and for $j = i$ it is an $\mathfrak{sl}(2)$-calculation that we have already done in Section 12.2.1.

Set $\tilde{L}(\lambda) \overset{\text{def}}{=} M(\lambda)/\sum \operatorname{Im}(\varphi_i)$, and let v denote the generating highest weight vector. By Theorem 13.1.3.22, $\tilde{L}(\lambda)$ is spanned by elements of the form $F_{i_1} \cdots F_{i_k} v$. We claim that all E_is and F_is act locally nilpotently on it. The statement about the E_is is automatic, since it holds for $M(\lambda)$. For the F_is, let $i \neq j$ and $N \geq 1 - a_{ij}$. Then

$$F_i^N F_j \in \operatorname{span}\{F_i^n F_j F_i^{N-n}\}_{0 \leq n \leq -a_{ij}}.$$

Indeed, for $N = 1 - a_{ij}$, this is just the Serre relation; for $N > 1 - a_{ij}$, you multiply the statement for $N - 1$ on the left by F_i, the apply the Serre relation to the $n = -a_{ij}$ summand. Then for very large N, we have some large $N' = N - (\text{constant})$ such that

$$F_i^N F_{i_1} \cdots F_{i_k} v = \sum (\dots) F_i^{N'} v,$$

which vanishes in $\tilde{L}(\lambda)$ by construction. This prove the claim.

Since E_i and F_i act locally nilpotently, we learn, again because we understand well the $\mathfrak{sl}(2)$-story, that $\tilde{L}(\lambda)$ decomposes as a direct sum of finite-dimensional $\mathcal{U}_{q^{d_i}}\mathfrak{sl}(2)$-modules. This implies that if we decompose $\tilde{L}(\lambda)$ into weight spaces, the set of inhabited weights is invariant under the ith simple reflection. It is therefore invariant under the whole Weyl group.

But the weights are bounded above by λ already for $M(\lambda)$, and so by Weyl invariance there are only finitely many inhabited weights. Each weight space is finite-dimensional already for $M(\lambda)$. It follows that $\dim \tilde{L}(\lambda) < \infty$, completing the proof. □

13.2.1.5 Proposition *The action of $\mathcal{U}_q\mathfrak{g}$ on the sum of all modules in $\mathrm{Rep}_+(\mathcal{U}_q\mathfrak{g})$ is faithful.*

Proof Suppose $u \in \mathcal{U}_q\mathfrak{g}$ vanishes on all modules in $\mathrm{Rep}_+(\mathcal{U}_q\mathfrak{g})$. We use the triangular decomposition from Theorem 13.1.3.22. Choose bases $\{x_I\}$ for $\mathcal{U}_q\mathfrak{n}_+$ and $\{y_J\}$ for $\mathcal{U}_q\mathfrak{n}_-$, each homogeneous for the action of $\mathcal{U}_q\mathfrak{h}$, so that we can talk about $\operatorname{wt}(x_I), \operatorname{wt}(y_J) \in Q$; Theorem 13.1.4.7 provides such a basis. Then $u = \sum_{I,J,\nu} a_{J\nu i} y_J K_\nu x_I$ where K_ν is a monomial in the K_is corresponding to $\nu \in Q$ the root lattice. The sum is finite; we will choose λ, μ large compared to $\operatorname{wt}(x_I), -\operatorname{wt}(y_J)$ for those i, j appearing in u.

Let $\tilde{L}(\lambda)$ be as in the proof of Proposition 13.2.1.4. Define the *Cartan involution* $\omega \in \operatorname{Aut}(\mathcal{U}_q\mathfrak{g})$ to be the automorphism sending $E_i \mapsto \mathcal{S}(F_i) = -K_i F_i$, $F_i \mapsto \mathcal{S}(E_i) = -E_i K_i^{-1}$, and $K_i \mapsto \mathcal{S}(K_i) = K_i^{-1}$. We can pull back modules along automorphisms, thereby defining $\tilde{L}(\mu)^\omega$, which just like $\tilde{L}(\lambda)$ is in $\mathrm{Rep}_+(\mathcal{U}_q\mathfrak{g})$. Unpacking the construction, we see that $\tilde{L}(\mu)^\omega$ is a lowest-weight module with lowest weight $-\mu$. Let $v_\lambda \in \tilde{L}(\lambda)$ and $w_\mu \in \tilde{L}(\mu)^\omega$ denote, respectively, the highest and lowest weight vectors. Then in particular

$$u(v_\lambda \otimes w_\mu) = 0. \tag{13.2.1.6}$$

Each x_I is a product of Es, and each y_J is a product of Fs. Inspecting the comultiplication, we find that $x_I(v_\lambda \otimes w_\mu) = v_\lambda \otimes x_I w_\mu$, since v_λ is a highest weight vector. Thus

$$y_J K_\nu x_I(v_\lambda \otimes w_\mu) = y_J K_\nu(v_\lambda \otimes x_I w_\mu) = q^{(\nu,\lambda+\mathrm{wt}(x_I)-\mu)} y_J(v_\lambda \otimes x_I w_\mu).$$

and so equation (13.2.1.6) unpacks to

$$\sum_{I,J,\nu} a_{J\nu i} q^{(\nu,\lambda+\mathrm{wt}(x_I)-\mu)} y_J(v_\lambda \otimes x_I w_\mu) = 0 \tag{13.2.1.7}$$

The action of the xs on the lowest-weight Verma $M(\mu)^\omega$ is free, and this remains approximately true for $\tilde{L}(\mu)^\omega$. By "approximately true," we mean when μ is large compared to $\mathrm{wt}(x_I)$. Choose I_0 to maximize $\mathrm{wt}(x_{I_0})$ such that $a_{J\nu I_0} \neq 0$. Then, with μ very large, $x_I w_\mu \neq 0$. Look at $y_J(v_\lambda \otimes x_{I_0} w_\mu)$. It is a sum of various terms like $(y'_J v_\lambda) \otimes (K_\alpha y''_J x_{I_0} w_\mu)$. One of the summands is $y_J v_\lambda \otimes x_{I_0} w_\mu$, and for all other summands the second tensorand has strictly lower weight. Thus equation (13.2.1.7) implies:

$$\sum_{J,\nu} a_{J\nu I_0} q^{(\nu,\lambda+\mathrm{wt}(x_{I_0})-\mu)} y_J v_\lambda \otimes x_{I_0} w_\mu = 0. \tag{13.2.1.8}$$

We can repeat the trick: for λ very large, the ys act freely on v_λ. We conclude, therefore, that for all large enough λ, μ, for the above I_0 and for all j,

$$\sum_{\nu} a_{J\nu I_0} q^{(\nu,\lambda+\mathrm{wt}(x_{I_0})-\mu)} = 0.$$

Changing of variables by $b_\nu = q^{(\nu,\mathrm{wt}(x_{I_0}))} a_{J\nu I_0}$, we find

$$\sum_{\nu} b_\nu q^{(\nu,\lambda-\mu)} = 0. \tag{13.2.1.9}$$

But a theorem of Artin's says that distinct characters of an abelian group are linearly independent. Since equation (13.2.1.9) holds for all sufficiently large λ, μ, we must conclude that b_ν, and hence $a_{J\nu I_0}$, vanishes for all ν, J. But we had chosen I_0 such that $a_{J\nu I_0} \neq 0$. The only way out is if $u = 0$ at the beginning. □

Given a weight $\lambda \in P^+$, we have been working with three modules: the Verma module $M(\lambda)$ with highest weight vector v_λ, its simple quotient $L(\lambda)$, and its maximal finite-dimensional quotient $\tilde{L}(\lambda) = M(\lambda)/\langle F_i^{m(i)+1} v\rangle$. All of these modules make sense when q is a fixed complex number. Indeed, they make sense already over $\mathbb{Q}[q^{\pm 1}, (q^d - q^{-d})^{-1}]$, where $d = \mathrm{lcm}(d_i)$. For any module V with a weight space decomposition, including for any finite-dimensional module, define the *character* by $\mathrm{ch}\, V \overset{\text{def}}{=} \sum_{\mu \in P} \dim V_\mu e^\mu \in \mathbb{Z}[e^P]$. Here $\mathbb{Z}[e^P]$ is the integral group algebra of P; we write e^P to emphasize that we are thinking of the abelian group P multiplicatively when we write the group algebra, and e^μ is the basis element of $\mathbb{Z}[e^P]$ corresponding to $\mu \in P$. If V is infinite-dimensional, then $\mathrm{ch}\, V$ is an infinite sum, living in $\mathbb{Z}[\![e^P]\!]$.

13.2.1.10 Theorem (Weyl character formula for $\mathcal{U}_q\mathfrak{g}$)
Suppose q is transcendental over $\mathbb{Q}$. Then $L(\lambda) = \tilde{L}(\lambda)$, and $\operatorname{ch} L(\lambda)$ is given by the Weyl character formula.

Proof Let $\mathbb{K} \stackrel{\text{def}}{=} \mathbb{Q}(q)$ and $A \stackrel{\text{def}}{=} \mathbb{Q}[q^{\pm 1}] \subseteq \mathbb{K}$. Let V be any quotient of $\tilde{L}(\lambda)$, for example $V = \tilde{L}(\lambda)$ or $V = L(\lambda)$. Then V is spanned by $F^I v_\lambda$ for finitely many Is. Let V_A be the A-submodule spanned by the $F^I v_\lambda$s.

Since A is a principle ideal domain, V_A is a free finitely-generated A module. Pick an "$\mathfrak{sl}(2)$ triple" $\{E_i, F_i, K_i\}$, and recall the symbol

$$[K; m]_q \stackrel{\text{def}}{=} \frac{Kq^m - K^{-1}q^{-m}}{q - q^{-1}}.$$

By working in $\mathcal{U}_q\mathfrak{sl}(2)$, we see that V_A is invariant under the operators E_i, F_i, $K_i^{\pm 1}$, and also $[K_i; m]_{q_i}$ for all m.

Clearly $V_A \otimes_A \mathbb{K} = V$, and in fact this holds for each weight space: $V_{A,\mu} \otimes_A \mathbb{K} = V_\mu$. But the point of introducing V_A is that it has a specialization at $q = 1$. Define $\bar{V} \stackrel{\text{def}}{=} V_A \otimes_A \mathbb{C}$ where $A \to \mathbb{C}$ sends $q \mapsto 1$; it is a $\mathbb{C}$-vector space. Let e_i, f_i, k_i, and h_i denote the operators on $\bar{V}$ corresponding respectively to E_i, F_i, K_i, and $[K_i; 0] = \frac{K_i - K_i^{-1}}{q_i - q_i^{-1}}$. Then $k_i \equiv 1$ and for any weight vector v_μ with weight μ, we find

$$h_i v_\mu = \left. \frac{q^{(\mu,\alpha_i)} - q^{-(\mu,\alpha_i)}}{q^{(\alpha_i,\alpha_i)/2} - q^{-(\alpha_i,\alpha_i)/2}} v_\mu \right|_{q=1} = \frac{2(\mu, \alpha)}{(\alpha, \alpha)} v_\mu$$

either by factoring or by applying L'Hospital's rule.

Thus $\{e_i, f_i, h_i\}$ is a classical $\mathfrak{sl}(2)$-triple. The quantum Serre relations specialize at $q \to 1$ to the classical Serre relations, and we discover that $\bar{V}$ is a $\mathfrak{g}$-module.

In particular, $\overline{\tilde{L}(\lambda)}$ and $\overline{L(\lambda)}$ are both finite-dimensional $\mathfrak{g}$-modules with heighest weight λ, and so have the same dimension and indeed the same character. But the character was not changed by the passage $V \rightsquigarrow V_A \rightsquigarrow \bar{V}$. We conclude that $\tilde{L}(\lambda) = L(\lambda)$, as they have the same dimension, and that their character agrees with the classical case. ☐

13.2.1.11 Corollary *If q is transcendental, then the category $\textsc{Rep}(\mathcal{U}_q\mathfrak{g})$ of finite-dimensional $\mathcal{U}_q\mathfrak{g}$-modules is semisimple.*

We employ a standard argument that works in many examples, including category $\mathcal{O}$, Kac–Moody algebras, etc.

Proof It suffices to prove that its subcategory $\textsc{Rep}_+(\mathcal{U}_q\mathfrak{g})$ is semisimple. For this, it suffices to prove any exact sequence $0 \to L(\mu) \to M \to L(\lambda) \to 0$ splits. Recall the standard ordering on the weights generated by $\mu < \mu + \alpha_i$. We consider various cases:

1. Suppose $\mu = \lambda$. Then $\dim M_\lambda = 2$, but the module is semisimple over $\mathcal{U}_q\mathfrak{h} = \mathbb{C}[K_1^{\pm 1}, \dots, K_r^{\pm 1}]$, so $M_\lambda = \mathbb{K}v_\lambda \oplus \mathbb{K}v'_\lambda$, where v_λ is the highest weight vector of $L(\mu) = L(\lambda) \hookrightarrow M$. So build the submodule in M_λ generated by v'_λ, and it must be isomorphic to $L(\lambda)$, and so we have the splitting.

2. If $\mu < \lambda$, then $\dim M_\lambda = 1$. The vector of weight λ is a highest weight vector in M, and so generates a copy of $\tilde{L}(\lambda) = L(\lambda)$ inside M.

3. If $\lambda < \mu$, then we go to the dual modules, and reduce to the previous case.

4. If λ, μ are incomparable, then λ is not a weight of $L(\mu)$, but it is a weight of M, and the corresponding vector is a highest weight vector. □

13.2.2 $\mathcal{Z}(\mathcal{U}_q\mathfrak{g})$

We turn now to the center of $\mathcal{U}_q\mathfrak{g}$ and the quantum Harish-Chandra formula. We studied the classical version of the Harish-Chandra formula in Theorem 9.4.1.14 and the quantum $\mathfrak{sl}(2)$ version in Theorem 12.2.1.8. We will work over $\mathbb{K} \overset{\text{def}}{=} \mathbb{Q}(q)$. We will abbreviate $\mathcal{U} \overset{\text{def}}{=} \mathcal{U}_q\mathfrak{g}$ and let $\mathcal{Z} \overset{\text{def}}{=} \mathcal{Z}(\mathcal{U})$ denote its center. Recall the root decomposition $\mathcal{U} = \bigoplus \mathcal{U}_\mu$, where $\mathcal{U}_\mu \subset \mathcal{U}$ is the subspace transforming under the Ks with weight μ. Let $\mathcal{U}^0 = \mathcal{U}_q\mathfrak{h} = \mathbb{K}[K_1^{\pm 1}, \dots, K_r^{\pm 1}]$. Then clearly $\mathcal{U}^0 \subseteq \mathcal{U}_0$, and indeed $\mathcal{U}_0$ is precisely the centralizer of $\mathcal{U}^0$. Thus $\mathcal{Z} \subseteq \mathcal{U}_0$. Moreover, each $u \in \mathcal{U}_0$ has a triangular decomposition of the form $u = \sum_{\mathrm{wt}\, I = -\mathrm{wt}\, J} a_{I\mu J} F^I K_\mu E^J$. Define

$$\pi(u) = \sum_\mu a_{\emptyset\mu\emptyset} K_\mu.$$

Then π is a projection $\mathcal{U}_0 \to \mathcal{U}^0$. Its restriction to $\mathcal{Z}$ is the *Harish-Chandra homomorphism* for $\mathcal{U} = \mathcal{U}_q\mathfrak{g}$. Given $\lambda \in P$, define $\lambda : \mathcal{U}^0 \to \mathbb{K}$ by $\lambda(K_\mu) = q^{(\lambda,\mu)}$.

13.2.2.1 Lemma *1. If $z \in \mathcal{Z}$, then $z|_{M(\lambda)} = \lambda(\pi(z))\mathrm{id}$.*

2. $\pi|_{\mathcal{Z}} : \mathcal{Z} \to \mathcal{U}^0$ is a ring homomorphism.

Proof 2. is immediate from 1. For 1., since z is central, it suffices to look at the highest vector, and $zv_\lambda = \sum a_{\emptyset\mu\emptyset} q^{\lambda,\mu} v_\lambda$. □

13.2.2.2 Lemma *$\pi|_{\mathcal{Z}}$ is injective.*

Proof Suppose that $\pi(z) = 0$. Then $z|_{L(\lambda)} = 0$, so by semisimplicity $z = 0$ on $\mathrm{R}\textsc{ep}_+(\mathcal{U})$, and the result follows from Proposition 13.2.1.5. □

Our goal now is to identify the image of $\pi|_{\mathcal{Z}}$. Based on the classical case, we expect this to involve the shifted Weyl group action. Thus it is convenient to define a "shift" by declaring that for $\nu \in P$, we set $\gamma_\nu(K_\mu) = q^{(\nu,\mu)} K_\mu$.

13.2.2.3 Lemma *$\gamma_{-\rho}(\pi(\mathcal{Z})) \subseteq (\mathcal{U}^0)^W$.*

Proof We copy the proof from the classical case. We have already constructed a nontrivial map $\varphi_i : M(s_i(\lambda + \rho) - \rho) \to M(\lambda)$ when we built the module $\tilde{L}(\lambda)$ in the proof of Proposition 13.2.1.4.

Thus $\lambda(\pi(z)) = \big(s_i(\lambda+\rho)-\rho\big)\pi(z)$, or equivalently $(\lambda+\rho)\big(\gamma_{-\rho}\circ\pi(z)\big) = s_i(\lambda+\rho)\big(\gamma_{-\rho}\circ\pi(z)\big)$. Suppose $\gamma_{-\rho}\pi(z) = \sum a_\mu K_\mu$ and $s_i\big(\gamma_{-\rho}\pi(z)\big) = \sum b_\mu K_\mu$. Then

$$\sum a_\mu q^{(\lambda+\rho,\mu)} = \sum b_\mu q^{(\lambda+\rho,\mu)}$$

for all $\lambda \in P^+$. This forces $a_\mu = b_\mu$. □

13.2.2.4 Lemma *Let $\mathcal{U}^0_{\mathrm{ev}} \subseteq \mathcal{U}^0$ denote the subring spanned by those K_μ with $\mu \in 2P \cap Q$. Then $\gamma_{-\rho}(\pi(\mathcal{Z})) \subseteq \big(\mathcal{U}^0_{\mathrm{ev}}\big)^W$.*

Proof Recall that each $\epsilon : Q \to \{\pm 1\}$ determines a one-dimensional representation $V_{0,\epsilon} : \mathcal{U} \to \mathbb{K}$ in which E_i and F_i act by 0 and K_μ acts by $\epsilon(\mu)$. Consider the map $\tilde\epsilon : \mathcal{U}^0 \to \mathcal{U}^0$ sending $K_\mu \mapsto \epsilon(\mu)K_\mu$. Then clearly $\tilde\epsilon(\mathcal{Z}) = \mathcal{Z}$ and $\tilde\epsilon$ commutes with the shifted Harish-Chandra map:

$$\gamma_{-\rho}\circ\pi\circ\tilde\epsilon = \tilde\epsilon\circ\gamma_{-\rho}\circ\pi.$$

Suppose that $\gamma_{-\rho}\pi(z) = \sum a_\mu K_\mu$, so that $\gamma_{-\rho}\pi\tilde\epsilon(z) = \sum a_\mu \epsilon(\mu) K_\mu$. Both of these must be W-invariant:

$$a_\mu = a_{w\mu}, \quad a_\mu\epsilon(\mu) = \epsilon(w\mu)a_{w\mu}.$$

Take $w = s_i$. Then we learn that

$$s_i(\mu) = \mu - 2\frac{(\mu,\alpha_i)}{(\alpha_i,\alpha_i)}.$$

This forces $\mu \in 2P$. □

Our aim is the following theorem:

13.2.2.5 Theorem (Harish-Chandra isomorphism for $\mathcal{U}_q\mathfrak{g}$)
The shifted Harish-Chandra map $\gamma_{-\rho}\circ\pi : \mathcal{Z} \to (\mathcal{U}^0_{\mathrm{ev}})^W$ is an isomorphism.

Recall our construction of $\mathcal{U} = \mathcal{U}_q\mathfrak{g}$ in Section 13.1 starting with the Drinfeld double of the upper Borel. Namely, we have Hopf subalgebras $\mathcal{U}^{\geq 0} \overset{\mathrm{def}}{=} \mathcal{U}_q\mathfrak{b}_+$ and $\mathcal{U}^{\leq 0} \overset{\mathrm{def}}{=} \mathcal{U}_q\mathfrak{b}_-$, and they are a dual pair of Hopf algebras: there is a nondegenerate pairing $\langle,\rangle : \mathcal{U}^{\leq 0}\otimes\mathcal{U}^{\geq 0} \to \mathbb{K}$ inducing algebra homomorphisms $\mathcal{U}^{\leq 0} \to (\mathcal{U}^{\geq 0})^*$ and $\mathcal{U}^{\geq 0} \to (\mathcal{U}^{\leq 0})^*$. Very explicitly, a typical element of $\mathcal{U}^{\geq 0}$ looks like $E^I K_\mu$ where $\mu \in Q$ and I is some sequence of is. The map $\mathcal{U}^{\leq 0} \to (\mathcal{U}^{\geq 0})^*$ sends F_i and K_i to the linear maps f_i, k_i defined by

$$f_i(E^I K_\mu) = \begin{cases} 0 & \text{if } I \neq \{i\}, \\ \frac{-1}{q^{d_i}-q^{-d_i}} & \text{if } I = \{i\}, \end{cases} \qquad k_i(E^I K_\mu) = \begin{cases} 0 & \text{if } I \neq \emptyset, \\ q^{-(\alpha_i,\mu)} & \text{if } I = \emptyset, \end{cases}$$

Let us confirm that:

13.2.2.6 Lemma *This pairing is nondegenerate.*

The following proof requires that q is transcendental, as we only proved the description of $L(\lambda)$ in that case. But in fact the statement holds whenever q is not a root of unity.

Proof Since we have the root decomposition, it suffices to show that for a fixed weight μ, $\mathcal{U}^{\leq 0}_{-\mu} \otimes \mathcal{U}^{\geq 0}_{\mu} \to \mathbb{K}$ is nondegenerate. Fix $y \in \mathcal{U}^{\leq 0}_{-\mu}$ and suppose that $\langle y, x\rangle = 0$ for every $x \in \mathcal{U}^{\geq 0}_{\mu}$. We want to show that $y = 0$.

We will do this by induction on μ. It is clear when $\mu = 0$ and $\mu = \alpha_i$ a simple root. Now suppose that μ is not a simple root. Then for all $x \in \mathcal{U}^{\geq 0}_{\mu-\alpha_i}$, we have

$$\langle y, E_\alpha x\rangle = \langle y, xE_\alpha\rangle = 0. \tag{13.2.2.7}$$

By induction, we know that

$$\Delta(y) = y \otimes K_{-\mu} + \sum_i r_i(y) \otimes F_i K_{-\mu} K_i + \cdots = 1 \otimes y + \sum_i F_i \otimes r'_i(y) K_i^{-1} + \ldots \tag{13.2.2.8}$$

for some functions $r_i, r'_i : \mathcal{U}^{\leq 0}_{-\mu} \to \mathcal{U}^{\leq 0}_{-\mu+\alpha_i}$.

Either by construction or computation, $\langle , \rangle$ is a Hopf pairing. Applying equation (13.2.2.8) to equation (13.2.2.7) and using our inductive hypothesis then forces:

$$r_i(y) = r'_i(y) = 0. \tag{13.2.2.9}$$

But

$$E_i y - y E_i = \frac{K_i r_i(y) - r'_i(y) K_i^{-1}}{q_i - q_i^{-1}} \tag{13.2.2.10}$$

and so $yE_i = E_i y$ for all i. Look at the actions of yE_i and $E_i y$ on $L(\lambda)$ for λ sufficiently large. If $y \neq 0$, then $yv_\lambda \neq 0$, but if $yE_i = E_i y$ for all i, then $E_i y v_\lambda = y E_i v_\lambda = 0$, and so yv_λ is another highest weight vector. So $y = 0$. □

We will extend the pairing $\langle , \rangle$ to a pairing $(,)$ on all of $\mathcal{U}$ by using the Killing form. Indeed, suppose $y \in \mathcal{U}^{\leq 0}_{-\nu}$, $y' \in \mathcal{U}^{\leq 0}_{-\nu'}$, $x \in \mathcal{U}^{\geq 0}_{\mu}$, $x' \in \mathcal{U}^{\geq 0}_{\mu'}$, and $\lambda, \lambda' \in Q$. Then define

$$\big(yK_\nu K_\lambda x, y'K_{\nu'}K_{\lambda'}x'\big) \stackrel{\text{def}}{=} \langle y', x\rangle\langle y, x'\rangle q^{(2\rho,\nu)} q^{(\lambda,\lambda')} \tag{13.2.2.11}$$

The funny K_νs are there to make the exponents nice on the right-hand side. Clearly the pairing is nonzero only when $\mu = \nu'$ and $\mu' = \nu$, and extends by linearity to all of $\mathcal{U} = \bigoplus_{\eta,\xi>0} \mathcal{U}^{\leq 0}_{-\eta}\mathcal{U}^0\mathcal{U}^{\geq 0}_{\xi}$.

The pairing in equation (13.2.2.11) is not symmetric, but it almost is:

$$(a, b) = q^{(2\rho,\mu-\nu)}(b, a), \quad a \in \mathcal{U}^{\leq}_{\nu}\mathcal{U}^0\mathcal{U}^{\geq 0}_{\mu} \text{ and } b \in \mathcal{U}^{\leq}_{\mu}\mathcal{U}^0\mathcal{U}^{\geq 0}_{\nu}.$$

Let M be a finite-dimensional $\mathcal{U}$-module. For $m \in M$ and $f \in M^*$, define the (f, m)th *matrix coefficient* of $u \in \mathcal{U}$ to be

$$c_{f,m}(u) \stackrel{\text{def}}{=} f(um). \tag{13.2.2.12}$$

13.2.2.13 Lemma *Suppose that the weights of M are in $\frac{1}{2}Q$. Then there exists a unique $u_{f,m} \in \mathcal{U}$ such that $(u_{f,m}, u') = c_{f,m}(u')$ for all $u' \in \mathcal{U}$.*

Proof Uniqueness is trivial because the pairing is nondegenerate. The important part is existence. For this it suffices to check when M is irreducible. Supposing semisimplicity, it suffices to check when f and m are weight vectors. So let's assume $f \in M^*_\gamma$ and $m \in M_\delta$.

Consider $u' = yx$ for $y \in \mathcal{U}^{\leq 0}_{-\mu}$ and $x \in \mathcal{U}^{\geq 0}$. Then the only way $c_{f,m}(u') \neq 0$ is if $\delta+\nu-\mu+\gamma = 0$. Set $\eta = -2(\delta+\nu) \in Q$. By nondegeneracy of the pairing and finite-dimensionality of the root decomposition spaces, we can find $u_0 \in \mathcal{U}^{\leq 0}_{-\nu} K_\eta \mathcal{U}^{\geq 0}_{\mu}$ such that $c_{f,m}(yx) = (u_0, yx)$. The challenge now is to show that this choice of u_0 is consistent if we change η. We have:

$$\langle f, yk_\lambda xm\rangle = q^{(\delta+\nu,\lambda)}\langle f, yxm\rangle \tag{13.2.2.14}$$

$$(u_0, yK_\lambda x) = q^{-(\eta,\lambda)/2}(u_0, yx) \tag{13.2.2.15}$$

In order to be consistent, we must have $(\delta+\nu,\lambda) = -(\eta,\lambda)/2$, which we do have. Thus we have constructed, for any weights μ, ν, an element $u^{\mu,\nu}_{f,m} \in \mathcal{U}$ such that $(u^{\mu,\nu}_{f,m}, u') = c_{f,m}(u')$ for all $u' \in \mathcal{U}^{\leq 0}_{-\mu}\mathcal{U}^0\mathcal{U}^{\geq 0}_{\nu}$. But for any finite-dimensional module, the set of weights for which $c_{f,m} \neq 0$ is finite. So we can set $u_{f,m} \overset{\text{def}}{=} \sum u^{\mu,\nu}_{f,m}$. □

Proof (of Theorem 13.2.2.5) Injectivity follows from Lemma 13.2.2.2. For surjectivity, we will copy the classical case: we will study invariant polynomials, and the shift by ρ will be very natural.

Lemma 13.2.2.13 automatically implies the following. Suppose $\lambda \in P^+$ such that $2\lambda \in Q$. Then there is a unique $z_\lambda \in \mathcal{U}$ such that $(u, z_\lambda) = \mathrm{tr}_{L(\lambda)}(uK_{2\rho})$. We will show that $z_\lambda \in \mathcal{Z}$.

Consider the map $\phi \mapsto \mathrm{tr}(\phi K_{2\rho})$ from $\mathrm{End}(L(\lambda)) \to \mathbb{K}$. We claim it is a $\mathcal{U}$-module homomorphism, where $\mathrm{End}(L(\lambda))$ is a $\mathcal{U}$-module via the adjoint action and $\mathbb{K}$ is a $\mathcal{U}$-module via the counit. Equivalently, we claim that for $u \in \mathcal{U}$ and $\phi \in \mathrm{End}(L(\lambda))$, we have

$$\mathrm{tr}\left(\sum u_{(1)}\,\phi\,\mathcal{S}(u_2)\,K_{2\rho}\right) = \epsilon(u)\ \mathrm{tr}\left(\phi K_{2\rho}\right). \tag{13.2.2.16}$$

It suffices to check on generators. For the Ks it is easy. For $u = E_i$, the left-hand side is

$$LHS = \mathrm{tr}\left((E_i\phi K_i^{-1} - \phi E_i K_i^{-1})K_{2\rho}\right) = \mathrm{tr}\left((E_i\phi - \phi E_i)K_i^{-1}K_{2\rho}\right).$$

Try to commute the $K_i^{-1}K_{2\rho}^{-1}$ past the E_i. You will pick up some power of q. Specifically, you will pick up a factor of $q^{(\alpha_i,\alpha_i)-(2\rho,\alpha_i)} = 1$. So

$$LHS = \mathrm{tr}\left(E_i(\phi K_i^{-1}K_{2\rho}) - (\phi K_i^{-1}K_{2\rho})E_i\right) = 0.$$

The case $u = F_i$ is analogous. So the shifted trace is the correct notion of trace in the quantum case, because the usual trace is not a homomorphism of $\mathcal{U}$-modules, but the shifted trace is.

The map $\mathcal{U} \to \mathrm{End}(L(\lambda))$ is a module map where $\mathcal{U}$ is given the adjoint action. Together with Exercise 14, we learn that

$$\epsilon(u)\,(x, z_\lambda) = (\mathrm{ad}(u)x, z_\lambda) = (x, \mathrm{ad}(\mathcal{S}u)z_\lambda)$$

for all x. By nondegeneracy of the pairing, we must have, for all u,

$$\mathrm{ad}(\mathcal{S}u)z_\lambda = \epsilon(u)z_\lambda.$$

But this happens only when $z_\lambda \in \mathcal{Z}$.

In summary, we have constructed a bunch of elements of $\mathcal{Z}$, one for each $\lambda \in P^+$ such that $2\lambda \in Q$. Let us now work out the image of z_λ under the Harish-Chandra homomorphism $\gamma_{-\rho} \circ \pi$. For any $K_\mu \in \mathcal{U}^0$, we have $(K_\mu, z_\lambda) = (K_\mu, \pi(z_\lambda))$. Suppose that $\pi(z_\lambda) = \sum_\nu a_\nu K_\nu$. Then

$$(K_\mu, z_\lambda) = \sum_\nu a_\nu (K_\mu, K_\nu) = \sum_\nu a_\nu q^{(\mu,\nu)}.$$

On the other hand

$$(K_\mu, z_\lambda) = \operatorname{tr}_{L(\lambda)} K_\mu K_{2\rho} = \sum_\nu (\dim L(\lambda)_\nu) q^{(\nu,\mu+2\rho)} = \sum_\nu (\dim L(\lambda)_\nu) q^{(\nu,2\rho)} q^{(\mu,\nu)}.$$

Assuming q is not a root of unity, we compare powers of q for infinitely many different μ, and conclude that $a_\nu = (\dim L(\lambda)_\nu) q^{(\nu,2\rho)}$. Hence:

$$\gamma_{-\rho}\pi(z_\lambda) = \gamma_{-\rho} \sum_\nu (\dim L(\lambda)_\nu) q^{(\nu,2\rho)} K_\nu = \sum_\nu (\dim L(\lambda)_\nu) q^{(\nu,\rho)} K_\nu.$$

As λ ranges over $P^+ \cap \frac{1}{2}Q$, these span $(\mathcal{U}^0_{\mathrm{ev}})^W$. □

13.2.2.17 Remark Recall that in Section 6.2.3 we defined the algebraic group G for a semisimple Lie algebra $\mathfrak{g}$ by declaring that its ring $\mathscr{C}(G)$ of algebraic functions was the ring of matrix coefficients. Having defined quantum matrix coefficients in equation (13.2.2.12), we can by the same logic construct the *quantum function algebra* $\mathscr{C}_q(G)$ of algebraic functions on the quantum group: $\mathscr{C}_q(G)$ is the subspace of $(\mathcal{U}_q\mathfrak{g})^*$ consisting of matrix coefficients.

Why is $\mathscr{C}_q(G)$ a vector subspace? You can rescale the matrix coefficient $c_{f,m}$ by rescaling f or m and you can add matrix coefficients by taking direct sums of modules. Why is it closed under multiplication? Because you can tensor finite-dimensional modules: let M_1, M_2 be finite-dimensional representations of $\mathcal{U}_q\mathfrak{g}$, and look at $M_1 \otimes M_2$; then for $f_i \in M_i^*$ and $m_i \in M_i$ and for $u \in \mathcal{U}$ we have

$$\begin{aligned} c_{f_1\otimes f_2, m_1\otimes m_2}(u) &= \left\langle f_1 \otimes f_2, \sum u_{(1)}m_1 \otimes u_{(2)}m_2 \right\rangle = \sum \left\langle f_1, u_{(1)}m_1 \right\rangle \left\langle f_2, u_{(2)}m_2 \right\rangle = \\ &= \sum c_{f_1,m_1}(u_{(1)})\, c_{f_2,m_2}(u_{(2)}) = c_{f_1,m_1} c_{f_2,m_2} \quad (\text{product in } (\mathcal{U}_q\mathfrak{g})^*). \end{aligned}$$

It is most natural to define $\mathscr{C}_q(G)$ using the matrix coefficients of all finite-dimensional modules. It has a subring $\mathscr{C}_{+q}(G)$ consisting of matrix coefficients of modules in the category $\mathrm{REP}_+(\mathcal{U}_q\mathfrak{g})$, i.e. those modules in which all the Ks act with eigenvalues in $q^{\mathbb{Z}}$ (and not $-q^{\mathbb{Z}}$). But in fact these give the same ring $\mathscr{C}_q(G)$ provided q is not a root of unity, more or less by Proposition 13.2.1.5. Indeed, $\mathscr{C}_q(G)$ is clearly generated by $\mathscr{C}_{+q}(G)$ together with the matrix coefficients of the modules $V_{0,\epsilon}$ for the possible signs $\epsilon : \{1, \dots, r\} \to \{\pm 1\}$, which is to say the functions that vanish on E, F and send $K_i \mapsto \epsilon(i)$. Look at the matrix coefficient c_j corresponding to the highest weight vector in the jth fundamental representation. It sends $K_i \mapsto q^{\delta_{ij}}$ and vanishes on the Es and Fs. Provided q is not a root of unity, these c_js are linearly independent and all maps $\epsilon : \{1, \dots, r\} \to \{\pm 1\}$ are in their span.

For example, when $\mathfrak{g} = \mathfrak{sl}(2)$, we can construct the ring $\mathscr{C}_q(\mathrm{SL}(2)) = \mathscr{C}(\mathrm{SL}_q(2))$ from Section 12.1 in this way, recovering Theorem 12.2.4.3. Indeed, the defining relations equations (12.1.1.3), (12.1.1.4), (12.1.1.6), (12.1.1.7), (12.1.1.9), and (12.1.1.10) for $\mathscr{C}_q(\mathrm{Mat}(2)) = \mathscr{C}(\mathrm{Mat}_q(2))$ come just from inspecting $V_{1,+}$, and the determinant relation $\det_q \stackrel{\text{def}}{=} qd - qbc = 1$ comes from finding a copy of the trivial module inside $V_{1,+}^{\otimes 2}$. ♢

13.2.3 The quantum R-matrix

What is so important about $\mathcal{U}_q\mathfrak{g}$? There are many important things, some of which we have already seen: the quantization refines and explains a lot of classical combinatorics. Another important aspect connects us back to the discussion of ribbon categories from Sections 12.3.3 and 12.3.4:

13.2.3.1 Proposition *$\mathcal{U}_\hbar\mathfrak{g}$ is quasitriangular with*

$$R = \exp\left(\frac{\hbar}{2}\sum_{i,j=1}^{r} b_{ij} H_i \otimes H_j\right)\left(1 + \sum_{i=1}^{r} \sinh\left(\frac{\hbar d_i}{2}\right) E_i \otimes F_i + \dots\right)$$

where r is the rank of $\mathfrak{g}$ and a_{ij} is the Cartan matrix, and $b = (da)^{-1}$, where $(da)_{ij} = d_i a_{ij}$ is the symmetrized Cartan matrix. And $\dots$ are the higher terms in E, F. Moreover, $\mathcal{U}_\hbar\mathfrak{g}$ is ribbon with ribbon element

$$\tau = 1 - \frac{\hbar}{2}\left(H_\rho + \sum_{ij} b_{ij} H_j H_i + \sum_i F_i E_i\right) + \dots$$

Proof Quasitriangularity and the formula for R are almost immediate from the construction of $\mathcal{U}_\hbar\mathfrak{g}$ as a quantum double. The formula for τ follows from Proposition 12.3.4.9 and Exercise 8, which together imply that any quasitriangular structure on $\mathcal{U}_\hbar\mathfrak{g}$ determines a ribbon structure. □

The formula for R in Proposition 13.2.3.1 does not converge in the algebraic tensor product $\mathcal{U}_\hbar\mathfrak{g} \otimes \mathcal{U}_\hbar\mathfrak{g}$. But it does converge in the $\hbar$-adic completed tensor product. This is good enough to imply:

13.2.3.2 Corollary *The category $\mathrm{Rep}(\mathcal{U}_\hbar\mathfrak{g})$ of finite-dimensional $\mathcal{U}_\hbar\mathfrak{g}$-modules is braided and in fact ribbon.* □

$\mathrm{Rep}(\mathcal{U}_\hbar\mathfrak{g})$ and $\mathrm{Rep}_+(\mathcal{U}_q\mathfrak{g})$ are essentially the same: the difference between $\mathcal{U}_\hbar\mathfrak{g}$ and $\mathcal{U}_q\mathfrak{g}$ is that in the former, $K_i - 1 = O(q-1)$, so a representation of $\mathcal{U}_q\mathfrak{g}$ extends to a representation of $\mathcal{U}_\hbar\mathfrak{g}$ exactly when all the K_is act by $+1$ when $q \to 1$. The braiding extends to all of $\mathrm{Rep}(\mathcal{U}_q\mathfrak{g})$ by declaring that the $V_{0,\epsilon}$s braid trivially with all modules. In fact, the formula for R is nicely algebraic. See, there is an infinite sum

$$1 + \sum_{i=1}^{r} \sinh\left(\frac{\hbar d_i}{2}\right) E_i \otimes F_i + \dots$$

where the kth term in the sum has expressions of the form $E^k \otimes F^k$ — it is the dual to the pairing $\langle,\rangle : \mathcal{U}_q\mathfrak{n}_- \otimes \mathcal{U}_q\mathfrak{n}_+$. The coefficients are Laurent polynomials in q. For example, $\sinh\left(\frac{\hbar d_i}{2}\right) =$

$\frac{1}{2}(q^{d_i} - q^{-d_i})$. For any finite-dimensional modules M_1, M_2, this sum converges, because for large enough k, $E^k|_{M_1}$ and $F^k|_{M_2}$ vanish. Indeed, as soon as both M_is are highest-weight modules, the sum will converge. On the other hand, suppose that $v_\mu \in M_1$ and $v_\nu \in M_2$ are weight vectors. Then

$$\exp\left(\frac{\hbar}{2}\sum_{i,j=1}^{r} b_{ij} H_i \otimes H_j\right)(v_\mu \otimes v_\nu) = q^{(\mu,\nu)}(v_\mu \otimes v_\nu).$$

So the braiding on $\mathrm{Rep}_+(\mathcal{U}_q\mathfrak{g})$ is reasonable and computable. Note, though, that it is sometimes valued in fractional powers of q, since μ, ν are weights, not roots.

13.2.3.3 Example We wrote down already the R-matrix in the case $\mathfrak{g} = \mathfrak{sl}(2)$ in Proposition 13.1.2.14; it was:

$$R = e^{\frac{\hbar}{4} H\otimes H} \sum_{n=0}^{\infty} \frac{1}{[n]_q!}\left(\frac{q-q^{-1}}{2} E \otimes F\right)^n.$$

As always, $q = e^{\frac{\hbar}{2}}$ and $[n]_q = \frac{q^n - q^{-n}}{q - q^{-1}}$ and $[n]_q! = [n]_q \cdots [1]_q$. Let us work out the ribbon element τ. According to Proposition 12.3.4.9 applied to Exercise 8, we first need to compute $u = m^{\mathrm{op}}((\mathrm{id} \otimes \mathcal{S})(R))$. Let $a_n = \frac{1}{[n]_q!}\sinh(\frac{\hbar}{2})^n$ denote the coefficient of $E^n \otimes F^n$ in the infinite sum. Then

$$\begin{aligned} u = m^{\mathrm{op}}(\mathrm{id}\otimes\mathcal{S})(R) &= m^{\mathrm{op}}(\mathrm{id}\otimes\mathcal{S})\left[\left(\sum_{n=0}^{\infty}\frac{(\hbar/4)^n}{n!}H^n\otimes H^n\right)\left(\sum_m a_m E^m \otimes F^m\right)\right] \\ &= m^{\mathrm{op}}(\mathrm{id}\otimes\mathcal{S})\left[\sum_{n,m}\frac{(\hbar/4)^n}{n!}a_m H^n E^m \otimes H^n F^m\right] \\ &= m^{\mathrm{op}}\left[\sum_{n,m}\frac{(\hbar/4)^n}{n!}a_m H^n E^m \otimes \mathcal{S}(F)^m(-H)^n\right] \\ &= \sum_{n,m}\frac{(-\hbar/4)^n}{n!}\mathcal{S}(F)^m H^{2n} E^m \end{aligned}$$

and so

$$\tau = e^{-\hbar H/2} u = e^{-\hbar H/2}\sum_{n,m}\frac{(-\hbar/4)^n}{n!}a_m \mathcal{S}(F)^m H^{2n}E^m. \tag{13.2.3.4}$$

How does τ act on the irreducible $\mathcal{U}_\hbar\mathfrak{sl}(2)$ module V_λ with highest weight λ? It suffices to compute the eigenvalue of τ acting on the highest weight vector v_λ, which enjoys $Ev_\lambda = 0$ and $Hv_\lambda = \lambda v_\lambda$. But since $Ev_\lambda = 0$, τv_λ has only summands with $m = 0$:

$$\begin{aligned} \tau v_\lambda &= e^{-\hbar H/2}\sum_{n,m}\frac{(-\hbar/4)^n}{n!}\mathcal{S}(F)^m H^{2n}E^m v_\lambda \\ &= e^{-\hbar H/2}\sum_{n}\frac{(-\hbar/4)^n}{n!}H^{2n}v_\lambda \\ &= e^{-\hbar H/2}e^{-\hbar H^2/4}v_\lambda \\ &= e^{-\hbar\lambda(\lambda+2)/4}v_\lambda \end{aligned}$$

Note that $\lambda(\lambda+2)$ is nothing but the value of the quadratic Casimir c_2 for $\mathcal{U}\mathfrak{sl}(2)$ acting on V_λ. $\Diamond$

Our goal for the remainder of this section is to give a formula for R for general $\mathfrak{g}$. We will do so by using a version of the Weyl group action on $\mathcal{U}_\hbar\mathfrak{g}$. Let us emphasize again that the Weyl group does not act canonically on $\mathfrak{g}$, but it almost does: the normalizer of the torus acts, and that normalizer is the Weyl together with a copy of the torus. In any case, we quantized the almost-action of W on $\mathfrak{g}$ in Theorem 13.1.4.2. We will use a different quantization: we will deform W so that its elements are no longer grouplike.

Let us start with $\mathfrak{g} = \mathfrak{sl}(2)$. Then $W = \mathbb{Z}/2$ with unique nontrivial element w. The classical action is canonical up to some signs, and might as well be $H \mapsto -H$, $E \mapsto -F$, and $F \mapsto -E$. We will quantize this by quantizing the semidirect product algebra $\mathcal{U}\mathfrak{g} \rtimes B$. Let $(\mathcal{U}_\hbar\mathfrak{sl}(2))_B$ denote the algebra generated by E, F, H, w with relations as in Proposition 13.1.2.14 together with

$$wEw^{-1} = -q^{-1}F, \quad wFw^{-1} = -qE, \quad wHw^{-1} = -H.$$

13.2.3.5 Proposition *$(\mathcal{U}_\hbar\mathfrak{sl}(2))_\mathcal{B}$ is a Hopf algebra with comultiplication as in Proposition 13.1.2.14 together with $\Delta w = R^{-1}(w\otimes w)$, where R is the R-matrix from Proposition 13.1.2.14. Let τ denote the ribbon element given in equation (13.2.3.4). Then $(\mathcal{U}_\hbar\mathfrak{sl}(2))_\mathcal{B}$ has a quotient Hopf algebra in which $w^2 = \tau$.*

The quotient Hopf algebra $(\mathcal{U}_\hbar\mathfrak{sl}(2))_\mathcal{B}/(w^2 = \tau)$ is sometimes called the *quantum Weyl group* of $\mathfrak{sl}(2)$. It doesn't quantize the Weyl group, but rather the braid group action.

Proof We must check that Δ is an algebra homomorphism, i.e. that

$$\Delta(wxw^{-1}) = \Delta(w)\Delta(x)\Delta(w^{-1}) = R^{-1}(w\otimes w)\Delta(x)(w^{-1}\otimes w^{-1})R$$

for all $x \in \mathcal{U}_\hbar\mathfrak{sl}(2)$. But it is easy to see that $(w\otimes w)\Delta(x)(w^{-1}\otimes w^{-1}) = \Delta^{\mathrm{op}}(wxw^{-1})$, and so the result is equivalent to the defining property of the R-matrix. Coassociativity of Δ is equivalent to the Yang–Baxter equation. Compatibility with $w^2 = \tau$ is equivalent to the defining property of τ: $\Delta(\tau) = R^{-1}R_{12}^{-1}\tau\otimes\tau$. $\square$

13.2.3.6 Lemma *1. $w \in (\mathcal{U}_\hbar\mathfrak{sl}(2))_\mathcal{B}$ commutes with $e^{-\hbar H^2/8}$.*

2. Let $\tilde{w} \overset{\text{def}}{=} e^{\hbar H^2/8}w$. Then $waw^{-1} = T(a)$ for all $a \in \mathcal{U}_\hbar\mathfrak{sl}(2)$, where T is Lusztig's braid group action from Theorem 13.1.4.2.

Proof 1. is obvious. For 2., note that $[H^2, E] = 4E(H+1)$, and so $e^{\hbar H^2/8}Ee^{\hbar H^2/8} = Ee^{\hbar(H+1)/2}$. Thus $\tilde{w}E\tilde{w}^{-1} = wEKqw^{-1} = -q^{-1}FK^{-1}q = T(E)$. The formula for $\tilde{w}F\tilde{w}^{-1}$ is analogous. $\square$

We now replace $\mathfrak{sl}(2)$ by $\mathfrak{g}$. Define $(\mathcal{U}_\hbar\mathfrak{g})_\mathcal{B} = \mathcal{U}_\hbar\mathfrak{g} \rtimes \mathcal{B}(\mathfrak{g})$ to be the algebra generated by $\mathcal{U}_\hbar\mathfrak{g}$ and symbols $\tilde{w}_1, \dots, \tilde{w}_r$ such that the $\tilde{w}_i$s satisfy the Artin braid relations from Definition 13.1.4.1 and $\tilde{w}_i x \tilde{w}_i^{-1} = T_i(x)$, where T_i is Lusztig's braid group action from Theorem 13.1.4.2. Set $w_i \overset{\text{def}}{=} e^{\hbar d_i H_i^2/8}\tilde{w}_i$. For each i, the ith $\mathfrak{sl}(2)$-triple $\{E_i, F_i, H_i\}$ generates a copy of $\mathcal{U}_{d_i\hbar}\mathfrak{sl}(2) \subseteq \mathcal{U}_\hbar\mathfrak{g}$.

13.2.3.7 Proposition *We can make $(\mathcal{U}_\hbar\mathfrak{g})_{\mathcal{B}}$ into a Hopf algebra such that for each i, the embedding $(\mathcal{U}_{d_i\hbar}\mathfrak{sl}(2))_{\mathcal{B}} \subset (\mathcal{U}_\hbar\mathfrak{g})_{\mathcal{B}}$ generated by $\{E_i, F_i, H_i, w_i\}$ is a Hopf embedding. Its comultiplication restricts to the usual one on $\mathcal{U}_\hbar\mathfrak{g}$, and enjoys $\Delta w_i = R(i)^{-1}(w_i \otimes w_i)$, where $R(i)$ is the R-matrix for $\mathcal{U}_{d_i\hbar}\mathfrak{sl}(2)$.* □

Let $w_0 \in W$ denote the longest word, choose a reduced expression $w_0 = s_{i_1} \cdots s_{i_N}$, and look at the corresponding lift $\dot{w}_0 = w_{i_1} \cdots w_{i_N} \in (\mathcal{U}_\hbar\mathfrak{g})_{\mathcal{B}}$. If we are lucky, maybe we could hope that

$$\Delta \dot{w}_0 \overset{?}{=} R^{-1}(\dot{w}_0 \otimes \dot{w}_0) \tag{13.2.3.8}$$

If this is true, then we would have a wonderful product formula:

$$\Delta(w_{i_1}) \cdots \Delta(w_{i_N}) \overset{?}{=} R^{-1}(w_{i_1} \cdots w_{i_N} \otimes w_{i_1} \cdots w_{i_N}) \tag{13.2.3.9}$$

But the left-hand side of equation (13.2.3.9) is

$$\Delta(w_{i_1}) \cdots \Delta(w_{i_N}) = R(i_1)^{-1}(w_{i_1} \otimes w_{i_1}) \cdots R(i_N)^{-1}(w_{i_N} \otimes w_{i_N}).$$

Move the $R(i)$s past the w_is and cancel. Then we would have a formula for R as a product of $\mathfrak{sl}(2)$-R-matrices, twisted by the W-action:

$$R^{-1} \overset{?}{=} R(\alpha_{i_1})^{-1}\, R(s_{i_1}(\alpha_{i_2}))^{-1} \cdots R(s_{i_1}s_{i_2}\cdots s_{i_{N-1}}(\alpha_{i_N}))^{-1} \overset{\text{def}}{=} \overrightarrow{\prod_{\alpha \in \Delta_+}} R(\alpha)^{-1} \tag{13.2.3.10}$$

By $\overrightarrow{\prod}$ we mean that the product is ordered using the parameterization Theorem 13.1.4.5.

So equation (13.2.3.8) would be wonderful, but it turns out to be false. However, it is almost true: it holds if we take out the part of R that goes with the Cartan. Instead, define

$$\tilde{R}(i) \overset{\text{def}}{=} \sum_{n=0}^{\infty} \frac{\sinh(d_i\hbar/2)^n}{[n]_{q^{d_i}}!} E_i^n \otimes F_i^n$$

so that $R(i) = e^{\hbar d_i H_i \otimes H_i/4}\tilde{R}(i)$.

13.2.3.11 Lemma $\Delta(\tilde{w}_i) = \tilde{R}(i)(\tilde{w}_i \otimes \tilde{w}_i)$. □

The $\tilde{w}$s are better anyway, because they are the ones that satisfy the Artin braid relations. And it is a general fact that, if you take any $w \in \mathcal{W}$ and choose any reduced word for it, then the corresponding element of the Artin braid group didn't depend on the choice of word. So in fact $\tilde{w}_0 \overset{\text{def}}{=} \tilde{w}_{i_1} \cdots \tilde{w}_{i_N}$ is well-defined without a choice of reduced expression.

13.2.3.12 Theorem (Multiplicative formula for the universal R-matrix) *Let $\tilde{R} \overset{\text{def}}{=} \exp\left(-\frac{\hbar}{2}\sum_{ij} b_{ij} H_i \otimes H_j\right) R$ denote the nilpotent part of R. Then we do have $\Delta(\tilde{w}_0) = \tilde{R}(\tilde{w}_0 \otimes \tilde{w}_0)$. In particular,*

$$R = \exp\left(\frac{h}{4}\sum_{ij} b_{ij} H_i \otimes H_j\right) \overrightarrow{\prod_{\alpha>0}} \tilde{R}(\alpha).$$ □

This is due to [KR90]. Of course, $\tilde{R}(\alpha)$ is nothing but the nilpotent part of the R-matrix for the αth $\mathcal{U}_{d_\alpha \hbar}\mathfrak{sl}(2) \hookrightarrow \mathcal{U}_\hbar \mathfrak{g}$, defined using Theorems 13.1.4.2 and 13.1.4.5 together with the fixed choice of reduced word for w_0 (which also determines the order of the product).

13.2.4 Quantum Schur–Weyl duality

Consider the action of $\mathrm{SL}(n)$ on $\mathbb{C}^n$. Then $\mathrm{SL}(n)$ acts diagonally on $(\mathbb{C}^n)^{\otimes N}$. The symmetric group S_N also acts on $(\mathbb{C}^n)^{\otimes N}$, commuting with the $\mathrm{SL}(n)$-action.

13.2.4.1 Theorem (Schur–Weyl duality)
The actions of $\mathrm{SL}(n)$ and S_N on $(\mathbb{C}^n)^{\otimes N}$ centralize each other in the following sense: the only automorphisms of $(\mathbb{C}^n)^{\otimes N}$ that commute with the S_N-action are those in $\mathrm{GL}(n)$, i.e. $\mathrm{SL}(n)$ and rescalings; the only automorphisms that commute with the $\mathrm{GL}(n)$ action are S_N and rescalings. □

13.2.4.2 Corollary *$(\mathbb{C}^n)^{\otimes N}$ decomposes as $\bigoplus_{\lambda \dashv N, |\lambda| \leq n} V_\lambda^{\mathrm{GL}(n)} \otimes W_\lambda^{S_N}$, where the sum Young diagrams with N boxes and at most n rows, i.e. partitions $N = \lambda_1 + \cdots + \lambda_k$ with $k \leq n$, and where $V_\lambda^{\mathrm{GL}(n)}$ and $W_\lambda^{S_N}$ are the irreps of $\mathrm{GL}(n)$ and S_N indexed by λ.* □

To define the S_N action on $(\mathbb{C}^n)^{\otimes N}$ requires the symmetric monoidal structure on $\mathrm{R{\scriptsize EP}}(\mathrm{GL}(n))$. If we quantize, we lose the symmetric monoidal structure, replacing it with a braiding. Let $V = \mathbb{C}^n$ denote the defining representation of $\mathcal{U}_q\mathfrak{sl}(n)$. Since $\mathrm{R{\scriptsize EP}}(\mathcal{U}_q\mathfrak{sl}(n))$ is braided monoidal, there is an action by the braid group B_N on $V^{\otimes N}$ commuting with the $\mathcal{U}_q\mathfrak{sl}(n)$-action. Recall the *Hecke–Iwahori algebra* from Definition 12.3.2.1:

$$\mathcal{H}_N(q) \stackrel{\text{def}}{=} \langle s_i,\, i = 1, \dots, N-1 \text{ s.t. } (s_i - q)(s_i + 1) = 0,\, s_i s_{i\pm1} s_i = s_{i\pm1} s_i s_{i\pm1},\, s_i s_j = s_j s_i, |i-j| > 1\rangle$$

It is a quotient of the group algebra of the braid group, and at $q = 1$ it is the group algebra of the symmetric group.

13.2.4.3 Theorem (Quantum Schur–Weyl duality)
The braid group action on $V^{\otimes N} \in \mathrm{R{\scriptsize EP}}(\mathcal{U}_q\mathfrak{sl}(n))$ factors through $\mathcal{H}_N(q)$. Moreover, the actions of $\mathcal{H}_N(q)$ and $\mathcal{U}_q(\mathfrak{sl}_n)$ centralize each other. □

That the braid group action factors through $\mathcal{H}_N(q)$ is an $\mathfrak{sl}(2)$ calculation carried out in Section 12.3.2. Just as in the classical case, when $N > n$ the action there are $\mathcal{H}_N(q)$-modules missed in the decomposition of $V^{\otimes N}$, meaning that the action of $\mathcal{H}_N(q)$ on $V^{\otimes N}$ has kernel. We studied already the most important example — the Temperley–Lieb algebra — in Section 12.3.2.

13.3 Kashiwara's crystal bases

In this last section, we build a canonical basis for every finite-dimesnional $\mathcal{U}_q\mathfrak{g}$-module. Suppose you didn't know about quantum groups. Representations of $\mathfrak{sl}(2)$ have somewhat-canonical bases given by equation (5.2.0.6). It's not perfect, because you have to choose some convention about normalizations, but it's pretty good. But for other $\mathfrak{g}$, it doesn't work. You could try to construct a basis for the classical module $L(\lambda)$ by starting with the highest weight vector and acting on it by F_αs, but $F_1F_2v_\lambda$ and $F_2F_1v_\lambda$ might have some nontrivial linear dependency. What we will do is to build a related module, with operators $\tilde{F}_\alpha, \tilde{E}_\alpha$, and the only way $\tilde{F}_1\tilde{F}_2v_\lambda$ and $\tilde{F}_2\tilde{F}_1v_\lambda$ can be linearly dependent is if they are either equal or one of them vanishes. Moreover, $\tilde{F}_\alpha$ and $\tilde{E}_\alpha$ will be inverses when they are nonzero, so you won't have to think about those pesky numbers in equation (5.2.0.6).

We restrict our attention to the case when q is transcendental over $\mathbb{Q}$: we will work over $\mathbb{K} \overset{\text{def}}{=} \mathbb{Q}(q)$. Inside $\mathbb{K}$ is a ring A which consists of all fractions $\frac{f(q)}{g(q)}$ where $f, g \in \mathbb{Q}[q]$ are polynomials and $g(0) \neq 0$. A is a local ring with unique maximal ideal (q) and $A/qA \cong \mathbb{Q}$, and $\mathbb{K}$ is the field of fractions of A.

Let M be a finite-dimensional $\mathcal{U} \overset{\text{def}}{=} \mathcal{U}_q\mathfrak{g}$ module, which we suppose is in the category $\textsc{Rep}_+(\mathcal{U}_q\mathfrak{g})$, meaning that the Ks all act with eigenvalues in $q^{\mathbb{Z}}$ (and not $-q^{\mathbb{Z}}$). For each $\alpha \in \Gamma$, consider the subalgebra $\mathcal{U}^\alpha = \mathcal{U}_{q_\alpha}\mathfrak{sl}(2) \subset \mathcal{U}_q\mathfrak{g}$. We completely understand M over $\mathcal{U}^\alpha$: it decomposes as $M = \bigoplus L(n_i)$, where each $L(n_i)$ is a chain of length n_i.

We will need the divided powers

$$F_\alpha^{[j]} \overset{\text{def}}{=} \frac{F_\alpha^j}{[j]_{q_\alpha}!}.$$

Continuing to fix $\alpha \in \Gamma$, it is easy to see that each $x \in M$ can be written uniquely as

$$x = \sum F_\alpha^{[j]} x_j$$

where $E_\alpha x_j = 0$ for each j. This decomposition does not depend on a basis for M. The *Kashiwara operators* are:

$$\tilde{F}_\alpha x \overset{\text{def}}{=} \sum F_\alpha^{(j+1)} x_j,$$
$$\tilde{E}_\alpha x \overset{\text{def}}{=} \sum F_\alpha^{(j-1)} x_j.$$

Then

$$\tilde{F}_\alpha \tilde{E}_\alpha x = x - x_0,$$

and, if x has weight μ,

$$\tilde{E}_\alpha \tilde{F}_\alpha x = x - F_\alpha^{(r)} x_r,$$

where $r = (\mu, \alpha^\vee)$ where $\alpha^\vee \overset{\text{def}}{=} 2\alpha/(\alpha,\alpha)$.

We will now define a lattice over A that will be invariant under these Kashiwara operators.

13.3.0.1 Definition *An A-submodule $\tilde{M} \subset M$ is an* admissible lattice *if:*

1. $\tilde{M} \otimes_A \mathbb{K} = M$
2. $\tilde{M} = \bigoplus \tilde{M}_\lambda$, *where* $\tilde{M}_\lambda \overset{\text{def}}{=} \tilde{M} \cap M_\lambda$
3. $\tilde{M}$ *is invariant with respect to* $\tilde{E}_\alpha, \tilde{F}_\alpha$.

The Kashiwara operators are natural in the sense that if $\phi : M \to N$ is a homomorphism of $\mathcal{U}$-modules, then $\tilde{F}_\alpha, \tilde{E}_\alpha$ commute with ϕ. It follows that $M_1 \oplus M_2 \subseteq \tilde{M}_1 \oplus \tilde{M}_2$ is admissible iff each $\tilde{M}_i$ is admissible. Also obvious: given $\phi : M \xrightarrow{\sim} N$, $\phi(\tilde{M})$ is admissible iff $\tilde{M}$ is admissible. But consider the map $\times q : M \to M$. It is a $\mathcal{U}$-isomorphism, so if $\tilde{M}$ is admissible, then so is $q\tilde{M}$. It follows that $\tilde{E}_\alpha, \tilde{F}_\alpha$ are well-defined on $\tilde{M}/q\tilde{M}$.

13.3.0.2 Definition *A* crystal base *of M consists of a pair $(\tilde{M}, B)$ where $\tilde{M}$ is an admissible lattice and B is a basis for $\tilde{M}/q\tilde{M}$ over $A/qA = \mathbb{Q}$. We require that:*

1. *$B = \bigcup_{\lambda \in P(M)} B_\lambda$, where $B_\lambda \overset{\text{def}}{=} B \cap (\tilde{M}_\lambda / q\tilde{M}_\lambda)$ and $P(M)$ are the weights of M.*
2. *$\tilde{F}_\alpha B \subseteq B \cup \{0\}$ and $\tilde{E}_\alpha B \subseteq B \cup \{0\}$ for all $\alpha \in \Gamma$.*
3. *If $b_1, b_2 \in B$, then $b_2 = \tilde{E}_\alpha b_1$ iff $\tilde{F}_\alpha b_2 = b_1$.*

It is worth emphasizing that being a crystal base is something you can test by restricting to all the $\mathcal{U}^\alpha$s inside $\mathcal{U}$.

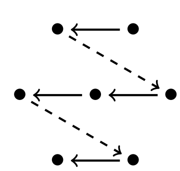
Suppose we have a crystal base $(\tilde{M}, B)$. We can describe the actions of $\tilde{E}, \tilde{F}$ on the base by drawing a *crystal graph*. The vertices of this graph are the basis element B. The edges are directed and colored: there are Γ many colors, and you draw an edge from b_1 to b_2 with color α exactly when $\tilde{F}_\alpha b_1 = b_2$.

13.3.0.3 Example For $\mathfrak{g} = \mathfrak{sl}(2)$, there is only one color. Irreducible graphs look like chains $\bullet \to \bullet \to \cdots \to \bullet$. $\Diamond$

13.3.0.4 Theorem (Construction of crystal bases)
Choose a highest weight vector $v_\lambda \in L(\lambda)$. Collect all nonzero vectors of the form $\tilde{F}_{\alpha_1} \cdots \tilde{F}_{\alpha_r} v_\lambda$ and let $\tilde{L}(\lambda)$ be their span. Let $\bar{v}_\lambda \in \tilde{L}(\lambda)/q\tilde{L}(\lambda)$ denote the image of v_λ, and set $B(\lambda) \overset{\text{def}}{=} \{\tilde{F}_{\alpha_1} \cdots \tilde{F}_{\alpha_r} \bar{v}_\lambda$ s.t. $\tilde{F}_{\alpha_1} \cdots \tilde{F}_{\alpha_r} \bar{v}_\lambda \neq 0\}$. Then $(\tilde{L}(\lambda), B(\lambda))$ is a crystal base for $L(\lambda)$. Moreover, it is unique up to rescaling.

Our strategy to prove Theorem 13.3.0.4 is as follows. Let us say that λ is *nice* if the recipe given in Theorem 13.3.0.4 produces a crystal base. We will prove by hand that small weights are nice. We will then prove that crystal bases play well with tensor products and direct sum decompositions. Theorem 13.3.0.4 is essentially automatic for $\mathfrak{g} = \mathfrak{sl}(2)$.

13.3.0.5 Example The 7-dimensional representation of G_2 looks like:



$\Diamond$

13.3.0.6 Example (Minuscule representation) The classical definition of *minuscule representation* is an irrep all of whose weights are a single Weyl orbit: $P(\lambda) \overset{\text{def}}{=} P(L(\lambda)) = W\lambda$. For example, for $\mathfrak{sl}(n)$, every fundamental representation is minuscule. The number of minuscule representations is exactly the number of elements in the quotient P/Q: for each class in P/Q, the minimal dominant weight in that class is minuscule. We claim that minuscule weights are nice.

Suppose λ is minuscule. Let $\mu \in P(\lambda)$ and $\alpha \in \Gamma$. Then $(\mu, \alpha^\vee) \in \{0, 1, -1\}$. Each weight space $L(\lambda)_\mu$ is one-dimensional, so it is easy to choose a basis, which is unique up to rescaling if we require that W act well. We construct the basis $\{x_\mu\}$ by setting $x_\lambda = v_\lambda$ and declaring:

$$E_\alpha x_\mu = \begin{cases} x_{\mu+\alpha}, & (\mu, \alpha^\vee) = -1 \\ 0, & (\mu, \alpha^\vee) = 0, 1 \end{cases}$$

$$F_\alpha x_\mu = \begin{cases} x_{\mu-\alpha}, & (\mu, \alpha^\vee) = 1 \\ 0, & (\mu, \alpha^\vee) = 0, -1 \end{cases}$$

Furthermore, $\tilde{E}_\alpha = E_\alpha$ and $\tilde{F}_\alpha = F_\alpha$. The A-span of the x_μs is an admissible lattice, and the $\bar{x}_\mu$s are a crystal base.

For example, for the standard representation of $\mathfrak{sl}(3)$, the picture is $\bullet \overset{\alpha_1}{\to} \bullet \overset{\alpha_2}{\to} \bullet$. $\Diamond$

Suppose that we show that crystal bases play well with tensor products and passing to direct summands. Then we will have proven Theorem 13.3.0.4 for $\mathfrak{g} = \mathfrak{sl}(n)$, since the minuscule representations generate all representations. This will not suffice, however, for G_2, F_4, E_8, for which the only minuscule representations are the trivial ones. Let us give ourselves one more representation:

13.3.0.7 Example (Dominant short root) Suppose that λ is the dominant short root. Then $P(L(\lambda)) = W\lambda \cup \{0\}$. If all the roots are the same length, $L(\lambda)$ is the adjoint representation. For B_n, $L(\lambda)$ is the standard $(2n+1)$-dimensional representation, whereas for C_n it is the exterior square of the $(2n)$-dimensional standard representation (minus a copy of the trivial representation).

We define Δ_{sh} to be the set of all short roots, and Γ_{sh} to be the set of simple short roots. We will construct our basis to be:

$$\{x_\beta \text{ s.t. } \beta \in \Delta_{\text{sh}}\} \cup \{h_\alpha \text{ s.t. } \alpha \in \Gamma_{\text{sh}}\} \tag{13.3.0.8}$$

For $\mathfrak{g} = \mathfrak{sl}(n)$ this is obviously a basis for the adjoint representation. In general, we have the relations:

$$E_\alpha x_\alpha = 0, \qquad F_\alpha x_\alpha = h_\alpha, \qquad F_\alpha^{(2)} x_\alpha = x_{-\alpha}, \qquad F_\alpha^{(3)} x_\alpha = 0.$$

If $\beta \neq \pm\alpha$, then $(\beta, \alpha^\vee) = 0, \pm 1$, and so $E_\alpha x_\beta, F_\alpha x_\beta$ are the same as in the minuscule case:

$$E_\alpha \left(h_\beta + \frac{(\beta, \alpha^\vee)}{[2]_{q_\alpha}} h_\alpha \right) = 0, \qquad F_\alpha \left(h_\beta + \frac{(\beta, \alpha^\vee)}{[2]_{q_\alpha}} h_\alpha \right) = 0.$$

To see the second equation, start with $h_\alpha = F_\alpha x_\alpha$ and work with it:

$$E_\alpha h_\alpha = E_\alpha F_\alpha x_\alpha = \frac{K_\alpha - K_\alpha^{-1}}{q_\alpha - q_\alpha^{-1}} x_\alpha = (q_\alpha + q_\alpha^{-1})^{-1} x_\alpha$$

because $K_\alpha x_\alpha = q_\alpha^2 x_\alpha$, and

$$E_\alpha h_\beta = E_\alpha F_\beta x_\beta = F_\beta E_\alpha x_\beta.$$

Since $E_\alpha x_\beta = 0$ if $(\alpha, \beta) = 0$ or $x_{\beta+\alpha}$ otherwise, $F_\beta E_\alpha x_\beta = x_\alpha$ or 0.

By a little calculation, $\frac{(\beta,\alpha^\vee)}{[2]_\alpha} = \frac{(\beta,\alpha^\vee)}{q_\alpha^2+1} q_\alpha \in qA$. Thus:

$$0 = \tilde{E}_\alpha \left(\bar{h}_\beta + \frac{(\beta, \alpha^\vee)}{[2]_{q_\alpha}} \bar{h}_\alpha \right) = \tilde{E}_\alpha \bar{h}_\beta$$

and so the set from equation (13.3.0.8) span an admissible lattice $\tilde{M}$ and their images in $\tilde{M}/q\tilde{M}$ are a crystal base. ♢

13.3.0.9 Lemma *Let $\tilde{M}$ be an admissible lattice and $x \in M$ decomposed as $x = \sum F_\alpha^{[j]} x_j$. If $x \in \tilde{M}$, then $x_j \in \tilde{M}$ for all j. Moreover, if $\tilde{E}_\alpha x \in q\tilde{M}$, then $x_j \in q\tilde{M}$ for all $j > 0$.*

Proof $x - \tilde{F}_\alpha \tilde{E}_\alpha x = x_0 \in \tilde{M}$. Now apply a raising operator: $\tilde{E}_\alpha x \in \tilde{M}$, and $\tilde{E}_\alpha x = \sum F_\alpha^{(j-1)} x_j$, so that $(\tilde{E}_\alpha x)_0 = x_1$. Thus $x_1 \in \tilde{M}$. Rinse and repeat to get the first statement.

For the second statement, use that $q\tilde{M}$ is admissible. By the first statement, if $E_\alpha x \in q\tilde{M}$, then $(\tilde{E}_\alpha x)_j = x_{j+1} \in q\tilde{M}$. □

We now wish to describe the highest weight vectors in the quotient $\tilde{M}/q\tilde{M}$. Suppose that $S \subseteq \tilde{M}/q\tilde{M}$, and we define $\mathrm{HW}(S) \stackrel{\text{def}}{=} \{x \in S \text{ s.t. } \tilde{E}_\alpha x = 0 \; \forall \alpha \in \Gamma\}$.

13.3.0.10 Lemma *Let $(\tilde{M}, B)$ be a crystal base. Then:*

1. *Any $b \in B$ can be written as*
$$b = \tilde{F}_{\alpha_1} \cdots \tilde{F}_{\alpha_r} b'$$
for some $b' \in \mathrm{HW}(B)$.

2. *For each $\lambda \in P(M)$, the corresponding weight space $\mathrm{HW}(\tilde{M}_\lambda / q\tilde{M}_\lambda)$ is generated as a vector space by $\mathrm{HW}(B_\lambda)$. Note that $B = \bigsqcup B_\lambda$ by definition.*

3. *If $\mathrm{HW}(B_\lambda) \neq \emptyset$, then λ is dominant.*

Proof

1. Start with b and apply $\tilde{E}_\alpha$s until you can't anymore: $b' = \tilde{E}_{\alpha_r} \cdots \tilde{E}_{\alpha_1} b \in \mathrm{HW}(B)$. On B, $\tilde{E}_\alpha$ and $\tilde{F}_\alpha$ are sort of inverses — they are inverses except when they are 0.

2. Suppose that $x \in \mathrm{HW}(\tilde{M}_\lambda / q\tilde{M}_\lambda)$. Then $x = \sum c_b b$, where the sum ranges over $b \in B_\lambda$. Now apply $\tilde{E}_\alpha$. Some of the bs are killed, but those that are not killed remain distinct. So throw away the zero ones: $0 = \tilde{E}_\alpha x = \sum c_b \tilde{E}_\alpha b$, and the nonzero $\tilde{E}_\alpha b$s are linearly independent. But therefore $c_b = 0$ if $\tilde{E}_\alpha b \neq 0$.

3. Finally, the condition that λ is dominant is just a condition on $\mathfrak{sl}(2)$. What we have to do is check that $(\lambda, \alpha^\vee) \geq 0 \; \forall \alpha \in \Gamma$, but we know that this is true for $\mathfrak{sl}(2)$. □

Recall that λ is *nice* if $L(\lambda)$ has a crystal base constructed as in Theorem 13.3.0.4. Examples 13.3.0.3, 13.3.0.6, and 13.3.0.7 show that all weights for $\mathfrak{sl}(2)$ and all small weights for $\mathfrak{g}$ are nice.

13.3.0.11 Proposition *Suppose λ is nice and let $\tilde{L}(\lambda), B(\lambda)$ be as in Theorem 13.3.0.4. Then $\mathrm{HW}(B(\lambda)) = B(\lambda)_\lambda$. Moreover, if $(\tilde{L}, B)$ is some other crystal base for $L(\lambda)$, then there exists $a \in \mathbb{K}$ such that $\tilde{L} = a\tilde{L}(\lambda)$ and $B = aB(\lambda)$.*

Proof For the first sentence, if you start with the highest weight vector, you can get all vectors Lemma 13.3.0.10 part 2, you can go back. For the second sentence, we know that $B_\lambda = \{\overline{av_\lambda}\}$ and $av_\lambda \in \tilde{L}$, because we obviously have a unique element in $\tilde{L}$ of weight λ. Supposing λ is nice, apply the Kashiwara operators to either v_λ or av_λ, and the result follows. □

How could Theorem 13.3.0.4 go wrong? We start with the highest vector, and then hope that we can get everywhere by applying Kashiwara operators, and if we can then the crystal base is unique up to endomorphism.

13.3.0.12 Proposition *Let M be a finite-dimensional representation of $\mathcal{U}_q\mathfrak{g}$ such that $M = \bigoplus L(\lambda_i)$ and all λ_i are nice. Suppose that $(\tilde{M}, B)$ is a crystal base for M. Then there exists $\phi_i : L(\lambda_i) \to M$ such that $\tilde{M} = \bigoplus \phi_i\big(\tilde{L}(\lambda_i)\big)$ and $B = \bigsqcup \phi_i\big(B(\lambda_i)\big)$.*

Proof We use induction on the number of components. Order the components so that $\lambda_1 \not< \lambda_2 \not< \dots$ in the standard order. In particular, λ_1 is weakly maximal. Then we know that B_{λ_1} contains only highest weight vectors: $\mathrm{HW}(B_{\lambda_1}) = B_{\lambda_1}$. Choose $b_1 \in B_{\lambda_1}$ and start generating. In particular, $b_1 = \overline{v_{\lambda_1}}$ for some highest weight vector with highest weight λ_1, and a choice of v_{λ_1} determines $\phi_1 : L(\lambda_1) \to M$. But $\{\tilde{F}_{\alpha_r} \cdots \tilde{F}_{\alpha_1} b_1\} \subseteq B$, and by Proposition 13.3.0.11 this is $\phi_1(B(\lambda_1))$. Consider the natural map $\tilde{M} \to \tilde{M}/q\tilde{M}$, and let $\tilde{M}'$ be the preimage of the sublattice generated by $B' = B \setminus \phi_1(B(\lambda_1))$. Then $(\tilde{M}', B')$ is a crystal base for M with $\phi_1(L(\lambda_1))$ subtracted off, and you can proceed by induction. □

Now we come to the interesting part, which is to study the behavior of crystal bases under tensor products. Our goal will be to show that if $(\tilde{M}_1, B_1)$ and $(\tilde{M}_2, B_2)$ are crystal bases for M_1, M_2, then $\big(\tilde{M}_1 \otimes \tilde{M}_2, B_1 \otimes B_2\big)$ is a crystal base for $M_1 \otimes M_2$. The way we've set it this is false. Instead, we need to modify the tensor product: we will write $M_1 \otimes' M_2$ for Kashiwara's tensor product, which is the usual one twisted by a certain antiautomorphism. Specifically, we define:

$$\Delta'(E_\alpha) = E_\alpha \otimes K_\alpha^{-1} + 1 \otimes E_\alpha \tag{13.3.0.13}$$
$$\Delta'(F_\alpha) = F_\alpha \otimes 1 + K_\alpha \otimes F_\alpha \tag{13.3.0.14}$$
$$\Delta'(K_\mu) = K_\mu \otimes K_\mu \tag{13.3.0.15}$$

and in fact we would like formulas for the divided powers

$$\Delta'(E_\alpha^{[r]}) = \sum_{i=0}^{r} q_\alpha^{-i(r-i)} E_\alpha^{[i]} \otimes E_\alpha^{[r-i]} K_\alpha^{-i} \tag{13.3.0.16}$$
$$\Delta'(F_\alpha^{[r]}) = \sum_{i=0}^{r} q_\alpha^{-i(r-i)} F_\alpha^{[i-r]} K_\alpha^{i} \otimes F_\alpha^{[i]} \tag{13.3.0.17}$$

Had we insisted on using the original tensor product, then we would have needed to work with lowest vectors rather than highest vectors. Equations (13.3.0.16) and (13.3.0.17) are easily checked by induction.

Given $b \in B_1$ or B_2, define $e_\alpha(b) \stackrel{\text{def}}{=} \max\{r \text{ s.t. } \tilde{F}_\alpha^r b \neq 0\}$, and $f_\alpha(b) = \max\{r \text{ s.t. } \tilde{E}_\alpha^r b \neq 0\}$.

13.3.0.18 Theorem (Tensor product formula for crystal bases)
Let $(\tilde{M}_1, B_1)$, $(\tilde{M}_2, B_2)$ be crystal bases. Then $(\tilde{M}_1 \otimes' \tilde{M}_2, B_1 \otimes B_2)$ is a crystal base for $M_1 \otimes' M_2$, where the tensor product is given by the actions:

$$\tilde{F}_\alpha(b_1 \otimes b_2) = \begin{cases} \tilde{F}_\alpha b_1 \otimes b_2, & f_\alpha(b_1) > e_\alpha(b_2) \\ b_1 \otimes \tilde{F}_\alpha b_2, & f_\alpha(b_1) \leq e_\alpha(b_2) \end{cases} \tag{13.3.0.19}$$

$$\tilde{E}_\alpha(b_1 \otimes b_2) = \begin{cases} \tilde{F}_\alpha b_1 \otimes b_2, & f_\alpha(b_1) \geq e_\alpha(b_2) \\ b_1 \otimes \tilde{F}_\alpha b_2, & f_\alpha(b_1) < e_\alpha(b_2) \end{cases} \tag{13.3.0.20}$$

13.3.0.21 Example Let $\mathfrak{g} = \mathfrak{sl}(3)$ and write V for the standard representation. The crystal base for V is terribly simple: $\bullet \longrightarrow \bullet \dashrightarrow \bullet$. For $V \otimes V$ we get:

$$\begin{tikzcd} \bullet \arrow[d] & \bullet \arrow[r, dashed] & \bullet \arrow[d] \\ \bullet \arrow[r] \arrow[d, dashed] & \bullet \arrow[d, dashed] & \bullet \\ \bullet \arrow[r] & \bullet \arrow[r, dashed] & \bullet \end{tikzcd}$$

The picture for V^* is $\bullet \dashrightarrow \bullet \longrightarrow \bullet$, and so for $V \otimes V^*$ you get:

$$\begin{tikzcd} \bullet \arrow[r, dashed] \arrow[d] & \bullet \arrow[d] & \bullet \\ \bullet \arrow[d, dashed] & \bullet \arrow[r] & \bullet \arrow[d, dashed] \\ \bullet \arrow[r, dashed] & \bullet \arrow[r] & \bullet \end{tikzcd}$$

These show hands-on the decompositions $V^{\otimes 2} = \mathcal{S}^2(V) \oplus \bigwedge^2(V)$ and $V \otimes V^* = \mathfrak{sl}(3) \oplus 1$. $\Diamond$

Proof (of Theorem 13.3.0.18) It is sufficient to prove it in $\mathfrak{sl}(2)$, since the properties of being a crystal base are tested by restricting to $\mathcal{U}_{q_\alpha}\mathfrak{sl}(2) \hookrightarrow \mathcal{U}_q\mathfrak{g}$. We will begin with the case when $M_1 = L(1)$ and $M_2 = L(m)$, and take $x \in L(1)$ and $y \in L(m)$ highest vectors. Then we have $M = M_1 \otimes' M_2 = L(m+1) \oplus L(m-1)$, where the highest vectors are $z_0 = x \otimes y$ and $z_1 = x \otimes Fy - q^m[m]Fx \otimes y$ — the latter is a simple calculation. Note that $q^m[m] \in qA$. Using equations (13.3.0.13) to (13.3.0.17) we see that:

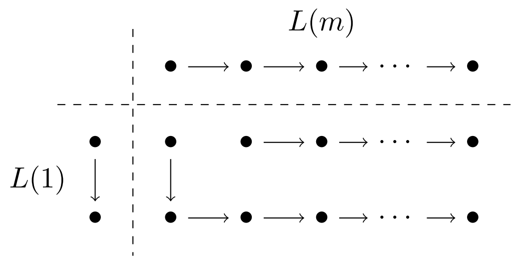
$$F^{[r]}z_0 = \begin{cases} q^r x \otimes F^{[r]}y + Fx \otimes F^{[r-1]}y, & 0 < r < m+1 \\ Fx \otimes F^{[m]}y, & r = m+1 \end{cases}$$

$$F^{[r]}z_1 = [r+1]q^r x \otimes F^{[r+1]}y - q^m[m-r]Fx \otimes F^{[r]}y$$

All the coefficients belong to A, and $[r+1]q^r = 1 \mod q$, and the terms with q^r vanish in the quotient by (q). It follows that $F^{(i)}x \otimes F^{(j)}y$ generate an admissible lattice, and moreover:

$$\tilde{F}\big(\bar{x} \otimes \tilde{F}^i \bar{y}\big) = \begin{cases} \tilde{F}\bar{x} \otimes \bar{y}, & i = 0 \\ \bar{x} \otimes \tilde{F}^{i+1}\bar{y}, & i > 0 \end{cases}$$
$$\tilde{F}\big(\tilde{F}\bar{x} \otimes \tilde{F}^i \bar{y}\big) = \tilde{F}\bar{x} \otimes \tilde{F}^{i+1}\bar{y}$$

In terms of pictures:



This proves the result for $M_1 = L(1)$ and $M_2 = L(m)$.

Now we consider $M_1 \otimes (M_2 \otimes M_3)$. Then we have:

$$\tilde{F}(b_1 \otimes b_2 \otimes b_3) = \begin{cases} \tilde{F}b_1 \otimes b_2 \otimes b_3, & f(b_1) > e(b_2 \otimes b_3), \\ b_1 \otimes \tilde{F}b_2 \otimes b_3, & f(b_1) \le e(b_2 \otimes b_3) \text{ and } f(b_2) > e(b_3), \\ b_1 \otimes b_2 \otimes \tilde{F}b_3, & f(b_1) \le e(b_2 \otimes b_3) \text{ and } f(b_2) \le e(b_3). \end{cases}$$

We also to consider $(M_1 \otimes M_2) \otimes M_3$. We again have three cases: $f(b_1 \otimes b_2) > e(b_3)$ and $f(b_1) > e(b_2)$; $f(b_1 \otimes b_2) > e(b_3)$ and $f(b_2) \le e(b_3)$; $f(b_1 \otimes b_2) \le e(b_3)$. These are readily calculated:

$$f(b_1 \otimes b_2) = \begin{cases} f(b_1) - e(b_2) + f(b_2), & f(b_1) > e(b_2), \\ f(b_2), & f(b_1) \le e(b_2), \end{cases}$$
$$e(b_1 \otimes b_2) = \begin{cases} e(b_1) - f(b_2) + e(b_2), & f(b_1) \ge e(b_2), \\ e(b_2), & f(b_1) < e(b_2). \end{cases}$$

The result then follows by induction. □

Let $\mathcal{U} = \mathcal{U}_q\mathfrak{g}$ and define $\sigma : \mathcal{U} \to \mathcal{U}^{\mathrm{op}}$ by $\sigma(E_\alpha) = q_\alpha F_\alpha K_\alpha^{-1}$, $\sigma(F_\alpha) = q_\alpha^{-1} K_\alpha E_\alpha$, and $\sigma(K_\mu) = K_\mu$. A bilinear form $(,) : M \times M \to \mathbb{K}$ is *σ-invariant* or *contragradient* if $(um, m') = (m, \sigma(u)m')$. Such forms have a number of properties:

1. $\ker(,)$ is an invariant subspace.
2. $(M_\mu, M_\nu) = 0$ if $\mu \ne \nu$.
3. If $M = \bigoplus M[\lambda_i]$ are the *isotypic components* $M[\lambda_i] = L(\lambda_i)^{\oplus r}$, then $(M[\lambda_i], M[\lambda_j]) = 0$ if $i \ne j$.

4. $L(\lambda)$ admits a unique-up-to-scalar σ-invariant form. Indeed, make the vector space M^* into a $\mathcal{U}$-module by declaring $\langle u\phi, m\rangle = \langle \phi, \sigma(u)m\rangle$ (meaning we use σ rather than $\mathcal{S}$). Then $L(\lambda)^*$ is irreducible and has the same character as $L(\lambda)$, and so there is a unique-up-to-scalar isomorphism $L(\lambda)^* \cong L(\lambda)$. In particular, after choosing a highest weight vector $v_\lambda \in L(\lambda)$, there is a unique form for which $(v_\lambda, v_\lambda) = 1$.

5. The comultiplication Δ' from equations (13.3.0.13) to (13.3.0.15) enjoys $\Delta' \circ \sigma = (\sigma \otimes \sigma) \circ \Delta'$. Thus if $(,)_1$ and $(,)_2$ are forms on M_1 and M_2, then you can construct $(,)$ on $M_1 \otimes M_2$ by declaring $(m_1 \otimes m_2, m_1' \otimes m_2') = (m_1, m_1')_1 \otimes (m_2, m_2')_2$.

6. Suppose $\tilde{M}$ is an admissible lattice and that $(,) : \tilde{M} \times \tilde{M} \to A$. Then:

$$(\tilde{E}_\alpha m, m') = (m, \tilde{F}_\alpha m') \mod qA$$

   So $(,)$ descends to $(,)_0 : (\tilde{M}/q\tilde{M})^{\times 2} \to \mathbb{Q}$.

13.3.0.22 Definition *A* polarization *of $(\tilde{M}, B)$ to be a σ-invariant form $(,)$ on M such that $(\tilde{M}, \tilde{M}) \subseteq A$, and such that B is orthonormal for $(,)_0$, i.e. $(b, b')_0 = \delta_{b,b'}$.*

13.3.0.23 Lemma *If $(\tilde{M}, B)$ admits a polarization, then*

1. $\tilde{M} = \{x \in M \text{ s.t. } (x, x) \in A\}$.

2. $\tilde{M} = \bigoplus \tilde{M} \cap M[\lambda_i]$, where the sum is over isotypic components.

Proof We will skip the proof of 1. For 2., note that if $x \in \tilde{M}$, we can write it as $\sum x_{\lambda_i}$, and so $(x_{\lambda_i}, x_{\lambda_i}) \in A$, and so $x_{\lambda_i} \in \tilde{M} \cap M[\lambda_i]$. □

Direct calculation — construct the form and check that it works — proves:

13.3.0.24 Proposition *Small representations admit polarizations. In particular, let $L(\lambda)$ be a small representation, i.e. one spanned by $\{x_\mu, h_\alpha\}$, where $\mu \in W \cdot \lambda$ and $\alpha \in \Gamma_{\mathrm{sh}}$ (you need the h_α only if $L(\lambda)$ is not minuscule). Define*

$$(x_\mu, x_\nu) \overset{\text{def}}{=} \delta_{\mu,\nu}, \qquad (h_\mu, x_\nu) \overset{\text{def}}{=} 0,$$

$$(h_\beta, h_\gamma) \overset{\text{def}}{=} \begin{cases} 1 + q_\beta^2, & \beta = \gamma \\ q_\beta, & (\beta, \gamma) \neq 0, \\ 0, & (\beta, \gamma) = 0. \end{cases}$$

And then $\{\bar{x}_\mu, \bar{h}_\alpha\} = B$ is an orthonormal base. □

What happens when you tensor? We use a classical formula:

13.3.0.25 Lemma *Let $L(\lambda_0)$ be small. Then $L(\lambda) \otimes L(\lambda_0)$ has the following components:*

1. $L(\lambda + \mu)$ with multiplicity 1 is $\mu \in W \cdot \lambda_0$ and $\lambda + \mu \in P^+$.

2. In the non-minuscule case: $L(\lambda)$ with multiplicity $\#\{\alpha \in \Pi_{\mathrm{sh}}$ s.t. $(\lambda, \alpha^\vee) > 0\}$.

Proof We use the Weyl character formula:

$$\operatorname{ch} L(\lambda) \cdot \operatorname{ch} L(\lambda_0) = \frac{1}{\mathcal{D}} \sum_{\substack{w \in W \\ \mu \in W\lambda_0}} \operatorname{sign}(w)\, e^{w(\lambda+\rho)+\mu} + \frac{1}{\mathcal{D}} \left|\Pi_{\mathrm{sh}}\right| \sum_{w \in W} \operatorname{sign}(w)\, e^{w(\lambda+\rho)} \tag{13.3.0.26}$$

In the minuscule case, if $w(\lambda+\rho)+\mu$ is not dominant, then $w(\lambda+\rho)+\mu$ lies on a wall and cancels.

In the non-minuscule case it could happen that $\lambda+\rho+\mu$ is not dominant, and so that means that $(\lambda+\rho+\alpha) = s_\alpha(\lambda+\rho)$. Then this guy in the first sum in equation (13.3.0.26) cancels with a guy in the second sum. So you get the cancelation, and that gives the stated multiplicities. □

Assume that $\big(\tilde{L}(\lambda), B(\lambda)\big)$ is a crystal base with polarization. If $L(\lambda_0)$ is small, then we already know that $L(\lambda_0) \otimes L(\lambda)$ admits a crystal base $\big(\tilde{L}(\lambda_0) \otimes \tilde{L}(\lambda), B(\lambda) \otimes B(\lambda_0)\big)$. Write $B = B(\lambda) \otimes B(\lambda_0)$ and $\tilde{M} = \tilde{L}(\lambda_0) \otimes \tilde{L}(\lambda)$. Then $\mathrm{HW}(B_\nu) =$ the multiplicity of $L(\nu)$ in $L(\lambda_0) \otimes L(\lambda)$.

Moreover, $\tilde{M} = \bigoplus M[\nu] \cap \tilde{M}$, where the sum is over isotypic components. Since B is orthonormal, it is also a sum of isotypic components: $B = \bigsqcup B[\nu]$. Then $\mathrm{HW}(B[\nu]) =$ multiplicity of $L(\nu)$ in $L(\lambda_0) \otimes L(\lambda)$. But then we must have $B[\nu]_\nu = \{b_1, \ldots, b_s\}$. Then we start applying $\tilde{F}_\alpha$s: we have $\tilde{F}_{\alpha_1} \ldots \tilde{F}_{\alpha_k} b_i$, and they are distinct because we can go back, and they generate everything because if not then there is another highest vector. And we know that every vector in the canonical base comes from some highest vector, but if it doesn't come form this one, then we must have another highest vector. And we know how many there are. So in particular $\{\tilde{F}_{\alpha_1} \ldots \tilde{F}_{\alpha_k} b_i\}$ generates a copy of $L(\nu)$ in $L(\lambda_0) \otimes L(\lambda)$, and so it gives us a canonical basis of $L(\nu)$ with polarization.

Proof (of Theorem 13.3.0.4) The statement holds for small weights by Examples 13.3.0.6 and 13.3.0.7. Small weights generate all weights under tensor products and passing to direct summands, so all representations admit crystal bases by Proposition 13.3.0.12 and Theorem 13.3.0.18. Why is it the nice one? The above discussion implies that we can in fact build polarized crystal bases. Thus if λ was nice, then it's still true for any ν inside the tensor product $L(\lambda) \otimes L(\lambda_0)$ for λ_0 small. So now we can do induction on the weights. Uniqueness is given in Proposition 13.3.0.11. □

Exercises

1. Suppose that A is an associative unital algebra. Show that every formal deformation of A is equivalent to a formal deformation with $\tilde{1} = 1$.

2. Let A be an associative algebra with center $\mathcal{Z}(A)$. Show that any formal deformation of A determines a Poisson structure on $\mathcal{Z}(A)$ which extends to an action of $\mathcal{Z}(A)$ on A by derivations.

3. Let A be a commutative algebra and $\tilde{m}, \hat{m}$ two formal deformations of A as associative algebras. Suppose that $\tilde{m}, \hat{m}$ are equivalent via an isomorphism $\phi = \mathrm{id} + \hbar\phi^{(1)} + \ldots$. Show that if $\phi^{(1)}$ is a derivation, then $\tilde{m}^{(1)} = \hat{m}^{(1)}$. Show that in general $\tilde{m}$ and $\hat{m}$ induce the same Poisson structure on A.

4. Suppose A is a cocommutative Hopf algebra and $A_\hbar$ a formal deformation in the sense of Hopf algebras. What structure on A is analogous to the Poisson structure in the case of formal deformations of commutative algebras? You should end up inventing the notion of *co-Poisson Hopf algebra*.

5. Suppose G and H are groups, and $G \curvearrowright H$ be automorphisms. Then there is a semidirect product $G \ltimes H$. Find Hopf algerba versions of $G \ltimes H$. There should be four versions: G and H can each be replaced by the group algebra $\mathbb{C}[G], \mathbb{C}[H]$ or my the algebra of functions $\mathscr{C}(G), \mathscr{C}(H)$.

6. Prove Proposition 13.1.2.11.

7. Prove Proposition 13.1.2.16. Specifically, set $\phi(H) = H$, $\phi(E) = E\, f(H)$, and $\phi(F) = g(H)\, F$, and find suitable functions f, g.

8. Compute the antipode $\mathcal{S}$ for $\mathcal{U}_\hbar\mathfrak{g}$. In particular, show that $\mathcal{S}^2(a) = e^{\hbar H_\rho/2} a e^{-\hbar H_\rho/2}$, the H_ρ is the Cartan element corresponding to $\rho = \frac{1}{2}\sum_{\alpha\in\Delta_+}\alpha$.

9. (a) Find a formula for $\Delta(E_i^{[n]})$, where $E \in \mathcal{U}_q\tilde{\mathfrak{b}}_+$ and $E_i^{[n]} = E_i^n/[n]_{q_i}!$ as in Section 13.1.3.
   (b) Verify Remark 13.1.3.8.
   (c) Rewrite Lusztig's braid group action on $\mathcal{U}_q\mathfrak{g}$ from Theorem 13.1.4.2 in terms of the adjoint action.

10. Prove Lemma 13.1.3.21. Hint: the only nontrivial case is $k = i$. Calculate $[F_i, E_i^n]$ by working in $\mathcal{U}_\hbar\mathfrak{sl}(2)$.

11. Show that the action of T_i from Theorem 13.1.4.2 preserves equation (13.1.3.16). Hint: you will need the Serre relation. Also calculate the action of T_i^2.

12. Check that the divided power form of $\mathcal{U}_q\mathfrak{sl}(2)$ from Remark 13.1.4.8 is closed under Δ.

13. Prove equation (13.2.2.10).

14. Show that the pairing $(,)$ from equation (13.2.2.11) satisfies $(\mathrm{ad}(u)a, b) = \big(a, \mathrm{ad}(\mathcal{S}(u))b\big)$. Hint: it suffices to check on generators.

15. Compute the counit and antipode for the Hopf algebra $(\mathcal{U}_\hbar\mathfrak{sl}(2))_W$ from Proposition 13.2.3.5.

16. Show that, if $\mathfrak{g}$ admits a nontrivial minuscule representation, then the sum of the minuscule representations is faithful G-representation, and so tensor-generates $\mathrm{R{\scriptsize EP}}(G)$.

Bibliography

[BDMGK01] Marcus Berg, Cécile DeWitt-Morette, Shangjr Gwo, and Eric Kramer. The Pin groups in physics: C, P and T. *Rev. Math. Phys.*, 13(8):953–1034, 2001.

[BG02] Ken A. Brown and Ken R. Goodearl. *Lectures on algebraic quantum groups*. Advanced Courses in Mathematics. CRM Barcelona. Birkhäuser Verlag, Basel, 2002.

[Bia] Bianchi classification. Wikipedia, the free encyclopedia. Wikimedia Foundation, Inc. http://en.wikipedia.org/wiki/Bianchi_classification.

[BK01] Bojko Bakalov and Alexander Kirillov, Jr. *Lectures on tensor categories and modular functors*, volume 21 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2001.

[BNLT03] Dror Bar-Natan, Thang T. Q. Le, and Dylan P. Thurston. Two applications of elementary knot theory to Lie algebras and Vassiliev invariants. *Geom. Topol.*, 7:1–31, 2003.

[Bor91] Armand Borel. *Linear algebraic groups*, volume 126 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[Bou05] Nicolas Bourbaki. *Lie groups and Lie algebras. Chapters 1–3, 4–6, and 7–9*. Elements of Mathematics (Berlin). Springer-Verlag, Berlin, 1989, 2002, 2005. Translated from the French 1968, 1975, and 1982 originals by Andrew Pressley.

[Car72] Roger W. Carter. *Simple groups of Lie type*. John Wiley & Sons, London-New York-Sydney, 1972. Pure and Applied Mathematics, Vol. 28.

[CE99] Henri Cartan and Samuel Eilenberg. *Homological algebra*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1999. With an appendix by David A. Buchsbaum, Reprint of the 1956 original.

[CP94] Vyjayanthi Chari and Andrew Pressley. *A guide to quantum groups*. Cambridge University Press, Cambridge, 1994.

[Dix96] Jacques Dixmier. *Enveloping algebras*, volume 11 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1996. Revised reprint of the 1977 translation.

[EL07] Sam Evens and Jiang-Hua Lu. Poisson geometry of the Grothendieck resolution of a complex semisimple group. *Mosc. Math. J.*, 7(4):613–642, 766, 2007.

[ES02] Pavel Etingof and Olivier Schiffmann. *Lectures on quantum groups.* Lectures in Mathematical Physics. International Press, Somerville, MA, second edition, 2002.

[FH91] William Fulton and Joe Harris. *Representation theory*, volume 129 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1991. A first course, Readings in Mathematics.

[Fuk86] D.B. Fuks. *Cohomology of Infinite-Dimensional Lie Algebras.* Monographs in Contemporary Mathematics. Springer, 1986.

[GM03] Sergei I. Gelfand and Yuri I. Manin. *Methods of homological algebra.* Springer Monographs in Mathematics. Springer-Verlag, Berlin, second edition, 2003.

[GOV97] V. V. Gorbatsevich, A. L. Onishchik, and E. B. Vinberg. *Foundations of Lie theory and Lie transformation groups.* Springer-Verlag, Berlin, 1997. Translated from the Russian by A. Kozlowski, Reprint of the 1993 translation [ıt Lie groups and Lie algebras. I, Encyclopaedia Math. Sci., 20, Springer, Berlin, 1993; MR1306737 (95f:22001)].

[GW04] K. R. Goodearl and R. B. Warfield, Jr. *An introduction to noncommutative Noetherian rings*, volume 61 of *London Mathematical Society Student Texts.* Cambridge University Press, Cambridge, second edition, 2004.

[HC49] Harish-Chandra. On representations of Lie algebras. *Ann. of Math. (2)*, 50:900–915, 1949.

[Hel01] Sigurdur Helgason. *Differential geometry, Lie groups, and symmetric spaces*, volume 34 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2001.

[Hoc66] G. Hochschild. An addition to Ado's theorem. *Proc. Amer. Math. Soc.*, 17:531–533, 1966.

[Hum08] James E. Humphreys. *Representations of semisimple Lie algebras in the BGG category $\mathscr{O}$*, volume 94 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2008.

[Iwa48] Kenkichi Iwasawa. On the representation of Lie algebras. *Jap. J. Math.*, 19:405–426, 1948.

[Jan03] Jens Carsten Jantzen. *Representations of algebraic groups*, volume 107 of *Mathematical Surveys and Monographs.* American Mathematical Society, Providence, RI, second edition, 2003.

[Kac69] V. G. Kac. Automorphisms of finite order of semisimple Lie algebras. *Funkcional. Anal. i Priložen.*, 3(3):94–96, 1969.

[Kac90] Victor G. Kac. *Infinite-dimensional Lie algebras*. Cambridge University Press, Cambridge, third edition, 1990.

[Kna02] Anthony W. Knapp. *Lie groups beyond an introduction*, volume 140 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, second edition, 2002.

[KR90] A. N. Kirillov and N. Reshetikhin. q-Weyl group and a multiplicative formula for universal R-matrices. *Comm. Math. Phys.*, 134(2):421–431, 1990.

[KS97] Anatoli Klimyk and Konrad Schmüdgen. *Quantum groups and their representations*. Texts and Monographs in Physics. Springer-Verlag, Berlin, 1997.

[Lam01] T. Y. Lam. *A first course in noncommutative rings*, volume 131 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 2001.

[Lan85] Serge Lang. $\mathrm{SL}_2(\mathbf{R})$, volume 105 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1985. Reprint of the 1975 edition.

[Lee09] Jeffrey M. Lee. *Manifolds and differential geometry*, volume 107 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2009.

[Lus93] George Lusztig. *Introduction to quantum groups*, volume 110 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, 1993.

[MR01] J. C. McConnell and J. C. Robson. *Noncommutative Noetherian rings*, volume 30 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, revised edition, 2001. With the cooperation of L. W. Small.

[Ser09] Vera Serganova. Math 274: Geometric invariant theory. Fall 2009. Edited notes by Anton Geraschenko available at `http://math.berkeley.edu/~anton/written/GIT/GIT.pdf`, 2009.

[Spr77] T. A. Springer. *Invariant theory*. Lecture Notes in Mathematics, Vol. 585. Springer-Verlag, Berlin, 1977.

[TL71] H. N. V. Temperley and E. H. Lieb. Relations between the "percolation" and "colouring" problem and other graph-theoretical problems associated with regular planar lattices: some exact results for the "percolation" problem. *Proc. Roy. Soc. London Ser. A*, 322(1549):251–280, 1971.

[Tri05] Todd Trimble. The super brauer group and super division algebras. `http://math.ucr.edu/home/baez/trimble/superdivision.html`, 2005.

[Var84] V. S. Varadarajan. *Lie groups, Lie algebras, and their representations*, volume 102 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1984. Reprint of the 1974 edition.

[Wei88] Alan Weinstein. Some remarks on dressing transformations. *J. Fac. Sci. Univ. Tokyo. Sect. 1A, Math.*, 35:163–167, 1988.

[Wei94] Charles A. Weibel. *An introduction to homological algebra*, volume 38 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1994.

[Wen85] Hans Gerd Wenzl. *REPRESENTATIONS OF HECKE ALGEBRAS AND SUBFACTORS*. ProQuest LLC, Ann Arbor, MI, 1985. Thesis (Ph.D.)–University of Pennsylvania.

Index