Lattices in semi-simple Lie groups

Rémi Boutonnet

This course aims to present various constructions and properties of lattices in Lie groups.

I thank Yves Benoist for allowing me to follow his lecture notes [**Ben08**]. Most of the time, I did nothing more than translate his notes into english. Sometimes I added minor details to proofs and I take full responsibility if mistakes or inconsistencies appear. No originality is claimed.

Contents

Part 1

Generalities

CHAPTER 1

Locally compact groups and lattices

1. The Haar measure

DEFINITION 1.1. A *topological group* is a group G endowed with a topology for which the map $(g,h) \in G \times G \to gh^{-1}$ is continuous. G is said to be *locally compact* if every point admits a compact neighborhood, or equivalently if the trivial element admits a compact neighborhood.

In the above terminology, the term compact includes the Hausdorff axiom, according to the French convention.

EXAMPLE 1.2. We will encounter many locally compact groups.

- All Lie groups are examples of locally compact groups.
- Any group can be made locally compact by considering its discrete topology. This is the topology we will usually use for countable groups.
- The additive group $(\mathbb{Q}_p, +)$ is locally compact (for the topology given by its ultrametric norm), and the subgroup $\mathbb{Z}_p \subset \mathbb{Q}_p$ is a compact neighborhood of 0.

DEFINITION 1.3. A *Haar measure* on a locally compact group is a (non-zero) Radon measure λ on G which is invariant under left translation, in the sense that for all $g \in G$, and all Borel set $A \subset G$, $\lambda(gA) = \lambda(A)$.

We recall that a Radon measure is a measure which is finite on compact sets and *regular*, meaning that for any Borel set A,

$$\lambda(A) = \sup\{\lambda(K) \mid K \subset A \text{ compact}\} = \inf\{\lambda(U) \mid U \text{ open set containing } A\}.$$

EXAMPLE 1.4. If $G = (\mathbb{R}, +)$, the Lebesgue measure is a Haar measure. The counting measure on a discrete group is a Haar measure.

The following observation follows from standard considerations on measurable functions.

LEMMA 1.5. *A Radon measure λ on a locally compact group is a Haar measure if and only if for every $f \in C_c(G)$ and every $g \in G$, we have $\int_G f(x)d\lambda(x) = \int_G f(gx)d\lambda(x)$.*

THEOREM 1.6. *If G is a locally compact group it always admits a Haar measure. Moreover, any two Haar measures on G are proportional.*

We will not prove the existence part in this theorem because, one, it is rather long to do and not much more instructive than the usual construction of the Lebesgue measure, and two, because for our examples a Haar measure can often be found by more concrete methods. For example, one can easily construct a left invariant volume form on a Lie group: just pick an n-form on the tangent space at the identity of the Lie group G (where $n = \dim(G)$) and propagate it using the left translations L_g, $g \in G$. We shall see how to concretely compute a Haar measure for $G = \mathrm{SL}_2(\mathbb{R})$. Moreover, for all discrete groups we already explained that the counting measure is a Haar measure.

PROOF. Let us prove the proportionality statement. Fix a Haar measure λ.

Claim 1. For all non-zero function $f \in C_c(G)$ such that $f \geq 0$ we have $\int_G f d\lambda \neq 0$.

To prove this claim, note that given such a function f, there exists $\varepsilon > 0$ such that $U := \{x \in G \mid f(x) \geq \varepsilon\}$ is non-empty, and of course we have $f \geq \varepsilon 1_U$. So it suffices to prove that $\lambda(U) > 0$ for any non-empty open set U in G. By assumption, λ is non-zero so there exists a Borel set A such that $\lambda(A) \neq 0$. Since λ is moreover regular we may actually find a compact subset $K \subset A$ such that $\lambda(K) \neq 0$.

If $U \subset G$ is a non-empty open set, the collection of translates gU, $g \in G$ is an open cover of G, and in particular, of K. By compactness of K, we may extract from it a finite sub-cover. In other words we find a finite set $F \subset G$ such that $K \subset \cup_{g \in F} gU$. But this leads to an inequality on measures:

$$0 < \lambda(K) \leq \sum_{g \in F} \lambda(gU) = |F| \lambda(U).$$

So we indeed arrive at the conclusion that $\lambda(U) \neq 0$, and Claim 1 follows.

Fix two non-zero functions $f, g \in C_c(G)$, such that $f, g \geq 0$.

Claim 2. The ratio $\int_G f d\lambda / \int_G g d\lambda$ makes sense thanks to Claim 1. It does not depend on λ.

Assume that μ is another Haar measure on G. We consider the function $h : G \times G \to \mathbb{R}$ given by the formula

$$h(x, y) = \frac{f(x) g(yx)}{\int_G g(tx) d\mu(t)}.$$

This is easily seen to be a compactly supported continuous function. Moreover, we have $\int_G h(x, y) d\mu(y) = f(x)$ for all $x \in G$. Therefore,

$$\int_G \int_G h(x, y) d\mu(y) d\lambda(x) = \int_G f(x) d\lambda(x).$$

On the other hand, since h is compactly supported and continuous, it is integrable, and Fubini Theorem applies. Combining it with the fact that λ and μ are Haar measures we get

$$\begin{aligned}
\int_G \int_G h(x, y) d\mu(y) d\lambda(x) &= \int_G \int_G \frac{f(x) g(yx)}{\int_G g(tx) d\mu(t)} d\lambda(x) d\mu(y) \\
&= \int_G \int_G \frac{f(y^{-1}x) g(x)}{\int_G g(ty^{-1}x) d\mu(t)} d\lambda(x) d\mu(y) \\
&= \int_G \int_G \frac{f(y^{-1}x) g(x)}{\int_G g(ty^{-1}x) d\mu(t)} d\mu(y) d\lambda(x) \\
&= \int_G \int_G \frac{f(y^{-1}) g(x)}{\int_G g(ty^{-1}) d\mu(t)} d\mu(y) d\lambda(x).
\end{aligned}$$

So we arrive at

$$\int_G f(x) d\lambda(x) = \int_G g(x) d\lambda(x) \int_G \frac{f(y^{-1})}{\int_G g(ty^{-1}) d\mu(t)} d\mu(y) = \int_G g(x) d\lambda(x) \times C,$$

where $C = \int_G \frac{f(y^{-1})}{\int_G g(ty^{-1}) d\mu(t)} d\mu(y)$, which does not depend on λ. This proves Claim 2.

Now if λ' is another Haar measure, we get for any two non-negative, non-zero functions $f, g \in C_c(G)$,

$$\int_G f d\lambda = \frac{\int_G g d\lambda}{\int_G g d\lambda'} \int_G f d\lambda'.$$

Fixing g once and for all and setting $\alpha := \int_G g d\lambda / \int_G g d\lambda'$, we arrive at $\int f d\lambda = \alpha \int f d\lambda'$. So λ and $\alpha\lambda$ coincide as linear functionals on $C_c(G)$, which implies that these two radon measures are equal. □

Example 1.7. Consider the group $H := \{M(a,b) \mid a \in \mathbb{R}^*, b \in \mathbb{R}\}$, where

$$M(a,b) = \begin{pmatrix} a & b \\ 0 & 1/a \end{pmatrix}.$$

This group is parametrized by $\mathbb{R}^* \times \mathbb{R}$. The Haar measure on this group is given by $da\, db/a^2$. One easily checks that indeed this formula defines an invariant measure. Observe that if we only allow positive a's then the group that we obtain acts transitively by homography on the upper half plane, and the stabilizer of every point is trivial. In fact the Haar measure that we gave above is exactly the one coming from the usual hyperbolic metric on the upper half plane (for which the group actually acts by isometries).

Example 1.8. The Haar measure on $G := \mathrm{SL}_2(\mathbb{R})$ can be described as follows. Observe that every element of G can be uniquely written as element of the form hk, where $h \in H := \{M(a,b) \mid a \in \mathbb{R}_+^*, b \in \mathbb{R}\}$ and $k \in K := \mathrm{SO}(2)$. This follow for instance by considering the action by homography of G on the upper half plane. The action is transitive and the stabilizer of the point i is K. So the decomposition follows from Example 1.7.

More explicitly, for any $g \in G$, we can find $a > 0$, $b \in \mathbb{R}$ and $\theta \in [0, 2\pi[$ such that

$$g = M(a,b) \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

With this parametrization, one checks that the measure $\frac{1}{a^2} d\theta\, da\, db$ is a Haar measure on G.

Let us observe that a (left) Haar measure on G as we defined above needs not be right invariant, that is, $\lambda(Ag)$ needs not be equal to $\lambda(A)$ for all $g \in G$, $A \subset G$. But there is a nice way to measure this defect.

Fixing $g \in G$, the Radon measure defined by $A \mapsto \lambda(Ag)$ is again left invariant, because $h(Ag) = (hA)g$ for all Borel set A. So it is again a Haar measure and by the previous theorem, it follows that there exists a constant $\Delta(g)$, depending on g such that

$$\Delta(g)\lambda(Ag) = \lambda(A) \text{ for all every set } A \subset G.$$

Applying the above formula to Ag^{-1} gives $\Delta(g)\lambda(A) = \lambda(Ag^{-1})$, and thus we see that $\Delta(g)$ is characterized by the formula

$$\int_G f(xg) d\lambda(x) = \Delta(g) \int_G f(x) d\lambda(x), \text{ for all } f \in C_c(G).$$

This equation shows that Δ is a continuous map from G to $\mathbb{R}_+^*$, and it is readily seen that it is in fact a group homomorphism. Moreover, It does not depend on a choice of the Haar measure λ.

Definition 1.9. The group homomorphism $\Delta : G \to R_+^*$ is called the *modular function* on the locally compact group G. G is called *unimodular* if this homomorphism is trivial.

EXAMPLE 1.10. We make the following observations.

- Discrete group are obviously unimodular since the counting measure is clearly both left and right invariant.
- Compact groups are unimodular. To see this observe that any Haar measure λ on such a group G is finite. Thus we have $\Delta(g)\lambda(G) = \lambda(Gg^{-1}) = \lambda(G)$ for all $g \in G$, showing that Δ has constant value 1.
- Since the modular function is a character $G \to R_+^*$, any group that does not admit a character is unimodular. For instance simple groups are unimodular. This also gives another proof that compact groups are unimodular.

REMARK 1.11. One should be careful that unimodularity doesn't pass to subgroups. In particular, the restriction of the modular function Δ_G of a group G to a subgroup H needs not be the modular function Δ_H of the subgroup.

For example, we deduce from the previous exemple that $\mathrm{SL}_2(\mathbb{R})$ is unimodular, since it has no character. On the other hand it contains the subgroup $H := \{M(a,b) \mid a \in \mathbb{R}^*, b \in \mathbb{R}\}$, where

$$M(a,b) = \begin{pmatrix} a & b \\ 0 & 1/a \end{pmatrix}.$$

But one verifies that the modular function on H is given by $\Delta_H(M(a,b)) = a^2$, which is non-trivial.

EXERCISE 1.12. We have only considered left Haar measures so far. One easily checks that analogous results (existence and proportionality) also hold for right Haar measure. Check that if λ is a left Haar measure then $A \mapsto \lambda(A^{-1})$ is a right Haar measure. Moreover, check the formula

$$\int_G f(x^{-1})d\lambda(x) = \int_G \Delta(x)f(x)d\lambda(x), \text{ for all } f \in C_c(G).$$

2. Lattices in locally compact groups

DEFINITION 1.13. If G is a topological group, a *discrete* subgroup $\Gamma < G$ is a subgroup which is discrete in G for the induced topology. This amounts to saying that there is a neighborhood $U \subset G$ of the trivial element $e \in G$ such that $U \cap \Gamma = \{e\}$.

When Γ is a discrete subgroup in a locally compact group G, the map $G \to G/\Gamma$ is a local homeomorphism, and it is even a covering map. So we may consider the Radon measure λ_X on $X := G/\Gamma$ obtained locally as the push forward of a right Haar measure λ on G. This can be done precisely by using a partition of unity argument, and is well defined because we chose a right invariant Haar measure on G.

DEFINITION 1.14. A discrete subgroup Γ in a locally compact group G is called a *lattice* if $\lambda_X(X) < \infty$.

We will restrict our attention to locally compact, second countable groups (i.e. those admitting a countable basis of open sets). We write l.c.s.c. for short.

PROPOSITION 1.15. *Given a discrete subgroup Γ in a l.c.s.c. group G. The following facts are equivalent.*

(i) Γ is a lattice in G;
(ii) There exists a set $\Omega \subset G$ with finite right Haar measure such that $\Omega\Gamma = G$;

(iii) There exists a Borel fundamental domain $\mathcal{F} \subset G$ for the right action of Γ on G with finite Haar measure. This means that F is a Borel subset of G such that $\mathcal{F}\Gamma = G$ and $\mathcal{F}g \cap \mathcal{F}h = \emptyset$ for all distinct elements $g, h \in \Gamma$.

LEMMA 1.16. *For any discrete group Γ in a l.c.s.c. group there always exists a Borel fondamental domain.*

PROOF. Fix a neighborhood $V \subset G$ of e such that $V \cap \Gamma = \{e\}$, and pick a neighborhood $U \subset G$ of e such that $U^{-1}U \subset V$. This is possible because the map $(g, h) \mapsto g^{-1}h$ is continuous on $G \times G$. Since G is second countable, there exists a countable set of elements $(g_n)_{n\geq 1}$ in G such that $G = \bigcup_{n\geq 1} g_n U$.

Define inductively a sequence of Borel sets $\mathcal{F}_n \subset G$, $n \geq 1$ as follows. Set $\mathcal{F}_1 = g_1 U$ and

$$\mathcal{F}_{n+1} = g_{n+1}U \setminus \left(g_{n+1}U \cap \bigcup_{k=1}^{n} g_k U\Gamma \right).$$

These are disjoint sets and better, for every $n \neq m$ we have $\mathcal{F}_n\Gamma \cap \mathcal{F}_m\Gamma = \emptyset$, while

$$\bigcup_{k=1}^{n} \mathcal{F}_k\Gamma = \bigcup_{k=1}^{n} g_k U\Gamma.$$

So the set $\mathcal{F} = \bigcup_n \mathcal{F}_n$ is a Borel set such that $\mathcal{F}\Gamma = G$. Assume now that $g, h \in \Gamma$ are two elements such that $\mathcal{F}g \cap \mathcal{F}h \neq \emptyset$. There exist two indices such that $\mathcal{F}_n g \cap \mathcal{F}_m h \neq \emptyset$. By construction this forces n and m to be equal. Since $\mathcal{F}_n \subset g_n U$ we can then find $x, y \in U$ such that $g_n x g = g_n y h$, which leads to $x^{-1}y = gh^{-1} \in V \cap \Gamma = \{e\}$. So we conclude that $g = h$.

Thus $\mathcal{F}$ is indeed a Borel fundamental domain. □

PROOF OF PROPOSITION 1.15. Clearly *(iii)* implies *(ii)*. Assuming *(ii)*, note that since the Haar measure is regular, we may enlarge Ω if necessary to assume that it is an open set with finite measure. One observe that the push forward measure of Ω on G/Γ majorizes the measure λ_X, because it does so locally. This forces $\lambda_X(X)$ to be finite, and *(i)* follows.

More generally, we claim that given any Borel fundamental domain $\mathcal{F}$ for the right Γ-action on G, the measure λ_X is exactly the push forward of the restriction of the right Haar measure λ to $\mathcal{F}$ by the projection map $\pi : G \to G/\Gamma$. This implies in particular $\lambda_X(X) < \infty$ if and only if $\lambda(\mathcal{F}) < \infty$ and hence, that *(iii)* $\Leftrightarrow$ *(i)*.

To prove this claim, it suffices to prove that these two measures coincide locally on G/Γ. More precisely, it suffices to check that for all open set $U \subset G$ such that $Ug \cap Uh = \emptyset$ for all $g, h \in \Gamma$, $g \neq h$, the two measures coincide on $\pi(U)$. Take a Borel subset $A \subset \pi(U)$. We need to check that $\lambda_X(A) = \lambda(\pi^{-1}(A) \cap F)$. By definition, we have

$$\lambda_X(A) = \lambda(U \cap \pi^{-1}(A)) = \sum_{g\in\Gamma} \lambda(U \cap \pi^{-1}(A) \cap \mathcal{F}g).$$

In the last inequality above, we used the fact that G is the disjoint union of the sets $\mathcal{F}g$, $g \in \Gamma$. Now since the measure λ is right-invariant and since the set $\pi^{-1}(A)$ is globally right Γ-invariant, we further find

$$\lambda_X(A) = \sum_{g\in\Gamma} \lambda(Ug^{-1} \cap \pi^{-1}(A)g^{-1} \cap \mathcal{F}) = \sum_{g\in\Gamma} \lambda(Ug^{-1} \cap \pi^{-1}(A) \cap \mathcal{F}).$$

Since the sets Ug, $g \in \Gamma$ are pairwise disjoint and $\pi^{-1}(A) \subset \bigcup_{g\in\Gamma} Ug$, we conlcude

$$\lambda_X(A) = \lambda(\bigcup_{g\in\Gamma} Ug^{-1} \cap \pi^{-1}(A) \cap \mathcal{F}) = \lambda(\pi^{-1}(A) \cap \mathcal{F}),$$

as desired. □

In the above proof we used σ-additivity of the measure for sums indexed by elements of Γ, without knowing a priori that Γ is countable. But this is indeed the case since G is second countable and Γ is discrete inside G.

EXAMPLE 1.17. It is trivial to see that $\mathbb{Z}^n$ is a lattice in $\mathbb{R}^n$, for which a fundamental domain is given by $[-1, 1[^n$. The quotient is the torus. In this case the measure λ_X is just the Lebesgue measure on the torus.

A less trivial example is that of $\mathrm{SL}_2(\mathbb{Z})$ inside $\mathrm{SL}_2(\mathbb{R})$. To see that this is indeed a lattice, one can use the fact that the action of $\mathrm{SL}_2(\mathbb{Z})$ on the hyperbolic half-plane admits a Borel fundamental domain with finite measure. We leave the details as an exercise. In the same spirit, the fundamental group of any compact surface with negative curvature embeds as a lattice inside $\mathrm{PSL}_2(\mathbb{R})$.

From the above facts, we derive the following corollary, which rules out existence of lattices in some pathologic cases.

COROLLARY 1.18. *An l.c.s.c. group G admitting a lattice Γ is unimodular, and the measure λ_X on the quotient space G/Γ is G-invariant.*

PROOF. If Γ is a lattice in G, then there exists a Borel fundamental domain $\mathcal{F}$ with finite right Haar measure: $\lambda(\mathcal{F}) < \infty$. Observe now that any two Borel fundamental domains $\mathcal{F}$ and $\mathcal{F}'$ actually have the same measure λ. Indeed,

$$\lambda(\mathcal{F}) = \sum_{g\in\Gamma} \lambda(\mathcal{F} \cap \mathcal{F}'g) = \sum_{g\in\Gamma} \lambda(\mathcal{F}g^{-1} \cap \mathcal{F}') = \lambda(\mathcal{F}').$$

Now if $g \in G$, we see that $g\mathcal{F}$ is again a fundamental domain for the right Γ-action. This implies that $\lambda(g\mathcal{F}) = \lambda(\mathcal{F})$, and hence λ is left g-invariant. So λ is both left and right G-invariant, showing that G is unimodular.

The second assertion then follows from the definition of λ_X. □

If G and H and two l.c.s.c. groups with lattices Γ and Λ, respectively, then $\Gamma \times \Lambda$ is a lattice in $G \times H$. This silly example motivates the following definition.

DEFINITION 1.19. A lattice Γ in a product of groups $\Pi_{i=1}^n G_i$ is said to be *irreducible* if its projects on each of the factors onto a dense subgroup.

We shall see later that $\mathrm{SL}_n(\mathbb{Z}[\sqrt{2}])$ embeds as an irreducible lattice inside $\mathrm{SL}_n(\mathbb{R}) \times \mathrm{SL}_n(\mathbb{R})$.

CHAPTER 2

Recalls on Lie groups

This chapter is partially built on Helgason's book [**Hel78**], Chapter II.

1. Main definitions and examples

DEFINITION 2.1. A *Lie group* over $k = \mathbb{R}$ or $\mathbb{C}$ is a smooth manifold G over k which admits a group structure such that the corresponding structure map $m : (x, y) \in G \times G \mapsto xy^{-1} \in G$ is smooth. A *morphism* between two Lie groups will be by definition a smooth group homomorphism.

We will most of the time focus on *real* Lie groups.

REMARK 2.2. As usual with groups, one checks that the smoothness of the map $(x, y) \mapsto xy^{-1}$ is equivalent to that of the product map and the inverse map. It also implies that for all $g \in G$, the translation maps $L_g : x \mapsto gx$ and $R_g : x \mapsto xg$ are smooth.

EXAMPLE 2.3. Besides the trivial examples of $\mathbb{R}^n$ and $\mathbb{T}^n := \mathbb{R}^n/\mathbb{Z}^n$, here are some of the standard examples to always keep in mind.

- The multiplicative group $\mathrm{GL}_n(k)$ is an open set inside $M_n(k)$. As such, it may be endowed with the corresponding smooth structure. It is of dimension n^2 as a manifold.
- *(Special linear groups)* The subgroup $\mathrm{SL}_n(k)$ consisting of matrices with determinant 1 is a k-submanifold (because the determinant map is a submersion), and it is invariant under the map m, so it is also a Lie group over k. Its dimension over k is $n^2 - 1$.
- *(Orthogonal groups)* For any non-degenerate quadratic form q over k^n, the corresponding orthogonal group $O(q, k)$ is clearly invariant under the product map. It is a submanifold of $M_n(k)$ because the map $g \in M_n(k) \mapsto g^T Q g \in S_n(k)$ is a submersion (here Q stands for the matrix of q, and $S_n(k)$ for the space of symmetric matrices). It is then a Lie group over k. Observe that if the two quadratic forms are conjugate, then the subgroups are conjugate, thus isomorphic as Lie groups. So in the complex setting $O(q, \mathbb{C})$ is always isomorphic with $O(n)$ while in the real case, we get the groups $O(p, q)$ indexed by the signature (p, q) of the quadratic form. We will denote by $\mathrm{SO}(q)$ the intersection of $O(q)$ with $\mathrm{SL}_n(k)$.
- *(Isometry groups)* The semi-direct product $\mathrm{SO}(n, \mathbb{R}) \ltimes \mathbb{R}^n$ is a Lie group (endowed with the product structure as a manifold). It is the Lie group of orientation preserving isometries of $\mathbb{R}^n$.
- *(Unitary groups)* The group $U(n) := \{g \in M_n(\mathbb{C}) \mid g^* g = 1\}$ is a *real* Lie group. But it is not a complex Lie group, although it lives inside $M_n(\mathbb{C})$. It is because the map $g \in M_n(\mathbb{C}) \mapsto g^* g \in S_n(\mathbb{R})$ is a submersion (only defined over $\mathbb{R}$).
- *(Symplectic groups)* The group $\mathrm{Sp}(2n, k) := \{g \in M_{2n}(k) \mid g^T J g = J\}$, where $J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$, is a Lie group over k.

The advantage with Lie group, is that they come with a so-called Lie algebra, giving all the tools from linear algebra to study them.

DEFINITION 2.4. Given an arbitrary field k, a *Lie algebra* over k is a finite dimensional k-vector space V, endowed with a bilinear map $[\cdot,\cdot] : V \times V \to V$ (the so-called *Lie bracket*) which satisfies the two axioms:

- $[X,Y] = -[Y,X]$ for all $X, Y \in V$ (anti-symmetry);
- $[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$ for all $X, Y, Z \in V$ (Jacobi identity).

Concretely, given any Lie group G, we define its Lie algebra $\mathfrak{g}$ as follows.

As a vector space, the Lie algebra is just the tangent space T_eG at the identity element. This tangent space can be identified with the vector space of *left invariant vector fields*.

Recall that a *vector field* on the manifold G is the derivation of the algebra $C^\infty(G)$, that is, a linear map $X : C^\infty(G) \to C^\infty(G)$ such that $X(f_1 f_2) = f_1 X(f_2) + f_2 X(f_1)$. On the other side, for any $g \in G$, the translation map $L_g : h \in G \mapsto gh \in G$ induces a map $\tau_g : f \in C^\infty(G) \to C^\infty(G)$. A vector field X on G is called *left-invariant* if it commutes with τ_g for all $g \in G$, i.e. $X(f \circ L_g) = (X(f)) \circ L_g$, for all $f \in C^\infty(G)$. Note that this formula is equivalent to

$$(X(f))(g) = (X(f \circ L_g))(e) \text{ for all } g \in G.$$

We write $X_g(f)$ instead of $(X(f))(g)$, so that the map $X \mapsto X_e$ is a linear isomorphism from the set of left invariant vector fields onto the tangent space of G at e. So we identify this way $\mathfrak{g}$ with the space of left invariant vector fields on G. Then the formula $[X,Y] := XY - YX$ (defined by composition of endomorphisms on $C^\infty(G)$) defines a derivation of $C^\infty(G)$, i.e. a vector field. The map $[\cdot,\cdot]$ is clearly bilinear and anti-symmetric and the Jacobi identity is easily checked.

DEFINITION 2.5. The Lie algebra $(\mathfrak{g}, [\cdot,\cdot])$ defined above is called the Lie algebra of G.

EXAMPLE 2.6. The Lie algebra of $G = \mathrm{GL}_n(k)$ is $M_n(k)$, endowed with the Lie bracket

$$[X,Y] = XY - YX.$$

Let us indicate how to prove this fact. First observe that since $\mathrm{GL}_n(k)$ is open inside $M_n(k)$, the tangent space at every point is naturally identified with $M_n(k)$. Now given a vector $X \in M_n(k)$, we view it as a left-invariant vector field by the formula $g \in G \mapsto gX \in M_n(k) \simeq T_gG$. Its effect on $C^\infty(G)$ is given by $X(f)(g) = (df)_g(gX)$ for all $f \in C^\infty(G)$, $g \in G$. Differentiating further, we see that for all $X, Y \in M_n(k)$, $Y(X(f))(g) = d^2(f)_g(gX, gY) + (df)_g(gYX)$. Since the second derivative $d^2(f)_g$ is a symmetric bilinear form, we get $(Y(X(f))(g) - X(Y)(f))(g) = (df)_g(gXY) - (df)_g(gYX) = (XY - YX)(f)(g)$, as desired.

Now recall that for any sub-manifold $N \subset M$ defined by a submersion $\phi : M \to M'$ as $N = \phi^{-1}(\{x\})$, the tangent space of N at a point is just the kernel of the derivative of ϕ at that point. In particular we may compute the Lie algebras of all the standard examples mentioned above.

EXAMPLE 2.7. We have the following computations. We only describe the underlying vector space, because the Lie bracket is simply the restriction of the Lie bracket on $M_n(k)$.

- Since the derivative of the determinant map is the trace, the Lie algebra of $\mathrm{SL}_n(k)$ is the vector space of trace 0 matrices in $M_n(k)$;

- The Lie algebra of $O(n)$ is the subspace of matrices such that $X^T + X = 0$ (anti-symmetric matrices);
- The Lie algebra of $U(n)$ are matrices such that $X^* + X = 0$ (anti-hermitian matrices);
- The Lie algebra of the symplectic groups $\mathrm{Sp}(2n)$ is the space of matrices satisfying $X^T J + JX = 0$.

2. Isomorphisms

In this section we explore the relationship between a Lie group and its Lie algebra. Observe that given any Lie group G and any finite group F, the product $G \times F$ is naturally a Lie group, with the same Lie algebra as G. So we cannot distinguish these two groups only with their Lie algebra. The following Lemma allows to get rid of that noise.

LEMMA 2.8. *The connected component G^0 of the identity of a Lie group G, is a normal subgroup, which is open (and hence closed) inside G. In particular G^0 is also a Lie group and G and G^0 have the same Lie algebra.*

PROOF. Since the map $(g, h) \mapsto gh^{-1}$ is continuous and equals e at (e, e), it maps the connected set $G^0 \times G^0$ into G^0. So G^0 is a subgroup of G, and the same argument shows that it is normal inside G (This is a general fact for locally compact groups).

Clearly since G is a manifold, it is locally connected, so G^0 is open. Observe that an open subgroup H in a locally compact group is always closed, since its complementary is the union of the open sets gH, $g \in G \setminus H$. □

The above definition allows to restrict to connected Lie groups. But there still exist distinct connected Lie groups with the same Lie algebra. For example one can show that SU(2) is a 2-covering of SO(3). In fact, this situation appears for every Lie group which is not simply connected.

LEMMA 2.9. *The universal cover $\widetilde{G}$ of a connected Lie group G is naturally endowed with a Lie group structure such that the covering map $\widetilde{G} \to G$ is a group homomorphism. Its kernel is contained in the center of $\widetilde{G}$.*

PROOF. Denote by $\pi : \widetilde{G} \to G$ the covering map. Fix a lift e of the identity element e_G. We define the product on $\widetilde{G}$ as follows. Fix $g, h \in \widetilde{G}$ and choose paths $t \mapsto g_t, h_t \in \widetilde{G}$ between e and g, h: $g_0 = h_0 = e$, $g_1 = g$, $h_1 = h$. The product path $t \mapsto \pi(g_t)\pi(h_t)$ in G is a path between e_G and $\pi(g)\pi(h)$. Lift it to a path γ inside $\widetilde{G}$ starting at e. We set $gh := \gamma(1)$. One checks that this definition is independent of the choices of paths that we made.

It is then easy to verify that this product law is associative, that e is a neutral element, and that the inverse g^{-1} of g is the end point of a lift starting at e of the path $t \mapsto \pi(g_t)^{-1}$. Moreover the covering map π is clearly a group homomorphism. Since it is also a local homeomorphism, we may define the analytic structure of $\widetilde{G}$ by declaring that π is locally analytic. The fact that G is a Lie group implies that the structure map $(g, h) \mapsto gh^{-1}$ is analytic.

Finally, take $g, h \in \widetilde{G}$, with $\pi(g) = e$. Take a path $t \mapsto g_t$ from e to g such that $g_t = g$ for all $t \geq 1/2$ and a path $t \mapsto h_t$ from e to h such that $h_t = e$ for all $e \leq 1/2$. We have:

- $\pi(g_t)\pi(h_t) = \pi(g_t) = \pi(h_t)\pi(g_t)$ if $t \leq 1/2$;

- $\pi(g_t)\pi(h_t) = \pi(g)\pi(h_t) = \pi(h_t) = \pi(h_t)\pi(g_t)$ if $t \geq 1/2$.

So $gh = hg$, as desired. □

A connected Lie group is always locally isomorphic to its universal cover in the following sense.

DEFINITION 2.10. Two Lie groups G and H are said to be locally isomorphic if there exist open neighborhoods $U \subset G$ and $V \subset H$ of the identity elements e_G and e_H and an analytic diffeomorphism φ from U onto V such that $\varphi(gh) = \varphi(g)\varphi(h)$ for all $g, h \in U$ such that $gh \in U$.

THEOREM 2.11. *Consider two Lie groups G and H with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$. Denote by $\widetilde{G}$ and $\widetilde{H}$ the universal covers of the identity components of G and H respectively. The following are equivalent.*

(i) G and H are locally isomorphic;
(ii) $\widetilde{G}$ and $\widetilde{H}$ are isomorphic;
(iii) $\mathfrak{g}$ and $\mathfrak{h}$ are isomorphic.

PROOF. To prove that $(i) \Leftrightarrow (ii)$ first observe that being locally isomorphic is an equivalence relation and that G (resp. H) is locally isomorphic with $\widetilde{G}$ (resp. $\widetilde{H}$). Then we leave it to the reader to check that two simply connected groups are locally isomorphic if and only if they are isomorphic.

We postpone the rest of the proof to the end of this section. □

The implication $(ii) \Rightarrow (iii)$ will rely on the following lemma.

LEMMA 2.12. *Consider two Lie groups G and H, with respective Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$. If $\phi : G \to H$ is a differentiable homomorphism then its derivative $d\phi : \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism.*

PROOF. Here the statement is ambiguous because the underlying vector space of the Lie algebra may be viewed as the tangent space at the identity, in which case we consider the derivative of ϕ at e, or as the vector space of left-invariant vector fields in which case we consider the push forward of vector fields, but we need to check that the push forward of a left-invariant vector field is again left invariant. This ambiguity is solved by the following observation.

Take tangent vector $X \in T_eG$, and denote by $X' := (d\phi)_e(X) \in T_eH$. Then the left invariant vector fields $\widetilde{X}$ and $\widetilde{X}'$ on G and H associated to X and Y satisfy the relation $\widetilde{X}'(f)_{\phi(g)} := \widetilde{X}(f \circ \phi)_g$, for all $f \in C^\infty(H)$, $g \in G$. Indeed, by definition, we have

$$\widetilde{X}'(f)_{\phi(g)} = X'(f \circ L_{\phi(g)}) = X(f \circ L_{\phi(g)} \circ \phi).$$

But since ϕ is a homomorphism, $\phi \circ L_g = L_{\phi(g)} \circ \phi$, and we deduce

$$\widetilde{X}'(f)_{\phi(g)} = X(f \circ \phi \circ L_g) = \widetilde{X}(f \circ \phi)_g,$$

as claimed.

To conclude the proof of the lemma, take $X, Y \in \mathfrak{g}$. Denote by $X' := d\phi(X)$ and $Y' := d\phi(Y)$, and by $\widetilde{X}, \widetilde{Y}, \widetilde{X}', \widetilde{Y}'$ the left invariant vector fields corresponding to X, Y, X', Y'. For all $f \in C^\infty(H)$ and $h \in H$, we have $\widetilde{X}'(\widetilde{Y}'(f))_{\phi(h)} = \widetilde{X}(\widetilde{Y}'(f) \circ \phi))_h = \widetilde{X}(\widetilde{Y}(f \circ \phi))(h)$ and hence

$$(\widetilde{X}'\widetilde{Y}' - \widetilde{Y}'\widetilde{X}')(f)_{\phi(h)} = (\widetilde{X}\widetilde{Y} - \widetilde{Y}\widetilde{X})(f \circ \phi)_h.$$

When $h = e$, the left hand side is by definition equal to $[d\phi(X), d\phi(Y)](f)$ and the right hand side is $[X, Y](f \circ \phi) = d\phi([X, Y])(f)$. Thus ϕ is a Lie algebra homomorphism. □

Applying the above lemma to the covering map $\widetilde{G} \to G$, we deduce that a connected Lie group and its universal cover have the same Lie algebra. So the implication $(ii) \Rightarrow (iii)$ readily follows.

The converse implication is based on the exponential mapping, which is the main tool to pass from the Lie algebra $\mathfrak{g}$ to the group G.

Fix a Lie group G and denote by $\mathfrak{g}$ its Lie algebra. The exponential mapping is an analytic map $\exp : \mathfrak{g} \to G$ whose differential is the identity on $\mathfrak{g}$. In order to properly define the exponential, we start with a lemma.

LEMMA 2.13. *Given a vector $X \in \mathfrak{g}$, denote by $\widetilde{X}$ the corresponding left invariant vector field. The differential equation*

$$\gamma'(t) = \widetilde{X}_{\gamma(t)} \text{ and } f(0) = e \tag{2.1}$$

admits a unique solution $\gamma : \mathbb{R} \to G$, which is a one parameter subgroup. We define $\exp(X) := \gamma(1)$. Then the map $\exp : \mathfrak{g} \to G$ is smooth, and its derivative at e is the identity.

PROOF. First observe that since $\widetilde{X}$ is left invariant, we have for all $g, h \in G$, $(dL_g(\widetilde{X}))_h = \widetilde{X}_{gh}$.

Existence and uniqueness of γ on a neighborhood of 0 follow from Cauchy Lipschitz theorem. So we obtain a solution γ defined on an open interval I. We may assume that I is the maximal open interval on which γ is defined. Observe that for all $s \in I$ and $t \in \mathbb{R}$ such that $s + t \in I$, the map $\sigma : t \mapsto \gamma(s)^{-1}\gamma(s + t)$ satisfies: $\sigma(0) = e$ and

$$\sigma'(t) = dL_{\gamma(s)^{-1}}(\gamma'(s + t)) = dL_{\gamma(s)^{-1}}\widetilde{X}_{\gamma(s+t)} = \widetilde{X}_{\gamma(s)^{-1}\gamma(s+t)} = \widetilde{X}_{\sigma(t)}.$$

In particular σ is a solution to (2.1) which is defined on $(I - s)$. By maximality of I, we must have $I = \mathbb{R}$. Moreover, σ and γ coincide for all $s, t \in \mathbb{R}$, showing that $\gamma(s + t) = \gamma(s)\gamma(t)$. So indeed γ is a one parameter subgroup of G.

The fact that exp is smooth also follows from the regularity of solutions of differential equations with a parameter. To emphasize the dependance on X, let us write $\gamma_t(X)$ instead of γ_t for the solution of (2.1). From uniqueness, we see that $\gamma_t(sX) = \gamma_{ts}(X)$ for all $s, t \in \mathbb{R}$. This implies that $\gamma_t(X) = \exp(tX)$. From the chain rule, we see that

$$(d\exp)_0(X) = \left\{ \frac{d}{dt}(t \mapsto \exp(tX)) \right\}_{t=0} = \gamma'(0) = X. \qquad \square$$

REMARK 2.14. In fact, the curve γ from the previous lemma is characterized as the unique one parameter subgroup in G such that $\gamma'(0) = X$.

DEFINITION 2.15. The map $\exp : \mathfrak{g} \to G$ is called the *exponential mapping* of G.

EXAMPLE 2.16. When $G = \mathrm{GL}_n(k)$ we recover the standard exponential of matrices. Indeed, for $X \in M_n(\mathbb{C})$, the curve $\gamma : t \mapsto \exp(tX) = \sum_{n \geq 0}(tX)^n/n!$ satisfies $\gamma'(t) = \gamma(t)X$, which is indeed the same as $\widetilde{X}(\gamma(t))$ with the description of the left invariant vector field as in Example 2.6.

PROPOSITION 2.17. *The exponential mapping satisfies the following properties.*

(1) *If $\Phi : G \to H$ is a differentiable homomorphism, and $\phi : \mathfrak{g} \to \mathfrak{h}$ is its derivative at the identity element e then*

$$\Phi \circ \exp = \exp \circ \phi.$$

(2) *For all $X \in \mathfrak{g}$, $g \in G$ and all function $f \in C^\infty(G)$, for all $t \in \mathbb{R}$ small enough and $N \in \mathbb{N}$, we have*

$$f(g \exp(tX)) = \sum_{n=0}^{N} \frac{1}{n!} (\widetilde{X})^n (f)_g + O(t^{N+1}).$$

(3) *For all $X, Y \in \mathfrak{g}$ and all t small enough, we have the two formulae*

$$\exp(tX)\exp(tY) = \exp\left(t(X+Y) + \frac{t^2}{2}[X,Y] + O(t^3)\right)$$

$$\exp(tX)\exp(tY)\exp(-tX)\exp(-tY) = \exp\left(t^2[X,Y] + O(t^3)\right).$$

Proof. (1) Fix $X \in \mathfrak{g}$ and define $\gamma : t \in \mathbb{R} \mapsto \Phi(\exp(tX))$. Denote by $Y := \phi(X) = (d\Phi)_e(X)$. Denote by $\widetilde{X}$ and $\widetilde{Y}$ the left invariant vector fields on G and H associated with X and Y. In the proof of Lemma 2.12 we saw that $\widetilde{Y}(f)_{\Phi(g)} = \widetilde{X}(f \circ \Phi)_g$ for all $f \in C^\infty(H)$, $g \in G$. This formula rephrases as $\widetilde{Y}(f)_{\Phi(g)} = ((d\Phi)_g(\widetilde{X}_g))(f)$. Hence $\widetilde{Y}_{\Phi(g)} = (d\Phi)_g(\widetilde{X}_g)$. Thus we get

$$\gamma'(t) = (d\Phi)_{\exp(tX)}(\widetilde{X}_{\exp(tX)}) = \widetilde{Y}_{\gamma(t)}.$$

By uniqueness, we see that $\gamma(t) = \exp(tY)$, proving the formula.
(2) Fix $X \in \mathfrak{g}$. Observe that for all $f \in C^\infty(G)$ and $g \in G$, we have

$$\widetilde{X}_g(f) = X(f \circ L_g) = \frac{d}{dt}\{f(g\exp(tX))\}_{t=0}.$$

In particular, we deduce that

$$[\widetilde{X}f](g\exp(uX)) = \frac{d}{dt}\{f(g\exp(uX)\exp(tX))\}_{t=0} = \frac{d}{du}\{f(g\exp(uX))\}.$$

By induction, we then see that

$$[\widetilde{X}^n f](g\exp(uX)) = \frac{d^n}{du^n}\{f(g\exp(uX))\}.$$

So the result follows from the smoothness of f.
(3) Let us only prove the first formula and leave the other one as an exercise. Fix a smooth function $f \in C^\infty(G)$ at e. From the computations in (2), we see that for all $n, m \geq 0$, we have

$$[\widetilde{X}^n \widetilde{Y}^m f](e) = \frac{d^n}{dt^n}\frac{d^m}{ds^m}\{f(\exp(tX)\exp(sY))\}_{s=0,t=0}.$$

So we get the following generalization of (2).

$$f(\exp(tX)\exp(sY)) = \sum_{n,m=0}^{N} \frac{t^n}{n!}\frac{s^m}{m!}[\widetilde{X}^n\widetilde{Y}^m f](e) + O(t^{N+1}s^{N+1}), \tag{2.2}$$

for t, s small enough. On the other hand, we know that there exists an analytic function $Z : I \to \mathfrak{g}$, defined on an open interval around 0 such that

$$\exp(tX)\exp(tY) = \exp(Z(t)) \text{ for all } t \text{ small enough.}$$

Since $Z(0) = 0$, we may find $Z_1, Z_2 \in \mathfrak{g}$ such that $Z(t) = tZ_1 + t^2 Z_2 + O(t^3)$ for all $t \in I$. The result easily follows from expanding $f(\exp(Z(t)))$ and comparing the coefficients with (2.2) applied to $s = t$. □

Note that (1) above implies in particular that the exponential mapping $\exp_H : \mathfrak{h} \to H$ on a subgroup H of G is the restriction of $\exp_G : \mathfrak{g} \to G$ to the Lie subalgebra $\mathfrak{h}$. In combination with Example 2.16, this gives that the exponential mapping on any linear group is the usual exponential for matrices. This fact can also be derived from (2) above.

Even if we don't provide the main tools to actually prove the implication $(iii) \Rightarrow (i)$ of Theorem 2.11, let us at least mention that it goes as one expects.

Idea of proof of Theorem 2.11, $(iii) \Rightarrow (i)$. Given a Lie algebra homomorphism $\phi : \mathfrak{g} \to \mathfrak{h}$, if ϕ is the derivative of a group homomorphism Φ from G to H, it has to satisfy $\Phi(\exp(X)) = \exp(\phi(X))$ for all $X \in \mathfrak{g}$, due to item (1) in Proposition 2.17.

So fix a neighborhood U_0 of 0 in $\mathfrak{g}$ such that exp is a diffeomorphism from U_0 to its image $U := \exp(U_0)$. Denote by $\log : U \to U_0$ the inverse map. Then for all $g \in U$, we set $\Phi(g) := \exp(\phi(\log(g)))$. In other words, we *define* Φ on U by the formula $\Phi(\exp(X)) = \exp(\phi(X))$ for all $X \in U_0$.

Of course, since ϕ is an isomorphism, we find that Φ is a diffeomorphism from U onto its image. The hard part is to prove that this actually defines a group homomorphism, i.e. that $\Phi(\exp(X)\exp(Y)) = \Phi(\exp(X))\Phi(\exp(Y))$, for all $X, Y \in U_0$ such that $\exp(X), \exp(Y) \in U$. One possibility for that is to show that the element $\log(\exp(X)\exp(Y))$ is obtained as an infinite (but convergent) sum of the form

$$\log(\exp(X)\exp(Y)) = X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}[X,[X,Y]] - \frac{1}{12}[Y,[X,Y]] + \dots$$

where each term in the sum is a scalar multiple of "admissible" terms in the following sense (with the involved scalars not depending on X and Y). The set of "admissible" elements in $\mathfrak{g}$ is the smallest subset of $\mathfrak{g}$ which contains X and Y and is stable under Lie bracket. In other words, an admissible term is expressed as a word in X and Y, together with a choice of placing the brackets.

This formula, called the Baker-Campbell Hausdorff formula is proved in [**Hal15**]. Note that since ϕ is a Lie algebra homomorphism, it maps an admissible term in X, Y to the the admissible term in $\phi(X)$, $\phi(Y)$, with the same expression. So clearly, ϕ maps $\log(\exp(X)\exp(Y))$ to $\log(\exp(\phi(X))\exp(\phi(Y)))$. Proving the result. □

3. Lie subgroups

Exactly as there are two kinds of "sub manifolds", there are two kinds of "Lie subgroups": the embedded subgroups, and the subgroups arising from an injective immersion. The term "Lie subgroup" will be used for the second, larger class of subgroups.

Definition 2.18. Given a Lie group G, a *Lie subgroup* $H \subset G$ is a Lie group such that the inclusion map is an immersion at every point.

Example 2.19. Pick a point in $a \in \mathbb{R}^2$ which is not a multiple of a point in $\mathbb{Q}^2$ e.g. $a = (1, \sqrt{2})$. In the torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$, the image of the line $\mathbb{R}a$ is a Lie subgroup. By our choice of a it is dense in $\mathbb{T}^2$.

PROPOSITION 2.20. *Given a Lie subgroup $H \subset G$, its Lie algebra $\mathfrak{h}$ is naturally identified with a Lie subalgebra of $\mathfrak{g}$. The embedding $\mathfrak{h} \subset \mathfrak{g}$ may be described as*

$$\mathfrak{h} = \{X \in \mathfrak{g} \mid \exp(tX) \in H, \text{ for all } t \in \mathbb{R}\}^1.$$

Moreover, any Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$ is the Lie algebra of exactly one connected Lie subgroup H of G.

PROOF. Since the inclusion map $i : H \hookrightarrow G$ is an analytic group homomorphism, its differential di is a Lie algebra homomorphism. Moreover, since i is an immersion, di is injective, thus giving the desired inclusion $\mathfrak{h} \hookrightarrow \mathfrak{g}$. We admit the other statements. □

While the above result gives a perfect correspondence between connected Lie subgroups and Lie subalgebras, there also exist a nice characterization of *embedded* Lie subgroups.

THEOREM 2.21 (Cartan, Von Neumann). *An embedded Lie subgroup is always closed. Conversely any closed subgroup of a Lie group is automatically an embedded Lie subgroup.*

This result is very strong because it allows to pass from a mere topological property to a much stronger Lie subgroup structure.

COROLLARY 2.22. *Consider two Lie groups G and H, with G connected, and a Lie group homomorphism $\Phi : G \to H$. Denote by $\phi : \mathfrak{g} \to \mathfrak{h}$ its derivative. We have*

(1) The kernel of Φ is a Lie subgroup of G. Its Lie algebra is the kernel of ϕ;
(2) The image of Φ is a Lie subgroup of H; its Lie algebra is the range of ϕ;

PROOF. (1) Since Φ is continuous, its kernel is a closed subgroup, hence it is an embedded Lie subgroup. Its Lie algebra is then $\{X \in \mathfrak{g} \mid \Phi(\exp(tX)) = e_H \text{ for all } t \in \mathbb{R}\}$. Since $\Phi(\exp(tX)) = \exp(t\phi(X))$ for all $X \in \mathfrak{g}$, we clearly see that the Lie algebra of $\mathrm{Ker}(\Phi)$ is the kernel of ϕ.

(2) Denote by $H_1 \subset H$ the Lie subgroup of H whose Lie algebra is $\phi(\mathfrak{g})$. This group is generated by the elements $\exp(\phi(X))$, $X \in \mathfrak{g}$. On the other hand $\Phi(G)$ is generated by the elements $\Phi(\exp(X))$, $X \in \mathfrak{g}$. So we see that $\Phi(G)$ coincides with H_1. □

4. The adjoint representation

There are in fact two adjoint representations: one for the Lie group, one for the Lie algebra. We will see how they are related and give some extra properties.

Lie group setting. Given a Lie group G, we may define for each $g \in G$ a smooth automorphism $I(g) : h \in G \mapsto ghg^{-1} \in G$. Such group automorphisms are called *inner automorphisms*. The differential of I at e is then an invertible endomorphism of the Lie algebra $\mathfrak{g}$ of G, denoted by $\mathrm{Ad}(g) \in \mathcal{L}(\mathfrak{g})$; it is even a Lie algebra automorphism. The mapping $\mathrm{Ad} : G \to \mathrm{GL}(\mathfrak{g})$ is then a linear representation of G, called the *adjoint representation*.

Lie algebra setting. Given a Lie algebra $\mathfrak{g}$ over k, a *derivation* D of $\mathfrak{g}$ is a k-linear map $D : \mathfrak{g} \to \mathfrak{g}$ such that $D([X, Y]) = [DX, Y] + [X, DY]$ for all $X, Y \in \mathfrak{g}$. One checks that if D and D' are two derivations, then so is $DD' - D'D$. This operation turns the vector space $\mathrm{Der}(\mathfrak{g})$ of all derivations of $\mathfrak{g}$ into a Lie algebra.

[1] In fact this description only holds if H has countably many connected components, which we will always assume.

It follows from the Jacobi identity that any $X \in \mathfrak{g}$, the endomorphism $\mathrm{ad}(X) : \mathfrak{g} \to \mathfrak{g}$ is actually a derivation, called an *inner derivation*. The map $\mathrm{ad} : \mathfrak{g} \to \mathrm{Der}(\mathfrak{g})$ is actually a Lie algebra homomorphism, called the *adjoint representation*. The term *representation* of a Lie algebra refers to a Lie algebra homomorphism from a given Lie algebra into the Lie algebra $\mathcal{L}(V)$ of all endomorphisms of a finite dimensional vector space V, endowed with the Lie bracket of $[X, Y] = XY - YX$. Since ad is a representation, its image is a Lie subalgebra of $\mathrm{Der}(\mathfrak{g})$.

LEMMA 2.23. *Take a Lie group G with Lie algebra $\mathfrak{g}$. The group $\mathrm{Aut}(\mathfrak{g})$ of all automorphisms of $\mathfrak{g}$ is a Lie group. Its Lie algebra is $\mathrm{Der}(\mathfrak{g})$. The subgroup $\mathrm{Int}(\mathfrak{g}) := \{\mathrm{Ad}(g) \mid g \in G\}$ is a Lie subgroup of $\mathrm{Aut}(\mathfrak{g})$. Its Lie algebra is $\mathrm{ad}(\mathfrak{g})$. Finally the map $\mathrm{ad} : \mathfrak{g} \to \mathrm{Der}(\mathfrak{g})$ is the derivative at e of the homomorphism $\mathrm{Ad} : G \to \mathrm{Aut}(\mathfrak{g})$.*

PROOF. $\mathrm{Aut}(\mathfrak{g})$ is clearly a closed subgroup of $\mathrm{GL}(\mathfrak{g})$ so it is an embedded Lie subgroup. Denote by $\exp : \mathcal{L}(\mathfrak{g}) \to \mathrm{GL}(\mathfrak{g})$ the exponential mapping (which coincides with the usual exponential of endomorphisms). We know that the Lie algebra of $\mathrm{Aut}(\mathfrak{g})$ is the set $\{D \in \mathcal{L}(\mathfrak{g}) \mid \exp(tD) \in \mathrm{Aut}(\mathfrak{g}) \text{ for all } g \in \mathbb{R}\}$. On the one hand, observe that if $\exp(tD)$ is a Lie algebra automorphism for all $t \in \mathbb{R}$, then

$$\exp(tD)[X, Y] = [\exp(tD)X, \exp(tD)Y], \text{ for all } X, Y \in \mathfrak{g}, t \in \mathbb{R}. \tag{4.1}$$

If we differentiate this expression at e we get exactly the condition that D is a derivation. Conversely, if D is a derivation, then we get by induction on $n \geq 1$ that

$$D^n[X, Y] = \sum_{i=0}^{n} \frac{n!}{i!(n-i)!}[D^i X, D^{n-i} Y], \text{ for all } X, Y \in \mathfrak{g}.$$

One easily deduces that D verifies (4.1). So the Lie algebra of $\mathrm{Aut}(\mathfrak{g})$ is indeed $\mathrm{Der}(\mathfrak{g})$.

Let us now compute the derivative of Ad. First of all, recall that for all $g \in G$, $\mathrm{Ad}(g)$ is itself the derivative at e of $I(g) : h \mapsto ghg^{-1}$. Thus we know that $\exp(\mathrm{Ad}(g)X) = I(g)(\exp(X)) = g \exp(X) g^{-1}$, for all $g \in G$, $X \in \mathfrak{g}$. This easily implies that $g \mapsto \mathrm{Ad}(g)$ is smooth. Proceeding as in Proposition 2.17 we find for all $t \in \mathbb{R}$ and $X, Y \in \mathfrak{g}$,

$$\exp(\mathrm{Ad}(\exp(tX))tY) = \exp(tX)\exp(tY)\exp(-tX) = \exp(tY + t^2[X, Y] + O(t^3)).$$

Hence for t small enough, $\mathrm{Ad}(\exp(tX))tY = tY + t^2[X, Y] + O(t^3)$. This shows that the derivative at $t = 0$ of $t \mapsto \mathrm{Ad}(\exp(tX))Y = Y + t[X, Y] + O(t^2)$ is $\mathrm{ad}(X)$. So indeed, $d(\mathrm{Ad})_e = \mathrm{ad}$.

Corollary 2.22 thus implies that $\mathrm{ad}(\mathfrak{g})$ is the Lie algebra of $\mathrm{Int}(\mathfrak{g}) = \mathrm{Ad}(G)$. □

COROLLARY 2.24. *The kernel of the adjoint representation of a connected Lie group is its center.*

PROOF. If $\mathrm{Ad}(g) = e_{\mathrm{GL}(\mathfrak{g})}$, then $\mathrm{Ad}(g)X = X$ for all $X \in \mathfrak{g}$. This amounts to $\exp(\mathrm{Ad}(g)X) = \exp(X)$, and further, $I(g)(\exp(X)) = \exp(X)$, which means that g commutes with all $\exp(X)$ for $X \in \mathfrak{g}$. Since G is connected, it is generated by the image of the exponential map. So g follows in the center of G. This argument can clearly be reversed to show the converse inclusion. □

Thanks to the above fact, we may also prove a nice correspondance between Lie groups and Lie algebras in the case of trivial center.

COROLLARY 2.25. *Two connected Lie groups with trivial center are isomorphic if and only if their Lie algebras are isomorphic.*

PROOF. We already know that if two Lie groups G and H are isomorphic, then so are their Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$. Conversely, assume that $\phi : \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra isomorphism. Then $\Psi : T \in \mathrm{Aut}(\mathfrak{g}) \mapsto \phi T \phi^{-1} \in \mathrm{Aut}(\mathfrak{h})$ is a group isomorphism, and even a Lie group isomorphism. Moreover, its derivative is given by the same formula $\psi : T \in \mathrm{Der}(\mathfrak{g}) \mapsto \phi T \phi^{-1} \in \mathrm{Der}(\mathfrak{h})$. It is easily checked that the inner derivation $\mathrm{ad}(X) \in \mathrm{ad}(\mathfrak{g})$ is mapped to the derivation $\mathrm{ad}(\phi(X)) \in \mathrm{ad}(\mathfrak{h})$. So ψ maps $\mathrm{ad}(\mathfrak{g})$ to $\mathrm{ad}(\mathfrak{h})$, which shows that Ψ maps $\mathrm{Ad}(G)$ onto $\mathrm{Ad}(H)$. Moreover, since $G \simeq \mathrm{Ad}(G)$ and $H \simeq \mathrm{Ad}(H)$, we are done. □

CHAPTER 3

Algebraic groups

In this chapter we go over some basic definitions of algebraic groups, over a field of characteristic 0, but not necessarily algebraically closed (typically $\mathbb{R}$ or $\mathbb{Q}$). We will use a rather naive and incomplete approach.

We follow Benoist's presentation [**Ben08**]. There are two more standard and comprehensive references on this topic: the book of Humphreys [**Hum81**] mostly deals with algebraically closed fields but it is good to get familiar with the main concepts related to algebraic groups. The book of Borel [**Bor91**] emphasizes more the field of definition and allows non-algebraically closed fields.

1. Affine and projective algebraic k-varieties

Let us fix an algebraically closed field K and an arbitrary subfield k. We recall here the definitions of affine and projective k-varieties.

Consider a k-vector space $\mathbb{V}_k$ and put $\mathbb{V} := K \otimes_k \mathbb{V}_k$. Denote by $K[\mathbb{V}]$ the ring of regular functions on $\mathbb{V}$ over K (i.e. polynomials), and by $k[\mathbb{V}]$ the k-subring of functions that map $\mathbb{V}_k$ into k. Note that $K[\mathbb{V}] = K \otimes_k k[\mathbb{V}_k]$.

For a subset $\mathbb{X} \subset \mathbb{V}$ denote by $I(\mathbb{X}) \subset K[\mathbb{V}]$ the ideal of functions vanishing on X and by $I_k(\mathbb{X}) := I(\mathbb{X}) \cap k[\mathbb{V}]$. Conversely if $I \subset K[\mathbb{V}]$ is an ideal we denote by $V(I) \subset \mathbb{V}$ the set of common zeroes of the elements of I. Recall that the sets $V(I)$ are the closed sets of the *Zariski topology* on $\mathbb{V}$. A Zariski closed subset is also called an *affine algebraic variety*. Note that we may restrict the Zariski topology to any affine variety.

An ideal $I \subset K[\mathbb{V}]$ is said to be *defined over* k if it is spanned over K by its intersection with $k[\mathbb{V}]$. This is equivalent to $I = K \otimes_k (I \cap k[\mathbb{V}_k])$. An affine variety $\mathbb{X} \subset \mathbb{V}$ is said to be *defined over* k if its ideal $I(\mathbb{X})$ is defined over k. In contrast, $\mathbb{X}$ is said to be k-closed if it is the zero set of *some* ideal I defined over k: $\mathbb{X} = V(I)$. These two definitions do not coincide in general, but they do in characteristic 0. This fact is actually very convenient in practice.

An algebraic variety $\mathbb{X} \subset \mathbb{V}$ comes with its ring of *regular functions* $K[\mathbb{X}] = K[\mathbb{V}]/I(\mathbb{X})$. If $\mathbb{X}$ is defined over k, then the ring of k-regular functions $k[\mathbb{X}] = k[\mathbb{V}]/I_k(X)$ is such that $K[\mathbb{X}] = K \otimes_k k[\mathbb{X}]$. In this case we denote by $\mathbb{X}_k := \mathbb{X} \cap k^d$ the set of *k-points* of X. Note that the algebras $k[X]$ appearing this way are finitely generated and reduced k-algebras.

Remark 3.1. In fact an affine algebraic variety $\mathbb{X}$ is homeomorphic (w.r.t. Zariski topologies) to the set of maximal ideals of its algebra $K[\mathbb{X}]$. If it is defined over k it is not necessarily the case that the set of k-points $\mathbb{X}_k$ corresponds to maximal ideals of $k[\mathbb{X}]$. In fact a k-point $x \in \mathbb{X}$ corresponds to an ideal $I = I_k(x)$ inside $k[\mathbb{X}]$ which is maximal and such that $k[\mathbb{X}]/I = k$. In particular, the notion of k-point is independent of the algebraically closed field K that contains k. We leave this as an exercise to those that followed the algebraic geometry course in the first semester.

A *morphism* (or *rational map*) between two affine varieties $\mathbb{X}$ and $\mathbb{Y}$ is a map $\varphi : \mathbb{X} \to \mathbb{Y}$ for which composition $f \mapsto f \circ \varphi$ defines a K-algebra morphism $f_* : K[\mathbb{Y}] \to K[\mathbb{X}]$ (this sloppy formulation can easily be made rigorous). When $\mathbb{X}$ and $\mathbb{Y}$ are defined over k, we say that f is defined over k if $f_*(k[\mathbb{Y}]) \subset k[\mathbb{X}]$. Not that any k-morphism maps k-points into k-points.

A k-variety is said to be *k-irreducible* if it can not be covered by two proper k-closed subsets (i.e. subvarieties). This amounts to saying that its ring of functions $k[\mathbb{X}]$ is an integral domain. In this case we denote by $k(Z)$ its fraction field, called the field of *k-rational functions*. Using the Noetherian property of $k[\mathbb{V}]$ one can show that any affine k-variety is a finite union of irreducible subvarieties. Its (finitely many) maximal irreducible subvarieties are called the *irreducible components* of $\mathbb{X}$.

The *dimension* of an irreducible k-variety $\mathbb{X}$ is the transcendence degree of the function field $k(\mathbb{X})$ over k. Here, $k(\mathbb{X})$ is the fraction field of $k[\mathbb{X}]$.

The *tangent space* T_x of $\mathbb{X}$ at a point x is the K-vector space of point derivations on the local ring, i.e. K-linear maps $d : K[\mathbb{X}]_x \to K$ such that $d(fg) = d(f)g(x)+f(x)d(g)$ for all $f, g \in K[\mathbb{X}]_x$. Here $K[\mathbb{X}]_x$ denotes the ring of rational functions P/Q with $P, Q \in K[\mathbb{X}]$, $Q(x) \neq 0$.

Exercise 3.2. More concretely, if $\mathbb{X} \subset \mathbb{V}$, check that we may identify T_x with the subspace of $\mathbb{V}$ defined as the intersection of kernels of differentials $dP_x \in \mathbb{V}$ as P varies in $I(X)$:

$$T_x = \bigcap_{P \in I(X)} \mathrm{Ker}(dP_x).$$

Deduce that if $\mathbb{X}$ is defined over k, and $x \in \mathbb{X}_k$ is a k-point then $T_x(k) := \mathfrak{g} \cap \mathbb{V}_k$ spans linearly T_x over K (better, $T_x = K \otimes_k T_x(k)$).

Although the exercise gives a concrete way of picturing the tangent space, our initial definition is more functional. In particular, we see that any K-morphism $\varphi : \mathbb{X} \to \mathbb{Y}$ admits a differential $d\varphi_x : T_x(\mathbb{X}) \to T_{\varphi(x)}(\mathbb{Y})$ at any point $x \in X$. Also it makes it easier to define the *tangent bundle* of $\mathbb{X}$ as the space of derivations of $K[\mathbb{X}]$, but we won't discuss its bundle structure precisely.

A point in an irreducible variety $\mathbb{X}$ is said to be *smooth* if its tangent space has minimal dimension. We admit a few facts:

(1) The set of smooth points of a variety $\mathbb{X}$ is then a non-empty Zariski open set (which is defined over k if $\mathbb{X}$ is).
(2) If $x \in \mathbb{X}$ is smooth then $\dim(T_x) = \dim(X)$.
(3) In this case $\mathbb{X} \subset \mathbb{V}$ coincides on a neighborhood of x with the zero set of polynomials $P_i \in I(\mathbb{X})$, $1 \leq i \leq \dim(\mathbb{V}) - \dim(\mathbb{X})$ such that $T_x = \cap_i \mathrm{Ker}(dP_i)$.

The variety $\mathbb{X}$ is said to be *smooth* if all of its points are smooth.

In the case where $k = \mathbb{R}$ or $\mathbb{C}$, the set of k-points of a smooth affine k-variety $\mathbb{X} \subset \mathbb{V}$ is an analytic submanifold of $\mathbb{V}_k$ with dimension $\dim(\mathbb{X})$ and the set of k-points of its algebraic tangent space is equal to its analytic tangent space.

More generally, one may define k-schemes and general k-varieties by gluing affine k-varieties. We won't need the full generality here, but only the case of projective varieties. We bypass the gluing procedure and define them as follows. A subset of the projective space $\mathbb{P}(\mathbb{V})$ is said to be Zariski closed if it is the set of common zero of a family of *homogeneous* polynomials P_i on $\mathbb{V}$. Note that such zero set is indeed well defined in the

projective space thanks to homogeneity. A Zariski closed set is also called a *projective variety*, and it is said to be k-closed if the polynomials P_i have coefficients in k. Since we only care about characteristic 0 we also say in this case that the variety is defined over k, or is a k-variety. The set of k-points of a k-variety $\mathbb{Z}$ is by definition the set $\mathbb{Z} \cap \mathbb{P}(\mathbb{V}_k)$. One can in this context also define the notion of a regular function on a projective variety, a k-morphism between two projective varieties, etc. This can all be made consistent with the affine case.

We finally define a *quasi-projective* k-variety to be a k-open subset of a projective k-variety. For example, the process of homogenization of a polynomial shows that any affine k-variety is quasi-projective.

This algebraic setting has to advantages. First it allows to keep track of fields of definition, while at the same time offering the possibility to work with points in the algebraic closure (Note however that a k-variety may have no k-point at all). The second advantage is that the image of an algebraic variety by a morphism needs not be closed again an algebraic variety, but it still behaves well:

THEOREM 3.3 (Chevalley). *Let $\varphi : \mathbb{X} \to \mathbb{Y}$ be a morphism between two algebraic K-varieties. Then the range $\varphi(\mathbb{X})$ contains an open subset of its closure, for the restriction of the Zariski topology on $\mathbb{Y}$.*

Caution: we are only dealing with K varieties. There is no analogous result for k-points in general. For example, the map $x \in \mathbb{R} \mapsto x^2 \in \mathbb{R}$ does not map $\mathbb{R}$ onto a Zariski open subset of $\mathbb{R}$.

The proof relies on the following lemma in commutative algebra.

LEMMA 3.4. *Consider two K-algebras $A \subset B$ such that B is a finitely generated A-algebra and an integral domain. Then for every non-zero element $b \in B$ there exists a non-zero element $a \in A$ such that every morphism $\phi : A \to K$ satisfying $\phi(a) \neq 0$ extends to a morphism $\widetilde{\phi} : B \to K$ satisfying $\widetilde{\phi}(b) \neq 0$.*

PROOF. We may proceed by induction on the number of generators of B over A, hence reducing to the case where B is generated by one element x. So B is a quotient of $A[T]$. If $B = A[T]$ the result is easy. If B is a proper quotient of $A[T]$, take $P \in A[T]$ to be a polynomial of minimal degree d such that $P(x) = 0$. Denote also by L the fraction field of A and observe that the ideal $I := \{Q \in L[T] \mid Q(x) = 0\}$ is generated by P. We thus have a commutative diagram:

$$\begin{array}{ccc} L[T] & \to & L[T]/I \\ \cup & & \cup \\ A[T] & \to & B \end{array}$$

In particular, the element $b \in B$ may be viewed inside $L[T]/I$ as the image of a polynomial $Q_0 \in L[T]$ of degree at most $d-1$. Multiplying Q_0 by an appropriate element of A we get a polynomial $Q \in A[T]$ such that b divides $Q(x)$ inside B and still $\deg(Q) \leq d-1$.

Note that P is irreducible over $L[\mathbb{X}]$ because B is an integral domain. Since we are in characteristic 0 we deduce that P and P' are co-prime. Thus we may find $r \in A$ and $R_1, R_2 \in A[T]$ such that $r = R_1 P + R_2 P'$. Take for q a non-zero coefficient of the polynomial Q, for p_d the coefficient of T^d in P and set $a := r p_d q$.

Take a morphism $\phi : A \to K$ such that $\phi(a) \neq 0$, and denote by ϕ^+ its extension $A[T] \to K[T]$. Since $\phi(a) \neq 0$, we see that $\phi^+(Q) \neq 0$. Moreover, the polynomial $\phi^+(P) \in K[T]$ has degree d, so it admits d roots in K. Since moreover $\phi(r) \neq 0$, we see

that these roots are distinct in K. Thus there exists one of these roots $\lambda \in K$ which is not a root of $\phi^+(Q)$ (because $\phi^+(Q)$ has degree at most $d-1$). Define now $\phi_1 : A[T] \to K$ to be the composition of $\phi^+ : A[T] \to K[T]$ with the evaluation at λ. We clearly see that $\phi_1(P) = \phi^+(P)(\lambda) = 0$.

Claim. ϕ_1 vanishes on $\{R \in A[T] \mid R(x) = 0\}$.

Here there is a minor subtlety: if we knew that ϕ_1 could be extended to $L[T]$ then the condition $\phi_1(P) = 0$ would imply the claim. But ϕ_1 may have zeroes on A. So ϕ_1 needs not extend to $L[T]$. But the claim follows easily from the following fact: if $R(x) = 0$ and $R \in A[T]$, then we may find an integer $n \geq 0$ and $R_0 \in A[T]$ such that $R = R_0 P/p_d^n$. Since $\phi(p_d) \neq 0$, we then get $\phi_1(R) = \phi_1(R_0 P)/\phi(p_d)^n = 0$ proving the claim. We leave the proof of the fact as an exercise.

Thanks to the claim, ϕ_1 then defines a morphism $\widetilde{\phi}$ on B. Since λ is not a root of $\phi(Q)$ we deduce that $\phi_1(Q) \neq 0$ and hence $\widetilde{\phi}(Q(x)) \neq 0$, which implies that $\widetilde{\phi}(b) \neq 0$, as desired. □

In fact we won't use the element b in the proof of Chevalley's Theorem, we will only use the fact that there exists $a \in A$ such that any morphism $\phi : A \to K$ that does not vanish on a extends to a morphism of B. (Observe that the element b was used in the proof of the lemma).

Proof of Theorem 3.3. We may assume that both $\mathbb{X}$ and $\mathbb{Y}$ are affine irreducible K-varieties, and that $\varphi(\mathbb{X})$ is Zariski dense in $\mathbb{Y}$. Since the range of φ is Zariski dense in $\mathbb{Y}$, the morphism $\varphi_* : K[\mathbb{Y}] \to K[\mathbb{X}]$ is injective. The above lemma applies to $A = K[\mathbb{Y}]$ and $B = K[\mathbb{X}]$. So we may find a function $a \in A = K[\mathbb{Y}]$ such that any morphism $\phi : A \to K$ that does not vanish on a extends to a morphism on B. But note that morphisms $\phi : A \to K$ are exactly evaluation maps at points of $\mathbb{Y}$. Saying that a morphism extends to B then amounts to saying that the point is in the range of φ. So we readily see that the set of points $y \in \mathbb{Y}$ such that $a(y) \neq 0$ is contained in the range of φ. This set is clearly an open set. □

In fact, a similar proof with a little more commutative algebra shows the following.

Corollary 3.5. *Assume that K has characteristic 0. Let $\phi : \mathbb{X} \to \mathbb{Y}$ be a morphism between two K-varieties, whose range is Zariski dense in $\mathbb{Y}$. Then there is a Zariski open subset U of $\mathbb{X}$ such that for all $x \in U$, the derivative $d\phi_x : T_x(\mathbb{X}) \to T_{\phi(x)}(\mathbb{Y})$ is onto.*

We omit the proof.

2. Algebraic groups and their Lie algebras

Keep the notation $k \subset K$ as before, and we assume moreover that they have characteristic 0.

Definition 3.6. An (affine) *algebraic group* over k is an affine algebraic variety over k whose set of K-points $\mathbb{G}$ is a group such that the product and inverse maps $(x, y) \mapsto xy$ and $x \mapsto x^{-1}$ are k-rational.

It is a fact that in a k-algebraic group $\mathbb{G}$, the neutral element e is always a k-point. Indeed $\{e\}$ is the range of $\mathbb{G}$ by the k-rational map $x \mapsto xx^{-1}$ and so it is defined over k (thanks to Galois theoretic considerations).

A *k-morphism* between two k-algebraic groups $\mathbb{G}$ and $\mathbb{H}$ is a k-morphism of k-varieties which is also a group homomorphism. The morphism is called a *k-isogeny* if it is surjective and has finite kernel. We warn that the surjectivity assumption is only made on the K-points, but needs not happen on the sets of k-points.

Example 3.7. Most Lie groups that we have encountered are algebraic groups. For example $\mathrm{GL}(d, K)$ is an algebraic groups defined over the prime field of K. We will see that any algebraic group as defined above is in fact isomorphic with a subgroup of $\mathrm{GL}(d, K)$. Let us describe such a representation for the two one dimensional algebraic groups. The *additive group* $\mathbb{G}_a := (K, +)$ is realized as a linear group via the embedding

$$x \in K \mapsto \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}.$$

The *multiplicative group* $\mathbb{G}_m := (K^*, \times)$ is identified with the linear group

$$\{\begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix} \mid xy = 1\}.$$

These two groups are defined over the prime field of K (over $\mathbb{Q}$ in our case). They are not isomorphic as algebraic varieties (since $k[\mathbb{G}_a] = k[X]$ while $k[\mathbb{G}_m] = k[X, X^{-1}]$) so nor as algebraic groups.

A *k-character* on a k-group $\mathbb{G}$ is a k-morphism $\mathbb{G} \to \mathbb{G}_m$ and a *k-cocharacter* is a k-morphism $\mathbb{G}_m \to \mathbb{G}$.

A *k-representation* is a k-morphism $\mathbb{G} \to \mathrm{GL}(\mathbb{W})$, where $\mathbb{W} = K \otimes_k \mathbb{W}_k$ and $\mathbb{W}_k$ is a k-vector space, and the k-structure on $\mathrm{GL}(\mathbb{W})$ is inherited from the one on $\mathrm{End}(\mathbb{W})$ as before. So a morphism $\pi : \mathbb{G} \to \mathrm{GL}(\mathbb{W})$ is a k-representation if and only if the coefficient functions $g \mapsto \pi(g)_{i,j}$ in a basis of $\mathbb{W}_k$ are k-rational functions on $\mathbb{G}$.

Proposition 3.8. *Consider a morphism $\varphi : \mathbb{G} \to \mathbb{H}$ between two algebraic groups. Then the range $\varphi(\mathbb{G})$ is closed inside $\mathbb{H}$.*

Proof. We may replace $\mathbb{H}$ by the closure of $\varphi(\mathbb{G})$. In this case Theorem 3.3 implies that $\varphi(\mathbb{G})$ is a subgroup of $\mathbb{H}$ that contains an open subset. By homogeneity, it is thus a Zariski open subgroup. Since an open subgroup is also closed, we conclude that $\varphi(\mathbb{G})$ is both closed and dense in $\mathbb{H}$, showing equality. □

Remark 3.9. There is a subtlety about the definition of algebraic groups. In general, an algebraic group is **not** a topological group for its Zariski topology. Namely the Zariski topology on the product $\mathbb{G} \times \mathbb{G}$ is not the product of the Zariski topology on $\mathbb{G}$ by itself in the usual topological sense. So the product map $\mathbb{G} \times \mathbb{G} \to \mathbb{G}$ is not continuous for the product topology in general. However, the translation functions $\lambda_g : x \in \mathbb{G} \mapsto gx$ are continuous, which ensures that the above proof is correct (check it).

We now prove that any affine group is linear. Let us start with some notation: the *left and right regular representations* λ and ρ of $\mathbb{G}$ on its function algebra $K[\mathbb{G}]$ are defined by $\lambda_x(f)(y) = f(x^{-1}y)$ and $\rho_x(f)(y) = f(yx)$ for all $x, y \in \mathbb{G}$, $f \in K[\mathbb{G}]$.

Lemma 3.10. *Fix a k-group $\mathbb{G}$. Given a finite dimensional vector subspace $E \subset K[\mathbb{G}]$, there exists a finite dimensional subspace $F \subset K[\mathbb{G}]$ defined over k, which contains E and which is globally invariant under the right regular representation. In particular, $k[\mathbb{G}]$ is an increasing union of finite dimensional ρ-invariant vector spaces.*

PROOF. We may enlarge E if needed and assume that E is defined over k. In this case it suffices to verify it for 1-dimensional k-subspaces of E. So without loss of generality, we may assume that E is one dimensional, spanned by an element $f \in k[\mathbb{G}]$.

By definition of products of algebraic varieties, the algebra of k-regular functions on the variety $\mathbb{G} \times \mathbb{G}$ is $k[\mathbb{G}] \otimes_k k[\mathbb{G}]$. The product map induces a map $\mu_* : k[\mathbb{G}] \to k[\mathbb{G}] \otimes_k k[\mathbb{G}]$. Write $\mu_*(f) = \sum_i a_i \otimes b_i$ with $a_i, b_i \in k[\mathbb{G}]$. Then for all $g, h \in \mathbb{G}$, we have $\rho_g(f)(h) = f(hg) = \mu_*(f)(h, g) = \sum_i a_i(h)b_i(g)$. Hence, $\rho_g(f) = \sum_i b_i(g)a_i \in F$ and we see that $\rho(G)f$ is contained in the linear span of the a_i's. So we have found a finite dimensional vector space defined over k which contains $\rho(G)E$. The intersection of all such vector spaces is then globally invariant and still defined over k. □

COROLLARY 3.11. *Any k-group $\mathbb{G}$ is isomorphic to a k-subgroup of* $\mathrm{GL}(d, K)$.

PROOF. Denote by $E \subset K[\mathbb{G}]$ a finite generating set of $K[\mathbb{G}]$ and by $F \subset K[\mathbb{G}]$ a finite dimensional vector space which is defined over k and globally $\rho(\mathbb{G})$-invariant and which contains E. Let us check that the restriction of ρ to F is a k-rational representation π of $\mathbb{G}$.

Since F is $\rho(\mathbb{G})$-invariant, we have that $\mu_*(F) \subset F \otimes_K K[\mathbb{G}]$. Take a basis $f_1, \ldots, f_n \in k[\mathbb{G}]$ of F. Then we may find functions $m_{i,j} \in k[\mathbb{G}]$ such that $\mu_*(f_i) = \sum_j f_j \otimes m_{i,j}$. Then we deduce that

$$\rho_g(f_i) = \sum_j m_{i,j}(g) f_j.$$

Denoting by F_k the k-span of $f_1, \ldots, f_n$, we get a k-structure $\mathrm{End}_k(F_k)$ in $\mathrm{End}(F)$. Saying that π is a k-rational representation into $\mathrm{GL}(F)$ is exactly saying that $m_{i,j}$ are k-rational functions on $\mathbb{G}$.

The kernel of π is trivial, because if $\rho(g)$ acts trivially on $E \subset F$ then it acts trivially on the whole $K[\mathbb{G}]$. This is easily seen to imply that $g = e$. So the representation is injective. By Proposition 3.8 we know that the range $\pi(\mathbb{G})$ is a closed subgroup of $\mathrm{GL}(d, K)$. It is in fact k-closed because it is globally invariant under the Galois group of K over k. □

REMARK 3.12. In the above proof, we admit that a bijective algebraic morphism between two smooth K-varieties (e.g. algebraic groups) is in fact an isomorphism.

DEFINITION 3.13. The *Lie algebra* $\mathfrak{g}$ of a k-group $\mathbb{G}$ is the vector space of left invariant derivations on $K[\mathbb{G}]$, i.e. linear maps $D : K[\mathbb{G}] \to K[\mathbb{G}]$ such that $D(fg) = fD(g) + gD(f)$ and $D\lambda_x = \lambda_x D$ for all $f, g \in K[\mathbb{G}]$ and $x \in \mathbb{G}$. The Lie bracket on $\mathfrak{g}$ is the expected one: $[D_1, D_2] = D_1D_2 - D_2D_1$.

One can prove that the map $D \in \mathfrak{g} \mapsto D_e \in T_e(\mathbb{G})$ is a linear isomorphism. So the Lie algebra of $\mathbb{G}$ is naturally identified with its tangent space at the neutral element (as vector spaces). If $\mathbb{G}$ is defined over k, we denote by $\mathfrak{g}_k$ the set of k-points of $\mathfrak{g}$, i.e. the subset of derivations that map $k[\mathbb{G}]$ into itself. As we observed, when $k = \mathbb{R}$ or $\mathbb{C}$, $\mathfrak{g}_k$ is equal to the Lie algebra of the Lie group $\mathbb{G}_k$.

Note that for any ring $A \subset K$, the set $\mathbb{G}_A := \mathbb{G} \cap \mathrm{GL}(d, A)$ is again a group. This notation is consistent with the notation $\mathbb{G}_k$ for the set of k-points.

3. Actions of algebraic groups

DEFINITION 3.14. Fix a k-group $\mathbb{G}$ and a k-variety $\mathbb{X}$. An algebraic action of $\mathbb{G}$ on $\mathbb{X}$ is a group action of $\mathbb{G}$ on $\mathbb{X}$ such that the map $(g, x) \in \mathbb{G} \times \mathbb{X} \mapsto g \cdot x \in \mathbb{X}$ is rational. If this map is k-rational, we talk about a k-action.

From Theorem 3.3 we deduce the following result that generalizes the case of group homomorphisms.

PROPOSITION 3.15. *The orbits of an algebraic action are open in their closure.*

Let us now prove another result of Chevalley about homogeneous spaces. In particular the result gives to any homogeneous space $\mathbb{G}/\mathbb{H}$ the structure of an algebraic variety.

THEOREM 3.16 (Chevalley). *Take a k-group $\mathbb{G}$ and a k-subgroup $\mathbb{H} < \mathbb{G}$. Then there exists a k-representation of $\mathbb{G}$ on a vector space $\mathbb{V} = K \otimes_k \mathbb{V}_k$ and an element $v \in \mathbb{V}_k$ such that the stabilizer in $\mathbb{G}$ of the line Kv is exactly $\mathbb{H}$, i.e. $\mathbb{H} = \{g \in \mathbb{G} \mid gv \in Kv\}$.*

PROOF. Denote by $I := \{f \in K[\mathbb{G}] \mid f_{|\mathbb{H}} = 0\}$ and by $I_k := I \cap k[\mathbb{G}]$ the ideals corresponding to $\mathbb{H}$. By Lemma 3.10 we may find a finite dimensional subspace $F \subset K[\mathbb{G}]$ defined over k, which is invariant under $\rho(\mathbb{G})$ and such that $F \cap I_k$ generates I as an ideal inside $K[\mathbb{G}]$.

As in Corollary 3.11, the restriction of ρ to F is a representation π of $\mathbb{G}$ defined over k. Moreover, for $g \in \mathbb{G}$, $\pi(g)$ preserves $F \cap I$ if and only if $\rho(g)$ preserves I. The latter is equivalent to $g \in \mathbb{H}$.

Denote by $d := \dim(F \cap I)$ and consider the representation $\widetilde{\pi}$ of $\mathbb{G}$ on $\mathbb{V} := \bigwedge_d F$ defined by $\widetilde{\pi}(g)v_1 \wedge \cdots \wedge v_d = (\pi(g)v_1) \wedge \cdots \wedge (\pi(g)v_d)$ for $g \in \mathbb{G}$, $v_i \in F$. This representation is again algebraic, and defined over k, for the k-form $\mathbb{V}_k := \bigwedge_d(F \cap k[\mathbb{G}]) \subset \mathbb{V}$. Then for a basis $v_1, \ldots, v_k$ of $F \cap I_k$, we see that the k-point $v = v_1 \wedge \cdots \wedge v_k$ is a k-point in $\mathbb{V}$. Moreover, one checks that $\widetilde{\pi}(g)v \in Kv$ if and only if $g(F \cap I) \subset F \cap I$. We have seen that this amounts to $g \in \mathbb{H}$. □

The above result gives to $\mathbb{G}/\mathbb{H}$ the structure of a quasi-projective k-variety. Indeed, denote by $x = Kv \in \mathbb{P}(\mathbb{V})$. Then Proposition 3.15 shows that the orbit $\mathbb{G}x \subset \mathbb{P}(\mathbb{V})$ is open in its closure, so it is a quasi-projective variety which is identified as a set with $\mathbb{G}/\mathbb{H}$. Because this orbit is invariant under the Galois group of K over k it is actually a quasi-projective k-variety. Observe also that the action of $\mathbb{G}$ on $\mathbb{G}/\mathbb{H}$ is k-rational.

PROPOSITION 3.17. *Let $k = \mathbb{R}$ or $\mathbb{C}$, and let $\mathbb{G}$ be a k-group.*

1. *If $\mathbb{G}$ is Zariski connected, the group $\mathbb{G}_k$ is Zariski dense inside $\mathbb{G}$.*
2. *Take a k-action of $\mathbb{G}$ on a k-variety $\mathbb{X}$. Then for all $v \in \mathbb{X}_k$, the $\mathbb{G}_k$-orbits inside $(\mathbb{G}v)_k = \mathbb{G}v \cap \mathbb{X}_k$ are open and closed for the analytic topology.*

PROOF. (1) $\mathbb{G}_k$ is a Lie group over k, whose Lie algebra is $\mathfrak{g}_k$. In particular the Zariski closure of $\mathbb{G}_k$ is a k-subgroup $\mathbb{H}$ of $\mathbb{G}$ whose Lie algebra contains $\mathfrak{g}$. Since $\mathbb{G}$ is connected, this forces $\mathbb{G} \subset \mathbb{H}$, and hence $\mathbb{H} = \mathbb{G}$.

(2) We may assume that $\mathbb{X} = \mathbb{G}v$, in which case, by homogeneity, every point of $\mathbb{X}$ is smooth. So $\mathbb{G}_k$ and $\mathbb{X}_k$ are analytic k-manifolds whose tangent spaces identify with the k-points of their algebraic tangent spaces. The orbit map $g \in \mathbb{G} \mapsto gv \in \mathbb{X}$ is onto, so its differential is again onto on a non-empty Zariski open subset of $\mathbb{G}$ by Corollary 3.5. By $\mathbb{G}$-invariance, the differential has to be onto at every point of $\mathbb{G}$, and hence at e. So the differential $d\rho_e : T_e(\mathbb{G}) \to T_v(\mathbb{X})$ is onto and maps k-points onto k-points. By the normal form theorem for sub-immersions, we know that ρ_e has open range inside $\mathbb{X}_k$. This argument shows that in fact all the $\mathbb{G}_k$ orbits are open in $\mathbb{X}_k$. Thus they are also closed. □

4. Jordan decomposition

We show here that the Jordan decomposition for matrices actually fits in the framework of algebraic groups. We continue to denote by k a subfield of an algebraically closed field K of characteristic 0.

Recall that an endomorphism of a k vector space $\mathbb{V}_k$ is called *semi-simple* if it is diagonalisable over K, i.e. as an endomorphism of $\mathbb{V} = K \otimes_k \mathbb{V}_k$. An endomorphism u of $\mathbb{V}_k$ is *unipotent* if its only eigenvalue over K is 1. This amounts to require that $u - 1$ is nilpotent.

Recall that there are two versions of the Jordan decomposition.

- The *additive* version states that any endomorphism X of $\mathbb{V}_k$ can be decomposed as the sum of a semi-simple endomorphism $S \in \mathrm{End}(\mathbb{V}_k)$ and a nilpotent endomorphism $N \in \mathrm{End}(\mathbb{V}_k)$ such that $[S, N] = 0$. This decomposition is unique and S and N are expressed as polynomials in X (with coefficients in k).
- The *multiplicative* version states that any invertible endomorphism $g \in \mathrm{GL}(\mathbb{V}_k)$ writes as the product of a semi-simple endomorphism $s \in \mathrm{GL}(\mathbb{V}_k)$ and a unipotent endomorphism $u \in \mathrm{GL}(\mathbb{V}_k)$ such that $[s, u] = 1$. This decomposition is unique and s and u are expressed as polynomials in g.

To pass from one decomposition to the other observe that if $g \in \mathrm{GL}(\mathbb{V}_k)$ and if $g = S + N$ is its additive Jordan decomposition then S is invertible. Hence we may set $s := S$ and $u := 1 + S^{-1}N$, so that $g = su$.

Proposition 3.18. *Let $\mathbb{V}_k$ be a k-vector space, set $\mathbb{V} = K \otimes_k \mathbb{V}_k$ and take a k-subgroup $\mathbb{G} \subset \mathrm{GL}(\mathbb{V})$. Take $g \in \mathbb{G}$ and write $g = su$ for its Jordan decomposition in $\mathrm{GL}(\mathbb{V})$. We have:*

(1) Any subspace $\mathbb{W} \subset \mathbb{V}$ which is g-invariant is also s and u-invariant.
(2) We have $s, u \in \mathbb{G}$.
(3) If $g \in \mathbb{G}_k$ then $s, u \in \mathbb{G}_k$.
(4) g is semi-simple (resp. unipotent) if and only if ρ_g is semi-simple (resp. unipotent) on $K[\mathbb{G}]$, in the sense that its restriction to any finite dimensional $\mathbb{G}$-invariant subspace of $K[\mathbb{G}]$ is semi-simple (resp. unipotent).

Proof. (1) This is obvious since s and u are expressed as polynomials in g.
(2) For $m \geq 1$ denote by π_m the linear representation of $\mathrm{GL}(\mathbb{V})$ on $K_m[\mathrm{End}(\mathbb{V})] := \{P \in K[\mathrm{End}(\mathbb{V})] \mid \deg(P) \leq m\}$ via $\pi_m(g)(P) : x \mapsto P(xg)$. One checks that if g is semi-simple or unipotent on $\mathbb{V}$, then so is $\pi_m(g)$ on $K_m[\mathrm{End}(\mathbb{V})]$. In particular $\pi_m(g) = \pi_m(s)\pi_m(u)$ is the Jordan decomposition of $\pi_m(g)$ on $K_m[\mathrm{End}(\mathbb{V})]$.

Now denote by $I(\mathbb{G}) \subset K[\mathrm{End}(\mathbb{V})]$ the ideal corresponding to $\mathbb{G}$, and set $I_m(\mathbb{G}) := I(\mathbb{G}) \cap K_m[\mathrm{End}(\mathbb{V})]$. Assume that we chose m such that $I_m(\mathbb{G})$ generates $I(\mathbb{G})$ as an ideal of $K[\mathrm{End}(\mathbb{V})]$. We see that an element $g \in \mathrm{GL}(\mathbb{V})$ belongs to $\mathbb{G}$ if and only if $\pi_m(g)$ leaves $I_m(\mathbb{G})$ invariant. So the statement follows from item (1).
(3) We saw above that this fact is true for the ambient group $\mathrm{GL}(\mathbb{V})$ so it remains true for $\mathbb{G}$ since $\mathbb{G}_k = \mathbb{G} \cap \mathrm{GL}(\mathbb{V}_k)$.
(4) If g is semi-simple then we saw in the proof of (2) that $\pi_m(g)$ is a semi-simple endomorphism of $K_m[\mathrm{End}(\mathbb{V})]$, which preserves $I_m(\mathbb{G})$, so the quotient representation ρ_m on $K_m[\mathbb{G}] := K_m[\mathrm{End}(\mathbb{V})]/I_m(\mathbb{G})$ is such that $\rho_m(g)$ is semi-simple as well. But note that $K_m[\mathbb{G}]$ are finite dimensional $\mathbb{G}$-invariant subspaces of $K[\mathbb{G}]$ which exhaust $K[\mathbb{G}]$ as m

goes to infinity, and ρ_m is the restriction of ρ. So indeed ρ_g is semi-simple on $K[\mathbb{G}]$. The same argument also shows that if g is unipotent, then ρ_g is unipotent on $K[\mathbb{G}]$.

Conversely, assume that ρ_g is a semi-simple endomorphism of $K[\mathbb{G}]$. Write $g = su$ for its Jordan decomposition, $s, u \in \mathbb{G}$. Since g and s commute and ρ_g and ρ_s are semi simple endomorphisms of $K[\mathbb{G}]$ this is also the case of ρ_u. And at the same time we know that ρ_u is unipotent, so we conclude that $\rho_u = \mathrm{id}$ on $K[\mathbb{G}]$. This implies that $u = 1$, and thus g is semi-simple. The same argument also shows that if ρ_g is unipotent then so is g. □

Remark 3.19. Observe that when $k = \mathbb{R}$ and $g \in \mathbb{G}_{\mathbb{R}} \subset \mathrm{GL}(\mathbb{V})$ is a semi-simple element whose eigenvalues are real and positive, then we may write $g = \exp(X)$ for some endomorphism X of $\mathbb{V}$ with real eigenvalues. Then we see that for all $t \in \mathbb{R}$, the element $g^t = \exp(tX)$ is also in $\mathbb{G}_{\mathbb{R}}$, because $\pi_m(g^t)$ has the same eigenspaces as $\pi_m(g)$, and hence also preserves $I_m(\mathbb{G})$.

Corollary 3.20. *If $\mathbb{G}$ and $\mathbb{H}$ are two algebraic groups and $\phi : \mathbb{G} \to \mathbb{H}$ is a rational morphism then ϕ maps semi-simple (resp. unipotent) elements of $\mathbb{G}$ to semi-simple (resp. unipotent) elements in $\mathbb{H}$. In particular, the Jordan decomposition makes sense in $\mathbb{G}$ independently of a rational embedding $\mathbb{G} \subset \mathrm{GL}(\mathbb{V})$.*

Proof. Take two rational embeddings $\mathbb{G} \subset \mathrm{GL}(\mathbb{V})$, $\mathbb{H} \subset \mathrm{GL}(\mathbb{W})$. So when we say semi-simple or unipotent it will be with respect to these representations.

By Proposition 3.8, we know that the range of ϕ is a closed subgroup of $\mathbb{H}$. So we may replace $\mathbb{H}$ by this subgroup and assume without loss of generality that ϕ is surjective. The algebra homomorphism $\phi_* : K[\mathbb{H}] \to K[\mathbb{G}]$ is then an embedding, which intertwines the right regular representations, i.e. $\rho_g \phi_* = \phi_* \rho_{\phi(g)}$ for all $g \in G$. This implies easily that if ρ_g is semi-simple or unipotent on $K[\mathbb{G}]$ then so is $\rho_{\phi(g)}$ on $K[\mathbb{H}]$. So the first part of the corollary follows from Proposition 3.18, (4).

In particular we deduce that the notions of semi-simple and unipotent elements in $\mathbb{G}$ don't depend on the choice of a rational embedding $\mathbb{G} \subset \mathrm{GL}(\mathbb{V})$. Moreover if $g = su$ is the Jordan decomposition of $g \in \mathbb{G}$, then it is clear that $\phi(g) = \phi(s)\phi(u)$ is the Jordan decomposition of $\phi(g)$. □

Exercise 3.21. Deduce from the above corollary that there is no rational morphism from $\mathbb{G}_m$ into $\mathbb{G}_a$, nor from $\mathbb{G}_a$ into $\mathbb{G}_m$. Give also a direct proof of this fact.

Part 2

Structure of Lie algebras, semi-simple groups

CHAPTER 4

General structure of Lie algebras

In this section, we study Lie algebras over a field k of characteristic 0. We will also sometimes consider an algebraically closed field K containing K (hence K also has characteristic 0). The presentation is inspired from Yves Benoist's notes (Section 2 in [**Ben08**]).

1. Nilpotent and solvable Lie algebras

An *ideal* in a Lie algebra $\mathfrak{g}$ is a vector subspace $\mathfrak{a} \subset \mathfrak{g}$ such that $[\mathfrak{a}, \mathfrak{g}] \subset \mathfrak{a}$. It is in particular a Lie subalgebra of $\mathfrak{g}$. A Lie algebra is called *abelian* if its bracket is identically 0.

DEFINITION 4.1. A k-Lie algebra $\mathfrak{g}$ is said to be *nilpotent* (resp. *solvable*) if there exists an increasing family of ideals $\mathfrak{g} = \mathfrak{g}_0 \supset \cdots \supset \mathfrak{g}_n = \{0\}$ such that $[\mathfrak{g}, \mathfrak{g}_i] \subset \mathfrak{g}_{i+1}$ (resp. $[\mathfrak{g}_i, \mathfrak{g}_i] \subset \mathfrak{g}_{i+1}$) for all $0 \leq i \leq n-1$.

Exactly as is done with groups one checks that $\mathfrak{g}$ is nilpotent (resp. solvable) if and only if the non-increasing sequence of ideals $\mathfrak{g}_i$ defined by $\mathfrak{g}_0 = \mathfrak{g}$ and $\mathfrak{g}_{i+1} = [\mathfrak{g}, \mathfrak{g}_i]$ (resp. $\mathfrak{g}_{i+1} = [\mathfrak{g}_i, \mathfrak{g}_i]$) vanishes after some finite i.

It is obvious that a nilpotent Lie algebra is always solvable, but the converse in not true in general.

EXAMPLE 4.2. The prototypical examples are the following ones.

- The Lie algebra $\mathfrak{a}_d$ of diagonal matrices in $M_d(k)$ is abelian;
- The Lie algebra $\mathfrak{u}_d^+$ of (strictly) upper triangular matrices in $M_d(k)$ is nilpotent;
- The Lie algebra $\mathfrak{p}_d^+ = \mathfrak{a}_d + \mathfrak{u}_d^+$ of all upper triangular matrices is solvable.

We now present several results of Lie and Engel that relate all solvable and nilpotent Lie algebras to the above examples. First observe that every Lie algebra is represented on a vector space, thanks to the adjoint representation: $\mathrm{ad} : \mathfrak{g} \to \mathrm{Der}(\mathfrak{g}) \subset \mathcal{L}(\mathfrak{g})$. The representation needs not be faithful, since we observed that its kernel is precisely the center of $\mathfrak{g}$.

For nilpotent Lie algebras, we have the following two results.

THEOREM 4.3. *A Lie algebra $\mathfrak{g}$ is nilpotent if and only if $\mathrm{ad}(X)$ is nilpotent as an endomorphism of $\mathfrak{g}$ (meaning that there exists some k, possibly depending on X such that $\mathrm{ad}(X)^k = 0$).*

THEOREM 4.4 (Engel). *Let V be a k-vector space and consider a Lie subalgebra $\mathfrak{g} \subset \mathcal{L}(V)$ for which every element is nilpotent. Then there exists a basis of V for which $\mathfrak{g} \subset \mathfrak{u}_d^+$, where $d = \dim(V)$. (In particular $\mathfrak{g}$ is nilpotent.)*

REMARK 4.5. Note that the Lie algebra $\mathfrak{g} = \mathrm{span}_k \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is abelian and hence nilpotent. However its elements are not nilpotent and it cannot be represented inside $\mathfrak{u}_2^+$ in any basis of $\mathbb{C}^2$. This does not contradict the above two results.

The proof of the two theorems is based on the following lemma.

LEMMA 4.6. *Let $V \neq \{0\}$ be a k-vector space and consider a Lie subalgebra $\mathfrak{g} \subset \mathcal{L}(V)$ for which every element is nilpotent. Then there exists a non-zero vector $v \in V$ such that $X(v) = 0$ for all $X \in \mathfrak{g}$.*

PROOF. We prove by induction on the dimension of the Lie algebra $\mathfrak{g}$ that for any embedding of $\mathfrak{g}$ into some $\mathcal{L}(V)$ with $V \neq \{0\}$ such that every element of $\mathfrak{g}$ is a nilpotent endomorphism of V, there exists a non-zero vector $v \in V$ such that $X(v) = 0$ for all $X \in \mathfrak{g}$.

If $\mathfrak{g}$ has dimension 1, it is spanned by a single element X. If X acts as a nilpotent element on a vector space there clearly it has a non-trivial kernel. So in this case the statement is obvious.

Fix now a Lie algebra $\mathfrak{g}$ and an embedding $\mathfrak{g} \subset \mathcal{L}(V)$, $V \neq 0$ as in our assumptions, and assume that the result holds true for all Lie algebra with dimension less that $\dim(\mathfrak{g})$. Take $\mathfrak{h} \subset \mathfrak{g}$ a proper Lie subalgebra of $\mathfrak{g}$ with maximal dimension. (Observe that any Lie algebra $\mathfrak{g}$ with dimension at least two always admits a non-zero proper Lie subalgebra: any $\mathrm{span}_k(X)$ for $X \in \mathfrak{g}$ is indeed a non-zero Lie subalgebra with dimension 1).

Claim. $\mathfrak{h}$ is an ideal in $\mathfrak{g}$ with co-dimension 1.

Consider the adjoint representation $\mathrm{ad} : \mathfrak{g} \to \mathcal{L}(\mathfrak{g})$. Being a Lie subalgebra, $\mathfrak{h}$ preserves the subspace $\mathfrak{h} \subset \mathfrak{g}$. So we may consider the quotient representation $\rho : \mathfrak{h} \to \mathcal{L}(\mathfrak{g}/\mathfrak{h})$ given by $\rho(X)(Y + \mathfrak{h}) = [X, Y] + \mathfrak{h}$, for all $X \in \mathfrak{h}$, $Y \in \mathfrak{g}$. Since $\dim \rho(\mathfrak{h}) < \dim(\mathfrak{g})$, we may apply our induction assumption to $\rho(\mathfrak{h})$ and get a vector $Y + \mathfrak{h}$, $Y \in \mathfrak{g} \setminus \mathfrak{h}$, such that $\rho(X)(Y + \mathfrak{h}) = \mathfrak{h}$ for all $X \in \mathfrak{h}$. This amounts to saying that $[X, Y] \subset \mathfrak{h}$ for all $X \in \mathfrak{h}$. Thus the vector space $\mathfrak{h}'$spanned by $\mathfrak{h}$ and Y is a Lie subalgebra of $\mathfrak{g}$ containing strictly $\mathfrak{h}$. We conclude that it is equal to $\mathfrak{g}$, showing that indeed $\mathfrak{h}$ has co-dimension 1 in $\mathfrak{g}$. Moreover $\mathfrak{h}$ is clearly an ideal in $\mathfrak{h}' = \mathfrak{g}$. This proves our claim.

Now, apply the induction hypothesis to $\mathfrak{h}$, to deduce that the following subspace $W \subset V$ is non-zero:

$$W = \{v \in V \mid X(v) = 0 \text{ for all } X \in \mathfrak{h}\}.$$

Using the claim, we see that this subspace is $\mathfrak{g}$ invariant. Indeed, for any $X \in \mathfrak{g}$, $v \in W$ and $Y \in \mathfrak{h}$, we have

$$Y(X(v)) = X(Y(v)) - [X, Y](v) = 0,$$

so that $X(v) \in W$.

Now fix $Y \in \mathfrak{g} \setminus \mathfrak{h}$. Since Y acts as a nilpotent element on V this is also true for its restriction to W. So it has a non-trivial kernel in W: there exists a non-zero vector $v \in W$ such that $Y(v) = 0$. Since Y and $\mathfrak{h}$ span $\mathfrak{g}$, the vector v satisfies the desired conclusion. □

PROOF OF ENGEL'S THEOREM. This follows from an induction on the dimension of V. The result is clearly true if $\dim(V) \leq 1$. Assume now that V is arbitrary. Thanks to Lemma 4.6 we may find a vector $v_1 \in V$ such that $X(v_1) = 0$ for all $X \in \mathfrak{g}$. Setting $W := \mathrm{span}_k(v_1)$, the representation of $\mathfrak{g}$ on V yields a representation on the quotient space $\rho : \mathfrak{g} \to \mathcal{L}(V/W)$. Since $\dim(V/W) = \dim(V) - 1$, we may apply the induction assumption to find vectors $v_2, \ldots, v_d \in V$ such that $V = \mathrm{span}_k(\{v_1, \ldots, v_d\})$ and $\rho(X)(v_i + W) \in \mathrm{span}_k(\{v_j + W \mid j < i\})$ for all $X \in \mathfrak{g}$ (with the convention that $\mathrm{span}_k(\emptyset) = \{0\}$). This means exactly that $[X, v_i] \in \mathrm{span}_k(\{v_1, \ldots, v_{i-1}\})$for all i and all $X \in \mathfrak{g}$, i.e. that $\mathfrak{g} \subset \mathfrak{u}_d^+$. □

PROOF OF THEOREM 4.3. If $\mathfrak{g}$ is nilpotent, pick a sequence of ideals $\mathfrak{g} = \mathfrak{g}_0 \supset \cdots \supset \mathfrak{g}_n = \{0\}$ such that $[\mathfrak{g}, \mathfrak{g}_i] \subset \mathfrak{g}_{i+1}$ for all i. Then we observe that for all $X, Y \in \mathfrak{g}$, and all integer $i \geq 0$, $\mathrm{ad}(X)^i(Y) \in \mathfrak{g}_i$. Thus $\mathrm{ad}(X)^n = 0$, showing that $\mathrm{ad}(X)$ is nilpotent.

Conversely assume now that each $\mathrm{ad}(X)$, $X \in \mathfrak{g}$ is nilpotent. By Engel's theorem, we know that $\mathfrak{h} = \mathrm{ad}(\mathfrak{g})$ is a nilpotent Lie algebra: there exist ideals $\mathfrak{h} = \mathfrak{h}_0 \supset \cdots \supset \mathfrak{h}_n = \{0\}$ such that $[\mathfrak{h}, \mathfrak{h}_i] \subset \mathfrak{h}_{i+1}$ for all i. For each i dente by $g_i := \mathrm{ad}^{-1}(\mathfrak{h}_i)$, the pre-image of $\mathfrak{h}_i$. Then clearly these are ideals in $\mathfrak{g}$, such that $\mathfrak{g} = \mathfrak{g}_0 \supset \cdots \supset \mathfrak{g}_n = \ker(\mathrm{ad}) = \mathcal{Z}(\mathfrak{g})$ such that $[\mathfrak{g}, \mathfrak{g}_i] \subset \mathfrak{g}_{i+1}$. Since $[\mathfrak{g}, \mathfrak{g}_n] = \{0\}$, this shows that $\mathfrak{g}$ is nilpotent. □

In fact the second part of the proof above shows that a Lie algebra $\mathfrak{g}$ is nilpotent if and only if $\mathrm{ad}(\mathfrak{g})$ is nilpotent. The same argument applies to solvability.

For solvable groups, we have the following theorem. Let us emphasize that it is only valid on an algebraically closed field, as Remark 4.5 shows (for $k = \mathbb{R}$ for instance). This result generalizes the fact that a complex matrix is conjugate to a triangular matrix.

THEOREM 4.7 (Lie). *Let K be an algebraically closed field of characteristic 0 and V a K-vector space. Consider a solvable Lie subalgebra $\mathfrak{g} \subset \mathcal{L}(V)$ over K. Then there exists a basis of V in which $\mathfrak{g} \subset \mathfrak{p}_d^+$, where $d = \dim(\mathfrak{g})$.*

Exactly as for the nilpotent case, Lie's theorem follows from the following lemma.

LEMMA 4.8. *If V a K-vector space and $\mathfrak{g} \subset \mathcal{L}(V)$ is a solvable Lie subalgebra over K, then there exists a non-zero vector $v \in V$ which is an eigenvector for all elements of $\mathfrak{g}$.*

PROOF. Once again, we proceed by induction on the dimension of $\mathfrak{g}$. The result being trivial if $\dim(\mathfrak{g}) \leq 1$, let us assume that $\dim(\mathfrak{g}) \geq 2$. We note that $[\mathfrak{g}, \mathfrak{g}]$ is a proper ideal of $\mathfrak{g}$, and better, any intermediate vector space $[\mathfrak{g}, \mathfrak{g}] \subset \mathfrak{h} \subset \mathfrak{g}$ is also automatically an ideal of $\mathfrak{g}$. So we may find an ideal $\mathfrak{h}$ of codimension 1 in $\mathfrak{g}$.

We apply the induction hypothesis to the vector subspace $\mathfrak{h}$ and find a non-zero vector $v \in V$ such that $Xv = \alpha(X)v$ for all $X \in \mathfrak{h}$ and some scalar $\alpha(X) \in K$ depending on X. It is important to observe that the map $\alpha : \mathfrak{h} \to K$ is a linear functional on $\mathfrak{h}$. This observation will also play a key role in later sections.

Claim. For all $X \in \mathfrak{g}$ and $Y \in \mathfrak{h}$, we have $\alpha([X, Y]) = 0$.

Fix $X \in \mathfrak{g}$. Consider a chain of subspaces $\{0\} = V_0 \subset V_1 \subset \cdots \subset V_n$ where $V_i = \mathrm{span}_K(\{v, Xv, \ldots, X^{i-1}v\})$ and n is the smallest i such that $V_i = V_{i+1}$. Note that $\dim(V_n) = n$ and $XV_n \subset V_n$. By induction one verifies that $Y(X^i v) = \alpha(Y)X^i v + Z$ for some $Z \in V_i$. Indeed this fact is trivial for $i = 0$, and if it is true for some $i < n$, then

$$\begin{aligned} Y(X^{i+1}v) = YX(X^i v) &= XY(X^i v) - [X, Y](X^i v) \\ &= X(\alpha(Y)X^i v + Z) - (\alpha([X, Y])X^i v + Z'), \end{aligned}$$

for some elements $Z, Z' \in V_i$. Since $Z - (\alpha([X, Y])X^i v + Z')$ belongs to V_{i+1} we find that the assertion is still true for $i + 1$.

Thus the vector space V_n is globally invariant under the elements of $\mathfrak{h}$, and in the basis $v, Xv, \ldots, X^{n-1}v$, any element $Y \in \mathfrak{h}$ can be written as an upper triangular matrix, with diagonal coefficients all equal to $\alpha(Y)$. Taking the trace on this subspace V_n, we find $\mathrm{Tr}(Y) = n\alpha(Y)$ for all $Y \in \mathfrak{h}$. Applying this to an element of the form $[X, Y]$ for $Y \in \mathfrak{h}$, we find

$$n\alpha([X, Y]) = \mathrm{Tr}([X, Y]) = \mathrm{Tr}(XY - YX) = 0.$$

This proves the claim since K has characteristic 0.

Now, denote by $W \subset V$ the subspace $W := \{w \in V \mid Xw = \alpha(X)w, \text{ for all } X \in \mathfrak{h}\}$. Since $v \in W$ we have that $W \neq \{0\}$. Moreover, the claim implies that W is globally $\mathfrak{g}$ invariant. Indeed if $w \in W$, $X \in \mathfrak{g}$ and $Y \in \mathfrak{h}$, we have

$$Y(X(w)) = X(Y(w)) - [X,Y](w) = \alpha(Y)X(w) - \alpha([X,Y])(w) = \alpha(Y)X(w),$$

so $X(w) \in W$. Now pick any $X \in \mathfrak{g} \setminus \mathfrak{h}$. Since K is algebraically closed, we may find an element $w \in W$ which is an eigenvector of X. Since $\mathfrak{h}$ has codimension 1 in $\mathfrak{g}$, X and $\mathfrak{h}$ linearly span $\mathfrak{g}$. This shows that w is a common eigenvector for all elements in $\mathfrak{g}$. □

We leave as an exercise to complete proof of Lie's theorem from the above lemma. To conclude this section we provide another criterion, due to Cartan, to check that a Lie algebra is solvable.

THEOREM 4.9. *Consider a k-vector space V, and a Lie subalgebra $\mathfrak{g} \subset \mathcal{L}(V)$. Then $\mathfrak{g}$ is solvable if and only if*

$$\operatorname{Tr}(XY) = 0 \text{ for every } X \in \mathfrak{g} \text{ and } Y \in [\mathfrak{g},\mathfrak{g}]. \tag{1.1}$$

In the proof we will need the following two facts.

EXERCISE 4.10. If $\alpha_1, \ldots, \alpha_n \in k$ are elements such that $\sum_{i=1}^n \alpha_i \phi(\alpha_i) = 0$ for all $\mathbb{Q}$-linear map $\phi \in \mathcal{L}_{\mathbb{Q}}(k)$, then $\alpha_1 = \cdots = \alpha_n = 0$.

EXERCISE 4.11. If $X \in \mathcal{L}(V)$ is a nilpotent element, check that $\operatorname{ad}(X) : Y \mapsto XY - YX$ is nilpotent in $\mathcal{L}(\mathcal{L}(V))$. Hint: for all $n \geq 1$, check that $ad(X)^n(Y) \in \operatorname{span}(\{X^k Y X^{n-k} \mid 0 \leq k \leq n\})$.

PROOF OF THEOREM 4.9. We first note that it suffices to prove the theorem in the case where k is algebraically closed. Indeed, if the equivalence holds for the algebraic closure K of k, we may apply it to the K-Lie algebra $\mathfrak{g} \otimes_k K \subset \mathcal{L}_K(V \otimes_k K)$. Observing that $\mathfrak{g}$ is solvable if and only if $\mathfrak{g} \otimes_k K$ is, and that equation (1.1) holds for $\mathfrak{g}$ if and only if it holds for $\mathfrak{g} \otimes_k K$, we then conclude that the theorem holds for $\mathfrak{g}$. This way we reduce to the case where $k = K$ is algebraically closed.

If $\mathfrak{g}$ is solvable, then we may apply Lie's theorem to represent its elements by triangular matrices in a fixed basis of V. Then one sees that elements of $[\mathfrak{g},\mathfrak{g}]$ are written as strictly upper triangular matrices in this basis (with zeros on the diagonal). Computing the trace in this basis clearly gives (1.1).

The converse is more delicate, although the general idea is simple: in order to show that $\mathfrak{g}$ is solvable, one only needs to checks that $[\mathfrak{g},\mathfrak{g}]$ is nilpotent. By Engel's theorem, it suffices to check that all elements $X \in [\mathfrak{g},\mathfrak{g}]$ are nilpotent. Apply (1.1) to find that $\operatorname{Tr}(X^2) = \sum_i \alpha_i^2 = 0$, where $\alpha_1, \ldots, \alpha_n$ are the eigenvalues of X. This "shows" that each α_i vanishes, and hence that X is nilpotent.

Unfortunately, the eigenvalues of X need not be real in general, so $\sum_i \alpha_i^2 = 0$ does not imply that each α_i vanishes. To get around this issue, we use Jordan's decomposition of X inside $\mathcal{L}(V)$. This will create another issue, because it forces us to get out of the Lie algebra $\mathfrak{g}$ so we will have to be careful to apply equation (1.1).

Given $X \in [\mathfrak{g},\mathfrak{g}]$, we may write $X = S + N$ for its Jordan decomposition: $S, N \in \mathcal{L}(V)$ are obtained as polynomials in X, S and N commute with each other, S is diagonalisable over K and N is nilpotent. Representing $S = \operatorname{diag}(\alpha_1, \ldots, \alpha_n)$ by a diagonal matrix in a suitable basis, for any $\mathbb{Q}$-linear map $\phi : k \to k$ denote by $\phi(S) := \operatorname{diag}(\phi(\alpha_1), \ldots, \phi(\alpha_n)) \in \mathcal{L}(V)$.

STEP 1. To prove that X is nilpotent it suffices to check that $\mathrm{Tr}(X\phi(S)) = 0$ for all $\phi \in \mathcal{L}_{\mathbb{Q}}(k)$.

Indeed note that $\mathrm{Tr}(X\phi(S)) = \sum_{i=1}^{n} \alpha_i \phi(\alpha_i)$ so it follows from Exercise 4.10 that $\alpha_i = 0$ for all i, hence $S = 0$ and $X = N$ is nilpotent.

Unfortunately, we cannot apply equation (1.1) to deduce that $\mathrm{Tr}(X\phi(S)) = 0$ because we don't know that $\phi(S)$ belongs to $\mathfrak{g}$. The following claim fixes this gap.

STEP 2. For all $\phi \in \mathcal{L}_{\mathbb{Q}}(k)$, to prove that $\mathrm{Tr}(X\phi(S)) = 0$ it suffices to check that $\mathrm{ad}\,\phi(S) \in \mathrm{ad}(\mathcal{L}(V)) \subset \mathcal{L}(\mathcal{L}(V))$ maps $\mathfrak{g}$ into $[\mathfrak{g}, \mathfrak{g}]$.

Indeed, if $\mathrm{ad}(\phi(S))$ maps $\mathfrak{g}$ into $[\mathfrak{g}, \mathfrak{g}]$, we get that for all $Y, Z \in \mathfrak{g}$, $\mathrm{Tr}(Y[Z, \phi(S)]) = 0$, by (1.1). Thus $\mathrm{Tr}([Y, Z]\phi(S)) = \mathrm{Tr}(Y[Z, \phi(S)]) = 0$, and hence $\mathrm{Tr}(X'\phi(S)) = 0$ for all $X' \in [\mathfrak{g}, \mathfrak{g}]$ and in particular for $X' = X$.

STEP 3. We show that $\mathrm{ad}(X) = \mathrm{ad}(S) + \mathrm{ad}(N)$ is the Jordan decomposition of $\mathrm{ad}(X)$. In particular $\mathrm{ad}(S)$ maps $\mathfrak{g}$ into $[\mathfrak{g}, \mathfrak{g}]$.

Observe that $\mathrm{ad}(S)$ is indeed diagonalisable with eigenvalues $\alpha_i - \alpha_j$ and eigenvectors the canonical matrices $E_{i,j}$. It moreover follows from Exercise 4.11 that $\mathrm{ad}(N)$ is nilpotent. Moreover $[\mathrm{ad}(S), \mathrm{ad}(N)] = \mathrm{ad}([S, N]) = 0$, so indeed $\mathrm{ad}(X) = \mathrm{ad}(S) + \mathrm{ad}(N)$ is the Jordan decomposition of $\mathrm{ad}(X)$. In particular $\mathrm{ad}(S)$ can be expressed as a polynomial in $\mathrm{ad}(X)$. But since $X \in \mathfrak{g}$, $\mathrm{ad}(X)$ maps $\mathfrak{g}$ into $[\mathfrak{g}, \mathfrak{g}]$, and any polynomial in $\mathrm{ad}(X)$ does so.

STEP 4. We finally prove that $\mathrm{ad}(\phi(S))$ maps $\mathfrak{g}$ into $[\mathfrak{g}, \mathfrak{g}]$.

We observed that $\mathrm{ad}(S)$ is a diagonal element in the basis $E_{i,j}$, with diagonal values $\alpha_i - \alpha_j$. In the same basis $\mathrm{ad}(\phi(S))$ is the diagonal element with diagonal values $\phi(\alpha_i - \alpha_j)$. Using an interpolation polynomial, we may find a polynomial $P \in k[X]$ such that $P(\alpha_i - \alpha_j) = \phi(\alpha_i - \alpha_j)$. Hence $\mathrm{ad}(\Phi(S))$ can be expressed as a polynomial in $\mathrm{ad}(S)$. So again, it maps $\mathfrak{g}$ into $[\mathfrak{g}, \mathfrak{g}]$. This finishes the proof of the theorem. □

2. Semi-simple Lie algebras

We continue to assume that our Lie algebras are defined over the field k of characteristic 0.

PROPOSITION 4.12. *Any Lie algebra $\mathfrak{g}$ admits a greatest solvable ideal $\mathfrak{h}$, in the sense that $\mathfrak{h}$ is a solvable ideal of $\mathfrak{g}$, which contains all other solvable ideals of $\mathfrak{g}$. We call $\mathfrak{h}$ the* solvable radical *of $\mathfrak{g}$.*

PROOF. Note that any solvable ideal of $\mathfrak{g}$ is contained in a maximal solvable ideal of $\mathfrak{g}$. To show the proposition, it thus suffices to show that any two maximal solvable ideals $\mathfrak{h}$ and $\mathfrak{h}'$ in $\mathfrak{g}$ coincide. But by maximality, we only need to check that $\mathfrak{h} + \mathfrak{h}'$ is again a solvable ideal in $\mathfrak{g}$. Observe that we have a short exact sequence

$$0 \to \mathfrak{h} \to \mathfrak{h} + \mathfrak{h}' \to (\mathfrak{h} + \mathfrak{h}')/\mathfrak{h} = \mathfrak{h}'/(\mathfrak{h} \cap \mathfrak{h}') \to 0.$$

We leave as an exercise to check that given any short exact sequence, the extension Lie algebra is solvable if and only if the quotient and the subalgebra both are solvable (we also point out that the nilpotent analogue of this fact is not true). □

DEFINITION 4.13. A Lie algebra is called *semi-simple* if its solvable radical is $\{0\}$.

In particular a semi-simple Lie algebra has trivial center.

EXERCISE 4.14. Observe that if $\mathfrak{h}$ is an ideal inside $\mathfrak{g}$, then so is $[\mathfrak{h}, \mathfrak{h}]$. Using this, check that a Lie algebra is semi-simple if and only if it has no abelian ideal.

Alternatively, semi-simple Lie algebras may be characterized in terms of their *Killing form*. By definition, the Killing form on a Lie algebra $\mathfrak{g}$ is the symmetric bilinear form B on $\mathfrak{g}$ defined by the formula

$$B(X,Y) = \mathrm{Tr}(\mathrm{ad}(X)\,\mathrm{ad}(Y)), \text{ for all } X, Y \in \mathfrak{g}.$$

This form satisfies the invariance property $B([X,Y],Z)+B(Y,[X,Z]) = 0$ for all $X, Y, Z \in \mathfrak{g}$.

THEOREM 4.15. *A Lie algebra is semi-simple if and only if its Killing form is non-degenerate.*

PROOF. Consider $\mathfrak{u} \subset \mathfrak{g}$ the subspace $\mathfrak{u} := \{X \in \mathfrak{g} \mid B(X,Y) = 0 \text{ for all } Y \in \mathfrak{g}\}$. Since $B([X,Y],Z) + B(Y,[X,Z]) = 0$, for all $X, Y, Z \in \mathfrak{g}$, we see that $\mathfrak{u}$ is an ideal in $\mathfrak{g}$. Since $\mathrm{Tr}(\mathrm{ad}(X)\,\mathrm{ad}(Y)) = 0$ for all $X, Y \in \mathfrak{u}$, it follows from Cartan's criterion that $\mathrm{ad}_{\mathfrak{g}}(\mathfrak{u})$ is solvable. Since $\mathfrak{u}$ is the extension of $\mathrm{ad}_{\mathfrak{g}}(\mathfrak{u})$ by the center of $\mathfrak{g}$, it follows that $\mathfrak{u}$ is solvable. So obviously, if $\mathfrak{g}$ is semi-simple, then $\mathfrak{u} = 0$, that is, B is non-degenerate.

Conversely, let us assume that B is non-degenerate and show that $\mathfrak{g}$ is semi-simple. By Exercise 4.14, it suffices to show that $\mathfrak{g}$ has no abelian ideal. Take an abelian ideal $\mathfrak{a}$ of $\mathfrak{g}$. Take a supplementary space $\mathfrak{a}'$ of $\mathfrak{a}$ in $\mathfrak{g}$: $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}'$. In this decomposition, the elements $\mathrm{ad}(X)$ and $\mathrm{ad}(Y)$, $X \in \mathfrak{a}$, $Y \in \mathfrak{g}$ are written as bloc matrices

$$\mathrm{ad}(X) = \begin{pmatrix} 0 & * \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathrm{ad}(Y) = \begin{pmatrix} * & * \\ 0 & * \end{pmatrix}.$$

So it is obvious that $\mathrm{Tr}(\mathrm{ad}(X)\,\mathrm{ad}(Y)) = 0$ for all $X \in \mathfrak{a}$, $Y \in \mathfrak{g}$. Hence $\mathfrak{a} \subset \mathfrak{u} = \{0\}$. □

The above characterization shows that semi-simplicity does not depend on the field of definition: if $\mathfrak{g}$ is a k-Lie algebra then for all field extension $k \subset k'$, $\mathfrak{g}$ is semi-simple if and only if $\mathfrak{g} \otimes_k k'$ is semi-simple over k'. In fact it can be proved that the solvable radical of $\mathfrak{g} \otimes_k k'$ is $\mathfrak{r} \otimes_k k'$, where $\mathfrak{r}$ is the solvable radical of $\mathfrak{g}$.

If $\mathfrak{g}$, $\mathfrak{h}$ and $\mathfrak{k}$ are Lie algebras, we write $\mathfrak{k} = \mathfrak{g} \oplus \mathfrak{h}$ to mean that $\mathfrak{k}$ is the direct sum of $\mathfrak{g}$ and $\mathfrak{h}$ as vector spaces, and that $[X,Y] = 0$ for all $X \in \mathfrak{g}$, $Y \in \mathfrak{h}$.

PROPOSITION 4.16. *Let $\mathfrak{g}$ be a semi-simple Lie algebra. Then for any ideal $\mathfrak{a}$, its orthogonal $\mathfrak{a}^\perp$ w.r.t the Killing form is also an ideal in $\mathfrak{g}$ and we have $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^\perp$.*

PROOF. Using the invariance property of the Killing form, one easily shows that $\mathfrak{a}^\perp$ is also an ideal in $\mathfrak{g}$. Using Cartan's solvability criterion, one shows that $\mathfrak{a} \cap \mathfrak{a}^\perp$ is solvable (exactly as we showed in the previous proof that $\mathfrak{g} \cap \mathfrak{g}^\perp$ was a solvable ideal of $\mathfrak{g}$). Since $\mathfrak{g}$ is semi-simple, it follows that $\mathfrak{a} \cap \mathfrak{a}^\perp = 0$, and hence $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{a}^\perp$. Note moreover that this direct sum is not only a vector space direct sum but actually a Lie algebra direct sum, since a $[X,Y] \in \mathfrak{a} \cap \mathfrak{a}^\perp = \{0\}$, for all $X \in \mathfrak{a}$, $Y \in \mathfrak{a}^\perp$. □

A *simple* Lie algebra is one that admits no non-trivial ideals. The previous proposition easily implies the following description of semi-simple Lie algebras.

COROLLARY 4.17. *A Lie algebra is semi-simple if and only if it is a direct sum of simple Lie algebras.*

Let us derive some important consequences of the above facts. The following property of semi-simple Lie algebras will be important for us, because it characterizes them algebraically. It will imply that every connected semi-simple Lie group is almost an algebraic group.

PROPOSITION 4.18. *Every derivation of a semi-simple Lie algebra is inner. Equivalently, the map* $\mathrm{ad} : \mathfrak{g} \to \mathrm{Der}(\mathfrak{g})$ *is a Lie algebra isomorphism.*

PROOF. The map ad is injective because a semi-simple Lie algebra has trivial center. To show that it is surjective, we first prove several facts.

Claim 1. The range $\mathrm{ad}(\mathfrak{g})$ of ad is an ideal in $\mathrm{Der}(\mathfrak{g})$. Denote this ideal by $\mathfrak{a} \subset \mathrm{Der}(\mathfrak{g})$.

Indeed, for all $X, Y \in \mathfrak{g}$ and $D \in \mathrm{Der}(\mathfrak{g})$, we have

$$(\mathrm{ad}(X)D - D\,\mathrm{ad}(X))(Y) = [X, DY] - D([X, Y]) = -[DX, Y] = -\,\mathrm{ad}(DX)(Y),$$

showing that $[\mathrm{ad}(X), D] = -\,\mathrm{ad}(DX) \in \mathfrak{a}$.

Claim 2. The restriction of the Killing form of a Lie algebra $\mathfrak{h}$ to an ideal $\mathfrak{b}$ is the Killing form of $\mathfrak{b}$.

If $X, Y \in \mathfrak{b}$, note that both $\mathrm{ad}_{\mathfrak{h}}(X)$ and $\mathrm{ad}_{\mathfrak{h}}(Y)$ map $\mathfrak{h}$ into $\mathfrak{b}$. The claim then follows easily by writting the elements as bloc matrices.

Claim 3. We have $\mathrm{Der}(\mathfrak{g}) = \mathfrak{a} \oplus \mathfrak{a}^{\perp}$, where the orthogonal is taken with respect to the Killing form on $\mathrm{Der}(\mathfrak{g})$.

This follows from Claim 2 and from the fact that the Killing form on $\mathfrak{a}$ is non-degenerate.

Claim 4. $\mathfrak{a}^{\perp} = 0$.

Take $D \in \mathfrak{a}^{\perp}$ and $X \in \mathfrak{g}$. As we so in the proof of Claim 1, we have $\mathrm{ad}(DX) = [D, \mathrm{ad}(X)]$. So $\mathrm{ad}(DX) \in \mathfrak{a} \cap \mathfrak{a}^{\perp} = \{0\}$. Since ad is injective, we conclude that $DX = 0$, and hence $D = 0$. The proof is now complete. □

DEFINITION 4.19. An element X in a semi-simple Lie algebra is called *nilpotent* (resp. semi-simple) if $\mathrm{ad}(X)$ is a nilpotent (resp. semi-simple) endomorphism of $\mathfrak{g}$.

PROPOSITION 4.20. *In a semi-simple algebra every element X admits a unique decomposition $X = S + N$ with S semi-simple, N nilpotent, and $[S, N] = 0$.*

PROOF. The uniqueness follows from the uniqueness of the Jordan decomposition of $\mathrm{ad}(X)$ inside $\mathcal{L}(\mathfrak{g})$, and from the fact that ad is injective.

To prove the existence, first perform the Jordan decomposition of $\mathrm{ad}(X)$ inside $\mathcal{L}(\mathfrak{g})$: $\mathrm{ad}(X) = \widetilde{S} + \widetilde{N}$, with $\widetilde{S}, \widetilde{N} \in \mathcal{L}(\mathfrak{g})$, $\widetilde{S}$ semi-simple, $\widetilde{N}$ nilpotent and $[\widetilde{S}, \widetilde{N}] = 0$. By Proposition 4.18, we only need to check that $\widetilde{S}$ is a derivation of $\mathfrak{g}$, i.e. that

$$\widetilde{S}([Y, Z]) = [\widetilde{S}Y, Z] + [Y, \widetilde{S}Z], \text{ for all } Y, Z \in \mathfrak{g}. \tag{2.1}$$

To check this statement, we may assume that $k = K$ is algebraically closed. Then the endomorphism $\widetilde{S}$ is explicit. Decompose $\mathfrak{g}$ into a sum of vector subspaces $\mathfrak{g} = \oplus_{\lambda \in K} \mathfrak{g}_{\lambda}$, where $\mathfrak{g}_{\lambda} = \cup_{n \geq 1} \mathrm{Ker}((\mathrm{ad}(X) - \lambda)^n)$. Then $\widetilde{S}$ acts on each $\mathfrak{g}_{\lambda}$ by multiplication by λ. By linearity, it suffices to check (2.1) for $Y \in \mathfrak{g}_{\lambda}$ and $Z \in \mathfrak{g}_{\mu}$ for some $\lambda, \mu \in K$. But in this case the formula amounts to $[\mathfrak{g}_{\lambda}, \mathfrak{g}_{\mu}] \subset \mathfrak{g}_{\lambda+\mu}$. This later fact is a consequence of the formula

$$(\mathrm{ad}(X) - \lambda - \mu)^n([Y, Z]) \in \mathrm{span}(\{[(\mathrm{ad}(X) - \lambda)^i Y, (\mathrm{ad}(X) - \mu)^{n-i} Z] \mid 0 \leq i \leq n\}),$$

which is easily checked by induction. □

3. Representations of $\mathfrak{sl}_2$

In this section, we continue to denote by k a field of characteristic 0. We study the representations of the Lie algebra $\mathfrak{sl}_2(k)$ consisting of 2×2 matrices over k, with trace 0. This example is fundamental in two respects. First it is the smallest example of a simple Lie algebra, being of dimension 3. Moreover, this Lie algebra is embedded in many ways in any other semi-simple Lie algebra (over an algebraically closed field).

Observe that $\mathfrak{sl}_2(k)$ is spanned by the matrices H, X, Y given below, satisfying the relations $[X, Y] = H$, $[H, X] = 2X$, $[H, Y] = -2Y$:

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

A representation of a semi-simple Lie algebra $\mathfrak{g}$ on a vector space V is called *irreducible* if its only $\mathfrak{g}$-invariant subspaces are the trivial ones $\{0\}$ and V. The following result is in fact true for general semi-simple Lie algebras, and the proof is based on the same unitary trick of Weyl.

THEOREM 4.21. *Every finite dimensional representation of $\mathfrak{s} := \mathfrak{sl}_2(k)$ is a direct sum of irreducible representations.*

PROOF. Fix a representation $\pi : \mathfrak{s} \to \mathcal{L}(V)$ over k. We only need to show that for every $\mathfrak{s}$-invariant subspace $W \subset V$ is complemented in V, in the sense that it admits an $\mathfrak{s}$-invariant complement W'.

Claim 1. We may assume that $k = \mathbb{C}$.

This follows from a standard, but very useful, argument. Fix a basis $e_1, \dots, e_n \in V$ of V such that $e_1, \dots, e_m$ forms a basis of W. In this basis, $V \simeq k^n$, and the endomorphisms $\pi(X)$, $\pi(Y)$, $\pi(H)$ are represented as matrices with coefficients in k. Denote by k' the subfield generated by these coefficients. It is a finitely generated subfield of k, and setting $V_0 = \mathrm{span}_{k'}(e_1, \dots, e_n)$, $W_0 = W \cap V_0$, the elements of $\pi(\mathfrak{sl}_2(k'))$ leave V_0 and W_0 globally invariant. If we manage to find a supplementary W_0' of W_0 in V_0 that is $\mathfrak{sl}_2(k')$-invariant then we will get that $W' := W_0 \otimes_{k'} k$ is an $\mathfrak{s}$-invariant supplementary of W. This way we may assume that $k = k'$, and hence, that k is finitely generated.

Now observe that any finitely generated field of characteristic 0 embeds into $\mathbb{C}$. So the representation π gives rise to a complex representation of $\mathfrak{sl}_2(\mathbb{C})$ on $\widetilde{V} = V \otimes_k \mathbb{C}$, and leaves the subspace $\widetilde{W} = W \otimes_k \mathbb{C}$ globally invariant. Denote by $p : V \to V/W$ the projection map over k. Finding an $\mathfrak{s}$-invariant complement of W in V amounts to finding a k-linear map $\phi : V/W \to V$ such that

$$\phi(\pi(Z)(v + W)) = \pi(Z)\phi(v + W) \quad \text{and} \quad p \circ \phi = \mathrm{id},$$

for all $v \in V$ and $Z \in \{H, X, Y\}$. This is a linear system defined over k (as can be seen by expressing it in appropriate bases of V and V/W), so it admits a solution over k if and only if it admits a solution over $\mathbb{C}$. The claim follows.

Claim. 2 The group $K := \mathrm{SU}(2)$ is simply connected and its Lie algebra $\mathfrak{su}_2$ satisfies $\mathfrak{sl}_2(\mathbb{C}) = \mathfrak{su}_2 \otimes_{\mathbb{R}} \mathbb{C}$.

As a topological space, the group SU(2) is homeomorphic to the sphere $\mathbb{S}^3$. So it is indeed simply connected. Now its Lie algebra is the Lie algebra of complex anti-hermitian matrices of trace 0: $\mathfrak{su}_2 = \{X \in \mathfrak{sl}_2(\mathbb{C}) \mid X^* = -X\}$. Observe that any matrix can be uniquely written as a sums of an hermitian matrix and an anti-hermitian. Moreover X is anti-hermitian if and only if iX is hermitian. Thus $\mathfrak{sl}_2(\mathbb{C}) = \mathfrak{su}_2 \oplus i\mathfrak{su}_2$.

With this claim in hand, any representation of $\mathfrak{sl}_2(\mathbb{C})$ gives rise to a linear representation $\mathfrak{su}_2 \to \mathcal{L}_{\mathbb{C}}(V)$. By Theorem 2.11, this representation gives rise to a Lie group morphism $K \to \mathrm{GL}_{\mathbb{C}}(V)$. Since K is compact, we may find a K-invariant inner product on V. Indeed fixing an arbitrary inner product $\langle \cdot, \cdot \rangle$ on V, the following inner product is K-invariant:

$$\langle v, w \rangle_K := \int_K \langle gv, gw \rangle d\lambda(k), \quad v, w \in V.$$

Here, λ is a Haar measure on K, which is finite because K is compact.

Since $W \subset V$ is $\mathfrak{sl}_2(\mathbb{C})$ invariant, it is $\mathfrak{su}_2$-invariant, and hence K-invariant. Therefore its orthogonal W' for the K-invariant inner product is again K-invariant. So W' is an $\mathfrak{su}_2$-invariant complex subspace of V. So it is also $\mathfrak{sl}_2(\mathbb{C})$-invariant, as wanted. □

For any integer $d \geq 0$, denote by V_d the vector space of homogeneous polynomials of degree d in two variables x, y: $V_d = \mathrm{span}_k(\{x^i y^{d-i} \mid 0 \leq i \leq d\})$. Then we have a representation $\pi_d : \mathfrak{sl}_2(k) \to \mathcal{L}(V_d)$ defined by

$$\pi_d(X) = x\frac{\partial}{\partial y}, \quad \pi_d(Y) = y\frac{\partial}{\partial x}, \quad \pi_d(H) = x\frac{\partial}{\partial x} - y\frac{\partial}{\partial y}.$$

PROPOSITION 4.22. *For each $d \geq 0$, π_d is an irreducible representation of $\mathfrak{s} := \mathfrak{sl}_2(k)$. Moreover, every finite dimensional irreducible representation of $\mathfrak{s}$ is isomorphic to some π_d.*

PROOF. We first show that V_d is irreducible. We ignore the notation π_d and for instance write freely Hv instead of $\pi_d(H)v$, $v \in V_d$. First note that for all $0 \leq i \leq d$, $x^i y^{d-i}$ is an eigenvector of H, with eigenvalue $2i - d$. So V_d can be expressed as the direct sum of (one dimensional) subspaces $V = \bigoplus_i V_i$, where V_i consists entirely of eigenvectors of H, with eigenvalue $2i - d$. Moreover we observe that $X(V_i) \subset V_{i+1}$ with equality if and only if $i < d$. Likewise $Y(V_i) \subset V_{i-1}$ with equality if and only if $i \geq 1$.

Take a non-zero invariant subspace $W \subset V_d$ and fix a non-zero vector $v \in W$. We may write v as a sum of elements v_i, with each $i \in V_i$, and take i_0 to be the largest integer such that $v_{i_0} \neq 0$. Then $Y^{i_0} v$ belongs to W and it is equal to $Y^{i_0}(v_i)$ since all the other components will be killed. Moreover $Y^{i_0}(v_i) \in V_0 \setminus \{0\}$. So we conclude that $V_0 \subset W$, and applying X^i sufficiently many times we show that $V_i \subset W$ for all $i \leq d$. Hence $W = V_d$.

Now fix an arbitrary irreducible representation $\pi : \mathfrak{sl}_2(k) \to \mathcal{L}(V)$. We need to recover the above structure. Assume first that k is algebraically closed. In particular, we may decompose V into a direct sum of characteristic spaces of H: $V = \bigoplus_{\lambda \in k} V_\lambda$, where $V_\lambda = \cup_n \mathrm{Ker}(\pi(H) - \lambda)^n$.

Using the fact that $[H, X] = 2X$, observe that for all $\lambda \in k$, we have

$$(\pi(H) - 2 - \lambda)\pi(X) = \pi(X)\pi(H) - \lambda\pi(X) = \pi(X)(\pi(H) - \lambda).$$

So by induction, it follows $(\pi(H) - 2 - \lambda)^n \pi(X) = \pi(X)(\pi(H) - \lambda)^n$ for all $n \geq 0$. So we conclude that for all $\lambda \in k$,

$$\pi(X)V_\lambda \subset V_{\lambda+2} \quad \text{and likewise} \quad \pi(Y)V_\lambda \subset V_{\lambda-2}.$$

Since V is finite dimensional, we may find $\lambda \in k$ such that $V_\lambda \neq \{0\}$ but $V_{\lambda+2} = \{0\}$. Pick a non-zero vector $v \in V_\lambda$ which is a λ-eigenvector of $\pi(H)$ and for all $i \geq 0$, define $v_i := \pi(Y)^i v \in V_{\lambda-2i}$. Since V is finite dimensional, there exists an integer n such that $v_i \neq 0$ for all $0 \leq i \leq n$ and $v_{n+1} = 0$.

Claim. The linear span of $v_0, v_1, \ldots, v_n$ is $\pi(\mathfrak{s})$-invariant (hence equal to V), and for all $i \geq 0$ we have

$$\pi(X)v_{i+1} = (i+1)(\lambda - i)v_i.$$

Denote by $W = \mathrm{span}_k(\{v_0, v_1, \ldots, v_n\})$. Clearly since each v_i is an eigenvector of $\pi(H)$, W is globally invariant under $\pi(H)$. By definition it is also globally invariant under $\pi(Y)$, so we only need to check that it is globally invariant under $\pi(X)$, which clearly follows from the stated formula. We prove this formula by induction. Note that $\pi(X)v_0 \in V_{\lambda+2} = \{0\}$. So we see that

$$\lambda v_0 = \pi(H)v_0 = \pi(X)\pi(Y)v_0 - \pi(Y)\pi(X)v_0 = \pi(X)v_1.$$

So the formula holds true for $i = 0$. Assuming that it holds for some $i \geq 0$, since $v_{i+1} \in V_{\lambda-2i-2}$, we have

$$(\lambda - 2i - 2)v_{i+1} = \pi(H)v_{i+1} = \pi(X)v_{i+2} - \pi(Y)\pi(X)v_{i+1} = \pi(X)v_{i+2} - (i+1)(\lambda - i)v_{i+1}.$$

This shows that $\pi(X)v_{i+2} = \alpha v_{i+1}$, with $\alpha = \lambda - 2(i+2) + (i+1)(\lambda - i) = (i+2)\lambda - (i+1)(i+2)$. This proves the claim.

The claim implies in particular that $\pi(H)$ is actually diagonalizable over k, with eigenvalues $\lambda - 2i$, $i = 1, \ldots, n$. Moreover the formula applied to $i = n$ shows that $0 = (n+1)(\lambda - n)v_n$. Since $v_n \neq 0$, this shows that necessarily, $\lambda = n$ and the eigenvalues of $\pi(H)$ are real. Then one easily verifies that the representation π is isomorphic with π_n.

It remains to treat the general case, where k is not necessarily algebraically closed. If π is a representation of $\mathfrak{sl}_2(k)$ into $\mathcal{L}_k(V)$, we may view it as a representation into $\mathcal{L}_K(V \otimes_k K)$, where K is the algebraic closure of k. Then we reproduce the above argument to deduce in particular that H is diagonalizable over K, with eigenvalues inside $\mathbb{Q} \subset k$. So it is diagonalizable over k, and now we see that the whole argument is applicable on k, giving the result. □

As we saw in the above proofs, the eigenspaces of the semi-simple element H play a crucial role in understanding a given irreducible representation. And since X and Y shift these eigenspaces, the data of a single eigenvalue matters. The largest eigenvalue of H in V is called the *weight* of the representation. So V_d is the (unique) representation of weight d.

We shall see a similar description for irreducible representations of arbitrary semi-simple Lie algebras.

4. $\mathfrak{sl}_2$-triples and Jacobson-Morozov theorem

As we saw above, in *any* representation π of $\mathfrak{sl}_2(k)$, the element $\pi(X)$ is nilpotent. Conversely, we will show that any nilpotent element in a semi-simple Lie algebra $\mathfrak{g}$ over k is of the form $\pi(X)$ for some Lie algebra homomorphism $\pi : \mathfrak{sl}_2(k) \to \mathfrak{g}$. Note that such a Lie algebra homomorphism is completely described by the images $\pi(H)$, $\pi(X)$ and $\pi(Y)$. So existence of π amounts to finding $\mathfrak{sl}_2$-triples in $\mathfrak{g}$ in the following sense.

Definition 4.23. An $\mathfrak{sl}_2$-*triple* in a Lie algebra $\mathfrak{g}$ is a triple of elements (H, X, Y) such that $H = [X, Y]$, $[H, X] = 2X$ and $[H, Y] = -2Y$. In other words H, X and Y span a copy of $\mathfrak{sl}_2(k)$ in $\mathfrak{g}$ and the notations H, X, Y are consistent with the previous section.

Theorem 4.24 (Jacobson, Morozov). *For any nilpotent element X in a semi-simple Lie algebra $\mathfrak{g}$, there exists $H, Y \in \mathfrak{g}$ such that (H, X, Y) is an $\mathfrak{sl}_2$-triple.*

Observe that in an $\mathfrak{sl}_2$-triple, (H, X, Y), $-2X = [X, H] = \mathrm{ad}(X)^2(Y)$. So a first step towards constructing H and Y will be to prove that X belongs to the range of $\mathrm{ad}(X)^2$. We start with two lemmas.

LEMMA 4.25. *Take two endomorphisms A, B of a finite dimensional vector space V, with A nilpotent. If $[A, [A, B]] = 0$ then the product AB is nilpotent.*

PROOF. This purely algebraic lemma seems very hard to prove using only elementary algebraic manipulations. Instead, we will use eigenvalues considerations. Note that we may pass to the algebraic closure of k to solve it. We need to show that the only eigenvalue of AB is 0.

Claim 1. The commutator $C := [A, B] \in \mathcal{L}(V)$ is nilpotent.

Observe that for all $i \geq 0$, we have $C^{i+1} = C^i(AB - BA) = A(C^iB) - (C^iB)A$, because C commutes with A. Hence C^{i+1} is expressed as a commutator, showing that it has trace 0. Since this is true for all i, we conclude that C is nilpotent.

Claim 2. For all $i \geq 1$, we have $[A^i, B] = iCA^{i-1}$.

This is easily verified by induction.

Now take an eigenvalue λ of AB and a corresponding eigenvector $v \in V$. Denote by $n \geq 0$ the smallest integer such that $A^{(n+1)}v = 0$. Applying claim 2 with $i = n + 1$, we get:

$$(n + 1)CA^nv = A^{n+1}Bv - BA^{n+1}v = A^n(\lambda v).$$

This shows that $\lambda/(n + 1)$ is an eigenvalue of C, and hence that $\lambda = 0$, by Claim 1. □

LEMMA 4.26. *Assume that $H, X, Y' \in \mathcal{L}(V)$ are three endomorphisms of a finite dimensional vector space V such that $[H, X] = 2X$ and $[X, Y'] = H$. Then H leaves globally invariant the kernel of X and on this subspace all its eigenvalues are non-negative integers.*

PROOF. Since $[H, X] = 2X$, it is obvious that the subspace $W = \mathrm{Ker}(X)$ is H invariant. Assume now that λ is an eigenvalue of H, admitting an eigenvector $v \in W\backslash\{0\}$. Denote by $n \geq 0$ the greatest integer such that $v \in \mathrm{Img}(X^n)$.

Claim 1. Since $v \in \mathrm{Ker}(X)$ such an integer n exists.

Indeed, for each integer i such that $\mathrm{Img}(X^i) \cap \mathrm{Ker}(X) \neq \{0\}$, the surjection $X : \mathrm{Img}(X^i) \to \mathrm{Img}(X^{i+1})$ is not injective. So $\mathrm{rk}(X^{i+1}) < \mathrm{rk}(X^i)$. Since V is finite dimensional this can only happen for finitely many i's.

Claim 2. We have $[Y', X^{i+1}] = -(i + 1)(H - i)X^i$ for all $i \geq 0$.

This is proved by induction. Write $v = X^nu$, for some $u \in V$. Applied to $i = n$, the above formula gives:

$$(n+1)(\lambda-n)v = (n+1)(H-n)X^nu = -[Y', X^{n+1}]u = -Y'Xv+X^{n+1}(Y'v) = X^{n+1}(Y'v).$$

By definition of n, we know that v is not in the range of X^{n+1} so above quantity must be 0, and in particular, $\lambda = n$, proving the lemma. □

We are now ready to prove Jacobson-Morozov theorem.

PROOF OF THEOREM 4.24. We first prove that X is in the range of $\mathrm{ad}(X)^2$. Denote by B the Killing form of $\mathfrak{g}$. Observe that for all $Y, Z \in \mathfrak{g}$, we have

$$B(\mathrm{ad}(X)^2Y, Z) = -B(\mathrm{ad}(X)Y, \mathrm{ad}(X)Z) = B(Y, \mathrm{ad}(X)^2Z).$$

This shows that $\mathrm{ad}(Z)^2$ is self adjoint for the non-degenerate symmetric form B. So its range equal to the orthogonal of its kernel. Take $Z \in \mathrm{Ker}(\mathrm{ad}(X)^2)$. By Lemma 4.25, we know that $\mathrm{ad}(X)\,\mathrm{ad}(Z)$ is nilpotent, and hence $B(X,Z) = \mathrm{Tr}(\mathrm{ad}(X)\,\mathrm{ad}(Z)) = 0$. So we deduce that X belongs to the range of $\mathrm{ad}(X)^2$.

Therefore, we may find $Y' \in \mathfrak{g}$ such that $-2X = \mathrm{ad}(X)^2 Y'$. Setting $H = [X, Y']$, we get $[H, X] = 2X$. It now remains to arrange Y' to ensure the extra relation $[H, Y] = -2Y$.

Denote by $u := [H, Y'] + 2Y'$ and observe that $u \in \mathrm{Ker}(\mathrm{ad}(X))$, thanks to Jacobi identity. Applying Lemma 4.26 to $\mathrm{ad}(X)$, $\mathrm{ad}(Y')$ and $\mathrm{ad}(H)$, we know that $\mathrm{ad}(H) + 2$ is invertible on $\mathrm{Ker}(\mathrm{ad}(X))$. So we may find $Z \in \mathfrak{g}$ such that $[X, Z] = 0$ and $-u = [H, Z] + 2Z$.

The element $Y := Y' + Z$ then satisfies $[X, Y] = [X, Y'] + [X, Z] = H$ and $[X, Y] + 2Y = u - u = 0$. So (H, X, Y) is the desired $\mathfrak{sl}_2$-triple. □

A first immediate consequence of Jacobson-Morozov Theorem is the following.

COROLLARY 4.27. *If $\mathfrak{g}$ is a semi-simple Lie algebra and X is a nilpotent element in $\mathfrak{g}$ then $\pi(X)$ is a nilpotent endomorphism for any finite dimensional representation of $\mathfrak{g}$.*

CHAPTER 5

Real and complex semi-simple Lie algebras

1. Cartan subalgebras and Roots spaces; complex case

In order to understand a semi-simple Lie algebra, it is sufficient to understand how it acts on itself, since we saw that the adjoint representation is injective. As we saw for $\mathfrak{sl}_2(k)$, a key aspect in understanding a given representation is to diagonalize some semi-simple elements. Since we can diagonalize simultaneously semi-simple elements that commute with one another, we introduce the following notion.

DEFINITION 5.1. Given a complex semi-simple Lie algebra $\mathfrak{g}$, a *Cartan subalgebra* is an abelian Lie subalgebra consisting only of semi-simple elements, and which is maximal for these properties.

EXAMPLE 5.2. For $\mathfrak{g} = \mathfrak{sl}_d(\mathbb{C})$, the subalgebra of diagonal matrices with trace 0 is a Cartan subalgebra (of dimension $d-1$).

By a simple dimension argument, we see that Cartan subalgebras always exist. We will see later that any two Cartan subalgebras in a given semi-simple Lie algebra are always conjugate by an automorphism. This will be very important because it implies that all the structure that we unravel from a given Cartan subalgebra is canonical; it depends only on the Lie algebra, up to isomorphism.

Let us fix a Cartan subalgebra in subalgebra $\mathfrak{h}$ in a semi-simple Lie algebra $\mathfrak{g}$.

DEFINITION 5.3. For all linear functional $\alpha : \mathfrak{h} \to \mathbb{C}$, denote by $\mathfrak{g}_\alpha \subset \mathfrak{g}$ the vector subspace

$$\mathfrak{g}_\alpha := \{X \in \mathfrak{g} \mid [H, X] = \alpha(H)X, \text{ for all } H \in \mathfrak{h}\}.$$

When it is non-zero and $\alpha \neq 0$, we say that α is a *root* of $\mathfrak{h}$ and we call $\mathfrak{g}_\alpha$ the *root space* associated with α. We will denote by $\Delta \subset \mathfrak{h}^*$ the set of roots of $\mathfrak{h}$.

We then have the following root space decomposition.

PROPOSITION 5.4. *We have* $\mathfrak{g} = \mathfrak{g}_0 \oplus (\bigoplus_{\alpha\in\Delta} \mathfrak{g}_\alpha)$. (This sum is only meant as vector spaces)

PROOF. This is obvious, since each $\mathrm{ad}(H)$, $H \in \mathfrak{h}$ is diagonalisable, and since $\mathfrak{h}$ is abelian, we may find a linear basis of $\mathfrak{g}$ in which each $\mathrm{ad}(H)$ is diagonal. For each element e of this basis, we thus have $[H, e] = \mathrm{ad}(H)(e) = \alpha(H)e$, for some $\alpha(H) \in \mathbb{C}$, which is uniquely determined, and depends linearly on H. In other words, $e \in \mathfrak{g}_\alpha$, and α is a root of $\mathfrak{h}$ (or $\alpha = 0$)

So we see that the set of $\mathfrak{g}_\alpha$ span $\mathfrak{g}$, for $\alpha \in \Delta \cup \{0\}$. Now let us check that the sum is direct. By contradiction, take elements $X_\alpha \in \mathfrak{g}_\alpha$, $\alpha \in \Delta$ such that $\sum_\alpha X_\alpha = 0$, but not all X_α vanish. Then for all $H \in \mathfrak{h}$, and all integer $n \geq 1$, we have

$$\sum_\alpha \alpha(H)^n X_\alpha = 0.$$

Using Vandermonde determinants, we see that this can happen only if for all $H \in \mathfrak{h}$, there exists two roots $\alpha \neq \beta$ such that $\alpha(H) = \beta(H)$. In other words, we must have $\mathfrak{h} = \bigcup_{\alpha \neq \beta} \mathrm{Ker}(\alpha - \beta)$. This would imply that $\mathfrak{h}$ is a finite union of hyperplanes, a contradiction. □

LEMMA 5.5. *For all $\alpha, \beta \in \{0\} \cup \Delta$, we have $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$.*

PROOF. Take $X \in \mathfrak{g}_\alpha$, $Y \in \mathfrak{g}_\beta$ and $H \in \mathfrak{h}$. By Jacobi identity, we have

$$[H, [X, Y]] = -[X, [Y, H]] - [Y, [H, X]] = \beta(H)[X, Y] + \alpha(H)[X, Y]. \qquad \square$$

LEMMA 5.6. *A Cartan subalgebra $\mathfrak{h}$ in a semi-simple Lie algebra $\mathfrak{g}$ is maximal abelian inside $\mathfrak{g}$. In other words, $\mathfrak{g}_0 = \mathfrak{h}$.*

PROOF. Take $X \in \mathfrak{g}$ an element that commutes with $\mathfrak{h}$. We want to show that $X \in \mathfrak{h}$.

Writing $X = S+N$ for the Jordan decomposition, we observe that both S and N commute with $\mathfrak{h}$. Indeed we know that $\mathrm{ad}(X) = \mathrm{ad}(S) + \mathrm{ad}(N)$ is the Jordan decomposition of $\mathrm{ad}(X)$ inside $\mathcal{L}(\mathfrak{g})$. So $\mathrm{ad}(S)$ and $\mathrm{ad}(N)$ are polynomials in $\mathrm{ad}(X)$ and hence they commute with $\mathrm{ad}(H)$ for all $H \in \mathfrak{h}$. Hence $\mathrm{ad}([H, S]) = \mathrm{ad}([H, N]) = 0$ for all $H \in \mathfrak{h}$. Since ad is injective, we indeed find that $\mathrm{ad}(S)$ and $\mathrm{ad}(N)$ both commute with $\mathfrak{h}$. But the maximality property of $\mathfrak{h}$ implies that $S \in \mathfrak{h}$. So we may as well assume that X is nilpotent.

By Jacobson-Morozov theorem, we know that X is part of an $\mathfrak{sl}_2$-triple (H, X, Y). Decompose $\mathfrak{g} = \mathfrak{g}_0 \oplus (\bigoplus_{\alpha \in \Delta} \mathfrak{g}_\alpha)$ into root spaces, and write $H = \sum_{\alpha \in \{0\} \cup \Delta} H_\alpha$. We have $[H, X] = 2X \in \mathfrak{g}_0$ while

$$[H, X] = \sum_{\alpha \in \{0\} \cup \Delta} [H_\alpha, X].$$

By Lemma 5.5 we know that $[H_\alpha, X] \in \mathfrak{g}_\alpha$. Since the spaces $\mathfrak{g}_\alpha$ are in direct sum this shows that each $[H_\alpha, X]$, $\alpha \in \Delta$ vanishes. Hence $2X = [H, X] = [H_0, X]$. Now write $H_0 = H_s + H_n$ for the Jordan decomposition. Then as before, we know that H_s and H_n both commute with H, and hence $H_s \in \mathfrak{h}$, by maximality of $\mathfrak{h}$. Since by assumption X commutes with $\mathfrak{h}$, we deduce that $[H_s, X] = 0$ and thus $[H_n, X] = 2X$. But if $X \neq 0$, this shows that 2 is an eigenvalue of the nilpotent endomorphism $\mathrm{ad}(H_n)$, a contradiction. □

Since the Killing form of $\mathfrak{g}$ is non-degenerate, we may use it to identify $\mathfrak{h}^*$ with $\mathfrak{h}$. Then each root α may be represented by an element of $\mathfrak{h}$. The following lemma gives details and properties of these representing elements.

LEMMA 5.7. *The restriction of the Killing form B to $\mathfrak{h}$ is non-degenerate, and $\Delta = -\Delta$. For $\alpha \in \Delta$, denote by H_α the element of $\mathfrak{h}$ such that $\alpha(H) = B(H_\alpha, H)$ for all $H \in \mathfrak{h}$. We have the following facts.*

a) For $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_{-\alpha}$, we have $[X, Y] = B(X, Y)H_\alpha$;
b) For all $\alpha \in \Delta$, $\alpha(H_\alpha) \neq 0$;
c) Put $c_\alpha := 2/\alpha(H_\alpha)$ and $H'_\alpha := c_\alpha H_\alpha$. Pick $X_\alpha \in \mathfrak{g}_\alpha$ and $Y_\alpha \in \mathfrak{g}_{-\alpha}$ such that $B(X_\alpha, Y_\alpha) = c_\alpha$. Then $(H'_\alpha, X_\alpha, Y_\alpha)$ is an $\mathfrak{sl}_2$-triple.

PROOF. Take $\alpha, \beta \in \Delta \cup \{0\}$ and $X \in \mathfrak{g}_\alpha$, $Y \in \mathfrak{g}_\beta$. If $\alpha + \beta \neq 0$ then we may find $H \in \mathfrak{h}$ such that $\alpha(H) + \beta(H) \neq 0$. Then by the invariance property of the Killing form we have

$$(\alpha(H) + \beta(H))B(X, Y) = B([H, X], Y) + B(X, [H, Y]) = 0.$$

This shows that $B(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$ as soon as $\alpha + \beta \neq 0$.

Take $\alpha \in \{0\} \cup \Delta$ and $X \in \mathfrak{g}_\alpha$ a non-zero element. Since B is non-degenerate, there exists $Y \in \mathfrak{g}$ such that $B(X,Y) \neq 0$. Decomposing Y in the root space decomposition of $\mathfrak{g}$, the above fact shows that necessarily the component of Y on $\mathfrak{g}_{-\alpha}$ is non-zero. This shows that $\Delta = -\Delta$ and that the restriction of B to $\mathfrak{h}$ is non-degenerate. This proves the first assertion.

a) By Lemma 5.5 we know that $[X,Y] \in \mathfrak{g}_0 = \mathfrak{h}$. For $H \in \mathfrak{h}$, we compute

$$B([X,Y],H) = -B(Y,[X,H]) = \alpha(H)B(X,Y).$$

Since the restriction of B to $\mathfrak{h}$ is non-degenerate this gives the result.

b) Assume by contradiction that $\alpha(H_\alpha) = 0$. Since B is non-degenerate, we may find $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_{-\alpha}$ such that $B(X,Y) = 1$. Then $[H_\alpha, X] = 0 = [H_\alpha, Y]$ and $[X,Y] = H_\alpha$. So H_α, X and Y span a nilpotent Lie subalgebra of $\mathfrak{g}$. By Lie's Theorem, it can be represented as a Lie algebra of upper triangular matrices. Then H_α is the commutator $[X,Y]$, so it is a strictly upper triangular matrix, in this representation. In particular, H_α is a nilpotent element. Since $H_\alpha \in \mathfrak{h}$, it is also semi-simple, so it is equal to 0, a contradiction.

c) Observe that $\alpha(H'_\alpha) = 2$, by our choice of renormalization. Thus $[H'_\alpha, X_\alpha] = 2X_\alpha$, $[H'_\alpha, Y_\alpha] = -2Y_\alpha$ and $[X_\alpha, Y_\alpha] = H'_\alpha$. □

We deduce the following facts.

COROLLARY 5.8. *The roots spaces are one-dimensional. For $\alpha, \beta \in \Delta$ such that $\alpha + \beta \in \Delta$, we have $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \mathfrak{g}_{\alpha+\beta}$.*

PROOF. Take $\alpha \in \Delta$. Denote by $(H'_\alpha, X_\alpha, Y_\alpha)$ the $\mathfrak{sl}_2$-triple given by Lemma 5.7, and denote by $\mathfrak{s}_\alpha \subset \mathfrak{g}$ the Lie subalgebra it generates.

Assume that $\dim(\mathfrak{g}_\alpha) \geq 2$. Then we may find some $X \in \mathfrak{g}_\alpha \setminus \{0\}$ such that $B(X, Y_\alpha) = 0$. In particular, property *a*) of Lemma 5.7 tells us that $[Y_\alpha, X] = 0$. Denote by $E \subset \mathfrak{g}$ the $\mathfrak{s}_\alpha$-submodule generated by X. Since $X \in \mathfrak{g}_\alpha$, X is an eigenvector of H'_α, with eigenvalue 2. At the same time it is in the Kernel of $\mathrm{ad}(Y_\alpha)$. By the classification of representations of $\mathfrak{sl}_2(\mathbb{C})$ (Proposition 4.22), we know that this is impossible. So $\dim(\mathfrak{g}_\alpha) = 1$. This proves the first part of the statement.

For the second part, we know that $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$. Assume by contradiction that $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = 0$ and $\mathfrak{g}_{\alpha+\beta} \neq \{0\}$. Then the direct sum $E' := \bigoplus_{n\in\mathbb{N}} \mathfrak{g}_{\beta-n\alpha}$ is an $\mathfrak{s}_\alpha$-submodule of $E := \bigoplus_{n\in\mathbb{Z}} \mathfrak{g}_{\beta-n\alpha}$. The highest eigenvalue of H'_α in E' is $\beta(H'_\alpha)$. By our classification of representations of $\mathfrak{sl}_2$, we deduce that $\beta(H'_\alpha) \geq 0$. But we may also consider the $\mathfrak{s}_\alpha$-representation on the quotient space E/E'. On this space the smallest eigenvalue of H'_α is $\beta(H'_\alpha) + 2$, with eigenvectors in the (non-trivial) image of $\mathfrak{g}_{\beta+\alpha}$. Proposition 4.22, this forces $\beta(H'_\alpha) + 2 \leq 0$. This is impossible. □

2. Cartan subalgebras and Roots spaces; real forms

In this section we investigate the *real* structure of a complex semi-simple Lie algebra. Of course this will allow us to study real semi-simple Lie algebras. But it will also be useful to derive results on the complex case. For instance, the uniqueness of Cartan subalgebras in complex semi-simple Lie algebras will follow from real considerations.

DEFINITION 5.9. Given a real Lie algebra $\mathfrak{g}_0$, we may construct its *complexification* $\mathfrak{g} = (\mathfrak{g}_0)_\mathbb{C} := \mathfrak{g}_0 \otimes_\mathbb{R} \mathbb{C}$, with the natural $\mathbb{C}$-bilinear bracket. We may also change points of view and start with $\mathfrak{g}$ as initial data; we then say that $\mathfrak{g}_0$ is a *real form* of $\mathfrak{g}$.

If $\mathfrak{g}$ is a complex semi-simple Lie algebra, we say that a real form $\mathfrak{u}$ is a *compact* if its Killing form is negative definite.

REMARK 5.10. Observe that if $\mathfrak{g}_0$ is a real form of a complex Lie algebra $\mathfrak{g}$, then the Killing form B_0 of the real Lie algebra $\mathfrak{g}_0$ is equal to the restriction to $\mathfrak{g}_0$ of the Killing form B of the complex Lie algebra $\mathfrak{g}$. Indeed the trace needed to define the Killing form may be used in an arbitrary basis. So we may use a basis of $\mathfrak{g}_0$ over $\mathbb{R}$, which is at the same time a basis of $\mathfrak{g}$ over $\mathbb{C}$. This simple observation will be useful later. It also shows that $\mathfrak{u}$ is a compact real form if and only if $B(X,X) < 0$ for all $X \in \mathfrak{u} \setminus \{0\}$.

In this section we will prove the following result.

THEOREM 5.11. *Consider a complex semi-simple Lie algebra $\mathfrak{g}$ with a Cartan subalgebra $\mathfrak{h}$. Then $\mathfrak{g}$ admits a compact real form $\mathfrak{u}$ such that $\mathfrak{h} \cap \mathfrak{u}$ is a real form of $\mathfrak{h}$.*

Fix a semi-simple complex Lie algebra $\mathfrak{g}$ and a Cartan subalgebra $\mathfrak{h}$. We freely use the notation from the previous section. In particular we consider the elements H_α, $\alpha \in \Delta$ introduced in Lemma 5.7.

LEMMA 5.12. *The following facts hold true.*

a) The Killing form B is positive definite on $\mathfrak{h}_\mathbb{R} := \sum_{\alpha\in\Delta} \mathbb{R}H_\alpha$.
b) We have $\mathfrak{h} = \mathfrak{h}_\mathbb{R} \oplus i\mathfrak{h}_\mathbb{R}$.

PROOF. *a*) Fix $H \in \mathfrak{h}_\mathbb{R}$. Then we may write $\mathrm{ad}(H)$ as a diagonal matrix in the roots space decomposition $\mathfrak{g} = \mathfrak{g}_0 \oplus (\bigoplus_{\alpha\in\Delta} \mathfrak{g}_\alpha)$. We get

$$B(H,H) = \mathrm{Tr}(\mathrm{ad}(H)^2) = \sum_{\alpha\in\Delta} \alpha(H)^2 = \sum_{\alpha\in\Delta} B(H_\alpha, H)^2. \tag{2.1}$$

Let us check that $B(H_\beta, H_\alpha) = \beta(H_\alpha)$ is real for all $\alpha, \beta \in \Delta$. This will imply the positivity of B on $\mathfrak{h}_\mathbb{R}$. Since $(H'_\alpha, X_\alpha, Y_\alpha)$ as defined in Lemma 5.7.c is an $\mathfrak{sl}_2$-triple, Proposition 4.22 shows that the eigenvalues $\beta(H'_\alpha)$ of H'_α are integers. In particular $\beta(H'_\alpha) \in \mathbb{R}$ for all $\beta \in \Delta$. Since $H'_\alpha = 2H_\alpha/\alpha(H_\alpha)$, it remains to check that $\alpha(H_\alpha)$ is real.

Since $\beta(H'_\alpha) \in \mathbb{R}$ for all $\alpha, \beta \in \Delta$, we deduce from (2.1) that $B(H'_\alpha, H'_\alpha) = \sum_{\beta\in\Delta} \beta(H'_\alpha)^2 \in \mathbb{R}$. Hence

$$\alpha(H_\alpha) = B(H_\alpha, H_\alpha) = \frac{1}{4}\alpha(H_\alpha)^2 B(H'_\alpha, H'_\alpha).$$

We conclude that indeed, $\alpha(H_\alpha) \in \mathbb{R}$. So B is positive on $\mathfrak{h}_\mathbb{R}$.

Moreover if an element $H \in \mathfrak{h}$ satisfies $B(H_\alpha, H) = 0$ for all $\alpha \in \Delta$, then α is in the kernel of all roots, and hence $\mathrm{ad}(H) = 0$. Since $\mathfrak{g}_\mathbb{C}$ has trivial center, we conclude that $H = 0$. This implies two things: B restricts to a positive definite bilinear form on $\mathfrak{h}_\mathbb{R}$, and $\mathfrak{h}_\mathbb{R}$ generates $\mathfrak{h}$ over $\mathbb{C}$.

b) Since B is positive definite on $\mathfrak{h}_\mathbb{R}$ it is negative definite on $i\mathfrak{h}_\mathbb{R}$. Hence $\mathfrak{h}_\mathbb{R} \cap i\mathfrak{h}_\mathbb{R} = \{0\}$. □

For each $\alpha \in \Delta$, we choose $E_\alpha \in \mathfrak{g}_\alpha$ and $E_{-\alpha} \in \mathfrak{g}_{-\alpha}$ such that $[E_\alpha, E_{-\alpha}] = H_\alpha$. By Corollary 5.8, for all $\alpha, \beta \in \Delta$ such that $\alpha + \beta \neq 0$, we may find scalars $N_{\alpha,\beta} \in \mathbb{C}$ such that

$$[E_\alpha, E_\beta] = N_{\alpha,\beta} E_{\alpha+\beta}$$

(and we set $N_{\alpha,\beta} = 0$ is $\alpha + \beta \notin \Delta$).

LEMMA 5.13. *The numbers $N_{\alpha,\beta}$ satisfy the following relations, for all $\alpha, \beta, \gamma, \delta \in \Delta$.*

a) $N_{\beta,\alpha} = -N_{\beta,\alpha}$.
b) If $\alpha+\beta+\gamma=0$ *then* $N_{\alpha,\beta} = N_{\beta,\gamma} = N_{\gamma,\alpha}$*;*
c) If $\alpha+\beta$ *is still a root, then* $N_{\alpha,\beta}N_{-\alpha-\beta} < 0$.
d) If $\alpha,\beta,\gamma,\delta$ *are pairwise non-proportional and* $\alpha+\beta+\gamma+\delta=0$*, then*

$$N_{\alpha,\beta}N_{\gamma,\delta} + N_{\beta,\gamma}N_{\alpha,\delta} + N_{\gamma,\alpha}N_{\beta,\delta} = 0.$$

PROOF. Fact *a*) is straightforward.

b) If $\alpha+\beta+\gamma=0$ then $H_\alpha + H_\beta + H_\gamma = 0$.

$$\begin{aligned} 0 &= [E_\alpha,[E_\beta,E_\gamma]] + [E_\beta,[E_\gamma,E_\alpha]] + [E_\gamma,[E_\alpha,E_\beta]] \\ &= N_{\beta,\gamma}[E_\alpha,E_{-\alpha}] + N_{\gamma,\alpha}[E_\beta,E_{-\beta}] + N_{\alpha,\beta}[E_\gamma,E_{-\gamma}] \\ &= N_{\beta,\gamma}H_\alpha + N_{\gamma,\alpha}H_\beta + N_{\alpha,\beta}H_\gamma. \end{aligned}$$

So we find that $(N_{\beta,\gamma}-N_{\gamma,\alpha})H_\beta + (N_{\beta,\gamma}-N_{\alpha,\beta})H_\gamma = 0$. We will see later that the only roots proportional to a given root α are α and $-\alpha$. So the relation $\alpha+\beta+\gamma=0$ implies that α,β,γ are pairwise not proportional. We conclude that $N_{\alpha,\beta} = N_{\beta,\gamma} = N_{\gamma,\alpha}$.

c) We will use the $\mathfrak{sl}_2$-triple $(H'_\alpha, X_\alpha, Y_\alpha)$ constructed in Lemma 5.7.c.

We derive from the proof of Proposition 4.22 that for each representation π of $\mathfrak{sl}_2(k)$ and each eigenvector v of the semi-simple element $\pi(H)$, we have $\pi(X)\pi(Y)v = cv$ where c is a non-negative integer. Indeed the claim from that proof states that $\pi(X)\pi(Y)v_i = (i+1)(n-i)v_i$ for all $0 \le i \le n$. Moreover, if $\pi(Y)v_i \neq 0$, then $i < n$ and the constant c is strictly positive in this case.

In our situation, since we know that $\alpha+\beta$ is still a root, so is $\gamma := -\alpha-\beta$. This shows that $E_{-\beta}$ is an eigenvector of $\mathrm{ad}(H'_\alpha)$ and $[Y_\alpha, E_{-\beta}] \in \mathfrak{g}_\gamma$ is non-zero, by Corollary 5.8. So we conclude that $[X_\alpha,[Y_\alpha,E_{-\beta}]] = cE_{-\beta}$ for some strictly positive integer c.

Now since the root saves are one dimensional, we may find scalars $a,b \in \mathbb{C}$ such that $E_\alpha = aX_\alpha$ and $E_{-\alpha} = bY_\alpha$. Note moreover that $B(E_\alpha,E_{-\alpha}) = 1$, because of the relation $[E_\alpha,E_{-\alpha}] = H_\alpha$ and Lemma 5.7.a. On the other hand $B(X_\alpha,Y_\alpha) = 2/\alpha(H_\alpha) = 2/B(H_\alpha,H_\alpha) > 0$ (by Lemma 5.12). In conclusion, we see that the product ab is positive.

Thus $[E_\alpha,[E_{-\alpha},E_{-\beta}]] = abcE_{-\beta}$, with $abc > 0$. So we get $N_{-\alpha,-\beta}N_{\alpha,\gamma} > 0$. By a and b we know that $N_{\alpha,\gamma} = -N_{\gamma,\alpha} = -N_{\alpha,\beta}$. So we indeed find $N_{\alpha,\beta}N_{-\alpha-\beta} < 0$.

d) First assume that $\alpha+\beta \in \Delta$. Then we have $[[E_\alpha,E_\beta],E_\gamma] = N_{\alpha,\beta}N_{\alpha+\beta,\gamma}E_{-\delta}$. Apply item b to the roots $\alpha+\beta$, γ, δ, so that $N_{\alpha+\beta,\gamma} = N_{\gamma,\delta}$. We conclude

$$[[E_\alpha,E_\beta],E_\gamma] = N_{\alpha,\beta}N_{\gamma,\delta}E_{-\delta}.$$

Note that this relation remains true if $\alpha+\beta$ is not a root, because in this case both sides are equal to 0. We may permute cyclically the indices α,β,γ in this relation and use Jacobi identity to get the desired relation. □

LEMMA 5.14. *We may choose the elements* $E_\alpha \in \mathfrak{g}_\alpha$ *in such a way that* $N_{\alpha,\beta} = -N_{-\alpha,-\beta}$*. In this case the numbers* $N_{\alpha,\beta}$ *are reals.*

PROOF. Note that if $N_{\alpha,\beta} = -N_{-\alpha,-\beta}$, then Lemma 5.13.c implies that $N_{\alpha,\beta}^2 > 0$, and so $N_{\alpha,\beta}$ is a real number. So we only need to prove the first statement.

First we choose a total order on the dual space $\mathfrak{h}_\mathbb{R}^*$ such that the sum of two positive elements is still positive and such that $a < b$ if and only if $b-a > 0$ for all $a,b \in \mathfrak{h}_\mathbb{R}^*$. This can be done by choosing an arbitrary isomorphism $\mathfrak{h}_\mathbb{R}^* \simeq \mathbb{R}^d$ (i.e. by fixing a basis of $\mathfrak{h}_\mathbb{R}^*$) and by pulling back the lexicographic order on $\mathbb{R}^d$. By Lemma 5.12, we may view the set of roots Δ as a subset of $\mathfrak{h}_\mathbb{R}^*$, and we define the set of *positive roots* $\Delta_+ := \{\alpha \in \Delta \mid \alpha > 0\}$.

For any positive root $\rho \in \Delta_+$, we set $\Delta_\rho := \{\alpha \in \Delta \mid -\rho \leq \alpha \leq \rho\}$. We show by induction on the cardinality of Δ_ρ that we may choose the vectors $E_\alpha \in \mathfrak{g}_\alpha$ so that

$$N_{\alpha,\beta} = -N_{-\alpha,-\beta} \text{ for all } \alpha, \beta \in \Delta_\rho \text{ such that } \alpha + \beta \in \Delta_\rho. \tag{2.2}$$

Since Δ is totally ordered and finite, we know that there is some $\rho_0 \in \Delta_+$ such that $\Delta_{\rho_0} = \{\pm\rho_0\}$. In this case we set E_{ρ_0} arbitrarily and the vector $E_{-\rho_0}$ is automatically determined by the relation $[E_{\rho_0}, E_{-\rho_0}] = H_{\rho_0}$ (which amounts to $B(E_{\rho_0}, E_{-\rho_0}) = 1$).

Assume now that $\rho > \rho_0$ and that we have constructed all the vectors E_α, for $\alpha \in \Delta_{\rho'}$ with $\rho' < \rho$ in such a way that (2.2) holds true for ρ'. We need to construct E_ρ and $E_{-\rho}$.

Case 1. If ρ is not a sum of two positive roots, then we choose E_ρ arbitrarily and $E_{-\rho_0}$ is automatically determined by the relation $B(E_\rho, E_{-\rho}) = 1$.

Case 2. Otherwise, we may fix two positive roots $\gamma, \delta \in \Delta_+$ such that $\rho = \gamma + \delta$. Then we set $E_\rho := c[E_\gamma, E_\delta]$ and $E_{-\rho} := -c[E_{-\gamma}, E_{-\delta}]$, where c is chosen so that $B(E_\rho, E_{-\rho}) = 1$. In other words, we are forcing the relation $N_{-\gamma,-\delta} = -N_{\gamma,\delta}$. Note that this makes sense, since γ and δ are in Δ_ρ.

In both cases let us now verify equation (2.2). Take $\alpha, \beta \in \Delta_\rho$ such that $\alpha + \beta \in \Delta_\rho$. If none of $\alpha, \beta, \alpha + \beta$ equals $\pm\rho$ then the equation holds by induction.

If $\alpha + \beta = \rho$ then since $\alpha, \beta \leq \rho$, this forces both α and β to be positive. So this situation only appears in case 2. We then apply Lemma 5.13.d to $\alpha, \beta, -\gamma, -\delta$ and $-\alpha, -\beta, \gamma, \delta$, and get:

$$N_{\alpha,\beta}N_{-\gamma,-\delta} + N_{\beta,-\gamma}N_{\alpha,-\delta} + N_{-\gamma,\alpha}N_{\beta,-\delta} = 0.$$
$$N_{-\alpha,-\beta}N_{\gamma,\delta} + N_{-\beta,\gamma}N_{-\alpha,\delta} + N_{\gamma,-\alpha}N_{-\beta,\delta} = 0.$$

But by our induction assumption, we find $N_{\beta,-\gamma} = -N_{-\beta,\gamma}$, $N_{\alpha,-\delta} = -N_{-\alpha,\delta}$, $N_{-\gamma,\alpha} = -N_{\gamma,-\alpha}$ and $N_{\beta,-\delta} = -N_{-\beta,\delta}$. Since by construction we also forced the relation $N_{-\gamma,-\delta} = -N_{\gamma,\delta}$, there is no other choice but to have $N_{\alpha,\beta} = -N_{-\alpha,-\beta}$, as wanted.

The case where $\alpha + \beta = -\rho$ follows, since $-\alpha - \beta = \rho$.

Assume now that $\alpha = \rho$. Then we may write $\alpha' + \beta' = \rho$, with $\alpha' = \alpha + \beta$ and $\beta' = -\beta$. The following discussion gives $N_{\alpha',\beta'} = -N_{-\alpha',\beta'}$. We may apply Lemma 5.13.b to the triples $(\alpha', \beta', -\alpha)$ and $(-\alpha', -\beta', \alpha)$, to get $N_{\alpha',\beta'} = N_{\beta',-\alpha} = N_{-\beta,-\alpha}$ and $N_{-\alpha',-\beta'} = N_{-\beta',\alpha} = N_{\beta,\alpha}$. The desired equality follows. All the remaining cases $\alpha = -\rho$, and $\beta = \pm\rho$ follow by taking opposite or buy symmetry. So the induction step is verified, which concludes our proof. □

We can now easily deduce Theorem 5.11.

Proof of Theorem 5.11. We choose the elements E_α given by the previous lemma. We set

$$\mathfrak{u} := i\mathfrak{h}_\mathbb{R} \oplus \left(\bigoplus_{\alpha\in\Delta} \mathbb{R}i(E_\alpha + E_{-\alpha})\right) \oplus \left(\bigoplus_{\alpha\in\Delta} \mathbb{R}(E_\alpha - E_{-\alpha})\right).$$

This is a real Lie subalgebra of $\mathfrak{g}$. For instance we have

$$[i(E_\alpha + E_{-\alpha}), i(E_\beta + E_{-\beta})] = -N_{\alpha,\beta}(E_{\alpha+\beta} - E_{-\alpha-\beta}) - N_{\alpha,-\beta}(E_{\alpha-\beta} - E_{-\alpha+\beta}) \in \mathfrak{u}.$$
$$[iH_\alpha, i(E_\beta + E_{-\beta})] = -\beta(H_\alpha)(E_\beta - E_{-\beta}) \in \mathfrak{u},$$

because we saw that $\beta(H_\alpha) = B(H_\beta, H_\alpha) \in \mathbb{R}$. etc.

The Killing form is negative definite on $\mathfrak{u}$. Indeed, we know that B is negative definite on $i\mathfrak{h}_\mathbb{R}$. As we observed above, $B(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$ if $\alpha + \beta \neq 0$. And we also have $B(E_\alpha, E_{-\alpha}) = 1$ by construction. So one easily verifies that B is definite negative on each summand

$\mathbb{R}i(E_\alpha+E_{-\alpha})$ or $\mathbb{R}(E_\alpha-E_{-\alpha})$ and that all these direct summand are pairwise orthogonal. So we conclude that B is definite negative on the whole of $\mathfrak{u}$. In particular, we have that $\mathfrak{u}\cap i\mathfrak{u}=\{0\}$. We conclude that $\mathfrak{g}=\mathfrak{u}\oplus i\mathfrak{u}$, and $\mathfrak{u}$ is indeed a compact real form of $\mathfrak{g}$. We moreover have $\mathfrak{h}\cap\mathfrak{u}=i\mathfrak{h}_{\mathbb{R}}$, which is a real form in $\mathfrak{h}$. □

EXAMPLE 5.15. For the Lie algebra $\mathfrak{g}=\mathfrak{sl}_2(\mathbb{C})$, we may proceed to the above construction for the Cartan subalgebra consisting of diagonal matrices with trace 0. Note that the condition from Lemma 5.14 is empty since there are only two roots α and $-\alpha$, which are opposite from one another. So we may take $E_\alpha=X$ and $E_{-\alpha}=Y$, where (H,X,Y) is the standard $\mathfrak{sl}_2$-triple. We get the compact real form $\mathfrak{su}(2)$, with basis

$$iH=\begin{pmatrix}i&0\\0&-i\end{pmatrix},\quad i(X+Y)=\begin{pmatrix}0&i\\i&0\end{pmatrix},\quad X-Y=\begin{pmatrix}0&1\\-1&0\end{pmatrix}.$$

3. Cartan decompositions, compact Lie groups

In this section we shall prove that in a semi-simple complex Lie algebra any two compact real forms are always conjugate. Our main tool will be Cartan involutions. A Cartan involution contains similar information as a compact real form but since it is an automorphism it is more handy to use (for instance we may apply functional calculus on it).

DEFINITION 5.16. A *Cartan involution* of a real semi-simple Lie algebra $\mathfrak{g}_0$ is an automorphism $\theta\in\mathrm{Aut}(\mathfrak{g}_0)$ such that $\theta^2=\mathrm{id}$ and the symmetric bilinear form $B_\theta:(X,Y)\mapsto -B_0(\theta(X),Y)$ is positive definite. Here B_0 is the Killing form of $\mathfrak{g}_0$.

LEMMA 5.17. *To any Cartan involution θ on a real semi-simple Lie algebra $\mathfrak{g}_0$ corresponds a* Cartan decomposition $\mathfrak{g}_0=\mathfrak{k}\oplus\mathfrak{q}$, *where* $\mathfrak{k}=\{X\in\mathfrak{g}\mid\theta(X)=X\}$ *and* $\mathfrak{q}=\{X\in\mathfrak{g}\mid\theta(X)=-X\}$.

PROOF. Since θ is a linear involution, this follows from basic linear algebra. □

EXAMPLE 5.18. If $\mathfrak{g}=\mathfrak{sl}_d(\mathbb{R})$, then we may take $\theta(X)=-{}^tX$, and we see that $\mathfrak{k}$ consists of the skew-symmetric matrices and $\mathfrak{q}$ consists of symmetric matrices.

Assume that $\mathfrak{g}$ is a complex semi-simple Lie algebra. Fix a compact real form $\mathfrak{u}$ of $\mathfrak{g}$, so that $\mathfrak{g}=\mathfrak{u}\oplus i\mathfrak{u}$. View $\mathfrak{g}$ as a real Lie algebra and define $\theta_{\mathfrak{u}}\in\mathrm{Aut}_{\mathbb{R}}(\mathfrak{g})$ to be the identity on $\mathfrak{u}$ and $\theta=-\mathrm{id}$ on $i\mathfrak{u}$. For all $X\in\mathfrak{g}\neq\{0\}$ we may write $X=a+ib$, with $a,b\in\mathfrak{u}$ and get

$$B(\theta_{\mathfrak{u}}(X),X)=B(a-ib,a+ib)=B(a,a)+B(b,b)<0,$$

where B denotes the Killing form of $\mathfrak{g}$. In other words, the hermitian form $(X,Y)\mapsto -B(\theta_{\mathfrak{u}}(X),Y)$ is positive definite on $\mathfrak{g}$.

As we said above, Cartan involutions contain similar information as compact real forms. More precisely,

LEMMA 5.19. *Consider a semi-simple real Lie algebra $\mathfrak{g}_0$ and an involutive automorphism $\theta\in\mathrm{Aut}_{\mathbb{R}}(\mathfrak{g}_0)$. Then θ is a Cartan involution if and only if there exists a compact real form $\mathfrak{u}$ of the complexification $\mathfrak{g}=\mathfrak{g}_0\oplus i\mathfrak{g}_0$ such that $\theta_{\mathfrak{u}}$ as defined above leaves $\mathfrak{g}_0$ globally invariant and $\theta_{\mathfrak{u}|\mathfrak{g}_0}=\theta$. Moreover, $\mathfrak{u}$ is uniquely determined by θ.*

PROOF. Assume first that θ is a Cartan involution and denote by $\mathfrak{g}_0=\mathfrak{k}\oplus\mathfrak{q}$ the corresponding Cartan decomposition. Then we know that $\mathfrak{k}$ is a Lie subalgebra of $\mathfrak{g}_0$, $[\mathfrak{k},\mathfrak{q}]\subset\mathfrak{q}$, $[\mathfrak{q},\mathfrak{q}]\subset\mathfrak{k}$. Therefore, $\mathfrak{u}:=\mathfrak{k}\oplus i\mathfrak{q}$ is a Lie subalgebra. Moreover, since the

Killing form B_0 of $\mathfrak{g}_0$ is positive on $\mathfrak{q}$ and negative on $\mathfrak{k}$, we have $B_0(\mathfrak{k},\mathfrak{q})=0$. Denote by B the Killing form of $\mathfrak{g}$. In view of Remark 5.10, for all $X\in\mathfrak{k}$, $Y\in\mathfrak{q}$ we have

$$B(X+iY,X+iY)=B_0(X,X)-B_0(Y,Y).$$

Thus B is negative definite on $\mathfrak{u}$, which shows that $\mathfrak{u}$ is a compact form of $\mathfrak{g}$. Finally the involution $\theta_{\mathfrak{u}}$ is equal to id on $\mathfrak{k}$ and to $-\mathrm{id}$ on $\mathfrak{q}$. So it leaves $\mathfrak{g}$ globally invariant and coincide with θ on it.

Conversely assume that θ is the restriction of some $\theta_{\mathfrak{u}}$ to $\mathfrak{g}_0$. We use Remark 5.10 and denote by B_0 (resp. B) the Killing form of $\mathfrak{g}_0$ (resp. $\mathfrak{g}$, $\mathfrak{u}$). Fix $X\in\mathfrak{g}_0$ and view it inside $\mathfrak{g}$. We have $B_0(\theta(X),X)=B(\theta_{\mathfrak{u}}(X),X)$. As we already observed, this quantity is always negative if $X\neq 0$. Hence θ is a Cartan involution of $\mathfrak{g}_0$.

Note that $\mathfrak{u}$ is uniquely determined by $\theta_{\mathfrak{u}}$ (it is its fixed point set). Moreover, $\theta_{\mathfrak{u}}$ is the unique anti-linear extension of θ to $\mathfrak{g}$. This proves the moreover part. □

THEOREM 5.20. *Any semi-simple real Lie algebra $\mathfrak{g}_0$ admits a Cartan involution. Moreover, any two Cartan involutions θ and θ' on $\mathfrak{g}_0$ are conjugate by an automorphism: there exists $\phi\in\mathrm{Aut}(\mathfrak{g}_0)$ such that $\theta'=\phi\theta\phi^{-1}$.*

PROOF. By Theorem 5.11, we know that the complexification $\mathfrak{g}$ of $\mathfrak{g}_0$, admits a compact real form $\mathfrak{u}$. The problem is that the corresponding involution $\theta:=\theta_{\mathfrak{u}}$ needs not preserve the real form $\mathfrak{g}_0$. So the task is to find an appropriate conjugate θ' of θ which satisfies this condition.

Denote by τ the complex conjugation on $\mathfrak{g}$ with respect to the real form $\mathfrak{g}_0$. In other words $\tau=\mathrm{id}$ on $\mathfrak{g}_0$ and $\tau=-\mathrm{id}$ on $i\mathfrak{g}_0$. Define $N:=\tau\theta$ and observe that N is in fact $\mathbb{C}$-linear because both τ and θ are anti-linear. As usual, denote by B the Killing form of $\mathfrak{g}$. For all $X,Y\in\mathfrak{g}$, we have

$$B_\theta(NX,Y)=-B(\theta NX,Y)=-B(N^{-1}\theta X,Y)=-B(\theta X,NY)=B_\theta(X,NY).$$

Thus N is self-adjoint with respect to the positive definite hermitian form B_θ. Hence it is diagonalisable, with real eigenvalues. In particular, $P:=N^2$ is an automorphism of the complex Lie algebra $\mathfrak{g}$ and there exists a basis $X_1,\dots,X_n$ of $\mathfrak{g}$ such that $PX_i=\lambda_iX_i$ for some $\lambda_1,\dots,\lambda_n\in\mathbb{R}_+^*$.

For each i,j, the Lie bracket $[X_i,X_j]$ may be expressed as a linear combination of the X_k's:

$$[X_i,X_j]=\sum_{k=1}^n c_{i,j}^kX_k.$$

Since P is an automorphism of $\mathfrak{g}_{\mathbb{C}}$, we have $\lambda_i\lambda_jc_{i,j}^k=\lambda_kc_{i,j}^k$ for all i,j,k. This relation implies that for all $t\in\mathbb{R}$, we also have $\lambda_i^t\lambda_j^tc_{i,j}^k=\lambda_k^tc_{i,j}^k$ for all i,j,k. In other words, for all $t\in\mathbb{R}$, the linear endomorphism P^t of $\mathfrak{g}_{\mathbb{C}}$ is again an automorphism of $\mathfrak{g}_{\mathbb{C}}$. We know that P^t commutes with N and since $\theta N\theta^{-1}=N^{-1}$ we have $\theta P\theta^{-1}=P^{-1}$ and further $\theta P^t\theta^{-1}=P^{-t}$ for all $t\in\mathbb{R}$.

We set $\phi:=P^{1/4}$, and $\theta':=\phi\,\theta\,\phi^{-1}$. We have $\theta'\tau=P^{1/4}\theta P^{-1/4}\tau=P^{1/2}\theta\tau=P^{1/2}N^{-1}$. But observe that $P^{1/2}N^{-1}=NP^{-1/2}$, because this endomorphism is diagonal with eigenvalues ±1. So we obtain

$$\theta'\tau=P^{1/2}N^{-1}=NP^{-1/2}=\tau\theta'.$$

Therefore θ' commutes with τ, so it leaves its eigenspaces $\mathfrak{g}_0$ and $i\mathfrak{g}_0$ globally invariant. Moreover $\theta'=\phi\,\theta_{\mathfrak{u}}\,\phi^{-1}=\theta_{\phi(\mathfrak{u})}$. Since ϕ is an automorphism of $\mathfrak{g}$, $\phi(\mathfrak{u})$ is a compact real form of $\mathfrak{g}$, and so the restriction $\theta'_{|\mathfrak{g}_0}$ is a Cartan involution by Proposition 5.19.

For the moreover part, assume that θ and θ' are two Cartan involutions of $\mathfrak{g}_0$. Extend these two maps to anti-linear maps on the complexification $\mathfrak{g}$, still denoted by θ and θ'. By Lemma 5.19, these maps are of the form $\theta_{\mathfrak{u}}$ for some compact real form $\mathfrak{u}$ of $\mathfrak{g}$. Thus the hermitian form $B_\theta : (X, Y) \mapsto -B(\theta X, Y)$ is positive definite on $\mathfrak{g}$. As above, we may define an automorphism $P = (\theta\theta')^2$ which is positive with respect to this positive definite form. The same computation gives that $\phi := P^{1/4}$ is still an automorphism of $\mathfrak{g}$ and θ and $\theta_1 := \phi\,\theta'\,\phi^{-1}$ commute.

In particular we may simultaneously diagonalize these two involutions, and get

$$\mathfrak{g} = (\mathfrak{a} \cap \mathfrak{a}_1) \oplus (\mathfrak{a} \cap \mathfrak{b}_1) \oplus (\mathfrak{b} \cap \mathfrak{a}_1) \oplus (\mathfrak{b} \cap \mathfrak{b}_1),$$

where $\mathfrak{a} = \{X \in \mathfrak{g} \mid \theta(X) = X\}$, $\mathfrak{b} = \{X \in \mathfrak{g} \mid \theta(X) = -X\}$ and $\mathfrak{a}_1$, $\mathfrak{b}_1$ are defined similarly for θ_1. In particular, the Killing form B is positive definite on $\mathfrak{a}$ and $\mathfrak{a}_1$ and negative definite on $\mathfrak{b}$ and $\mathfrak{b}_1$. This shows that $(\mathfrak{a} \cap \mathfrak{b}_1) = (\mathfrak{b} \cap \mathfrak{a}_1) = \{0\}$, i.e. $\theta = \theta_1$.

This conjugation was taking place inside $\mathfrak{g}$, but note that ϕ preserves $\mathfrak{g}_0$, because θ and θ' do so, as well as P. So we may restrict this conjugacy to $\mathfrak{g}_0$ to get the result. □

The above proof has the following corollary.

COROLLARY 5.21. *On a complex semi-simple Lie algebra there is a unique compact real form, up to automorphism.*

We can now state one of the main consequences of compact real forms.

THEOREM 5.22. *The map $U \mapsto \mathfrak{u}_{\mathbb{C}} := Lie(U)_{\mathbb{C}}$ is a bijection from the set of (isomorphism classes of) compact connected real Lie groups with trivial center onto the set of (isomorphism classes of) semi-simple complex Lie algebras.*

PROOF. First we prove that if U is a compact connected Lie group with trivial center then its Lie algebra $\mathfrak{u}$ is a compact, semi-simple real Lie algebra. The adjoint group $\mathrm{Ad}(U)$ is again compact inside $\mathcal{L}(\mathfrak{u})$. Thus it preserves an inner product B_0 of $\mathfrak{u}$. We may then differentiate the relation $B_0(gY, gZ) = B_0(Y, Z)$ with respect to $g \in U$ and find $B_0(\mathrm{ad}(X)Y, Z) + B_0(Y, \mathrm{ad}(X)Z) = 0$ for all $X, Y, Z \in \mathfrak{u}$. In other words, $\mathrm{ad}(X)$ is anti-symmetric with respect to B_0, and thus $\mathrm{Tr}(\mathrm{ad}(X)^2) < 0$ for all $X \in \mathfrak{u} \setminus \{0\}$. This shows that the Killing form of $\mathfrak{u}$ is negative definite, and thus $\mathfrak{u}$ is a compact semi-simple real Lie algebra. In particular $\mathfrak{u}_{\mathbb{C}}$ is indeed semi-simple.

We know that two connected Lie groups with trivial center having isomorphic Lie algebras are isomorphic. So the map $U \mapsto \mathfrak{u}$ is injective. Moreover Corollary 5.21 implies that the map $\mathfrak{u} \mapsto \mathfrak{u}_{\mathbb{C}}$ is injective.

It remains to check that every semi-simple complex Lie algebra $\mathfrak{g}$ arises as the complexification of the Lie algebra of a compact Lie group with trivial center. By Theorem 5.11 we know that $\mathfrak{g}$ admits a compact real form $\mathfrak{u}$. Denote by $U := \mathrm{Aut}(\mathfrak{u})_e$ the connected component of the identity of the Lie group $\mathrm{Aut}(\mathfrak{u})$. By Lemma 2.23, its Lie algebra is $\mathrm{Der}(\mathfrak{u})$, and since $\mathfrak{u}$ is semi-simple we know that this Lie algebra is isomorphic to $\mathfrak{u}$ via the adjoint representation. Finally, observe that U is compact because it preserves the Killing form of $\mathfrak{u}$. □

4. Uniqueness of Cartan subalgebras

We now turn to the questions of uniqueness of Cartan subalgebras. The notion of a Cartan subalgebra may also be considered in the real setting.

DEFINITION 5.23. A Cartan subalgebra $\mathfrak{h}_0$ of a semi-simple real Lie algebra $\mathfrak{g}_0$ is a maximal subalgebra among abelian subalgebras consisting of semi-simple elements. Exactly as in the complex case, a Cartan subalgebra is in fact maximal abelian inside $\mathfrak{g}_0$.

LEMMA 5.24. *Take a Cartan subalgebra $\mathfrak{h}_0$ of a semi-simple real Lie algebra $\mathfrak{g}_0$. Then*

- *the complexification $\mathfrak{h} = \mathfrak{h}_0 \oplus i\mathfrak{h}_0$ is a Cartan subalgebra of the complexification $\mathfrak{g}$ of $\mathfrak{g}_0$.*
- *There exists a Cartan involution of $\mathfrak{g}_0$ that preserves $\mathfrak{h}$.*

PROOF. The first fact is easy, since $\mathfrak{h}$ consists in semi-simple elements and is maximal abelian. To prove the second fact, take a compact real form $\mathfrak{u}$ of $\mathfrak{g}$ such that $\mathfrak{h} \cap \mathfrak{u}$ is a real form of $\mathfrak{h}$. Then it follows that the involution $\theta_{\mathfrak{u}}$ of $\mathfrak{g}$ preserves $\mathfrak{h}$.

We repeat the procedure of conjugating θ_u to an involution θ of $\mathfrak{g}$ that preserves $\mathfrak{g}_0$. The complex conjugation τ of $\mathfrak{g}$ with respect to $\mathfrak{g}_0$ (i.e. $\tau(X + iY) = X - iY$ for $X, Y \in \mathfrak{g}_0$) also preserves $\mathfrak{h}$. So the operator $P = (\tau\theta_{\mathfrak{u}})^2$ also preserves $\mathfrak{h}$ and hence, this is also the case of $\phi := P^{1/4}$. We conclude that $\theta := \phi\theta_{\mathfrak{u}}\phi^{-1}$ leaves $\mathfrak{h}$ and $\mathfrak{g}_0$ globally invariant. As we saw in the proof of Theorem 5.20, the restriction of θ to $\mathfrak{g}_0$ is a Cartan involution, which preserves $\mathfrak{h}_0 = \mathfrak{h} \cap \mathfrak{g}_0$. □

In a real Lie algebra it makes sense to emphasize semi-simple elements X such that $\mathrm{ad}(X)$ is diagonalizable over $\mathbb{R}$. This leads to the following variant of Cartan subalgebras.

DEFINITION 5.25. A *Cartan subspace* of a semi-simple real Lie algebra $\mathfrak{g}_0$ is a maximal subalgebra among all abelian subalgebras consisting of semi-simple elements admitting real eigenvalues.

Fix a semi-simple real Lie algebra $\mathfrak{g}_0$ and a Cartan subspace $\mathfrak{a}$. Similarly to the complex case, we may decompose $\mathfrak{g}_0$ as a direct sum

$$\mathfrak{g}_0 = \mathfrak{l} \oplus (\bigoplus_{\lambda \in \Sigma} \mathfrak{g}_\lambda),$$

where $\mathfrak{l}$ is the centralizer of $\mathfrak{a}$ in $\mathfrak{g}_0$ and Σ is the set of so-called *restricted roots*, i.e. non-zero linear functionals $\lambda \in \mathfrak{a}^*$ for which the following root space $\mathfrak{g}_\lambda$ is non-zero.

$$\mathfrak{g}_\lambda := \{Y \in \mathfrak{g}_0 \mid \mathrm{ad}(X)Y = \lambda(X)Y, \text{ for all } X \in \mathfrak{a}\}.$$

Fix a Cartan involution θ of $\mathfrak{g}_0$ and denote by $\mathfrak{g}_0 = \mathfrak{k} \oplus \mathfrak{q}$ the corresponding Cartan decomposition. We denote by $K := \{g \in \mathrm{Aut}(\mathfrak{g}_0) \mid \theta g = g\theta\}$. It preserves the positive definite bilinear form B_θ so it is a compact group. Its Lie algebra is $\mathrm{ad}_{\mathfrak{g}_0}(\mathfrak{k}) \subset \mathrm{Der}(\mathfrak{g}_0)$. Denote by K_e its connected component of the identity.

PROPOSITION 5.26. *The following facts hold.*

a) The Cartan subspace $\mathfrak{a}$ in $\mathfrak{g}_0$ is conjugate to one contained in $\mathfrak{q}$
b) Any two Cartan subspaces of $\mathfrak{g}_0$ contained in $\mathfrak{q}$ are conjugate by an element of K_e.

PROOF. a) Take a Cartan subalgebra $\mathfrak{h}_0$ of $\mathfrak{g}_0$ which contains $\mathfrak{a}$. By Lemma 5.24, we may replace $\mathfrak{h}_0$ with one of its conjugates to assume that it is θ-stable. In other words, $\mathfrak{h}_0 = (\mathfrak{h}_0 \cap \mathfrak{k}) \oplus (\mathfrak{h}_0 \cap \mathfrak{q})$. Since every element $\mathrm{ad}(X)$, $X \in \mathfrak{g}$ is anti-symmetric with respect to the Killing form we find that $\mathrm{ad}(X)$, has only imaginary eigenvalues when $X \in \mathfrak{k}$ while it has real eigenvalues when $X \in \mathfrak{k}$.

Take $X \in \mathfrak{a} \subset \mathfrak{h}_0$, and write it as $X = A + B$, with $A \in \mathfrak{h}_0 \cap \mathfrak{k}$, $B \in \mathfrak{h}_0 \cap \mathfrak{q}$. Since $A, B, X \in \mathfrak{h}_0$, they are simultaneously diagonalizable. But as we said, B and X have real eigenvalues, while A has imaginary eigenvalues. So $A = 0$ and $\mathfrak{a} \subset \mathfrak{q}$.

b) Fix two Cartan subspaces $\mathfrak{a}_1$ and $\mathfrak{a}_2$ of $\mathfrak{g}_0$, contained in $\mathfrak{q}$. Take $X_1 \in \mathfrak{a}_1$ and $X_2 \in \mathfrak{a}_2$ to be *regular elements*. This means that X_i does not belong to the kernel of a restricted root of $\mathfrak{a}_i$ (note that such elements exist, because $\mathfrak{a}_i$ is not equal to the union of finitely many hyperplanes). This way, we know that any element $X \in \mathfrak{g}$ which commutes to X_i actually commutes with $\mathfrak{a}_i$.

Consider the real valued fonction $g \in K_e \mapsto B(X_1, gX_2)$. Since K_e is compact, this function attains its maximum at some element $g \in K_e$. Replacing $\mathfrak{a}_2$ by $g(\mathfrak{a}_2)$ and X_2 by gX_2 if necessary, we may assume that $g = e$. So the differential of this function at $g = e$ vanishes, which shows that for all $X \in \mathfrak{k}$, we have $B(X_1, [X, X_2]) = 0$. This rewrites as $-B(X_1, \mathrm{ad}(X_2)X) = 0$ and hence $B([X_1, X_2], X) = 0$ for all $X \in \mathfrak{k}$. Since $X_1, X_2 \in \mathfrak{q}$, we have that $[X_1, X_2] \in \mathfrak{k}$ and hence $[X_1, X_2] = 0$. Since X_2 is regular w.r.t. $\mathfrak{a}_2$, this shows that $[X_1, \mathfrak{a}_2] = 0$ and further, $[\mathfrak{a}_1, \mathfrak{a}_2] = 0$. By maximality, we must have $\mathfrak{a}_1 = \mathfrak{a}_2$. □

We can now prove our main result.

THEOREM 5.27. *The following facts are true.*

a) In a compact real Lie algebra, any two Cartan subalgebras are conjugate.
b) In a semi-simple complex Lie algebra, any two Cartan subalgebras are conjugate.
c) In a semi-simple real Lie algebra, any two Cartan subspaces are conjugate, and there are only finitely many conjugacy classes of Cartan subalgebras.

PROOF. *a*) We may reproduce the proof of Proposition 5.26.b in this setting, but we can also use directly that result as follows. If $\mathfrak{u}$ is a compact real Lie algebra, then the map $\theta_{\mathfrak{u}}$ is a Cartan involution of its complexification $\mathfrak{g}$. Indeed we already observed that the hermitian form $(X, Y) \mapsto -B(\theta_{\mathfrak{u}}(X), Y)$ is positive definite. However Cartan involutions are about real Lie algebras while B is the Killing form of the complex Lie algebra $\mathfrak{g}$, so we leave as an exercise to check that the Killing form $B_{\mathbb{R}}$ of $\mathfrak{g}$ viewed as a real Lie algebra is $B_{\mathbb{R}} = 2\Re(B)$.

So $\theta_{\mathfrak{u}}$ is indeed a Cartan involution of $\mathfrak{g}$ and the associated Cartan decomposition is the decomposition $\mathfrak{g} = \mathfrak{u} \oplus i\mathfrak{u}$. Now if $\mathfrak{h}_1$, $\mathfrak{h}_2$ are two Cartan subalgebras of $\mathfrak{u}$ then $i\mathfrak{h}_1$ and $i\mathfrak{h}_2$ are Cartan subspaces of $\mathfrak{g}$ contained in $i\mathfrak{u}$ so we may apply Proposition 5.26.b to conjugate them by an automorphism of $\mathfrak{u}$.

b) If $\mathfrak{h}_1$ and $\mathfrak{h}_2$ are two Cartan subalgebras of a semi-simple complex Lie algebra $\mathfrak{g}$, then we may apply Theorem 5.11 to find two compact real forms $\mathfrak{u}_1$ and $\mathfrak{u}_2$ of $\mathfrak{g}$ such that $\mathfrak{u}_i \cap \mathfrak{h}_i$ is a Cartan subalgebra of $\mathfrak{u}_i$ for $i = 1, 2$. By Theorem 5.21, we may find assume that $\mathfrak{u}_1 = \mathfrak{u}_2$. We may then apply item *a*) to conclude and further assume that $\mathfrak{u}_1 \cap \mathfrak{h}_1 = \mathfrak{u}_2 \cap \mathfrak{h}_2$. Which finishes the proof.

c) The first assertion follows from Proposition 5.26. Let us prove the second assertion. We fix a semi-simple real Lie algebra $\mathfrak{g}_0$, a Cartan involution θ and denote by $\mathfrak{g}_0 = \mathfrak{k} \oplus \mathfrak{q}$ the corresponding Cartan decomposition. Fix moreover a Cartan subspace $\mathfrak{a} \subset \mathfrak{q}$.

Any Cartan subalgebra $\mathfrak{h}_0$ of $\mathfrak{g}_0$ may be conjugate to a θ-stable Cartan subalgebra $\mathfrak{h}_1$. We may embed $\mathfrak{a}_1 := \mathfrak{h}_1 \cap \mathfrak{q}$ in a Cartan subspace of $\mathfrak{g}_0$ contained in $\mathfrak{q}$, and applying Proposition 5.26.b, we may assume that $\mathfrak{a}_1 \subset \mathfrak{a}$.

Claim 1. There are only finitely many possibilities for $\mathfrak{a}_1$.

Denote by $\Sigma_1 := \{\lambda \in \Sigma \mid \lambda(\mathfrak{a}_1) = 0\}$, the set of restricted roots vanishing on $\mathfrak{a}_1$. Since Σ is a finite set, there are only finitely many possibilities for the set Σ_1. So we only need to check that $\mathfrak{a}_1$ is completely determined by Σ_1, and more precisely we will show that $\mathfrak{a}_1$ is equal to $\mathfrak{a}_1' := \{X \in \mathfrak{a} \mid \lambda(X) = 0, \text{ for all } X \in \Sigma_1\}$. It is clear that $\mathfrak{a}_1 \subset \mathfrak{a}_1'$, by definition of Σ_1. Take now $X \in \mathfrak{a}_1'$ and $Y \in \mathfrak{h}_1$. Since $\mathfrak{h}_1$ is abelian, Y commutes with $\mathfrak{a}_1$, and we find that

$$Y \in \mathfrak{l} \oplus \left(\bigoplus_{\lambda \in \Sigma_1} \mathfrak{g}_\lambda \right).$$

It follows that $[X, Y] = 0$. Hence X commutes with $\mathfrak{h}_1$ and it is a semi-simple element. By maximality of $\mathfrak{h}_1$, we get that $X \in \mathfrak{h}_1$. Since moreover $X \in \mathfrak{a}_1' \subset \mathfrak{a} \subset \mathfrak{q}$, we conclude that $X \in \mathfrak{h}_1 \cap \mathfrak{q} = \mathfrak{a}_1$, showing that $\mathfrak{a}_1' = \mathfrak{a}_1$.

So we are left to prove the following claim.

Claim 2. If $\mathfrak{h}_2$ is another θ-stable Cartan subalgebra such that $\mathfrak{h}_2 \cap \mathfrak{q} = \mathfrak{h}_1 \cap \mathfrak{q}$ then $\mathfrak{h}_2$ is conjugate to $\mathfrak{h}_1$.

Denote by $\mathfrak{m}$ the centralizer of $\mathfrak{a}_1 = \mathfrak{h}_1 \cap \mathfrak{q}$ inside $\mathfrak{k}$ and by $\mathfrak{z}$ its center. We first show that $\mathfrak{m}/\mathfrak{z}$ is a compact real Lie algebra. Indeed the Killing form $B_\mathfrak{k}$ of $\mathfrak{k}$ restricts to a negative definite symmetric bilinear form on $\mathfrak{m}$. As usual, the elements $\mathrm{ad}_\mathfrak{m}(X)$, $X \in \mathfrak{m}$, are anti-symmetric with respect to this form, showing that $\mathrm{Tr}(\mathrm{ad}_\mathfrak{m}(X)^2) < 0$ for all $X \in \mathfrak{m} \setminus \mathrm{Ker}(\mathrm{ad}_\mathfrak{m})$. Since $\mathfrak{z}$ is the Kernel of $\mathrm{ad}_\mathfrak{m}$, we easily deduce that $\mathfrak{m}/\mathfrak{z}$ is indeed a compact real Lie algebra.

Note that $\mathfrak{h}_1 \cap \mathfrak{k}/\mathfrak{z}$ and $\mathfrak{h}_2 \cap \mathfrak{k}/\mathfrak{z}$ are Cartan subalgebras inside $\mathfrak{m}/\mathfrak{z}$. Using item a), we know that they are conjugate by an automorphism ϕ of $\mathfrak{m}/\mathfrak{z}$. We observe by the proof of a) that we may actually choose ϕ in the connected component of the identity of $\mathrm{Aut}(\mathfrak{m}/\mathfrak{z})$. In order to conclude we need to be able to lift ϕ to an appropriate automorphism of $\mathfrak{g}_0$.

Denote by M the identity component of the automorphism group of $\mathfrak{g}_0$ that commute with θ, and that fix all elements of $\mathfrak{a}_1 = \mathfrak{h}_1 \cap \mathfrak{q}$. The compact Lie group M admits $\mathfrak{m}$ as its Lie algebra (in fact its Lie algebra is $\mathrm{ad}_{\mathfrak{g}_0}(\mathfrak{m})$, which we identify with $\mathfrak{m}$). Further $\mathfrak{m}/\mathfrak{z} = \mathrm{ad}_\mathfrak{m}(\mathfrak{m})$ is the Lie algebra of $\mathrm{Ad}_\mathfrak{m}(M)$ and one concludes that the identity component of $\mathrm{Aut}(\mathfrak{m}/\mathfrak{z})$ may be identified $\mathrm{Ad}_\mathfrak{m}(M)$. Thus we may find $\psi \in M$ such that $\phi = \mathrm{Ad}_\mathfrak{m}(\psi)$. Then one checks that ψ maps $\mathfrak{h}_1 \cap \mathfrak{k}$ onto $\mathfrak{h}_2 \cap \mathfrak{k}$. Since ψ is an automorphism of $\mathfrak{g}_0$ that fix $\mathfrak{a}_1$, we conclude that it conjugates $\mathfrak{h}_1$ onto $\mathfrak{h}_2$. □

5. Some words on classification

Now that we know that there exists a unique Cartan subalgebra up to conjugacy in a semi-simple complex Lie algebra, we get for free many invariants for such Lie algebras.

Definition 5.28. The *rank* of a complex semi-simple Lie algebra is the dimension of its Cartan subalgebras. The *real rank* of a real semi-simple Lie algebra is the dimension of its Cartan subspaces.

In fact simple Lie algebras over $\mathbb{R}$ or $\mathbb{C}$ are completely classified. This classification is not the point of this course, so we will only say a few words on how this works, and refer to the book of Helgason, [**Hel78**, Chapter X] for more details (the point of the whole book is essentially to classify the real semi-simple Lie algebras, and their corresponding symmetric spaces).

Let us focus only on the complex case. If we are given a semi-simple complex Lie algebra $\mathfrak{g}$, then we may pick a Cartan subalgebra $\mathfrak{h}$ in it. Then we consider the corresponding

set of roots $\Delta \subset \mathfrak{h}^*$. We have seen in the first two sections of this chapter that we could use the Killing form B on $\mathfrak{g}$ to represent Δ inside $\mathfrak{h}$, as a set $\{H_\alpha \mid \alpha \in \Delta\}$. We even saw that the restriction of the Killing form B on the real vector space $\mathfrak{h}_\mathbb{R}$ generated by Δ was positive definite. So, $(\mathfrak{h}_\mathbb{R}, B_{|\mathfrak{h}_\mathbb{R}})$ is a Euclidean space containing the finite set Δ, and it can be checked that it is a reduced abstract root system in the following sense.

DEFINITION 5.29. Let $(E, \langle\cdot,\cdot\rangle)$ be a Euclidean vector space, and for every $x \in E$, denote by $s_x : E \to E$ the orthogonal symmetry with respect to $(\mathbb{R}x)^\perp$, i.e.

$$s_x(y) = y - 2\frac{\langle y, x\rangle}{\langle x, x\rangle}x, \text{ for all } y \in E.$$

We say that a finite subset $\Delta \subset E$ is an *abstract root system* if it spans E and $2\frac{\langle\beta,\alpha\rangle}{\langle\alpha,\alpha\rangle} \in \mathbb{Z}$ and $s_\alpha(\beta) \in \Delta$ for all $\alpha, \beta \in \Delta$. We say that Δ is *reduced* if $\alpha \in \Delta$ implies $2\alpha \notin \Delta$.

This abstract root system is, up to isometry, canonically associated with $\mathfrak{g}$ since there is only one Cartan subalgebra in $\mathfrak{g}$ up to conjugacy. Conversely, one can prove that

- the abstract root system associated to a semi-simple complex Lie algebra characterizes it, up to isomorphism, and
- any reduced abstract root system actually comes from a semi-simple complex Lie algebra.

If the Lie algebra is simple, then the root system is *irreducible*, which means that it is not the union of two pairwise orthogonal non-empty subsets. So the remaining task is to classify all the irreducible, reduced, abstract root systems. This part is essentially of combinatorial nature (and goes through the so-called Dynkin diagrams).

This study allows to give a complete list of all simple complex Lie algebras. Recall moreover that any semi-simple Lie algebra is the direct sum of simple ones.

THEOREM 5.30. *Over $\mathbb{C}$, there are four infinite families of* classical *simple Lie algebras, and five* exceptional *Lie algebras:*

$$A_r := \mathfrak{sl}(r+1, \mathbb{C}) \quad r \geq 1,$$
$$B_r := \mathfrak{so}(2r+1, \mathbb{C}) \quad r \geq 2,$$
$$C_r := \mathfrak{sp}(r, \mathbb{C}) \quad r \geq 3,$$
$$D_r := \mathfrak{so}(2r, \mathbb{C}) \quad r \geq 4,$$
$$E_6, E_7, E_8, F_4, G_2.$$

Then we know that any real simple Lie algebra is a real form of one of the above Lie algebras. But the classification is again non-trivial.

EXERCISE 5.31. Compute the real rank of each of the above Lie algebras. Exhibit one of their Cartan subalgebras.

CHAPTER 6

Semi-simple Lie groups

In this chapter, we make use of our study of Lie algebras to derive further structure of Lie groups.

1. Recap

DEFINITION 6.1. We say that a Lie group is semi-simple if its Lie algebra is semi-simple. It is called simple if its Lie algebra is simple and of dimension at least two.

We saw that there are two settings in which Lie groups are characterized by their Lie algebras: the simply connected setting and the case of groups with trivial center.

LEMMA 6.2. *Given any semi-simple Lie algebra $\mathfrak{g}$ there exists a connected Lie group G with Lie algebra $\mathfrak{g}$, namely, the identity component of* $\mathrm{Aut}(\mathfrak{g})$*. Its universal cover is then the unique simply connected Lie group G whose Lie algebra is $\mathfrak{g}$. Moreover, since $\mathfrak{g}$ has trivial center,* $\mathrm{Ad}(G)$ *is a Lie group with trivial center whose Lie algebra is also $\mathfrak{g}$.*

This remark is easily seen to imply the following fact.

PROPOSITION 6.3. *Any semi-simple, simply connected (connected) Lie group is the product of finitely many simple, simply connected Lie groups. Any semi-simple, connected Lie group with trivial center is the product of finitely many simple ones.*

In fact, for simply connected groups it is even better: *there is an equivalence of categories between the Category of simply connected semi-simple Lie groups and the category of semi-simple Lie algebras over the same field.* Note that for groups with trivial center, this is not always the case that a morphism between Lie algebras gives rise to a morphism between the Lie groups; it may only be a local morphism.

2. Cartan decomposition

We fix a connected semi-simple *real* Lie group G, with finite center, and we denote by $\mathfrak{g}$ its Lie algebra.

There always exists a Cartan involution θ of G(i.e. an involutive smooth automorphism whose differential is a Cartan involution of $\mathfrak{g}$). This is obvious if G is simply connected, and easily deduced if G has trivial center, but it is in fact true in full generality. We denote by $\theta_0 = d\theta_e$ the Lie algebra Cartan involution.

Denote by $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{q}$ the corresponding Cartan decomposition and choose a Cartan subspace $\mathfrak{a} \subset \mathfrak{q}$. Denote by $K < G$ the subgroup of fixed points of θ and set $A := \exp(\mathfrak{a})$.

THEOREM 6.4 (Cartan decomposition). *The map $\phi : (k, X) \in K \times \mathfrak{q} \mapsto k\exp(X) \in G$ is an onto diffeomorphism. Moreover any element $g \in G$ can be written as a product $g = kak'$, with $k, k' \in K$ and $a \in A$. In other words, we have*

$$G = KAK.$$

PROOF. First things first, before showing that ϕ is a diffeomorphism, let us check that exp is an injective immersion.

Claim 1. $\exp : \mathfrak{q} \to G$ is an immersion.

To prove this claim we admit that $\exp : \mathfrak{g} \to G$ is differentiable at every point of $\mathfrak{g}$, with derivative

$$d(\exp)_X(Y) = d(L_{\exp(X)})_e \circ \frac{\mathrm{id} - e^{-\operatorname{ad}(X)}}{\operatorname{ad}(X)}(Y), \text{ for all } X, Y \in \mathfrak{g}.$$

Here the exponential $e^{-\operatorname{ad}(X)}$ is the usual exponential for endomorphisms of $\mathfrak{g}$. This formula is proved in [**Hel78**], Chapter II.1.4. Alternatively, one may observe that to prove this formula in the case of semi-simple Lie groups, one only needs to consider the case of a Lie group with trivial center, and hence, of a linear Lie group. So it suffices to verify the statement for the exponential of matrices, which is an easy computation. Let us check that $d(\exp)_X$ is injective for all $X \in \mathfrak{q}$. We warn that this fact is not true for an arbitrary $X \in \mathfrak{g}$.

Assume that $Y \in \mathfrak{g}$ is such that $d(\exp)_X(Y) = 0$. This implies in particular that $Y = e^{-\operatorname{ad}(X)}(Y)$. But since $X \in \mathfrak{q}$, we know that X is self adjoint for the positive definite form B_{θ_0}. In particular $\operatorname{ad}(X)$ is diagonalisable on $\mathfrak{g}$ and the relation $Y = e^{-\operatorname{ad}(X)}(Y)$ implies that $\operatorname{ad}(X)Y = 0$. But one checks that when $[X, Y] = 0$, the above formula becomes

$$d(\exp)_X(Y) = d(L_{\exp(X)})_e \circ (-\sum_{k \geq 1} \frac{1}{k!}(-\operatorname{ad} X)^{k-1})(Y) = d(L_{\exp(X)})_e(Y).$$

This thus forces $Y = 0$, proving that exp is an immersion on $\mathfrak{q}$.

Claim 2. $\exp : \mathfrak{q} \to G$ is injective.

This is again a consequence of the fact that elements $\operatorname{ad}(X)$, $X \in \mathfrak{q}$ are diagonalisable on $\mathfrak{g}$. Indeed if $X, Y \in \mathfrak{q}$ are such that $\exp(X) = \exp(Y)$, then in particular we have $\operatorname{Ad}(\exp(X)) = \operatorname{Ad}(\exp(Y))$, and so $\exp(\operatorname{ad}(X)) = \exp(\operatorname{ad}(Y))$. This implies that $\operatorname{ad}(X)$ and $\operatorname{ad}(Y)$ have the same eigenvalues, with the same eigenspace decomposition. So, $\operatorname{ad}(X) = \operatorname{ad}(Y)$, which implies $X = Y$, because $\mathfrak{g}$ is semi-simple.

Claim 3. $\exp(\mathfrak{q}) = \{\theta(g)^{-1}g \mid g \in G\}$.

Denote by $S := \{\theta(g)^{-1}g \mid g \in G\}$. Observe that for all $X \in fq$, we have $\exp(X)\theta(g)^{-1}g$, with $g = \exp(X/2)$. So $\exp(\mathfrak{q}) \subset S$. Conversely take $g \in S$. We have

$$\begin{aligned} B_{\theta_0}(\operatorname{Ad}(\theta(g)^{-1}g)X, X) &= B(\theta_0(\operatorname{Ad}(\theta(g)^{-1}g)X), X) \\ &= B(\operatorname{Ad}(g^{-1}\theta(g))\theta_0(X), X) \\ &= B(\operatorname{Ad}(\theta(g))\theta_0(X), \operatorname{Ad}(g)X) \\ &= B_{\theta_0}(\operatorname{Ad}(g)X, \operatorname{Ad}(g)X). \end{aligned}$$

This shows that $\operatorname{Ad}(\theta(g)^{-1}g)$ is a positive endomorphism with respect to the positive definite form B_{θ_0}. So it may be written as $\exp(X)$ for some $X \in \mathcal{L}(\mathfrak{g})$. By Remark 3.19, we know that for all $t \in \mathbb{R}$, $\exp(tX)$ belongs to the real algebraic group $\operatorname{Aut}(\mathfrak{g})$. Since the map $t \mapsto \exp(tX) \in \mathcal{L}(\mathfrak{g})$ is continuous, we know that it actually ranges into the identity component of $\operatorname{Aut}(\mathfrak{g})$. By Proposition 4.18, this identity component is equal to $\operatorname{Ad}(G)$. Moreover, X is the derivative of $t \mapsto \exp(tX)$ at $t = 0$. It thus belongs to $\mathfrak{g}$ and is self-adjoint with respect to B_{θ_0}. This forces $X \in \mathfrak{q}$, and thus proves the claim.

We can now deduce that ϕ is an onto diffeomorphism. It follows from Claim 1 that ϕ is an immersion. So we only need to check that it is bijective. Note that if $g \in K \cap \exp(\mathfrak{q})$, then $\theta(g) = g = g^{-1}$, so $g^2 = 1$. Writing $g = \exp(X)$, this shows that $\exp(2X) = 1$, and

so $X = 0$ by Claim 2. This proves that ϕ is injective. Let us now fix $g \in G$ and try to write it as $g = k\exp(X)$ with $k \in K$, $X \in \mathfrak{q}$. By Claim 3, he element $x := \theta(g)^{-1}g$ is of the form $x = \exp(X)$ for some $X \in \mathfrak{q}$. Set now $k := g\exp(-X/2)$. We see that $\theta(k)^{-1}k = \exp(-X/2)\theta(g)^{-1}g\exp(-X/2) = e$. Hence $\theta(k) = k$, proving that $k \in K$.

To derive the second statement, it suffices to use Proposition 5.26 to deduce that any element of $\mathfrak{q}$ can be conjugate by an element of K to an element in $\mathfrak{a}$. □

REMARK 6.5. Actually the Cartan decomposition $G = KAK$ admits a finer version, which appeals to an ordering of the set of roots. This finer version imposes extra conditions on a appearing in a decomposition $g = kak'$, which then makes this element A unique (but k, k' need not be unique). We refer [**Hel78**] for details.

PROPOSITION 6.6. *K is a maximal compact subgroup of G. Its Lie algebra is $\mathfrak{k}$. In fact, any maximal compact subgroup of G is conjugate K.*

PROOF. The Lie algebra of K identifies with $\{X \in \mathfrak{g} \mid \exp(tX) \in K, \text{ for all } t \in \mathbb{R}\}$. This is easily since to coincide with the fixed point algebra of θ_0, which is $\mathfrak{k}$. Since $\mathfrak{k}$ is compactly embedded inside $\mathfrak{g}$, we then know that the connected component of the identity of K is compact. But since G is connected, the first statement of Theorem 6.4 also shows that K is connected. So K is compact.

The fact that any maximal compact subgroup is conjugate to K (which implies in particular that K is itself maximal compact) can be derived from a fixed point theorem. One shows that any compact group $L < G$ fixes a point inside G/K. This is a consequence of the fact that G/K can be endowed with a Riemannian metric of non-positive curvature for which the action $G \curvearrowright G/K$ is by isometries. We will not prove this. □

There are other decompositions for Lie groups; the Iwasawa decomposition and the Bruhat decomposition. We don't present them here.

Part 3

Lattices: Constructions, structure and applications

CHAPTER 7

Arithmetic lattices

The next two chapters are based on Benoist notes [**Ben08**], Chapter 1 and 5.

The point is to define arithmetic subgroups of Lie groups and prove that they are lattices in semi-simple situations. Let us first provide more definitions on algebraic groups.

1. More on algebraic groups

We fix a field k of characteristic 0 and K an algebraically closed field containing k. A k-group $\mathbb{G}$ is said to be *connected* if it is Zariski-connected.

We say that $\mathbb{G}$ is a *k-torus* if it is abelian connected, and all its elements are semi-simple. In other words, $\mathbb{G}$ is a k-torus if and only if it is defined over k, and is isomorphic over K with $(\mathbb{G}_m)^r$ for some $r \geq 1$. The torus is said to be *k-split* if it is isomorphic over k with $(\mathbb{G}_m)^r$. This amounts to say that the elements of $\mathbb{G}_k$ are diagonalisable over k (for all k-representation).

Example 7.1. The $\mathbb{R}$-group $\mathrm{SO}(2,\mathbb{C})$ is a torus but it is not split over $\mathbb{R}$. In contrast, the group $\mathrm{SO}(1,1)$ is an $\mathbb{R}$-split torus. Note that these two algebraic groups are isomorphic over $\mathbb{C}$.

A k-group $\mathbb{G}$ is said to be *semi-simple* if it does not contain an abelian connected normal k-subgroup. It is a fact that $\mathbb{G}$ is semi-simple if and only if its Lie algebra $\mathfrak{g}$ is semi-simple, which is also equivalent to saying that $\mathfrak{g}_k$ is semi-simple. This notion does not depend on the field of definition k.

In the same spirit a k-group $\mathbb{G}$ is said to be k-simple if it is connected and its only proper normal k-subgroup is the trivial subgroup. The following example shows that although semi-simplicity does not depend on the field of definition k, simplicity does.

As for Lie groups, we may define the *adjoint representation* of $\mathbb{G}$ on its Lie algebra $\mathfrak{g}$, by requiring that $\mathrm{Ad}(g) : \mathfrak{g} \to \mathfrak{g}$ is the differential of the inner automorphism $I(g) : h \in \mathbb{G} \mapsto ghg^{-1} \in \mathbb{G}$. This representation is k-rational if $\mathbb{G}$ is a k-group. This can be easily seen for $\mathbb{G} = \mathrm{GL}(\mathbb{V})$, and then it follows for arbitrary $\mathbb{G}$ by restriction. The differential of Ad is the adjoint representation ad of the Lie algebra $\mathfrak{g}$. In characteristic 0 if $\mathbb{G}$ is connected the Kernel of Ad is the center of $\mathbb{G}$ and its Lie algebra is the kernel of ad.

A semi-simple k-group $\mathbb{G}$ is *adjoint* if it has trivial center. If it is connected, then this amounts to saying that its adjoint representation is faithful.

Proposition 7.2. *If $\mathbb{G}$ is a connected adjoint semi-simple k-group then it is the direct product of finitely many simple k-groups. This decomposition is unique.*

Proof. Since $\mathbb{G}$ is adjoint, it is isomorphic with its image in $\mathrm{Aut}(\mathfrak{g})$, and since it is semi-simple, we know that its Lie algebra $\mathfrak{g}$ is semi-simple, hence $\mathrm{ad}(\mathfrak{g}) = \mathrm{Der}(\mathfrak{g})$. By connectedness, we must have that $\mathbb{G} \simeq \mathrm{Ad}(\mathbb{G}) = \mathrm{Aut}(\mathfrak{g})^0$. Since $\mathbb{G}$ is defined over k, the Lie algebra $\mathfrak{g}$ is of the form $\mathfrak{g} = \mathfrak{g}_k \otimes_k K$ and $\mathfrak{g}_k$ is semi-simple. So we may write the

later as a direct sum of finitely many simple k-subalgebras $\mathfrak{g}_k = \bigoplus_i \mathfrak{g}_i$. Then it is easily seen that $\mathrm{Aut}(\mathfrak{g})$ is equal to the direct product of the subgroups $\mathrm{Aut}(\mathfrak{g}_i \otimes_k K)$, up to finite index (a finite group permuting some of the K-simple factors of the $\mathfrak{g}_i \otimes_k K$'s may appear). When taking the connected component of the identity, the finite index noise disappears and we have a genuine direct product decomposition

$$\mathbb{G} \simeq \mathrm{Aut}(\mathfrak{g})^0 = \Pi_i \, \mathrm{Aut}(\mathfrak{g}_i \otimes_k K)^0.$$

It remains to check that each k-subgroup $\mathbb{G}_i := \mathrm{Aut}(\mathfrak{g}_i \otimes_k K)^0$ is simple if $\mathfrak{g}_i$ is k-simple. Since $\mathfrak{g}_i$ is simple we see that any normal k-subgroup $\mathbb{H}$ of $\mathbb{G}_i$ must have trivial Lie algebra. So $\mathbb{H}$ must be a finite normal subgroup of $\mathbb{G}_i$. Since a finite normal subgroup in a connected group is contained in the center of that group, we see that $\mathbb{H}$ must be trivial, because $\mathbb{G}$ has trivial center. This finishes the proof of the existence statement. We leave the uniqueness as an exercise. □

A k-group $\mathbb{G}$ is said to be k-split if it contains a maximal k-torus which is k-split. More generally $\mathbb{G}$ is called *k-isotropic* if it contains some non-trivial k-split torus. It is called *anistropic* if it not isotropic. The *rank* of $\mathbb{G}$ is the dimension of a maximal k-torus in $\mathbb{G}$, while the *k-rank* of $\mathbb{G}$ is the dimension of a maximal *k-split* torus.

Example 7.3.

- The real group $\mathbb{G} = \mathrm{SL}_d(\mathbb{C})$ is $\mathbb{R}$-split, and of rank $d-1$. Indeed the subgroup of diagonal matrices of determinant 1 is an $\mathbb{R}$-split torus which is maximal among all $\mathbb{R}$-tori of $\mathbb{G}$.
- The real group $\mathrm{SO}(d) = \{A \in \mathrm{SL}_d(\mathbb{C}) \mid A^t A = \mathrm{id}\}$ has real rank 0, while it contains an $\mathbb{R}$-torus, given described blocwise by

$$\begin{pmatrix} \mathrm{SO}(2,\mathbb{C}) & 0 & \cdots & 0 \\ 0 & \mathrm{SO}(2,\mathbb{C}) & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \mathrm{SO}(2,\mathbb{C}) \end{pmatrix}.$$

  The above description makes sense when d is even. If d is odd, then add a column and a line of 0's except for the coefficient (d,d) which is set to 1. In fact this torus is maximal in $\mathrm{SO}(d)$, showing that the rank of $\mathrm{SO}(d)$ is the integer part of $d/2$.
- The $\mathbb{R}$-group $\mathrm{SO}(p,q)$ is of real rank is $\min(p,q)$. Because it admits an $\mathbb{R}$-split torus, described in a suitable basis by a bloc diagonal matrix, whose diagonal blocs are copies of $\mathrm{SO}(1,1)$.

Proposition 7.4. *If $\mathbb{G}$ is semi-simple, then $\mathbb{G}$ is k-isotropic if and only if $\mathbb{G}_k$ contains non-trivial unipotent elements.*

Proof. Assume first that $\mathbb{G}_k$ contains a non-trivial unipotent element u. Then as seen in the first graded homework, we may define its logarithm $X = \log(u)$, which is a nilpotent element in $\mathfrak{g}_k$. By Jacobson-Morozov Theorem, X is part of an $\mathfrak{sl}_2$-triple (H, X, Y) in $\mathfrak{g}_k$. Moreover, from the description of the representations of $\mathfrak{sl}_2(k)$, we know that H is diagonalisable with integral eigenvalues n_i, $H = \mathrm{diag}(n_i)$ in a suitable k-basis of $\mathbb{V}$. Then we see that the homomorphism $x \in \mathbb{G}_m \mapsto \mathrm{diag}(x^{n_i}) \in \mathrm{GL}(\mathbb{V})$ is a k-rational group homomorphism. The image of this homomorphism is a connected k-subgroup $\mathbb{H}$ of $\mathrm{GL}(\mathbb{V})$ whose Lie algebra is equal to the span of H. From the homework, this is seen to imply that $\mathbb{H}$ is contained in $\mathbb{G}$. Since $\mathbb{H}$ is a non-trivial k-split torus, we conclude that $\mathbb{G}$ is k-isotropic.

Conversely, assume that $\mathbb{G}$ contains a k-split torus $\mathbb{H}$. Then we may diagonalise $\mathbb{H}$ in the adjoint representation of $\mathbb{G}$ on $\mathfrak{g}$. Hence we may decompose $\mathfrak{g} = \mathfrak{g}_0 \oplus_{\lambda\in\Delta} \mathfrak{g}_\lambda$, where Δ is a set of non-trivial k-characters on $\mathbb{H}$ (the set of *roots* of the torus $\mathbb{H}$), and

$$\mathfrak{g}_\lambda = \{X \in \mathfrak{g} \mid \mathrm{Ad}(h)(X) = \lambda(h)X, \text{ for all } h \in \mathbb{H}\}.$$

As in the Lie algebra case, we can check that $[g_\lambda, g_\mu] \subset g_{\lambda\mu}$, where $\lambda\mu : \mathbb{H} \to \mathbb{G}_m$ is the product character. Since only finitely many $\mathfrak{g}_\lambda$ are non-zero, we see that any element $\mathfrak{g}_\lambda$ is nilpotent. Take a root $\lambda \in \Delta$ such that $\mathfrak{g}_\lambda \cap \mathfrak{g}_k \neq \{0\}$. Then any non-zero element $X \in \mathfrak{g}_\lambda \cap \mathfrak{g}_k$ gives the desired non-trivial unipotent element $u = \exp(X)$. □

Finally, we mention another Proposition that we will need later.

Proposition 7.5. *Consider a subfield $k_0 \subset k$ and an adjoint k_0-simple k_0-group $\mathbb{G}$. Assume that $\mathbb{G}$ is k_0-isotropic. Viewing $\mathbb{G}$ as a k-group, Proposition 7.2 implies that we may write $\mathbb{G}$ as a direct product of k-simple groups $\mathbb{G}_i$.*

Then each $\mathbb{G}_i$ is isotropic.

Proof. We have a splitting $\mathfrak{g} = \mathfrak{g}_{k_0} \otimes_{k_0} K$. We may decompose the k-Lie algebra $\mathfrak{g}_k := \mathfrak{g}_{k_0} \otimes_{k_0} k$ as a direct sum of finitely many simple k-subalgebras $\mathfrak{g}_k = \bigoplus_{i\in I} \mathfrak{g}_i$. By the previous proposition, we just need to show that the Lie algebra $\mathfrak{g}_i$ contains a non-trivial unipotent element for all i.

By assumption $\mathfrak{g}_{k_0}$ contains a non-zero unipotent element X. Viewing X as an element of $\mathfrak{g}_k$, we may decompose it as a sum $X = \sum_i X_i$. Since X is nilpotent, each of the X_i's are nilpotent, so it suffices to check that each X_i is non-zero. For this we would like to use the Galois group of the inclusion $k_0 \subset k$, but we don't know that this is a Galois extension. However, since K is algebraically closed, the extension $k_0 \subset K$ is satisfies the Galois property.

Write $\mathfrak{g} = \mathfrak{g}_k \otimes_k K = \bigoplus_i \mathfrak{g}_i \otimes_k K$. Since each $\mathfrak{g}_i$ is simple over k we deduce that $\mathfrak{g}_i \otimes_k K$ is semi-simple over K. So it may be decomposed as a direct sum of finitely many K-simple ideals:

$$\mathfrak{g}_i \otimes_k K = \bigoplus_{j\in J_i} \mathfrak{g}_{i,j}, \text{ for all } i \in I.$$

In this decomposition, we may write $X_i = \sum_j X_{i,j}$, so that $\mathfrak{g} = \bigoplus_{i,j} \mathfrak{g}_{i,j}$ and $X = \sum_{i,j} X_{i,j}$. Here $\mathfrak{g}_{i,j}$ are the unique K-simple factors of $\mathfrak{g}$. So any Galois automorphism of K over k_0 permutes these simple factors, and hence induces a permutation of the index set $\widetilde{I} := \bigsqcup_{i\in I} J_i$. This action is transitive because if $\mathcal{O} \subset \widetilde{I}$ is an orbit under the Galois group, then $\sum_{(i,j)\in\mathcal{O}} \mathfrak{g}_{i,j}$ is globally invariant under the Galois group, and hence it is a subalgebra defined over k_0 (and an ideal). As $\mathfrak{g}_{k_0}$ is simple, this ideal must be either trivial or everything, so we conclude that $\mathcal{O}$ is either empty or equal to $\widetilde{I}$, proving that the action is indeed transitive.

Now, since X is non-zero, we know that some coefficient $X_{i,j}$ is non-zero. Moreover $X \in \mathfrak{g}_{k_0}$ is invariant under the Galois group of K over k_0. Hence its coordinates $X_{i,j}$ are permuted under the Galois action. Since the action is transitive on $\widetilde{I}$, we conclude that all the coordinates of X are non-zero. Gathering these coordinates appropriately we conclude that $X_i \neq 0$ for all $i \in I$. □

2. Arithmetic groups

We start by giving a strict definition of an arithmetic group, as the set of integer points of a $\mathbb{Q}$-group. So here $k = \mathbb{Q}$ and K continues to denote an algebraic closed field containing k, e.g. $K = \mathbb{C}$.

DEFINITION 7.6. Consider a linear algebraic $\mathbb{Q}$-group $\mathbb{G} \subset \mathrm{GL}(d, K)$. Then the group $\mathbb{G}_{\mathbb{Z}} := \mathbb{G} \cap \mathrm{GL}(d, \mathbb{Z})$ is called an *arithmetic subgroup* of $\mathbb{G}$.

This notion is a priori not satisfactory because it depends on the choice of a $\mathbb{Q}$-embedding $\mathbb{G} \subset \mathrm{GL}(d, K)$, that is, on the choice of a faithful $\mathbb{Q}$-representation $\mathbb{G} \hookrightarrow \mathrm{GL}(\mathbb{V})$, and on the choice of a $\mathbb{Q}$-basis of $\mathbb{V}_{\mathbb{Q}}$, in order to identify $\mathbb{V}$ with K^d.

Let us prove that, in fact, the notion of an arithmetic subgroup is not too sensitive to these choices.

DEFINITION 7.7. Two subgroups Γ_1 and Γ_2 of a group G are *commensurable* if $\Gamma_1 \cap \Gamma_2$ has finite index inside both Γ_1 and Γ_2 (meaning that the coset space $\Gamma_i/\Gamma_1 \cap \Gamma_2$ is finite for $i = 1, 2$).

EXERCISE 7.8. Check that the commensurability relation is an equivalence relation on the set of subgroups of a given group.

PROPOSITION 7.9. *If $\mathbb{G} \subset \mathrm{GL}(d, K)$ is a $\mathbb{Q}$-subgroup and if $\pi : \mathbb{G} \to \mathrm{GL}(d', K)$ is another faithful $\mathbb{Q}$-representation then $\pi(\mathbb{G}_{\mathbb{Z}})$ and $\pi(\mathbb{G})_{\mathbb{Z}}$ are commensurable subgroups inside $\pi(\mathbb{G})$. In particular the commensurability class of the arithmetic subgroup $\mathbb{G}_{\mathbb{Z}}$ depends only on $\mathbb{G}$.*

The proposition is based on the study of lattices in vector spaces. Below, we will consider vector spaces over $\mathbb{Q}$ or $\mathbb{R}$ (later). A *lattice* in a real vector space is a discrete subgroup with finite covolume. Classical structure theorems for lattices tell us that they are generated as additive groups by a basis of the vector space. By analogy, in the case of a $\mathbb{Q}$-vector space $\mathbb{V}_{\mathbb{Q}}$ a lattice is then defined to be the additive group generated by a fixed basis of $\mathbb{V}_{\mathbb{Q}}$. This amounts to saying that the subgroup is a lattice inside the real vector space $\mathbb{R} \otimes_{\mathbb{Q}} \mathbb{V}_{\mathbb{Q}}$.

LEMMA 7.10. *Let $\mathbb{V}_{\mathbb{Q}}$ be a $\mathbb{Q}$-vector space. The sum $\Delta = \Delta_1 + \Delta_2$ of two lattices $\Delta_1, \Delta_2 \subset \mathbb{V}_{\mathbb{Q}}$ is again a lattice.*

PROOF. Since $\Delta_1, \Delta_2 \subset \mathbb{V}_{\mathbb{Q}}$ are two lattices, we may find two bases $f_1, \dots, f_d$, $g_1, \dots, g_d$ of $\mathbb{V}_{\mathbb{Q}}$ such that $\Delta_1 = \mathbb{Z}f_1 \oplus \cdots \oplus \mathbb{Z}f_d$ and $\Delta_2 = \mathbb{Z}g_1 \oplus \cdots \oplus \mathbb{Z}g_d$. Denote by $M \in M(d, \mathbb{Q})$ the matrix expressing the coefficients of $g_1, \dots, g_d$ in the basis $f_1, \dots, f_d$. Denote by $a \in \mathbb{Z} \setminus \{0\}$ the least common multiple of the denominators of the entries of M. We see that $g_1, \dots, g_d$ belong to $\frac{1}{a}\Delta_1$. So both Δ_1 and Δ_2 are contained in $\frac{1}{a}\Delta_1$, which is a lattice in $\mathbb{V}_{\mathbb{Q}}$. So $\Delta_1 + \Delta_2$ is discrete inside $\mathbb{R} \otimes_{\mathbb{Q}} \mathbb{V}_Q$, and hence is a lattice. □

LEMMA 7.11. *Consider a $\mathbb{Q}$-subgroup $\mathbb{G} \subset \mathrm{GL}(d, K)$ and a $\mathbb{Q}$-representation $\pi : \mathbb{G} \to \mathrm{GL}(\mathbb{V})$ for some $\mathbb{Q}$-vector space $\mathbb{V} = K \otimes \mathbb{V}_{\mathbb{Q}}$. Then there exists a lattice in $\mathbb{V}_{\mathbb{Q}}$ which is globally invariant under $\mathbb{G}_{\mathbb{Z}}$. Moreover any lattice of $\mathbb{V}_{\mathbb{Q}}$ is globally invariant under some finite index subgroup of $\mathbb{G}_{\mathbb{Z}}$.*

PROOF. Fix a lattice $\Delta \subset \mathbb{V}_{\mathbb{Q}}$ and fix a basis of $\mathbb{V}_{\mathbb{Q}}$ generating this lattice as an additive group.

Observe that the coefficients of $\pi(g)$ in this basis can be expressed as polynomials over $\mathbb{Q}$ of the coefficients of g as a matrix in $\mathrm{GL}(d, K)$. Indeed, any coefficient function

$\phi_{i,j} : x \in \mathrm{GL}(\mathbb{V}) \mapsto x_{i,j} \in K$ is $\mathbb{Q}$-rational over $\mathrm{GL}(\mathbb{V})$ and we see that the function $g \mapsto \pi(g)_{i,j}$ is nothing but $\phi_{i,j} \circ \pi_* \in \mathbb{Q}[\mathbb{G}]$. So, like any rational function on $\mathbb{G}$, it is the restriction to $\mathbb{G}$ of a rational function on $\mathrm{GL}(d, K)$, i.e. of a polynomial in the coefficients of the matrix $g \in \mathrm{GL}(d, K)$.

In fact we will rather use the fact that the coefficients of $\pi(g) - 1$ can be expressed as polynomials over $\mathbb{Q}$ of the coefficients of $g - 1$ as a matrix in $\mathrm{GL}(d, K)$. The constant terms of these polynomials are 0. So we may find an integer $m \geq 1$ such that if $g = 1$ modulo m then $\pi(g)$ has integer coefficients. So we see that the lattice Δ is globally invariant under the congruence subgroup $\Gamma_m = \{g \in \mathbb{G}_\mathbb{Z} \mid g = 1 \mod m\}$. Moreover it is obvious that Γ_m has finite index inside $\mathbb{G}_\mathbb{Z}$ since the quotient embeds inside the finite group $\mathrm{GL}(\mathbb{Z}/m\mathbb{Z})$. This proves the second statement.

Now take a set of representatives of $\mathbb{G}_\mathbb{Z}/\Gamma_m$, i.e. a finite subset $F \subset \mathbb{G}_\mathbb{Z}$ such that the quotient map $\mathbb{G}_\mathbb{Z} \to \mathbb{G}_\mathbb{Z}/\Gamma_m$ restricts to a bijection from F onto $\mathbb{G}_\mathbb{Z}/\Gamma_m$. By the previous lemma, we see that $\tilde{\Delta} := \sum_{g \in F} \pi(g)\Delta$ is also a lattice inside $\mathbb{V}_\mathbb{Q}$. Moreover, if $h \in \mathbb{G}_\mathbb{Z}$ and $g \in F$, then we may find $g' \in F$ such that $hg \in g'\Gamma_m$. Since Γ_m preserves Δ, we see that $\pi(h)\pi(g)\Delta = \pi(g')\Delta \in \tilde{\Delta}$. This shows that $\tilde{\Delta}$ is $\mathbb{G}_\mathbb{Z}$-invariant, proving the first statement. □

Proof of Proposition 7.9. By assumption, $\pi(\mathbb{G})_\mathbb{Z}$ is the stabilizer of the lattice $\mathbb{Z}^{d'}$ inside $\pi(\mathbb{G})$. Then by the lemma, we know that there is a finite index subgroup Γ of $\mathbb{G}_\mathbb{Z}$ that preserves this lattice, i.e., $\pi(\Gamma) \subset \pi(\mathbb{G})_\mathbb{Z}$. Now, conversely, we use π^{-1} to find a finite index subgroup $\Lambda < \pi(\mathbb{G})_\mathbb{Z}$ such that $\Lambda \subset \pi(\mathbb{G}_\mathbb{Z})$ (here we use the fact that a bijective $\mathbb{Q}$-morphism between two smooth varieties is a $\mathbb{Q}$-isomorphism, so $\pi^{-1}$is indeed again a $\mathbb{Q}$-representation of $\pi(\mathbb{G})$). Then we conclude that $\pi(\Gamma) \cap \Lambda$ has finite index inside both $\pi(\mathbb{G}_\mathbb{Z})$ and $\pi(\mathbb{G})_\mathbb{Z}$. □

One of the main results that we will prove is the following one, asserting that arithmetic subgroups often give rise to lattices. We state it only for semi-simple groups, but it is completely understood for which $\mathbb{Q}$-groups $\mathbb{G}$ the arithmetic subgroup $\mathbb{G}_\mathbb{Z}$ is a lattice inside $\mathbb{G}_\mathbb{R}$. The first assertion is due to Borel and Harish-Chandra, while the second statement is the so-called Godement co-compactness criterion. It is worth recalling that the set of real points of an algebraic $\mathbb{R}$-group (and *a fortiori* of a $\mathbb{Q}$-group) is a Lie group.

Theorem 7.12. *Let $\mathbb{G}$ be a semi-simple $\mathbb{Q}$-group. The following facts hold true.*

1. *$\mathbb{G}_\mathbb{Z}$ is a lattice inside $\mathbb{G}_\mathbb{R}$.*
2. *$\mathbb{G}_\mathbb{Z}$ is co-compact inside $\mathbb{G}_\mathbb{R}$ if and only if $\mathbb{G}_\mathbb{Z}$ doesn't admit a nontrivial unipotent element.*

Concrete examples will be provided in class.

We will only prove the result in the case where $\mathbb{G}$ has trivial center. Note that the fact that $\Gamma < G$ is a lattice or a co-compact lattice only depends on the commensurability class of Γ inside G.

It is obvious that $\mathbb{G}_\mathbb{Z}$ is discrete inside $\mathbb{G}_\mathbb{R}$. The whole point of the proof is to check that it has finite co-volume in $\mathbb{G}_\mathbb{R}$ or is co-compact in the suitable cases. The proof requires preliminary work regarding the *space of lattices* and a general dynamical criterion to prove that a measure is finite. We discuss the space of lattice in this chapter and prove item 2. We postpone the proof of the first item to the next chapter.

3. The space of lattices

The set X^+ of all lattices in $\mathbb{R}^d$ identifies with the homogeneous space $\mathrm{GL}(d,\mathbb{R})/\mathrm{SL}^{\pm}(d,\mathbb{Z})$ as follows.

Observe that $\mathrm{GL}_d(\mathbb{R})$ acts on $\mathbb{R}^d$ and hence on the set of all its lattices. The action is *transitive*: if $\Delta = \mathbb{Z}f_1 \oplus \cdots \oplus \mathbb{Z}f_d$ for some basis $(f_1,\dots,f_d)$ of $\mathbb{R}^d$, then $\Delta = g\cdot\mathbb{Z}^d$, where $g \in \mathrm{GL}(d,\mathbb{R})$ is the matrix of the basis $(f_1,\dots,f_d)$ in the canonical basis $(e_1,\dots,e_d)$. So the set of all lattices is identified with a single orbit $\mathrm{GL}(d,\mathbb{R})\cdot\mathbb{Z}^d$ of this action. The stabilizer of the lattice $\mathbb{Z}^d$ is equal to the subgroup $\mathrm{SL}^{\pm}(d,\mathbb{Z})$ if integer matrices whose inverse also has integral coefficients.

We conclude that X^+ is identified with $\mathrm{GL}(d,\mathbb{R})/\mathrm{SL}^{\pm}(d,\mathbb{Z})$, and so we endow it with the corresponding quotient topology. So, a sequence $\Delta_n = g_n\cdot\mathbb{Z}^d \in X^+$ converges to a lattice $\Delta = g\cdot\mathbb{Z}^d \in X^+$ if and only if there exists a sequence $\gamma_n \in \mathrm{SL}^{\pm}(d,\mathbb{Z})$ such that $g_n\gamma_n$ converges to g inside G.

Exercise 7.13. Check that the latter is equivalent to finding a basis $(f_{n,1},\dots,f_{n,d})$ of Δ_n for each n, that converges to a basis of Δ.

Definition 7.14. The *covolume* d_Δ of a lattice $\Delta \subset \mathbb{R}^d$ is the absolute value of the determinant of any of its basis. In other words, the covolume of $g\cdot\mathbb{Z}^d$ is the determinant of g.

We denote by X the set of lattices of covolume 1. It is a (closed) subset of X^+ and it is naturally identified with $\mathrm{SL}(d,\mathbb{R})/\mathrm{SL}(d,\mathbb{Z})$.

3.1. An embedding. Take a $\mathbb{Q}$-subgroup $\mathbb{G} \subset \mathrm{GL}(d,K)$, and the corresponding arithmetic subgroup $\mathbb{G}_\mathbb{Z}$. Denote by $\mathbb{H} := \mathrm{GL}(d,K)$. Then we have an injection

$$i : \mathbb{G}_\mathbb{R}/\mathbb{G}_\mathbb{Z} \hookrightarrow \mathbb{H}_\mathbb{R}/\mathbb{H}_\mathbb{Z} = X^+.$$

The following proposition shows that this inclusion behaves well topologically. We will use it later to reduce our study of the quotient $\mathbb{G}_\mathbb{R}/\mathbb{G}_\mathbb{Z}$ to a study on the space of lattices in $\mathbb{R}^d$.

Proposition 7.15. *Assume that $\mathbb{G}$ does not admit a non-trivial $\mathbb{Q}$-character $\mathbb{G} \to \mathbb{G}_m$. Then the injection i is proper, in the sense that the pre-image of a compact set is a compact set.*

Proof. Observe that it suffices to show that if $g_n \in \mathbb{G}_\mathbb{R}$ is a sequence such that $g_n\mathbb{H}_Z$ converges inside X^+ then $g_n\mathbb{G}_\mathbb{Z}$ converges in $\mathbb{G}_\mathbb{R}/\mathbb{G}_\mathbb{Z}$.

We apply Chevalley's Theorem 3.16 to the subgroup $\mathbb{G} < \mathbb{H}$: we find a $\mathbb{Q}$-representation of $\mathbb{H}$ on a $\mathbb{Q}$-vector space $\mathbb{W}$ and a point $w \in \mathbb{W}_\mathbb{Q}$ such that $\mathbb{G}$ is exactly the set of elements in $\mathbb{H}$ which stabilize the line Kw. In particular, for all $g \in \mathbb{G}$, there exists $\alpha(g) \in K$ such that $g\cdot w = \alpha(g)w$. Then we see that α is a group homomorphism, and since it is a coefficient map with respect to an element in $\mathbb{W}_\mathbb{Q}$, it is $\mathbb{Q}$-rational. Thus α is a $\mathbb{Q}$-character on $\mathbb{G}$, and must be trivial by assumption. So we conclude that $\mathbb{G}$ is actually the stabilizer of w.

By Lemma 7.11, we know that $\mathbb{H}_\mathbb{Z}$ preserves a lattice Δ of $\mathbb{W}_\mathbb{Q}$. For an appropriate integer $n \geq 1$, we have that $\Delta' := \frac{1}{n}\Delta$ is also preserved by $\mathbb{H}_\mathbb{Z}$, and it contains w. Thus, the $\mathbb{H}_\mathbb{Z}$-orbit of w is contained in Δ': it is discrete in $\mathbb{W}_\mathbb{R}$.

Since $g_n\mathbb{H}_Z$ converges inside X^+ , we may find elements $\gamma_n \in \mathbb{H}_\mathbb{Z}$ such that $g_n\gamma_n$ converges to some $g \in \mathbb{H}_\mathbb{R}$. The sequence $\gamma_n^{-1}w = \gamma_n^{-1}g_n^{-1}w$ converges to $g^{-1}w$, and belongs to a

discrete set, so it is eventually constant. So we may find an index n_0 such that $\gamma_n^{-1}w = \gamma_{n_0}^{-1}w$ for all $n \geq n_0$. Then for all $n \geq n_0$, the element $a_n := \gamma_n\gamma_{n_0}^{-1}$ preserves w for large n, and hence belongs to $\mathbb{G}_{\mathbb{R}} \cap \mathbb{H}_{\mathbb{Z}} = \mathbb{G}_Z$. Further, we see that the sequence $g_n a_n$ converges inside $\mathbb{G}_{\mathbb{R}}$ towards $g\gamma_{n_0}^{-1}$. This proves that indeed, $g_n\mathbb{G}_{\mathbb{Z}}$ converges in $\mathbb{G}_{\mathbb{R}}/\mathbb{G}_{\mathbb{Z}}$. □

3.2. A compactness criterion. The following criterion, due to Mahler, allows to determine when a subset of X or X^+ is relatively compact.

THEOREM 7.16 (Mahler). *A subset $Y \subset X^+$ is relatively compact if and only if there exist constants $C, c > 0$ such that*

$$d_\Delta < C \text{ and } \inf_{v \in \Delta \setminus \{0\}} \|v\| > c, \quad \text{for all } \Delta \in Y.$$

In other words, the subset is relatively compact if its elements don't have arbitrarily large covolume, and if they don't admit arbitrarily small vectors. In particular, if $Y \subset X$, only the second condition matters.

Let us start the proof with the following inequality of Hermite-Minkowski that tells us that a lattice with a given covolume d cannot have only large vectors.

LEMMA 7.17 (Hermite-Minkowski). *Any lattice Δ in $\mathbb{R}^d$ contains a vector $v \neq 0$ such that $\|v\| \leq 2(d_\Delta/v_d)^{1/d}$, where v_d denotes the volume of the unit ball inside $\mathbb{R}^d$.*

PROOF. Denote by r the minimal length of a vector in $\Delta \setminus \{0\}$. Then any two distinct vectors of Δ are at least r-apart, which tells us that the quotient map $\mathbb{R}^d \to \mathbb{R}^d/\Delta$ is injective on any open ball of diameter r (i.e. of radius $r/2$).This is easily seen to imply that the volume of such a ball, equal to $(r/2)^d v_d$ is at most equal to the measure of a fundamental domain for Δ, equal to d_Δ. So $(r/2)^d v_d \leq d_\Delta$, which clearly implies the result. □

PROOF OF THEOREM 7.16. First we note that the functions $g \in \mathrm{GL}(d, \mathbb{R}) \mapsto d_{g\cdot\mathbb{Z}^d} = |\det(g)|$ and $g \mapsto (\min_{v \in g\cdot\mathbb{Z}^d \setminus \{0\}} \|v\|)^{-1}$ are continuous, so the corresponding functions on X^+ are also continuous. Hence they must be bounded on any relatively compact subset of X^+.

Conversely, let us show by induction on the dimension d that any subset $Y \subset X^+$ for which there exists constants $C, c > 0$ such that $d_\Delta < C$ and $\|v\| \geq c$ for all $\Delta \in Y$ and all $v \in \Delta \setminus \{0\}$, is relatively compact.

Observe that if $d = 1$ this is trivial. Assume now that $d > 1$ and that $Y \subset X^+$ is such a subset. Let us take a sequence $\Delta_n \in Y$ and try to extract a subsequence that converges in X^+. Since the covolumes d_{Δ_n} are bounded by C, we may apply Hermite-Minkowski's inequality to find for each n a vector $v_n \in \Delta_n \setminus \{0\}$ with norm at most $2(C/v_d)^{1/d}$. We may assume that v_n is a vector of minimal norm of $\Delta_n \setminus \{0\}$. By assumption on Y, we know that $\|v_n\| \geq c$ for all n. So after passing to a subsequence of Δ_n, we may find a non-zero vector $v_\infty \in \mathbb{R}^d$ such that v_n converges to v_∞.

For each n, denote by $g_n \in \mathrm{GL}(d, \mathbb{R})$ an element such that $g_n(v_n) = v_\infty$, and we assume moreover that $g_n \to \mathrm{id}$. Note that if we find a convergent subsequence of $(g_n\Delta_n)_n$, then the corresponding subsequence of $(\Delta_n)_n$ will also converge to the same limit. So without loss of generality we may assume that v_∞ actually belongs to each Δ_n. Note that if we chose for g_n the product of a dilation and a rotation, then we can in fact assume that v_∞ is of minimal norm inside $\Delta_n \setminus \{0\}$.

Now let us look at the euclidean subspace $v_\infty^\perp \subset \mathbb{R}^d$. Denote by Δ_n' the projection of Δ_n onto $v_\infty^\perp$.

Claim 1. Δ_n' is a lattice in $v_\infty^\perp$ with covolume $d_{\Delta_n}/\|v_\infty\|$.

This is classical. The fact that v_∞ is a vector of minimal norm inside Δ_n tells us that Δ_n' is discrete inside $v_\infty^\perp$, and it clearly spans $v_\infty^\perp$ as a vector space. So it is indeed a lattice. Taking a basis $(f_1, \dots, f_{d-1})$ of Δ_n' we may thus find constants α_i such that $(v_\infty, f_1 + \alpha_1 v_\infty, \dots, f_{d-1} + \alpha_{d-1} v_\infty)$ is a basis of Δ. The determinant of this basis is equal to $\|v_\infty\| \det(f_1, \dots, f_{d-1})$, as desired.

Claim 2. Any non-zero vector of Δ_n' has norm at least $c\sqrt{3}/2$.

Denote by $r := \|v_\infty\| \geq c$. If $x \in \Delta_n'$ is a non-zero element, then we may find $\alpha \in \mathbb{R}$ such that $x + \alpha v_\infty \in \Delta$. Subtracting an integer multiple of v_∞, we may assume that $|\alpha| \leq r/2$. Since $x \perp v_\infty$ and since v_∞ has minimal norm in $\Delta_n \setminus \{0\}$, we get

$$\|x\|^2 + |\alpha|^2 = \|x + \alpha v_\infty\|^2 \geq r^2.$$

Thus, $\|x\| \geq r^2 - r^2/4 \geq c\sqrt{3}/2$.

These two claims and the induction hypothesis imply that we may replace (Δ_n') by a subsequence to assume that it converges toward a lattice Δ' in $v_\infty^\perp$. So we may find a basis $(f_{n,1}, \dots, f_{n,d-1})$ of Δ_n' that converges to a basis $(f_1, \dots, f_{d-1})$ of Δ'. So we may find $\alpha_{n,1}, \dots, \alpha_{n,d-1} \in \mathbb{R}$ such that $|\alpha_{n,i}| \leq r/2$ for all i and $\mathcal{B}_n := (v_\infty, f_{n,1} + \alpha_{n,1} v_\infty, \dots, f_{n,d-1} + \alpha_{n,d-1} v_\infty)$ is a basis of Δ_n for each n. Then we may extract a subsequence along which $\alpha_{n,i}$ converges. Then we see that the basis $\mathcal{B}_n$ converges to a basis of $\mathbb{R}^d$, and hence Δ_n converges to some lattice of $\mathbb{R}^d$. This finishes the proof. □

3.3. Δ-rational subspaces. We finish our discussion on the space of lattices with some facts about lower dimensional subspaces and their position relative to a fixed lattice $\Delta \subset \mathbb{R}^d$. This will be used later.

Definition 7.18. Fix a lattice $\Delta \subset \mathbb{R}^d$. A subspace $L \subset \mathbb{R}^d$ is Δ-rational if it is spanned over $\mathbb{R}$ by its intersection with Δ.

Note that L is Δ-rational if and only if $L \cap \Delta$ is a lattice in L. As a subspace of $\mathbb{R}^d$, L is a euclidean space, we may then naturally speak of the covolume of $\Delta \cap L$ in L for this euclidean structure. We denote by $d(L) = d_\Delta(L)$ this covolume. We set $d(\{0\}) = 1$.

Lemma 7.19. *Assume that $\Delta = \mathbb{Z}^d \subset \mathbb{R}^d$ and take a subspace $L \subset \mathbb{R}^d$. The following are equivalent*

(i) L-is Δ-rational;
(ii) L is spanned over $\mathbb{R}$ by vectors in $\mathbb{Q}^d$;
(iii) L is defined as the zero set in $\mathbb{R}^d$ of linear equations with rational coefficients.

Proof. $(i) \Rightarrow (ii)$. By assumption L is spanned by $L \cap \Delta \subset \mathbb{Q}^d$.
$(ii) \Rightarrow (i)$. Take a basis $v_1, \dots, v_\ell$ of L made of rational vectors. Then some integer multiple $nv_1, \dots, nv_\ell$ is still a basis of L, but is made of integer vectors, i.e. contained in $\Delta \cap L$.
$(ii) \Leftrightarrow (iii)$. Observe that (iii) amounts to saying that $L^\perp$ is spanned by vectors with rational coefficients, i.e. is Δ-rational. So we need to verify that L is Δ-rational if and only if $L^\perp$ is Δ-rational. This easily follows from the fact that the canonical inner product on $\mathbb{R}^d$ is defined over $\mathbb{Q}$. □

LEMMA 7.20. *Take a lattice $\Delta \subset \mathbb{R}^d$ and two Δ-rational subspaces $L, M \subset \mathbb{R}^d$. Then $L+M$ and $L \cap M$ is still Δ-rational and we have the following relation on the covolumes:*

$$d(L \cap M)d(L+M) \leq d(L)d(M).$$

PROOF. To check the rationality statements, we may change the basis in $\mathbb{R}^d$ in order to assume that $\Delta = \mathbb{Z}^d$. In this case we may apply the previous lemma. Then $L+M$ is spanned by vectors in $\mathbb{Q}^d$ so it is Δ-rational. Moreover, $L \cap M$ is the zero set of linear equations with coefficients in $\mathbb{Q}$, so it is Δ-rational.

Let us now prove the inequality about the covolumes. Take a basis $u_1, \dots, u_k$ of $\Delta \cap L \cap M$, and add vectors $v_1, \dots, v_\ell$ to get a basis $u_1, \dots, u_k, v_1, \dots, v_\ell$ of $\Delta \cap L$ and vectors $w_1, \dots, w_m$ to get a basis $u_1, \dots, u_k, w_1, \dots, w_m$ of $\Delta \cap M$. Then the family of vectors $\mathcal{B} := (u_1, \dots, u_k, v_1, \dots, v_\ell, w_1, \dots, w_m)$ spans a lattice inside $L+M$, which has finite index inside $\Delta \cap (L+M)$. In particular the determinant of this family $\mathcal{B}$ is greater than or equal to the covolume of $\Delta \cap (L+M)$ inside $L+M$. So the stated inequality will follow once we prove that

$$|\det(u_1, \dots, u_k)||\det(\mathcal{B})| \leq |\det(u_1, \dots, u_k, v_1 \dots, v_\ell)||\det(u_1, \dots, u_k, w_1, \dots, w_m)|.$$

Note that this inequality does not involve the lattice Δ anymore, but is purely of euclidean nature (note that the determinant is always taken with respect of an orthonormal basis of the subspace we consider). Using Schmidt orthogonalization algorithm, we may assume that the families $(u_1, \dots, u_k, v_1, \dots, v_\ell)$ and $(u_1, \dots, u_k, w_1, \dots, w_m)$ are both orthogonal families. Since the inequality is insensitive to rescaling each vector, we may assume that these bases are orthonormal. Besides, we may find vectors $w'_1, \dots, w'_m \in L+M$ so that $(u_1, \dots, u_k, v_1, \dots, v_\ell, w'_1, \dots, w'_m)$ is an orthonormal basis of $L+M$. In this basis, the matrix of the family $\mathcal{B} = (u_1, \dots, u_k, v_1, \dots, v_\ell, w_1, \dots, w_m)$ is

$$\begin{pmatrix} \mathrm{id}_k & 0 & 0 \\ 0 & \mathrm{id}_\ell & * \\ 0 & 0 & A \end{pmatrix},$$

where $*$ designate possibly non-zero terms. Thus $\det(\mathcal{B}) = \det(A)$, and the matrix A is the matrix of the vectors $(Pw_1, \dots, Pw_m)$, where P is the orthogonal projection onto the linear span of $(w'_1, \dots, w'_m)$. Since P has norm at most 1, as an operator on the euclidean space $\mathbb{R}^d$, it is also the case of its restriction to the linear span of $w_1, \dots, w_m$. This restriction thus has determinant at most 1. Hence $|\det(A)| \leq |\det(P)| \leq 1$. This proves the desired inequality. □

LEMMA 7.21. *Fix a lattice $\Delta \subset \mathbb{R}^d$. For $0 \leq i \leq d$, denote by $S_i(\Delta)$ the set of Δ-rational subspaces of $\mathbb{R}^d$ of dimension i. Then for each i and $C > 0$ the set $\{L \in S_i(\Delta) \mid d_\Delta(L) \leq C\}$ is finite. Moreover the following map is continuous:*

$$\theta_i : \Delta \in X \mapsto \min_{L \in S_i(\Delta)} d_\Delta(L).$$

PROOF. Note that the set of elements $v_1 \wedge \dots \wedge v_i$, where $v_1, \dots, v_i \in \Delta$ is discrete inside $\bigwedge^i(\mathbb{R}^d)$. Take $L \in S_i(\Delta)$, and take an arbitrary basis $v_1, \dots v_i$ of $\Delta \cap L$. Then the map $L \mapsto v_1 \wedge \dots \wedge v_i \in \bigwedge^i(\mathbb{R}^d)$ is injective, so there are only many $L \in S_i(\Delta)$ such that $d_\Delta(L) = \|v_1 \wedge \dots \wedge v_i\| \leq C$. This proves the first statement.

The second statement follows from the inequalities $\|g^{-1}\|^{-d} d_\Delta(L) \leq d_{g\Delta}(gL) \leq \|g\|^d d_\Delta(L)$, for all $g \in SL(d, \mathbb{R})$. □

3.4. Proof of the co-compactness criterion. We fix a semi-simple $\mathbb{Q}$-group $\mathbb{G}$ and we aim to prove the second statement of Theorem 7.12. We start with a criterion regarding unipotent elements. Denote by $\mathfrak{g}$ the Lie algebra of $\mathbb{G}$. Since $\mathbb{G}$ is defined over $\mathbb{Q}$, so is $\mathfrak{g}$, i.e. $\mathfrak{g} = \mathfrak{g}_{\mathbb{Q}} \otimes_{\mathbb{Q}} K$.

LEMMA 7.22. *The following assertions are equivalent.*

(i) $\mathbb{G}_{\mathbb{Z}}$ *has non-trivial unipotent elements;*
(ii) $\mathbb{G}_{\mathbb{Q}}$ *has non-trivial unipotent elements;*
(iii) $\mathfrak{g}_{\mathbb{Q}}$ *has non-zero nilpotent elements.*

PROOF. We observed in the first graded homework that if we embed $\mathbb{G}$ as a $\mathbb{Q}$-subgroup of $\mathrm{GL}(\mathbb{V})$, then the exponential of matrices restricts to a $\mathbb{Q}$-rational map from the set of nilpotent matrices inside $\mathfrak{g}$ (viewed inside $\mathrm{End}(\mathbb{V})$) onto the set of unipotent elements in $\mathbb{G}$. This proves the equivalence between (ii) and (iii). Moreover since the coefficients of this exponential map are in $\mathbb{Q}$, and the constant term is $1 \in \mathbb{Z}$, then we see that if $A \in \mathfrak{g}_{\mathbb{Q}}$ is nilpotent then $\exp(nA) \in \mathrm{GL}(\mathbb{V})$ has integral coefficients for an appropriate n. So (iii) implies (i). □

PROOF OF THEOREM 7.12.2. We assume that $\mathbb{G}$ has trivial center. In particular, $\mathbb{G}$ is faithfully represented as a $\mathbb{Q}$-subgroup of $\mathbb{H} := \mathrm{GL}(\mathfrak{g})$ via its adjoint representation. Since $\mathbb{G}$ is semi-simple, $\mathrm{Ad}(\mathbb{G})$ is contained in $\mathrm{SL}(\mathfrak{g})$.

Assume first that $\mathbb{G}_{\mathbb{Z}}$ is not cocompact inside $\mathbb{G}_{\mathbb{R}}$. By Proposition 7.15, the image of $\mathbb{G}_{\mathbb{R}}/\mathbb{G}_{\mathbb{Z}}$ inside the space of lattices $\mathbb{H}_{\mathbb{R}}/\mathbb{H}_{\mathbb{Z}}$ is not compact. By Mahler compactness criterion, this means that we may find a sequence $g_n \in \mathbb{G}_{\mathbb{R}}$ such that the sequence of lattices $\mathrm{Ad}(g_n) \cdot \mathfrak{g}_{\mathbb{Z}} \subset \mathfrak{g}$ has no convergent subsequence in $\mathbb{H}_{\mathbb{R}}/\mathbb{H}_{\mathbb{Z}}$. Since all these lattices have the same covolume, Mahler compactness criterion tells us in fact that there exists a sequence of vectors $X_n \in \mathfrak{g}_{\mathbb{Z}} \setminus \{0\}$ such that $\mathrm{Ad}(g_n)(X_n)$ converges to 0. This implies that $\mathrm{Tr}(X_n^i)$ converges to 0 for all $i \geq 0$. Since the set $\mathrm{Tr}(\mathfrak{g}_{\mathbb{Z}})$ is discrete in $\mathbb{R}$, this shows that for n large enough $\mathrm{Tr}(X_n^i) = 0$ for all $i \leq \dim(\mathfrak{g})$. So for n large enough, X_n is a non-zero nilpotent element in $\mathfrak{g}_{\mathbb{Z}}$. So the previous lemma shows that $\mathbb{G}_{\mathbb{Z}}$ has a non-trivial unipotent element.

Conversely assume that $\mathbb{G}_{\mathbb{Z}}$ is co-compact inside $\mathbb{G}_{\mathbb{R}}$ and take a unipotent element $u \in \mathbb{G}_{\mathbb{Z}}$. The next two claims will show that u is necessarily trivial. We denote by $X := \log(u)$ which is a nilpotent element of $\mathfrak{g}$.

Claim 1. There exists a sequence $(g_n)_{n\geq 1}$ in $\mathbb{G}_{\mathbb{R}}$ such that $g_n u g_n^{-1}$ converges to the trivial element $e_{\mathbb{G}}$, in the analytic topology.

To prove this claim we don't use that fact that $u \in \mathbb{G}_{\mathbb{Z}}$, the proof relies only on the Lie group structure of $\mathbb{G}_{\mathbb{R}}$. We denote by $\exp : \mathfrak{g}_{\mathbb{R}} \to \mathbb{G}_{\mathbb{R}}$ the analytic exponential map. We may apply Jacobson-Morozov Theorem to find a semi-simple element $H \in \mathfrak{g}_{\mathbb{R}}$ such that $[H, X] = 2X$. For all $n \geq 1$, denote by $g_n := \exp(-nH)$. By Lemma 2.17 we have $\mathrm{Ad}(g_n)(X) = \exp(\mathrm{ad}(-nH))(X)$, where the later exponential refer to the exponential of endomorphisms $\mathrm{End}(\mathfrak{g}_{\mathbb{R}}) \to \mathrm{GL}(\mathfrak{g}_{\mathbb{R}})$. The relation $\mathrm{ad}(H)(X) = 2X$ then implies that $\exp(\mathrm{ad}(-nH))(X) = \exp(-2n)X$. In particular, we deduce that $\mathrm{Ad}(g_n)(X)$ converges to 0 inside $\mathfrak{g}_{\mathbb{R}}$ as $n \to \infty$. Applying the exponential map we get the desired convergence:

$$\lim_n g_n u g_n^{-1} = \lim_n \exp(\mathrm{Ad}(g_n)X) = e_{\mathbb{G}}.$$

Claim 2. The sequence (g_n) of Claim 1 may be taken inside $\mathbb{G}_{\mathbb{Z}}$. Hence $g_n u g_n^{-1} = e_{\mathbb{G}}$ for n large enough.

Since $\mathbb{G}_{\mathbb{R}}/\mathbb{G}_{\mathbb{Z}}$ is compact, we may find elements $\gamma_n \in \mathbb{G}_{\mathbb{Z}}$ such that $g_n\gamma_n$ converges to some $h \in \mathbb{G}_{\mathbb{R}}$. Then we see that $\gamma_n^{-1}u\gamma_n = (g_n\gamma_n)^{-1}g_nug_n^{-1}(g_n\gamma_n)$ converges to $h^{-1}e_{\mathbb{G}}h = e_{\mathbb{G}}$. Now this sequence is contained inside $\mathbb{G}_{\mathbb{Z}}$, which is discrete inside $\mathbb{G}_{\mathbb{R}}$, so it is eventually equal to $e_{\mathbb{G}}$. So some conjugate of u is trivial, proving that u is trivial. □

CHAPTER 8

Finiteness of measures

In this chapter we prove a dynamical criterion ensuring that a Radon measure on a locally compact space X which is invariant under some continuous action of a group G is finite. We then apply this criterion to complete the proof of Theorem 7.12.

1. An abstract criterion

Let G be an l.c.s.c. group and X be a locally compact topological space on which G acts continuously. We provide here a general criterion ensuring that any G-invariant Radon measure ν on X is finite.

We recall that a continuous function $f : X \to [0, +\infty[$ is said to be proper if the inverse image of a compact set is compact (i.e. $f(x)$ tends to infinity as x leaves compact subsets of X).

Definition 8.1. Given a probability measure μ on G, we define the *averaging operator* A_μ, by the formula

$$A_\mu(f) : x \in X \mapsto \int_G f(gx) d\mu(g),$$

for all functions f on X for which the integral makes sense (e.g. for all continuous functions $f : X \to [0, +\infty[$).

Observe that if $A \subset X$ is a measurable subset of X and if $x \in X$, then $A_\mu(\mathbf{1}_A)(x)$ is equal to the measure of the set of elements $g \in G$ such that $gx \in A$, i.e. to the probability that gx lies in A, when g is distributed according to μ.

We consider the following condition of contraction of some averaging operator:

CA There exists a proper continuous function $f : X \to [0, +\infty[$, a probability measure μ on G and constants $a < 1$, $b > 0$ such that $A_\mu(f) \leq af + b$.

Proposition 8.2. *Assuming condition* **CA**, *every G-invariant Radon measure ν on X is finite.*

The proof is based on the following lemma, which states that the random walk on X, whose starting point is x and which moves at each step according to the law μ, is essentially contained, after a long time, in a compact space K.

Lemma 8.3. *Assume condition* **CA**. *For all $\varepsilon > 0$ there exists a compact set $K \subset X$ such that for all $x \in X$ there exists an integer $M = M(x)$ satisfying*

$$A_\mu^n(\mathbf{1}_K)(x) \geq 1 - \varepsilon, \text{ for all } n \geq M.$$

Moreover, we may choose $M(x)$ to be bounded on compact sets.

PROOF. Note that the operator A_μ is linear and acts trivially on constant functions. By induction, we have, for all $n \geq 1$,

$$A_\mu^n(f) \leq a^n f + b(1 + a + \cdots + a^{n-1}) \leq a^n f + \frac{b}{1-a}.$$

Since f is proper, the following set K is compact:

$$K := \left\{ x \in X \mid f(x) \leq \frac{2b}{\varepsilon(1-a)} \right\}.$$

Moreover $\mathbf{1}_{X\setminus K} \leq \frac{(1-a)\varepsilon}{2b} f$, and it follows that for all $x \in X$,

$$A_\mu^n(\mathbf{1}_{X\setminus K})(x) \leq \frac{(1-a)\varepsilon}{2b} A_\mu^n(f)(x) \leq \frac{(1-a)\varepsilon}{2b} \left(a^n f(x) + \frac{b}{1-a} \right) \leq \varepsilon,$$

if n is large enough. Passing to the complementary, we get the result. The moreover part is clear since f is continuous (hence bounded on compact sets). □

PROOF OF PROPOSITION 8.2. Take $\varepsilon = 1/2$ and denote by K the corresponding compact set given by the previous lemma. Then for any compact set $L \subset X$, we may find an integer M so that

$$A_\mu^n(\mathbf{1}_K)(x) \geq 1/2, \text{ for all } x \in L, n \geq M.$$

In particular we have,

$$\nu(L) \leq 2 \int_L A_\mu^n(\mathbf{1}_K)(x) d\nu(x) \leq 2 \int_X A_\mu^n(\mathbf{1}_K)(x) d\nu(x) = 2 \int_X \mathbf{1}_{\mathbf{K}}(x) d\nu(x) = 2\nu(K),$$

because $\int_X A_\mu(F)(x) d\nu(x) = \int_X F d\nu$ for every non-negative function F, since ν is G-invariant. Since ν is a regular measure, we conclude that $\nu(X) = \sup_L \nu(L) \leq 2\nu(K)$, proving the result. □

2. Expansion in linear random walks

In this section we collect properties of linear actions on vector spaces and the corresponding actions on projective spaces.

We endow $\mathbb{R}^d$ with an inner product $\langle \cdot, \cdot \rangle$ and denote by $\| \cdot \|$ the corresponding euclidean norm. Then we also denote by $\| \cdot \|$ the corresponding operator norm on $M(d, \mathbb{R})$.

Denote by $G := \mathrm{GL}(d, \mathbb{R})$, and by $S^+ \subset G$ the subset of positive definite matrices. Take a probability measure μ on G, denote by $\mathrm{supp}(\mu)$ its support and by Γ_μ the closure of the subgroup generated by $\mathrm{supp}(\mu)$. We say that μ is symmetric if $\mu(A) = \mu(A^{-1})$ for all subset $A \subset G$, where $A^{-1} = \{g^{-1} \mid g \in A\}$.

The following result is a special case of a theorem of Furstenberg.

THEOREM 8.4. *Assume that μ is symmetric, with $\mathrm{supp}(\mu) \subset S^+$ and that it has a finite first moment: $\int_G |\log(\|g\|)| d\mu(g) < \infty$. Then for every non-zero vector $v \in \mathbb{R}^d$, we have*

$$\int_G \log(\frac{\|gv\|}{\|v\|}) d\mu(g) \geq 0.$$

This inequality is an equality if and only if Γ_μ stabilizes the line $\mathbb{R}v$.

PROOF. Using the fact the μ is symmetric, we see that

$$\int_G \log(\|gv\|)d\mu(g) = \frac{1}{2}\int_G \log(\|gv\|\|g^{-1}v\|)d\mu(g).$$

Now using Cauchy-Shwartz inequality and the fact that μ is supported on symmetric matrices we get

$$\int_G \log(\|gv\|)d\mu(g) \geq \frac{1}{2}\int_G \log(\langle gv, g^{-1}v\rangle)d\mu(g) = \int_G \log(\|v\|)d\mu(g).$$

This proves the desired inequality. Observe that this is an equality if and only if we have equality in the Cauchy-Schwartz inequality used above. This amounts to saying that for μ-almost every $g \in G$, gv and $g^{-1}v$ are proportional, i.e. the v is an eigenvector of g^2. Since g is almost surely positive definite this implies that v is almost surely an eigenvector of g. Hence Γ_μ stabilizes the line $\mathbb{R}v$ □

For each $1 \leq i \leq d$, we endow the exterior product $\bigwedge^i \mathbb{R}^d$ with the inner product such that the vectors $e_{n_1} \wedge \cdots \wedge e_{n_i}$ form an orthonormal basis as the i-tuple $1 \leq n_1 < \cdots < n_i \leq d$ vary and $e_1, \ldots, e_d$ denotes an orthonormal basis of $\mathbb{R}^d$. Keep denoting by $\|\cdot\|$ the associated Euclidean norm. The group $G = \mathrm{GL}(d, \mathbb{R})$ acts linearly on $\bigwedge^i \mathbb{R}^d$ by the formula $g \cdot (v_1 \wedge \cdots \wedge v_i) = (gv_1) \wedge \cdots \wedge (gv_i)$. It preserves the cone of non-zero *elementary vectors*

$$W_i := \{v = v_1 \wedge \cdots \wedge v_i \neq 0 \mid v_i \in \mathbb{R}^d\}.$$

The quotient of W_i by dilations is the *Grassmanian variety* $\mathbb{G}_i^d$ of i-dimensional subspaces of $\mathbb{R}^d$. We denote by $\varphi_i : W_i \to]0, +\infty[$ the function given by

$$\varphi_i(v) := \|v\|^{-1}, \text{ for all } v \in W_i.$$

The next corollary shows that these functions are contracted under averaging operators (under soft conditions). It will be used to construct some functions on the space of lattices that satisfy assumption **CA** in the suitable setting.

The conditions that we require on $\mu \in \mathrm{Prob}(G)$ are the following

ISC Γ_μ acts irreducibly on $\mathbb{R}^d$, in the sense that there is no non-zero Γ_μ-invariant proper subspace of $\mathbb{R}^d$; μ is symmetric and the support $\mathrm{supp}(\mu)$ is compact and contained in S^+.

COROLLARY 8.5. *Assume that $\mu \in \mathrm{Prob}(G)$ satisfies condition* **ISC**. *Then there exists $\delta > 0$ and $a_0 < 1$ such that for all $0 < i < d$ we have $A_\mu(\varphi_i^\delta) \leq a_0\varphi_i^\delta$.*

PROOF. We may apply Theorem 8.4 to the linear action of G on $\bigwedge^i \mathbb{R}^d$. We find that for all $v \in W_i$, $I_v := \int_G \log(\frac{\|gv\|}{\|v\|})d\mu(g) \geq 0$. Moreover, since Γ_μ acts irreducibly on $\mathbb{R}^d$, it does not stabilize a line in $\mathbb{R}v$, for any $v \in W_i$. So $I_v > 0$ for all $v \in W_i$. In fact note that the map $v \in W_i \mapsto I_v$ is continuous and invariant under dilations. So it factors through a continuous map $\mathbb{G}_i^d \to \mathbb{R}$. Since $\mathbb{G}_i^d$ is compact (being a closed subspace of the projective space $\mathbb{P}(\bigwedge^i \mathbb{R}^d)$), we conclude that there exists a constant $c > 0$ such that $I_v \geq 2c$ for all $v \in W_i$.

Observe that for all $0 < i < d$, and for all $v \in W_i$, we have

$$\|gv\| \leq \|g\|^i\|v\| \leq \max(\|g\|^d, 1)\|v\|.$$

Set $M := \sup\{\log(\max(\|g\|^d, 1)) \mid g \in \mathrm{supp}(\mu)\}$ and $\delta := \min(1/M, c/M^2)$. Using the fact that $e^t \leq 1 + t + t^2$ for all $t \in [-1, 1]$ we compute

$$\begin{aligned}\frac{A_\mu(\varphi_i^\delta)(v)}{\varphi_i^\delta(v)} &= \int_G \exp(-\delta \log(\frac{\|gv\|}{\|v\|}))d\mu(g) \\ &\leq 1 - \delta \int_G \log(\frac{\|gv\|}{\|v\|})d\mu(g) + \delta^2 \int_G \left(\log(\frac{\|gv\|}{\|v\|})\right)^2 d\mu(g) \\ &\leq 1 - 2c\delta + M^2\delta^2 \leq 1 - c\delta.\end{aligned}$$

We may thus set $a_0 := 1 - c\delta < 1$. □

3. Proof of the main Theorem

We are now ready to prove the first item of Theorem 7.12

First there are several reductions. Let $\mathbb{G}$ be a semi-simple $\mathbb{Q}$-group with trivial center. Note that $\mathbb{G}^0$ has finite index inside $\mathbb{G}$, so we may easily reduce to the case where $\mathbb{G}$ is connected. By Proposition 7.2, we may then write $\mathbb{G}$ as a direct product of finitely many simple $\mathbb{Q}$-groups $\mathbb{G}_i$, which are connected and have finite center. Then the set of integral (resp. real) point also splits as the product of the set of integral (resp.real) points Γ_i inside $\mathbb{G}_i$. Moreover, it is easily seen that if Γ_i is a lattice inside $\mathbb{G}_i$ for all i, then the finite product $\Pi_i\Gamma_i$ is a lattice inside $\Pi_i\mathbb{G}_i$. So it suffices to prove the theorem for $\mathbb{Q}$-simple groups.

Finally, we already proved the second item of the theorem which implies in particular that if $\mathbb{G}$ is $\mathbb{Q}$-anisotropic then $\mathbb{G}_\mathbb{Z}$ is a lattice in $\mathbb{G}_\mathbb{R}$.

In conclusion, we may assume that $\mathbb{G}$ is a $\mathbb{Q}$-simple, $\mathbb{Q}$-isotropic group, which is connected and has trivial center.

3.1. First step: ensuring condition ISC. We construct a nice representation of $\mathbb{G}$ and a probability measure on $\mathbb{G}$ satisfying condition **ISC**.

Lemma 8.6. *Let $\mathbb{G}$ be as above. There exists a faithful $\mathbb{Q}$-representation of $\mathbb{G}$ which is irreducible over $\mathbb{R}$.*

In the proof of this lemma we will need the following exercise.

Exercise 8.7. Prove that if G and H are two arbitrary groups and $\pi : G \to \mathrm{GL}(V)$ and $\rho : H \to \mathrm{GL}(W)$ are two irreducible representations of G and H on finite dimensional complex vector spaces then the representation $\widetilde{\pi} : G \times H \to \mathrm{GL}(V \otimes W)$ given by $\widetilde{\pi}(g, h)(v \otimes w) = (\pi(g)v) \otimes (\rho(h)w)$ for all $v \in V$, $w \in W$ is again irreducible.
Hint. Use Burnside theorem which state that if $\pi : G \to \mathrm{GL}(V)$ is irreducible then the linear span of $\pi(G)$ is the whole of $\mathrm{End}(V)$.

Proof of Lemma 8.6. By assumption, $\mathbb{G}$ is adjoint, so it is faithfully represented over $\mathfrak{g}$. Decompose this adjoint representation into irreducible representations over $K = \mathbb{C}$: $\mathfrak{g} = \bigoplus_{i=1}^n \mathfrak{g}_i$, where each $\mathfrak{g}_i$ is simple over $\mathbb{C}$. For each i, denote by ρ_i the restriction of the adjoint representation of $\mathbb{G}$ to $\mathfrak{g}_i$.

The Galois group of $\mathbb{C}$ over $\mathbb{Q}$ permutes the subalgebras $\mathfrak{g}_i$, and intertwines the representations ρ_i, so the tensor product representation $\rho := \rho_1 \otimes \cdots \otimes \rho_n$ is defined over $\mathbb{Q}$. In order to check that this representation is irreducible, note that the decomposition $\mathfrak{g} = \bigoplus_{i=1}^n \mathfrak{g}_i$ also admits a counterpart a the group level, namely, $\mathbb{G} = \prod_{i=1}^n \mathbb{G}_i$, where

each $\mathbb{G}_i$ is $\mathbb{C}$-simple. Moreover, every representation ρ_i is obtained by composing the projection map $\mathbb{G} \to \mathbb{G}_i$ with the adjoint representation of $\mathbb{G}_i$, so it is in fact a representation of $\mathfrak{g}_i$ and the representation ρ is then of the form presented in the exercise. Since each ρ_i is irreducible, we conclude that ρ is an irreducible representation of $\mathbb{G}$, over $\mathbb{C}$. In particular ρ is irreducible over $\mathbb{R}$, and it is clearly faithful. □

It is a fact that since $\mathbb{G}$ is a connected algebraic group, the analytic group $\mathbb{G}_\mathbb{R}$ has only finitely many connected components. Denote by $(\mathbb{G}_\mathbb{R})^0$ the analytic connected component of $\mathbb{G}_\mathbb{R}$.

LEMMA 8.8. *The semi-simple Lie group $(\mathbb{G}_\mathbb{R})^0$ has no compact factor.*

PROOF. One can check that if $\mathbb{G}$ is connected and has trivial center, then $(\mathbb{G}_\mathbb{R})^0$ is connected and has trivial center. So it splits uniquely as a direct product of simple Lie groups, and by uniqueness, each factor in this decomposition is the identity component of $(\mathbb{G}_i)_\mathbb{R}$, where $\mathbb{G}_i$ is one of the $\mathbb{R}$-simple factors of $\mathbb{G}$ viewed as an $\mathbb{R}$-group. By Proposition 7.5, we see that each $\mathbb{G}_i$ is $\mathbb{R}$-isotropic, which amounts to saying that there is a non-zero nilpotent element in its real Lie algebra $\mathfrak{g}_i$. In particular if $\mathfrak{g}_i = \mathfrak{k} \oplus \mathfrak{q}$ is a Cartan decomposition of $\mathfrak{g}_i$, then we have that $\mathfrak{q} \neq 0$. So Theorem 6.4 shows that $(\mathbb{G}_i)_\mathbb{R}$ is non-compact. □

LEMMA 8.9. *Let $\mathbb{G} \subset \mathrm{GL}(d, \mathbb{C})$ be a connected semi-simple $\mathbb{R}$-subgroup with trivial center. Set $G := (\mathbb{G}_\mathbb{R})^0 \subset \mathrm{GL}(d, \mathbb{R})$. Then*

- *There exists an inner product on $\mathbb{R}^d$ such that if $g \in G$ then $g^T \in G$.*
- *If moreover G has no-compact factor, then $G \cap S^+$ generates a dense subgroup of G, where S^+ denotes the set of positive definite matrices for this inner product.*

PROOF. To construct the inner product, write $\mathfrak{g}_\mathbb{R} = \mathfrak{k} \oplus \mathfrak{q}$ for the Cartan decomposition of the Lie algebra of G. Then $\mathfrak{u} := \mathfrak{k} \oplus i\mathfrak{q}$ is a compact real form of the Lie algebra $\mathfrak{g}$ of $\mathbb{G}$. As such, it is the Lie algebra of a compact Lie subgroup U of $\mathrm{GL}(d, \mathbb{C})$. Then we may find a complex inner product (i.e. a positive definite hermitian form) $\langle \cdot, \cdot \rangle$ on $\mathbb{C}^d$ which is U-invariant. So for all $v, w \in \mathbb{C}^d$, $X \in \mathfrak{k} \oplus i\mathfrak{q}$, and $t \in \mathbb{R}$, we have $\langle \exp(tX)v, \exp(tX)w \rangle = \langle v, w \rangle$. Derivating this relation gives

$$\langle X(v), w \rangle + \langle v, X(w) \rangle = 0.$$

So $X = -X^*$ for all $X \in \mathfrak{k} \oplus i\mathfrak{q}$. This implies that $X = X^*$ for all $X \in \mathfrak{q}$. Now look we define the real inner product $\langle \cdot, \cdot \rangle_\mathbb{R}$ as the real part of the above complex inner product, restricted to $\mathbb{R}^d$. Since $\mathfrak{g}_\mathbb{R} \subset \mathrm{GL}(d, \mathbb{R})$, we see that $X = -X^T$ for all $X \in \mathfrak{k}$ and $X = X^T$ for all $X \in \mathfrak{q}$. In particular, $\mathfrak{g}_\mathbb{R}$ is invariant under the transpose map, and so is G, by connectedness.

In fact the above construction is made so that the Cartan involution with respect to which we made the Cartan decomposition is just the map $X \mapsto -X^T$. In this way, the elements of $\mathfrak{q}$ are symmetric matrices, and their exponential are symmetric positive definite elements of G. Now, note that $[\mathfrak{q}, \mathfrak{q}] \oplus \mathfrak{q}$ is an ideal $\mathfrak{h}$ in $\mathfrak{g}_\mathbb{R}$. Since $\mathfrak{g}_\mathbb{R}$ is semi-simple, we have $\mathfrak{g}_\mathbb{R} = \mathfrak{h}^\perp \oplus \mathfrak{h}$, whereas $\mathfrak{h}^\perp$ is contained in $\mathfrak{k}$. This leads to the fact that $\exp(\mathfrak{h}^\perp)$ is a compact normal subgroup of G. Since G has no compact factor this implies that $\mathfrak{h}^\perp = \{0\}$, and hence $\mathfrak{g}_\mathbb{R} = [\mathfrak{q}, \mathfrak{q}] \oplus \mathfrak{q}$. Then we deduce that the closure of the group generated by $\exp(\mathfrak{q})$ is dense in G. □

Denote by G the connected component of the identity of $\mathbb{G}_\mathbb{R}$. Combining the above facts, we deduce the following.

COROLLARY 8.10. *The exists a faithful irreducible representation of G inside some $\mathrm{SL}(d,\mathbb{R})$ and a measure $\mu \in \mathrm{Prob}(G)$ satisfying the condition* **ISC**. *Moreover this representation comes from a $\mathbb{Q}$-representation of $\mathbb{G}$, so $\Gamma := G_{\mathbb{Z}}$ is identified (up to commensurability) with $G \cap \mathrm{SL}(d,\mathbb{Z})$.*

PROOF. We know that $\mathbb{G}$ admits a faithful $\mathbb{Q}$-representation which is irreducible over $\mathbb{R}$. This representation gives an embedding $G \subset \mathrm{GL}(d,\mathbb{R})$. Since G is semi-simple and connected, we actually have $G \subset \mathrm{SL}(d,\mathbb{R})$. Denote by $\langle \cdot, \cdot \rangle$ an inner product as in the previous lemma. Then we know that $G \cap S^+$ generates a dense subgroup of G. Write $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{q}$ for the Cartan decomposition of $\mathfrak{g}$ corresponding to the Cartan involution $X \mapsto -X^t$. Take a compact neighborhood U of 0 in $\mathfrak{q}$ such that $U = -U$. Then we know that $\exp(U)$ generates a dense subgroup of G and it is a compact symmetric set contained inside S^+. Take for μ_0 the restriction of the Lebesgue measure on $\mathfrak{q}$ to U, normalized to be a probability measure, and denote by $\mu \in \mathrm{Prob}(G)$ the push-forward of μ_0 with respect to the exponential map. This measure satisfies condition **ISC**. □

In this situation, we know by Proposition 7.15 that the quotient G/Γ embeds into the space of lattices X^+ in a proper way. In fact, since $G \subset \mathrm{SL}(d,\mathbb{R})$, we even have that G/Γ is contained inside X. The G-invariant measure ν on G/Γ gives a G-invariant measure $\widetilde{\nu}$ on X. Since ν was a Radon measure and the embedding into X is proper, $\widetilde{\nu}$ is a radon measure on X. We need to show that $\widetilde{\nu}$ is a finite measure. For that we will prove that the action of G on X satisfies property **CA**. Note that this has nothing to do with Γ anymore, it is only a property of G.

So in order to conclude the proof of Theorem 7.12, we only need to find a proper function on X which is contracted by the averaging operator A_μ corresponding to the above μ, in the sense of condition **CA**. This is achieved in the next section.

3.2. Second step: finding a nice proper function. Let us recall some notation from a previous section. If Δ denotes a lattice in $\mathbb{R}^d$ then we say that a subspace $\subset \mathbb{R}^d$ is Δ-rational if $\delta \cap L$ is a lattice in L. In this case, we denote by $d_\Delta(L)$ the covolume of $\Delta \cap L$ in L, with respect to the Euclidean structure on L inherited from the one on $\mathbb{R}^d$.

For each $0 < i < d$, denote by $\alpha_i : X \to \mathbb{R}$ the function defined by

$$\alpha_i(\Delta) := \sup\{d_\Delta(L)^{-1} \mid L \subset \mathbb{R}^d, \Delta - \text{rational}, \dim(L) = i\}.$$

By Lemma 7.21 we know that each α_i is continuous on X. Set moreover $\alpha_0 := 1$. Note that the function α_1 coincides the function $\Delta \mapsto \min_{v \in \Delta \setminus \{0\}} \|v\|^{-1}$ use in Mahler criterion. By Mahler criterion, we thus know that α_1 is proper on X.

LEMMA 8.11. *Let $\mu \in \mathrm{Prob}(\mathrm{GL}(d,\mathbb{R}))$ be a probability measure satisfying condition* **ISC**. *Then there exists $\delta > 0$, $a_0 < 1$ and $b_0 > 0$ such that*

$$A_\mu(\alpha_i^\delta) \leq a_0 \alpha_i^\delta + b_0 \max_{0 < j \leq \min(i, d-i)} (\alpha_{i-j}^\delta \alpha_{i+j}^\delta)^{1/2}, \text{ for all } 0 < i < d.$$

PROOF. Let δ and a_0 given by Corollary 8.5. For $\Delta \in X$, recall that $A_\mu(\alpha_i^\delta)(\Delta) = \int_G \alpha_i^\delta(g\Delta) d\mu(g)$. We want to find an upper bound for this integral.

Denote by $r := \sup\{\max(\|g\|^d, \|g^{-1}\|^d) \mid g \in \mathrm{supp}(\mu)\}$. So if L is any Δ-rational subspace of $\mathbb{R}^d$, we have, for μ-almost every $g \in G$,

$$r^{-1} d_\Delta(L) \leq d_{g\Delta}(gL) \leq r d_\Delta(L). \tag{3.1}$$

Take a Δ-rational space $L_i \subset \mathbb{R}^d$ with dimension i such that $\alpha_i(\Delta) = d_\Delta(L_i)^{-1}$. It exists by Lemma 7.21. Consider the finite set

$$\Psi_i := \{L \subset \mathbb{R}^d \mid \Delta - \text{rational}, \dim(L) = i, d_\Delta(L) \leq r^2 d_\Delta(L_i)\}.$$

Two cases may occur:

Case 1. L_i is the only element of Ψ_i.

In this case, for every Δ-rational subspace $L \subset \mathbb{R}^d$ with dimension i, and μ-almost every $g \in G$, we have

$$d_{g\Delta}(gL) \geq d_{g\Delta}(gL_i).$$

Indeed, it is obvious if $L = L_i$ and otherwise we use inequality (3.1) and the fact that $L \notin \Psi_i$. It then follows that $\alpha_i(g\Delta) = (d_{g\Delta}(gL_i))^{-1}$. We thus have the following equality

$$A_\mu(\alpha_i^\delta)(\Delta) = \int_G \frac{1}{d_{g\Delta}(gL_i)^\delta} d\mu(g) = \int_G \varphi_i^\delta(gw) d\mu(g),$$

where $w = e_1 \wedge \cdots \wedge e_i \in W_i$ and $e_1, \ldots, e_i$ is a basis of $L_i \cap \Delta$. By Corollary 8.5, we deduce that $A_\mu(\alpha_i^\delta)(\Delta) \leq a_0 \alpha_i^\delta(\Delta)$.

Case 2. There exists another element L_i' in Ψ_i.

In this case, put $j := \dim(L + i + L_i') - i > 0$. By Lemma 7.20, we have for μ-almost every $g \in G$,

$$\begin{aligned}
\alpha_i(g\Delta) &\leq r\alpha_i(\Delta) = r/d_\Delta(L_i) \\
&\leq r^2 (d_\Delta(L_i) d_\Delta(L_i'))^{-1/2} \\
&\leq r^2 (d_\Delta(L_i \cap L_i') d_\Delta(L_i + L_i'))^{-1/2} \\
&\leq r^2 (\alpha_{i-j}(\Delta)\alpha_{i+j}(\Delta))^{1/2}.
\end{aligned}$$

So if we set $b_0 := r^{2\delta}$, we get

$$A_\mu(\alpha_i^\delta)(\Delta) \leq b_0 \max_{0<j\leq \min(i,d-i)} (\alpha_{i-j}^\delta \alpha_{i+j}^\delta)^{1/2}.$$

In both cases, we get the desired estimate. □

Corollary 8.12. *If $\mu \in \mathrm{Prob}(\mathrm{GL}(d,\mathbb{R}))$ satisfis condition* **ISC** *then there exists $\delta > 0$ and $\varepsilon > 0$ such that the function $f : X \to [0, \infty[$ defined as follows satisfies condition* **CA***:*

$$f := \sum_{0<i<d} \varepsilon^{(d-i)i} \alpha_i^\delta.$$

Proof. For any choice of $\varepsilon > 0$ and $\delta > 0$, a function f as above is proper, because we observe that Mahler criterion shows that α_1 is proper. For each i, set $\beta_i := \varepsilon^{(d-i)i}\alpha_i^\delta$, so that $f = \sum_i \beta_i$. By the previous lemma, we have

$$A_\mu(f) \leq a_0 \sum_{0<i<d} \beta_i + b_0 \sum_{0<i<d} \varepsilon^{(d-i)i} \max_{0<j\leq \min(i,d-i)} (\alpha_{i-j}^\delta \alpha_{i+j}^\delta)^{1/2}.$$

Observe now that we have

$$2(d-i)i = (d-i-j)(i+j) + (d-i+j)(i-j) + 2j^2.$$

Combining this with the arithmetic-geomoetric inequality, we get, assuming that $\varepsilon < 1$,

$$\begin{aligned} A_\mu(f) &\leq a_0 f + b_0 \sum_i \max_{0<j\leq \min(i,d-i)} \varepsilon^{j^2}(\beta_{i-j} + \beta_{i+j})/2 \\ &\leq a_0 f + b_0 \varepsilon \sum_i \max_{j=1,\dots,d} \beta_j \\ &\leq (a_0 + b_0 \varepsilon d) f + b_0 \varepsilon d. \end{aligned}$$

This shows that $A_0(f) \leq af + b$ with $a = a_0 + b_0\varepsilon d$ and $b = b_0 \varepsilon d > 0$. If ε is small enough, we indeed get that $a < 1$, concluding the proof. □

Bibliography

[Ben08] Yves Benoist, *Réseaux des groupes de lie*, Cours de Master, Orsay, 2007-2008.

[Bor91] Armand Borel, *Linear algebraic groups. 2nd enlarged ed.*, 2nd enlarged ed. ed., vol. 126, New York etc.: Springer-Verlag, 1991.

[Hal15] Brian Hall, *Lie groups, Lie algebras, and representations. An elementary introduction. 2nd ed.*, 2nd ed. ed., vol. 222, Cham: Springer, 2015 (English).

[Hel78] Sigurdur Helgason, *Differential geometry, Lie groups, and symmetric spaces.*, Pure and Applied Mathematics, 80. New York-San Francisco-London: Academic Press. XV, 628 p. $ 27.00 (1978)., 1978.

[Hum81] James E. Humphreys, *Linear algebraic groups. Corr. 2nd printing.*, vol. 21, Springer, New York, NY, 1981 (English).