Chapter 3
Rationality

In this chapter we generalize some of the phenomena hinted at in Section 1.2. We will define the notion of algebraic groups defined over $\mathbb{Q}$, and show how these often give rise to closed (and sometimes even compact) orbits on $\mathrm{SL}_d(\mathbb{Z})\backslash \mathrm{SL}_d(\mathbb{R})$. We motivate this discussion by studying orthogonal groups, unipotent groups, and orbits arising from number fields. Finally, we will turn this discussion around by proving the Borel density theorem, which implies that finite volume orbits typically arise from algebraic groups defined over $\mathbb{Q}$. For this we also introduce some more basic concepts and results concerning algebraic groups without developing this important theory very far (which cannot be done in a couple of pages).

3.1 Quadratic Forms, Stabilizer Subgroups, and Orbits

3.1.1 Orthogonal Groups

Let

$$Q(u_1,\dots,u_d)=(u_1,\dots,u_d)A_Q(u_1,\dots,u_d)^{\mathrm{t}}$$

be a rational quadratic form defined by a symmetric matrix $A_Q\in \mathrm{Mat}_d(\mathbb{Q})$. We show now how any such quadratic form gives rise to a closed orbit of its associated special orthogonal subgroup

$$\mathrm{SO}(Q)=\{g\in \mathrm{SL}_d \mid Q\left((u_1,\dots,u_d)g\right)=Q(u_1,\dots,u_d) \text{ for all } u_1,\dots,u_d\}. \tag{3.1}$$

Proposition 3.1 (Closed orbits). *If Q is a rational quadratic form, then the orbit*

$$(\mathrm{SL}_d(\mathbb{Z})I_d)\,\mathrm{SO}(Q)(\mathbb{R})\subseteq \mathsf{X}_d$$

of the identity coset under the real points of $\mathrm{SO}(Q)$ is closed.

Notice that the notation $\mathrm{SO}(Q)$ and SL_d used in (3.1) deliberately does not specify any field or ring, and therefore leaves somewhat undetermined the group being discussed; in particular does not specify whether the group is countable or uncountable, for example. For now we should think of this as a convenient shorthand, or a macro, which defines many different groups at once. For example, if we specify the real points, then the notation denotes the closed linear subgroup of $\mathrm{SL}_d(\mathbb{R})$ defined by

$$\begin{aligned}\mathrm{SO}(Q)(\mathbb{R}) = \{g \in \mathrm{SL}_d(\mathbb{R}) \mid Q\left((u_1,\dots,u_d)g\right) = Q(u_1,\dots,u_d)&\\ \text{for all } u_1,\dots,u_d\}.&\end{aligned}$$

Similarly, we may specify the integer points to obtain a discrete subgroup

$$\begin{aligned}\mathrm{SO}(Q)(\mathbb{Z}) = \{g \in \mathrm{SL}_d(\mathbb{Z}) \mid Q\left((u_1,\dots,u_d)g\right) = Q(u_1,\dots,u_d)&\\ \text{for all } u_1,\dots,u_d\}.&\end{aligned}$$

of $\mathrm{SO}(Q)(\mathbb{R})$. More generally, for any ring R we obtain the group $\mathrm{SO}(Q)(R)$ of R-points of $\mathrm{SO}(Q)$ (or any similar expression) by taking the R-points of the ambient group, here SL_d, in its definition.

PROOF OF PROPOSITION 3.1. Notice that $Q\left((u_1,\dots,u_d)g\right)$ is the quadratic form defined by gA_Qg^{t} and that the symmetric matrix A_Q is in a one-to-one correspondence to Q. Therefore, we may also write

$$\mathrm{SO}(Q) = \{g \in \mathrm{SL}_d \mid gA_Qg^{\mathrm{t}} = A_Q\}.$$

Multiplying A_Q by its common denominator if necessary, we may assume that $A_Q \in \mathrm{Mat}_d(\mathbb{Z})$ (without changing $\mathrm{SO}(Q)$). Now suppose that

$$\mathrm{SL}_d(\mathbb{Z})h_n \longrightarrow \mathrm{SL}_d(\mathbb{Z})g = x \tag{3.2}$$

as $n \to \infty$ with $h_n \in \mathrm{SO}(Q)(\mathbb{R})$ and $g \in \mathrm{SL}_d(\mathbb{R})$. In order to show that the orbit is closed, we need to show that

$$x \in \mathrm{SL}_d(\mathbb{Z})I_d\,\mathrm{SO}(Q)(\mathbb{R}). \tag{3.3}$$

Notice that (3.2) simply means that there exist sequences (γ_n) in $\mathrm{SL}_d(\mathbb{Z})$ and (ε_n) in $\mathrm{SL}_d(\mathbb{R})$ with $\varepsilon_n \to I_d$ as $n \to \infty$, such that $\gamma_n h_n = g\varepsilon_n$ for all $n \geqslant 1$. Applying these matrices to A_Q gives

$$\gamma_n A_Q \gamma_n^{\mathrm{t}} = \gamma_n h_n A_Q (\gamma_n h_n)^{\mathrm{t}} = g\varepsilon_n A_Q (g\varepsilon_n)^{\mathrm{t}} \longrightarrow gA_Qg^{\mathrm{t}}$$

as $n \to \infty$.

However, $\gamma_n A_Q \gamma_n^{\mathrm{t}} \in \mathrm{Mat}_d(\mathbb{Z})$, so the convergent sequence $(\gamma_n A_Q \gamma_n^{\mathrm{t}})$ has to stabilize: there exists some n_0 such that

$$\gamma_{n_0} A_Q \gamma_{n_0}^{\mathrm{t}} = \gamma_n A_Q \gamma_n^{\mathrm{t}} = gA_Qg^{\mathrm{t}}$$

for all $n \geqslant n_0$. This implies that $\gamma_{n_0}^{-1} g \in \mathrm{SO}(Q)(\mathbb{R})$, giving (3.3). □

In some cases it is also relatively straightforward to combine the previous statement with Mahler's compactness criterion (Theorem 1.17) and so deduce compactness of orbits.

Proposition 3.2 (Compact orbits). *If Q is a rational quadratic form such that*[†]

$$0 \notin Q(\mathbb{Q}^d \smallsetminus \{0\}),$$

then the orbit $(\mathrm{SL}_d(\mathbb{Z}) I_d)\, SO(Q)(\mathbb{R})$ is compact. Equivalently,

$$\mathrm{SO}(Q)(\mathbb{Z}) = \{g \in \mathrm{SL}_d(\mathbb{Z}) \mid g A_Q g^{\mathrm{t}} = A_Q\}$$

is a uniform lattice in $\mathrm{SO}(Q)(\mathbb{R})$.

Proof. Just as in the proof of Proposition 3.1, we may assume that A_Q lies in $\mathrm{Mat}_d(\mathbb{Z})$. We need to show that there exists some $\delta > 0$ such that

$$\mathrm{SL}_d(\mathbb{Z}) I_d \,\mathrm{SO}(Q)(\mathbb{R}) \subseteq \mathsf{X}_d(\delta). \tag{3.4}$$

Then Theorem 1.17 and Proposition 3.1 together show that the orbit is compact.

As $Q : \mathbb{R}^d \to \mathbb{R}$ is continuous, there exists some $\delta > 0$ such that $\|x\| < \delta$ implies that $|Q(x)| < 1$. Now suppose that (3.4) does not hold for δ. Then there exists some $h \in \mathrm{SO}(Q)(\mathbb{R})$ such that $\mathbb{Z}^d h$ contains a non-zero δ-short vector mh with $m \in \mathbb{Z}^d$. However, this shows that

$$|Q(m)| = |Q(mh)| < 1 \tag{3.5}$$

which implies that $Q(m) = 0$ since $A_Q \in \mathrm{Mat}_d(\mathbb{Z})$, contradicting our assumption and completing the proof. □

Example 3.3. These examples describe some of the possibilities that may arise in low dimensions.

(1) If $Q_1(u_1, u_2) = u_1 u_2$, then Proposition 3.1 shows that $\mathrm{SL}_2(\mathbb{Z})A$ is closed since

$$\mathrm{SO}(Q_1)(\mathbb{R}) = A$$

is simply the diagonal subgroup of $\mathrm{SL}_2(\mathbb{R})$ (see Section 1.2.2). However, the orbit is not compact, it is the divergent orbit mentioned on page 25.

(2) If $Q_2(u_1, u_2) = u_1^2 - u_1 u_2 - u_2^2$, then Proposition 3.2 applies (see Exercise 3.1.1), and gives a compact orbit $\mathrm{SL}_2(\mathbb{Z})\,\mathrm{SO}(Q_2)(\mathbb{R})$. As we will see later (in Theorem 3.5), there exists some $g \in \mathrm{SL}_2(\mathbb{R})$ for which

$$Q_2(u_1, u_2) = Q_1\left((u_1, u_2) g\right),$$

[†] Q is then called anisotropic over $\mathbb{Q}$.

which in turn implies that

$$\mathrm{SO}(Q_2)(\mathbb{R}) = g\,\mathrm{SO}(Q_1)(\mathbb{R})g^{-1}.$$

To see this notice that $h \in \mathrm{SO}(Q_1)(\mathbb{R})$ and $u = (u_1, u_2)$ gives

$$Q_2(ughg^{-1}) = Q_1((ughg^{-1})g) = Q_1(ug) = Q_2(u).$$

Hence

$$\mathrm{SL}_2(\mathbb{Z})\,\mathrm{SO}(Q_2)(\mathbb{R})g = \mathrm{SL}_2(\mathbb{Z})gA$$

is also compact. In fact $g = g_{\text{golden}}$ from Section 1.2.2 can be used, recovering the claim made on page 25.

(3) If $Q_3(u_1, u_2, u_3) = 2u_1u_3 - u_2^2$ then Proposition 3.1 applies, and shows that

$$\mathrm{SL}_3(\mathbb{Z})\,\mathrm{SO}(Q_3)(\mathbb{R}) \subseteq \mathsf{X}_3$$

is closed. However, it is not compact (see Exercise 3.1.2).

(4) If $Q_4(u_1, u_2, u_3) = u_1^2 + u_2^2 - 3u_3^2$ then Proposition 3.2 applies. To see this, assume for the purposes of a contradiction (and without loss of generality by clearing denominators as usual) that $Q_4(m_1, m_2, m_3) = 0$ for some primitive[†] integer vector $(m_1, m_2, m_3) \in \mathbb{Z}^3$. Then using congruences modulo 4 shows that

$$m_1^2 + m_2^2 - 3m_3^2 \equiv m_1^2 + m_2^2 + m_3^2 \pmod 4,$$

is a sum of three squares modulo 4. However, the only squares modulo 4 are 0 and 1, which forces m_1, m_2, m_3 to all be even, contradicting the assumption. Hence the orbit

$$\mathrm{SL}_3(\mathbb{Z})\,\mathrm{SO}(Q_4)(\mathbb{R})$$

is compact.

We now recall some of the basic theory of quadratic forms over the reals$^{(10)}$. Any symmetric matrix $A \in \mathrm{Mat}_d(\mathbb{R})$ can be diagonalized in the sense that there is an orthogonal matrix k for which kAk^{t} is diagonal. In the associated coordinate system $(v_1, \dots, v_d)$ we then have

$$Q'(v_1, \dots, v_d) = Q\left((v_1, \dots, v_d)k\right) = \sum_{i=1}^{d} c_i v_i^2.$$

The form Q is *non-degenerate* if $c_i \neq 0$ for $i = 1, \dots, d$ (equivalently, if $\det A_Q \neq 0$), is *indefinite* if there exist i, j with $c_i > 0$ and $c_j < 0$, and is *positive-definite* if $c_i > 0$ for all $i = 1, \dots, d$.

[†] An integer vector is *primitive* if the entries are co-prime.

Taking the square roots of the absolute values of the entries in the diagonal matrix kA_Qk^{t}, we may define a diagonal matrix a for which

$$a^{-1}kA_Qk^{\mathrm{t}}a^{-1}$$

is diagonal with entries in $\{0, \pm 1\}$. Assuming that Q is non-degenerate (so that the entries lie in $\{\pm 1\}$), write p for the number of $+1$s and q for the number of -1s; the *signature*$^{(11)}$ of Q is (p, q). We usually assume that $p \geqslant q$ (this can always be achieved by replacing the form Q with the form $-Q$).

The discussion above shows that if Q and Q' are non-degenerate and of the same signature, then there exists some $g \in \mathrm{GL}_d(\mathbb{R})$ such that

$$Q'(u_1, \dots, u_d) = Q\left((u_1, \dots, u_d)g\right).$$

Moreover, we also have

$$Q'(u_1, \dots, u_d) = \lambda Q\left((u_1, \dots, u_d)g'\right)$$

for $g' \in \mathrm{SL}_d(\mathbb{R})$ and $\lambda \neq 0$, which implies that $\mathrm{SO}(Q)$ and $\mathrm{SO}(Q')$ are conjugate in $\mathrm{SL}_d(\mathbb{R})$.

Example 3.4. The quadratic forms (from Example 3.3) Q_1 and Q_2 have signature $(1, 1)$; the quadratic forms Q_3 and Q_4 have signature $(2, 1)$. It follows that the orthogonal groups $\mathrm{SO}(Q_1)(\mathbb{R})$ and $\mathrm{SO}(Q_2)(\mathbb{R})$ are conjugate (as claimed earlier), and the orthogonal groups $\mathrm{SO}(Q_3)(\mathbb{R})$ and $\mathrm{SO}(Q_4)(\mathbb{R})$ are conjugate.

We summarize and strengthen our discussion as follows.

Theorem 3.5 (Signature of quadratic forms). *Any non-degenerate quadratic form Q on $\mathbb{R}^d$ can be assigned a signature (p, q) with $p + q = d$. Given a form Q of signature (p, q), the set of quadratic forms of the form Q' with*

$$Q'(u_1, \dots, u_d) = Q\left((u_1, \dots, u_d)g\right)$$

obtained from Q by some $g \in \mathrm{GL}_d(\mathbb{R})$, is precisely the set of quadratic forms of signature (p, q). The group of $\mathbb{R}$-points of two orthogonal groups for non-degenerate quadratic forms of the same signature are conjugate in $\mathrm{SL}_d(\mathbb{R})$.

In the following we will always (and sometimes implicitly) assume that the quadratic forms are non-degenerate. Fixing, for a given signature (p, q), some real quadratic form Q of this signature, we define $\mathrm{SO}(p, q) = \mathrm{SO}(Q)$. If $p = d$, then

$$\mathrm{SO}(p, q) = \mathrm{SO}(d)$$

is compact, and if $0 < p < d$ it is not†. Our discussion above (and Example 3.3(3),(4)), shows that there are various matrices $g \in \mathrm{SL}_d(\mathbb{R})$ for which

---

† Since, for example, it contains at least one copy of $\mathrm{SO}(1, 1) \cong A$ as a closed subgroup.

$$\mathrm{SL}_d(\mathbb{Z})g\,\mathrm{SO}(p,q)(\mathbb{R})$$

is closed or even compact – these orbits correspond[†] to rational quadratic forms with signature (p,q).

3.1.2 Rational Stabilizer Subgroups

It is straightforward to generalize Proposition 3.1. However, setting up the language of linear groups, in which the generalization is naturally phrased, requires more work than does the generalization itself. We start this introduction to linear algebraic groups here, discuss other classes of examples in Sections 3.2 and 3.3, and return to the theory of linear algebraic groups in Section 3.4 and Chapter 7. For a detailed account of algebraic geometry, we refer to the monographs of Hartshorne [75] or Shafarevich [168], and for linear algebraic groups we refer to those of Borel [8], Humphreys [80], and Springer [176].

An *affine variety* is a subset Z of $\mathbb{C}^n$ or, more generally, of $\overline{\mathbb{K}}^n$ for another field $\mathbb{K}$ with $\overline{\mathbb{K}}$ an algebraic closure, defined by the vanishing of polynomial equations[‡]. We will write both Z and $Z(\overline{\mathbb{K}})$ for this variety, so that

$$Z = Z(\overline{\mathbb{K}})$$

will always consist of all solutions to the polynomial equations over the algebraic closure. An important example for us is

$$\mathrm{SL}_d = \{g \in \mathrm{Mat}_d \mid \det g - 1 = 0\},$$

where Mat_d is the d^2-dimensional vector space of $d \times d$ matrices.

A *regular function* is simply the restriction of a polynomial to the variety[§]. In order to be able to work with this definition, and in particular to have a way to uniquely describe a regular function, we need to know when a polynomial vanishes on the variety. The description of the set of polynomials that vanish on an affine variety is given by the Hilbert Nullstellensatz,(12) which we now recall. We refer to Eisenbud [56, Th. 1.6] or Hungerford [81, Prop. VIII 7.4] for the proof.

[†] At this stage we only know one direction of this correspondence. The second direction will be obtained from the Borel density theorem, see Exercise 3.5.1 and Exercise 4.4.1.

[‡] We apologize to the expert for this barbaric and old-fashioned definition, but as our focus will usually be on rather concrete groups comprising $\mathbb{R}$-points, this approach is appropriate here. We will on occasion (indeed, are just about to) avoid mentioning the field we are working over, but we still wish to avoid talking about schemes, spectrum, and using the language of modern algebraic geometry.

[§] Once again we must apologize for avoiding a more general definition, our excuse being that this is adequate for affine varieties.

Theorem 3.6 (Hilbert Nullstellensatz). *Let $\mathbb{K}$ be an algebraically closed field, and let $\mathcal{J} \subseteq \mathbb{K}[x_1, \dots, x_n]$ be an ideal defining the affine variety*

$$Z(\mathcal{J}) = \{x \in \mathbb{K}^n \mid f(x) = 0 \textit{ for all } f \in \mathcal{J}\}.$$

Then $f \in \mathbb{K}[x_1, \dots, x_n]$ vanishes on $Z(\mathcal{J})$ if and only if there exists a power f^m, $m \geqslant 1$, of f that belongs to $\mathcal{J}$.

The ideal

$$\operatorname{rad}(\mathcal{J}) = \{f \in k[x_1, \dots, x_n] \mid f^m \in \mathcal{J} \text{ for some } m \geqslant 1\}$$

is called the *radical* of the ideal $\mathcal{J}$. If we now write $\mathbb{K}[Z]$ for the ring of regular functions on the variety $Z = Z(\mathcal{J})$ defined by the ideal $\mathcal{J}$, then we can reformulate the Nullstellensatz by the formula

$$\mathbb{K}[Z(\mathcal{J})] = \mathbb{K}[x_1, \dots, x_n]/\operatorname{rad}(\mathcal{J}).$$

Returning to our example

$$\mathrm{SL}_d = Z(\det(\cdot) - 1) \subseteq \mathrm{Mat}_d,$$

we need to establish what the radical of the ideal generated by the polynomial $\det(\cdot) - 1$ in d^2 variables is in order to talk about regular functions. This is explained by the following result.

Lemma 3.7 (SL_d **is Zariski connected**). *For any $d \geqslant 1$ the polynomial $\det(g) - 1$ is irreducible as a polynomial in the variables g_{ij}, $1 \leqslant i, j \leqslant d$, with coefficients in $\mathbb{C}$ (or in any other field).*

Before proving this, we note that additional background in the theory of algebraic groups would make it almost immediate by the following argument. Notice first that the group $\mathrm{SL}_d(\mathbb{C})$ of complex solutions to $\det(g) = 1$ is connected (this may be seen, for example, as a consequence of Lemma 1.24), and that every point of this variety is smooth†. If now $\det(g) - 1 = p_1^{\ell_1}(g) p_2^{\ell_2}(g) \cdots p_\ell^{\ell_k}(g)$ is the decomposition into irreducible polynomials with multiplicities ℓ_i, then the group $\mathrm{SL}_d(\mathbb{C})$ would be the union of the varieties defined by $p_1, \dots, p_k$. By connectedness of $\mathrm{SL}_d(\mathbb{C})$ these varieties would have to intersect, but this contradicts smoothness of the variety SL_d at the intersection points. The only possibility that remains is that $\det(g) - 1 = p^\ell(g)$ for some irreducible polynomial p and $\ell > 1$. This is clearly impossible by degree considerations (see the first part of the more elementary proof below).

PROOF OF LEMMA 3.7. Suppose that $\det(g) - 1 = p(g)q(g)$, where p, q are polynomials in the independent variables g_{ij}, $1 \leqslant i, j \leqslant d$. Now notice that

---

† This is a general property of algebraic groups: Every variety has a smooth point, and as usual for a group any two points in the group have similar properties as the group acts transitively, see Section 3.4.

the determinant is linear in each of its rows, so for every pair i, j the polynomial $\det(g) - 1$ is of degree one in the variable g_{ij}. It follows that for any i, j either p or q is independent of g_{ij} (that is, of degree zero in the variable g_{ij}). As this holds for every pair i, j, we obtain a partition

$$P \sqcup Q = \{(i, j) \mid 1 \leqslant i, j \leqslant d\}$$

of the indices so that

$$p(g) \in \mathbb{C}[g_{ij} \mid (i, j) \in P]$$

and

$$q(g) \in \mathbb{C}[g_{ij} \mid (i, j) \in Q].$$

If P (or Q) is empty, then $p \in \mathbb{C}$ (respectively $q \in \mathbb{C}$) is a scalar — which is the desired conclusion.

With deg denoting the total degree,

$$d = \deg(\det(g) - 1) = \deg(p(g)q(g)) = \deg(p(g)) + \deg(q(g)). \tag{3.6}$$

Assuming that P and Q are both non-empty, we derive a contradiction by defining

$$\deg_P(g_{ij}) = \begin{cases} 1 & \text{if } (i, j) \in P; \\ -1 & \text{if } (i, j) \in Q, \end{cases}$$

which extends to monomials m in the usual way, and to polynomials by defining

$$\deg_P\left(\sum c_k m_k\right) = \max\{\deg_P(m_k) \mid c_k \neq 0\}.$$

Just as in (3.6), we find that

$$\deg_P(pq) = \deg_P(p) + \deg_P(q).$$

Now q must have a constant term (since $\det(g) - 1$ has a constant term), so $\deg_P(q) = 0$. It follows that $p(g)q(g)$ contains monomials in the variables g_{ij} with $(i, j) \in P$ of total degree $\deg_P(p) = \deg(p)$ only. However, this is a contradiction as $\det(g) - 1$ contains a constant term, and all other monomials have total degree d. □

Let $\mathbb{K}$ be any field. We will often be interested not in the whole variety consisting of all points in $\overline{\mathbb{K}}^n$ defined by an ideal over the algebraic closure of a field, but in fact only in the $\mathbb{K}$-*points* of the variety, meaning those vectors in $\mathbb{K}^n$ on which the polynomials all vanish. In general this set may be empty because $\mathbb{K}$ is not assumed to be algebraically closed, and even if it is non-empty it may not resemble the whole variety. In particular, there is no reason for the set of $\mathbb{K}$-points to remember the ideal at all (in other words, Theorem 3.6 does not hold without the requirement that the field be algebraically closed). Nonetheless, we may define for any affine variety Z its $\mathbb{K}$-points as the set

$$Z(\mathbb{K}) = Z \cap \mathbb{K}^n,$$

where as before $Z = Z(\overline{\mathbb{K}})$ by definition. Moreover, we are often interested in regular functions with 'coefficients' in $\mathbb{K}$, which we formally define as the *ring of* $\mathbb{K}$*-regular functions*

$$\mathbb{K}[Z] = \mathbb{K}[x_1, \dots, x_n]/\mathcal{J} \cap \mathbb{K}[x_1, \dots, x_n],$$

under the assumption that Z is *defined over* $\mathbb{K}$, meaning that $\mathcal{J} = \operatorname{rad}(\mathcal{J})$ defines Z and $\mathcal{J} \cap \mathbb{K}[x_1, \dots, x_n]$ generates $\mathcal{J} \subseteq \overline{\mathbb{K}}[x_1, \dots, x_n]$. We will return to these notions in Section 3.4.

Let us return to our main example SL_d which is defined over any field $\mathbb{K}$, since the coefficients of the polynomial $\det(\cdot) - 1$ are integers. Hence it makes sense to consider the ring of $\mathbb{K}$-regular functions

$$\mathbb{K}[\mathrm{SL}_d] = \mathbb{K}[g_{11}, \dots, g_{1d}, g_{21}, \dots, g_{2d}, \dots, g_{d1}, \dots, g_{dd}]/\langle \det(g) - 1 \rangle, \quad (3.7)$$

where $\mathbb{K}$ is the field of coefficients allowed in the polynomials. For us the field $\mathbb{K}$ will often be $\mathbb{R}$, $\mathbb{Q}_p$, or $\mathbb{Q}$.

A D-dimensional *algebraic representation of* SL_d *over* $\mathbb{K}$ is a D^2-tuple of polynomials

$$\phi_{ij}(g) \in \mathbb{K}[\mathrm{SL}_d]$$

for $1 \leqslant i, j \leqslant D$, which we think of as a matrix

$$\phi \in \mathrm{Mat}_D\left(\mathbb{K}[\mathrm{SL}_d]\right)$$

with the properties that $\phi(I_d) = I_D$ and

$$\phi(g)\phi(h) = \phi(gh) \quad (3.8)$$

for all $g, h \in \mathrm{SL}_d$. Equivalently, (3.8) could be required to hold as an abstract identity in the variables $g_{k\ell}, h_{k\ell}$ satisfying the polynomial condition

$$\det(g) = \det(h) = 1.$$

This equivalence follows from Hilbert's Nullstellensatz (Theorem 3.6) and Lemma 3.7.

An example of such a representation has been mentioned: If $A \in \mathrm{Mat}_d$ is symmetric and $g \in \mathrm{SL}_d$, then the map

$$A \longmapsto gAg^{\mathrm{t}} \quad (3.9)$$

is linear in A and polynomial in g. In fact, by identifying the space of symmetric matrices in $\mathrm{Mat}_d(\mathbb{Q})$ with the vector space $\mathbb{Q}^D$, where $D = \frac{d(d+1)}{2}$, we obtain a matrix representation $\phi(g) \in \mathrm{Mat}_D$ of (3.9) for which the matrix entries $\phi_{ij}(g)$ are polynomials of total degree 2 with coefficients in $\mathbb{Q}$. Moreover, $\phi(g)\phi(h)$ is the matrix corresponding to the composition

$$A \longmapsto hAh^{\mathrm{t}} \longmapsto g\left(hAh^{\mathrm{t}}\right)g^{\mathrm{t}} = (gh)A(gh)^{\mathrm{t}},$$

which is also represented by $\phi(gh)$. Therefore, (3.8) holds by uniqueness of matrix representations.

Let us give another representation of SL_d, which will be important in Section 3.3. The conjugation representation is defined by

$$\mathrm{Mat}_d \ni v \longmapsto gvg^{-1}$$

for $g \in \mathrm{SL}_d$. Since $\det(g) = 1$, the matrix g^{-1} has entries which are regular functions (since the inverse is calculated by taking the matrix consisting of the determinants of the minor matrices multiplied by the inverse of the determinant). Therefore, we can again chose a basis and get a $D = d^2$-dimensional representation† (defined over any field $\mathbb{K}$).

Proposition 3.8 (Rational stabilizer groups of points have closed orbits). *Let $\phi : \mathrm{SL}_d \to \mathrm{GL}_D$ be an algebraic representation over $\mathbb{Q}$, and let $v \in \mathbb{Q}^D$. Then the (rational) stabilizer subgroup*

$$\mathrm{Stab}_{\mathrm{SL}_d}(v) = \{g \in \mathrm{SL}_d \mid \phi(g)v = v\}$$

gives rise to a closed *orbit*

$$\mathrm{SL}_d(\mathbb{Z})I_d \,\mathrm{Stab}_{\mathrm{SL}_d}(v)(\mathbb{R}) \subseteq \mathsf{X}_d$$

through the identity coset.

Notice that $\mathrm{Stab}_{\mathrm{SL}_d}(v)$ is itself a subgroup defined by polynomial equations (and hence will be seen to be an algebraic subgroup defined over $\mathbb{Q}$, once we define this notion in Section 3.4). The proof of Proposition 3.8 is much quicker than the discussion above, which was included to familiarize the notion of algebraic representations of SL_d.

PROOF OF PROPOSITION 3.8. Notice that there are finitely many coefficients in (a representation of) the polynomials in $\phi(g)$. Let N be their common denominator, so that $\phi(\gamma) \in \frac{1}{N}\mathrm{Mat}_D(\mathbb{Z})$ for all $\gamma \in \mathrm{SL}_d(\mathbb{Z})$. Let M be the common denominator of the entries in v. Suppose that

$$\mathrm{SL}_d(\mathbb{Z})h_n \longrightarrow \mathrm{SL}_d(\mathbb{Z})g = x, \tag{3.10}$$

with $h_n \in \mathrm{Stab}_{\mathrm{SL}_d}(v)(\mathbb{R})$ and $g \in \mathrm{SL}_d(\mathbb{R})$. We wish to show that

$$x \in \mathrm{SL}_d(\mathbb{Z})\,\mathrm{Stab}_{\mathrm{SL}_d}(v)(\mathbb{R}). \tag{3.11}$$

As in the proof of Proposition 3.1, we may rewrite (3.10) as $\gamma_n h_n = g\varepsilon_n$ with $\gamma_n \in \mathrm{SL}_d(\mathbb{Z})$, $\varepsilon_n \in \mathrm{SL}_d(\mathbb{R})$, and $\varepsilon_n \to I_d$. Applying these matrices to v

---

† As will become more and more clear, part of the art in discussing algebraic groups and their representations will be to not really write down any concrete polynomials or regular functions (as these quickly become quite complicated).

via the representation ϕ shows that the sequence $(\phi(\gamma_n)v)$ lies in $\frac{1}{MN}\mathbb{Z}^D$ and converges,

$$\phi(\gamma_n)v = \phi(\gamma_n h_n)v = \phi(g)\phi(\varepsilon_n)v \longrightarrow \phi(g)v.$$

Therefore this sequence must stabilize, and so $\phi(\gamma_n)v = \phi(g)v$ for some n, which shows that $\gamma_n^{-1}g \in \mathrm{Stab}_{\mathrm{SL}_d}(v)(\mathbb{R})$, giving (3.11). $\square$

Although the following is not needed for the proof above, let us try to understand a little more about $\mathrm{SL}_d(\mathbb{K})$ and algebraic representations of SL_d over any field $\mathbb{K}$.

(1) As shown in Lemma 1.24, $\mathrm{SL}_d(\mathbb{K})$ is generated by the elementary unipotent subgroups

$$U_{ij}(\mathbb{K}) = \{u_{ij}(t) = I + tE_{ij} \mid t \in \mathbb{K}\}$$

with $i \neq j$ and E_{ij} being the elementary matrix with (i,j)th entry 1 and all other entries 0.

(2) $\mathrm{SL}_d(\mathbb{K})$ coincides with its commutator subgroup

$$[\mathrm{SL}_d(\mathbb{K}), \mathrm{SL}_d(\mathbb{K})] = \langle [g,h] \mid g,h \in \mathrm{SL}_d(\mathbb{K})\rangle\,,$$

where $[g,h] = g^{-1}h^{-1}gh$. To see this, notice that if we choose an appropriate diagonal matrix a then

$$[u_{ij}(t), a] = u_{ij}(\alpha t)$$

for some $\alpha \neq 0$. Hence $[\mathrm{SL}_d(\mathbb{K}), \mathrm{SL}_d(\mathbb{K})] \supseteq U_{ij}(\mathbb{K})$ for all $i \neq j$, and the result follows by the remark above.

(3) It follows that $\mathrm{SL}_d(\mathbb{K})$ (resp. $\mathrm{SL}_d(\overline{\mathbb{K}})$) cannot have any abelian factors, and so $\det\phi(g) = 1$ for every algebraic representation over $\mathbb{K}$. By Theorem 3.6 and Lemma 3.7 this must therefore also hold as an identity in

$$\mathbb{K}[\mathrm{SL}_d] = \mathbb{K}[g_{ij} : i,j = 1,\dots,d]/\langle \det g - 1\rangle.$$

Exercises for Section 3.1

Exercise 3.1.1. Prove that $u_1^2 - u_1u_2 - u_2^2 \neq 0$ for $(u_1,u_2) \in \mathbb{Q}^2\smallsetminus\{0\}$ (a fact used in Example 3.3(2)).

Exercise 3.1.2. Prove the claim made in Example 3.3(3), by showing that the closed orbit $\mathrm{SL}_3(\mathbb{Z})\,\mathrm{SO}(Q_3)(\mathbb{R}) \subseteq \mathsf{X}_3$ has unbounded height.

Exercise 3.1.3. Let $A = \mathrm{SO}(1,1)(\mathbb{R}) \subseteq \mathrm{SL}_2(\mathbb{R})$. Show that every closed A-orbit corresponds (as indicated after Theorem 3.5) to a binary quadratic form with rational coefficients. Notice that this cannot hold for $K = \mathrm{SO}(2)(\mathbb{R})$.

Exercise 3.1.4. For any subspace $V \subseteq \mathbb{R}^d$ we define

$$L_V = \{g \in G \mid V = Vg \text{ and } g|_V \text{ preserves the volume}\}.$$

(1) Show that $\mathrm{SL}_d(\mathbb{Z})L_V \subseteq \mathsf{X}_d$ is closed if V is a rational subspace.
(2) More generally, let $x_0 = \mathrm{SL}_d(\mathbb{Z})g_0$ and let V be a $\mathbb{Z}^d g_0$-rational subspace. Show that $x_0 L_V$ is closed.
(3) Let x_0 and V be as in (2). Let $G < \mathrm{SL}_d(\mathbb{R})$ be a closed subgroup such that $x_0 G$ is closed. Show that $x_0(G \cap L_V)$ is closed.

3.2 Rational Unipotent Subgroups

[†]In this section we will construct lattices in certain[‡] connected, simply connected nilpotent Lie groups. By Ado's theorem (see Ado [1] or Knapp [103, Th. B.8]) and Engel's theorem (see Knapp [103, Th. 1.35],) such a group can be embedded into the upper triangular subgroup[§]

$$N = \left\{ \begin{pmatrix} 1 & * & * & \dots & * \\ & 1 & * & \dots & * \\ & & & \ddots & \vdots \\ & & & & 1 \end{pmatrix} \right\} \subseteq \mathrm{SL}_d(\mathbb{R}) \tag{3.12}$$

for some $d > 1$. A subgroup $G < \mathrm{SL}_d(\mathbb{R})$ is called *unipotent* if it is conjugated to a subgroup of N.

Theorem 3.9 (Lattices and Mal′cev basis for unipotent $\mathbb{Q}$-groups). *Let $G \leqslant \mathrm{SL}_d(\mathbb{R})$ be a connected unipotent subgroup whose Lie algebra $\mathfrak{g}$ is a rational subspace of $\mathfrak{sl}_d(\mathbb{R})$. Then*

$$\mathbb{G}(\mathbb{Z}) = G \cap \mathrm{SL}_d(\mathbb{Z})$$

is a uniform lattice in G. Moreover, writing $\ell = \dim \mathbb{G}$, there exist elements

$$v_1, \dots, v_\ell \in \mathfrak{g} \cap \mathfrak{sl}_d(\mathbb{Q})$$

for which

$$\mathbb{G}(\mathbb{Z}) = \{\exp(k_1 v_1)\exp(k_2 v_2)\cdots\exp(k_\ell v_\ell) \mid k_1, \dots, k_\ell \in \mathbb{Z}\},$$

$$G = \mathbb{G}(\mathbb{R}) = \{\exp(s_1 v_1)\exp(s_2 v_2)\cdots\exp(s_\ell v_\ell) \mid s_1, \dots, s_\ell \in \mathbb{R}\},$$

[†] This section gives more examples of compact quotients of nilpotent groups, but otherwise is not essential for most of what follows. It will, however, become part of our proof of the Borel Harish-Chandra theorem in Section 7.4.

[‡] Once we have discussed these notions it will be easy to see that the groups we will discuss here are of the form $G = \mathbb{G}(\mathbb{R})$ for a connected unipotent algebraic group $\mathbb{G}$ defined over $\mathbb{Q}$. As the theorem and its proof does not require this language we leave this fact to the reader.

[§] Ado's and Engel's theorems are usually stated for a nilpotent Lie algebra instead of for the corresponding simply connected group, but the former implies the latter, see Exercise 3.2.1.

and

$$F = \{\exp(s_1 v_1)\exp(s_2 v_2)\cdots\exp(s_\ell v_\ell) \mid s_1, \dots, s_\ell \in [0,1)\}$$

is a fundamental domain for $\mathbb{G}(\mathbb{Z})$ *in* G*. Moreover, the map*

$$s_1, \dots, s_\ell \longmapsto \exp(s_1 v_1)\exp(s_2 v_2)\cdots\exp(s_\ell v_\ell)$$

is a (polynomial) diffeomorphism between $\mathbb{R}^\ell$ *and* G*. The vectors* $v_1, \dots, v_\ell$ *in* $\mathfrak{g}$ *are called a* Mal′cev basis*.*

Proof. As $\mathfrak{g} \subseteq \mathfrak{sl}_d(\mathbb{R})$ is, by assumption, both a nilpotent Lie algebra and a rational subspace, the same holds for all the elements of the lower central series. In particular, $\mathfrak{g}_1 = [\mathfrak{g}, \mathfrak{g}]$ is a rational subspace. By assumption, $\mathfrak{g}$ can be conjugated into the Lie algebra of N. Therefore, the exponential map

$$\exp(v) = I + v + \tfrac{1}{2}v^2 + \cdots + \tfrac{1}{(d-1)!}v^{d-1}$$

is actually a polynomial map on $\mathfrak{g}$ with the logarithm map

$$\log(g) = g - I - \tfrac{1}{2}(g-I)^2 + \cdots + (-1)^d \tfrac{1}{d-1}(g-I)^{d-1}$$

as a polynomial inverse (which is defined on all of G). From this it follows that the linear group G is isomorphic to its Lie algebra $\mathfrak{g}$, if we equip the latter with the polynomial group operation $v * w = \log(\exp(v)\exp(w))$.

Recall that there is a — possibly immersed — Lie subgroup $G_1 \lhd G$ with Lie algebra $\mathfrak{g}_1$. This shows that for sufficiently small $v, w \in \mathfrak{g}_1$ the product $v * w$ lies in $\mathfrak{g}_1$. However, using the fact that the group product $v * w$ for $v, w \in \mathfrak{g}$ is a polynomial in v and w, we can now conclude[†] that $\mathfrak{g}_1 * \mathfrak{g}_1 \subseteq \mathfrak{g}_1$. Indeed, if ψ is a linear function vanishing on $\mathfrak{g}_1$ and $v \in \mathfrak{g}_1$ is sufficiently small, then $w \mapsto \psi(v * w)$ is a polynomial on $\mathfrak{g}_1$ which vanishes on all sufficiently small w. Hence $\psi(v * w) = 0$ for all $w \in \mathfrak{g}_1$. Reversing the roles of v and w, and using the fact that a linear subspace is defined by the collection of all linear functions that vanish on it, we see that $\mathfrak{g}_1 * \mathfrak{g}_1 \subseteq \mathfrak{g}_1$. However, this shows that $G_1 = \exp(\mathfrak{g}_1)$ is simply the isomorphic image of the Lie ideal $\mathfrak{g}_1$ and so is a normal closed connected subgroup of G. Note furthermore that the Lie algebra of G/G_1 equals $\mathfrak{g}/\mathfrak{g}_1$. Hence G/G_1 is abelian and can be identified with its Lie algebra under the exponential map.

As $m = \dim(G_1) < \ell = \dim(G)$ and the Lie algebra $\mathfrak{g}_1$ of G_1 is rational, we may assume that the theorem already holds for the unipotent subgroup G_1. So let $v_1, \dots, v_m$ be the Mal′cev basis for G_1 and the uniform lattice

$$G_1(\mathbb{Z}) = G_1 \cap \mathrm{SL}_d(\mathbb{Z}).$$

[†] Once we have introduced the notion of Zariski density we will see that this argument uses the fact that the Hausdorff (that is, standard) neighborhood of $(0,0) \in \mathfrak{g}_1 \times \mathfrak{g}_1$ is Zariski dense in $\mathfrak{g}_1 \times \mathfrak{g}_1$

Let $F_1 \subseteq G_1$ be a fundamental domain as in the theorem for $G_1(\mathbb{Z})$ in G_1. Let $V \subseteq \mathfrak{g}$ be a rational linear complement to $\mathfrak{g}_1 < \mathfrak{g}$. We claim that the image of $G(\mathbb{Z})$ in the abelian group $G/G_1 \simeq \mathfrak{g}/\mathfrak{g}_1 \simeq V$ is discrete. To see this, suppose that $K \subseteq G/G_1$ is a compact neighborhood of the identity and $G_1\gamma \in K \cap (G(\mathbb{Z})/G_1) \subseteq G/G_1$. Then we may modify the representative γ by elements of $G_1(\mathbb{Z})$ on the left to ensure that $\gamma \in F_1 \exp(V)$ belongs to a fixed compact set. As $G(\mathbb{Z})$ is discrete, this shows that there are only finitely many possibilities for γG_1, and so the image of $G(\mathbb{Z})$ in G/G_1 is discrete.

Next we claim that the image of $G(\mathbb{Z})$ modulo G_1 is a lattice in V. To see this, we have to find $\ell - m = \dim V$ linearly independent elements in the image of $G(\mathbb{Z})$ in $G/G_1 \simeq V$. This follows in turn since for every rational element $v \in V$ we have

$$\exp(Nv) = 1 + Nv + \tfrac{1}{2}N^2v^2 + \cdots + \tfrac{1}{(d-1)!}N^{d-1}v^{d-1} \in G(\mathbb{Z})$$

for a sufficiently divisible N.

We now choose $v_{m+1}, \dots, v_\ell \in \mathfrak{g}$ so that

$$\exp(v_j) \in G(\mathbb{Z})$$

for $j = m+1, \dots, \ell$ and the elements

$$G_1 \exp(v_{m+1}), \dots, G_1 \exp(v_\ell)$$

are a basis of the lattice obtained from $G(\mathbb{Z})$ in G/G_1. The elements

$$v_1, \dots, v_m, v_{m+1}, \dots, v_\ell$$

are now a Mal′cev basis. To see this, let $\gamma \in G(\mathbb{Z})$. Considering γG_1 we find $k_{m+1}, \dots, k_\ell \in \mathbb{Z}$ such that $\gamma G_1 = G_1 \exp(k_{m+1}v_{m+1}) \cdots \exp(k_\ell v_\ell)$, or equivalently

$$\gamma_2 = \gamma(\exp(k_{m+1}v_{m+1}) \cdots \exp(k_\ell v_\ell))^{-1} \in G_1.$$

Applying the inductive assumption it follows that $\gamma = \exp(k_1v_1) \cdots \exp(k_\ell v_\ell)$ for some $k_1, \dots, k_\ell \in \mathbb{Z}$. If $g \in G$ is arbitrary we may argue similarly to obtain unique $s_1, \dots, s_\ell \in \mathbb{R}$ with $g = \exp(s_1v_1) \cdots \exp(s_\ell v_\ell)$. Furthermore, if we consider g as a representative of a coset $G(\mathbb{Z})g$ we may define $k_j = \lfloor s_j \rfloor$ for $j = m+1, \dots, \ell$ and multiply g on the left with the lattice element $(\exp(k_{m+1}v_{m+1}) \cdots \exp(k_\ell v_\ell))^{-1}$ to obtain

$$g' = (\exp(k_{m+1}v_{m+1}) \cdots \exp(k_\ell v_\ell))^{-1} g = g_1' \exp(s_{m+1}'v_{m+1}) \cdots \exp(s_\ell' v_\ell')$$

with $g_1' \in G$. This ensures for g' that we have $s_{m+1}', \dots, s_\ell' \in [0,1)$. Applying the inductive assumption to $g_1' \in G_1$, we deduce that the set F is indeed a fundamental domain. □

Exercises for Section 3.2

Exercise 3.2.1. In Knapp [103, Th. B.8, Th. 1.35] it is shown that any nilpotent Lie algebra can be embedded into the Lie algebra $\mathfrak{n}$ of N for some $d > 1$ (where N is defined by (3.12)). Use this (and the discussions regarding the exponential map of this chapter applied to $G = N$) to show that every connected, simply connected nilpotent Lie group can be embedded into N.

Exercise 3.2.2. Let G be a unipotent subgroup of $\mathrm{SL}_d(\mathbb{R})$ (with a rational Lie algebra). Show that G can be defined using polynomial equations (with rational coefficients).

3.3 Dirichlet's Unit Theorem and Compact Torus Orbits

[†]In this section we study another class of examples of orbits of rational stabilizer groups, which will also lead to a proof of Dirichlet's unit theorem.(13) Let

$$K = \mathbb{Q}(\zeta) \cong \mathbb{Q}[T]/\langle m(T)\rangle$$

be an algebraic number field generated by ζ, with minimal polynomial m of degree $d = [K : \mathbb{Q}] = \deg m(T)$. We may assume that $m(T)$ is monic. Let $\mathcal{O} \subseteq K$ be an *order* (a subring of K that is isomorphic to $\mathbb{Z}^d$ as a group). Replacing ζ by $n\zeta$ has the effect of multiplying the non-leading coefficients of $m(T)$ by powers of n. Thus we may assume that $m(T) \in \mathbb{Z}[T]$, so that ζ is an algebraic integer[‡], and, for example, $\mathcal{O} = \mathbb{Z}[\zeta]$ is an order. Even though K can be embedded into $\mathbb{R}$ or $\mathbb{C}$, we prefer not to think of K as a subfield of $\mathbb{C}$ but rather as an abstract field, for instance as $K = \mathbb{Q}[T]/\langle m(T)\rangle$.

Theorem 3.10 (Dirichlet unit theorem). *Let $\mathcal{O}$ be an order in an algebraic number field K. The group $\mathcal{O}^\times$ of units is isomorphic to $F \times \mathbb{Z}^{r+s-1}$, where $F \subseteq K$ is a finite group of roots of unity, r is the number of real embeddings $K \hookrightarrow \mathbb{R}$, and s is the number of pairs of complex embeddings $K \hookrightarrow \mathbb{C}$.*

The numbers r and s may be described as follows. Splitting $m(T)$ over $\mathbb{C}$ gives

$$m(T) = (T - \zeta_1)\cdots(T - \zeta_r)(T - \zeta_{r+1})(T - \overline{\zeta_{r+1}})\cdots(T - \zeta_{r+s})(T - \overline{\zeta_{r+s}}),$$

with $\zeta_1, \dots, \zeta_r \in \mathbb{R}$ and $\zeta_{r+1}, \dots, \zeta_{r+s} \in \mathbb{C}\smallsetminus\mathbb{R}$. As $\mathbb{Q}[T]/\langle m(T)\rangle \cong K$, the real embeddings $\phi_i : K \to \mathbb{R}$ are then all of the form

---

[†] This section provides interesting examples of algebraic groups (more precisely, of torus subgroups) and compact orbits, and connects these to algebraic number theory. It is not essential for most of the later chapters. It will, however, become part of our proof of the Borel Harish-Chandra theorem in Section 7.4.

[‡] An *algebraic integer* is an algebraic number for which the monic minimal polynomial has integer coefficients.

$$\phi_i(f(T)) = f(\zeta_i)$$

for some $i = 1, \dots, r$, and the complex embeddings are all of the form

$$\phi_{r+i}(f(T)) = f(\zeta_{r+i}),$$

respectively

$$\overline{\phi_{r+i}}(f(T)) = f(\overline{\zeta_{r+i}}),$$

for $i = 1, \dots, s$ and $f \in \mathbb{Q}[T]$.

Another point of view is given by studying the multiplication by T map

$$\begin{aligned} \cdot T \colon \mathbb{Q}[T]/\langle m(T)\rangle &\longrightarrow \mathbb{Q}[T]/\langle m(T)\rangle \\ f(T) + \langle m(T)\rangle &\longmapsto T f(T) + \langle m(T)\rangle, \end{aligned}$$

or, equivalently, the multiplication map $r \mapsto r\cdot\zeta$ on $K = \mathbb{Q}(\zeta)$. Considered as a linear map over $\mathbb{Q}$, the characteristic polynomial of the linear map $r \mapsto r\cdot\zeta$ is a rational polynomial which annihilates the map. It follows that $m(T)$ is the characteristic and also the minimal polynomial of the map. Therefore, the linear map $r \mapsto r\cdot\zeta$ has eigenvalues

$$\zeta_1, \dots, \zeta_r, \zeta_{r+1}, \overline{\zeta_{r+1}}, \dots, \zeta_{r+s}, \overline{\zeta_{r+s}}.$$

More generally, if $\cdot b$ is the multiplication by $b \in K = \mathbb{Q}(\zeta)$ map, then its eigenvalues (considered as a $\mathbb{Q}$-linear map on the vector space K over $\mathbb{Q}$) are again[†]

$$\phi_1(b), \dots, \phi_r(b), \phi_{r+1}(b), \overline{\phi_{r+1}(b)}, \dots, \phi_{r+s}(b), \overline{\phi_{r+s}(b)}.$$

We now discuss how to obtain a concrete matrix representation of K, which will allow us to use the results of Section 3.1. This is quite similar to how one can consider $\mathbb{C}$ as a field of 2×2 matrices using the correspondence

$$a + \mathrm{i}b \longleftrightarrow \begin{pmatrix} a & b \\ -b & a \end{pmatrix},$$

and it is helpful to view the construction below simply as an analogue of this. In order to make the construction a bit more flexible we start with another definition.

A *proper $\mathcal{O}$-ideal* $\mathcal{J} \subseteq \mathcal{O}$ is an additive subgroup isomorphic to $\mathbb{Z}^d$ and for which

$$\mathcal{O} = \{b \in K \mid b\mathcal{J} \subseteq \mathcal{J}\}.$$

[†] This follows since $b = f(\xi)$ for some polynomial $f(T)$. If $b \in K\smallsetminus\mathbb{Q}$ then none of the eigenvectors are in $\mathbb{Q}$. In that case the eigenvectors only appear after 'extending the scalars', for example replacing $K \cong \mathbb{Q}^d$ by $K \otimes \mathbb{C} \cong \mathbb{C}^d$.

Clearly $\mathcal{J} = \mathcal{O}$ is a proper ideal.[†] For Theorem 3.10 $\mathcal{J} = \mathcal{O}$ would suffice, but it is of independent interest to note that any proper ideal can be used as the basis of the following construction.

So let $a_1, \dots, a_d$ be a $\mathbb{Z}$-basis of a proper $\mathcal{O}$-ideal $\mathcal{J}$. With this basis in mind, we may now identify the linear map $\cdot b$ on K with a matrix

$$\psi(b) \in \mathrm{Mat}_d(\mathbb{Q}).$$

We are again using row vectors so that $\cdot b\colon K \to K$ corresponds to multiplying row vectors $v \in \mathbb{Q}^n$ on the right by $\psi(b)$. By assumption, for $b \in K$ we have

$$b \in \mathcal{O} \iff (\cdot(b))\,(a_i) \in \mathcal{J} \text{ for all } i \iff \psi(b) \in \mathrm{Mat}_d(\mathbb{Z}),$$

and so also

$$b \in \mathcal{O}^\times \iff \psi(b) \in \mathrm{GL}_d(\mathbb{Z}) = \{g \in \mathrm{Mat}_d(\mathbb{Z}) \mid \det(g) = \pm 1\}. \tag{3.13}$$

Below we will be studying the subgroup

$$\mathcal{O}^1 = \{b \in \mathcal{O}^\times \mid \psi(b) \in \mathrm{SL}_d(\mathbb{Z})\};$$

this is either $\mathcal{O}^\times$ or an index two subgroup of $\mathcal{O}^\times$, and so it suffices to show the desired description for $\mathcal{O}^1$.

Proposition 3.11 (Compact torus orbit). *Let $v_{\mathcal{J}} = \psi(\zeta) \in \mathrm{Mat}_d(\mathbb{Z})$ and consider the stabilizer subgroup*

$$\mathbb{T}_{\mathcal{J}} = \{g \in \mathrm{SL}_d \mid g v_{\mathcal{J}} g^{-1} = v_{\mathcal{J}}\}$$

for the conjugation action (that is, the centralizer *of $v_{\mathcal{J}}$). Then the orbit*

$$(\mathrm{SL}_d(\mathbb{Z}) I_d)\, \mathbb{T}_{\mathcal{J}}(\mathbb{R})$$

is compact, and the corresponding uniform lattice $\mathbb{T}_{\mathcal{J}}(\mathbb{Z}) < \mathbb{T}_{\mathcal{J}}(\mathbb{R})$ satisfies

$$\mathbb{T}_{\mathcal{J}}(\mathbb{Z}) = \mathrm{SL}_d(\mathbb{Z}) \cap \mathbb{T}_{\mathcal{J}}(\mathbb{R}) = \psi(\mathcal{O}^1).$$

In more technical language, the subgroup $\mathbb{T}_{\mathcal{J}}$ is a special case of an *algebraic torus* (it is in fact a $\mathbb{Q}$-anisotropic $\mathbb{Q}$-torus). Moreover, the algebraic group $\mathbb{T}_{\mathcal{J}}$ is closely related to the group $\mathrm{Res}_{K|\mathbb{Q}}\,\mathbb{G}_m$ obtained by applying *restriction of scalars* to the multiplicative group $\mathbb{G}_m$ — it is the kernel of the $\mathbb{Q}$-split character $N_{K|\mathbb{Q}}$ on $\mathrm{Res}_{K|\mathbb{Q}}\,\mathbb{G}_m$. Minding our language we will not use these words often, but we will give a short introduction to these terms in Chapter 7.

PROOF OF PROPOSITION 3.11. By Proposition 3.8, we know that the orbit is closed. We prove compactness along the lines of the proof of Proposition 3.2.

[†] Moreover, if $\mathcal{O}$ is the maximal order, then any ideal is a proper ideal.

For this we need a replacement for the quadratic form, and this is provided by the *norm form*

$$N(b) = N_{K|\mathbb{Q}}(b) = \det \psi(b)$$

which is originally defined on K. Since K is a field, $N_{K|\mathbb{Q}}(b) = 0$ for $b \in K$ if and only if $b = 0$, which is similar to the hypothesis in Proposition 3.2. Let us write

$$\iota(v) = v_1 a_1 + \cdots + v_d a_d$$

for $v \in \mathbb{Q}^d$, so that by assumption ι gives an isomorphism between $\mathbb{Z}^d$ and $\mathcal{J}$ as well as between $\mathbb{Q}^d$ and K. We also note that $\psi \circ \iota : \mathbb{Q}^d \to \mathrm{Mat}_d(\mathbb{Q})$ is linear, and so we can extend it to a linear map

$$\Psi : \mathbb{R}^d \to \mathrm{Mat}_d(\mathbb{R})$$

which also extends the norm form to the polynomial map $x \mapsto \det(\Psi(x))$ in d variables $x_1, \dots, x_d$ of total degree d.

Now suppose that $(\mathrm{SL}_d(\mathbb{Z}) I_d)\, \mathbb{T}_{\mathcal{J}}(\mathbb{R})$ is unbounded. Then for some m in $\mathbb{Z}^d \smallsetminus \{0\}$ and $h \in \mathbb{T}_{\mathcal{J}}(\mathbb{R})$ the vector mh is very small. This implies that

$$|\det \Psi(mh)| < 1.$$

We claim that

$$\Psi(mh) = \Psi(m)h. \tag{3.14}$$

Assuming this for now, and recalling that $h \in \mathrm{SL}_d(\mathbb{R})$, we obtain that (in analogy to (3.5) on page 77) $|\det \Psi(mh)| = |\det \Psi(m)| < 1$, which forces $\det \Psi(m) = 0$ (since $\det \Psi(m) \in \mathbb{Z}$). However, $m \in \mathbb{Z}^d \smallsetminus \{0\}$ corresponding to some $b = \iota(m) \in \mathcal{J} \smallsetminus \{0\}$ cannot have

$$N_{K|\mathbb{Q}}(b) = \det \Psi(m) = 0,$$

proving that $(\mathrm{SL}_d(\mathbb{Z}) I_d) \mathbb{T}_{\mathcal{J}}(\mathbb{R})$ is bounded, and hence compact.

To prove the claim (3.14), and the statement $\mathbb{T}_{\mathcal{J}}(\mathbb{Z}) = \psi(\mathcal{O}^1)$ in the proposition, we would like to understand $\mathbb{T}_{\mathcal{J}}$ better. Notice that

$$\{g \in \mathrm{Mat}_d \mid g v_{\mathcal{J}} = v_{\mathcal{J}} g\} \tag{3.15}$$

is a linear subspace defined by the requirement to commute with $v_{\mathcal{J}}$. To analyze the dimension[†] of this subspace we may conjugate $v_{\mathcal{J}}$ over $\mathbb{C}$ to the diagonal matrix v_{diag} with eigenvalues

$$\zeta_1, \dots, \zeta_r, \zeta_{r+1}, \overline{\zeta_{r+1}}, \dots, \zeta_{r+s}, \overline{\zeta_{r+s}}.$$

[†] As the subspace in question is defined by rational equations, the dimension of it as a subspace of $\mathrm{Mat}_d(\mathbb{Q})$ over $\mathbb{Q}$ equals the dimension of it as a subspace of $\mathrm{Mat}_d(\mathbb{R})$ over $\mathbb{R}$ (and similarly for $\mathbb{C}$).

As these are all different, the only matrices that commute with v_{diag} are diagonal matrices. This shows that the dimension of the subspace in (3.15) is d. Hence

$$\{g \in \mathrm{Mat}_d(\mathbb{Q}) \mid g v_{\mathcal{J}} = v_{\mathcal{J}} g\} = \psi(K)$$

and taking the $\mathbb{R}$-linear hull we get

$$\{g \in \mathrm{Mat}_d(\mathbb{R}) \mid g v_{\mathcal{J}} = v_{\mathcal{J}} g\} = \langle \psi(\mathbb{K}) \rangle_{\mathbb{R}} = \Psi(\mathbb{R}^d). \tag{3.16}$$

The first of these equations implies that

$$\mathbb{T}_{\mathcal{J}}(\mathbb{Z}) = \psi\left(\{b \in K \mid \psi(b) \in \mathrm{SL}_d(\mathbb{Z})\}\right) = \psi(\mathcal{O}^1)$$

by (3.13).

Also notice that

$$\psi(ab) = \psi(a)\psi(b) \tag{3.17}$$

for $a, b \in K$, since ψ is giving the matrix representation of multiplication[†] by elements of K in the given basis. This may also be phrased as

$$\psi(\iota(mh)) = \psi(\iota(m))h \tag{3.18}$$

for $m \in \mathbb{Z}^d$ and $h \in \psi(K)$. Indeed, $a = \iota(m) \in \mathcal{J}$ and $h = \psi(b)$ is the matrix which sends m corresponding to a to mh corresponding to ab, so that the left-hand sides of (3.17) and (3.18) agree. The right-hand sides agree tautologically, and so (3.18) follows. Equivalently, we have shown

$$\Psi(mh) = \Psi(m)h$$

for $m \in \mathbb{Z}^d$ and $h \in \psi(K)$. However, this is a linear equation in h which therefore also holds for $h \in \Psi(\mathbb{R}^d)$. In summary, we obtain (3.14) and the proposition follows. □

To finish the proof of Theorem 3.10, we need to analyze the structure of $\mathbb{T}_{\mathcal{J}}(\mathbb{R})$.

Proposition 3.12 ($\mathbb{R}$-points of the torus subgroup). *With the notation as above,*

$$\mathbb{T}_{\mathcal{J}}(\mathbb{R}) \cong M \times \mathbb{R}^{r+s-1},$$

where M is a compact linear group with connected component of the identity isomorphic to $(\mathbb{S}^1)^s$.

The pair of numbers (r, s) play a similar role for $\mathbb{T}_{\mathcal{J}}$ as the signature of the associated quadratic form does for an orthogonal group. In this sense, the result above is an analogue of Theorem 3.5.

[†] As K is commutative, we do not have to worry about the order of multiplication of the matrices and the fact that we are using row vectors instead of column vectors.

Proof of Proposition 3.12. We already did most of the work for this in the proof of Proposition 3.11. In fact, as in that proof, the group

$$\mathbb{T}_{\mathcal{J}}(\mathbb{R}) = \{g \in \mathrm{SL}_d(\mathbb{R}) \mid g v_{\mathcal{J}} = v_{\mathcal{J}} g\}$$

is conjugate to[†]

$$\{g \in \mathrm{SL}_d(\mathbb{R}) \mid g v_{\zeta,\mathbb{R}} = v_{\zeta,\mathbb{R}} g\}$$

where $v_{\zeta,\mathbb{R}}$ is the block-diagonal matrix

$$v_{\zeta,\mathbb{R}} = \begin{pmatrix} \zeta_1 & & & & & \\ & \ddots & & & & \\ & & \zeta_r & & & \\ & & & \imath(\zeta_{r+1}) & & \\ & & & & \ddots & \\ & & & & & \imath(\zeta_{r+s}) \end{pmatrix} \in \mathrm{Mat}_d(\mathbb{R})$$

and $\imath$ is the map defined by

$$\imath\colon x + \mathrm{i}y \to \begin{pmatrix} x & y \\ -y & x \end{pmatrix}.$$

We use $v_{\zeta,\mathbb{R}}$ (instead of v_{diag}) to ensure that the conjugation takes place over $\mathbb{R}$, which is needed to analyze $\mathbb{T}_{\mathcal{J}}(\mathbb{R})$. It is easy to check (for example, by a dimension argument as in the proof of Proposition 3.11) that

$$\{g \in \mathrm{Mat}_d(\mathbb{R}) \mid g v_{\zeta,\mathbb{R}} = v_{\zeta,\mathbb{R}} g\}$$

$$= \left\{ \begin{pmatrix} a_1 & & & & & \\ & \ddots & & & & \\ & & a_r & & & \\ & & & \imath(b_1) & & \\ & & & & \ddots & \\ & & & & & \imath(b_s) \end{pmatrix} \mid a_1, \dots, a_r \in \mathbb{R}, b_1, \dots, b_s \in \mathbb{C} \right\}.$$

Therefore $\mathbb{T}_{\mathcal{J}}(\mathbb{R})$ is isomorphic to the multiplicative group

$$\{(a_1, \dots, a_r, b_1, \dots, b_s) \in \mathbb{R}^r \times \mathbb{C}^s \mid a_1 \cdots a_r |b_1|^2 \cdots |b_s|^2 = 1\}$$

[†] Just as in the theory of Jordan normal forms, this follows quickly from consideration of $\mathbb{R}^d$ as an $\mathbb{R}[T]$-module, where T acts via $v_{\mathcal{J}}$, which gives

$$\mathbb{R}^d \cong \mathbb{R}[T]/\langle T-\zeta_1\rangle \times \cdots \times \mathbb{R}[T]/\langle T-\zeta_r\rangle \times \mathbb{R}[T]/\langle p_{\zeta_{r+1}}(T)\rangle \times \cdots \times \mathbb{R}[T]/\langle p_{\zeta_{r+s}}(T)\rangle,$$

where $p_{\zeta_{r+1}}(T), \dots, p_{\zeta_{r+s}}(T)$ are the quadratic real minimal polynomials of $\zeta_{r+1}, \dots, \zeta_{r+s} \in \mathbb{C}$. We refer to Hungerford [81, Ch. VII] for the details.

which contains the non-compact part

$$\{(\mathrm{e}^{t_1},\dots,\mathrm{e}^{t_r},\mathrm{e}^{t_{r+1}},\dots,\mathrm{e}^{t_{r+s}}) \mid t_1+\cdots+t_r+2t_{r+1}+\cdots+2t_{r+s}=0\},$$

and this is isomorphic (as a Lie group) to $\mathbb{R}^{r+s-1}$. The subgroup $M \subseteq \mathbb{T}_{\mathcal{J}}(\mathbb{R})$ is then the subgroup isomorphic to the 'group of signs'

$$\{(\varepsilon_1,\dots,\varepsilon_r,z_1,\dots,z_s) \mid \varepsilon_i \in \{\pm 1\}, |z_i|=1, \varepsilon_1\cdots\varepsilon_r=1\}.$$

□

PROOF OF THEOREM 3.10. By Proposition 3.11, $\mathcal{O}^1$ is isomorphic to a (uniform) lattice in $\mathbb{T}_{\mathcal{J}}(\mathbb{R})$, which by Proposition 3.12 is isomorphic to the abelian group $M \times \mathbb{R}^{r+s-1}$. Taking the quotient by M we obtain a uniform lattice in $\mathbb{R}^{r+s-1}$, which must be generated by $r+s-1$ elements. Suppose that $b_1,\dots,b_{r+s-1} \in \mathcal{O}^1$ are elements that give rise to a $\mathbb{Z}$-basis of the lattices in $\mathbb{R}^{r+s-1}$. Then $b_1,\dots,b_{r+s-1}$ generate $\mathcal{O}^1$ up to the kernel of the map from $\mathcal{O}^1$ to $\mathbb{R}^{r+s-1}$. However, this kernel F maps under ψ and the isomorphism to $M \times \mathbb{R}^{r+s-1}$ to the compact group M (with discrete image) and so must be finite. □

3.3.1 More on Compact Orbits for the Diagonal Subgroup

[†]The proof above, while it fits naturally into our discussion, is certainly not the shortest proof of Dirichlet's unit theorem. However, the set-up used here can be used further to discuss interesting distribution properties of compact orbits arising from number fields. For this we define the *complete Galois embedding*

$$\phi=(\phi_1,\dots,\phi_r,\phi_{r+1},\dots,\phi_{r+s}) : K \to \mathbb{R}^r \times \mathbb{C}^s \cong \mathbb{R}^{r+2s} \tag{3.19}$$

(which clearly is an embedding, since each ϕ_i is injective). We define (r,s) to be the *type* of the number field (as mentioned this plays the role of the signature of a quadratic form), and define $\mathbb{T}_{r,s}$ to be the centralizer of a regular matrix

$$v_{r,s}=\begin{pmatrix}\alpha_1 &&&&& \\ & \ddots &&&& \\ && \alpha_r &&& \\ &&& \imath(\beta_1) && \\ &&&& \ddots & \\ &&&&& \imath(\beta_s)\end{pmatrix}$$

[†] The remainder of Section 3.3 will not be needed again.

with pairwise different $\alpha_j \in \mathbb{R}$ and $\beta_j \in \mathbb{C}\smallsetminus\mathbb{R}$.

Proposition 3.13 (Ideal classes and torus orbits). *Let K be a number field of type (r,s), and let $\mathcal{O}$ be an order in K. Then for any proper $\mathcal{O}$-ideal $\mathcal{J} \subseteq K$ the normalized lattice*

$$x_{\mathcal{J}} = \frac{1}{\operatorname{covol}(\phi(\mathcal{J}))^{1/d}}\phi(\mathcal{J}) \in \mathsf{X}_d$$

has compact orbit under $\mathbb{T}_{r,s}(\mathbb{R})$. Two ideals $\mathcal{J}_1, \mathcal{J}_2$ give rise to the same orbit if and only if they are ideals in the same number field (and order), and are equivalent (that is, there exists some $a \in K\smallsetminus\{0\}$ with $\mathcal{J}_1 = a\mathcal{J}_2$).

For the relationship above between ideal classes and compact orbits, we need to agree on the following convention. For a given field K and all its isomorphic copies we agree to pick one representative, say K, and one complete Galois embedding ϕ as in (3.19).

Proof of Proposition 3.13. Let $K = \mathbb{Q}(\zeta)$, $\mathcal{O}$, and $\mathcal{J} \subseteq \mathcal{O}$ be given. We will use the same notation as used in Proposition 3.11. Recall that $\{a_1, \dots, a_d\}$ is a basis of $\mathcal{J}$. Taking the image of this basis under ϕ, we obtain a basis of $\mathbb{R}^d$. Indeed, if this were not the case then we could find non-zero elements $b \in \mathcal{O}$ for which $\phi(b)$ is arbitrarily small (see Exercise 1.3.1). However, this also implies that $|N(b)| = |\det\psi(b)| < 1$ and so with $b \in \mathcal{O}$ a contradiction. Replacing a_1 with $-a_1$ if necessary, we may assume that

$$g_{\mathcal{J}} = \frac{1}{\operatorname{covol}(\phi(\mathcal{J}))^{1/d}}\begin{pmatrix}\phi(a_1)\\ \vdots \\ \phi(a_d)\end{pmatrix}$$

has determinant one. By construction, $x_{\mathcal{J}} = \mathbb{Z}^d g_{\mathcal{J}}$; also notice that $g_{\mathcal{J}}$ is up to the scalar the matrix representation of the map ϕ from K (with the basis $\{a_1, \dots, a_d\}$) to $\mathbb{R}^r \times \mathbb{C}^s$ (with the standard basis). Furthermore, recall that $v_{\mathcal{J}} = \psi(\zeta)$ is the matrix representation of multiplication by ζ on K (with basis $a_1, \dots, a_d$). In $\mathbb{R}^r \times \mathbb{C}^s$ multiplication by ζ corresponds to multiplying the various coordinates by $\phi_1(\zeta), \dots, \phi_r(\zeta)$ and to applying the matrices corresponding to the complex numbers $\phi_{r+1}(\zeta), \dots, \phi_{r+s}(\zeta)$ respectively; that is, to an application of a block-diagonal matrix $v_{\zeta,\mathbb{R}}$. This shows (as we are using row vectors) that

$$v_{\mathcal{J}} g_{\mathcal{J}} = g_{\mathcal{J}} v_{\zeta,\mathbb{R}}. \tag{3.20}$$

Now $v_{\zeta,\mathbb{R}}$ is of the same type as $v_{r,s}$ and defines the same centralizer $\mathbb{T}_{r,s}$. Therefore,

$$\mathbb{T}_{r,s} = g_{\mathcal{J}}^{-1}\mathbb{T}_{\mathcal{J}} g_{\mathcal{J}}$$

since (3.20) gives for instance for $g \in \mathbb{T}_{\mathcal{J}}$ that

$$g_{\mathcal{J}}^{-1} g g_{\mathcal{J}} v_{\zeta,\mathbb{R}} = g_{\mathcal{J}}^{-1} g v_{\mathcal{J}} g_{\mathcal{J}} = g_{\mathcal{J}}^{-1} v_{\mathcal{J}} g g_{\mathcal{J}} = v_{\zeta,\mathbb{R}} g_{\mathcal{J}}^{-1} g g_{\mathcal{J}}.$$

Moreover,

$$\mathrm{SL}_d(\mathbb{Z})g_{\mathcal{J}}\mathbb{T}_{r,s}(\mathbb{R}) = \mathrm{SL}_d(\mathbb{Z})\mathbb{T}_{\mathcal{J}}(\mathbb{R})g_{\mathcal{J}}$$

is compact by Proposition 3.11.

Notice that if we choose a different basis of $\mathcal{J}$, then this does not change the point $x_{\mathcal{J}} \in \mathsf{X}_d$. Also notice that if $\mathcal{J}' = b\mathcal{J}$ for some $b \in K^\times$ then $ba_1, \dots, ba_d$ is a basis of $\mathcal{J}'$, and using this basis we see by (3.20) (which by the same argument also holds for $\cdot b$ instead of $\cdot\zeta$) that

$$g_{\mathcal{J}'} = \psi(b)g_{\mathcal{J}} = g_{\mathcal{J}}v_{b,\mathbb{R}}.$$

Since $v_{b,\mathbb{R}} \in \mathbb{T}_{r,s}(\mathbb{R})$ this shows that

$$x_{\mathcal{J}'} \in x_{\mathcal{J}}\mathbb{T}_{r,s}(\mathbb{R}),$$

which is the first direction of the second claim in the proposition.

Let now $\mathcal{J}$ (and $\mathcal{J}'$) be a proper $\mathcal{O}$ (respectively $\mathcal{O}'$)-ideal in a number field K (respectively K'), let $x_{\mathcal{J}}$, $x_{\mathcal{J}'}$ be the corresponding elements of X_d, and assume that

$$x_{\mathcal{J}'} = x_{\mathcal{J}}t$$

for some $t \in \mathbb{T}_{r,s}(\mathbb{R})$. By the definition of properness for an $\mathcal{O}$-ideal $\mathcal{J}$ we have

$$\begin{aligned} \mathcal{O} &= \{a \in K \mid a\mathcal{J} \subseteq \mathcal{J}\} \\ &\cong \{v \in \langle\psi(\mathbb{K})\rangle_{\mathbb{R}} \mid \mathbb{Z}^d v \subseteq \mathbb{Z}^d\} && \text{(via } v = \psi(a)\text{)} \\ &= \{v \in \mathrm{Mat}_d(\mathbb{R}) \mid vv_{\mathcal{J}} = v_{\mathcal{J}}v \text{ and } \mathbb{Z}^d v \subseteq \mathbb{Z}^d\} \\ &\cong \{v \in \mathrm{Mat}_d(\mathbb{R}) \mid vv_{r,s} = v_{r,s}v \text{ and } x_{\mathcal{J}}v \subseteq x_{\mathcal{J}}\}, \end{aligned}$$

via conjugation by $g_{\mathcal{J}}$. The latter set comprises all block diagonal matrices with entries $\phi(a)$ for all $a \in \mathcal{O}$. For the lattices $x_{\mathcal{J}'}$ and $x_{\mathcal{J}}$, this implies that $\mathcal{O} \cong \mathcal{O}'$ and hence $K \cong K'$. By the convention discussed just after the proposition, this means that $K = K'$, and that the same complete Galois embedding ϕ is used. By the argument above, this also implies that we have $\mathcal{O} = \mathcal{O}'$. Suppose that $a_1, \dots, a_d$ is a basis of $\mathcal{J}$, so that $x_{\mathcal{J}} = \mathbb{Z}^d g_{\mathcal{J}}$ as before. Choosing the basis $a'_1, \dots, a'_d$ of $\mathcal{J}'$ correctly gives $x_{\mathcal{J}'} = \mathbb{Z}^d g_{\mathcal{J}'}$ and $g_{\mathcal{J}'} = g_{\mathcal{J}}t$. This shows that $\phi_i(a'_j) = \phi_i(a_j)t_i$ where t_i (in $\mathbb{R}$ or $\mathbb{C}$) is the ith entry of the block-diagonal matrix $t \in \mathbb{T}_{r,s}(\mathbb{R})$. This implies that

$$t_i = \phi_i\left(\frac{a'_j}{a_j}\right)$$

is independent of j. Hence there exists some $b \in K$ with

$$t_i = \phi_i(b)$$

for $i = 1, \dots, r+s$, and it follows that $\mathcal{J}' = b\mathcal{J}$. □

We remark that for a given order $\mathcal{O}$ there are only finitely many inequivalent ideal classes of proper $\mathcal{O}$-ideals. This observation makes the following folklore conjecture (generalizing results and conjectures of Linnik) well-formulated.

Conjecture 3.14. For a given order $\mathcal{O}$ in an algebraic number field K of type (r, s), let $\mu_{\mathcal{O}}$ be the probability measure on X_d obtained from normalizing the sum of the $\mathbb{T}_{r,s}(\mathbb{R})$-invariant probability measures on $x_{\mathcal{J}}\mathbb{T}_{r,s}(\mathbb{R})$ for the various ideal classes of proper $\mathcal{O}$-ideals. Then, as the *discriminant* $D = (\operatorname{covol}(\phi(\mathcal{J})))^2$ goes to infinity, all of the weak*-limit of the measures $\mu_{\mathcal{O}}$ are Haar measures of finite volume orbits xL for some closed linear subgroup $L \subseteq \mathrm{SL}_d(\mathbb{R})$.

This conjecture has been shown for $d = 2$ by Duke [38] (using subconvexity of L-functions, building on a breakthrough of Iwaniec [84]), and for $d = 3$ and type $r = 3$, $s = 0$ by Einsiedler, Lindenstrauss, Michel and Venkatesh [51] (by combining subconvexity bounds for L-functions with ergodic methods). More accessible but weaker results are contained in [46] and [52].

Exercises for Section 3.3

Exercise 3.3.1. (a) Let $d \geqslant 2$. Show that the compact orbits of $\mathbb{T}_{(d,0)}(\mathbb{R})$ (of type $(d, 0)$) in X_d are all of the form $x_{\mathcal{J}}\mathbb{T}$ for some proper $\mathcal{O}$-ideal and some order $\mathcal{O} \subseteq K$ in a totally real number field.
(b) Show that this is not the case for the type $(0, d/2)$ (with d even).
(c) Decide the same question for the remaining cases.

3.4 Linear Algebraic Groups

In this section and in Chapter 7 we will introduce linear algebraic groups, and will link this concept to the theory of linear Lie groups, pointing out the obvious similarities as well as some of the more subtle differences between the theories. We start with the basic definitions, but in order to avoid being too diverted by this important (and large) theory, we will be brief at times.

3.4.1 Basic Notions of Algebraic Varieties

Let $\mathbb{K}$ be a field[†] and let $\overline{\mathbb{K}}$ denote an algebraic closure of $\mathbb{K}$. A subset $S \subseteq \overline{\mathbb{K}}^d$ is called *Zariski closed* if $S = Z(\mathcal{J})$ is the variety $Z(\mathcal{J})$ defined by a subset

---

[†] We will generally be interested in the cases $\mathbb{R}$, $\mathbb{Q}_p$ and $\mathbb{Q}$, but will only assume that the field has characteristic zero a little later.

or, without loss of generality, an ideal $\mathcal{J} \subseteq \overline{\mathbb{K}}[x_1, \dots, x_d]$. A subset $S \subseteq \overline{\mathbb{K}}^d$ is also called Zariski $\mathbb{K}$-closed if $\mathcal{J}$ can be chosen in $\mathbb{K}[x_1, \dots, x_d]$. The Zariski closed subsets are the closed sets of a topology, which is called the *Zariski topology.* This is easily checked:

- If $S_1 = Z(\mathcal{J}_1)$ and $S_2 = Z(\mathcal{J}_2)$ then $S_1 \cup S_2 = Z(\mathcal{J}_1 \mathcal{J}_2)$.
- If $S_\alpha = Z(\mathcal{J}_\alpha)$ for $\alpha \in A$, then

$$\bigcap_{\alpha \in A} S_\alpha = Z\left(\bigcup_{\alpha \in A} \mathcal{J}_\alpha\right).$$

If $\mathbb{K} = \mathbb{R}$, $\mathbb{K} = \mathbb{C}$, or $\mathbb{K} = \mathbb{Q}_p$, then clearly every Zariski closed (or Zariski open) subset is also closed (or open) in the usual sense. For most of the derived properties (density, connectedness) this is not clear and indeed is often false. We will always say Zariski open, Zariski closed, Zariski dense, and so on, if we refer to properties of the Zariski topology. When we use the words open, closed, dense, and so on, then this will refer to the metric (often also referred to as the *Hausdorff*) topology of $\mathbb{R}^d$, $\mathbb{C}^d$, or $\mathbb{Q}_p^d$ derived from the norms on these spaces.

A variety (equivalently, a Zariski closed set) is called *Zariski connected*[†] or *irreducible* if it is not a union of two proper Zariski closed subsets. Equivalently, a variety Z is irreducible if its ring of regular functions

$$\overline{\mathbb{K}}[Z] = \overline{\mathbb{K}}[x_1, \dots, x_d]/\mathcal{J}(Z)$$

is a principal ideal domain (that is, without zero divisors).

Assume now that $Z = Z(\mathcal{J})$ is a connected variety. Then we can form the field of rational functions $\overline{\mathbb{K}}(Z)$ comprising all quotients $\frac{f}{g}$ with $f, g \in \overline{\mathbb{K}}[Z]$ and $g \neq 0$. The *transcendence degree*[‡] (see Hungerford [81, Sec. VI.1]) of $\overline{\mathbb{K}}(Z)$ is called the *dimension* $\dim(Z)$ of the variety Z. Notice that if $Z = \overline{\mathbb{K}}^d$ then the dimension of Z is d, and if Z is defined by a single irreducible polynomial

$$f \in \overline{\mathbb{K}}[x_1, \dots, x_d]$$

(in which case Z is called a *hypersurface*), then the dimension of Z is $(d-1)$. The following lemma further reinforces our intuition concerning this notion of dimension.

[†] This definition does not match the topological definition of connectedness, but it will come closer to doing so in the context of algebraic subgroups.

[‡] A field extension $\mathbb{F}|\mathbb{K}$ has transcendence degree n if $\mathbb{F}$ contains n mutually transcendental elements $f_1, \dots, f_n \in \mathbb{F}$ (that is, elements with the property that the evaluation map

$$\mathbb{K}[T_1, \dots, T_n] \ni g \mapsto g(f_1, \dots, f_n)$$

is injective) but does not contain $n+1$ mutually transcendental elements.

Lemma 3.15 (Strict monotonicity of dimension). *Suppose that* $Z_2 \subseteq Z_1$ *is a proper connected subvariety of a connected variety* $Z_1 \subseteq \overline{\mathbb{K}}^d$*. Then*

$$\dim Z_2 < \dim Z_1.$$

Proof. By definition

$$\overline{\mathbb{K}}[Z_1] = \overline{\mathbb{K}}[x_1, \dots, x_d]/\mathcal{J}_1,$$

with $\mathcal{J}_1 = \mathcal{J}(Z_1)$, has transcendence degree $k = \dim Z_1$. By reordering the variables if necessary, we may assume that

$$x_1 + \mathcal{J}_1, \dots, x_k + \mathcal{J}_1 \in \overline{\mathbb{K}}[Z_1] \tag{3.21}$$

are algebraically independent, and

$$x_{k+1} + \mathcal{J}_1, \dots, x_d + \mathcal{J}_1$$

are algebraically dependent on the elements in (3.21). All other regular or rational functions in $\overline{\mathbb{K}}(Z_1)$ are then algebraically dependent on the elements in (3.21). It follows that

$$\overline{\mathbb{K}}(Z_1) \cong \overline{\mathbb{K}}(x_1, \dots, x_k)\,[x_{k+1} + \mathcal{J}_1, \dots, x_d + \mathcal{J}_1]$$

is a finite field extension of the field of rational functions in the first k variables.

Since $Z_2 \subseteq Z_1$ is a proper subvariety, there exists some $f \in \mathcal{J}(Z_2) \smallsetminus \mathcal{J}(Z_1)$. As $f + \mathcal{J}_1$ is non-zero in $\overline{\mathbb{K}}(Z_1)$, there exists some

$$g + \mathcal{J}_1 \in \overline{\mathbb{K}}(x_1, \dots, x_k)\,[x_{k+1} + \mathcal{J}_1, \dots, x_d + \mathcal{J}_1]$$

such that $fg + \mathcal{J}_1 = 1 + \mathcal{J}_1$. Clearing the denominators (which belong to $\overline{\mathbb{K}}[x_1, \dots, x_k]$) in this relation, we find that there exists some $g_1 \in \overline{\mathbb{K}}[Z_1]$ such that

$$fg_1 + \mathcal{J}_1 = h + \mathcal{J}_1$$

for some non-zero $h \in \overline{\mathbb{K}}[x_1, \dots, x_k] \cap \mathcal{J}(Z_2)$. This shows that the transcendence degree of $\overline{\mathbb{K}}(Z_2)$ is less than or equal to $k - 1$. □

Assume again that $Z \subseteq \overline{\mathbb{K}}^d$ is a connected k-dimensional variety. A point $x^{(0)} \in Z$ is called *smooth* if the 'tangent space' in the variables $u_1, \dots, u_d$ defined by

$$\sum_{j=1}^{d} u_j \partial_{x_j} f(x^{(0)}) = (u_1, \dots, u_d) \cdot \nabla f(x^{(0)}) = 0$$

for all $f \in \mathcal{J}(Z)$, is k-dimensional. The partial derivatives are defined as abstract linear maps on the space of polynomials (so that the definition matches the usual maps if $\mathbb{K}$ is $\mathbb{R}$ or $\mathbb{C}$). It satisfies the usual properties (the product and chain rules, for example) over any field $\mathbb{K}$. The reader may quickly decide which points of the variety defined by the equation $y^2 = x^3$ are smooth in this sense (and thus see why the definition makes sense, and that it accords with geometrical intuition; see also Lemma 3.18). A variety is called *smooth* if every point of the variety is a smooth point.

Lemma 3.16 (Most points are smooth). *Let $Z \subseteq \overline{\mathbb{K}}^d$ be a connected variety and suppose the characteristic $\operatorname{char}\mathbb{K}$ of the field $\mathbb{K}$ is zero. Then the set of smooth points of Z is a non-empty Zariski open subset of Z. Moreover, the tangent space has at no point of Z a dimension smaller than $\dim Z$.*

The lemma should indeed be interpreted as saying that most points of a connected variety are smooth. This is because a non-empty Zariski open subset of a connected variety is automatically Zariski dense. Moreover, Zariski dense and Zariski open subsets of any variety have a nice intersection property[†]: every finite intersection of Zariski dense and open subsets is again Zariski dense and open.

Proof of Lemma 3.16. Let $k = \dim Z$, and assume again that

$$x_1 + \mathcal{J}(Z), \dots, x_k + \mathcal{J}(Z) \in \overline{\mathbb{K}}(Z)$$

are algebraically independent while

$$x_{k+1} + \mathcal{J}(Z), \dots, x_d + \mathcal{J}(Z)$$

are algebraically dependent on

$$x_1 + \mathcal{J}(Z), \dots, x_k + \mathcal{J}(Z).$$

Thus there exists, for every $\ell \in \{k+1, \dots, d\}$ a non-zero polynomial

$$f_\ell \in \overline{\mathbb{K}}[x_1, \dots, x_\ell] \cap \mathcal{J}(Z)$$

of minimal degree in x_ℓ for which (viewed as a polynomial in x_ℓ) the non-zero coefficients do not belong to $\overline{\mathbb{K}}[x_1, \dots, x_{\ell-1}] \cap \mathcal{J}(Z)$. Since $\operatorname{char}\mathbb{K} = 0$, we get[‡]

$$g_\ell = \partial_{x_\ell} f_\ell \notin \mathcal{J}(Z).$$

Using the derivative $\nabla(f_\ell)$ (for $\ell = k+1, \dots, d$) of these polynomials (as equations that define the tangent space) we see that every point outside the proper subvariety defined by the ideal

---

[†] For a connected variety this is easy to see. For a general variety this follows for example from the decomposition discussed in Lemma 3.17.

[‡] If $\operatorname{char}\mathbb{K} = p$ and it so happens that f_ℓ is a polynomial in $x_1, \dots, x_{\ell-1}, x_\ell^p$ then $\partial_{x_\ell} f_\ell = 0$. With more care this problem can be dealt with — we refer to Hartshorne [75] for the details.

$$\langle g_{k+1}\cdots g_d, \mathcal{J}(Z)\rangle$$

(that is, every point in a non-empty Zariski open subset O) has a tangent space of dimension less than or equal to k. To see that these points are smooth points of the variety we have to show that the tangent space is indeed k-dimensional. We show this first[†] on an even smaller Zariski open subset O'.

We claim that there exists some non-zero $h \in \overline{\mathbb{K}}[x_1,\dots,x_d]\smallsetminus\mathcal{J}(Z)$ with

$$h\mathcal{J}(Z) \subseteq (f_{k+1},\dots,f_d).$$

Assuming the claim for any $f\in\mathcal{J}(Z)$, we see that $hf = g_1' f_{k+1}+\cdots+g_{d-k}' f_d$, so that

$$\nabla(hf) = \nabla(h)f + h\nabla(f) = \\ \nabla(g_1')f_{k+1} + g_1'\nabla(f_{k+1}) + \cdots + \nabla(g_{d-k}')f_d + g_{d-k}'\nabla(f_d).$$

After evaluation at any point $x\in Z$ we then get

$$h(x)\nabla(f)(x) = g_1'(x)\nabla(f_{k+1})(x) + \cdots + g_{d-k}'(x)\nabla(f_d)(x)$$

which expresses $\nabla(f)(x)$ as a linear combination of $\nabla(f_j)(x)$ for

$$j = k+1,\dots,d$$

if only $h(x)\neq 0$. This shows that on the Zariski open set

$$O' = Z\smallsetminus Z(hg_{k+1}\cdots g_d)$$

every tangent space is exactly k-dimensional.

We now prove the claim. As $\mathcal{J}$ is finitely generated and prime, we only have to show that for every $f\in\mathcal{J}$ there is some $h\notin\mathcal{J}$ with $hf\in(f_{k+1},\dots,f_d)$. If $f\in\overline{\mathbb{K}}[x_1,\dots,x_{k+1}]\cap\mathcal{J}$, then we can take h to be a power of the leading coefficient of f_{k+1} (considered as a polynomial in x_{k+1} with coefficients in $\overline{\mathbb{K}}[x_1,\dots,x_k]$). In fact, with this choice of h we ensure that we can apply division with remainder[‡] to obtain $hf = af_{k+1}+b$ where $b=0$ as it has smaller degree in x_{k+1} than f_{k+1} does and belongs to $\mathcal{J}$. By induction on ℓ the same argument applies for any $f\in\overline{\mathbb{K}}[x_1,\dots,x_{\ell+1}]\cap\mathcal{J}$ (where we will have $b\in\overline{\mathbb{K}}[x_1,\dots,x_\ell]\cap\mathcal{J}$ by the same argument).

It remains to show that the set of smooth points is Zariski open and that at no point of Z the tangent space has dimension $<k$. If now

$$x^{(0)} = \left(x_1^{(0)},\dots,x_d^{(0)}\right)\in Z$$

[†] We use this step below to show that we can never have a tangent space of dimension strictly less than k, hence we cannot rely on this fact here.

[‡] Formally we apply division with remainder in the Euclidean domain $\overline{\mathbb{K}}(x_1,\dots,x_k)[x_{k+1}]$, and later in the argument in the Euclidean domain $\overline{\mathbb{K}}(x_1+J,\dots,x_\ell+J)[x_{\ell+1}]$.

is an arbitrary smooth point, or more generally a point whose tangent space has dimension $K \leqslant k$, then we may reorder the variables so that the tangent space projects onto the subspace spanned by the first K basis vectors, and so that for each $\ell \in \{K+1, \dots, d\}$ there exists some $f_\ell \in \mathcal{J}(Z)$ such that

$$(\nabla f_\ell)_\ell \neq 0$$

but

$$(\nabla f_\ell)_j = 0$$

for $j \in \{K+1, \dots, d\} \smallsetminus \{\ell\}$. It follows that the determinant

$$g = \det (\nabla f_\ell)_j \,,$$

where $\ell, j \in \{K+1, \dots, d\}$, does not vanish at the point $x^{(0)}$. Unfolding the definition shows that any other point

$$x \in O_g = Z \smallsetminus Z(g)$$

is also a point at which the tangent space has dimension less than or equal to K, which is less than or equal to k.

If $K < k$ at some point $x^{(0)}$, then we have found a non-empty Zariski open subset O_g on which all points have tangent spaces of dimension less than or equal to K. However, as Z is irreducible this set would have to intersect the non-empty Zariski open subset O' (on which the tangent spaces are known to be k-dimensional) nontrivially, which would give a contradiction.

Therefore, there is no point where the tangent space has dimension strictly less than k, and so applying the argument for $K = k$ we see that the set of smooth points is Zariski open (and Zariski dense). □

To generalize the notion of smoothness to general varieties we need another lemma.

Lemma 3.17 (Decomposition into Zariski connected components). *Let Z be a variety. Then Z is a finite union*

$$Z = \bigcup_{i=1}^{n} Z_i$$

of connected varieties $Z_1, \dots, Z_n$, where we may and will assume that $Z_i \not\subseteq Z_j$ for $i \neq j$. We will refer to $Z_1, \dots, Z_n$ as the Zariski connected components*. We claim furthermore that the decomposition into Zariski connected components is (up to their order) unique.*

We note that if Z is a hypersurface, then the claimed existence and uniqueness follow quickly from the statement that $\overline{\mathbb{K}}[x_1, \dots, x_d]$ is a unique factorization domain.

Sketch of Proof of Lemma 3.17. The existence of the decomposition follows from the fact that $\overline{\mathbb{K}}[x_1,\dots,x_d]$ is Noetherian. We sketch the argument. If $\mathcal{J}=\mathcal{J}(Z)\subseteq\overline{\mathbb{K}}[x_1,\dots,x_d]$ is not a prime ideal, then there exist

$$f_1,f_2\in\overline{\mathbb{K}}[x_1,\dots,x_d]\smallsetminus\mathcal{J}$$

with $f_1f_2\in\mathcal{J}$. We may define $\mathcal{J}_1=\langle\mathcal{J},f_1\rangle$ and $\mathcal{J}_2=\langle\mathcal{J},f_2\rangle$. Notice that $\mathcal{J}_1\mathcal{J}_2\subseteq\mathcal{J}\subseteq\mathcal{J}_1\cap\mathcal{J}_2$. If both of these are prime ideals, then we are done (see below). If not, then we may assume that $\mathcal{J}_1$ is not a prime ideal, and repeating the argument gives ideals $\mathcal{J}_{1,1},\mathcal{J}_{1,2}$. We do the same for $\mathcal{J}_2$ if $\mathcal{J}_2$ is not a prime ideal, and repeat as necessary. By the Noetherian property this construction has to terminate after finitely many steps. In other words, we can always find a finite tree with $\mathcal{J}$ at the top and prime ideals at the bottom, as illustrated in Figure 3.1.

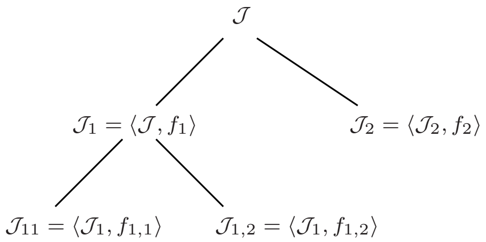

Fig. 3.1: Ideals inside $\mathcal{J}$.

If the prime ideals found are denoted $P_1,\dots,P_n$, then we have (by construction of the prime ideals) that

$$P_1\cdots P_n\subseteq\mathcal{J}\subseteq\bigcap_{i=1}^n P_i. \tag{3.22}$$

This translates to the statement

$$Z=\bigcup_{i=1}^n Z(P_i).$$

If the list of prime ideals has repetitions, we simply remove the repetitions. Also, if $P_i\subseteq P_j$ for $i\neq j$ then $Z(P_i)\supseteq Z(P_j)$ and we remove P_j from the list. Using $\mathcal{J}=\operatorname{rad}(\mathcal{J})$, we can now show that (3.22) still holds for the shortened list. Finally, uniqueness follows directly from the definitions: If $P_1,\dots,P_n$ and $P_1',\dots,P_m'$ both satisfy (3.22) (and are minimal lists), then for every

$$j\in\{1,\dots,m\}$$

we have

$$P_1 \cdots P_n \subseteq P'_j$$

and since P'_j is a prime ideal there exists some $i(j)$ with $P_{i(j)} \subseteq P_j$. Similarly, there exists for every $i \in \{1, \dots, n\}$ some $j(i)$ with $P'_{j(i)} \subseteq P_i$. Since now $P_{i(j(i))} \subseteq P_i$ for every i and $P'_{j(i(j))} \subseteq P'_j$ for every j, it follows that $i(\cdot)$ and $j(\cdot)$ are inverses of each other, $m = n$, and $P'_{j(i)} = P_i$. □

A point $x^{(0)} \in Z$ of a (not necessarily connected) variety is *smooth* if $x^{(0)}$ belongs to precisely one of the connected varieties $Z_i \subseteq Z$ as above, and $x^{(0)}$ is a smooth point of Z_i. Lemma 3.18 now says that inside every variety Z the subset of points that are smooth points of Z is a Zariski open and dense subset of Z.

3.4.2 Properties Concerning the Field

One smooth $\mathbb{K}$-point of a variety already gives rise to many other $\mathbb{K}$-points, if $\mathbb{K}$ is a local field.

Lemma 3.18 (Neighborhoods of smooth points). *Let $Z \subseteq \mathbb{C}^d$ be a k-dimensional connected variety defined over $\mathbb{R}$. Let $x^{(0)} \in Z(\mathbb{R})$ be a smooth point. Then there exists an analytic function defined on an open subset in $\mathbb{R}^k$ which is a homeomorphism to a neighborhood of $x^{(0)} \in Z(\mathbb{R})$. The same holds over $\mathbb{C}$ or over $\mathbb{Q}_p$ for a prime $p < \infty$.*

PROOF. Choose some $f_1, \dots, f_{d-k} \in \mathcal{J}(Z)$ such that $\nabla(f_j)(x^{(0)})$ are linearly independent for $j = 1, \dots, d-k$. By choosing a new coordinate system $x_1, \dots, x_k, y_1, \dots, y_{d-k}$ (which we will abbreviate to x, y) we can assume without loss of generality that

$$\partial_{y_i}(f_j)(x^{(0)}, y^{(0)}) = \delta_{ij}$$

for $i, j = 1, \dots, d-k$, and furthermore

$$\partial_{x_i}(f_j)(x^{(0)}, y^{(0)}) = 0$$

for $i = 1, \dots, k$ and $j = 1, \dots, d-k$.

Applying the implicit function theorem (over $\mathbb{R}$, $\mathbb{C}$, or$^{(14)}$ $\mathbb{Q}_p$) on a neighborhood of $(x^{(0)}, y^{(0)})$ to the equations $f_1(x, y) = \cdots = f_{d-k}(x, y) = 0$, we obtain $(d-k)$ analytic functions $\phi_1(x), \dots, \phi_{d-k}(x)$ which are all defined on a neighborhood U of $x^{(0)}$ such that

$$f_j\left(x, \phi_1(x), \dots, \phi_{d-k}(x)\right) = 0$$

for $j = 1, \dots, d-k$. It remains to see why the points

$$(x, \phi_1(x), \dots, \phi_{d-k}(x))$$

belong to Z (this is in question because we do not know whether $f_1, \dots, f_{d-k}$ generate $\mathcal{J}(Z)$) in some possibly smaller neighborhood $U' \subseteq U$.

Let

$$\mathcal{J}' = \langle f_1, \dots, f_{d-k} \rangle \subseteq \mathcal{J} = \mathcal{J}(Z)$$

and $Z(\mathcal{J}') = Z \cup Z'$, where Z' is the union of all connected components of $Z(\mathcal{J}')$ other than Z. Here Z cannot be contained properly in a connected component of $Z' \subseteq Z(\mathcal{J}')$ since the tangent space of $Z(\mathcal{J}')$ at $\left(x^{(0)}, y^{(0)}\right)$ has dimension k, which would contradict Lemma 3.15 and Lemma 3.16.

We claim that $(x^{(0)}, y^{(0)}) \notin Z'$. Assuming this claim, there exists some polynomial $g \in \mathcal{J}(Z')$ with $g(x^{(0)}, y^{(0)}) \neq 0$. Suppose now that $f \in \mathcal{J}$. Then the product $f \cdot g$ vanishes on

$$Z \cup Z' = Z(\mathcal{J}'),$$

and so there exists some ℓ with $(fg)^\ell \in \mathcal{J}'$ by Hilbert's Nullstellensatz (Theorem 3.6). Let

$$U' = \{x \in U \mid g(x, \phi_1(x), \dots, \phi_{d-k}(x)) \neq 0\},$$

and suppose that $x \in U'$. Then

$$(fg)^\ell \left(x, \phi_1(x), \dots, \phi_{d-k}(x)\right) = 0$$

since all elements in $\mathcal{J}'$ vanish on such vectors (by definition of $\phi_1, \dots, \phi_{d-k}$). However, since $x \in U'$, this shows that

$$f\left(x, \phi_1(x), \dots, \phi_{d-k}(x)\right) = 0$$

for all $f \in \mathcal{J}$ and $x \in U'$, as required.

To prove the claim[†] we will show that after removing all connected components of Z' that do not contain $(x^{(0)}, y^{(0)})$ from $Z(\mathcal{J}') = Z \cup Z'$ we obtain a connected variety Z'' (see below for a more formal definition). Since Z is not removed, this then implies that $Z'' = Z$ and that Z' does not contain $(x^{(0)}, y^{(0)})$ as claimed.

Let $\mathcal{R} = \mathbb{R}[x, y]$ be the ring of polynomials, let

$$\mathcal{M} = \{f \in \mathcal{R} \mid f(x^{(0)}, y^{(0)}) = 0\}$$

be the maximal ideal in $\mathcal{R}$ corresponding to $(x^{(0)}, y^{(0)})$, and let

---

[†] A slight warning is in order. The remainder of this argument is surprisingly long, and quite algebraic. The reader who wishes to only get a glimpse of the algebraic background may decide to skip it, we will not need this type of argument again.

$$\mathcal{R}_{\mathcal{M}} = \left\{ \frac{f}{g} \mid f, g \in \mathbb{R}[x, y] \text{ with } g \notin \mathcal{M} \right\}$$

be the local ring[†] corresponding to $\mathcal{M}$ (consisting of rational functions that are well-defined at $(x^{(0)}, y^{(0)})$).

Let $Z_1, \dots, Z_a$ be the connected components of $Z \cup Z' = Z(\mathcal{J}')$ that contain $(x^{(0)}, y^{(0)})$, and let $Z_1', \dots, Z_b'$ be those that do not contain $(x^{(0)}, y^{(0)})$. Making our definition above more precise, we set $Z'' = Z_1 \cup \cdots \cup Z_a$. We now show that

$$\mathcal{J}'' = (\mathcal{J}'\mathcal{R}_{\mathcal{M}}) \cap \mathcal{R} = \left\{ f = \frac{p}{q} \in \mathcal{R} \mid p \in \mathcal{J}', q \notin \mathcal{M} \right\}$$

defines the variety Z''. Pick a polynomial $q_j \in J(Z_j')$ with $q_j(x^{(0)}, y^{(0)}) \neq 0$ for $j = 1, \dots, b$. Choose any polynomials $F_i \in \mathcal{J}(Z_i)$ for $i = 1, \dots, a$. Then by Hilbert's Nullstellensatz (Theorem 3.6) there exists some $\ell \geqslant 1$ with

$$(F_1 \cdots F_a q_1 \cdots q_b)^\ell \in \mathcal{J}'.$$

Using only the definition of $\mathcal{J}''$, this implies that $(F_1 \cdots F_a)^\ell \in \mathcal{J}''$. Using the Noetherian property of $\mathcal{R}$ we find some $\ell \geqslant 1$ with $\mathcal{J}(Z_1)^\ell \cdots \mathcal{J}(Z_a)^\ell \subseteq \mathcal{J}''$ and so $Z(\mathcal{J}'') \subseteq Z'' = Z_1 \cup \cdots \cup Z_a$.

For the opposite inclusion, fix some $i \in \{1, \dots, a\}$ and notice that by definition any $f \in \mathcal{J}''$ is of the form $f = \frac{p}{q}$ with

$$p \in \mathcal{J}' \subseteq \mathcal{J}(Z_i)$$

and

$$q \notin \mathcal{M} \supseteq \mathcal{J}(Z_i),$$

which gives $f \in \mathcal{J}(Z_i)$. This shows that $\mathcal{J}'' \subseteq \mathcal{J}(Z_i)$ and so $Z_i \subseteq Z(\mathcal{J}'')$ for $i = 1, \dots, a$.

We will show the claim by showing that $\mathcal{J}''$ is a prime ideal (which then gives the claim that $a = 1$ and $Z'' = Z$). For this we prove that $f \in \mathcal{J}''$ if and only if there exists a neighborhood O of $(x^{(0)}, y^{(0)})$ in

$$M = \{(x, \phi_1(x), \dots, \phi_{d-k}(x)) \mid x \in U\}$$

such that the restriction of f to O is zero. By the properties of $\phi_1, \dots, \phi_{d-k}$ any such restriction can be identified with an analytic function on a neighborhood of $x^{(0)}$ inside U, and so the restriction is uniquely determined by its

---

[†] A *local ring* is a ring with a unique maximal ideal. The ring $\mathcal{R}_{\mathcal{M}}$ is the *localization* of $\mathcal{R}$ at $\mathcal{M}$, it is a local ring with maximal ideal $\mathcal{M}\mathcal{R}_{\mathcal{M}}$.

Taylor expansion at $x^{(0)}$. Since the Cauchy product of Taylor series has no zero-divisors[†], this equivalence then shows that $\mathcal{J}''$ is a prime ideal.

Suppose first that $f \in \mathcal{J}''$. Then $f = \frac{p}{q}$, where $p \in \mathcal{J}'$ vanishes on M (by definition of M), and q does not vanish at $(x^{(0)}, y^{(0)})$. This shows that there is a neighborhood O on which f is well-defined and identical to zero.

Now suppose that f is a polynomial for which there exists a neighborhood O of $(x^{(0)}, y^{(0)})$ in M on which f vanishes. By our assumptions from the beginning of the proof we have $f_j(x,y) \in y_j - y_j^{(0)} + \mathcal{M}^2$ (where $\mathcal{M}^2$ consists of all polynomials that vanish with order 2 or more at $(x^{(0)}, y^{(0)})$). Let $n \geqslant 1$ be arbitrary. We can now use the polynomials f_j to express the polynomial f as above in the form

$$f \in \sum_{\ell=0}^{n} F_\ell(x - x^{(0)}) + R_{n+1}(x,y) + \mathcal{J}',$$

where F_ℓ is a homogeneous polynomial of degree ℓ for $\ell = 0, \dots, n$ and the polynomial $R_{n+1}(x,y) \in \mathcal{M}^{n+1}$ only has terms that vanish of order $n+1$ or higher. This shows that $\sum_{\ell=0}^{n} F_\ell(x - x^{(0)})$ is the Taylor approximation of $f(x, \phi_1(x), \dots, \phi_{d-k}(x))$ at $x^{(0)}$ of degree n. Since f vanishes in a neighborhood of $(x^{(0)}, y^{(0)})$ in M we have $\sum_{\ell=0}^{n} F_\ell = 0$. This shows that $f \in \mathcal{J}' + \mathcal{M}^n$ for all $n \geqslant 1$.

A corollary of Nakayama's lemma states that

$$\bigcap_{n=1}^{\infty} (\mathcal{I} + \mathcal{M}^n) = \mathcal{I}$$

in any local Noetherian ring with an ideal $\mathcal{I}$ and a maximal ideal $\mathcal{M}$. We refer to Hungerford [81, Cor. VIII.4.7] and Matsumura [129] for convenient sources for this result. Switching from the ring $\mathcal{R}$ to the local ring $\mathcal{R}_{\mathcal{M}}$ we see that

$$f \in \bigcap_{n=1}^{\infty} \left(\mathcal{J}'\mathcal{R}_{\mathcal{M}} + \mathcal{M}^n\mathcal{R}_{\mathcal{M}}\right),$$

which gives $f \in \mathcal{J}'\mathcal{R}_{\mathcal{M}}$ and so $f \in \mathcal{J}''$. This establishes the above equivalence, and hence shows that $\mathcal{J}''$ is a prime ideal, the claim, and so also the lemma. □

In Section 3.1 we considered two notions of '$\mathbb{K}$ varieties': A variety Z is *defined over* $\mathbb{F}$, for some subfield[‡] $\mathbb{F} \subseteq \overline{\mathbb{K}}$, if its complete ideal of relations (as in the Hilbert Nullstellensatz Theorem 3.6) is generated by polynomials with

---

[†] The Cauchy product of $\sum_{n=0}^{\infty} a_n x^n$ and $\sum_{n=0}^{\infty} b_n x^n$ is the series $\sum_{n=0}^{\infty} c_n x^n$ with coefficients $c_n = \sum_{j=0}^{n} a_j b_{n-j}$ for $n \geqslant 0$; viewed either as formal power series or as functions where they converge, the product will only vanish if one of the series vanishes.

[‡] We introduce this extra field for example in order to set $\mathbb{K} = \mathbb{R}$, $\overline{\mathbb{K}} = \mathbb{C}$, and $\mathbb{F} = \mathbb{Q}$.

coefficients in $\mathbb{F}$. On the other, a variety is $\mathbb{F}$-*closed* if it can be defined by polynomials with coefficients in $\mathbb{F}$.

As in any topological space, we can define a notion of *closure*: the Zariski closure of a subset $S \subseteq \overline{\mathbb{K}}^d$ is the smallest Zariski closed subset $Z \subseteq \overline{\mathbb{K}}^d$ containing S. This notion has many convenient properties, including good behaviour with regards to subfields. Note however, that the Zariski closure of a subset in $\mathbb{R}^d$ is frequently much bigger than the closure in the Hausdorff topology.

Lemma 3.19 (Closures of subsets of $\mathbb{F}^d$). *Let $\mathbb{F} \subseteq \overline{\mathbb{K}}$ be any subfield and $S \subseteq \mathbb{F}^d$. Then the Zariski closure of S is defined over $\mathbb{F}$.*

PROOF. Suppose that f is a polynomial in $x_1, \dots, x_d$ that vanishes on S. Let V be the vector space generated by the coefficients of f over $\mathbb{F}$. Let

$$a_1, \dots, a_n$$

be a basis of V over $\mathbb{F}$, and write

$$f = \sum_{i=1}^{n} f_i a_i$$

with $f_i \in \mathbb{F}[x_1, \dots, x_d]$. For any $x \in S$ we now have

$$f(x) = \sum_{i=1}^{n} \underbrace{f_i(x)}_{\in \mathbb{K}} a_i = 0,$$

and so $f_i(x) = 0$ for $i = 1, \dots, n$. This shows that the ideal of polynomials that vanish on S is generated by those that have coefficients in $\mathbb{F}$. □

Clearly a variety that is defined over $\mathbb{K}$ is also $\mathbb{K}$-closed. In general the converse is not true, but fortunately this problem only manifests itself over fields of positive characteristic.

Lemma 3.20 ($\mathbb{K}$-closed vs. defined over $\mathbb{K}$). *Suppose that $\mathbb{K}$ has characteristic zero. Then a $\mathbb{K}$-closed variety (or a variety that is stable under all Galois automorphisms of $\overline{\mathbb{K}}|\mathbb{K}$) is also defined over $\mathbb{K}$.*

PROOF. Let $Z = Z(f_1, \dots, f_n)$ be the variety defined by the polynomials

$$f_1, \dots, f_n \in \mathbb{K}[x_1, \dots, x_d],$$

and suppose that $f \in \overline{\mathbb{K}}[x_1, \dots, x_d]$ vanishes on Z (that is, suppose that f lies in $\mathcal{J}(Z)$). Then there exists a finite Galois field extension $\mathbb{L}|\mathbb{K}$ such that f has coefficients in $\mathbb{L}$.

Let σ be any Galois automorphism of the extension $\mathbb{L}|\mathbb{K}$. We now claim that the polynomial $\sigma(f)$ obtained by applying σ to all coefficients of f also belongs to $\mathcal{J}(Z)$. This is straightforward to check as follows. Since Z is $\mathbb{K}$-closed, any Galois automorphism of $\overline{\mathbb{K}}|\mathbb{K}$ maps $Z = Z(\overline{\mathbb{K}})$ onto Z. Extending the automorphism σ of $\mathbb{L}|\mathbb{K}$ in some way to an automorphism of $\overline{\mathbb{K}}|\mathbb{K}$ we get

$$(\sigma(f))\,(x) = (\sigma(f))\,\bigl(\sigma(\sigma^{-1}(x))\bigr) = \sigma\,\Bigl(\underbrace{f(\underbrace{\sigma^{-1}(x)}_{\in Z})}_{=0}\Bigr) = 0$$

for all $x \in Z$.

The claim now implies that $\operatorname{tr}(f) = \sum_\sigma \sigma(f)$, where the sum is taken over the finite list of Galois automorphisms of $\mathbb{L}|\mathbb{K}$, belongs to $\mathcal{J}(Z)$. Clearly $\operatorname{tr}(f)$ has coefficients in $\mathbb{L}$ and is fixed by all Galois automorphisms of $\mathbb{L}|\mathbb{K}$. Therefore, $\operatorname{tr}(f) \in \mathbb{K}[x_1, \dots, x_d]$ (this requires the assumption that $\operatorname{char}(\mathbb{K}) = 0$).

We next claim that there exist elements

$$a_1, \dots, a_{[\mathbb{L}|\mathbb{K}]} \in \mathbb{L}$$

and

$$a_1^*, \dots, a_{[\mathbb{L}|\mathbb{K}]}^* \in \mathbb{L}$$

that are dual bases in the sense that

$$\operatorname{tr}(a_i^* a_j) = \delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases}$$

for all i, j. We then have

$$a = \sum_i \operatorname{tr}(a_i^* a) a_i,$$

which also holds for the polynomial f instead of $a \in \mathbb{L}$. Since

$$\operatorname{tr}(a_i^* f) \in \mathcal{J}(Z) \cap \mathbb{K}[x_1, \dots, x_d]$$

by the argument above, the lemma follows from the claim.

It remains to construct the dual basis. Let $a_1, \dots, a_{[\mathbb{L}:\mathbb{K}]} \in \mathbb{L}$ be any basis of $\mathbb{L}$ over $\mathbb{K}$. By linear algebra, there exists a dual basis for the dual vector space $\mathbb{L}^*$ over $\mathbb{K}$. We claim that the map sending $a \in \mathbb{L}$ to $\phi(a) \in \mathbb{L}^*$ defined by

$$\phi(a)(b) = \operatorname{tr}(ab)$$

is an isomorphism of vector spaces. This may be seen as follows:

- $\phi(1)(1) = \operatorname{tr}(1) = [\mathbb{L} : \mathbb{K}]$, so ϕ is non-trivial (again since $\operatorname{char}(\mathbb{K}) = 0$);
- if $\phi(a) = 0$ then also $\phi(aa')(b) = \operatorname{tr}\left(a(a'b)\right) = 0$ for all $a', b \in \mathbb{L}$, so the kernel of ϕ is an ideal, and the field $\mathbb{L}$ has no non-trivial ideals.

Thus the pre-image under ϕ of the dual basis in $\mathbb{L}^*$ gives a dual basis in the above sense in $\mathbb{L}$.

If the variety Z is only assumed to be invariant under all Galois automorphisms, then once more $\mathcal{J}(Z)$ is invariant under all Galois automorphisms and so the above argument shows again that Z is defined over $\mathbb{K}$. $\square$

In the arguments above there is always an implied coordinate system in $\overline{\mathbb{K}}^d$ (corresponding to the variables $x_1, \dots, x_d$). We note that it is customary to write $\mathbb{A}^d$ for the d-dimensional affine space without a preferred origin, coordinate system, or base field (so that $\mathbb{A}^d(\mathbb{L}) \simeq \mathbb{L}^d$ for any field $\mathbb{L}$). For us the ambient affine space will be $\mathrm{Mat}_d \simeq \mathbb{A}^{d^2}$, and on this space very few coordinate changes make sense with regards to the existing (and to us important) multiplicative structure. For that reason and also because we are often interested in subgroups of SL_d (and the orbits of the group of their $\mathbb{R}$-points), we are happy with choosing one coordinate system and discussing subvarieties and algebraic subgroups of SL_d instead of general varieties and general algebraic groups. We will however, switch frequently from one field to another, and as before will write $Z(\mathbb{K}) = Z(\overline{\mathbb{K}}) \cap \mathrm{Mat}_d(\mathbb{K})$ for the $\mathbb{K}$-points of a subvariety $Z < \mathrm{Mat}_d$ defined over $\mathbb{K}$.

3.4.3 Linear Algebraic Groups

A variety $\mathbb{G} \subseteq \mathrm{SL}_d$ is a *(linear) algebraic subgroup* (of SL_d) if $\mathbb{G}(\overline{\mathbb{K}}) \subseteq \mathrm{SL}_d(\overline{\mathbb{K}})$ is a subgroup. Notice that for any subvariety $Z \subseteq \mathrm{SL}_d$ and $g \in \mathrm{SL}_d(\overline{\mathbb{K}})$ we can define the translated variety gZ by the ideal

$$\lambda(g)\mathcal{J}(Z) = \{f(g^{-1}x) \mid f \in \mathcal{J}(Z)\}.$$

Here λ denotes the left representation

$$\lambda(g)f(x) = f(g^{-1}x)$$

on the space of all polynomials.

Lemma 3.21 (Smoothness). *Every point of a linear algebraic subgroup is smooth.*

The tangent space at the identity is called the *Lie algebra* of the algebraic subgroup.

Proof of Lemma 3.21. Suppose that $g \in \mathbb{G}(\overline{\mathbb{K}})$ is a smooth point of the variety $\mathbb{G}$. Then one can quickly check that $I = g^{-1}g$ is a smooth point of the left-translate variety $g^{-1}\mathbb{G}$. However, since $g^{-1}\mathbb{G} = \mathbb{G}$ we see that I is a smooth point of $\mathbb{G}$. By the same argument, any other point is also smooth. $\square$

Lemma 3.22 (Connected components). *Let $\mathbb{G} \subseteq \mathrm{SL}_d$ be an algebraic subgroup. The connected component $\mathbb{G}^o < \mathbb{G}$ is by definition the unique*

Zariski connected component of $\mathbb{G}$ *that contains the identity, it is an algebraic normal subgroup. There are points* $g_1, \dots, g_n \in \mathbb{G}$ *where* $n = [\mathbb{G}(\overline{\mathbb{K}}) : \mathbb{G}^o(\overline{\mathbb{K}})]$ *with*

$$\mathbb{G} = \bigsqcup_{i=1}^{n} g_i \mathbb{G}^o.$$

If $\mathbb{G}$ *is defined over* $\mathbb{K}$*, and* $\mathbb{K}$ *has zero characteristic, then* $\mathbb{G}^o$ *is also defined over* $\mathbb{K}$*.*

As a corollary of the lemma we mention that it makes sense to talk about the dimension of a (not necessarily Zariski connected) algebraic subgroup. Since all Zariski connected components are translates of the connected component $\mathbb{G}^o$, they all have the same dimension.

Proof of Lemma 3.22. The first statement is essentially an extension of the argument in the previous lemma. If

$$\mathbb{G} = \bigcup_{i=1}^{n} Z_i$$

is the decomposition into connected components, then there exists a point which is contained in only one component. Translating $\mathbb{G}$ by one of its elements $g \in \mathbb{G}(\overline{\mathbb{K}})$ we may permute the connected components

$$\mathbb{G} = g^{-1}\mathbb{G} = \bigcup_{i=1}^{n} g^{-1} Z_i.$$

but would leave the subvariety $\bigcup_{i \neq j} Z_i \cap Z_j$ consisting of all points that are contained in more than one of the connected components invariant. Therefore, we have

$$\mathbb{G} = \bigsqcup_{i=1}^{n} Z_i.$$

Suppose that $Z_1 = \mathbb{G}^o$. If now $g \in \mathbb{G}^o$ then $I \in g^{-1}\mathbb{G}^o$, which by uniqueness of the decomposition gives $\mathbb{G}^o = g^{-1}\mathbb{G}^o$ for all $g \in \mathbb{G}^o$.

We have shown that $\mathbb{G}^o$ is a linear algebraic subgroup. If now $g \in Z_i$ for $i > 1$, then the same argument gives $g^{-1}Z_i = \mathbb{G}^o = Z_i g^{-1}$. In other words,

$$Z_i = g\mathbb{G}^o = \mathbb{G}^o g$$

is a coset of $\mathbb{G}^o$ in $\mathbb{G}$.

Now suppose that $\mathbb{G}$ is defined over $\mathbb{K}$, and let σ be a Galois automorphism of $\overline{\mathbb{K}}|\mathbb{K}$. Then σ induces a permutation of the cosets $g_i\mathbb{G}^o(\overline{\mathbb{K}})$ with

$$\sigma\left(\mathbb{G}^o(\overline{\mathbb{K}})\right) = \mathbb{G}^o(\overline{\mathbb{K}})$$

since $\sigma(I) = I$. As this holds for all Galois automorphisms we see that $\mathbb{G}^o$ is defined over $\mathbb{K}$ if $\mathbb{K}$ has characteristic zero by Lemma 3.20. □

For completeness we mention another (more general but, up to isomorphisms, equivalent) definition: A *linear algebraic group* is an affine variety equipped with multiplication and inverse maps such that

- the multiplication and inverse maps are regular functions (from the group to the group);
- the variety is isomorphic to a linear algebraic subgroup of SL_d for some d such that the multiplication and inverse maps correspond to multiplication and inversion for matrices.

We note that the standard definition does not make the second requirement above, and instead derives this property from the first via a construction similar to the proof of Chevalley's theorem in Section 3.4.5.

Example 3.23. We list some standard examples of linear algebraic groups.

(a) $\mathbb{G}_a$ denotes the additive group structure of the field. This is a linear algebraic group because (for example) it is isomorphic to the algebraic subgroup $U < \mathrm{SL}_2$ with

$$U(\overline{\mathbb{K}}) = \left\{ \begin{pmatrix} 1 & * \\ & 1 \end{pmatrix} \mid x \in \overline{\mathbb{K}} \right\},$$

which we saw earlier is associated to the horocycle flow if $\mathbb{K} = \mathbb{R}$.

(b) $\mathbb{G}_m$ stands for the multiplicative group structure of the field. This is a linear algebraic group because (for example) it is isomorphic to the algebraic subgroup $A < \mathrm{SL}_2$ with

$$A(\overline{\mathbb{K}}) = \left\{ \begin{pmatrix} a & \\ & a^{-1} \end{pmatrix} \mid a \in \overline{\mathbb{K}} \right\},$$

which we saw earlier is associated to the geodesic flow if $\mathbb{K} = \mathbb{R}$.

3.4.4 $\mathbb{K}$-points of Linear Algebraic Groups

As noted before, a variety Z defined over a field $\mathbb{K}$ does not have to contain any $\mathbb{K}$-points (that is, $Z(\mathbb{K})$ may be empty[†]), and even if it is non-empty it may not be Zariski dense in the variety. Since a subgroup always contains

---

[†] A trivial example to have in mind here is the variety defined by the equation

$$x^2 + y^2 = -1,$$

defined over $\mathbb{R}$, and a less trivial example is the variety defined by the equation

$$x^3 + y^3 = 1,$$

defined over $\mathbb{Q}$.

the identity the former problem cannot arise for linear algebraic subgroups. Even more is true, as a result of the following lemma, which relies on the fact that $\mathbb{G}$ is smooth at the identity.

Lemma 3.24 (Density of $\mathbb{R}$-points and $\mathbb{Q}_p$-points). *If $\mathbb{G} \subseteq \mathrm{SL}_d$ is a Zariski connected linear algebraic subgroup defined over $\mathbb{R}$, then $\mathbb{G}(\mathbb{R})$ is Zariski dense in $\mathbb{G}$. The same holds for $\mathbb{K} = \mathbb{Q}_p$ for a prime number $p < \infty$.*

We note that the above holds much more generally, see [8, Th. 18.3]. We will come back to this problem for the special case $\mathbb{K} = \mathbb{Q}$ later.

Proof of Lemma 3.24. By Lemma 3.21, $x^{(0)} = I \in \mathbb{G}$ is a smooth point. By Lemma 3.18 $\mathbb{G}(\mathbb{R})$ contains the image of an analytic function of the form

$$\Phi\colon U \ni (x_1, \dots, x_k) \longmapsto (x_1, \dots, x_k, \phi_{k+1}(x_1, \dots, x_k), \dots, \phi_{d^2}(x_1, \dots, x_k)) . \tag{3.23}$$

Let Z be the Zariski closure of these real points. By Lemma 3.17 we may write

$$Z = \bigcup_{i=1}^{n} Z_i$$

as a union of irreducible varieties. By Lemma 3.15, either $Z = \mathbb{G}$ or

$$\dim Z_i < k = \dim \mathbb{G}$$

for $i = 1, \dots n$. However, the latter case cannot happen since a finite union of varieties of dimension strictly less than k cannot contain all points

$$\Phi(x_1, \dots, x_k)$$

in (3.23). Specifically, in this case each $\mathcal{J}(Z_i)$ must contain some non-zero polynomial

$$f_i \in \mathbb{C}[x_1, \dots, x_k],$$

so that every point $\Phi(x_1, \dots, x_k)$ in (3.23) would have to satisfy the equation $f_1 \cdots f_n = 0$. This is a contradiction, since every non-empty open subset of $\mathbb{R}^k$ in the Hausdorff topology is Zariski dense (since all the partial derivatives, including the 0th, of a polynomial at a point determine the polynomial). The p-adic case is similar. □

Clearly, the group $\mathbb{G}(\mathbb{R})$ of $\mathbb{R}$-points of an algebraic subgroup $\mathbb{G} \subseteq \mathrm{SL}_d$ is a linear Lie group with a real Lie algebra $\mathfrak{g}_{\mathbb{G}(\mathbb{R})}$. For the algebraic subgroup $\mathbb{G}$ in SL_d we have already defined a Lie algebra $\mathfrak{g}$ which, by definition, is a complex vector space. Assuming that $\mathbb{G}$ is defined over $\mathbb{R}$, this complex vector space

$$\mathfrak{g} \subseteq \mathfrak{sl}_d(\mathbb{C})$$

can be defined by linear equations with real coefficients, so is invariant under complex conjugation and, in particular,

$$\mathfrak{g}(\mathbb{R}) = \mathfrak{g} \cap \mathfrak{sl}_d(\mathbb{R})$$

has the same dimension over $\mathbb{R}$ as $\mathfrak{g} = \mathfrak{g}(\mathbb{C})$ has over $\mathbb{C}$.

Lemma 3.25 (Lie algebras of Lie groups and algebraic groups). *Let $\mathbb{G} \subseteq \mathrm{SL}_d$ be an algebraic subgroup defined over $\mathbb{R}$. Then the $\mathbb{R}$-points*

$$\mathfrak{g}(\mathbb{R}) = \mathfrak{g} \cap \mathfrak{sl}_d(\mathbb{R})$$

of the Lie algebra $\mathfrak{g}$ of the algebraic subgroup comprise precisely the Lie algebra of the Lie subgroup $\mathbb{G}(\mathbb{R}) \subseteq \mathrm{SL}_d(\mathbb{R})$. The same holds over $\mathbb{C}$ or $\mathbb{Q}_p$ for a prime $p < \infty$.

PROOF. Using the same notation and setup as in the proofs of Lemmas 3.18 (with d replaced by d^2) and Lemma 3.24, we see that the tangent space of $\mathbb{G}(\mathbb{R})$ (in the sense of manifolds or of Lie groups) is the image of the total derivative of Φ at $(x_1, \ldots, x_k)$. However, by the implicit function theorem, this image is precisely the real subspace defined by the equations

$$(u_1, \ldots, u_{d^2}) \cdot \nabla f_j(I) = 0$$

for $j = 1, \ldots, d^2 - k$. This proves the lemma in the real case, and the complex and p-adic cases are similar. □

The following discussion is not essential for later developments, but it may be useful to bear it in mind. By [170, Ch. VII, Sect. 2.2, Th. 1] the set of $\mathbb{C}$-points $Z(\mathbb{C})$ of a Zariski connected variety Z is connected in the Hausdorff topology. For the $\mathbb{R}$-points $Z(\mathbb{R})$ of a Zariski connected variety Z over $\mathbb{R}$ this is not true. However, for algebraic groups $\mathbb{G}$ defined over $\mathbb{R}$, the connected component $\mathbb{G}(\mathbb{R})^o$ (in the Hausdorff topology) only has finite index. We will discuss this again for particular algebraic subgroups later (where it will usually be easy to see). For now, notice that $A(\mathbb{R})^o < A(\mathbb{R})$ for A as in Example 3.23(b) has index two. Over $\mathbb{Q}_p$ the analogous question does not make sense, so Zariski connected is *a priori* the only sensible notion of connectedness.

3.4.5 Chevalley's theorem subgroups and representations

Clearly, every algebraic representation gives rise to many algebraic subgroups by defining stabilizer subgroups (as in Section 3.1.2). Chevalley's theorem(15) almost turns this construction around: Given an algebraic subgroup there exists an algebraic representation so that the subgroup can be defined via the representation as a stabilizer of a line (instead of a point as in Section 3.1.2).

Theorem 3.26 (Chevalley). *Let $\mathbb{H} < \mathrm{SL}_d$ be an algebraic subgroup. Then there exists an algebraic representation $\rho : \mathrm{SL}_d \to \mathrm{SL}_D$ and a D-dimensional* *vector v such that*

$$\mathbb{H} = \{g \in \mathrm{SL}_d \mid \rho(g)v \sim v\},$$

where $\sim$ denotes proportionality[†]*. If H is defined*[‡] *over $\mathbb{K}$, then the algebraic representation ρ is also defined over $\mathbb{K}$, and we may choose $v \in \mathbb{K}^D$.*

As we will see, the theorem is proved by transforming the defining ideal of $\mathbb{H}$ (which is finitely-generated) into a single vector in a high-dimensional vector space.

Proof of Theorem 3.26. For any $g \in \mathbb{H}$ we have $g\mathbb{H} = \mathbb{H}$ and equivalently

$$\lambda(g)\mathcal{J}(\mathbb{H}) = \mathcal{J}(\mathbb{H}).$$

Moreover, we also have that $\lambda(g)\mathcal{J}(\mathbb{H}) = \mathcal{J}(\mathbb{H})$ for some $g \in \mathrm{SL}_d$ implies that $g \in \mathbb{H}$. As the ideal is infinite dimensional we cannot use it directly. However, by the Noetherian property we know that $\mathcal{J}(\mathbb{H}) \subseteq \overline{\mathbb{K}}[\mathrm{Mat}_d]$ is finitely generated (as an ideal). Thus we can assume it is generated by polynomials of degree less than or equal to m for some m. Write $\mathcal{P}_{\leqslant m}$ for the space of all polynomials in $\overline{\mathbb{K}}[\mathrm{Mat}_d]$ of degree $\leqslant m$, and define

$$\mathcal{J}_{\leqslant m} = \mathcal{J}(\mathbb{H}) \cap \mathcal{P}_{\leqslant m}.$$

Now notice that $\lambda(g)\mathcal{P}_{\leqslant m} = \mathcal{P}_{\leqslant m}$ for all $g \in \mathrm{SL}_d$ and that

$$\lambda(g)\mathcal{J}_{\leqslant m} = \mathcal{J}_{\leqslant m}$$

is equivalent to $g \in \mathbb{H}$ (since $\mathcal{J}_{\leqslant m}$ generates $\mathcal{J}(\mathbb{H})$). In other words, we have found a finite-dimensional representation of SL_d and a subspace so that $\mathbb{H}$ is precisely the subgroup of SL_d that sends the subspace into itself. The representation is also an algebraic representation (which the reader can quickly check).

What is not quite as in the theorem is that the subspace might not be a single line. However, even that can quickly be rectified. Let $\ell = \dim \mathcal{J}_{\leqslant m}$ and define $V = \bigwedge^\ell \mathcal{P}_{\leqslant m}$ and let $v \in \bigwedge^\ell \mathcal{J}_{\leqslant m} \smallsetminus \{0\}$. The algebraic representation of SL_d on $\mathcal{P}_{\leqslant m}$ induces an algebraic representation ρ on V (check this) and for any $g \in \mathrm{SL}_d$ the condition $\rho(g)v \sim v$ is equivalent to $\lambda(g)\mathcal{J}_{\leqslant m} = \mathcal{J}_{\leqslant m}$ and hence to $g \in \mathbb{H}$.

If $\mathbb{H}$ is now additionally defined over $\mathbb{K}$, then $\mathcal{J}_{\leqslant m} \cap \mathbb{K}[\mathrm{Mat}_d]$ generates $\mathcal{J}(\mathbb{H})$ and we can choose v as the wedge of ℓ elements in $\mathcal{J}_{\leqslant m} \cap \mathbb{K}[\mathrm{Mat}_d]$. Since the regular representation (and its ℓth wedge power) are defined over any field, this proves the last claim of the theorem. □

[†] Notice that proportionality is itself a polynomial condition, defined by requiring the vanishing of all 2×2 determinants corresponding to pairs of components of $\rho(g)v$ and of v.

[‡] That is, if the ideal of relations that are satisfied on $\mathbb{H}$ is generated by polynomials with coefficients in $\mathbb{K}$.

Lemma 3.27 (Zariski closures of groups). *If $S \subseteq \mathrm{SL}_d(\mathbb{K})$ is a subgroup, then the Zariski closure $\mathbb{G} = \overline{S}^Z$ is a linear algebraic subgroup defined over $\mathbb{K}$.*

Proof. By Lemma 3.19 we know that $\mathbb{G}$ is defined over $\mathbb{K}$, so it is enough to show that $\mathbb{G}(\overline{\mathbb{K}})$ is a subgroup.

For any $g \in S$ we have $gS = S$ by the assumption on S, so $g\mathbb{G} = \mathbb{G}$ for all $g \in S$. However, as in the proof of Theorem 3.26 this property of preserving the variety is equivalent to the property of preserving the ideal $\mathcal{J}(\mathbb{G})$ of relations defining $\mathbb{G}$ or equivalently a particular line inside an algebraic representation of SL_d.

As this is a polynomial condition (see one of the footnotes to Theorem 3.26) which holds for all $g \in S$, it must also hold for all $g \in \mathbb{G}$. In other words, we have shown that $g\mathbb{G} = \mathbb{G}$ also holds for g in the Zariski closure of S, that is for all $g \in \mathbb{G} = \mathbb{G}(\overline{\mathbb{K}})$. □

3.4.6 Jordan Decomposition, Algebraic Subgroups and Representations

Algebraic groups and algebraic representations have some striking differences to the theory of Lie groups, which we will now start to discuss.

Let ρ be an algebraic representation of SL_d (or more generally of an algebraic subgroup $\mathbb{H}$). Then we have the following facts:

- if $u \in \mathrm{SL}_d$ ($u \in \mathbb{H}$) is nilpotent, then so is $\rho(u)$;
- if $a \in \mathrm{SL}_d(\mathbb{R})$ ($a \in \mathbb{H}$) is diagonalizable (when considered as an element $a \in \mathrm{SL}_d$) and has only real and positive eigenvalues, then the same holds for $\rho(a)$.

The first property is readily proved for the case SL_d and $\mathbb{K} = \mathbb{Q}$ or $\mathbb{K}$ a local field. Indeed, if $u \in \mathrm{SL}_d(\overline{\mathbb{K}})$ is unipotent, then there exists some $a \in \mathrm{SL}_d(\overline{\mathbb{K}})$ with $a^n u a^{-n} \to I$ as $n \to \infty$, which implies that

$$\rho(a)^n \rho(u) \rho(a)^{-n} = \rho(a^n u a^{-n}) \longrightarrow I$$

as $n \to \infty$, so the eigenvalues of $\rho(u)$ (which are not changed by conjugation) must all be 1, and hence $\rho(u)$ must be unipotent.

The second property requires a bit more work. We also note that if the algebraic representation is only defined on the subgroup $\mathbb{H}$ then neither claim would be correct in the context of Lie theory. For this notice that the Lie groups $U(\mathbb{R})$ and $A(\mathbb{R})$ are not that much different. On the one hand, the former is connected and the latter is not, so they are not isomorphic. However, there is a surjective group homomorphism from $A(\mathbb{R})$ onto $U(\mathbb{R})$, and an injective homomorphism from $U(\mathbb{R})$ into $A(\mathbb{R})^o < A(\mathbb{R})$. This does not contradict the above claims, since the two maps are basically the logarithm and the exponential map, which are not algebraic homomorphisms.

Recall from linear algebra that every matrix $g \in \mathrm{SL}_d(\overline{\mathbb{K}})$ has a *Jordan decomposition*

$$g = g_{\mathrm{ss}} g_{\mathrm{u}}$$

into a $\overline{\mathbb{K}}$-*diagonalizable* or *semi-simple* matrix $g_{\mathrm{ss}} \in \mathrm{SL}_d(\overline{\mathbb{K}})$ and a *unipotent* $g_{\mathrm{u}} \in \mathrm{SL}_d(\overline{\mathbb{K}})$. The two components g_{ss} and g_{u} commute with each other, and under this requirement the decomposition is unique. If $\mathbb{K}$ is $\mathbb{R}$ or $\mathbb{C}$, then $g_{\mathrm{ss}} = g_{\mathrm{pos}} g_{\mathrm{comp}}$ can be further decomposed into a product of two commuting semi-simple elements $g_{\mathrm{pos}}, g_{\mathrm{comp}} \in \mathrm{SL}_d(\mathbb{C})$, where the *positive semi-simple* part g_{pos} has only real and positive eigenvalues, and all the eigenvalues of the *compact semi-simple* part g_{comp} have absolute value one. This decomposition is also unique, and if $g \in \mathrm{SL}_d(\mathbb{R})$ then $g_{\mathrm{u}}, g_{\mathrm{pos}}, g_{\mathrm{comp}}$ lie in $\mathrm{SL}_d(\mathbb{R})$ (see Exercise 3.4.1). If $\mathbb{K} = \mathbb{Q}_p$, then a similar decomposition can be shown, and the following results hold in that case also (see Exercise 3.4.2).

The following two results contain the claims made in the beginning of this section in greater generality.

Proposition 3.28 (Jordan decomposition and subgroups). *Let $\mathbb{H}$ be an algebraic subgroup of SL_d, and let g be an element of $\mathbb{H}$. If $g = g_{\mathrm{ss}} g_{\mathrm{u}}$ is the Jordan decomposition of g in $\mathrm{SL}_d(\overline{\mathbb{K}})$, then $g_{\mathrm{ss}}, g_{\mathrm{u}} \in \mathbb{H}$ also. If $\mathbb{H}$ is defined over $\mathbb{K} = \mathbb{R}$ (or $\mathbb{K} = \mathbb{C}$) and $g_{\mathrm{ss}} = g_{\mathrm{pos}} g_{\mathrm{comp}}$ is the decomposition into positive semi-simple and compact semi-simple parts, then once again $g_{\mathrm{pos}}, g_{\mathrm{comp}} \in \mathbb{H}$.*

Proposition 3.29 (Jordan decomposition and representations). *Let $\mathbb{H}$ be an algebraic subgroup of SL_d, and let $\rho\colon \mathbb{H} \to \mathrm{GL}_D$ be an algebraic representation. Then $\rho(g)_{\mathrm{u}} = \rho(g_{\mathrm{u}})$ and $\rho(g)_{\mathrm{ss}} = \rho(g_{\mathrm{ss}})$ for all $g \in \mathbb{H}$. If $\mathbb{K}$ is $\mathbb{R}$ or $\mathbb{C}$, then we also have $\rho(g)_{\mathrm{pos}} = \rho(g_{\mathrm{pos}})$ and $\rho(g)_{\mathrm{comp}} = \rho(g_{\mathrm{comp}})$.*

The proof of these results is intertwined. We will first prove Proposition 3.29 in a special case, then prove Proposition 3.28, and finally obtain Proposition 3.29 as a corollary.

PROOF OF PROPOSITION 3.29 FOR A CHEVALLEY REPRESENTATION. Suppose that ρ is the representation of SL_d obtained in the proof of Theorem 3.26 for a subgroup $\mathbb{H} \leqslant \mathrm{SL}_d$. Let $g = g_{\mathrm{ss}}$ be semi-simple, and assume (without loss of generality, by applying any necessary conjugation to $\mathbb{H}$ and g) that g is diagonal. Then it is easy to see[†] that $\lambda(g)$ restricted to $\mathcal{P}_{\leqslant m}$ is diagonal, with eigenvectors given by monomials in the standard variables. Therefore all eigenvalues of $\lambda(g)$ are simply products of powers of eigenvalues of g. Taking the ℓth wedge representation, the same holds for $\rho(g) = \wedge^{\ell} \lambda(g)$. Let $g = g_u$ be unipotent. If $\mathbb{K}$ is $\mathbb{Q}$ or a local field (which is where our main interest lies), we have already shown that $\rho(g)$ is unipotent. In general we may argue again step by step as above. First, show that $\lambda(g)$ restricted to $\mathcal{P}_{\leqslant m}$ is unipotent by considering monomials corresponding to the eigendirections

---

[†] We use the notation from the proof of Theorem 3.26.

(resp. generalized eigendirections). Then we can show that $\rho(g) = \wedge^{\ell}\lambda(g)$ is also unipotent.

If now $g = g_{\mathrm{ss}}g_{\mathrm{u}}$ is any element of SL_d, then $\rho(g_{\mathrm{ss}})$ is semi-simple, $\rho(g_{\mathrm{u}})$ is unipotent, $\rho(g) = \rho(g_{\mathrm{ss}})\rho(g_{\mathrm{u}})$, and $\rho(g_{\mathrm{ss}})$, $\rho(g_{\mathrm{u}})$ commute with each other. This proves the claim.

If $\mathbb{K}$ is $\mathbb{R}$ or $\mathbb{C}$, then the argument above also shows that the eigenvalues of $\rho(g_{\mathrm{pos}})$ are positive and the eigenvalues of $\rho(g_{\mathrm{comp}})$ have absolute value one, giving the theorem. □

PROOF OF PROPOSITION 3.28. Let $\mathbb{H} \leqslant \mathrm{SL}_d$ be an algebraic subgroup and let $\rho, v \in \overline{\mathbb{K}}^D$ be as in Theorem 3.26. Let $g \in \mathbb{H}$ so that $v \in \overline{\mathbb{K}}^D$ is an eigenvector of $\rho(g)$ for the Chevalley representation. By the properties of the Jordan decomposition, v is therefore also an eigenvector of $\rho(g)_{\mathrm{ss}} = \rho(g_{\mathrm{ss}})$ and of $\rho(g)_{\mathrm{u}} = \rho(g_{\mathrm{u}})$. It follows that $g_{\mathrm{ss}}, g_{\mathrm{u}} \in \mathbb{H}$. If $\mathbb{K}$ is $\mathbb{R}$ or $\mathbb{C}$, and $g_{\mathrm{ss}} = g_{\mathrm{pos}}g_{\mathrm{comp}}$ then $\rho(g)_{\mathrm{pos}} = \rho(g_{\mathrm{pos}})$ has v as an eigenvalue. Thus $g_{\mathrm{pos}}, g_{\mathrm{comp}} \in \mathbb{H}$ as well. □

PROOF OF PROPOSITION 3.29. Let $\mathbb{H} \leqslant \mathrm{SL}_d$ and let $\rho\colon \mathbb{H} \to \mathrm{GL}_D$ be an arbitrary algebraic representation. Then

$$\mathbb{L} = \mathrm{Graph}(\rho) \subseteq \mathbb{H} \times \mathrm{GL}_D \subseteq \mathrm{SL}_{d+D+1}$$

is an algebraic subgroup in the following way. We require the elements of $\mathbb{L}$ to be of block form

$$\begin{pmatrix} h & & \\ & g & \\ & & \det(g)^{-1} \end{pmatrix}$$

with $h \in \mathrm{SL}_d$ and $g \in \mathrm{GL}_D$ (by using linear equations, the condition that $\det h = 1$, and the polynomial equation that the last entry should be the inverse of the determinant of the middle block), require $h \in \mathbb{H}$ (by the known relations of $\mathbb{H}$), and finally $g = \rho(h)$ (which is a polynomial condition by assumption on ρ).

Now let $h = h_{\mathrm{ss}} \in \mathbb{H}$ be semi-simple, so that

$$g = \begin{pmatrix} h & & \\ & \rho(h) & \\ & & \det(\rho(h))^{-1} \end{pmatrix} \in \mathbb{L}$$

and hence by Proposition 3.28 we also have

$$g_{\mathrm{u}} = \begin{pmatrix} h_{\mathrm{u}} & & \\ & \rho(h)_{\mathrm{u}} & \\ & & 1 \end{pmatrix} \in \mathbb{L}.$$

However, since $h_{\mathrm{u}} = I_d$ and $\mathbb{L}$ is a graph of a homomorphism, we also have $\rho(h)_{\mathrm{u}} = I_D$. This shows that $\rho(h)$ is semi-simple if h is semi-simple.

The same argument also applies to unipotent elements (respectively, to positive or compact semi-simple elements if $\mathbb{K}$ is $\mathbb{R}$ or $\mathbb{C}$). The proposition follows from the uniqueness of the Jordan decomposition. □

Exercises for Section 3.4

Exercise 3.4.1. Let $g \in \mathrm{SL}_d(\mathbb{C})$ be diagonalizable. Define two commuting matrices

$$g_{\mathrm{pos}}, g_{\mathrm{comp}} \in \mathrm{SL}_d(\mathbb{C})$$

with $g = g_{\mathrm{pos}} g_{\mathrm{comp}}$ such that g_{pos} only has positive eigenvalues and g_{comp} only has eigenvalues of absolute value one. Show that these are also uniquely determined by these properties, and that $g \in \mathrm{SL}_d(\mathbb{R})$ implies that $g_{\mathrm{pos}}, g_{\mathrm{comp}} \in \mathrm{SL}_d(\mathbb{R})$.

Exercise 3.4.2. Let $\mathbb{K} = \mathbb{Q}_p$ and let $P < \overline{\mathbb{Q}}_p^{\times}$ be a subgroup isomorphic to $\mathbb{Q}$ and containing p. Show that every matrix $g \in \mathrm{SL}_d(\overline{\mathbb{K}})$ is the product of commuting elements $g_{\mathrm{pos}}, g_{\mathrm{comp}} \in \mathrm{SL}_d(\overline{\mathbb{K}})$ where the eigenvalues of g_{pos} are elements of P, and the eigenvalues of g_{comp} have absolute value one. Generalize the results of Section 3.4.6 to include this p-adic case.

3.5 Borel Density Theorem

We will show in this section a version of the Borel density theorem,(16) which will show another relationship between finite volume orbits and rationally defined subgroups. It is the generalization of the basic observation that a lattice $\Lambda < \mathbb{R}^d$ cannot be contained in a proper subspace to the setting of lattices in linear algebraic groups.

For the proof we will need two basic theorems, each of them fundamental to its own subject. However, the two subjects concerned are often — in the context of this book, wrongly — considered far from each other. Concretely, we will need Poincaré recurrence from ergodic theory (in some sense the pigeonhole principle for ergodic theory, see Theorem 1.8 and Exercise 1.1.7), and Chevalley's theorem from the theory of algebraic groups (see Theorem 3.26), and will combine these with the facts derived in Section 3.4.6. This approach goes back to work of Furstenberg [68] and Dani [19].

Theorem 3.30 (Borel density theorem). *Suppose that $\mathbb{H} < \mathrm{SL}_d$ is an algebraic subgroup defined over $\mathbb{R}$ and suppose that $\Gamma < \mathbb{H}(\mathbb{R})$ is a lattice. Then*

(1) *If $\mathbb{H}$ is semi-simple*[†] *such that $\mathbb{H}(\mathbb{R})^o$ has no compact factors then Γ is Zariski dense in $\mathbb{H}$. If $\mathbb{H}$ is only assumed to be semi-simple then the Zariski*

[†] A linear algebraic group $\mathbb{H}$ is semi-simple if it is Zariski connected and its Lie algebra is semi-simple. Notice that this does not imply that $\mathbb{H}(\mathbb{R})$ is connected as a manifold.

closure of Γ contains all non-compact factors of $\mathbb{H}(\mathbb{R})^o$ (and possibly some or all of the compact factors).

(2) *In the general case, the Zariski closure $\mathbb{L} < \mathbb{H}$ of Γ contains all unipotent elements of $\mathbb{H}(\mathbb{R})$ and more generally all elements of $\mathbb{H}(\mathbb{R})$ that only have positive real eigenvalues.*

For the proof we will also need the following simple observation from linear algebra.

Lemma 3.31 (Convergence to some eigenvector). *Let $g \in \mathrm{SL}_d(\mathbb{R})$ have the property that all its eigenvalues are real and positive, and let*

$$\rho : \mathrm{SL}_d(\mathbb{R}) \to \mathrm{SL}_D(\mathbb{R})$$

be a finite-dimensional algebraic representation (obtained, for example, from Chevalley's theorem). Then for any $w \in \mathbb{R}^D \smallsetminus \{0\}$ there is some $v \in \mathbb{R}^D$ with

$$\frac{1}{\|\rho(g^n)w\|}\rho(g^n)w \longrightarrow v \in \mathbb{R}^D$$

as $n \to \infty$, and v is an eigenvector of $\rho(g)$.

Proof. By Proposition 3.28 if g is unipotent then $\rho(g)$ is also, and if g has only positive eigenvalues then the same holds for $\rho(g)$. Given $w \in \mathbb{R}^D \smallsetminus \{0\}$, we may write

$$w = \sum_{\lambda > 0} w_\lambda \neq 0,$$

where each w_λ is a generalized eigenvector for the eigenvalue λ and the map $\rho(g)$. Then there is some largest eigenvalue λ_L with $w_{\lambda_L} \neq 0$ (and hence $w_\lambda = 0$ for any $\lambda > \lambda_L$). Also notice that $\|\rho(g^n)w_\lambda\|$ is asymptotic to $\lambda^n n^{k(\lambda)}$ for some $k(\lambda) \geqslant 0$ (this may be seen by looking at the Jordan normal form of $\rho(g)$, see also the argument below). Thus

$$\frac{1}{\|\rho(g^n)w\|}\rho(g^n w) - \frac{1}{\|\rho(g^n)w_{\lambda_L}\|}\rho(g^n w_{\lambda_L}) \longrightarrow 0$$

as $n \to \infty$. This reduces the problem to the case of a single eigenvalue, and hence (by canceling the eigenvalue) to the case of a unipotent matrix

$$A = \frac{1}{\lambda_L}\rho(g)|_{V_{\lambda_L}}$$

acting on the generalized eigenspace V_{λ_L} of $\rho(g)$ for the eigenvalue λ_L. Choosing a Jordan basis of A, we may assume that A is a block matrix

$$A = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_\ell \end{pmatrix}$$

where each

$$A_i = \begin{pmatrix} 1 & 1 & & \\ & \ddots & \ddots & \\ & & 1 & 1 \\ & & & 1 \end{pmatrix}.$$

We split $w = w_{\lambda_L}$ into components $\sum_i w^{(i)}$ corresponding to this block decomposition, and apply A_i to the vector

$$w^{(i)} = \begin{pmatrix} w_1^{(i)} \\ \vdots \\ w_k^{(i)} \end{pmatrix}$$

to obtain

$$A_i^n \begin{pmatrix} w_1^{(i)} \\ \vdots \\ w_k^{(i)} \end{pmatrix} = \begin{pmatrix} w_1^{(i)} + w_2^{(i)} n + w_3^{(i)} \binom{n}{2} + \cdots + w_k^{(i)} \binom{n}{k} \\ \vdots \\ w_{k-1}^{(i)} + w_k^{(i)} n \\ w_k^{(i)} \end{pmatrix}.$$

If now $w^{(i)} \neq 0$, then the above is a vector-valued polynomial whose entry of highest degree is in any case the first row corresponding to the eigenspace of A_i. Since this holds for each i, the lemma follows. □

PROOF OF THEOREM 3.30, PART (2). Let $g \in \mathbb{H}(\mathbb{R})$ have positive real eigenvalues, let $\mathbb{L}$ be the Zariski closure of $\Gamma \leqslant \mathbb{H}(\mathbb{R}) \leqslant \mathrm{SL}_d(\mathbb{R})$ and let

$$\rho : \mathrm{SL}_d \to \mathrm{SL}_D$$

and $w \in \mathbb{R}^D$ be the Chevalley representation for $\mathbb{L} = \mathrm{Stab}_{\mathrm{SL}_d}(\mathbb{R}w)$ as in Theorem 3.26. By Poincaré recurrence we have for almost every $x \in \Gamma\backslash\mathbb{H}(\mathbb{R})$ a sequence $n_k \to \infty$ with $xg^{n_k} \to x$ as $k \to \infty$. We now switch this convergence to the group level as follows: for almost every $h \in \mathbb{H}(\mathbb{R})$ there exist sequences $n_k \to \infty$, $\varepsilon_k \to e$, and $\gamma_k \in \Gamma$ with $\gamma_k h g^{n_k} = h\varepsilon_k$ for all $k \geqslant 1$, or equivalently with

$$\gamma_k = \underbrace{h\varepsilon_k h^{-1}}_{\to e} h g^{-n_k} h^{-1}.$$

Applying this group element to w gives

$$\frac{1}{\|w\|} w = \frac{1}{\|\rho(\gamma_n)w\|}\rho(\gamma_n)w = \lim_{k\to\infty} \frac{1}{\|\rho(hg^{-n_k}h^{-1})w\|}\rho(hg^{-n_k}h^{-1})w = v_h,$$

where we have used the fact that $\Gamma \leqslant \mathbb{L}(\mathbb{R})$ fixes $\mathbb{R}w$ by definition, and Lemma 3.31. It follows by the same lemma that w is an eigenvector of $\rho(hgh^{-1})$ for almost every h. Taking $h \to e$ shows that w is an eigenvector of $\rho(g)$ also and so $g \in \mathbb{L}(\mathbb{R})$. $\square$

PROOF OF THEOREM 3.30, PART (1). Let $H^o = \mathbb{H}(\mathbb{R})^o$ be the connected component of the set of real points of $\mathbb{H}$. Let F be a non-compact almost direct simple factor of H^o. Then F contains a one-parameter unipotent subgroup U, and we can apply Part (2) of the theorem to U and to all its conjugates, which together generate a normal connected subgroup of F (and hence all of F). Thus $\mathbb{L}(\mathbb{R})$ contains F. We may apply this for all non-compact almost direct factors of $\mathbb{H}$, which then proves the second claim in Part (1).

This also proves the first claim in Part (1) since by the above $\mathbb{L}$ and $\mathbb{H}$ have the same Lie algebra and hence have the same dimension. However, $\mathbb{H}$ is by assumption connected and so $\mathbb{L} = \mathbb{H}$ follows. $\square$

Exercises for Section 3.5

Exercise 3.5.1. Let Q be a real non-degenerate quadratic form of signature (p,q) in d variables with $p \geqslant q \geqslant 1$. Suppose the orbit $\mathrm{SL}_d(\mathbb{Z})\,\mathrm{SO}(Q)(\mathbb{R})$ has finite volume. Show that a multiple of Q has integer coefficients.

3.6 Irreducible Quotients

In this section we classify lattices in semi-simple groups into reducible and irreducible lattices, and derive interesting density results (in the standard topology) from the Borel density theorem (which gives weak Zariski density).

Definition 3.32. Let G be a connected semi-simple Lie group. A lattice $\Gamma < G$ is called *reducible* if $G = H_1 \cdot H_2$ can be written as an almost direct product of nontrivial connected Lie subgroups $H_1, H_2 \leqslant G$ such that $\Gamma_1 = \Gamma \cap H_1$ is a lattice in H_1 and $\Gamma_2 = \Gamma \cap H_2$ is a lattice in H_2. The lattice is called *irreducible* if it is not reducible.

Examples of reducible lattices are of course very easy to find, for example $\mathrm{SL}_2(\mathbb{Z}) \times \mathrm{SL}_2(\mathbb{Z})$ is a reducible lattice in $\mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$. Irreducible lattices are a bit more difficult to find[†], but for now we only note that, for example, $\mathrm{SL}_2(\mathbb{Z}[\sqrt{2}])$ can be made into an irreducible lattice in $\mathrm{SL}_2(\mathbb{R})\times\mathrm{SL}_2(\mathbb{R})$, see Exercise 3.6.1.

Corollary 3.33 (Dense projections of irreducible lattices). *Let*

[†] By definition any lattice in a simple group is irreducible, but let us discuss a more interesting example.

$$G = H_1 H_2 \subseteq \mathrm{SL}_d(\mathbb{R})$$

be an almost direct product of the connected components of the groups of $\mathbb{R}$-*points* $H_1, H_2 \subseteq \mathrm{SL}_d(\mathbb{R})$ *of two semi-simple algebraic groups defined over* $\mathbb{R}$. *We assume furthermore that* G *has no compact factors. Let* $\Gamma < G$ *be an irreducible lattice in* G, *and suppose that* H_2 *is non-trivial. Then the projection of* Γ *to*

$$G/C(G)H_2 \cong H_1/C(H_1)$$

is dense in $H_1/C(H_1)$.

Proof. We note that G is also the connected component of the group of $\mathbb{R}$-points of its Zariski closure. In fact if $\mathbb{H}_1, \mathbb{H}_2$ are the algebraic groups giving rise to H_1, H_2 then $\mathbb{G} = \mathbb{H}_1\mathbb{H}_2$ is an algebraic group defined over $\mathbb{R}$ with the property that $G = \mathbb{G}(\mathbb{R})^o$. Also, if $F \lhd G$ is any connected normal subgroup, then $F = \mathbb{F}(\mathbb{R})^o$ for a normal algebraic subgroup $\mathbb{F} \lhd \mathbb{G}$. In fact, if $\mathfrak{g} = \mathfrak{f} + \mathfrak{f}'$ is a decomposition of the Lie algebra $\mathfrak{g}$ of G into the Lie algebra $\mathfrak{f}$ of F and a transversal Lie ideal $\mathfrak{f}'$ of $\mathfrak{g}$, then $\mathbb{F} = C_{\mathbb{G}}(\mathfrak{f}')^o$. Therefore we may apply the Borel density theorem (Theorem 3.30) for $\mathbb{G}$ or any of its normal subgroups.

Write

$$\pi_1 \colon G \longrightarrow G/C(G)H_2 \cong H_1/C(H_1)$$

for the projection map. There are two cases to consider: either $\pi_1(\Gamma)$ is discrete or it is not.

Discrete image implies reducibility. If $\pi_1(\Gamma)$ is discrete and $B_1 \subseteq H_1$ is a fundamental domain for the discrete pre-image of $\pi_1(\Gamma)$ in H_1 and $B_2 \subseteq H_2$ is a fundamental domain for $\Gamma \cap H_2$ in H_2, then we claim that $B_1B_2 \subseteq G$ is an injective domain for Γ. Indeed, if $\gamma \in \Gamma$, $b_1, b_1' \in B_1$, and $b_2, b_2' \in B_2$ satisfy $\gamma b_1 b_2 = b_1' b_2'$, then this identity modulo $H_2 \lhd G$ gives

$$(\gamma H_2)\,(b_1 H_2) = b_1' H_2.$$

Interpreting this in H_1 gives

$$\eta b_1 = b_1'$$

for some η with $\pi_1(\eta) = \pi_1(\gamma)$. By our assumption that B_1 is a fundamental domain, it follows that $b_1 = b_1'$. Multiplying $\gamma b_1 b_2 = b_1' b_2'$ with b_1^{-1} we get $\gamma b_2 = b_2'$ and $\gamma \in H_2$. Now $b_2 = b_2'$ and $\gamma = I$ by the injectivity assumption on B_2. Hence $B_1B_2 \subseteq G$ is an injective domain for Γ, and has finite Haar measure since Γ is a lattice by assumption. This also implies that[†] each of B_1 and B_2 has finite Haar measure. This implies that $\Gamma \cap H_2$ is a lattice in H_2.

By the Borel density theorem (Theorem 3.30) applied to $\Gamma \cap H_2 \subseteq H_2$ there is a finite collection $\{\gamma_1, \ldots, \gamma_n\} \subseteq \Gamma \cap H_2$ such that

---

[†] As G is the almost direct product of H_1 and H_2 the Haar measure m_G is, in the same sense, also almost the product of the Haar measures $m_{H_1} \times m_{H_2}$.

$$C(\gamma_1,\dots,\gamma_n)=\{h\in H_2 \mid \gamma_i h=h\gamma_i \text{ for } i=1,\dots,n\}$$

is the center $C(H_2)$ of H_2. In fact, we may choose $\gamma_1 \in \Gamma\cap H_2\smallsetminus C(H_2)$ and then successively choose $\gamma_2,\dots$ so that at each stage

$$C(\gamma_1,\dots,\gamma_m)\subsetneq C(\gamma_1,\dots,\gamma_{m-1}).$$

By the Noetherian property, we must find some n with

$$C(\gamma_1,\dots,\gamma_n)=C(\Gamma\cap H_2).$$

Since $\Gamma\cap H_2$ is Zariski dense in H_2 we deduce that

$$C(\gamma_1,\dots,\gamma_n)=C(H_2)$$

as required.

We claim that this implies that

$$\pi_2(\Gamma)\subseteq H_2/C(H_2)$$

must be discrete as well. In fact, if $\pi_2(\gamma)$ is sufficiently small but non-trivial, then by construction

$$[\pi_2(\gamma),\pi_2(\gamma_m)]\neq I$$

for some $m\in\{1,\dots,n\}$, and then

$$[\gamma,\gamma_m]\in H_2\cap\Gamma$$

is very close to an element of $C(H_2)$ but does not belong to $C(H_2)$. However, $C(H_2)$ is finite (it is zero-dimensional because its Lie algebra is trivial). This contradicts the assumed discreteness of Γ, so $\pi_2(\Gamma)$ must be discrete as claimed.

The claim establishes a symmetry between H_1 and H_2 in the above discussion. Applying the argument above to H_1 we also see that $\Gamma\cap H_1$ is a lattice in H_1. In other words, we have shown that Γ is a reducible lattice.

Showing density. This shows that we may assume that $\pi_1(\Gamma)$ is not discrete. Let

$$F=\pi_1^{-1}\left(\overline{\pi_1(\Gamma)}\right)\cap H_1$$

be the pre-image in H_1 of the closure of $\pi_1(\Gamma)$. Clearly Γ stabilizes the Lie algebra $\mathfrak{f}$ of F. By the Borel density theorem (Theorem 3.30) applied to the lattice Γ in G, the same holds for $G\geqslant H_1$. It follows that $\mathfrak{f}\lhd\mathfrak{h}_1$ is a Lie ideal in the Lie algebra $\mathfrak{h}_1$ of H_1.

If $\mathfrak{f}=\mathfrak{h}_1$, then we get the desired density of $\pi_1(\Gamma)$ in $H_1/C(H_1)$. So suppose that $\mathfrak{f}\neq\mathfrak{h}_1$, and define H_1' to be the almost direct product of all factors of H_1 whose Lie algebra is not contained in $\mathfrak{f}$. Also define H_2' to be the almost direct product of H_2 and all factors of H_1 whose Lie algebra is contained in $\mathfrak{f}$. Since $\mathfrak{f} \neq \mathfrak{h}_1$, the group H_1' is non-trivial. If π_1' denotes the analogous projection for the almost direct product $G = H_1'H_2'$ then we see that

$$\pi_1'(\Gamma) \subseteq H_1'/C(H_1')$$

is discrete. By the first argument in the proof, this implies that Γ is a reducible lattice. Therefore irreducibility of the lattice implies that $\mathfrak{f} = \mathfrak{h}_1$ and the result follows. □

While Corollary 3.33 gives interesting results for irreducible lattices, it can also apply in a weaker (potentially trivial) form to reducible lattices. This is because every reducible lattice can be 'reduced', or 'almost decomposed' into irreducible lattices as follows. If $\Gamma < H_1H_2$ is a reducible lattice such that $\Gamma \cap H_i < H_i$ is a lattice for $i = 1, 2$, then

$$(\Gamma \cap H_1)(\Gamma \cap H_2) \subseteq \Gamma$$

is also a lattice in H_1H_2 and so has finite index in Γ. Studying now

$$\Gamma \cap H_i < H_i$$

we may obtain an irreducible lattice, and if not we may repeat the decomposition step as before. Ultimately we find finitely many irreducible lattices (that are potentially lattices in simple groups).

In this context the following notion is useful. Let $\Gamma, \Lambda < G$ be two subgroups. Then we say that Γ and Λ are *commensurable* if $\Gamma \cap \Lambda$ has finite index in both Γ and Λ.

Our interest in the notion of irreducibility is clearly explained in the following corollary.

Corollary 3.34 (Mixing of semi-simple groups). *Let G be the connected component of the group of $\mathbb{R}$-points of a semisimple algebraic group defined over $\mathbb{R}$. Suppose that G has no compact factors. Let $X = \Gamma\backslash G$ be the quotient by an irreducible quotient of G. Then every almost direct factor of G acts ergodically and the action of G is mixing with respect to the Haar measure m_X on X.*

Proof. By the Howe–Moore theorem for semi-simple groups (Theorem 2.15), it is sufficient to show that every simple factor acts ergodically.

So let $F \lhd G$ be a (non-trivial) simple factor of G, and suppose that

$$F{\scriptstyle\bullet}B = B \subseteq X$$

is F-invariant[†].

Let

[†] By [53, Prop. 8.3], we may assume the strict invariance $F{\scriptstyle\bullet}B = B$ rather than the *a priori* weaker invariance in the measure algebra $m_X(g{\scriptstyle\bullet}B \triangle B) = 0$ for all $g \in F$.

$$\pi_X\colon G \longrightarrow X = \Gamma\backslash G$$

be the natural factor map, and let $B_G = \pi_X^{-1}(B) \subseteq G$ be the set in G corresponding to B. By the properties of B we have $B_G F = B_G$, or equivalently $B_G = \pi^{-1}(\pi(B_G))$ if $\pi : G \to G/F$ denotes the projection map. By construction, $\Gamma B_G = B_G$ and so $\pi(\Gamma)\pi(B_G) = \pi(B_G)$.

Recall from [53, Prop. 8.6] that, for any two Borel sets $B_1, B_2 \subseteq G/F$ with $m_{G/F}(B_1)m_{G/F}(B_2) > 0$, the set

$$\bigl\{gF \in G/F \mid m_{G/F}\left(gFB_1 \cap B_2\right) > 0\bigr\}$$

is non-empty and open.

We may apply this to the sets $B_1 = \pi(B_G)$ and $B_2 = G/F \smallsetminus \pi(\Gamma)$. Since $\pi(\Gamma)$ is dense in G/F by Corollary 3.33, we deduce that either $\pi(B_G)$ has zero measure or its complement does. Since G is the almost direct product of F and G/F, we see that either B_G or its complement has zero measure in G. It follows that either $m_X(B) = 0$ or $m_X(X \smallsetminus B) = 0$ as required. □

Exercises for Section 3.6

Exercise 3.6.1. Let $D > 1$ be a non-square integer, and for

$$\alpha = a + b\sqrt{D} \in \mathbb{Q}(\sqrt{D})$$

let $\overline{\alpha} = a - b\sqrt{D}$ denote its Galois conjugate. Now let

$$\mathrm{SL}_2(\mathbb{Z}[\sqrt{D}]) = \left\{ g = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \mid \alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22} \in \mathbb{Z}[\sqrt{D}], \det(g) = 1 \right\},$$

and consider $\mathrm{SL}_2(\mathbb{Z}[\sqrt{D}])$ as a subgroup of $\mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ using the *diagonal embedding*

$$\begin{aligned} \imath\colon\ \mathrm{SL}_2(\mathbb{Z}[\sqrt{D}]) &\longrightarrow \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R}) \\ g = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} &\longmapsto (g, \overline{g}) \end{aligned}$$

where

$$\overline{g} = \begin{pmatrix} \overline{\alpha}_{11} & \overline{\alpha}_{12} \\ \overline{\alpha}_{21} & \overline{\alpha}_{22} \end{pmatrix}.$$

(a) Show that $\Gamma = \imath\bigl(\mathrm{SL}_2(\mathbb{Z}[\sqrt{D}])\bigr) \leqslant \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ is a discrete subgroup.
(b) Show that Γ is a lattice in $\mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$.

Notes to Chapter 3

(10) (Page 78) Almost any algebra text will cover this material, for example Gerstein [69] or, for the more sophisticated aspects of the algebraic theory, see Lam [109].

(11)(Page 79) The word signature is used in various ways, all meaning that the number of $+1$s, -1s (and in the degenerate case 0s) can be reconstructed from the signature (and the dimension). The fact that the signature is a property of the form itself is Sylvester's law of inertia [177].
(12)(Page 80) Hilbert [78] proved this in his development of invariant theory.
(13)(Page 89) This was shown by Dirichlet [35] in 1846 for the ring $\mathbb{Z}[\zeta]$ (the understanding that this may not be the ring of integers in $\mathbb{Q}(\zeta)$ for an algebraic integer ζ came later, and of course the rank is not affected as $\mathbb{Z}[\zeta]$ has finite index in the ring of integers). We refer to the paper of Elstrodt [58] for an account of the history.
(14)(Page 105) The history, and various generalizations, of the implicit function theorem may be found in the account by Krantz and Parks [108]. The p-adic implicit function theorem may be found in the notes of Serre [167, p. 83].
(15)(Page 115) A modern proof from a sophisticated point of view is given by Conrad [17], and the original proof in Chevalley [14]. Any book on algebraic groups will contain a version of the theorem (possibly not under this name).
(16)(Page 120) Borel [7] proved this for semi-simple Lie groups without compact factors; generalizations and simplifications have been provided by Furstenberg [68], Moskowitz [139] and Dani [19] among others.